Summer 7-2013

Assessing the Impact of Speed Limit Reduction near Signalized High Speed Intersections Equipped with Advance Warning Flashers: A Case Study in Nebraska

Shefang Wang
University of Nebraska-Lincoln, sfwang2@gmail.com

ASSESSING THE IMPACT OF SPEED LIMIT REDUCTION NEAR SIGNALIZED HIGH SPEED INTERSECTIONS EQUIPPED WITH ADVANCE WARNING FLASHERS: A CASE STUDY IN NEBRASKA

By

Shefang Wang

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Civil Engineering

Under the Supervision of Professor Anuj Sharma

Lincoln, Nebraska

July, 2013

ASSESSING THE IMPACT OF SPEED LIMIT REDUCTION NEAR SIGNALIZED HIGH SPEED INTERSECTIONS EQUIPPED WITH ADVANCE WARNING FLASHERS: A CASE STUDY IN NEBRASKA

Shefang Wang, M.S.

University of Nebraska, 2013

Adviser: Anuj Sharma

This thesis evaluated the effects of 5 mph and 10 mph speed limit reductions in the vicinity of high-speed, signalized intersections equipped with Advance Warning Flashers (AWF). The selected methodology involved a field study of the impact of speed limit reduction at 7 high-speed, signalized intersections with AWF, using quantile regression models developed for speed. The quantile regression models for speed indicated that reduction of the speed limit from 60 mph to 55 mph did not have significant impact on observed speed during the green time. However, it was found that speed limit reduction from 65 mph to 55 mph led to statistically significant reductions in observed speed during the green period. The conclusions of this study, however, were limited by the low number of intersections where speed limits were reduced. Only two intersections with 10 mph reductions were available for observation where speed limit was reduced from 65 mph to 55 mph. Based on the available dataset, for a highway with a speed limit of 65 mph, a reduction to 55 mph at intersections equipped with AWF was found to be statistically significant in terms of reducing speeds over all speed percentiles during the green time. It is recommended that future research include other speed limit combinations, such as 5 mph reductions from 65 mph to 60 mph, and utilize larger

datasets to provide for improved generalizability and transferability of results. A before-and-after study could also provide partially controlled conditions to isolate the impacts of speed limit reduction.

ACKNOWLEDGMENTS

To my parents Zhongcheng Wang and Jianhua Pan.

To my friends Amber and Damon.

To Dr. Anuj Sharma, Dr. Laurence Rilett, and Dr. Aemal Khattak.

To all Nebraska Transportation Center fellows.

Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Background

According to the National Highway Traffic Safety Administration's (2010) FARS Data Tables (1), 12,504 (28%) of 44,713 vehicles were involved in fatal crashes took place at intersections and intersection-related locations in 2010. From an economic standpoint, the total cost of vehicle crashes (estimated for reported and unreported crashes) was U.S. $230.6 billion, of which intersection collisions accounted for about 30% (2). Focusing on the state of Nebraska, 385 traffic fatalities occurred at Nebraska intersections between 2006 and 2010, among which 249 (65%) fatalities were identified as occurring in rural areas (3). Safety concerns involving signalized intersections are critical for rural and rural highways, since high-speed aggravates the severity of crashes.

The Nebraska Department of Roads (NDOR) has at least partial maintenance responsibility for 224 vehicle signals. However, following an exhaustive literature search and numerous personnel contacts, it was found that there existed no consistent documented policy for assigning speed limits on highways in the vicinity of traffic signals. This inconsistency can also be seen in different speed limit reduction schemes. For example, on certain sections of Highway 77, the speed limit decreases from 65 mph to 55 mph in the vicinity of a signalized intersection; however, on Highway 34, northwest of Lincoln, the speed limit is 60 mph and there is no speed limit reduction at the intersection of Highway 34 and Highway 79. The operational impacts of speed limit reduction in the vicinity of high speed intersections have not been adequately studied for Nebraska-specific conditions.

A compounding issue in Nebraska is that until July 2009, 35 intersections on the Nebraska state highway system were equipped with AWF and a dilemma zone protection algorithm(4); AWF—a safety-enhancing device—could potentially detract from the safety benefits provided by an additional reduced speed limit sign.

1.2 Problem Statement

The Manual of Uniform Control Devices (MUCTD) states that "Advance warning signs and other traffic control devices to attract the motorist's attention to a signalized intersection are usually more effective than a reduced speed limit zone" (5). However, MUCTD is silent regarding recommendations for speed limit reduction in conjunction with AWF. For the past several years, NDOR has used AWF at high-speed rural intersections. The speed limit may or may not be reduced at such intersections, and the decision is made based upon case-by-case engineering judgments.

The current research aimed to verify the effectiveness of speed limit reduction at rural, high-speed intersections equipped with the NDOR AWF system. This was done by conducting a case study to examine the operational impacts of speed limit reduction during the green period.

Speed limits are reduced in the vicinity of high-speed intersections with the expectation of enhancing safety. Speed limit reduction can enhance road user safety in two ways: a) through a limiting function; b) through a coordinating function (6). The limiting function consists of setting a speed limit along the road, which forces drivers to reduce their speeds, thereby reducing the probability and severity of collisions (7). With the coordinating function, a speed limit can reduce the variance of speeds along the road, which can make speed more uniform, thereby increasing road safety (7). For example, suppose the speed limit in transition zone (two roadway segments with different speed limits) is reduced in the vicinity of a high-speed intersection; one possible consequence is the separation of drivers into two subsets consisting of a) those who drive accordingly and at lower speeds; b) those who choose their own speeds, which are likely higher than the reduced limit. The resulting variance in driving speeds could be a potential trap for highway safety.

In this thesis, a case-control study was performed by observing the impacts of 5 mph reduction and 10 mph reduction on near-intersection versus away from the intersection speed. These data were compared to changes in near and far speed distribution at intersections lacking any speed limit reduction. Compounding factors were controlled for using the quantile regression models to account for changes based on any factor other than the factor of interest, i.e., the presence of a speed limit reduction sign. Table 1-1 lists the hypotheses that were tested based on the study design.

Table 1-1 Hypotheses Tested (during the green time)

Models based on Speed Limit Reduction Group	Null Hypothesis	Alternative Hypothesis
0 mph	Speed remains constant when vehicles approach an intersection at the intersection with constant speed limit.	A significant change in speed occurs when vehicles approach an intersection with constant speed limit.
5 mph	Speed remains constant when vehicles approach an intersection with 5 mph speed limit reduction.	A significant change in speed occurs when vehicles approach an intersection with 5 mph speed limit reduction.
10 mph	Speed remains constant when vehicles approach an intersection with 10 mph speed limit reduction.	A significant change in speed occurs when vehicles approach an intersection with 10 mph speed limit reduction.

1.3 Thesis Organization

This thesis contains 6 chapters. Chapter 1 introduces the problem and the objectives of the current study. Chapter 2 reviews literature related to the topic, including the effects of speed limit changes, safety factors, and speed analyses. Chapter 3 describes the data collection plan, including data collection equipment and site selection. Chapter 4 introduces the conducted sensor validation and data processing using the machine learning technique. Chapter 5 provides a comprehensive data analysis based on the processed dataset. Finally, chapter 6 summarizes the findings of the research.

CHAPTER 2 LITERATURE REVIEW

2.1 Standards of Speed Limit

Speed limit is used primarily to enhance road safety, which can be achieved through either limiting or coordinating functions (6). The limiting function establishes a maximum speed along roads, which can reduce the likelihood of a crash and the severity of accidents. The coordinating function reduces speed variance along roads, which results in a more uniform traffic flow.

Two statutory national speed limits have been imposed throughout U.S. history. The first national speed limit was established during World War II, and was 35 mph. The second national speed limit was the National Maximum Speed Limit (NMSL), which allowed a maximum speed of 55 mph. Both were aimed at reducing energy consumption rather than saving on transportation costs (6). The NMSL has been amended several times. In 1974, it remained at 55 mph. In 1987, it increased to 65 mph on some qualified sections of interstate highways as mandated by Congress. Most recently, in 1995, the NMSL was repealed, allowing states to set their own speed limits. Nearly all states increased their speed limits at that time (6).

According to the MUTCD (2009), speed zone refers to “a section of highway with a speed limit that is established by law or regulation, but which might be different from a legislatively specified statutory speed limit” (5). The appropriate speed limit within speed zones is the maximum (or minimum) speed determined on the basis of specific road conditions. The posted speed limit is recommended to be within 5 mph of the 85th percentile distribution of roadway speeds (5). AASHTO’s “A Policy on Geometric

Design of Highways and Streets" mentions that traffic engineering studies on posted speed limits should coincide with prevailing conditions along the road, and should be capable of reasonable enforcement (8).

2.2 Studies of Driver Compliance

Many previous studies have examined the effectiveness of speed limit changes. In 1997, a study conducted by the Federal Highway Administration (FHWA) regarding the effects of raising and lowering speed limits reported that changing the speed limit had little effect on driver behavior (9). In that study, the speed limit was raised by between 0 and15 mph at some locations, while at control locations it was lowered by 5 to 20 mph. The before-after analysis showed that the resulting differences in mean and 85th percentile speeds were generally less than 2 mph.

In 2007, Kentucky enacted a law permitting the increase of the speed limit from 65 mph to 70 mph for specific sections of roadway. A before-after analysis found that the speed limit change resulted in only a small change in actual travel speeds. On rural interstates, the 85th percentile speed was 1.3 mph faster for passenger cars and 0.6 mph faster for trucks. As for the 85th percentile speed along rural four-lane parkways, car speeds increased by 2.0 mph and truck speeds increased by 1.2 mph (10, 11).

Similarly, in 2004 Virginia passed new legislation to raise the statutory maximum speed limit from 55 mph to 65 mph on limited access primary roads. A before-and-after study concluded that average speed increased by only 1.7-4.3 mph at all test sites. However, speed limit violation decreased from over 80% to approximately 50%. Also, variance in traffic speeds remained fairly constant (12). The consistent conclusion drawn from these studies is that the arbitrary determination of speed limits without accounting for driver tendencies has a limited effect on observed speeds.

Variable Speed Limit (VSL) has been applied to improve roadway safety under different conditions such as severe weather, unexpected changes in roadway geometrics, and traffic congestion (13, 14). VSL provides a changeable posted speed limit coinciding with changes in the characteristics of the speed zone.

Buddenmeyer et al. (13) conducted research concerning VSL along a section of I-80 in Wyoming. The major goal of the project was to reduce speed variability along the corridor and improve safety under adverse weather conditions. The dataset was collected by Wavetronix SmartSensorHD, and included traffic volume, vehicle speed, average speed, 85th percentile speed, average headway and gap, lane occupancy, and vehicle classification. The modeling results found that surface status, subsurface temperature, wind speed, dewpoint, and visibility were the most consistently significant variables to impact observed speeds. The final results indicated that VSL signs caused an actual speed reduction of 0.47 to 0.75 mph for every mile per hour in posted speed reduction.

Dynamic Message Signs (DMS) can provide drivers with direct messages regarding the detected speeds of approaching vehicles. Monsere et al. (15) studied the advanced curve warning DMS system, demonstrating its strong performance in speed reduction within the speed transition zone. In this project, speed limit was dropped from 65 mph to 45 mph prior on a curved section of road. The DMS system's effectiveness at reducing mean speed was examined in a before-after study, and demonstrated statistically significant results. Moreover, in an attitude survey, most drivers exhibited positive responses to the system.

Table 2-1 displays a summary of the literature discussed above, illustrating that changes in actual speed in these studies were found to be significantly smaller than changes in posted speed limits. Figure 2-1 illustrates this comparison, where the x-axis represents the change in the posted speed limit sign, and the y-axis represents the change in actual speed.

Table 2-1 Previous Research

Location	Before Speed Limit	After Speed Limit	Change	Mean Change	Change in 85^{th} Percentile Speed
Campbell County, KY(9)	55	45	-10	NA	-0.9
Franklin County, KY(9)	55	45	-10	NA	-3.8
Graves County, KY(9)	55	45	-10	NA	-0.8
Boone County, KY(9)	35	45	+10	NA	1.4
Rural Interstates, KY(10)	65	70	+5	NA	1.3 (PC) 0.6 (Trucks)
Four-lane parkways, KY(10)	65	70	+5	NA	2.0 (PC) 1.2 (Trucks)
Virginia (12)	55	65	+10	1.7~4.3	NA
I-5 SB, Douglas ,OR(15)	65 (PC), 55(Truck)	45	-20 (PC), -10(truck)	-3	NA
I-5 NB, Douglas, OR(15)	65 (PC), 55(Truck)	45	-20 (PC), -10(truck)	-2	NA

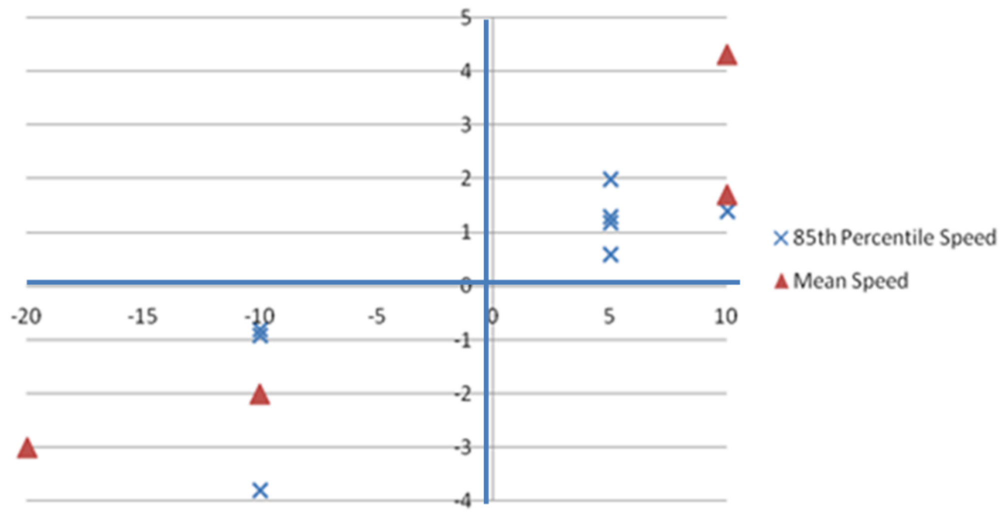


Figure 2-1 Comparison of Speed Limit Change and Actual Change

2.3 Speed Analysis

The two major speed limit related factors usually involved in an accident are average speed and speed variance (7). Kinetic energy increases with the square of speed; thus, higher speeds produce greater impact forces in crashes, and a higher chance of causing serious injury. If the speed limit is changed along a roadway section, it follows that the average speed at impact in traffic accidents might also change. The second possible effect of speed limit change is the disturbance of speed variance. If all vehicles are traveling at the same speed, there will be no overtaking, braking, or lane changing, which reduces the probability that accidents will occur. Such phenomenon has already been studied by Solomon (16), who noted that the relationship between accidents and travel speed was a U-shaped curve. This curve illustrates that higher variations around the average travel speed (both above and below) are associated with a greater chance of being involved in an accident.

2.3.1 Studies of Crashes and Safety

One common misconception regarding speed limits is that lowering the speed limit will increase road user safety and reduce crashes rates, and vice versa (6). Researchers have indicated that, in actuality, variance in speed poses a threat to safety. As an FHWA publication states, "the potential of being involved in a crash is highest when traveling at a speed much lower or much higher than the majority of motorists" (9). The U-shaped relationship between motorist speeds and the chance of being in a crash invalidates the notion that lowering speed limits will necessarily increase safety (16).

2.3.2 Advisory Speed for Transition Speed Zone

Special road characteristics, such as high-speed intersections, may favor an advisory speed limit different from, and probably lower than, that of other highway segments. However, prior to the current study, there were few studies to support any standard on how to set advisory speed limits for high-speed signalized intersections with AWF, though studies did exist for horizontal curves.

In order to avoid obtaining skewed results for the 85^{th} percentile speed, MUTCD requires that speed studies for signalized intersection approaches be undertaken outside the influence area of the traffic control signal, which is generally considered to be approximately 1/2 mile (5). However, the 85^{th} percentile speed may not represent road conditions in the vicinity of signalized intersections equipped with AWF. A reduced speed limit specific to the signalized intersection could reduce crash severity that results from high highway speeds. Arbitrary reductions, however, may result in the violation of driver expectations, leading to lower compliance; consequently the resulting increased variance in driving speeds would increase the probability of crashes. Thus, the establishment of a reduced transitional speed limit in advisory speed zones, such as at high-speed intersections, requires special engineering studies to demonstrate its effectiveness.

2.3.3 Factors Affecting Driving Speed

Many factors influence a driver's choice of speed. Researchers Ivette and Eric studied rural highways to provide a model for determining factors that influenced driver speed decisions (17). The model showed that, aside from speed limit signs, other road design features were significant, including the width of lanes and shoulders, the number of lanes, and the presence of warning signs. USLIMITS (18), a web-based system, can recommend speed limits after analyzing traffic characteristic parameters such as 85th percentile speed, 50th percentile speed, AADT, crash rate, and other related data.

These data, however, do not encompass factors related to signalized intersections with different speed limit reductions and equipped with AWD—which comprise a major object of study in the present paper. Also considered in the current research are other factors that may contribute to the speed model, such as distance from the intersection, time of day, and intersection indicator.

2.4 Summary

Chapter 2 presented a literature review of speed studies. It was found that speed limit signs had little effect on driver speed choice. Figure 2-1 illustrated that changes in actual speed in the reviewed studies were smaller than the changes in the posted speed limit, suggesting that drivers choose a comfortable speed based on roadway conditions, rather than speed limit signs. This thesis will conduct speed studies to analyze speed characteristics under different speed limit reduction schemes. Changes in the speed limit reduction and some other related factors will be tested through the modeling analysis in this thesis.

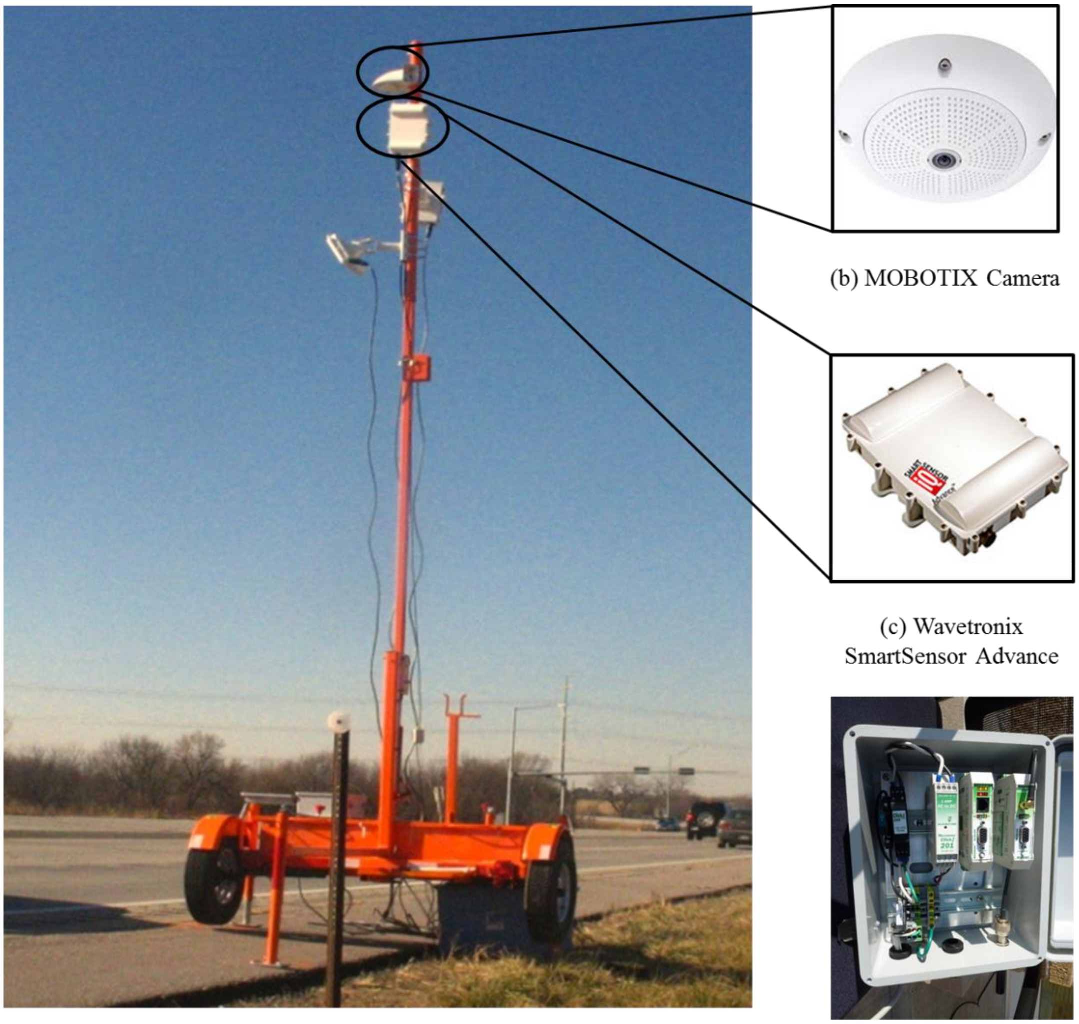
CHAPTER 3 DATA COLLECTION

3.1 Data Collection Equipment

A portable trailer (Figure 3-1 a) with a signal phase reader (Figure 3-1 d) was utilized during data collection. Data were collected on days having no precipitation and with wind gusts lower than 10 mph. The data collection trailer was equipped with a Wavetronix SmartSensor Advance sensor (Figure 3-1 c).

The Wavetronix SmartSensor Advance sensor installed on the trailer utilized digital wave radar technology to track vehicles upstream of the trailer and record their distance and speed up to a range of 500 ft. The recorded video by MOBOTIX camera (Figure 3-1 b) was used for ground truth validation.

The signal phase reader shown in Figure 3-1 d was placed at the traffic signal cabinet (C in Figure 3-2), which communicated the signal phase status via radio to the sensor trailer (A in Figure 3-2). Wavetronix Click! products are a set of power and communication modules that can connect various traffic components into a single, unified system. Within the signal phase reader, Click! 200 Serious devices convert AC to DC and provide surge protection. Click! 500 devices are customized modules that can process I/O data to support Wavetronix SmartSensor Advance applications. The signal status information is picked up by an inductance detector, which serves as an input of Click! 500 (19).



(b) MOBOTIX Camera

(c) Wavetronix SmartSensor Advance

(a) Mobile data Collection Trailer

(d) Signal Phase Reader

Figure 3-1 Equipment for data collection

Time synchronization with the portable system was maintained with reference to the trailer's Click! 500 real-time clock. The phase-reading Click! 500 received updates from the trailer's Click! 500 via the wireless link. When both of these systems were synchronized, drift between the two clocks was less than 70 ms. The entire system had a time resolution accuracy of at least 0.1 sec, as reported by the manufacturer. The data was pushed from the Click! 500, using the device's serial port and a serial to USB converter that connected to a laptop. MATLAB opened the serial port and saved the data in *.txt files.

The overall data collection schematic is shown in Figure 3-2**.** The MOBOTIX camera on the top (A2 in Figure 3-2, and Figure 3-1 b) was able to record live traffic with a 180° field of vision. Figure 3-3 displays the view from the camera. The data collected by MATLAB, as show in Figure 3-4, included date, time, ID, range, speed, and phase status.

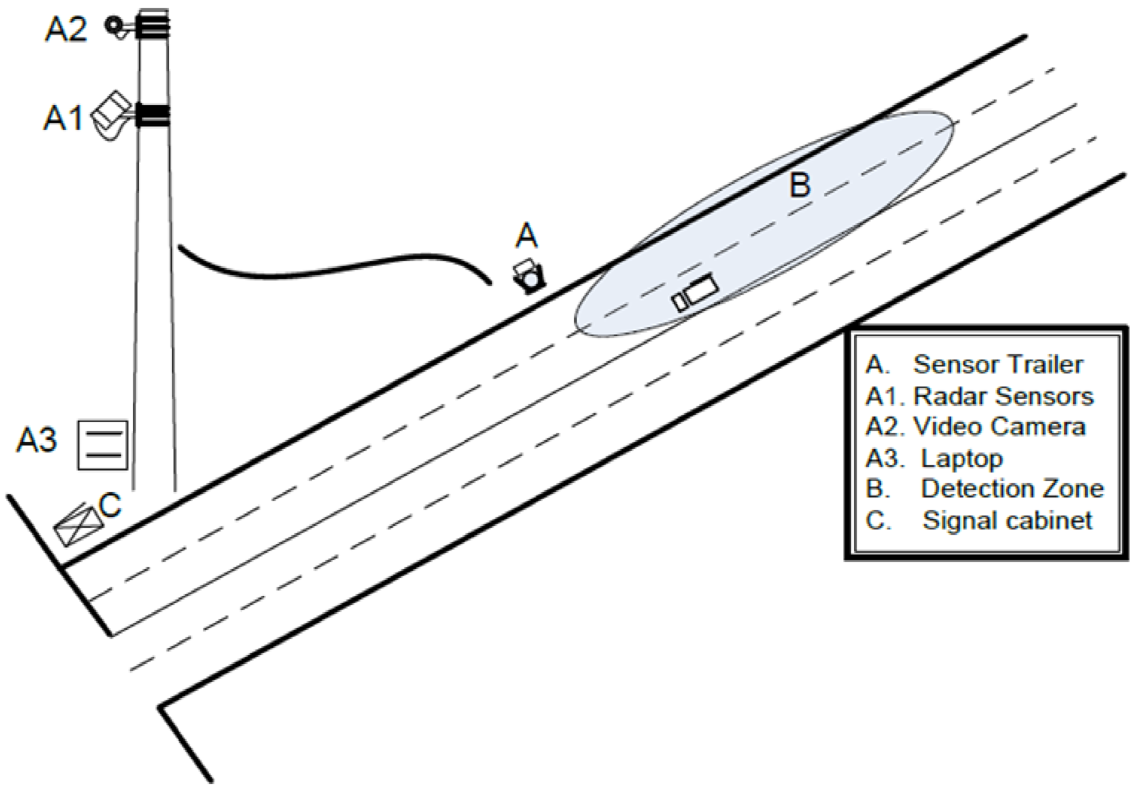


Figure 3-2 Portable Equipment Platform Trailer Setup for Data Collection

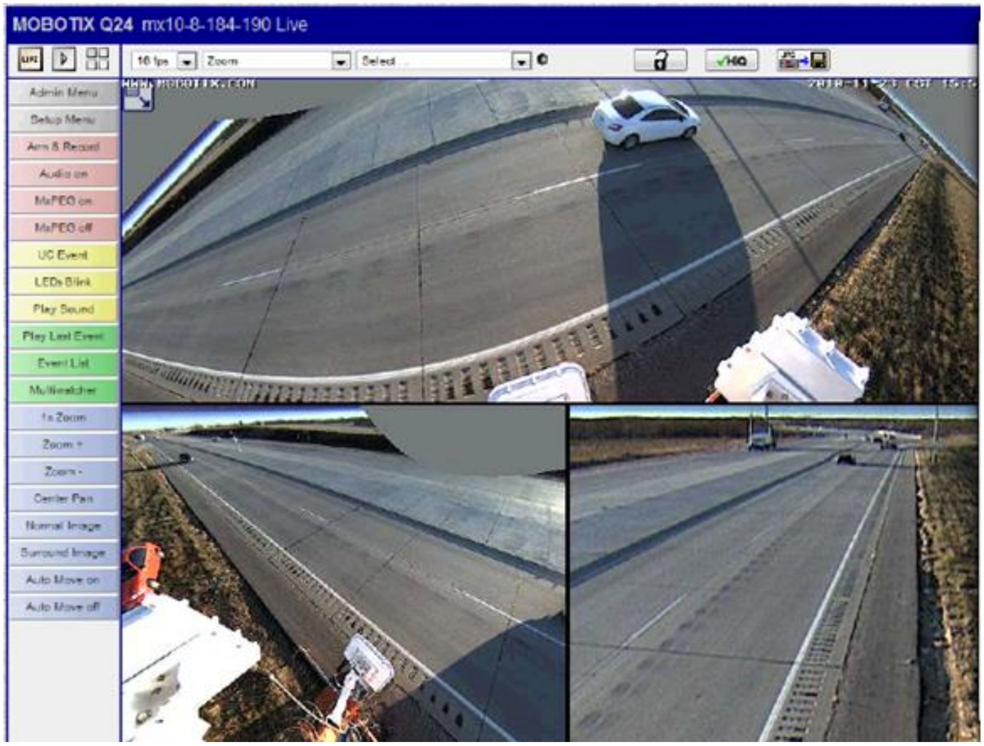


Figure 3-3 View from the MOBOTIX Fisheye Camera

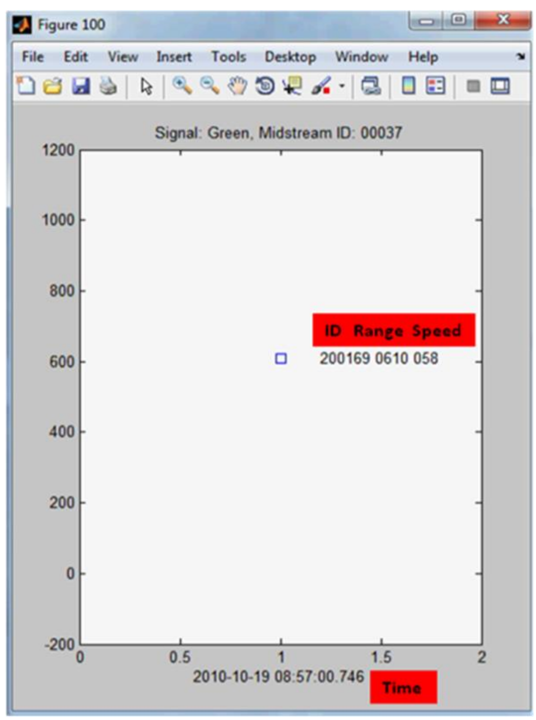


Figure 3-4 Data in MATLAB

3.2 Site Selection

Figure 3-5 gives an example of the trailer location at the US 77 and Saltillo Road test site. It can be seen that the mobile trailer was placed upstream (near the vicinity of the upstream speed limit reduction sign) on one day, and then downstream (approximately 500 ft in advance of the stop bar) on another day. Detailed layout information for all 7 intersections is included in appendix B.

The objective for placing the upstream detector was to place it as close to the beginning of the speed transition zone (i.e., the speed limit sign displaying a lower speed limit for the transition zone) as possible. Note that the beginning of the transition zones for all sites was more than 1,000 ft away from the intersection. The goal for placing the downstream detector was to place it approximately 500 ft from the stop bar. This was done in order to provide enough distance for the vehicle to decelerate upon seeing the speed limit reduction sign, and to avoid any influence of dilemma zone boundaries (5.5 sec).

The precise trailer location in the field varied by the location of the speed limit sign, the feasibility of parking the trailer, and the line of sight from the cabinet. The locations are listed in the last column of Table 3-1. Using this layout which includes two speed detection locations, a consecutive speed pattern along the road was outlined for a vehicle approaching the intersection. Note that the last column in Table 3-1 gives the physical location of the trailer. The Wavetronix SmartSensor Advance sensor is capable of detecting data from up to 500 ft away. Figure 3-6 shows the data collection range for each intersection, where the x-axis is the indicator for each intersection and the y-axis is the distance between the detected vehicle and the stop bar. The solid line denotes the detection range (~500 ft) covered by the Away from Stop Bar location, while the dash line indicates the data collection range (~500 ft) covered by the Near Stop Bar location. In addition, the location of speed limit sign and AWF are also marked in the figure.

Table 3-1 Information of Study Sites

#	Near-Stop Bar Speed Limit	Away-from Stop Bar Speed Limit	Speed Limit Drop Group	Site Location	Trailer Location	Trailer's Distance to Stop Bar (ft)
1	60	60	0 MPH Speed Limit Drop Group	US-34 & N-79 Lincoln(WB)	Near Stop Bar	1545
					Away from Stop Bar	495
2	55	55		US77 & Pioneers Blvd. Lincoln (SB)	Near Stop Bar	1380
					Away from Stop Bar	535
3	55	55		N-133 & N-36 Omaha (NB)	Near Stop Bar	1025
					Away from Stop Bar	505
4	60	55	5 MPH Speed Limit Drop Group	US75 & Platteview Rd Bellevue (SB)	Near Stop Bar	1560
					Away from Stop Bar	520
5	60	55		US-81& Lincoln Ave York (SB)	Near Stop Bar	930
					Away from Stop Bar	500
6	65	55	10 MPH Speed Limit Drop Group	US-77 & Saltillo Road. Lincoln (NB)	Near Stop Bar	1150
					Away from Stop Bar	500
7	65	55		US281& Platte River Doniphan (SB)	Near Stop Bar	2130
					Away from Stop Bar	740

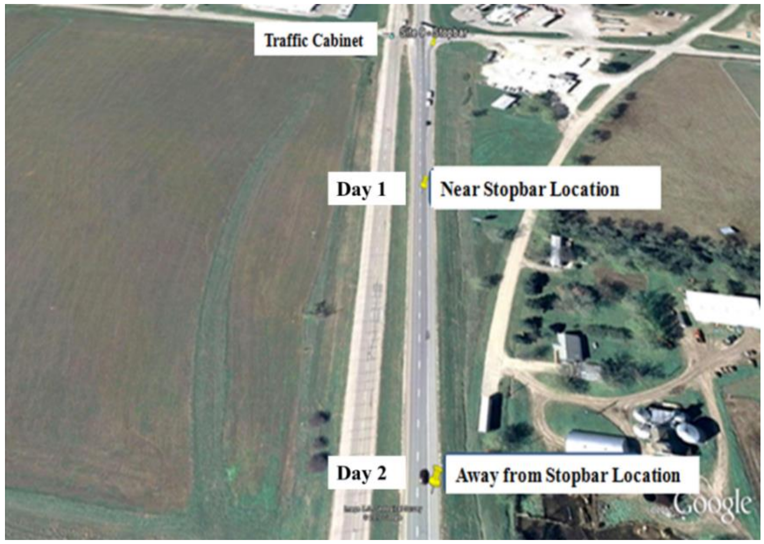


Figure 3-5 Trailer Layout at US Highway 77 and Saltillo Road

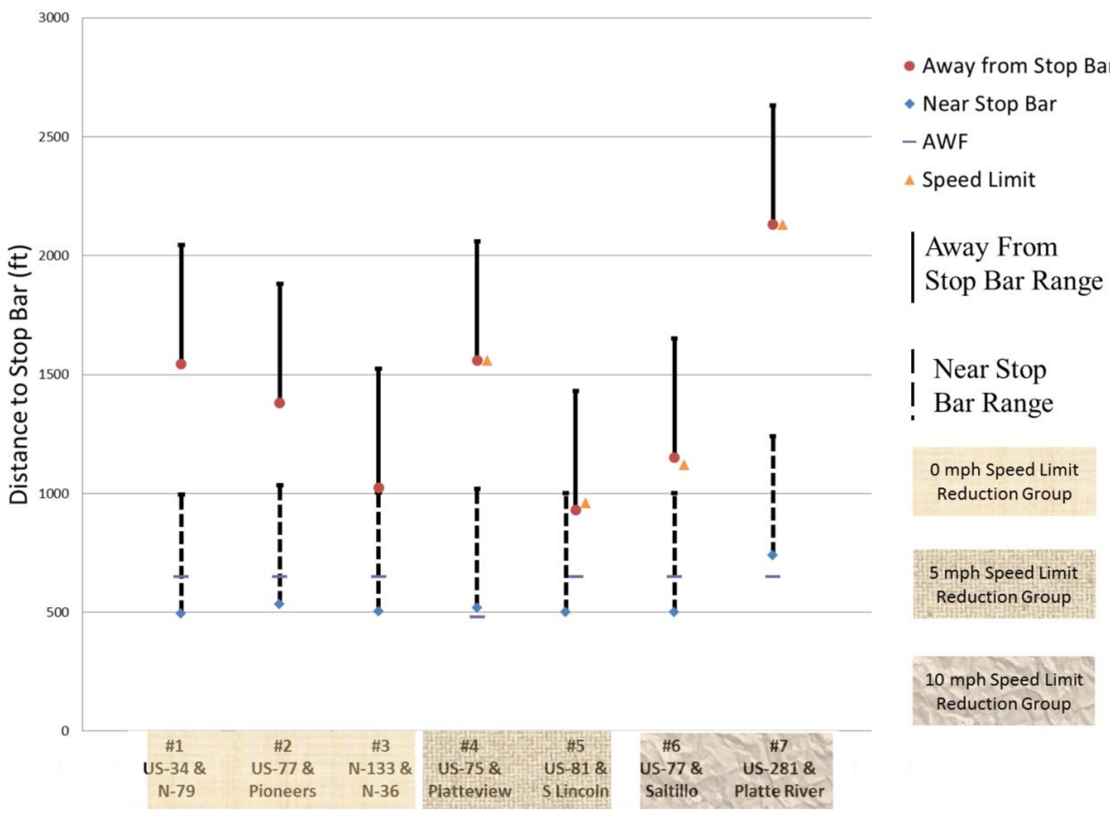


Figure 3-6 Data Collection Range for Each Intersection

3.3 Raw Data Format

The data from the Wavetronix SmartSensor Advance sensor was logged in *.txt files, as shown in Table 3-2. For signal status, 0 indicates red, 1 indicates yellow, and 2 represents green. In this study only the vehicles arriving during the middle of the green interval were included. Vehicles arriving during the red interval, yellow interval, or the first and the last 10 sec of the green interval were removed. This was done to ensure that only free-flowing vehicles were used to generate the speed profile.

Table 3-2 Sample of Wavetronix Raw Data

Time	ID	Range(feet)	Speed (mph)	Signal Status
12:14:59.246	202576	355	42	2
12:14:59.246	202575	320	50	2
12:14:59.337	202576	350	42	2
12:14:59.337	202575	315	51	2
12:14:59.437	202576	345	43	2
12:14:59.437	202575	305	51	2
12:14:59.538	202576	335	43	2
12:14:59.538	202575	300	52	2
12:14:59.638	202576	330	42	2
12:14:59.638	202575	290	52	2
12:14:59.745	202576	325	42	2
12:14:59.745	202575	285	55	2

A green threshold of 10 sec was selected due to the fact that in Nebraska, among 35 high speed signalized intersections equipped with AWF, flasher time is set from 6 to 10 sec before the onset of yellow, whereas, at the 7 select target intersections included in this study, flasher duration was between 6 and 8 sec (4). In order to avoid the influence of the AWF, a conservative time period of 10 sec was chosen; hence the last 10 sec were excluded from the entire green time.

As for the first 10 sec, queue length estimation was performed. Among all cycles examined, approximately 99% had a queue length smaller than 5 vehicles per lane. The queue length distribution chart is listed in Figure 3-7. Assuming a 2-sec discharge headway, it was assumed that 10 sec would be sufficient to clear the starting queue at the intersection.

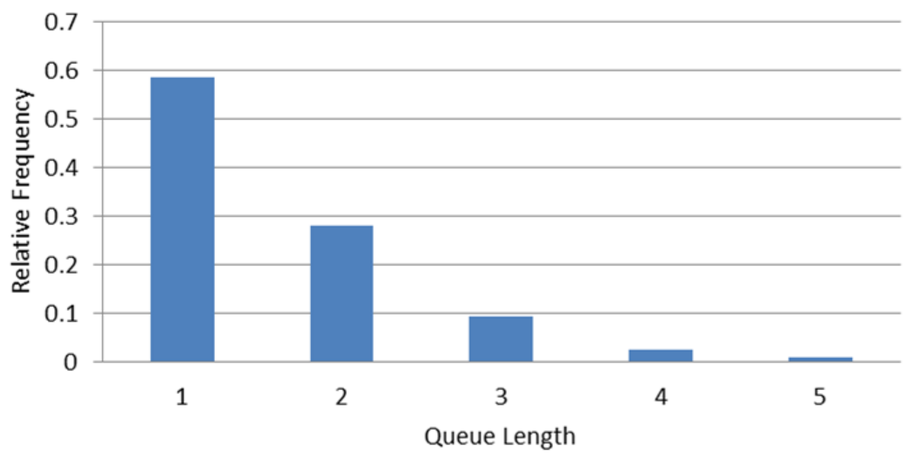


Figure 3-7 Queue Length Distribution

3.4 Summary

Chapter 3 described the data collection method using the Wavetronix SmartSensor Advance sensor at 7 intersections and the format of raw data. However, the raw data are subject to many undesirable entities; therefore, in the next chapter 4, it will describe the detailed data processing before further data analysis.

CHAPTER 4 DATA PROCESSING

4.1 Individual Vehicle Performance

Sharma et al. (20) described a procedure to evaluate WAD (Wide Area Detector) speed and location detection using a probe vehicle with a GPS handheld device. The speed obtained from the GPS device was validated against an Onboard Diagnostic Device (OBD), which recorded a vehicle's built-in speed sensor. Results confirmed that the GPS device accurately detected vehicle speed. As for position, the WASS-enabled GPS device had an accuracy of $\pm$6 feet. Therefore, GPS was used to validate the accuracy of the Wavetronix SmartSensor Advance sensor in the current study. A total of 55 test runs were conducted to verify Wavetronix SmartSensor Advance sensor's performance in individual vehicle detection at all 7 test sites.

The speed difference was defined as the difference in speeds detected by the Wavetronix SmartSensor Advance sensor and the GPS device at the same location. Figure 4-1 shows the speed difference histogram obtained by combining data obtained from 55 runs. The overall speed difference had a median of 0 mph and mean 0.01 mph, with a standard deviation of 1.39 mph.

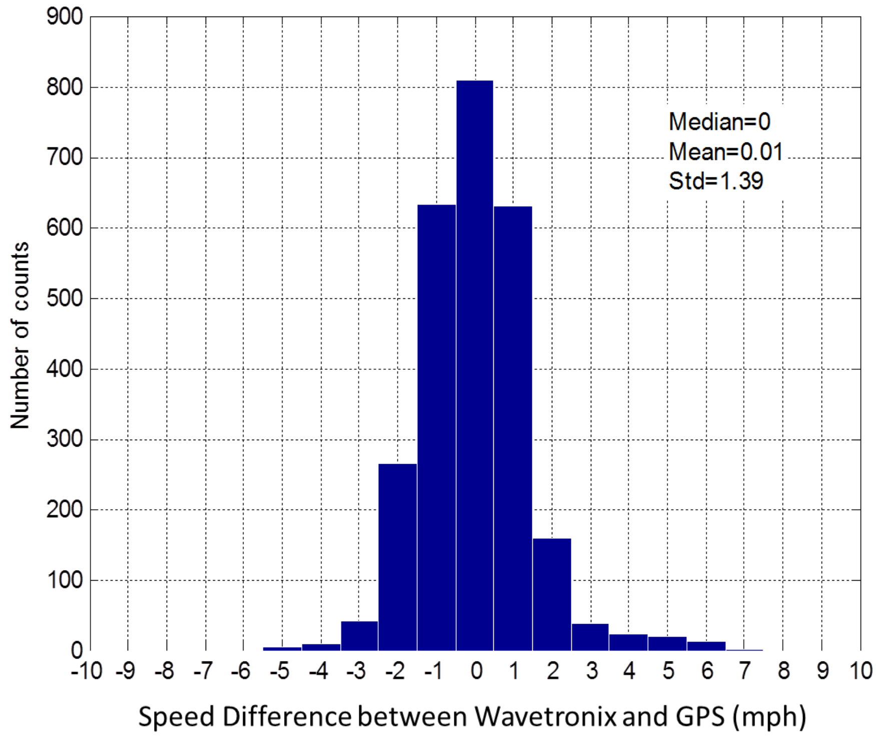


Figure 4-1 Speed Difference between Wavetronix and GPS Histogram

4.2 Data Processing for Eliminating Ambiguous Calls

In section 4.1, the Wavetronix SmartSensor Advance sensor's individual vehicle performance was evaluated using a GPS device. The next step in data processing was to minimize ambiguous calls generated by the sensor. Some common ambiguous calls observed based on ground truth validation are described in the following paragraphs:

Shared ID: Ideally, MATLAB data requisition programs assign a unique ID to each detected vehicle; however, once the program restarts, the ID assignment is also reset, which may generate duplicate IDs for separate vehicles. This type of ambiguous call can be addressed using the time stamp. If the time gap between two detection points is greater than an assigned threshold (7 sec), a new ID will be assigned. The threshold is set at 7 sec due to the fact that it takes about 7 seconds for a vehicle pass through the 500 ft detection zone under the observed speed. Figure 4-2 gives example of the Shared ID. The distance and time diagram is plotted based on the ID to see vehicle's trajectory, however Figure 4-2 illustrates two vehicles trajectories sharing the same ID.

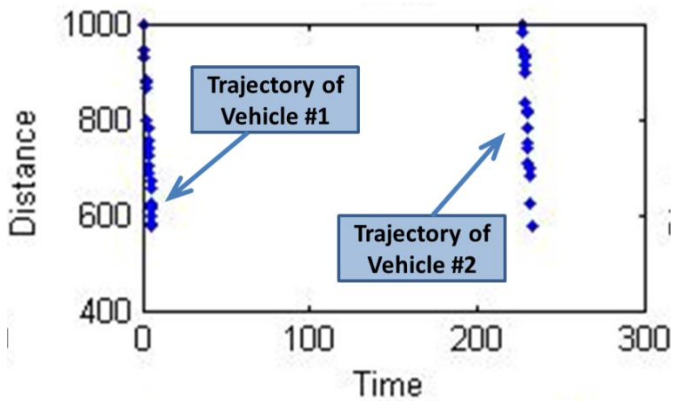


Figure 4-2 Shared ID

Stuck Call: Figure 4-3 shows a step-like trajectory where a vehicle was detected at the same point at multiple times. The data collection was conducted under free-flowing traffic conditions during the green time period, so it was highly improbable for a vehicle to have a stop and go profile. This trajectory was therefore classified as an ambiguous call. A similar example is illustrated in Figure 4-4, which shows a vehicle stopped at one location (805 ft location) for over 3 sec.

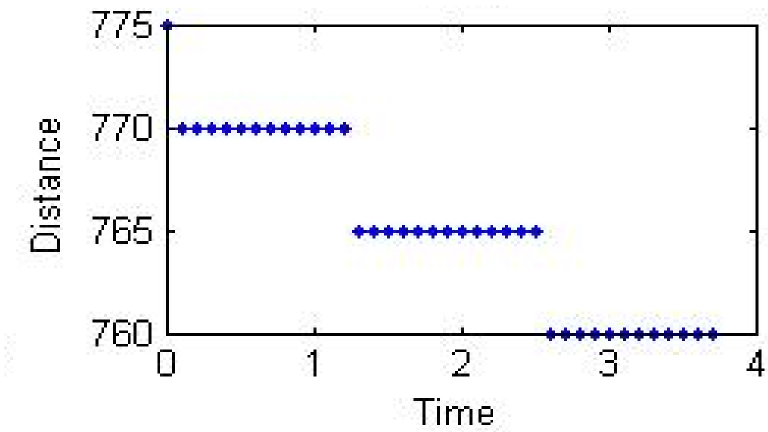


Figure 4-3 Step Like Detection

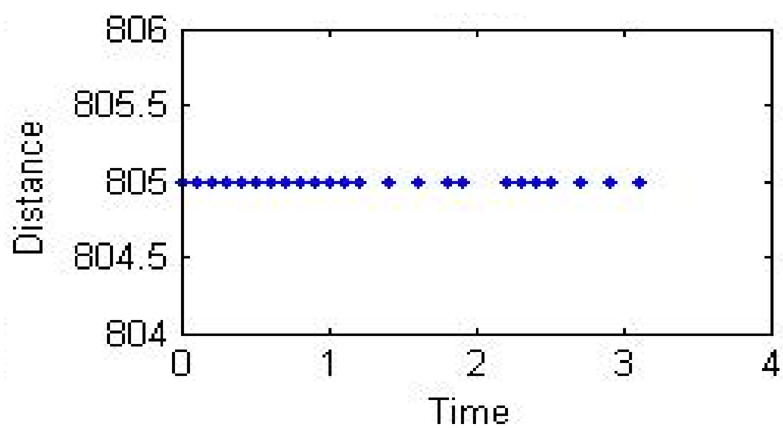


Figure 4-4 Stuck Detection at the Same Point

Single Detection: The Wavetronix SmartSensor Advance sensor has the ability to track vehicles continuously; however, the example in Figure 4-5 shows a single detection point when it was verified that there was no vehicle present at this location.

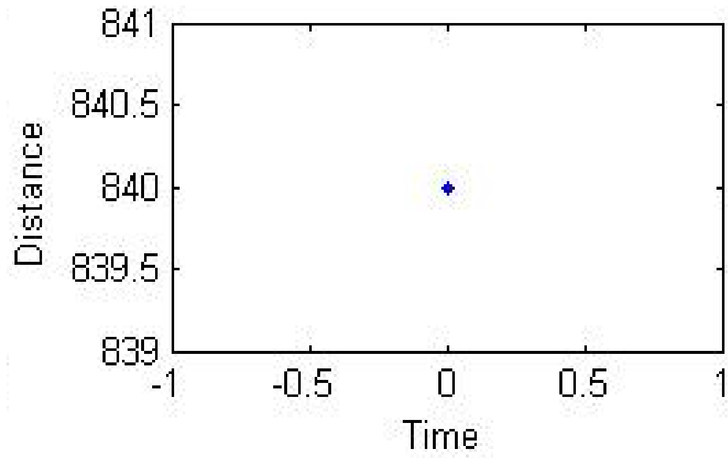


Figure 4-5 Single Detection

Dropped Call: Figure 4-6 is an example of a dropped detection. The trajectory illustrates that the sensor detected this vehicle for only 60 ft, then lost the detection beyond that range. Dropped calls were excluded from the processed dataset because they might represent a trailer towed by a pickup, the large size of truck, etc.

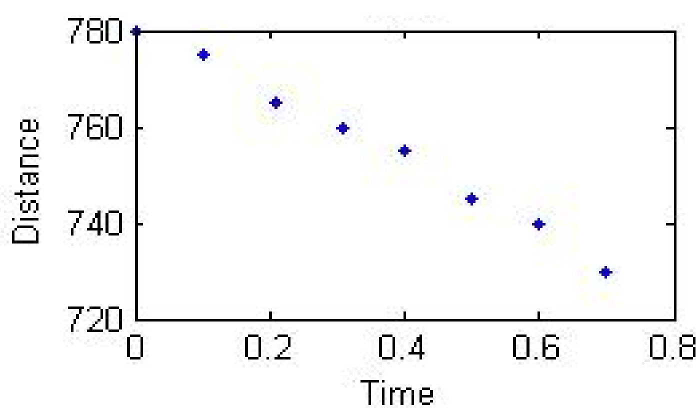


Figure 4-6 Dropped Detection

The box plot in Figure 4-7 shows the preliminary analysis of these ambiguous calls. The x-axis indicates the type of ambiguous call, and the y-axis is the relative frequency of each type of ambiguous call for each data collection day.

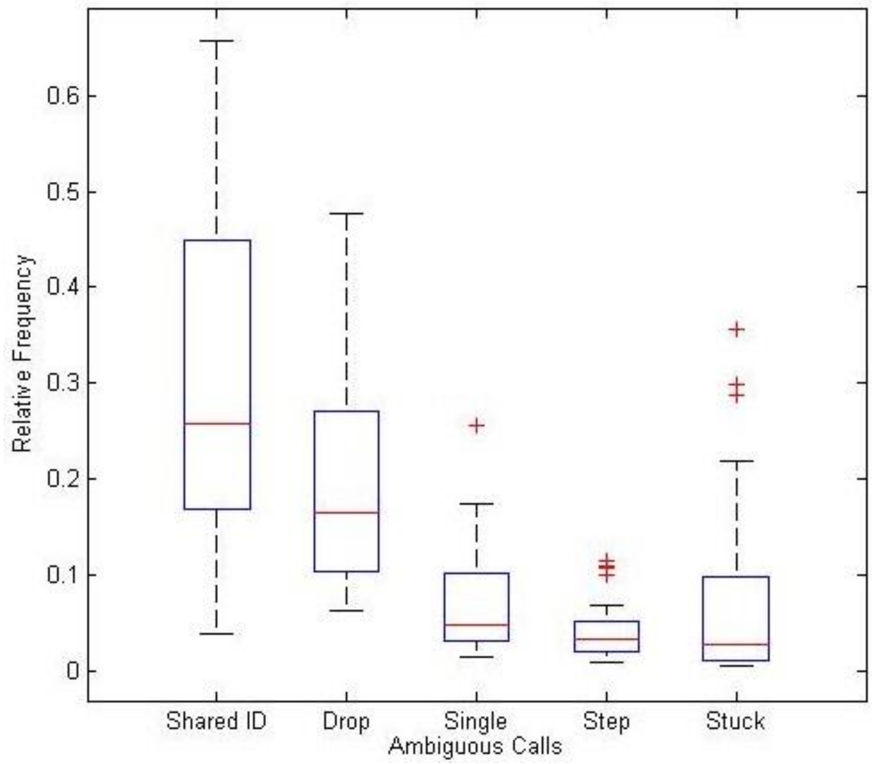


Figure 4-7 Preliminary Study of Ambiguous Calls

4.3 Machine Learning

Professor Andrew Ng from Stanford University's online course CS229 provided a detailed introduction to machine learning (21). In a given machine learning algorithm, one can use $x^{(i)}$ to denote the input variables, which are also called input features, and $y^{(i)}$ to denote the output variable that one is trying to predict. A pair of $(x^{(i)},y^{(i)})$ is called a training example. The dataset that will be used to learn is called a training set, which is a list of *m* training examples $\{ (x^{(i)},y^{(i)}) ; i=1, 2, 3, \ldots, m\}$. The goal is to build a hypothesis function *h(x)* that is a good predictor for the corresponding value *y*. Figure 4-8 gives an overview of a typical machine learning algorithm.

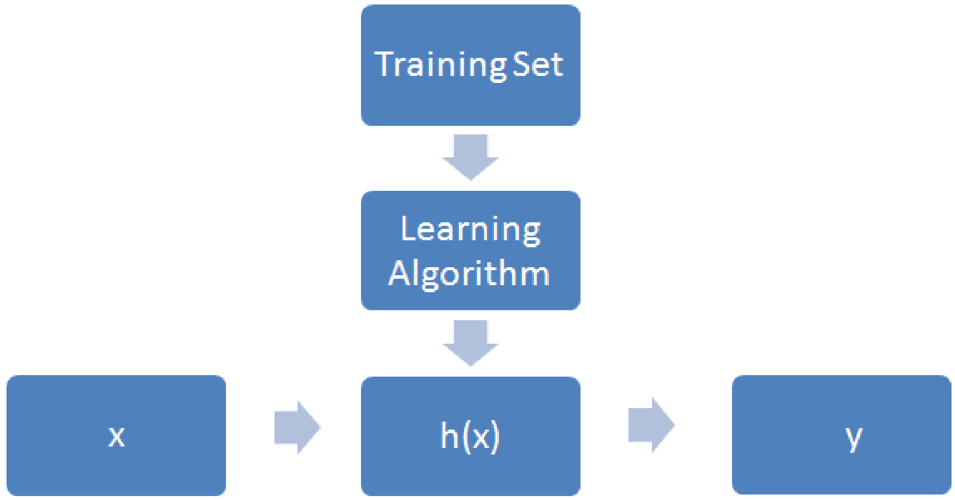


Figure 4-8 Machine Learning Concept

Based on the features of the output value *y*, machine learning can be categorized into two problems: the first is a regression problem when *y* is continuous; the second is a classification problem, when *y* is a small number of discrete values. As for this thesis, the focus will be upon the binary classification problem in machine learning. The goal is to classify the entire dataset into desirable calls and ambiguous calls. Here, *y* can take on only two values, 0 and 1, where 0 represents ambiguous calls and 1 represents desirable calls.

4.3.1 Logistic Function

When dealing with the binary classification problem, the form of hypothesis *h* can be written as:

$$h\,(x) = g(\beta x) = \frac{1}{1 + e^{-\beta x}},$$

where,

$$g(z) = \frac{1}{1 + e^{-z}}$$

g(z) is called the logistic function, and its plot is shown in Figure 4-9. It can be noticed that when z approaches + ∞, g(z) tends towards 1, and when z approaches - ∞, g(z) tends towards 0. Therefore, the hypothesis *h* is bounded between 0 and 1.

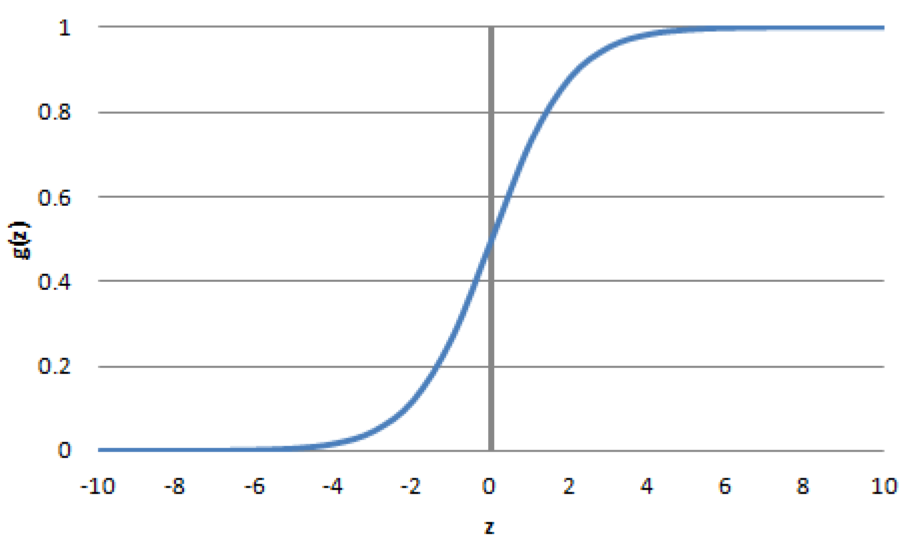


Figure 4-9 Logistic Function Plot

The *z* is defined as the classifier if z > 0, g(z) > 0.5, which means that it is 50% certain that *y* will be predicted as 1. If z < 0, *y* is predicted as 0. Suppose that the input *x* has two features (two dimensions with x1 and x2); then the classifier z can be visualized in

Figure 4-10.

A linear classifier is a linear combination of features (x1 and x2), which can be described as:

$$Linear\ Classifier(z) = K + \boldsymbol{x} \times \boldsymbol{L}$$

where:

- K = constant term of the boundary equation
- L = linear coefficients of the boundary equation
- x = input features

Whereas, with a quadratic classifier, the boundary equation is based on quadratic combination of input features, which can be described as:

$$Quadratic\ Classifier(z) = K + \boldsymbol{x} \times \boldsymbol{L} + \boldsymbol{x} \times \boldsymbol{Q} \times \boldsymbol{x}^{\boldsymbol{T}}$$

where,

- K = constant term of the boundary equation
- L = linear coefficients of the boundary equation
- Q = quadratic coefficient matrix of the boundary equation
- **x** = input features

The classifier *z* divides the entire dataset region into two zones. One zone is for y equals to 0 (if z < 0), and the other zone is for y equals to 1 (if z > 0).

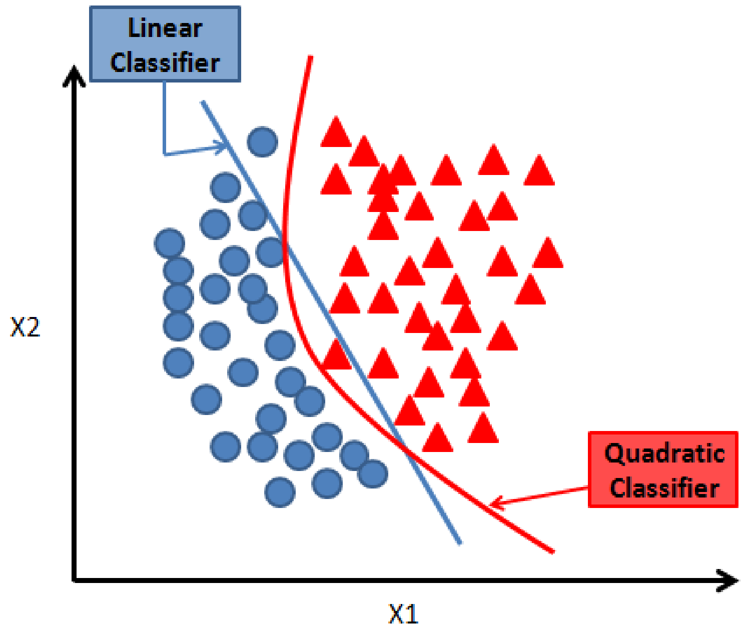


Figure 4-10 Classifier

4.3.2 Evaluation of Classifier

The performance evaluation table, as shown in Table 4-1, was used to evaluate the performance of the classifier.

Table 4-1 Performance Evaluation Table

		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	True Desirable	False Desirable
	Prediction Outcome: Ambiguous Calls:	False Ambiguous	True Ambiguous

Based on Table 4-1, two terms were defined: Precision (P) and Recall (R), where,

$$Precision\ (P) = \frac{No. of\ True\ Desirable\ Calls}{No. of\ True\ Desirable\ Calls\ +\ No. of\ False\ Desirable\ Calls}$$

$$Recall\ (R) = \frac{No.of\ True\ Desirable\ Calls}{No.of\ True\ Desirable\ Calls\ +\ No.of\ False\ Ambiguous\ Calls}$$

The F score (the higher the better) is the statistic used for performance measurement (21), which can be calculated by Precision (P) and Recall (R).

$$F = \frac{2PR}{P+R}$$

4.3.3 Training Set

Figure 4-8 indicates that in order to build the machine learning algorithm, one should first have a training set. Within the training set date, the input variable x and output results y are known. Based on the known output y and input variable x, one can estimate the algorithm *h*, which can be later used to predict the unknown outcome *y* based on the new variable *x.*

In this thesis, the training dataset had 549 unique ids (vehicles) obtained from the intersection of U.S. Highway 77 and Saltillo Road. These unique ids were classified as ambiguous or unambiguous calls by manual ground-truth validation using the video overlain with the detector information.

A training set { $(x^{(i)},y^{(i)})$; i=1, 2, 3, …, m} was available where *m* was equal to 549. The output *y* equaled either 1 or 0, where 1 indicated a prediction outcome as a desirable call, and 0 indicated a prediction outcome as an ambiguous call. As for the input variable $x^{(i)}$, it had up to 4 different features ($x^{(i)} = [x_1^i,\ x_2^i, x_3^i, x_4^i]$) in this study. Each of the features is described below, and includes Detection Range, Number of Actuations, Mean Speed, and Speed Variance.

Detection Range: Detection Range is the distance between the start and end points of the trajectory for an observed vehicle. A desirable call would continuously track the vehicle over approximately 500 ft, as per the sensor specifications. An ambiguous call might have a lower value for Detection Range, as in the case of single call, dropped call, stuck call, etc.

Number of Actuations: Number of Actuations is the total number of actuations that are registered for a unique ID (or a single vehicle). Ideally, the velocity and location of each individual vehicle is updated at every 5 ft. There is a priori expectation to observe approximately 100 Number of Actuations for each unique ID, which represents a single vehicle that is being tracked over a distance of 500 ft. Thus, an ID group with unreasonably few points is highly likely to be an ambiguous call.

Mean Speed and Speed Variance: In this study, all the data used for analysis occurred under free flow conditions during the green time. The Wavetronix SmartSensor Advance sensor's capacity ensured that each vehicle could be continually tracked over a range, thus, multiple actuations points were available for each vehicle, and the speed mean and variance for each individual vehicle over the distance for which the vehicle was tracked could be calculated.

4.3.4 Classifier Selection

Based on the training set { $(x^{(i)},y^{(i)})$; i=1, 2, 3, …, 549}, several different input variables were tested to determine a good classifier to differentiate between manually observed ambiguous call versus desirable calls.

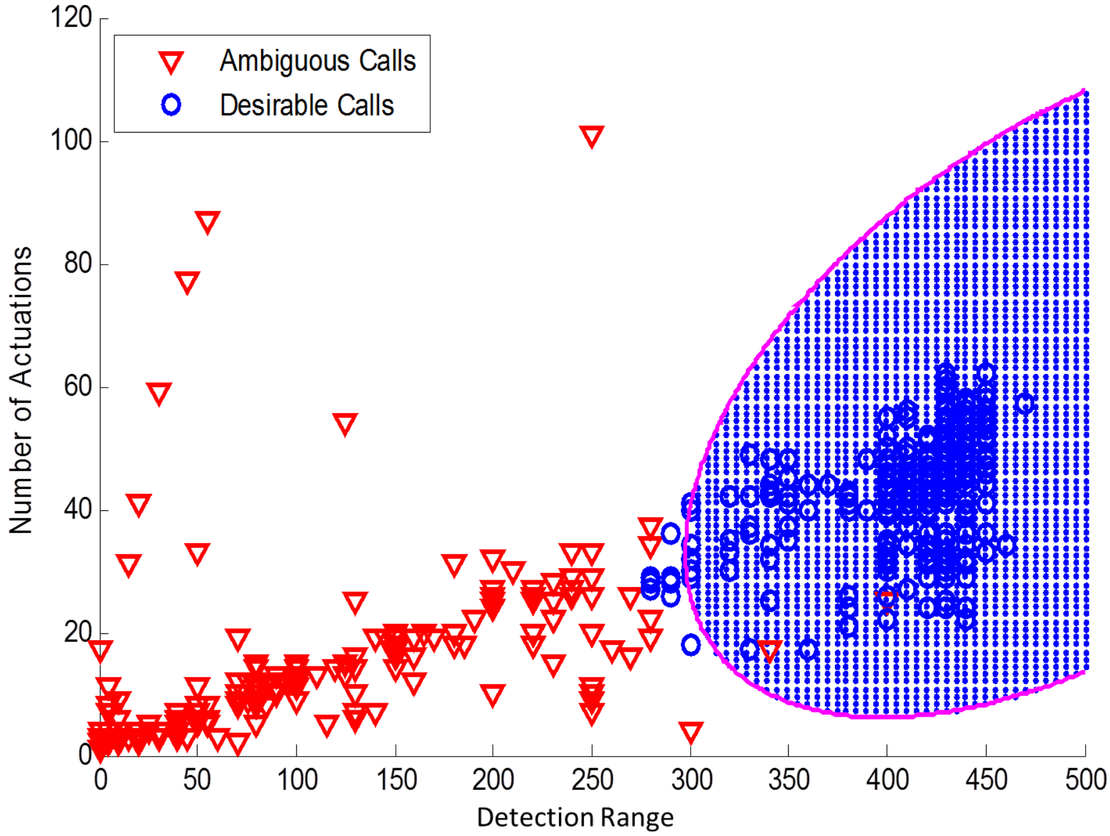


Figure 4-11 Quadratic Classifier based on Detection Range & Number of Actuations

Figure 4-11 is an example of a quadratic classifier based on Detection Range and Number of Actuations. In the figure, the circle represents desirable calls, and the triangle represents ambiguous calls. The curved line is the classifier boundary, which divides the whole dataset region into two zones, one for y = 0 (ambiguous calls), and the other (shaded area) for y = 1 (desirable calls).

Several combinations of the input variables described in section 4.3.3 were tested to find the best classifier with optimal classification accuracy, i.e., a higher F score. The test result is listed in Table 4-2. Detailed information on each tested classifier is listed in appendix C.

Table 4-2 Performance Evaluation for Each Classifier Based on Training Set

Input Features	P	R	F=2PR/(P+R)
Quadratic Classifier Detection Range & Number of Actuations	0.993	0.971	0.982
Linear Classifier Detection Range & Number of Actuations & Mean Speed & Speed Variance	0.963	1.000	0.981
Quadratic Classifier Detection Range & Number of Actuations & Mean Speed & Speed Variance	0.997	0.987	0.992
Quadratic Classifier Mean Speed & Speed Variance	0.727	0.959	0.827
Linear Classifier Detection Range & Number of Actuations	0.960	1.000	0.980
Quadratic Classifier Mean Speed & Number of Actuations	0.950	0.965	0.957

The first 3 classifiers had the highest F score: 1) quadratic classifier based on Detection Range and Number of Actuations had an F score of 0.982; 2) linear classifier based on all 4 features had an F score of 0.981, and 3) quadratic classifier based on all 4 features had an F score of 0.992.

The first 3 classifiers are then chosen to be validated by another dataset, which contained 15 minutes of data from all remaining intersections. The validation data set contained 456 examples — $\{(x^{(i)},y^{(i)})\ ; i=1, 2, 3, \ldots, 456\}$. The results are listed in Table 4-3.

Table 4-3 Performance Evaluation for Each Classifier Based on Validation Set

Input Features	P	R	F=2PR/(P+R)
Quadratic Classifier Detection Range & Number of Actuations	0.986	0.981	0.984
Linear Classifier Detection Range & Number of Actuations & Mean Speed & Speed Variance	0.947	0.991	0.968
Quadratic Classifier Detection Range & Number of Actuations & Mean Speed & Speed Variance	0.991	0.972	0.981

The F scores showed that only the Quadratic Classifier based on Detection Range & Number of Actuations continued to perform well. The remaining two classifiers, based on all 4 features, exhibited a decreased F score, which may suggest over-fitting. Additionally, using a complex classifier, the computation could be expensive with a higher dimension of matrix calculation; therefore, the quadratic classifier based on Detection Range & Number of Actuations was chosen for the final data reduction. Its parameters were:

$$Quadratic\ Classifier = K + [x_1, x_2] \times \boldsymbol{L} + [x_1, x_2] \times \boldsymbol{Q} \times [x_1, x_2]^{\boldsymbol{T}}$$

$$K = 49.0278, \quad \boldsymbol{L} = \begin{bmatrix} -0.2424 \\ 0.0743 \end{bmatrix}, \quad \boldsymbol{Q} = \begin{bmatrix} 3.1462 \times 10^{-4} & -6.1271 \times 10^{-4} \\ -6.1271 \times 10^{-4} & 0.0044 \end{bmatrix}$$

- x_1 = Detection Range
- x_2 = Number of Actuations
- K = Constant term of the classifier
- L = Linear coefficients of the classifier
- Q = Quadratic coefficient matrix of the classifier

4.4 Summary

Chapter 4 presented Wavetronix SmartSensor Advance sensor's performance of individual vehicle detection and a machine learning method for massive data processing. The performance comparison against GPS showed that the mean speed difference between GPS and the Wavetronix SmartSensor Advance sensor was 0.01 mph, with a standard deviation of 1.39 mph. Later, a machine learning based classifier was developed to purge the data of ambiguous calls. The F score of the quadratic classifier as tested on the validation dataset was 0.984. This processed data was then used for further data analysis and modeling, as discussed in the following chapters.

CHAPTER 5 DATA ANALYSIS

Chapter 5 presents a preliminary statistical analysis and detailed modeling description of the data collected at 7 test sites. The following paragraph presents the amount of data collected at 7 test sites, and evaluates the confidence interval for the mean using the boot-strapping technique. Quantile regression models for speed were estimated to study the impacts of speed limit reduction. The results of the quantile regression models are discussed, and the primary insights gained are highlighted.

5.1 Mean Speed

Researchers, including the California DOT, commonly use a sample size of 100 vehicles to conduct speed studies (17, 22). For the current thesis, after eliminating ambiguous calls, each site had a sample size (number of vehicles) higher than 100. Table 5-1 shows the sample size for each intersection.

Table 5-1 Sample Size for Each Intersection

#	Site Name	Sample Size(No. of Vehicles)	
		Away from Stop Bar	Near Stop Bar
1	US-34 & N-79	539	1284
2	US-77 & Pioneers	876	2264
3	N-133 & N-36	321	828
4	US-75 & Platteview	1685	3095
5	US-81 & S Lincoln	337	386
6	US-77 & Saltillo Rd.	661	656
7	US-281 & PlatteRiver	857	435

The speed of each vehicle over the observed trajectory was calculated, then used to estimate the overall mean speed for each location (away from the stop bar and near the stop bar) for each intersection. However, the distribution of vehicle speed may not be normally distributed, and it is important to note that the t distribution requires that the population from which sample are drawn is normal (23). In addition, Sawilowsky and Blair found that the t test is relatively robust to violation of the normality assumption when the following conditions hold: equal variances and sample sizes, sample size of 25 or more, and two-tailed test. However, unequal sample sizes are common and variances are often heterogeneous (24). Therefore, the bootstrap technique was applied to estimate the population mean. Detailed speed distribution plots for each location are listed in appendix D.

The bootstrap is a data simulation method with no the normality condition and without the restriction to comparison of means (25). It can be achieved with repeated samples that are the same size of the original sample, and when it is resampled with replacement (25, 26).

Figure 5-1 illustrates this process.

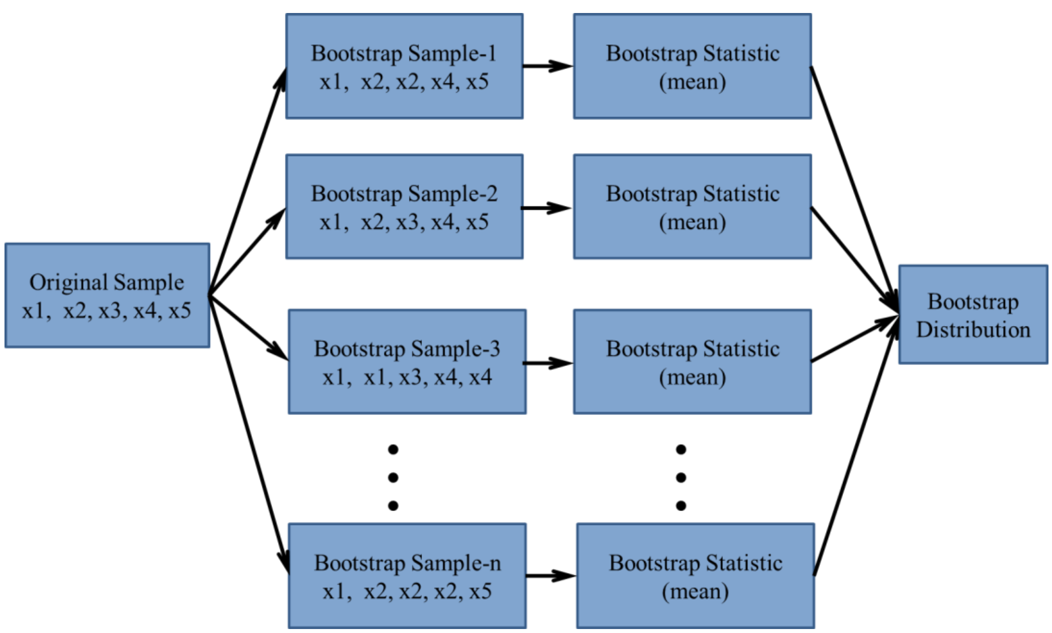


Figure 5-1 Illustration of the Bootstrap Method

Take

Figure 5-1 for example. The original dataset contained 5 observations (x1, x2, x3, x4, and x5). It was then randomly sampled with replacement *n* times, with each bootstrap sample containing exactly 5 observations. Note that applying bootstrap replication meant that an individual observation from the original dataset could be included several times, while other observations may not have been included at all. To find the standard error of the mean, the mean for each of the bootstrap samples can be calculated, followed by estimation of the standard deviation of the bootstrap means. The more bootstrap replications that are used, the more accurate the results. Generally, 10,000 replications is the recommended quantity (26).

One method of calculating the bootstrap confidence interval is the percentile interval method. By calculating a 95% confidence interval of the mean, one can select the bootstrap statistic which lies on the 2.5th percentile and 97.5th percentile (25, 26).

The results contained in Table 5-2 are based upon 10,000 replications, as recommended (26). The "Lower" column is the lower bounds of a 95% confidence interval, while the "Higher" column indicates the higher bounds of a 95% confidence interval. S.E. represents the standard error of the mean, which is the standard deviation of all the bootstrap sample means. For example, the figure shows that there was a 95% chance that the true mean speed calculated by different samples lay between 57.0 and 57.9 mph for the Away from Stop Bar location at intersection #1.

Table 5-2 95 % Confidence Bounds for Mean Speed

#	Site Name	Away from Stop Bar				Near Stop Bar			
		Mean	S.E.	Lower	Higher	Mean	S.E.	Lower	Higher
1	US-34 & N-79	57.4	0.2	57.0	57.9	59.2	0.2	58.9	59.5
2	US-77 & Pioneers	59.7	0.1	59.5	60.0	57.2	0.1	57.0	57.4
3	N-133 & N-36	58.7	0.3	58.2	59.2	56.1	0.2	55.7	56.6
4	US-75 & Platteview	57.9	0.1	57.6	58.1	57.5	0.1	57.4	57.7
5	US-81 & S Lincoln	55.9	0.4	55.2	56.6	56.2	0.3	55.5	56.8
6	US-77 & Saltillo	61.7	0.2	61.2	62.1	56.7	0.3	56.1	57.3
7	US-281& Platte River	61.6	0.2	61.2	61.9	57.3	0.2	56.9	57.8

Based on Table 5-2, the figures Figure 5-2, Figure 5-3, and Figure 5-4 were made to illustrate the speed statistics for each intersection at two different locations. Within the figure, the x-axis is the intersection number and the y-axis is the speed in mph. For each intersection, the mean speeds at the Away from Stop Bar location and Near Stop Bar location are marked as circles and diamonds, respectively. The 95% confidence interval boundary for mean speed is also marked.

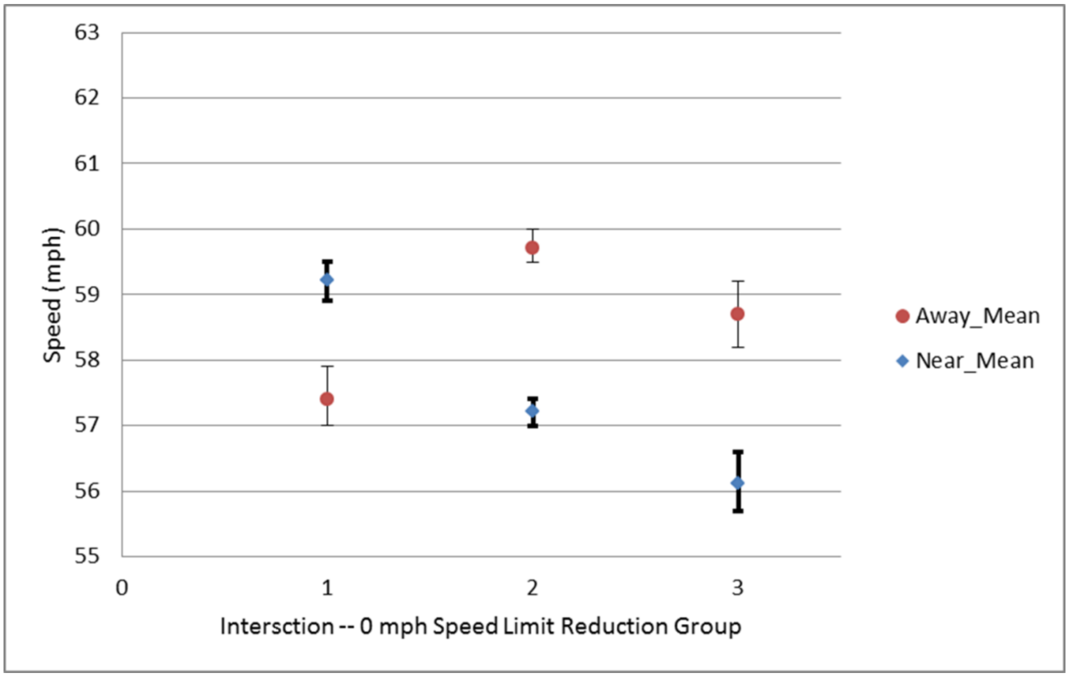


Figure 5-2 Speed Statistics for 0 mph Speed Limit Drop Group

Figure 5-2 illustrates speed statistics for intersections with constant speed limits. Intersections #2 (US-77 & Pioneer Blvd) and #3 (N-133 & N-36) depict a speed reduction between the Away from Stop Bar location and the Near Stop Bar location. The mean speeds dropped from 59.7 mph to 57.2 mph (a drop of 2.5 mph), and from 58.7 mph to 56.1 mph (a drop of 2.6 mph) as vehicles approached closer to the stop bar at the intersection of US-77 & Pioneer Blvd and the intersection of N-133 & N-36, respectively.

However, intersection #1 (US-34 & N-79) exhibited a speed increase as vehicles approached closer to the stop bar. The mean speeds increased from 57.4 mph to 59.2 mph (a gain of 1.8 mph) as vehicles approached closer to the stop bar at the intersection of US-34 & N-79. The intersection of US-34 & N-79 has a speed limit of 60 mph, while the rest of the intersections have a speed limit of 55 mph. In addition, this intersection is a T intersection, whereas the remainders are normal, 4-approach intersections. These unique intersection features could explain why mean speed increased when vehicles approached closer to the intersection.

Based on the intersection group with the constant speed limit, it can be deduced that, in the case of the absence of any speed limit reduction, there was no consistent behavior with respect to actual speeds observed away from and near the stop bar. A decrease in mean speed was observed at two sites, whereas an increase in mean speed was observed at one site.

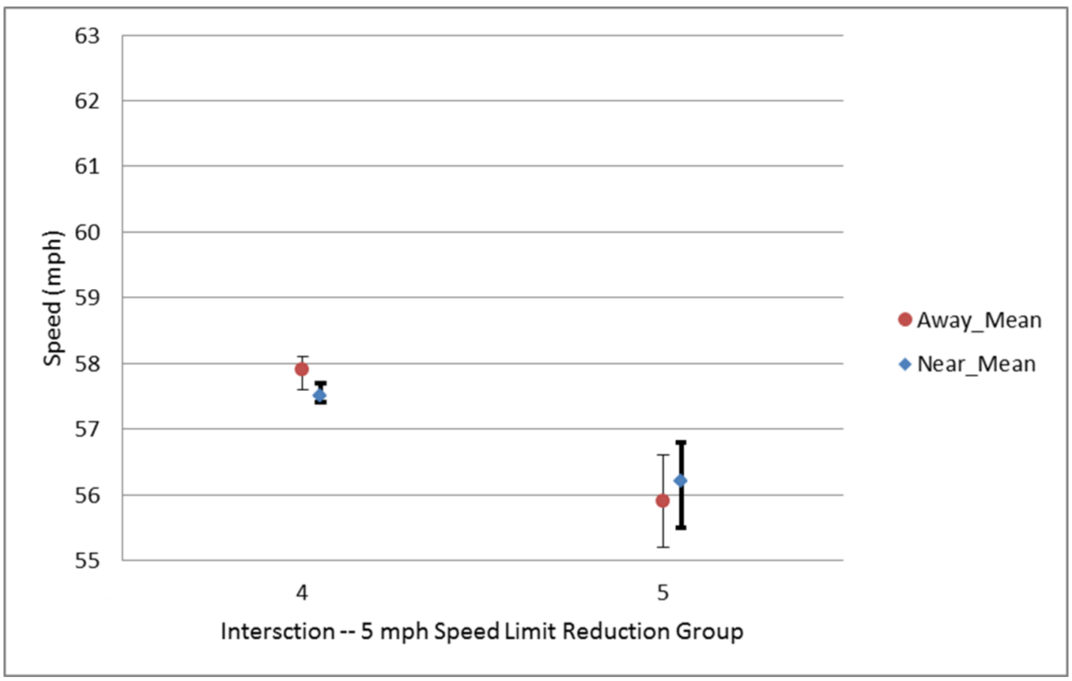


Figure 5-3 Speed Statistics for 5 mph Speed Limit Drop Group

Figure 5-3 shows that overall speeds remained, as evidenced by the almost overlapped speed markers at two intersections. The mean speeds changed from 57.9 mph to 57.5 mph (a drop of 0.4 mph) and from 55.9 mph to 56.2 mph (a gain of 0.3 mph) as vehicles approach closer to the stop bar at the Intersection of US-75 & Platteview and the Intersection of US81 & S Lincoln Ave., respectively. It can be seen that, in the case of 5 mph speed limit reduction, the overall speed difference between the Away from Stop Bar location and Near Stop Bar location was not statistically significant due to the overlapped confident interval.

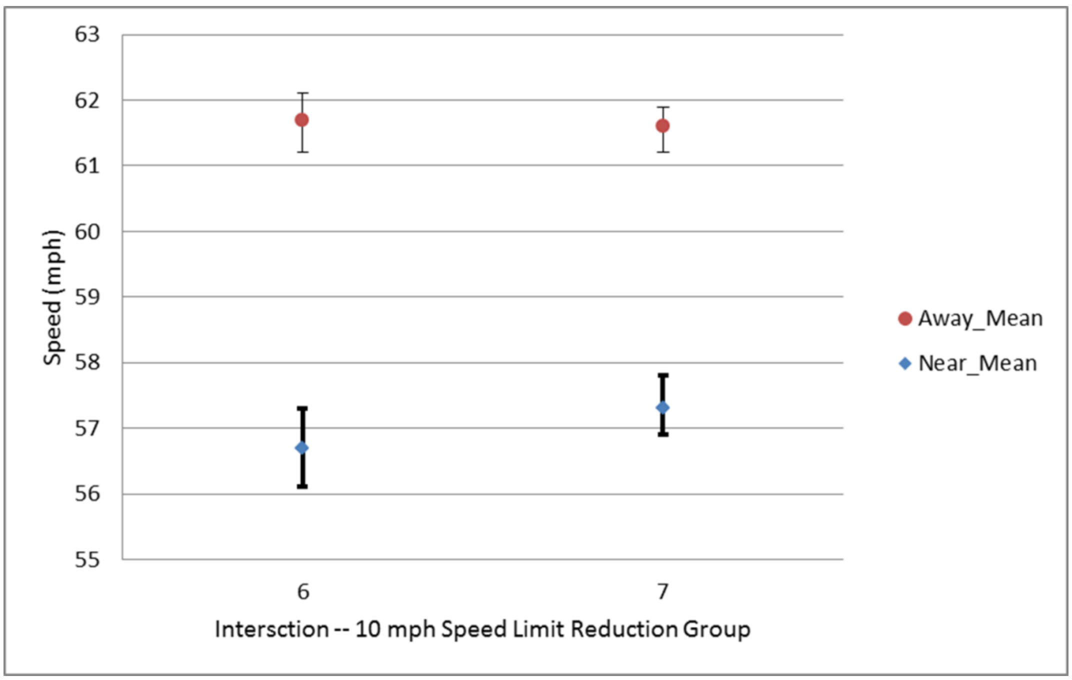


Figure 5-4 Speed Statistics for 10 mph Speed Limit Drop Group

Figure 5-4 describes the speed distribution at intersections with 10 mph speed limit reductions. Compared to Figure 5-2 and Figure 5-3, Figure 5-4 exhibited a stronger difference in mean speed, which demonstrated that drivers tended to slow down when close to the intersection. Mean speeds changed from 61.7 mph to 56.7 mph (a drop of 5.0 mph) and from 61.6 mph to 57.3 mph (a drop of 4.3 mph) as vehicles approached closer to the stop bar at the intersection of US-77 & Saltillo Road and the intersection of US-

281 & West Platte River, respectively. It can be seen that, in the case of 10 mph speed limit reductions, overall speed slowed down near the intersection.

Table 5-3 summarizes the descriptive statistics for the speed measurements obtained near and away from the stop bar for all the test sites. Using the Away from Stop Bar location at intersection # 1 (US-34 & N79) as an example, the speed limit difference was 0 mph, and the mean speed difference was 1.8 mph. Similar to the results obtained from the literature reviews documented in Table 2-1, this thesis demonstrates that the magnitude of changes in mean speeds was not the same as the magnitude of change in speed limit reduction.

Table 5-3 Speed Change between Near Stop Bar and Away from Stop Bar

#	Location	Speed Limit Drop Group	Speed Change	
			Speed Limit	Mean
1	US-34 & N-79	0 MPH Speed Limit Drop Group	0	1.8
2	US-77 & Pioneers		0	-2.5
3	N-133 & N-36		0	-2.6
4	US-75 & Platteview	5 MPH Speed Limit Drop Group	-5	-0.4
5	US-81 & S Lincoln		-5	0.3
6	US-77 & Saltillo	10 MPH Speed Limit Drop Group	-10	-5.0
7	US-281 & Platte River		-10	-4.3

The preliminary analysis of the mean speed suggested an insignificant and inconsistent impact on mean speed resulting from 0 mph and 5 mph speed limit reduction. A speed limit reduction of 10 mph showed a consistent mean speed reduction of 4 to 5 mph. The next section presents the models developed for speed mean and variance to assess the statistical impacts of speed limit reduction while controlling for site specific parameters and distance.

5.2 Speed Analyses Based on Quantile Regression

Wavetronix SmartSensor Advance sensor can continuously track vehicles for a range of 500 ft, triggering multiple detection points for each vehicle. For modeling purposes, only one detection point was randomly selected among the trajectory for each vehicle in order to avoid the problem of auto correlation. In addition, by random selection, the distance factor was also available and could be used as a continuous variable for building models.

5.2.1 Quantile Regression

Quantiles are points taken at regular intervals from the cumulative distribution function of a random dataset. If an ordered dataset is divided into 100 equal-sized data subsets, the k^{th} quantile is the value *x* that the probability that the random variable will be less than *x* is at most k/100, while the probability that the random variable larger than *x* is at most (100-k)/100 (27). For example, 85^{th} quantile of speed (v_85^{th}) indicates that 85% of drivers would choose a speed slower than v_85^{th}.

Quantile regression was first introduced in 1978 (28). Conventional modeling deals with the conditional mean of the dependent variable *y* against the independent variable *x*, while quantile regression can reveal the relationship between an independent variable x and conditional quantiles of a response variable y. Compared to normal linear regression, quantile regression has several advantages (28):

- No distribution assumptions. Roger and Gilbert (28) argued that the conventional least square estimator might be seriously deficient in linear models if the error terms do not normally distribute.
- Robust: Quantile regression is robust in handling extremes and outliers.
- Comprehensive: Traditional linear regression provides information only about the mean, whereas quantile regression can reveal the relationship between an independent variable x and conditional quantiles of a response variable y. Hence, quantile regression can provide a full picture of a dependent variable y. In addition, the speed limit setup is based on 85^{th} percentile speed; a higher speed might result in a more severe accident. These concerns necessitate a complete modeling of the entire dataset, especially for higher quantile speeds.

The basic idea of estimating quantile regression parameters as researched in a paper by the SAS Institute (29) is as described below. For a random variable y with the probability distribution function:

$$F(y) = Prob\ (Y \leq y)$$

The τ th ($0 < \tau < 1$) quantile of y is defined as the inverse function

$$Q(\tau) = inf\ \{y : F(y) \geq \tau\}$$

For a random sample {y1,….,yn} of y, the sample median is the minimizer of the sum of absolute deviations:

$$min_{\xi \in R} \sum_{i=1}^{n} |y_i - \xi|$$

Similar to the sample median, the general τ^{th} sample quantile $\xi(\tau)$ can be calculated as the optimization problem,

$$min_{\xi \in R} \sum_{i=1}^{n} \rho_\tau |y_i - \xi|$$

where $\rho_\tau(z) = z\big(\tau - I(z < 0)\big)$ and $0 < \tau < 1$; and I(.) is the indicator function. The sample mean minimizes the sum of squared residuals by

$$\hat{\mu} = argmin_{\beta \in R} \sum_{i=1}^{n} (y_i - \mu)^2$$

It can be applied in the linear conditional mean function $E(Y|X = x) = x'\beta$ based on the following equation:

$$\hat{\beta} = argmin_{\beta \in R^P} \sum_{i=1}^{n} (y_i - x_i'\beta)^2$$

Then, the linear conditional quantile function $Q(\tau|X = x) = x'\beta(\tau)$ can be estimated by:

$$\hat{\beta} = argmin_{\beta \in R^P} \sum_{i=1}^{n} \rho_\tau (y_i - x_i'\beta)$$

$$= argmin_{\beta \in R^P} \{ \sum_{i \in \{i | y_i \geq x_i'\beta\}} \tau |y_i - x_i'\beta| + \sum_{i \in \{i | y_i < x_i'\beta\}} (1 - \tau) |y_i - x_i'\beta| \}$$

To evaluate the model's goodness of fit, an analog of the R squared statistic can be developed for quantile regression models (30).

$$V^1(\tau) = \sum_{i=1}^{n} \rho_\tau (y_i, \hat{y}_i) = \sum_{i \in \{\boldsymbol{i} | \boldsymbol{y_i} \geq \mathrm{x}_\mathrm{i}'\beta\}} \boldsymbol{\tau} |\mathrm{y_i} - \mathrm{x}_\mathrm{i}'\beta| + \sum_{i \in \{i | y_i < \mathrm{x}_\mathrm{i}'\beta\}} (1-\tau)|\mathrm{y_i} - \mathrm{x}_\mathrm{i}'\beta|$$

$$V^0(\tau) = \sum_{i=1}^{n} \rho_\tau \left(y_i, \hat{Q}^{(\tau)}\right) = \sum_{i \in \{i | y_i \geq \bar{\mathrm{y}}\}} \tau \left|\mathrm{y_i} - \hat{Q}^{(\tau)}\right| + \sum_{i \in \{i | y_i < \bar{\mathrm{y}}\}} (1-\tau)\left|\mathrm{y_i} - \hat{Q}^{(\tau)}\right|$$

For the model that has only a constant term, the fitted constant is the sample p^{th} quantile $\hat{Q}^{(\tau)}$ for the sample [y_1, y_2, …y_n]. The goodness of fit is then defined as,

$$R(\tau) = 1 - \frac{V^1(\tau)}{V^0(\tau)}$$

$R(\tau)$ is within the range between 0 and 1, and a larger $R(\tau)$ indicates a better model fit.

5.2.2 Speed Model

Quantile regression can reveal the relationship between independent variables *x* and each of the conditional quantiles of the response variable *y* (both lower and upper, or all quantiles). As shown in the following section, several variables were tested in the speed model, and the description for each is described below.

Dependent Variable Speed: Wavetronix SmartSenor Advance senor's capacity enables chosen spot speed randomly for each vehicle to avoid the problem of auto correlation. In total, 14,169 vehicles were detected. There were 6,096 data points for the constant speed limit group (0 mph speed limit drop); 5,485 vehicles were detected in the group with the 5 mph speed limit drop; and there were 2,588 observations for the 10 mph reduction group.

Distance from Vehicle to Stop Bar: Aside from speed detection, Wavetronix SmartSenor Advance sensor can track vehicle location. This factor is the distance in feet between the detected vehicle and the stop bar.

5-minute Volume: Wavetronix sensor can log each vehicle and its associated timestamp; therefore, 5 minutes of volume data could be derived. Figure 5-5, Figure 5-6, and Figure 5-7 illustrate 5-minute volume distribution with the function of time of day. The x-axis is time of day and the y-axis is the average 5-minute volume over one hour.

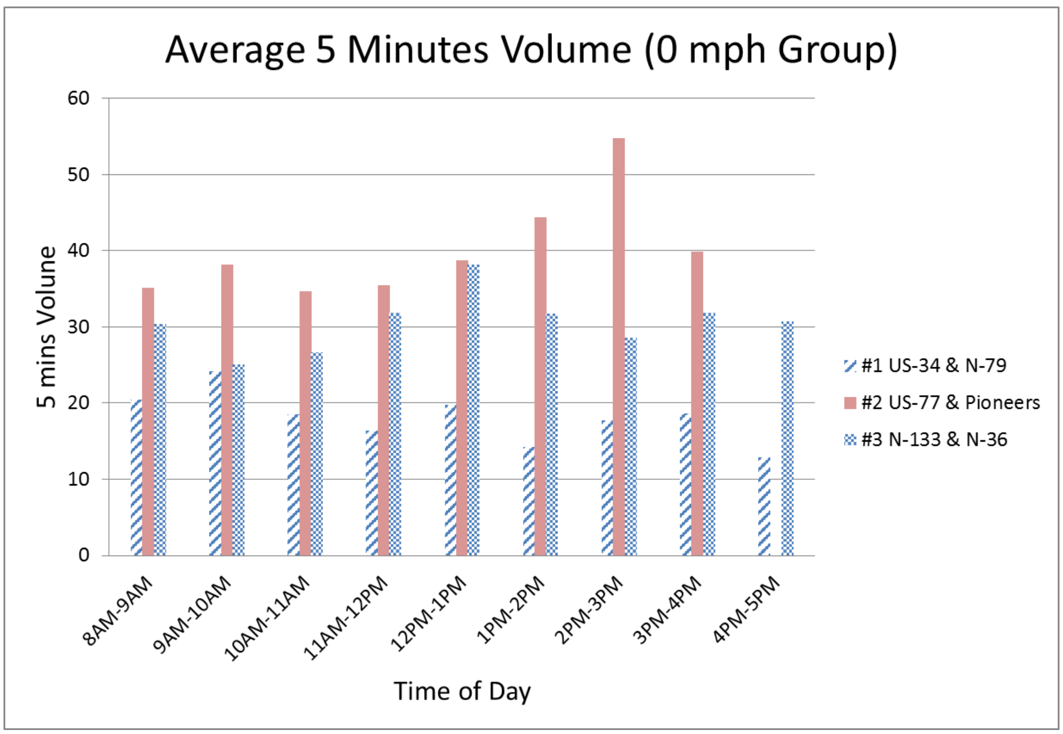


Figure 5-5 Distribution of Average 5 Minutes Volume for 0 mph Reduction Group

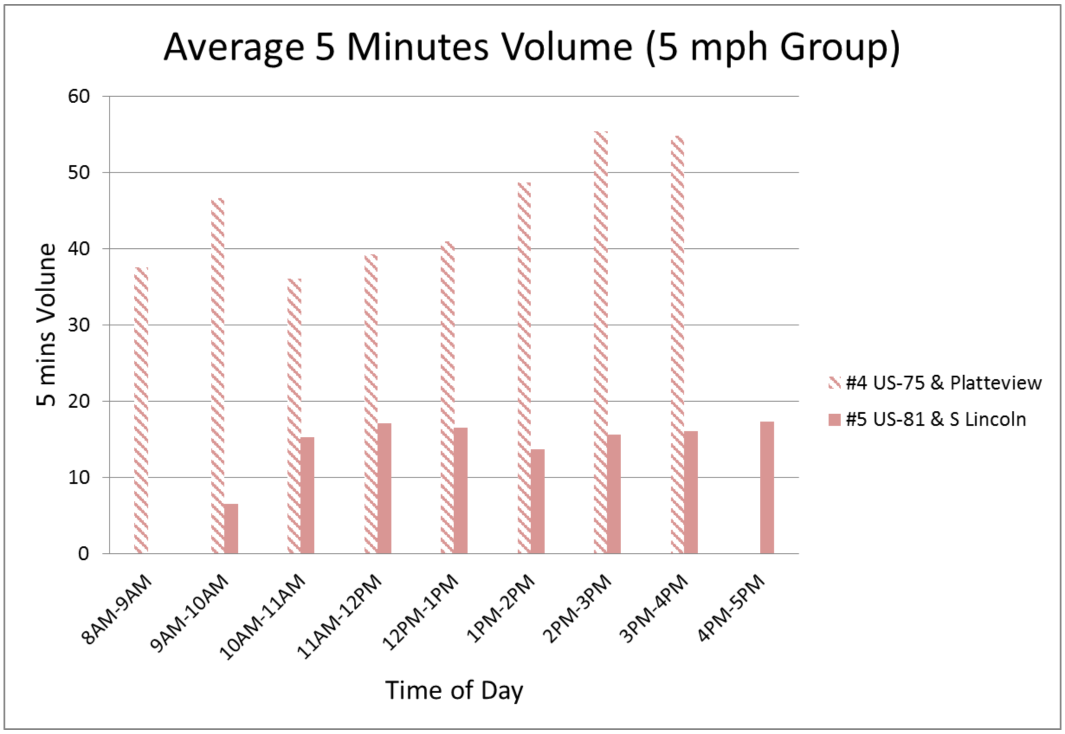


Figure 5-6 Distribution of Average 5 Minutes Volume for 5 mph Reduction Group

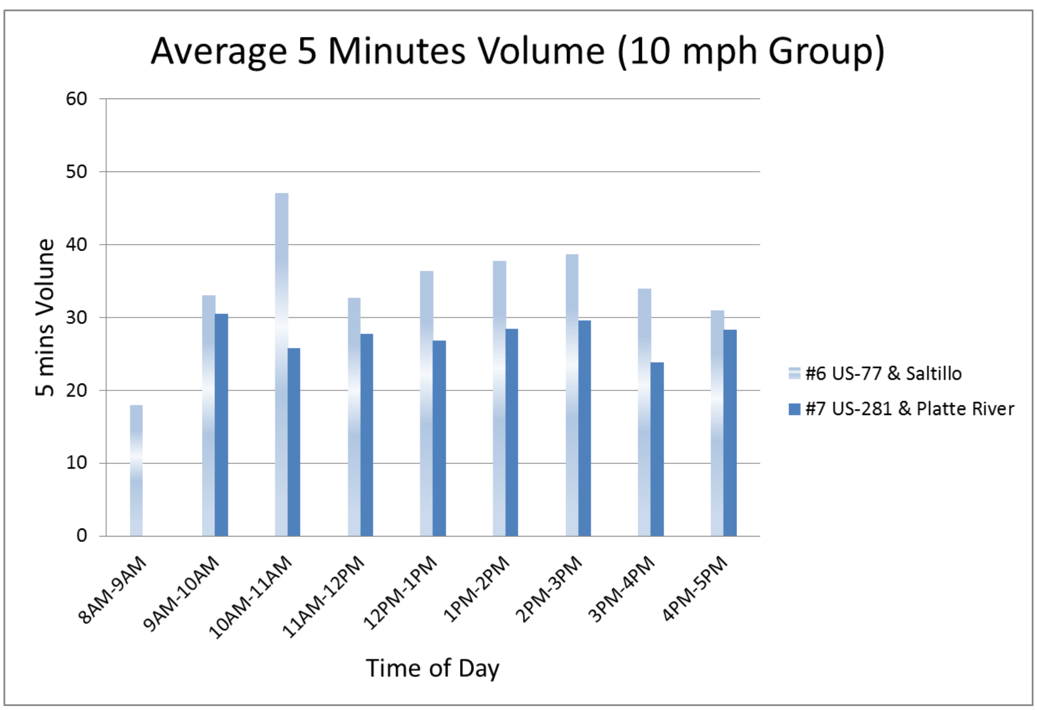


Figure 5-7 Distribution of Average 5 Minutes Volume for 10 mph Reduction Group

Aside from the continuous independent variable introduced above (Distance from Vehicle to Stop Bar and 5-minute volume), there are other indicator variables where the value is either 1 or 0.

Time of day indicator: Most of the data collection were conducted between 8 AM and 5 PM (Intersection #5 and # 7 do not have data available before 9 AM; intersection #2 and #4 do not have data available after 4 PM). Three time slots were evaluated in the model, before11 AM, 11AM ~ 3 PM, and 3PM~5 PM. For example, the 3PM~5 PM indicator is equal to 1 if vehicles were collected between 3 PM and 5 PM, 0 for other times of day.

Intersection Indicator: Each intersection was assigned an intersection indicator factor. Among the 7 targets, two intersections are noteworthy. One was the intersection of US34 & N79, where the near stop bar speed limit was 60 mph while the remaining 6 intersections had a near stop bar speed limit of 55 mph. In addition, this intersection (US34 & N79) was a T intersection, as was the intersection of US81 & S Lincoln. Features of 60 mph speed limit and T intersections may affect driver speed choices; as such, Intersection Indicator was used in the modeling analysis.

5.2.2.1 Interpretation of Results for 0 mph Speed Limit Reduction Group

The first model was constructed for intersections with a constant speed limit. 3 quantiles should be noted. The first is 85th quantile speed, which was related to the speed limit setup; next is 50th quantile speed, which was close to mean speed; the third is 15th quantile speed. Associated with 85th quantile speed, the difference between 15th and 85th speed can give an estimation of the speed variance. Therefore, the numerical models for these three quantiles were also provided separately. Table 5-4 lists the variables for this model, as well as each variable's statistics, including mean, standard deviation, minimum, and maximum.

Table 5-4 List of Variables for the Speed Model (0 mph Speed Limit Reduction Group)

Variable	Definition	Mean	Standard Deviation	Min	Max
Dependent Variable					
Speed	6,096 vehicles were detected for the 0 mph reduction group	58.0	5.9	16	99
Independent Variables					
Distance from Vehicle to Stop Bar	This factor is the real distance in feet between the detected vehicle and the intersection stop bar.	1035.8	411.7	535	2040
5-minute Volume	5-minutes volume associated with each detected vehicle. Not the average volume over one hour.	32	14.7	2	81
Before 11 AM Indicator	Equal to 1 if vehicles were collected between 8 AM and 11 AM	0.3	0.5	0	1
11AM-3 PM indicator	Equal to 1 if vehicles were collected between 11 AM and 3 PM	0.5	0.5	0	1
3 PM–5 PM Indicator	Equal to 1 if vehicles were collected between 3 PM and 5 PM	0.2	0.4	0	1
US-34 & N-79 Indicator	T intersection, with a higher speed limit of 60 mph, whereas the remainder were 55 mph.	0.3	0.5	0	1
US-77 & Pioneers Indicator	4-leg intersection, with a speed limit of 55 mph	0.5	0.5	0	1
N-133 & N-36 Indicator	4-leg intersection, with a speed limit of 55 mph	0.2	0.4	0	1

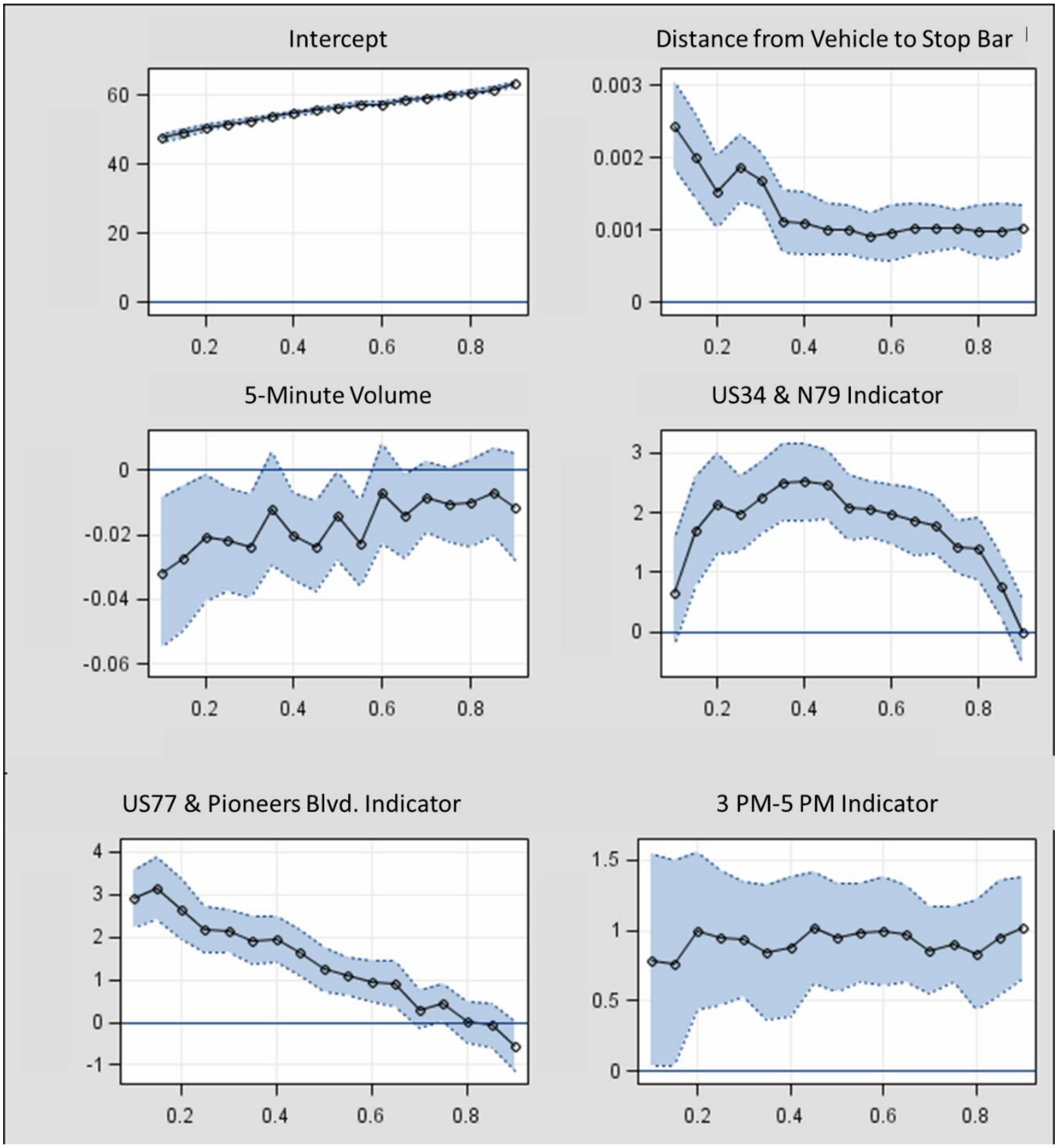


Figure 5-8 Speed Model of 0 MPH Group

Figure 5-8 shows the quantile plots of the speed model based on the 0 mph speed limit reduction group. These plots were built for the estimated coefficient (y-axis) of the tested variable and desired quantiles (x-axis). The 95% confidence interval bands are shaded. The shaded bands can be used to determine whether a variable was statistically significant parameter for a given speed quantile or not. Only the variables which were statistically significant for at least one quantile (10th, 15th, 50th, 85th, 90th) are chosen in the model. A variable was not significant for a given quantile if the confidence interval covered the zero-line, and vice versa. For the 0 mph model, 5 variables had a significant impact on speed —Distance from Vehicle to Stop bar, 5-minute volume, 2 Intersection indicators, and 3 PM-5 PM indicator.

The coefficient of Distance from Vehicle to Stop Bar was positive, showing that vehicles slowed down when they were close to the stop bar. Supposing two locations are tested, location A is 1,000 ft away from the stop bar, while location B is 0 feet away from the stop bar. In this case, the speed difference between location A (away from stop bar) and B (near stop bar) can be calculated by multiplying the coefficient by 1,000 (ft). The result showed that for a distance of 1,000 ft, 15th quantile speed was reduced to approximately 2.0 mph, whereas the speed reduction for the 50th and 85th quantiles was close to 1.0 mph. This suggests that lower-speed drivers tended to slow down even more than higher speed drivers. Such a difference in behavior between lower-speed drivers and higher-speed drivers as they approach closer to the intersection leads to a higher variability in speed closer to the intersection. The increased variance closer to the intersection in the absence of any speed limit sign could lead to a higher safety risk.

The coefficient of 5-minute Volume was negative, indicating that vehicles slowed down with a heavier traffic condition. This factor has significant effect over the lower quantile speed (smaller than 60th quantile) and has no statistical effect for the higher speed drivers.

The indicators for the intersection of US34 & N79 and intersection US77 & Pioneers Blvd. were significant, and had a positive effect on speed over most quantiles.

The positive effect indicates that at the same location away from the stop bar, the speeds at intersection US34 & N79 and intersection US77 & Pioneers Blvd. are higher than the speed at intersection #3 N133 & N36.

The last significant variable was the 3 PM-5 PM indicator. The positive effect indicated that drivers tended to choose a higher speed after 3 PM.

The coefficients of the quantile regression model estimated by SAS software are listed in Table 5-5. Based on Table 5-5, the numerical speed models for the 15th, 50th, and 85th quantiles are presented below.

Table 5-5 Coefficient Estimation based on Quantiles for Speed Model of 0 MPH Group

Variables	Quantiles				
	10th	15th	50th	85th	90th
Intercept	47.6601	49.0723	56.4936	61.6723	63.3497
Distance to Stop Bar	0.0024	0.0020	0.0010	0.0010	0.0010
5-minute Volume	-0.0318	-0.0272	-0.0139	*	*
US 34 & N79 Indicator	*	1.6886	2.0939	0.7699	*
US-77 & Pioneers Indicator	2.9046	3.1398	1.2484	*	*
3 PM – 5 PM Indicator	0.7873	0.7635	0.9470	0.9500	1.0162
R^2	0.9999	0.9998	0.9996	0.9997	0.9998
* Not significant at 95% level of significance					

$$Speed_{15th} = 49.0723 + 0.002 * Distance\ to\ Stopbar - 0.0272 * 5minuteVolume + 1.6886 * US34\&N79\ Indicator + 3.1398 * US77\&Pioneers\ Indicator + 0.7635 * 3PM{\sim}5PM\ Indicator$$

$$Speed_{50th} = 56.4936 + 0.001 * Distance\ to\ Stopbar - 0.0139 * 5minuteVolume + 2.0939 * US34\&N79\ Indicator + 1.2484 * US77\&Pioneers\ Indicator + 0.947 * 3PM{\sim}5PM\ Indicator$$

$$Speed_{85th} = 61.6723 + 0.001 * Distance\ to\ Stopbar + 0.7699 * US34\&N79\ Indicator + 0.95 * 3PM{\sim}5PM\ Indicator$$

5.2.2.2 Interpretation of Results for 5 mph Speed Limit Reduction Group

Preliminary study (section 5.1) showed an insignificant impact on mean speed resulting from a reduction of 5 mph in the speed limit. The following model was built to control for the variables of site-specific effects and distance. Table 5-6 lists the variables for 5 mph speed limit reduction model, as well as each variable's statistics.

Table 5-6 List of Variables for the Speed Model (5 mph Speed Limit Reduction Group)

Variable	Definition	Mean	Standard Deviation	Min	Max
Dependent Variable					
Speed	5485 vehicles were detected for the 5 mph reduction group	57.5	6.0	27	96
Independent Variables					
Distance from Vehicle to Stop Bar	This factor is the real distance in feet between the detected vehicle and the intersection stop bar.	1145.3	494.6	555	2055
5-minute Volume	5-minutes volume associated with each detected vehicle. Not the average volume over one hour	41.5	16.6	3	110
Before 11 AM Indicator	Equal to 1 if vehicles were collected between 8 AM and 11 AM	0.2	0.4	0	1
11 AM ~3 PM indicator	Equal to 1 if vehicles were collected between 11 AM and 3 PM	0.7	0.5	0	1
3 PM-5PM Indicator	Equal to 1 if vehicles were collected between 3 PM and 5 PM	0.1	0.3	0	1
US-75 & Platteview	4-leg intersection, with a speed limit of 55 mph	0.9	0.3	0	1
US-81 & S Lincoln	T intersection, with a speed limit of 55 mph	0.1	0.3	0	1

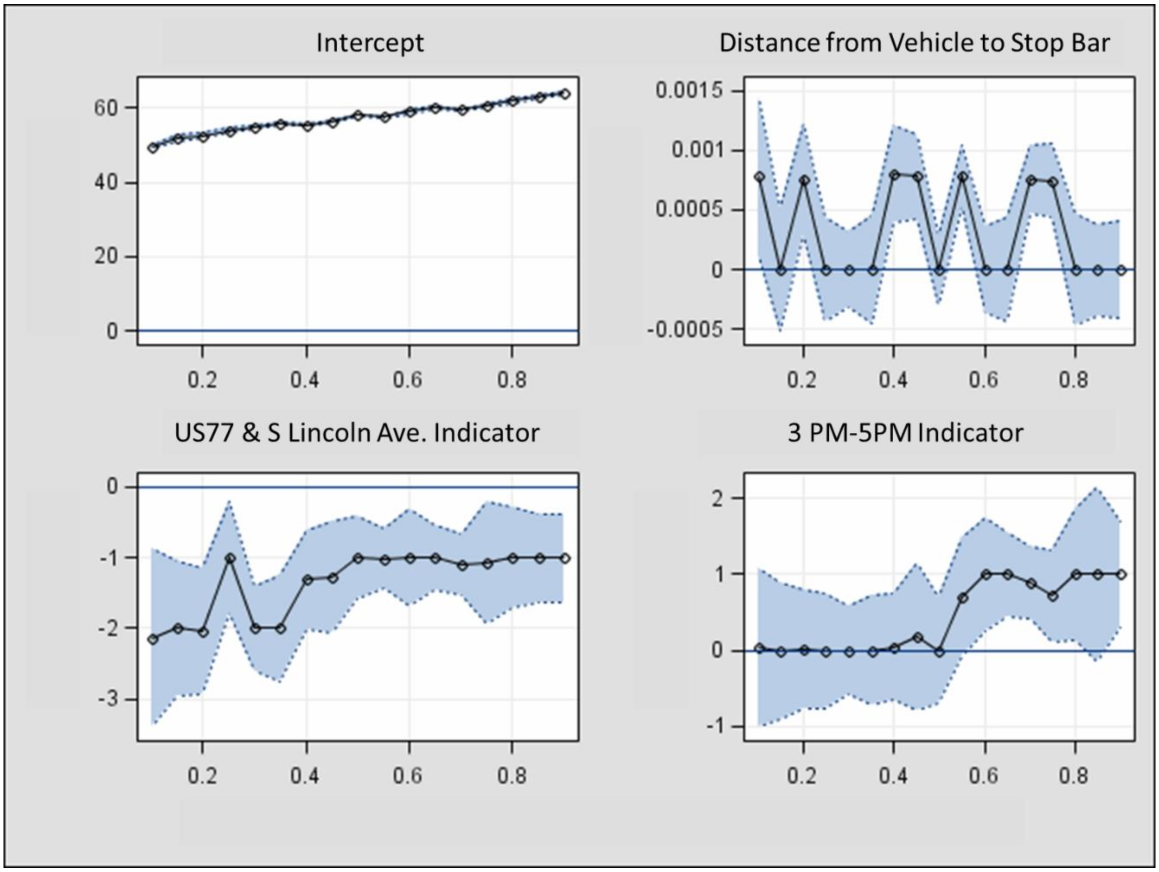


Figure 5-9 Speed Model of 5 MPH Group

Figure 5-9 presents the quantile plots of the speed model based on the 5 mph speed limit reduction group. The y-axis is the estimated coefficient of the tested variable, and the x-axis is the desired quantiles. The 95% confidence interval bands are shaded. The shaded bands are used to determine whether a variable was a statistically significant parameter for a given speed quantile. Only the variables that were statistically significant for at least one quantile (10th, 15th, 50th, 85th, 90th) are shown in the model. A variable was not significant for a given quantile if the confidence interval covered the zero-line, and vice versa. In the 5 mph group speed model, 2 factors were found to have a significant effect on driver speed. As for the distance variable, it had little effect (a significant effect was demonstrated for the 10th quantile only).

Distance from Vehicle to Stop Bar was not significant, as most of the confidence interval band crossed the value 0. This finding was in agreement with the mean speed analysis in section 5.1, indicating that little speed variation occurred along the roadway.

The US 77 & Lincoln Ave indicator was negative, which meant that drivers chose to slow down at this intersection. This site was a T intersection, where the left turn lane led to York City, Nebraska. Some vehicles probably made a left turn here, slowing down the overall speed.

Similar to the 0 mph group, the 3 PM-5 PM indicator was positive; however, this had a significant effect only upon the higher speed drivers. The estimated coefficients for the 5 mph group are listed in Table 5-7. Based on Table 5-7, numerical speed models for 15^{th}, 50^{th}, and 85^{th} quantiles were also listed.

Table 5-7 Coefficient Estimation based on Quantiles for Speed Model of 5 MPH Group

Variables	Quantiles				
	10^{th}	15^{th}	50^{th}	85^{th}	90^{th}
Intercept	49.5039	52.0000	58.0000	63.0000	64.0000
Distance to Stop Bar	0.0008	*	*	*	*
US77&Lincoln Ave Indicator	-2.1378	-2.0000	-1.0000	-1.0000	-1.0000
3 PM-5PM Indicator	*	*	*	1.0000	1.0000
R^2	0.9988	0.9980	0.9987	0.9997	0.9997
*Not significant at 95% level of significance					

$$Speed_{15th} = 52 - 2 * US77\&Lincoln\ Ave\ Indicator$$

$$Speed_{50th} = 58 - US77\&Lincoln\ Ave\ Indicator$$

$$Speed_{85th} = 63 - US77\&Lincoln\ Ave\ Indicator + 3\ PM{\sim}5PM\ Indicator$$

5.2.2.3 Interpretation of Results for 10 mph Speed Limit Reduction Group

Following the development of the models for the 0 mph and 5 mph speed limit reduction groups, the model for the 10 mph speed limit reduction group could be described as below. Table 5-8 lists the variables for 10 mph speed limit reduction model, as well as each variable's statistics.

Table 5-8 List of Variables for the Speed Model (10mph Speed Limit Reduction Group)

Variable	Definition	Mean	Standard Deviation	Min	Max
Dependent Variable					
Speed	2588 vehicles were detected for the 10 mph reduction group	59.6	7.1	13	96
Independent Variables					
Distance from Vehicle to Stop Bar	This factor is the real distance in feet between the detected vehicle and the intersection stop bar.	1518.5	681.8	545	2625
5-minute Volume	5-minutes volume associated with each detected vehicle. Not the average volume over one hour.	31.9	10.2	7	69
Before 11 AM Indicator	Equal to 1 if vehicles were collected between 8 AM and 11 AM	0.2	0.4	0	1
11 AM ~3 PM indicator	Equal to 1 if vehicles were collected between 11 AM and 3 PM	0.7	0.5	0	1
3 PM-5PM Indicator	Equal to 1 if vehicles were collected between 3 PM and 5 PM	0.1	0.3	0	1
US-77 & Saltillo Rd. Indicator	4-leg intersection, with a speed limit of 55 mph	0.5	0.5	0	1
US-281&PlatteRiver Indicator	4-leg intersection, with a speed limit of 55 mph	0.5	0.5	0	1

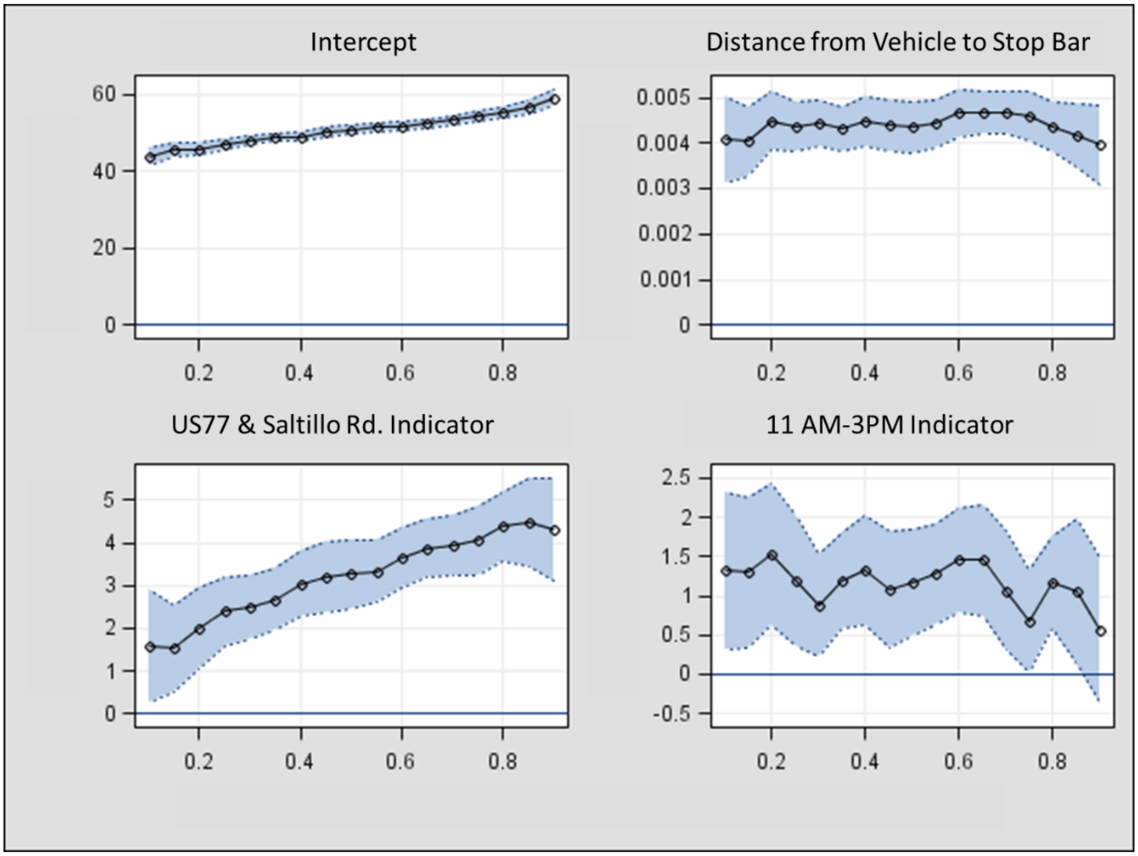


Figure 5-10 Speed Model of 10 MPH Group

Figure 5-10 shows the quantile plots of the speed model based on the 10 mph speed limit reduction group. The y-axis is the estimated coefficient of the tested variable, and the x-axis is the desired quantiles. The 95% confidence interval bands are shaded, which can be used to determine whether a variable was a statistically significant parameter for a given speed quantile. Only the variables which were statistically significant for at least one quantile (10^{th}, 15^{th}, 50^{th}, 85^{th}, 90^{th}) are shown in the model. A variable was not significant for a given quantile if the confidence interval covered the zero-line, and vice versa. In this model, 3 factors were found to have a significant effect on speed—Distance from Vehicle to Stop Bar, intersection indicator, and 11 AM to 3 PM indicator.

Distance to Stop bar was positive. Once timed by 1,000 ft, it shows that the 15th and 50th quantile speeds dropped by 4.0 mph and 4.3 mph after vehicles travelled 1,000 ft. As for 85th quantile speed, the reduction was 4.2 mph after vehicles travelled for 1,000 ft. Compared with the 0 mph group, this factor had a stronger influence in that drivers tended to adjust their speed to a lower level once they encountered the reduced speed limit sign. In addition, the trend line was fairly constant. Hence, the speed difference between higher speed drivers and lower speed drivers was relatively smaller than at the 0 mph site, suggesting a safer condition.

Intersection US77 & Saltillo Rd. indicator is significant, suggesting that the speed is higher at this intersection than the speed at intersection US281 & W Plate River Dr. At first, this is counterintuitive based on the bootstrap plot Figure 5-4; however, notice that the model is built based on the distance to stop bar, where the data collection location at intersection #7 is farther away from the stop bar than that at intersection #6. Hence, this indicator illustrate that at the same location away from the stop bar, the speed at intersection US77 & Saltillo Rd is faster than the speed at intersection US281 & W Plate River Dr.

The 11 AM to 3 PM indicator was significant for most of the speed quantiles, illustrating that during the time between 11 AM and 3 PM, the speed is higher than the other time of day. Table 5-9 shows the detailed statistics for each variable and numerical speed models for quantiles of 15th, 50th, and 85th are listed as well.

Table 5-9 Coefficient Estimation based on Quantiles for Speed Model of 10 MPH Group

Variables	Quantiles				
	10th	15th	50th	85th	90th
Intercept	43.7755	45.6263	50.8696	56.7708	59.0595
Distance to Stop Bar	0.0041	0.0040	0.0043	0.0042	0.0040
US77&Saltillo Indicator	1.5918	1.5253	3.2609	4.4792	4.3056
11AM-3 PM Indicator	1.3265	1.3030	1.1739	1.0625	*
R^2	0.9980	0.9986	0.9994	0.9999	0.9999
*Not significant at 95% level of significance					

$$Speed_{15th} = 45.6263 + 0.004 * Distance\ to\ Stopbar + 1.5253 * \text{US77\&Saltillo Indicator} + 1.303 * 11AM \sim 3\ PM\ Indicator$$

$$Speed_{50th} = 50.8696 + 0.0043 * Distance\ to\ Stopbar + 3.2609 * \text{US77\&Saltillo Indicator} + 1.1739 * 11AM \sim 3\ PM\ Indicator$$

$$Speed_{85th} = 56.7708 + 0.0042 * Distance\ to\ Stopbar + 4.4792 * \text{US77\&Saltillo Indicator} + 1.0625 * 11AM \sim 3\ PM\ Indicator$$

5.2.2.4 Summary of Speed Model

The speed model demonstrated that two groups (the 0 mph and 10 mph speed limit reduction groups) showed significant speed reduction at the proposed quantiles (15^{th}, 50^{th}, and 85^{th}) at locations close to the intersection. As for the 5 mph speed limit reduction group, there was no significant speed reduction as vehicles approached the intersection.

In addition, for the constant speed limit group, it was suggested that speed variance increased between higher speed drivers and lower speed drivers, as evidenced by the decreasing coefficient curve seen in Figure 5-8. However, for the 10 mph speed limit reduction group, the distance variable had a stronger influence, and the trend line (Figure 5-10) was fairly constant; hence, the speed variance was relatively smaller than that of the 0 mph site, and suggested a safer condition.

5.3 Summary

In this chapter, different methods were applied to the analysis of roadway speed characteristics. The preliminary mean speed analysis suggested an insignificant impact on mean speed by a reduction of 5 mph in the speed limit. On the other hand, the group with a 10 mph speed limit reduction showed a consistent mean speed reduction from 4 to 5 mph. Models were developed for speed to assess the statistical impacts of speed limit reduction while controlling for site specific parameters and distance.

This study is most intersected in the impact of speed limit reduction when it is close to the intersection. In the speed model, the continuous variable Distance from Vehicle to Stop Bar was found to have the significant effect on speed reduction when vehicles approaching an intersection. By timing the estimated coefficient of this variable at 1,000 ft, speed reduction (within 1000 ft) curves were created in the function of quantiles. Figure 5-11, Figure 5-12, and Figure 5-13 illustrate the amount of speed reduced within 1,000 ft as a function of different quantiles. The y-axis represents the quantile from 10^{th} to 90^{th}, and the x-axis represents the speed at two locations. The Near Stop bar location curve is marked by the blue circle and the Away from Stop Bar location (1,000 ft away from Near Stop Bar Location) is marked by the red diamond. In addition, the difference between 85^{th} and 15^{th} quantile speeds is also calculated, because this difference can be used to roughly estimate the speed dispersion.

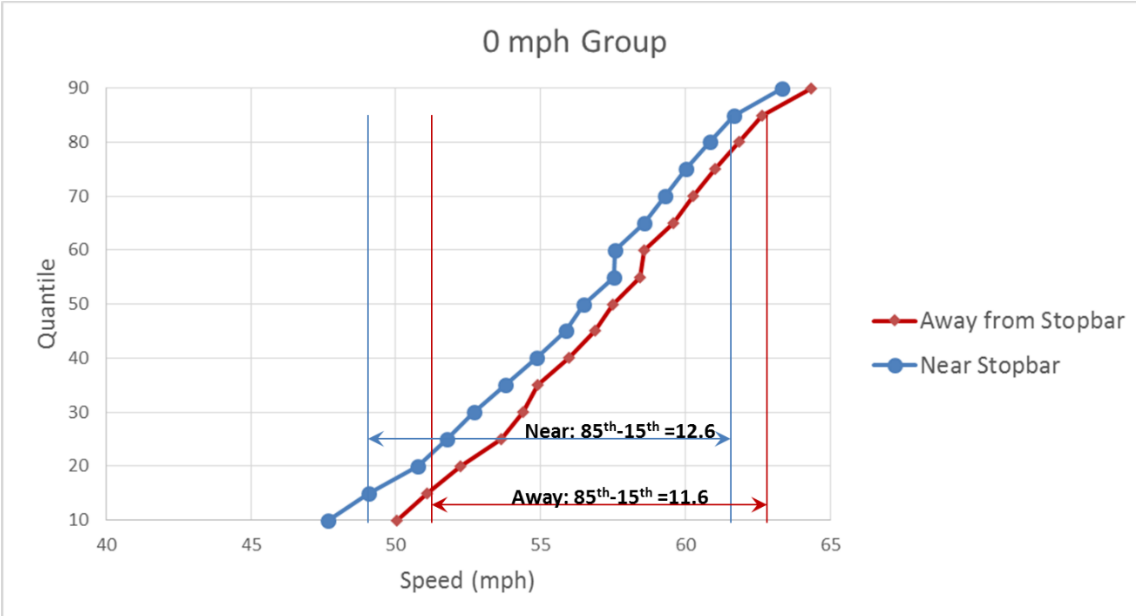


Figure 5-11 Speed Reduction Curve based on Quantiles within 1,000 ft (0 mph Group)

For the 0 mph speed limit reduction curve (Figure 5-11), the sites with a constant speed limit showed a speed drop of around 1 mph for high speed drivers (higher quantile >50th). However, this location factor had a stronger effect on lower quantile speed (i.e. slow speed drivers, 10th to 15th quantile), where the reduction is around 2.4 mph. The uneven speed reductions suggested that high speed drivers still tended to maintain a relatively high level of speed, while slow speed drivers tended to exhibit a greater decrease in speed when they were approaching an intersection. This further enlarged the speed variance, which is an unsafe factor on the roadway.

In addition, this unsafe roadway condition can be also proved by the speed difference between 85th and 15th quantiles speeds as marked in Figure 5-11. The difference increased from 11.6 mph (away from stop bar) to 12.6 mph (near stop bar), which suggested an increased speed variance when it is close to the stop bar.

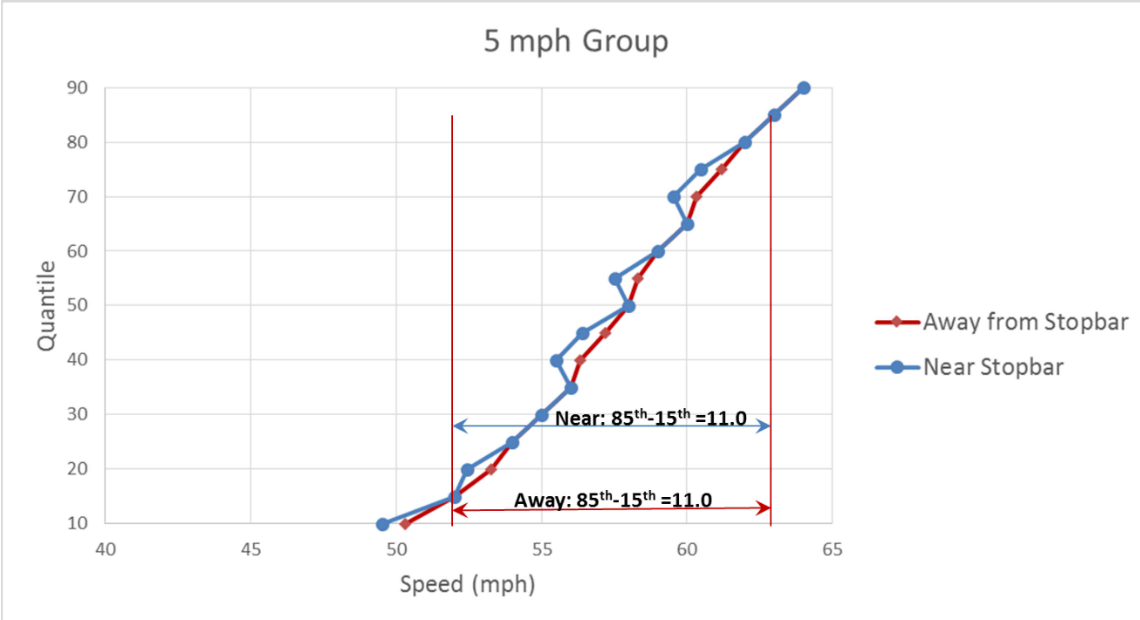


Figure 5-12 Speed Reduction Curve based on Quantiles within 1,000 ft (5 mph Group)

As for the 5 mph reduction group, the effect of speed reduction was less significant as evidenced by the almost overlapped curves between two locations (Figure 5-12). This suggested that most drivers tended to maintain a constant speed when they were close to the intersection.

In addition, the speed differences between 85th and 15th quantiles speeds at two locations remain the same, suggesting that the speed variance is stable when it is close to the intersection.

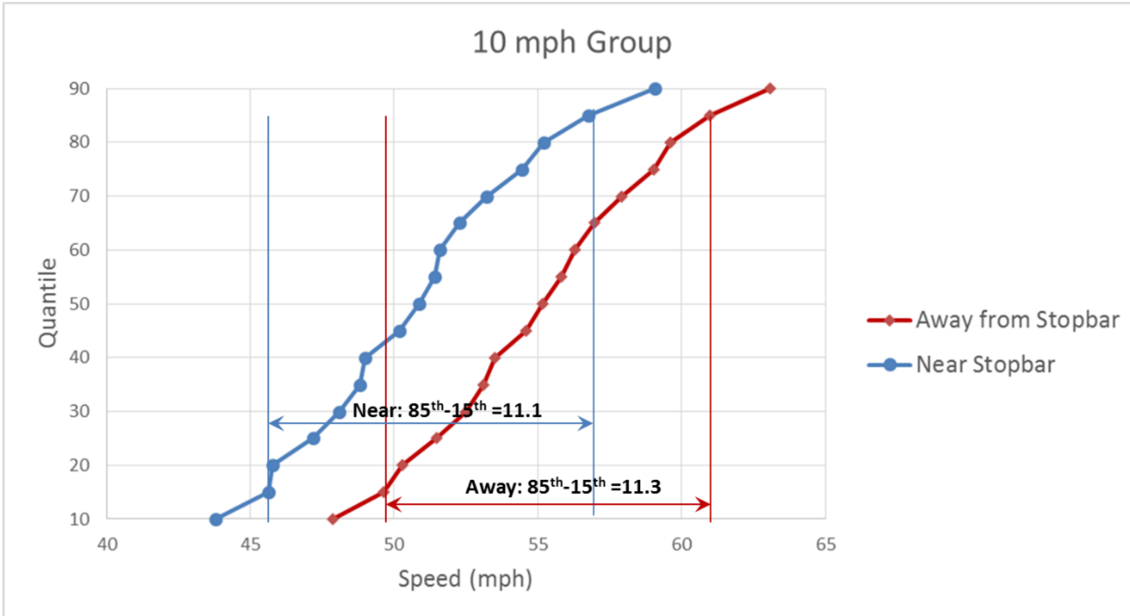


Figure 5-13 Speed Reduction Curve based on Quantiles within 1,000 ft (10 mph Group)

Intersections with a 10 mph speed limit reduction demonstrated the strongest reduction effect on driver speed, where overall speed was reduced by approximately 4 mph (Figure 5-13). Moreover, compared to sites with a constant speed limit, the speed reduction for higher speed drivers and lower speed drivers is fairly constant, illustrates a smaller speed variance among vehicles, resulting in safer roadway conditions.

Compared with 0 mph group, the 85th and 15th quantiles speed differences at two locations remain relative the same, with a little drop from 11.3 mph (away from stop bar) to 11.1 mph (near stop bar). The decreased 85th-15th-quantile-difference suggests a safer roadway condition when it is close to intersection.

The aforementioned speed models demonstrate that drivers tended to decelerate when they drove toward intersections having a constant speed limit or a 10 mph speed limit reduction. In the 5 mph speed limit reduction group, however, drivers tended to maintain their speeds when they were approaching the intersection. Table 5-10 summarizes the results from the models.

Table 5-10 Summary of the Models

Performance Measure Modeled		Quantiles		
		15th	50th	85th
Speed	0 mph	↓	↓	↓
	5 mph	*	*	*
	10 mph	↓	↓	↓
*Not significant at 95% level of significance				

CHAPTER 6 CONCLUSIONS

Speed studies are important in transportation engineering because they provide data to inform speed limit setup and safety analyses. The latest ITS application, the Wavetronix SmartSensor Advance, is useful for data collection to serve this purpose. As opposed to traditional loop detector and manual data collection methods, the Wavetronix SmartSensor Advance sensor can collect data more effectively and precisely. However, ITS technology may yet not be fully developed, and its performance is subject to many factors. Hence, evaluating the performance of Wavetronix before applying the data was a must.

The GPS speed comparison method confirmed the capability of the Wavetronix SmartSensor Advance to detect a single vehicle. However, many ambiguous calls were observed during data collection. This study applied a machine learning technology to obtain desirable data, consequently addressing this issue.

Once the desirable datasets were available, basic statistics on speed characteristics over the 7 selected intersections could be derived. Bootstrap mean speed analysis indicated results similar to those obtained in previous research, i.e., regardless of the change in speed limit, overall speed differences between the Away from Stop Bar and Near Stop Bar locations were fairly small. Drivers adjusted their vehicle speeds based mainly upon roadway conditions, not speed limit signs.

In chapter 1, two hypotheses were stated: a null hypothesis (speed remain constant when vehicles approach an intersection during the green time) and an alternative hypothesis (when driving toward an intersection, driver speed change during the green time). Quantile regression provided an analytical framework of testing these hypotheses. For the speed model, the null hypothesis was rejected, because speed changed when vehicles approached an intersection during the green time for the sites with 0 and 10 mph speed limit reduction. However, this factor was less significant for the 5 mph speed limit reduction group, and the result showed that vehicles maintained constant speeds when close to the intersection at the 15th, 50th, and 85th quantiles.

Future research should combine speed and speed variance studies with safety data to derive a better understanding of traffic safety. The impacts were also dependent on high-speed versus low-speed vehicles, suggesting that vehicle type might also be a crucial factor compounding the impacts.

REFERENCES

1. http://www-fars.nhtsa.dot.gov/Vehicles/VehiclesAllVehicles.aspx Accessed on November 3rd 2012.
2. National Highway Traffic Safety Administration. *Traffic Safety Facts 2009*. Accessed on March 1st 2012. http://www-nrd.nhtsa.dot.gov/cats/listpublications.aspx?Id=E&ShowBy=DocType
3. Nebraska Interagency Safety Committee. *Nebraska Strategic Highway Safety Plan Guidance for 2012-2016.* September 2012.
4. Justice Appiah, Laurence R. Rilett, Zifeng Wu. *Evaluation of NDOR's Actuated Advance Warning Systems*. Nebraska Transportation Center, 2011
5. FHWA. *Manual on Uniform Traffic Control Devices for Streets and Highways*. U.S. Department of Transportation, 2009
6. *Special Report 254Managing Speed—Review of Current Practice for Setting and Enforcing Speed Limit*. Transportation Research Board, National Research Council, Washington, D.C. 1998
7. Alexander C. Wagenaar, Frederic M. Streff, and Robert H. Schultz. *Effects of the 65 MPH Speed Limit on Injury Morbidity and Mortality*. Accident Analysis & Prevention Volume 22, Issue 6, December 1990
8. AASHTO. *A Policy on Geometric Design of Highways and Streets*. 2004
9. U.S. Department of Transportation, *Effects of Raising and Lowering Speed Limits on Selected Roadway Sections*. Publication FHWA-RD-97-084, FHWA, 1997
10. K.R. Agent. *Evaluation of 70 mph Speed Limit in Kentucky*. Kentucky Transportation Center, May 2008

11. K.R. Agent, Jerry G. Pigman, Joel M. Weber, *Evaluation of Speed Limits in Kentucky*. Kentucky Transportation Center, Kentucky Transportation Cabinet, Commonwealth of Kentucky, and FHWA, April 1997

12. Hojun Son, Michael D. Fontaine, and Byungkyu Park, *Long-Term Speed Compliance and Safety Impacts of Rational Speed Limits*. Journal of Transportation Engineering. Aug 2009

13. Jenna Buddemeyer, Rhonda K. Young, Brendan Dorsey. Rural Variable Speed Limit System for Southeast Wyoming. Transportation Research Record: Journal of the Transportation Research Board. No.2189, Washington, D.C.

14. Mohsen Towliat, Helena Svensson, Gunnar Lind, et al. *"Variable Speed Limits at Intersections-Effects and Experience"*, European Transportation Conference 2006.

15. Christopher M. Monsere, Casey Nolan, Robert L. Bertini, Edward L. Anderson, Tarek Abou El-Seoud. *Measuring the Impacts of Speed Reduction Technologies: A Dynamic Advanced Curve Warning System Evaluation*, 84th Annual Meeting of the Transportation Research Board. January, 2005.

16. David Solomon, Accidents on Main Rural Highways Related to Speed, Driver, and Vehicle. U.S. Department of Commerce and Bureau of Public Roads. July 1964

17. Ivette Cruzado1 and Eric T. Donnell, Factors Affecting Driver Speed Choice along Two-Lane Rural Highway Transition Zones. JOURNAL OF TRANSPORTATION ENGINEERING. Vol. 136, No. 8, 2010.

18. Raghavan Srinivasan, Martin Parker, David Harkey, Dwayne Tharpe, Roy Sumner. *Expert System for Recommending Speed Limits in Speed Zones* National Cooperative

Highway Research Program , Transportation Research Board, National Research Council, November 2006.

19. Wavetronix LLC. *Click! Series User Guide*.2008

20. Sharma, A., M. Harding, B. Giles, D. Bullock, J. Sturdevant, and S. Peeta. Performance Requirements and Evaluation Procedures for Advance Wide Area Detectors. Transportation Research Board for publication and presentation at the 87th Annual Meeting, Washington, D.C., 2008

21. Andrew Ng. CS 229 Machine Learning. www.cs229.stanford.edu/

22. California Department of Transportation. Realistic Speed Zoning, Why & How.

23. Simon P. Washington, Matthew G.Karlaftis, Fred L. *Mannering. Statistical and Econometric Methods for Transportation Data Analysis*. CHAPMAN & HALL / CRC. 2003

24. David M. Erceg-Hurn and Vikki M. Mirosevich. *Modern Robust Statistical Methods-An Easy Way to Maximize the Accuracy and Power of Your Research*. American Psychologist, October 2008.

25. Tim Hesterberg, David S. Moore, Shaun Monaghan, Ashley Clipson, and Rachel Epstein. *Bootstrap Methods and Permutation Tests*

26. Nancy Barker, *A Practical Introduction to the Bootstrap Using the SAS System*. Oxford Pharmaceutical Sciences, Wallingford, UK

27. http://en.wikipedia.org/wiki/Quantile. Accessed on 2/9/2013.

28. Roger Kornker, Gilbert Bassett Jr. *Regression Quantiles*. Econometrica, Vol.46, No. 1, 1978

29. Colin Chen, *An Introduction to Quantile Regression and the QUANTREG Procedure*, SAS Institute Inc. Cary, NC

30. Lingxin Hao, Daniel Q. Naiman. *Quantitative Applications in the Social Sciences*. SAGE Publications.

31. Anuj Sharma, Zifeng Wu, Shefang Wang, and Laurence R. Rilett. Speed Limit Recommendation in Vicinity of Signalized, High-Speed Intersection. Nebraska Transportation center SPR-P1 (11) M307, 2012.

APPENDICES

Appendix A : Acronyms

AADT	Annual Average Daily Traffic
AWF	Advance Warning Flashers
CDF	Cumulative Distribution Function
DMS	Dynamic Message Signs
DOT	Department of Transportation
FARS	Fatality Analysis Reporting Systems
ITS	Intelligent Transportation System
MATLAB	Matrix Laboratory
MUTCD	Manual on Uniform Traffic Control Devices
NDOR	Nebraska Department of Roads
NMSL	National Maximum Speed Limit
NTC	Nebraska Transportation Center
VSL	Variable Speed Limit
WASS	Wide Area Augmentation System

Appendix B Field Layout of Seven Intersections

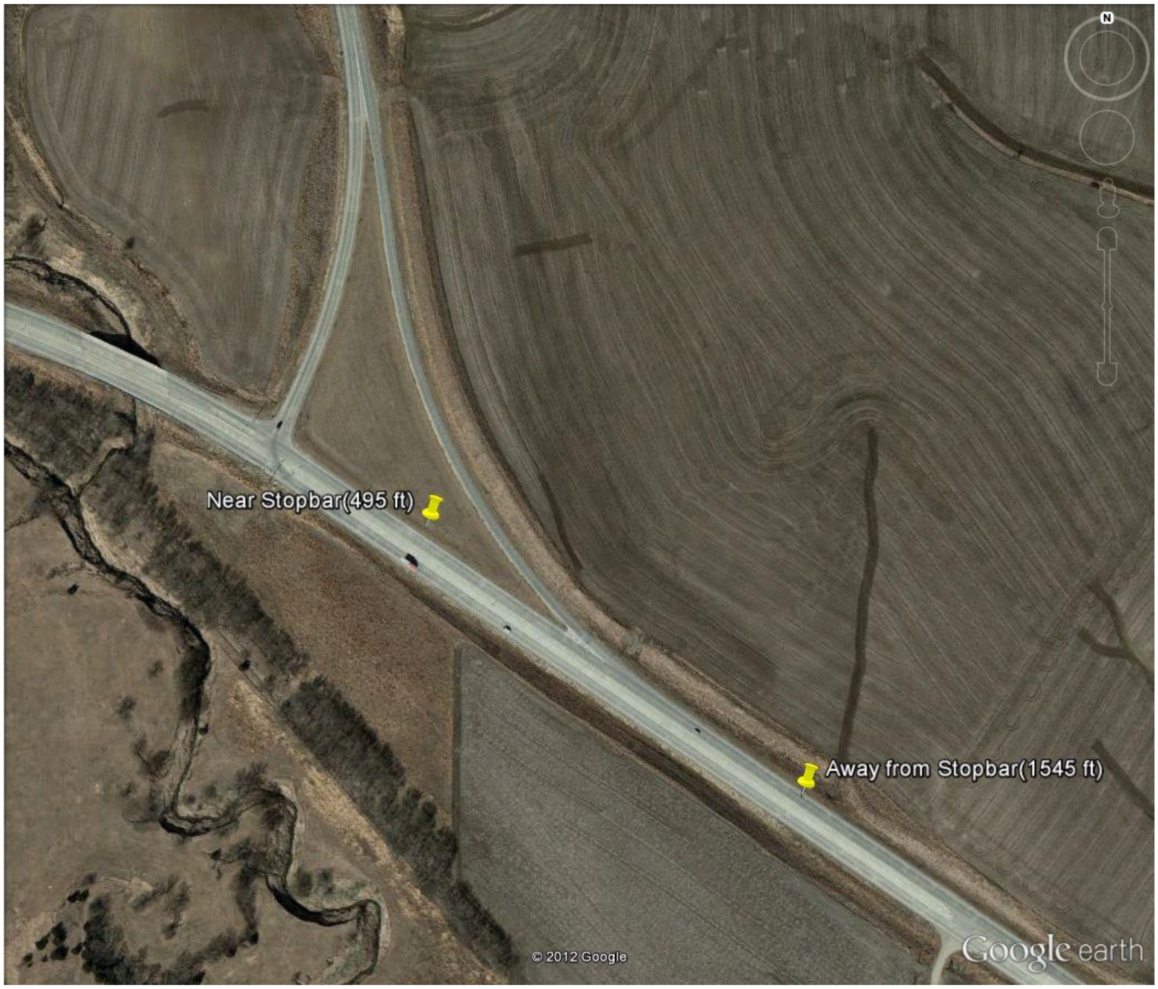


Figure B-1 US-34 & N-79

Near Stop Bar Speed Limit: 60 mph

Away from Stop Bar Speed Limit: 60 mph

Approach: Westbound

Number of Lanes: 2 Through Lanes + 1 Shared Right Turn Lane

Figure B-2 US77 & Pioneers Blvd

Near Stop Bar Speed Limit: 55 mph

Away from Stop Bear Speed Limit: 55 mph

Approach: Southbound

Number of Lanes: 2 Through Lanes + 1 Shared Right Turn Lane + 1 Left Turn Lane

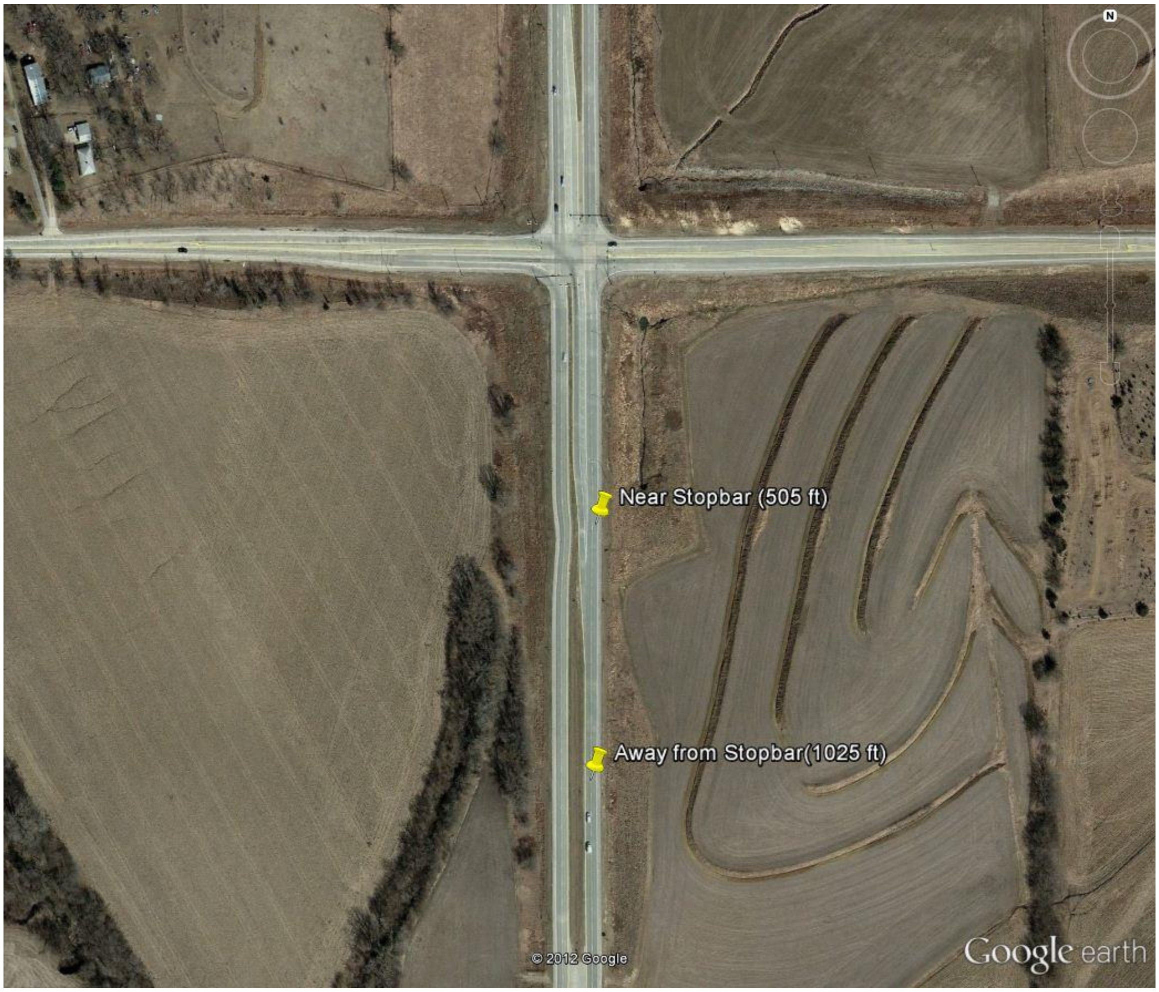


Figure B-3 N-133 & N-36

Near Stop Bar Speed Limit: 55 mph

Away from Stop Bar Speed Limit: 55 mph

Approach: Southbound

Number of Lanes: 2 Through Lanes + 1 Shared Right Turn Lane + 1 Left Turn Lane

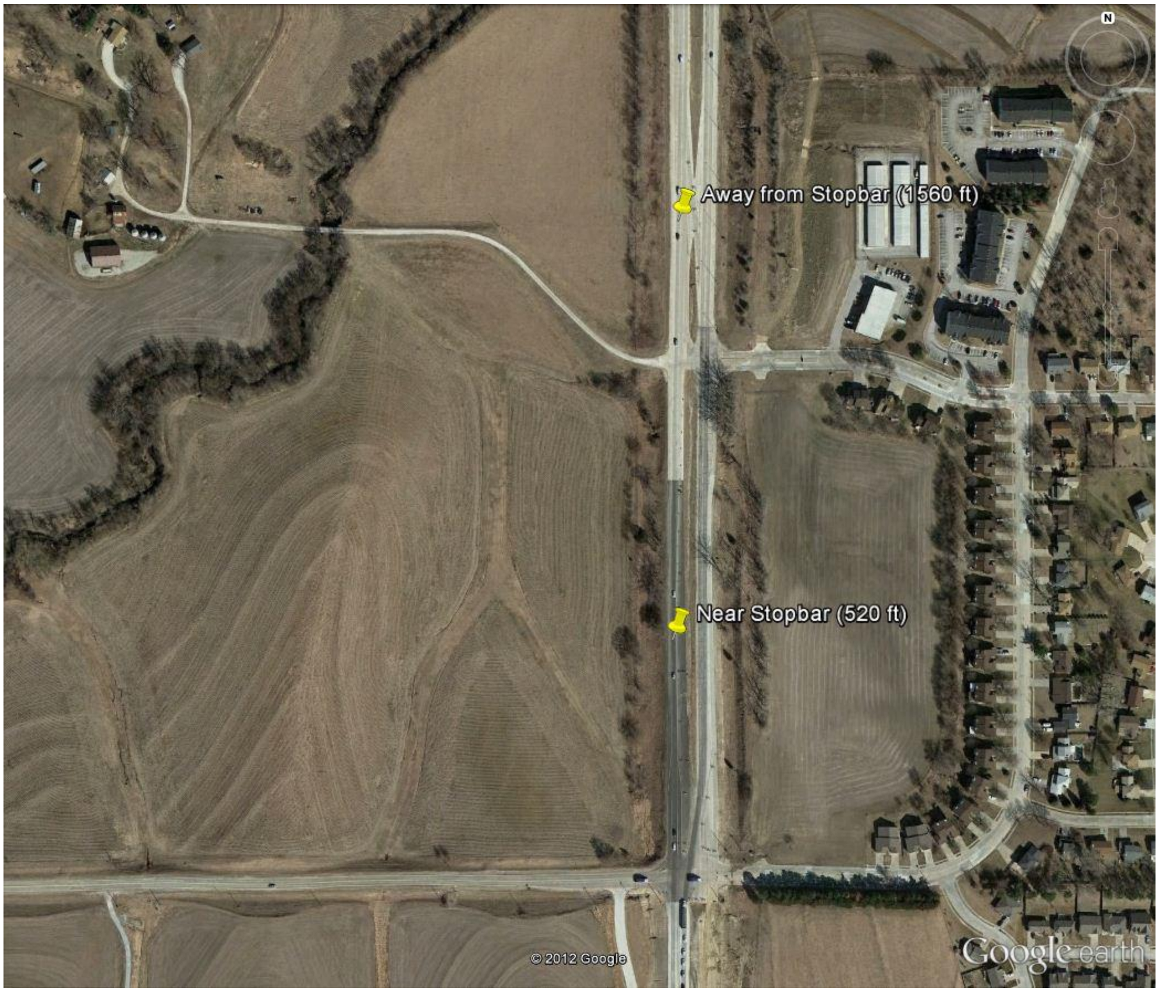


Figure B-4 US75 & Platteview Road

Near Stop Bar Speed Limit: 55 mph

Away from Stop Bar Speed Limit: 60 mph

Approach: Southbound

Number of Lanes: 2 Through Lanes + 1 Right Turn Lane + 1 Left Turn Lane

Figure B-5 US-81& Lincoln Ave

Near Stop Bar Speed Limit: 55 mph

Away from Stop Bar Speed Limit: 60 mph

Approach: Southbound

Number of Lanes: 2 Through Lanes + 1 Left Turn Lane

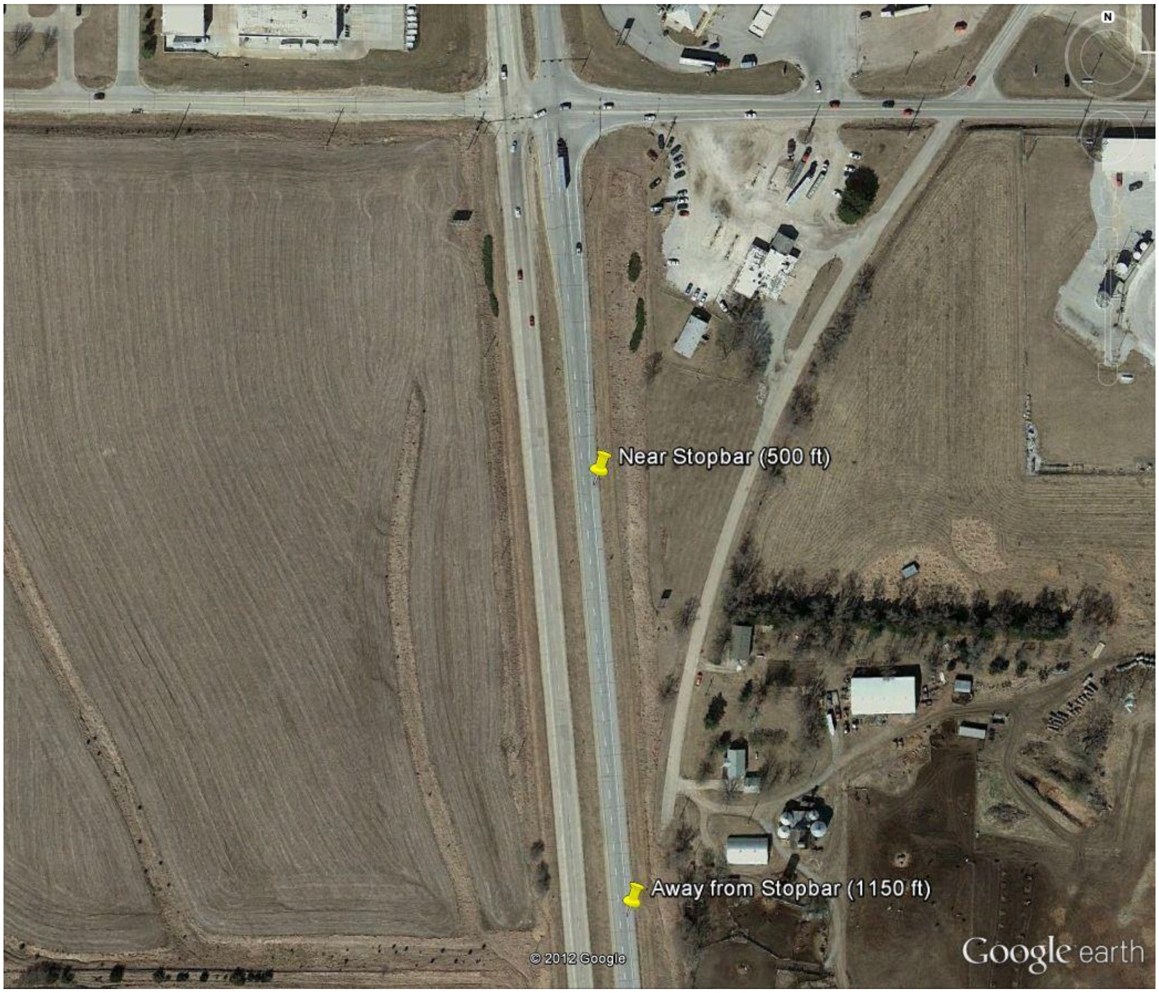


Figure B-6 US-77 & Saltillo Road

Near Stop Bar Speed Limit: 55 mph

Away from Stop Bar Speed Limit: 65 mph

Approach: Northbound

Number of Lanes: 2 Through Lanes + 1 Right Turn Lane + 1 Left Turn Lane

Figure B-7 US281& Platte River

Near Stop Bar Speed Limit: 55 mph

Away from Stop Bar Speed Limit: 65 mph

Approach: Southbound

Number of Lanes: 2 Through Lanes + 1 Right Turn Lane + 1 Left Turn Lane

Appendix C Machine Learning— Classifiers

Classifier's Performance Based on Training Dataset

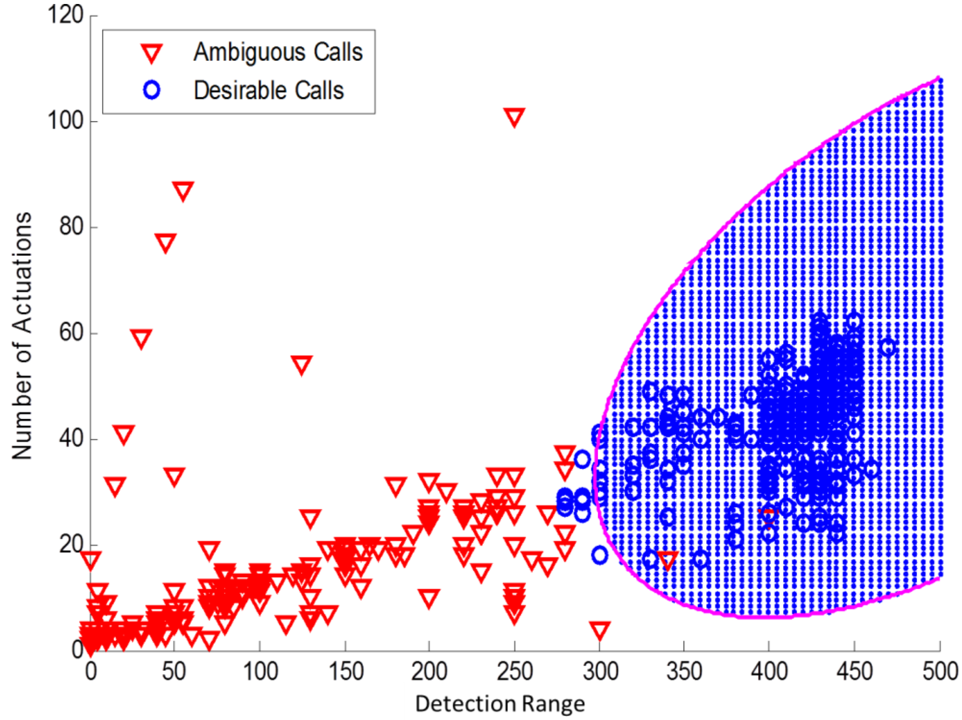


Figure C-1 Quadratic Classifier based on Difference of Range & Number of Actuations

Table C-1 Performance Evaluation Table for Quadratic Classifier based on Difference of Range & Number of Actuations

Total Training Examples: 549		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	305	2
	Prediction Outcome: Ambiguous Calls:	9	233
Precision (P) = 0.993; Recall(R) = 0.971			
F Score = 2PR/(P+R) = 0.982			

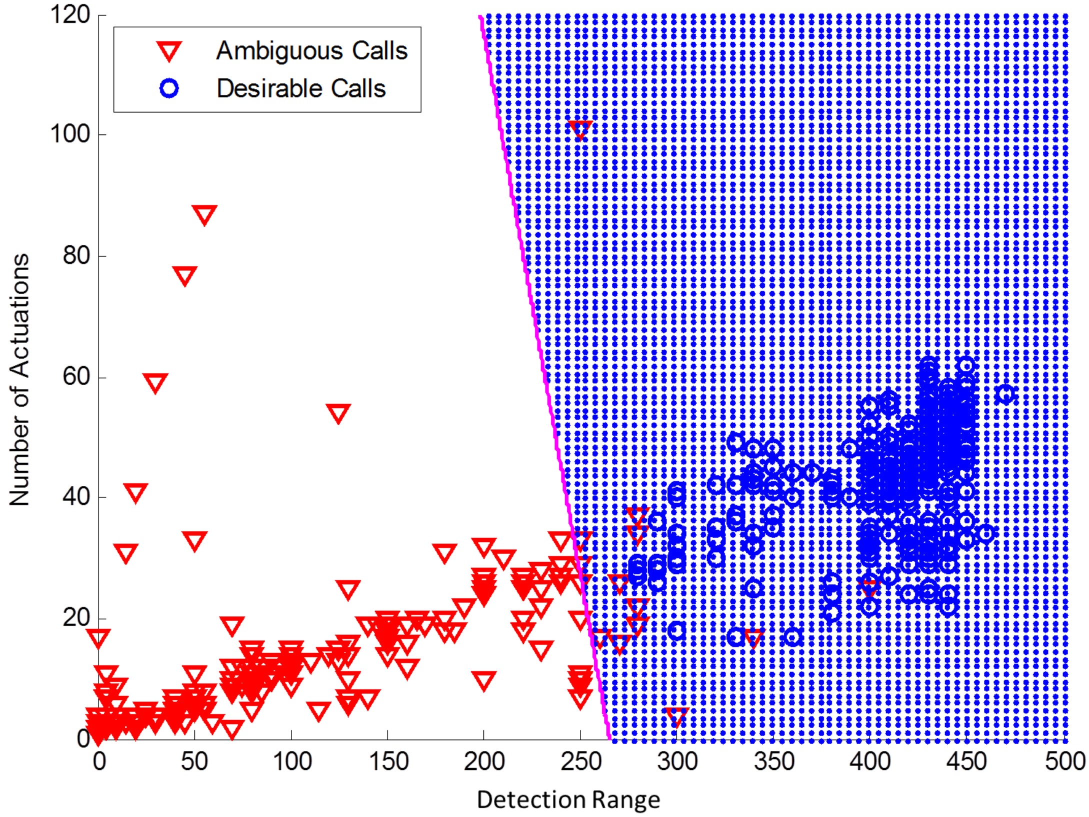


Figure C-2 Linear Classifier based on Difference of Range & Number of Actuations

Table C-2 Performance Evaluation Table for Quadratic Classifier based on Difference of Range & Number of Actuations

<table>
<tr><td colspan="2" rowspan="2">Total Training Examples: 549</td><td colspan="2">Manually Classified</td></tr>
<tr><td>Desirable Calls</td><td>Ambiguous Calls</td></tr>
<tr><td rowspan="2">Classifier</td><td>Prediction Outcome: Desirable Calls:</td><td>314</td><td>13</td></tr>
<tr><td>Prediction Outcome: Ambiguous Calls:</td><td>0</td><td>222</td></tr>
<tr><td colspan="4">Precision (P) = 0.960; Recall(R) = 1.000</td></tr>
<tr><td colspan="4">F Score = 2PR/(P+R) = 0.980</td></tr>
</table>

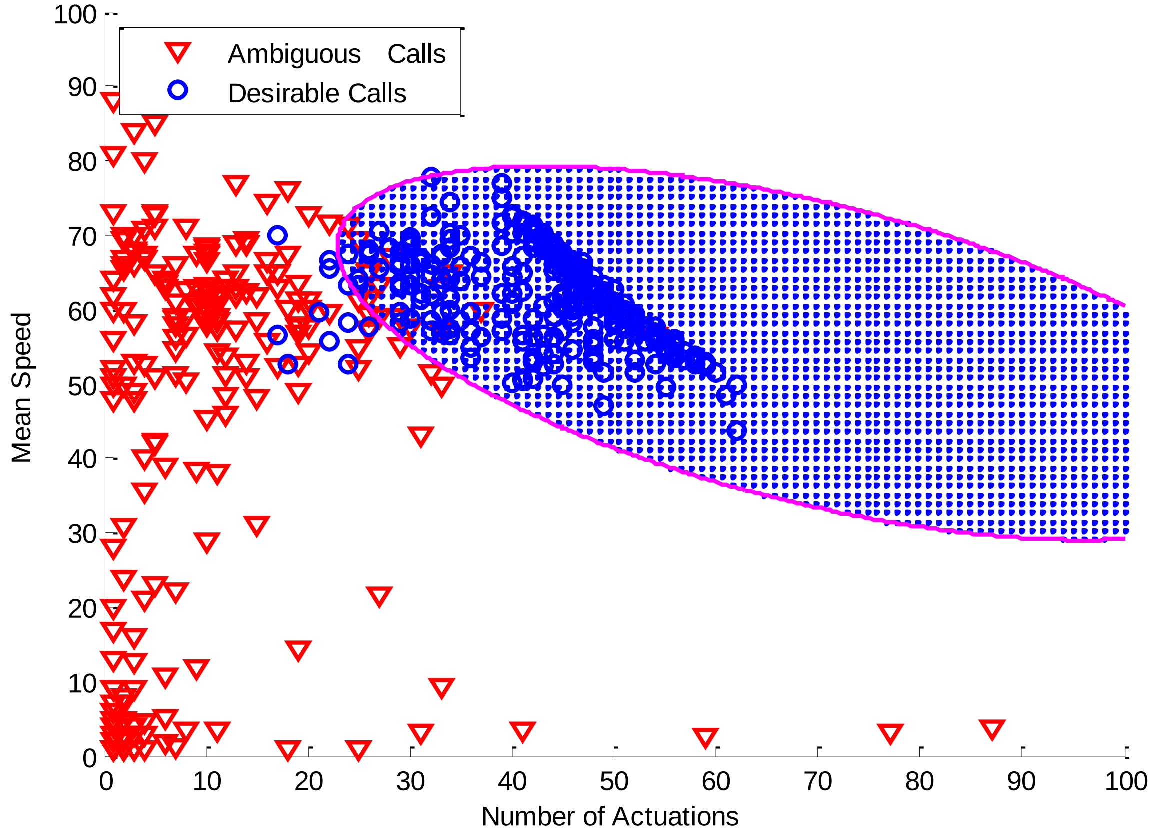


Figure C-3 Quadratic Classifier based on Number of Actuations & Mean Speed

Table C-3 Performance Evaluation Table for Quadratic Classifier based on Number of Actuations & Mean Speed

<table>
<tr><td colspan="2" rowspan="2">Total Training Examples: 549</td><td colspan="2">Manually Classified</td></tr>
<tr><td>Desirable Calls</td><td>Ambiguous Calls</td></tr>
<tr><td rowspan="2">Classifier</td><td>Prediction Outcome: Desirable Calls:</td><td>303</td><td>16</td></tr>
<tr><td>Prediction Outcome: Ambiguous Calls:</td><td>11</td><td>219</td></tr>
<tr><td colspan="4">Precision (P) = 0.950; Recall(R) = 0.965</td></tr>
<tr><td colspan="4">F Score = 2PR/(P+R) = 0.957</td></tr>
</table>

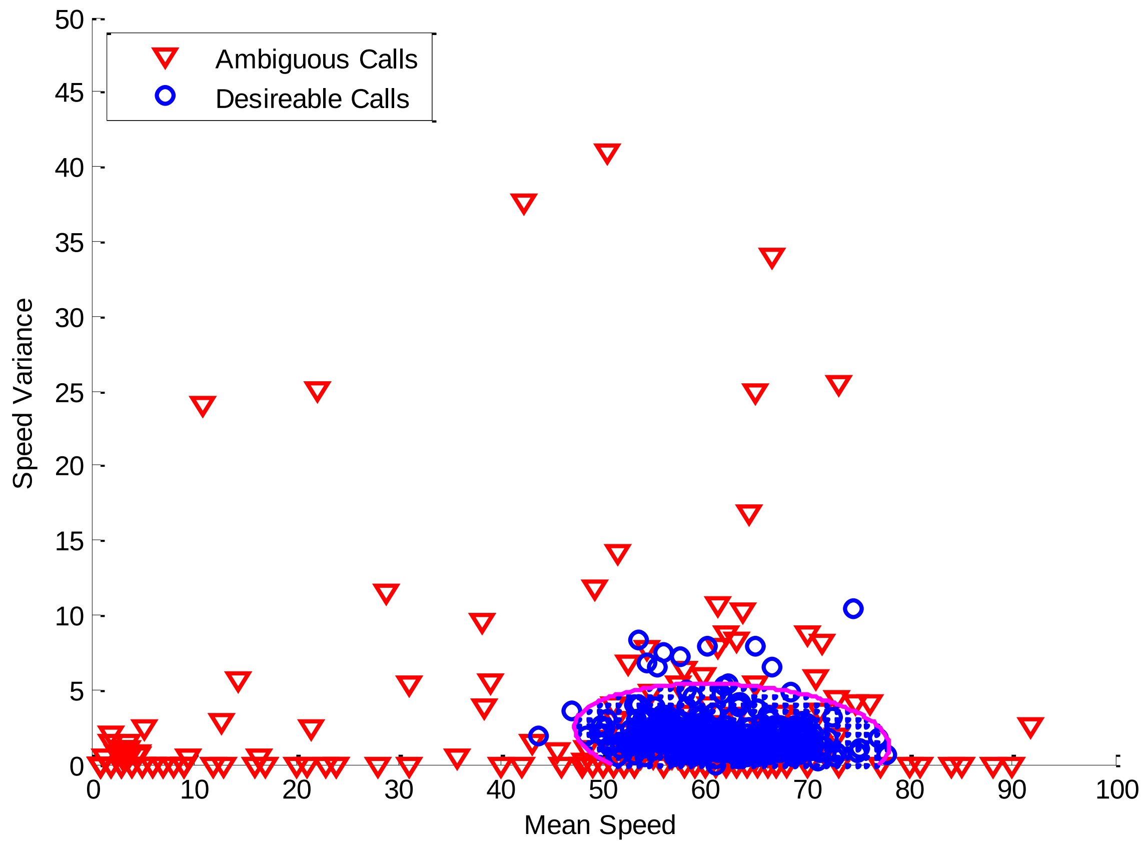


Figure C-4 Quadratic Classifier based on Mean Speed & Speed Variance

Table C-4 Performance Evaluation Table for Quadratic Classifier based on Mean Speed & Speed Variance

Total Training Examples: 549		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	301	113
	Prediction Outcome: Ambiguous Calls:	13	122
Precision (P) = 0.727; Recall(R) = 0.959			
F Score = 2PR/(P+R) = 0.827			

Table C-5 Performance Evaluation Table for Linear Classifier based on all 4 Features

(Difference of Range, Number of Actuations, Mean Speed, and Speed Variance)

Total Training Examples: 549		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	314	12
	Prediction Outcome: Ambiguous Calls:	0	223
Precision (P) = 0.963; Recall(R) = 1.000			
F Score = 2PR/(P+R) = 0.981			

Table C-6 Performance Evaluation Table for Quadratic Classifier based on all 4 Features

(Difference of Range, Number of Actuations, Mean Speed, and Speed Variance)

Total Training Examples: 549		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	310	1
	Prediction Outcome: Ambiguous Calls:	4	234
Precision (P) = 0.997; Recall(R) = 0.987			
F Score = 2PR/(P+R) = 0.992			

Classifier's Performance Based on Validation Dataset

Table C-7 Performance Evaluation Table for Quadratic Classifier based on Difference of Range & Number of Actuations

Total Validation Examples: 456		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	211	3
	Prediction Outcome: Ambiguous Calls:	4	238
Precision (P) = 0.986; Recall(R) = 0.981			
F Score = 2PR/(P+R) = 0.983			

Table C-8 Performance Evaluation Table for Quadratic Classifier based on all 4 Features (Difference of Range, Number of Actuations, Mean Speed, and Speed Variance)

Total Validation Examples: 456		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	209	2
	Prediction Outcome: Ambiguous Calls:	6	239
Precision (P) = 0.991; Recall(R) = 0.972			
F Score = 2PR/(P+R) = 0.981			

Table C-9 Performance Evaluation Table for Linear Classifier based on all 4 Features (Difference of Range, Number of Actuations, Mean Speed, and Speed Variance)

Total Validation Examples: 456		Manually Classified	
		Desirable Calls	Ambiguous Calls
Classifier	Prediction Outcome: Desirable Calls:	209	2
	Prediction Outcome: Ambiguous Calls:	6	239
Precision (P) = 0.947; Recall(R) = 0.991			
F Score = 2PR/(P+R) = 0.968			

Appendix D Speed Distribution

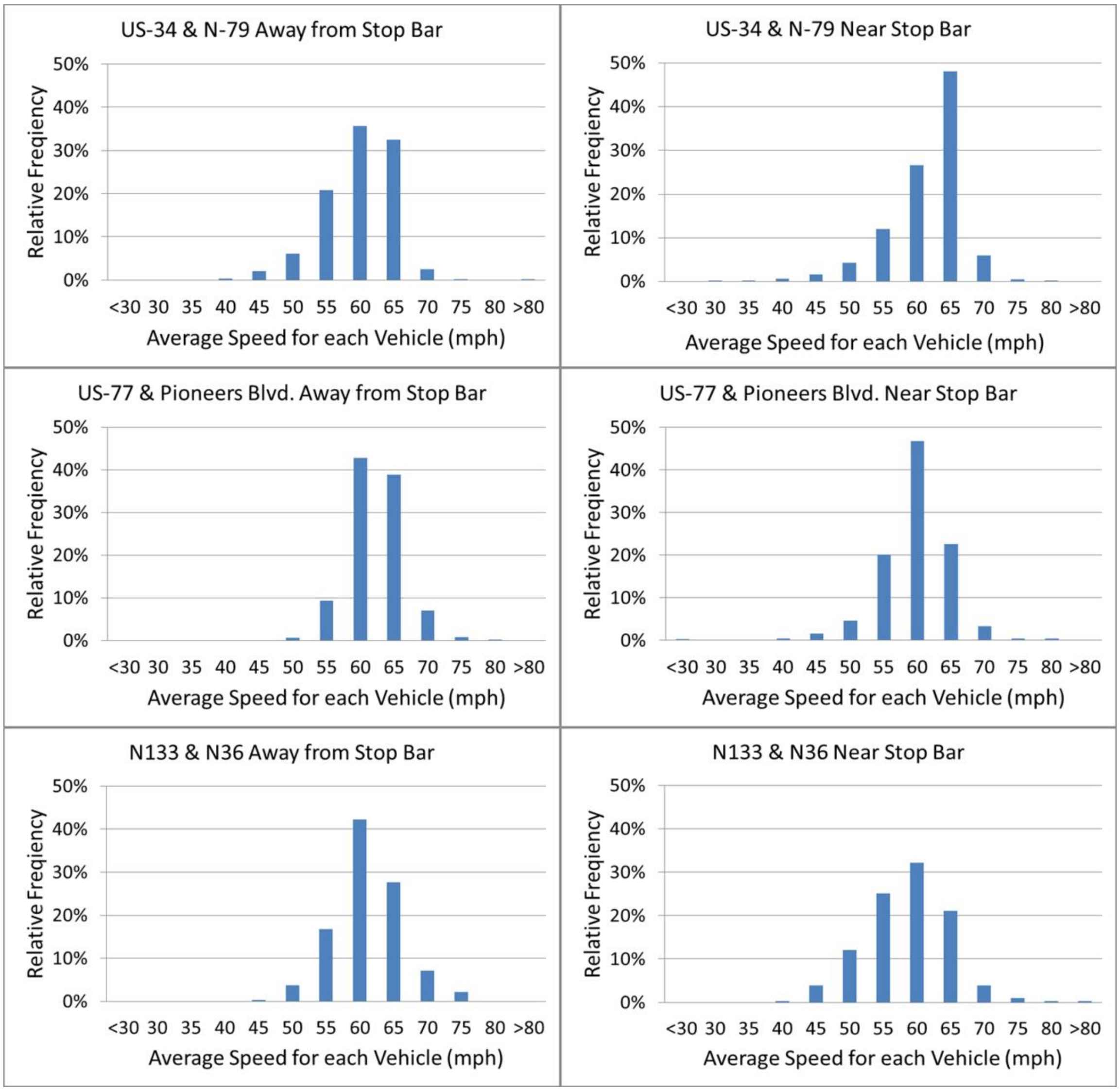


Figure D-1 Speed Distribution for 0 mph Speed Reduction Group

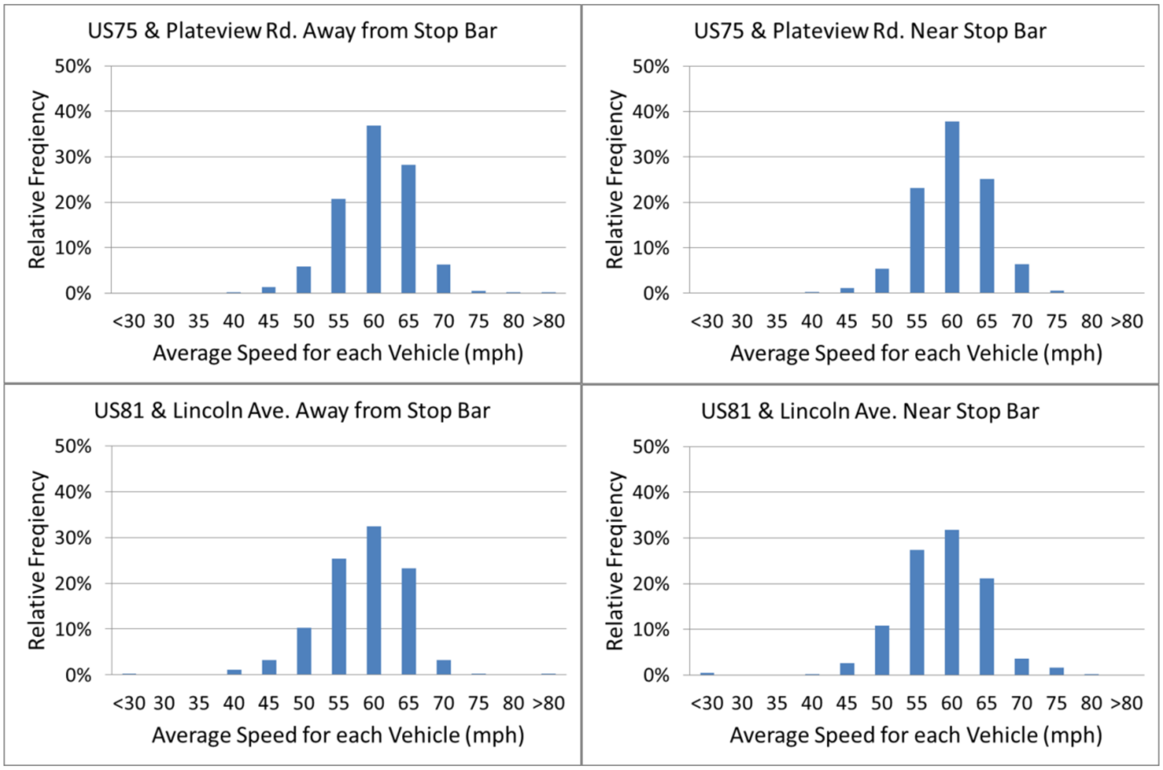


Figure D-2 Speed Distribution for 5 mph Speed Reduction Group

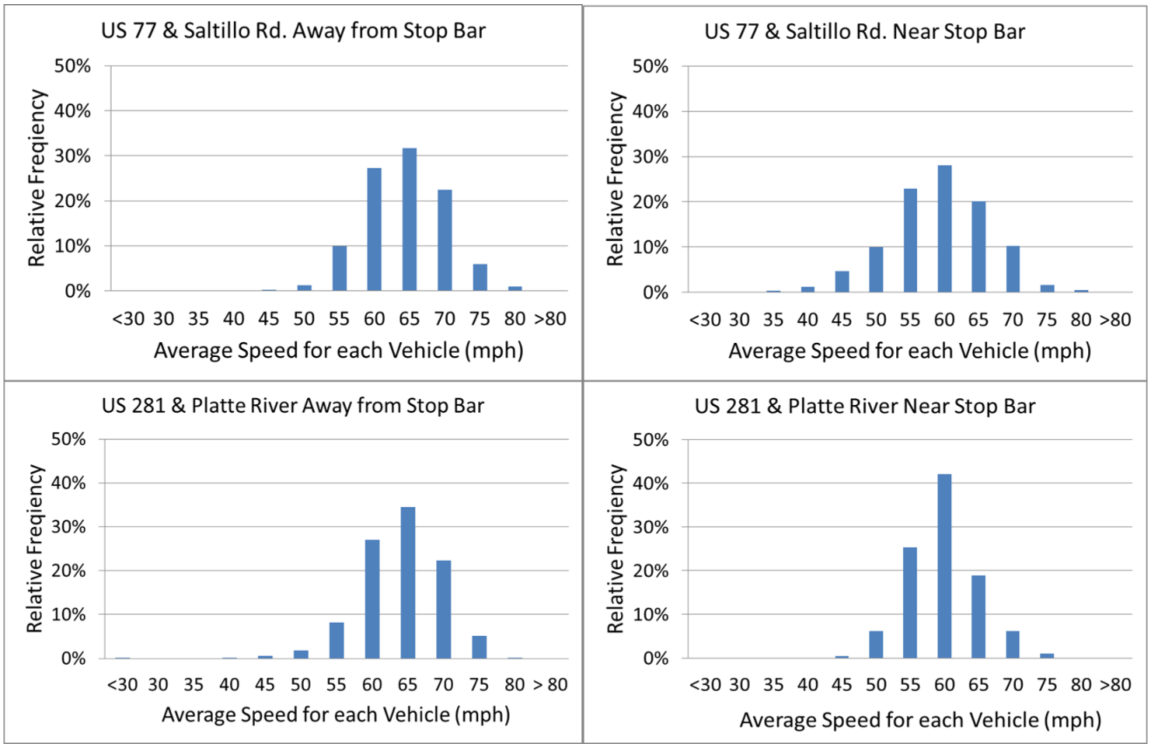


Figure D-3 Speed Distribution for 10 mph Speed Reduction Group

The relative frequency histograms Figure D-1, Figure D-2, and Figure D-3 are plotted, where the x-axis presents average speed for each vehicle, and the y-axis is the relative frequency.

To measure the asymmetry of a distribution, skewness of sample can be used, calculated by the equation:

$$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2)^{3/2}}, \text{ where}$$

- m_3 is the sample third central moment.
- m_2 is the sample variance.
- N is the number of observations.
- x_i is the value for i^{th} observation, and
- $\bar{x}$ is the mean of the values for all observations.

A negative skewness has a longer left tail and has a relatively few low values, while a positive skewness has a longer right tail and has a relatively few high values. Table D-1 lists the skewness for each intersection.

Table D-1 Skewness for Each Intersection

#	Location	Speed Limit Drop Group	Away-from-Stop bar	Near-Stop bar
1	US-34 & N-79	0 MPH Speed Limit Drop Group	-0.3742	-1.3659
2	US-77 & Pioneers		0.2332	-1.2972
3	N-133 & N-36		0.2326	0.4298
4	US-75 & Platteview	5 MPH Speed Limit Drop Group	-0.1242	-0.5586
5	US-81 & S Lincoln		-1.4783	-1.7145
6	US-77 & Saltillo	10 MPH Speed Limit Drop Group	-1.5144	-0.1587
7	US-281 & Platte River		-0.6572	0.2232

Besides the bootstrap method applied in the chapter 5, the t test to compare two population mean is also applied here. Based on chapter 8.3 in Statistics for Research 3rd Edition (Shirley Dowdy, Stanley Weardon, and Daniel Chilko), it provide a methodology about the inference about two means. When sample size of two populations are greater than 30 (n_1 and $n_2 \geq 30$), the confidence interval on $\mu_1 - \mu_2$ can be calculated by:

$$CI_{1-\alpha}: \bar{y}_1 - \bar{y}_2 \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Use S_1^2 and S_2^2 to estimate σ_1^2 and σ_2^2 if the population variance are unknown. The test statistic can be calculated by:

$$z = \frac{\bar{y}_1 - \bar{y}_2 - (\mu_1 - \mu_2)_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Table D-2 is the detailed statistics of the mean difference test, the Away column is the mean speed at the away stop bar location, and Near column represents mean speed at the near stop bar location. Lower and higher column is the lower boundary and higher boundary of the speed difference under the 95% confidence interval. If the interval covers the value of 0, it means the difference is not significant, and vice versa. The last column z is the test statistic of the mean speed difference. The result is almost the same drawn from the bootstrap method, except the intersection # 4 US-75 & Platte View.

Table D-2 Statistic Test Table

#	Site Name	Away Mean	Near Mean	Lower	Higher	z
1	US-34 & N-79	57.4	59.2	-2.3	-1.2	-6.3319
2	US-77 & Pioneers	59.7	57.2	2.2	2.9	14.6459
3	N-133 & N-36	58.7	56.1	1.9	3.3	7.3275
4	US-75 & Platte View	57.9	57.5	0.0	0.6	1.9728
5	US-81 & S Lincoln	55.9	56.2	-1.2	0.8	-0.423
6	US-77 & Saltillo	61.7	56.7	4.3	5.7	13.5491
7	US-281 & Platte River	61.6	57.3	3.7	4.9	14.1887