Integration of Microscopic Big Traffic Data in Simulation-Based Safety Analysis

SAFETY RESEARCH USING SIMULATION

UNIVERSITY TRANSPORTATION CENTER

Mohamed Abdel-Aty, PhD, PI
Pegasus Professor, Chair
Department of Civil, Environmental and Construction Engineering
University of Central Florida

Qi Shi, PhD
Ling Wang, PhD Candidate
Yina Wu, PhD Student
Essam Radwan, PhD, Professor
Binya Zhang, MS
Department of Civil, Environmental and Construction Engineering
University of Central Florida

Integration of Microscopic Big Traffic Data in Simulation-Based Safety Analysis

Mohamed Abdel-Aty, PhD, PI
Pegasus Professor, Chair
Department of Civil, Environmental and Construction Engineering
University of Central Florida

Yina Wu, PhD Student
Graduate Research Assistant
Department of Civil, Environmental and Construction Engineering
University of Central Florida

Qi Shi, PhD
Postdoctoral Research Associate
Department of Civil, Environmental and Construction Engineering
University of Central Florida

Essam Radwan, PhD
Professor
Department of Civil, Environmental and Construction Engineering
University of Central Florida

Ling Wang, PhD Candidate
Graduate Research Assistant
Department of Civil, Environmental and Construction Engineering
University of Central Florida

Binya Zhang, MS
Graduate Research Assistant
Department of Civil, Environmental and Construction Engineering
University of Central Florida

A Report on Research Sponsored by SAFER-SIM

March 2016

Table of Contents

List of Figures

List of Tables

Abstract

The advent of the Big Data era has transformed the outlook of numerous fields in science and engineering. The transportation arena expects to take advantage of Big Data through the popularization of Intelligent Transportation Systems (ITS). Judging by the data size, richness of information, and collection speed of data, ITS traffic detection systems serve as ideal sources of big traffic data. Collected from various sources, the data provide insight about the facilities at a microscopic level in real-time. Consequently, efficient integration and utilization of such data for better performance of transportation systems becomes a critical issue for traffic operators. In this project, different applications of the real-time microscopic traffic data were explored with a focus on operation efficiency and traffic safety. The applications range from direction measurement of operation efficiency and safety using the real-time traffic data to indirect use of the data in simulation tools. To achieve this goal, multiple tasks were set up and the corresponding efforts carried out.

Currently, three major types of traffic detection technologies exist in field practice. In-roadway detectors are a relatively mature technology with a long history of application. Nevertheless, they could disrupt traffic during installation and maintenance. Over-roadway detectors have an advantage over in-roadway detection systems because they are installed along the roadside and will not affect traffic during installation and maintenance. Off-road detectors, especially probe vehicles, are becoming popular yet have not been widely deployed because they require in-vehicle devices. This project used traffic data from two point-based over-roadway detection systems, namely the Microwave Vehicle Detection System (MVDS) and the Video Image Processing (VIP) system, and data from a segment-based probe-vehicle system, the Automatic Vehicle Identification (AVI) system, to conduct operation and safety evaluations.

Traffic safety is an important indicator of highway performance. This project carried out real-time safety evaluation for the studied expressways using MVDS data. To differentiate crash and non-crash conditions and identify the significant factors contributing to crash occurrence, logistic regressions were tested. Crash precursors for crashes on the mainline and ramps were identified. Real-time traffic flow parameters were found to significantly contribute to crash occurrence. The safety investigation of ramp crashes utilized not only real-time traffic data, but also weather data from nearby airports. It was found that both traffic and weather conditions would affect the crash likelihood in real-time. Both mainline and ramp crash prediction achieved relatively high accuracy, indicating the usefulness of real-time traffic data in traffic safety evaluation.

Traffic operation was evaluated from two aspects: congestion measurement and travel time reliability. The AVI system showed its capability in both tasks. Travel-time-based congestion measures using AVI data provided detailed congestion evaluation on spatial-temporal dimensions. By continuously monitoring the traffic conditions, traffic operators could have accurate information about their system at both a general and a detailed level and make well-informed decisions. Travel time reliability was also evaluated using AVI data. The strengths of AVI application in travel time reliability lie in that it reflects travel time in a straightforward manner and detects travel time at an individual vehicle level (though not a total vehicles level). Different reliability measures were introduced and the same conclusions were drawn. Based on performance, AVI data could adequately identify travel time reliability conditions on expressways.

Traffic data can also be used in micro-simulation. Microscopic simulation allows traffic operators to evaluate traffic conditions or test the impact of a specific management strategy in a simulated environment. However, to achieve convincing results, the traffic input has to be as close as possible to real traffic conditions. This project incorporated real-time traffic data corresponding to fog conditions into the simulation to evaluate the

changes in traffic flow. By using real-time traffic data at short-time intervals, validation and calibration of the microscopic simulation model could be significantly enhanced. The Surrogate Safety Assessment Model (SSAM) was then introduced to evaluate traffic conflicts under the specific adverse weather type (fog). Through simulation, potential countermeasures could be proposed and their effects could be estimated without actual deployment of expensive devices. It was hoped that the real-time field traffic data would facilitate an understanding of the effects of adverse weather and other factors on driver behavior in the future.

In urban areas, safety at signalized intersections is a critical issue. Drivers' behavior in dilemma zones is especially worth exploration. Cellular Automaton (CA) models are simulation tools to simulate drivers' behavior in dilemma zone analysis. This project evaluated the viability of real-time traffic data from a video processing system in CA simulation. The use of real-time traffic data was not common in previous studies. Multiple variables regarding vehicle speed, distance to intersection, relative position to other vehicles, and more were extracted from the video data. These data were then matched with drivers' stop/go decisions at intersections. Rear-end risks and red-light-running risks were evaluated. Appropriate countermeasures (e.g., pavement marking, flashing, and the combination of the two) to cope with dilemma zone problems and their effects were identified.

Overall, multiple tasks in this project demonstrated the great potential to apply microscopic traffic data to traffic safety studies and simulation applications.

CHAPTER 1: Introduction

The abundance of information is a distinguishing feature of the 21st century. Nearly all aspects of social activities have experienced some degree of transformation due to the insights introduced by this information. In the transportation arena, the rapid popularization of Intelligent Transportation Systems (ITS) in the past few decades has brought new perspectives to traffic operations and management. The ITS traffic data often contain microscopic traffic flow information and are collected continuously from different sources over a vast geographical scale. Huge in size and rich in information, the seemingly disorganized data have great potential for operation efficiency and traffic safety evaluation and improvement.

Operation efficiency and traffic safety have long been deemed as priorities among highway system performance measurements. While efficiency could be evaluated in terms of traffic congestion and travel time reliability, safety is often improved by identifying how traffic flow parameters affect crash occurrence. Recently, simulations also serve as a cost-effective way to evaluate traffic efficiency and safety in simulated scenarios allowing researchers and practitioners to test proposed improvement strategies.

The main objectives of this study were to explore different vehicle detection systems, focusing on the most widely deployed infrastructure-based sensing technologies. Applications of these types of data for efficiency and safety evaluation, and simulation, were investigated.

To achieve the proposed objectives, several tasks were carried out.

- Task 1: Identify and collect continuous measurements of traffic conditions at different levels: segment-based, point-based, and vehicle based;
- Task 2: Validate the use of different microscopic data sources;

- Task 3: Measure traffic efficiency: congestion measurement and travel time reliability in real-time;
- Task 4: Evaluate traffic safety: relationship between traffic flow (speed, flow, density, congestion) and crash occurrence, crash precursors in real-time;
- Task 5: Simulate traffic flow under adverse weather conditions in micro-simulation. Real-time traffic data (combined as needed with weather data) could be used to tune traffic flow under different conditions in micro-simulation;
- Task 6: Investigate the use of video-based data;
- Task 7: Develop a simulation model using video-based parameters;
- Task 8: Use the new simulation model to investigate dilemma zone decisions at signalized intersections.

Following this chapter, this report begins with a literature review on existing sensing technologies in Chapter 2. Chapter 3 describes the data that were collected for the purpose of this study (Tasks 1 and 2). Chapter 4 focuses on applications of the microscopic data in traffic efficiency measurement, which included congestion measurement and travel time reliability in real-time (Task 3). Chapter 5 demonstrates how these data could be used for proactive traffic safety evaluation in real-time (Task 4). Chapter 6 offers insights about micro-simulation utilizing the traffic data under different conditions (Task 5). Chapters 7 and 8 discuss the utilization of the video-based traffic data and develop a simulation tool to study drivers' decision-making at dilemma zones near signalized intersections (Tasks 6-8). Conclusions are summarized in Chapter 9.

CHAPTER 2: Literature Review

A wide variety of traffic detection technologies have been tested and applied to improve the performance of automatic traffic monitoring. According to Martin et al. (2003), state-of-the-art detection technologies fall into three categories as shown in Figure 2.1. in-roadway detectors, over-roadway detectors, and off-roadway technologies. The terms “in-roadway” and “over-roadway” are used in the Traffic Detector Handbook (Klein et al., 2006). In-roadway detectors are commonly known as intrusive detectors and over-roadway detectors are known as non-intrusive detectors. Off-roadway detector technologies include probe vehicles and remote sensing. Probe vehicles could be further broken down into GPS, cellular phone, Bluetooth, Automatic Vehicle Location (AVL), and Automatic Vehicle Identification (AVI) technologies.

An in-roadway sensor is one that is embedded in the pavement of the roadway, embedded in the subgrade of the roadway, or taped or otherwise attached to the surface of the roadway (Mimbela and Klein, 2000). The most common in-roadway detector technologies include inductive loop detectors, weigh-in-motion sensors, magnetometers, tape switches, micro-loops, pneumatic road tubes, and piezoelectric cables. In-roadway detector technologies have been implemented since the early stage of automatic traffic surveillance, thus they are applications of relatively mature technologies. However, they have several drawbacks including disruption of traffic for installation and repair. They also have high failure rates in certain conditions (Martin et al., 2003), such as poor road surfaces and adverse weather conditions.

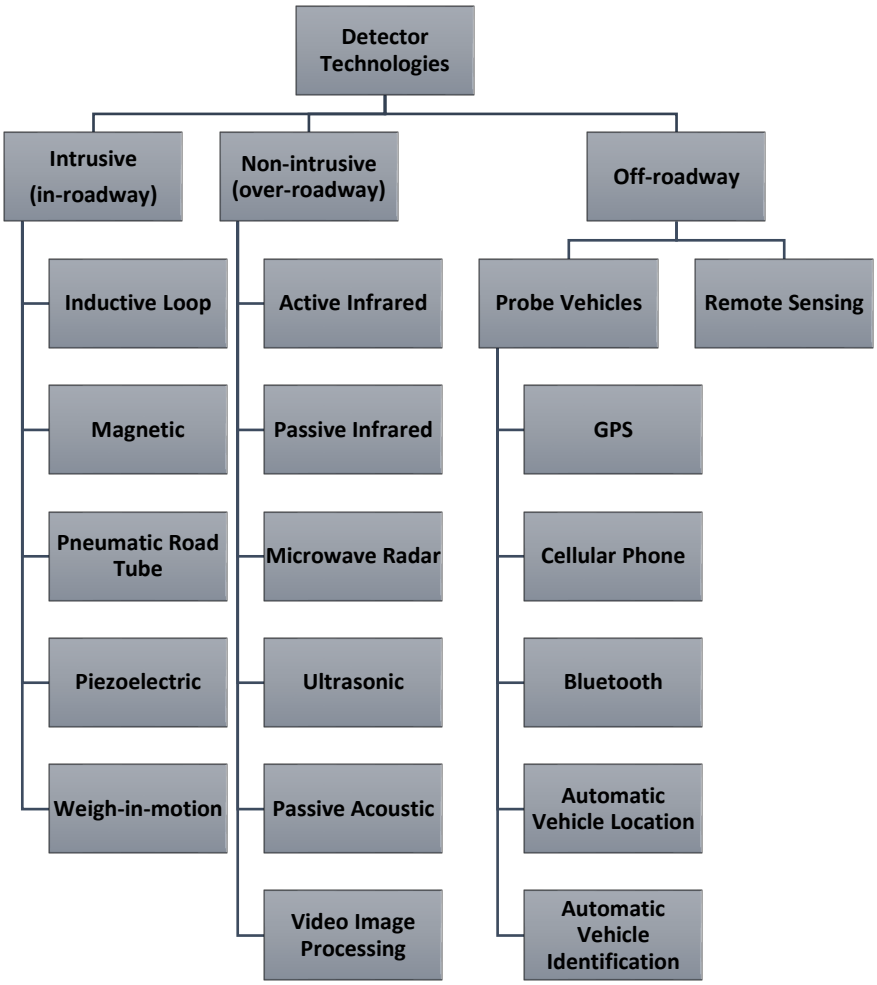


Figure 2.1 - Automatic traffic detection technologies

An over-roadway sensor is one that is mounted above the roadway itself or alongside the roadway, offset from the nearest traffic lane by some distance. Existing over-roadway sensors include video image processors, microwave radar sensors, ultrasonic sensors, passive infrared sensors, laser radar sensors, and passive acoustic sensors (Mimbela and Klein, 2000). Compared with in-roadway sensors, over-roadway sensors have a significant advantage in that they minimize the disruption of traffic during installation and maintenance.

In addition to the intrusive and non-intrusive traffic detection technologies, off-roadway technologies are developing quickly. Probe vehicles and remote sensing are

currently two new sensing technologies. Probe vehicles require in-vehicle devices and have opened new fields for traffic researchers because they track individual vehicle movement. Aircraft or satellites perform remote sensing, and the technology applies aircraft or satellite images to analyze and extract traffic information (Martin et al., 2003). However, for real-time traffic monitoring its utilization is quite limited.

Probe vehicle technologies include GPS, cellular phones, Bluetooth, AVI, and AVL. They collect real-time traffic data for operation monitoring, incident detection, and route guidance. Although probe vehicle systems require high implementation costs and fixed infrastructure, they have the advantages of continuous data collection and no disruption to traffic (Martin et al., 2003).

Of the detector systems introduced above, three systems that served as traffic data sources in this research will be discussed in detail: microwave radars, AVI, and video image processing.

Microwave radars are normally mounted adjacent to roadways. They are typically insensitive to inclement weather. In addition, they offer direct measurement of speed and multiple lane operation. Two types of microwave detectors exist: Doppler microwave detectors and Frequency-Modulated Continuous Wave (FMCW) detectors (also referred to as true-presence microwave detectors). Both types of detectors can detect volume, occupancy, classification, and speed. The true-presence vehicle detectors can detect stopped vehicles, whereas Doppler microwave detectors can only recognize vehicles above a minimum speed.

Automatic Vehicle Identification (AVI) refers to various components and processes that allow for the identification of vehicles for the purposes of charging a toll or providing data for various traffic-management strategies. Currently, two major types of AVI technology are deployed: laser and Radio Frequency (RF). Laser systems utilize a bar-coded sticker attached to the vehicle. However, detection is sensitive to weather and

dirt. RF systems use a transponder attached to vehicles and tag readers to identify the unique tag ID for Electronic Toll Collection. Compared to other intrusive or non-intrusive detector systems, AVI systems have the ability to provide space mean speed information and travel time information (Riley, 1999). Nevertheless, the data collection capability of the AVI system depends on the coverage area of AVI infrastructures (Martin et al., 2003).

Video image processing (VIP) systems normally consist of one or several video cameras, microprocessor-based equipment for processing the imagery, and software for interpreting the images and outputting traffic data (Martin et al., 2003). VIP detectors can monitor rich traffic information. In contrast to AVI and microwave detectors, VIP detectors can detect traffic on multiple lanes over a limited zone area. Passing vehicles in the images cause variations in the gray levels of the black-and-white pixel groups, so vehicle passage can be determined based on the variations. With image processing software, classification by length, speed, and headway, vehicle presence, volume, and more can be detected. Nevertheless, one limitation that could lower data quality is environmental factors, especially poor light conditions.

With the fast development of traffic detection technologies, their applications in traffic management become a significant topic for traffic operators. Considering each type of traffic detection technology has a unique functionality, specific traffic management applications can exploit each system (Antoniou et al., 2011). Based on the potential applications of these detection systems, they can be classified as point, point-to-point, and area-wide systems.

Inductive loop detectors, microwave detectors, and video image detectors all belong to the point-based sensors classification since they reflect traffic conditions at the installed locations. The AVI system is point-to-point (segment) based. The same vehicle is identified at various locations, and that data is used to calculate travel time and space mean speed. GPS and cellular phones can be deemed area-wide (vehicle-based)

sensors. They collect travel information from individual vehicles continuously over a large area as long as the vehicles are equipped with these devices. These data collection technologies create new opportunities in dynamic traffic management as well as other aspects of traffic simulation and prediction (Antoniou et al., 2011).

The benefits of traffic detection technologies include direct and indirect applications. Direct applications include congestion reduction, automatic incident detection (AID), and travel time estimation. Indirect applications are carried out through enhancement of traffic modeling in the model development, calibration and validation processes, and traffic simulation.

CHAPTER 3: Data Collection and Comparison

This study collected traffic data from different sources. The MVDS and AVI traffic data were collected from the urban expressways operated by the Central Florida Expressway Authority (CFX). Video data were recorded at an intersection in Orlando, Florida.

When CFX converted mainline toll plazas to open tolling express lanes, they adopted the AVI system for Electronic Toll Collection. If vehicles traveling on CFX's expressways are equipped with E-PASS or SunPass, they do not have to stop or slow down to pay the tolls. The AVI detectors will keep records of vehicle information and calculate tolls according to the distance that the vehicles traveled. Although AVI detectors can archive traffic information, they are not designed for this objective because only vehicles equipped with transponders can be detected. Starting in 2012, CFX introduced MVDS to their expressway network. These detectors are installed specifically for traffic monitoring. The two systems exhibit substantial differences. However, both can be leveraged to provide traffic professionals with valuable traffic information. In this study, both types of data were collected from three expressways (SR 408, SR 417 and SR 528) operated by CFX, as illustrated in Figure 3.1.

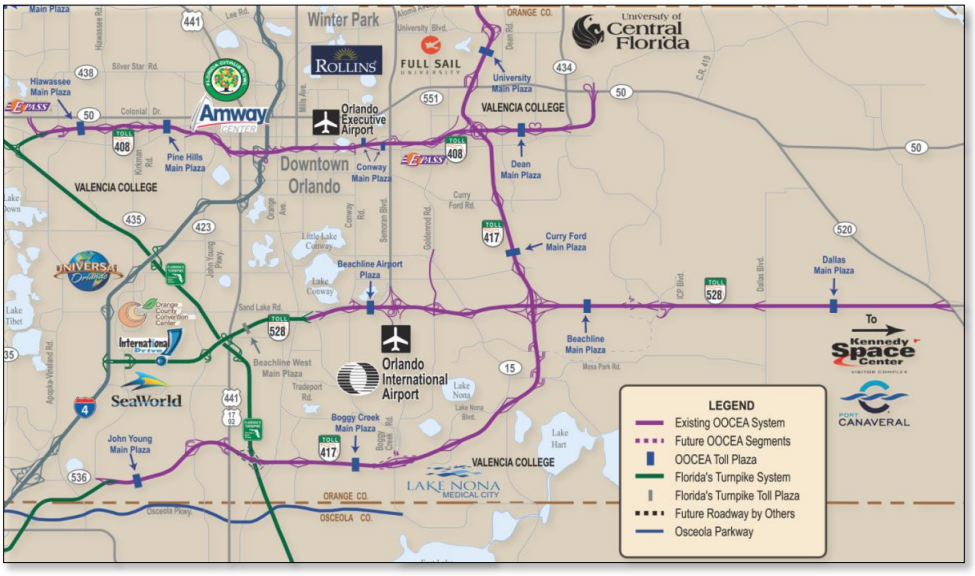


Figure 3.1 - Urban expressway system involved in the study (CFX, 2014)

3.1 AVI Traffic Data

The AVI traffic data were collected from September 2012 to December 2014. AVI detectors are installed at toll plazas for Electronic Toll Collection and at other locations for travel time estimation. Table 3.1 summarizes the active detector numbers for each expressway from July 2014 to December 2014.

Table 3.1 - Number of active AVI detectors

Expressway	Length	Direction	July	August	September	October	November	December
SR 408	22 miles	EB	23	23	21	21	21	21
		WB	18	18	17	18	18	17
SR 417	32 miles	NB	13	13	14	14	14	15
		SB	17	17	17	13	13	18
SR 528	22 miles	EB	7	6	8	8	9	9
		WB	8	8	8	10	10	10

SR 408 has the most detectors and the smallest average spacing between AVI detectors among the three expressways. SR 528 has relatively short AVI segments near

the international airport and west to SR 417. However, on the suburban segments leading to the coast area, the distance between adjacent AVI sensors could be greater than seven miles. The distance between adjacent sensors is determined on two basic criteria: the need for toll collection and the need for travel time estimation. In urban areas, the accessibility of the expressway has to accommodate the travelers' demand of entering and exiting the expressways. This makes the toll collection for a relatively short spacing necessary on SR 408 since it travels through the downtown Orlando area.

The raw AVI dataset records a vehicle's encrypted ID and timestamp when the vehicle passes an AVI detector. By ranking a vehicle's ID and its timestamp, the timestamp when two adjacent upstream and downstream detectors recorded a vehicle separately can be obtained. The travel time for a vehicle across a link is the timestamp difference. The link length is calculated based upon upstream and downstream detectors' mileposts. Employing travel time and link length, a vehicle's speed is determined using Equation 3.1.

$$\text{speed} = \frac{Link\ Length}{Travel\ Time} = \frac{|\text{milepost}_{\text{upstream}} - \text{milepost}_{\text{downstream}}|}{\text{timestamp}_{\text{downstream}} - \text{timestamp}_{\text{upstream}}} \quad (3.1)$$

Supposing speed limit is the free-flow speed, the free-flow travel time is the time a vehicle spent across a link at the free-flow speed. Comparing free -flow travel time with the timestamp difference between upstream and downstream, we can get the travel time index (TTI) by using Equation 3.2.

$$\text{TTI} = \frac{\text{timestamp}_{\text{downstream}} - \text{timestamp}_{\text{upstream}}}{Free\ Flow\ Travel\ Time} \quad (3.2)$$

The AVI data are based on individual vehicles equipped with electronic transponders; therefore, the traffic count using AVI sensors is not the complete traffic volume. Existing studies utilizing the AVI traffic speed data have shown that, in many cases, the AVI data are capped at the speed limit. This might be out of safety concerns to discourage speeding when providing motorists with estimated travel time. However, uncapped speed (see Figure 3.2) could reflect real traffic conditions and is more desirable when focusing on exploration of AVI data performance in traffic monitoring.

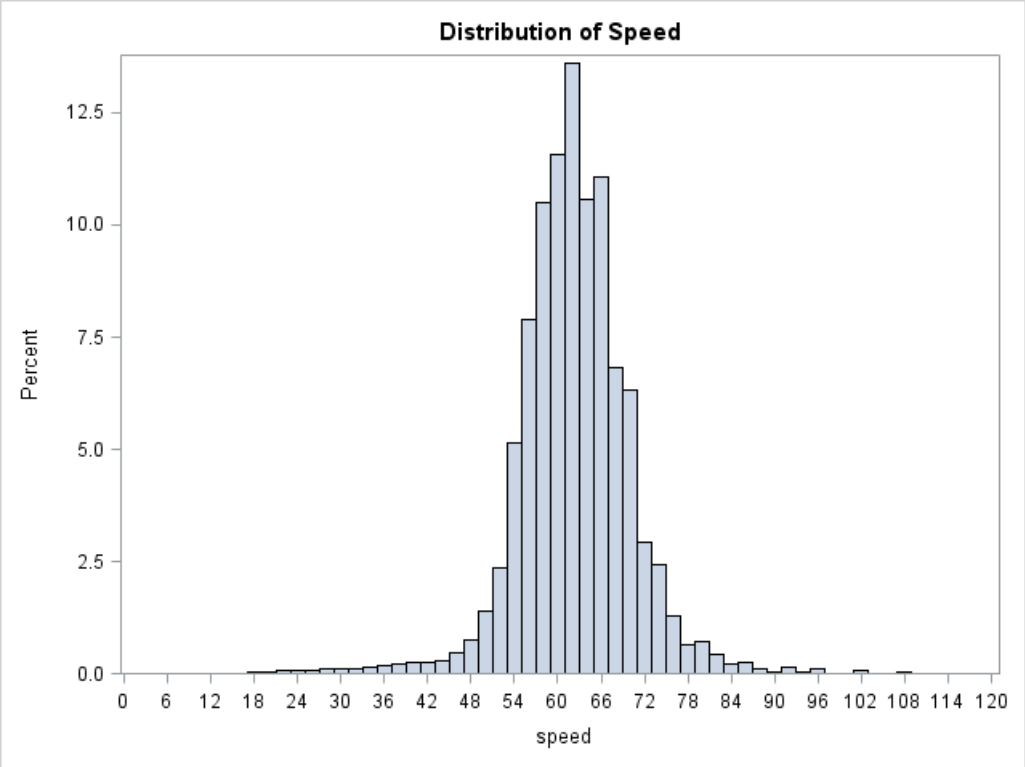


Figure 3.2 - SR 408 eastbound uncapped AVI data, December 2014

3.2 MVDS Traffic Data

This project uses MVDS data collected from July 2013 to December 2014. MVDS was initially introduced for traffic monitoring to CFX's expressways in 2012. MVDS does not identify individual vehicles. It returns aggregated traffic flow parameters at one-

minute intervals on each lane of the cross-section where the MVDS detector is installed. Archived traffic parameters by the MVDS sensors include traffic volume, time mean speed, lane occupancy, and traffic volume by vehicle length.

The MVDS traffic data also includes the timestamp when the sensor is polled. It was mentioned above that the sensors are polled every one minute. In addition, a unique sensor identifier and lane identifier are contained within the data. The sensor identifier consists of the roadway (i.e., SR 408, SR 417, or SR 528), milepost, and direction. The lanes are counted from the roadway median to the outside lane and fall into four categories: Mainline, Ramp, Mainline TP Express, and Mainline TP Cash (as shown in Table 3.2). Mainline TP Express indicates express lanes at mainline toll plazas; vehicles equipped with tags do not need to slow down on these lanes when they pass the toll plazas. Mainline TP Cash specifies a tollbooth at mainline toll plazas; vehicles need to stop and pay tolls. On expressways, these two types of lanes are physically separated.

Table 3.2 - MVDS on the CFX system

Route	Length (mi)	Direction	MVDS Detectors							
			Total	Mainline (including TP Express)	TP Cash	Ramp	Distance between adjacent detectors			
							Mean	*SD*	Min	Max
SR 408	21.4	EB	102	55	8	39	0.38	0.18	0.1	1
		WB	102	55	8	39	0.39	0.18	0.1	1
SR 417	31.5	NB	93	55	7	31	0.58	0.28	0.2	1.3
		SB	94	55	7	32	0.58	0.28	0.2	1.2
SR 528	22.4	EB	49	26	4	19	0.84	0.79	0.1	3
		WB	51	29	4	18	0.84	0.82	0.1	3.1

Compared with AVI traffic data, MVDS data reflect traffic states at their installed locations instead of at a segment. They also have several advantages over AVI data.

One advantage is traffic data for different types of lanes from MVDS sensors. Given that MVDS sensors monitor traffic conditions on each traveling lane, traffic data at toll plazas and on ramps can be collected. AVI data only provide traffic information of a cross-section on the mainline. To have a general understanding about expressway performance, analysis of toll plazas and ramps is necessary as well. Another advantage is the richness of traffic information from MVDS data. The traffic count from uncapped AVI data is not the complete traffic volume, whereas MVDS data includes speed, complete traffic volume, and lane occupancy as a surrogate measure of traffic density, as well as the volume by vehicle lengths.

3.3 Video Data

The video data collection was conducted in Orlando, Florida. The studied intersection is located at the northwest corner of the University of Central Florida, which plays a major role in stimulating economic and residential development in Orlando. This intersection is a four-legged signalized intersection between Corporate Blvd. and Gemini Blvd. running east-west and Alafaya Trail running north-south (see Figure 3.3) with a yellow interval of 4.3 seconds and an all-red interval of 1 second.

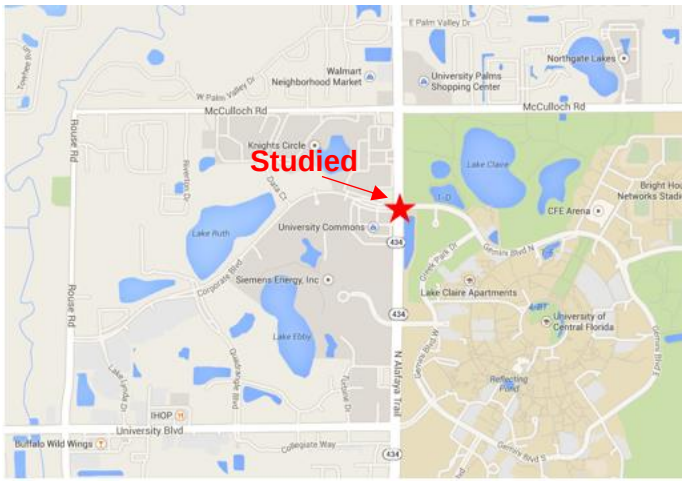


Figure 3.3 - Site location map (Google Maps, 2014)

This study specifically considered the northbound approach (see Figure 3.4).

(a)

(b)

Figure 3.4 - Studied intersection

(a) view from the intersection and (b) view into the intersection

Alafaya Trail southbound consists of five-lane divided traffic: three lanes for through movements, one exclusive left turn lane, and one exclusive right turn lane. Alafaya Trail northbound is three-lane divided traffic. Corporate Blvd. eastbound and Gemini Blvd. northbound are two-lane divided traffic. The existing posted speed limit on Alafaya Trail is 45 miles per hour (mph, see Figure 3.5). Thirty-six hours of video, which

includes 28 off-peak hours (1:30pm-4:30pm) and 8 peak hours (4:30pm-6:00pm), were filmed during the weekdays. Using Adobe Premiere Pro software to extract data from videos, 1292 vehicles' behavior was recorded. This does not include vehicles forced to stop by the vehicle in front. Due to the small sample size of light truck vehicles, and the fact that their crossing behaviors at intersections were different from other vehicles (Elmitiny et al., 2010), data from this vehicle type were excluded from the database.

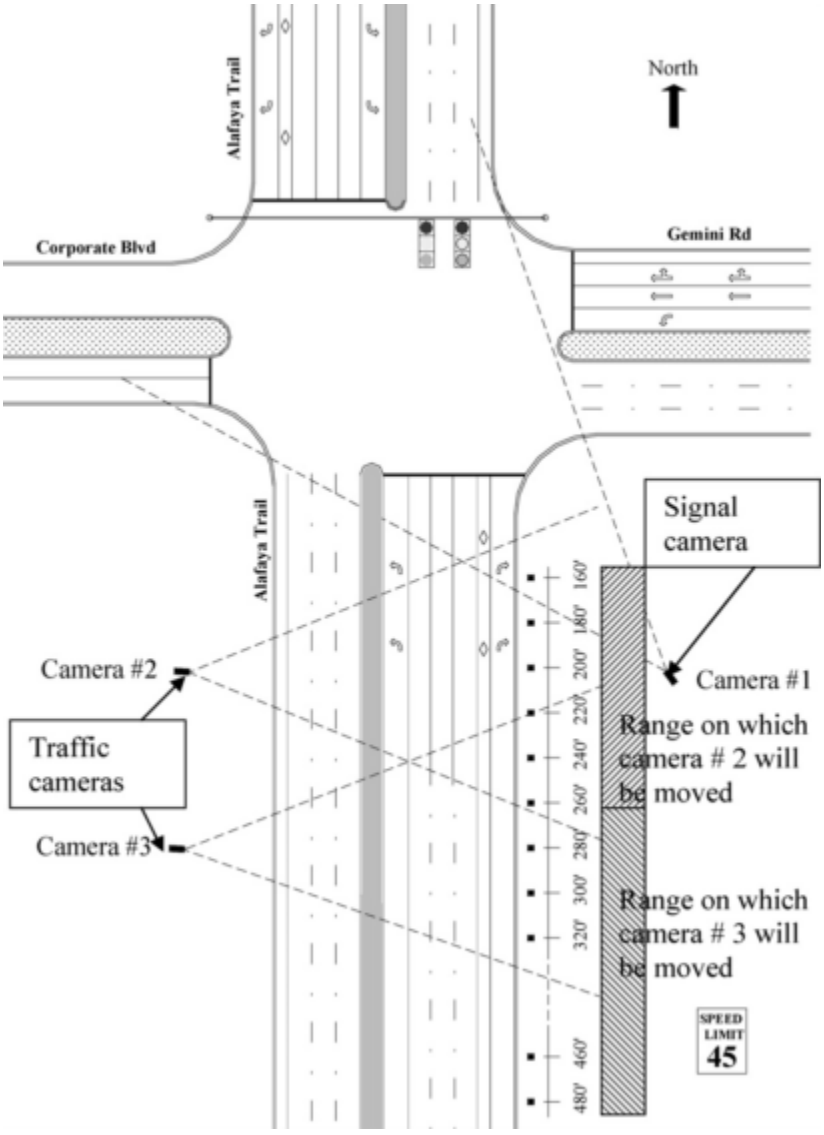


Figure 3.5 - Video-based system for data collection

Eight variables were obtained from the video:

- DISTANCE (in ft): the car's distance from the intersection at the onset of the yellow indication;
- SPEED (in mph): the car's operating speed at the onset of the yellow indication;
- ST_GO: the driver's stop/go decision (stop = 0; go = 1);
- Y_TIME (in seconds): if the car crossed the intersection, the value is the time elapsed from the onset of the yellow indication until the car entered the intersection; otherwise, the value is missing;
- RLR: whether or not the car ran a red light (no = 0; yes = 1);
- LD_FL: whether the car was in a leading position or a following position in the traffic flow (leading = 0; follow = 1); if headway was shorter than 1 s the car was considered following in the platoon;
- L_POSITION: the car's lane position (left lane = 0; middle lane = 1; right lane = 2);
- V_TYPE: vehicle type (passenger car (PC) = 0; light truck vehicle (LTV) = 1; larger size vehicle (LSV) = 2).

CHAPTER 4: Measurement of Traffic Efficiency

To evaluate how well the automatic traffic detection systems could reflect traffic efficiency, AVI and MVDS data on CFX's expressways were implemented to efficiency analysis. Performance was measured from three perspectives: congestion measurement, queue end detection, and travel time reliability.

4.1 Congestion Measurement

Traffic operation on expressways focuses on providing motorists with efficient movement to their destinations. To achieve this goal, reducing congestion is the most important task. Measuring congestion accurately is a prerequisite in congestion management. Traditionally, transportation authorities have implemented volume-to-capacity (V/C) ratios and level of service (LOS) as indicators of congestion intensity (Grant et al., 2011). Traffic demand can vary considerably in both temporal and spatial dimensions and roadway capacity can be reduced by incidents. In such cases, V/C ratios and LOS lack the capability to capture the variability of congestion. With the fast development of ITS technology, real-time congestion measurement is becoming an urgent call. On expressways, AVI traffic detection systems are employed to archive traffic data in a real-time manner. Congestion measures based on the AVI system were introduced and compared in this study.

Congestion can be measured using three characteristics, specifically traffic density, travel time, and travel speed. Data from the AVI system can be used to calculate the travel time of vehicles on a segment. Therefore, this study measured congestion by travel time. Travel time index (TTI) is the commonly accepted measure for evaluation of traffic congestion. It is defined as the ratio of actual travel time to an ideal (free-flow) travel time (Cambridge Systematics Inc., 2005), as shown in Equation 4.1

$$\text{TTI} = \frac{\text{actual travel time}}{\text{free-flow travel time}} \quad (4.1)$$

TTI indicates the additional time spent on a trip made during peak traffic hours compared to an ideal trip on the same corridor. On the CFX system, free-flow travel time is calculated based on segment length and average speed limit. The average speed limit of a segment recognizes that speed limits may vary within the segment. The levels of congestion and the corresponding TTI for the studied expressways are listed in Table 4.1, drawn from the Enhancing Expressway Operations Using Travel Time Performance Data (Griffin, 2011).

Table 4.1 - Travel time index and congestion levels

Functional Class	Travel Time Index for Different Congestion Levels		
	Below congestion threshold	Moderate Congestion	High Congestion
Freeway	Less than 1.25	1.25 – 1.99	Higher than 2.00

4.2 Expressway Congestion Evaluation

Two major efforts have been made to evaluate the performances of proposed congestion measures on the expressways. One is the evaluation of spatial-temporal distribution of congestion on the mainline. The other effort is to identify the ramps that experience congestion. The congestion conditions on SR 408 in December 2014 were examined and plotted. Then, the performances of different traffic detection systems were compared using this information.

To measure current expressway congestion conditions, the traffic data were aggregated at five-minute intervals and averaged by the weekdays for each month. Filled contour plots were generated to illustrate the spatial-temporal properties of the

congestion. The TTI congestion plots shown in Figure 4.1 and Figure 4.2 are the distributions of congestion on SR 408 eastbound and westbound in December 2014. According to the figure, congestion appears on eastbound lanes during the evening peak hours and on westbound lanes during the morning peak hours. In addition, the congestion occurs only at specific locations. The segment around MP 18 going eastbound experiences a high level of congestion, while the segment around MP 9 going eastbound and the segments from MP 11 to MP 17 going westbound have moderate congestion.

Mainline congestion conditions on SR 408 in December 2014 can be summarized based on TTI. However, it should be noticed that the congestion intensity changes with time. When it is approaching peak hours, the congestion intensity gradually increases. Once the peak time has passed, the congestion becomes less severe. The congested area for SR 408 is approximately from MP 09 to MP 10 and MP 17 to MP 18 going eastbound and from MP 10 to MP 17 going westbound. It can be seen that the congested area on westbound lanes is much longer than that going eastbound. In addition, the congestion intensity going westbound is higher. The reason might be that westbound lanes experience congestion during morning peak hours, whereas eastbound experiences congestion during evening peak hours. SR 408 is an urban expressway that carries a large amount of commuters. As a result, the departure time from home to work in the morning is less flexible and is concentrated during 8:00am to 9:00am. In contrast, in the evening, the departure time from work to home is more flexible and therefore congestion is relieved to some extent. The westbound congested segments in the morning are toward downtown Orlando, where land use is chiefly for offices and businesses. In the evening, congestion is concentrated in the opposite direction at interchanges with two other major corridors: Interstate-4 and SR 417.

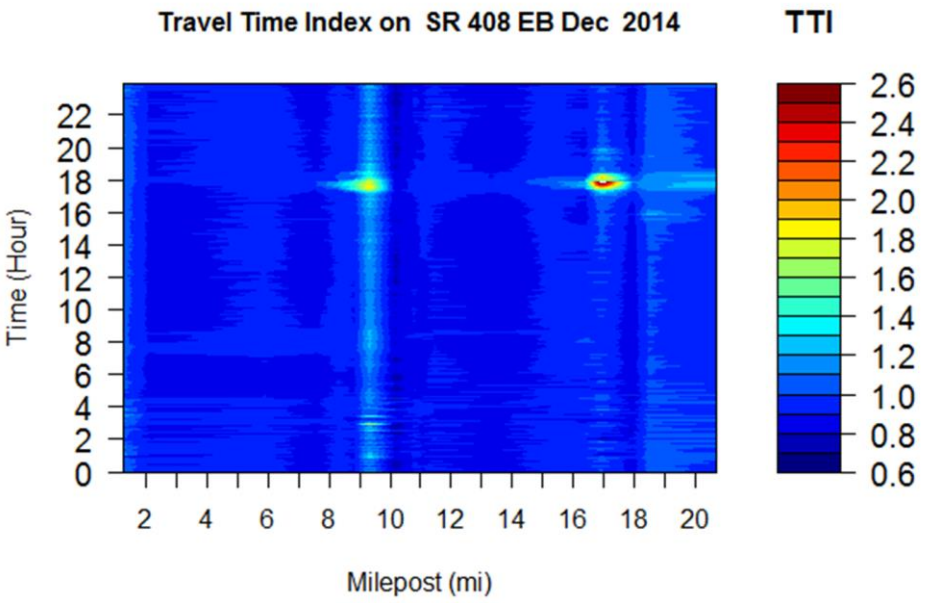


Figure 4.1 - Mainline weekday travel time index of SR 408 eastbound, December 2014

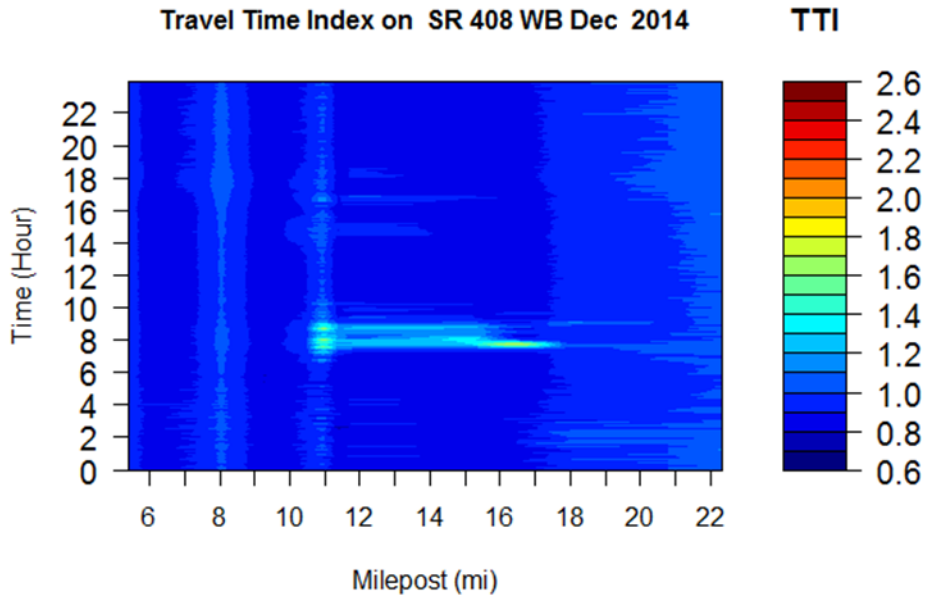


Figure 4.2 - Mainline weekday travel time index of SR 408 westbound, December 2014

4.3 Travel Time Reliability

Most travelers in metropolitan areas expect peak hour congestion. However, unexpected delays can cause severe consequences (Texas Transportation Institute, 2006). To reduce unexpected delays, improvement of travel time reliability is necessary. Travel time reliability is defined as the consistency or dependability in travel times, as measured from day-to-day and/or across different times of the day (Texas Transportation Institute, 2006).

Travel time can be measured directly or estimated indirectly. Travel time measurement can be based on floating car technique, GPS, or Bluetooth. However, the performance of in-vehicle-based travel time measurements is restricted by the sample size of probed vehicles. AVI technology, which is often integrated with electronic tolling, is another method to provide direct measurement of travel time. Although AVI can only detect vehicles equipped with tags, with the current relatively high adoption rate, a sufficient sample size can be captured by the system. Estimation of travel time varies from simple estimation, based on posted speed limit, to more complicated algorithms using point-based detector data (i.e., loop detectors, radar detectors) (Martchouk, 2009).

Various approaches have been proposed to measure travel time reliability. They range from statistical range measures and buffer measures to tardy trip indicators (Martchouk, 2009). Statistical measures include travel window and percent variation (see Equations 4.2 and 4.3).

$$\text{Travel Window} = \text{Average Travel Time} \pm \text{Standard Deviation} \tag{4.2}$$

$$\text{Percent Variation} = \frac{\text{Standard Deviation}}{\text{Average Travel Time}} \times 100\% \tag{4.3}$$

However, travel time window and percent variation have some drawbacks, including understandability for a nontechnical audience. In addition, they treat early and late arrivals with equal weight, despite the fact that the public cares much more about late arrivals (Texas Transportation Institute, 2006).

Another category of approach is buffer measures, which are an assessment of how much extra time needs to be allowed for uncertainty in travel conditions (Lomax et al., 2003). Buffer time is the difference between the 95th percentile travel time and the average travel time (see Equation 4.4). An interpretation of the 95th percentile is that in one month, a traveler could be late to work for one day out of that month.

$$\text{Buffer Time} = \text{95th Percentile Travel Time} - \text{Average Travel Time} \quad (4.4)$$

Buffer index uses the buffer time concept. Currently two expressions about buffer index exist: Equation 4.5, which is used by Martchouk (2009), and Equation 4.6, in which travel rate (in minutes per mile) rather than travel time is used to identify an average trip (Lomax et al., 2003).

$$\text{Buffer Index} = \left[\frac{\text{95th Percenteile Travel Time} - \text{Average Travel Time}}{\text{Average Travel Time}}\right] \times 100\% \quad (4.5)$$

$$\text{Buffer Index} = \begin{pmatrix}\text{Average of}\\ \text{All Sections}\\ \text{Using VMT to}\\ \text{Weight the Section}\end{pmatrix} \left[\frac{\text{95th Percenteile Travel Rate} - \text{Average Travel Rate}}{\text{Average Travel Rate}}\right] \times 100\% \quad (4.6)$$

An extension of buffer index is planning time index, which is the 95th percentile travel time index. It could be used as a trip-planning factor for trips that require on-time arrivals (Lomax et al., 2003).

$$\text{Planning Time Index} = \frac{\text{95th Percentile Travel Time}}{\text{Travel Time Based on Free-Flow Speed}} \quad (4.7)$$

Another approach, tardy trip indicators, implies the amount of late trips. On-time arrival rate (Equation 4.8) and misery index (Equation 4.9) can be used as tardy trip indicators.

$$\text{On-Time Arrival} = 100\% - \left(\begin{matrix}\text{Percent of Travel Time Greater than} \\ \text{110\% of the Average Travel Time}\end{matrix}\right) \quad (4.8)$$

Here, as suggested by (Schrank, 2001), the 10% of average travel time is used as a "lateness threshold." However, this 10% late arrival may be relatively conservative for some applications (Lomax et al., 2003).

$$\text{Misery Index} = \frac{\begin{matrix}\text{Average of the Travel Time for} \\ \text{the Longest 20\% of the Trips}\end{matrix} - \text{Average Travel Time for All Trips}}{\text{Average Travel Time for All Trips}} \quad (4.9)$$

The misery index compares the worst 20% of trips against the normal conditions to show the negative aspect of trip reliability. Another interpretation of misery index is "how bad are the worst days?"

In this project, the AVI system measured travel time directly for different segments on the expressways. Thus, there was no need for travel time estimation from point detectors. Using the travel time collected by the AVI system, travel time reliability was evaluated using various measures discussed above, including percent variation, buffer index, planning time index, and misery index. In addition, travel time reliability was

investigated at different time intervals, such as weekday weekend, day of the week, different time of day, and 5-minute intervals.

Figure 4.3 and Figure 4.4 show that during weekdays, the percent variation is significantly higher than that of weekends on the segments proven to experience congestion in previous sections. On segments without congestion, the percent variations for weekday and weekend are comparable. The figures imply that travel time reliability is reduced during weekdays. A potential cause of the finding might be the high traffic demand and congestion on weekdays.

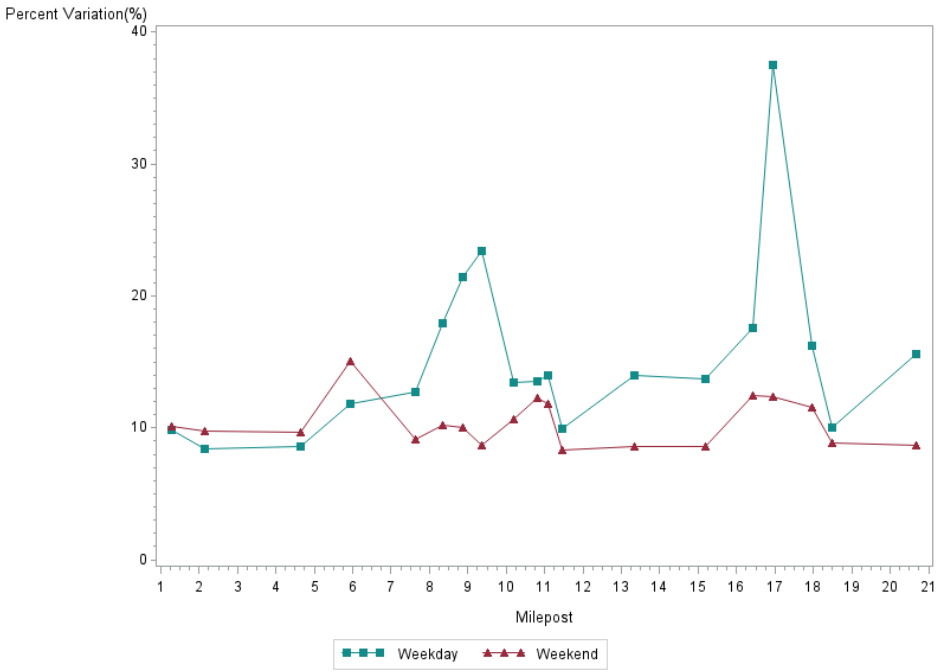


Figure 4.3 - Weekday and weekend percent variation on SR 408 eastbound, July 2014

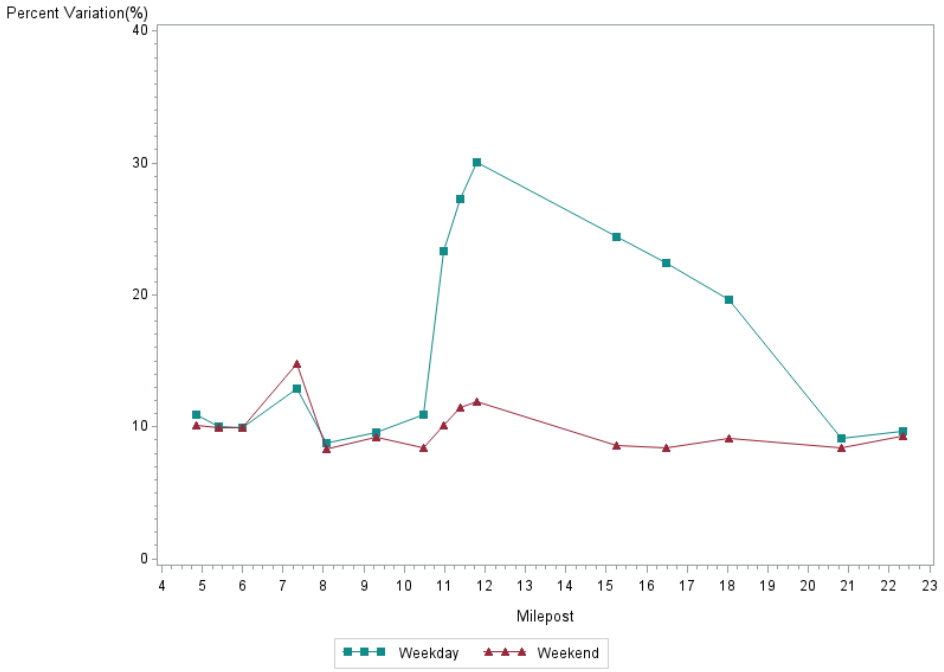


Figure 4.4 - Weekday and weekend percent variation on SR 408 westbound, July 2014

Figure 4.5 and Figure 4.6 illustrate the percent variation by day of the week. Conclusions are consistent with the percent variation by weekday and weekend. One finding is that Friday has a much lower percent variation compared to Monday through Thursday. The difference between Friday and other weekdays could be caused by a variety of reasons. One possible explanation is that the traffic pattern on Friday is more comparable to weekends. Although Saturday and Sunday have the lowest percent variation (as expected based on weekend findings), the profiles of the percent variation of the two days is similar to that of Friday. In the future, when evaluating travel time reliability on different days, Friday might be considered with Saturday and Sunday. Regardless, whether this assumption is valid also depends on conclusions from other reliability measures.

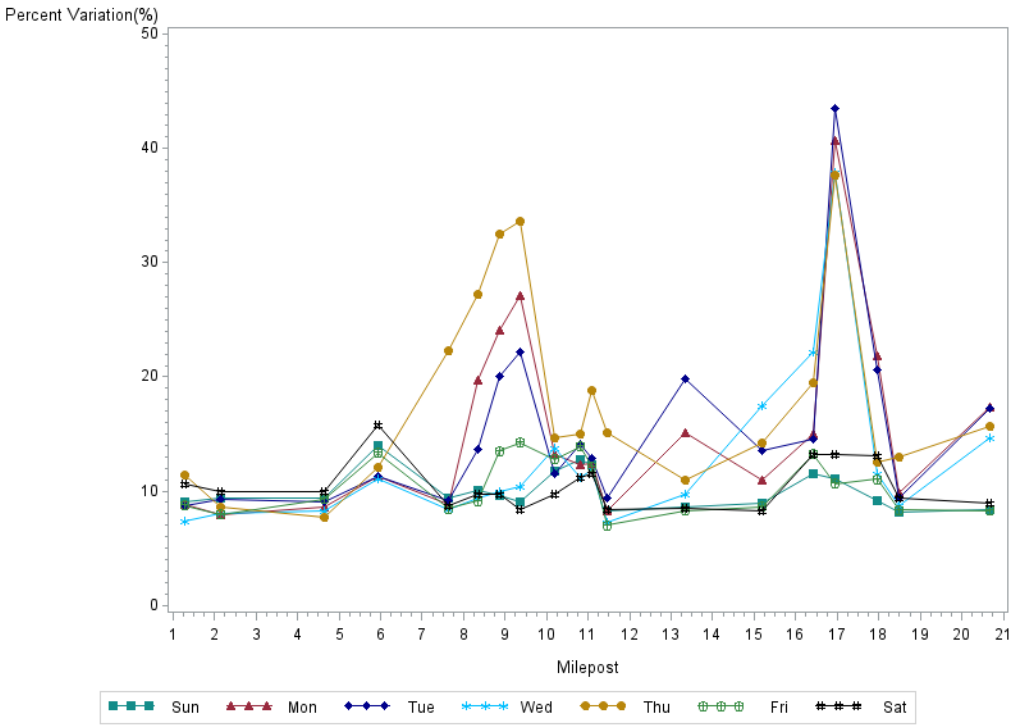


Figure 4.5 - Day of week percent variation on SR 408 eastbound, July 2014

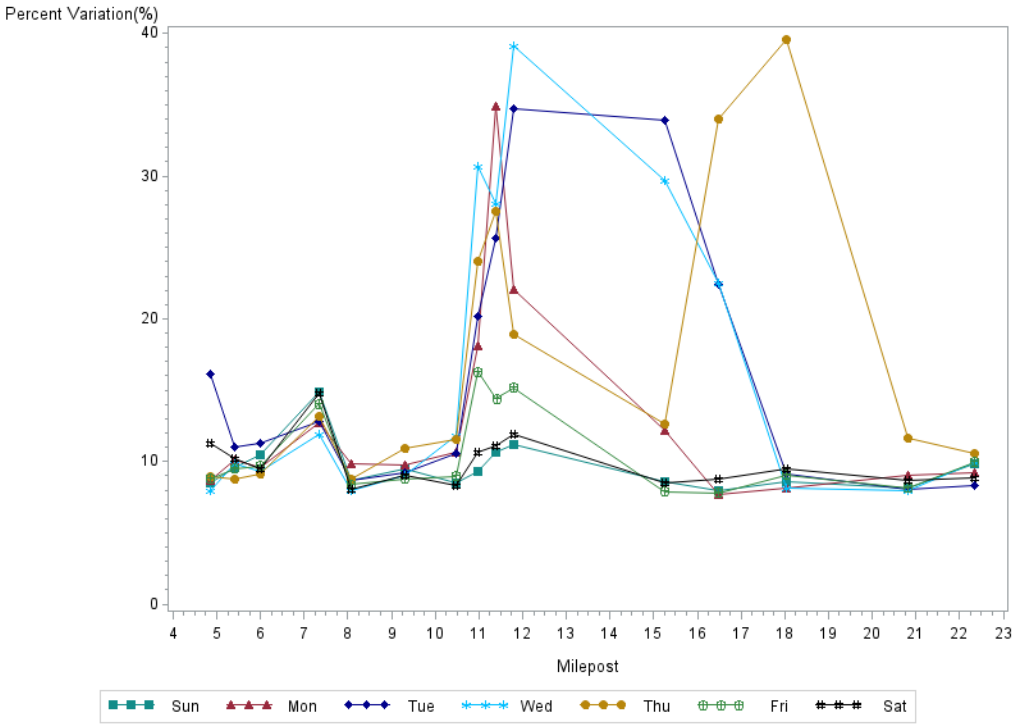


Figure 4.6 - Day of week percent variation on SR 408 westbound, July 2014

The percent variation by different time of day implies that peak hours have significantly lower travel time reliability (higher percent variation). In addition, the figures show clearly that not both morning and evening peak hours will experience deterioration on the same traveling bound, affected by the traffic demand at different times of day. In Figure 4.7 and Figure 4.8, SR 408 eastbound has high values of percent variation during evening peak hours near MP 9.0 and MP 18, while westbound has higher percent variation from MP 11 to MP 19 during the morning peak hours. This is generally in agreement with the findings from congestion evaluation.

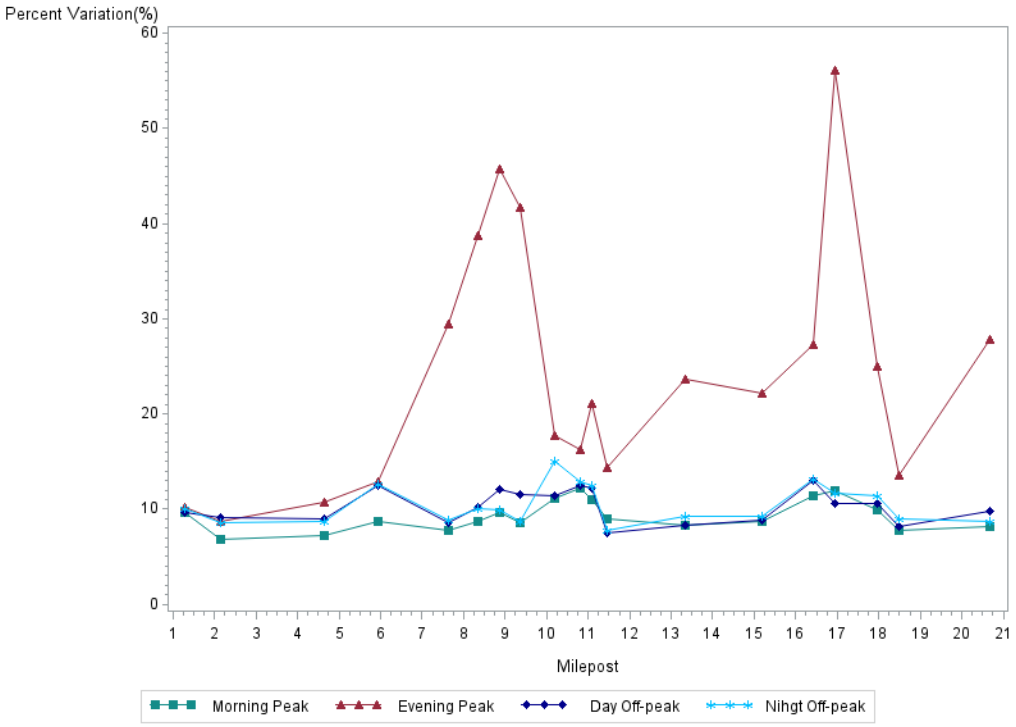


Figure 4.7 - Time of day percent variation on SR 408 eastbound, July 2014

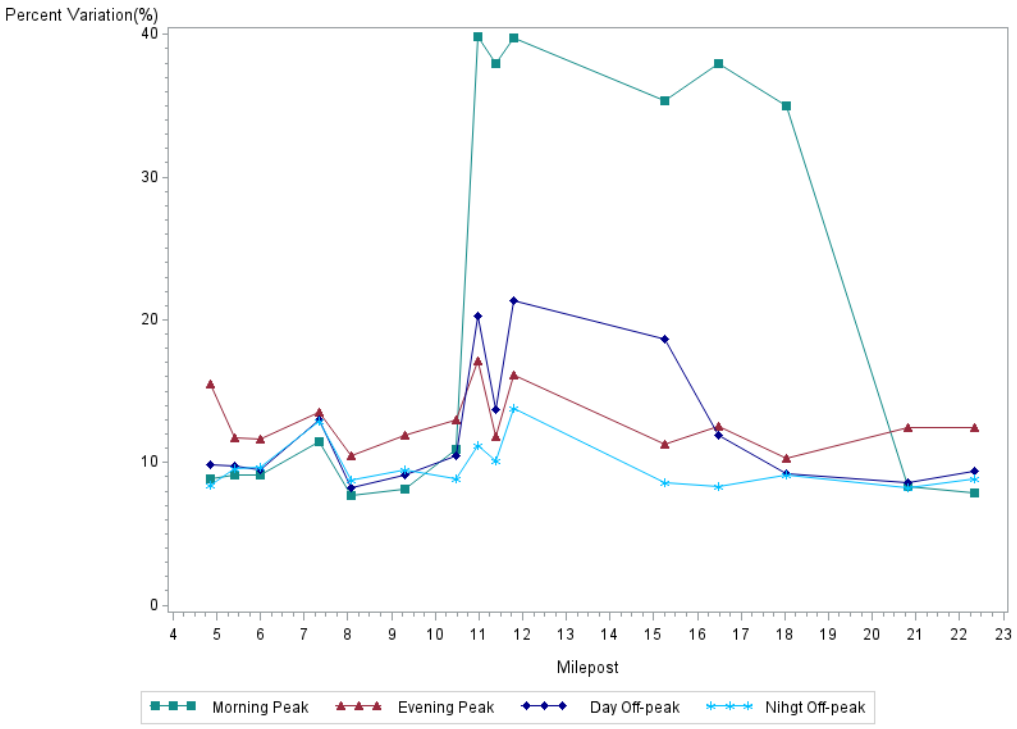


Figure 4.8 - Time of day percent variation on SR 408 westbound, July 2014

The filled contour plots in Figure 4.9 and Figure 4.10 are highly comparable to the plots of congestion on SR 408 (Figure 4.1 and Figure 4.2). The distribution of percent variation for weekdays in spatial-temporal dimension evaluated at 5-minute intervals reveals the travel time reliability at a more detailed level. Although percent variation does not treat late and early arrival differently, it is straightforward and easy for calculation because only average speed and its standard deviation are needed. Consequently, for evaluation of real-time travel time reliability, percent variation serves as a possible candidate.

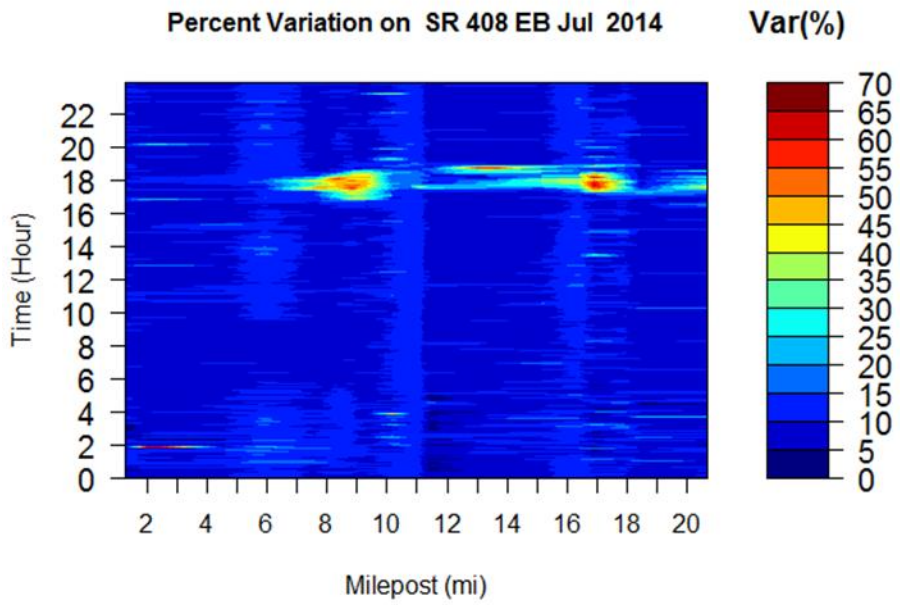


Figure 4.9 - Weekday percent variation evaluated at 5-min intervals on SR 408 eastbound, July 2014

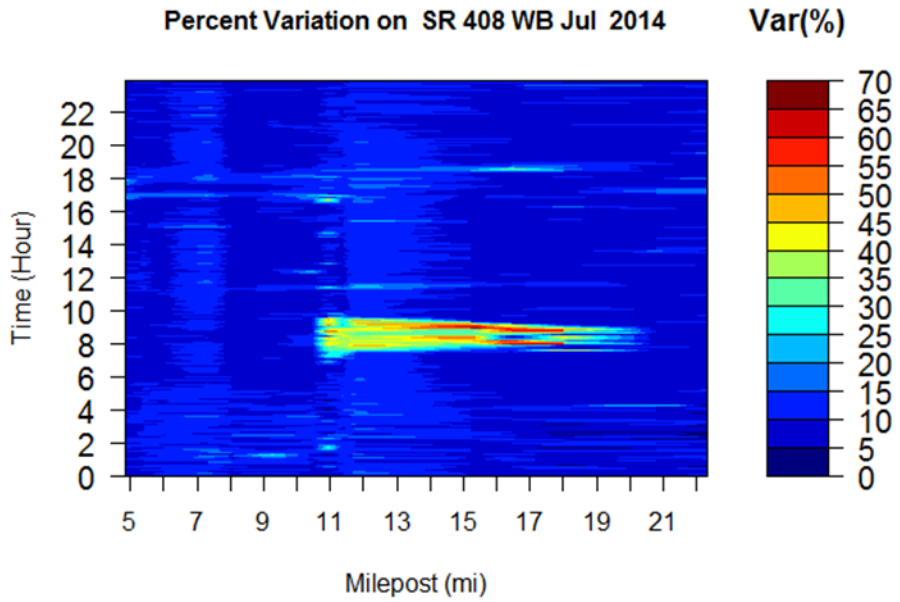


Figure 4.10 - Weekday percent variation evaluated at 5-min intervals on SR 408 westbound, July 2014

Travel time reliability was also evaluated using buffer measures (buffer index and planning time index). Results are shown in Figure 4.11 through Figure 4.20. Conclusions are similar to those of percent variation. While buffer index and planning time index both incorporate 95th percentile travel time, they are different in that buffer index uses average travel time and planning time index uses free-flow travel time. For real-time evaluation, the buffer index and planning time index have one disadvantage – they are based on 95th percentile travel time, which would be difficult to determine in real-time. The difficulty is that most traffic authorities only archive traffic data aggregated at 1-minute or other time intervals. However, to get the 95th percentile travel time, traffic time for individual vehicles needs to be archived and processed in real-time. In this regard, traffic detection at an individual vehicle level is recommended for real-time estimation of travel time reliability.

The tardy trip indicator misery index (Figure 4.21 – Figure 4.28) performs similarly to statistical range measures and buffer measures. To be used for travel time reliability in real-time, this measurement requires travel time information of individual vehicles to get the longest 20% of trips.

In conclusion, the AVI system could provide data to evaluate travel time reliability. In this study, the AVI data used were at an individual vehicle level. However, most agencies do not currently archive and use AVI data at this level. In most cases, for travel time estimation, the AVI data are aggregated at 1-minute intervals, meaning there is only one reading per minute. If travel time information for each vehicle equipped with tags could be collected in practice, real-time calculation of 95th percentile speed and the longest 20% of trips is made possible, and therefore the real-time estimation of travel time reliability.

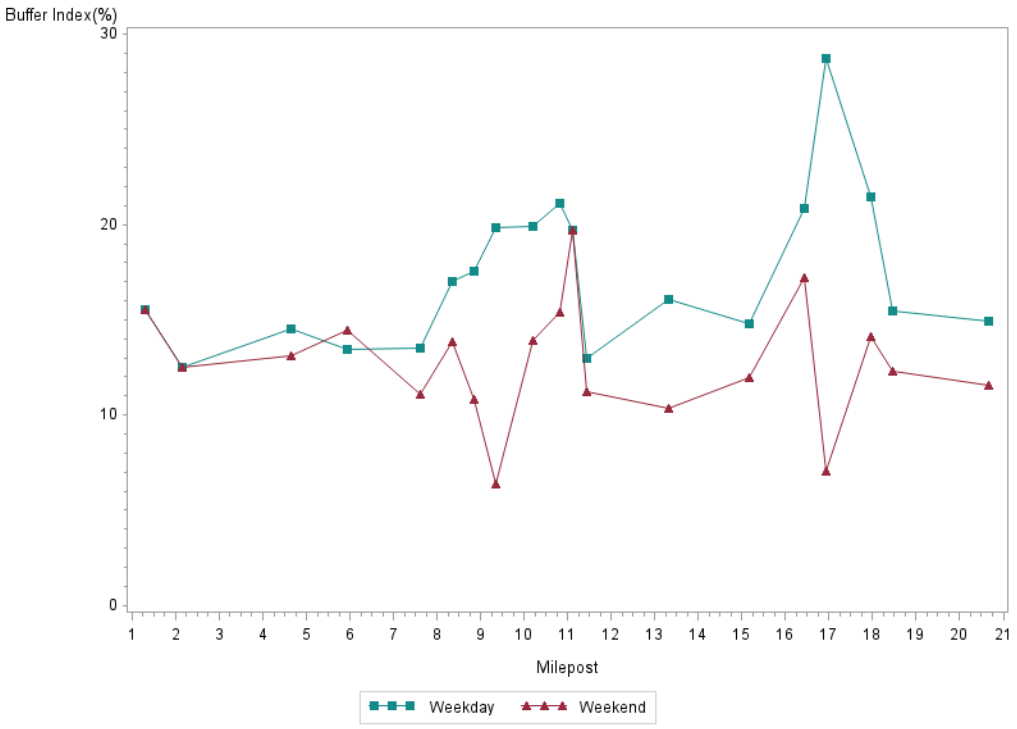


Figure 4.11 - Weekday and weekend buffer index on SR 408 eastbound, July 2014

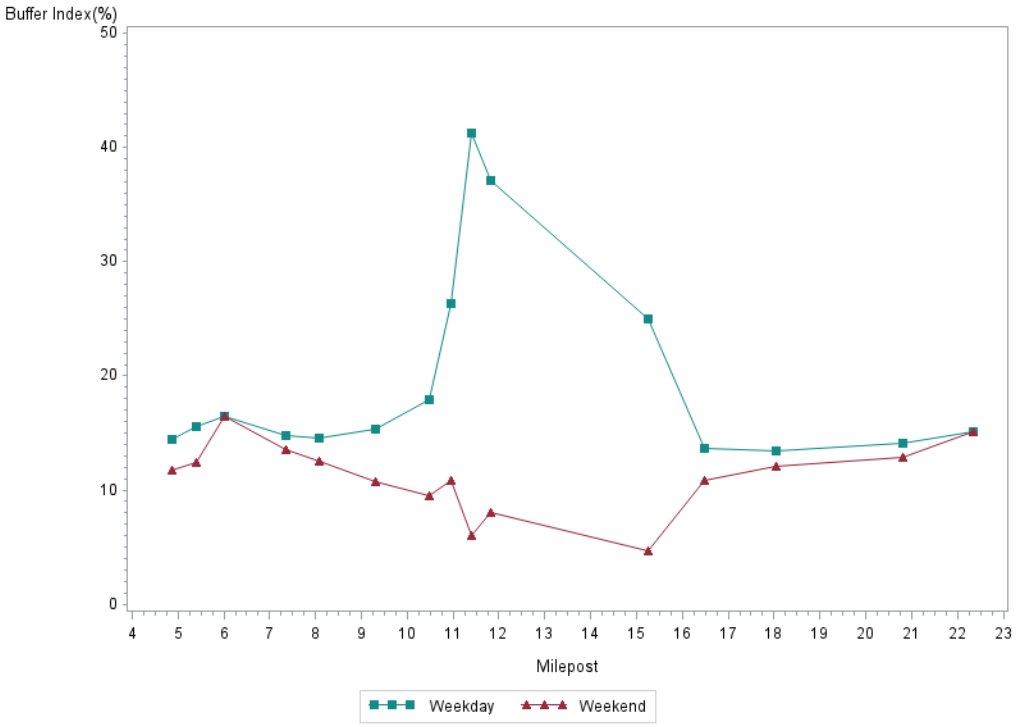


Figure 4.12 - Weekday and weekend buffer index on SR 408 westbound, July 2014

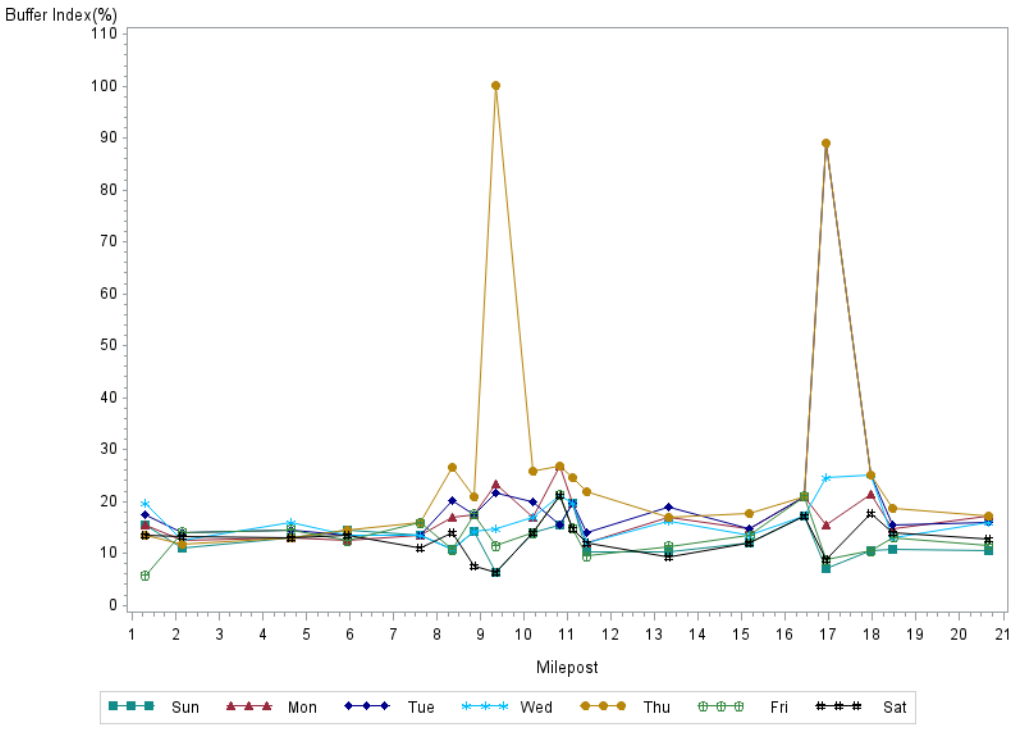


Figure 4.13 - Day of week buffer index on SR 408 eastbound, July 2014

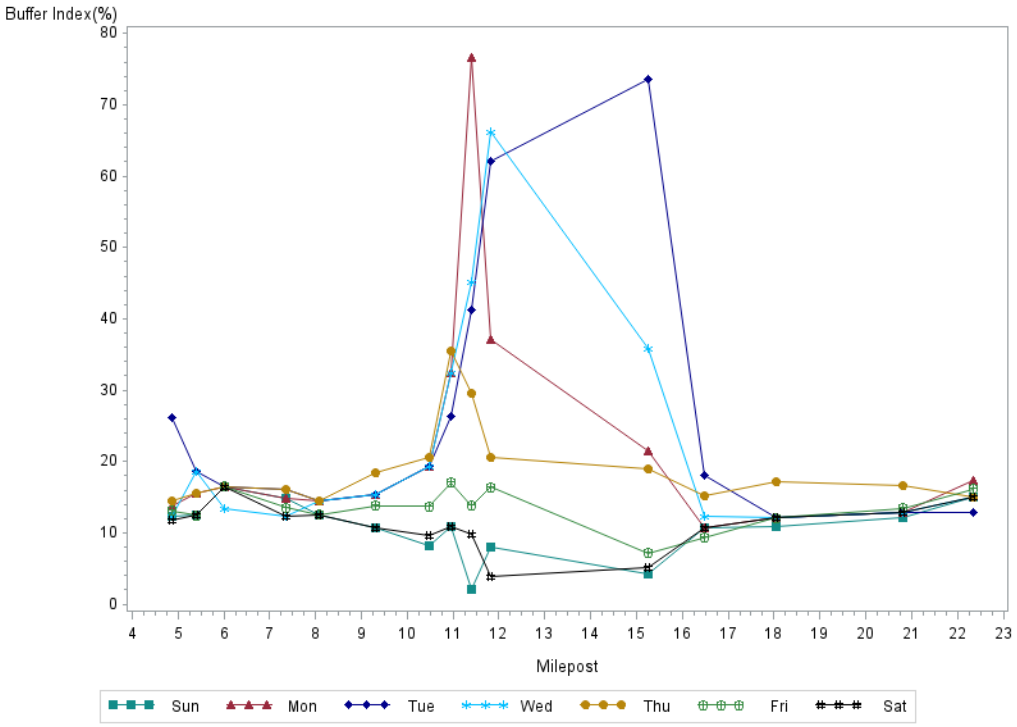


Figure 4.14 - Day of week buffer index on SR 408 westbound, July 2014

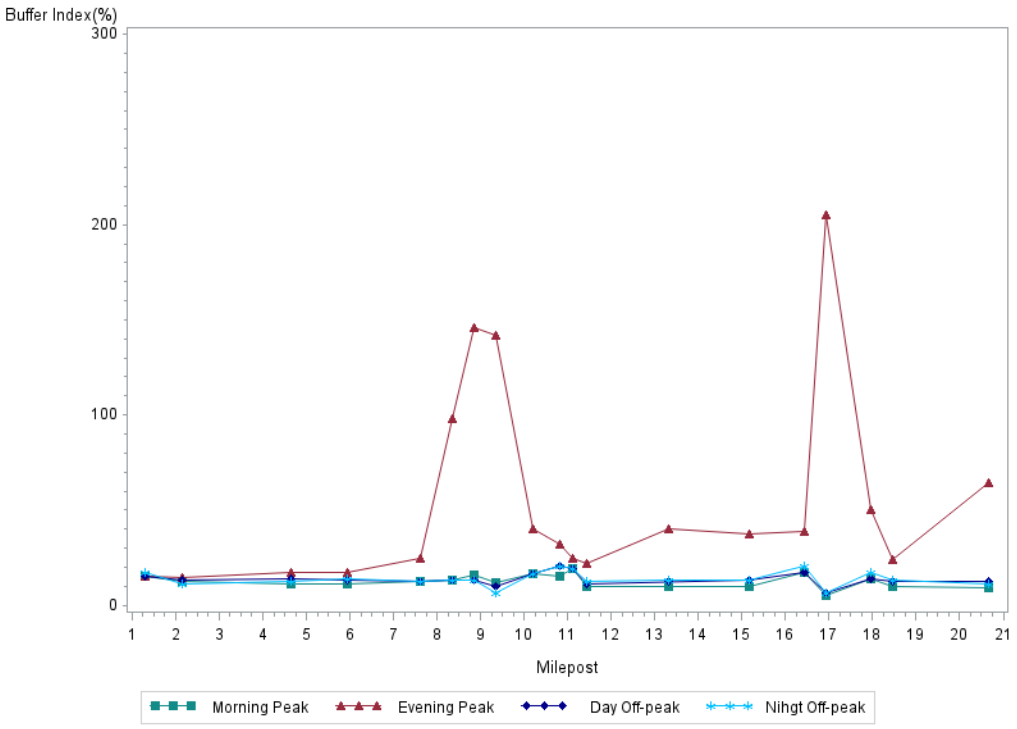


Figure 4.15 - Time of day buffer index on SR 408 eastbound, July 2014

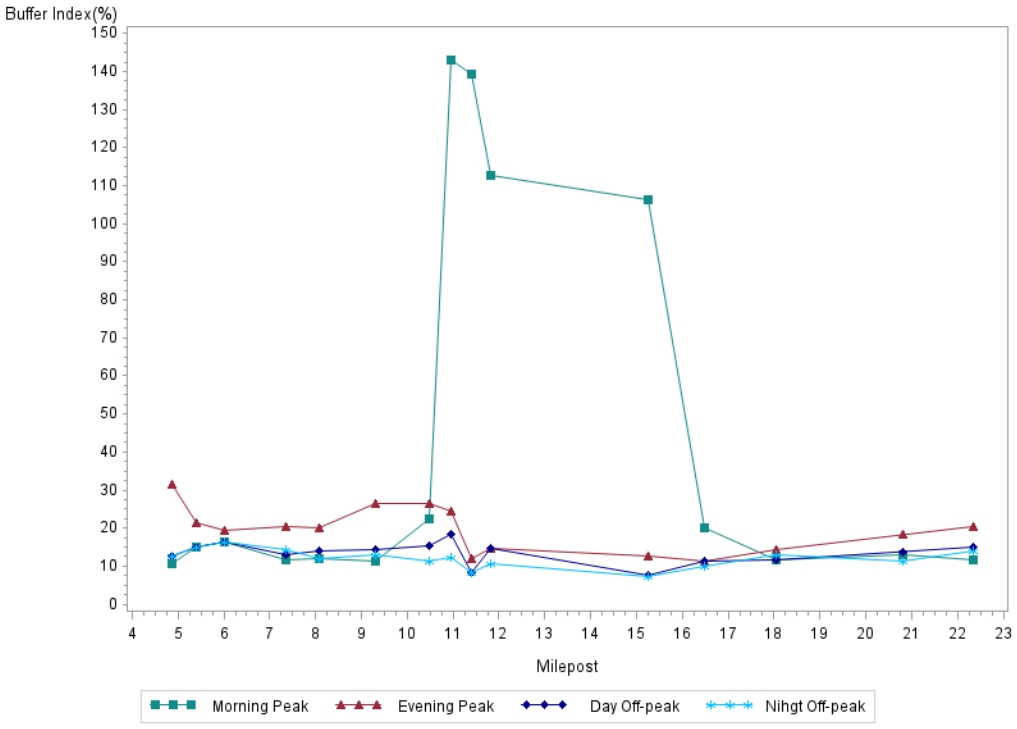


Figure 4.16 - Time of day buffer index on SR 408 westbound, July 2014

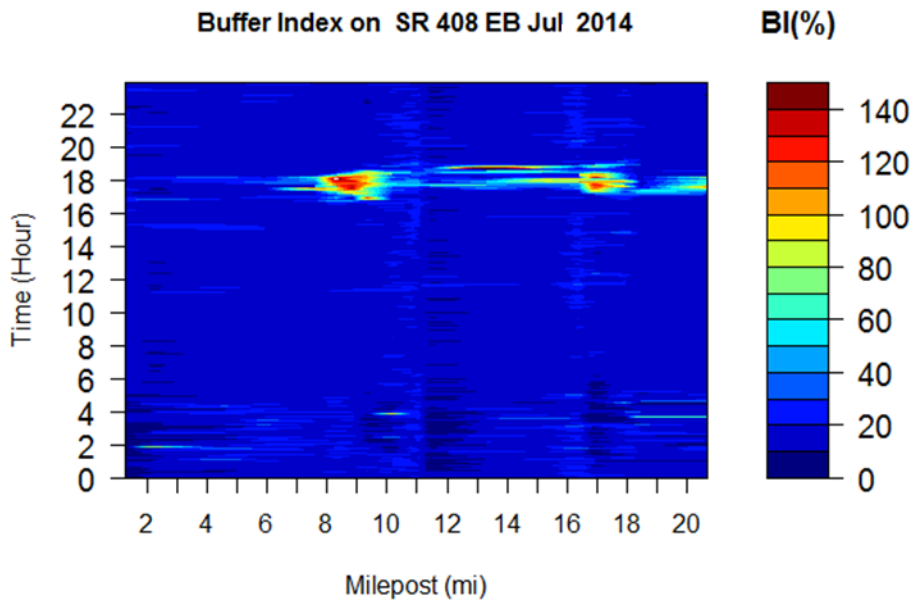


Figure 4.17 - Weekday buffer index evaluated at 5-min intervals on SR 408 eastbound, July 2014

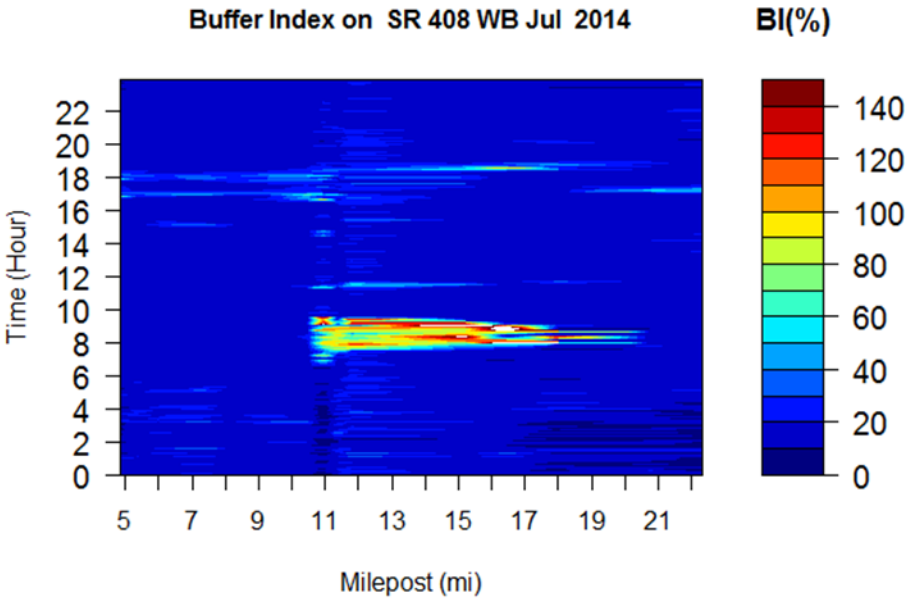


Figure 4.18 - Weekday buffer index evaluated at 5-min intervals on SR 408 westbound, July 2014

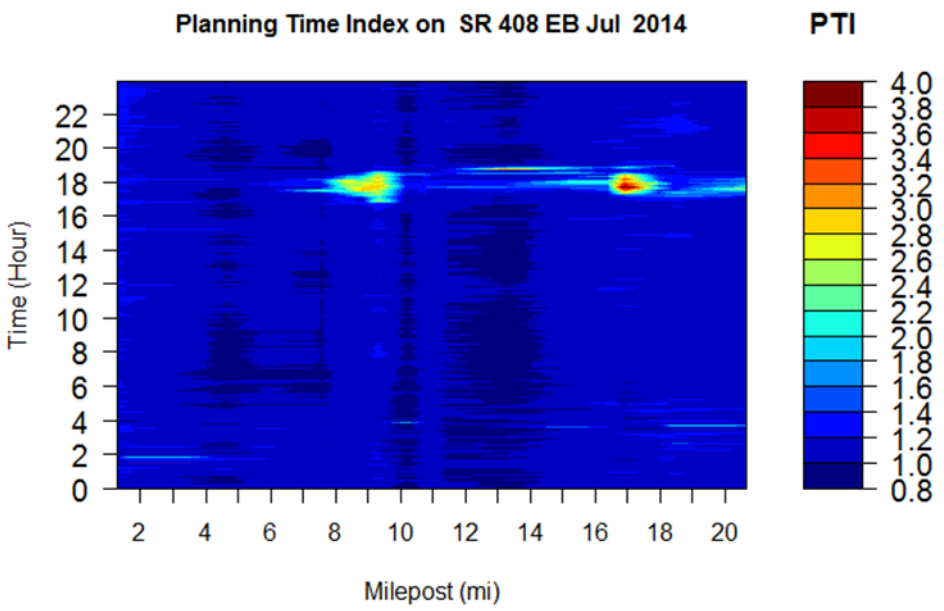


Figure 4.19 - Weekday planning time index evaluated at 5-min intervals on SR 408 eastbound, July 2014

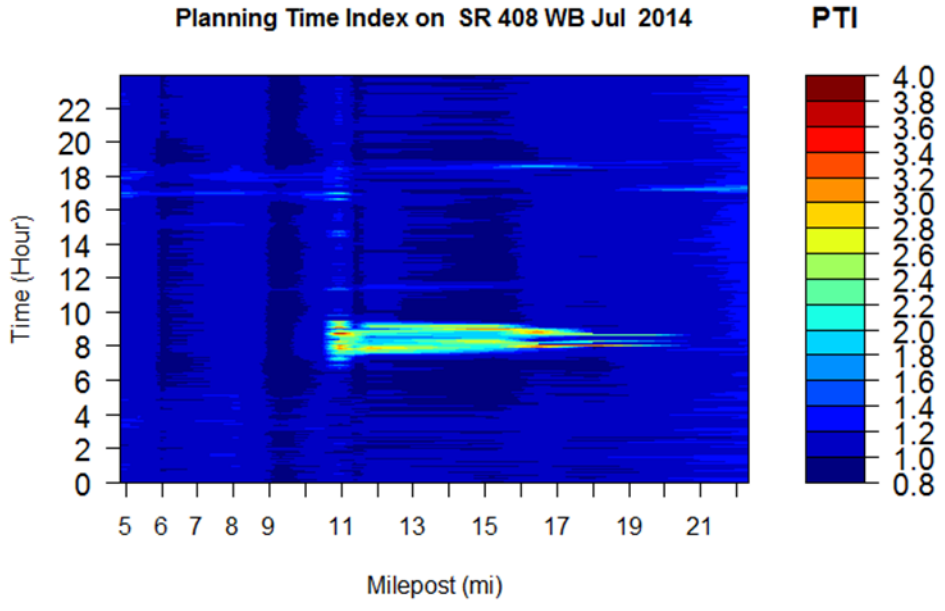


Figure 4.20 - Weekday planning time index evaluated at 5-min intervals on SR 408 westbound, July 2014

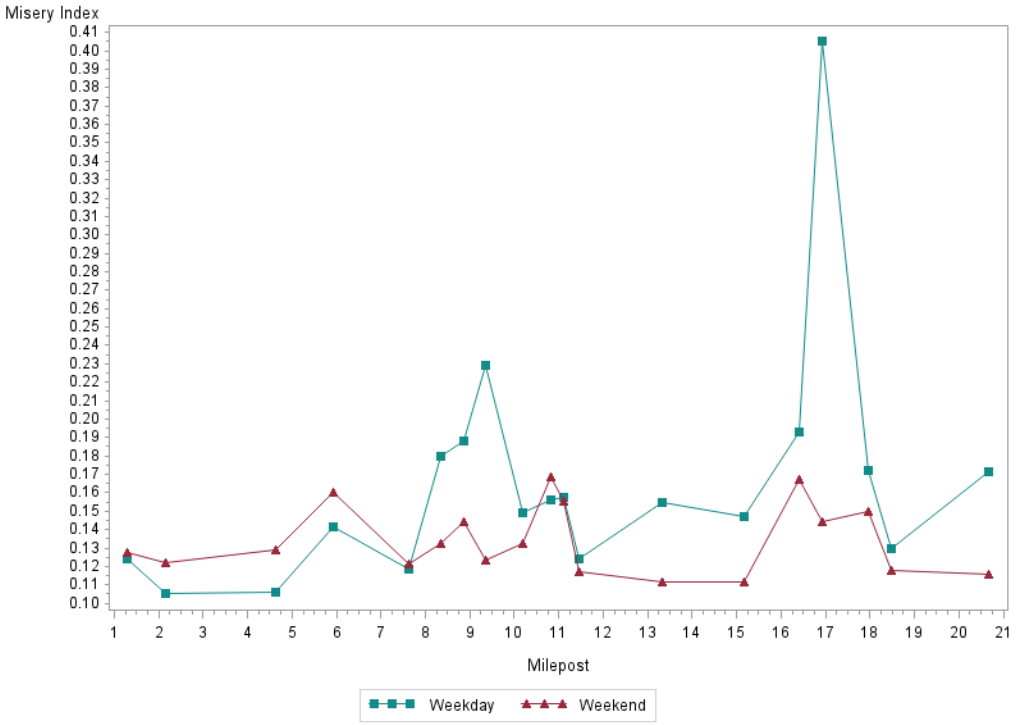


Figure 4.21 - Weekday and weekend misery index on SR 408 eastbound, July 2014

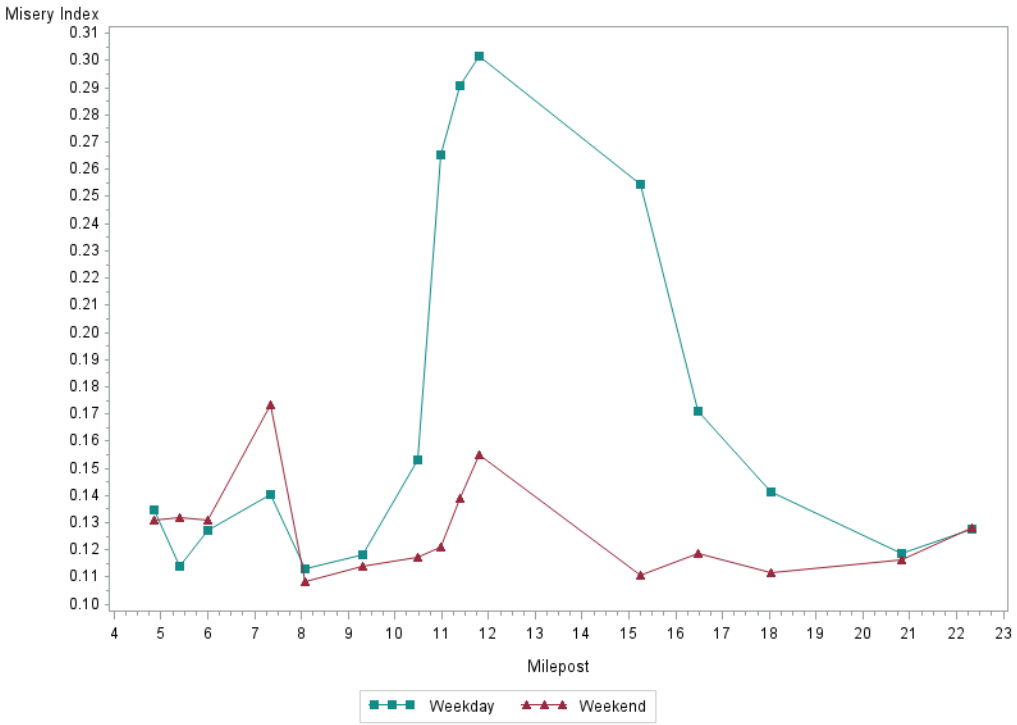


Figure 4.22 - Weekday and weekend misery index on SR 408 westbound, July 2014

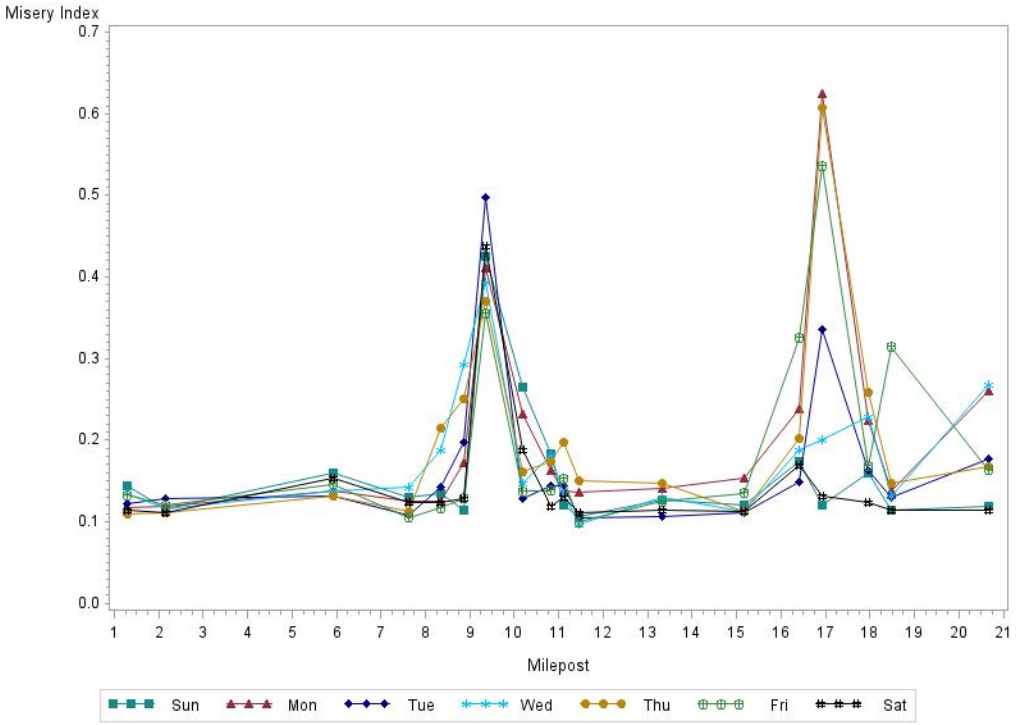


Figure 4.23 - Day of week misery index on SR 408 eastbound, July 2014

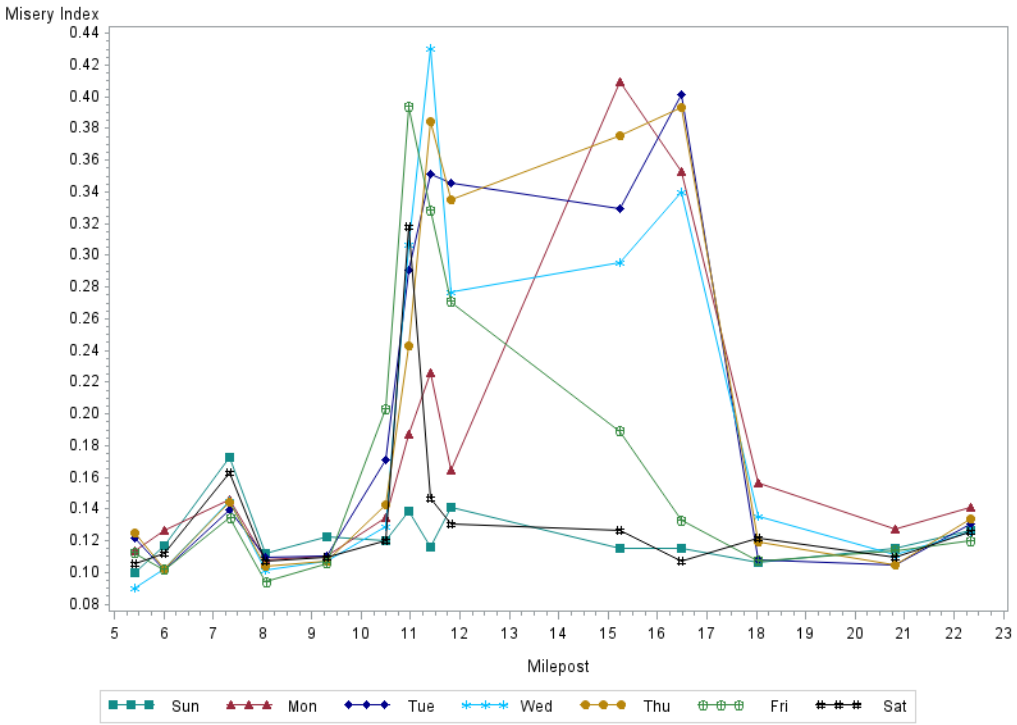


Figure 4.24 - Day of week buffer index on SR 408 westbound, July 2014

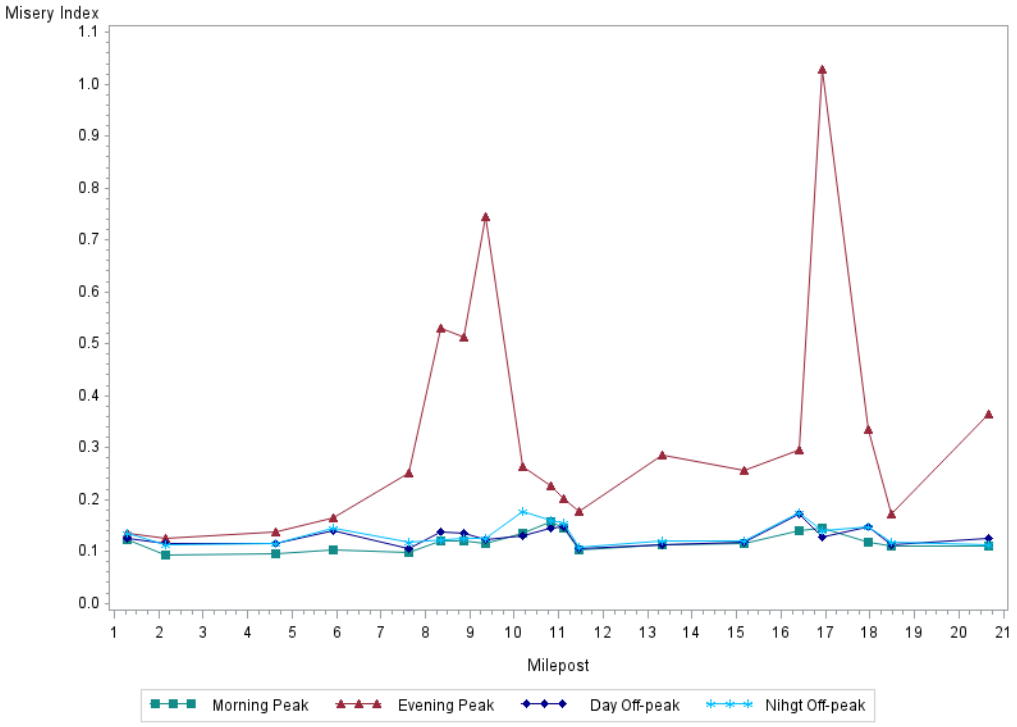


Figure 4.25 - Time of day misery index on SR 408 eastbound, July 2014

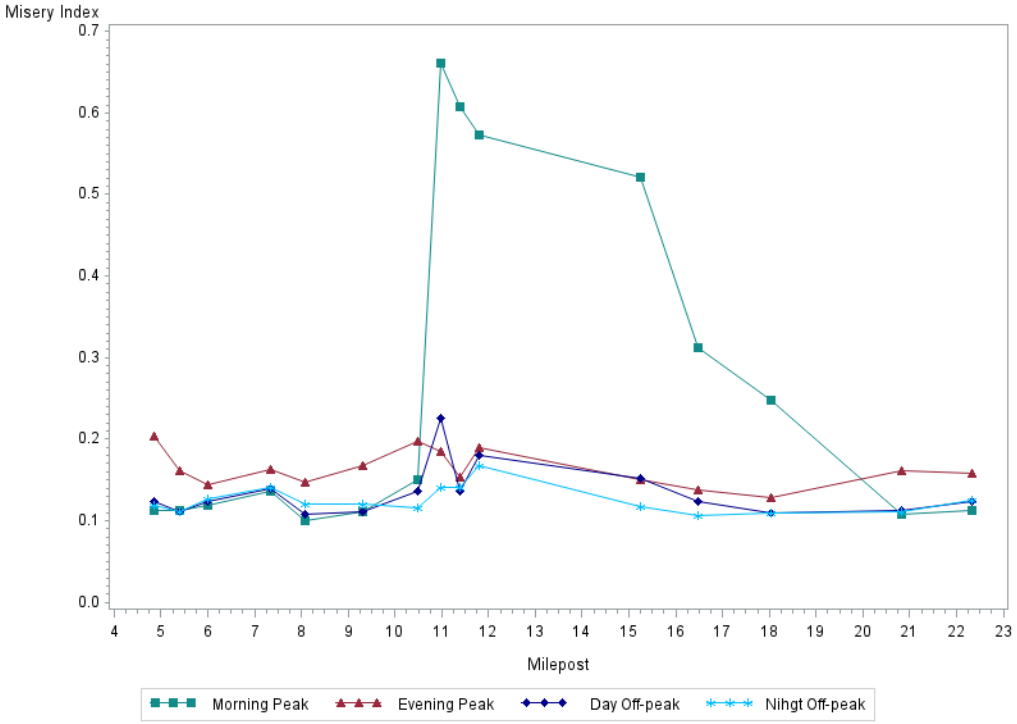


Figure 4.26 - Time of day misery index on SR 408 eastbound, July 2014

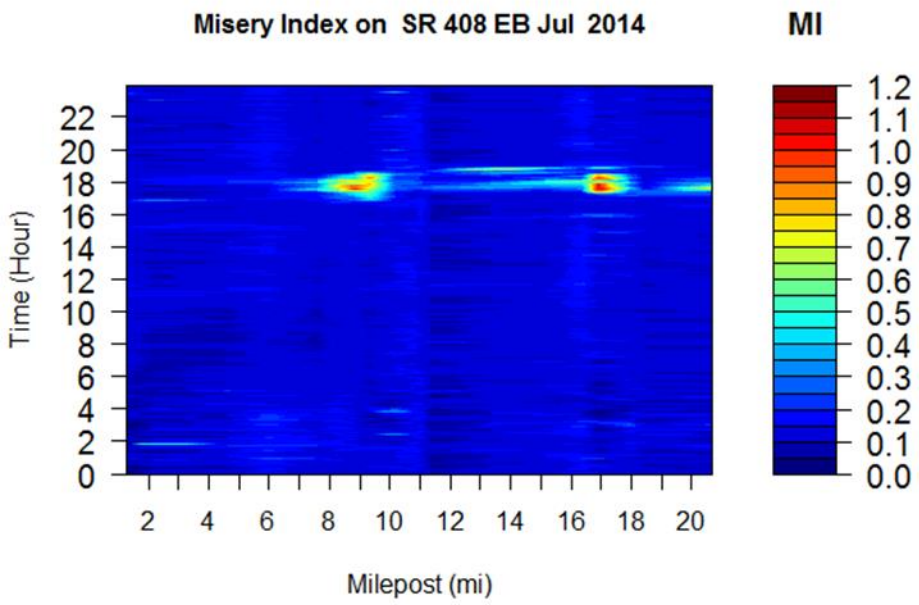


Figure 4.27 - Weekday misery index evaluated at 5-min intervals on SR 408 eastbound, July 2014

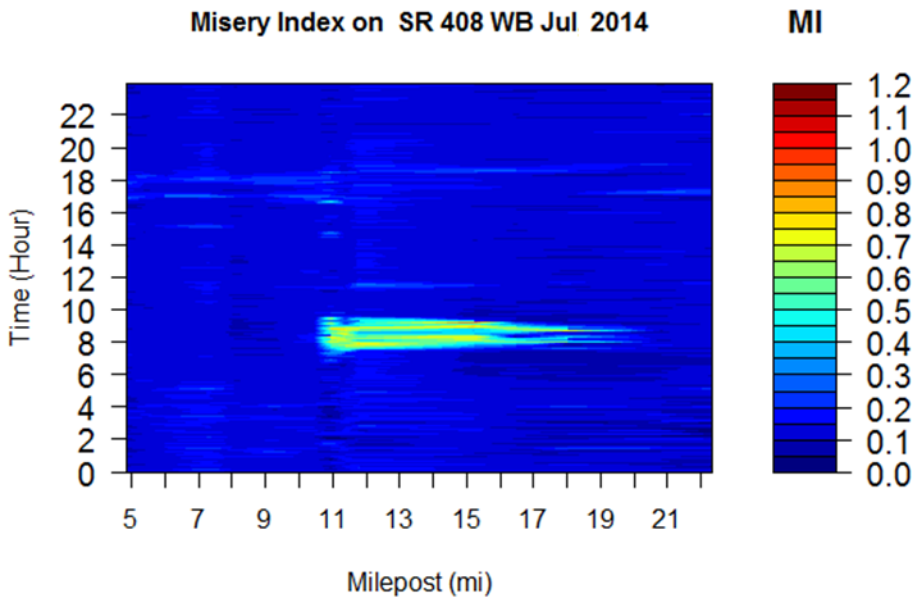


Figure 4.28 - Weekday misery index evaluated at 5-min intervals on SR 408 westbound, July 2014

CHAPTER 5: Traffic Safety Evaluation

In addition to traffic mobility, traffic safety is another crucial indicator of freeway/expressway performance. In recent decades, with the access to ITS traffic data, real-time safety evaluations have been widely accepted in addition to the traditional crash frequency studies to identify the crash precursors. Compared with traditional crash frequency studies, real-time safety evaluations treat individual crashes as study objects. By revealing the crash contributing factors, there could be possibilities for more proactive traffic management strategies.

In this project, real-time safety evaluations were carried out based on the expressway data. The evaluations were conducted for total crashes on mainlines and ramps, respectively. For each mainline crash case, two upstream and two downstream MVDS detectors nearest to the crash were selected to represent the traffic conditions nearest the crash location, as shown in Figure 5.1. For each crash case, traffic data from these detectors were extracted from 5-10 minutes prior to the crash occurrence to represent the traffic conditions immediately prior to the crashes. The geometric characteristics at the crash locations were also included in the analysis.

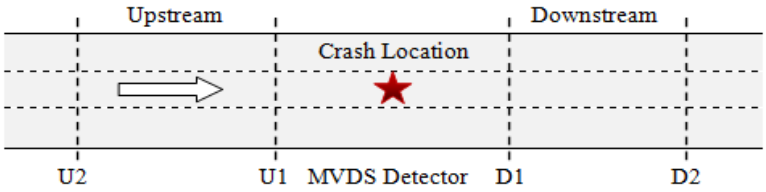


Figure 5.1 - Assignment of MVDS detectors to crash location

Crash data were collected from the 18-month period of July 2013 to December 2014. There were 1553 crashes identified on the three expressways in this period. However, since the MVDS underwent a system upgrade in April 2014, no traffic data

were available during that month. Consequently, the crash data were matched with 17-month traffic data and geometric data. After matching the data, SR 408 had the most crashes matched (699 crashes), followed by SR 528 (361 crashes), and SR 417 (334 crashes). For each matched crash case, ten non-crash cases were randomly selected from where no crash occurred.

Two general analysis approaches are employed in real-time traffic safety studies: statistical methods and data mining based methods. Statistical methods include logistic regression (Hourdos et al., 2006), matched case-control logistic regression (Abdel-Aty et al., 2004; Abdel-Aty et al., 2005; Zheng et al., 2010), and Bayesian statistics (Abdel-Aty et al., 2012; Ahmed et al., 2012; Yu and Abdel-Aty, 2013). Data mining based methods include neural networks (Pande and Abdel-Aty, 2006), classification trees (Pande and Abdel-Aty, 2006), random forests (Pande et al., 2011; Ahmed and Abdel-Aty, 2012; Hossain and Muromachi, 2012) and support vector machines (Yu et al., 2013), among others. Statistical methods present the effects of variables in an interpretable way. However, they assume a linear relationship between the dependent and independent variables. In contrast, while data mining methods often make predictions with very high accuracy, they are questioned because the analyzing process is similar to a black box.

A simple logistic regression model was used in this study to identify the contributing factors for different types of crashes. The log odds of the outcome is modeled as a linear combination of the predictor variables. The dichotomous responses, crash and non-crash, were converted into the probabilities p(y=1) and 1-p, respectively. The logistic regression model is found in Equation 5.1,

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_j x_j \quad (5.1)$$

where β_0 is the intercept, β_j is the coefficient of variable, and x_j is the value of explanatory variable (e.g., speed, truck percentage, road surface conditions).

5.1 Real-Time Safety Evaluation for Total Mainline Crashes

For total crashes, several variables were found to significantly affect crash occurrence, as shown in Table 5.1. During peak hours, traffic demand was higher than off-peak hours. A large number of motorists traveling at the same time could increase crash likelihood considerably. Traffic flow parameters that contributed to crash occurrence included higher traffic volume per lane and greater speed differences between inner and outer lanes on a cross-section at U1 station (shown in Figure 5.1) and a higher truck percentage and congestion index at D1 station (shown in Figure 5.1). The higher volume per lane and congestion index implied more congested traffic conditions. Speed differences between lanes indicated the speed variation on a segment. In addition, the presence of trucks could cause higher risks for motorists. The effects of three geometric variables were significant as well. On the studied expressways, a cross-section contained two to five lanes. Given the proportion of five lane segments within the cross-section was quite small, the effect on safety was studied together with four lane segments in the safety evaluation. If a cross-section had more than three lanes, there was a higher probability of a crash on the segment. A narrower shoulder and median width would also increase the probability of crash occurrence. The effects are straightforward for interpretation. With a wider shoulder or median width, a driver has more leeway to adjust their behavior and avoid a crash.

Table 5.1 - Parameter estimates and model fitting for total crashes

Parameter	Description	Estimate	Standard Error	Wald Chi-Square	*p*-value
Intercept		-3.1420	0.1318	567.9733	<.0001
Peak	Peak hour indicator: Peak = 1: weekday 7:00 - 9:00 & 17:00 – 19:00, Peak = 0: otherwise	0.1659	0.0888	3.4933	0.0616
U1_lanevol	Average traffic volume by lane at U1 station	0.0130	0.000891	212.6196	<.0001
U1_spddiff	Speed difference between the inner and outer lanes at U1 station	0.0228	0.00598	14.5063	0.0001
D1_trkpct	Truck percentage at D1 station	1.2891	0.2388	29.1463	<.0001
D1_ci	Congestion index at D1 station	4.6351	0.3374	188.7165	<.0001
Lane45	Number of lanes on cross section per direction: Lane45 = 1: 4 or 5 lanes at detection location, Lane45 = 0: otherwise Base level: 2 lanes	0.3196	0.0906	12.4456	0.0004
Median	Median width (ft)	-0.00505	0.00178	8.0038	0.0047
Shoulder	Shoulder width: Shoulder = 1: shoulder width ≥10ft, Shoulder = 0: shoulder with < 10ft	-0.5613	0.0900	38.9195	<.0001
Model Performance					
AIC		7983.006			
-2 Log L		7963.006			
ROC		0.7095			

The cross-validation method was used to validate the model of logistic regression for total crashes. The ROC (Receiver Operating Curve) value in this case was 0.7095, which indicated a decent validity of the model. The ROC profile is shown in Figure 5.2.

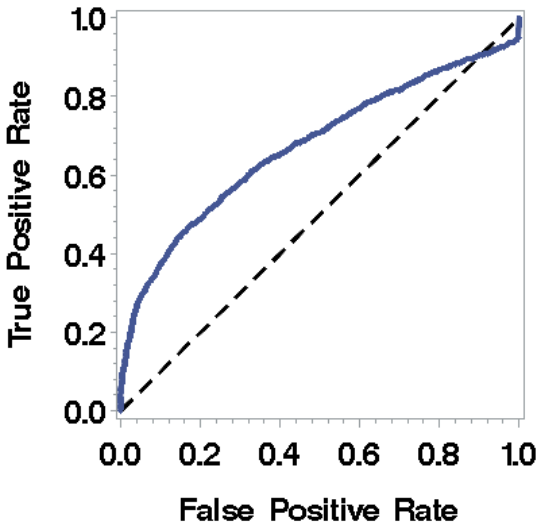


Figure 5.2 - ROC profile for total crash modeling

5.2 Real-Time Safety Evaluation for Ramps

Ramps are an important component of expressways. They connect expressways with service roads or other expressways. The safety of ramps is a big concern. Compared to mainlines, ramps generally have smaller radii and/or are steeper. The crash risk on ramps may be much higher than on mainlines. The occurrence of crashes may block ramps, so vehicles cannot exit or enter an expressway from the desired ramp. One effective way to improve ramp safety is to explore ramp crash precursors in real-time to identify hazardous conditions and reduce the risk of crashes by ITS traffic control.

In general, crash factors can be environmental, traffic, vehicle, and driver. The former two factors are important and are more easily collected by non-intrusive sensors than the latter two. Environmental factors include geometric design and weather, and traffic factors include volume, speed, lane occupancy, truck percentage, and more. The impacts of geometric design and traffic factors are universally studied. However, weather should also be addressed. On average, in the United States from 2002 to 2012, twenty-

three percent (23%) of crashes were weather-related, and seventy-four percent (74%) of weather-related crashes happened on wet pavement (Federal Highway Administration, 2014). Meanwhile, weather-related crashes caused 94 million to 272 million hours of delay each year (Goodwin, 2002). Hence, in addition to geometric design and traffic factors, the weather was also included in the real-time crash analysis.

The project chose 14.2 miles of SR 408, 26.9 miles of SR 417, and 7.6 miles of SR 528. They were in International Airport's (MCO) and Orlando Executive Airport's (ORL) 7.0-mile coverage buffer. Within this buffer, the weather provided by the airport stations could highly represent weather on expressways (Ahmed et al., 2013). Four datasets were collected: detailed information for every crash, MVDS traffic data, ramp geometrics, and weather information. Detailed information is described below.

The raw ramp crash data were obtained from Signal Four Analytics. The dataset contained detailed information for all reported crashes in the period from July 2013 to June 2014. The information included exact time of crash, crash coordinates, crash street and intersecting street, number of vehicles involved, type and severity of the crash, number of injuries and/or fatalities involved, weather conditions, road surface and light conditions, and more.

The traffic flow data were provided by CFX. The traffic data (e.g., volume, speed, and lane occupancy), were calculated automatically every minute by MVDS. MVDS additionally recognized the lengths of passing vehicles and categorized them into four groups: vehicles less than 10 feet long belonged to group 1, 10-24 feet long to group 2, 24.1-54 feet long to group 3, and greater than 54 feet long to group 4. The term "passenger cars" was used for groups 1, 2, and 3, and "trucks" for group 4. The extracted traffic data were collected from 5-10 minutes prior to crash. In order to reduce data noise, the traffic data were aggregated into 5-minute intervals. For example, if a crash occurred at 8:00am, the traffic data were extracted from 7:55am to 8:00am and

from 7:50am to 7:55am. Twenty non-crash cases were randomly chosen for each crash observation. At the same time, if any crash happened within two hours of the time of a non-crash data point, then this non-crash data point was excluded to ensure the purity of the non-crash traffic flow data.

The geometric data of ramps were collected manually using ArcGIS map and the Roadway Characteristics Inventory (RCI). There were 141 ramps, and each ramp had two variables: ramp type and ramp configuration. Of the 141 ramps, 70 were off-ramps and 71 were diamond ramps.

Airport weather data were collected from the National Climate Data Center (NCDC). The dataset included sky condition, weather type, wind direction and speed, pressure, humidity, temperature, visibility, and hourly precipitation. Visibility, weather type, and hourly precipitation were used in this study. Weather type and hourly precipitation were integrated into one variable called road surface condition. If hourly precipitation was higher than zero, or weather type contained TS (thunderstorm), RA (rain), or DZ (drizzle), it was assumed that the road surface condition of an observation was wet in the following hour.

Integrating crash, traffic, geometric, and weather data was important work. Every ramp was first assigned an ID for the geometric data. Then, for every crash, its location in ArcGIS was identified and an ID variable (the same as the geometric data ID) was manually added to stand for the ramp where the crash happened. All traffic data at a particular ramp would have the same ID, which was the same ID as geometric and crash data. Based on this ID variable, crash, geometric, and traffic data were combined. The last step was adding weather data into the combined data. The weather data for the ramps was from the airport closest to the ramps.

After combing these four datasets, 165 crashes and 3300 non-crashes with complete information were filtered out. Each observation had ten variables. The definitions and acronyms of these variables are shown in Table 5.2.

Table 5.2 - Variables considered for real-time ramp safety evaluation

Data	Symbol	Description
Traffic Flow	Spd	Average speed in a 5-minute interval (mile/h)
	Std_spd	Standard deviation of speed in a 5-minute interval (mile/h)
	Log(Vehcnt)	Logarithm of vehicle count in a 5-minute interval (veh/5minutes)
	Occ	Average lane occupancy in a 5-minute interval (%)
	Std_occ	Standard deviation of occupancy in a 5-minute interval (%)
	P_truck	Percentage of trucks in a 5-minute interval (%)
Ramp Geometrics	Type	1 = if the ramp is an off-ramp; 0 = otherwise
	Configuration	1 = if the ramp is a diamond-ramp; 0 = otherwise
Weather	Visibility	The distance at which an object or light can be clearly discerned (mile)
	Surface	1 = if the road surface condition is wet; 0 = otherwise

In order to prevent high correlations between independent variables for crash prediction models, a Pearson correlation test was done before model building. The results showed that occupancy and the standard deviation of occupancy were highly correlated with the logarithm of vehicle count and with speed. Meanwhile, occupancy was also highly correlated with its standard deviation with a 0.82 correlation coefficient. Hence, occupancy and standard deviation of occupancy were excluded from the model building. Only logarithm of vehicle count, speed, speed standard deviation, and truck percentage were built into the model.

Estimation results are shown in Table 5.3. Six variables were found to be significant in the model at the 95% confidence interval. The overall accuracies for training and validation were 0.761 and 0.759, respectively, with a cutoff-point of 0.035.

Table 5.3 - Real-time safety evaluation for ramp crashes

Parameter	Estimate	Standard Error	Wald Chi-Square	*p*-value
Intercept	-7.2613	0.8153	79.3212	<.0001
Log(Vehcnt)	0.8588	0.1183	52.7051	<.0001
Spd	0.0564	0.0109	26.9808	<.0001
P_truck	-2.6582	1.2851	4.2787	0.0386
Configuration	-0.7371	0.197	14.0039	0.0002
Visibility	-0.1634	0.0291	31.4986	<.0001
Surface	1.9536	0.2153	82.3353	<.0001
	AUC	Sensitivity	Specificity	Accuracy
Training	0.850	0.727	0.762	0.761
Validation	0.832	0.733	0.76	0.759

The logarithm of vehicle count in 5-minute intervals was positive, indicating that a high volume might increase the crash risk on a ramp. When there were more vehicles, the crash likelihood increased since exposure was high. Meanwhile, the high volume also increased the interactions between vehicles. More interactions could result in more multi-vehicle crashes or single vehicle crashes in which a driver intended to avoid a multi-vehicle crash. Speed was found to be significant in the positive direction. High speed increased both the braking distance and the reaction distance. Hence, a vehicle travelling at a higher speed was more likely to have a collision with other objects. In addition to vehicle count and speed, truck percentage was also significant. When truck percentage increased by 1%, the odds ratio decreased by 93%. Truck was defined as a vehicle whose length was more than 40 ft. These vehicles were usually commercial trucks in good condition, and the truck drivers were more experienced and drove carefully. Hence, the higher truck percentage decreased the crash likelihood.

Ramp configuration was significant and negatively related to crashes. The odds of a crash on diamond ramps were 47.8% of that on a curved ramp. Curved ramps have smaller turning radii, which could lead to a loss of vehicle control and result in crashes. For the weather variables, both road surface condition and visibility were significant. The odds of a crash on a wet road surface were 7.0 times the odds on a dry road surface. Wet road surfaces have less friction and result in longer braking distances than on dry surfaces. The longer braking distance could result in a vehicle running into a vehicle ahead or into fixed objects. The wet road surface also could result in vehicles spinning out of control. Visibility was significant in the negative direction. Under poor visibility, car following and lane changing are much harder, so vehicles could have rear-end or sideswipe crashes.

The crash hazard for ramps can be identified based on the proposed model. ITS can be implemented to decrease crash risk in real-time, e.g., reducing the speed limit by using Dynamic Message Signs (DMS).

5.3 Conclusions

According to the analysis, ITS traffic data have great potential for a more proactive safety management system. In this project, the data proved their usefulness for expressway mainline and ramp safety evaluations. However, beyond the scope of current discussion, with more traffic detectors deployed on the roadway network and new detection technologies, it is expected that in the future more precise monitoring from different data sources would better serve the traffic authorities and help improve the safety performance of their system.

CHAPTER 6: Real-Time Traffic Data in Micro-Simulation

Fog has a significant impact on traffic safety. It reduces visibility and consequently affects drivers' vision, perception, and judgments. It is necessary to study the relationship between fog and safety. However, because fog is not a frequent event and the occurrence of crashes during foggy condition is rare, collecting enough field samples to get valid results would take a very long time. One solution for this problem is simulation, which can duplicate field traffic performance and provide the safety performance for a specific traffic condition. Among the simulation tools, VISSIM is one of the most widely used.

In this project, VISSIM simulated traffic under foggy conditions based on the data from MVDS and the Fog Monitor System (FMS). The input volume was divided into low, middle, and high levels, and speed limit had two values (50 mph or 70 mph), hence there were six combinations. For each combination, rear-end and lane-change conflicts were collected by the Surrogate Safety Assessment Model (SSAM), which was developed by the Federal Highway Administration (FHWA, Gettman et al., 2008). By comparing the number of conflicts, we could quantify the impact of certain factors, such as volume and speed limit, on road safety in the poor visibility area.

For the purpose of this study, two sources of data were collected. One was fog data, which provided the visibility distance and other weather information. The other was traffic data, which collected the detailed traffic parameters of the selected road segment. The study area was located on Interstate Highway 4 (I-4) from milepost 19 to milepost 23 (a 4-mile length), roughly situated between State Highway 559 and State Highway 557. This area is surrounded by several small lakes, so the possibility of morning fog is high. In January 2008, about 70 vehicles were involved in a large pileup in this area, caused by thick morning fog combined with smoke. Four people were killed and 38 were injured. Given this, simulation of this area is worth it from a practical point of view.

6.1 Data Source

Since fog cannot be used directly to describe weather's effect on driving behaviors, visibility was taken as a surrogate parameter. To measure the value of visibility, the FMS was applied to the selected road segment to collect the visibility related parameters. Detailed information about these parameters is shown in Table 6.1.

Table 6.1 - Field weather and visibility data and units

Name	Unit
Visibility	m
Air temperature	°F
Humidity	%
Barometric pressure	Kpa
Wind direction	/
Wind speed	mph
Solar radiation	W/m^2
Dew point	°F
Subsurface moisture	VWC
Rain fall	inches

There are nine FMS monitors along the 4-mile study segment, with an average spacing of 0.25 mile. Installation sites are shown in Figure 6.1 (Abdel-Aty et al., 2014).

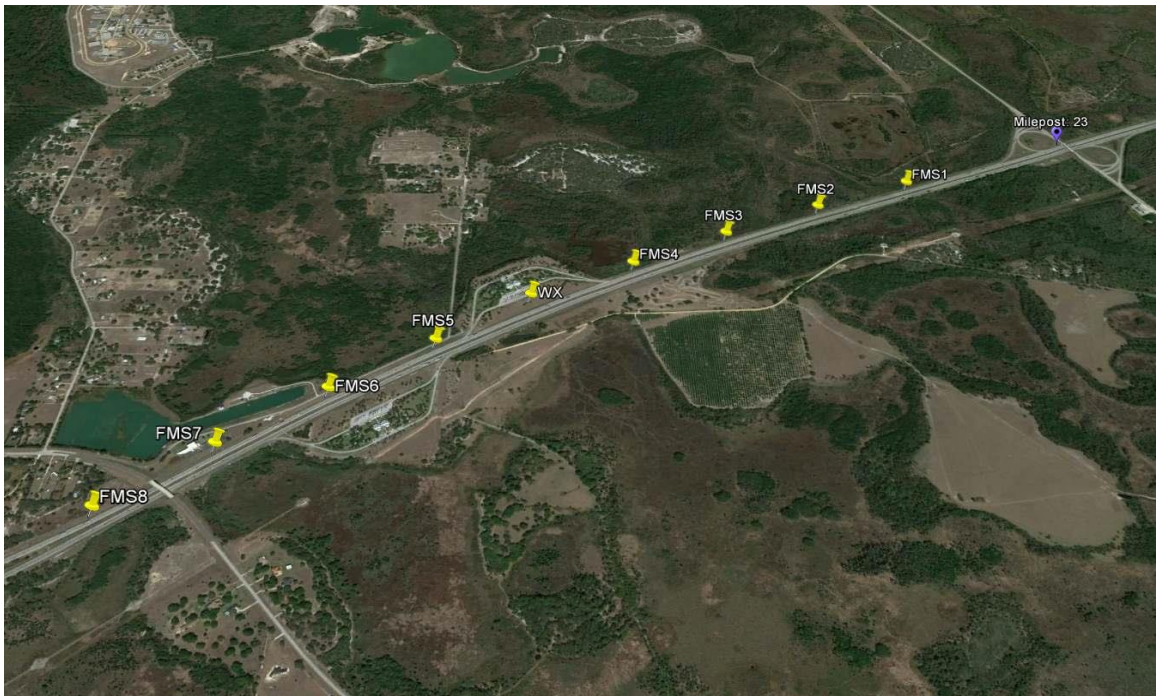


Figure 6.1 - FMS installment locations

Visibility was defined as the distance at which an object or light can be clearly discerned. In VISSIM, the look-ahead distance was defined as the distance at which a vehicle could see forward in order to react to other vehicles in front or to the side. It contained the maximum and the minimum look-ahead distance. Hence, the maximum and minimum visibilities were set as the maximum and minimum look- ahead distances in VISSIM.

In order to investigate the characteristics of traffic flow at this study area, a microwave vehicle detector was installed on I-4. The yellow circle, shown in Figure 6.2, indicates its location. The detector could record eight important variables from every vehicle, including speed, vehicle length, classification, and lane assignment. If the value in the classification column was 1, it indicated the vehicle was a small vehicle; otherwise, the vehicle was defined as a heavy vehicle. By aggregating vehicle count in an hour and

percentage of heavy vehicles, a VISSIM traffic input including hourly volume and truck percentage was obtained.

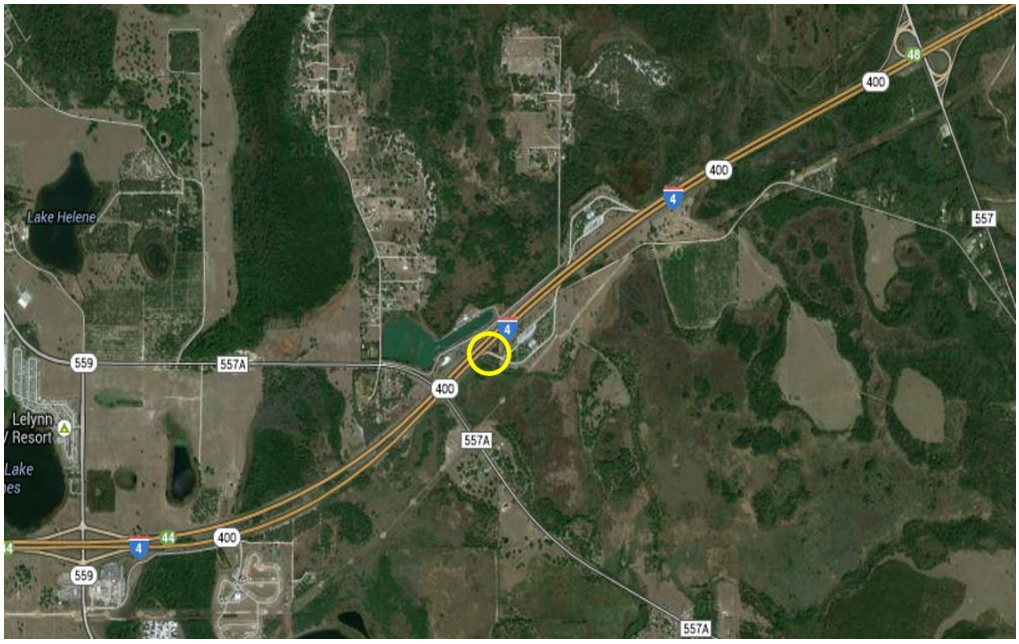


Figure 6.2 - Traffic flow detector location

6.2 VISSIM Calibration and Validation

There is one rest area each on eastbound and westbound, and few vehicles in the rest area exit to or enter from I-4. Because the traffic volume is low, exiting or entering should have an insignificant impact on traffic on the mainline. Only the 4-mile mainline, with three lanes for each direction, was built in VISSIM simulation. Its geometry is shown in Figure 6.3, marked in yellow.

Figure 6.3 - VISSIM network

According to the field data, fog-related low-visibility situations mainly appeared between 5:00am and 10:00am. After 12:00pm, the fog disappeared, which resulted in better visibility for drivers, with visibility values of almost all 1.24 miles. VISSIM intended to simulate the traffic under foggy conditions; fog-related low-visibility data should be chosen. On February 4, 2014 from 7:00am to 8:00am, there was fog with maximum and minimum visibility of 0.17 miles and 0.09 miles, respectively. Therefore, these two values were put in the simulation model. The data of traffic volume, truck percentage, and speed taken from the same date and time were chosen for the fog condition simulation model's calibration and validation.

In order to simulate road traffic conditions, three driving behavior parameters were selected for model validation. These parameters were all parameters in Wiedemann's 99 model, which was designed for freeway traffic. Table 6.2 shows detailed information for these parameters.

Table 6.2 - Parameter definitions in the basic model

Symbol	Name	Definition	Unit
CC1	Headway time	The distance between vehicles in seconds that a driver wants to maintain at a certain speed	s
CC2	Following variation	Restricts the longitudinal oscillation or how much more distance than the desired safety distance a driver allows before intentionally moving closer to the vehicle in front	mile
CC6	Speed dependency of oscillation	Indicates the impact of distance on speed oscillation when a vehicle is in a following process; a large value indicates greater speed oscillation with an increasing space. CC6 = 0 means distance won't affect speed oscillation.	---

In order to find the optimal combination of these parameters and build the VISSM network that can best duplicate the field traffic conditions, each parameter had four different values. This information is shown in Table 6.3.

Table 6.3 - Parameter values for calibrating the basic model

Symbol	Name	Value			
CC1	Headway time	0.9	1.2	1.5	1.8
CC2	Following variation	0.0025	0.0050	0.0075	0.0099
CC6	Speed dependency of oscillation	0	8.00	11.44	14.00

The four values of each parameter generated $4^3 = 64$ combinations. Three runs with different random seeds for each combination were conducted in VISSIM, resulting in 192 runs. For each different random seed, the stochastic functions in VISSIM were assigned a different value sequence and the traffic flow changed. Generally, in order to eliminate the impact of stochastic functions on traffic and produce meaningful results, several runs with different random seeds are needed. The VISSIM output was based on the arithmetic mean of the results obtained from multiple simulation runs with different random seeds.

The field speed was compared with mean simulated speed. The results showed that the group “Headway time = 0.9 s, Following variation = 0.0025 miles, Speed dependency of oscillation = 11.44” satisfied the speed validation requirement because the speed difference between field speed and simulated speed was less than 5 mph at the 95% confidence interval. Detailed information for the speed validation is in Table 6.4.

Table 6.4 - Speed validation results

Direction	Name	Mean	95% Confidence Interval
EB	Speed in the field data (mph)	69.772	(-0.847, -0.065)
	Speed in the calibration data (mph)	70.229	
	Difference	-0.456	
WB	Speed in the field data (mph)	70.117	(-1.429, -0.602)
	Speed in the calibration data (mph)	71.132	
	Difference	-1.015	

6.3 Experimental Design and Results

The validated VISSIM network was used to test the influence of traffic volume and speed limit on road safety, measured by the number of conflicts. Three different volume levels along with two speed limits were set up in this model. The relationship between levels and values is shown in Table 6.5.

Table 6.5 - Experimental design

Traffic volume level	Traffic volume (veh/hour)	Speed limit in the fog area (mph)
Low	4000	50
		70
Middle	8000	50
		70
High	12000	50
		70

One-mile length conflict data were used for analysis, located near the center of the VISSIM network and not impacted by entering and exiting vehicles. Normally entering and exiting vehicles are not influenced by the driver parameters or by traffic conditions (e.g., congestion or no congestion) set in VISSIM. The location of this analysis area is shown in Figure 6.4

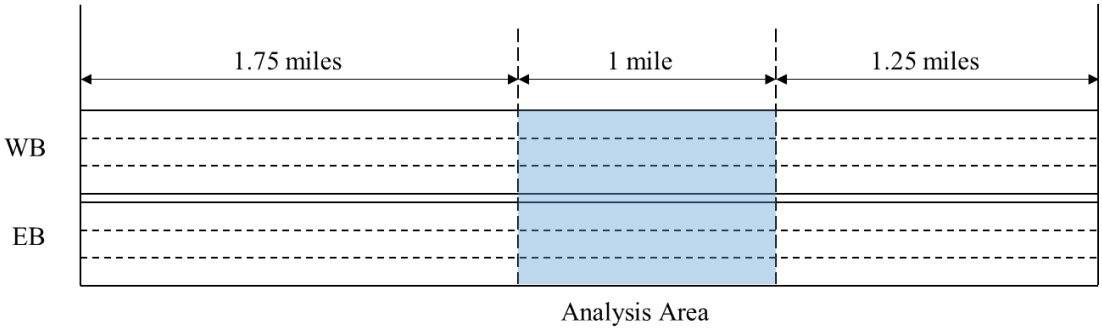


Figure 6.4 - Analysis area

Each simulation run resulted in a corresponding trajectory file. The file was then analyzed via SSAM to obtain conflict information, including conflict number, type, and location. SSAM used several parameters to measure conflicts and to describe the conflict locations and characteristics. The main conflict measure parameters were time-to collision (TTC) and post encroachment time (PET). When TTC was less than 1.5 s or PET was less than 5.0 s, it was considered a conflict.

In SSAM, all conflicts involved more than two vehicles and were summarized into the following four types: unclassified, crossing, rear-end, and lane-change. The conflict angle was used to classify a crash into one of the four types. A very small angle meant that the vehicles' trajectories were close to each other and a potential rear-end collision could occur, while a large angle implied a potential right-angle or head-on collision of the two vehicles (Pu and Joshi, 2008). When a vehicle was approaching from the left, a

negative angle was recorded, whereas a positive angle was recorded when the second vehicle approached from the right. The conflict angles ranged from -180° to 180°.The conflict angles of the four conflict types are shown in Table 6.6.

Table 6.6 - Conflict angles of different conflict types

Conflict Type	Conflict Angle
Unclassified conflict	No conflict angle information provided
Crossing conflict $\boldsymbol{\theta_c}$	$\|\|\theta_c\|\| > 85°$
Rear-end crash $\boldsymbol{\theta_r}$	$\big\|\|\theta_r\|\big\| < 30°$
Lane-change conflict $\boldsymbol{\theta_l}$	$30° < \|\|\theta_l\|\| < 85°$

Not all types of conflict were analyzed in this project. Referencing Table 6.6, the crossing conflict type has an angle of the two vehicles' heading of greater than 85°. This is rare for field highway segments and for VISSIM simulation. In addition, it is difficult to analyze an unclassified conflict, and the sample size of unclassified conflicts was very small. Hence, crossing conflicts and unclassified conflicts were excluded. Table 6.7 shows the number of lane-change and rear-end conflicts from the analysis area, which included eastbound and westbound lanes.

Table 6.7 – Number of conflicts

Volume (veh/h)	Speed Limit (mph)	Number of Conflicts		
		Lane-change	Rear-end	Total
4000	50	25	3	28
	70	134	48	182
8000	50	104	56	160
	70	292	271	563
12000	50	198	131	329
	70	309	270	579

Generally, the total number of conflicts, both lane-change and rear-end, increased with traffic volume. The more vehicles in the study area, the more crashes. It is easy to understand. The higher volume meant higher crash exposure and more interactions between vehicles.

In addition to traffic volume, speed limit was found to be a significant factor on number of conflicts. For the same traffic volume, there were fewer conflicts in the fog area when the speed limit was lower. Lower speed decreases stopping distance, which is the combination of braking distance and reaction distance. A crash occurs when the visible distance between vehicles is shorter than stopping distance. Fog decreases the visible distance between vehicles. If the stopping distance does not change due to fog, it may be longer than the visible distance and the road may have more crashes. Hence, lower speed limits in the fog area should be implemented.

As Table 6.7 indicates, when the volume was low, the speed limit had more impact on conflict. When volume was 4000 veh/h, the total number of conflicts at 70 mph was 6.50 times the number at 50 mph. However, when volume was 12000 veh/h, the ratio was only 1.76. When volume was high, the concentration of vehicles increased and the distance between them decreased. The reduction of visible distance due to fog combined with the reduction in distance due to high volume worsened the safety conditions. The speed limit alone cannot improve conditions at a large magnitude. In this case, warning drivers to drive much more carefully than usual, or advising them to change route or take a break until visibility improves, may be helpful.

CHAPTER 7: Dilemma Zone Simulation

At the onset of a traffic signal's yellow indication, drivers approaching the intersection must make a quick decision to either stop or cross the intersection. Among all the intersection-related crashes, yellow-phase-related crashes caused by these dilemma zones are of significant concern to transportation engineers. The dilemma zone, which is also known as the 'indecision period', describes the region which begins at the position where most people choose to stop and ends at the position where most people choose to cross the intersection at the onset of the yellow indication. Using traditional traffic models to describe the microscopic behavior of vehicles can be a very time consuming and complex process. Recently, with the rapid development of computation technologies, many Cellular Automaton (CA) based simulation models were developed to improve the dilemma zone problem.

7.1 Video Data Analysis

The ST_GO variable describes drivers' decisions at the onset of the yellow indication. "Stop" means the driver chose to stop during the yellow interval, while "go" means the driver decided to cross the intersection at the yellow interval. Five hundred and eighty-five go decisions as well as six hundred and seventy-nine stop decisions were observed. Both of these decisions account for about 50% of all observations.

A logistic model was employed to analyze the importance of different independent variables for the drivers' stop/go decisions (see Table 7.1). Five variables were considered during the variable selection, including SPEED, DISTANCE, LD_FL, L_POSITION and V_TYPE. The variable definitions were explained in section 3.3.

Table 7.1 - Independent variables for the stop/go decision

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
SPEED	0.2118	0.0183	134.302	<.0001
DISTANCE	-0.0246	0.00145	286.449	<.0001
LD_FL	1.0327	0.1622	40.5204	<.0001
L_POSITION	0.0201	0.0991	0.0413	0.8389
V_TYPE	0.2209	0.1577	1.9624	0.1613

According to the results of the logistic regression analysis, the SPEED, DISTANCE, and LD_FL variables had a significant impact on drivers' stop/go decisions at the 0.05 significance level; the other two variables (L_POSITION & V_TYPE) did not have significant influence on drivers' stop/go decisions. Thus, three significant variables (SPEED, DISTANCE, and LD_FL) were chosen to be the main factors for predicting drivers' decisions. The results were consistent with the results of the previous study (Elmitiny et al., 2010).

7.1.1 Speed Variable

The speed limit at the north approach was 45 mph. The mean speed was 48.2 mph, which was slightly higher than the speed limit and lower than the lead vehicles' mean speed of 49.0 mph. The range of the operating speed was 25 mph to 63 mph, and the standard deviation of the speed was 5.0 mph. Most of the operating speeds of the vehicles were at the 45 mph to 55 mph interval, which accounted for 73.2% of the observations (see Figure 7.1). The leading vehicle speeds followed a normal distribution ~$N(49.5, 4.9^2)$, which is considered the expected speed distribution for leading vehicles in this simulation research.

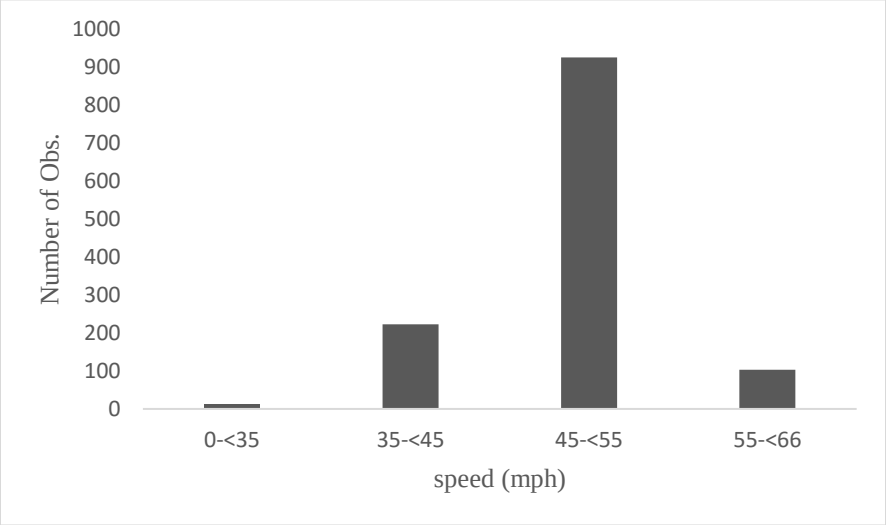


Figure 7.1 - Number of observations in different speed intervals

The mean operating speed of vehicles that made go decisions (M = 49.93, SD = 4.99) was statistically higher (p = 0.000) than the mean operating speed of vehicles that made stop decisions (M = 47.79, SD = 4.83).

In this study, the speed variable was divided into 3 groups (mph): Group 1 [0, 45], Group 2 [45, 53], and Group 3 [53, 66]. Table 7.2 lists the descriptive statistics of drivers' stop/go decisions at the onset of the yellow signal by speed factor. Statistical results showed that people who drove in different speed groups made stop/go decisions differently ($\chi_{3,1264} = 407{,}173$, p = 0.000). With an increase in speed, the probability that the driver chose to stop also increased.

Table 7.2 - Descriptive statistics of stop/go decision by speed factor

Level	Statistics	Stop/Go Decision		
		Stop	Go	Total
0 - <45	Count	158	78	236
	% Within SPEED	66.95	33.05	100.00
	% Within ST_GO	23.27	13.33	18.67
45 - <53	Count	456	369	825
	% Within SPEED	55.27	44.73	100.00
	% Within ST_GO	67.16	63.08	65.27
53 - <66	Count	65	138	203
	% Within SPEED	32.02	67.98	100.00
	% Within ST_GO	9.57	23.59	16.06
Total	Count	679	585	1264
	% Within SPEED	53.72	46.28	100.00
	% Within ST_GO	100.00	100.00	100.00

In Group 1, only 33.05% drivers chose to go. The percentage increased to 44.72 when the driver traveled at 45 mph to 53 mph. When the vehicle's speed was 53 mph to 66 mph, the probability to go was 67.49%, which was significantly higher than the probabilities of Group 1 and Group 2. Regardless of the distance factor and other factors, about 50% of the drivers who drove at the speed of 45 mph to 53mph chose to stop.

7.1.2 Distance Variable

The mean distance of the vehicles to the stop line was 319.3 ft at the beginning of the yellow phases with a standard deviation of 80.2 ft. The minimum distance was 160.0 ft and the maximum speed was 480.0 ft. As can be seen in Figure 7.2, most of the observations were in the 200 ft to 400 ft distance region.

The distance to the stop line had a negative effect on the percentage of drivers who decided to cross the intersection. About 90% drivers chose to go if they were within 220 ft of the stop line. Furthermore, when the distance was more than 400 ft, the probability dropped to below 10% (see Figure 7.3).

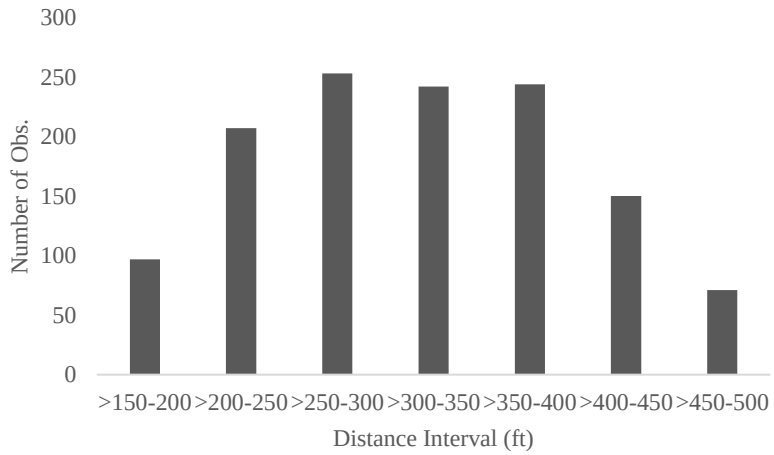


Figure 7.2 - Number of observations in different distance intervals

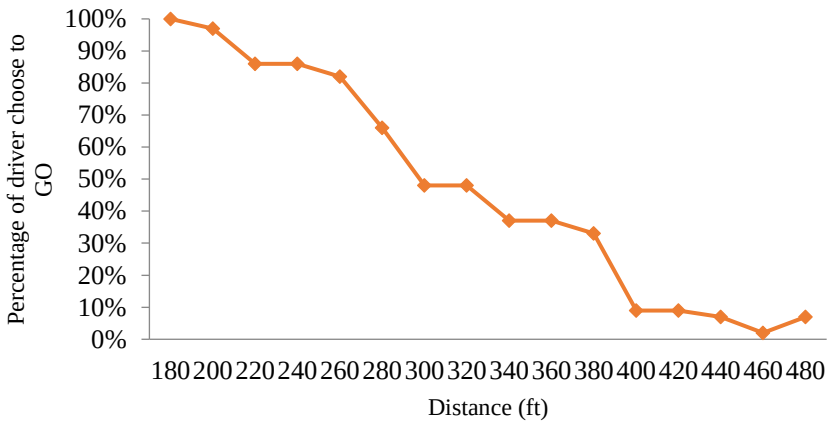


Figure 7.3 - Stop/go decisions with different distances to the stop line

The vehicles' distance to the stop line was divided into 4 groups (ft): Group 1 [0, 280], Group 2 [280, 390], Group 3 [390, 430], and Group 4 [430, 480]. Table 7.3 lists the descriptive statistics of drivers' stop/go decisions at the onset of the yellow signal by distance factor. The statistical test demonstrated a significant difference for the different speed groups ($\chi_{2,1264} = 55.863$, $p = 0.000$). In Group 1, most of the drivers chose to cross the intersection instead of stop. The number dropped significantly in Group 2, in

which only about 40% of drivers decided to go. The go percentage of drivers who were more than 390 ft away from the stop line was below 10%. The trend is logical in that the driver was more likely to cross if s/he was closer to the stop line at the onset of the yellow signal.

Table 7.3 - Descriptive statistics of stop/go decision by distance factor

Level	Statistics	Stop/Go Decision		
		Stop	Go	Total
0 - <280	Count	75	349	424
	% Within DISTANCE	17.69	82.31	100.00
	% Within ST_GO	11.05	59.66	33.54
280 - <390	Count	348	217	565
	% Within DISTANCE	61.59	38.41	100.00
	% Within ST_GO	51.25	37.09	44.70
390 - <430	Count	120	11	131
	% Within DISTANCE	91.60	8.40	100.00
	% Within ST_GO	17.67	1.88	10.36
430 - <480	Count	136	8	144
	% Within DISTANCE	94.44	5.56	100.00
	% Within ST_GO	20.03	1.37	11.39
Total	Count	679	585	1264
	% Within DISTANCE	53.72	46.28	100.00
	% Within ST_GO	100.00	100.00	100.00

7.1.3 LD_FL Variable

Driver behavior of the leading and following vehicles was different. Hurwitz (2009) analyzed driver responses and pointed out the difference between lead and follow vehicles. During the data collection, 565 leading vehicles and 699 following vehicles were recorded.

Table 7.4 lists the descriptive statistics of drivers' stop/go decisions at the onset of the yellow signal by LD_FL factor. The position in platoons also had a significant effect on driver behavior ($\chi_{1,1264} = 93.104$, $p = 0.000$). Table 7.4 also indicates that the

following vehicles were more prone to cross the intersection compared with the leading vehicles.

Table 7.4 - Descriptive statistics of stop/go decision by LD_FL factor

Level	Statistics	Stop/Go Decision		
		Stop	Go	Total
Lead	Count	388	176	564
	% Within LD_FL	68.79	31.21	100.00
	% Within ST_GO	57.14	30.09	44.62
Follow	Count	291	409	700
	% Within LD_FL	41.57	58.43	100.00
	% Within ST_GO	42.86	69.91	55.38
Total	Count	679	585	1264
	% Within LD_FL	53.72	46.28	100.00
	% Within ST_GO	100.00	100.00	100.00

The vehicle speeds at different positions were significantly different. The mean speed of leading vehicles (M = 49.52, SD = 4.93) was statistically higher than the following vehicles (M = 48.19, SD = 5.01) with p = 0.000.

7.1.4 Other Factors

If the vehicle chose to enter the intersection at the onset of the yellow indication, the time elapsed from the onset of the yellow until the car entered the intersection was recorded (Y_TIME). The mean time was 3.9 seconds with a standard deviation of 0.8 seconds. The minimum time was 2.1 seconds, while the maximum time was 7.2 seconds.

There were 538 light trucks and 726 passenger cars observed during the data collection (V_TYPE, see Figure 7.4). The statistical analysis did not show a significant difference between passenger cars and light truck vehicles ($\chi_{1,1264} = 1.576$, p = 0.209).

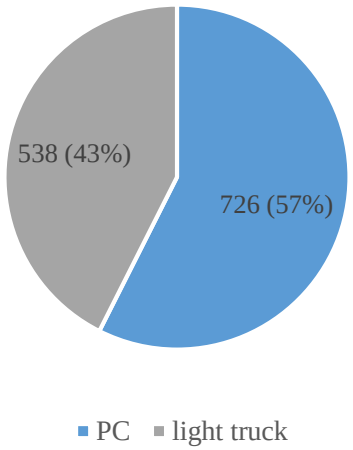


Figure 7.4 - Number of observations by different types of vehicles

The mean speed of passenger cars was 49.0 mph and the mean speed of light trucks was 48.4 mph (Figure 7.5). A significant difference was found at the 0.05 significance level (p = 0.031).

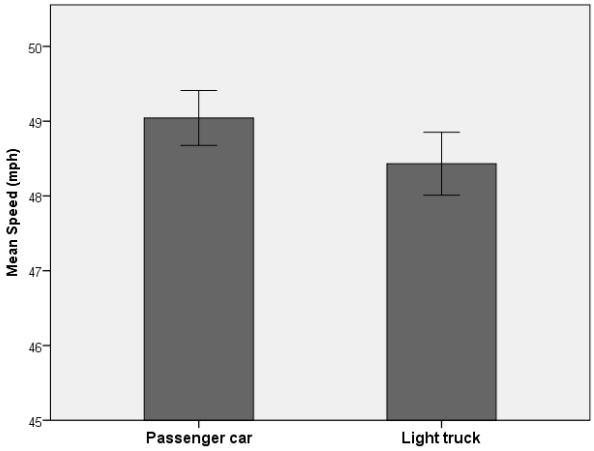


Figure 7.5 - Mean speed by different vehicle types

There were three lanes at the studied approach (L_Position). The middle lane had the highest mean speed, while the left lane had the lowest mean speed. However, the right lane had the lowest standard deviation of the operating speed (see Table 7.5).

Table 7.5 - Speed by different lane position

	Left Lane	Middle Lane	Right Lane
Mean (mph)	47.59	49.77	49.07
Standard Deviation (mph)	5.28	4.80	4.63

Table 7.6 shows the contingency table of different lane positions. There was no significant difference of the drivers' stop/go decisions among vehicles in different lanes ($\chi_{2,1264} = 1.287$, $p = 0.525$).

Table 7.6 - Contingency table of stop/go decisions by different lane positions

	Go	Stop
Left Lane	202	255
Middle Lane	215	235
Right Lane	168	189

Typically, there are two types of red-light running violations (RLR, Federal Highway Administration). The first type of RLR, which is referred to as the "permissive yellow" rule, means the driver can enter the intersection legally during the yellow interval. The second type is called the "restrictive yellow" rule and forbids the driver to enter or be in the intersection during the red interval. The first type of rule is more common in the United States.

Two hundred and seventeen red-light running violations were observed during the data collection. Table 7.7 shows the descriptive statistics of RLR violations. The speed Group 2, distance Group 2, following vehicles, and vehicles in the right lane had a relatively higher percentage of RLR violations. Most of the vehicles that had RLR violations needed about 4 to 5 seconds to cross the intersection at the onset of the

yellow indication, which indicated drivers were more prone to make false stop/go decisions when they had 4 to 5 seconds elapsed time to enter the intersection.

Table 7.7 - Descriptive statistics of RLR violations by different factors

Factor		Statistics	RLR
Speed Group	0 - <45	Count	35
		Percentage (%)	16.1
	45 - <53	Count	129
		Percentage (%)	58.9
	53 - <66	Count	53
		Percentage (%)	24.4
Distance Group	0 - <280	Count	39
		Percentage (%)	18.0
	280 - <390	Count	160
		Percentage (%)	73.7
	390 - <430	Count	10
		Percentage (%)	4.6
	430 - <480	Count	8
		Percentage (%)	3.7
Lead/Follow	Lead	Count	52
		Percentage (%)	24.0
	Follow	Count	165
		Percentage (%)	76.0
Lane Position	Left Lane	Count	71
		Percentage (%)	32.7
	Middle Lane	Count	88
		Percentage (%)	40.6
	Right Lane	Count	58
		Percentage (%)	26.7
Vehicle Type	Passenger Car	Count	124
		Percentage (%)	57.1
	Light Truck	Count	93
		Percentage (%)	42.9
Elapsed Time	>4-5	Count	163
		Percentage (%)	75.1
	>5-6	Count	52
		Percentage (%)	24.0
	>6-7	Count	2
		Percentage (%)	0.9

Table 7.8 illustrates the logistic regression analysis for the RLR violations. Five factors were considered. Three factors (lead/follow, speed, and distance) showed a significant impact on the presence of RLR violations. Distance had the highest impact on the presence of RLR violations, while the lead or follow position had the least impact. The vehicles' lane position and vehicle type did not show a significant relationship with the presence of RLR violations. The parameter estimates indicated that follow vehicles were more prone to have RLR violations. In addition, distance had a negative effect on reducing RLR violations, while speed had a positive effect on reducing RLR violations. Having compared the RLR violations of the leading vehicles and following vehicles, a statistically significant difference was observed ($\chi_{1,1264} = 45.239$, p = 0.000), indicating that following vehicles were more prone to have RLR violations than the leading vehicles.

Table 7.8 - Parameter estimates of the logistic model for RLR violations

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
SPEED	0.0578	0.0164	12.4733	0.0004
DISTANCE	0.0033	0.0010	10.2341	0.0014
LD_FL	1.3737	0.1836	55.9846	0.0000
L_POSITION	0.0486	0.0980	0.2453	0.6204
V_TYPE	0.0527	0.1558	0.1143	0.7353

The distance interval of 340 ft to 370 ft had the most RLR violations. Nearly 60 red-light running violations were observed in that distance region. The number of RLR violations showed an increasing trend in the distance interval of 220 ft to 370 ft, and showed a decreasing trend when the distance to the stop line was larger than 370 ft away from the stop line (Figure 7.6). However, lane position did not make a significant difference on RLR violations ($\chi_{2,1264} = 2.873$, p = 0.238).

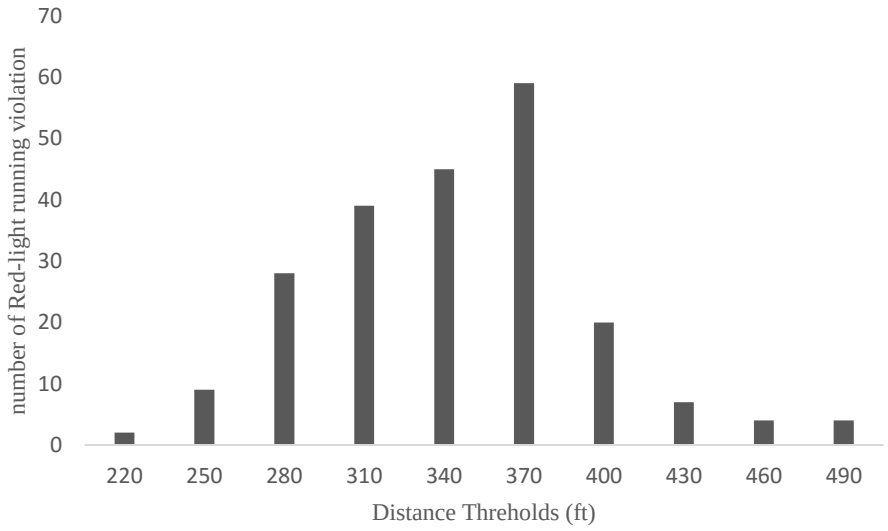


Figure 7.6 - Distribution of RLR violations by distance interval

The RLR vehicles (*M* = 49.67, *SD* = 5.42) had a higher mean speed (*p* = 0.008) than the vehicles without RLR violations (*M* = 48.60, *SD* = 4.91). The results indicated that drivers with an RLR violation could be more aggressive with speed than other drivers (see Figure 7.7). There was a significant difference between the lead vehicles and the following vehicles ($\chi_{1,1264} = 32.532$, $p < 0.001$) when the presence of RLR violations was compared with the different positions in a traffic platoon.

Figure 7.8 illustrates the elapsed time and the distance to the stop line of the vehicles that had RLR violations. Referencing Figure 7.8, it can be observed that most of the violations occurred at 4 s to 6 s elapsed time to enter the intersection and 250 ft to 370 ft away from the stop line. The results were consistent with previous studies (Bonneson et al., 2002).

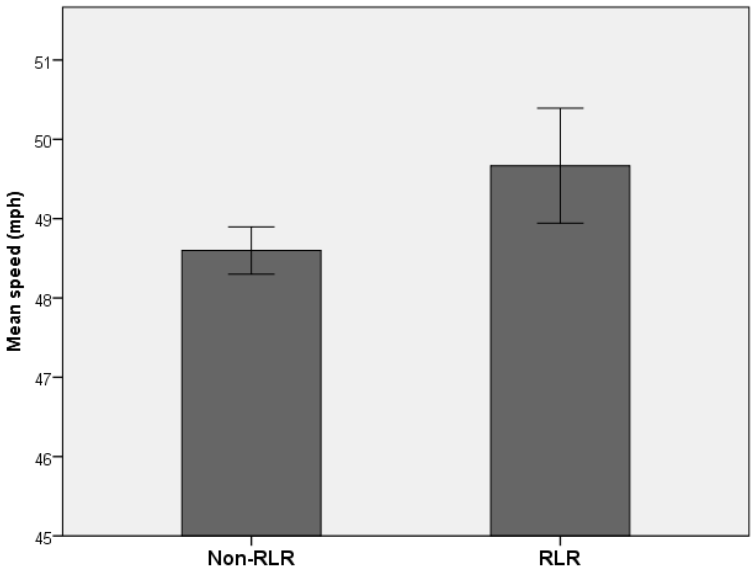


Figure 7.7 - Mean speed by presence of RLR violations

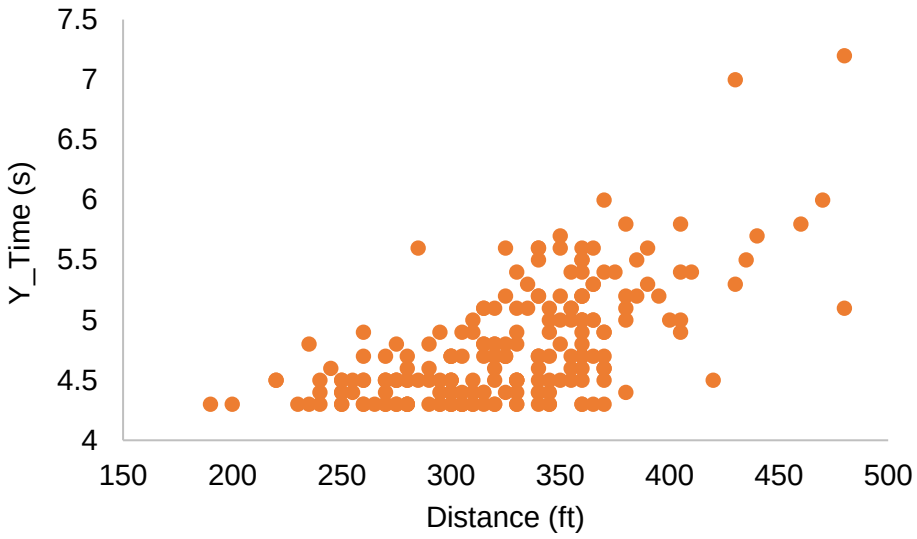


Figure 7.8 - Red-light running violations

7.2 Stop/Go Decision Rules

In statistics, a logistic regression is a type of statistical classification model to predict a binary or categorical dependent variable based on one or more independent variables. A logistic regression analysis can be employed to describe the relationship

between explanatory variables and a response variable. Previous studies have appropriately applied logistic regression analyses to test the significance of observable factors and drivers' characteristics and grouped drivers into different categories (Papaioannou, 2007).

A binary logistic regression is appropriate to use to explain drivers' stop/go decisions as a function of several factors. A logistic model can be used to predict driver behavior. Three factors, which included speed group, distance group, and lead/follow position, were used as variables in the logistic regression analysis to predict drivers' stop/go decisions.

The probability that a driver will decide to cross the intersection was modeled as a logistic distribution in Equation 7.1, where $g(x) = 0$ for stopping and $g(x) = 1$ for crossing.

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \tag{7.1}$$

The logit of the logistic regression model is given by Equation 7.2.

$$g(x) = \ln\frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots\ldots + \beta_n x_n \tag{7.2}$$

Statistical analysis was performed using SAS and hypothesis testing was based on a 0.05 significance level (see Table 7.9). The logistic model was found to be appropriate for the data (Hosmer-Lemeshow goodness of fit $\chi^2 = 2.7349$, $df = 8$, $p = 0.9499$). The ROC area of 0.874 indicated that 87.4% of (go, stop) decision pairs were classified correctly by the model, which meant that predictive accuracy was good. The odds ratio of the lead/follow vehicles meant the odds of go decisions for follow vehicles

were 2.547 times the odds of the go decisions for lead vehicles. Meanwhile, referencing Table 7.9, odds ratios for distance Group 2 and Group 3 relative to distance Group 1 were 4.479 and 26.629, respectively. Further, the odds of go decisions for speed Group 2, speed Group 3, and speed Group 4 were 0.090, 0.011, and 0.005 times the odds of go decisions for speed Group 1, respectively.

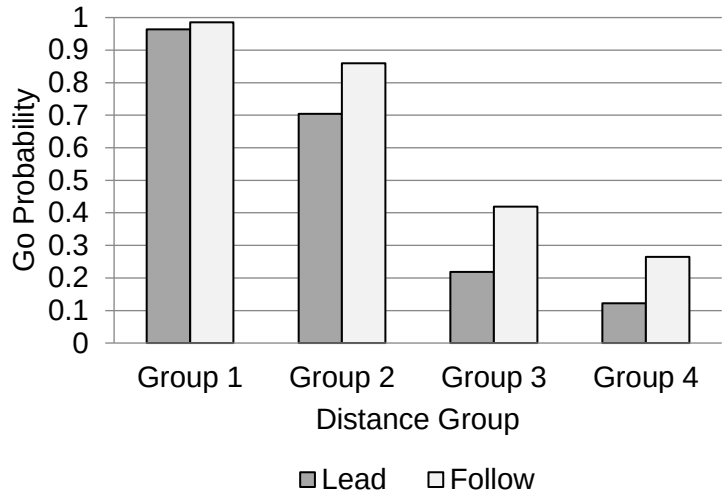
Assuming that the speed of a vehicle was 45 mph to 53 mph (Group 2), the probability of the lead vehicle driver choosing to go was always lower than the following vehicle, and dropped more quickly than the follow group with increased distance, as Table 7.9 demonstrates. When the lead position car was more than 430 ft from the stop line, the probability of the driver choosing to go was only about 10%.

Table 7.9 - Model estimation and odds ratios

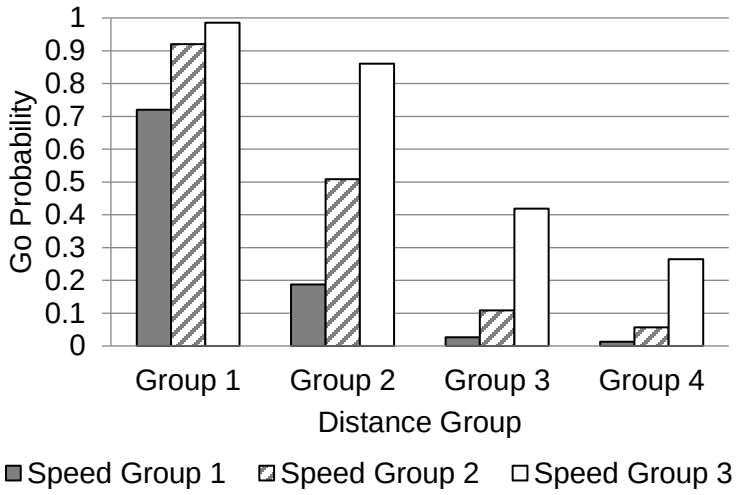
Parameter	Estimate	Odds Ratio	95% Wald Confidence Limits		Wald Chi-Square	Pr > Chi-Square
Follow vs. Lead	0.9458	2.547	1.870	3.469	35.8336	<.0001
Speed Group						
2 vs. 1	1.4994	4.479	2.974	6.746	51.4817	<.0001
3 vs. 1	3.2820	26.629	14.837	47.793	120.9566	<.0001
Distance Group						
2 vs. 1	-2.4108	0.090	0.063	0.128	174.7836	<.0001
3 vs. 1	-4.5557	0.011	0.005	0.022	141.3785	<.0001
4 vs. 1	-5.2498	0.005	0.002	0.013	122.3220	<.0001

Figure 7.9 shows the go probability for the following vehicle in different speed and distance groups. The driver preferred to go when s/he traveled at a higher speed. Even if the car was only 390 ft away from the stop line, the relative probability for vehicles below 45 mph was only near 20%. Vehicles in the following position in the platoon were more prone to go compared to the leading vehicles. Meanwhile, drivers who were in speed Group 1 were prone to choose to stop if s/he was more than 280 ft

away from the stop line. If the vehicles were in distance Group 4, the drivers were prone to stop no matter how fast the vehicle was traveling (Figure 7.9 (b)).



(a)



(b)

Figure 7.9 - Drivers' stop/go decisions

7.3 Cellular Automaton Model

Previously, microscopic simulation of driver behavior was very complex and time consuming. With the rapid development of computer technology, a number of simulation systems have been developed, including different types of Cellular Automaton (CA) models. Due to characteristics of the CA model, it was widely used for traffic flow simulation once it was introduce to the traffic field.

7.3.1 Simulation Environment

During the simulation, the lane was composed of cells that could be empty or occupied by one car. Each cell corresponded to 1.5 m and each car occupied 5 cells (standard value in CA models). The simulation environment was set up as an open boundary one-dimensional lattice.

7.3.2 Model Parameters and Variables

A series of parameters were defined in the CA model as shown in

Table 7.10. Some variables were from the literature, and some variables were calibrated by the field data. The length of the road was set as 5000 cells, or 7500 m. The number of time steps was indicated by T, with 1 time step representing 1 second during the simulation. The simulation covered 1500 seconds (time steps). The maximum acceleration was 1.5 m/s^2 (1 cell/s^2), and the maximum deceleration was 3 m/s^2 (2 cell/s^2). The initial operating speed of the vehicles followed a normal distribution, which was calibrated by the field data. The expected mean speed of the leading vehicles was set as the input.

Table 7.10 - CA model parameters

Parameter	Description	Value
L	Road length	5000 (cells)
T	Number of time steps	1500 (s)
Acc	Maximum acceleration	1 (cell/s^2)
Dec	Maximum deceleration	2 (cell/s^2)

v_{int}	Initial speed of vehicle	~ Normal(49.5, 4.9^2)

7.3.3 Driver Behavior

7.3.3.1 General Rules

Response time referred to the time interval between the signal changing and the brake (or acceleration) response. Wortman and Matthias (1983) found that the average reaction time was between 1.09 s and 1.55 s, which was consistent with the results of other studies (Chang et al., 1985; Newton et al., 1997; Gates et al., 2007). The Institute of Transportation Engineers (1989) recommended 1.0 second as the brake-response time for yellow intervals. In the simulation, the default value for the yellow interval reaction time was 1.0 second.

Drivers' stop/go decisions were based on the probabilities calculated by the logistic regression. The sensitivity and specificity reached the same value when the probability was equal to 0.48 (Figure 7.10). Therefore, if the probability value was larger than 0.48, the driver chose to go. Otherwise, the driver chose to stop. Other driver behavior rules followed the CA model rules (Jia et al., 2007; Ding et al., 2014).

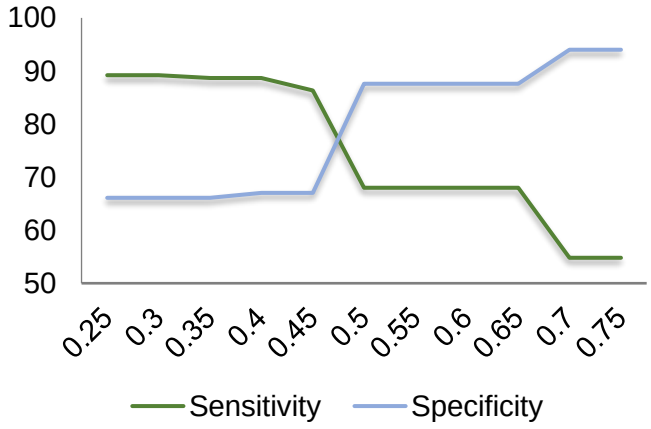


Figure 7.10 - Relationship between sensitivity and specificity in the logistic model

7.3.3.2 Randomization

For drivers that did not have an obvious speed up or slow down behavior, it was assumed that the slowing was caused by the randomization. The acceleration of the non-stopping cars revealed a normal distribution. Ding et al. (2014) calibrated the randomization probability at the intersection at p = 0.16.

7.3.3.3 Updating Rules

The position and velocity of the vehicles were updated according to the following transition rules:

- Step1 - Acceleration: If $v_n < v_{max}$, the speed is advanced by one, unless the distance to the next vehicle ahead is smaller than $v_n + 1$.

$$v_n \rightarrow \min(v_n+1, v_{max}) \tag{7.3}$$

- Step 2 - Deceleration: If the n^{th} vehicle's speed will exceed the front vehicle at the next time step (Δt), the velocity of the n^{th} vehicle is reduced by 1.

$$v_n \rightarrow \min(\, v_n,\, d_n/\Delta t - 1) \tag{7.4}$$

- Step 3: The velocity of each vehicle (if $v_n > 0$) is decreased by one with probability p.

$$v_n \rightarrow (v_n - 1, 0) \tag{7.5}$$

- Step 4: Update vehicle movement.

$$x_n \rightarrow x_n + v_n \times \Delta t \tag{7.6}$$

7.3.4 Signal Timing

A fixed time signal control program was set at the intersection. To simplify the simulation process, the intersection signal was set as a fixed timing program with a relatively short circle of 60 s that included a 25 s green phase, a 4 s yellow phase, and a 31 s red phase. In the fixed signal time program, when a signal changed from green to yellow, drivers made stop/go decision and behaved differently under different countermeasures.

7.3.5 Risky Situations

Two types of dangerous situations were considered during the simulation: rear-end crashes and red-light running violations. There were no rear-end crashes during the simulation based on the CA model, thus a concept of risky situation was proposed to describe the rear-end crashes caused by the behavior of the drivers (see Figure 7.11). Sometimes, the drivers' behaviors were caused by the false expectations of other drivers. Risky situations were divided into two types, those caused by a stopped car and those caused by non-stopped cars. A criterion of "slam on the brake" was defined to describe situations in which the former car encountered emergencies (such as the signal turning from green to yellow) and a risky situation presented because of the inefficient response time. Meanwhile, the presence of RLR violations was also caused by the false decisions of the drivers. Thus, another criterion that described the percentage of drivers' false go decisions was proposed to compare the potential RLR risk.

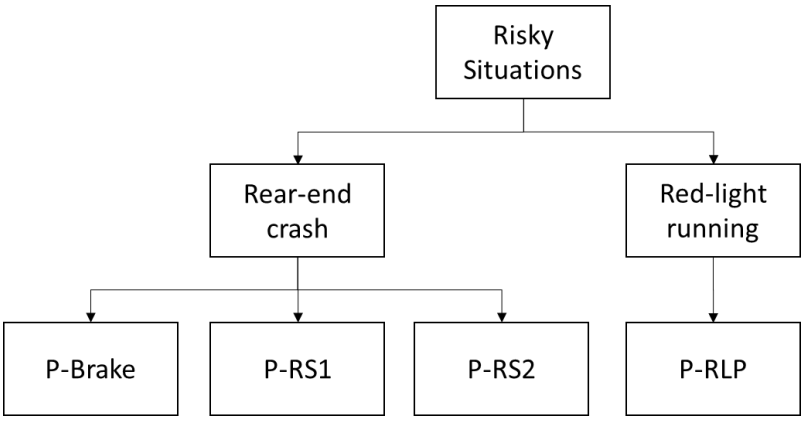


Figure 7.11 - Risky situations

Four risky situations were analyzed: slam on the brake, situations caused by a stopped car, situations caused by non-stopped cars, and RLR rate.

- Slam on the brake (BRAKE)

$$(a1)v_n^t - v_n^{t+1} > 2 \quad (7.7),$$

- RS caused by stopped cars (RS1)

$$(a1)v_n^t - v_n^{t+1} > 2 and\ d_n^{t+1} = 0, (b1)v_{n+1}^t > 0, (c1)\ v_{n+1}^{t+1} = 0 \quad (7.8),$$

- RS caused by non-stopped cars (RS2)

$$(a1)v_n^t - v_n^{t+1} > 2 and\ d_n^{t+1} = 0, (b1)v_{n+1}^t > 0, (c1)\ v_{n+1}^{t+1} \neq 0 \quad (7.9),$$

- False go decision (RLR)

$$v_n^t * t_Y < \mathrm{x_n^t} \quad (7.10),$$

where t_Y is the yellow interval and $\mathrm{x_n^t}$ is the gap of the vehicle to the vehicle in front.

The probabilities of occurrence of rear-end crashes caused by stopped cars, non-stopped cars, or slam on the brake were denoted as $P-RS1$, $P-RS2$, and $P-BRAKE$, respectively. The probability of occurrence of RLR was $P-RLR$.

7.3.6 Simulation Output

The simulation was conducted based on C#. The output contained six documents. Brake1 contained data of emergency braking (BRAKE). Each number in the Brake1 document represented the $P-BRAKE$ during each simulation process (see Figure 7.12).

0.000257556	0.000308188	7.69E-05	8.77E-05	0.000237444	0.000213352	0.000164864	5.07E-05	0.000124063	0.000192165	0.000206933	6.05E-05	0.000271738	8.97E-05	0.000109318

Figure 7.12 - Output of Brake1 document

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
0.000190502	0.000238869	6.40E-05	6.67E-05	0.000176143	0.000146088	0.00012577	4.47E-05	8.63E-05	0.000153652	0.000161928	4.62E-05	0.000203579	4.63E-05	8.05E-05	0	9.48E-05	1.43E-05	0.000255412

Figure 7.13 - Output of DSZ1 document

DSZ1 contained data of risky situations caused by stopped cars (RS1). Each number in the DSZ1 document represented the $P-RS1$ during each simulation process (see Figure 7.13). DSZ2 contained data of risky situations caused by non-stopped cars (RS2). Each number in the DSZ2 document represented the $P-RS2$ during each simulation process (see Figure 7.14).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
0	7.69E-06	8.00E-06	0	0	6.45E-06	0	0	0	7.69E-06	7.14E-06	0	8.00E-06	7.41E-06	0	0	8.00E-06	0	8.00E-06	0	2.36E-05	1.43E-05	0	0	7.69E-06

Figure 7.14 - Output of DSZ2 document

The tl1 document contained the spatial and temporal information of RS1 and RS2 (see Figure 7.15). The first line represented the time when the RS1 or RS2 happened, the second line represented the location of the risky situation, the third line recorded in which simulation process the risky situation presented, and the last line showed the expected speed of the vehicles.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
34	35	36	38	40	42	36	38	39	41	44	34	35	38	40	41	43	36	39	41	42	44	45	35	37	38	40	42	43	45
4990	4990	4985	4980	4975	4970	4990	4985	4985	4980	4975	4990	4990	4985	4980	4980	4975	4990	4985	4980	4980	4975	4975	4990	4985	4980	4975	4970	4965	4960
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02
13	13	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	13	13	16	16	16	16	16

Figure 7.15 - Output of tl1 document

The tl2 document contained the spatial and temporal information of emergency brake (see Figure 7.16). The first line represented the time when the emergency brake happened, the second line represented the location of the risky situation, the third line recorded in which simulation process the risky situation presented, and the last line showed the expected speed of the vehicles.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM
27	30	31	33	34	35	37	39	41	32	35	37	38	40	43	26	27	30	31	33	34	37	39	40	42	32	35	38	40	41	43	44	31	34	36	37	39	41	42
4990	4989	4995	4987	4990	4983	4978	4975	4968	4994	4989	4981	4985	4978	4975	4992	4995	4986	4995	4982	4990	4985	4973	4980	4973	4994	4989	4983	4975	4980	4970	4975	4994	4990	4985	4978	4975	4970	4964
0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02
13	13	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	13	13	16	16	16	16	16	16	13	13	13	13	15	15	15	15	15

Figure 7.16 - Output of tl2 document

The stgo-error document contained the information about drivers' false decisions. Each number represented the $P - RLR$ during each simulation process (see Figure 7.17).

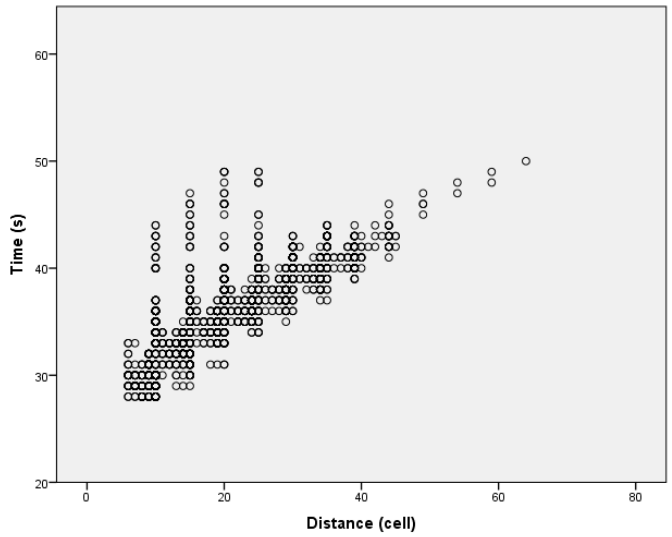
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
0	0.5	0.5	0	0.5	0.5	0.333333	0	0.5	0.333333	0	0.5	0.333333	0.25	0	0.5

Figure 7.17 - Output of stgo-error document

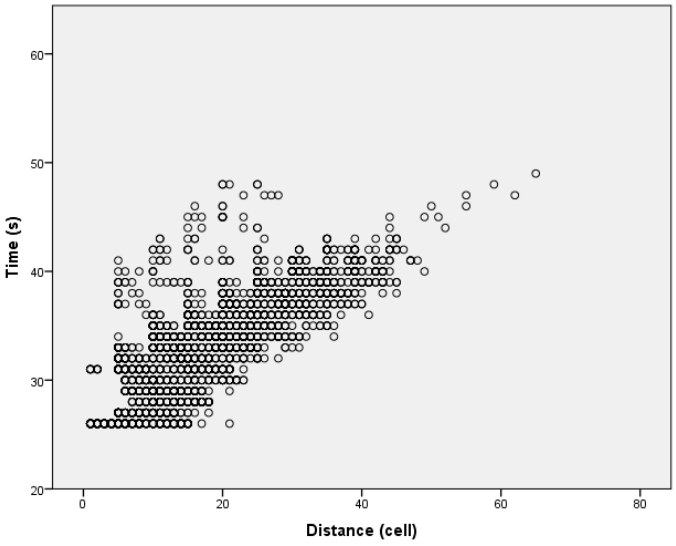
CHAPTER 8: Comparative Analysis of Different Dilemma Zone Countermeasures

8.1 Typical Intersection

A typical simulation will follow the general rules described in Chapter 7.3.3.1. The spatial and temporal information of the risky situations when the expected mean speed of the leading vehicles followed a normal distribution ~$N(50, 5)$ is depicted in Figure 8.1. Figure 8.1(a) shows the RS1 and RS2 risky situations, which represented the possible rear-end crashes. Most of this type of risky situation was present at the end of yellow phases and the beginning of red phases. One possible reason for these situations was the difference of the driver behavior during the yellow interval and their false judgments of other drivers' decisions (Yan et al., 2005). Meanwhile, most of the risky situations were located 10 ft to 45 ft away from the intersection. Figure 8.1(b) describes the distribution of the presence of the emergency brake. Different from rear-end crash risky situations, most of this type of risky situation was present closer to the intersection and began to present soon after the onset of the yellow indication. Thus, most risky situations happened at the beginning of the yellow indication until 10 seconds after the onset of the red indication, and were located at 7.5 m (5 cells) to 60 m (40 cells) away from the stop line.



(a)



(b)

Figure 8.1 - Spatial and temporal distribution of risky situations at the typical intersection when the expected speed of lead vehicles followed ~*N*(50, 5) (a) RS1 and RS2 and (b) emergency brake

When increasing the standard deviation of the leading vehicles, the risk probabilities also showed an increasing trend, especially when the standard deviation increased from 2 mph to 5 mph. If the standard deviation had already increased to 5

mph, the probability of the risky situations was more sensitive to the increase of the expected mean speed (Table 8.1).

Table 8.1 - Impact of standard deviation of leading vehicles on risky situations at the typical intersection

Speed Distribution / Risky Situation ($*10^{-5}$)	(50, 2)	(50, 5)	(50, 10)
P-BRAKE	9.32	64.41	67.58
P-RS1	6.44	43.97	43.09
P-RS2	0.31	2.88	2.75

Table 8.2 shows the potential RLR violations by different expected mean speeds of the leading vehicles at the typical intersection. When the expected mean speed was lower than 50 mph, there was an increasing trend in the percentage of false decisions with the increase in speed. If the expected mean speed increased from 50 mph to 60 mph, the percentage was reduced. As demonstrated in Chapter 7.2, vehicles traveling at a speed of about 50 mph were more prone to have a false go decision. Thus, the simulation scenarios with more vehicles driving at this operating speed had a higher percentage of false decisions, which led to high P-RLR.

Table 8.2 - RLR risk probabilities by different expected mean speed of leading vehicles at the typical intersection

Speed Distribution	(30, 5)	(40, 5)	(50, 5)	(60, 5)	(50, 2)	(50, 10)
RLR Risk Probability ($*10^{-4}$)	37	56	166	52	323	134

8.2 Intersection with Flashing Green Signal

8.2.1 Scenario Construction

In some countries, flashing green signals are implemented at the end of green phases to give drivers advanced warning of the upcoming yellow indication; these are still part of the green phases. Newton et al. (1997) found that about 80% of drivers made acceleration or deceleration decisions during the flashing phases in the Change Anticipation System (TLCAS) program. Many of them decelerated, but some of the drivers chose to accelerate. The TLCAS maximum deceleration value (2.5 m/s^2) was significantly different from the regular program (3.1 m/s^2) that did not include flashing indication phases. The maximum acceleration was 1.6 m/s^2 in the TLCAS program in contrast to 2.0 m/s^2 in the regular program. As for reaction time, the mean value for the TLCAS was 2.05 s, which was much larger than the regular program. Based on results from the TLCAS program, the simulation process of an intersection with flashing green signals is shown in Figure 8.2.

Four seconds before the onset of the yellow indication, the flashing green signal began. To simplify the simulation, 75% of the drivers decelerated at 2.0 m/s^2, 5% of the drivers accelerated at 1.0 m/s^2, and the other 20% of drivers approached the intersection with the same speed. The default value for the flashing green reaction time in this simulation was 2 seconds. The rules of driver behavior after the onset of the yellow indication were the same as the typical intersection (section 8.1)

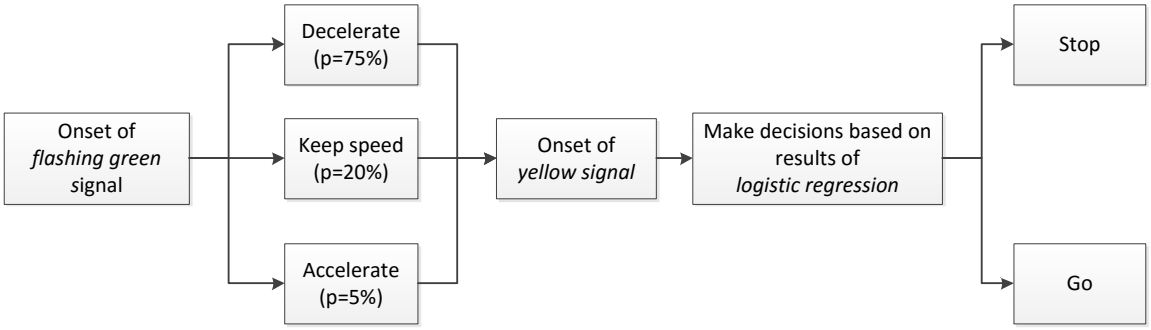
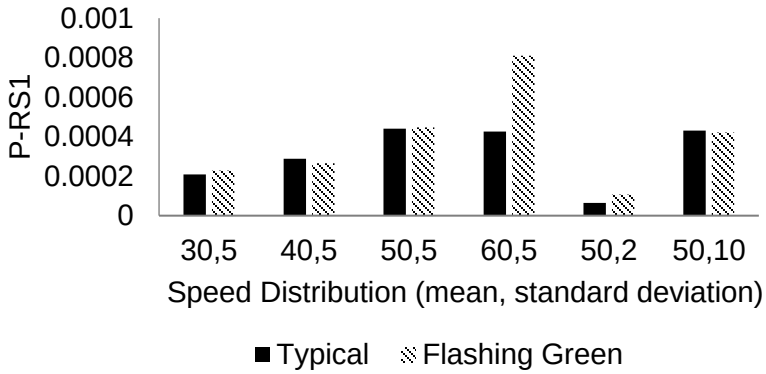


Figure 8.2 – Simulation process of an intersection with flashing green signals

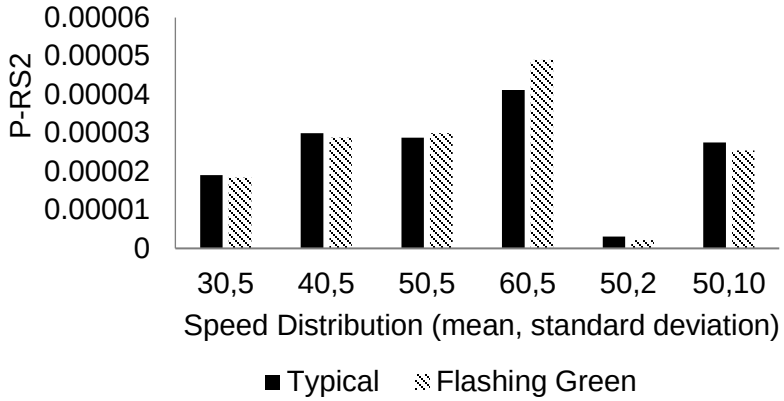
8.2.2 Simulation Results

Figure 8.3 shows the impact of increasing expected mean speed of leading vehicles on the presence of risky situations. Increasing trends were found. When the expected mean speed increased from 40 mph to 50 mph, risky situations caused by a stopped car significantly increased, which indicated that many of the risky situations were due to drivers' different stop/go decisions. When the expected mean speed increased from 50 mph to 60 mph, the percentage of drivers who chose to stop dropped significantly because of the relatively larger parameter estimate of speed Group 3 in the model (equal to 3.2820). Thus, there was no significant increase in risky situations caused by stopped cars when the expected speed increased from 50 mph to 60 mph.

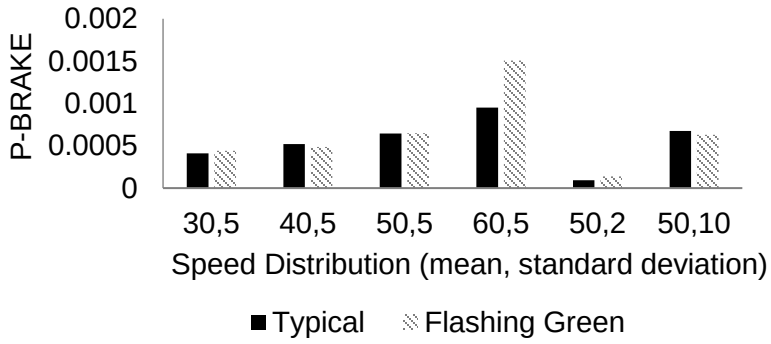
Additionally, Figure 8.3 shows the comparison of the intersection with the flashing green signal and the typical intersection. The flashing green led to a longer driver indecision interval, and drivers' behaviors were more varied, which made drivers' behaviors harder to predict and could lead to rear-end crashes. No significant improvement was found. However, when the vehicles had high operating speed or there was low variance of the speed between vehicles, higher risk probabilities presented. The different acceleration or deceleration decisions during the flashing green phases caused the increasing risky probabilities in some cases. The results were consistent with the results of the previous study (Factor et al., 2012).



(a)



(b)



(c)

Figure 8.3 - Impact of expected mean speed of leading vehicles on risky situations at the typical intersection and the intersection with flashing green signal (a) P-RS1, (b) P-RS2, and (c) P-BRAKE

Figure 8.4 shows the spatial and temporal information of the risky situations of the intersection flashing green signals. When the expected mean speed was 30 mph, most of the risky situations presented during 25-45 seconds at 7.5 m (5 cells) to 60 m (40 cells) away from the stop line. When the expected mean speed increased from 30 mph to 60 mph, the time range increased to 25-50 seconds and the distance range increased to 7.5 m (5 cells) to 75 m (50 cells). Thus, both the time and the distance range of the risky situations' distributions were increased with the increase of the expected mean speed.

The RLR risk of the flashing green did not have a significant effect compared to the typical intersection because in both scenarios the drivers made their stop/go decisions based on the logistic model, which was affected by speed, distance, or the lead/follow position of the vehicles (see Table 8.3).

Table 8.3 - P-RLR of the typical intersection and the intersection with flashing green signal

Expected Speed Distribution / Scenario	(30, 5)	(40, 5)	(50, 5)	(60, 5)	(50, 2)	(50, 10)
Typical Intersection Scenario (*10^{-4})	37	56	166	52	323	134
Flashing Green Scenario (*10^{-4})	37	84	205	85	300	138

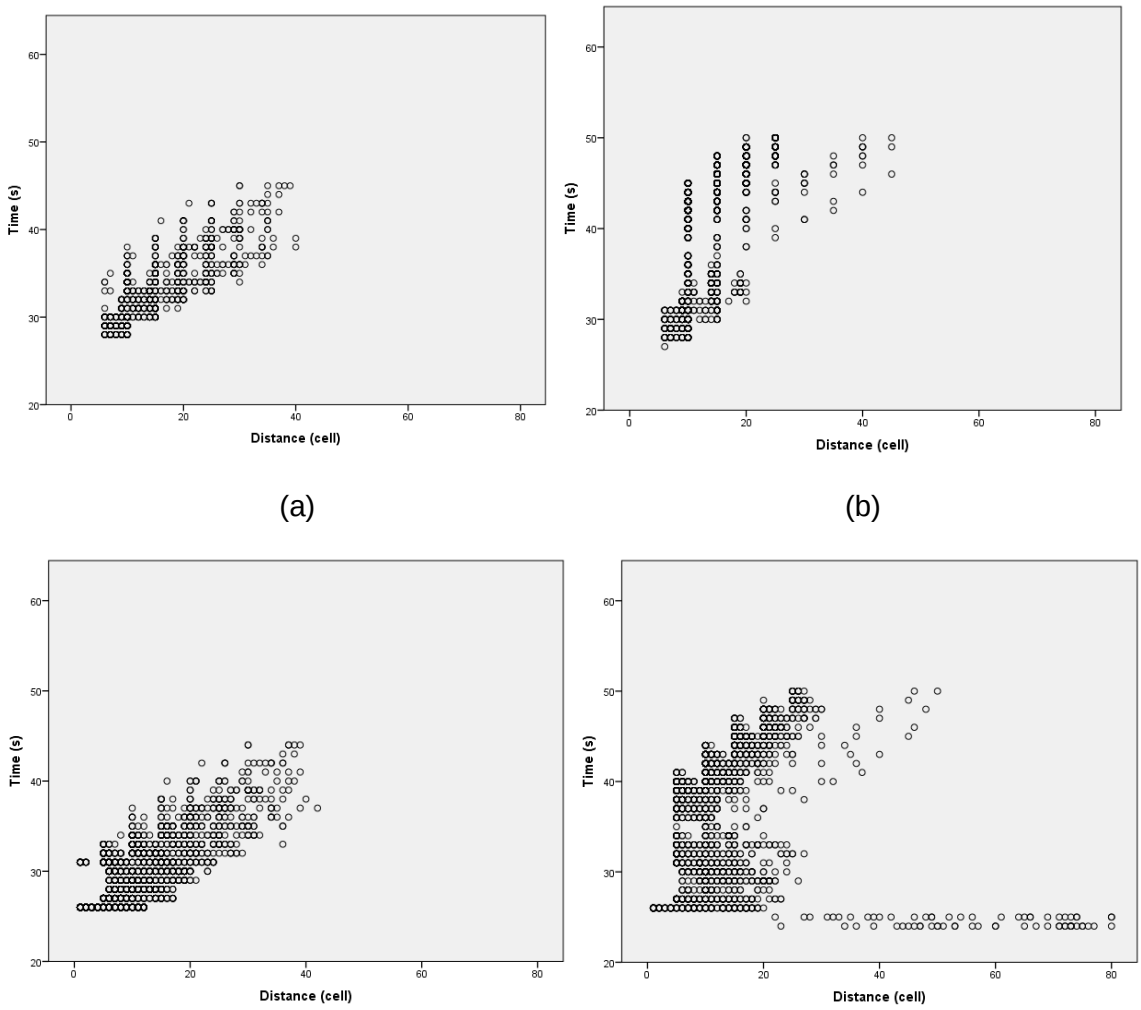


Figure 8.4 - Spatial and temporal distribution of risky situations at the intersection with flashing green signals

(a) RS1 and RS2 when the expected mean speed was 30 mph, (b) RS1 and RS2 when the expected mean speed was 60 mph, (c) emergency brake situations when the expected mean speed was 30 mph, and (d) emergency brake situations when the expected mean speed was 60 mph (standard deviation = 5 mph)

8.3 Intersection with Pavement Marking

8.3.1 Scenario Construction

The rules of pavement marking are depicted in Figure 8.5. Two cars (*A* & *B*) face the change of yellow indication. Since car *B* has passed the pavement marking, s/he should choose to go while car *A* should choose to stop.

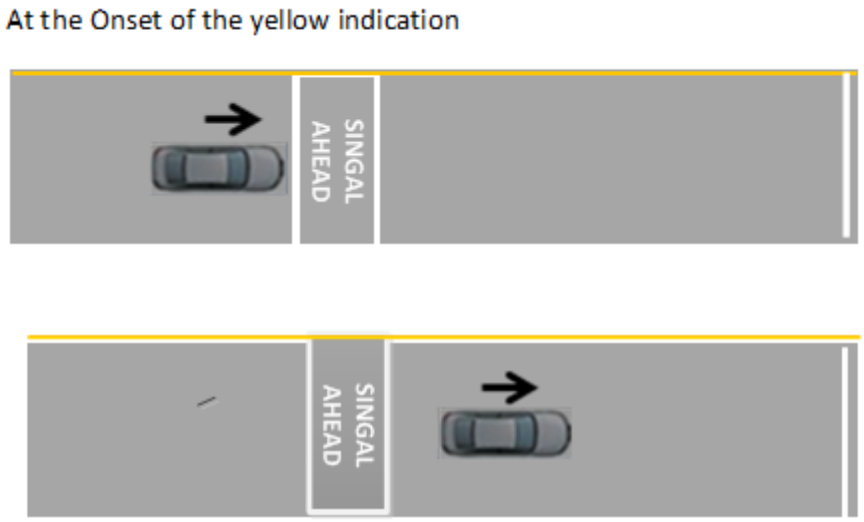


Figure 8.5 - Scenario of intersection with pavement marking

ITE's Engineering handbook (Pline, 1999) suggests the distance from the marking to stop bar is

$$X = V_0 t + \frac{{V_0}^2}{2a+19.6g} \quad (8.1),$$

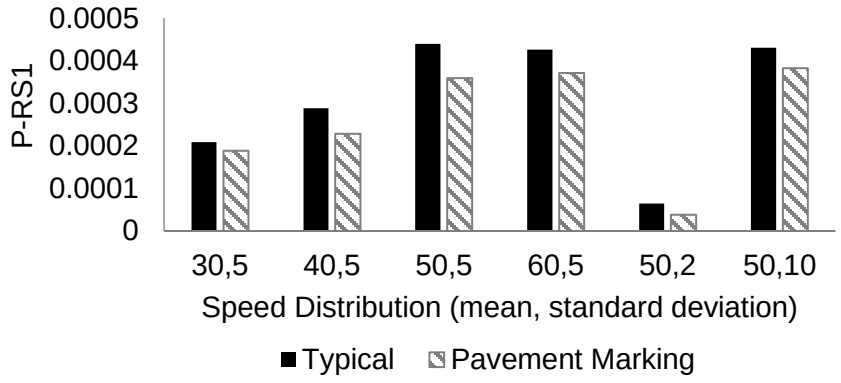
where V_0 is the 85th percentile speed or speed limit, *t* is the reaction time, *a* is the average deceleration rate, and *g* is the grade of the intersection. In this study, the pavement marking was designed with a 45 mph speed limit.

8.3.2 Simulation Results

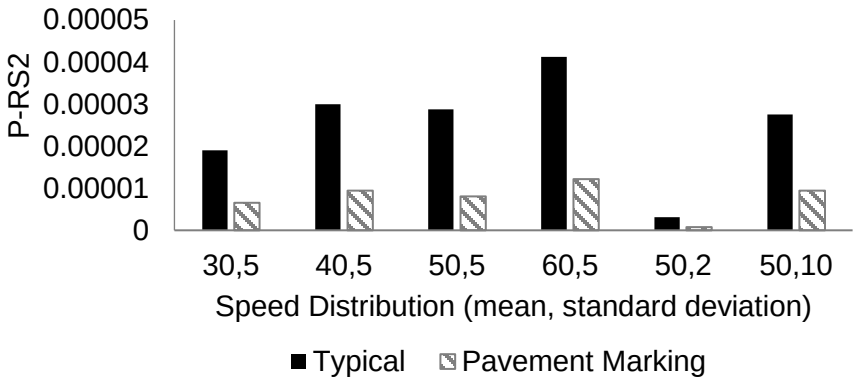
Figure 8.6 demonstrates the risky situations probabilities of pavement-marking scenarios. Significant improvement in the three different types of risky situations was observed, especially in risky situations caused by non-stopped vehicles. The pavement marking countermeasures had a more positive effect when the operating speeds of the vehicles were relatively higher. When the expected mean speed was 30 mph, little effect was found in reducing the risk probabilities. The probabilities of risky situations under the pavement-marking scenario increased with the increasing of the standard deviation and the expected mean speed.

Figure 8.7 depicts the presence of risky situations at the intersection with the pavement-marking countermeasure. Compared with the simulation with expected mean speed of 30 mph and the simulation with expected mean speed of 60 mph, the pavement-marking countermeasure effectively prevented the expansion of the distance range.

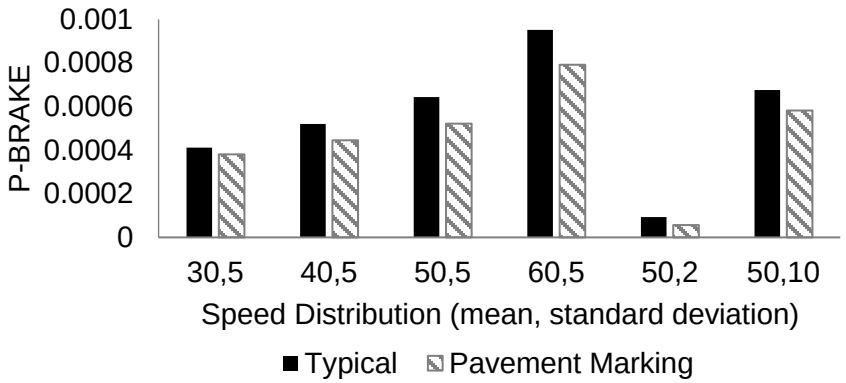
The pavement marking significantly increased the risk of RLR violations in the vehicles with a low operating speed or a high variance. The reason for this phenomenon was that the design speed for the pavement marking was 45 mph. If the vehicle's speed was lower than the design speed, there was a high risk of RLR violations. Thus, the pavement-marking countermeasure had positive effects on reducing rear-end crash risks. When the vehicles had low operating speed, the safety of the intersection was dramatically decreased (Figure 8.8).



(a)



(b)



(c)

Figure 8.6 - Comparative risky situations probabilities analysis of the typical intersection and of the intersection with pavement marking

(a) P-RS1, (b) P-RS2, and (c) P-BRAKE

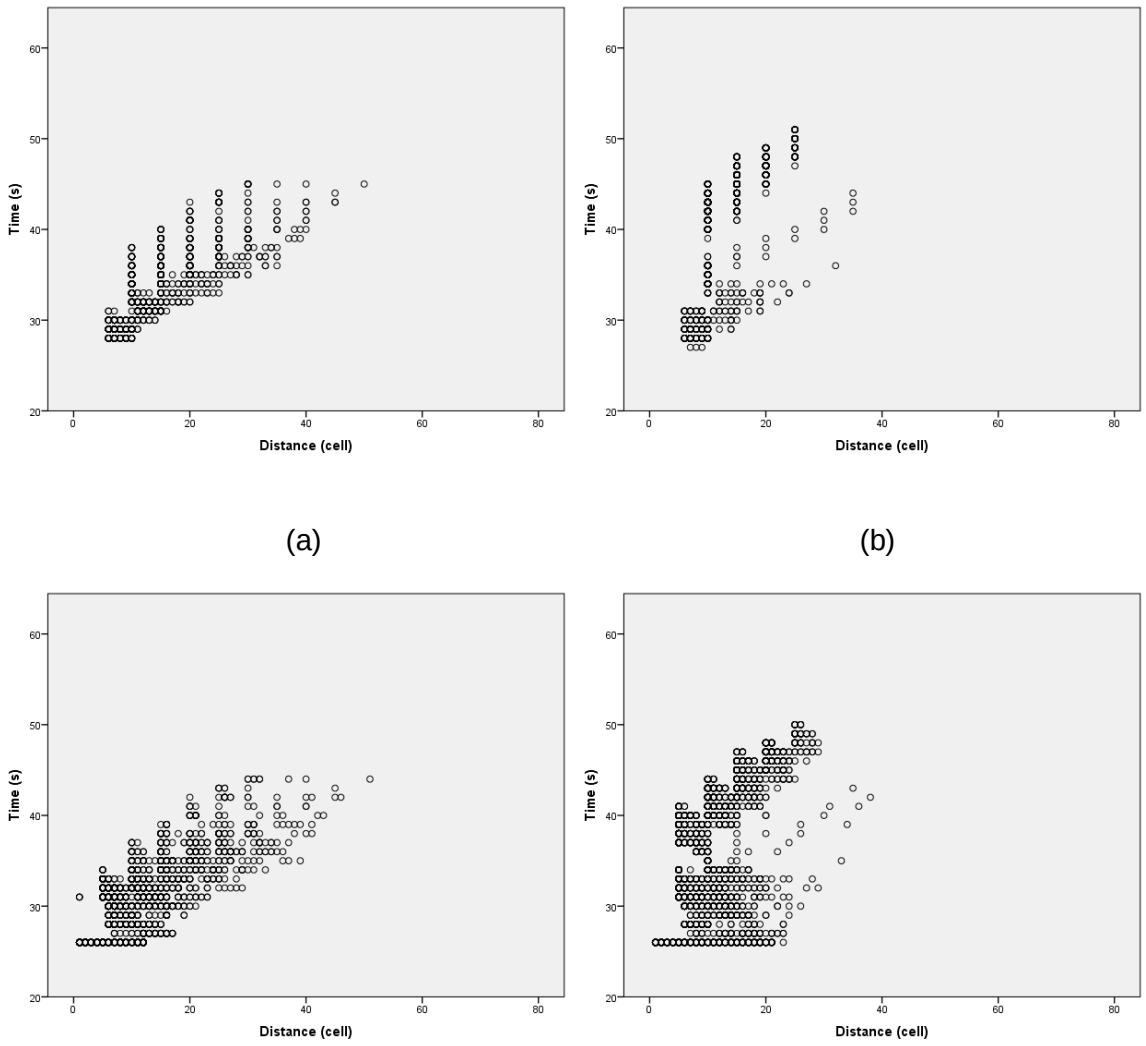


Figure 8.7 - Spatial and temporal distribution of risky situations at the intersection with the pavement marking and an auxiliary indication countermeasure (a) RS1 and RS2 when the expected mean speed was 30 mph, (b) RS1 and RS2 when the expected mean speed was 60 mph, (c) emergency brake situations when the mean speed was 30 mph, and (d) emergency brake situations when the expected mean speed was 60 mph (standard deviation = 5 mph)

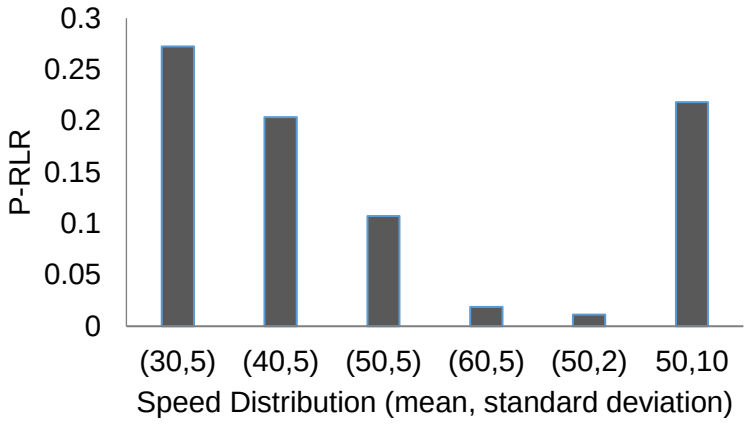


Figure 8.8 - Probability of RLR violation at the intersection with pavement marking

8.4 Intersection with Pavement Marking and an Auxiliary Indication Countermeasure

8.4.1 Scenario Construction

The flashing yellow signal beside the pavement marking was a warning signal, which was onset if the speed of the vehicle was below the speed limit. The flashing yellow signal began to flash a few seconds (about 1-3 s) before the onset of the yellow indication and continued until the end of the red interval. Given this disadvantage of pavement marking, a new countermeasure was proposed to solve this problem – install an auxiliary flashing yellow signal next to the pavement marking. Thus, we proposed a fourth scenario with both pavement marking and an auxiliary flashing yellow indication (PMAIC).

In Figure 8.9 (a), when vehicles were approaching the intersection with a speed lower than the speed limit of the intersection approach (45 mph), the auxiliary flashing yellow indication began flashing at *n* seconds before the yellow phase. If the vehicle had not passed the pavement marking before the onset of the auxiliary flashing yellow indication and could see the flashing indication, the driver should have chosen to stop during the yellow interval. Otherwise, the driver should have chosen to go at the yellow duration. In Figure 8.9 (b), when the operation speed of the vehicle was not below the

speed limit (45 mph), the driver should have followed the rules, which were similar to the pavement-marking scenario.

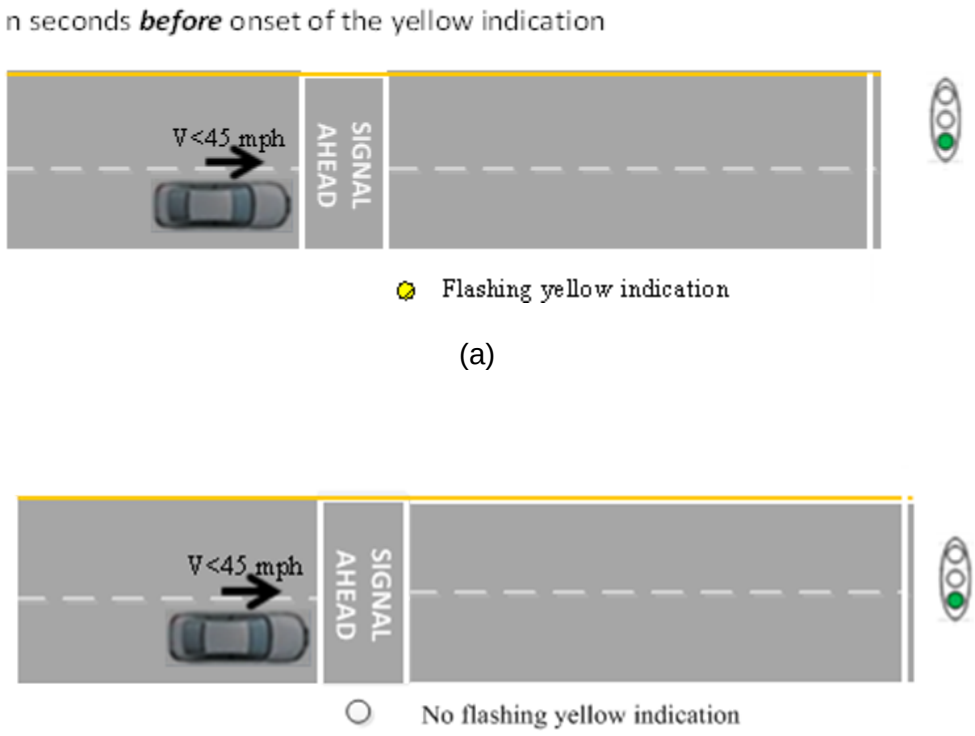


Figure 8.9 - Scenario of the PMAIC

(a) approaching the intersection with a speed below the speed limit and (b) approaching the intersection with a speed not below the speed limit

The processes of the PMAIC scenario simulation are shown in Figure 8.10. The value t_F was based on the vehicles' current speed at 5 seconds before the onset of the yellow indication.

$$V_c t_F = V_0 t_Y - V_c t_Y \quad (8.2)$$

then,

$$t_F = \frac{V_0 t_Y - V_c t_Y}{V_c} \quad (8.3)$$

where V_0 denotes the speed limit or 85th percentile speed, V_c denotes speed *n* seconds before the onset of the yellow indication, and t_Y denotes the yellow interval. The value for the judgment time *n* is

$$n = \frac{V_0 t_Y - V_{min} t_Y}{V_{min}} + 1 \quad (8.4)$$

where V_{min} denotes the expected minimum speed of vehicles. For example, if a vehicle approached the intersection at 30 mph (13.4 m/s), which was lower than the speed limit 45 mph (20 m/s), then 2.3 seconds before the yellow phase, the auxiliary flashing yellow indication next to the pavement marking began to flash.

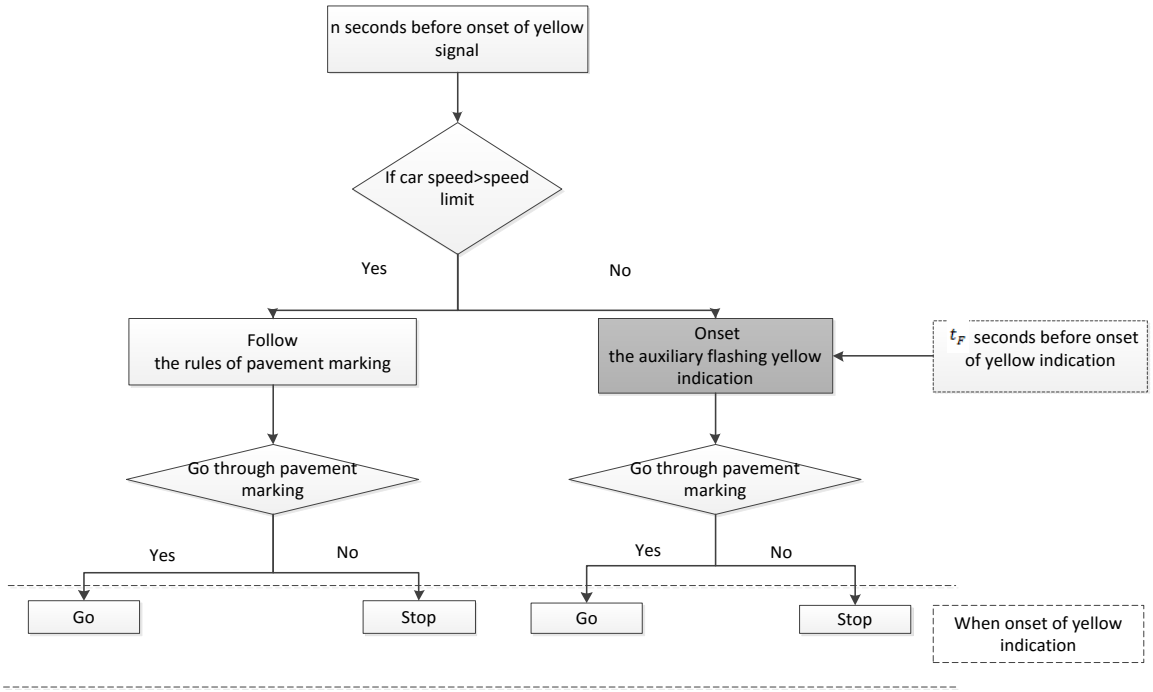
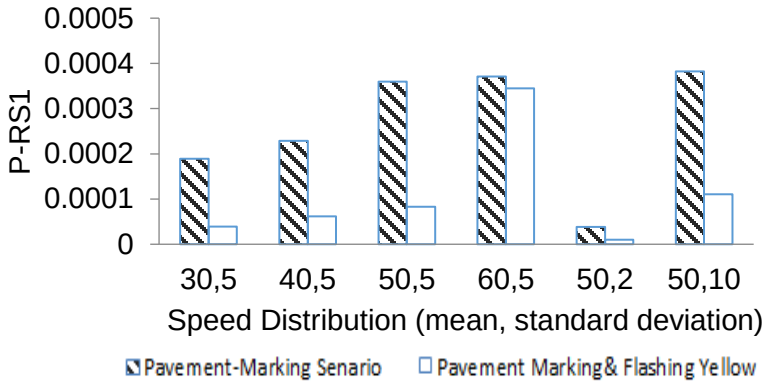


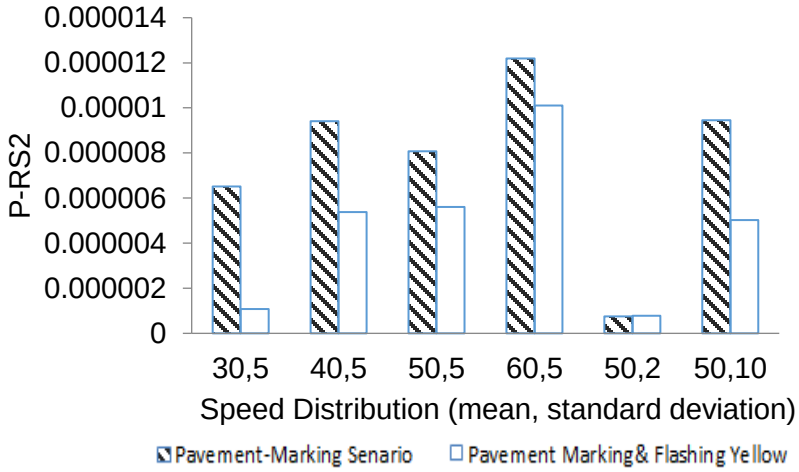
Figure 8.10 - Simulation processes of pavement marking and auxiliary flashing indication

8.4.2 Simulation Results

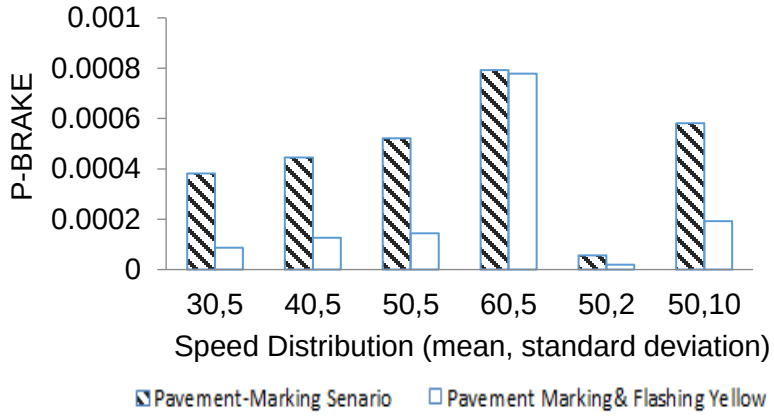
The PMAIC scenario simulation results are shown in Figure 8.11. Compared with the pavement marking scenarios, the PMAIC intersection had less rear-end risk situations. The decrease in probability of rear-end RS under the PMAIC became much less apparent when the average speed was above 45 mph because the cars followed the rules of pavement marking if speeds were larger than 45 mph. Like other scenarios, the probability of risky situations under the PMAIC scenario increased with the standard deviation of speed. However, compared with other scenarios, it also decreased the probability of rear-end crash risks under the situation of high variance operating speed.



(a)



(b)



(c)

Figure 8.11 - The probabilities of different types of rear-end RS under the pavement-marking scenario and the PMAIC scenario

(a)RS1, (b) RS2, and (c) BRAKE

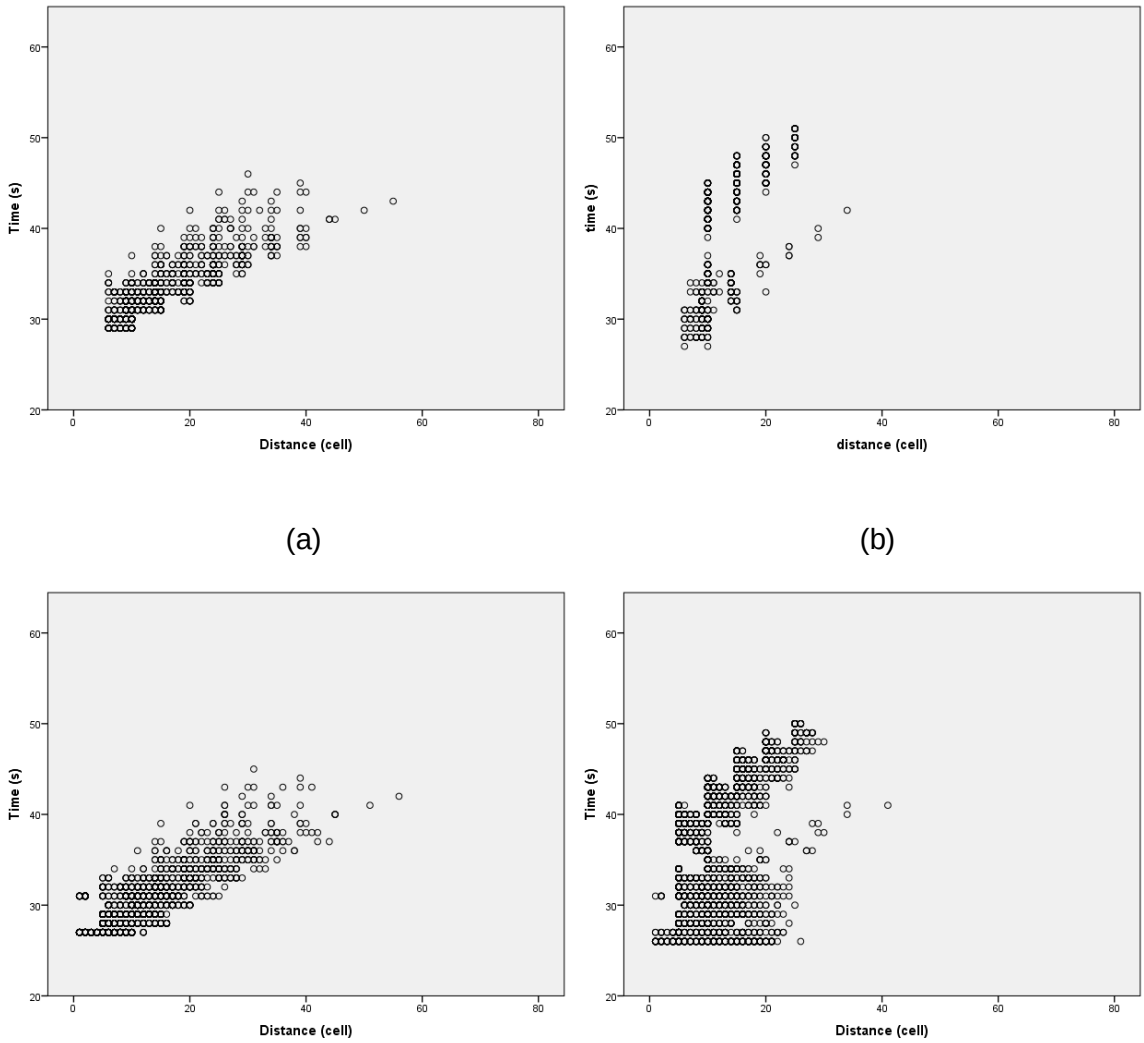


Figure 8.12 - Spatial and temporal distribution of risky situations at the intersection with the PMAIC

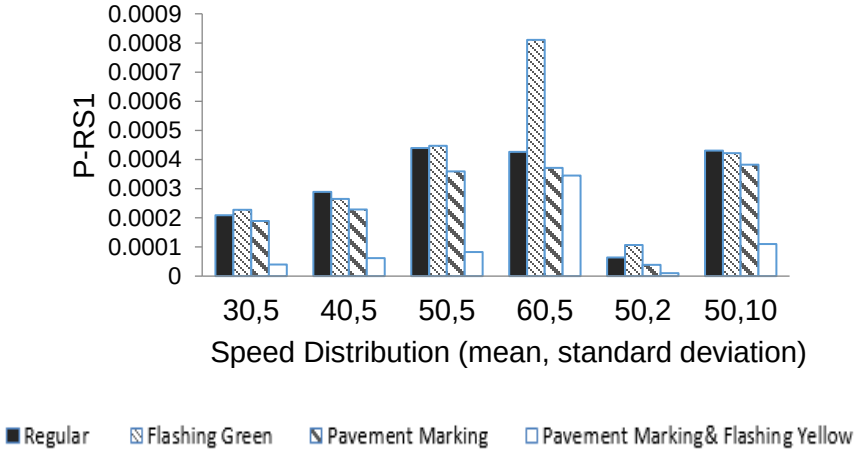
(a) RS1 and RS2 when the expected mean speed was 30 mph, (b) RS1 and RS2 when the expected mean speed was 60 mph, (c) emergency brake situations when the mean speed was 30 mph, and (d) emergency brake situations when the expected mean speed was 60 mph (standard deviation = 5mph)

Figure 8.12 demonstrates the risky situations with different expected mean speeds. The number of risky situations increased with the increase of expected mean speed. Similar to the pavement-marking countermeasure, the PMAIC also effectively prevented the increase of the distance range. In addition, during the new-countermeasure scenario simulation, rare RLR violations happened. Therefore, the PMAIC effectively reduced both the rear-end risky situations and the RLR violations.

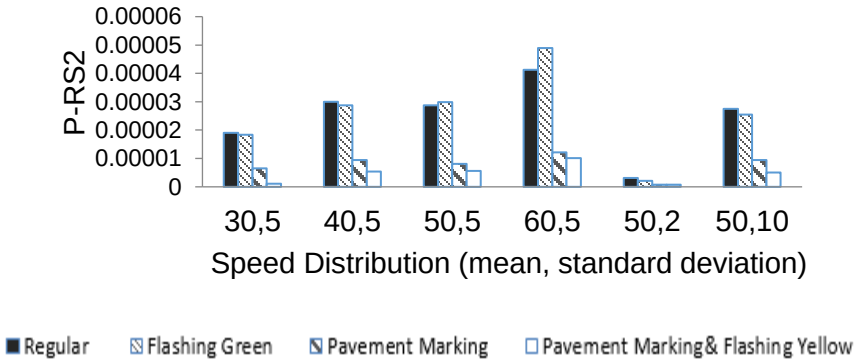
8.5 Comparative Analysis

8.5.1 Rear-End Risks

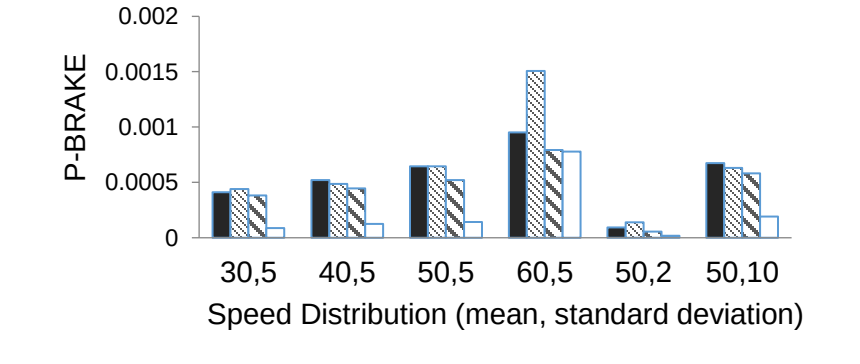
Figure 8.13 demonstrates the rear-end crash risks of the different intersection scenarios. A clear trend was found in that the mean speed or standard deviation can influence the BRAKE and RS2 risk probabilities. The RS1 risks for scenarios did not increase obviously when the speed rose from 50 mph to 60 mph. Low standard deviation showed a positive effect on safety improvement. In all of the four scenarios, the (50, 2) speed distribution simulations had the lowest risk probabilities. The results revealed that the main contributing causes to accident risk were high mean speed and high standard deviation of the speed distribution. Figure 8.13 also illustrates that the standard deviation did not have a significant impact on rear-end RS when greater than 5 mph.



(a)



(b)



(c)

Figure 8.13 - The probability of different kinds of rear-end RS under the scenarios of typical intersection, intersection with flashing green signals, intersection with pavement marking, and intersection with PMAIC

(a) RS1, (b) RS2, and (c) BRAKE

From the comparison, the results also showed that the flashing green countermeasure did not improve safety significantly, especially under the situations of a high mean speed or a low standard deviation of speed distribution. The rear-end risk probabilities of the flashing green countermeasure were even higher than the probabilities of the typical intersection. Distinction between the typical intersection scenario and the flashing green scenario was probably due to the increase of the indecision period when drivers behaved differently. Even though previous studies found that many drivers decelerated during the flashing green interval and somehow decreased the probability of rear-end crash risky situations, some of the flashing green intersections still suffered from high-risk probabilities for all three types of risk situations. Köll et al. (2004) found that the flashing green phases are associated with a substantial increase of early stop. However, it also produced larger indecision zones and led to longer periods of uncertainty, where following drivers cannot easily predict the front vehicles' stop/go decision (Factor et al., 2012). With respect to the pavement-marking scenario, it decreased the rear-end crash risks. All of the risk probabilities of the intersection with pavement marking were lower than the typical intersection and the intersection with flashing green phases, especially rear-end crashes caused by non-stopped cars, which meant the front vehicles in the crashes chose to cross the intersection during the yellow interval. The PMAIC effectively reduced the rear-end crash risk probabilities, especially at the low expected mean speed and the scenario with a high standard deviation.

Compared with the typical intersection scenario, the flashing green scenario had a larger range of the emergency brake risk at both time and distance distribution. Because of the difference of the driver behavior during the flashing green phases, there were more risky situations present at the flashing green interval (see Figure 8.14).

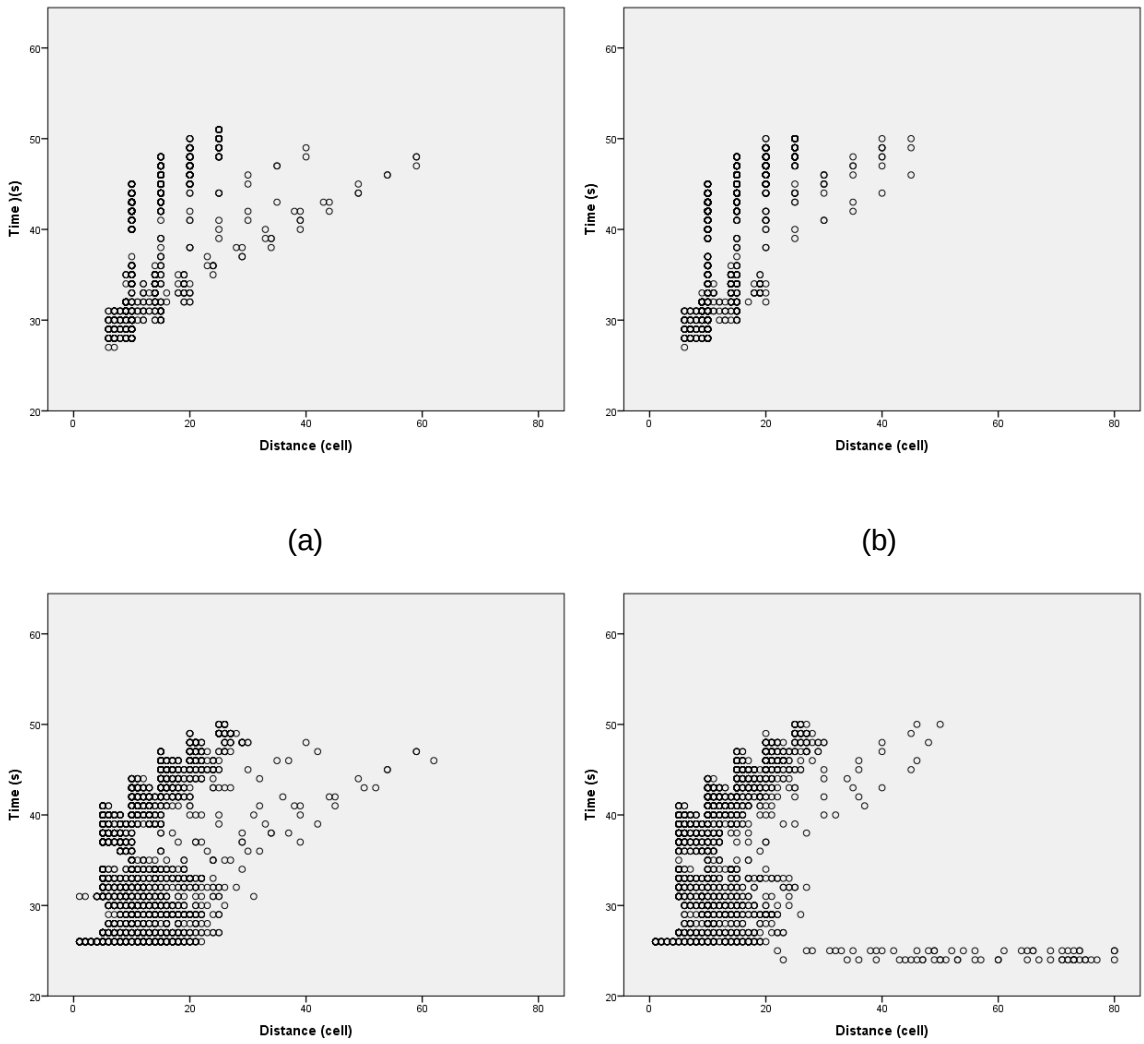


Figure 8.14 - Spatial and temporal distribution of risky situations when the mean expected speed was 60 mph

(a) RS1 and RS2 of the typical intersection scenario, (b) RS1 and RS2 of the flashing green scenario, (c) emergency brake situations of the typical intersection scenario, and (d) emergency brake situations of the flashing green scenario (standard deviation = 5 mph)

Compared with the spatial and temporal distribution of pavement marking, fewer risky situations were observed at the PMAIC scenario, especially during the yellow

interval and the beginning of the red interval. The results indicated the PMAIC effectively mitigated the problem of the dilemma zone (see Figure 8.15).

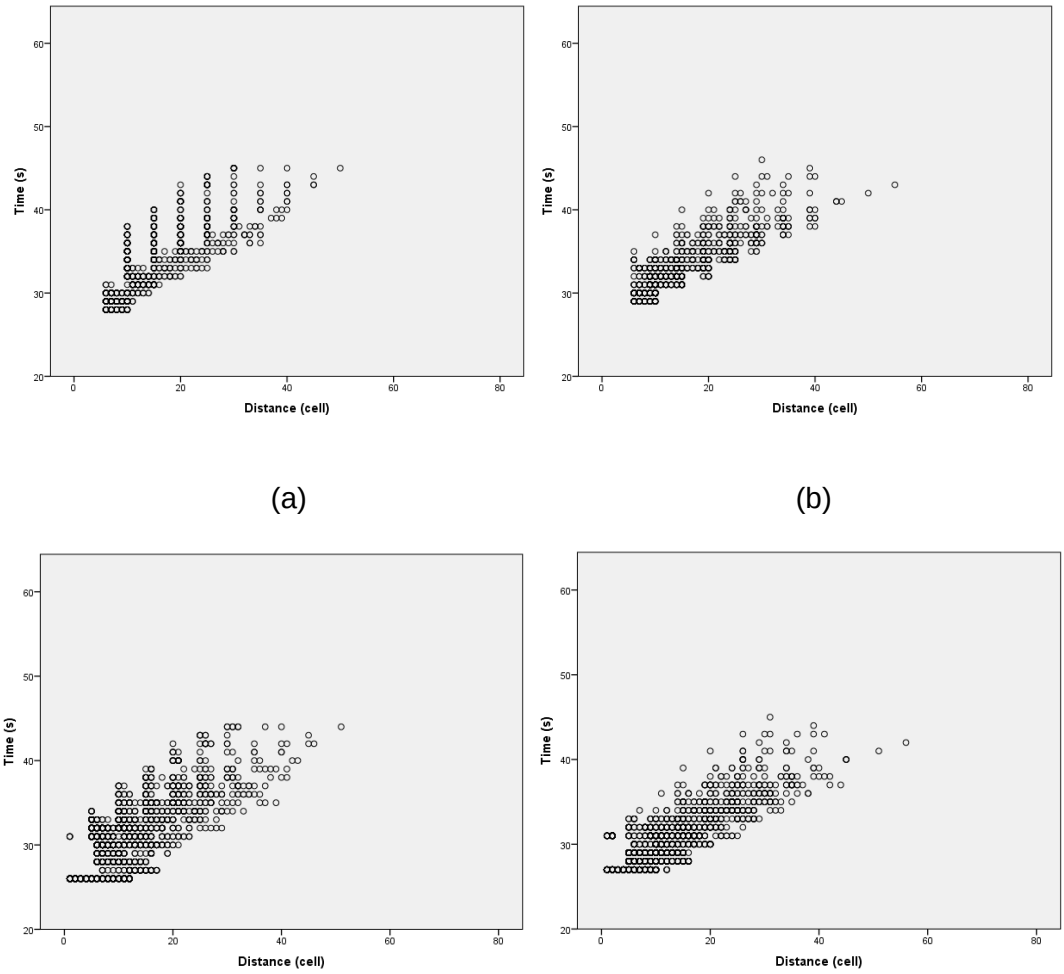


Figure 8.15 - Spatial and temporal distribution of risky situations when the mean expected speed was 30 mph

(a) RS1 and RS2 of the pavement marking scenario, (b) RS1 and RS2 of the PMAIC scenario, (c) emergency brake situations of the pavement marking scenario, and (d) emergency brake situations of the PMAIC scenario (standard deviation = 5 mph)

8.5.2 Red–Light Running Risk

At the yellow duration, drivers' false stop/go decisions can lead to red-light running (RLR) violations. During the simulation, the percentage of false decisions, or the potential for red-light running violations, was calculated. The results are shown in Figure 8.16. In the scenarios of typical intersection and intersection with flashing green signal, the comparison between different lead vehicles' speed distributions illustrated that mean speed was an important factor for the decision-making. The highest risk probabilities occurred at the 50 mph mean speed distribution. The standard deviation was also found to have little impact on the drivers' typical stop/go decision.

Previous studies pointed out that drivers would be more prone to stop at the intersection with the flashing green/yellow phases (Newton et al., 1997). This was because many of the drivers would decelerate during the flashing green phases, which would lead to lower speed at the onset of the yellow indication. These driver behaviors would increase the stop decisions according to the logistic regression model of stop/go decisions. However, drivers would still make decisions based on their own judgment, so the percentage of false go decisions would not decrease. The simulation results suggested that the flashing green phase measure could not effectively decrease the percentage of false decisions by drivers.

From the analysis of rear-end crash risk, it seemed that the pavement marking was an effective countermeasure to improve rear-end risk. However, Figure 8.16 demonstrates that the RLR violations were significant when the mean speed of the leading vehicles' speed distribution was lower than 50 mph or the standard deviation of the speed distribution was high. The disadvantage of the pavement marking was that if a driver encountered the yellow signal at a speed lower than the speed limit, even though s/he had passed the pavement marking, there was still a high chance for RLR and s/he could not execute the go maneuver safely. Such a negative situation might have resulted

in RLR due to the lower approaching speed. Therefore, the pavement marking effectively improved the intersection's safety only when the vehicles were approaching the intersection with high speed and low speed differences between vehicles. Otherwise, the pavement-marking countermeasure led to a high chance of RLR violations. Rare RLR violations happened during the simulation of the PMAIC scenario.

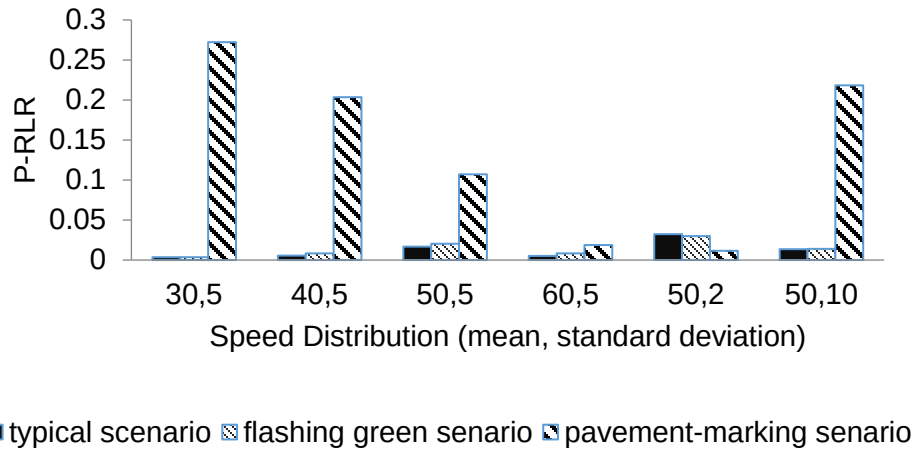


Figure 8.16 - Probability of RLR RS under different scenarios

8.6 Conclusions

Driver behavior during the yellow interval was influenced by the operating speed, the distance to the stop line, and the lead/follow position of the vehicle in a platoon. A logistic regression model was used to predict the stop/go decision of drivers as a function of distance to the stop line, the operating speed, and the lead/follow position. Most of the RLR violations were caused by drivers' false stop/go decisions during the yellow interval.

From the simulation results, the mean speed and the standard deviation played a significant role in rear-end crash risk situations, where a lower speed and lower standard

deviation led to less rear-end crash risk situations at the same intersection. High differences in speed were more prone to cause rear-end crashes.

According to the comparative analysis (Table 8.4), the flashing green countermeasure had little influence on rear-end crash risk reduction. The difference found between drivers' deceleration and acceleration decisions might be the major reason for the presence of the accident risk in the flashing green scenarios. Meanwhile, the pavement-marking countermeasure could effectively decrease the rear-end crash risk in most situations, especially the rear-end crashes caused by non-stopped cars, which meant the front vehicles in the crashes chose to cross the intersection during the yellow interval. The PMAIC, or the addition of an auxiliary flashing yellow indication next to the pavement marking, could further reduce the rear-end crash risks when the expected mean speed of the leading vehicles was relatively low.

Table 8.4 - Comparative analysis of different dilemma zone countermeasures

Risky Situations	Flashing Green Countermeasure	Pavement-Marking Countermeasure	PMAIC
Rear-end crash risk	Increased the risk of rear-end crash	Decreased the risk of rear-end crash	Decreased the risk of rear-end crash, especially when the expected mean speed was relatively lower
RLR violations	No significant impact on reducing RLR risk	Significantly increased the RLR with low expected mean speed	Effectively reduced the probabilities of RLR violations'

With respect to RLR violations, the RLR risk analysis showed that the mean speed of the leading vehicles had important influence on RLR risk in the typical intersection simulation scenarios as well as in the intersection with flashing green signal simulation scenarios. The results indicated that the flashing green phases could not reduce the percentage of false go decisions because the drivers made stop/go decisions

based on their own speed and position instead of other drivers' approaching speeds. The pavement marking could effectively reduce RLR risk situations when the vehicles were approaching the intersection with high speed and low speed differences with other vehicles. Otherwise, the intersection suffered from a high potential of RLR violations. The PMAIC had rare RLR violations. Therefore, PMAIC could effectively improve safety at signalized intersections.

Spatial and temporal analyses of the risky situations indicated that both the distance and the time range of the risky situations would increase with the increase of the operating speed of the vehicles. The pavement-marking countermeasure and the PMAIC could effectively prevent the increase of the distance range with the increase of the operating speed. Comparative analyses of different scenarios demonstrated the effectiveness of the PMAIC in reducing the probabilities of risky situations.

CHAPTER 9: Conclusions

The rapid development of ITS systems in the past few decades has catalyzed the implementation of Big Data in the transportation arena. To harness the power of Big Data for better traffic system performance, it is vital to take full advantage of its real-time nature. In this project, real-time traffic data from various sources, ranging from Automatic Vehicle Identification (AVI) systems and Microwave Vehicle Determination Systems (MVDS) to Video Image Processing (VIP), were implemented to explore their viability to improve the operation efficiency and traffic safety. To achieve the goal, multiple applications of the data were tested, including efficiency evaluation, safety analysis, microscopic simulation, and cellular automaton simulation.

Currently, three major types of automatic traffic detection technologies are widely used, namely in-roadway, over-roadway, and off-roadway detection technologies. In this project, real-time traffic data from point-based MVDS, video processing systems, and segment-based AVI system were collected on urban expressways and at a signalized intersection in the Orlando area. To ensure the validity of research findings, data quality were evaluated, especially for the MVDS and AVI systems. It was found that despite the two systems generating a large amount of data in real-time, the quality of the data was high enough (normally with a percentage of good data above 98%). Moreover, appropriate data processing helped identify the abnormal data to exclude in order to improve the data quality. The AVI system could provide segment speed and travel time information at an individual vehicle level. However, it could not capture the total traffic on expressways. In contrast, MVDS returns traffic information per lane aggregated at one-minute intervals. MVDS could be used to monitor traffic volume, speed, and occupancy at the detection locations. Video data provided other insights about driver behavior, such as stop/go decision red light running at intersections.

For traffic efficiency measurements, two major efforts were made. Real-time congestion measures were introduced based on the AVI system. The congestion measures evaluated congestion on the studied urban expressways from the perspective of travel time. With the real-time traffic data, congestion conditions were evaluated at 5-minute intervals, thus rendering detailed descriptions about the spatial-temporal distribution of congestion. It was found that congestion on urban expressways occurred mainly during morning and evening peak hours. Nevertheless, the segments experiencing congestion were location-specific. On segments with high traffic demand and near interchanges with other major corridors, congestion was more likely to occur.

Some previous studies have pointed out that it is the unexpected delays rather than the everyday delays that cause more frustration to motorists. Consequently, travel time reliability based on real-time traffic data was investigated. In the absence of large quantities of probe vehicles for the foreseeable future, the AVI system served as an ideal tool to evaluate travel time reliability for two reasons. One advantage was that it measured travel time directly. The other was that real-time evaluation was made possible because individual vehicles could be detected by AVI detectors. Statistical range measures, buffer indicators, and tardy trip indicators were all tested for performance in travel time reliability measurement. The conclusions from different measures agreed with each other. In addition, although unexpected delays were different from everyday congestion, it was found that the congested segments had lower travel time reliability compared with other segments whose traffic was closer to free-flow conditions. Improving recurrent congestion was also likely to have positive effects on travel time reliability improvement.

In addition to traffic efficiency, traffic safety is another critical aspect of highway performance. In this project, real-time safety evaluations for both the mainline and ramps were conducted based on the MVDS traffic data. For ramp safety analyses, real-time

information from nearby airports was incorporated into the evaluation. For mainline safety evaluations, crash precursors for total crashes were identified. It turned out that peak hours had a major impact on crash occurrence. Real-time traffic volume per lane, speed differences between inner and outer lanes, congestion index, and truck percentage were significant traffic variables affecting crash occurrence. Geometric characteristics such as number of lanes, median width, and shoulder width were also found to be significantly tied with mainline crashes. For ramp crashes, in addition to the traffic parameters (volume, speed, and truck percentage), ramp configuration, real-time weather conditions, and road surface conditions also had an impact on crash occurrence in real-time. According to the safety analyses, both mainline and ramp crash prediction models achieved good estimation accuracy. Consequently, the automatic traffic detection data showed great potential for the construction of a more proactive traffic safety management system.

Microscopic simulation is a cost-effective way for traffic researchers to evaluate traffic operation on existing or proposed facilities. However, the performance of simulation heavily depends on the input traffic data. In this project, simulation of traffic under foggy conditions using field traffic data was tested in the VISSIM environment. The surrogate safety assessment model (SSAM) for rear-end and lane-change conflicts under reduced visibility was estimated via simulation. To achieve the objective of this task, fog data and traffic detection data at a specific location on I-4 were collected. In the simulation, real traffic flow observations corresponding to fog conditions were input to the system. By using real-time traffic data, validation and calibration of the microscopic simulation model could be significantly enhanced by providing traffic data at short-time intervals. Meanwhile, the traffic flow corresponding to a specific condition (adverse weather, or fog) could be captured, thus the evaluation was closer to reality. The simulation results indicated that the total number of conflicts, lane-change conflicts, and

rear-end conflicts increased along with the traffic volume and speed limit. Given foggy conditions decrease the visible distance for motorists, a lower speed limit was recommended to ensure that vehicles could slow down in time to avoid potential collisions. In addition, if traffic volume was high, reducing speed limit alone in the face of fog would have only limited effects on traffic flow. In such cases, other countermeasures such as warning messages, detour strategies, or even closure of roadway facilities might be needed.

The above efforts were focused on freeway/expressway safety and operation. In urban areas, safety conditions at signalized intersections also call for attention. Drivers' stop/go decisions in dilemma zones at intersections could significantly affect the safety level. Traditional traffic models for the dilemma zone could be cumbersome in terms of time consumption and complexity of the modeling process. In contrast, Cellular Automaton (CA) models are simulation based models that could more efficiently deal with the dilemma zone issue. Nevertheless, studies seldom take advantage of real-time traffic data. In this project, real-time traffic data at intersections were extracted from video processing systems to generate the rule of drivers' stop/go decisions. Speed variables, distance variables, leading or following position, and some other factors from the video data were prepared. Based on the data, rear-end risks and red-light-running risks were evaluated. From the simulation results, the mean speed and the standard deviation played a significant role in rear-end crash risk situations. Appropriate countermeasures to cope with dilemma zone problems were tested and identified (e.g., pavement marking, flashing yellow, and the combination of the two). Pavement-marking countermeasures could effectively decrease the rear-end crash risk in most situations, especially the rear-end crashes caused by non-stopped cars. The PMAIC, or the addition of an auxiliary flashing yellow indication next to the pavement marking, could further reduce the rear-end crash risks when the expected mean speed of the leading

vehicles was relatively low. For RLR violations, simulation scenarios reflected that mean speed of the leading vehicles had an important influence on RLR risks. The pavement marking could also effectively reduce RLR risk situations when the vehicles were approaching the intersection with high speed and low speed differences with other vehicles. On the other hand, the PMAIC had rare RLR violations. Therefore, PMAIC could effectively improve safety at signalized intersections.

In this project, different applications of microscopic big traffic data were evaluated. It was confirmed through evaluation that these data could efficiently aid professionals' understanding in traffic operation and safety and help traffic authorities to come up with more proactive traffic management strategies. Furthermore, different simulation tests also proved that real-time traffic data helped to better reflect the traffic conditions in a specific simulated environment in which more accurate countermeasures could be suggested. In summary, microscopic big traffic data have great potential in future simulation-based traffic studies.

References

Abdel-Aty, M., Oloufa, A., Peng, Y., Shen, T.-W., Yang, X., Lee, J., 2014. Real time monitoring and prediction of reduced visibility events on florida's highways. Florida Department of Transportation.

Abdel-Aty, M., Uddin, N., Pande, A., 2005. Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways. Transportation research record: journal of the transportation research board 1908 (1), 51-58.

Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, F.M., Hsia, L., 2004. Predicting freeway crashes from loop detector data by matched case-control logistic regression. Transportation Research Record: Journal of the Transportation Research Board 1897 (1), 88-95.

Abdel-Aty, M.A., Hassan, H.M., Ahmed, M., Al-Ghamdi, A.S., 2012. Real-time prediction of visibility related crashes. Transportation research part C: emerging technologies 24, 288-298.

Ahmed, M.M., Abdel-Aty, M., Lee, J., Yu, R., 2013. Exploring the feasibility of using airport data in real-time risk assessment. In: Proceedings of the Transportation Research Board 92nd Annual Meeting.

Ahmed, M.M., Abdel-Aty, M., Yu, R., 2012. Assessment of interaction of crash occurrence, mountainous freeway geometry, real-time weather, and traffic data. Transportation Research Record: Journal of the Transportation Research Board 2280 (1), 51-59.

Ahmed, M.M., Abdel-Aty, M.A., 2012. The viability of using automatic vehicle identification data for real-time crash prediction. Intelligent Transportation Systems, IEEE Transactions on 13 (2), 459-468.

Antoniou, C., Balakrishna, R., Koutsopoulos, H.N., 2011. A synthesis of emerging data collection technologies and their impact on traffic management applications. European Transport Research Review 3 (3), 139-148.

Bonneson, J.A., Zimmerman, K., Brewer, M.A., 2002. Engineering countermeasures to reduce red-light-running. Texas Transportation Institute, Texas A & M University System.

CFX, 2014. Expressway map. Central Florida Expressway Authority.

Chang, M.-S., Messer, C.J., Santiago, A.J., 1985. Timing traffic signal change intervals based on driver behavior.

Dias, C., Miska, M., Kuwahara, M., Warita, H., Year. Relationship between congestion and traffic accidents on expressways: An investigation with bayesian belief networks. In: Proceedings of the Proceedings of 40th Annual Meeting of Infrastructure Planning (JSCE), Japan.

Ding Y., Wu Y., Abdel-Aty M. (2014). Studying rear-end crash risk under countermeasures of amber dilemma at signalized intersections based on cellular automaton model. Manuscript under review in Accident Analysis & Prevention.

Elmitiny, N., Yan, X., Radwan, E., Russo, C., Nashar, D., 2010. Classification analysis of driver's stop/go decision and red-light running violation. Accident Analysis & Prevention 42 (1), 101-111.

Factor, R., Prashker, J.N., Mahalel, D., 2012. The flashing green light paradox. Transportation research part F: traffic psychology and behaviour 15 (3), 279-288.

Federal Highway Administration, 2014. How do weather events impact roads? Federal Highway Administration.

Gates, T.J., Noyce, D.A., Laracuente, L., Year. Analysis of dilemma zone driver behavior at signalized intersections. In: Proceedings of the Transportation Research Board 86th Annual Meeting.

Gerlough, D.L., Huber, M.J., 1975. Traffic flow theory: A monograph Transportation Research Board, National Research Council.

Gettman, D., Pu, L., Sayed, T., Shelby, S.G., 2008. Surrogate safety assessment model and validation: Final report.

Goodwin, L., 2002. Weather impacts on arterial traffic flow. Mitretek systems inc.

Grant, M., Bowen, B., Day, M., Winick, R., Bauer, J., Chavis, A., Trainor, S., 2011. Congestion management process: A guidebook.

Griffin, L.A., Year. Enhancing expressway operations using travel time performance data. In: Proceedings of the Facilities Management and Maintenance Workshop, Nashville, TN.

Hall, F.L., 1996. Traffic stream characteristics. Traffic Flow Theory. US Federal Highway Administration.

Hamad, K., Kikuchi, S., 2002. Developing a measure of traffic congestion: Fuzzy inference approach. Transportation Research Record: Journal of the Transportation Research Board 1802 (1), 77-85.

Hossain, M., Muromachi, Y., 2012. A bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. Accident Analysis & Prevention 45, 373-381.

Hourdos, J.N., Garg, V., Michalopoulos, P.G., Davis, G.A., 2006. Real-time detection of crash-prone conditions at freeway high-crash locations. Transportation Research Record: Journal of the Transportation Research Board 1968 (1), 83-91.

Hurwitz, D.S., 2009. Application of driver behavior and comprehension to dilemma zone definition and evaluation. Open Access Dissertations, 112.

Institute of Transportation Engineers, 1989. Determining vehicle change intervals: A proposed recommended practice Institute of Transportation Engineers, Washington D.C.

Jia, B., Gao, Z.-Y., Li, K., Li, X., 2007. Models and simulations of traffic system based on the theory of cellular automaton. Science, Beijing.

Khan, A.M., 2007. Intelligent infrastructure-based queue-end warning system for avoiding rear impacts. Intelligent Transport Systems, IET 1 (2), 138-143.

Klein, L.A., Mills, M.K., Gibson, D.R., 2006. Traffic detector handbook: -volume ii.

Köll, H., Bader, M., Axhausen, K., 2004. Driver behaviour during flashing green before amber: A comparative study. Accident Analysis & Prevention 36 (2), 273-280.

Lomax, T., Schrank, D., Turner, S., Margiotta, R., 2003. Selecting travel reliability measures. Texas Transportation Institute monograph (May 2003).

Martchouk, M., 2009. Travel time reliability Purdue University.

Martin, P.T., Feng, Y., Wang, X., 2003. Detector technology evaluation. Mountain-Plains Consortium.

Mimbela, L.E.Y., Klein, L.A., 2000. Summary of vehicle detection and surveillance technologies used in intelligent transportation systems.

Newton, C., Mussa, R.N., Sadalla, E.K., Burns, E.K., Matthias, J., 1997. Evaluation of an alternative traffic light change anticipation system. Accident Analysis & Prevention 29 (2), 201-209.

Pande, A., Abdel-Aty, M., 2006. Assessment of freeway traffic parameters leading to lane-change related collisions. Accident Analysis & Prevention 38 (5), 936-948.

Pande, A., Das, A., Abdel-Aty, M., Hassan, H., 2011. Estimation of real-time crash risk. Transportation Research Record: Journal of the Transportation Research Board 2237 (1), 60-66.

Papaioannou, P., 2007. Driver behaviour, dilemma zone and safety effects at urban signalised intersections in greece. Accident Analysis & Prevention 39 (1), 147-158.

Pline, J.L., 1999. Institute of transportation engineers. Traffic engineering handbook.

Pu, L., Joshi, R., 2008. Surrogate safety assessment model (ssam): Software user manual.

Texas Transportation Institute, 2006. Travel Time Reliability: Making it there on time, all the time. Federal Highway Administration, USA.

Riley, J.D., 1999. Evaluation of travel time estimates derived from automatic vehicle identification tags in san antonio, tx. Virginia Polytechnic Institute and State University.

Schrank, D., Year. The keys to estimating mobility in urban areas. In: Proceedings of the ITE 2001 Annual Meeting and Exhibit.

Cambridge Systematics Inc., 2005. Traffic congestion and reliability: Trends and advanced strategies for congestion mitigation. Federal Highway Administration.

Wortman, R.H., Matthias, J.S., 1983. Evaluation of driver behavior at signalized intersections (discussion).

Yan, X., Radwan, E., Abdel-Aty, M., 2005. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. Accident Analysis & Prevention 37 (6), 983-995.

Yu, R., Abdel-Aty, M., 2013. Multi-level bayesian analyses for single-and multi-vehicle freeway crashes. Accident Analysis & Prevention 58, 97-105.

Yu, R., Shi, Q., Abdel-Aty, M., 2013. Feasibility of incorporating reliability analysis in traffic safety investigation. Transportation Research Record: Journal of the Transportation Research Board 2386 (1), 35-41.

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