Assessing the Impact of Active Signage Systems on Driving Behavior and Traffic Safety

Matin Giahi Foomani

A Thesis

In the Department

of

Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy (Civil Engineering) at

Concordia University

Montreal, Quebec, Canada

August 2022

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: *Matin Giahi Foomani*

Entitled: *Assessing the Impact of Active Signage Systems on Driving Behavior and Traffic Safety*

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (*Civil Engineering*)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Rolf Wuthrich	Chair
Dr. Ciprian Alecsandru	Thesis Supervisor
Dr. Anjali Awasthi	Thesis Supervisor
Dr. Subhash Rakheja	Examiner
Dr. Mohamed Ouf	Examiner
Dr. Luis Amador	Examiner
Dr. Sybil Derrible	External Examiner

Approved by	Dr. Michelle Nokken	Graduate Program Director
August 2022	Dr. Mourad Debbabi	Dean

Abstract

Assessing the Impact of Active Signage Systems on Driving Behavior

and Traffic Safety

Matin Giahi Foomani

Concordia University, 2022

Unsignalized Stop-Controlled Intersections (SCI) are widely used in North America, and account for one out of every ten collisions. Understanding how drivers and pedestrians behave at unsignalized intersections is critical for public safety. Drivers who do not obey the stop-sign's indication by not coming to a complete stop or miss or fail to stop at SCI create a substantial safety risk. For decades, visibility and placement of road alignments and signage at intersections have been a concern among transportation safety specialists. Deployment of backlit Light-Emitting Diode (LED) or other illuminated signs (also known as active road signs) has been increased especially at hot-spots and locations with known safety problems, or potential collision risks. While these signs are expected to improve safety measures by regulating safe travelers' passage, their performance is not yet fully understood. Although environmental factors such as intersection type, location, and road design are playing a major role, compositional variables such as driver behaviour, which can be explained in terms of carelessness, lack of attention, or overconfidence, is resulting in a failure to comply with the law of making a complete stop at SCI.

Previous empirical research demonstrated some correlation between several variables such as traveller compliance with road signs and alignments, direct and indirect road safety measures, collision/conflict frequency, and road/traffic characteristics. These studies commonly employ before-after or cross-reference analyses to determine the long-term effects of various countermeasures at SCI. A few studies also utilized calibrated micro-simulations models to evaluate the surrogate safety measures at SCI.

This thesis defines a methodology to evaluate the safety performance of a new and untested signage without putting traffic at long risk. To evaluate the performance of the signs, the suggested methodology investigates multiple parameters and identifies influencing variables in a conflict-based collision-prediction model at SCI. The proposed methodology is applied to a real-world network in the city of Montreal, with several three-leg SCI equipped with different countermeasures. The experiment was designed in a fashion which isolates the influence of several variables, allowing the focus to be on the impact of the target variable (signage type). Field experiments have been performed to study the driver's behavior in terms of approaching speed as well as quantitative analysis on reactions to various signs, using different sample groups from the same population. This research sets up a microsimulation model that captures drivers' behaviour with respect to signage according to the observed data. A genetic algorithm was deployed to calibrate the microsimulation model in terms of turning movement counts and the critical conflicts were calculated at each intersection using vehicle trajectories. Collision-prediction regression models was then developed for the intersections under investigation, using traffic volume and conflict.

The results demonstrated a high correlation among countermeasures and drivers' speed and compliance. The relationship between critical conflicts computed in microsimulation models and actual collisions was found to be statistically significant. The model which includes drivers' compliance in collision-prediction regression was also found to fit the collision data better. However, the results of this study do not support the previous assumption that the conflict-based collision-prediction models fit the collision data better than the volume-based collision-prediction models at SCI, especially with drivers' compliance supplementary data. Finally, while the backlit signs' performance was marginally better than that of a normal LED active sign, the difference was not statistically significant.

The methodology suggested in this thesis has the potential to be implemented in safety performance evaluation of a countermeasure without placing traffic at danger for an extended period. For instance, when there is apprehension about an adverse effect. Future research could investigate leveraging drivers' behaviour to countermeasures, to improve the performance of collision-prediction regression models like the one proposed in this thesis. Finally, the results from the performance assessment of the LED active signs can assist transportation specialists in deciding whether or not to deploy these countermeasures.

Résumé Analytique

ÉVALUATION DE L'IMPACT DES SYSTÈMES DE SIGNALISATION ACTIVE

SUR LE COMPORTEMENT AU VOLANT ET LA SÉCURITÉ ROUTIÈRE

Matin Giahi Foomani

Université Concordia, 2022

Les intersections contrôlées par au moins un panneau d'arrêt sans signalisation sont largement utilisées en Amérique du Nord et sont responsables d'une collision sur dix. Pour la sécurité du public, il est primordial de comprendre le comportement des conducteurs et des piétons aux intersections non signalisées. Les conducteurs qui enfreignent les panneaux d'arrêt en ne s'arrêtant pas complètement, manquent l'arrêt ou omettent de s'arrêter aux intersections contrôlées par un panneau d'arrêt créent un risque de sécurité important. La visibilité et l'emplacement des tracés routiers et de la signalisation aux intersections, sont une préoccupation des spécialistes en sécurité de transport depuis des décennies. L'utilisation des diodes électroluminescente (DEL) rétroéclairées ou d'autres panneaux routiers illuminés (aussi nommés active road signs en anglais) a augmenté, surtout dans les zones à risque et intersections avec des problèmes de sécurité connus ou à risques de collision potentiels. Bien que ces panneaux routiers soient censés améliorer les mesures de sécurité en réglementant le passage sécuritaire des usagers de la route, leur performance n'est pas encore entièrement comprise. Alors que des facteurs environnementaux comme les types d'intersections, l'endroit et la conception des routes jouent un grand rôle, d'autres variables soit les comportements des conducteurs, tels que l'imprudence, le manque d'attention ou l'excès de confiance, entraînent un non-respect de la loi sur l'arrêt complet aux intersections contrôlées par un panneau d'arrêt.

La recherche empirique précédente a eu tendance à découvrir des liens entre plusieurs variables dont la conformité des conducteurs aux panneaux routiers et tracés routiers, les mesures de sécurité directe et indirecte de la route, la fréquence des conflits et collisions ainsi que les caractéristiques du trafic et de la route. Ces études utilisent fréquemment des analyses avant-après ou des analyses à références croisées pour déterminer les effets à long terme de plusieurs contre-mesures aux intersections contrôlées par un panneau d'arrêt. Quelques études ont aussi utilisé des modèles de micro-simulations calibrés pour évaluer les mesures de sécurité de remplacement aux intersections contrôlées par un panneau d'arrêt.

Cette thèse propose une méthodologie pour évaluer la performance de sécurité de nouvelle signalisation routière non testée sans faire courir de risques au trafic à long termes. Pour évaluer le rendement des panneaux routiers, la méthodologie suggérée examine plusieurs paramètres et identifie des variables influentes dans un modèle de prédiction de collisions causées par des conflits aux intersections contrôlées par un panneau d'arrêt. La méthodologie proposée est appliquée à un réseau réel dans la ville de Montréal, avec plusieurs intersections contrôlées par un panneau d'arrêt à trois voies comprenant différentes contre-mesures. Cette expérience a été conçue de manière à isolé l'influence de plusieurs variables, permettant de se concentrer sur l'impact de la variable ciblée (type de signalisation routière). Des expériences sur le terrain ont été réalisées pour étudier le comportement des conducteurs en termes de vitesse d'approche ainsi qu'une analyse quantitative de sa réaction face aux diverses signalisations routières, en utilisant différents sous-groupes d'une même population. Cette étude met en place un modèle de micro-simulation qui imite le comportement des conducteurs en matière de signalisation routière selon les données observées. Un algorithme génétique a été utilisé pour calibrer le modèle de micro-simulation en comptant les changements de direction. Quant à eux, les conflits critiques ont été calculés à chaque intersections en utilisant la trajectoire des véhicules. Un modèle de régression pour la prédiction des collisions a ensuite été développé pour les intersections sous enquête en utilisant le volume et les conflits du trafic.

Les résultats démontrent une grande corrélation parmi les contre-mesures, la vitesse et la conformité des conducteurs. La relation entre les conflits critiques compilés dans les modèles de micro-simulation et des collisions réelles était statistiquement significative. Le modèle qui incluait la conformité des conducteurs en régression pour la prédiction des collisions s'est également avéré mieux adapté aux données de collisions. Toutefois, cette étude ne supporte pas l'ancienne hypothèse que les modèles de prédiction des collisions basés sur les conflits, s'adaptent mieux aux données des collisions que les modèles de prédiction des collisions basés sur le volume des intersections contrôlées par un panneau d'arrêt, surtout avec les données supplémentaires relatives à la conformité des conducteurs. Enfin, alors que la performance des panneaux routiers rétroéclairées était légèrement mieux que celle des panneaux actifs normaux à DEL, la différence n'était pas statistiquement significative.

La méthodologie suggérée dans cette thèse a le potentiel d'être utilisée dans l'évaluation de la performance de sécurité d'une contre-mesure sans mettre le trafic en danger pendant une période prolongée, par exemple, lorsqu'il y a une crainte d'une conséquence grave. Des recherches futures pourraient porter sur une meilleure compréhension des comportements des conducteurs vis-à-vis les contre-mesures, afin d'améliorer les performances des modèles de régression pour la prédiction des collisions, comme celui proposé dans cette thèse. En conclusion, les résultats de l'évaluation de la performance des panneaux actifs à DEL peuvent aider les spécialistes en transport, à déterminer l'utilisation appropriée d'une contre-mesure.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors Dr. Ciprian Alecsandru and Dr. Anjali Awasthi. Though only my name is printed on the cover, this was not possible without their thoughtful comments, encouragements, and recommendations. Dr. Alecsandru not only extended compassionate support, consistent optimism, and academic guidance, he was also my caring coach and a mentor during the ups and downs of this long journey. Working under the supervision of Dr. Awasthi, a hardworking, brilliant woman in engineering, is an absolute honor.

Besides my advisors, I would like to thank the rest of my thesis committee from Concordia university: Prof. Rakheja, Prof. Amador, and Prof. Ouf, for their insightful comments, encouragement, and challenging questions which encouraged me to look at the research from the right perspective. I am also grateful to have Prof. Derrible, from University of Illinois as part of the final examination.

I am proud to acknowledge that the support for this study was provided through MITACS Contract IT04576 and in-kind contributions from Orange Traffic in building two prototypes, providing radar equipment and help with installation of the treatments. I would like to thank the Borough of Lachine and the local police for permitting the research team to set up the test bed in their jurisdiction and grant access to the traffic records. Mr. Paul Bourque this study was impossible without your support.

Prof. Davis and Prof. Gates from TRB ACP55 Committee have been an important influence on me throughout my studies. I am forever thankful for their inspirational literatures.

I am grateful to all my dear colleagues at transportation lab; Shohel, Huizhu, Raphaël, Amir, Chao li, Ehsan Nateghinia, Hamed, and of course my best friend Behzad (Doktor Jan!). To my dear niece, Summer, who passionately made the flawless translations in French, and Dr. William (Bill) Fletcher for his insightful editorial comments and efforts - I thank you.

Most importantly, none of this would have been possible without the love and patience of my family and my talented wife Amélie, in particular. She has been a constant source of love, concern, support, and strength. She put the needs of me ahead of her own, especially the last intense academic year to make the time and space available for this research. I would also like to express my heartfelt gratitude to my mom and dad and my extended family who have aided and encouraged me throughout this endeavor.

This thesis is dedicated to my baba Soren,
who has been my inspiration since his tumultuous debut!

Soren; this is for YOU!

Preface and Disclaimer

In this thesis, portions of three chapters (chapters 2, 3 and 4) have been based on published/under review material.

Transportation Research Board:

Investigating Safety of Different Treatments at Stop-Controlled Intersection Using Driver Behavior Analyses. *publication at 95th Annual Meeting (Washington, DC 2016); M, Foomani, C Li, C Alecsandru*

Safety Assessment of Backlit Pedestrian Crossing Sign on Stopping Characteristics at Unsignalized Intersection. *poster, and publication at 96th Annual Meeting (Washington, DC 2017); M Foomani, C Alecsandru, A. Awasthi*

Journal of Advanced Transportation (Under review):

Using Driver Behaviour and A Micro-Simulation Approach for Predicting Collision and Conflict at Un-Signalized Intersections with LED Signs: A Case Study of Lachine area In Montreal. *(JAT 2022); M Foomani, C Alecsandru, A. Awasthi*

Conference Presentations:

- Predicting Collision at Unsignalized Intersections with Active Signage Systems using Traffic Conflicts. *M, Foomani, C Alecsandru; A. Awasthi Canadian Institute of Transportation Engineers CITE 2021*
- Investigating Safety Impact of Backlit Pedestrian Crossing Sign Using Driver Behavior Analysis. *M, Foomani, C Alecsandru; P. Lauziere Canadian Transportation Research Forum CTRF 2016*
- Safety Performance Assessment of Stop-Operated Intersection Equipped with Active Road Sign. *M, Foomani, C Alecsandru; A. Awasthi Canadian Transportation Research Forum CTRF 2015*

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Description
AASHTO	American Association of Highway and Transportation Officials
ADAS	Advanced Driver-Assistance Systems
ADT/AADT	Average Daily Traffic / Annual Average Daily Traffic
AIC	Akaike Information Criterion
ANN	Artificial Neural Networks
ARS	Active Road Sign
AV	Autonomous Vehicle
BIC	Bayes Information Criterion
CMF	Crash (Collision) Modification Factor
COM	Component Object Model
CRF	Crash Reduction Factors
DR	Deceleration Rate
FARS	Fatality Analysis Reporting System
FHWA	Federal Highway Administration
GA	Genetic Algorithm
GEH	Geoffrey E. Havers (Test)
GT	Gap Time
HD	High Definition
HPSM	Highway Performance Monitoring System
HSM	Highway Safety Manual
ICWS	Intersection Conflict Warning Systems
IHSDM	Interactive Highway Safety Design Model
ITE	Institute of Transportation Engineers
ITS	Intelligent Transportation Systems
LED	Light-Emitting Diode
LLR-T	Log-Likelihood Ratio Test
MCMIS	Motor Carrier Management Information System
MNL	Multinomial Logistic Regression

MSE	Mean Square of Errors
MUTCD	Manual on Uniform Traffic Control Devices
NB	Negative Binomial
NCDB	National Collision Data Base
NCHRP	National Cooperative Highway Research Program
NCUTCD	National Committee on Uniform Traffic Control Devices
NEISS	National Electronic Injury Surveillance System
PET	Post Encroachment Time
RTM	Regression To the Mean
SCI	Stop-Controlled Intersections
SHSP	Strategic Highway Safety Plan
SPF	Safety Performance Function
SSAM	Surrogate Safety Assessment Model
SSE	Sum of Squared Error
TC	Transport Canada
TMC	Turning Movement Counts
TRB	Transportation Research Board
TTC	Time To Collision
TTI	Texas Transportation Institute
VAS	Vehicle Activated Sign
VHP	Volume Per Hour

Table 1 List of abbreviations

1 Chapter 1: Introduction

The purpose of traffic engineering is to create a transportation system that is as efficient, comfortable and as safe as possible for all road users. Focusing on traffic safety, it is well understood that geometric design and traffic operations improvements can decrease risks, reducing collision, and their severity. Road safety is more than an engineering problem and the term “safety” needs more clarification. Beside engineering, there are 7 others ‘E’s that are necessary in a comprehensive safety approach including: Education, Enforcement, Emergency Medical Services, Environment, Economics, Evaluation, and Everyone.[1] Road collisions are mainly influenced by the interaction of human, vehicle, and environmental factors.[2]

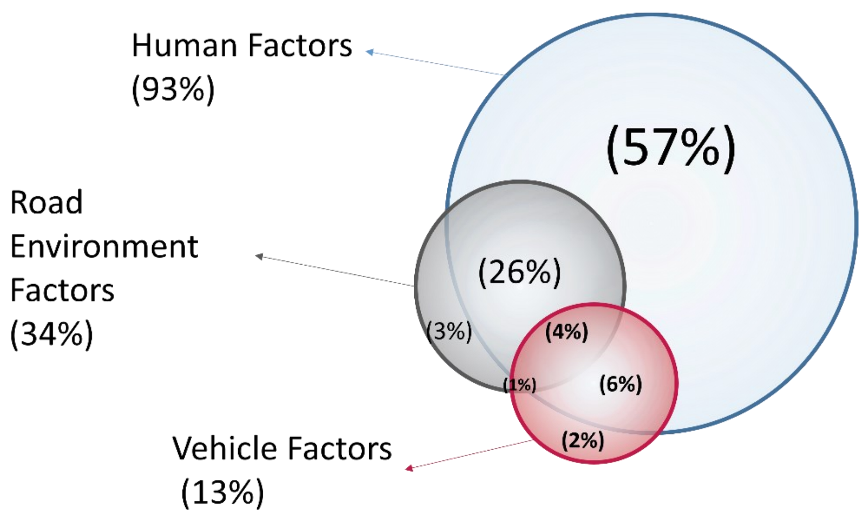


Figure 1 Road environment, human, and vehicle interaction

Compliance to a road sign is a factor within “Human” and “Road Environment” which accounts for 26% of the collisions. The actual long-term or predicted safety performance of these factors is referred to as substantive safety. The quantitative measurement of substantive safety is related to collision frequency, severity, and type. These measurements are over a long enough period to provide a high level of confidence that the recorded collision is a fair representation for a safety performance prediction.

1.1 Problem Statement

Developing sustainable road and highway safety preventive systems is a challenge for transportation professionals and it necessitates continued improvement efforts. Among the road facilities, intersections are known to be the most vulnerable infrastructure with 45% of total collisions [3]. While only a quarter of all driving occurs at night, the Highway Safety Information System (HSIS) reports that nighttime collisions account for 55 percent of all collisions. Furthermore, when all types of crossings with various control systems are considered, stop-sign operated intersections are the most vulnerable facilities, with the highest collision rate resulting in fatal or serious injuries. [4]. Twenty-one percent of collisions at SCI are known to be due to the environment surrounding the driver and "inadequate and poorly maintained signs" is often cited as a contributing factor [3]. About 8% of fatalities and 11% of total collisions occur at SCI. In Canada 1,739 individuals have lost their lives at SCI between 2008 and 2017 [5]. As a result, the SCI environment at night is a challenging and complex environment for drivers to navigate, and it is well-known as a primary contributor to traffic injuries and fatalities. Improved signage visibility at an intersection might hypothetically lower the likelihood of a collision, which is why transportation agencies propose LED stop-signs for areas with elevated risk [6]. Hence the assumption made in this study was that a "new" sign with even superior conspicuity features could cut the danger more. According to the Strategic Highway Safety Plans (SHSP) [7], the performance of highway safety is required to be evaluated for any "new" safety component or treatment.

Classic road safety analysis requires historical data for evaluation. The challenge rests in safety related data availability to conduct analysis for a new or untested treatment and understanding their impact to road safety. The classic fashion for safety impact assessment of a new treatment with observing, monitoring, and recording collisions could put traffic at risk, while it is transportation professionals' duty to prevent future collisions and not cause them. Hence, there is hesitation among transportation specialists to install a not fully known treatment or making an alteration to signage, with uncertainty in their consequences.

As of this date, we identified just a few research attempts, to evaluate the safety performance of a new regulatory sign without holding into the collision records. The main reason that historical data was not used was the limitation in data accuracy and availability [8,9]. Another challenge identified in these studies is the length of time and the magnitude of change necessary for the results to be statistically significant[10]. Understanding the limitations in surrogate safety analysis (see section 2.3.2 for full discussion), they are still widely used to evaluate traffic signage performance. For instance, some studies were successful in identifying indicators that fit the collision prediction models better than the conventional traffic volume [11–13]. These studies have been focused on conflicts pattern, speed distribution, braking habits, and deceleration, while some other studies looked at road users' behavior in compliance to the active signage as a surrogate safety measure[8,14]. Meanwhile, inconsistency in the results has been identified [8] and previous research has paid little attention to urban areas and in particular to unsignalized intersections [13].

The following questions were raised from gaps in the reviewed literatures on the topic:

A. Is it possible to assess road safety in a more effective and timely manner?
B. How to improve the safety of an experiment and avoid long term potential risk from testing a new treatment?
C. Would a change in sample combinations affect the results of drivers' behaviour assessment in a before-after, and cross-reference analysis at SCI? How may sample selection be improved?
D. How to control the effects of independent variables in treatment performance analysis?

E. Are there any additional indications for preventative SCI safety analysis that suits the collision prediction models better than traffic volume?

F. How to correlate and model road users' behavior and surrogate safety indicators for a new or tested treatment?

1.2 Research Aims and Objectives

Traffic engineers have installed flashing lights on signs for several years with the hope that a flashing lit-emitted sign increases the conspicuity of the signage and is expected to reduce the frequency of collisions at SCI. Previous research has determined these installations had a positive impact on road safety (see section 2.4.2 for full discussion). However, there is inconsistency in the findings and it's unclear if the improvement was attributable only to the signs or perhaps undetected affecting factors. This research attempts to continue the work in preventive safety analysis of active signage to indicate and define impactful parameters through a methodology which allows assessment of these types of signage.

A prototype of new and untested active signages was provided by a Montreal-based traffic signage company, Orange Traffic Co., to researchers associated with Concordia University for evaluation. In response to SHSP direction on necessity of testing any "New" safety component or treatment, a series of empirical studies was needed to evaluate this countermeasure and benchmark the behavior of drivers approaching SCI against classic active, and conventional signages.

As for statistical study, the estimate of collision (as a safety performance measure), is a function of many parameters, including traffic volume, geometry, environment as well as frequency of conflict. By leveraging the observed driver behavior from the empirical study, a new variable could be introduced to the estimation model. Therefore, this research aims to assess the value of this new variable for a better estimation of collision and conflict models.

Previous statistical studies on collision frequency at SCI have suggested that the conflict-collision based models fits the collision data better than the traffic volume-collision based models (see section 2.3.2 for full discussion). By including various independent variables, this study attempted to explore and compare the performance of both conflict and volume-based models.

The simulated conflict data can be generated indirectly using a microscopic traffic simulation, such as Vissim, Paramics, or Aimsun. However, the use of such models is justified only when they can adequately represent the real driving behaviour of vehicles on the road. Another aim of this research is to introduce driving reaction to signage into the model and look at the level of calibration required in a model, so that it can be used to produce critical conflicts under different scenarios.

Finally, the aim of this study is to help transportation professionals in improvement of public safety by understanding the performance of active signage. Even though the new active signage is conceptually similar to the former ones, the research team was obligated to conduct the experiment in a short and effective manner, avoiding waiting and collecting collision data and risking travelers in the study.

To realise these aims, the following objectives have been identified:

1. Design an experiment to assess the performance of active signage in a fashion where most variables are controlled (e.g., geometry, environment, driving population), and identify and check the statistical significance of the remaining variables (e.g., light, absence of opposing traffic, pavement surface condition). Compare the outcomes of the studies while using different sample populations. This objective will enable experiment bias to be reduced (in response to questions ‘C’ and ‘D’).

2. Examine changes in the behaviour of drivers approaching several stop-controlled intersections with various countermeasures to revalidate the performance of the active signs. This objective would provide qualitative metrics for compliance and speed patterns which then can be incorporated into the safety performance analysis and traffic microsimulation (in response to questions 'E' and 'B').

3. Extend calibration methods for the microsimulation model to incorporate observed road counts, speed, and compliance to signage; and analyze how well the driving behaviour is represented in the microsimulation model. This objective will provide reliable critical conflict information which then can be used in statistical study (in response to question 'A').

4. Establish a method of safety performance analysis for a new treatment using surrogate safety measures rather than historical collision by deploying information populated in objective '2' and '3'. This objective will enable researchers to improve the safety performance function of a certain class of SCI (i.e., three-legs) by deployment of variables related to drivers' behavior (In response to question 'E').

5. Benchmark the performance of BLS and understand the ability of a new treatment in commanding drivers' attention within a short-term trail. This objective would enable traffic engineers to understand the safety effectiveness of these types of signs in term of number of conflicts reduced in different light/traffic volume settings (In response to questions 'F').

1.3 Method and Scope

To achieve the goals set above, a research methodology including a three-pronged investigation of the safety related effects of active signs was designed: a field study, a traffic microsimulation model and a statistical study as shown in Figure 2. A detailed research framework is presented in Figure 6.

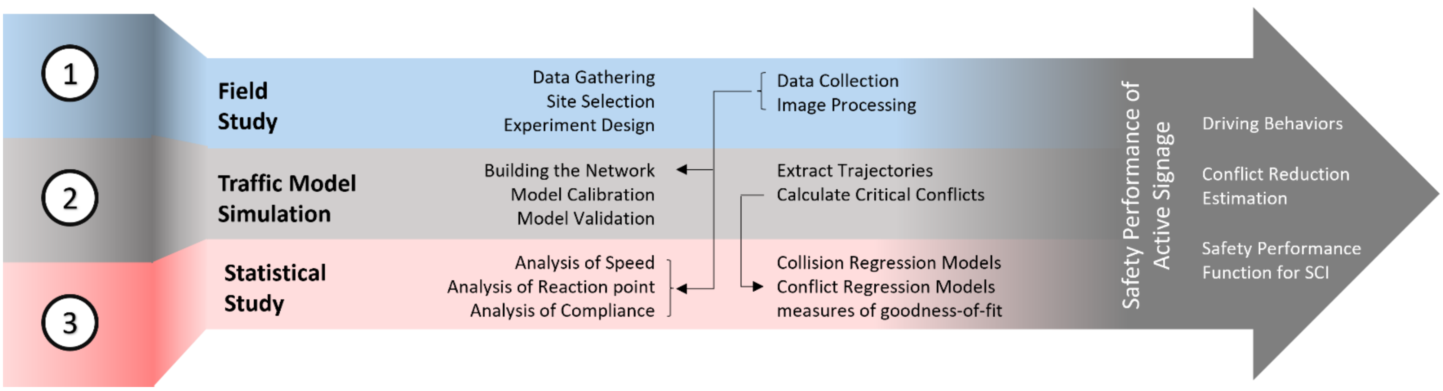


Figure 2 Research Conceptual Framework

Based on the research methodology, the following tasks were performed:

1- Select the site and design the experiment; Previous studies on active signage assessment were performed either in long term before-after installation of a flashing sign or cross reference between SCI equipped and not-equipped with active signage. To meet the first objective, it was necessary to select identical SCIs in geometry design, where the study was taking place during the same time intervals while the experiment samples went through several intersections. A link in urban areas in the city of Montreal was identified and experiments were designed in a fashion so the same traffic were exposed to several types of active signage. Data collection was conducted before and after installation of BLS to complement the cross-reference analysis. The BLS was removed within two months.

2- Data collection and processing from several sources: this includes data gathered with high-definition (HD) radars and 189 hours of video recording during 10 days. By deploying several image processing techniques, volume, speed, trajectories, compliance, and reaction to signage was captured. Supplementary data was gathered from several sources including local and national collision databases, weather conditions and rainfall, previous vehicle counts, and intersection drawings provided by the city. Several databases were created to provide concurrent access to the collected data.

3- Developing and calibrating microscopic simulation models in Vissim: this task was to fulfill the requirement of objective 3 to emulate target conflicts. Traffic conflicts are interactions between travellers that occur when one or more of them take evasive efforts to avoid a collision and, unlike collisions, they have the advantage of giving a bigger sample size due to higher frequency of occurrence. Data from compliance, vehicle speed, and vehicle count were enforced into the model set-up. Genetic Algorithm was used to optimize the models against Turning Movement Counts (TMC)s. Four models were calibrated for morning, afternoon, and evening peak traffic as well as 4 hours of twilight and dusk. Calculated conflict was validated by sample real-world conflict to guarantee the integrity of conflict data.

4- Evaluate compliance data using qualitative analysis. Several statistical models were developed for qualitative analysis including ANOVA for analysis of speed and Binomial and Multinomial Logistic Regression (MNL) for reaction and compliance. The main goal of the statistical modeling was to determine if, after controlling for several independent variables, the distribution over the three degrees of stopping compliance (i.e., 1– full-stop, 2– roll-through, and 3– blow-through) was different for each intersection and treatment.

5- Developing Statistical models for conflict and collision. For the collision model, the collision data was extracted, and cross referenced between databases. Using the hourly traffic volume captured in the step 2 and compliance data from step 4, several statistical models were developed with Poisson regression and Negative Binomial regression models. The Conflict data from steps 2 and 3 then were used to build collision-conflict regression models. The compliance data from step 4 was also used to evaluate the potential of improvement of these models.

6- Evaluate the performance of the models: To meet the expectation of objective 4, measure of goodness-of-fit for collision and conflict models was performed using a variety of comparison indicators, including AIC. The estimated conflict results, from the best model for SCI with standard signage, was then compared to the performance of intersections utilizing BLS from step 2 to achieve final objective.

1.4 Thesis Contribution

The first and second objectives of this study are evaluating the performance of tested and, untested regulatory signs with higher conspicuity. This study was able to reenforce the previous findings on the subjective benefits of active signage on road safety metrics. Beside the ‘compliance’ and ‘presence of opposing traffic’, other independent variables were assessed, such as pavement surface condition (e.g., wet, dry), turning maneuver, and ambient lights in the qualitative analysis. The same intersection geometry allowed for bias control as the result of design variance. The design of the experiment provided a unique opportunity to have same drivers attend several intersections and the benefit of this set-up was identified. Furthermore, based on our research in the available literature, the correlation between collision and compliance to the active signage at SCI has not been covered before and is the motive of the third and fourth objectives. The research achieved the goal of enriching the collision prediction models by inclusion of the driver compliance and conflict. This study was conducted in urban areas where other studies focused in sub-urban or rural areas. Finally, two types of regulatory BLS signs were assessed for the first time and the findings are being shared. The assessment exposed the traffic to this new treatment for only two months and the proposed method warrants less "potential" risk to the public.

1.5 Thesis Structure

As indicated in Figure 3, this thesis is divided into five chapters. The current chapter outlines the research's context and motivation, as well as a glance into the rest of the thesis. In chapter 2, some of the most relevant literature is reviewed to present the existing methods adopted by agencies and research by academics for road signage improvement and evaluation in transportation and road safety. The different approaches and methods are presented and investigated while each influence, limitations and gaps are discussed. The main area of focus was on the investigation of treatments' effects on road safety.

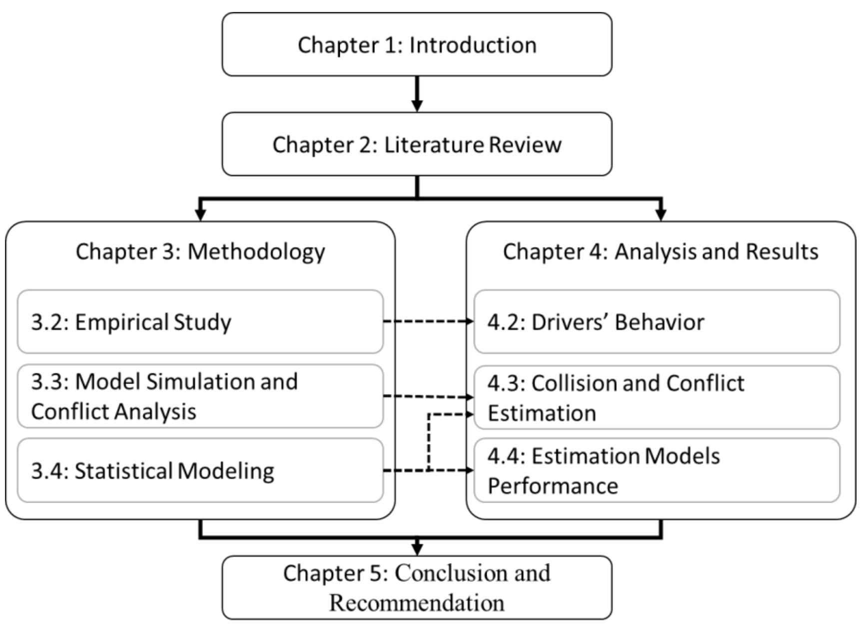


Figure 3 Thesis Roadmap

Chapter 3 describes the design of the empirical research, qualitative study, and statistical study. The method and process of building the microsimulation network using a genetic algorithm is also outlined in the chapter. Chapter 4 elaborates and presents the results of the analysis of driving behaviour in terms of speed and sign compliance under three scenarios. In the same chapter, the analysis of conflict data and the construction of collision-conflict regression models are detailed, followed by an evaluation of the regression models' performance. Finally, chapter 5 summarises the findings, offers the conclusions, and outlines potential future study directions.

2 Chapter 2: Literature Review

This chapter consists of a focused summary of published research literature related to road safety. The review starts with an assessment of the correlation between sight distance and vehicle speed as a key design element for road alignment, marking and signage. It also contains a brief overview of road safety analysis using collision and/or conflict data as a safety measure. After this section, an overview of studies on microscopic traffic simulation and surrogate safety assessment models is being presented with a focus on conflict analysis at un-signalized intersections. Then, the next section of the review entails the efforts and studies underlining the limitations of traffic safety simulation models with respect to safety treatment effects, and some prior attempts on traffic signage evaluation is being presented. The final segment is focusing on the subject under consideration of this thesis, providing an in-depth overview of recent studies on the enhanced traffic signs and describes them here after.

2.1 Introduction

Road signs are visual graphics, symbols or messages that are created to display information and communicate with the road users. Traffic signs are defined as part of traffic control devices by the Federal Highway Administration (FHWA) and tailored by jurisdictional authorities around the globe to be used and regulate, warn, or guide traffic. For example, these guidelines enable traffic engineers, to estimate distance required according to the sign size and design considering the posted speed [6,15]. The traffic signs are among the countermeasures intended to decrease crash frequency and/or severity [15].

Over the past several decades, studies determined the time interval requirements for the drivers to detect, recognize, read, and respond to the roadside posted signs [16–18]. There are many factors influencing drivers' perception and reaction to these signs (e.g., general lighting, drivers' awareness or distraction level, quality of the sign, etc.) [19]. As a result, transportation professionals and specialists seek to incorporate all these variables into design to ensure that a sign fulfils its primary function of conveying the message, at all times, and in a variety of settings. The challenge of interpreting the conventional signs in adverse nighttime conditions under adverse road weather such as rain, fog, snow or storm has been addressed in several studies [20,21]. Some studies have identified that low-contrast, low-diffusion, ambiguity, visual noise, clutter, information load, and complexity of signs could be root cause of some drivers' failure to comprehend them [20,22]. As a result, through past decades, engineers have introduced new materials to improve retro-reflectivity of traffic signs' face, hence enhance their visibility under various conditions. Efforts was put in place by authorities to improve visibility, boosting conspicuity and, legibility of these type of traffic control devices. Advances in electronics and miniaturisation is improving signs' attraction when needed. Signs with flashing lights, actuated signs and illuminated signs are some examples of these kind of efforts. Different guidelines recommend that light emitting diode (LED) units might be used in regulatory or warning signs (also known as Active Road Sign), to improve the conspicuity of signs[6]. These enhancements are expected to improve safety and reduce the frequency of collisions.[23] An overview of recent studies on this matter is performed and categorized as described here after.

2.2 Sight Distance:

The American Association of State Highway and Transportation Officials (AASHTO) has a set of recommendation for sight distances and associated design speeds, to meet safety requirements in geometric designs for highway ramps, intersections, horizontal and vertical alignments. The recommended values are based on driver's perception and reaction to different stimuli such as the ability of drivers to recognize the need to take a decision and react with an appropriate response (e.g., turning, braking, etc.). For example, the recommended perception reaction time for horizontal and vertical alignments of uninterrupted facilities is 2.5 seconds [24,25]. The design values should be conservative and typically accommodating at least 85% of the driving population. There are recommendations based on grades, vehicle types, geometry of intersections, number and capacity of approaches. For instance, in an intersection with stop-sign, AASHTO provides detail analysis for sight distance requirement while considering the time gap needed for vehicles performing both crossing and turning maneuvers. The sight distance for a passenger car, approaching a stopped controlled intersection is equal to $7.5\ s$ of travel time at the design speed on the major road to the stop line[25,26]. This is a generic rule of thumb to estimate design speed while other parameters such as visibility and geometry might being ignored. The other driving behaviors to which the driver must attend and the amount of time these actions require are also not considered.

AASHTO recommend Minimum Required Stopping Sight Distance (MRSSD) from the conflict line (for instance stop-line at intersection) to be calculated as follow [24].

$$MRSSD = V_s T_r + \frac{V_s^2}{2a \pm 2gG} \qquad \text{Eq. 1}$$

V_s is the velocity or 'set' speed for side or traffic-controlled road (m/s), T_r the driver's perception–reaction time (s), default $2.5\ s$, g the gravitational constant ($\sim 9.81 m/s^2$), a is the deceleration rate(m/s^2), default 3.4 m/s^2, and G the grade in decimal.

For SCI, the stop and/or YIELD signs AASHTO recommendation is that they should be visible from an enough distance to command drivers' attention. The Eq. 1 is being used by transportation specialists to determine that distance. This is the same requirements for all other regulatory or warning signs such as chevron markings on sharp curves or pedestrian crossing and school zones. Therefore, it is critical that the Available Sight Distance (ASD) for any traffic control sign to be greater or at least equal to the minimum required stopping sight distance mentioned above [17]. Some studies investigated the proper ASD with respect to ambient light conditions and considered human factors such as drivers' eyesight. In one study, the maximum distance for sign identification and the appropriate installation distance was evaluated under nighttime condition. Traffic sign background luminance, luminance contrast, complexity of information, size of the symbols, and geometry of the facility are among the factors that was considered in the study [16,26] .

Another study investigated both visibility (ability to observe the presence of signs) and legibility (ability to distinguish and read/interpret signs). Schnell et al. -2004, have deployed simulation to evaluate the required ASD in laboratory environment [27]. The ASD for a sign may be determined via field testing experiments using drivers with minimum acceptable visual acuity. Awadallah et al. -2009, suggest testing these properties by measuring the sign retro-reflectivity[17]. Jones et al. -2012, proposed a technique to perform the test in a normal open road environment using digital cameras and stereoscopic image processing techniques[28]. .Altamira et al. -2010, proposed a tool to analyse availability of sight distance based on three dimensional (3D) visualization. All these studies had come to the same conclusion that visibility and legibility alone may not guarantee the performance of the sign since poor reaction times of drivers (decision/ command) and compliance to the sign may create a threat to the environment [18]. Limitations in the lab environment causes lack of confidence on safety performance measure since field and lab environment are principally different. It’s not possible to evaluate the safety matrix such as collision and conflict in a lab-controlled environment, therefore empirical studies are needed to fill these gaps.

2.3 Road Safety Analysis

Road safety analysis is used to advance our understanding of the state of practice in road safety, and consequently, improve the public safety within a given transportation network or facility. Through traffic safety analysis, one seeks to improve road safety by reducing vehicular conflicts and, implicitly, collision. Therefore, information about collisions frequencies or collisions severities, or both, is used to assess and predict the safety performance of an existing or a new transportation facility. According to Highway Safety Manual (HSM) a "crash" (referred to as "collision" in Canadian literature) is defined as a set of events that results in injury or property damage due to at least one motorized vehicle with other motorized, un-motorized or stationary objects [29]. Collision type, geometry, operation, weather, and visibility conditions along with human factors are some of the most frequently used indicators in collision analysis [29].

For level of injury HSM is recommending the American national injury standard which divides injury levels into categories according to KABCO where K represents the most severe injury and O means no injury, only property damage [30]. Subjective safety analysis is a form of safety assessment that represents the opinion of the road users, public or professionals to evaluate the level of safety of a given facility. Hence these kinds of analysis are more qualitative. Nevertheless, road safety management programs use rather objective safety analysis methodologies, which requires quantitative measures.

The quantitative methods using collision data typically require large sample size from reference sites to develop reliable Safety Performance Functions (SPF). This can be costly or even impractical (e.g., not applicable to a new infrastructure or safety countermeasure). According to several studies a considerable number of collisions with injuries are not being reported in police database [31]. The underreported collisions are those with lower severity and/or without property damage (e.g., to avoid increasing insurance premium). Researchers have addressed the presence of underreporting in collisions[14,32–34]. The main challenge in many safety studies is lack of collision data or data quality with respect to data quality measures [35–37]. For example, in Québec, *Société de l'Assurance Automobile* (SAA) has defined a minimum damage threshold which, if not exceeded, allows these minor collisions to go unrecorded. Besides lack of data and sometimes limited availability, another major problem with objective safety analysis pertains to benchmarking new road safety treatments/implementations [38,39].

Non-collision-based research are widely considered as an alternative approach to the classical collision studies. These research entail studying specific recordable traffic parameters that can be associated with near-collision events, which are called surrogate safety measures. Using more eventful data in a shorter time period is expected to provide the grounds for reliable traffic safety analysis, without focusing on collision which is a rare event. Studying the surrogate measures has enabled traffic specialists to leverage computing and simulation power to mimic the real-world traffic conditions. These micro-simulation traffic models attempt to close the gap between insufficient events such as collision and limitation in observation period. In addition, it is important to highlight that these models are unable to measure safety performance directly by estimating "collision" . In this section a brief review of previous research on collision frequency analysis and traffic conflict assessment methods is presented.

2.3.1 Collision Based Analysis

Road safety assessment using collisions data needs collecting and processing the pertinent data. In North America, there are several sources that tracks and store this information, such as: Highway Performance Monitoring System (HPSM), Fatality Analysis Reporting System (FARS), National Electronic Injury Surveillance System (NEISS), and finally, Motor Carrier Management Information System (MCMIS). Transport Canada has also made this information available to public under national collision database (NCDB)[40]. The depth of data varies based on facility type and the location. Unfortunately, not all transportation facilities have enough data to meet HSM minimum three years of data requirement for collision analysis [29]. Collision data consist of three elements; collision types (based on the injury type) and initial impact type on the represented clusters, collision's location using the location's GEOID, and date and time. Beside the collision data, secondary data are essential for safety analysis. The shape and characteristics of the infrastructure under the study as well as the operational data would be also necessary [41]. The physical characteristics includes length of the segment, area, road classification, intersection type, and traffic control devices. For operational data usually average daily traffic (ADT) or average annual daily traffic (AADT) are incorporated[10].

There are two safety performance measures often being used: collision frequency and collision rate. Collision frequency represents the number of collisions observed in a given period, typically one year, and can be calculated using Eq. 2. Collision rate represents collision frequency with respect to a measured traffic volume (e.g., number of collisions per one thousand or one million vehicles).[38]

$$Collisions\ Frequency\ = \frac{Number\ of\ Collisions}{Period\ in\ Years\ or\ facility\ segment} \qquad \textit{Eq. 2}$$

Collisions are nonnegative and discrete events which with deployment of some tools, are used to indicate the safety performance of the transportation facility (intersection, railroad link, terminal, …). There are several tools that are used which are mainly Poisson based. All these tools have the objective to find an estimates for collision frequency or severity and the variance. Figure 4 presents the extensions of Poisson-based models. The negative binomial (NB) is the most widely used among Poisson-based models due to its ability to address overdispersion[42].

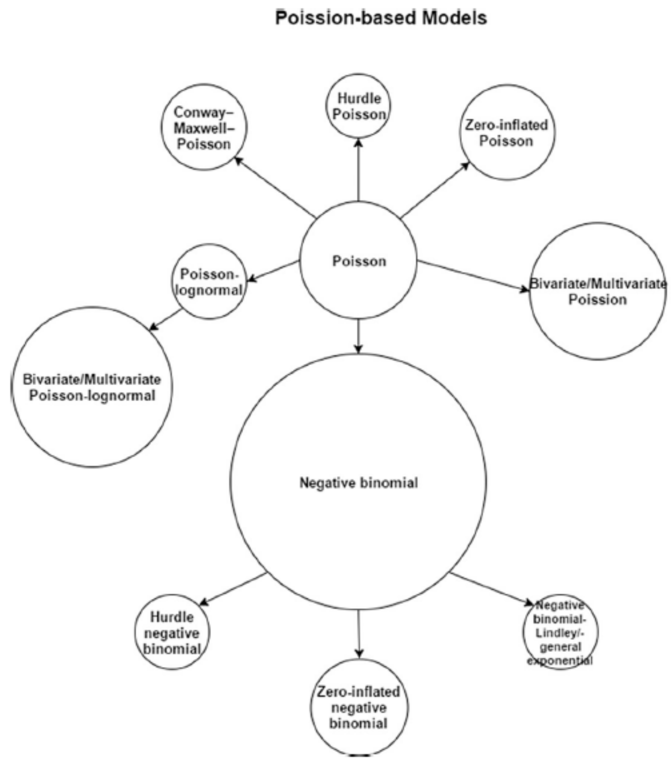


Figure 4 Family tree of crash-frequency methodologies (Lord. D.2021)- Circle size indicates the relative popularity of one methodology.

A simple regression model to estimate collision for a SCI using a single variable with consideration of AADT following Poisson distribution model, can be estimated from this regression model [43–45]:

$$\frac{Collisions}{year} = \ln\left[\alpha(Major\ approaches\ AADT)^{\beta 1}(Minor\ approaches\ AADT)^{\beta 2}\right] \qquad Eq.\ 3$$

The parameters α is the value of intercept, β_1, and β_2 are the regression parameters. These functions relate the collision number and types of collisions at a site to traffic volume and in more comprehensive form other road or drivers' characteristics. They have been developed for each section of road (e.g., midblock, intersection), and different collision types (angles, rear-end), based on historical collision records. The regression model used for SPF assumes that collisions are following Poisson distribution [38]. The Binomial distribution is approximated by a Poisson distribution as this is commonly assumed in counts of rare events occurring in each time period. However, the Poisson distribution assumption leads to equality between mean and variance. Therefore, it does not allow for the variance to be adjusted independently from the mean. As it is shown by many studies, over-dispersion (or rarely under-depression) of collision data makes this assumption obsolete. Literature has identified two possible reasons for over-dispersion: heterogeneity (i.e., inaccuracy in observed data, overlooking an important attributes) and high proportion of zero accident periods [46].

Some conjugate models such as two stage mixed Poisson models can address the random variations in the mean of the Poisson model. Among those, Poison-Gamma (Known as Negative Binomial) and its extensions and Zero-Inflated Regression Models are more well-recognized in road safety analysis [47]. According to Persaud et al., the NB method is worth considering for situations where the main constrain is having access to enough reference group to calibrate required EB's SPF [48].

2.3.1.1 Observational Before-After Studies

The performance measures are being used to benchmark a setting or a segment to another similar cases or understand the impact of a change in a "before" and "after" analysis. Statistical analyses are being employed to recognise these magnitudes. Hypothesis testing like t-test, z-test and H-test are being used to assess the statistical significance of estimate from an added treatments to an existing road section compared to the similar road sections [49]. The assumption that population distribution is normal is not always correct for the collisions [37,38,44]. The t-test is regularly deployed whenever two compared proportions are dependent, while z-test can be applied with null hypothesis of independent extents. H-test (one-way ANOVA) is another hypothesis testing tool, but nonparametric, through which the combined distribution of the populations is being tested rather than between distributions [50]. H-test is widely used in safety assessment in comparison between populations, and the distribution of test statistics can be determined using the Eq. 4.

$$H = {}^{12}\!/_{n(n+1)} + \sum \frac{R_j^2}{n_j} - 3(n+1)$$ *Eq. 4*

Where:

n_j= number of measurements of the j^{th} sample

n= the total sample size from all combinations

R_j= rank sum of sample j

The changes observed in comparison may be in part due to the spontaneous *Regression To the Mean* (RTM) or site selection bias and not as the effect of treatment or change in countermeasure [39]. Hence, it is naïve (i.e., not statistically sound) to infer difference from a comparison on a non-random selection of a site to be only due to countermeasure. Hauer addressed this potential bias of RTM phenomenon by proposing a four-step improving approach for naive before-after study. In this method, a matching comparison group is used to control RTM where some (or all) of the changes are being removed due to any factors other than the treatment. He also identified five groups of factors, and consequently, making the conventional assumption (i.e., the observed change is due to the effect of the treatment) questionable [39].

Statistical inference such as Empirical Bayes (EB) Method was originally deployed for observational before-after safety assessment of treatments and countermeasures [38]. EB is known to eliminate the effects of RTM better than the conventional comparison methods. For several decades this method has been the core of traffic safety studies used by transportation safety specialists and agencies around the world in conducting before and after studies either by deploying traffic flow factor, comparison groups or other techniques[39,51–56] . This method is used widely to provide general feedback about effect of a countermeasure. The results of a survey conducted by National Cooperative Highway Research Program (NCHRP) in United States and Canada reported EB as the most common method in before-after evaluations [57]. EB comprised a prediction and an estimation; Prediction of the expected collision rate during the "after" period, without treatment and, estimation of the collision rate of the treated site "after" the conversion being deployed. To find the "predicted" collisions rate, two sets of information are being combined; the annual number of collisions that would be expected at the case study based on characteristics (i.e., traffic volume with an EB procedure in which regression model is being used), and the count of collisions in n years before conversion.

These comparison sites should have similar geometric and traffic characteristics as treatment is applied. The variance and the mean from the regression model along with the collision distribution can deliberate the dispersion factor properties. There are several techniques which address the over/under dispersion phenomenon in EB [46]. Some of them will be explained in next sections.

To estimate the expected collision per year during the "before" period (considering RTM effects the collision rate) Eq. 3 is being used with the "before" AADT combined with "before" collision count. For "after" period two steps need to be completed:

Step 1 - **Predict** safety of an entity in the "after" period, had the treatment not been applied. For the "after" period, in order to estimate the expected collision per year 'A', the collision rate can be calculated using same equation while "after" AADT value is being deployed. The ratio between before and after will then give a "predict" of what would have been the collision rate without the treatment. The summation of all 'A' is then the predicted number of collisions in the "after" period. The variance of the collisions will depend on the method used for predicting the value and there are several methods available for the analysis [38].

Step 2 - **Estimate** the treated entity's safety associated with the "after" period. The mean value of collisions and the variance in "after" period, denoted by 'B', are estimated to be the same, since it is assumed that follows a Poisson distribution. The summation of all 'B' values then will represent the 'estimate' number of collisions after the treatment- Eq. 8 denotes the mean and variance of 'prediction' as well as 'estimation' for the after period.

$$\hat{\mu} = \sum A \qquad \textit{Eq. 5}$$

$$var\ (\hat{\mu}) = \sum var\ A \qquad \textit{Eq. 6}$$

$$\hat{\lambda} = \sum B \ \ var \qquad \textit{Eq. 7}$$

$$var\hat{\lambda} = \hat{\lambda} \qquad \textit{Eq. 8}$$

Where: $\hat{\mu}$ is the "prediction" of collisions frequency for all sites/year A and $\hat{\lambda}$ is the "Estimation" of collisions frequency for all sites/year.

The target collision reduction value [δ] and its rate [θ] after treatment is installed will be:

$$\hat{\delta} = \hat{\lambda} - \hat{\mu} \qquad \textit{Eq. 9}$$

$$\hat{\theta} = \hat{\lambda}/\hat{\mu} \qquad \textit{Eq. 10}$$

As mentioned in step 1 and 2, both $\hat{\lambda}$ and $\hat{\mu}$ are random variables (have a mean and variance) and hence $\hat{\delta}$ and $\hat{\theta}$ would also become a random variable. The step 3 and 4 in the EB method provide the $\hat{\delta}$ and $\hat{\theta}$ estimates and their variances using [Eq. 11-Eq. 14].[29,38]. $\hat{\theta}$ indeed, is the Collision Modification Factor (CMF) which offer great assistance to road designers and used in the safety performance function to identify the proper treatment in making preventive decisions. The FHWA provides a searchable database of CMFs along with guidance and resources on using CMFs in road safety practice[58].

$$\theta = \frac{\frac{\lambda}{\mu}}{\mu\left\{1 + \left[\frac{var(\mu)}{\mu^2}\right]\right\}}, \qquad \textit{Eq. 11}$$

$$\delta = \mu - \lambda \qquad \textit{Eq. 12}$$

$$var(\theta) = \frac{\theta^2\left\{\left[\frac{var(\lambda)}{\lambda^2}\right] + \left[\frac{var(\mu)}{\mu^2}\right]\right\}}{\left\{1 + \left[\frac{var(\mu)}{\mu^2}\right]\right\}^2}, \qquad \textit{Eq. 13}$$

$$var(\delta) = \mathrm{var}(\mu) + \lambda \qquad \textit{Eq. 14}$$

The FHWA has also included a safety analysis module in the Interactive Highway Safety Design Model tool (IHSDM) which can be calibrated to estimate collision rates using the EB method. The efficiency of the method is recognized in literature and has been suggested under Highway Safety Manual (HSM) Part C - Predictive methods [56,59]. The more generic form of the prediction function mentioned earlier – known as Safety Performance Function (SPF) has been described in five steps by Garber [60].

Summary

Conventional before-and-after comparison of collision rates has limitations such as ignoring complementary data (e.g., drivers' behaviour such as compliance to traffic control device). Even though, having control groups might improve the estimation it has not been recognized like statistical inference method, such as Empirical Bayes (EB). EB method has a proven record on literature reviewed and have been validated [39]. Some issues and drawbacks of these method has been addressed that may affect the validity of some results mainly caused by over-dispersion [56].

In some studies, FB approach has been suggested to tackle some of the key issues with EB method, such site selection bias. A set-up of FB with zero-inflated NB modeling for prediction is an appropriate method to address data limitation on the control group as well as excess zero counts [61]. The sophistication cost of NB makes the EB method more attractive among traffic safety analysts[62].

Finally, it is important to re-acknowledge that observational collision studies are opportunistic endeavours. Nevertheless, according to Hauer, to detect reliable effectiveness of a treatment the number of samples for collisions need to be enough, such that the standard deviation of the estimate should be 2 to 3 times smaller than the expected effect [39]. This requires tens of thousands of collisions to be detected for a minor change in safety of only a few percentage points. This requirement is not feasible in most of the studies.

2.3.2 Conflict Based Analysis

As previously mentioned, many collision datasets often present data quality challenges (e.g. inaccurate collision reporting, uncertainty of data, limited availability, etc.). In addition to data limitation, some research points out other uncertainties in Bayes methods to give the best collision estimate [63]. Elvik et al. -2004, suggests that road safety improvements are implemented for a variety of reasons and that a randomly high collision count is not among the key selected criteria. This study was conducted in Norway for both treated and untreated sites, over several years, and revealed that the random effect of fluctuation in collision rates were equally likely.

In principle, collisions are rarely occurring events, while there are other more frequently types of events which can be correlated to the road safety performance. Estimation and appraisal of safety based on proxy information also known as “Surrogate Safety Measures” (e.g., time to collision, post-encroachment time, deceleration rate, maximum speed, and speed differential) has been explored in many studies to overcome the shortage of collision data. Traffic conflicts represent the cornerstone of these measures to objectively assess the collision potential of a location without having to wait for the collision history to accumulate.

Traffic conflicts are defined as interactions between travellers that occur when one or more take evasive actions (e.g., braking or weaving) to avoid a collision (i.e., if the travellers’ movement remains unchanged, that move may potentially lead to a collision [64]). Unlike collisions, conflicts have a distinct advantage as they are providing a larger sample size due to higher frequency of occurrence. This enables researchers and agencies to assess the safety of facilities more promptly. The concept of traffic conflicts first developed in late 60’s by industry (General Motors engineers) to identify safety issues related to vehicle fabrication. In that approach, a set of definitions and procedures was formed for observing traffic conflicts at intersections as an alternative to collisions [65]. Hydén demonstrated the relationship between collision and conflict in a safety pyramid where, the basis of the pyramid is undisturbed passages followed by disturbed and interrupted passage. In the middle of the Hydén pyramid we have conflict with layers from light to critical conflict and “missed collision”. At the summit of the pyramid, a collision occurs with the severity KABCO, where "K" denotes a fatal injury [64].

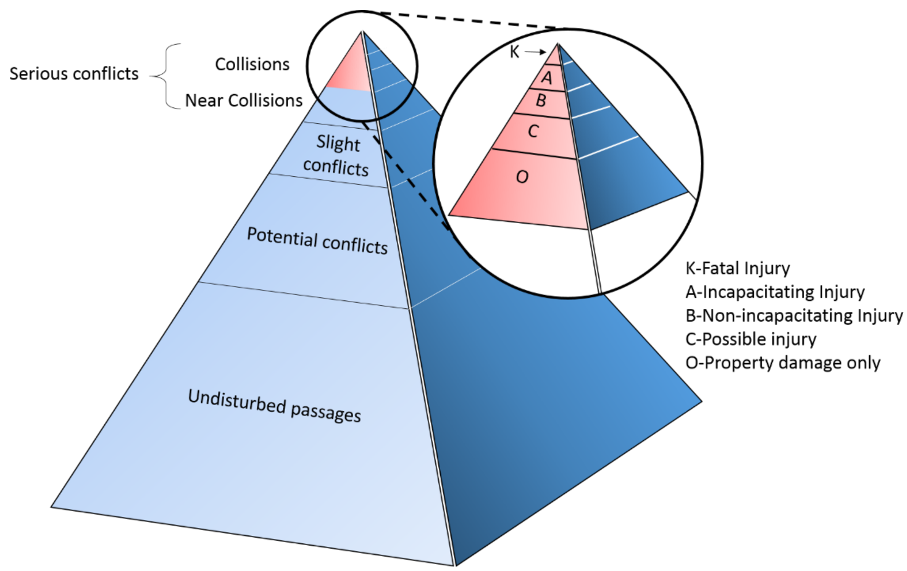


Figure 5 Hydén Safety Pyramid (Hydén, 1987) - collision scaled to KABCO

Over the years, transportation professionals have attempted to better understand the amorphous relationship between several independent variables such as speed, traffic volume, conflict, and collisions, as the “dependent variable”. Numerous regression-based approaches have been used to develop predictive models that correlate conflicts and road and environment parameters [65–68]. These studies attempted to identify specific conflict patterns instead of simply using collision rates as diagnostic tool for traffic safety assessment. Rao et al. -1997 introduced a probabilistic model for SCI to determine the number of conflicts assuming a vehicle crossing time at the intersection is less than the average inter-arrival time of vehicles on the opposite approaches [69]. The time to collision (TTC) first was introduced by Hydén -1987 whom attempted to relate TTC with speeds and determine conflict severity. The author investigated 50 intersections and demonstrated a correlation between reported collisions with injury and observed serious conflicts[64]. In another study, Clennon and Thorson -1975 revealed a relative correlation among the intersection traffic flow and injury severity [70]. Caliendo & Guida -2007 investigated the relationship between simulated conflicts and recorded collisions at two-way SCI in Italy[66,71]. The authors also tested whether collisions correlate better with conflicts than traffic flow. They introduced a conflict prediction function to model the expected number of collisions. The authors also discovered that their safety analysis approach outperformed the traditional traffic volume-based model by a small margin.

The observation of conflicts can be performed either manually (human observers), semi-automatically or fully-automatically. FHWA introduced a user guide as a training aid and reference for practitioners and engineers for conflict observation [72]. Computer aid packages with image processing capabilities are frequently employed, as are telematics and sensor technologies, as well as vehicle trajectories generated from microsimulation models. The main challenge in conflict analysis is that the method is subjective and relies either on observers' judgment, quality of image processing, accuracy of sensors or performance of simulation model. Several studies have yet criticized manual collection of conflict data for potential lack of consistency [70,73]. Nevertheless, Parker and Zageer study shows that there can be 80% agreement between different observers[74], while Davis et al. 2014 recommend improving the manual data collection properties by setting up an observer panel and scoring system [9]. The surrogate safety measures, outlined in FHWA guideline included in conflict analysis are [75] [21]:

- "Time to Collision", represents the expected time for two or multiple road users to collide if they remain at their current speed and direction. When vehicles are on a collision course, the collision severity and risk of conflict increase when time to collision values become smaller.
- "Post Encroachment Time", is the time lapse between end of encroachment of a vehicle and the time that another vehicle (or road user) subsequently arrives and occupied the same spot.
- "Deceleration Rate", represents the rate at which a crossing vehicle must decelerate to avoid a collision.

- “Time Headway”, is the time laps between the passing of two consecutive vehicles past a known, fixed location.
- “Conflict Type” ,describes whether the conflict is the result of a rear end, lane change, or crossing movement.
- Speed related parameters such as “Maximum Speed” of either vehicle throughout the conflict, and “Delta Speed” which is the difference in vehicle speeds as observed with respect to direction and approach.

2.3.2.1 Microscopic Simulation and SSAM

Although the early computer-based modeling techniques were ineffective to replicate the real-world scenarios, the advancement in computational power and simulation algorithm has allowed for more traffic modeling opportunities. Gettman and Head –2003 investigated and defined the process in evaluation of road safety using microscopic traffic simulation software packs. This process was denoted as the Surrogate Safety Assessment Methodology (SSAM) which includes detailed guidelines for simulation development, calibration procedure, data extraction, and presenting the results [75]. Sayed-et al, have developed a similar tool that captures TTC (Time to Collision) as the critical conflict event from simulations [73]. Over time, SSAM evolved and became the foundation for future conflict analysis in objective safety assessment. In 2008, the Federal Highway Administration (FHWA) released a software package that generates surrogate measures from trajectory data files recorded in traffic microsimulation models. Most of commercially available microsimulation software generate vehicle trajectory as an output file[21] and several studies have deployed SSAM in their objective safety analysis for different road facilities such as intersections, homogeneous road segments, roundabouts [11,66,76–78]. The following section reviews the studies that considered conflict analysis at SCI facilities utilising SSAM.

2.3.2.2 Micro-simulation Approach Conflict Analysis at Intersections

Intersections, in general, pose a safety risk due to the increase in crossing paths, as well as the increased likelihood for unsafe driver maneuvers due to complexity of the task. There are conflicts not only in vehicle trajectories of each approach, but also exposes risk for not-motorized traffic. SSAM has sparked a growing interest in constructing microsimulation traffic models to investigate the traffic behavior and safety assessment of un-signalized intersections.

Pirdavani et al. -2010, used Post Encroachment Time (PET) as the primary proxy safety indicator[79]. In this study the microsimulation software package used was PARAMICS and two parameters, speed limit and traffic volume, were selected as the calibrating parameters for this surrogate safety study. Four loop detectors were set at each approach of a four-leg intersection to capture data from the experiment, which defined four conflict zones inside the intersection. The findings of this study showed a significant correlation between the selected parameters and critical conflicts. Speed limit and traffic volume had a high impact on PET behavior. However, this method could only capture crossing conflicts and no other conflict types. The authors didn't indicate if any calibration of the microscopic simulation was performed. Their approach was reused at a T-intersection in Mumbai, India. Similarly, the PET measure was used as the proxy conflict indicator and the two parameters of speed and volume extracted from the video recording was used to calibrate the microscopic simulation model, this research used another microsimulation software package namely VISSIM [80]. In order to detect the conflicts, the authors employed video recording along with image processing to identify unsafe maneuvers. The authors suggested breaking down the conflict area of the intersection under the study to small pieces using small rhombus tiles (unlike rectangular grid) for a better transformation from the camera screen coordinates to real-world coordinates. Depending on the combination of tiles occupied by the vehicles, the authors distinguished between safe and unsafe cases however this was not clear that how this could impact the

PET or TTC. There was no benchmark between their methodology and the regular rectangular grids however, by changing the traffic volume or vehicle trajectories they generated graphical correlations between PET and the frequency of conflicts.

In another effort, El-Basyouny-2011 aimed to find the relationships between collisions and conflicts for signalized intersection. The authors compared volume-based to conflict-based SPF's [81]. The data was collected from 51 signalized intersections in British Columbia, Canada and the authors deployed FB models for collision statistical analysis. As an alternative to Poisson, NB modeling based on average hourly conflicts (AHC) was proposed for the development of SPFs. The authors developed conflict-based regression models and found a significant correlation between collisions and conflicts upon adjusting for measurement errors compared with the ADT NB model. The proposed model fitted the collision data equally well. However, the regression model for the collision analysis did not contain any other independent variables, and the author did not specify whether geometric attributes of intersections were considered. This research focused on signalised intersections, which differ from SCI in terms of traffic operations.

Another study investigated the estimate of collisions at SCI in Italy. The authors employed simulated conflicts for two-way SCI by calibrating an AIMSUN model instead of using observed conflict counts. Their finding was supporting the results from the earlier Canadian study, and also demonstrated a significant correlation between collisions and critical conflicts. The authors suggested that conflict-based models are a better alternative to conventional traffic volume-based model [66]. The collision model proposed by the authors has been derived from SCI regression model (Eq. 15) with volume per hour (VHP) and a dummy variable (DU) was introduced to unbiased rush hours' time intervals.

$$Estimated\ Collisions = e^{\beta_0+\beta_1 DU}(VHP_{Minor})^{\beta_2}(VHP_{Minor})^{\beta_3} \qquad \textit{Eq. 15}$$

Where:

β is Vector of regression parameters ($\beta_0 - \beta_3$)

DU is a dummy variable to control the effect of intersections with significant higher volume on Major in rush hours

VHP_{Minor} is hourly volume on Minor links

VHP_{Major} is hourly volume on Major links

The collision-based regression model proposed by the authors, included three types of conflicts; crossing, rear-end and lane changing on both PET and TTC parameters. The regression model from Eq. 16 estimates the frequency of collisions based on the independent variable, hourly conflict(cf).

$$Estimated\ Collisions = e^{\beta_0 + \beta_1 DU}(cf)^{\beta_2} \qquad \textit{Eq. 16}$$

The authors conclude conflict-based SPF fits better to the collision data than the classic volume-based regression model. They also explained that due to close correlation between conflicts and traffic volume, when conflict variable is included in the volume-based regression model, it does not significantly contribute toward performance of the model and make traffic volume variable insignificant. This study has been developed later [82] in another SCI conflict investigation to develop collision prediction models based on traffic volume and the correlation with two other safety measures, conflicts, and delay. The authors included 133 two-way SCI (78 three leg and 55 four leg) in greater Toronto region. Vehicular traffic at each site was modeled with Synchro to estimate the intersection delay, and microscopic traffic simulations were conducted with VISSIM. The delay at intersection includes both left turn maneuver delays as well as delay at entering the intersection. This study also presented the value of developing conflict-collision prediction modeling compared to the traditional collision-volume modeling approach.

Lorion and Persaud research results both from the delays of each movement and conflict-based collision prediction models were shown to be reliable for types of intersection analyzed (i.e. three and four leg intersection, respectively). Neither the site attributes - geometry of these intersections were considered in this study nor the drivers' behavior. The calibration process and the reliability of the micro-simulation model was also missing in the report.

Several researchers have raised two main concerns about using simulated conflicts [66] [78]. First, unlike natural driver behavior, simulation behavior in the models has been designed in a fashion that avoids unsafe vehicle interactions. Second, there are many model parameters in micro-simulation models and several ways to model traffic in micro-simulation models (e.g., priority rules). These parameters can have a significant impact on the result in simulated conflicts. Hence in order to derive valid results, the selection of the right parameters and proper calibration of the model are essential.

Essa and Sayed in a study intended to find out to what extent simulated traffic conflicts can accurately represent field-measures. They have identified a four-leg signalized intersection in BC to conduct the investigation. Two cameras on each approach were used to record vehicular movements for 60 hours. Automated conflict detection was carried out by utilizing image processing tools. After observing real world conflicts, the authors proposed to calibrate VISSIM in two steps using a genetic algorithm (GA) procedure. The authors concluded that there is a reasonable correlation between simulated and real-world conflicts (especially after appropriate calibration). Their study included one type of conflict (rear-end), and one surrogate safety measure indicator (TTC). This study was also conducted at signalized intersections, however driving behaviour is expected to be different from SCI [83].

2.3.2.3 Microsimulation Calibration

According to studies mentioned in previous section, traffic conflicts provide useful insight into the traffic safety failure mechanism that leads to road collisions. They are more common and have a lower social cost. However, the fundamental constraint remains, that simulation is not easily replicating real world driving conditions. Despite the limitations in traffic microsimulations, the flexibility in modeling, being a low-cost process, and integrability in iterative operations, makes them an alternative tool in proactive safety studies. A microsimulation model still can be used to emulate the geometric alignment, traffic characteristics and driving behavior of the real-world network, if it goes through a proper calibration process. This emulation is highly dependent on how accurately the field conditions are represented by the embedded model parameters. Most of the traffic microsimulation software such as *VISSIM, CORSIM, AIMSUN* and *PARAMICS* have a large array of parameters that can be adjusted in the process of calibration to realistically reproduce road conditions and observe driving behavior. The calibration process usually starts with a default set of values as an input parameter and a search-based heuristic or meta heuristic algorithms deploying artificial intelligence to identify best fitting parameters. Different branches of genetic algorithms are widely used in research to find the near optimal simulation models parameters. The two primary categories that are often employed are Artificial Neural Networks (ANN) and genetic algorithms (GA). While GA performs well on traffic measurements with discrete values such as traffic counts, ANN usually performs well on continuous ones such as speed. [83–85].

2.4 Treatment Effects Analysis

According to the Strategic Highway Safety Plans (SHSP) [7], the determination of highway safety improvement requires any component after implementation to be evaluated. The results of these evaluations would enable agencies to develop countermeasures that will reduce the rate and/or severity of collisions. In this section, the research studies on the evaluation methods of traffic signs are presented.

2.4.1 Traffic Signs Evaluation

According to the Manual on Uniform Traffic Control Devices (MUTCD) one of the main conditions traffic signs must meet is to command drivers' attention. The first traffic signs have been introduced in 1895. Over the past 125 years the sheeting material has been evolved radically. Depending on their use and color, new sheeting, pigments, and film being manufactured, and new retroreflective prisms being employed to ensure road signs reflect an appropriate amount of light on an appropriate angle. In 1988 3M company introduced diamond grade reflective sheeting with micro-cube corners which reflect a significantly higher percentage of light than previous reflective sheeting materials. This was followed by fluorescent yellow and green signs, and, in the past few decades, the regulatory-warning signs are equipped with overhead or embedded lightings (also known as active signs). The Traffic Association of Canada (TAC) characterise Active (illuminated) Signs as a type of road sign which is do not rely on vehicle headlights for illumination [86] . These types of signs are being visible from a greater distance and can emit light so they can be seen from further, typically over a kilometer. The Institute of Transportation Engineers (ITE) has provided guidelines and minimum performance and photometric requirements [22,87]. The flashing frequency recommended by FHWA is that all LED units shall flash simultaneously at a rate between 50 and 60 pulses per minute (~1hz). [6]

The illuminated signs should be considered wherever reflectorized signs are not being effective. For example, where background light sources or other uncontrollable distractions reduces visibility, at decision points on high speed/high volume facilities or where vehicle headlights may not adequately illuminate the signs. Use of prismatic lens retro-reflective sheeting shall not be considered a substitute for sign illumination, especially in urban areas[86]. Some studies investigated the effectiveness of these kind of materials and compared them under different road conditions. In this section we investigate different methods in evaluation of such treatments.

Several empirical studies indicated the effects of different materials on sign conspicuity. Zwahlen and Schnell -1997 studied the benefits of fluorescent versus non-fluorescent color signs [88]. The authors reviewed a range of literature on peripheral vision aspects including same experiment on visual performance of fluorescent retroreflective signs. The attempts were to assess the visual performance in terms of detection distances, colour recognition and legibility [89]. The authors also conducted an experiment with participation of 18 individuals screening the color signs on a field test. The test variables were target sizes, colors and peripheral angles and the participants were trained to respond by indicating if they detected the colors and recognized the sign on a road tour. As it was expected, the fluorescent colors were detected with a higher frequency at greater distances due to higher luminance contrast. The experiment result on scatterplot has been approximated with linear regression on the detection and recognition percentages. The author suggested regression models to obtain an estimate for the detection and recognition percentages for the three fluorescent colors: orange, yellow, and yellow-green. However, the test has been conducted in an isolated environment, with all participants being college students. The sample size didn’t represent all driving population and the participants were alerted and instructed to look for the signs. The experimenter was seated in the front passenger seat and hence not engaged with driving tasks and finally the regression model deals with generous error due to dispersion in observation results [88,89].

There are a few studies about measuring illumination intensity reflected from signs using electrical instruments such as handheld devices and mobile asset data collection. The Texas Transportation Institute (TTI) has introduced an advanced mobile asset collection method and conducted a study on the accuracy of the real-time reflectivity measurements [90]. However, measuring luminance in the field is not an easy effort and required quite sophisticated and expensive equipment. Even though federal highway administration recommendation is to use reflectivity measures to set the minimum visibility standards, however, the limitations are also outlined on the recommendation, such as the variation in geometry of the facility and other influencing factors mentioned earlier under sight distance challenges[90].

2.4.2 Enhanced Traffic Signs

Gates et al. -2014, has evaluated signs with enhanced conspicuity properties. The study focused on driver's behavior when they are exposed to the signs with the different material characteristics. Unlike previous efforts, this experiment was unbiased to peripheral vision aspects or illumination intensity of the signage. The signs performance was evaluated directly from driver's indirect behavior. The mean vehicular speed was collected and considered as the indicator of drivers' respond to the sign message. The authors used analysis of variance (ANOVA) to assess the differences between speed mean values, under different sign categories. In the model, multiple independent variables (i.e., sign treatments, light conditions) was presented, while the mean vehicle speed at the control points was the selected dependent variable. With only one independent variable, the study performed the one-way ANOVAs and the *t-tests* for 8 facilities operated with different treatments (e.g., fluorescent red stop-signs, flashing red LED stop-signs). Fluorescent, prismatic sign sheeting, embedded LED was concluded to have statistically significant impact on the mean speed. According to the result from the study, the embedded LED stop-sign can reduce 28.9% and 52.9% the number of vehicles not fully stopping, and the number of vehicles moving through the intersection without slowing down (blow-through traffic), respectively. This study was conducted in a rural area and no other independent variable was considered. Only one type of treatment was displayed at any given time to the test participants. While traffic, weather condition, geometry and driving population might be different at each intersection. Test on statistical significance on the report indicates some of the findings are not reliable enough and some misinterpretation on Nominal Logistic Regression was on the report [8].

Safety performance of flashing warning sign on human factors at rural intersections has been evaluated by Minnesota Department of Transportation[91]. Rural, two-way stop-controlled intersections with three types of improved treatments have been investigated. This study included before-after treatment installation analysis both on collision history and human factors. A survey from drivers was supplementing the data of the study. The result from this survey presented that, drivers tend to not believe in positive impact of the treatment on their speed since they indicated that they “reduce speeds equally” with or without the flashing lights. Three years’ worth of study collision data consisting of 12 rural, two-way SCI, was collected for the before and after analysis with 40% improvement in favor of after treatment being upgraded to flashing light. Drivers’ speeds were measured by pavement-based magnetic sensors on both major and minor approaches. Despite the rather dramatic decrease in collision experience, the authors could not provide unequivocal evidence for safety benefits of overhead flashers from survey data and the observed speeds. The field study result and statistical collision test were inconsistent. However, the authors indicated sensor failures and complexity of the test as limitation. Control for possible regression-to-mean effects wasn’t included in the report as well as statistical significance of effects were not testified.The effect of influencing factors which possibly changes over time; traffic volume, weather, road user behavior, vehicle fleet, inclination in report collisions and other potential factors was ignored.

As for collision reduction impact, a more recent study has investigated overhead flashing beacons efficiency at SCI in North Carolina [92]. This study focused on rural intersections and the treatment was an overhead beacon with flashing yellow on major approaches and flashing red beacon on stop-controlled approaches. Thirty-four SCI in North Carolina on two-lane roads, with no turn lanes, and two-way stop control have been used in the study. The authors had access to at least three years of post-installation collision data, as well as average daily traffic (ADT) for the selected SCI. The authors deployed Empirical Bayes (EB) before-after method in order to control possible Regression-To-Mean

(RTM) effects. The authors have selected a calibration procedure to find adjusting factors for local collisions data for and safety performance function [93]. Hence for each treated intersection, five not-treated intersections were chosen to comprise the reference group. The authors reported a 12% (± 6%) average decrease in all collisions. Despite of this finding, the rare occurrence of "ran stop-sign" collisions in the police reports and the rather high cost of the tested treatment made agency to conclude that overall collision reduction effects from the improvement were modest. The right-angle collision was not included in the study since selected SCI had no turning lane. The study didn't investigate drivers' behavior and it failed to justify if the sample size was enough to reflect the change in collision modification factor.

Another report published by Federal Highway Administration on safety evaluation of flashing Beacons at SCI has included both rural and urban intersections [52]. A wide range of intersections including multi-way SCI, on either two-lane or four-lane approaches, have been investigated. Not only overhead flashing beacon but also pole-mounted flashing red beacons above stop-signs was in the study. A fairly large number of SCI equipped with this treatment has been identified in North and South Carolina (100 locations). The study has been designed with generic before-after safety analysis using the EB method. Not only collisions frequency (all, rear-end, and at-angle) but also collision severity (fatal 'K' and injury 'A-C') has been investigated in this study. The author followed the method introduced by Hauer [39] to estimate the minimum required sample size. Based on minimum and optimum sample sizes, several hundred additional SCI were selected in the reference group in aggregated and disaggregated analyses. The results considering all intersections, revealed 13.3% reduction in at-angle collisions, and a 10.2% reduction in injury collisions with 4.6% and 4.9% standard error respectively. Sites with standard pole-mounted beacon stop-sign (58.2% ±32.6%) seemed to show more reduction rather than overhead beacons (11.9% ±10.8%). The agency cautiously (due to the limitation in sample size) recommended the active signage of this kind, for the rural areas.

Some studies were centered around the drivers' behavior to evaluate safety performance of the active stop-sign. In one study conducted by Arnold and Lantz, a T-SCI with a stop-sign on the minor approach was selected. This intersection had 14 reported collisions in only two years. Half of the collisions involved injuries. The drivers' interaction with a stop-sign equipped with flashing LED at the corners has been investigated. The observation started from a week before and continued after installation of the sign. Three data collection stations (i.e., traffic counters) were used on the approach to collect approaching speed in three corresponding sessions. Data collection included three seven-days sessions: session 1 - before installation, session 2 - within a week after installation and session 3 - after a full season of the deployed treatment. For driver compliance, a site survey was performed at 15-minute intervals during the morning, lunch, afternoon, and evening peak periods. For compliance, data comprised under four categories with the number of motorists who came to a full stop (voluntarily or because of conflicting traffic), rolled through (0 to 3 mph), or blow through the stop-sign (over 3 mph). Overall, a decrease in average vehicle speed was reported in the range of 1.3 mph to 3.4 mph. The study has recorded a slight increase between the two last sessions on the average speed between stations 2 and 3 (further from sign). The speed at the closest station (station one) remains unchanged on sessions two and three. Another finding was the positive impact of the flashing stop-sign on the speed reduction at nighttime - from 2% for daytime to more than twice at nighttime (4.2%-7.3%). The intersection under the investigation was selected due to high collision frequency with ADT of $7{,}100$ veh/h however, the correlation between speed reduction, volume and collision frequency was not reported. There was no mentioning of enough warm-up period after installation for session two. Calibration of the test equipment has not been performed after first session, statistical significance was not in the report and biased observers' verdict might impacted the result [94].

Another effort deployed a unique non-standard animated LED (scanning eyes) on top of a regular stop-sign rather than active sign to attract drivers' attention. The study was on a hotspot (4 collision in 3 years) three-way, bidirectional SCI in St. Petersburg in FL. One hundred and twenty-six sessions were recorded and the result of the visual analysis shown the number of right-angle conflicts decreased from 4 to 1.4 per session [95].

Kwon et al. -2013, have suggested an advanced LED warning System. The proposed system actively detects vehicles presence on all approaches (specially with sight restrictions) and triggers LED blinker warning signs for the conflicting movements. Two phases of video data collection have been performed before and after installation. The drivers' compliance to the traffic control devices was measured using video data analyzer. Speed mean and standard deviations was calculated. The result of this study showed the average vehicle speed after installation of the treatment on the main approach was slightly increased during the daytime and on the contrary significantly decrease (4.2 mph) at nighttime. However, the occurrence of the roll-through maneuver increased from 13% to 24% after the installation of the warning system. Statistical significance was not included in the report [96]. After few iterations, the authors replicate the surrogate traffic safety on the advanced LED warning system for rural intersections (ALERT). The before-after analysis revalidated the previous conclusion while the mean speeds after installation of ALERT for conflict and non-conflict periods were 47.9 mph and 51.8 mph, respectively. The authors also selected wait time for vehicles on the minor road to be used as the surrogate measure for drivers' recognition of acceptable gap. The average wait time was 3.9 seconds during the conflict periods and 2.5 seconds during the non-conflict periods. The roll-through maneuver increased again after installation of ALERT [97]. The location was selected in a rural, low volume link and hence the treatment may not perform the same in a more active urban setting with higher light pollution. The solution is also complex to be widely deployed and has high upfront and maintenance cost. Despite of the complexities, The National Committee on Uniform Traffic Control Devices (NCUTCD) have proposed Intersection Conflict Warning Systems (ICWS) and recommended to be included in the Manual [98].

Monsere et al. -2015 and Jomaa et al. 2005, have evaluated two other applications of the actuated LED warning sign on speed alert [99] and sharp curve with chevron [100] . Both studies were being able to demonstrate an improvement in the mean speed however the impact on the safety performance measure remained uncertain.

Some other studies strived to estimate the collision reduction for SCI with active signs. Davis et al, conducted a two-step comprehensive study on active stop-sign at SCI with respect to collision frequency, conflict patterns as well as drivers' behavior. For the first step, a statistical procedure, based on Empirical Bayes (EB) method, was defined, and accounted for RTM bias using the reference sites. Traffic volume and other necessary data (e.g., geometry) were collected at different sites. Three-year worth of reported collision data for 255 SCI in Minnesota was included in the study. Fifteen SCI locations had the treatment installed, while the rest of the sites was grouped as the reference. The authors suggested an enhanced FB method - previously developed [33], to estimate the target collision at each SCI with an assumption that the treatment has not been installed. This method considers other covariates to ameliorate the estimation of conventional Full Bayes. In comparison to conventional EB method outlined in the HSM, the collision estimate was consistent with the result from the proposed method. The estimation on average right-angle collisions at SCI with active signs was 41.5% less than the estimation with assumption that the treatment has not been installed - between 0% and 70.8%, with 95% confidence. The result of this study reported a CMF of 0.585 for the countermeasure (Replace standard stop-sign with flashing LED stop-sign). For the second step, an empirical study was designed to evaluate stopping compliance, estimating speed of approaching vehicle to active sign, and finally, comparing the braking deceleration. Video data was collected and processed to estimate deceleration and speed. Compliance to sign was manually captured. Multinomial Logistic Regression was used to summarize how the degree of compliance varies with respect to a condition. For instance, the treatment increased the odds for "clear stops" vs "clear non-stops" 6.4 times, whenever minor approach drivers encounter opposing traffic (ODDs with and without treatment was respectively 4.2 and 10.6). However, the result was not disaggregated enough to present nighttime or road condition impacts.

The effects of artificial ambient light on driving behaviour at 490 SCI in rural areas were explored in the state of Iowa. The results were compared to the same number of SCI without the "destination light". At treated sites, there was a 50% drop in the ratio of night-to-day collisions. According to Goswamy et al. 2018, the SCI with proper lighting can reduce property damage, injury, and overall collisions by 18 percent, 24 percent, and 33 percent, respectively, at night. [12]

Some other research examined the safety benefits of a sensor actuated, embedded LED lighting system at pedestrian crossing pavement markings. The authors have limited the focused only on the speed and deceleration rates with and without the presence of pedestrians. Both speed and deceleration rate were significantly reduced. The number of measurements was limited (200 total sample). Although no statistical test results were reported, a 19.3% and 16.4% drop in mean speed at zebra crossings was observed when the sponsor module was enabled or disabled, respectively. According to the authors this reduction "corresponds to a lower fatal accident and serious injuries risk" *Patela-2020* which had lack of evidence, [101]

The influence of the other factors in performance of the active signage was studied in a recent study by Rista et al. 2020. In a sampling of 13 sites, this study discovered a wide variety of driver yield rates, ranging from 5% to 88 percent. According to the authors, intersection configuration, crossing distance, traffic count, posted speed limit, location, sign face, approach side (near or distant), and other factors all may play a role in driver yielding. To determine the relationship between the dependent variable (Yield/not Yield) and the independent variable, the author utilised a logistic regression model. Hourly volumes, operational speed (as a binary), lane width deviation, and sidewalk presence were found to be predictors for log-odds of driver yielding probability. The authors adopted "a staged pedestrian approach" which means that the study was obliged to increase a pedestrian traffic count. this approach will not be acceptable in an objective safety assessment. The independent variables chosen were all connected to the driving environment, while the dependent variable (Yield/not Yield) was the only driver behaviour variable in the study. The research was carried out in Texas, USA, and focused on active pedestrian crossing signs in rural area. There was no effort put into improving safety performance or analysing conflicts. [102]

Another recent study conducted by TTI in 2021 strive to identify statistical interaction between treatment type and site characteristic variables with a larger sample size and staged pedestrian using ANCOVA models for site mean yield rates and logistic regression that considered the individual driver response to crossing pedestrian. This study has considered Pedestrian hybrid beacon, rectangular rapid flashing beacon and Light emitting diode signs. This study found higher yielding being present at night. The presence of yield lines as, speed limit group, number of lanes, lane with and hourly volume were also among variables impacting driver reaction. Like previous study, the result was on rural with tempered traffic count and not looking into the safety performance metrics. Using binomial logistic regression was limiting the study with more than two categorized which makes it difficult for continues variables like speed.[103]

Another recent study utilized driving simulator to evaluate Stop sign configurations on driving Speed on simulation videos in rural Intersection using 32 rather young volunteers. The comparison was between regular Stop-sign, Stop-sign with advance notice and fluorescent stop-sign and drivers right of way. Driver speed was compared using ANOVA and conclusion was that the fluorescent sign reduced the average speed for 12.6km/h and 9.2km/h type 1>3 and 2>3 respectively. [104]

A recent study in the city of Montreal, evaluated the performance of SCI compared to uncontrolled approaches. Video collection from 100 intersections and 130,000 vehicles was used for to determine the within-site and within-approach correlations deploying linear mixed-effects regression models. The authors concluded that presence of stop-signs significantly reduced speed approach and variance, however there was not enough statistical evidence supporting the vehicle–pedestrian interactions. The authors recommended "implementing stop-signs to reduce pedestrian crashes may be less effective than other interventions" -*Stipancic et al. 2021*. surrogate measures such as TTC and PET was extracted through an automated process of image processing. Several covariates were identified in the study including control device configuration, geometric variables, time of day, etc. Three regression models were developed for speed, TTC and PET. TTC and PET was improved somehow for vehicle-vehicle however vehicle–pedestrian models had very few significant variables limiting the authors to make a conclusion for pedestrians' safety. This study did not consider the stop-sign types (active signs vs regular) however the contribution on PET and TTC analysis and meaningful conclusion on general assumption of performance of stop-sign to reduce risk for pedestrians was significant.[105]

2.5 Chapter Summary

This chapter presented the review of available literature related to the proposed research topic. For decades, road safety analysis and modeling has been practiced by transportation specialists. Several assessment methods, tools and models were discussed that tend to evaluate quantitative and quantitative road safety measures. The collisions record is the primary safety measure that gives a sense for future safety performance of a facility. In order to evaluate the impact of a countermeasure or treatment collision data is necessary for before and after period. However, Data availability and quality of available data is a challenge. Due to the rare occurrence of traffic collisions, regression to the mean is essential in collision analysis. Several type of regression models are being used by traffic safety analyst including

Poisson and extensions of NB. Different branches of Empirical Bayes methods are a commonly accepted and explored for treatment before and after safety analysis. Other methods, such as Full-Bayes, were shown to improve collision estimate under certain conditions, as described in this chapter.

The inability to resolve data quality concerns for unrecorded and poorly recorded collisions is frequently mentioned in various studies. The surrogate safety assessment, and conflict analysis, are used as alternative solutions to fill-up this gap, at the cost of not providing direct collision estimation. As presented in this chapter, there is strong correlation between conflicts and collisions. Transportation specialists have deployed manual and automated methods to detect and observe conflicts. Some of these methods was explored in this chapter. Computer-based modelling tools and data processing techniques have widened the window of opportunity for traffic engineers to emulate infrastructure modifications prior to deployment. However, the quality of a developed traffic model determines the success of safety impact assessment. Microsimulation model calibration and tools to extract collision observations are essential for computer-based, road safety conflict analysis.

To evaluate safety effectiveness of a treatment, both changes in collision record (CMF) and surrogate safety assessment being suggested in literature. Before and after analysis of the collision history (with or without using a comparison group) have been frequently used to estimate safety effectiveness of a treatment or countermeasure. However, the same challenges for data availability remains. Deployment of a new untested treatment might have negative consequences and exposes travellers to risks which are difficult to quantify. Stop controlled intersections are vulnerable facilities when inadequate design is used, or when visibility is hindered by obstacles. Improving signage visibility is one of the first amelioration solutions sought by transportation agencies. Traffic signs have evolved to improve drivers' interaction in a dynamic driving environment. Regulatory and warning signs, (pedestrian crossing and stop-signs), have been cites in the literature as a risk element. New materials and design mechanisms are utilized to increase visibility of road signage. The enhancements have been evaluated in some non-collision and collision-based studies covered in this section. However, the results were either equivocal or not in agreement from report to report on stopping compliance or approaching speed. In the next chapters a case study for a newly introduced treatment will be presented and the safety impact using conflict-based models for 3-ways SCI will be assessed.

3 Chapter 3: Methodology

The literature review was included in the chapter 2 of this thesis and the inadequacies of the reported findings were discussed. This chapter describes the methodology and clarifies the steps and sequence of tasks to be carried out in order to make a safety assessment of new, or existing treatments with no (or insufficient) safety performance history. After a brief introduction, this chapter will present the design of an empirical study including selection of a testbed, design of an experiment and data collection. Following that, the micro-simulation modeling and calibration will be detailed, with underlining the data validation process from simulation representing observation. Then, the next section of the chapter entails the approaches in statistical modeling from base to top of the Hydén pyramid, from driver reaction to collision. The final segment will summarize the proposed methodology and introduces the next steps of the research.

3.1 Introduction

Previous studies on active signage assessment were performed either under a long term before-after analysis of a flashing sign or as a cross reference between SCI equipped and not-equipped with the active signage [8,9,14,33,102,106]. Drivers' reactions to signage could be affected by a combination of factors that aren't necessarily related to the type of signage such as: intersection configuration, weather condition, neighbourhood type, crossing distance, traffic counts, traffic mix, posted speed limits, location, and other factors. Some of these factors may change over time or from one site to another.

In this thesis the author strived to control some of these variables and evaluate the impact of the remaining ones. The site selection criteria were to find several SCIs with:[1] similar geometry design, [2] same traffic mix and count [3] to make concurrent experiments for observation. A link in urban area in the city of Montreal was identified and an experiment was design in a fashion that same traffic was exposed to several types of active signs within few minutes' intervals. Two signs in the experiment were of BLS. During the "before" analysis, some intersections were standard stop-signs where two intersections were converted to BLS for the purpose of this study. Data collection was also conducted before and after installation of the BLS to complement the cross-reference analysis. Data collection included speed/count-read, using 3 stations of high-definition (HD) radars and 189 hours of video recording within 10 days. With deploying several image processing techniques, volume, speed, trajectories, compliance, and the reaction to the active signs was captured[107].

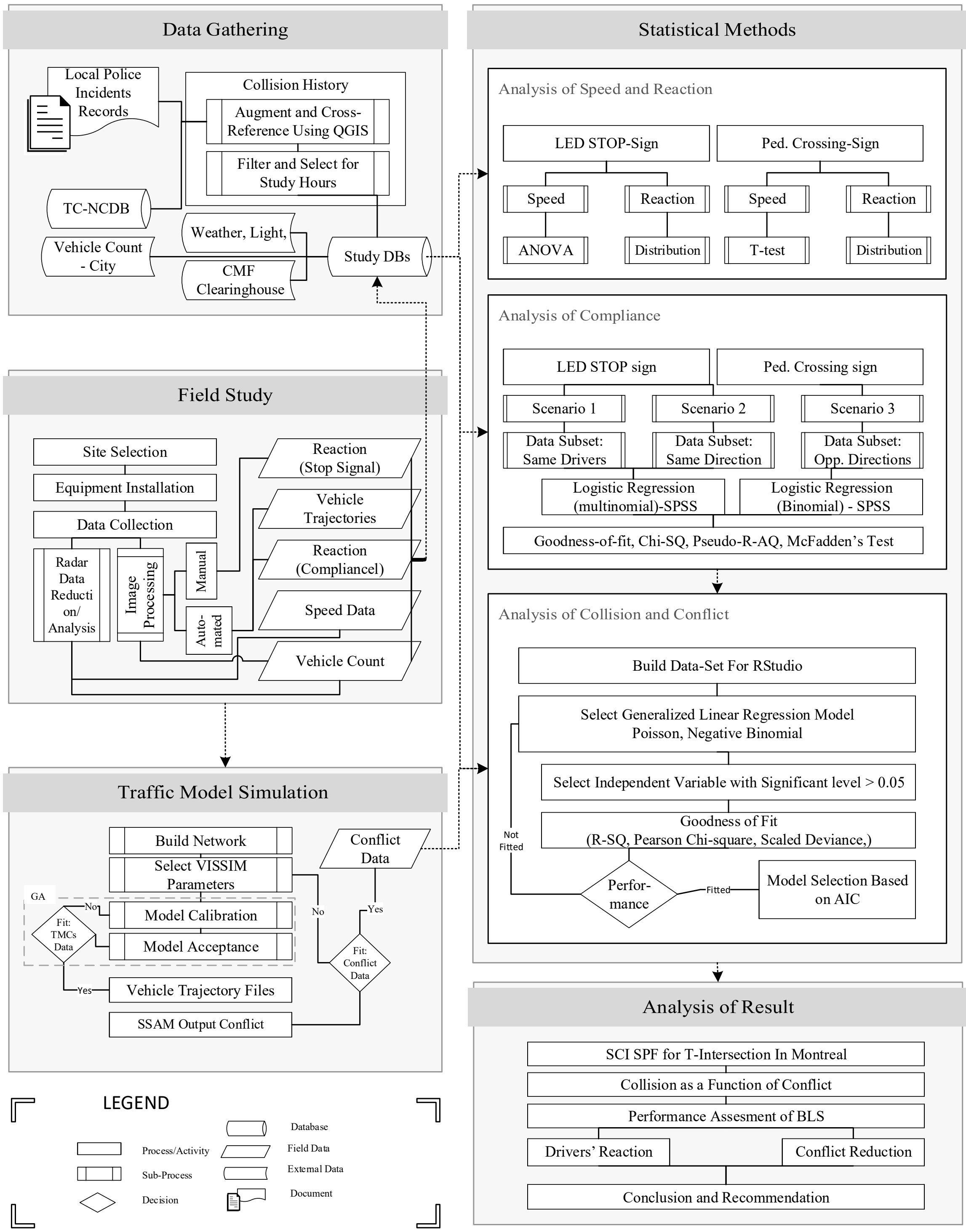


Figure 6 Detailed Framework of the Study

Supplementary data was gathered from several sources including local and national collision data bases, weather condition and rainfall, previous vehicle counts, and intersection drawings provided by the city. Several databases were created to process the data collected. The next three section in this chapter will walk through the three analysis stages shown in Figure 6, which are “Field Study”, “Traffic Model Simulation” and “Statistical Method”. The analysis of the results is included in the subsequent chapters.

3.2 Empirical Study

In order to re-evaluate surrogate safety measures of a newly installed traffic control device the following objectives have been set out:

- To classify the drivers’ compliance to the traffic control device.
- To observe performance of each treatment under low light ambiance(daytime/twilight).
- To observe interactions for both motorized, and mixed motorized with non-motorized traffic.
- To record the approaching vehicles’ speed and traffic volume.
- To estimate vehicles average deceleration magnitude.
- To capture independent variables such as pavement, weather, and light condition.
- To observe the speed variance.

Two regulatory signs, Pedestrian crossing and stop-signs is investigated and several SCI for the test bed is identified for this study. As mentioned in the introductory chapter, about 8% of fatalities and 11% of total collisions occur at SCI.

3.2.1 Experiment Site

Several potential sites in the greater Montreal area were identified as possible study locations and being filtered using the following criteria.

1. Intersections with similar structural and operational characteristics with respect to geometry, traffic flow, driver population mix and various transportation modes.
2. SCI equipped with enhanced treatment (overhead beacon and post-mounted LED stop-sign)
3. Ability to safely attach, mount and install data collection equipment.
4. Accessible location to visit in an urban setting (i.e., within Montreal metropolitan community area).

5. Availability of at least 5 years of collision history for safety performance measures.

A section under jurisdiction of Lachine borough (in the Greater Montréal Area) was found to meet the criteria listed above. The east-west arterial of Victoria Avenue includes four intersections along two-lane bidirectional road. The five SCI between avenue 18th to the 33rd was identified for the empirical analysis and an extension of four more intersection for the simulation and/or statistical analysis. The study area concerns the section of Victoria Street (east/west axis) from 15th to 34th Avenue (north/south axis). Victoria Street itself is an extension of another major street; Notre-Dame Street.

The intersections under study are:

1. rue Victoria/15e Avenue (intersection with stop in all directions),
2. rue Victoria/16e Avenue (intersection with stop in all directions),
3. rue Victoria/18e Avenue (intersection with stop in all directions),
4. rue Victoria/19e Avenue (non-stop, one-way intersection on secondary),
5. rue Victoria/21e Avenue (intersection with stop in all directions and one-way on secondary),
6. rue Victoria/25e Avenue (intersection with stop in all directions),
7. rue Victoria/28e Avenue (intersection with stop in all directions),
8. rue Victoria/32e Avenue (crossroads with traffic lights),
9. rue Victoria/33e Avenue (intersection with stop in all directions and flashing lights on rue Victoria, one-way on secondary),
10. rue Victoria/34e Avenue (intersection with stop on secondary).

It should be noted that 32nd Avenue is a main artery according to the hierarchy of the road network of the city of Montreal. Hence intersection “8” was not stop controlled, the data was only used for microsimulation calibration and not the statistical analysis. The traffic light signal timing has a protected phase for each of the pedestrian crossings (all red phase) to allow safe pedestrian crossing. Pedestrian pushbuttons are thus positioned at each corner of the intersection to allow users to call the protected phase.

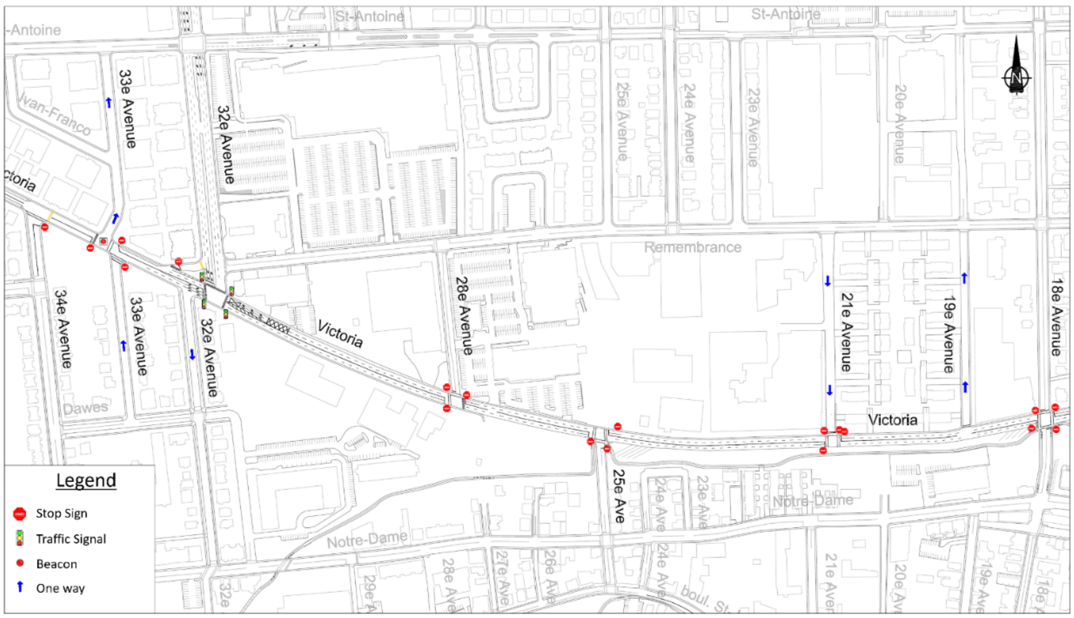


Figure 7 Study area and intersection control devices

The intersection has the pedestrian traffic lights as well. Intersections “1” and “2” were also excluded from the simulation and conflict analysis, however the collision and volume data was used to enhance the quality of the collision prediction model for the link.

Intersections before intersection “1” were ignored since the geometry and number of lanes were not consistent with the rest of the road segment (rue Victoria/1e Avenue: rue Victoria/8e Avenue -3 lanes, rue Victoria/8e Avenue: rue Victoria/14e Avenue -2 lanes). On street parking was not permitted on the study section. There are minor roads access and egress traffic along the segment, with a posted speed limit of 40 km/h and a negligible longitudinal grade.

Two intersections downstream from 18th avenue (25th and 28th) have same type of control devices. Both westbound approaches are provided with Active stop-signs.

Figure 8 Treatment types

There is a small commercial plaza with two access driveways, between avenue 25th and 28th (north side of Victoria Avenue). The subsequent intersection is a three-leg signalized intersection with a fixed cycle time. Avenue 33rd is a dog-leg intersection with a 30-m offset between northbound and southbound approaches. The eastbound approach passes southbound before stop-line (a one-way, left turn with local access only). The intersection has an overhead beacon supplementing standard pedestal-mounted retro-reflective sheeting stop-sign. All signs meet regulation for stop (ARRÉT) P-010 described under section 2B-05/4L.05 of MUTCD. Figure 9 presents the visual properties of two BLS signs and the regular LED sign and shows BLS sign is more legible specially in dark and from distance. All intersections pass the stop sight distance test described in section 2.2 with no obstacle preventing the drivers to visually see the signage from the recommended distance.

Figure 9 Visual comparison between BLS and LED (night and day)

3.2.2 Experiment Design

A hybrid data collection method was used in this study to measure per-vehicle speed, by combining information processed from digital video cameras and high-definition radars. The data comprises per drivers' speed, correlative speed, headway, encroachment time, trajectory, maneuver, and traffic volume.

A pole mounting unit attached the radar and camera(s) to a telescopic mast (Figure 10). The radar projection was perpendicular to the street centerline and parallel to the stop-line, while the video camera was tilted at a suitable vantage point to cover the conflict zone and capture vehicle trajectories.

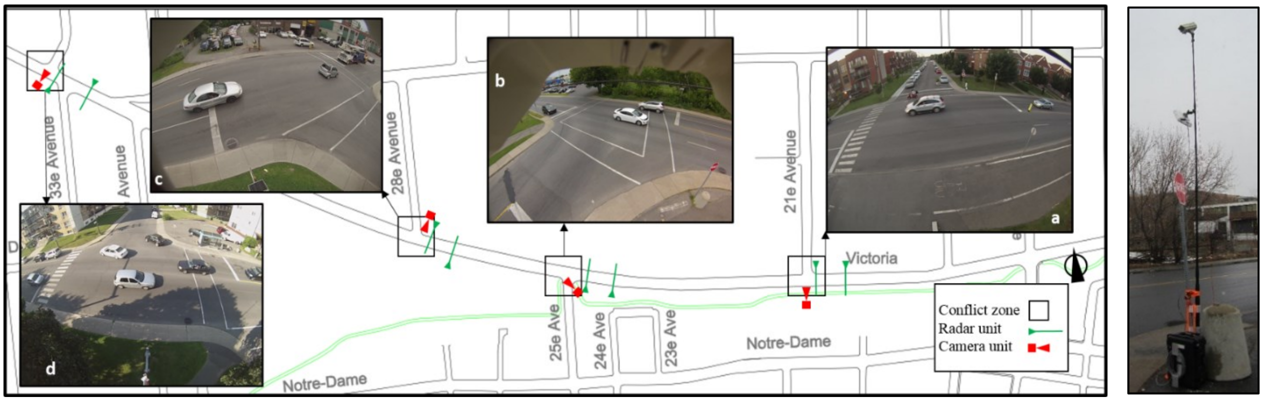


Figure 10 Left-Camera positions and snapshot. Right- hybrid data collection station

The research team used three mobile data collection points that were set up in advance to record morning peak and/or evening peaks. The equipment was powered via external batteries to complete up to 6hrs of data collection.

		2014					2015					
		December		**April**		**Installation and Warm-up**	**July**					
		Before BLS installation					After BLS installation and warm-up					
		Tue	Wed	Mon	Wed		Mon	Wed	Fri	Tue	Wed	Thu
Minor	**Location** / **Date**	2	3	6	8		6	8	10	14	15	16
18th	Victoria& Ave. 18			5	5			5				
	30m WB Ave. 18			5				5				
	30m EB Ave. 18				5			5				
21st	Victoria& Ave. 21	7					7		6	7		
	30m EB Ave. 21										5	
25th	Victoria& Ave. 25	7	7	5			5		5		7	
	30m EB Ave. 25			5			5					5
28th	Victoria& Ave. 28	7	7		5				7		7	
	30m EB Ave. 28				5							
32nd	Victoria& Ave. 32									7		
33rd	Victoria& Ave. 33		7							7		7
	30m EB Ave. 33											5

Table 2 Recording hours

The data collection points were either installed at a sub-set of the intersections under the study link or thirty meters upstream to the stop-line on WB or EB approaches. Data collection was conducted during month of December of 2014 and two months of April, July of 2015. Table 2 summarizes the sessions.

In total 32 sessions, 189 hours of usable recording was done in 10 days. The radar was set up at 30 meters upstream of the intersections 18th, 21st, 25th,28th and 33rd to analyze the variance in speed for all types of treatments. The sessions and data are presented in Appendix I. The summary of the other relevant environmental data is shown in the table below.

Date	Sundown		Weather			
	Sunset	Civil	7:30 8:30	12:00 13:00	16:30 17:30	18:00 22:00
2014, Tue, Dec 02	16:13	16:46	Clear	Mostly Cloudy	Cloudy	Snow Showers
2014, Wed, Dec 03	16:12	16:46	Cloudy	Snow	Snow	Cloudy
2015, Mon, Apr 06	19:28	19:59	Cloudy	Snow	Snow	Snow, Fog
2015, Wed, Apr 08	19:30	20:01	Cloudy	Cloudy	Cloudy	Snow
2015, Mon, Jul 06	20:45	21:22	Cloudy	Clear	Clear	Mainly Clear
2015, Wed, Jul 08	20:44	21:21	Cloudy	Mostly Cloudy	Mainly Clear	Cloudy
2015, Fri, Jul 10	20:43	21:20	Mainly Clear	Mainly Clear	Mainly Clear	Mainly Clear
2015, Tue, Jul 14	20:40	21:17	Mostly Cloudy	Cloudy	Cloudy	Cloudy
2015, Wed, Jul 15	20:40	21:16	Mainly Clear	Mainly Clear	Mainly Clear	Mainly Clear
2015, Thu, Jul 16	20:39	21:15	Mainly Clear	Mainly Clear	Clear	Mainly Clear

Table 3 Light and weather characteristics

3.2.2.1 Video footage from 30m upstream:

The data collection points at 30 metres upstream was used to capture drivers' compliance and first reaction. Video footage with perpendicular angle with the median line of the link was captured (see Figure 11). In transition from three-dimensional workspace into two-dimensional image, calibration and verification was essential. Detailed description of the manual video analysis is presented in previous work[13,108]. A measuring line was marked and annotated on the pictures. These grid lines were then converted into a transparent layer and overlaid into the video using the available software. After this step, the video was run at 60fps rate (monitoring resolution for every 0.017 seconds) and whenever a driver first pressed brake (based on brake light), the vehicle location was marked as driver reaction point. (The front bumper line was set as the reaction line). The extracted data from the footage was displaying drivers' reaction distribution.

Figure 11 Image annotation and labeling at 30m point - (Victoria and 18th avenue)

3.2.2.2 Video footage from the intersections:

At the stop-line the speed detected by the radar and footage was correlated with compliance to the control device. The drivers' compliance to the signage is correlated to speed of the driver at the stop line by ITE in the manual of transportation engineering studies. Three categories are defined; full stop ($0-1\ km/h$), rolled through ($1-\ 5\pm 1\ km/h$) and blew through the stop-sign (more than $5\ km/h$) [109]. The high-definition radar used in this study has 95% accuracy in speed detection with $\pm 1\ km/h$. For video-based data collection, the number of frames is correlated with compliance. For instance, for a vehicle with a length l and speed s, the total number of frames in which the object crosses a virtual line (Figure 12) can be calculated as follow:

$$\#\text{frames} = \frac{l(m)}{s\left(\frac{m}{s}\right)} resolution\ rate\ (fps) \qquad \textit{Eq. 17}$$

For a passenger car with an average length of 4.5 m, a blow-through occurs when the total number of frames in which the vehicle occupies the pixels designated by the virtual lines is less than 95 frames (i.e. 3 seconds).

Figure 12 Left, image annotation and labeling at intersection - (westbound Victoria and 28th avenue), right, examples of video recording presenting road users' tracks, identification number, and speed overlaid (Victoria and 21st avenue)

If the front bumper covers two adjacent lines in 210 frames or more, it can be interpreted as a full stop. Subsequently, between these two thresholds falls under roll through. A manual visual cross-reference of the video recordings was undertaken in addition to image processing to group the number of motorists who came to a full stop willingly and those who came to a full stop due to conflicting traffic. The method for feature tracking analysis was recommended in previous studies [13,108,110]. A video analysis open-source tool was used to generate trajectories for all moving objects in the video frames (Figure 12) track the vehicle's trajectories[111]. The trajectory files were then imported into SSAM. This data was only used for validation of the model. This process is described under section 3.3.2.5 where the actual conflict data is extracted from the microsimulation model.

3.2.2.3 *Introducing the new treatment:*

For the purpose of this thesis, two new treatments were placed. The new treatments are a BLS standard pedestal-mounted stop (ARRÊT) and a BLS pedestrian crossing sign with 1Hz flashing backlit LED. The LED is embedded behind the sign-face.

The sign follows the Quebec P-010 and MUTCD sign description under Section 2A.07. The prototype of a solar powered control device was built in house for the purpose of this thesis by the industrial partner, Orange Traffic (Figure 13). The installation was done by and with the permission of the city to replace the standard stop-sign on the westbound approach of the major road at Victoria and 25th avenue.

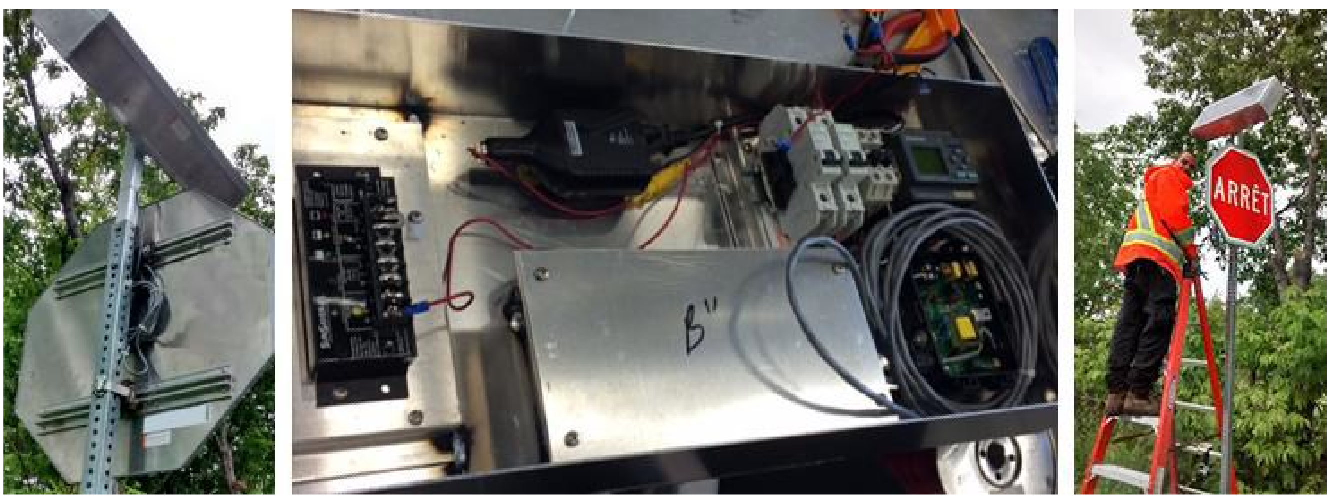


Figure 13 Prototype of the BLS (left), control box (middle), switch from battery to solar panel(right)

The BLS sheeting includes micro-prisms ranked III and IV under D4956-04 FHWA classifications with 50 cd/lx/m2 luminous intensity. This makes retro-reflectivity same as the replaced static sign. The thickness of the panel is 12mm. The sign radiates uniform light which makes them legible from a greater distance compared to standard LED. The LEDs in standard sign are placed at an equal distance from the center of the sign which forms a flashing circle shape at night. The four legs intersection of Victoria Avenue and 18th Avenue was equipped with four pedestal-mounted retro-reflective sheeting stop-sign, regulatory sign (ARRÉT) P-010. The east and west approach had an additional pedestrian-crossing sign (P-270-2). The eastbound approach was selected to have the new treatment installed while the westbound approach remained the same. The sign was mounted under the stop-sign like before.

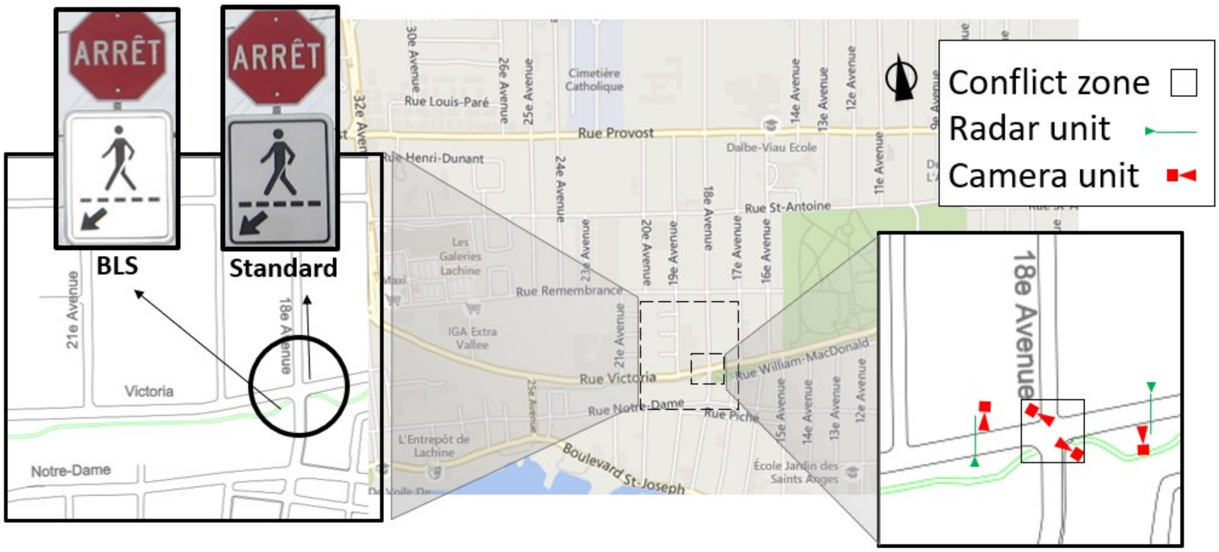


Figure 14 Site specification and data collection set-up

On May 29th, 2015, the BLS signs were setup, with one month of warm-up time. The set-up was reversed after the data collection data in late July. The traffic was exposed to the sign for less than 3 months.

3.3 Model Simulation and Conflict Analysis.

Existing microsimulation software, as described in section 02.3.2, are powerful tool that allows transportation specialists to simulate traffic behaviour, extract vehicle trajectories, and even pollutants. To the best of the author's knowledge, they do not, however, permit a direct assessment of road safety since aggressive behavior causing collision is inherently prohibited. The microsimulation approach has hitherto suffered from a lack of direct measures for evaluating safety using collisions.

Conflict analysis, on the other hand, has the significant advantage of being an alternative approach in safety analysis, and even more so when collision data is limited. This limitation imposed by the random variation in the number of collisions, as described in section 2.3.1 and forces researchers to consider alternate approaches, such as the use of micro-simulation and surrogate safety measures[83]. As previously determined, conflicts at unsignalized intersections are correlating better with collision than with traffic flow, hence they can be utilized in collision modelling [66,81]. In this section the methodology for building a network model to obtain calibrated conflict counts at intersections by using simulation and SSAM are presented, and a detailed procedure of model calibration is presented in Figure 15.

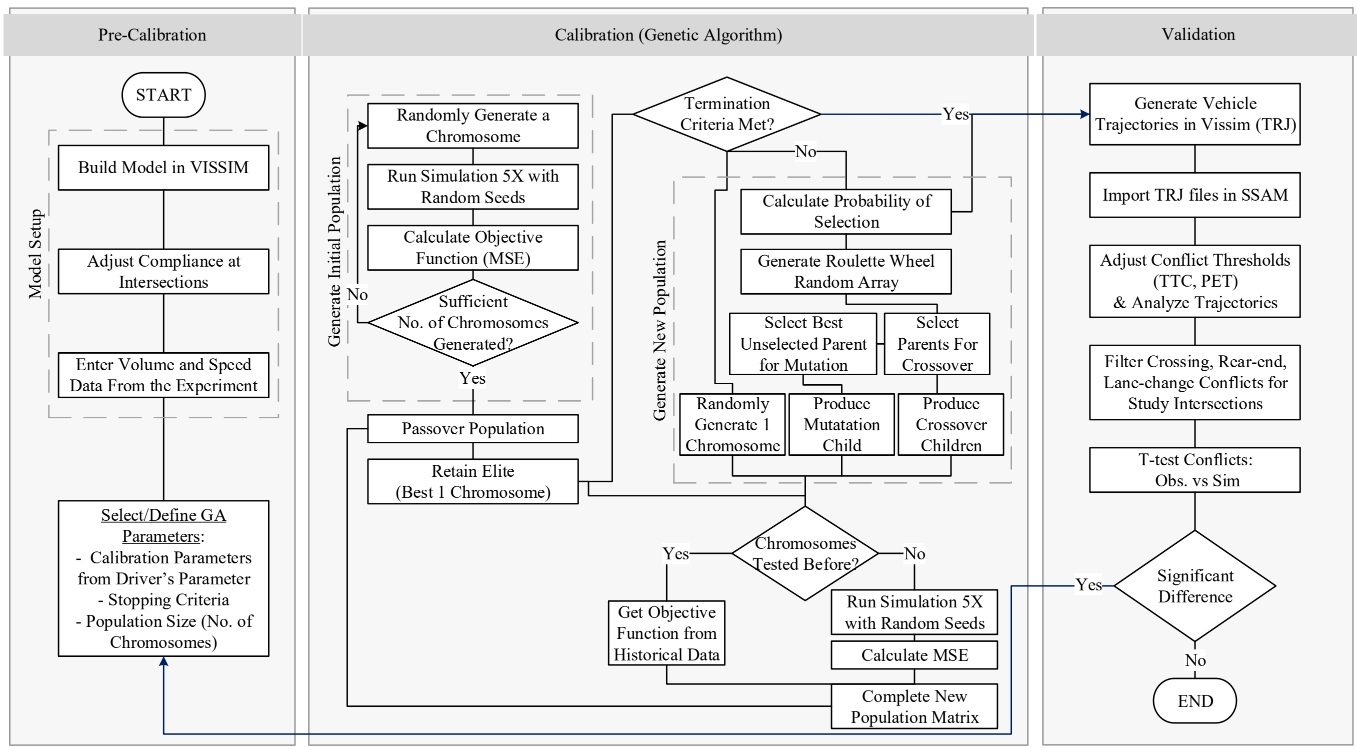


Figure 15 Microsimulation set-up procedure utilizing GA

Gettman et al. assessed various traffic simulation software compatibility with SSAM [21]. In the comparison study, VISSIM successfully presented comparable values of the "total number of conflicts". PTV VISSIM 2021 was used to conduct the analysis by keeping some of the default values of driving behavior to reduce the risk of generating unrealistic maneuvers in simulation. Gettman et al. also suggests eliminating TTC and PET results equal to zero values in the model to avoid inaccuracies generated by model.

3.3.1 Building the Simulation Model

City of Montreal maintains the as-built library of the drawings of the road sections. The section under study was imported into the software and the map overlay layer was used for validation and necessary adjustment. The result was a transition from drawings Figure 7 to a preliminary microsimulation model Figure 16.

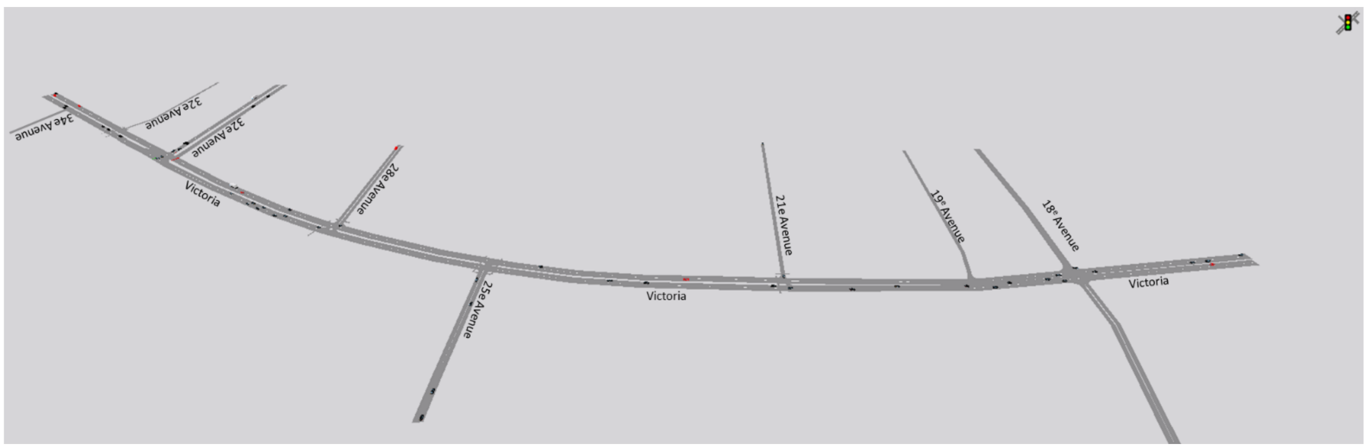


Figure 16 Snapshot of the built network

Aside from the network's physical form, other network characteristics such as, connectors, conflict zones, reduce speed zones, signalization, control devices, sensors, desired speed was applied using the data collected from the experiment. The most valuable information for the network is the traffic assignment which defines the movement of the travelers in the network. Consequently, the data from the experiment and vehicle count from the city with 4 different time intervals was examined and incorporated into the model as shown in Figure 17.

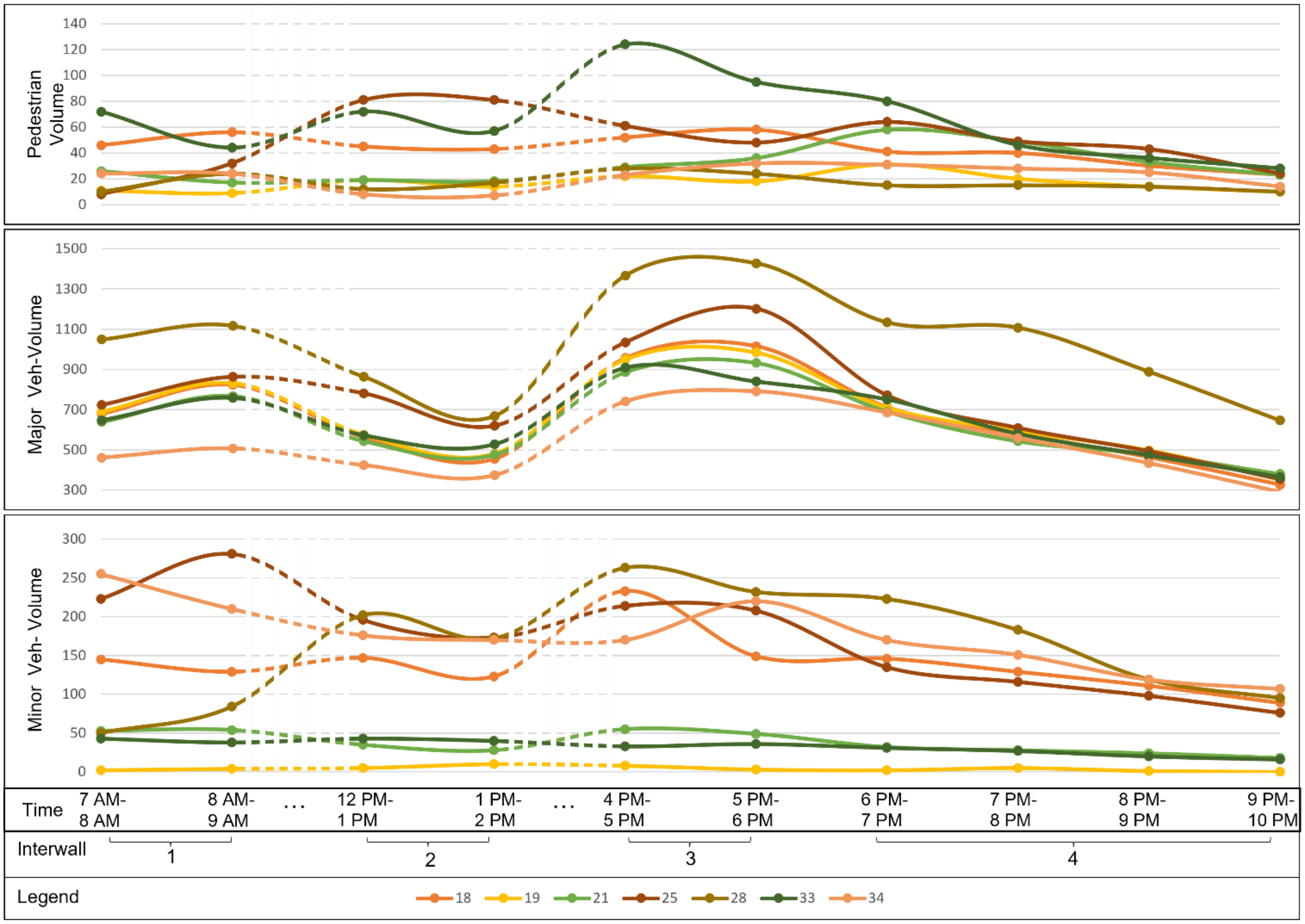


Figure 17 Hourly, traffic counts Major and Minor

The car-following model was implemented with the additional constraints for safe-to-stop braking distance. Lane-changing maneuvers are not allowed when vehicles are queuing up waiting to turn or advancing through the intersections.

A major limitation of the VISSIM model was the inability of the model to reflect the same driver behavior at the stop-sign as it was captured from the experiment. There is no explicit feature available in VISSIM that can simulate a roll-through or blow-through behavior. Essentially, neither vehicles would be able to burn the stop-sign, nor they behave differently when a stop-sign type changes. Thus, all the vehicles approaching to the stop-sign were making the expected, full-stop, before proceeding into the intersection.

To bypass this limitation, two dummy links was added to the approaches in a fashion that travelers violate the stop rules while reducing their speed. The odd ratio for distribution was taken from the qualitative analysis made for this reason under section 4.2.2.1.

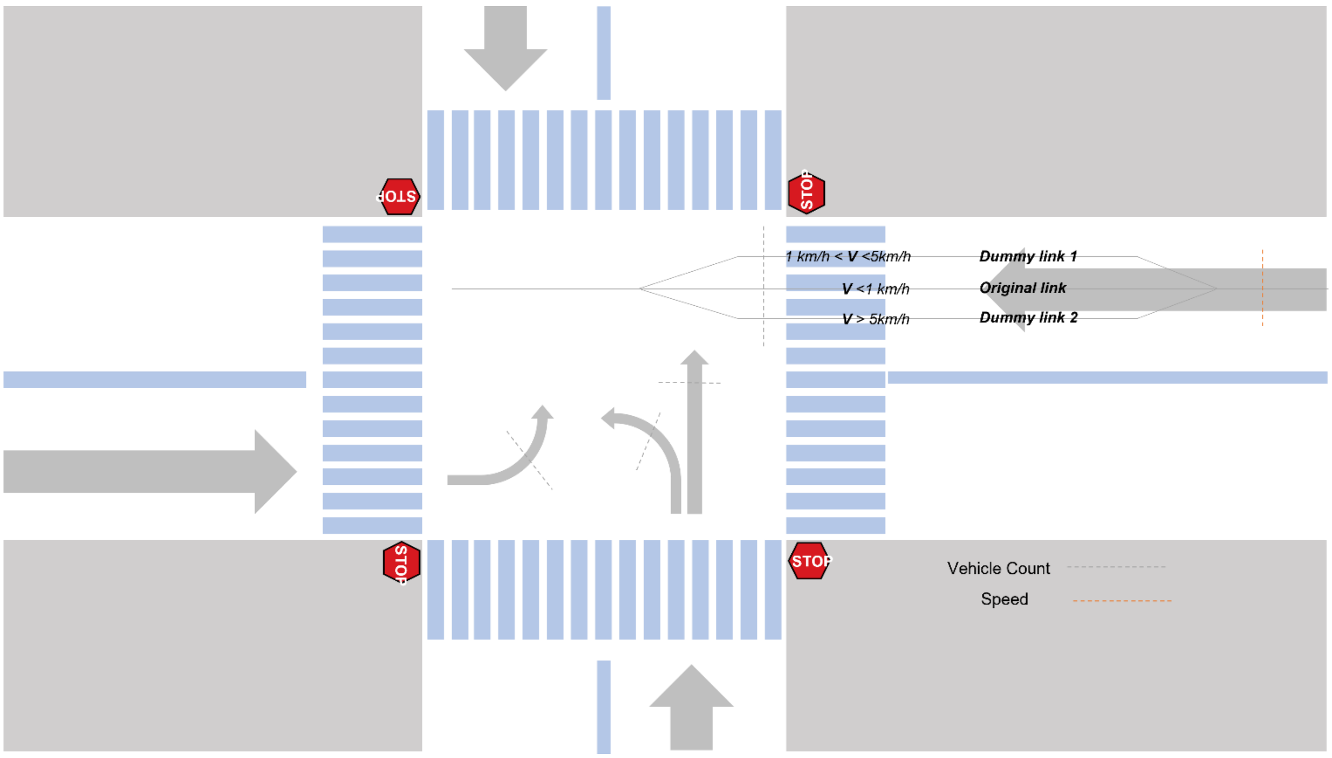


Figure 18 Dummy links and their position with respect to the stop line.

This technique yielded three alternative results at the stop line: full-stop, roll-through, and blow-through. The treatment type at each approach, for the time of the day, determines the ratio of utilization of these dummy links. The next section describes the processes of model calibration to represent the experiment, pattern in traffic, speed, density, queue, delay time, travel time, and most importantly vehicle trajectories and conflict.

3.3.2 Calibration of the Simulation Model

The success of a traffic simulation model in delivering accurate results is determined by how well it was calibrated. Typically, the calibration of microsimulation models is carried out by using vehicle counts and, in some cases, average speed of vehicles on roads. Given that the focus of this research was on conflicts, it was important that the model is also be calibrated against observed conflicts. To the best of the author's knowledge, available traffic microsimulation platforms (such as Paramics, VISSIM, or Aimsun) currently do not provide conflict estimates as a direct output of a model. In practice, conflict is being estimated by external modules (e.g., SSAM) that use the simulation output. Regrettably, the SSAM platform does not allow direct connectivity with any of the optimization platforms to conduct an iterative process. This along with the small size of the network, dictated the need for the model to be calibrated manually for the purposes of analysing conflicts in this research. However, as software packages expand and potentially produce conflict results, it is imperative to automate the calibration process using conflict in the objective function for the future research. Nonetheless, the model was still calibrated by carrying on one extra step in calibration. Due to the stochastic nature of the microsimulation, and the complex/non-linear relationship between model parameters and the objective function, heuristic methods are best for model calibration [112–114]

Compared to the trial-and-error technique, the heuristic methods are robust, and the computation is automated to avoid exhaustive computation in search of the global optimum. Among the heuristic models, according to Ma et al., "One cannot safely say that one particular method outperforms all others". In this research, a Genetic Algorithm (GA) was used because, compared to the other used optimization algorithms, GA is a multiple point search that reduces the number of search steps needed to determine optimal set of effective parameters. In addition, it has the advantage of not requiring gradient information without obvious disadvantages. While GA performs well on discrete data such as

TMC, other algorithms such as artificial neural network (ANN) usually performs better on continuous data such as speed. The technique has been widely used in similar research and the results have been shown to be stable in similarly formulated problems [85,112]. As such, the effort was spent to develop a Genetic Algorithm that could be expanded in the future for model calibration/validation.

VISSIM is offering an interface (API) which enables inter-process communication between software. The interface is based on the Component Object Model (COM) technology and creates a hierarchical model in which software can alter the simulator's functions and parameters. These are the parameters that a user can change. The GA was coded in MATLAB as: the interface integration was plausible with VISSIM and as a mathematical software package it has built-in functions for optimization. It should be noted that conflict was a better parameter in optimization than TMCs however the current code should be treated as a proof-of-concept demonstrating the GA capabilities with capacity to expand and include conflict as part of the objective function once conflict becomes a direct output of a simulation.

3.3.2.1 Calibration Parameters

VISSIM includes about 190 driving behavior, vehicular and other parameters[115]. In this study driving behavior parameters are the predominant factors under consideration and therefore the significance of their impact needed to be identified to select from these parameters. VISSIM also contains psycho-physical car-following model for longitudinal vehicle movement and a rule-based algorithm for lateral movements defined through the Wiedemann model 74. Based on the previous studies, the sensitivity analysis process was conducted by investigation of variance ANOVA for different parameters under "driving behaviors" and the following parameters were used [36,73,83] .

- Max and accepted deceleration (own and trailing) when vehicle can slow down safely without any dangerous with accepted deceleration.

- Maximum look ahead distance for what driver can see forward to react to other vehicles either in front or to the side.
- Average standstill distance for the average desired distance between stopped cars and stop-lines.
- Additive/multiple part of desired safety distance which determines the saturation flow rate.
- minimum distance (headway) to the vehicle in front that must be available for a lane change.

The table below summarizes the driving behaviour parameters that were used in this thesis as calibration parameters, their corresponding attribute name in VISSIM-COM, range and, the increments used in the genetic algorithm.

		Range and increment		
Parameter Name (GA Gene)	COM Attribute	Min	Max	Step
Max Deceleration Own (m/s2)	MaxDecelOwn	-4	-1	0.5
Accepted Deceleration Own (m/s2)	AccDecelOwn	-4	-1	0.5
Deceleration Reduction Distance (Own) (m) *	DecelRedDistOwn	50	100	2
Deceleration Reduction Distance (Trail) (m)	DecelRedDistTrail	50	100	2
Max Look Ahead Distance (m)	LookAheadDistMax	100	300	8
Avg Standstill Distance for Wiedemann 74 (m)	W74ax	0	5	0.5
Min clearance (front/rear) (m) (i.e., headway)	MinFrontRearClear	0.1	5	0.1
Additive part of safety distance	W74bxAdd	1	3	0.1
Multiplic. part of safety distance	W74bxMult	1	3	0.1

Table 4 Driving behaviour parameters used for network calibration.

*1m/s^2 per unit distance (DecelRedDistTrail and DecelRedDistDown): This reduces the Maximum deceleration with increasing distance from the emergency stop distance linearly by this value down to the Accepted deceleration.

3.3.2.2 Objective Function

Calibration of the parameters presented in Table 4 require the definition of goodness-of-fit measures to compare model outcomes to observations of the system being modelled. The Mean Square of Errors

(MSE), between model estimates and field measurements for turning movement counts (TMCs) is used in this case study. Some studies suggested alternative metrics such as traffic speed, or travel time/speed (2.3.2.3). Since speed distribution at multiple point was one of the experiment inputs which was employed in the model, hence there was no guarantee that it would be a proper metric. Vehicle turn at intersection, on the other hand, was deemed a key contributing factor in driver's reaction at the stop signs (4.2.2.1). There are two reasons for using the squared difference between the observed and simulated values: higher penalty is placed on larger errors and to avoid balancing out the errors of opposite signs[85].

$$SSE = \sum_{i=1}^{n} (C_i^E - C_i^M)^2 \qquad \textit{Eq. 18}$$

where: i is cross reference number for turn,
n is the number of turning movement counts
C^E and C^M are observed and simulated volumes for a turn, respectively.
GA is being used to define the function (f) between SSE and the parameters (j) suggested from Table 4 and minimise SSE.

$$\min SSE = f(x_j) \qquad \textit{Eq. 19}$$

3.3.2.3 Genetic Algorithm for Model Calibration

Driving behaviour parameters were calibrated using a genetic algorithm (GA) to reflect observed local conditions. The genetic algorithm, which is based on natural selection, the mechanism that drives biological evolution, is a technique for resolving both constrained and unconstrained optimization issues. In the next part, the chromosomes that make up the GA representation, the selection strategy, the chromosomal encoding, crossover, and mutation operators, as well as their parameters—population size, crossover frequency, and halting criteria is described.

Each of the GA-adjusted model parameters, such as headway, deceleration, and look ahead distance, represents a single gene on a chromosome. The real-value coding of the genes was chosen over binary coding because real-valued genes are more efficient, require less memory, have no loss of precision, and allow for higher mutation rates[116].

The size of the population is often limited by the time it takes to finish each simulation run. The assumption here was that the genetic algorithm would converge in an acceptable amount of time. A limited population size, on the other hand, could result in a rapid convergence to the local minimum. A significant population size, on the other hand, would operate as a random search, increasing the time to convergence to unfeasible limits.[117]. the population size is an important parameter which directly influences the ability to search an optimum solution in the search space. Some researchers have revealed that having a population size at least as large calibration parameters leads to the accuracy of getting an optimal solution. The initial population (with a population size of 9) was defined using random values (pseudorandom values drawn from the standard uniform distribution on the open interval 0 and1) for each parameter (gene). In order to evaluate the fitness of each chromosome, an average from the value of the objective function from 5 simulation runs is extracted while the same parameters are used in each simulation but different random seeds. After generation of the first population, in every iteration, a set of parents that have lower fitted values are being selected to breed the new offspring.
Four actions are being made in breading. First the best chromosome (the elite) is being selected to the next generation. The roulette wheel selection and cross-over operator mechanism is used to generate six new offspring. The crossover operator in this method underlies the process of producing two children form each two parents using two-point crossover. Each parent can participate in more than one crossover process. There is no constraint on not repeating same parents in the roulette wheel selection process.

From the remaining unselected parents not attending the cross-over operator (minimum 2 max 6), the best one is mutated to avoid premature convergence to local optimum. The selected chromosome will be mutated or not with probability of mutation, the direction of change for to increase or decrease is using a fair coin toss trial and finally up to how many steps the gene is mutated is randomly selected for between the lower and upper boundaries. So far, 8 new parents are counted which includes; the elite from previous population, 6 from the roulette wheel and the chromosome from the outcome of mutation. The final chromosome in each iteration is a randomly generated new chromosome which is generated the same way from the process of initial process. This to avoid the early convergence even more and expand the search area outside of the search areas between the elites (Figure 15).

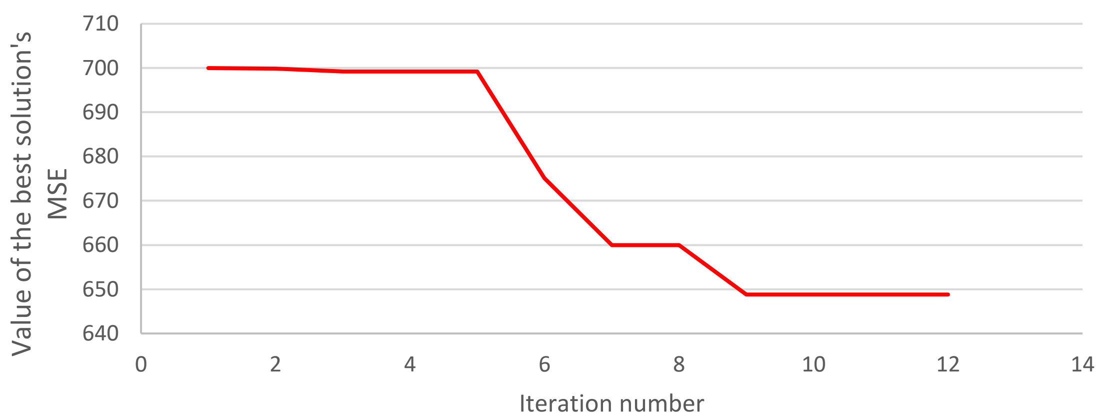
This process is repeated until a termination criterion is reached. According to Saleh et al., the most common stopping conditions are[118]:

- A solution is found that satisfies minimum criteria
- Fixed number of generations reached
- Allocated budget (computation time/money) reached
- The highest-ranking solution's fitness is reaching or has reached a plateau such that successive iterations no longer produce better results
- Manual inspection
- Combinations of the above

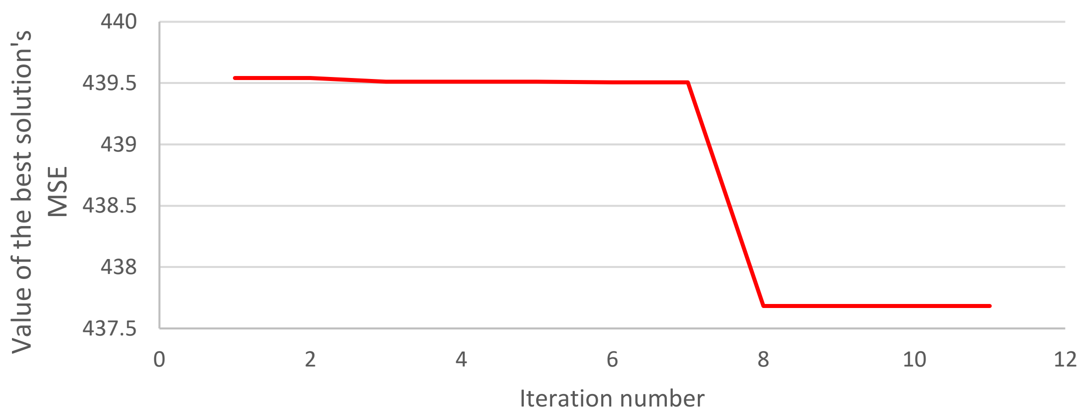
For calibrating the microsimulation, the desired value of the objective function is zero (simulation matches all the observed counts). But 100% calibration is never achievable, so some realistic criterion would need to be set. The plateau reaching termination clause, offers a good balance between computational effort and controlling the uncertainty of the quality of results as the number of iterations without observing improvements can be controlled. Thus, the GA is allowed to run until no further improvements could be made for four consecutive generations within reasonable computational effort (a maximum of 15 iterations as the network is small). The process of updating the configuration file and running VISSIM was automated using MATLAB and the GA algorithm codes are included under section 9.2.

3.3.2.4 GA Results

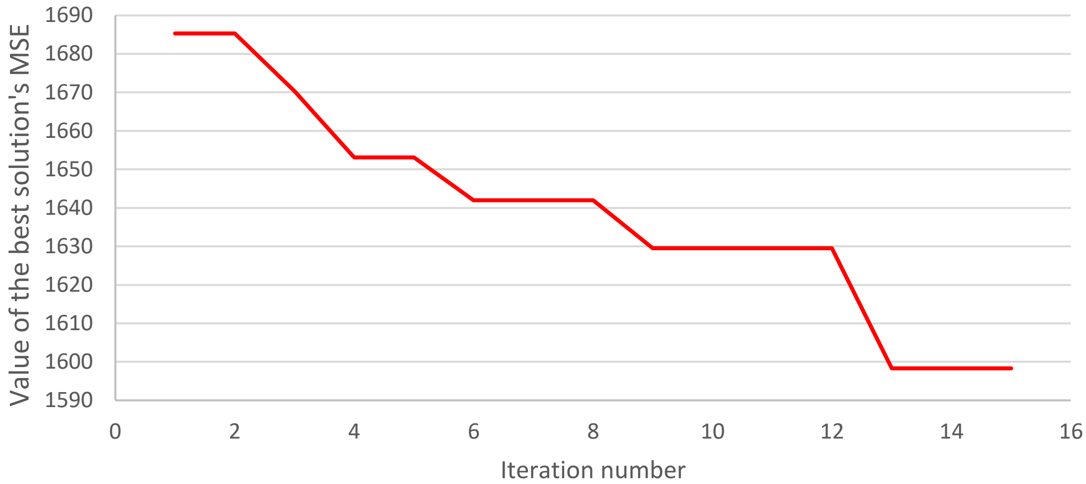
The network was calibrated for the AM, Mid-day, PM peak hours and the four hours evening traffic. The following figures showing the convergence of the best solution for each of the modelling periods. As can be seen from the Figure 19, calibration of the AM and Mid-day models stopped after 12 and 11 iterations, respectively as there was no improvement to the objective function. The PM model and the evening models, however, stops as the genetic algorithm reaches its maximum number iterations.



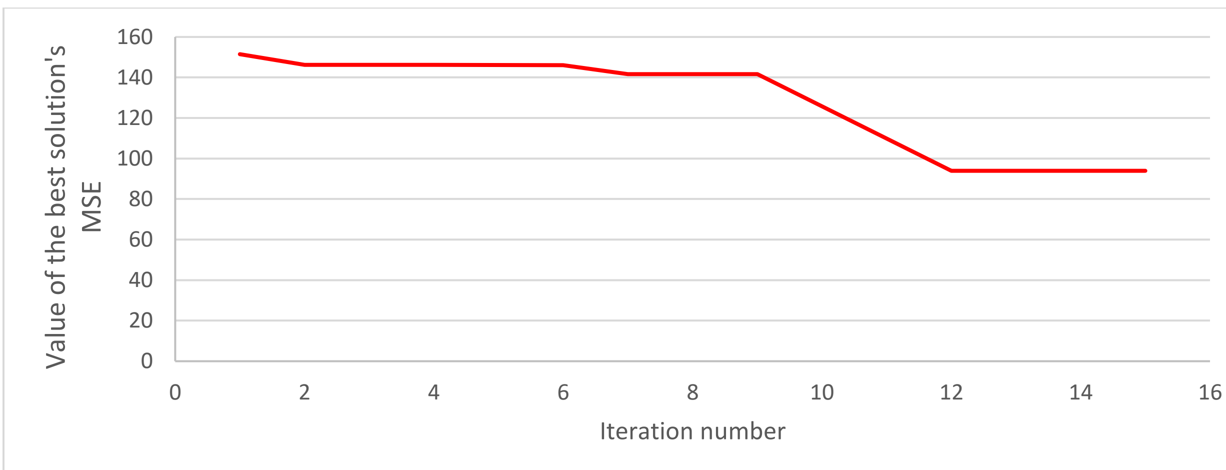
(a)



(b)



(c)



(D)

Figure 19 Improvement of the Value of the Best Solution's Objective Function for the a) AM Peak Hour, b) Mid-day Peak Hour, and c) PM Peak Hours d) Evening Peak Hour

Table 5 summarizes the results from the GA and identifies the final chromosome for each period.

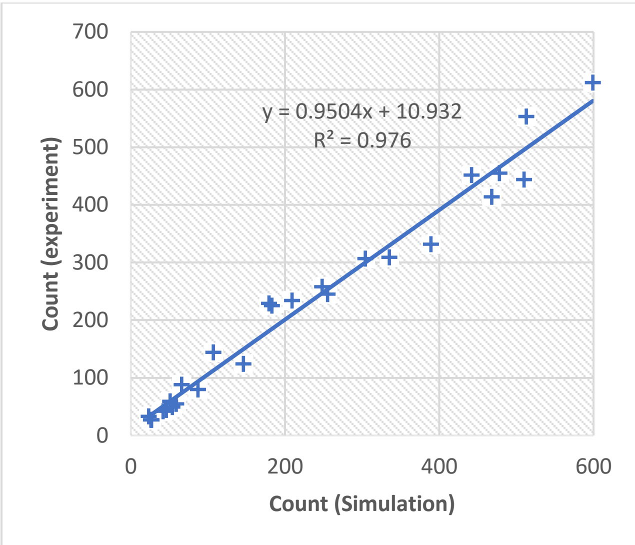
	Range	AM Calibrated Value	PM Calibrated Value	Mid-day Calibrated Value	Evening Calibrated Value
Max Deceleration Own (m/s2)	[-4,-1]	-2.5	-1.5	-3.5	-1
Accepted Deceleration Own (m/s2)	[-5,-1]	-2	-1.5	-5	-4
Deceleration Reduction Distance (Own) (m)	[50,100]	78	86	86	78
Deceleration Reduction Distance (Trail) (m)	[50,100]	92	56	62	50
Max Look Ahead Distance (m)	[100,300]	276	236	300	132
Avg Standstill Distance for Wiedemann 74 (m)	[0,5]	4	2.5	2	0
Min clearance (front/rear) (m) (i.e., headway)	[0.1,5]	0.1	4.9	1.1	3.9
Additive part of safety distance	[1,3]	2.9	2.3	2.8	2.4
Multiplic. part of safety distance	[1,3]	2.6	2.6	2.4	3.0

Table 5 Final Chromosome for each period

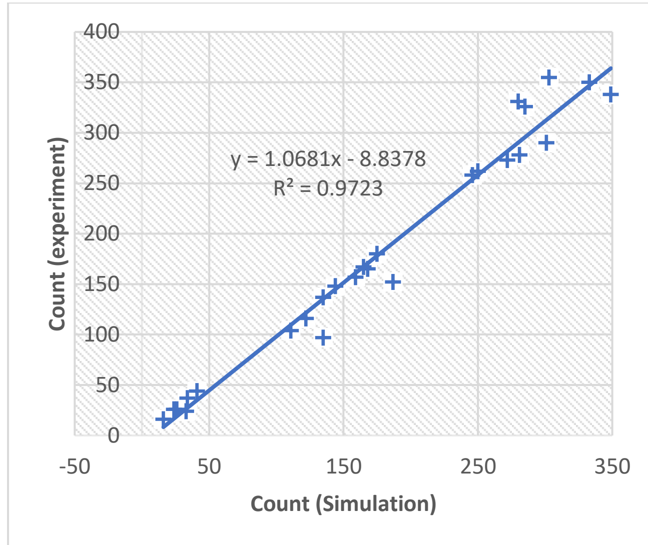
The GEH parameter - measuring the percent error with respect to the mean value of the observed and simulated counts was also calculated for the final solution. The GEH is calculated by:

$$GEH = \sqrt{\frac{(Obs - Sim)^2}{(Obs + Sim)/_2}} \qquad \textit{Eq. 20}$$

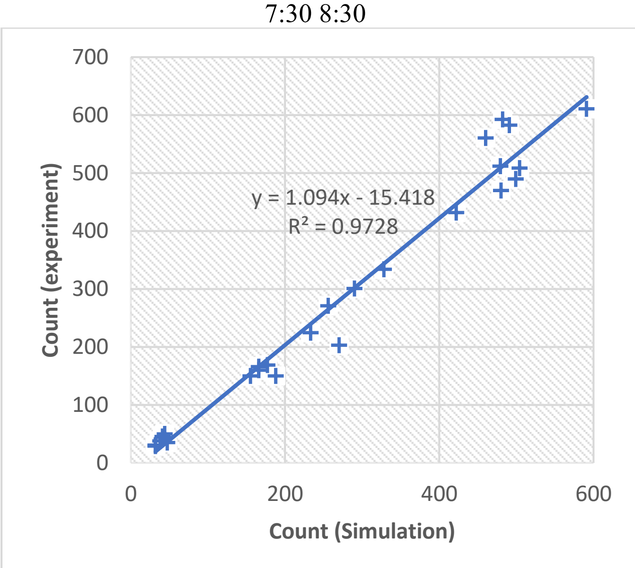
Where: 'Obs' is the observed turning movement count at a specific location, and 'Sim' is the simulated turning movement count at a specific location.



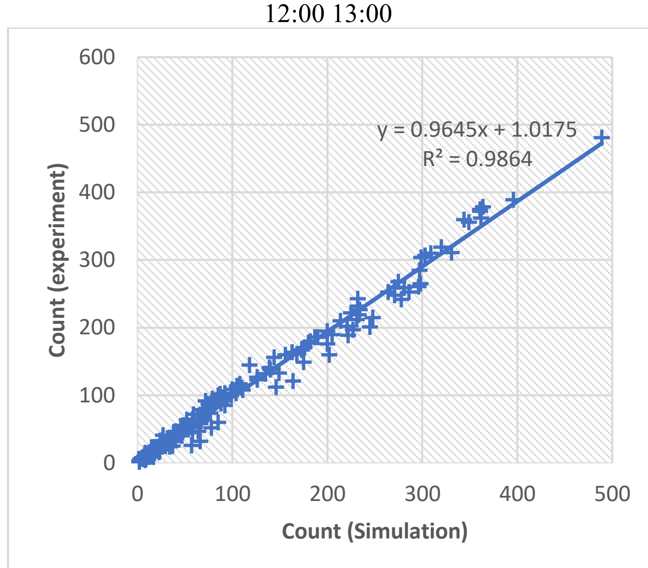
7:30 8:30



12:00 13:00



16:30 17:30



18:00 22:00

Table 6 Turn count trendline between simulation and experiment

GEH values below 5 are considered to be a good match between model volumes and observed counts. Common role for model calibration requires that at least 85% of the observed link volumes/turns in a traffic model have a GEH less than 5. In this model, 100% of the turns had a GEH of less than 5.

3.3.2.5 Validation of Conflict

As mentioned above, VISSIM was used to construct the trajectories of all the simulated vehicles, and the recorded trajectory data was then analysed using SSAM to determine the vehicle conflicts in the network. The VISSIM vehicle trajectory data was used in SSAM to analyse the detected vehicle conflicts in the study area. Figure 20 shows the coordinate of the conflicts captured from the trajectory output file. Three type of conflict is identified in the area's filtered in the software (intersections). Most studies use two surrogate metrics to assess traffic safety: Time to Collision (TTC) and Post Encroachment Time (PET). SSAM can estimate the TTC and PET values of each vehicle interaction automatically, allowing it to record all potential conflicts.

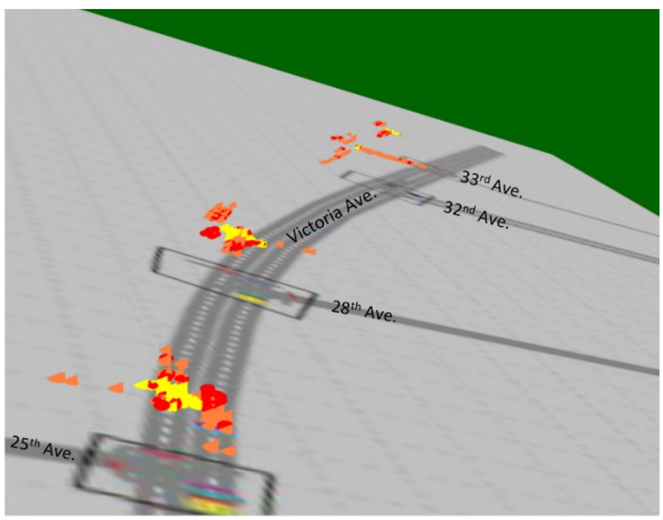


Figure 20 Three-dimensional visualization of hourly conflict in SSAM, yellow: crossing, orange: lanechange, red: rearend

Validation step would strive to looking into the conflict (cumulative both types) and the correlation between simulation and the conflicts from the experiment. Different conflict types: crossing, rear-end and lane changing can be detected by setting the threshold for TTC, and PET. The type of conflict in the performed analysis is selected based on the characteristics of facility. More details about using GA in simulation models calibration can be found in section 2.3.2.4. Based on previous studies, on a two-lane SCI, over 85% of the actual conflict percentile are happening with TTC ≤ 1.5 Sec and PET ≤ 5 Sec. Zero values would need to be removed as these conflicts are caused by errors in the simulation as well as irrelevant conflicts outside of a 50-meter diameter of the intersection center and north/south bound conflicts. Conflict angle thresholds remained as default with 30^ and 80^ for rear-end and crossing angle consecutively. A two-sample test was caried out for validation process. Since the conflict data is discrete and not normally distributed, Mann-Whitney U-Test, which is nonparametric equivalent to the t-test, was used to validate the performance of the conflict data from the simulation with following hypothesis:

- H_0: There is no difference in samples of observation conflict with respect to data from simulation,
- H_1: There is a statistically significant difference in the samples between the two.

Table 7 illustrates the observational and simulation conflicts, as well as their U-test rankings.

Intersection		Ave. 25th vs Victoria Ave.					Ave. 28th vs Victoria Ave.				
	Period	07:00 09:00	12:00 14:00	16:00 18:00	18:00 20:00	20:00 22:00	07:00 09:00	12:00 14:00	16:00 18:00	18:00 20:00	20:00 22:00
U_1 Simulation	Conflict	77	59	43	8	6	45	73	69	8	2
	Rank	20	15	11	5.5	4	12	19	18	5.5	1
U_2 Observation	Conflict	61	38	52	14	8	41	67	46	5	3
	Rank	16	9	15	8	5.5	10	17	13	3	2

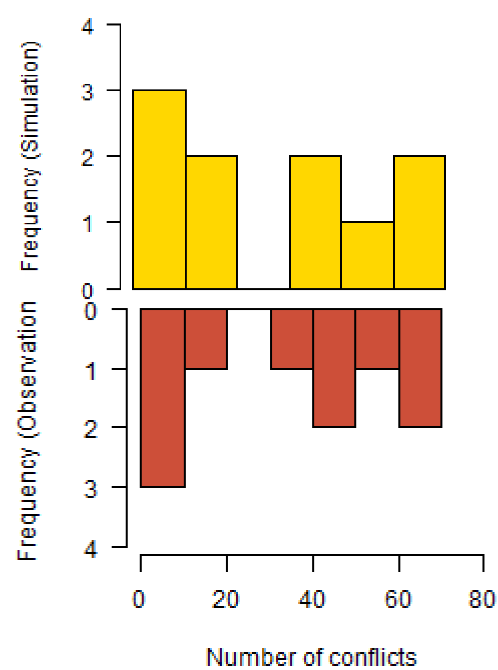


Table 7 Frequency of conflict simulation vs observation

The selected U_i is the smallest of the two U_1 and U_2 described below:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \qquad \textit{Eq. 21}$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \qquad \textit{Eq. 22}$$

Where, n_1 and n_2 are the number of samples in each group, and R_1 and R_2 are the sum of their ranks [119]. By plugging the lower $U_i = U_2$ in the Eq. 23, the Z value was calculated at 0.454, which is in the 95% region of acceptance: [-1.96: 1.96].

$$Z = (U_l - \frac{n_1 n_2}{2}) / \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \qquad \textit{Eq. 23}$$

Using R studio, the *P-Value* for a two-tailed test was extracted as 0.6499 which is greater than α, hence, H_0 cannot be rejected. It follows that the randomly selected value of the simulation-based population is assumed to be equal to the randomly selected value of the field-observation population. In other words, the difference between the randomly selected value of these two populations is not big enough to be statistically significant and validation process is complete without a need to investigate other driving behavior parameters than the literature has recommended.

3.4 Statistical Modeling

The two preceding sections of this chapter have demonstrated the data collection procedures from, [1] an experiment, [2] external resources as well as [3] a calibrated microsimulation model. In this section, several methodologies will be reviewed in alignment with the objective of the research established in first chapter. The methodology can be classified into three categories:

1. Investigating the effects of various countermeasures on variations in approaching vehicular speed using ANOVA.
2. Qualitative evaluation of drivers' responses to various countermeasures. The focus would not only be on drivers' compliance, but also on environmental factors (e.g., ambient light, pavement surface condition) and entails the following two methods:
    a. The braking habits as the first signal of drivers' reaction to plot samples distribution.
    b. Describe binomial and multinomial logistic regression models for drivers' reaction analysis at ultimate decision point, stop-line, with respect to the primary parameter, signage type, and the approach to control several parameters in the regression models. Finally, the assessment methods available for the performance of each parameter as well as a benchmark models.
3. The quantitative analysis section will walk through the process of the assessment of intersection safety performance. Several generalized linear models are presented for collision and conflict estimates and model evaluation performance and the measure of goodness-of-fit are being discusses.

3.4.1 Analysis of Speed

The speed and variance at each individual intersection were different. As a basic step, the differences between groups of data can be examined to see if they are statistically significant. It works by analyzing the levels of variance between the groups and variance within through samples taken from each of them. A t-test or pairwise t-test is a simple way to make the analysis of the variance, however, for more than two groups, the error of test would be a compound of each pairwise test. For instance, null hypothesis of t-test between two groups would be at a significant level $\alpha = 0.05$ with 95% confidence, while pairwise between three groups would be increased to $\alpha = 0.05^3$. which brings the confidence level to 0.857.

ANOVA, or analysis of variance, is a powerful statistical tool for demonstrating the difference between two or more means or components using significance tests (Eq. 24). It also demonstrates how to do numerous comparisons of the means of several populations. The ANOVA test compares two types of variation: variation between sample means and variation within each sample. One-way ANOVA test statistics are represented by the formula below[120]:

$$\text{ANOVA Coefficient(f)} = \frac{MST\ \left(Mean\ square\ due\ to\ \frac{treatments}{groups}\right)}{MSE\ \left(Mmean\ square\ due\ to\ error\ (within\ groups)\right)} \qquad \textit{Eq. 24}$$

$$MST = \frac{\sum_{i=1}^{k}\left(\frac{S_i^2}{n_i}\right) - \frac{T^2}{n}}{k-1} \qquad \textit{Eq. 25}$$

$$MSE = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} Y_{ij}^2 - \sum_{i=1}^{k}\left(\frac{S_i^2}{n_i}\right)}{n-k} \qquad \textit{Eq. 26}$$

Where; Y_{ij} is an observation j from group i, S_i is the sum of group i, T is the grand total of all observations, k is number of groups, n is the total number of observations and n_i is total observations for group i.

3.4.2 Analysis of Drivers' Compliance

The compliance to the traffic control device of a drivers approaching to a SCI can be captured in two timeframes; first reaction which occurs when a driver reads, recognizes, and reacts to a signage and second reaction which occurred when the decision is made by a driver to respond to the signage at the stop-line. In the following section the indications used to capture these moments are described.

3.4.2.1 First Reaction to the signage

Given that a driver is required to come to a complete stop at the stop-line, the deceleration rate is proportional to the time that the driver begins to brak. However, it's worth to mention that there are additional parameters that might violate this assumption, including but not limited to: the length of the links between the intersections, slope of the approach, drivers' intent to make a stop. Even though the link slope and distances between the intersections were nearly constant, the author would carefully consider whether the braking habit among the sample populations of drivers could be recognized as an indicator in performance of a treatment and commanding drivers to comply. As described in section 3.2.2.1 the collected data from the footage showed different distributions of drivers' reactions which will be discussed in section 4.2.2.

3.4.2.2 Second Reaction to the Signage

The study was carried out in three separate scenarios, each using a different data set from the collected samples (see section 3.2.2.2):

A. Same drivers (stop-sign only): In this scenario, only the vehicles traveling across the whole length of the link and participating in at least three approaches out of four, were included in the study - to make sure that the analysed data reflects the compositional variation from different drivers entering into the experiments.

B. Same direction (stop-sign only). In this scenario, all the vehicles despite of their departure from the study link was considered from all the samples.

C. Opposing directions (stop-sign and Pedestrian crossing). In this study the two approaches of a single intersection with different treatment were studied

Where $A \in B$ and $C \notin$ B

As previously stated (section 3.2.2), the Manual of Transportation Engineering Studies recommends the reaction of motorists at stop-lines to be classified into three categories: full stop ($0 - 1\ km/h$), rolled through ($1 - 5\ km/h$) and blew through (greater than $5\ km/h$).

If we consider driving compliance at the stop-line as the "dependent variable", then we have multiple outcome levels on this Variable (i.e., compliance to the stop-sign). These outcomes essentially representing different groups (e.g., drivers those who blow-through the SCI intersection). These outcomes are nominal with no ordering ranks. If the outcome was ordered (e.g., low, high) in terms of category membership it was reflecting, then binomial or multinomial was not appropriate method and perhaps ordinal logistic regression was a better fit. Although ordinal logistic regression might be a reasonable approach to use for the dependent variable (driver compliance), the proportional odds would be violated since input variable has a disproportionate effect on a specific level of the outcome variable [9]. For predicting the collision and conflicts as they are continuous dependent variable, using a given set of independent features, linear regressions are used as it's described in section 3.4.3 whereas logistic regression is only used to predict the categorical outcome [120].

For 3 category dependent variables, K-1 (3-1=2) logistic regression models were developed using SPSS Statistics version 28.01.1 (14). That means one outcome is chosen as "Reference/Pivot" category and all the other K-1 outcomes are separately regressed against the pivot outcome. Last category is usually the "reference category".

The three outcomes would be:

- Category A: Blow through,
- Category B: Roll through,
- Category C: Full stop.

Two models developed for:

1. Category A with reference to category C,

$$\ln\left(\frac{p(A)}{p(C)}\right) = a_1 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_m x_{im} \qquad \textit{Eq. 27}$$

Where α is the regression constant and β_i are regression coefficients, x_i $i = 1, \dots, n$ are the observations with *m* independent or explanatory variables.

The Eq. 27 can be written for $p(A)$:

$$p(A) = p(C) * e^{(a_1+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_m x_{im})} \qquad \textit{Eq. 28}$$

2. Category B with reference to category C.

$p(B)$ for second model is given by:

$$p(B) = p(C) * e^{(a_2+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_m x_{im})} \qquad \textit{Eq. 29}$$

Finally, $p(C)$ is equal to $1 - [p(A) + p(B)]$ and hence:

$$p(C) = \frac{1}{1 + e^{(a_1+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_m x_{im})} + e^{(a_2+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_m x_{im})}} \qquad \textit{Eq. 30}$$

The independent (to compliance) categories varied with respect to:

- Categorical:
  - Treatment type (i.e. Standard, LED, BLS, and Beacon).
- Dichotomous:
  - Pavement surface condition (wet or dry),
  - Conflict potential (absence of opposing traffic),
  - Natural ambient light (Before civic dawn),
  - Maneuver (move through approach or not).

For each comparison on the two models (with 3 categories), the estimating parameters (e.g., treatment type) are predicting whether a driver reaction falls under reference category relative to target category (Category A or B). Table 8 shows the features and descriptions of the variables.

The 'Treatment type' variable is represented by three dummy variables as follows: the first variable to represent comparisons between drivers passing through intersection with regular stop-sign (this serves as the baseline/reference category & is coded 0 across all dummy variables) and those passing through intersection with LED stop-sign ('LED STOP' variable) and set to 1. Another variable to represent drivers passing through intersection with BLS stop ('BLS STOP' variable). And the third variable to represent drivers passing through intersection with overhead beacon ('Beacon STOP' variable) – with the latter driver groups coded 1 on their respective dummy variable.

Variable Category	Variable Name	Variable Type	# of outcomes	Variable (Description)
Dependent	Driver Complacence	Nominal - Category	3	FS (Full Stop)
				RT (Roll-Through)
				BT (Blow-Through)
Independent	Treatment type	Nominal - Category	4	STOP (Standard STOP)
				LED (LED STOP)
				BLS (Backlit STOP)*
				BEACON (STOP with over head Beacon)
Independent	Pavement surface condition	Boolean	2	0 (Dry)
				1 (Wet)
Independent	Conflict potential	Boolean	2	0 (No opposing traffic/Ped)
				1 (Opposing traffic/Ped)
Independent	Natural ambient light	Boolean	2	0 (Prior to civic dawn)
				1 (After civic dawn)
Independent	Maneuver	Boolean	2	0 (Through)
				1 (Turn)
Independent	LED STOP	Boolean (Dummy)	2	0 (Standard STOP)
				1 (LED STOP)
Independent	BLS STOP*	Boolean (Dummy)	2	0 (Standard STOP)
				1 (Backlit STOP)*
Independent	Beacon STOP	Boolean (Dummy)	2	0 (Standard STOP)
				1 (STOP with over head Beacon)
* Treatment performance on the study				

Table 8 Variable description and characteristics

The goodness-of-fit of a model in a logistic regression can be tested using the likelihood ratio statistic, where the results of a likelihood ratio (such as the chi-square test) reflect the fit of the model with full complement of predictors compared to the NULL model with no predictor. If this test indicates statistical significance, it means that there is a significant improvement in fit of the model relative to the base line.

One of the more preferred alternatives of the pseudo-R-square is the McFadden's test[121]. This can be considered as an index of the proportionate improvement in model fit relative to the null model[122].

$$McFadden = \frac{Deviance_{null} - Deviance_{full}}{Deviance_{null}} \quad \textit{Eq. 31}$$

Although the indices from pseudo-R-square can be descriptive for rough analogy for the proportion of improvement between posterior and NULL model, but it cannot be interpreted as the actual proportion of variance accounted for as it is used in the context of least square regression [122].

The parameter estimator in MNL can provide the predicted change in the log-odds of the target group in the dependent variable relative to the baseline group. This means for every 1 unit increase of the independent variable, while everything else stays the same, the relative odd for dependent variable will change with respect to direction of the coefficient (increase if $\beta > 1$ or decrease if $\beta < 1$). However, each predictor's likelihood ratio test had to be reviewed to ensure that they were statistically significant for the model before making such assumption.

The result from this test would indicate (if the estimator parameter is statistically significant), what would be the log-odds for a driver to roll through instead of blow through if the stop-sign type changes from Standard to BLS knowing every other element remains the same.

Previous studies on active signage system (section 2.4.2), the countermeasure was set up for a somewhat long period of time (few years) to compare the drivers' behavior in a before and after analysis. In this study, the countermeasure was installed for only two months. two scenarios was studied during that period with regard to the compliance at the stop-sign and one scenario for Pedestrian crossing. The performance of each scenario with respect to goodness-of-fit is also presented in the next chapter.

3.4.3 Analysis of Conflict and Collision Model

While a qualitative analysis of speed and drivers' behaviour would provide valuable insight into the efficacy of various safety treatments at unsignalized intersections, collision is the best-known indicator for any road safety assessment (check section 2.3). The first part of this section will present the methodology of analysing road safety at SCI using the historical collision data. In the absence of the collision data for after deployment of a treatment, conflict is used as a surrogate measure. In this section the approach of using a validated critical traffic conflict assessment derived from a microsimulation is presented as a proxy for road safety assessment. The process of extraction of the 'target collisions' from the observation (8 years period) is being presented. Poisson and NB estimation method are also presented along with the study plan to incorporate the experiment and simulation output data into the estimate models for collision and conflict.

3.4.3.1 Collision History

The collision data was extracted, and cross referenced between Transport Canada National Collision Database (TC-NCDB) and hard copies from the police records in borough of Lachine. The dataset combined includes all the required attributes for the collision study. That includes georeferenced, time tagged and collisions type. Despite acceptable data size and complete collision attributes for the network under study, there was no collision record for the BLS SCIs to allow a conventional safety performance evaluation. This was attributable to the relative short timeframe of the study after installation of BLS (3 months) and subsequent unavailability of relevant "after" collision data. However, the frequency of collision data and the period was significant enough for an attempt to build a safety performance function for the SCI. This would give an opportunity to consider the effect of the observed and simulated data from drivers' behaviour, on the performance of not only conflict model but an actual SPF. Collision count was not sufficient to attempt severity analysis or analysis by collision type. The inquiries made to the municipality and borough of Lachine, revealed that neither the average annual daily traffic of these intersections nor the SPF were available. Hence, only the average short-term count was used, as explained in the previous section.

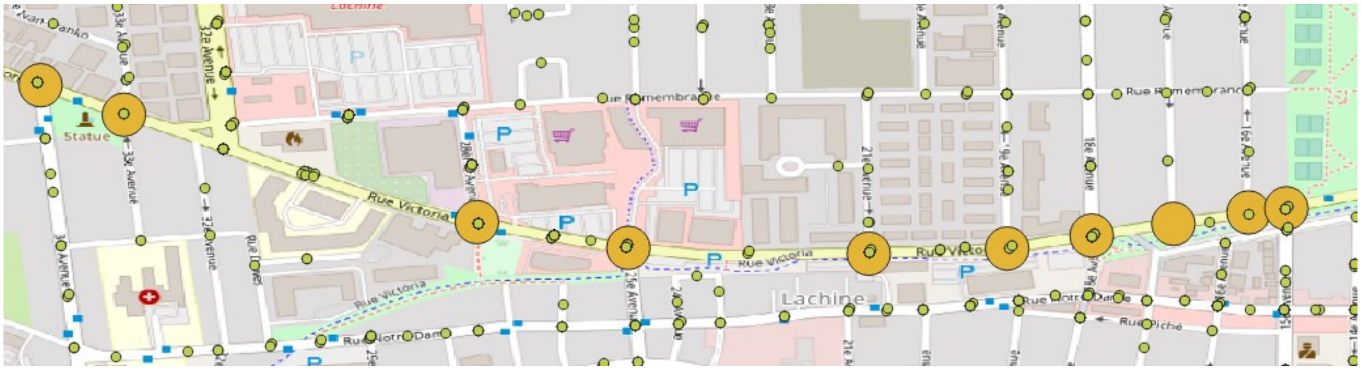


Figure 21 Spatial distribution of the observed collision (QGIS)

In order to view, filter and analyze the geospatial collision data an open-source geographic information system application was used. The available data was loaded on QGIS 3.22.2 and by several tasks, including the addition of a buffer to the link and filters, the collision data was extracted and cross referenced with the manual police report in excel. To make sure that no collision is being unrecorded from the data set, a manual validation of the collisions outside the buffer zone was done since there are instances that location of an accident is attributed to the closest civic address (e.g., a house number) while the detailed police report indicates the intersection.

Intersection vs Victoria Ave	**Total Collision (collision used for the study)**			
	Day	**Night**	**Twilight**	**Grand Total**
18th	25 (17)	8 (5)	()	**33 (22)**
19th	5 (1)	2 ()	()	**7 (1)**
21st	9 (5)	6 (3)	1 ()	**16 (8)**
25th	17 (14)	7 (3)	()	**24 (17)**
28th	12 (10)	7 (5)	2 (2)	**21 (17)**
33rd	4 (1)	3 (1)	()	**7 (2)**
34th	3 (1)	2 (2)	1 (1)	**6 (4)**
Total	**75 (49)**	**35 (19)**	**4 (3)**	**114 (71)**

Table 9 Collision count at intersections (2012-2018)

A more detailed analysis of the accident reports showed that the collisions occurred at specific periods. Over 62 percent of collisions at these intersections occurred during the three peak-hours of the day and the four-hour time periods during the night. Because these periods corresponded to peak traffic volumes and high traffic at night, it was considered worthwhile to use this sub-set as the "target collisions" for the study.

3.4.3.2 Collision model selection

As mentioned in section 2.3.1 the Poisson distribution approximate rare-event count data such as collisions. Poisson distribution is a series of Bernoulli trials which falls under the Binomial Distribution.

In Bernoulli trials the collision is a 'success' with probability of p and q is the 'failure' probability ($q = 1 - p$). In the context of an intersection with N passages (trials), a random variable Y can be considered that records the number of 'success' events out of N trials. Hence a series of Bernoulli trials which is a Binomial distribution is given as follow:

$$p(Y = y) = \binom{N}{y} p^y (1 - p)^{N-y} \qquad \textit{Eq. 32}$$

Where, y is a non-negative integer defined as number of collisions. Since it's unlikely and rare for occurrence of 'success' with a large number of trials, it can be demonstrated that a binomial distribution can approximate a Poisson distribution. Let $p = \frac{\lambda}{N}$ (λ mean of the Poisson distribution) so that a large trial size offset by the diminution of p to produce a constant λ for all values of p.

$$p(Y=y)=\binom{N}{Y}\frac{\lambda}{N}^{y}\left(1-\frac{\lambda}{N}\right)^{N-y}$$ *Eq. 33*

While the binomial coefficient given as follow:

Eq. 34

$$\binom{N}{Y}=\frac{N!}{(N-y)!\,y!}$$

and Eq. 33 given Eq. 34 will be:

$$p(Y=y)=\left(\frac{N!}{(N-y)!\,y!}\right)\lambda^{y}\left(\frac{1}{N}\right)^{y}\left(\frac{N-\lambda}{N}\right)^{N-y}=\frac{\lambda}{y!}^{y}\left(\frac{N!\,(N-\lambda)^{N-y}}{(N-y)!\,N^{N}}\right)$$ *Eq. 35*

Hence as N → ∞

Eq. 35 by approximation can be shown as:

Eq. 36

$$p(Y=y)\cong\frac{\lambda}{y!}^{y}\,e^{-\lambda}$$

Eq. 36 means that the probability of number of crashes per time interval to be equal to y while if we considering Poisson as process in the interval i, then this equation can be shown as: [42,123]

$$P(y_i)\cong\frac{EXP(-\lambda_i)\lambda_i^{\,y_i}}{y_i!}$$ *Eq. 37*

Where, y_i is a nonnegative integer which represents the number of collisions on the intersection $i, (i=1,2,3,\dots,n)$ and λ_i is the Poisson parameter for that intersection. The safety performance function λ_i gives the “expected” number of collisions per period (year) which indeed is the Poisson parameter, and it can be defined as a function of explanatory variables (e.g., traffic flow), as shown below:

$$\lambda_i = f(X_i; \beta) = EXP(\beta X_i) \quad \textit{Eq. 38}$$

Where: X_i is a vector of explanatory variable, i is the facility (e.g., intersection or road section) and β is a vector of regression parameters. Since flow and collision frequency have a proven non-linear relationship at intersections hence the Eq. 38 can be expanded as [124,125]:

$$\lambda_i = F_{1i}^{\beta_1} F_{2i}^{\beta_2} e^{\beta_0 + \beta_3 X_{3i} + \cdots + \beta_n X_{ni}} \quad \textit{Eq. 39}$$

Where: F_{1i} and F_{2i} are vectors of flow (e.g., major road, minor road, motorist, non-motorist, …)

From rearranging the exponential terms from Eq. 39 we get the following linear function which a regression model can then estimate the value of β while the assumption is that collision at n facilities are independent from each other and that the number of collisions at site i represented by λ_i follows a Poisson distribution.

$$\lambda_i = EXP(\beta_0 + \beta_1 \ln(F_{1i}) + \beta_2 \ln(F_{2i}) + \beta_3 X_{3i} + \cdots + \beta_n X_{ni}) \quad \textit{Eq. 40}$$

By taking log from both sides:

$$\ln(\lambda_i) = \beta_0 + \beta_1 \ln(F_{1i}) + \beta_2 \ln(F_{2i}) + \beta_3 X_{3i} + \cdots + \beta_n X_{ni} \quad \textit{Eq. 41}$$

We are interested to maximize the probability of collisions given β (parameter of the model) and find a vector that best describes and models the probability function. The Maximum likelihood estimate generates Poisson parameters that are consistent and asymptotically normal and efficient Using and rearranging the

Eq. 36 and Eq. 37 the likelihood function for β is shown as:

$$L(\beta) = \prod_i \frac{EXP(-EXP(\beta X_i)(EXP(\beta X_i)^{y_i}}{y_i!} \quad \textit{Eq. 42}$$

And ultimately a log from the likelihood function in Eq. 42 can construct a linear function for β that is easy to manipulate for the estimation of the β.

$$LL(\beta) = \sum_{i=1}^{n} (-EXP(\beta X_i) + y_i(\beta X_i) - ln(y_i!)) \qquad \text{Eq. 43}$$

The following variable was considered in the Poisson model for the link under the study:

Variable, (name)	**Description**	**Data source**
λ_i ,(Collision)	Dependant variable for Collision model	collision history 3.4.3.1
F_1,(F_{ma})	Hourly Flow of traffic in the Major Road (Victoria Avenue)	Experiment and historical data
F_2,(F_{mi})	Hourly Flow of traffic in the Minor Road (extensions of Victoria Avenue)	
F_3,(F_{pe})	Hourly count of pedestrians	
X_4, (FS_{BT})	Compliance to the sign Full-stop ratio to blow-through	Qualitative study under section 4.2.2.1
X_5, (FS_{RT})	Compliance to the sign Full-stop ration to roll-through	
X_6, ($FS_{Non\text{-}FS}$)	Compliance to the sign Full-stop to non-full stop	
X_7(Conf)	3.3.2 number of all conflicts at each intersection	Simulation model in section 3.3.2

Table 10 Collision frequency model covariate list

After the calibration of the model with β_i, the estimate λ_i will be enhanced for each intersection with an updated $\widehat{\lambda_i}$ from the calibration of all explanatory variable. If all $\beta_i = 0$ the result of the estimate defines the null model (known as naïve estimation).

3.4.3.3 Conflict model selection

Given the data set of collision counts and section traits, the first step is to test the presence of "overdispersion", to discriminate between the Poisson model and the NB or other models. It was mentioned in the section 2.3.1 that the Poisson model is vulnerable to over-dispersion ($\sigma^2 = \mu$, where the mean, μ, is denoted by λ) due to heterogeneity or high proportion of zero accident[46] . Poison-Gamma (known as Negative Binomial - NB) method is more suitable [48], because the number of zero accidents are significant.

Considering the above comments, in this research the collision data was better fitted Poisson properties model (see section 4.3.2). However, unlike the collision, conflicts are better modelled through NB models [126] , as was the case for this study. The likelihood tests conducted was also supported a better fit of the NB model in comparison to Poisson (see section 4.3.3.1). When the equality $\sigma^2 = \mu$ does not hold then NB will remove the restriction and variance can be different from the mean. The gamma variable in NB is a random effect with a mean equal to one hence it has no effect on the estimation of mean of NB while the variance of gamma is not equal to one ($VAR = \alpha$) and hence can improve the variance estimate in NB. NB model is derived by rewriting the Eq. 38. such that, for each observation "i":

$$\lambda_i = EXP(\beta X_i + \varepsilon_i) \qquad \textit{Eq. 44}$$

The $EXP(\varepsilon_i)$ is the gamma probability distribution term with both parameters $\alpha > 0$ and will help the variance to differ from the estimate of the mean.

$$EXP(\varepsilon_i) \sim \Gamma(\alpha, \alpha) \qquad \textit{Eq. 45}$$

Hence the estimate for the variance of the Eq. 44 is given:

$$\text{Var}[y_i|\lambda_i,\alpha\,] = \sigma^2 \; = \lambda_i + \; 1/\alpha\lambda_i^2 \qquad \textit{Eq. 46}$$

The parameter α is the over-dispersion parameter, which can define the selection criteria between Poisson or NB. And finally, the probability for the dependent variable for each observation *"i"* is estimated as follow[123]:

$$P\,(y_i) = y_i|\lambda_i,\alpha = \frac{\Gamma\left(\frac{1}{a}\right) + y_i}{\Gamma\frac{1}{a}y_i!}\left(\frac{1/a}{\left(\frac{1}{a}\right) + \lambda_i}\right)^{1/\alpha}\left(\frac{\lambda_i}{\left(\frac{1}{a}\right) + \lambda_i}\right)^{y_i} \qquad \textit{Eq. 47}$$

Consequently, the log-likelihood of the NB model is given as:

$$LL(\beta,\alpha) = ln\left[\prod_i \frac{\Gamma\left(\frac{1}{a}\right) + y_i}{\Gamma\frac{1}{a}y_i!}\left(\frac{1/a}{\left(\frac{1}{a}\right) + \lambda_i}\right)^{1/\alpha}\left(\frac{\lambda_i}{\left(\frac{1}{a}\right) + \lambda_i}\right)^{y_i}\right] \qquad \textit{Eq. 48}$$

This maximum likelihood estimation can be estimated using iterative optimization methods or numerical methods (e.g., Newton-Raphson) to find the best answer for β and α that approximately solve this equation. For this purpose, a R language codes was implemented that maximizes the log-likelihood function. R-Studio version 2021 as an integrated development environment for running R script was used with the addition of the following libraries: MASS, stats, ggplot2, GGally and lmtest. Some libraries were used for statistical analysis and some for improving the visualization. The codes for all the statistical analysis can be found under Appendix III.

For conflict regression model, we employed the same independent variables as in the Table 10 for the regression model, but the dependent variable was switched from collision to conflict.

Zero-inflated NB model was not considered since there was no intersection with "Zero-count" state, where the likelihood of the collision is extremely rare compared to "normal-count" state[127]. The same can be said about Generalized NB model since there was no over-dispersion for the model hence, we expect no over-dispersion for any individual intersection. In the next section, the goodness of fit will describe the importance of having less parameters in the model. In AIC, the log-likelihood is not the only factor impacting the performance of a models, but the number of parameters will have an effect.

3.4.3.4 Measures of Goodness-of-fit

There are numerous goodness-of-fit statistics used to assess the overall fit of Poisson and NB regression models. Chi-square Pearson (*Eq. 49*) and Deviance (*Eq. 50*) statistics test are used to assess model's overall performance [47,123].

$$X^2 = \sum_{i=1}^{n} \frac{(y_i - \widehat{\lambda_i})^2}{\widehat{\lambda_i}} \qquad \textit{Eq. 49}$$

$$G^2 = 2\sum_{i=1}^{n} \{ y_i \ln y_i - y_i ln\widehat{\lambda_i} - (y_i - \widehat{\lambda_i}) \} \qquad \textit{Eq. 50}$$

The result from these two tests can be compared against the degree of freedom (number of "n" observations from which "k" number of parameters "β_i" substructed) with the ideal condition that degree of freedom matches Pearson statistics. Overdispersion occurs when the Pearson statistics $\chi^2/(n-k) > 1$ - which indicates that the variance is over dispersed. This implies that the model is not fitting with the Poisson and using NB could be a better option (as observed on this study for conflict model). For NB, the Chi-square test would include the overdispersion parameter "$\hat{\varphi} = \frac{1}{a}$" estimate and hence (Eq. 49) and (Eq. 50) can be presented as:

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - \widehat{\lambda_\iota})^2}{\widehat{\lambda_\iota}\,(1 + \hat{\varphi})} \qquad \textit{Eq. 51}$$

$$G^2 = 2\sum_{i=1}^{n} (y_i \, ln \, y_i - y_i ln\widehat{\lambda_\iota} - (y_i + \hat{\varphi}) ln(\frac{(y_i + \hat{\varphi})}{(\widehat{\lambda_\iota} + \hat{\varphi})} \qquad \textit{Eq. 52}$$

The log-likelihood ratio test (LLR-T) is an appropriate approach to compare the prior analysis with no explanatory variable and the posterior (since the posterior is somehow correlated to the prior with a likelihood function).The equation bellow show how to estimate the difference log-likelihood between the prior and posterior mode with LLR-T which represent the statistical significance of the new estimators added to the model.

$$\mathrm{LLR - T = -2(LL\ (Prior) - LL\ (Posterior))} \qquad \textit{Eq. 53}$$

Likelihood represents a probability with a negative value for log-likelihood (for instance a 0.1 likelihood would give a log-likelihood of -1) hence, the log-likelihood approaching 0 is the ideal. A pseudo$-R^2$ (such as McFadden's test which was an approximate to the R^2 in MNL) is proposed using the Pearson Chi-square for the Poisson model (Eq. 49):

$$R^2 = 1 - \frac{Pearson\ (Posterior)}{Peasrson\ (Prior)} \qquad \textit{Eq. 54}$$

To estimate the pseudo$-R^2$ in NB model, the overdispersion parameters from prior model ($\hat{\varphi}_0$) is compared to the overdispersion parameters of the posterior model ($\hat{\varphi}$) and is presented as:

$$R^2 = 1 - \frac{\hat{\varphi}_0 n - 1}{\hat{\varphi} n - k} \qquad \textit{Eq. 55}$$

Where "n" total number of observations and "k" is the total number of parameters in the model.

In Eq. 54 if we substitute the Pearson with log-likelihood, the ρ^2 statistics test can be done. Both R^2, ρ^2 values are between 0 and 1, with a higher value indicating a better performance.

AIC is another commonly employed test where performance of several model is compared. This test is not only considering the log-likelihood of the model (e.g., in LLR-T), but also adding a panelizing term for the number of independent variable or coefficients (β, α) that are employed. This will balance the goodness of fit versus the inclusion of variables in the model. Lower AIC means a better performance of the model [42].

$$\mathrm{AIC} = -2(\mathrm{LL} - k) \qquad \textit{Eq. 56}$$

Since the sample size remain the same between the models under investigation in this study, there was no need to explore other measure of fit such as Bayes information criterion (BIC) to associate the penalty term with "n".

3.5 Chapter Summary

Traffic data collection and analysis are critical components of every study on transportation system. The ability of any research to produce favourable outcomes, depends on the data quality, integrity, and consistency. Three data sources have been recognised to be necessary for this thesis to achieve its goals. Data from the empirical study was assembled with external data such as intersection geometry, historical vehicle counts, and environmental data. This chapter addressed the concern that a poorly designed experiment, relying solely on deriver reaction to signage, is prone to biased results. This chapter builds upon previous research efforts with an experiment design that could address some of those limitations. By utilizing the collected data from the experiment, a traffic microsimulation model was built, and the calibration process was outlined. Solution to a minimization objective function for TMCs using a genetic algorithm, on a simple study area network in the Montreal was presented. It was also discussed how to validate the calibration effort's results by comparing attributes of simulated and observed conflict. All of these efforts were made to ensure that the necessary data for both qualitative and quantitative statistical analysis was available for this research. The statistical analysis section of this chapter described several predictive modelling techniques such as regression analysis to determine the relationship between a dataset's dependent (i.e., compliance, collision, conflict) and independent variables. For discrete data analysis (such as compliance), the logistic regression model was rectified with respect to the need of the study. The application of the generalized linear models and in particular Poisson regression and its extension, negative binomial was also presented. Evaluation techniques between models was also discussed, allowing for a more in-depth review of the results in the following section of this thesis.

4 Chapter 4: Analysis and Results

4.1 Introduction

In the previous chapter, several sources of information were presented. It has been discussed that the essential data was acquired through observation from an experiment, inquiries from historical data, and extractions from a reliable traffic microsimulation. Several methodologies were presented in alignment with the objective of the research. The focus on this chapter is the link between the information available for the study and road safety measures for the treatments under investigation at SCI. The analysis can be classified into two categories:

1. Qualitative analysis for (a) variance in approaching vehicular speed (ANOVA) with respect to various countermeasures, (b) distribution of braking habits and (c) drivers' reaction at stop-line

2. Quantitative analysis with generalized linear models for collision and conflict estimates.

The performance of the models is assessed in the final part of this chapter. Next and final chapter of the thesis summarizes the results and concludes with some findings and recommendations pertinent to the study.

4.2 Drivers' Behavior

As previously indicated, the empirical testing was performed on two types of new and untested regulatory backlit LED based signs: a stop-sign and a Pedestrian crossing sign (Figure 9). This section summarises the data from each treatment's field research, as specified in the manual on uniform traffic control devices (MUTCD) as a critical step in evaluating a control device's performance.

4.2.1 Analysis of Speed

Speed was recommended as a major contributing surrogate collision measure in the previous studies assessing the effect of a treatment at SCI [8,9,67]. As a basic procedure, we used ANOVA as described in section 3.4.1 and a pairwise analysis was conducted between the observed variance and the average value of the approaching speeds. The HD radar was employed in multiple sessions to capture the speed of the approaching traffic at upstream, east bound of rue Victoria. The average speed and variance captured by the radar is summarised in Table 13.

Intersection & Victoria Ave.	Treatment type	Number of samples	Average speed (km/h)	Standard deviation (km/h)	95% Confidence Interval
21st Ave.	Flex panel	100	35.79	5.3	±1.05
25th Ave.	STOP (Standard)	100	36.31	4.61	±0.91
25th Ave.	BLS	100	32.82	4.28	±0.85
28th Ave.	LED	100	30.97	5.63	±1.12
33rd Ave.	Beacon	100	35.01	6.18	±1.22
Total		500	34.17	5.59	±0.5

Table 11 Approaching speed average and variance

As it is shown in graphs below, 75% percentile of the speed values for both LED and BLS signs are lower when compared to the other three treatments. The presence of the BLS signs also exhibits more uniform speed distribution.

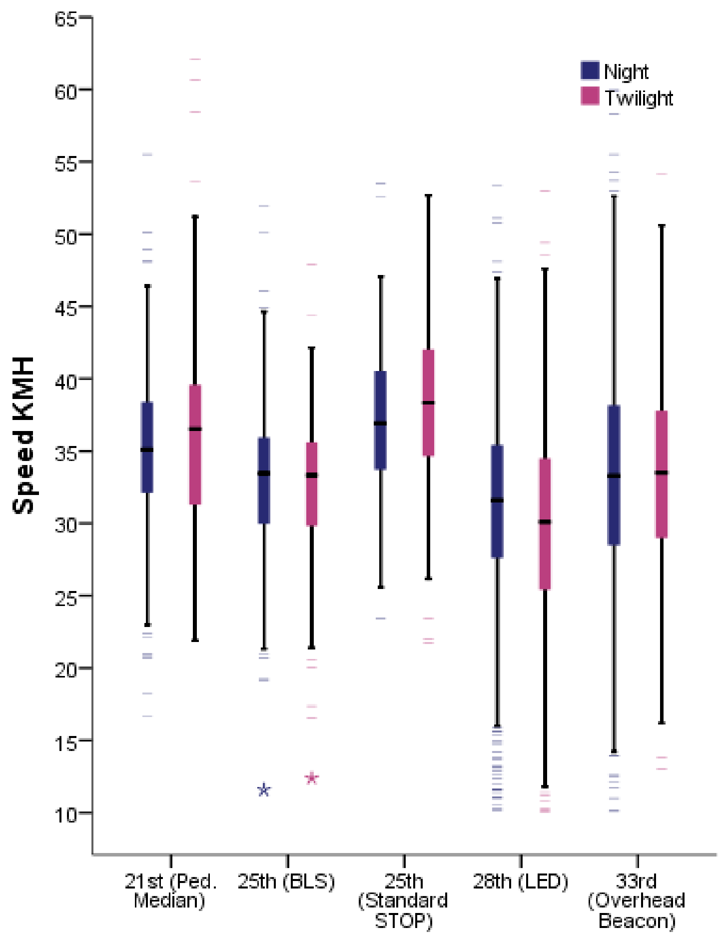


Figure 22 Boxplot speed comparison

Honestly significant difference Tukey-HSD test was applied as a post-hoc method for a pairwise contrasts analysis [128,129]. The assumption of homogeneity of the variance from Levene's test result was not statistically significant (0.08 at confidence level α=0.05). The significance *p-value* of ANOVA was 0.001, which is also less than 0.05 - meaning there is a statistically significant difference somewhere among the speed means on our dependant variable in the 5 sample groups (Table 11).

Levene Statistic	df1	df2	Sig.
.528	4	495	.080

Table 12 Test of Homogeneity of Variances

The multiple comparison contrast between groups is presented in Table 14. As expected, difference among groups occurred as the average speed is statistically significant from at an alpha level of 0.05. There was not statistically significant in comparison between BLS and LED. Statistically significant reduction in upstream speed mean values were observed only for BLS and LED signs (%3.5 ±1.8 and %5.3 ±2.0 respectively). The other two treatment did not show a significant improvement on speed patterns in comparison to the standard stop-sign. Another interesting finding was the insignificant effects of the overhead flashing Beacon in comparison to standard stop (Std. STOP) sign which was unexpectedly insignificant. Previous reports have shown significant improvement in safety performance for this countermeasure [52,92]. This is perhaps attributable to the fact that this test is conducted in urban area and lower minor volume at intersection 33rd .

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1973.58	4	493.39	17.94	.001
Within Groups	13611.84	495	27.50		
Total	15585.43	499			

Table 13 Analysis of Variance, ANOVA

Intersection Victoria Ave. & "*Minor (treatment type)*"		Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
21st Ave. *(Std. STOP)*	*25th (Std. STOP)*	-.590	.7416	.932	-2.621	1.440
	25th Ave. (BLS)	**2.898***	.7416	**.001**	.868	4.929
	28th Ave. (LED)	**4.747***	.7416	**.000**	2.717	6.777
	33rd Ave (Beacon)	.705	.7416	.877	-1.325	2.735
25th (Std. STOP)	*21st Ave. (Std. STOP)*	.590	.7416	.932	-1.440	2.621
	25th Ave. (BLS)	**3.489***	.7416	**.000**	1.458	5.519
	28th Ave. (LED)	**5.337***	.7416	**.000**	3.307	7.368
	33rd Ave (Beacon)	1.295	.7416	.406	-.735	3.326
25th Ave. (BLS)	*21st Ave. (Std. STOP)*	**-2.898***	.7416	**.001**	-4.929	-.868
	25th (STD. STOP)	**-3.489***	.7416	**.000**	-5.519	-1.458
	28th Ave. (LED)	1.849	.7416	.094	-.182	3.879
	33rd Ave (Beacon)	**-2.193***	.7416	**.027**	-4.224	-.163
28th Ave. (LED)	*21st Ave. (Std. STOP)*	**-4.747***	.7416	**.000**	-6.777	-2.717
	25th (STD. STOP)	**-5.337***	.7416	**.000**	-7.368	-3.307
	25th Ave. (BLS)	-1.8488	.7416	.094	-3.879	.182
	33rd Ave (Beacon)	**-4.0420***	.7416	**.000**	-6.072	-2.012
33rd Ave (Beacon)	*21st Ave. (Std. STOP)*	-.7050	.7416	.877	-2.735	1.325
	25th (STD. STOP)	-1.2953	.7416	.406	-3.326	.735
	25th Ave. (BLS)	**2.1932***	.7416	**.027**	.163	4.224
	28th Ave. (LED)	**4.0420***	.7416	**.000**	2.012	6.072

*. The mean difference is significant at the 0.05 level.

Table 14 Multiple comparison of the treatments

The study at 18th avenue and Victoria Avenue yielded a speed result that included both approaches (Figure 9). The eastbound approach had BLS, whereas the westbound approach had a standard stop-sign (i.e., with retroreflective surface).Table 15 presents the summary of the data Table 15 captured from over 1500 vehicles passing through the intersection. More samples were collected because, unlike the previous test, these are two different and independent test samples. In this case, an independent t-test was used to assess the variance between the two groups. Table 16 shows that homogeneity of variance is not significantly different (≥0.1), and that the *p-value* is near 0.05. Therefore, the null hypothesis (i.e., no difference between the means) is rejected, so it can be assumed that the difference in average speed between drivers exposed to the BLS and standard sign is statistically significant.

	Direction	N	Mean	Std. Deviation	Std. Error Mean
Avg. speed	Eastbound BLS	790	28.36	5.020	.179
	Westbound Standard	760	31.30	5.566	.202

Table 15 Speed mean data at 30 m upstream intersection 18th Avenue

	Levene's Test		t-test for Equality of Means							
	F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
					One-Sided p	Two-Sided p			Lower	Upper
Equal variances assumed	3.39	0.066	-10.93	1548	<.001	<.001	-2.939	0.269	-3.467	-2.412
Equal variances not assumed			-10.9	1517.6	<.001	<.001	-2.939	0.27	-3.468	-2.411

Table 16 Independent samples t-test

The average speed on the west approach was somewhat lower (2.94 km/h), while the links on both sides of the intersection were nearly identical. Another finding showed that on the west approach, the speed dispersion was slightly better. When drivers move toward the BLS, the speed distribution became more uniform. The variability (STD 0.55) was lower on the eastbound side than on the westbound side, implying that drivers approaching the BLS sign were reacting faster and in a milder and more even fashion.

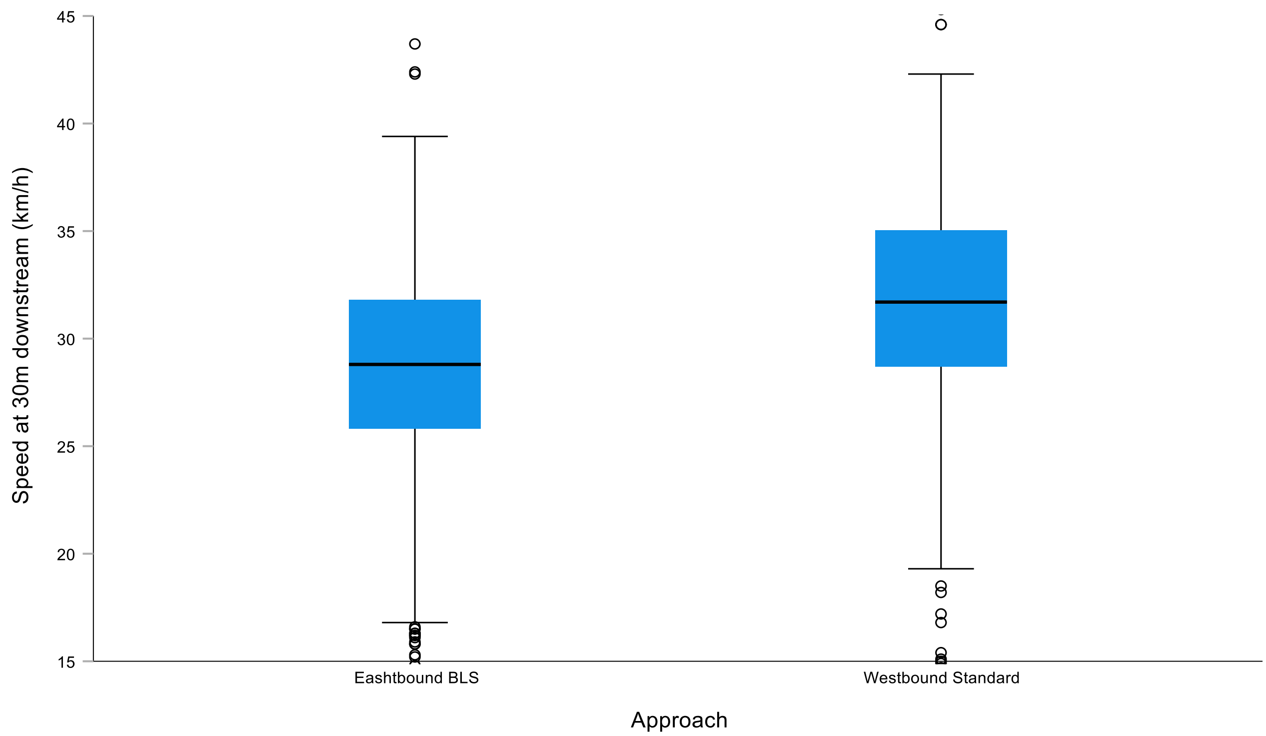
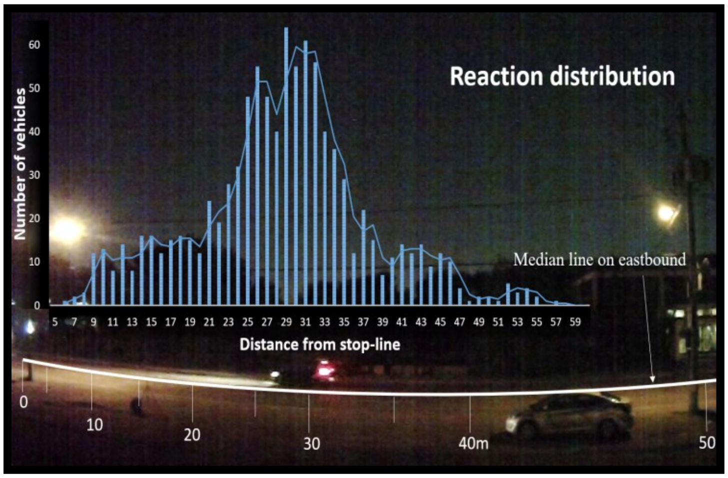
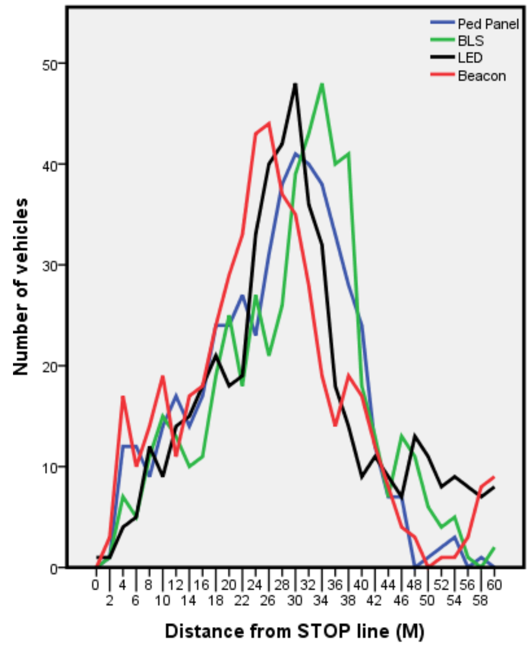


Table 17 Boxplot speed vs approach for intersection 18th Avenue and Victoria Avenue

4.2.2 Qualitative Analysis

An example of the driver's initial reaction points as they are approaching each crossing is shown in Figure 23(.a) . The reaction of the drivers to the downstream stop-sign was monitored from a fixed vantage location (by observing the vehicle's brake lamp), as described in section 3.2.2.1. Figure 23(.b) is the distribution profile of the detected drivers' reactions for all three signs. Even though drivers begin deceleration further from the stop-line when approaching an active signage, no marginal improvement was detected from this observation. The information presented below need to be carefully interpreted since the overall performance of the speed is a function of more factors to be considered than the treatment type as described in section 3.4.2.1 .





a. b.

Figure 23 Reaction points for drivers according to the brake signal

4.2.2.1 Stopping Compliance

The purpose of the statistical modelling described in section 3.4.2.2 is to measure the compliance of the drivers at the decision stop-line and evaluate if the distribution of three degrees of stopping compliance changes for different treatments, while controlling for weather conditions, presence of nighttime lighting, maneuver, and the presence of opposing traffic. The outcomes of the three scenarios discussed before are reported in this section.

4.2.2.1.1 Scenario 1

For scenario one, more than 6,000 vehicles passed through west-bound of the link. Only 1,156 vehicle samples satisfied the requirements needed for this scenario after excluding turning vehicles and those which were not in a sequence or whose compliance was unclear (i.e., pedestrian waiting at intersection causing confusion of potential conflict, outlier data from video analysis), whereas more than half were enrolled in the targeted treatment (BLS) intersection.

Table 18 summarizes the 3,468 observations recorded in the scenario one with two thirds of the drivers fully complying to the sign.

		N	Marginal Percentage
Driver Complacence	Blow-Through	143	4.1%
	Full Stop	2288	66.0%
	Roll-Through	1037	29.9%
Total		3468	100.0%

Table 18 Distribution of driver compliance in scenario one

Based on the likelihood ratio test, we may conclude that at least one population's slope of predictors is non-zero, and the model incorporating the full set of predictors represents a significant improvement in fit relative to the null model $[LR_{X^2}(14) = 350.168,\ p < .001]$.

	Model Fitting Criteria			Likelihood Ratio Tests		
Model	AIC	BIC	-2 Log-Likelihood	Chi-Square	df	Sig.
Intercept Only	1267.168	1279.471	1263.168			
Final	945.000	1043.422	913.000	350.168	14	<.001

The result from the other tests were in agreement with above and the McFadden Pseudo R-Square test. It can be concluded that the full model containing our predictors represents a 6.6% improvement in fit relative to the null model. The R^2 limitations for non-linear models such as NML was described in section 3.4.2.23.4.2.2.

The LR chi-square test didn't warrant the effect of "pavement surface condition (wet vs. dry)" on the final result since it had no statistically significant effect on the dependant variable (p> 0.05).

Table 19 illustrates the scenario one final model's results, which included effects from whether an approaching car yielded to opposing traffic, whether the approach took place at night, whether the driver turned and an interaction effect for yielding for four types of stop-signs.

Driver Compliance [a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B) Lower Bound	Upper Bound
Blow-Through	Intercept	-2.274	.165	190.349	1	<.001			
	Natural ambient light	-.673	.257	6.867	1	.009	.510	.309	.844
	Conflict potential	.641	.192	11.198	1	<.001	1.898	1.304	2.763
	Maneuver	-.462	.215	4.605	1	.032	.630	.413	.961
	LED STOP	-.672	.231	8.455	1	.004	.511	.324	.803
	BLS STOP	-1.713	.346	24.513	1	<.001	.180	.091	.355
	Beacon STOP	**-.419**	**.245**	**2.923**	**1**	**.087**	**.657**	**.407**	**1.063**
Roll-Through	Intercept	-.423	.076	30.665	1	<.001			
	Natural ambient light	-.560	.105	28.207	1	<.001	.571	.465	.702
	Conflict potential	.312	.089	12.289	1	<.001	1.367	1.148	1.628
	Maneuver	**-.065**	**.088**	**.551**	**1**	**.458**	**.937**	**.788**	**1.113**
	LED STOP	-.516	.101	25.921	1	<.001	.597	.489	.728
	BLS STOP	-.827	.116	51.221	1	<.001	.437	.349	.548
	Beacon STOP	-.431	.113	14.446	1	<.001	.650	.521	.812

a. The reference category is: Full-Stop.

Table 19 Scenario 1, MNL parameter estimates

From the dependant variable, 'full-stop' is the desired response. Hence, the other two categories were compared to this reference category in the analysis. The negative regression slopes (B) for enhanced signs warrants the improvement of the full-stop vs blow-through, and the full-stop vs roll-through in almost all instances from a treated site to a standard stop-sign. When a coefficient B is negative, then this indicates that with increasing values on a predictor {0, Standard stop to 1, Treated stop-sign}, the chance of being in the reference category (Full-stop) to non-reference category (Blow-Through and Roll-Through) is increasing, and vice-versa. For conflict potential with positive coefficient B with increasing value {0, non-Opposing Traffic: 1, Opposing Traffic} the risk of being in the non-reference category (Blow-Through and Roll-Through) is decreasing.

The coefficient β (*Exp(B)*) represents the predicted change in the log-odds of driver behavior falling under target group (blow through/roll through) relative to the baseline group which is the full stop. This means for every 1 unit increase of the independent variable, while everything else stays the same, the relative odd for dependent variable will change with respect to direction of the coefficient (i.e., increases if $\beta > 1$ or decreases if $\beta < 1$). For instance, motorists at intersection with LED sign are at lower risk to blow-through than full-stop (negative slope). However, the odds of full-stop to blow-through will decreases from LED stop-sign to normal stop-sign (1 unit increase from 0) for $exp^{0.511}$ times if all other variables remain constant.

Table 20 and Table 21 present the odd ratio for drivers' reaction as well as the percentage for each compliance category respectively.

Compliance Ratio	Treatment			
	STD	LED	BLS	BCN
$[\frac{FS}{RT}]$ **	1.64	2.74	3.04	2.26
$[\frac{FS}{BT}]$ **	10.81	20.40	27.64	***~~19.21*~~***
* Rejected in MNL analysis ** With and without the presence of opposing traffic or pedestrian crossing				

Table 20 Drivers' compliance odds with respect to treatment types under scenario 1

Compliance category		Treatment			
		STD	LED	BLS	BCN
Full stop	No opposing traffic	39.96%	56.53%	57.34%	48.45%
	With opposing traffic	18.76%	14.22%	15.91%	18.44%
	With and without Opposing traffic	***58.72%***	***70.75%***	***73.25%***	***66.90%***
Roll through		***35.85%***	***25.78%***	***24.10%***	***29.62%***
Blow through		***5.43%***	***3.47%***	***2.65%***	***~~3.48%*~~***
* Rejected in MNL analysis					

Table 21 Percentage of drivers falling into different categories under scenario 1

As mentioned in section 2.4.2, the performance of the enhanced traffic signage is better at nighttime. The risk of accident is also increases at intersection in the presence of opposing traffic with crossing path. Arguably the most interesting effect of the enhanced signage is when there is a risk for the traveling drivers with opposing and turning traffic. To better understand the magnitude of this effect, Table 22 displays what is the odds ratio between reference category and other two categories, under nighttime vs. daytime conditions.

Categories	Ambient light	Odd ratio			
		STOP	LED	BLS	BCN
Full-stop to Blow-through	Light	9.929	13.300	20.557	***18.167****
	Dark	10.143	31.001	30.544	***21.861****
Full-stop to Roll-through	Light	1.853	1.928	2.717	***1.964****
	Dark	2.536	3.263	3.290	***1.734****
* Rejected in MNL analysis					

Table 22 Drivers' compliance odds with respect to ambient light for different treatments

For instance, at nighttime, at an approach equipped with a regular stop-sign, with an incoming turning vehicle, the likelihood for a driver to make a full-stop is 10 times more than blowing through the intersection. This odd is ~3 times more in the same setting at a BLS or LED stop-sign. Another interesting finding is the performance of the BLS sign during daytime compared to the other treatments in commanding drivers' behavior. This was not a surprise since the sign was harder to miss in daytime.

The NML model performed well in predicting who would fall into the 'Full Stop' category (odd ration 96.7%). However, it underperformed overall (73.7%) as predictor of group membership across all three levels of the dependent variable (Appendix II).

4.2.2.1.2 Scenario 2

Vehicle tracking was no longer a selection criterion in scenario 2, which meant that all vehicles flowing through the link, regardless of their entrance/exit intersection were considered. Near 9,000 samples were selected from over 11,500 vehicles passed through the link during the experiment. The samples with ambiguous compliance were eliminated from the study.

Table 23 highlights the 9,097 observations recorded in the scenario two, which have about the same marginal percentage outcome as scenario one with 0.4% increase in both non-full-stop parameters categories and 1.8% (%66 vs %64.2 decrease if full stop category).

		N	Marginal Percentage
Driver Complacence	Blow-Through	405	4.5%
	Full Stop	5842	64.2%
	Roll-Through	2850	31.3%
Total		9097	100%

Table 23 Distribution of driver compliance in scenario two

The likelihood ratio test for the model is suggesting that incorporating the full set of predictors represents a significant improvement in fit relative to the null model [$LR_{X^2}(12) = 331.737,\ p < .001$].

Driver Compliance[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B) Lower Bound	Upper Bound
Blow-Through	Intercept	-2.347	.088	704.283	1	<.001			
	Natural ambient light	-.304	.114	7.171	1	.007	.738	.591	.922
	Conflict potential	-.302	.123	5.989	1	.014	.739	.581	.942
	Maneuver	1.737	.131	175.690	1	<.001	5.681	4.394	7.345
	LED STOP	-.974	.146	44.827	1	<.001	.377	.284	.502
	BLS STOP	-.716	.176	16.614	1	<.001	.489	.347	.690
	Beacon STOP	-.666	.154	18.634	1	<.001	.514	.380	.695
Roll-Through	Intercept	-.414	.041	101.265	1	<.001			
	Natural ambient light	-.196	.049	15.749	1	<.001	.822	.746	.906
	Conflict potential	***.064***	***.055***	***1.370***	***1***	***.242***	***1.067***	***.957***	***1.188***
	Maneuver	.401	.074	28.996	1	<.001	1.493	1.291	1.728
	LED STOP	-.538	.061	76.955	1	<.001	.584	.518	.658
	BLS STOP	-.684	.076	81.540	1	<.001	.505	.435	.586
	Beacon STOP	-.371	.067	30.531	1	<.001	.690	.605	.787

a. The reference category is: Full Stop.

Table 24 Scenario 2 MNL Parameter estimates

The McFadden Pseudo R-Square test indicated that the full model containing the selected predictors represents a 2.3% improvement in fit relative to the null model, which indicate less improvement in comparison to scenario one. And finally, the "pavement surface condition (wet vs dry)" had no statistically significant effect on the dependant variable ($p > 0.05$) and hence was removed from the final model (Appendix II section 8.1). Table 24 demonstrates the results from the model for scenario two, which included effects from whether an approaching car yielded to opposing traffic, whether the approach took place at night, whether the driver turned and an interaction effect for yielding for four types of stop-signs.

The reference category is 'full-stop' and the other two categories were compared to the reference in the analysis. The treatments coefficients (B) are showing negative regression slopes for all enhanced signs, which means improvement from reference category to the other categories which is in the same direction as results presented in Table 26. The log-odds β (*Exp(B)*) of driver behavior falling under target group (blow through/roll through) relative to the base group category increased in comparison to "Scenario one" for all odds. This means for every 1-unit increase (1 unit increase from 0) of the treatment type as independent variable, while everything else stays the same, the relative odd for dependent variable (compliance) will more likely changes on "Scenario two" in comparison to "Scenario one". For instance, the odds of, full-stop to blow-through, will decreases from LED stop-sign to normal stop-sign for 0.377 under "Scenario two" compared to 0.349 times in "Scenario one" (given all other variables remain constant). Figure 24 is trying to illustrate the error mean between different independent variables for both scenarios. As it shows the odds ratios in both scenarios for 95% of the data is within small range for the independent variable, treatment type, while the ambient light, opposing traffic and maneuver is covering a wider range.

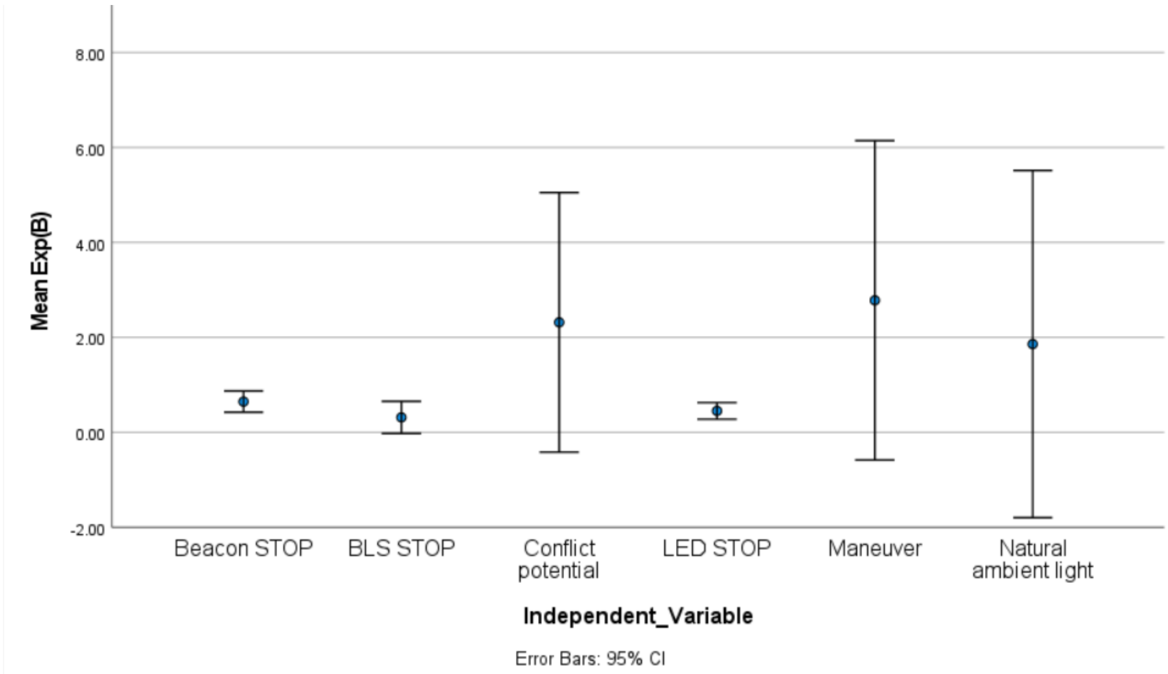


Figure 24 Simple error bar mean for Exp(B) both scenarios

Table 25 and Table 26 present the odd ratio for drivers' reaction as well as the percentage for each compliance category respectively. The percentage differences between each category **(δ)** in comparison to scenario 1 is showing increase pm all categories odd ratios in comparison to scenario one. This can be attributed to the fact that the overall FS vs non-FS slightly decreased by 1.8% under scenario two.

Compliance Ratio	**Treatment**			
	STD (δ)	**LED (δ)**	**BLS (δ)**	**BCN (δ)**
$[\frac{FS}{RT}]$ **	1.56 *(4.6%)*	2.52 *(8.2%)*	2.88 *(5.1%)*	2.20 *(2.6%)*
$[\frac{FS}{BT}]$ **	10.49 *(2.9%)*	18.14 *(11.1%)*	23.73 *(14.2%)*	~~**15.34 *(20.2%)***~~ *
* Rejected in MNL analysis ** With and without the presence of opposing traffic or pedestrian crossing				

Table 25 Drivers' compliance odds with respect to treatment types under scenario 2 and the delta with Scenario one (in parenthesis)

Compliance category		Treatment			
		STD (𝛿)	**LED (𝛿)**	**BLS (𝛿)**	**BCN (𝛿)**
Full-stop	No opposing traffic	40.96% *(1.01%)*	54.36% *(-2.17%)*	53.48% *(-3.86%)*	51.88% *(3.43%)*
	With opposing traffic	16.66% *(-2.10%)*	14.51% *(0.28%)*	18.52% (2.61%)	13.93% (-4.51%)
	With and without Opposing traffic	**57.62%** *(-1.10%)*	**68.87%** *(-1.89%)*	**72.00%** *(-1.25%)*	**65.81%** *(-1.09%)*
Roll through		**36.89%** *(1.04%)*	**27.34%** *(1.56%)*	**24.97%** (0.87%)	**29.90%** *(0.28%)*
Blow through		**5.49%** *(0.06%)*	**3.80%** *(0.33%)*	**3.03%** *(0.38%)*	~~4.29% (0.81%)**~~
* Rejected in MNL analysis					

Table 26 Percentage of drivers falling into different categories under scenario two and the delta with scenario one (in parenthesis)

To get some idea of the magnitude of this change, Figure 25 displays how the estimated odds for dependent variable varied over different possibilities for treatments.

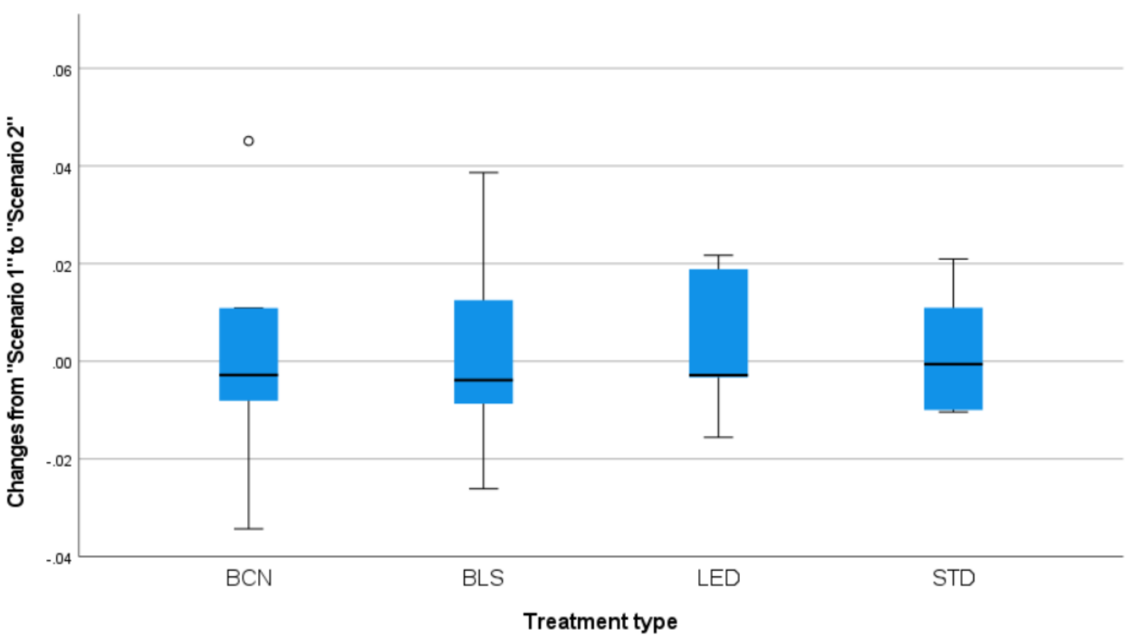


Figure 25 Boxplot of changes between two scenarios by Treatment type

Despite increase in odds of predicting category membership of the dependent variable in "scenario 2" arguably the performance of the model in predicting correctly a driver falling into the relative dependent variable ("Blow-through" and "Roll-through") was very poor. For instance, the model correctly predicted a driver falling into the 'Roll-through' group at a rate of 57/ (57+2793)*100% = 2%. The model correctly predicted a driver falling into the 'Full stop' at a rate of 98.2%. The overall classification accuracy for the model dropped significantly from 73.7% to 63.7% (Appendix II section 8.1).

4.2.2.1.3 Scenario 3

The "scenario 3" study involved 4504 motorists approaching the same intersection from two different directions on the same link (Ave. Victoria). Hence, each driver is most likely attending the experiment only once and motorists on the two approaches are not correlated. West approach is equipped with the BLS pedestrian crossing and East approach with conventional sign as described in section 3.2.2.3. Both signs are complemented with a stop-sign at the top. The geometry of the approaches is identical. The data was collected for 15 hours on 3 different sessions during the peak evening hours and after civic sunset. Previous studies have indicated that pedestrian crossings coupled with a stop-sign overall performs better on both vehicle speed and driver reaction[12,130] . The drivers' compliance at the stop line with and without traffic (pedestrians or motorists) for the two treatments (BLS and standard) is summarised in the table below.

			STD	BLS	Grand Total
Full-stop		**Total**	**65.49%**	**74.92%**	**71.71%**
	Opposing traffic	*Without*	*48.34%*	*54.12%*	*52.15%*
		With	*17.16%*	*20.80%*	*19.56%*
Roll- through		**Total**	**29.62%**	**22.92%**	**25.20%**
	Opposing traffic	*Without*	*19.50%*	*16.22%*	*17.34%*
		With	*10.11%*	*6.70%*	*7.86%*
Blow-through		**Total**	**4.89%**	**2.15%**	**3.09%**
	Opposing traffic	*Without*	*3.98%*	*1.88%*	*2.60%*
		With	*0.91%*	*0.27%*	*0.49%*

Table 27 Distribution of driver compliance in scenario three

The results from Multinomial logistic regression, determines if the distribution across the three stopping compliance categories varied with respect to the approaching driver requiring to yield to opposing traffic or the type of treatment they have been exposed to. Since there were only two types of treatment in scenario two, there was no need to introduce a dummy variable to allow a between group comparison. The ambient light and maneuver had no significant statistically effect on stopping compliance and hence both variables were excluded from the model.

Compliance [a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Roll-through	Intercept	-.849	.061	193.498	1	<.001			
	Treatment Type	-.390	.072	29.661	1	<.001	.677	.589	.779
	Conflict potential	.187	.076	6.151	1	.013	1.206	1.040	1.398
Blow-through	Intercept	-2.445	.126	376.247	1	<.001			
	Treatment Type	-.958	.175	30.088	1	<.001	.384	.273	.540
	Conflict potential	-.696	.236	8.672	1	.003	.499	.314	.792

a. The reference category is: Full-stop.

Table 28 Scenario 3 MNL Parameter estimates

Arguably, the response variable falling into “Full-stop” category increases when the explanatory variable “Treatment Type” is changing from Standard sign to BLS Ped-crossing (0 to 1).

The important finding from this experiment was the statistical significance of the model compared to the baseline or null (when all of the regression coefficients in the model are equal to zero). The Chi-Square test results $[LR_{X^2}(4) = 71.265,\ p < .001]$ was showing less statistical significance than the other two scenarios, while both Pearson and Deviance test failed (p=.091 and p=.090 respectively) this was the same issue observed in the before and after studies on the fit of the model. [33]

In other words, the χ^2statistic test is $[LR_{\chi^2}(3) = 53.384,\ p < .001]$, if there is in fact no effect of the predictor variables “Treatment Type”. Clearly the model performed well in predicting who would fall into the ‘Full-stop’ category. However, it performed poorly overall when predicting the category membership across all three levels of our dependent variable with 0% falling into the ‘Roll-through’ and ‘Blow-through’ category. Finally, the result from McFadden fitness test shows only 1.1% improvement with full set of predictors under Scenario 3.

4.3 Collisions and Conflict Estimations

The objective of the quantitative analysis in the present section is to determine the estimated expected number of rare events (e.g., collision), on a given intersection during the four focused intervals. In the section 3.4.3 the process of extracting historical data from the collision database and amending into a geospatial software was described. Poisson and Poisson-gamma models, maximizing their likelihood functions and test of goodness of fit was also presented.

The first part of this section gives an overview of the data being used in the GLM models, followed by the outlines for finding a vector (β) that best describes and models the collision probability functions (maximizes the probability of collision, given the parameters of the model). In the last part of this section, 'conflict' is being considered as the 'dependent variable' whereas traffic flow and drivers' behaviour are utilized as the predictors.

4.3.1 Data Description and Correlations

The following parameters was necessary for the analysis:

- **Collision:** For each collision data in Table 9, a variety of details is recorded in the NDCB. The number of total collision used for each time period including both 'property damage' and 'injury/fatal' collisions. This was due to the fact that there were too few injury collisions available to make an analysis only on the basis of crash severity and to author's knowledge there is no defined method to link the severity of the collision within conflict analysis domain in urban areas.

- **Conflict:** The process of performing and calibrating microsimulation models was described in the section 3.3. These efforts were to identify the frequency of the critical conflicts ($0 < TTC \leq 1.5$ s and $0 < PET \leq 5$ s) for each intersection scenario. Collisions were assumed to be highly probable when conflicts were characterized for having TTC and PET values falling within the range.

- **Average traffic volume:** Total traffic entering each intersection was observed during empirical study which varied from 800 to 2795 vehicles per period on Ave. Victoria and no-entry to 620 on the minor approaches and 15 to 219 pedestrians/cyclists. The turning traffic volumes (right turn, left turn, and straight on) entering each of the six intersections in the four periods were recorded and the TMC at intersections 19th and 34th was extracted from the previous studies by the city.

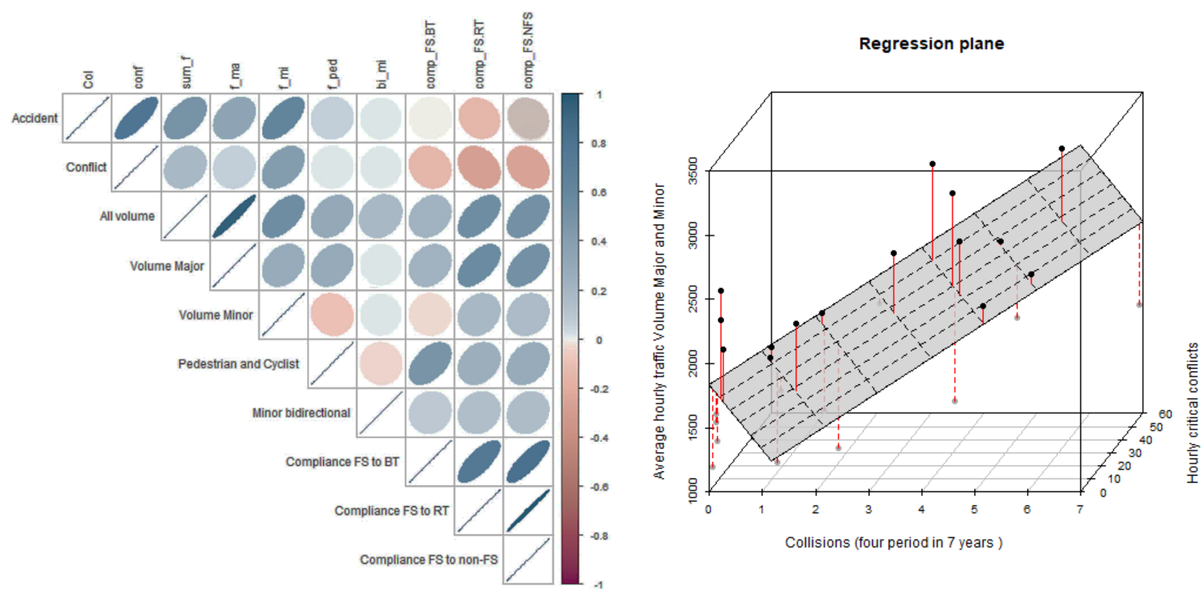
- **Compliance:** The three scenarios (described under section 4.2.2.1) combined have provided the necessary posterior information for the regression models in the quantitative analysis with respect to drivers' compliance to the signage. The odds for drivers' compliance to the treatment while there is a potential conflict with an opposing traffic was extracted for FS to RT, FS to BT and FS to non-FS for daytime and nighttime. The assumption for selecting this particular behaviour was that collision or conflict are more probable in the presence of opposing traffic.

The summary of the above-mentioned data is detailed in the Table 29 on the next page.

Treatment type	Bi-directional	Interval: peak-hours and ***Evening*** (7:30-8:30),(12:00-13:00), (16:30-17:30), ***(18:00-22:00)***					Drivers' behaviour odd ratio with respect to the ambient light when opposing traffic presented						
		7 years Collision	**Total critical conflict**	**Average traffic flow**			**Scenario**	$[\frac{FS}{RT}]$		$[\frac{FS}{BT}]$		$[\frac{FS}{non_{FS}}]$	
				$F_{ma}\ (\frac{veh}{h})$	$F_{mi}\ (\frac{veh}{h})$	$F_{Ped}\ (\frac{ped}{h})$		**Light**	**Dark**	**Light**	**Dark**	**Light**	**Dark**
BCN	N	(1),(0), (1), ***(0)***	(18),(8), (9),***(11)***	(1407),(1101), (1749),***(2170)***	(81),(83), (69),***(94)***	(116),(129), (219),***(190)***	2	24.33	30.50	2.47	2.98	2.25	2.71
BLS	Y	N/A	(47),(36), (28),***(14))***	(1609),(1140), (2091),***(2044)***	(310),(245), (388),***(441)***	(98),(75), (101),***(142)***	1	23.89	33.00	2.91	4.89	2.60	4.26
LED	Y	(2),(5), (6),***(4)***	(26),(41), (33),***(10)***	(2166),(1533), (2795),***(2778)***	(134),(374), (495),***(620)***	(34),(29), (52),***(54)***	1	16.22	28.00	2.50	4.20	2.17	3.65
STOP*	Y/N	(7),(7), (6),***(10)***	(132),(130), (137),***(101)***	(1371),(1069), (1824),***(2166)***	(358),(259), (329),***(346)***	(37),(41), (56),***(129)***	1	11.55	10.65	1.75	2.13	1.52	1.77
STOP-Ped	Y	(7),(4), (6),***(5)***	(57),(32), (48),***(25)***	(1504),(1020), (1973),***(2086)***	(274),(270), (382),***(475)***	(102),(88), (110),***(134)***	3	16.91	18.79	1.69	2.01	1.56	1.91
BLS-Ped	Y	N/A	(44),(28), (33),***(12***	(2081),(1378), (2237),***(2186)***	(597),(412), (589),***(425)***	(23),(51), (48),***(146)***	3	24.35	34.78	3.11	5.12	4.19	4.86

Table 29 Collison count, average conflict frequency from microsimulation, average traffic flow, compliance to the signage.

A significant correlation between recorded collisions and critical conflicts computed in microsimulation models was observed (0.792). This may not come as a surprise since when critical conflicts in the unit time increase, the expected number of collisions should logically increase too, and vice-versa. The traffic volumes also had a good level of conformity with both collision and conflict values (Figure 26-b). This was also consistent with previous studies (see section 2.3.1 for full discussion). Compliance on the other hand, had an inverse correlation with both collision and conflict, expecting an increase in compliance ratio decreases the frequency of both.



a. b.

Figure 26 a) correlation between all variables,(b) regression plane illustrating correlation between collisions at intersections with the traffic volume and critical conflict from the simulation

.

Figure 26-a illustrates the correlation and direction of all the covariates used in the GLM models. The correlation of variables is presented in the format of a matrix (horizonal variables are the codes used in rStudio while the vertical labels are the definition of that variable). The correlation of a variable with itself is indeed 1 and is displayed as a line while a low correlation is in the format of a pale circle. For

instance, variable conflict ($[conf]$) has a high direct correlation with f_{ma} and indirect correlation with $[\frac{FS}{non_{FS}}]$. Similar covariates may cancel out each other's effect if used in the same regression model, hence in the following two sections, several scenarios and models are shown, and their performance are discussed.

4.3.2 Collision Estimation

As mentioned in section 3.4.3.2, collision (Y) is considered "success" in Bernoulli trials, and Binomial distribution is the appropriate probability model that accounts for a set of Bernoulli trials. It was also shown that among the discrete distribution models, the binomial distribution can be approximated by a Poisson distribution or the extensions such as Poisson-gamma, zero-inflated and other widely used models to estimate collision (see section 2.3.1 for more detail). Furthermore, a suitable method to account for fluctuation of collision counts, which occur at a given intersection during a given unit of time (e.g., an interval in several years), is to assume that collision counts are random variables with the Poisson probability law. The 'expected' number of collisions per unit of time (λ) for the given intersection can then be estimated from Eq. 37. In section 3.4.3.3 it was mentioned that in some cases, collision counts appear to be "over-dispersed" or "under-dispersed" with respect to the theoretical variability consistent with the Poisson model ($\sigma 2 = \mathrm{VarY} = \lambda$). In order to account for this overdispersion, the expected number of accidents per unit of time in the Poisson model incorporates Gamma variate with $\mu = 1\ and\ \sigma^2 \neq 1$ that can help the variability of over-dispersion. Under this assumption, it was shown that the expected' number of collisions follows a Negative Binomial model, and the variance can be estimated from the Eq. 44, allowing for the variance of accident counts to be greater than the mean with utilization of the gamma parameter.

Before determining the vector (β) that captures the collision probability functions most reliably, the appropriate model has to be selected by investigating the signals of "over-dispersion and goodness of fit. The next section will describe a few models and check for their validity and performance in estimation of collision and its variance.

4.3.2.1 Naïve Model

The native model is basically the null model which is not utilizing any posterior information for the estimate. Hence the estimate is solely based on the recorded collisions variable as the prior data. The objective function in Eq. 43and Eq. 48 can be described as the maximizing log-likelihood function of Poison and NB Models for collision frequency at interactions under the study. By only maximization of the models with respect to the parameter μ, ultimately, the best approximator will return the average of collision frequency.

The following RStudio two function classes are being used to carry out the modeling hereafter:

- The function of GLM under the "Stat" Package for Poisson model
- the function of GLM.NB under the "MASS" Package for NB model

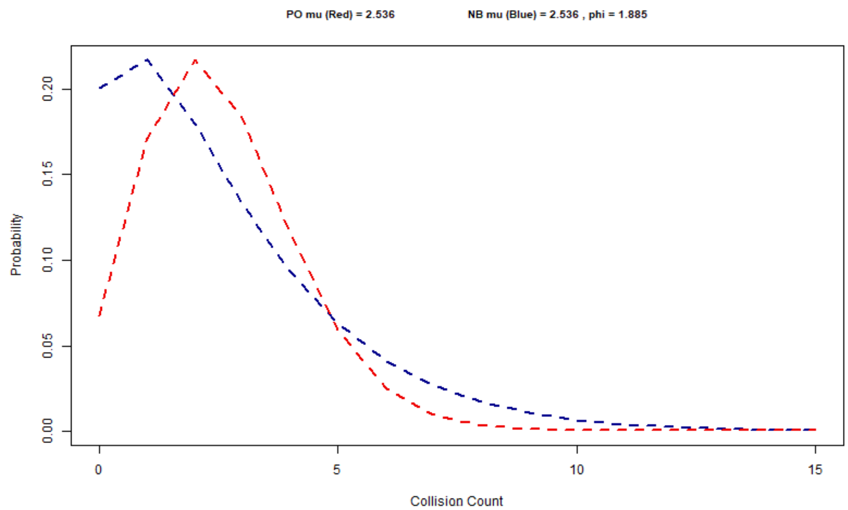


Figure 27 Poisson vs NB probability mass function

	Estimate β_0	Std. Error	z value	Pr(>|z|)	Degrees of freedom	Pearson Chi-Square	Over Dispersion Ratio	Log-Likelihood
Poisson	9.30E-01	0.119	7.84	4.49E-15	27	51.648	1.9129	-62.536
NB	9.30E-01	0.182	5.12	3.06E-07	26	22.024	0.847	-58.208

Table 30 NULL Poisson and NB regression models estimates for intercepts

Figure 27 shows a comparison between the PMF of the two null models and Table 30 provides the estimates for the coefficient of the intercept, which have significant p-values for both models. More especially, over-dispersion is obsevered in the Poisson model with a Pearson Chi-Square $[\chi^2(27) = 51.648\ at\ \ P > 0.5]$, hence the value for variance form Poisson can not be used. From the Eq. 39 the value Poisson estimate is:

$$\lambda_i = EXP(\beta_0) = EXP(\,0.9305) = 2.5357 \qquad \textit{Eq. 57}$$

These results indicate that for each intersection, the expected annual crash frequency is 2.536 for each interval which indeed is the average of the collision variable from all enteries "i". By using α (0.53), the variable of the NB model from Eq. 46 was found to be 2.829 which is somehow higher than the estimated variable in Poisson model from Eq. 57.

4.3.2.2 Alternative Models

The objective of alternative models is to estimate the expected number of collisions on a given intersection as a function of not only collision counts, but also traits such as traffic flow, geometric design variables, environmental conditions, drivers' behaviour etc. To put it another way, it is believed that collision frequency has a "systematic" component, and that explanatory variables may account for this non-random component. Subsequently, a regression model where the explanatory variables (and possibly combinations of them) act as "covariates" may provide a better estimate. Several alternative models are constructed and presented by including the posterior data into GLM. A traditional traffic volume-based collision-prediction model was first developed as the "Base model". This has allowed assessing the collision-prediction capabilities of traditional traffic volume-based model with the following alternative collision prediction models:

- Base model, Volume based (Eq. 58).
- Alternative 1, compliance-volume based (Eq. 59).
- Alternative 2, compliance-conflict-volume based (Eq. 60).
- Alternative 3, compliance-conflict based (Eq. 61).
- Alternative 4, conflict based (Eq. 62).

As mentioned in section 3.4.3.2, to obtain the estimates for vector β (β and α for NB), of the unknown parameters β, the likelihood function under the governing model (or equivalently its logarithm) is maximized. In Table 31 the estimates of the parameters of the five following regression models are presented.

$$\lambda_i = EXP(\beta_0 + \beta_1 \mathrm{D_v}).(f_{ma})^{\beta_2}.(f_{mi})^{\beta_3} \quad \textit{Eq. 58}$$

$$\lambda_i = EXP(\beta_0 + \beta_1 \mathrm{D_v} + \beta_2 FS_{BT} + \beta_3 FS_{RT}).(f_{ma})^{\beta_4}.(f_{mi})^{\beta_5} \quad \textit{Eq. 59}$$

$$\lambda_i = EXP(\beta_0 + \beta_1 \mathrm{D_v} + \beta_2 FS_{NFS}).(f_{ma})^{\beta_3}.(f_{mi})^{\beta_4}(conf)^{\beta_5} \quad \textit{Eq. 60}$$

$$\lambda_i = EXP(\beta_0 + \beta_1 FS_{BT} + \beta_2 FS_{NFS}).(conf)^{\beta_3} \quad \textit{Eq. 61}$$

$$\lambda_i = EXP(\beta_0).(conf)^{\beta_1} \quad \textit{Eq. 62}$$

The base model establishes a relationship between collision and traffic volumes. Furthermore, a dummy variable (D_v) has been considered which takes value ‘1’ for the evening time intervals and ‘0’ elsewhere. This was done to consider the fact that a larger “cumulative” traffic volume was used for the four hours evening time (18:00-22:00) compared to the lower volume during peak periods. The dummy variable was only used with models which incorporate flow covariate.

Variables	**BASE**				**Alternative 1**				**Alternative 2**				**Alternative 3**				**Alternative 4**			
	Coefficients	*std. Error*	*z-Value*	*p*	*Coefficients*	*std. Error*	*z-Value*	*P*	*Coefficients*	*std. Error*	*z-Value*	*p*	*Coefficients*	*std. Error*	*z-Value*	*p*	*Coefficients*	*std. Error*	*z-Value*	*p*
(Intercept)	-8.238	3.132	-2.63	**0.009**	-6.86	3.525	-1.946	**0.05**	-6.586	3.442	-1.913	0.06	-4.031	1.112	-3.624	**<0.001**	-2.123	0.661	-3.211	**0.001**
Dummy	-0.386	0.304	-1.268	**0.025**	0.745	0.304	2.453	**0.01**	0.309	0.342	0.902	0.37								
ln flow ma	0.873	0.43	2.031	**0.042**	0.847	0.504	1.682	0.09	0.476	0.496	0.961	0.34								
ln flow mi	0.518	0.143	3.634	**<0.001**	0.361	0.152	2.378	**0.02**	0.303	0.149	2.036	**0.04**								
ln conflict									0.644	0.269	2.397	**0.02**	1.172	0.226	5.194	**<0.001**	0.958	0.191	5.011	**<0.001**
Compliance FS_{BT}					0.029	0.036	0.793	0.43					-0.026	0.039	-0.669	0.503				
Compliance FS_{RT}					-0.618	0.369	-1.676	0.09												
Compliance $FS_{Non\text{-}FS}$									-0.069	0.341	-0.203	0.84	-0.874	0.417	-2.093	**0.036**				

Performance Metrics	BASE	Alternative 1	Alternative 2	Alternative 3	Alternative 4
Degree of Freedom	24	22	22	24	26
Deviance	32.067	23.746	18.282	28.618	34.146
Pearson Chi Square	28.4509	20.5838	15.6242	23.3293	28.8755
Over Dispersion Ratio	1.1855	0.9356	0.7102	0.9721	1.1106
Log Likelihood	-46.893	-42.732	-40	-45.168	-47.932
AIC	101.785	97.4641	91.9995	98.3357	99.8634

Table 31 Regression Model for collision as a function of traffic volumes, drivers' compliance, and total conflicts

The Pearson-Chi-square test (χ^2) for the base and alternative models was less than χ^2 [$p = 0.05$] with respect to their degree of freedom. For instance, this value for base model is between '0.1' and '0.05' [$\chi^2(24) = 36.415$ and $\chi^2(24) = 33.196$] which suggest statistically significant similarity between observations and estimations. Therefore, there is no strong reason to reject the hypothesis that the data follow a Poisson model. The ratio between Pearson and degree of freedom ($\chi^2/df \cong 1$) is also close to 1 which is an expected outcome from Poisson model.

When using the 'alternative 1' model it was found that the *p-value* of the traffic flow on the major approach (f_{ma}) and compliance ($comp$) covariates are between '0.05', and '0.1' level of significance. With somewhat less confidence than the base model, the alternative model may be carefully accepted. That means that even though the addition of the drivers' behavior and FS to non-FS compliance slightly improved performance of the null model in Eq. 58 however the model itself is less favorable.

The covariates of flow in the alternative model '2' on the other hand, were not found to be significant. This was expected due to the high correlation between traffic volumes and critical conflict and was reported in previous studies. That simply implies that the conflict together with the traffic volume is not contributing to the performance of the model. As the result, the Eq. 60 is not an acceptable alternative model to be proposed for safety performance function due to unacceptable parameters β for the flow on the major road (f_{ma}).

The last two alternative models imply not using the traffic flow but deploying conflict for posterior model. While 'alternative 3' have incorporated the drivers' reaction to the signage, 'alternative 4' is simply using critical conflict as the estimator for the collision estimation. The performance metrics of both models rejects the null hypothesis as there is statistically significant relationship between the observation and fitted values from the models.

While there have been some valid alternative models optional to the null model, however, their goodness of fit and performance, in line with section 3.4.3.4 with respect to null and base on one hand and among themselves will be assessed in the following section.

4.3.2.3 Performance Analysis of the Models

Three measures to assess the significance of the covariates entering the regression model are provided; Pseudo-R square (Eq. 54), likelihood ratio test (Eq. 53) and AIC with (Eq. 56). Even though the comparison of pseudo-R-square would be valid assumption since they are the same data, predicting the same outcome, they won't provide solid evidence for non-linear regression models such as Poisson and NB. Hence, in this situation, the higher pseudo-R-squared cautiously indicates which model better predicts the outcome.

		Null	Base	Alternative 1	Alternative 2	Alternative 3	Alternative 4
Log likelihood ratio test	**Null**	(0), 0, *[1]*	(3), 31.286, *[7.399e-07]*	(5), 39.607, *[1.792e-07]*	(5), 45.072, *[1.403e-08]*	(3), 34.736, *[1.385e-07]*	(1), 29.208, *[6.502e-08]*
	Base		(0), 0, *[1]*	(2), 8.321, *[1.56e-02]*	(2), 13.786, *[1.02e-03]*	(1), 3.4494, *[2.2e-16]*	(2), 2.0783, *[3.5372e-01]*
	Alternative 1			(0), 0, *[1]*	(0), 5.4646, *[2.2e-16]*	(2), 4.8716, *[8.7532e-02]*	(4), 10.399, *[3.421e-2]*
	Alternative 2				(0), 0, *[1]*	(2), 10.336, *[5.695e-03]*	(4), 15.864, *[3.207 e-03]*
	Alternative 3					(0), 0, *[1]*	(2), 5.5277, *[6.31e-02]*
	Alternative 4						(0), 0, *[1]*

	AIC	120.417	101.785	97.4641	91.9995	98.3357	99.8634
	Pseudo-R^2	0	0.149	0.384	0.533	0.302	0.136

p<0.001	p<0.01	p<0.05	p<0.1	p>0.1

Table 32 Comparison of the performance between null, base and alternative models. (DF), χ^2, *[p-value]*

LLR-T and AIC are commonly used to report performance of safety models with respect to the significance of the covariates entering the model as an alternative to R^2 as coefficient of determination. AIC has the advantage of considering the number of covariates into the account.

The performance of all the models is summarized in Table 32 with the following key findings:

- The LLR-T found a statistically significant difference between the null model and all five alternative models. Hence, the null model is being rejected compared to the base model and alternative models. This indicates the objectivity of introducing the covariates into the model to improve the estimates.

- Moreover, alternative 1 compared to the base model found moderate evidence against the null hypothesis in favor of the alternative. That implies that introducing the compliance covariates into the base model can contribute to a more reliable estimate. The AIC value suggests better performance of the alternative 1 model.

- Even though the conflict-based model (alternative 4) is showing a slight AIC improvement (99.9 vs 101.8), the LLR-T didn't find evidence against the null hypothesis. That suggests the similarity between these two models and the conflict-based model does not have a statistically significant advantage compared to the collision-based model.

- In this study, once compliance is factored into the conflict-based model (alternative 3), the performance of the model exceeds the classic volume-based models (base and alternative 1) with very strong evidence against the null hypothesis in favor of the alternative for base model and weaker evidence against the null hypothesis in favor of the alternative for alternative 1. This is an interesting finding in line with previous reports on the potential of the conflict-based collision prediction model and demonstrates the value of driver's behavior parameters in the regression models.

- The AIC value between the models suggests that the SPF models either volume based, or conflict based perform better when drivers' compliance is included as the covariates.

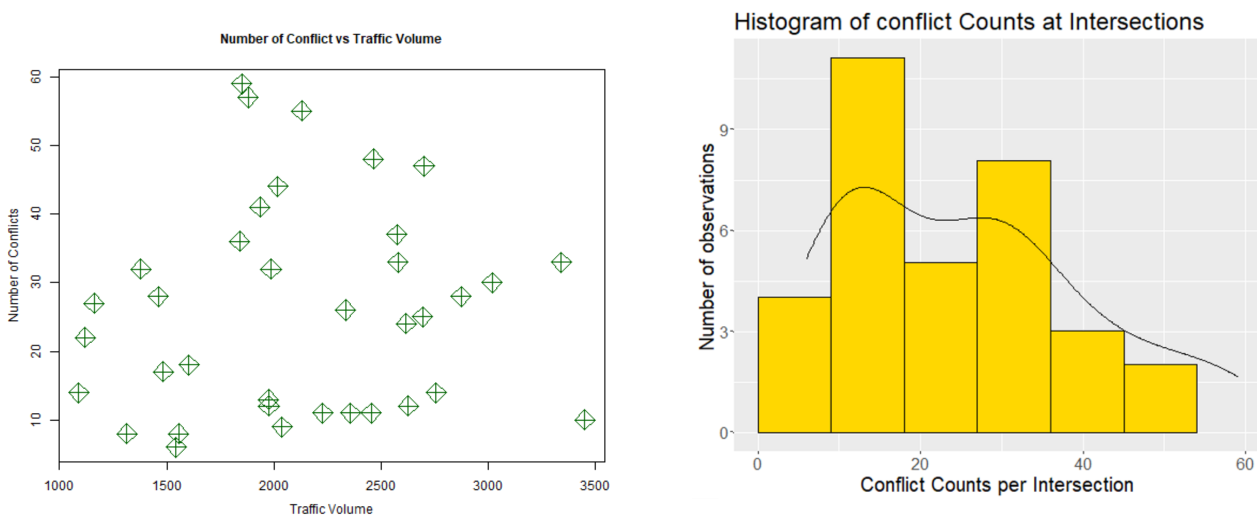
This quantitative study hence recommends Eq. 59 or Eq. 61(depending on the availability of either traffic volume or conflict, respectively) with the following estimates for parameters:

$$\lambda_i = EXP(-6.86 + 0.745 * \mathrm{d_v} + 0.029 * FS_{BT} - 0.618 * FS_{RT}).(f_{ma})^{0.847}.(f_{mi})^{0.361} \quad \textit{Eq. 63}$$

$$\lambda_i = EXP(-4.031 - 0.026FS_{BT} - 0.874FS_{NFS}).(conf)^{1.172} \quad \textit{Eq. 64}$$

4.3.3 Conflict Estimation

As mentioned in section 2.3.2, collision is assumed to be very probable when conflicts are characterized by a time to collision (TTC) and post-encroachment time (PET) below the threshold values of 1.5 and 5 seconds, respectively. In the absence of the collision data, conflicts are a considered to be reliable alternative for quantitative safety analysis. As mentioned in the section 3.3.2 a calibrated model, which is compatible with the applied microsimulation software, VISSIM, was used for identifying the number of critical conflicts at the intersections of the study. That included the critical conflict recorded for intersections equipped with 'new' (BLS) and already 'tested' (LED) treatments. In order to address the last key question of this research - *"to correlate and model road users' behavior and surrogate safety indicators for a new or tested treatment",* three conflict estimation statistical models are presented henceforth. If critical conflict frequency is considered as the prior estimate, the explanatory variables or possibly a combination of them could act as "covariates" and may provide a better conflict frequency estimate utilizing regression models. Figure 28 depicts the characteristics of the critical conflicts and their correlation with traffic volume (minor, major and pedestrian combined).



a. b.

Figure 28 (a) Hourly traffic volume vs critical conflict (b), regression plane illustrating correlation between collisions at intersections traffic volume and critical conflict from simulation.

4.3.3.1 Conflict Estimation Models

The null model is using the prior data or "observed" conflict to estimate the conflicts at each intersection for the period of "observation". Model A tends to use the observed drivers' compliance odds-ratios (Table 31) for predicting conflicts. Model B would seek to enhance the estimations by entering combined traffic volumes covariate into the previous models. All three models reported overdispersion ratio in Poisson model equal to 8.842, 6.685 and 6.456, respectively. This was an indicator of overdispersion due to heterogeneity and perhaps presence of unobserved covariates. However, these three models were fitted in Poisson/Gamma or Negative Binomial (NB) model as the scaled deviance values and the

Pearson values indicating goodness of fit for all three models at the 95% level. The over-dispersion ration in the three instances was reported 1.02, 1.17 and 1.19, respectively, indicating that data follows a Poisson/Gamma model.

	Model Null		**Model A (Conflict vs compliance)**		**Model B (Conflict vs compliance & flow)**	
Variables	*Coefficients*	*p*	*Coefficients*	*p*	*Coefficients*	*p*
(Intercept)	3.26	**<0.001**	4.318	**<0.001**	-1.799	0.364
Dum					0.603	**<0.001**
LED, BLS			0.959	**0.002**	1.053	**<0.001**
ln flow					0.811	**0.002**
comp FS BT			0.02	0.139	0.0352	0.073
comp FS RT			-0.744	**<0.001**	-0.973	**<0.001**
Performance Metrics						
Degree of Freedom	34		31		29	
alpha	0.305		0.199		0.11	
phi	3.274		5.018		9.06	
Deviance	37.437		36.9		37.77	
Pearson Chi Square	34.547		36.118		34.523	
Over Dispersion Ratio	1.016		1.165		1.19	
Log Likelihood	-145.114		-138.22		-130.479	
AIC	294.228		286.441		274.96	

Table 33 Regression Model for critical conflicts as a function of traffic volumes and drivers' compliance

The AIC on model B indicates a better performance compared to model A, and the same goes with the lower log likelihood of model A. The log-likelihood test between the two models also presents statistically significant outcome $\mathrm{LLR} - \mathrm{T}(2) = 15.483,\ p = 4.343 * 10^{-4}$ indicating that model B and A are different, with very strong evidence against the null hypothesis that these two models are the same. Hence the estimate of the conflict '$\widehat{Cf}$' for intersection 'i' for the time period 't'can be derived from equation below with the variance using Eq. 46 with $\alpha = 0.11$ from Table 33.

$$\widehat{Cf}_{i,t} = EXP(-1.799 + 0.603\mathrm{D_v} + 1.053\,BLS + 0.0352\,FS_{BT} - 0.973\,FS_{RT}).(f\ sum)^{0.811} \qquad \text{Eq. 65}$$

4.3.3.2 Conflict Reduction Rates

For a plausible candidate intersection to be upgraded to a flashing LED stop-sign or flashing BLS, the benefits associated with the sign would require to be assessed. Hence, a prediction of the collision reduction is essential for decision making (e.g., cost-benefit or cost-effectiveness analysis). As mentioned in the section 2.3.1 , CMF is widely used in road safety practice, and the FHWA under crash modification factor clearinghouse provides a searchable database along with guidance and resources[58]. As of this writing, there are two CMF reported for the LED stop-signs (see Figure 29) [9,33]. For instance, the approximate CMF multiplicative factor proposed by one of these studies is 0.59 $(\sigma = 0.18)$ with collision reduction factor (CRF) of approximately 41.5% (with 95% confidence intervals for the safety effectiveness 0%-70.6%)[9].

Compare	CMF	CRF(%)	Quality	Crash Type	Crash Severity	Area Type	Reference	Comments
☐	0.585	41.5	★★★★☆	Angle	All	Not specified	DAVIS ET AL., 2014	The authors used a hierarchical ... [*READ MORE*]
☐	0.59	41.1	★★★☆☆	Angle	All	Not specified	HUI XIONG AND GARY DAVIS, 2012	The number of crashes in ... [*READ MORE*]

Figure 29 Countermeasure: Replace standard stop-sign with flashing LED stop-sign (http://www.cmfclearinghouse.org/)

Although lack of collision data for BLS prevented the quantitative analysis to provide a CMF, however, Table 34 summarizes conflict as the surrogate safety measure performance for different treatments and the odd ratio of conflict at intersection with regular sign vs LED and BLS at the estimation point.

	Ped-Crossing		stop-sign	
	Day	**Night**	**Day**	**Night**
Regular	47.55(45.7), ±22.3	25.15(25), ±12.3	48.23 (50.33) ±22.6	20.1 (30.1) ±10.01
BLS	33.04(35), ±15.8	12.02(12) ±6.4	36.47(37.1) ±17.3	8.06(14.2) ±4.58
LED			35.65(33.33) ±16.9	10(15.8) ±5.46
Reduction rate (BLS)	***1.44***	***2.09***	***1.32***	***2.49***
Reduction rate (LED)			***1.35***	***2.01***

Table 34 Conflict reduction for regular, LED and BLS stop-signs- estimated hourly conflict (observed hourly conflict), variance

By way of explanation, naïve interpretation using the prior information, intersection equipped with BLS stop-sign would anticipate having an average of a conflict for every 1.3 and 2.5 conflicts of the regular stop-sign during daytime and nighttime respectively. That is, a 24.4% reduction in estimated conflict during daytime and 59.9% reduction during nighttime. The overall weighted conflict reduction for an intersection upgraded to BLS and LED would be 46.2% and 40.9% accordingly (considering 9h 17m average time between civic twilight start and end for month of April and July).

In order to include the effect of the 'observed variable', (which in this case is the frequency of critical conflict from the simulation model) and trade-off with 'expected value' from the negative binomial model, Bayesian approach is deployed. This approach is to update the parameter of the model on the bases of 'observation' estimator which indeed is the EB estimator and can be estimated using the Eq. 11 and Eq. 13 . Hence, the estimate from Table 34 is improved and updated and the results is presented in Table 35. The estimation provided for the 'regular' signage is based on intersections which did not convert to BLS. In order to predict the expected frequency of critical conflicts of the intersection in the after period, if it would have not been converted to BLS, Eq. 65 can be used with 'after' period traffic volume and compliance ratios for regular stop-sign. This is the first step in the EB before-and-after analysis and will calculate the value of μ in equation Eq. 5, which is equal to the variance for the assumed Poisson distribution.

	Ped-Crossing				Stop-Sign			
	Day		Night		Day		Night	
Signage Type:	EB_{θ}	$EB_{\theta_{var}}$	EB_{θ}	$EB_{\theta_{var}}$	EB_{θ}	$EB_{\theta_{var}}$	EB_{θ}	$EB_{\theta_{var}}$
Regular	47.24	39.43	25.11	18.43	48.55	41.14	22.41	17.23
Regular adjusted (μ)	50.54	50.54	24.64	24.64	47.05	47.05	20.56	20.56
BLS	33.44	26.57	12.01	6.84	36.59	29.41	10.45	6.38
LED					35.15	27.64	12.11	7.70

Table 35 EB critical conflict estimator for intersections using regular, regular adjusted with traffic volume, BLS and LED signage

As mentioned in section 2.3.1, there are two essential tasks for a before-and-after study in road safety studies; A 'prediction' of what would have been the safety of an entity in the period after had the treatment not been applied. And the 'estimation' on what the safety of the treated entity in the after period was. While the 'Regular adjusted' is providing the 'prediction', the $\boldsymbol{EB_{\theta}}$ estimates the frequency of the conflict for the intersection with BLS after adjustment to prior and posterior information and indeed is the estimate $\hat{\boldsymbol{\lambda}}$.

	Day							Night						
	$\hat{\boldsymbol{\delta}}$	$\hat{\boldsymbol{\theta}}$	**var ($\hat{\boldsymbol{\theta}}$)**	**SE ($\hat{\boldsymbol{\theta}}$)**	**Rate**	CI (95%) -	CI (95%) +	$\hat{\boldsymbol{\delta}}$	$\hat{\boldsymbol{\theta}}$	**var ($\hat{\boldsymbol{\theta}}$)**	**SE ($\hat{\boldsymbol{\theta}}$)**	**Rate**	CI (95%) -	CI (95%) +
Regular Ped crossing ~ BLS	17.10	0.65	0.02	0.13	35.1%	9%	62%	12.63	0.47	0.04	0.20	53.2%	13%	94%
Regular stop ~ BLS	10.46	0.76	0.02	0.16	23.8%	0%	55%	10.11	0.48	0.02	0.15	51.5%	21%	82%

Table 36 Potential conflict effects of converting from regular stop-sign to BLS stop-sign at 3-legs intersection (night and day)

Once 'prediction' and 'estimation' steps are completed, the impact of a treatment then can be assessed by comparing $\hat{\boldsymbol{\lambda}}$ and $\hat{\mu}$ and their variances. The Eq. 11 and Eq. 13 are being used to estimate the delta for reduction $\hat{\boldsymbol{\delta}}$ and the reduction factor $\hat{\boldsymbol{\theta}}$. The reduction factor can be used in Eq. 65 to estimate the conflict decrease of a target three-leg intersection with a known average hourly flow with conversion from a regular sign to BLS. Based on the results from Table 36 the EB method adjusted the reduction in the estimated conflict frequency for the daytime to an estimate point of 23.8% and 51.5% during nighttime. With 95% confidence interval, the worst and best expected conflict frequency improvement could be between 0% - 54.8% and 21.3%-81.8% for daytime and nighttime, respectively.

4.4 Chapter Summary

In this chapter qualitative and quantitative safety analysis for several treatments, including the underlining target treatment (BLS), at SCI was presented. The analysis demonstrated the ability of the methodology in chapter 3 to improve some safety estimations. The set-up of the empirical test allowed a significant (10%) improvement in categorical outcome prediction models for drivers' compliance. Three scenarios were presented in the qualitative analysis, by using a combination of the results, drivers' behavioural driver covariate was supplied. This covariate subsequently is used in the GLM models. Several safety performance functions for three-leg SCI were discussed and their performance was compared with respect to collision estimation. So far only a few studies have looked at conflict-based collision estimation models. But those studies did not consider drivers behaviour in their models and the 'observed' conflict was also deployed without consideration of random variability nature of conflict. The table below summarizes the key estimates for BLS which can be benchmarked against LED stop-sign with a known CMF.

	Day		**Night**	
	Full stop to non full stop	Conflict	Full stop to non full stop	Conflict
Regular stop-sign to BLS	41.3%	24.4%	58.3%	59.9%
Regular stop-sign to LED	29.8%	26.1%	51.4%	50.3%

Table 37 Compliance ratio and conflict frequency reduction estimations for LED and BLS stop-sign treatments.

In the before-and-after safety assessment of the BLS, empirical bayes method was deployed with traffic flow factors. The result suggests a significant improvement at nighttime with a modification factor 0.48 with 95% CI [0.33, 0.64] for stop BLS and 0.468 with 95% CI [0.26, 0.66] for pedestrian crossing. Although there was a slight improvement on nighttime estimation point for BLS in comparison to LED, no significant improvement was observed for the daytime.

5 Chapter 5: Conclusion and Recommendations

5.1 Summary:

To improve safety at stop-operated intersections, traffic engineers use traffic signages with flashing lights. There are few studies evaluating the safety impact and the performance of these signages to the road users. On the other hand, MUTCD guidelines require the safety performance of any 'new and untested' signages to be assessed. Two new treatments with flashing lights for stop-controlled intersections: backlit flashing stop-signs, and pedestrian crossing signs are evaluated in the study. Although many analysis tools have been developed and applied in a wide variety of areas, the focus of the research beforehand was placed on statistical models tailored specifically for the characteristics of critical conflict at intersections with various stop-control devices and drivers' reactions to these signs.

For drivers' behaviour analysis, the influence made by unobserved covariates in the statistical models was controlled either by including them into the model or eliminating them in a tailored empirical study. It was shown in this thesis that the outcomes from the before-and-after analysis in prediction models with categorical outcome is not reliable for all results. Meanwhile, using the same sample-set in experiments warrant confidence in certain results. As for conflicts analysis, in line with the literature, collisions at unsignalized intersections are correlating better with conflicts than traffic-flow. Hence, they can be utilized in collision safety performance modelling [66,81].

To calculate the critical conflicts, microsimulation models were set up, and the models' parameters were captured from the Genetic Algorithm utilizing heuristic methods. The MSE, between model estimates and field measurements for TMC's, was used and ultimately the conflict data was validated with field measures. Several collision prediction models were presented and compared. It is shown that accident estimations obtained from conflicts and compliance models had a better performance in comparison to classic traffic flow-based models. The results show that reasonable collision estimates are obtained by using the conflicts and compliance rather than traffic flow, while the estimated collisions are in the range of 95% Poisson confidence intervals of the observed collisions. The estimates (fitted values) from the proposed model compared to the actual collision recorded also conformed better when conflict-compliance were the covariates, as shown in the Figure 30. The performance of the traffic flow-based model was shown to improve slightly by introducing drivers' compliance in the safety performance function.

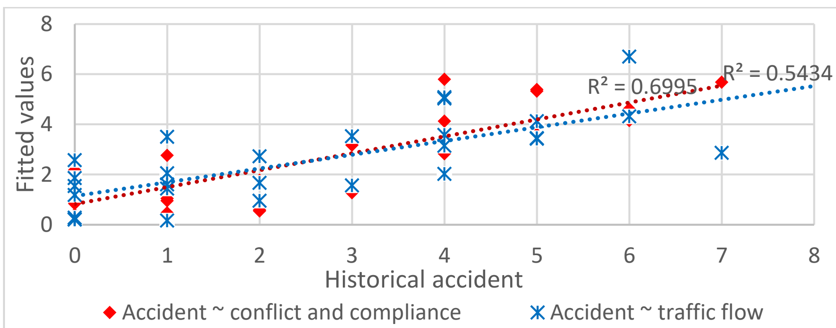


Figure 30 Fitted values from the models (estimate) vs recorded accident

Specifically, the AIC as the goodness-of-fit for the model with 'critical conflict only' improved to 99.9 compared to 'traffic flow only' with an AIC 101.9. This minor improvement has been reported in the previous studies [66,81]. However, the Log-Likelihood ratio test indicates strong similarity between these two models, and hence there was not enough evidence to promote 'critical conflict only' model compared to 'traffic flow only'. Lastly, the estimated conflict frequency reduction in daytime and nighttime was presented to be 23.8% and 51.5%, respectively, for the backlit flashing stop-sign. Similarly, the backlit pedestrian crossing sign showed 35.1% and 53.2% improvements for daytime and nighttime, respectively. To our knowledge, there are no examples in the literature on conflict-compliance models for unsignalized intersections, and no safety performance indicators are available for the backlit regulatory signage. Although this research represents a more robust verification approach for assessing safety at SCI using critical conflicts and compliance as a surrogate safety measure, it is yet premature to make a generalization that observed collision would be reduced with the same rate as reported conflict reduction.

5.2 Research Contributions

This work of research focused on the important topic of road safety assessment and active road signage with the goal to small contribution to the knowledge of safety as sustainability measure. Contributions were made toward design a methodology on road safety analysis which is more focused on drivers' reactions and interactions with active road signage. Some inconsistencies and ambiguity in performance of active road signage has been reported in the literature.

This study reaffirms the positive impact of active signage specially at nighttime. A step forward from the previous list of assessed active signage was made and backlit sings performance was assessed in this study. The proposed test methodology warrants less risk to public since the observation period was rather short.

Another contribution was to demonstrate collision-prediction models which deployed drivers' compliance in estimation and showing the added value of this predictor in the models. While past research had their attention on rural or sub-urban regions, this investigation was carried out in an urban setting.

5.2.1 Design an Unbiased Test

In the qualitative analysis of the active signages, the design of the experiment was unbiased for the following reasons:

1. **Same sample was being exposed to minimum 3 intersections:** This was to somewhat control the variation in human factors. The previous studies in the before and after setting were using different sample type and size entering the intersections. It was shown that for models using same samples they better passed the likelihood tests.

2. **Same road condition applied in each intersection:** For instance, if the road condition for the first intersection was 'wet', all the consecutive intersections for that sample would have been the same knowing the driver will exit them withing five minutes. Beside that, the pavement surface condition, wet/dry, was considered in the analysis and was insignificant after all.

3. **Ambient light was considered:** For an estimation of collision in Phoenix with less average nighttime-travel, using the CMF (based on the study in Minnesota) might not work. Inclusion of the odd likelihood compliance ratio on the other hand, for day and night would provide a more reliable SPF. It was shown in this research that the performance of the active signage decreases significantly during the daytime.

4. **Critical reaction was considered:** The odd ratios provided from this study was filtered for only the impactful maneuvers with consideration of the potential conflict with presence of opposing traffic. The method found significant correlation between drivers' compliance and conflicts.

5. **Active signage test in the city:** The light pollution is usually higher in the urban areas. A driver that missed to read or recognize a sign may not have deliberately violated the command, but simply missed the signages and that might increase when the conspicuity properties is less a reason of the failure. This study presents the urban area results without generalizing results for all urban intersections.

6. **Using intersections with similar properties:** In a cross-reference analysis, not only different samples are being used, but also there is a risk for an unobserved variable from undetected differences on the road properties. Same geometry and major links in the design of the experiment, reduces this variability.

5.2.2 Contributions in Modeling

A simulation model was calibrated, and several statistical models are presented in this study.

1. **A properly calibrated simulation model:** genetic algorithm was developed to carry out the calibration of the microsimulation model to extract critical conflict.
    a. The algorithm can thus be utilised by other researchers trying to calibrate other road networks based on compliance profiles.
    b. In section 3.3.2 it was shown how to set up a model to overcome limitations of reproducing drivers' compliance to signage at intersections with respect to different control devices.
    c. Since there is no direct output for conflict from the microsimulation, a process of validation based on limited observed conflict was presented and it was shown that using MSE, between model estimates and field measurements for TMCs is an appropriate approach, while compliance profile is given, in the calibration process.

2. **Better properties of the estimation models:**

    a. In chapter 4, the MNL models, although the design of the test was controlling some of the observed and potentially unobserved variables, several independent variables was deployed to improve the models' performance. A 10% improvement was shown between the MNL models.

    b. Conflict-compliance models for unsignalized intersections was presented and the SPF performances was compared. It was shown that the performance of this model was better than conflict or traffic flow only models.

    c. It was shown in this study that using the compliance properties as covariate, improves the estimation of the classic traffic volume-based model. Other researchers focusing on assessment of an intersection equipped with active signage system could use the outcome of the model from this research to inform their studies so that if drivers' compliance enters the regression model when traffic volume is already present it does significantly contribute toward reducing unexplained variability.

    d. This study didn't find evidence to accept previous assumption that estimations from a conflict-based model is necessarily better than volume-based models.

3. **Commentary on Performance Assessment of BLS**

Two types of regulatory BLS signs were assessed for the first time with following findings:

a. An EB method was carried out to find the estimated conflict for the before-and-after installation, where conflict frequency was adjusted based on the traffic volume and compliance profile. The estimations show installation of BLS would reduce conflict frequency as outlined in section 4.3.3.2.

b. One of the objectives of this research was to avoid ‘potential’ risk of installation of an equipment with unknown performance properties. Hence fortunately, there was no observed collision within the short period of installation of BLS. However, it was shown that the performance of BLS and LED signage on conflict analysis are alike with statistically insignificant improvement in nighttime performance for BLS. That cautiously infers the best possible CMF to be deployed for BLS signage would be the only available CMF for LED sign provided by FHWA. IT was shown in section 4.3.3 compliance covariate will improve the SPF and account for BLS effects.

c. While BLS performance was the focus of the evaluation process, the results from conflict estimation in section 4.3 reassures the better performance of both BLS and LED signs.

The presented methodology and reported results expanded the understanding of performance of the active signage systems, contribution was made to improve some of the modeling for better predictions and estimations, and finally, this was a validation on the positive impact of these type of control devices on road safety. The findings of this thesis open the door to a new scenario of research based on conflicts and compliance that is worth investigating specially in the absence of accident data.

5.2.3 Active Signage Performance

- While all types of active signage have shown improvement for approaching vehicular speed profiles, there was no statistically significant difference between BLS and LED. On the average speed point between the active signs versus the regular stop-sign, the speed was shown to be reduced by about $3.5km/h, 5.3km/h$ and $1.3km/h$ for BLS, LED and overhead beacon, respectively, on a 40 km/h urban link.

- For the pedestrian crossing signs, there was also a statistically significant speed reduction from 31.3 km/h to 28.4 Km/h, with better and more uniform approaches of vehicles to the BLS approach.

- For both LED and BLS, the findings of the qualitative study showed a statistically significant improvement in the drivers' full-stop to non-full-stop odds ratio. The overhead beacon couldn't statistically stand for this outcome. The full-stop to blow-through odds ratio for drivers with opposing traffic at night was improved three times (from 10 to 30) for LED and BLS. As for the roll-through behaviour, under the same condition, drivers made 24.2% more full-stop at approaches with LED or BLS than they did with the regular stop-sign (from 2.5 to 3.3).

- In the preceding situation, drivers approaching BLS made more full-stops during the daylight. BLS improved LED performance 35.1% for blow-through and 29.1% for roll-through. Both active signages performed better than regular stop-sign during daytime, while this was not as substantial as at night.

- The quantitative analysis for nighttime shown, the hourly estimated conflict frequency for BLS and LED are 35.2 ($\sigma = 5.9$) and 36.6 ($\sigma = 6.0$), respectively, while the projected frequency for regular stop-sing is 47.1 ($\sigma = 6.9$).

- The conversion of a regular stop-sign to BLS is anticipated to reduce the hourly conflict frequency between 0% - 54.8% and 21.3% - 81.8% for daytime and nighttime, respectively. The estimate points for daytime and nighttime are 23.8% and 51.5%, respectively.

- There was no collision data for BLS to estimate the safety performance for this type of signage. However, there was enough statistically significant evidence in similarity between LED and BLS on their performance with regard to conflict and drivers' behaviour. With an assumption which was made by many studies before about conflict and collision correlation, it can be cautiously concluded that the BLS is somewhat at the same performance as LED the sign for collision reduction with CMF of 0.585 and %41.5 CRF.

5.3 Research Limitation

The following are the key limitations of the proposed methodology and the outcomes:

1. Collision is the globally recognized road safety measure. Like many other non-collision-based studies, this limitation became the motive to look at other more frequent measures such as drivers' behaviour and critical conflict. Hence, the absence of a complete 'recorded' collisions data set is a major limitation in the safety assessment for BLS.

2. This research was based on the total collision numbers only so that the severity of collision has not been investigated due to data limitations and inability to conflict analysis in recognition of injury severity classification.

3. Unavailability of reference sites for BLS was another limitation that prevented this study to conduct before-and-after studies with traffic flow factors, rather than study with comparison group.

4. Calibration process for microsimulation was not against observed conflicts, this limitation is due to inability of traffic microsimulation platforms, which do not provide conflict estimates as a direct output of a model. It is imperative to automate the calibration process for future research, if this ability becomes available for iterative method using conflict as an output.

5.4 Recommendation and Future Work

a) Developing an API within the microsimulation that can calculate conflict will make it possible to remove the validation process described in section 3.3 and calibrate the model directly against conflict.

b) A larger simulation model can benefit from some improvement in the objective function. By using a multi-objective optimization function that calibrates to counts, speed and conflict information it will provide a reliable model.

c) This study used a subset of collision data to avoid limitations imposed in zero inflated model. In the future studies, if the collision frequencies are significant enough, the study can include more periods rather than the three peak periods and limited nighttime.

d) Use compliance in different settings of SPF with other predictors that may have correlation adding Performance of the signs with higher conspicuity is higher at nigh time. This is something that was missing in the previous studies and perhaps in future studies would be important to adjust the crash modification factor based on this variance.

e) Use of other drivers' behavior properties (e.g., deceleration rates, approaching speed) may offer great assistance to road designers and/or traffic operators in making preventive decisions based on robust estimation model.

f) This study was limited to only two regulatory signs. Investigation of other regulatory or warning signs would bring more insight to the performance of these active road signage. For instance, sharp curves on rural roads for can be considered as another infrastructure which benefits from the backlit chevron signs. At the time of writing this thesis there are no CMF is reported for this type of LED signages.

g) New and growing data have become more and more available. Much more data is captured through autonomous and connected technologies. This significantly greater information give advantage than that of traditional data sources which is basically relying of incomplete collision data. The use of data mining and machine learning techniques might be an alternative to the traditional road safety analysis used in this thesis, given the massive and complicated datasets expected to be available.

■

6 References

1. Groeger, J. A. (2011). How Many E's in Road Safety? *Handbook of Traffic Psychology*, 3–12. https://doi.org/10.1016/B978-0-12-381984-0.10001-3

2. Treat, J. R., Tumbas, N. S., McDonald S. T., Shinar, D., & Hume, R. D. (1977). Tri-level Study of the Causes of Traffic Accidents. Volume 1, Causal Factor Tabulations and Assessments. *DOT-HS-805-085;DOT-HS-034-3-535-77-TAC(1);NTIS-PB80121056;* DOI : https://doi.org/10.21949/1525022

3. *HSIS :: Data Request*. (n.d.). Retrieved August 16, 2020, from https://www.hsisinfo.org/datarequest.cfm

4. *AASHTO GREEN BOOK (GDHS-7) - A Policy on Geometric Design of Highways and Streets, 7th Edition: AASHTO Green Book*. (2018). American Association of State Highway and Transportation Officials.

5. Canada, T. (2018). *Canadian Motor Vehicle Traffic Collision Statistics: 2017*. Government of Canada. https://www.tc.gc.ca/eng/motorvehiclesafety/canadian-motor-vehicle-traffic-collision-statistics-2017.html

6. Federal Highway Administrator. (2012). *Manual on Uniform Traffic Control Devices (MUTCD)* (Issue May).

7. AASHTO. (2005). *AASHTO Strategic Highway Safety Plan (SHSP)*.

8. Gates, T., Carlson, P., & Hawkins, H. (2004). Field Evaluations of Warning and Regulatory Signs with Enhanced Conspicuity Properties. *Transportation Research Record*, *1862*(04), 64–76. https://doi.org/10.3141/1862-08

9. Davis, G. a, Hourdos, J., & Xiong, H. (2014). *Estimating the Crash Reduction and Vehicle Dynamics E ects of Flashing LED Stop Signs* (Issue January).

10. Hauer, E., & Garder, P. (1986). Research into the validity of the traffic conflicts technique. *Accident Analysis and Prevention*, *18*(6), 471–481. https://doi.org/10.1016/0001-4575(86)90020-5

11. Wang, C., & Stamatiadis, N. (2013). Surrogate Safety Measure for Simulation-Based Conflict Study. *Transportation Research Record: Journal of the Transportation Research Board*, *2386*(1), 72–80. https://doi.org/10.3141/2386-09

12. Goswamy, A., Hallmark, S., Litteral, T., & Pawlovich, M. (2018). Safety evaluation of destination lighting treatment at stop controlled cross-intersections. *Transportation Research Record*, *2672*(16), 113–121. https://doi.org/10.1177/0361198118774747

13. Tageldin, A., Sayed, T., Zaki, M. H., & Azab, M. (2014). A safety evaluation of an Adaptive Traffic Signal Control system using Computer Vision. *Advances in Transportation Studies*, *2*(SPECIAL ISSUE), 83–96. https://doi.org/10.4399/97888548735379

14. Gates, T. J., Hawkins, H. G., Chrysler, S. T., Paul, J., Holick, A. J., & Spiegelman, C. H. (2003). *Traffic Operational Impacts of Higher-Conspicuity Sign Materials*. *7*(2).

15. Highways, M. (n.d.). NCHRP Web-Only Document 126 - Methodology to Predict the Safety Performance of Rural Multilane Highways. *Development*, *February 2008*.

16. Schieber, F. (1987). Optimizing the Legibility of Symbol Highway Signs by 2D-Fourier Description of Images. In *Vision in Vehicles VI* (p. 163).

17. Awadallah, F. (2009). Intersection sight distance analysis and guidelines. *Transport Policy*, *16*(4), 143–150. https://doi.org/10.1016/j.tranpol.2009.04.001

18. Altamira, A. L., Marcet, J. E., Graffigna, A. B., & Gómez, A. M. (2010). Assessing available sight distance: an indirect tool to evaluate geometric design consistency. *4th International Symposium on Highway Geometric Design*, *1109*(5400), 1–23. http://www.4ishgd.valencia.upv.es/index_archivos/54.pdf

19. Montella, A., & Mauriello, F. (2012). Procedure for ranking unsignalized rural intersections for safety improvement. *Transportation Research Record*, *2318*, 75–82. https://doi.org/10.3141/2318-09

20. Freedman, M. (1985). TRAFFIC SIGNAL BRIGHTNESS REQUIREMENTS FOR THE NIGHTTIME DRIVING ENVIRONMENT. *Transportation Research Circular*, *297*, 22–23.

21. Gettman, D., Pu, L., Sayed, T., & Shelby, S. G. (2008). *Surrogate Safety Assessment Model and Validation : Final Report* (Issue June).

22. ITE. (2005). *Vehicle Traffic Control Signal Heads : Light Emitting Diode (LED) Circular Signal Supplement*.

23. Mazzotta, F., Vignali, V., & Irali, F. (2014). *Evaluation of the readability of road signs and roadside elements using Mobile Eye tracking device*. *5*. https://doi.org/10.6092/issn.2036-1602/5058

24. AASHTO. (2004). *a policy on Geometric Design of Highways and Streets*.

25. American Association of State Highway and Transportation Officials. (2009). AASHTO transportation glossary. LK - https://concordiauniversity.on.worldcat.org/oclc/427552642. In *AASHTO* (4th ed. NV). American Association of State Highway and Transportation Officials. http://app.knovel.com/hotlink/toc/id:kpAASHTOTD/aashto-transportation-glossary

26. Harwood, D., Mason, J., & Brydia, R. (2000). Sight Distance for Stop-Controlled Intersections Based on Gap Acceptance. *Transportation Research Record*, *1701*, 32–41. https://doi.org/10.3141/1701-05

27. Schnell, T., Aktan, F., & Li, C. (2004). Traffic Sign Luminance Requirements of Nighttime Drivers for Symbolic Signs. *Transportation Research Record*, *1862*(1862), 24–35. https://doi.org/10.3141/1862-04

28. Jones, K., Wood, J., & Bentley, B. (2012). A technique for on-road assessment of road sign visibility distances. In *Proceedings of Vision in Vehicles VIII*.

29. James A Bonneson. (2010). Highway Safety Manual. In *Highway Safety Manual* (1st ed., Vol. 1).

30. National Safety Council. (1997). *{ANSI D16.1--1996} {Manual} on Classification of Motor Vehicle Traffic Accidents*.

31. Elvik, R., & Mysen, A. (1999). Incomplete Accident Reporting: Meta-Analysis of Studies Made in 13 Countries. *Transportation Research Record*, *1665*(1), 133–140. https://doi.org/10.3141/1665-18

32. Hauer, E., & Hakkert, A. S. (1998). EXTENT AND SOME IMPLICATIONS OF INCOMPLETE ACCIDENT REPORTING. *Transportation Research Record*, *1185*, 1–10.

33. Davis, G. a. (2000). Estimating traffic accident rates while accounting for traffic volume measurement error: A Gibbs sampling approach. *Transportation Research Record Journal*, *1717*, 94–110.

34. Patil, S., Lord, D., & Reddy, S. (2012). Analysis of Crash Severities using Nested Logit Model - Accounting for the Underreporting of Crash Data. *Accident Analysis and Prevention*, *45*(July 2011), 646–653. https://doi.org/10.1016/j.aap.2011.09.034

35. Zahabi, S. A. H., Strauss, J., Manaugh, K., & Miranda-Moreno, L. F. (2011). Estimating Potential Effect of Speed Limits, Built Environment, and Other Factors on Severity of Pedestrian and Cyclist Injuries in Crashes: *Transportation Research Record: Journal of the Transportation Research Board*, *2247*, 81–90. https://doi.org/10.3141/2247-10

36. Tao, H., Foomani, M. G., & Alecsandru, C. (2015). *A TWO-STEP MICROSCOPIC TRAFFIC SAFETY EVALUATION MODEL OF RESERVED LANES FACILITIES : AN ARTERIAL CASE STUDY KEYWORDS :*

37. Bissell, D. (1980). Introductory statistics. In *Applied Statistics* (Vol. 29, Issue 1). https://doi.org/10.2307/2346426

38. Hauer, E. (1997). *observational before-after studies in road safety* (first). Emerlad.

39. Hauer, E. (1997). The naive before/after studies. In *observational before-after studies in road safety* (first edit, p. 74).

40. National Collision Database Online. (2014). *NCDB*. Transport Canada. http://wwwapps2.tc.gc.ca/Saf-Sec-Sur/7/NCDB-BNDC/p.aspx?l=en

41. Burns, P. C. (2005). International Harmonized Research Activities - Intelligent Transport Systems (IHRA- ITS) Working Group Report. *ESV -Paper No. 05-0461*, 1–9.

42. Dominique Lord, Xiao Qin, & Srinivas S. (2021). *Highway safety analytics and modeling.* (1st ed., Vol. 1). Elsevier.

43. Retting, R. A., Persaud, B. N., Hauer, E., Vallurapalli, R., & Mucsi, K. (1997). Crash reductions following traffic signal removal in Philadelphia. *Accident Analysis and Prevention*, *29*, 803–810.

44. Persaud, B. N., Retting, R. a, Garder, P. E., & Lord, D. (2001). Observational Before-After Study of the Safety Effect of U . S . Roundabout Conversions Using the Empirical Bayes Method. *Transporation Research Record*, *1751*, 1–8.

45. Lord, D. (2000). *The prediction of accidents on digital networks: characteristics and issues related*. University of Toronto.

46. Hauer, E. (2001). Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis and Prevention*, *33*(6), 799–808. https://doi.org/10.1016/S0001-4575(00)00094-4

47. Lord, D., Washington, S. P., & Ivan, J. N. (2005). Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident; Analysis and Prevention*, *37*(1), 35–46. https://doi.org/10.1016/j.aap.2004.02.004

48. Persaud, B., Lan, B., Lyon, C., & Bhim, R. (2010). Comparison of empirical Bayes and full Bayes approaches for before-after road safety evaluations. *Accident Analysis and Prevention*, *42*(1), 38–43. https://doi.org/10.1016/j.aap.2009.06.028

49. Hallmark, S. L., Hawkins, N., & Smadi, O. (2011). Assessment of the Effectiveness of Adding Reflectorized Treatments to Existing Chevrons on High-Crash Curves. *Transportation Research Board 90th Annual Meeting*, *2003*, 1–17.

50. Mann, P. S., & Lacke, C. J. (2010). Introductory Statistics. In B. Pearson (Ed.), *Nonparametric Methods* (seventh). wiley.

51. Hauer, E., & Persaud, B. (1996). *safety analysis of roadway geometric and ancillary features*.

52. Srinivasan, R., Carter, D., Persaud, B., Eccles, K., & Lyon, C. (2008). Safety Evaluation of Flashing Beacons at Stop-Controlled Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, *2056*(March), 77–86. https://doi.org/10.3141/2056-10

53. Shahdah, U., Saccomanno, F., & Persaud, B. (2013). *Estimates of countermeasure effects using crash-traffic conflict analysis*. *1989*, 1–9.

54. Shahdah, U. (2014). *Developing a Crash-Conflict Model for Safety Performance Analysis and Estimation of Crash Modification Factors for Urban Signalized Intersections*. *250*(14), 1–15.

55. Carlson, P. J., Lupes, M. S., Turner-Fairbank Highway Research Center., United States. Federal Highway Administration. Office of Research Development and Technology., & Texas Transportation Institute. (2007). *Methods for maintaining traffic sign retroreflectivity*. *November*. http://purl.fdlp.gov/GPO/gpo1124

56. Persaud, B., & Lyon, C. (2007). Empirical Bayes before-after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis and Prevention*, *39*(3), 546–555. https://doi.org/10.1016/j.aap.2006.09.009

57. GRIFFITH, M., HAYDEN, C., KONONOV, J., LEWIS, C. R., PAIN, R. F., SALZBERG, P., SCOPATZ, R., STEWART, D., & WASHINGTON, S. (2001). *Statistical Methods in Highway A Synthesis of Highway Practice*. National Cooperative Highway Research Program.

58. *Crash Modification Factors Clearinghouse*. (n.d.). Retrieved April 16, 2022, from http://www.cmfclearinghouse.org/

59. FHWA. (2014). *Interactive Highway Safety Design Model (IHSDM) software*. http://www.ihsdm.org/wiki/Welcome

60. Garber, N. J., & Hoel, L. a. (2009). Traffic and Highway FOURTH EDITION. In *America* (Fourth, p. 1249). CENGAGE learning.

61. Dong, C., Clarke, D. B., Yan, X., Khattak, A., & Huang, B. (2014). Multivariate random-parameters zero-inflated negative binomial regression model: An application to estimate crash frequencies at intersections. *Accident Analysis and Prevention*, *70*, 320–329. https://doi.org/10.1016/j.aap.2014.04.018

62. Lipinski, M. E., Seyfried, R. K., Seyfried, R. K., & Associates, I. (2016). *PROFESSIONAL TRAFFIC OPERATIONS ENGINEERS CERTIFICATION PROGRAM REFRESHER COURSE MODULE 6 TRAFFIC SAFETY*.

63. Elvik, R. (2004). To What Extent Is There Bias by Selection?: Selection for Road Safety Treatment in Norway. *Transportation Research Record*, *1897*(1), 200–205. https://doi.org/10.3141/1897-26

64. Hyden, C. (1987). *The development of a method for traffic safety evaluation: the Swedish traffic conflict technique*. Lund Instiiute of Technology.

65. Perkins, D., & Harris, J. (1967). *Criteria for traffic conflict characteristics*.

66. Caliendo, C., & Guida, M. (2012). A Micro-Simulation Approach for Predicting Crashes at Un-Signalized Intersections Using Traffic Conflicts. *Journal of Transportation Engineering*, *December*, 120822055938002. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000473

67. Perkins, D. D., & Bowman, B. L. (1983). EFFECTIVENESS EVALUATION BY USING NONACCIDENT MEASURES OF EFFECTIVENESS. *62nd Annual Meeting of the Transportation Research Board*, *905*. http://trid.trb.org/view.aspx?id=195833

68. William, T. (1972). AN EVALUATION OF THE TRAFFIC CONFLICTS TECHNIQUE. *51st Annual Meeting of the Highway Research Board*, *384*, 1–8.

69. Rao, V. T., & Rengaraju, V. R. (1997). Probabilistic Model for Conflicts at Urban Uncontrolled Intersection. *Journal of Transportation Engineering*, *123*(1), 81–84. https://doi.org/10.1061/(ASCE)0733-947X(1997)123:1(81)

70. Clennon, J. c., & Thorson, B. a. (1975). *Evaluation of the traffic conflicts technique* (Issues 3825-E).

71. Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. *Accident Analysis and Prevention*, *39*(4), 657–670. https://doi.org/10.1016/j.aap.2006.10.012

72. DOT-FHWA. (1989). *Traffic Conflict Techniques for Safety and Operations*.

73. Sayed, T. (2008). ESTIMATING THE SAFETY OF UNSIGNALIZED INTERSECTIONS USING TRAFFIC CONFLICTS. *Third National Access Management Conference*, *2*(6), 143–153.

74. Parker, M. r., & Zageer, C. V. (1989). Traffic Conflict Techniques for Safety and Operations. In *FHWA*.

75. Gettman, D., & Head, L. (2003). Surrogate Safety Measures from Traffic Simulation Models. *Transportation Research Record*, *1840*(1), 104–115. https://doi.org/10.3141/1840-12

76. Li, C., Karimi, M., & Alecsandru, C. (2018). Microscopic Simulation-Based High Occupancy Vehicle Lane Safety and Operation Assessment: A Case Study. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/5262514

77. Wang, C. (2015). *Sensitivity Analysis on New Simulation-based Conflict Metrics*. *July 2014*.

78. Saunier, N. (2014). Clustering Surrogate Safety Indicators to Understand Collision Processes. *Transportation Research Board 93rd Annual Meeting. January 12-16, Washington, D.C.*

79. Pirdavani, A., Brijs, T., Bellemans, T., & Wets, G. (2010). Evaluation of traffic safety at un-signalized intersections using microsimulation: A utilization of proximal safety indicators. In *Advances in Transportation Studies* (Issue 22, pp. 43–50).

80. Deepak, V. kill, & Vedagiri, P. (2014). Proactive Evaluation of Traffic Safety at An Unsignalized Intersection Using Micro- Simulation. *Journal of Traffic and Logistics Engineering*, *2*(2), 140–145. https://doi.org/10.12720/jtle.2.2.140-145

81. El-basyouny, K., & Sayed, T. (2011). Conflicts-Based Safety Performance Functions. *Transportation Research Record*, 1–12.

82. Lorion, A. C., & Persaud, B. (2015). INVESTIGATION OF SURROGATE MEASURES FOR SAFETY ASSESSMENT OF URBAN TWO-WAY STOP CONTROLLED INTERSECTIONS. *Transportation Esearch Board - 94th Annual Meeting*, *416*.

83. Essa, M., & Sayed, T. (2015). Simulated Traffic Conflicts: Do They Accurately Represent Field-Measured Conflicts? *Transportation Research Record*, *15–0707*.

84. Habtemichael, F. G., & Santos, L. D. P. (2012). Sensitivity Analysis of VISSIM Driver Behavior Parameters on Safety of Simulated Vehicles and Their Interaction with Operations of Simulated Traffic. *Transportation Research Board 92nd Annual Meeting*, *2013*, 1–17.

85. Shahrokhi, H. (2013). *An Efficient Soft Computing Based Method for Calibration of Vehicular Microscopic Simulation Models of Building , Civil and Environmental Engineering Presented in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science (Civil*. Concordia University.

86. TAC. (2000). *Manual of Standard Traffic Signs & Pavement Markings* (Issue September).

87. ITE. (1998). Vehicle Traffic Control Signal Heads – Part 2: Light Emitting Diode Traffic Control Signal Modules,. In *ITE*.

88. Zwahlen, H., & Schnell, T. (1997). Visual Detection and Recognition of Fluorescent Color Targets Versus Nonfluorescent Color Targets as a Function of Peripheral Viewing Angle and Target Size. *Transportation Research Record*, *1605*(1), 28–40. https://doi.org/10.3141/1605-05

89. Jenssen, G. D., BREKKE, B., Chrysler, S., & Burns, D. (1996). Visual Performance of Fluorescent Retroreflective Traffic Control Devices. *Elsevier VISION IN VEHICLES*, *VI*, 89–179.

90. Carlson, P. (2011). *Evaluation of Sign Retroreflectivity Measurements from the Advanced Mobile Asset Collection (AMAC) System EVALUATION OF SIGN RETROREFLECTIVITY*

MEASUREMENTS FROM THE ADVANCED MOBILE ASSET COLLECTION (AMAC) SYSTEM (Issue October).

91. Stirling Stackhouse, & Paul Cassidy. (1998). *Warning Flashers at Rural Intersections*.

92. Murphy, B. G., & Hummer, J. E. (2008). Development of Crash Reduction Factors for Overhead Flashing Beacons at Rural Intersections in North Carolina. *Transportation Research Record*, *2030*(1), 15–21. https://doi.org/10.3141/2030-03

93. Vogt, A., & Bared, J. G. (1998). Accident Models for Two-Lane Rural Roads. *FHWA Publications*, *FHWA-RD-98-133*.

94. ARNOLD, E. D., & LANTZ, K. E. (2013). *Evaluation of Best Practices in Traffic Operations and Safety: Phase I: Flashing LED Stop Sign and Optical Speed Bars* (Issue Final Report VTRC 07-R34).

95. van Houten, R., & Retting, R. a. (2001). Increasing motorist compliance and caution at stop signs. *Journal of Applied Behavior Analysis*, *34*(2), 185–193. https://doi.org/10.1901/jaba.2001.34-185

96. Kwon, T. M., Weidemann, R., Lund, V., & Boder, B. (2011). Determining the Effectiveness of an Advance LED Warning System for Rural Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, *2250*(1), 25–31. https://doi.org/10.3141/2250-04

97. Kwon, T. M., Lund, V., & Ismail, H. I. (2015). Evaluation of the ALERT System, a Rural Intersection Conflict Warning SystemTitle. *Transportation Research Board 95th Annual Meeting*.

98. Paddick, D., Heydel, T., Bott, M., Kuznicki, S., Woosley, D., Sullivan, J., & Kennedy, J. (2014). *Intersection Conflict Warning System* (Vol. 5, Issue 2).

99. Jomaa, D., Yella, S., & Dougherty, M. (2013). *Review of the Effectiveness of Vehicle Activated Signs*. *2013*(April), 123–130.

100. Monsere, C., Nolan, C., Bertini, R., Anderson, E., & El-Seoud, T. (2005). Measuring the Impacts of Speed Reduction Technologies: Evaluation of Dynamic Advanced Curve Warning System. *Transportation Research Record*, *1918*(November 2004), 98–107. https://doi.org/10.3141/1918-13

101. Patella, S. M., Sportiello, S., Carrese, S., Bella, F., & Asdrubali, F. (2020). The effect of a LED lighting crosswalk on pedestrian safety: Some experimental results. *Safety*, *6*(2). https://doi.org/10.3390/safety6020020

102. Rista, E., & Fitzpatrick, K. (2020). Comparison of LED-Embedded Pedestrian Crossing Signs with Rectangular Rapid Flashing Beacons and Pedestrian Hybrid Beacons. *Transportation Research Record*, *2674*(11), 856–866. https://doi.org/10.1177/0361198120941849

103. Fitzpatrick, K., & Park, E. S. (2021). Nighttime effectiveness of the pedestrian hybrid beacon, rectangular rapid flashing beacon, and LED-embedded crossing sign. *Journal of Safety Research*, *79*, 273–286. https://doi.org/10.1016/j.jsr.2021.09.009

104. Petrović, M., & Novačko, L. (n.d.). *EcoProduction. Environmental Issues in Logistics and Manufacturing Transformation of Transportation*. http://www.springer.com/series/10152

105. Stipancic, J., St-Aubin, P. G., Ledezma-Navarro, B., Labbe, A., Saunier, N., & Miranda-Moreno, L. (2021). Evaluating safety-influencing factors at stop-controlled intersections

using automated video analysis. *Journal of Safety Research*, *77*, 311–323. https://doi.org/10.1016/j.jsr.2021.03.006

106. Foomani, M. G., Alecsandru, C., & McConnell, L. (2017). *Safety Assessment of Backlit Pedestrian Crossing Sign on Stopping Characteristics at Unsignalized Intersection*. https://trid.trb.org/view/1437617

107. Giahi Foomani, M., Li, C., & Alecsandru, C. (2016). INVESTIGATING SAFETY OF DIFFERENT TREATMENTS AT STOP CONTROLLED INTERSECTION USING DRIVER BEHAVIOR ANALYSES Matin. *Transportation Esearch Board - 95th Annual Meeting*. https://doi.org/10.1111/bjd

108. St-Aubin, P., Saunier, N., Miranda-Moreno, L. F., & Ismail, K. (2013). Detailed Driver Behaviour Analysis and Trajectory Interpretation at Roundabouts Using Computer Vision Data. *Transportation Esearch Board - 92nd Annual Meeting*, 1–17. https://doi.org/10.3141/2389-07

109. Robertson, H. Douglas., Hummer, J. E., Nelson, D. C., & Engineers., I. of T. (1994). Manual of transportation engineering studies. In *TA - TT* -. Prentice Hall. http://www.gbv.de/dms/bowker/toc/9780130975690.pdf

110. Saunier, N., & Sayed, T. (2006). A feature-based tracking algorithm for vehicles in intersections. *Third Canadian Conference on Computer and Robot Vision, CRV 2006*, *2006*. https://doi.org/10.1109/CRV.2006.3

111. Saunier, N. (2013). *TrafficIntelligence*. Ecole Polytechnique de Montréal. https://bitbucket.org/Nicolas/trafficintelligence/wiki/Home

112. Ma, J., Dong, H., & Zhang, H. M. (2007). Calibration of Micro Simulation with Heuristic Optimization Methods. *TRB Committee on Traffic Flow Theory and Characteristics*, *45*(1), 1–18.

113. Fellendorf, M., & Vortisch, P. (2010). Microscopic traffic flow simulator VISSIM. In *International Series in Operations Research and Management Science* (Vol. 145, pp. 63–93). Springer New York LLC. https://doi.org/10.1007/978-1-4419-6142-6_2

114. Liu, Z., & Yu, X. (2013). *TRB 2013 Annual Meeting Paper revised from original submittal. 750*(216).

115. PTV AG. (2013). *PTV VISSIM 6 User Manual* (No. 6).

116. Wright, A. H. (1991). *Genetic Algorithms for Real Parameter Optimization*. *1*, 205–218. https://doi.org/10.1016/B978-0-08-050684-5.50016-1

117. Ma, T., & Abdulhai, B. (2001). Genetic algorithm-based combinatorial parametric optimization for the calibration of microscopic traffic simulation models. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 848–853. https://doi.org/10.1109/ITSC.2001.948771

118. Saleh, W., Kumar, R., Kirby, H., & Kumar, P. (2009). Real world driving cycle for motorcycles in Edinburgh. *Transportation Research Part D: Transport and Environment*, *14*(5), 326–333. https://doi.org/10.1016/J.TRD.2009.03.003

119. Montgomery, D. C., & Runger, G. C. (n.d.). *Applied statistics and probability for engineers*. 21.

120. McClave, J. T., Benson, P. G., & Sincich, T. (2015). *Statistics for Business and Economics*. Pearson Education. https://books.google.ca/books?id=ZdigBwAAQBAJ

121. Allison, P. D. (n.d.). *Paper 1485-2014 SAS Global Forum*.

122. Pituch, K. A., & Stevens, J. (James P. (n.d.). *Applied multivariate statistics for the social sciences : analyses with SAS and IBM's SPSS*.

123. Washington, S., Karlaftis, M., Mannering, F., & Anastasopoulos, P. (2020). *Statistical and Econometric Methods for Transportation Data Analysis*. https://doi.org/10.1201/9780429244018

124. Wagner, P., Hoffmann, R., & Leich, A. (2021). Observations on the relationship between crash frequency and traffic flow. *Safety*, *7*(1). https://doi.org/10.3390/safety7010003

125. Retallack, A. E., & Ostendorf, B. (2020). Relationship between traffic volume and accident frequency at intersections. *International Journal of Environmental Research and Public Health*, *17*(4). https://doi.org/10.3390/ijerph17041393

126. Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. In *Accident Analysis and Prevention* (Vol. 39, Issue 4, pp. 657–670). https://doi.org/10.1016/j.aap.2006.10.012

127. Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, *34*(1), 1–14. https://doi.org/10.1080/00401706.1992.10485228

128. Mason, R. L., Gunst, R. F., & Hess, J. L. T. A.-T. T.-. (2003). *Statistical design and analysis of experiments : with applications to engineering and science* (2nd ed. NV). J. Wiley.

https://doi.org/10.1002/0471458503 LK -
https://concordiauniversity.on.worldcat.org/oclc/52752095

129. Roger E Kirk. (2013). *Experimental design : procedures for the behavioral sciences LK - https://concordiauniversity.on.worldcat.org/oclc/907030601: Vol. Volume 1* (4th ed.). SAGE Publications. http://srmo.sagepub.com/view/experimental-design/SAGE.xml

130. Ni, Y., & Wang, M. (2015). Behavior-based Method on Pedestrian-Vehicle Conflict Analysis at Intersections. *Transportation Research Board 94th Annual Meeting*, *500*.

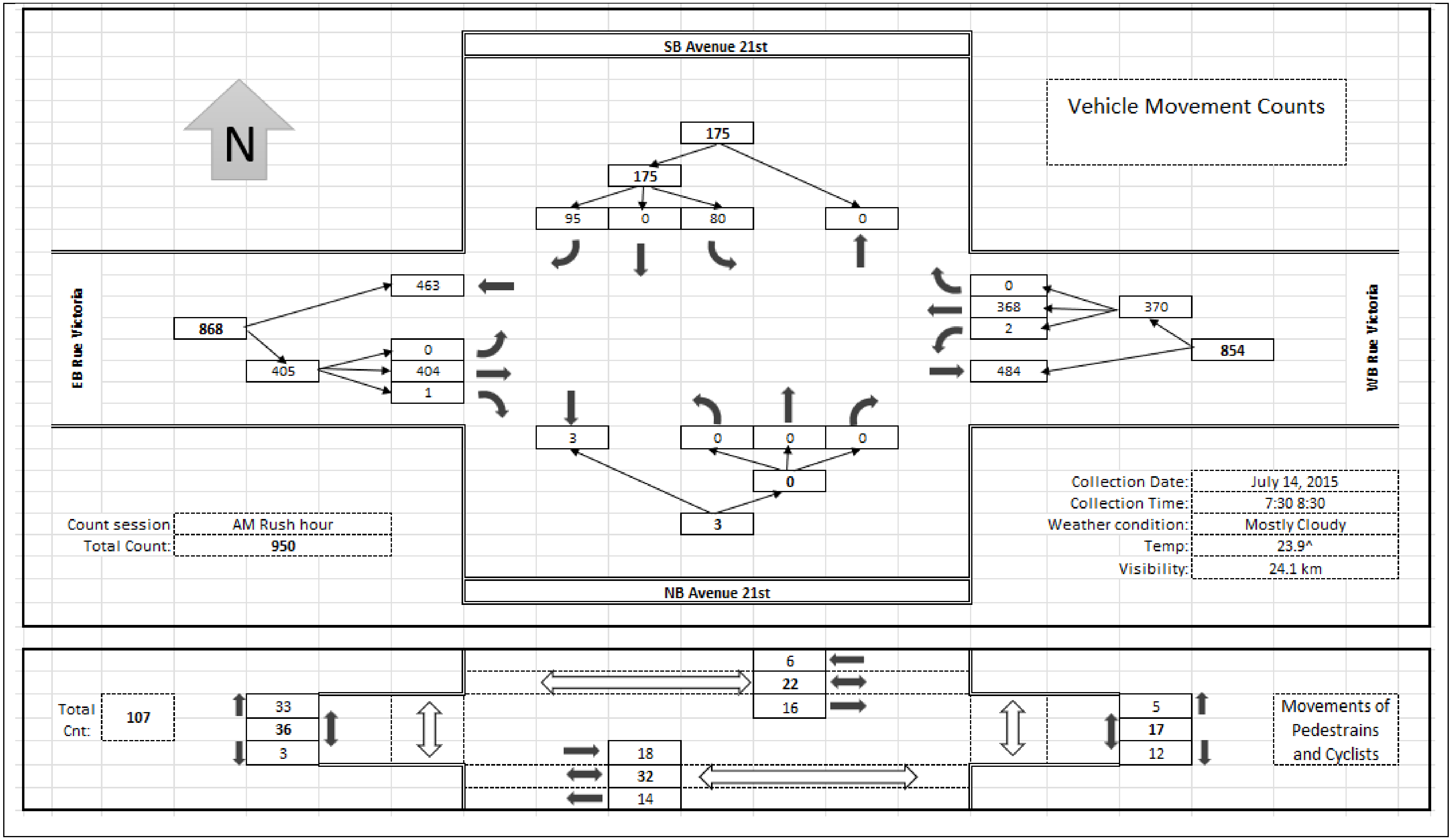
7 Appendix I (data)

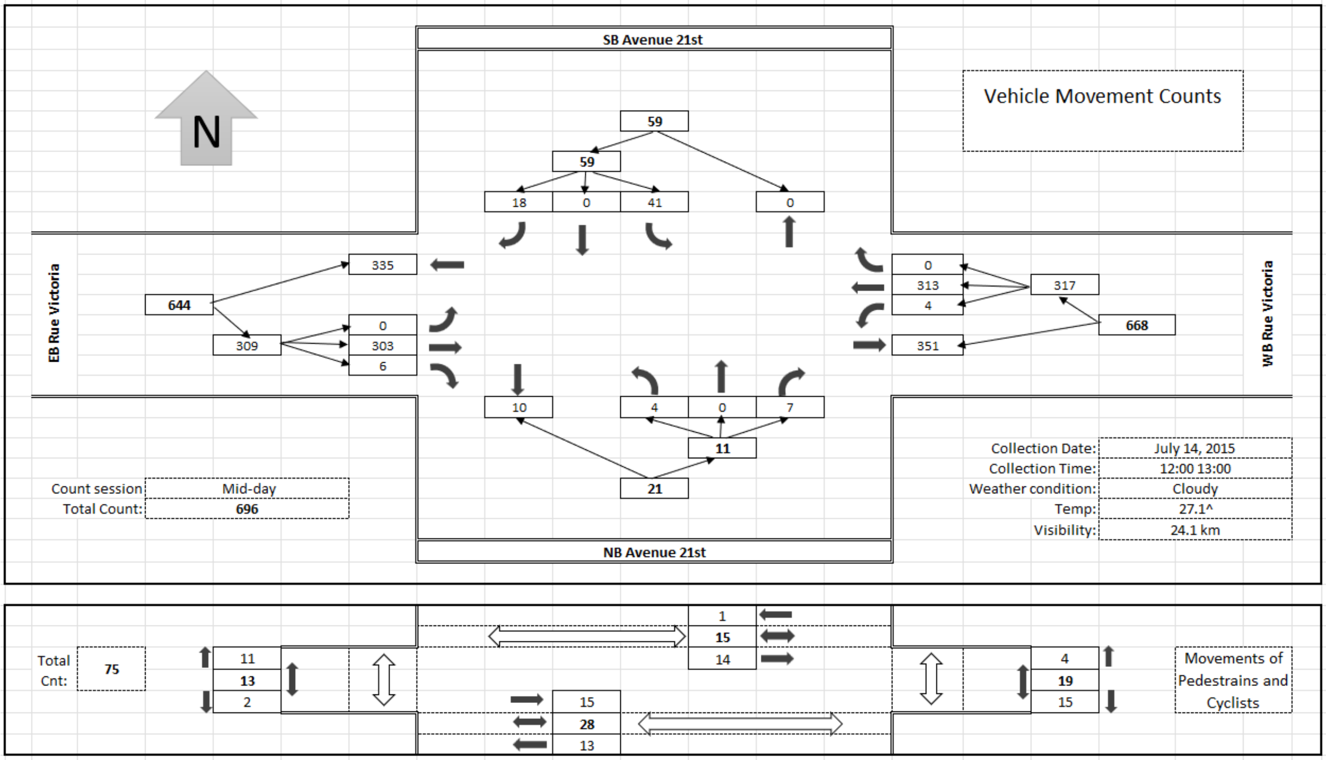
Ambient	Pavement	Weather	H	Date	Minor	18th: Victoria& Ave. 18	18th: 30m WB Ave. 18	18th: 30m EB Ave. 18	21st: Victoria& Ave. 21	21st: 30m EB Ave. 21
Day	wet	Clear	1	14, Tue,Dec 02	7:30 8:30				1	
Day	wet	Mostly	1	14, Tue,Dec 02	12:00 13:00				1	
TWI.	wet	Cloudy	1	14, Tue,Dec 02	16:30 17:30				1	
Night	wet	Snow	4	14, Tue,Dec 02	18:00 22:00				4	
Day	wet	Cloudy	1	14, Wed,Dec 03	7:30 8:30					
Day	wet	Snow	1	14, Wed,Dec 03	12:00 13:00					
TWI.	wet	Sno	1	14, Wed,Dec 03	16:30 17:30					
Night	wet	Cloudy	4	14, Wed,Dec 03	18:00 22:00					
Day	wet	Cloudy	1	15, Mon,Apr 06	7:30 8:30					
Day	wet	Snow	1	15, Mon,Apr 06	12:00 13:00					
Day	wet	Snow	1	15, Mon,Apr 06	16:30 17:30	1	1			
Night	wet	Snow, Fog	4	15, Mon,Apr 06	18:00 22:00	4	4			
Day	Dry	Cloudy	1	15, Wed,Apr 08	7:30 8:30					
Day	Dry	Cloudy	1	15, Wed,Apr 08	12:00 13:00					
Day	Dry	Cloudy	1	15, Wed,Apr 08	16:30 17:30	1		1		
Night	Dry	Snow	4	15, Wed,Apr 08	18:00 22:00	4		4		
Installation and Warm-up										
Day	Dry	Cloudy	1	15, Mon,Jul 06	7:30 8:30				1	
Day	Dry	Clear	1	15, Mon,Jul 06	12:00 13:00				1	
Day	Dry	Clear	1	15, Mon,Jul 06	16:30 17:30				1	
TWI.	Dry	Mainly Clear	4	15, Mon,Jul 06	18:00 22:00				4	
Day	Dry	Mainly Clear	1	15, Wed,Jul 08	7:30 8:30					
Day	Dry	Mainly Clear	1	15, Wed,Jul 08	12:00 13:00					
Day	Dry	Mainly Clear	1	15, Wed,Jul 08	16:30 17:30	1	1	1		
TWI.	Dry	Mainly Clear	4	15, Wed,Jul 08	18:00 22:00	4	4	4		
Day	Dry	Mainly Clear	1	15, Fri,Jul 10	7:30 8:30					
Day	Dry	Mainly Clear	1	15, Fri,Jul 10	12:00 13:00				1	
Day	Dry	Mainly Clear	1	15, Fri,Jul 10	16:30 17:30				1	
TWI.	Dry	Mainly Clear	4	15, Fri,Jul 10	18:00 22:00				4	
Day	Dry	Mostly	1	15, Tue,Jul 14	7:30 8:30				1	
Day	Dry	Cloudy	1	15, Tue,Jul 14	12:00 13:00				1	
Day	Dry	Cloudy	1	15, Tue,Jul 14	16:30 17:30				1	
TWI.	Dry	Cloudy	4	15, Tue,Jul 14	18:00 22:00				4	
Day	Dry	Mainly Clear	1	15, Wed,Jul 15	7:30 8:30					
Day	Dry	Mainly Clear	1	15, Wed,Jul 15	12:00 13:00					
Day	Dry	Mainly Clear	1	15, Wed,Jul 15	16:30 17:30					1
TWI.	Dry	Mainly Clear	4	15, Wed,Jul 15	18:00 22:00					4
Day	Dry	Mainly Clear	1	15, Thu,Jul 16	7:30 8:31					
Day	Dry	Mainly Clear	1	15, Thu,Jul 16	12:00 13:01					
Day	Dry	Clear	1	15, Thu,Jul 16	16:30 17:31					
TWI.	Dry	Mainly Clear	4	15, Thu,Jul 16	18:00 22:01					

25th	Victoria& Ave. 25	1	1	1	4	1	1	1	4			1	4								1	4							1	4					1	1	1	4				
	30m EB Ave. 25											1	4								1	4																			1	4
28th	Victoria& Ave. 28	1	1	1	4	1	1	1	4							1	4										1	1	1	4					1	1	1	4				
	30m EB Ave. 28															1	4																									
32nd	Victoria& Ave. 32																														1	1	1	4								
33rd	Victoria& Ave. 33					1	1	1	4																						1	1	1	4					1	1	1	4
	30m EB Ave. 33																																								1	4

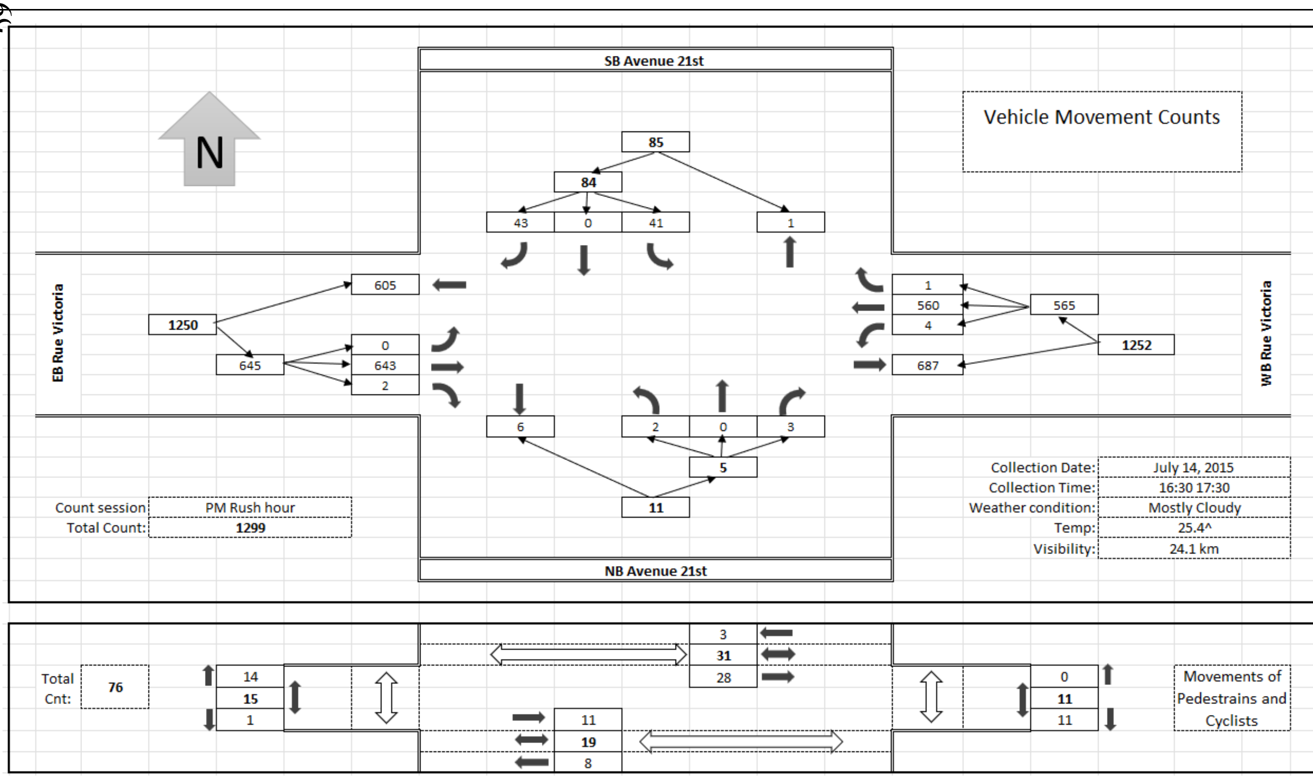
Table 38 Recording sessions and hours

Table 39 Sample recorded TMCs report sheets for intersection under study

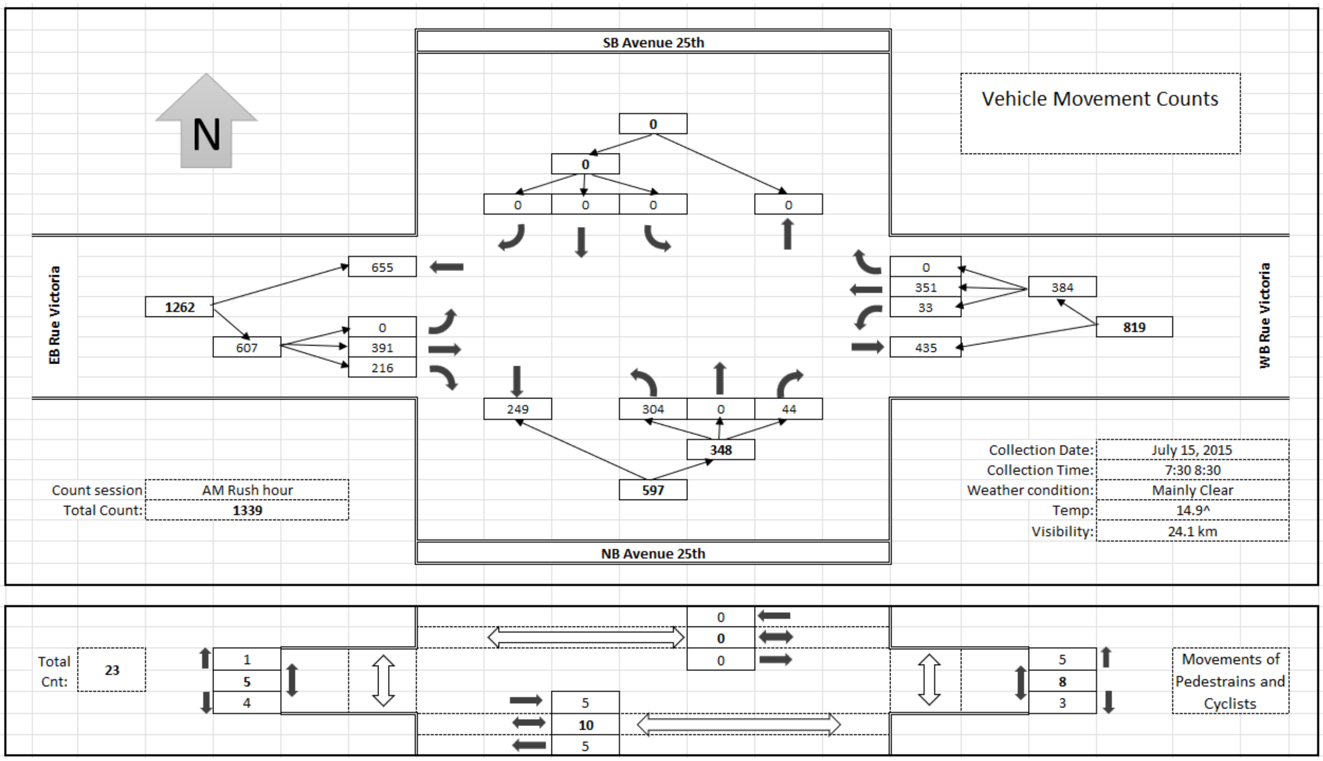


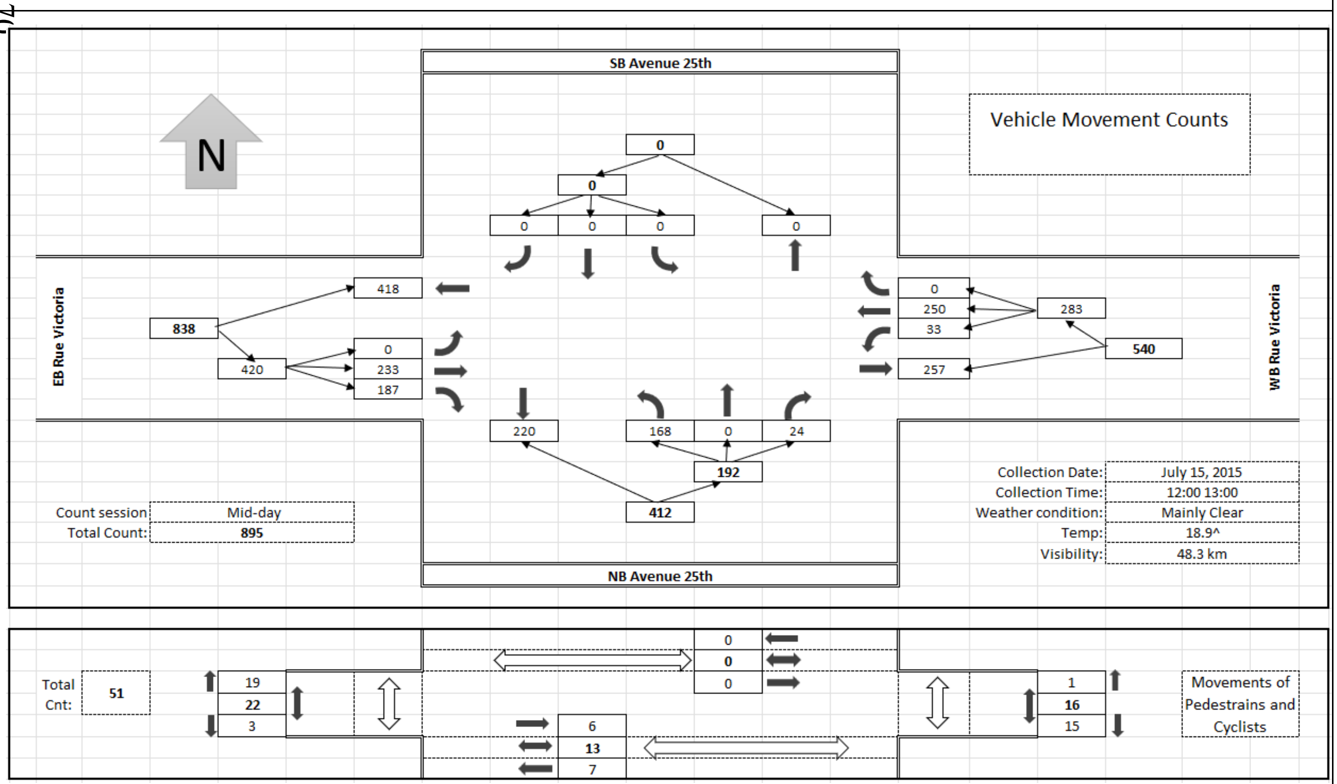


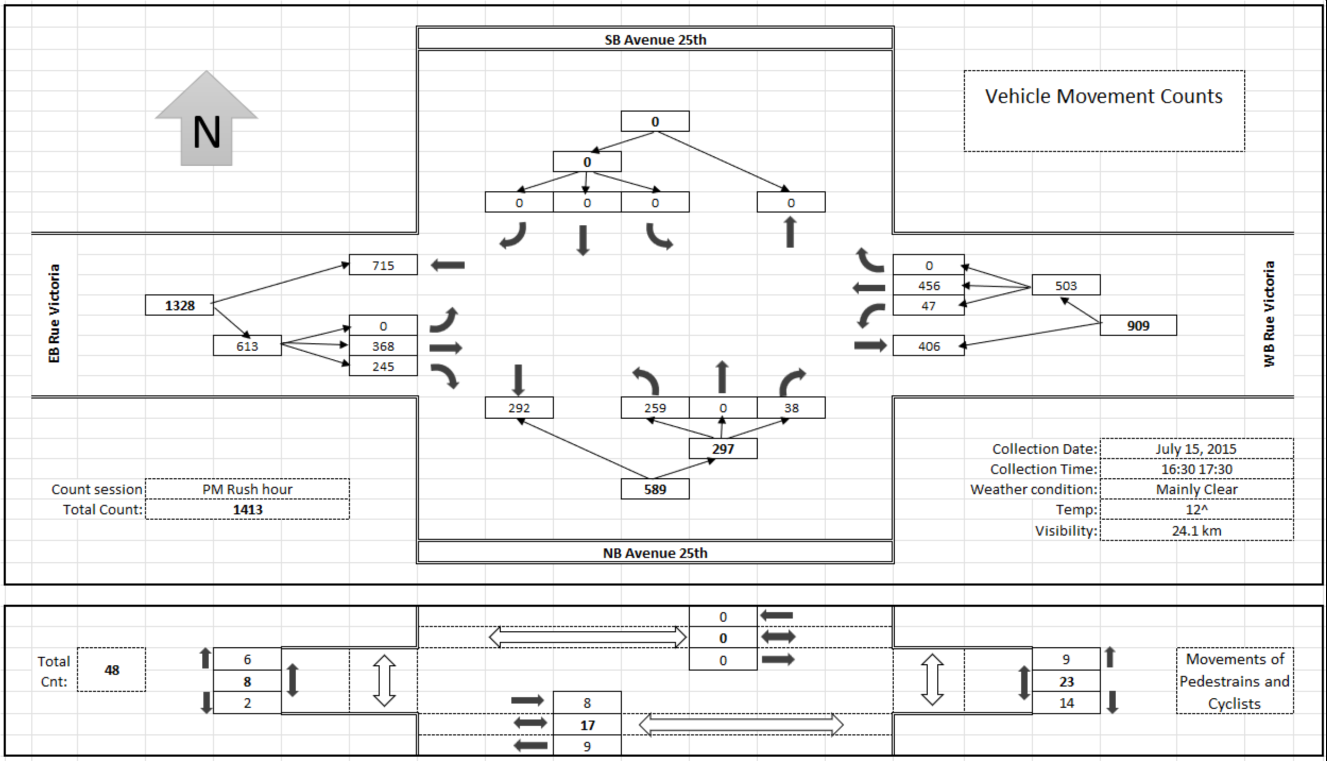
Vehicle Movement Counts
N
SB Avenue 21st
NB Avenue 21st
EB Rue Victoria
WB Rue Victoria
59
59
18
0
41
0
644
335
309
0
303
6
0
313
4
317
668
351
10
4
0
7
11
21
Count session: Mid-day
Total Count: 696
Collection Date: July 14, 2015
Collection Time: 12:00 13:00
Weather condition: Cloudy
Temp: 27.1^
Visibility: 24.1 km
Total Cnt: 75
11
13
2
1
15
14
15
28
13
4
19
15
Movements of Pedestrains and Cyclists

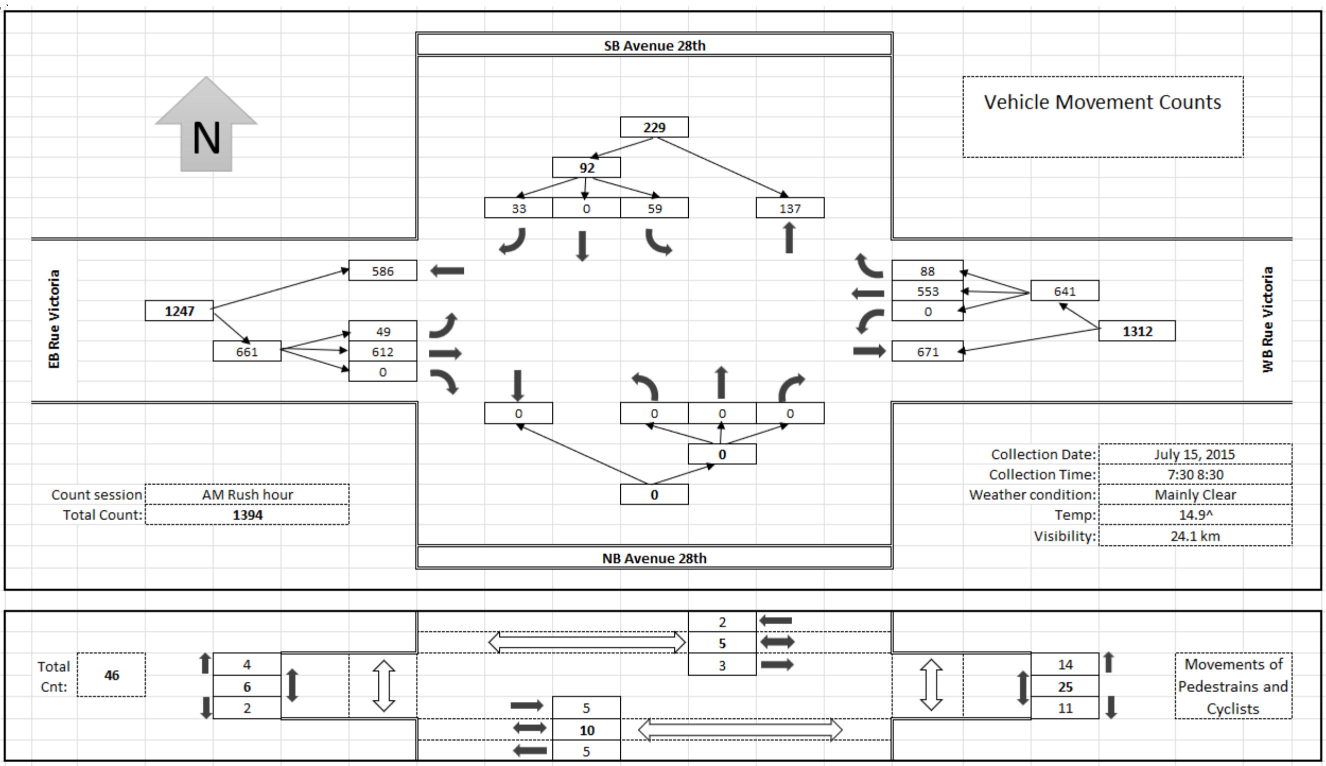


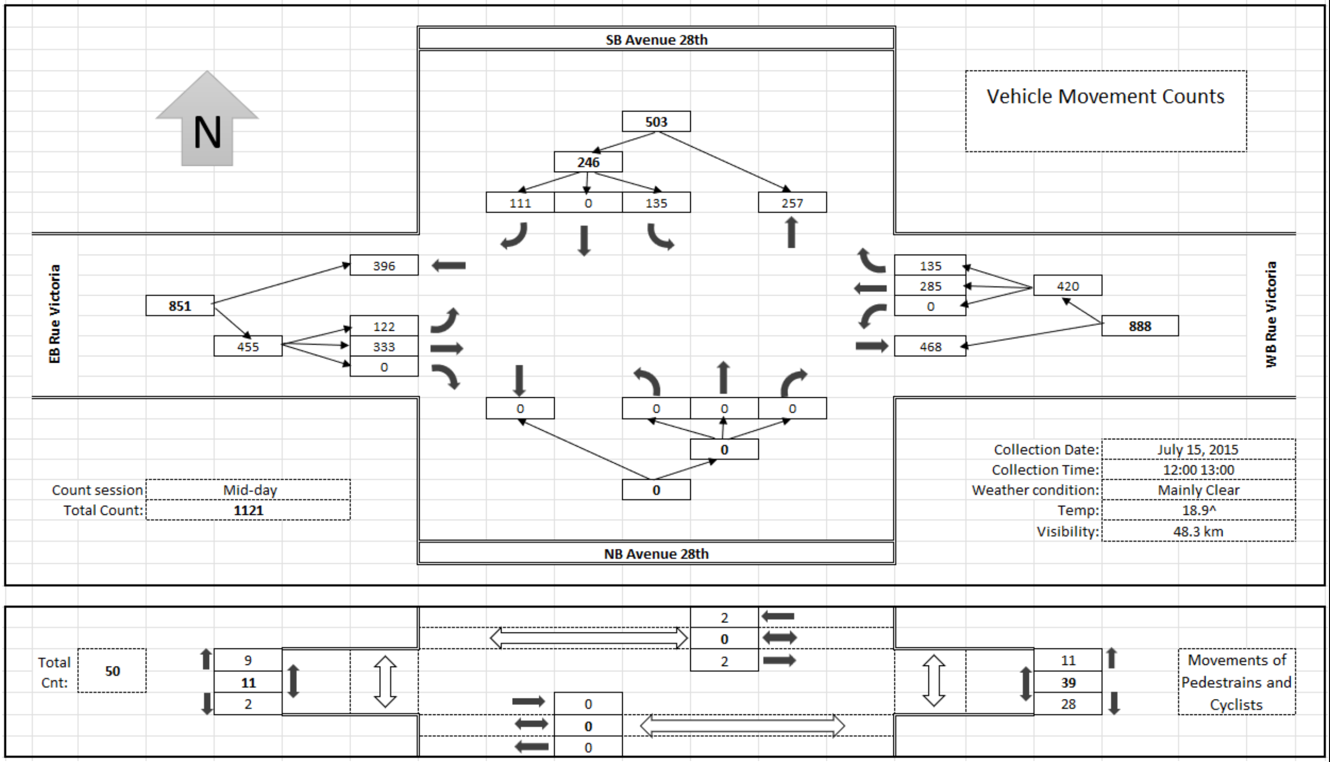
Vehicle Movement Counts
N
SB Avenue 21st
NB Avenue 21st
EB Rue Victoria
WB Rue Victoria
85
84
43
0
41
1
1250
605
645
0
643
2
1
560
4
565
1252
687
6
2
0
3
5
11
Count session: PM Rush hour
Total Count: 1299
Collection Date: July 14, 2015
Collection Time: 16:30 17:30
Weather condition: Mostly Cloudy
Temp: 25.4^
Visibility: 24.1 km
Total Cnt: 76
14
15
1
3
31
28
11
19
8
0
11
11
Movements of Pedestrains and Cyclists

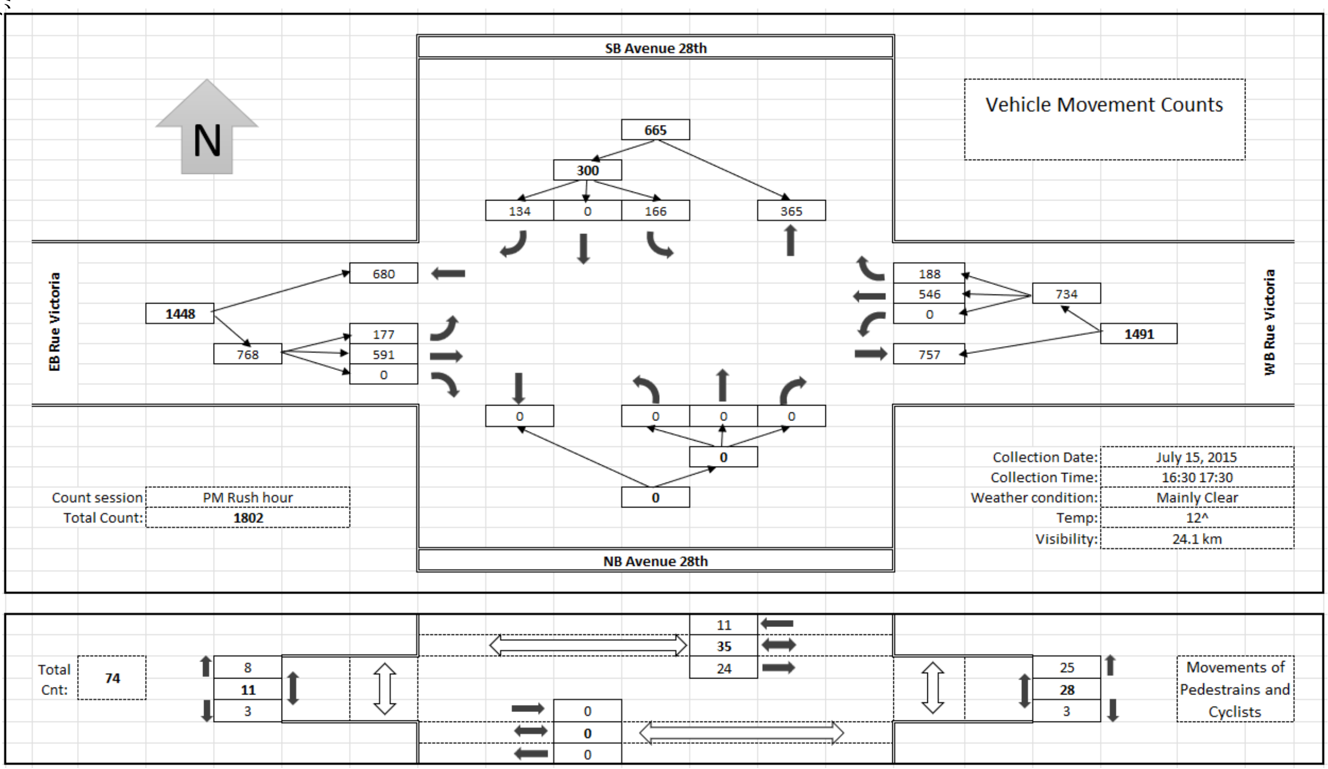

Vehicle Movement Counts
N
SB Avenue 25th
NB Avenue 25th
EB Rue Victoria
WB Rue Victoria
0
0
0 0 0 0
1262
655
607
0 391 216
819
384
0 351 33
435
249
304 0 44
348
597
Count session: AM Rush hour
Total Count: 1339
Collection Date: July 15, 2015
Collection Time: 7:30 8:30
Weather condition: Mainly Clear
Temp: 14.9^
Visibility: 24.1 km
Total Cnt: 23
1 5 4
0 0 0
5 10 5
5 8 3
Movements of Pedestrains and Cyclists

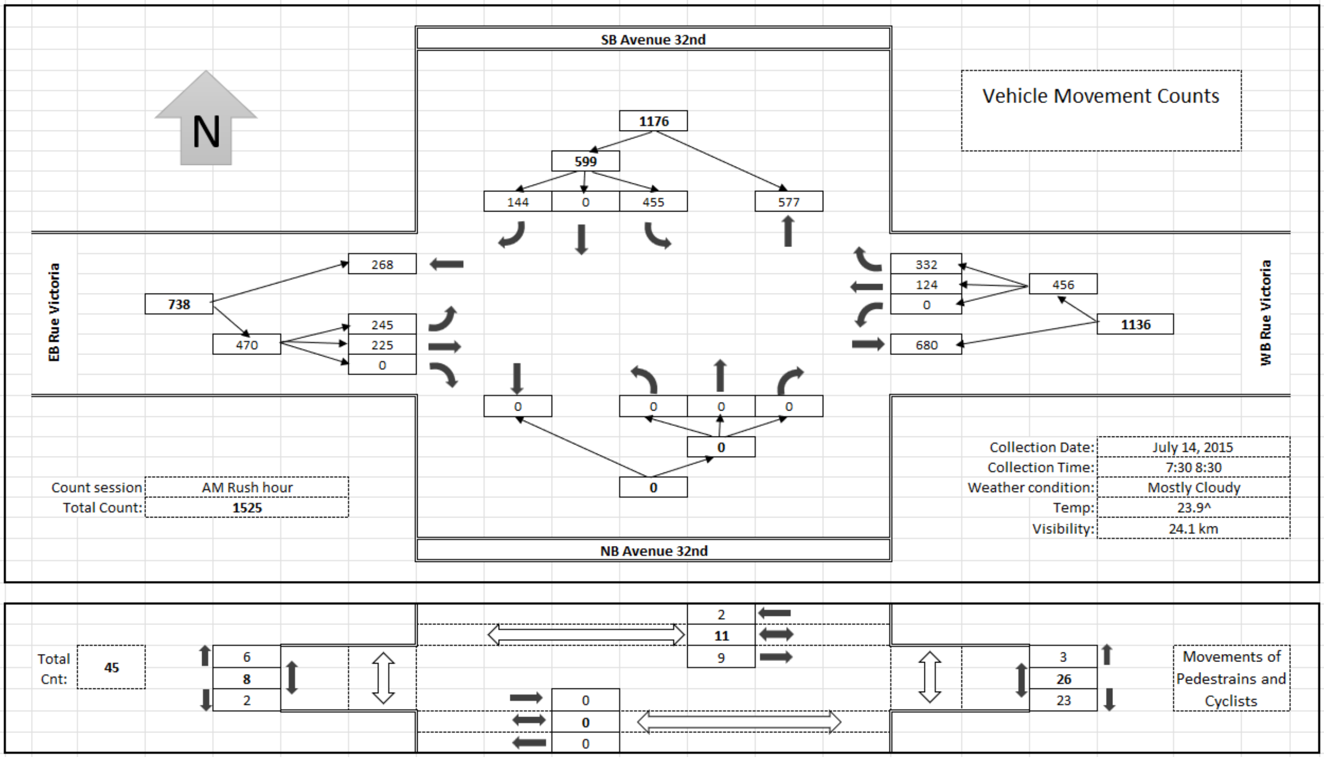

Vehicle Movement Counts
N
SB Avenue 25th
NB Avenue 25th
EB Rue Victoria
WB Rue Victoria
0
0
0 0 0 0
838
418
420
0 233 187
540
283
0 250 33
257
220
168 0 24
192
412
Count session: Mid-day
Total Count: 895
Collection Date: July 15, 2015
Collection Time: 12:00 13:00
Weather condition: Mainly Clear
Temp: 18.9^
Visibility: 48.3 km
Total Cnt: 51
19 22 3
0 0 0
6 13 7
1 16 15
Movements of Pedestrains and Cyclists

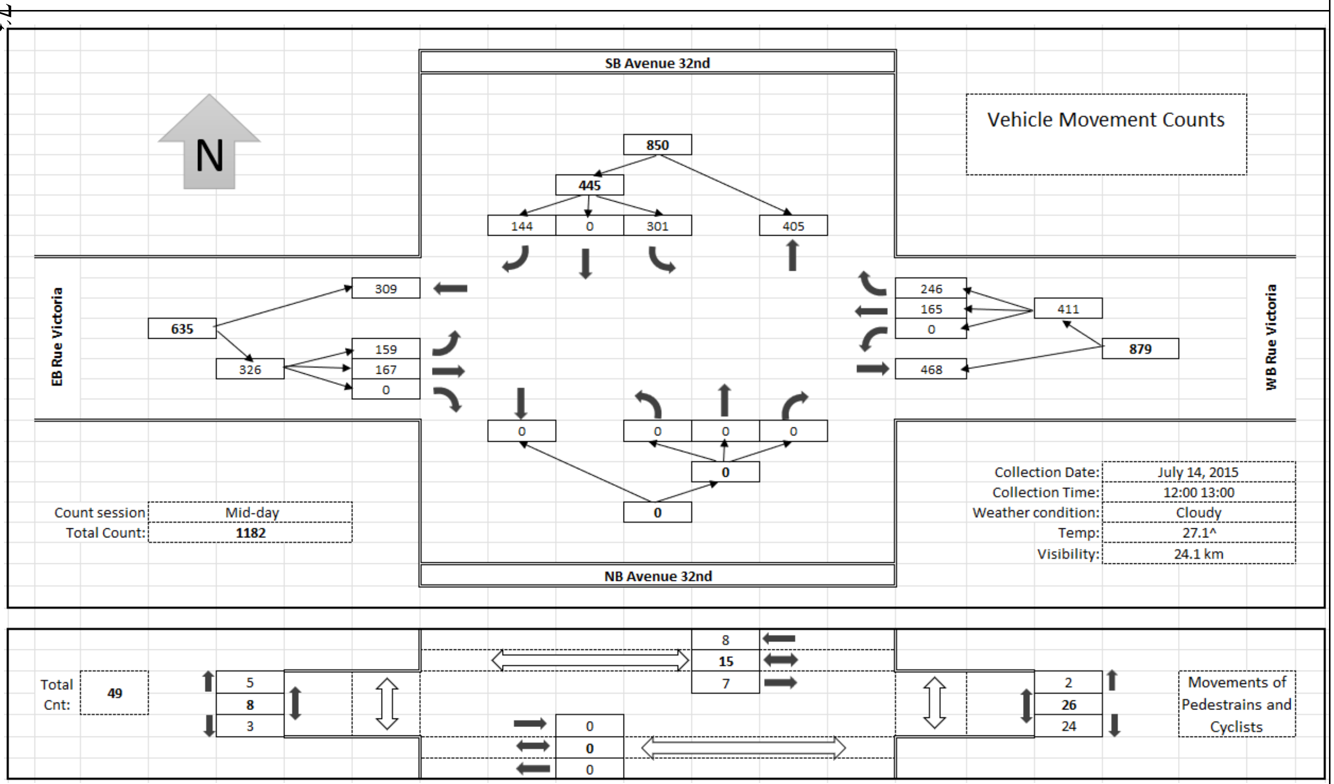

Vehicle Movement Counts
N
SB Avenue 25th
NB Avenue 25th
EB Rue Victoria
WB Rue Victoria
0
0
0
0
0
0
1328
715
613
0
368
245
909
503
0
456
47
406
292
259
0
38
297
589
Count session: PM Rush hour
Total Count: 1413
Collection Date: July 15, 2015
Collection Time: 16:30 17:30
Weather condition: Mainly Clear
Temp: 12^
Visibility: 24.1 km
Total Cnt: 48
6
8
2
0
0
0
8
17
9
9
23
14
Movements of Pedestrains and Cyclists

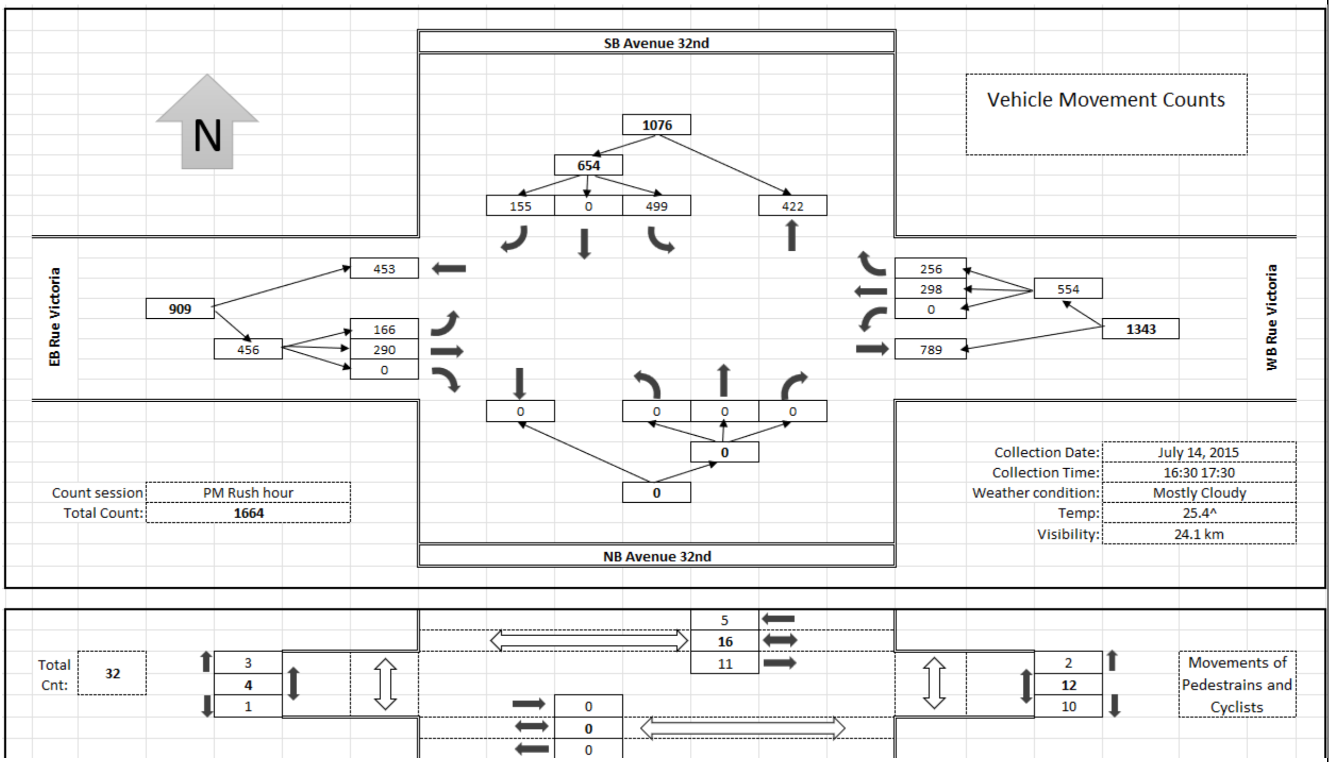

Vehicle Movement Counts
N
SB Avenue 28th
NB Avenue 28th
EB Rue Victoria
WB Rue Victoria
229
92
33
0
59
137
1247
586
661
49
612
0
1312
641
88
553
0
671
0
0
0
0
0
0
Count session: AM Rush hour
Total Count: 1394
Collection Date: July 15, 2015
Collection Time: 7:30 8:30
Weather condition: Mainly Clear
Temp: 14.9^
Visibility: 24.1 km
Total Cnt: 46
4
6
2
2
5
3
5
10
5
14
25
11
Movements of Pedestrains and Cyclists

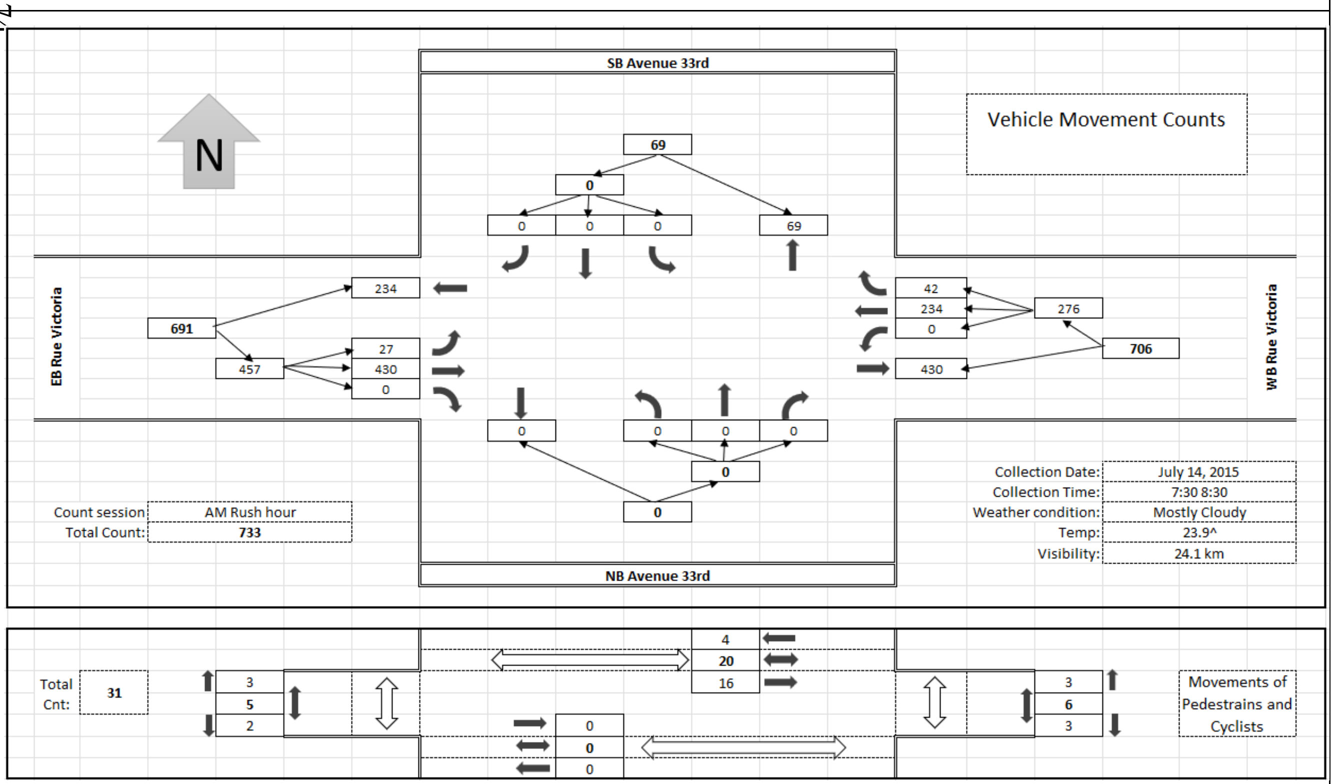

Vehicle Movement Counts
N
SB Avenue 28th
NB Avenue 28th
EB Rue Victoria
WB Rue Victoria
503
246
111
0
135
257
851
396
455
122
333
0
888
420
135
285
0
468
0
0
0
0
0
0
Count session: Mid-day
Total Count: 1121
Collection Date: July 15, 2015
Collection Time: 12:00 13:00
Weather condition: Mainly Clear
Temp: 18.9^
Visibility: 48.3 km
Total Cnt: 50
9
11
2
2
0
2
0
0
0
11
39
28
Movements of Pedestrains and Cyclists

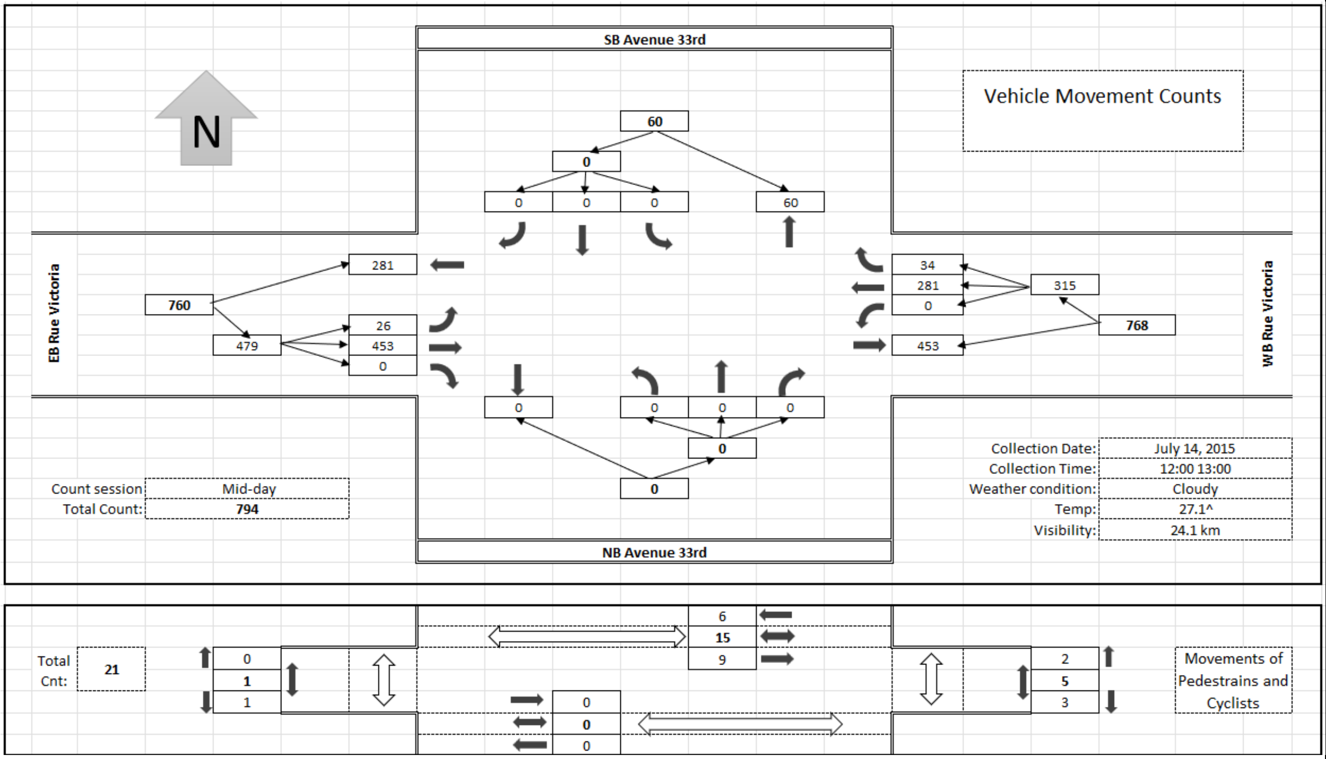

Vehicle Movement Counts
N
SB Avenue 28th
NB Avenue 28th
EB Rue Victoria
WB Rue Victoria
665
300
134
0
166
365
1448
680
768
177
591
0
1491
734
188
546
0
757
0
0
0
0
0
0
Count session: PM Rush hour
Total Count: 1802
Collection Date: July 15, 2015
Collection Time: 16:30 17:30
Weather condition: Mainly Clear
Temp: 12^
Visibility: 24.1 km
Total Cnt: 74
8
11
3
11
35
24
0
0
0
25
28
3
Movements of Pedestrains and Cyclists

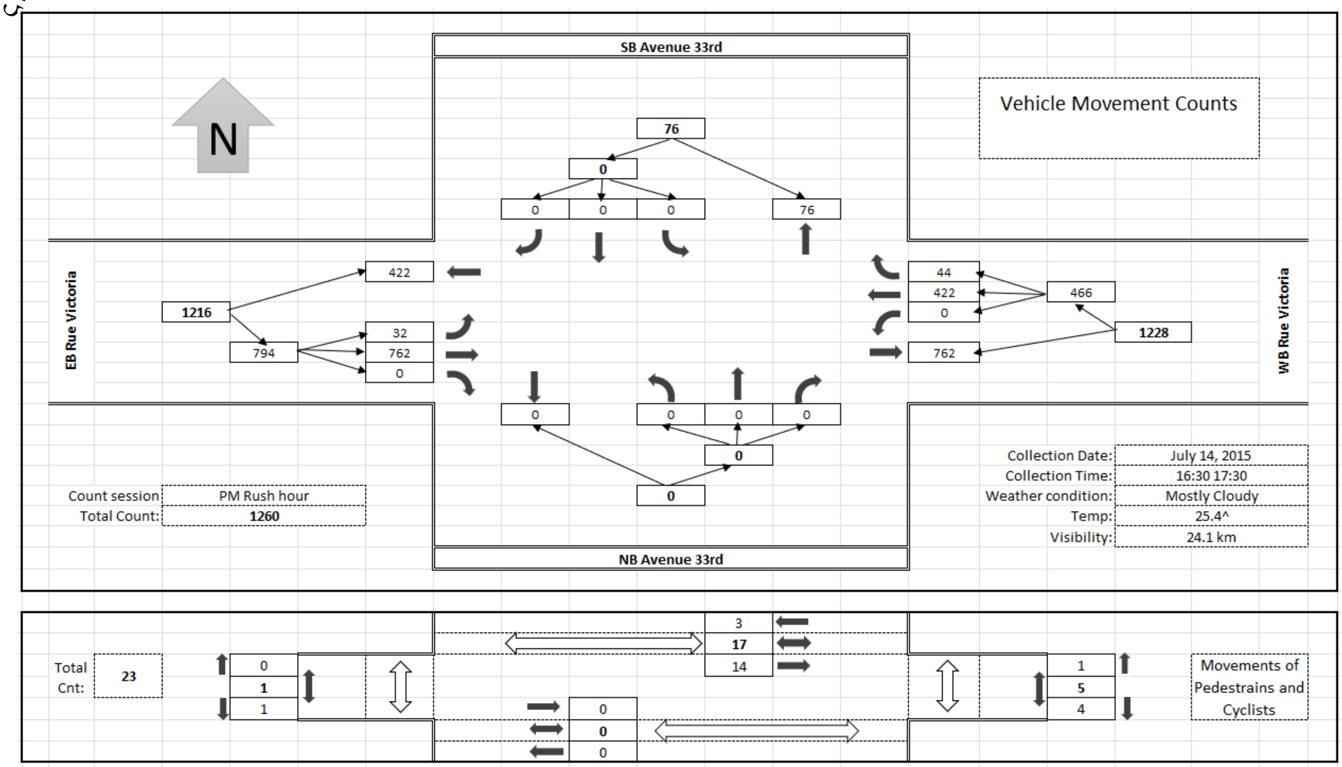

Vehicle Movement Counts
N
SB Avenue 32nd
NB Avenue 32nd
EB Rue Victoria
WB Rue Victoria
1176
599
144
0
455
577
738
268
470
245
225
0
332
124
0
456
1136
680
0
0
0
0
0
0
Count session: AM Rush hour
Total Count: 1525
Collection Date: July 14, 2015
Collection Time: 7:30 8:30
Weather condition: Mostly Cloudy
Temp: 23.9^
Visibility: 24.1 km
Total Cnt: 45
6
8
2
2
11
9
0
0
0
3
26
23
Movements of Pedestrains and Cyclists

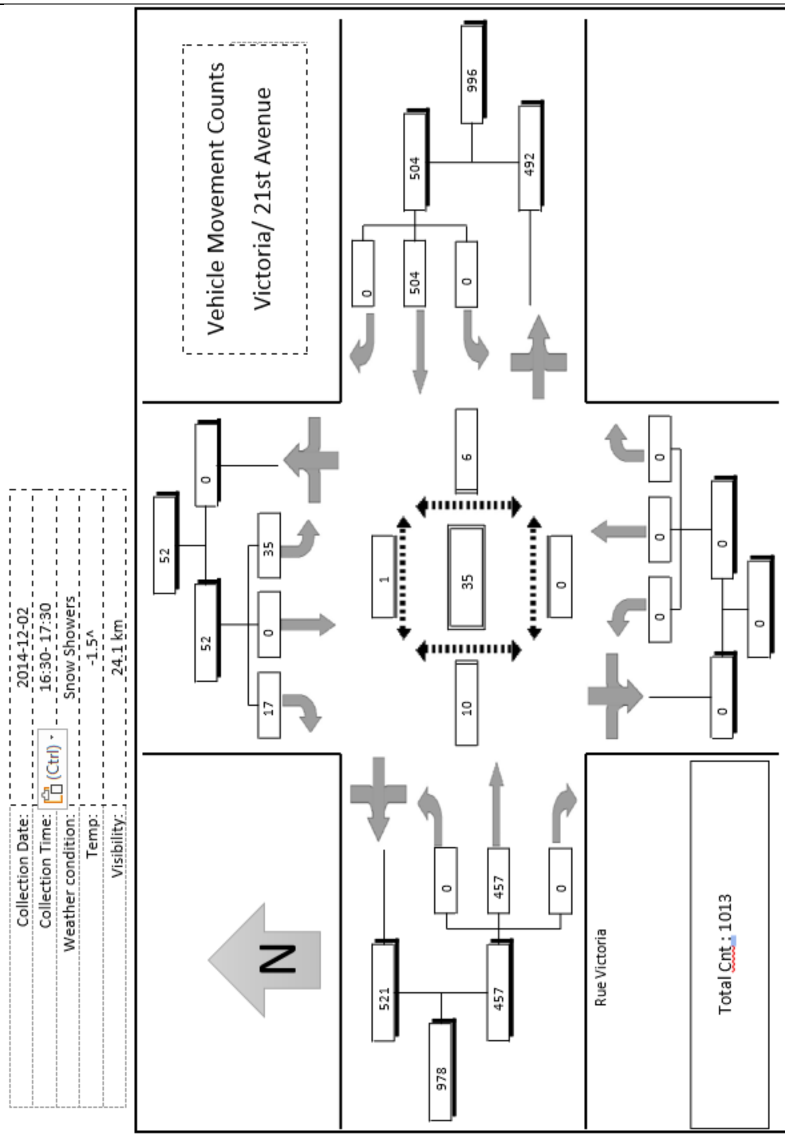

Vehicle Movement Counts
N
SB Avenue 32nd
NB Avenue 32nd
EB Rue Victoria
WB Rue Victoria
850
445
144
0
301
405
635
309
326
159
167
0
246
165
0
411
879
468
0
0
0
0
0
0
Count session: Mid-day
Total Count: 1182
Collection Date: July 14, 2015
Collection Time: 12:00 13:00
Weather condition: Cloudy
Temp: 27.1^
Visibility: 24.1 km
Total Cnt: 49
5
8
3
8
15
7
0
0
0
2
26
24
Movements of Pedestrains and Cyclists

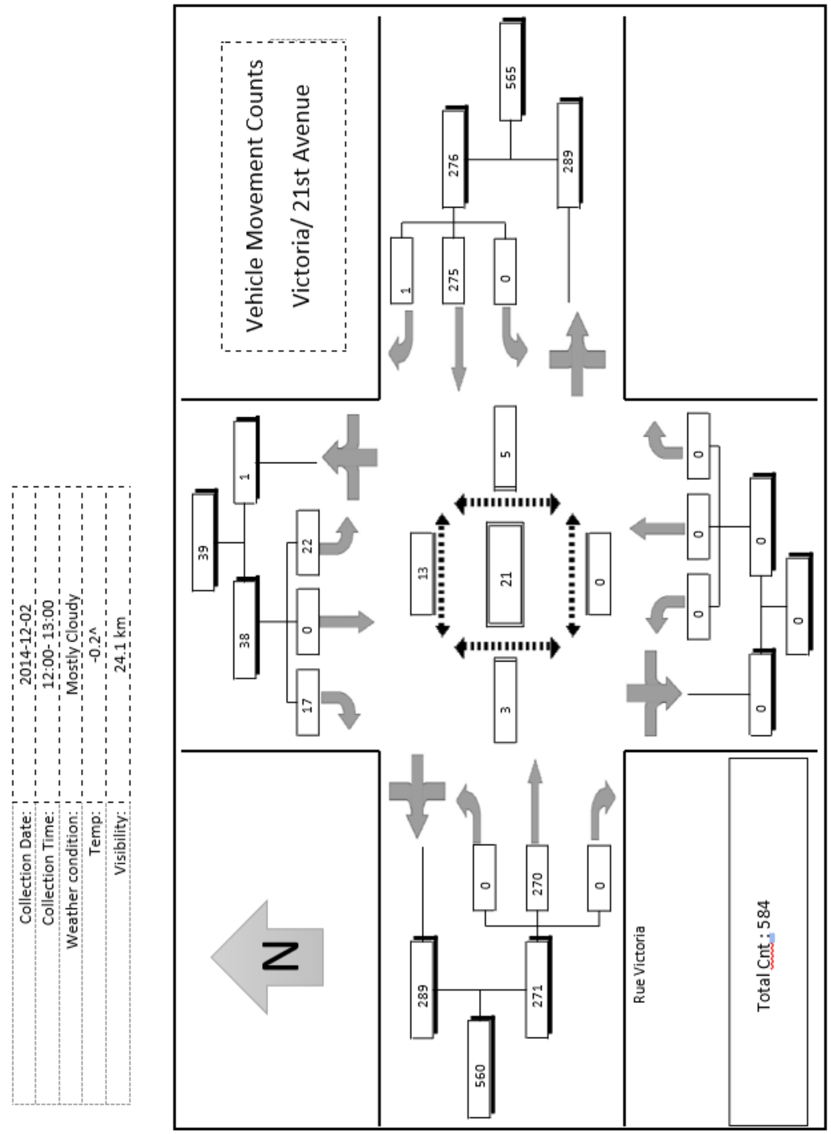

SB Avenue 32nd
Vehicle Movement Counts
N
1076
654
155
0
499
422
EB Rue Victoria
909
453
456
166
290
0
WB Rue Victoria
256
298
0
554
1343
789
0
0
0
0
0
0
Count session: PM Rush hour
Total Count: 1664
Collection Date: July 14, 2015
Collection Time: 16:30 17:30
Weather condition: Mostly Cloudy
Temp: 25.4^
Visibility: 24.1 km
NB Avenue 32nd
Total Cnt: 32
3
4
1
5
16
11
0
0
0
2
12
10
Movements of Pedestrains and Cyclists

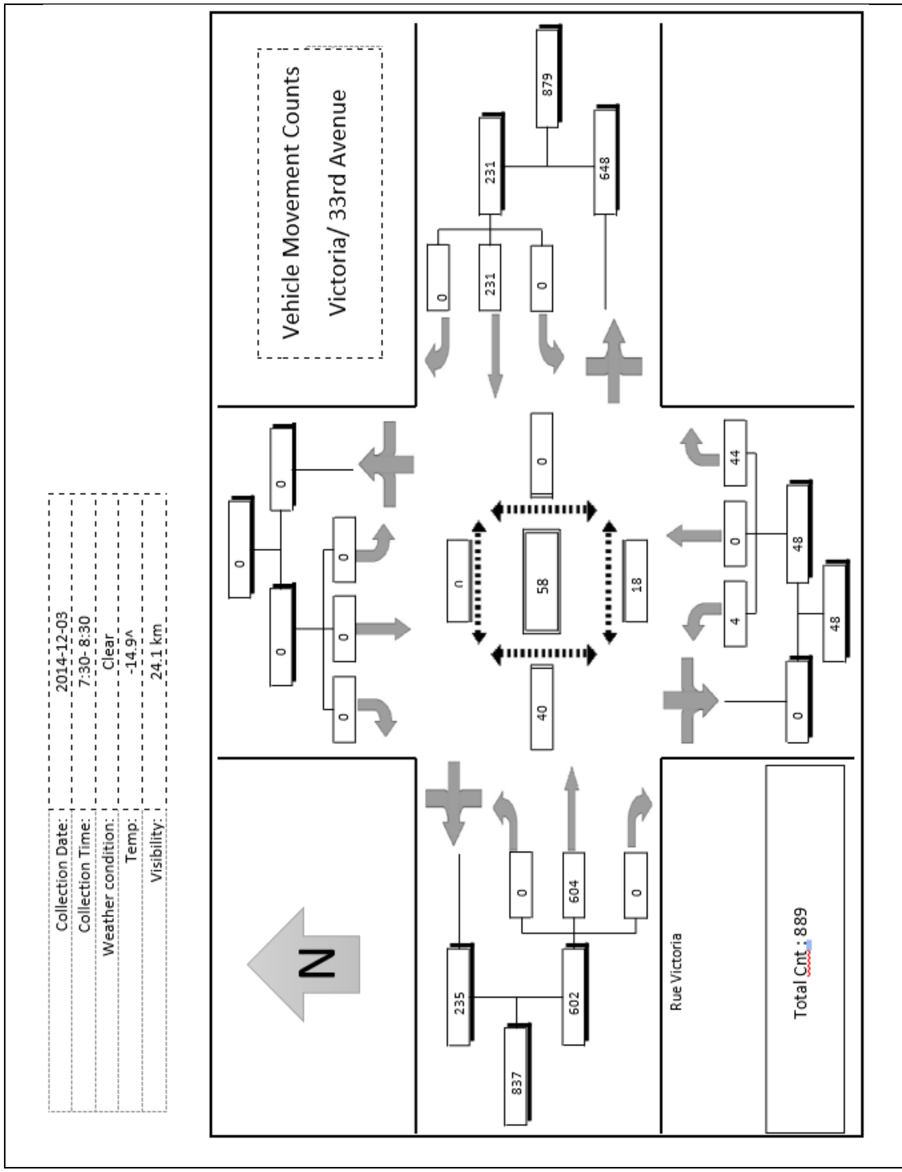

SB Avenue 33rd
Vehicle Movement Counts
N
69
0
0
0
0
69
EB Rue Victoria
691
234
457
27
430
0
WB Rue Victoria
42
234
0
276
706
430
0
0
0
0
0
0
Count session: AM Rush hour
Total Count: 733
Collection Date: July 14, 2015
Collection Time: 7:30 8:30
Weather condition: Mostly Cloudy
Temp: 23.9^
Visibility: 24.1 km
NB Avenue 33rd
Total Cnt: 31
3
5
2
4
20
16
0
0
0
3
6
3
Movements of Pedestrains and Cyclists

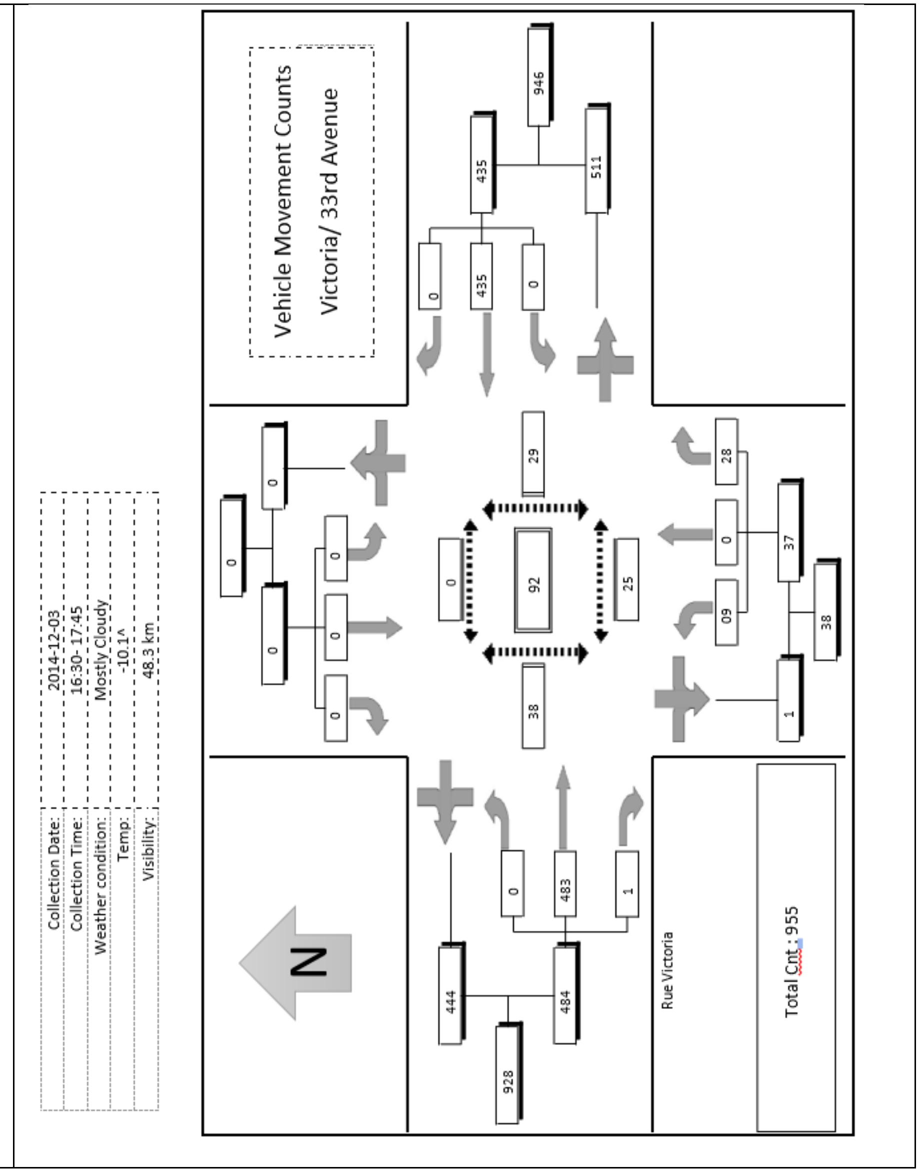

Vehicle Movement Counts
N
SB Avenue 33rd
NB Avenue 33rd
EB Rue Victoria
WB Rue Victoria
60
0
0
0
0
60
760
281
479
26
453
0
315
768
34
281
0
453
0
0
0
0
0
0
Count session: Mid-day
Total Count: 794
Collection Date: July 14, 2015
Collection Time: 12:00 13:00
Weather condition: Cloudy
Temp: 27.1^
Visibility: 24.1 km
Total Cnt: 21
0
1
1
6
15
9
0
0
0
2
5
3
Movements of Pedestrains and Cyclists


Vehicle Movement Counts
N
SB Avenue 33rd
NB Avenue 33rd
EB Rue Victoria
WB Rue Victoria
76
0
0
0
0
76
1216
422
794
32
762
0
466
1228
44
422
0
762
0
0
0
0
0
0
Count session: PM Rush hour
Total Count: 1260
Collection Date: July 14, 2015
Collection Time: 16:30 17:30
Weather condition: Mostly Cloudy
Temp: 25.4^
Visibility: 24.1 km
Total Cnt: 23
0
1
1
3
17
14
0
0
0
1
5
4
Movements of Pedestrains and Cyclists


Vehicle Movement Counts
Victoria/ 21st Avenue
Collection Date: 2014-12-02
Collection Time: 16:30- 17:30
Weather condition: Snow Showers
Temp: -1.5^
Visibility: 24.1 km
N
Rue Victoria
Total Cnt : 1013
996
504
492
521
457
978
52
52
35
17
35
1
0
6
10


Vehicle Movement Counts
Victoria/ 21st Avenue
Collection Date: 2014-12-02
Collection Time: 12:00- 13:00
Weather condition: Mostly Cloudy
Temp: -0.2^
Visibility: 24.1 km
N
Rue Victoria
Total Cnt : 584
565
276
289
275
289
271
270
560
39
38
22
17
21
13
5
3


Vehicle Movement Counts
Victoria/ 33rd Avenue
Collection Date: 2014-12-03
Collection Time: 7:30- 8:30
Weather condition: Clear
Temp: -14.9^
Visibility: 24.1 km
N
Rue Victoria
Total Cnt : 889
879
231
648
0
231
0
0
58
18
40
0
0
0
0
0
0
0
44
0
4
48
48
0
0
604
0
235
602
837


Vehicle Movement Counts
Victoria/ 33rd Avenue
Collection Date: 2014-12-03
Collection Time: 16:30- 17:45
Weather condition: Mostly Cloudy
Temp: -10.1^
Visibility: 48.3 km
N
Rue Victoria
Total Cnt : 955
946
435
511
0
435
0
29
0
92
25
38
0
0
0
0
0
0
28
0
09
37
38
1
0
483
1
444
484
928

Collection Date:	2014-12-03
Collection Time:	12:00- 13:00
Weather condition:	Mostly Cloudy
Temp:	-10.7^
Visibility:	24.1 km

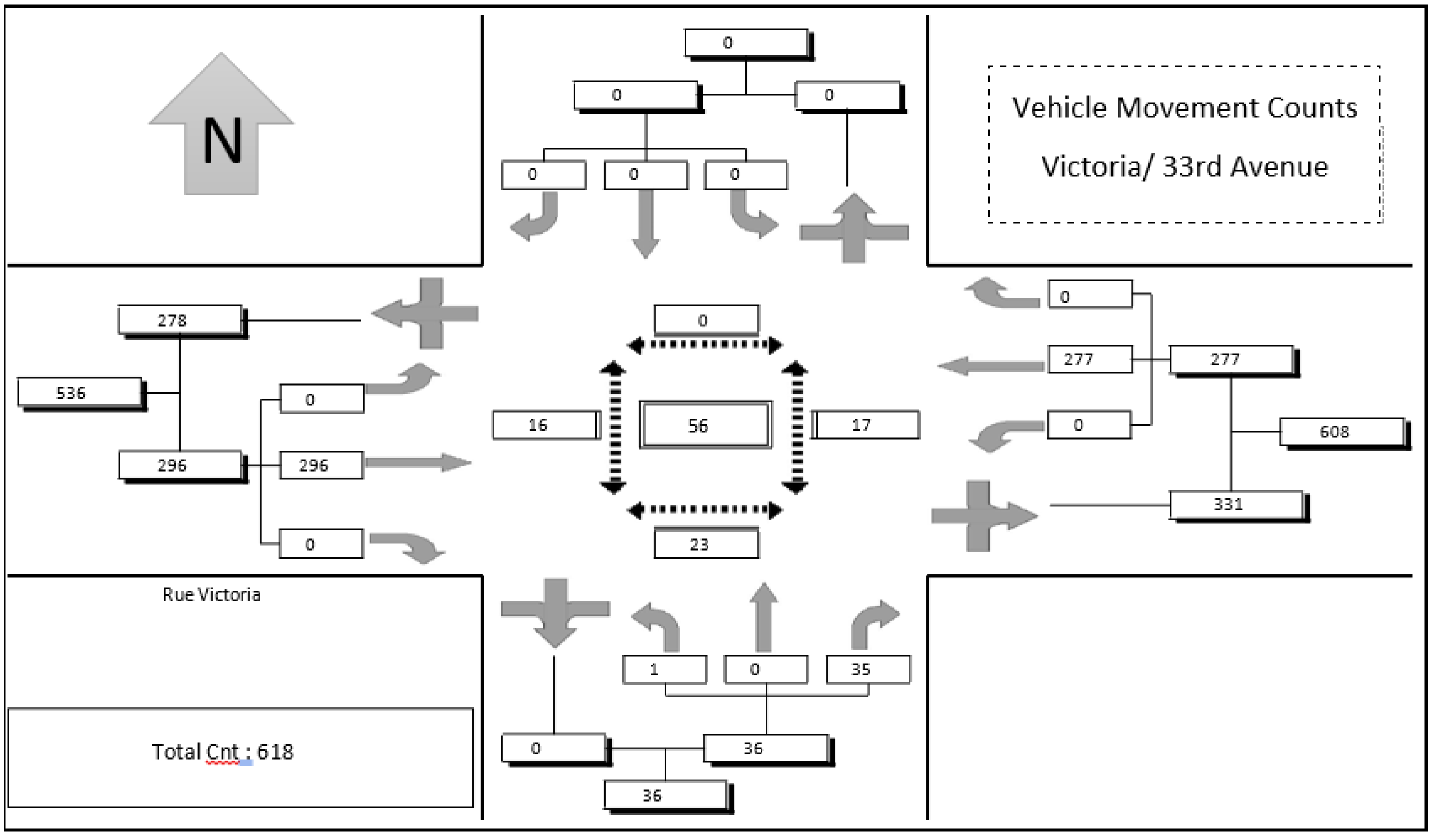


Table 40 External volume data provided by borough of Lachine

COMPTAGE VILLE DE LACHINE VICTORIA/34EME AVENUE

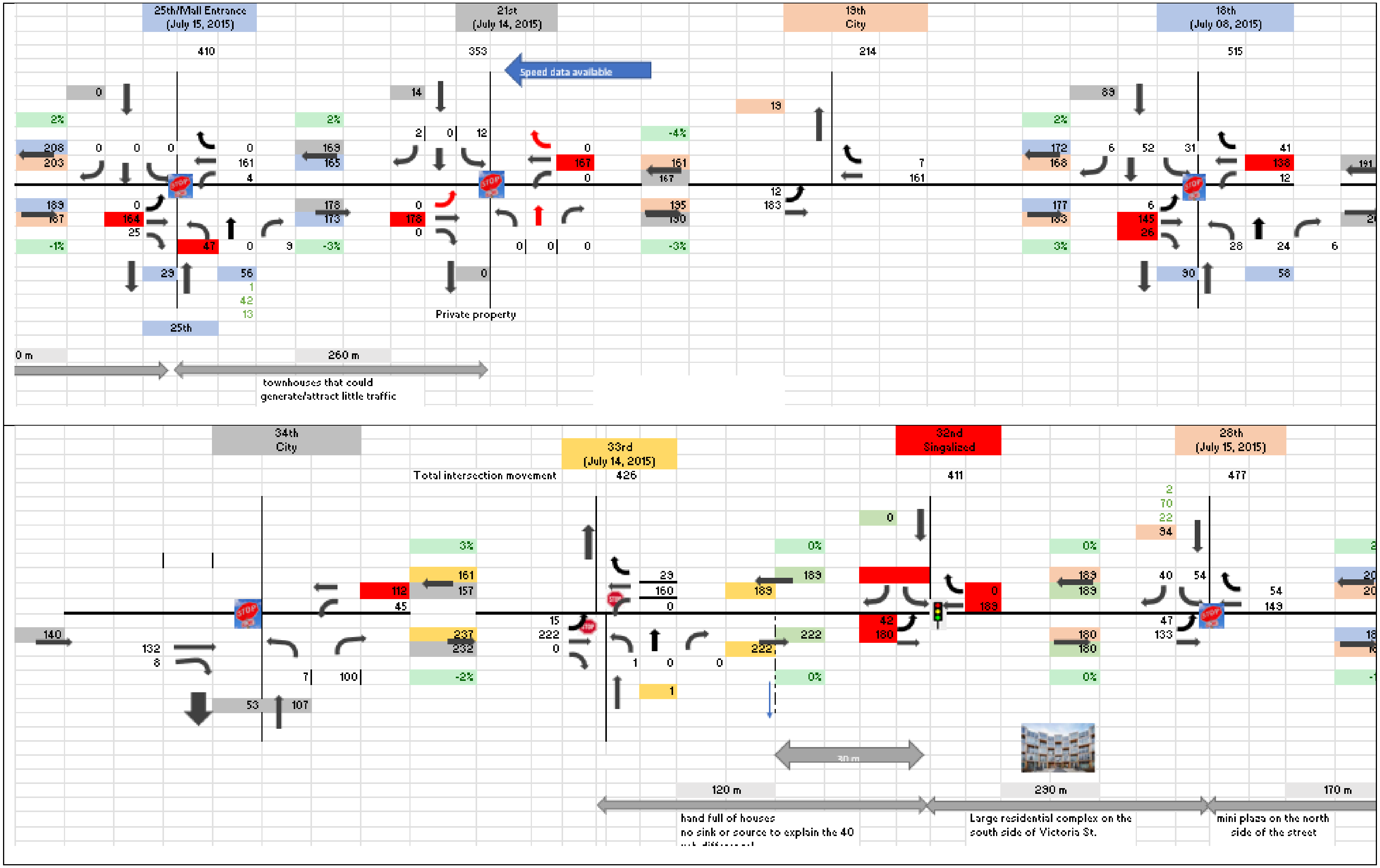
AM	Approche Nord				Approche Est				Approche Sud				Approche Ouest			
Start Time	Left	Thru	Right	Peds	Left	Thru	Right	Peds	Left	Thru	Right	Peds	Left	Thru	Right	Peds
06:00	0	0	0	0	5	4	0	1	0	0	11	0	0	8	0	2
06:15	0	0	0	1	12	6	0	0	0	0	18	0	0	13	1	0
06:30	0	0	0	0	14	9	0	0	1	0	25	3	0	18	1	2
06:45	0	0	0	2	33	12	0	0	3	0	30	2	0	28	4	0
07:00	0	0	0	0	29	16	0	0	5	0	40	4	0	42	5	0
07:15	0	0	0	1	22	26	0	0	2	0	58	3	0	55	6	2
07:30	0	0	0	2	29	32	0	0	6	0	69	3	0	69	13	4
07:45	0	0	0	1	19	23	0	0	3	0	72	0	0	65	11	4
08:00	0	0	0	1	31	36	0	0	8	0	45	3	0	75	6	3
08:15	0	0	0	1	17	23	0	0	4	0	52	2	0	61	6	0
08:30	0	0	0	0	23	31	0	0	5	0	46	3	0	68	7	4

MIDI	Approche Nord				Approche Est				Approche Sud				Approche Ouest			
Start Time	Left	Thru	Right	Peds	Left	Thru	Right	Peds	Left	Thru	Right	Peds	Left	Thru	Right	Peds
11:00	0	0	0	0	36	25	0	0	1	0	29	0	0	28	0	0
11:15	0	0	0	1	30	21	0	0	2	0	32	1	0	46	4	3
11:30	0	0	0	0	25	36	0	0	2	0	41	1	0	37	3	0
11:45	0	0	0	0	28	40	0	0	3	0	50	0	0	48	2	2
12:00	0	0	0	0	34	29	0	0	3	0	36	0	0	33	4	1
12:15	0	0	0	0	31	30	0	0	2	0	35	0	0	31	3	3
12:30	0	0	0	1	24	43	0	0	5	0	47	0	0	42	3	0
12:45	0	0	0	0	35	32	0	0	5	0	43	0	0	44	7	3

COMPTAGE VILLE DE LACHINE VICTORIA/34EME AVENUE

PM	Approche Nord				Approche Est				Approche Sud				Approche Ouest			
Start Time	Left	Thru	Right	Peds	Left	Thru	Right	Peds	Left	Thru	Right	Peds	Left	Thru	Right	Peds
15:30	0	0	0	1	31	36	0	0	2	0	44	2	0	57	7	6
15:45	0	0	0	1	33	58	0	0	4	0	44	3	0	72	6	2
16:00	0	0	0	0	34	32	0	0	1	0	40	0	0	62	7	1
16:15	0	0	0	4	28	43	0	0	0	0	42	0	0	94	5	0
16:30	0	0	0	0	49	46	0	0	2	0	38	3	0	96	12	3
16:45	0	0	0	0	33	55	0	0	0	0	47	4	0	137	8	8
17:00	0	0	0	1	25	52	0	0	9	0	56	0	0	111	7	5
17:15	0	0	0	1	30	76	0	0	4	0	53	4	0	116	6	1
17:30	0	0	0	0	35	62	0	0	3	0	56	3	0	97	7	6
17:45	0	0	0	2	34	55	0	0	5	0	34	5	0	71	7	4

Table 41 Flow synchronization for microsimulation model (21h00-22h00 in period 4)



2014	Sunrise/Sunset		Daylength		Astronomical Twilight		Nautical Twilight		Civil Twilight		Solar Noon	
Dec	Sunrise	Sunset	Length	Diff.	Start	End	Start	End	Start	End	Time	Mil. km
1	7:14 am ↘ (121°)	4:13 pm ↙ (239°)	8:59:03	−1:34	5:28 am	5:59 pm	6:03 am	5:23 pm	6:40 am	4:46 pm	11:43 am (22.7°)	147.510
2	7:15 am ↘ (121°)	4:12 pm ↙ (239°)	8:57:32	−1:30	5:29 am	5:58 pm	6:04 am	5:23 pm	6:41 am	4:46 pm	11:44 am (22.6°)	147.485
3	7:16 am ↘ (121°)	4:12 pm ↙ (238°)	8:56:05	−1:26	5:30 am	5:58 pm	6:05 am	5:23 pm	6:42 am	4:45 pm	11:44 am (22.5°)	147.462

2015	Sunrise/Sunset		Daylength		Astronomical Twilight		Nautical Twilight		Civil Twilight		Solar Noon	
Apr	Sunrise	Sunset	Length	Diff.	Start	End	Start	End	Start	End	Time	Mil. km
6	6:26 am → (80°)	7:28 pm ← (280°)	13:02:41	+3:08	4:41 am	9:13 pm	5:19 am	8:35 pm	5:56 am	7:59 pm	12:57 pm (51.1°)	149.698
7	6:24 am → (79°)	7:30 pm ← (281°)	13:05:49	+3:08	4:39 am	9:15 pm	5:17 am	8:36 pm	5:54 am	8:00 pm	12:56 pm (51.5°)	149.742
8	6:22 am → (79°)	7:31 pm ↖ (281°)	13:08:57	+3:07	4:37 am	9:17 pm	5:15 am	8:38 pm	5:52 am	8:01 pm	12:56 pm (51.8°)	149.785

Jul	Sunrise	Sunset	Length	Diff.	Start	End	Start	End	Start	End	Time	Mil. km
6	5:13 am ↗ (56°)	8:45 pm ↖ (304°)	15:32:06	−1:03	2:41 am	11:16 pm	3:46 am	10:11 pm	4:35 am	9:22 pm	12:59 pm (67.2°)	152.093
10	5:16 am ↗ (56°)	8:43 pm ↖ (304°)	15:27:11	−1:19	2:47 am	11:10 pm	3:50 am	10:08 pm	4:39 am	9:20 pm	1:00 pm (66.8°)	152.088
14	5:19 am ↗ (57°)	8:40 pm ↖ (303°)	15:21:16	−1:34	2:54 am	11:04 pm	3:55 am	10:04 pm	4:43 am	9:17 pm	1:00 pm (66.2°)	152.071
15	5:20 am ↗ (57°)	8:40 pm ↖ (302°)	15:19:38	−1:37	2:56 am	11:03 pm	3:56 am	10:03 pm	4:44 am	9:16 pm	1:00 pm (66.1°)	152.064
16	5:21 am ↗ (58°)	8:39 pm ↖ (302°)	15:17:57	−1:41	2:58 am	11:01 pm	3:57 am	10:02 pm	4:45 am	9:15 pm	1:00 pm (65.9°)	152.056

Figure 31 Ambient light during the study days

Date/Time (LST)	Temp (°C)	Visibility (km)	Weather
2014-12-02 0:00	-7.6	24.1	NA
2014-12-02 1:00	-8.7	24.1	Clear
2014-12-02 2:00	-10.4	24.1	NA
2014-12-02 3:00	-11.2	24.1	NA
2014-12-02 4:00	-12.6	24.1	Clear
2014-12-02 5:00	-14.1	24.1	NA
2014-12-02 6:00	-14.8	24.1	NA
2014-12-02 7:00	-14.9	48.3	Clear
2014-12-02 8:00	-14.4	48.3	NA
2014-12-02 9:00	-13.1	24.1	NA
2014-12-02 10:00	-12.6	24.1	Mainly Clear
2014-12-02 11:00	-11.9	24.1	NA
2014-12-02 12:00	-10.7	24.1	NA
2014-12-02 13:00	-10	24.1	Mostly Cloudy
2014-12-02 14:00	-9.3	24.1	NA
2014-12-02 15:00	-9.8	48.3	NA
2014-12-02 16:00	-10	48.3	Cloudy
2014-12-02 17:00	-9.7	24.1	NA
2014-12-02 18:00	-10	24.1	NA
2014-12-02 19:00	-9.7	24.1	Cloudy
2014-12-02 20:00	-9	24.1	Snow
2014-12-02 21:00	-9.4	1.6	Snow
2014-12-02 22:00	-9	1.6	Snow
2014-12-02 23:00	-8.2	2	Snow
2014-12-03 0:00	-5.4	2.4	Snow
2014-12-03 1:00	-3.2	4.8	Snow
2014-12-03 2:00	0.2	9.7	Snow
2014-12-03 3:00	1.2	24.1	Snow
2014-12-03 4:00	1.4	24.1	Snow
2014-12-03 5:00	1.8	24.1	Rain,Snow
2014-12-03 6:00	2	24.1	Rain
2014-12-03 7:00	2.8	24.1	Cloudy
2014-12-03 8:00	2.8	24.1	Drizzle
2014-12-03 9:00	2.8	24.1	NA
2014-12-03 10:00	3.1	4.8	Drizzle,Fog
2014-12-03 11:00	3.1	8.1	Rain,Fog
2014-12-03 12:00	3.2	11.3	Rain
2014-12-03 13:00	3.3	16.1	Cloudy
2014-12-03 14:00	2.8	8.1	Rain
2014-12-03 15:00	1.8	8.1	Rain,Snow,Fog
2014-12-03 16:00	0.6	3.2	Snow
2014-12-03 17:00	0.3	3.2	Snow
2014-12-03 18:00	0.8	4.8	Snow
2014-12-03 19:00	1.2	11.3	Snow
2014-12-03 20:00	1.5	16.1	Snow Showers
2014-12-03 21:00	1	24.1	NA
2014-12-03 22:00	0.8	24.1	Mostly Cloudy
2014-12-03 23:00	0.2	6.4	Snow Showers
2015-04-06 0:00	-4.1	24.1	NA
2015-04-06 1:00	-5	24.1	Clear
2015-04-06 2:00	-5.5	24.1	NA
2015-04-06 3:00	-5.4	24.1	NA

Date/Time (LST)	Temp (°C)	Visibility (km)	Weather
2015-04-08 8:00	-0.5	48.3	NA
2015-04-08 9:00	1.9	48.3	NA
2015-04-08 10:00	3	48.3	Cloudy
2015-04-08 11:00	4.5	48.3	NA
2015-04-08 12:00	4.2	48.3	NA
2015-04-08 13:00	5.5	48.3	Cloudy
2015-04-08 14:00	5.3	48.3	NA
2015-04-08 15:00	5.3	48.3	NA
2015-04-08 16:00	4.7	48.3	Cloudy
2015-04-08 17:00	3.7	48.3	NA
2015-04-08 18:00	3.2	19.3	Snow
2015-04-08 19:00	0.5	2.4	Snow
2015-04-08 20:00	0.4	4	Snow
2015-04-08 21:00	0.5	4.8	Snow
2015-04-08 22:00	0.5	4	Snow
2015-04-08 23:00	0.6	8.1	Snow
2015-07-06 0:00	19.8	24.1	NA
2015-07-06 1:00	19.4	24.1	Mainly Clear
2015-07-06 2:00	18.7	24.1	NA
2015-07-06 3:00	18.4	24.1	NA
2015-07-06 4:00	17.5	24.1	Mostly Cloudy
2015-07-06 5:00	17.4	24.1	NA
2015-07-06 6:00	19.1	24.1	NA
2015-07-06 7:00	20.4	24.1	Cloudy
2015-07-06 8:00	21.4	24.1	NA
2015-07-06 9:00	22.4	24.1	NA
2015-07-06 10:00	23.9	24.1	Clear
2015-07-06 11:00	24.9	24.1	NA
2015-07-06 12:00	26	24.1	NA
2015-07-06 13:00	27.3	24.1	Mainly Clear
2015-07-06 14:00	27.7	24.1	NA
2015-07-06 15:00	27.9	24.1	NA
2015-07-06 16:00	27.9	24.1	Clear
2015-07-06 17:00	27.7	24.1	NA
2015-07-06 18:00	26.8	24.1	NA
2015-07-06 19:00	26.2	24.1	Mostly Cloudy
2015-07-06 20:00	25.4	24.1	NA
2015-07-06 21:00	24.9	24.1	NA
2015-07-06 22:00	23.8	24.1	Mainly Clear
2015-07-06 23:00	21.7	24.1	NA
2015-07-08 0:00	22.8	11.3	Rain Showers
2015-07-08 1:00	21.3	9.7	Rain Showers
2015-07-08 2:00	17.9	16.1	Rain Showers
2015-07-08 3:00	16.9	24.1	NA
2015-07-08 4:00	16	24.1	Mainly Clear
2015-07-08 5:00	15.2	24.1	NA
2015-07-08 6:00	15.8	24.1	NA
2015-07-08 7:00	16.6	24.1	Mainly Clear
2015-07-08 8:00	17.4	24.1	NA
2015-07-08 9:00	18.3	24.1	NA
2015-07-08 10:00	18.1	24.1	Mostly Cloudy
2015-07-08 11:00	19.4	24.1	NA

Date/Time (LST)	Temp (°C)	Visibility (km)	Weather
2015-07-10 16:00	28.1	48.3	Mainly Clear
2015-07-10 17:00	27.5	48.3	NA
2015-07-10 18:00	26.9	48.3	NA
2015-07-10 19:00	25.7	48.3	Clear
2015-07-10 20:00	24	24.1	NA
2015-07-10 21:00	22.6	24.1	NA
2015-07-10 22:00	22	24.1	Mostly Cloudy
2015-07-10 23:00	21.8	24.1	NA
2015-07-14 0:00	22.5	24.1	NA
2015-07-14 1:00	22.5	24.1	Mainly Clear
2015-07-14 2:00	22.5	24.1	NA
2015-07-14 3:00	22.1	24.1	NA
2015-07-14 4:00	22	24.1	Mostly Cloudy
2015-07-14 5:00	22	24.1	NA
2015-07-14 6:00	22.7	24.1	NA
2015-07-14 7:00	23.9	24.1	Mostly Cloudy
2015-07-14 8:00	24.5	24.1	NA
2015-07-14 9:00	25.7	24.1	NA
2015-07-14 10:00	26.5	24.1	Mainly Clear
2015-07-14 11:00	26.8	22.5	NA
2015-07-14 12:00	27.1	24.1	NA
2015-07-14 13:00	25.1	24.1	Cloudy
2015-07-14 14:00	25.7	24.1	NA
2015-07-14 15:00	25.2	22.5	NA
2015-07-14 16:00	25.4	24.1	Cloudy
2015-07-14 17:00	25	24.1	NA
2015-07-14 18:00	22.8	24.1	NA
2015-07-14 19:00	22.4	24.1	Cloudy
2015-07-14 20:00	21.2	20.9	NA
2015-07-14 21:00	21.7	24.1	NA
2015-07-14 22:00	21.7	24.1	Cloudy
2015-07-14 23:00	18.1	24.1	NA
2015-07-15 0:00	16.9	24.1	NA
2015-07-15 1:00	15.6	24.1	Cloudy
2015-07-15 2:00	15.1	24.1	NA
2015-07-15 3:00	14	24.1	NA
2015-07-15 4:00	13.9	24.1	Mostly Cloudy
2015-07-15 5:00	13.8	24.1	NA
2015-07-15 6:00	13.7	24.1	NA
2015-07-15 7:00	14.1	24.1	Mainly Clear
2015-07-15 8:00	14.9	24.1	NA
2015-07-15 9:00	15.8	24.1	NA
2015-07-15 10:00	16.7	48.3	Mostly Cloudy
2015-07-15 11:00	18.5	48.3	NA
2015-07-15 12:00	18.9	48.3	NA
2015-07-15 13:00	20	48.3	Mainly Clear
2015-07-15 14:00	20.9	48.3	NA
2015-07-15 15:00	21.8	48.3	NA
2015-07-15 16:00	23	48.3	Clear
2015-07-15 17:00	22.6	48.3	NA
2015-07-15 18:00	21.9	48.3	NA
2015-07-15 19:00	20.5	48.3	Clear

2015-04-06 4:00	-4	24.1	Mainly Clear
2015-04-06 5:00	-4	24.1	NA
2015-04-06 6:00	-3.5	24.1	NA
2015-04-06 7:00	-2.8	24.1	Cloudy
2015-04-06 8:00	-1.5	24.1	NA
2015-04-06 9:00	0.2	24.1	NA
2015-04-06 10:00	1	48.3	Cloudy
2015-04-06 11:00	1.8	48.3	NA
2015-04-06 12:00	2.6	48.3	NA
2015-04-06 13:00	2.5	19.3	Snow
2015-04-06 14:00	1.1	4.8	Snow
2015-04-06 15:00	0.6	19.3	Snow
2015-04-06 16:00	0.7	4.8	Snow
2015-04-06 17:00	0.4	2.8	Snow
2015-04-06 18:00	0.3	2.4	Snow,Fog
2015-04-06 19:00	0.2	4.8	Snow,Fog
2015-04-06 20:00	0.3	4.8	Fog
2015-04-06 21:00	0.2	8.1	Fog
2015-04-06 22:00	-0.2	4.8	Fog
2015-04-06 23:00	-0.8	2.4	Fog
2015-04-08 0:00	-3	24.1	NA
2015-04-08 1:00	-2.5	24.1	Mainly Clear
2015-04-08 2:00	-1.6	24.1	NA
2015-04-08 3:00	-1.9	24.1	NA
2015-04-08 4:00	-2.5	24.1	Mostly Cloudy
2015-04-08 5:00	-2.6	24.1	NA
2015-04-08 6:00	-2.6	48.3	NA
2015-04-08 7:00	-1.6	48.3	Cloudy

2015-07-08 12:00	19.8	24.1	NA
2015-07-08 13:00	21.3	24.1	Mostly Cloudy
2015-07-08 14:00	22.2	24.1	NA
2015-07-08 15:00	22.8	24.1	NA
2015-07-08 16:00	22.7	24.1	Mainly Clear
2015-07-08 17:00	22.7	24.1	NA
2015-07-08 18:00	22.1	24.1	NA
2015-07-08 19:00	21.2	24.1	Mostly Cloudy
2015-07-08 20:00	20.2	24.1	NA
2015-07-08 21:00	19.5	24.1	NA
2015-07-08 22:00	18.3	24.1	Mainly Clear
2015-07-08 23:00	17.5	24.1	NA
2015-07-10 0:00	17.7	24.1	NA
2015-07-10 1:00	17.1	24.1	Mainly Clear
2015-07-10 2:00	17	24.1	NA
2015-07-10 3:00	16.8	24.1	NA
2015-07-10 4:00	16.6	24.1	Mainly Clear
2015-07-10 5:00	16.4	24.1	NA
2015-07-10 6:00	16.8	24.1	NA
2015-07-10 7:00	18.2	24.1	Mainly Clear
2015-07-10 8:00	20.3	24.1	NA
2015-07-10 9:00	22.6	24.1	NA
2015-07-10 10:00	24.3	24.1	Mainly Clear
2015-07-10 11:00	25.9	24.1	NA
2015-07-10 12:00	26	24.1	NA
2015-07-10 13:00	27.8	24.1	Mainly Clear
2015-07-10 14:00	28	24.1	NA
2015-07-10 15:00	28.1	24.1	NA

2015-07-15 20:00	18.6	24.1	NA
2015-07-15 21:00	17.2	24.1	NA
2015-07-15 22:00	16.1	24.1	Mainly Clear
2015-07-15 23:00	15.3	24.1	NA
2015-07-16 0:00	15	24.1	NA
2015-07-16 1:00	13.7	24.1	Clear
2015-07-16 2:00	13	24.1	NA
2015-07-16 3:00	12.2	24.1	NA
2015-07-16 4:00	11.8	24.1	Mainly Clear
2015-07-16 5:00	12	48.3	NA
2015-07-16 6:00	13.6	48.3	NA
2015-07-16 7:00	15.4	48.3	Mainly Clear
2015-07-16 8:00	17.1	48.3	NA
2015-07-16 9:00	18.5	48.3	NA
2015-07-16 10:00	19.9	48.3	Mainly Clear
2015-07-16 11:00	20.3	48.3	NA
2015-07-16 12:00	20.9	48.3	NA
2015-07-16 13:00	21.9	48.3	Mostly Cloudy
2015-07-16 14:00	22.6	48.3	NA
2015-07-16 15:00	22.5	48.3	NA
2015-07-16 16:00	22.7	48.3	Clear
2015-07-16 17:00	23.1	48.3	NA
2015-07-16 18:00	22.3	24.1	NA
2015-07-16 19:00	21.4	48.3	Clear
2015-07-16 20:00	19.9	48.3	NA
2015-07-16 21:00	19.1	24.1	NA
2015-07-16 22:00	18.8	24.1	Mainly Clear
2015-07-16 23:00	17.6	24.1	NA

Table 42 Weather during study period

	Inter.	Col	comp_FS.BT	comp_FS.RT	comp_FS.NFS	conf	f_ma	f_mi	f_ped
1	18^{th}	7	16.91	1.69	1.56	57	1504	274	102
2	18^{th}	4	16.91	1.69	1.56	32	1020	270	88
3	18^{th}	6	16.91	1.69	1.56	48	1973	382	110
4	18^{th}	5	18.79	2.1	1.91	25	2086	475	134
5	(BLS)18^{th}		24.35	3.11	4.19	44	1609	310	98
6	(BLS)18^{th}		24.35	3.11	4.19	28	1140	245	75
7	(BLS)18^{th}		24.35	3.11	4.19	33	2091	388	101
8	(BLS)18^{th}		34.78	5.12	4.86	12	2044	441	142
9	19^{th}	0	11.55	1.75	1.52	6	1522	0	20
10	19^{th}	1	11.55	1.75	1.52	14	1056	0	33
11	19^{th}	0	11.55	1.75	1.52	13	1934	0	40
12	19^{th}	0	10.65	2.13	1.77	11	2152	0	75
13	21^{st}	2	11.55	1.75	1.52	8	1407	107	43
14	21^{st}	2	11.55	1.75	1.52	22	1020	63	37
15	21^{st}	1	11.55	1.75	1.52	32	1818	104	65
16	21^{st}	3	10.65	2.13	1.77	11	2092	102	163

17	25th	5	11.55	1.75	1.52	55	1586	504	40
18	25th	4	11.55	1.75	1.52	59	1401	369	81
19	25th	4	11.55	1.75	1.52	37	2014	495	66
20	25th	4	10.65	2.13	1.77	30	2450	389	180
21	(BLS)25th		23.88	2.91	2.6	47	2081	597	23
22	(BLS)25th		23.88	2.91	2.6	36	1378	412	51
23	(BLS)25th		23.88	2.91	2.6	28	2237	589	48
24	(BLS)25th		33.1	4.89	4.26	14	2186	425	146
25	28th	2	16.22	2.5	2.17	26	2166	134	34
26	28th	5	16.22	2.5	2.17	41	1533	374	29
27	28th	6	16.22	2.5	2.17	33	2795	495	52
28	28th	4	28	4.2	3.65	10	2778	620	54
29	33rd	1	24.33	2.47	2.25	18	1407	81	116
30	33rd	0	24.33	2.47	2.25	8	1101	83	129
31	33rd	1	24.33	2.47	2.25	9	1749	69	219
32	33rd	0	30.5	2.98	2.71	11	2170	94	190
33	34th	0	11.55	1.75	1.52	17	970	465	48

34	34th	0	11.55	1.75	1.52	27	800	346	15
35	34th	1	11.55	1.75	1.52	12	1532	390	55
36	34th	3	10.65	2.13	1.77	24	1972	547	98
	Min.	**0**	**10.65**	**1.69**	**4.86**	**6**	**800**	**0**	**15**
	1st Qu.	**0.75**	**11.55**	**1.75**	**2.6**	**12**	**1405.5**	**100**	**42.25**
	Median	**2**	**16.22**	**2.13**	**1.77**	**25.5**	**1783.5**	**357.5**	**70.5**
	Mean	**2.535**	**17.873**	**2.412**	**2.25**	**26.055**	**1743.722**	**295.638**	**83.333**
	3rd Qu.	**4**	**24.33**	**2.91**	**1.52**	**33.75**	**2091.25**	**447**	**111.5**
	Max.	**7**	**34.78**	**5.12**	**1.52**	**59**	**2795**	**620**	**219**

Table 43 Collision, Flow, Conflict and log of flow and Conflict

8 Appendix II (results)

8.1 Qualitative Analysis (SPSS)

- Scenario 1

Case Processing Summary

		N	Marginal Percentage
Driver Complacence	Blow-Through	143	4.1%
	Full Stop	2288	66.0%
	Roll-Through	1037	29.9%
Valid		3468	100.0%
Missing		0	
Total		3468	
Subpopulation		38[a]	

a. The dependent variable has only one value observed in 11 (28.9%) subpopulations.

Model Fitting Information

Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	1267.168	1279.471	1263.168			
Final	945.000	1043.422	913.000	350.168	14	<.001

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	921.510	60	<.001
Deviance	765.622	60	<.001

Pseudo R-Square

Cox and Snell	.096
Nagelkerke	.122
McFadden	.066

Likelihood Ratio Tests

Effect	Model Fitting Criteria	Likelihood Ratio Tests

	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	1455.310	1541.428	1427.310	514.309	2	<.001
Pavement condition	**947.176**	**1033.295**	**919.176**	**6.176**	**2**	**.056**
Natural ambient light	1065.966	1152.084	1037.966	124.966	2	<.001
Conflict potential	1013.991	1100.110	985.991	72.991	2	<.001
Maneuver	952.475	1038.594	924.475	11.475	2	.003
LED STOP	996.207	1082.326	968.207	55.207	2	<.001
BLS STOP	1074.755	1160.874	1046.755	133.755	2	<.001
Beacon STOP	968.501	1054.620	940.501	27.501	2	<.001

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Parameter Estimates

Driver Complacence[a]	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound

Blow-Through	Intercept	-2.274	.165	190.349	1	<.001			
	Natural ambient light	-.673	.257	6.867	1	.009	.510	.309	.844
	Conflict potential	.641	.192	11.198	1	<.001	1.898	1.304	2.763
	Maneuver	-.462	.215	4.605	1	.032	.630	.413	.961
	LED STOP	-.672	.231	8.455	1	.004	.511	.324	.803
	BLS STOP	-1.713	.346	24.513	1	<.001	.180	.091	.355
	Beacon STOP	-.419	.245	2.923	1	.087	.657	.407	1.063
Roll-Through	Intercept	-.423	.076	30.665	1	<.001			
	Natural ambient light	-.560	.105	28.207	1	<.001	.571	.465	.702
	Conflict potential	.312	.089	12.289	1	<.001	1.367	1.148	1.628
	Maneuver	-.065	.088	.551	1	.458	.937	.788	1.113
	LED STOP	-.516	.101	25.921	1	<.001	.597	.489	.728
	BLS STOP	-.827	.116	51.221	1	<.001	.437	.349	.548
	Beacon STOP	-.431	.113	14.446	1	<.001	.650	.521	.812

a. The reference category is: Full Stop.

Classification

Observed	Predicted: Blow-Through	Full Stop	Roll-Through	Percent Correct
Blow-Through	41	85	17	28.7%
Full Stop	0	2204	84	96.3%
Roll-Through	0	645	392	37.8%
Overall Percentage	1.2%	84.6%	14.2%	73.7%

Table 44 Scenario one Model Classification

- Scenario 2

Likelihood Ratio Tests with all independent variables

Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	1941.886	2041.506	1913.886	690.918	2	<.001
Pavement condition	1272.009	1371.629	1251.322	3.378	2	.185
Natural ambient light	1271.563	1371.183	1243.563	20.596	2	<.001
Conflict potential	1260.592	1360.212	1232.592	9.624	2	.008
Maneuver	1408.744	1508.364	1380.744	157.776	2	<.001
LED STOP	1348.091	1447.711	1320.091	97.123	2	<.001
BLS STOP	1351.894	1451.514	1323.894	100.926	2	<.001
Beacon STOP	1294.430	1394.050	1266.430	43.462	2	<.001

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

After removing independent variables (Pavement condition)

Case Processing Summary

		N	Marginal Percentage
Driver Complacence	Blow-Through	405	4.5%
	Full Stop	5842	64.2%
	Roll-Through	2850	31.3%
Valid		9097	100.0%
Missing		0	
Total		9097	
Subpopulation		32[a]	

a. The dependent variable has only one value observed in 2 (6.3%) subpopulations.

Model Fitting Information

	Model Fitting Criteria			Likelihood Ratio Tests		
Model	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	1027.195	1041.426	1023.195			

Final	719.458	819.078	691.458	331.737	12	<.001

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	415.784	50	<.001
Deviance	448.824	50	<.001

Pseudo R-Square

Cox and Snell	.036
Nagelkerke	.045
McFadden	.023

Likelihood Ratio Tests

	Model Fitting Criteria			Likelihood Ratio Tests		
Effect	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	1860.603	1945.992	1836.603	1145.145	2	<.001
Natural ambient light	735.787	821.175	711.787	20.329	2	<.001

Conflict potential	724.135	809.524	700.135	8.677	2	.013
Maneuver	882.503	967.891	858.503	167.045	2	<.001
LED STOP	826.403	911.792	802.403	110.945	2	<.001
BLS STOP	807.933	893.321	783.933	92.475	2	<.001
Beacon STOP	759.535	844.923	735.535	44.076	2	<.001

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Parameter Estimates

Driver Complacence[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B) Lower Bound	Upper Bound
Blow-Through	Intercept	-2.347	.088	704.283	1	<.001			
	Natural ambient light	-.304	.114	7.171	1	.007	.738	.591	.922
	Conflict potential	-.302	.123	5.989	1	.014	.739	.581	.942
	Maneuver	1.737	.131	175.690	1	<.001	5.681	4.394	7.345

	LED STOP	-.974	.146	44.827	1	<.001	.377	.284	.502
	BLS STOP	-.716	.176	16.614	1	<.001	.489	.347	.690
	Beacon STOP	-.666	.154	18.634	1	<.001	.514	.380	.695
Roll-Through	Intercept	-.414	.041	101.265	1	<.001			
	Natural ambient light	-.196	.049	15.749	1	<.001	.822	.746	.906
	Conflict potential	.064	.055	1.370	1	.242	1.067	.957	1.188
	Maneuver	.401	.074	28.996	1	<.001	1.493	1.291	1.728
	LED STOP	-.538	.061	76.955	1	<.001	.584	.518	.658
	BLS STOP	-.684	.076	81.540	1	<.001	.505	.435	.586
	Beacon STOP	-.371	.067	30.531	1	<.001	.690	.605	.787

a. The reference category is: Full Stop.

Classification

Observed	Predicted Blow-Through	Full Stop	Roll-Through	Percent Correct
Blow-Through	0	381	24	0.0%
Full Stop	0	5739	103	98.2%
Roll-Through	0	2793	57	2.0%
Overall Percentage	0.0%	98.0%	2.0%	63.7%

- Scenario 3

Case Processing Summary

		N	Marginal Percentage
Compliance	Full-stop	3230	71.7%
	Roll-through	1135	25.2%
	Blow-through	139	3.1%
Valid		4504	100.0%
Missing		0	
Total		4504	
Subpopulation		4	

Model Fitting Information

	Model Fitting Criteria	Likelihood Ratio Tests		
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	124.418			
Final	53.153	71.265	4	<.001

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	4.797	2	.091
Deviance	4.808	2	.090

Pseudo R-Square

Cox and Snell	.016
Nagelkerke	.021
McFadden	.011

Likelihood Ratio Tests

	Model Fitting Criteria	Likelihood Ratio Tests		
Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	778.146	724.994	2	<.001
Treatment Type	106.537	53.384	2	<.001
Opposit Traffic	71.058	17.905	2	<.001

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Parameter Estimates

Compliance[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B) Lower Bound	95% Confidence Interval for Exp(B) Upper Bound
Roll-through	Intercept	-.849	.061	193.498	1	<.001			
	Treatment Type	-.390	.072	29.661	1	<.001	.677	.589	.779
	Opposit Traffic	.187	.076	6.151	1	.013	1.206	1.040	1.398
Blow-through	Intercept	-2.445	.126	376.247	1	<.001			
	Treatment Type	-.958	.175	30.088	1	<.001	.384	.273	.540
	Opposit Traffic	-.696	.236	8.672	1	.003	.499	.314	.792

a. The reference category is: Full-stop.

Classification

Observed	Predicted Full-stop	Predicted Roll-through	Predicted Blow-through	Predicted Percent Correct
Full-stop	3230	0	0	100.0%
Roll-through	1135	0	0	0.0%
Blow-through	139	0	0	0.0%
Overall Percentage	100.0%	0.0%	0.0%	71.7%

8.2 Quantitative Analysis (rStudio)

Table 45 Collision, Flow, Conflict and log of flow and Conflict

Coefficient Table

Poisson-Null

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	0.930	0.119	0.698 – 1.163	7.840	**<0.001**
Observations	28				
Deviance	63.353				
AIC	127.071				
log-Likelihood	-62.536				

Performance Metrics

Measure	Value
Degree of Freedom	27.0000
Pearson Chi Square	51.6479
Over Dispersion Ratio	1.9129
Log Likelihood	-62.5356
AIC	127.0712
RMSE	0.8529

Coefficient Table

NB-NULL

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	0.930	0.182	0.574 – 1.287	5.120	**<0.001**
Observations	28				
Deviance	33.426				
AIC	120.417				
log-Likelihood	-58.208				

Performance Metrics

Measure	Value
alpha	0.5305
phi	1.8852
Degree of Freedom	26.0000
Pearson Chi Square	22.0238
Over Dispersion Ratio	0.8471
Log Likelihood	-58.2085
AIC	120.4170
RMSE	0.8529

```
Model 1: Col ~ 1
Model 2: Col ~ 1
  #Df  LogLik Df  Chisq Pr(>Chisq)
1   1 -62.536
2   2 -58.208  1 8.6542   0.003263 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Coefficient Table

Poisson-BASE

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	-8.238	3.132	-14.460 – -2.167	-2.630	**0.009**
ln f ma	0.873	0.430	0.041 – 1.728	2.031	**0.042**
ln f mi	0.518	0.143	0.270 – 0.830	3.634	**<0.001**
dum	-0.386	0.304	-0.997 – 0.203	-1.268	0.205
Observations	28				
Deviance	32.067				
AIC	101.785				
log-Likelihood	-46.893				

Performance Metrics

Measure	**Value**
Degree of Freedom	24.0000
Pearson Chi Square	28.4509
Over Dispersion Ratio	1.1855
Log Likelihood	-46.8925
AIC	101.7851
RMSE	1.1879

Coefficient Table

Poisson-Alternative 1

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	-6.860	3.525	-13.976 – -0.110	-1.946	0.052
ln f ma	0.847	0.504	-0.118 – 1.864	1.682	0.093
ln f mi	0.361	0.152	0.095 – 0.698	2.378	**0.017**
comp FS BT	0.029	0.036	-0.044 – 0.098	0.793	0.428
comp FS RT	-0.618	0.369	-1.361 – 0.096	-1.676	0.094
dum	0.745	0.304	0.168 – 1.364	2.453	**0.014**
Observations	28				
Deviance	23.746				
AIC	97.464				
log-Likelihood	-42.732				

Performance Metrics

Measure	**Value**
Degree of Freedom	22.0000
Pearson Chi Square	20.5838
Over Dispersion Ratio	0.9356
Log Likelihood	-42.7320
AIC	97.4641
RMSE	0.9369

Coefficient Table

Poisson-Alternative 2

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	-6.586	3.442	-13.572 – -0.040	-1.913	0.056
ln f ma	0.476	0.496	-0.486 – 1.463	0.961	0.336
ln f mi	0.303	0.149	0.039 – 0.630	2.036	**0.042**
ln conf	0.644	0.269	0.142 – 1.201	2.397	**0.017**
comp FS NFS	-0.069	0.341	-0.723 – 0.627	-0.203	0.839
dum	0.309	0.342	-0.341 – 1.006	0.902	0.367
Observations	28				
Deviance	18.282				
AIC	91.999				
log-Likelihood	-40.000				

Performance Metrics

Measure	Value
Degree of Freedom	22.0000
Pearson Chi Square	15.6242
Over Dispersion Ratio	0.7102
Log Likelihood	-39.9997
AIC	91.9995
RMSE	0.8192

Coefficient Table

Poisson-Alternative 3

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	-4.031	1.112	-6.333 – -1.951	-3.624	**<0.001**
ln conf	1.172	0.226	0.750 – 1.637	5.194	**<0.001**
comp FS NFS	-0.874	0.417	-1.683 – -0.031	-2.093	**0.036**
comp FS BT	-0.026	0.039	-0.104 – 0.048	-0.669	0.503
Observations	28				
Deviance	28.618				
AIC	98.336				
log-Likelihood	-45.168				

Performance Metrics

Measure	Value
Degree of Freedom	24.0000
Pearson Chi Square	23.3293
Over Dispersion Ratio	0.9721
Log Likelihood	-45.1678
AIC	98.3357
RMSE	0.8309

Coefficient Table

Poisson-Alternative 4

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	-2.123	0.661	-3.491 – -0.892	-3.211	**0.001**
ln conf	0.958	0.191	0.595 – 1.347	5.011	**<0.001**
Observations	28				
Deviance	34.146				
AIC	99.863				
log-Likelihood	-47.932				

Performance Metrics

Measure	Value
Degree of Freedom	26.0000
Pearson Chi Square	28.8755
Over Dispersion Ratio	1.1106
Log Likelihood	-47.9317
AIC	99.8634
RMSE	0.8535

Coefficient Table

Conf_Null_Po

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	3.260	0.033	3.196 – 3.324	99.850	**<0.001**
Observations	36				
Deviance	308.327				
AIC	487.583				
log-Likelihood	-242.791				

Performance Metrics

Measure	Value
Degree of Freedom	35.0000
Pearson Chi Square	309.4883
Over Dispersion Ratio	8.8425
Log Likelihood	-242.7914
AIC	487.5828
RMSE	0.5744

Coefficient Table

Conf_Null_NB

Variables	Coefficients	std. Error	CI	z-Value	p
	Crash Data at Intersections				
(Intercept)	3.260	0.098	3.069 – 3.452	33.361	**<0.001**
Observations	36				
Deviance	37.437				
AIC	294.228				
log-Likelihood	-145.114				

Performance Metrics

Measure	Value
alpha	0.3054
phi	3.2740
Degree of Freedom	34.0000
Pearson Chi Square	34.5472
Over Dispersion Ratio	1.0161
Log Likelihood	-145.1138
AIC	294.2275
RMSE	0.5744

Coefficient Table

Conf_Compliance_NB

Variables	Coefficients	std. Error	CI	z-Value	p
	Crash Data at Intersections				
(Intercept)	4.318	0.313	3.718 – 4.912	13.814	**<0.001**
comp FS RT	-0.744	0.241	-1.190 – -0.292	-3.087	**0.002**
comp FS BT	0.020	0.024	-0.027 – 0.068	0.846	0.398
LED	0.959	0.260	0.436 – 1.498	3.681	**<0.001**
Observations	36				
Deviance	36.930				
AIC	286.441				
log-Likelihood	-138.220				

Performance Metrics

Measure	Value
alpha	0.1993
phi	5.0176
Degree of Freedom	31.0000
Pearson Chi Square	36.1182
Over Dispersion Ratio	1.1651
Log Likelihood	-138.2204
AIC	286.4408
RMSE	0.4917

Coefficient Table

Conf_flow_compliance_NB

	Crash Data at Intersections				
Variables	*Coefficients*	*std. Error*	*CI*	*z-Value*	*p*
(Intercept)	-1.799	1.980	-5.548 – 1.978	-0.909	0.364
ln Flow	0.811	0.267	0.303 – 1.317	3.035	**0.002**
comp FS BT	0.035	0.020	-0.004 – 0.074	1.795	0.073
comp FS RT	-0.973	0.205	-1.364 – -0.583	-4.741	**<0.001**
dum	0.603	0.160	0.301 – 0.914	3.766	**<0.001**
LED	1.053	0.210	0.634 – 1.480	5.003	**<0.001**
Observations	36				
Deviance	37.770				
AIC	274.957				
log-Likelihood	-130.479				

Performance Metrics

Measure	Value
alpha	0.1104
phi	9.0595
Degree of Freedom	29.0000
Pearson Chi Square	34.5226
Over Dispersion Ratio	1.1904
Log Likelihood	-130.4786
AIC	274.9572
RMSE	0.3939

*Table 46 Collision models performances - likelihood ratio tests - Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Base and Alternative vs Null	Base vs Alternative	Alternatives with each other
> #Base compared to Null Model 1: Col ~ 1 Model 2: Col ~ ln_f_ma + ln_f_mi + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 1 -62.536 2 4 -46.893 3 31.286 7.399e-07 ***	> #Alternative 1 compared to Base Model 1: Col ~ ln_f_ma + ln_f_mi + Dum Model 2: Col ~ ln_f_ma + ln_f_mi + comp_FS.BT + comp_FS.RT + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 4 -46.893 2 6 -42.732 2 8.321 0.0156 *	> #Alternative 2 compared to Alternative 1 Model 1: Col ~ ln_f_ma + ln_f_mi + comp_FS.BT + comp_FS.RT + Dum Model 2: Col ~ ln_f_ma + ln_f_mi + ln_conf + comp_FS.NFS + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 6 -42.732 2 6 -40.000 0 5.4646 < 2.2e-16 ***
> #Alternative 1 compared to Null Model 1: Col ~ 1 Model 2: Col ~ ln_f_ma + ln_f_mi + comp_FS.BT + comp_FS.RT + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 1 -62.536	> #Alternative 3 compared to Base Model 1: Col ~ ln_f_ma + ln_f_mi + Dum Model 2: Col ~ ln_f_ma + ln_f_mi + ln_conf + comp_FS.NFS + Dum #Df LogLik Df Chisq Pr(>Chisq)	> #Alternative 1 compared to Alternative 3 Model 1: Col ~ ln_conf + comp_FS.NFS + comp_FS.BT Model 2: Col ~ ln_f_ma + ln_f_mi + comp_FS.BT + comp_FS.RT + Dum

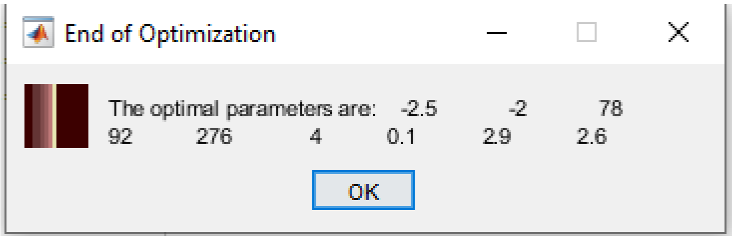
2 6 -42.732 5 39.607 1.792e-07 ***	1 4 -46.893 2 6 -40.000 2 13.786 0.001015 **	#Df LogLik Df Chisq Pr(>Chisq) 1 4 -45.168 2 6 -42.732 2 4.8716 0.08753 .
> #Alternative 2 compared to Null Model 1: Col ~ 1 Model 2: Col ~ ln_f_ma + ln_f_mi + ln_conf + comp_FS.NFS + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 1 -62.536 2 6 -40.000 5 45.072 1.403e-08 ***	> #Alternative 3 compared to Base Model 1: Col ~ ln_f_ma + ln_f_mi + Dum Model 2: Col ~ ln_conf + comp_FS.NFS + comp_FS.BT #Df LogLik Df Chisq Pr(>Chisq) 1 4 -46.893 2 4 -45.168 0 3.4494 < 2.2e-16 ***	> #Alternative 2 compared to Alternative 3 Model 1: Col ~ ln_conf + comp_FS.NFS + comp_FS.BT Model 2: Col ~ ln_f_ma + ln_f_mi + ln_conf + comp_FS.NFS + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 4 -45.168 2 6 -40.000 2 10.336 0.005695 **
> #Alternative 3 compared to Null Model 1: Col ~ 1 Model 2: Col ~ ln_conf + comp_FS.NFS + comp_FS.BT #Df LogLik Df Chisq Pr(>Chisq) 1 1 -62.536 2 4 -45.168 3 34.736 1.385e-07 ***	> #Base compared to Alternative 4 Model 1: Col ~ ln_conf Model 2: Col ~ ln_f_ma + ln_f_mi + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 2 -47.932 2 4 -46.893 2 2.0783 0.3537	> #Alternative 4 compared to Alternative 2 Model 1: Col ~ ln_conf Model 2: Col ~ ln_f_ma + ln_f_mi + ln_conf + comp_FS.NFS + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 2 -47.932 2 6 -40.000 4 15.864 0.003207 **
> #Alternative 4 compared to Null Model 1: Col ~ 1 Model 2: Col ~ ln_conf #Df LogLik Df Chisq Pr(>Chisq) 1 1 -62.536 2 2 -47.932 1 29.208 6.502e-08 ***		> #Alternative 4 compared to Alternative 1 Model 1: Col ~ ln_conf Model 2: Col ~ ln_f_ma + ln_f_mi + comp_FS.BT + comp_FS.RT + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 2 -47.932 2 6 -42.732 4 10.399 0.03421 *
		> #Alternative 4 compared to Alternative 2 Model 1: Col ~ ln_conf Model 2: Col ~ ln_f_ma + ln_f_mi + Dum #Df LogLik Df Chisq Pr(>Chisq) 1 2 -47.932 2 4 -46.893 2 2.0783 0.3537
		> #Alternative 4 compared to Alternative 3 Model 1: Col ~ ln_conf Model 2: Col ~ ln_conf + comp_FS.NFS + comp_FS.BT #Df LogLik Df Chisq Pr(>Chisq) 1 2 -47.932 2 4 -45.168 2 5.5277 0.06305 .

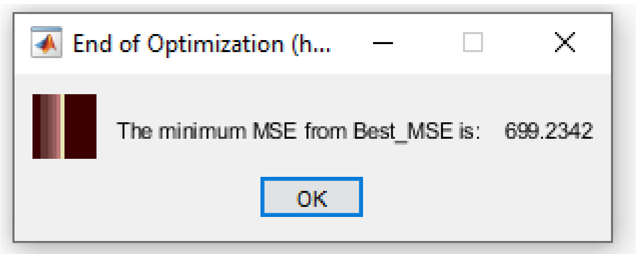
*Table 47 Conflict models performances- likelihood ratio tests - Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

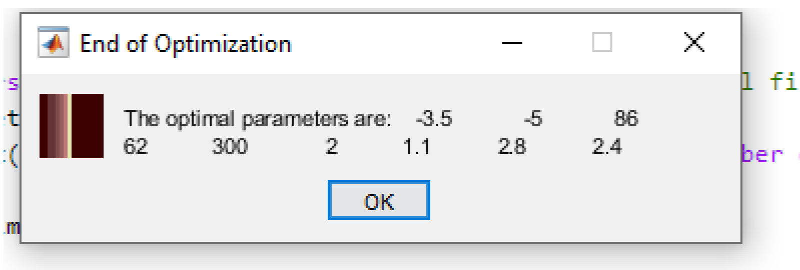
> #Compare Null NB and Conf_Compliance_NB Model 1: conf ~ 1	> #Compare Null NB and Conf_Compliance_NB Model 1: conf ~ 1	> #Compare Null NB and Conf_Compliance_NB Model 1: conf ~ 1

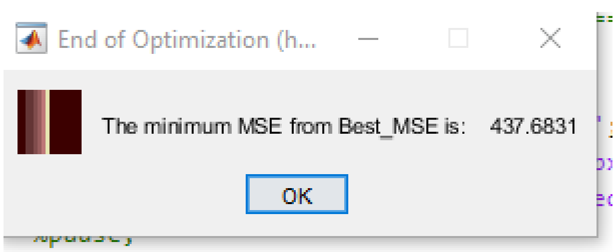
Model 2: conf ~ 1 #Df LogLik Df Chisq Pr(>Chisq) 1 1 -242.79 2 2 -145.11 1 195.36 < 2.2e-16 ***	Model 2: conf ~ comp_FS.RT + comp_FS.BT + LED #Df LogLik Df Chisq Pr(>Chisq) 1 2 -145.11 2 5 -138.22 3 13.787 0.00321 **	Model 2: conf ~ comp_FS.RT + comp_FS.BT + LED #Df LogLik Df Chisq Pr(>Chisq) 1 2 -145.11 2 5 -138.22 3 13.787 0.00321 **
> #Compare Null NB and NB Conf_flow_compliance_NB Model 1: conf ~ 1 Model 2: conf ~ ln_Flow + comp_FS.BT + comp_FS.RT + dum + LED #Df LogLik Df Chisq Pr(>Chisq) 1 2 -145.11 2 7 -130.48 5 29.27 2.052e-05 ***	> #Compare Conf_Compliance_NB and NB Conf_flow_compliance_NB Model 1: conf ~ comp_FS.RT + comp_FS.BT + LED Model 2: conf ~ ln_Flow + comp_FS.BT + comp_FS.RT + dum + LED #Df LogLik Df Chisq Pr(>Chisq) 1 5 -138.22 2 7 -130.48 2 15.483 0.0004343 ***	

8.3 Optimization (MATLAB)









9 Appendix III (Codes)

9.1 Statistical Analysis in "R" language

```
#Thesis "ASSESSING THE IMPACT OF ACTIVE SIGNAGE SYSTEMS ON DRIVING BEHAVIOR AND TRAFFIC SAFETY"
#Matin Foomani  Foomani[at] gmail . com

# Environment preparation------------------------------------------------------------

rm(list=ls())#clearing environment and remove variables
current_directory = dirname(rstudioapi::getSourceEditorContext()$path) #Getting the directory (same path as code)
setwd(current_directory)#Setting the current directory of the files as the active directory of R-studio

#Loading Data sets

db1 <- read.csv("Col_data.csv", header = TRUE)#DB1 is the set for Collision (no collision records fro BLS)
db2 <- read.csv("Con_data.csv", header = TRUE)#DB2 is the set for conflict (with more records for BLS)

#Import Packages used throughout the program
library("MASS")
library("stats")
library("sjPlot")
library("ggplot2")
library("GGally")
library("lmtest")
library("FactoClass")
library("scatterplot3d")
library("corrplot")

# Build subset of data amending to data-------------------------------------------------

#Making a subset of DB
db1_sub = subset(db1, select = -c(record,Site,Time,Dum,legs,bi_mi))
db2_sub = subset(db2, select = -c(record,Site.1,Site,Time,Dum, LED,legs,bi_mi))
#sapply(db1_sub,mean) # the average of each column
#sapply(db1_sub,sd)   # the standard deviation of each column
sapply(db1_sub,summary) #all summary
sapply(db2_sub,summary) #all summary

write.csv(db1_sub, file = "results/general/db_sum_collision.csv") # Export Results
write.csv(db2_sub, file = "results/general/db_sum_conflict.csv") # Export Results

#*** calculating log of flow and Conflict (10^-10 for values = 0)
#*For Collision DB1
db1$ln_f_ma = log(db1$f_ma+exp(10^-10))   # log of Major Flow
db1$ln_f_mi = log(db1$f_mi+exp(10^-10))   # log of minor Flow
db1$ln_f_ped  = log(db1$f_ped+exp(10^-10))# log of Ped Flow
db1$ln_conf  = log(db1$conf+exp(10^-10))# log of Conflict
db1$ln_Flow = log(db1$f_ma + db1$f_mi + db1$f_ped)# log of Traffic Flow (combined)
db1$sum_f= db1$f_mi + db1$f_ma + db1$f_ped # Sum of flows
#*For Conflict DB2
db2$ln_f_ma = log(db2$f_ma+exp(10^-10))   # log of Major Flow
db2$ln_f_mi = log(db2$f_mi+exp(10^-10))   # log of minor Flow
```

```
db2$ln_f_ped  = log(db2$f_ped+exp(10^-10))# log of Ped Flow
db2$ln_conf  = log(db2$conf+exp(10^-10))# log of Conflict
db2$ln_Flow = log(db2$f_ma + db2$f_mi + db2$f_ped)# log of Traffic Flow (combined)
db2$sum_f= db2$f_mi + db2$f_ma + db2$f_ped # Sum of flows

#View(columns name)
colnames_db1 <- colnames(db1) #Column names (kept the same lables)
print(colnames_db1) #Print lable

colnames_db2 <- colnames(db2) #Column names
print(colnames_db2) #Print lable

# Views and Visualization ------------------------------------------------

#Collision Counts VS conflict
options(device = "window")

png(file="results/general/Coll_Conf.png", width=600, height=500)

plot(db1$conf, db1$Col, main = "Number of Collisions vs Conflicts",
    ylab="", xlab = "", pch = 9, cex = 2.5 , col = "darkgreen", cex.main=1, cex.axis=1)

title(xlab = "Number of Conflicts", ylab = "Number of Collisions", line=2.5, cex.lab=1)


dev.off()

#Collision Counts VS total volume
png(file="results/general/Coll_flow.png", width=600, height=500)

plot(db1$f_mi + db1$f_ma + db1$f_ped, db1$Col, main = "Number of Collisions vs Traffic Volume",
    ylab="", xlab = "", pch = 9, cex = 2.5, col = "darkgreen", cex.main=2, cex.axis=2)

title(xlab = "Traffic Volume", ylab = "Number of Collisions", line=2.5, cex.lab=1)

dev.off()

#Conflict Counts VS total volume
png(file="results/general/Conf_flow1.png", width=600, height=500)

plot( db2$sum_f, db2$conf , main = "Number of Critical Conflict vs Traffic Volume",
    ylab="", xlab = "", pch = 9, cex = 2.5 , col = "darkgreen", cex.main=1, cex.axis=1)

title(xlab = "Traffic Volume", ylab = "Number of Conflicts", line=2.5, cex.lab=1)

dev.off()


#Histogram for Collision
png(file="results/general/Histogram_conf.png", width=600, height=500)

#GGplot library
library(ggplot2)
#Plot of the histogram


ggplot(db2, aes(conf)) +
  geom_histogram(breaks=seq(0, 60, by=9), color = "black", fill = "gold" ,linetype="dashed") +
  labs(title="Histogram of conflict Counts at Intersections",
      x="Conflict Counts per Intersection", y= "Number of observations") +
  theme(text = element_text(size=20))
```

```
ggplot(db2, aes(conf)) +
  geom_histogram(breaks=seq(0, 60, by=9),
             col="red",linetype="dashed",
             aes(fill=..count..)) +
  scale_fill_gradient("Count", low="green", high="red")



ggplot(db2, aes(conf)) +
  geom_histogram(aes(y =..density..),
             breaks=seq(0, 60, by = 9),
             col="black",
             fill="gold",
             alpha=1) +

  geom_density(col=1) +
  labs(title="Histogram of conflict Counts at Intersections",
     x="Conflict Counts per Intersection", y= "Number of observations") +
  theme(text = element_text(size=20))



dev.off()


#3D plots Choose one scatterplot3d or plotly Show Dr. Alecsandru first
png(file="results/general/matrix_3d.png", width=500, height=500)
db_sub_Var_3d = subset(db1, select = c(  Col, sum_f, conf))
var_matrix <- as.matrix(db_sub_Var_3d)
x<-var_matrix[,1]
y<-var_matrix[,2]
z<-var_matrix[,3]
df <- data.frame(x, y, z)
LM <- lm(y ~ x + z, df)
s3d <-scatterplot3d (x, z, y,pch=19 ,  type = "p",scale.y=.4 , grid=TRUE , box=TRUE, angle=46, color = "darkgrey",
                main = "Regression plane", xlab="Collisions (four period in 7 years )",ylab="Hourly critical conflicts",zlab="Average hourly
traffic Volume Major and Minor")


#addgrids3d(x, y, z, grid = c("xy", "xz", "yz"), col.grid = "grey",scale.y=0.4 , lty.grid=par("lty"), angle=46) #Used grid on xy

# compute locations of segments
orig     <- s3d$xyz.convert(x, z, y)
plane    <- s3d$xyz.convert(x, z, fitted(LM))
i.negpos <- 1 + (resid(LM) > 0) # which residuals are above the plane?

# draw residual distances to regression plane
segments(orig$x, orig$y, plane$x, plane$y, col = "red", lty = c(2, 1)[i.negpos],
       lwd = 1.5)
# draw the regression plane
s3d$plane3d(LM, draw_polygon = TRUE, draw_lines = TRUE,
         polygon_args = list(col = rgb(0.8, 0.8, 0.8, 0.8)))

# redraw positive residuals and segments above the plane
wh <- resid(LM) > 0
segments(orig$x[wh], orig$y[wh], plane$x[wh], plane$y[wh], col = "red", lty = 1, lwd = 1.5)
s3d$points3d(x[wh], z[wh], y[wh], pch = 19)
```

```
dev.off()

#Correlation Matrix

#Create a new subset with the dependent and ind-variable
print(colnames_db1)
db2_sub_Var = subset(db2, select = c(Col, conf, sum_f  ,f_ma , f_mi, f_ped,bi_mi, comp_FS.BT,comp_FS.RT,comp_FS.NFS))
cor_matrix <- cor(db2_sub_Var)#~ computing pairwise correlations of variables
write.csv(cor_matrix, file = "results/general/cor_matrix.csv") #saving the results in a CSV file

#computing the absolute value of correlations(the direction doesn't matter)
abs_cor_matrix <- abs(cor_matrix)


#saving in PNG file:
png(file="results/general/correlation2.png", width=800, height=600)


#Using CORRPLOT to change labels (only rows)

rownames(cor_matrix) <- c("Accident", "Conflict", "All volume", "Volume Major", "Volume Minor", "Pedestrian and Cyclist ", "Minor bidirectional",
                    "Compliance FS to BT","Compliance FS to RT","Compliance FS to non-FS")
corrplot(cor_matrix, method = 'ellipse', type = 'upper', tl.col="Black",)
dev.off()


#PMF for NB and Poisson models remained under each section since models needed to be created first before starting the plot.

# Collision Analysis -----------------------------------------------------

## Null Models ==============================================================
### Step 1, NULL model with Poisson (only intercept)

glm_po_0 <- stats::glm(Col ~ 1, data = db1, family = poisson(link = "log"))
# link function is log, because we are using ln(lmb) in SPF (Check section 3.4.3 in thesis)
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_0, "Poisson-Null")

#Data for probability mass function (Poisson Only):
beta_0 = glm_po_0[["coefficients"]][["(Intercept)"]] #Beta for intercept in NULL
Lmb_po_0 = exp(beta_0)#SPF of NULL Model: also mean value of the Poisson distribution
variance_po_0 = Lmb_po_0 #Variance of Poisson (which is equal to mean value)
crash_count <- 0:15 #Range of crash counts to plot the Poisson distribution:
Pr_Y_PO_0 <- dpois(crash_count, lambda=Lmb_po_0) #Probability of collisoin using Poisson distribution

#create file and plot
png(file="results/poisson/Poisson_PMF_NULL_Model.png", width=600, height=500)
plot(crash_count, Pr_Y_PO_0, type='l', lwd=6, lty = 3, col = "brown",
    cex.main=1, cex.lab=0.8, ylab='Probability',  xlab ='Collision Count',
    main=paste("Poisson  PMF of NULL model (mu =",  round(Lmb_po_0, digits = 3), ")"))

dev.off()



### Step 2, NULL model with NB (only intercept)

#MASS library
glm_nb_0 <-  MASS::glm.nb(Col ~ 1, data = db1, link = "log")
source("Customized_Writing_Functions.R")
```

```
generating_NegativeBinomial_results(glm_nb_0, "NB-NULL")


#Data for probability mass function (NB)
beta_nb_0 = glm_nb_0[["coefficients"]][["(Intercept)"]]#beta_0 for intercept
phi_nb_0 = glm_nb_0$theta#PHI
alpha_nb_0 = 1/phi_nb_0#alpha
mu_nb_0 = exp(beta_nb_0)#SPF for NULL Model NB
variance_nb_0 = mu_nb_0 + (1/phi_nb_0)*mu_nb_0^2 # Variance of NB
crash_count <- 0:15
Pr_Y_NB_0 <- dnbinom(crash_count, mu=mu_nb_0, size=phi_nb_0, log = FALSE)#Probability of Collision using NB

#create file and plot
png(file="results/negative_binomial/PO_NB_PMF_NULL_Model.png", width=800, height=500)
plot(crash_count, Pr_Y_NB_0, type='l', lwd=2, lty = 2, col = "blue4", cex.main=.8, cex.lab=1,
    ylab='Probability',xlab ='Collision Count',
    main=paste("                                                   NB mu (Blue) =",  round(mu_nb_0, digits = 3),
            ", phi =", round(phi_nb_0, digits = 3)," "))

dev.off()


#create combined plot

png(file="results/general/PO_NB_PMF_NULL_Model.png", width=800, height=500)
plot(crash_count, Pr_Y_NB_0, type='l', lwd=2, lty = 2, col = "blue4", cex.main=.8, cex.lab=1,
    ylab='Probability',xlab ='Collision Count',
    main=paste("                                                   NB mu (Blue) =",  round(mu_nb_0, digits = 3),
            ", phi =", round(phi_nb_0, digits = 3)," "))
par(new=TRUE)
plot(crash_count, Pr_Y_PO_0, type='l', lwd=2, lty = 2, col = "red2",
    cex.main=.8, cex.lab=0.1,  axes=FALSE,
    main=paste(" PO mu (Red) =",  round(Lmb_po_0, digits = 3),"                                                        "))

dev.off()


#Compare Null Poisson and NB (STEP 1&2)
lmtest::lrtest(glm_po_0, glm_nb_0)

## Alternative Models ===============================================================

### Step 1 Base Model (flow only) Poisson
glm_po_1 <- stats::glm(Col ~ ln_f_ma + ln_f_mi  +Dum , data = db1, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_1, "Poisson-BASE")

### Step 2 Base (flow only) NB
glm_nb_1 <- MASS::glm.nb(Col ~ ln_f_ma + ln_f_mi  +Dum , data = db1,  link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_1, "NB-BASE")


### Step 3 Alternative 1 Poisson
glm_po_2 <- stats::glm(Col ~ ln_f_ma + ln_f_mi  +comp_FS.BT +comp_FS.RT +Dum, data = db1, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_2, "Poisson-Alternative 1")

### Step 4 Alternative 1 NB
glm_nb_2 <- MASS::glm.nb(Col ~ ln_f_ma + ln_f_mi  +comp_FS.BT +comp_FS.RT +Dum, data = db1,  link = "log")
source("Customized_Writing_Functions.R")
```

```
generating_NegativeBinomial_results(glm_nb_2, "NB-Alternative 1")


### Step 5 Alternative 2 Poisson
glm_po_3 <- stats::glm(Col ~ ln_f_ma + ln_f_mi +ln_conf +comp_FS.NFS +Dum , data = db1, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_3, "Poisson-Alternative 2")

### Step 6 Alternative 2 NB
glm_nb_3 <- MASS::glm.nb(Col ~ ln_f_ma + ln_f_mi +ln_conf +comp_FS.NFS +Dum , data = db1,  link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_3, "NB-Alternative 2")


### Step 7 Alternative 3 Poisson
glm_po_4 <- stats::glm(Col ~ ln_conf +comp_FS.NFS +comp_FS.BT, data = db1, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_4, "Poisson-Alternative 3")

### Step 8 Alternative 3 NB
glm_nb_4 <- MASS::glm.nb(Col ~ ln_conf +comp_FS.NFS +comp_FS.BT,  data = db1,  link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_4, "NB-Alternative 3")


### Step 9 Alternative 4 Poisson
glm_po_5 <- stats::glm(Col ~ ln_conf  , data = db1, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_5, "Poisson-Alternative 4")

### Step 10 Alternative 5 NB
glm_nb_5 <- MASS::glm.nb(Col ~ ln_conf , data = db1,  link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_5, "NB-Alternative 4")

### Step 11 Base and Alternative vs Null
#Base compared to Null
lmtest::lrtest(glm_po_0, glm_po_1)
#Alternative 1 compared to Null
lmtest::lrtest(glm_po_0, glm_po_2)
#Alternative 2 compared to Null
lmtest::lrtest(glm_po_0, glm_po_3)
#Alternative 3 compared to Null
lmtest::lrtest(glm_po_0, glm_po_4)
#Alternative 4 compared to Null
lmtest::lrtest(glm_po_0, glm_po_5)


### Step 12 Base vs Alternative
#Alternative 1 compared to Base
lmtest::lrtest(glm_po_1, glm_po_2)
#Alternative 3 compared to Base
lmtest::lrtest(glm_po_1, glm_po_3)
#Alternative 3 compared to Base
lmtest::lrtest(glm_po_1, glm_po_4)
#Base compared to Alternative 4
lmtest::lrtest(glm_po_5, glm_po_1)

### Step 13 Alternatives with each other
#Alternative 2 compared to Alternative 1
lmtest::lrtest(glm_po_2, glm_po_3)
```

```
#Alternative 1 compared to Alternative 3
lmtest::lrtest(glm_po_4, glm_po_2)
#Alternative 2 compared to Alternative 3
lmtest::lrtest(glm_po_4, glm_po_3)
#Alternative 4 compared to Alternative 2
lmtest::lrtest(glm_po_5, glm_po_3)
#Alternative 4 compared to Alternative 1
lmtest::lrtest(glm_po_5, glm_po_2)
#Alternative 4 compared to Alternative 2
lmtest::lrtest(glm_po_5, glm_po_1)
#Alternative 4 compared to Alternative 3
lmtest::lrtest(glm_po_5, glm_po_4)


# Conflict Analysis -------------------------------------------------------

##%%%%%%%%%%%%%%%%%% Step 1 Conflict Null Poisson
glm_po_6 <- stats::glm(conf ~ 1, data = db2, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_6, "Conf_Null_Po")

#%%%%%%%%%%%%%%%%%% Step 2 Conflict Null NB
glm_nb_6 <- MASS::glm.nb(conf ~ 1, data = db2, link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_6, "Conf_Null_NB")


##%%%%%%%%%%%%%%%%%% Step 3 Conflict and compliance variables Poisson
glm_po_7 <- stats::glm(conf ~ comp_FS.RT +comp_FS.BT +LED, data = db2, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_7, "Conf_compliance_Poisson")


#%%%%%%%%%%%%%%%%%% Step 4 Conflict and compliance variables NB
glm_nb_7 <- MASS::glm.nb(conf ~ comp_FS.RT +comp_FS.BT +LED, data = db2, link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_7, "Conf_Compliance_NB")


##%%%%%%%%%%%%%%%%%% Step 5 Conflict, flow and compliance variables Poisson
glm_po_8 <- stats::glm(conf ~ ln_Flow  +comp_FS.NFS  +comp_FS.RT + Dum , data = db2, family = poisson(link = "log"))
source("Customized_Writing_Functions.R")
generating_Poisson_results(glm_po_8, "Conf_Flow_compliance_Poisson")

#%%%%%%%%%%%%%%%%%% Step 6 Conflict, flow and compliance variables NB
glm_nb_8 <- MASS::glm.nb(conf ~  ln_Flow  +comp_FS.BT  +comp_FS.RT  + Dum +LED, data = db2, link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_8, "Conf_flow_compliance_NB")


#%%%%%%%%%%%%%%%%%% Step 6 Conflict and flow only NB
glm_nb_9 <- MASS::glm.nb(conf ~  ln_Flow  + Dum +LED, data = db2, link = "log")
source("Customized_Writing_Functions.R")
generating_NegativeBinomial_results(glm_nb_9, "Conf_flow_NB")



#Compare Null Poisson and NB (STEP 1&2)
lmtest::lrtest(glm_po_6, glm_nb_6)
#Compare Null NB and Conf_Compliance_NB
lmtest::lrtest(glm_nb_6, glm_nb_7)
```

```
#Compare Null NB and NB Conf_flow_compliance_NB
lmtest::lrtest(glm_nb_6, glm_nb_8)
#Compare Conf_Compliance_NB and NB Conf_flow_compliance_NB
lmtest::lrtest(glm_nb_7, glm_nb_8)


# Extract fitted Values ---------------------------------------------------

#write fitted Values (all of collision)
fitted_col <- data.frame (glm_po_1_2$fitted.values,glm_po_2$fitted.values,glm_po_1_1$fitted.values)
write.csv(fitted_col, file = "results/general/fitted_Col.csv") # Export Results

#write fitted Values (all of Conflict)
fitted_con <- data.frame (glm_nb_7$fitted.values)
write.csv(fitted_con, file = "results/general/fitted_Con.csv") # Export Results


# Mann-Whitney U-Test (Conflict Simulation and recorded) ------------------


#Mann-Whitney U-Test to asse the correlation between Observed and simulated
#Data only from intersection 25th (first five read) and 28th, (second five read)
# Time series: 07:00 09:00 -  12:00 14:00 - 16:00 18:00 - 18:00 20:00 - 20:00 22:00
x1<-c(55,59,37,17,13,6,41,33,8,2) #Simulation Conflict
x2<-c(44,38,40,14,8,4,38,23,5,3) #Observed conflict
wilcox.test(x1, x2, alternative = "two.sided", paired = FALSE, exact = FALSE, correct = TRUE)

#mirror histogram
png(file="Sim1.png", width=300, height=200)
par(mar=c(0,5,3,3))
hist(x1 , main="" , xlim=c(0,60), ylab="Frequency (Simulation)", xlab="", ylim=c(0,4) , xaxt="n", las=1 , col="Gold", breaks=5)
dev.off()

png(file="Obs1.png", width=300, height=200)
par(mar=c(5,5,0,3))
hist(x2 , main="" , xlim=c(0,60), ylab="Frequency (Observation)", xlab="Number of conflicts", ylim=c(4,0) , las=1 , col="tomato3"  ,
breaks=5)
dev.off()
```

9.2 Genetic Algorithm in MATLAB

```
%#Thesis "ASSESSING THE IMPACT OF ACTIVE SIGNAGE SYSTEMS ON DRIVING BEHAVIOR AND TRAFFIC SAFETY"
%Matin Foomani  Foomani[at] gmail . com
clear; %clear the environment
clc;
%%
%{
 ========================================================================
 General info
 The line below lists Methods for class COM.Vissim_Vissim:
 Vissim.methods
 to get the value type of each method us:
 Vissim.invoke
 below command reveal the properties of an object
 Vissim.fields
 ========================================================================
%}
% Load traffic simulation Model
Vissim = actxserver('Vissim.Vissim');
Path='G:\My Drive\Education\PhD\Analysis\MATLAB';
Vissim.LoadNet([Path '\vissim_MD_01.inpx']) %file location.
uiwait(msgbox('Please Increase simulation speed in Vissim GUI, and number of runs to 1, then enter
any key to continue',"Attention","help"));
%pause;
sim=Vissim.Simulation; %defines the ISimulation Object

% ******* Load Observed TMCs
load("Obs_AM_TMCs.mat");
Obs_TMCs_Vol=cell2mat(Obs_AM_TMCs(:,2)); %to change from 'cell' (text) to 'double' (number)

% **********************************************************************
% ****************** Set handels for Calibration Parameter/Genes******************
% **********************************************************************

vnet=Vissim.Net; %defined the network object
DriveBeh=vnet.DrivingBehaviors; %link to the parapters that are related to driver's behavior
RoadTypes=DriveBeh.GetMultiAttValues('Name');%to test if the connection is wokring. this is to get
the "roadtype". The roadtype used for this study is called 'Urban (Lachine)' Item(6) in roadtype.
%******************* Checking COM attributes (internal check, can be removed)
Max_Dcc_O=DriveBeh.ItemByKey(6).get('AttValue','MaxDecelOwn'); %Max Deceleration Own (m/s2)
Acc_Dcc_O=DriveBeh.ItemByKey(6).get('AttValue','AccDecelOwn');% Accepted Deceleration Own (m/s2)
DecelRedDistOwn=DriveBeh.ItemByKey(6).get('AttValue','DecelRedDistOwn');%DecelRedDistOwn
DecelRedDistTrail=DriveBeh.ItemByKey(6).get('AttValue','DecelRedDistTrail');%DecelRedDistTrail
Max_LAH_Dis=DriveBeh.ItemByKey(6).get('AttValue','LookAheadDistMax');%Max Look Ahead Distance (m)
Avg_SS_Dis=DriveBeh.ItemByKey(6).get('AttValue','W74ax'); %Avg Standstill Distance for Wiedemann 74
(m)
Min_head=DriveBeh.ItemByKey(6).get('AttValue','MinFrontRearClear'); %Min clearance (front/rear) (m)
Add_SS_Dis=DriveBeh.ItemByKey(6).get('AttValue','W74bxAdd'); %Additive part of safety distance
[Value for the determination of the desired safety distance d. Enables the adaption of the time
needs.]
Multi_SS_Dis=DriveBeh.ItemByKey(6).get('AttValue','W74bxMult'); %Multiplic. part of safety distance
[Value for the determination of the desired safety distance d. Enables the adaption of the time
needs. Greater value = greater distribution (standard deviation) of safety distance.]
%uiwait(msgbox('COM help says Avg_SS_Dist should be between -1 and 1, but here is
2!!!',"Attention","help"));
```

```
InitialChromosome=[Max_Dcc_O,Acc_Dcc_O,DecelRedDistOwn,DecelRedDistTrail,Max_LAH_Dis,Avg_SS_Dis,Min_
head,Add_SS_Dis,Multi_SS_Dis];%read all the reading from above (all 9 parameters into one variable).

% ***********************************************************************
% ****************** Generate Initial Population**********************
% ***********************************************************************

Pop_size=9; %% 9 (the same as number of calibration parameters rule of thumb)
[~,NoOfGenes]=size(InitialChromosome); %number of parameters being calibrated (returnd 9).
Tested_mat=zeros(1,NoOfGenes); % This matrix keeps track of all the combinations tested
Initial_pop=[]; %Empty matrix
Summary_matrix=[]; %Empty matrix
[m,n]=size(Initial_pop); %read for stopping criteria
Steps=[0.5,0.5,2,2,8,0.5,0.1,0.1,0.1]; %Step for each of the 9 parameters.
ULs=[-1,-1,100,100,300,5,5,3,3]; %uper limits
LLs=[-4,-4,50,50,100,0,0.1,1,1]; %lower limits
DCC_Range=[-4:0.5:-1]; %Max_Dcc and Acc_Dcc (the first two parameters)
DCC_Dist_Range=[50:2:100];%DecelRedDistOwn and DecelRedDistTrail(Next two parameters)
LH_Dist_Range=[100:8:300];%Max_LAH_Dis
SS_Dis_Range=[0:0.5:5];%Avg_SS_Dis
H_Range=[0.1:0.1:5];%Min_head
Safety_Dis_Range=[1:0.1:3];%Multi_SS_Dis and Add_SS_Dis (last two paramters)

%Start with initial polulation
while m<Pop_size %size of the matrix generated in line 82


[Max_Dcc_O,Acc_Dcc_O,DecelRedDistOwn,DecelRedDistTrail,Max_LAH_Dis,Avg_SS_Dis,Min_head,Add_SS_Dis,Mu
lti_SS_Dis]=Vis_CreatChromosome(DCC_Range,DCC_Dist_Range,LH_Dist_Range,SS_Dis_Range,H_Range,Safety_D
is_Range); %will take you to "Vis_CreatChromosome" function

chromosome=[Max_Dcc_O,Acc_Dcc_O,DecelRedDistOwn,DecelRedDistTrail,Max_LAH_Dis,Avg_SS_Dis,Min_head,Ad
d_SS_Dis,Multi_SS_Dis];%create chromosome from the out put of the function 'Vis_CreatChromosome'
    if(sum(ismember(Tested_mat,chromosome,'rows'))==0) %to check is yjere any chromosome in
"tested_mat" which is exactly the same as the generated chromosome. This is to avoid counting a
repeated chromosome in "M" matrix.

        %Function to set the variables in Vissim
        Vis_SetDST_2(chromosome,vnet,DriveBeh); %call Vis_SetDST_2 and set variables from chromosome
into VISSIM
        sim.RunContinuous(); %which is the same as Vissim.Simulation.RunContinuous() %runs
simulation
        [MSE,GEH0to5,GEH5to10,GEHover10]=Vis_Calculate_MSE_2(vnet,Obs_TMCs_Vol);%this function is
taking the output from VISSIM and make the calcualtion for MSE and GEH
        Summary=[chromosome,MSE,GEH0to5,GEH5to10,GEHover10];%keep the output of the 3 steps above in
a summary Matix
        Initial_pop=[Initial_pop;Summary];
        Tested_mat=[Tested_mat;chromosome];
        Summary_matrix=[Summary_matrix;Summary];
        [m,n]=size(Initial_pop);
     elseif (sum(ismember(Tested_mat,chromosome,'rows'))>1)
        msgbox('choromosome repeated more than once!!!')
    end
end
Tested_mat=Tested_mat(2:Pop_size+1,:); %removes the first line from the TestedMat (this was all 0)


% ***********************************************************************
% ****************** RUN Genetic Algorithm ****************************
% ***********************************************************************

% ****** 1- select best from current run, 2-Cross-ver, 3- Mutate 4-Generate
% one new (random gens)
```

```
% ***** NOTE: for proof of concept, we are using MSE as the objective
% function to be minimized (not GEH)
% In the future, if conlict could be generated directly through
% microsimulation or through a 3rd party application which work in
% itterative manner, it would be better to use conflict than MSE.

dim=NoOfGenes+1; % Column number in the InitialPop (i.e., parent_mat) that represents the objective
function (change if needed (here we are using MSE, therefore column 13)
Delta_best_MSE=0;
Iteration_num=1;
Max_itirations=15; % FUTURE Research: with the SSAM connection, change to 30
Stopping_criteria=4; % can be increased with SSAM conenction and conflict
num_of_delta_zero=0;
Best_MSE=[];

while (Iteration_num<=Max_itirations && num_of_delta_zero<Stopping_criteria)  %stopping
criteria,(iteration number=15, four times with 0 improvement)
    [Parents_selected,New_generation,best_index]=Vis_Run_GA_2(Initial_pop,dim,Steps,ULs,LLs); %start
the fuction 'Vis_Run_GA_2' to generate parents
    Best_MSE=[Best_MSE;[Iteration_num,Initial_pop(best_index,dim)]];
    if Iteration_num>1
        Delta_best_MSE=Best_MSE(Iteration_num-1,2)-Best_MSE(Iteration_num,2);
    end
    if Delta_best_MSE<0
        msgbox('Change in the MES of the best solution is NEGATIVE!!!!',"Error??","error") %just in
case!
        pause;
    end
    if Delta_best_MSE==0
        num_of_delta_zero=num_of_delta_zero+1
    else
        num_of_delta_zero=0
    end

    if mod(Iteration_num,2)==0
        fprintf([sprintf('%d',Iteration_num),'  iterations completed']);
    end
    Iteration_num=Iteration_num+1
    Parents_mat=New_generation; %new generation is created with new parents
    [mm,nn]=size(Parents_mat);
    Initial_pop=[];
    for k=1:mm
        chromosome=Parents_mat(k,:);
        if(sum(ismember(Tested_mat,chromosome,'rows'))==0)
            %Function to set the variables in Vissim
            Vis_SetDST_2(chromosome,vnet,DriveBeh);
            sim.RunContinuous(); %which is the same as Vissim.Simulation.RunContinuous() %runs
simulation
            [MSE,GEH0to5,GEH5to10,GEHover10]=Vis_Calculate_MSE_2(vnet,Obs_TMCs_Vol);
            Summary=[chromosome,MSE,GEH0to5,GEH5to10,GEHover10];
            Initial_pop=[Initial_pop;Summary];
            Tested_mat=[Tested_mat;chromosome];
            Summary_matrix=[Summary_matrix;Summary];

        else
            Org_index=find(ismember(Tested_mat,chromosome,'rows')>0); %if it was used before do not
run again just read from summary.
            Summary=[chromosome,Summary_matrix(Org_index,dim:dim+3)];
            Initial_pop=[Initial_pop;Summary];
        end

    end
end
```

```
% *********************************************************************
% **************************** End of GA ******************************
% *********************************************************************

[I,c] = (min(Summary_matrix));
Best=c(dim);
Idata = (-pi/4):0.1:(pi/4);
Idata = (cos(Idata)-cot(Idata));
uiwait(msgbox(['The optimal parameters are:    ',num2str(Summary_matrix(Best,1:dim-1))],'End of Optimization','custom',Idata,pink(size(Idata,2))));
uiwait(msgbox(['The minimum MSE from Best_MSE is:    ',num2str(Best_MSE(Iteration_num-1,2))],'End of Optimization (MatinFoomaniEnd)','custom',Idata,pink(size(Idata,2))));

%Popup the results!

datapoints=vnet.DataCollectionMeasurements;
Decs=vnet.Detectors;
```

Function +++++++++++++++Vis_CreatChromosome+++++++++++++++++

```
function
[Max_Dcc_O,Acc_Dcc_O,DecelRedDistOwn,DecelRedDistTrail,Max_LAH_Dis,Avg_SS_Dis,Min_head,Add_SS_Dis,Multi_SS_Dis]=Vis_CreatChromosome(DCC_Range,DCC_Dist_Range,LH_Dist_Range,SS_Dis_Range,H_Range,Safety_Dis_Range)

[~,DCC_Range_Size]=size(DCC_Range);
Max_Dcc_O=DCC_Range(floor(1+DCC_Range_Size*rand)); %Rand function in matlab Uniformly distributed pseudorandom numbers.
Acc_Dcc_O=DCC_Range(floor(1+DCC_Range_Size*rand));

[~,Dcc_Dist_Size]=size(DCC_Dist_Range);
DecelRedDistOwn=DCC_Dist_Range(floor(1+Dcc_Dist_Size*rand));
DecelRedDistTrail=DCC_Dist_Range(floor(1+Dcc_Dist_Size*rand));

[~,LH_Dist_Size]=size(LH_Dist_Range);
Max_LAH_Dis=LH_Dist_Range(floor(1+LH_Dist_Size*rand));

[~,SS_Dist_Size]=size(SS_Dis_Range);
Avg_SS_Dis=SS_Dis_Range(floor(1+SS_Dist_Size*rand));

[~,H_Size]=size(H_Range);
Min_head=H_Range(floor(1+H_Size*rand));

[~,Safety_Dist_Size]=size(Safety_Dis_Range);
Add_SS_Dis=Safety_Dis_Range(floor(1+Safety_Dist_Size*rand));
Multi_SS_Dis=Safety_Dis_Range(floor(1+Safety_Dist_Size*rand));
```

Function +++++++++++++++ Vis_SetDST+++++++++++++++++

```
%****************** Setting COM attributes
vnet.DrivingBehaviors.ItemByKey(6).set('AttValue','MaxDecelOwn',chromosome(1)); %Max Deceleration Own (m/s2)
DriveBeh.ItemByKey(6).set('AttValue','AccDecelOwn',chromosome(2));% Accepted Deceleration Own (m/s2)
```

```
DriveBeh.ItemByKey(6).set('AttValue','DecelRedDistOwn',chromosome(3));
DriveBeh.ItemByKey(6).set('AttValue','DecelRedDistTrail',chromosome(4));
DriveBeh.ItemByKey(6).set('AttValue','LookAheadDistMax',chromosome(5));%Max Look Ahead Distance (m)
DriveBeh.ItemByKey(6).set('AttValue','W74ax',chromosome(6)); %Avg Standstill Distance for Wiedemann
74 (m)
DriveBeh.ItemByKey(6).set('AttValue','MinFrontRearClear',chromosome(7)); %Min clearance (front/rear)
(m)
DriveBeh.ItemByKey(6).set('AttValue','W74bxAdd',chromosome(8)); %Additive part of safety distance
[Value for the determination of the desired safety distance d. Enables the adaption of the time
needs.]
DriveBeh.ItemByKey(6).set('AttValue','W74bxMult',chromosome(9));
%vnet.ReducedSpeedAreas.ItemByKey(RSA_No(k,2)).set('AttValue','DesSpeedDistr(11)',chromosome(Counter
));
end
```

Function +++++++++++++++ Vis_Calculate +++++++++++++++++

```
function [MSE,GEH0to5,GEH5to10,GEHover10]=Vis_Calculate_MSE(vnet,Obs_TMCs)

Nodes=vnet.Nodes;
NameOfNodes=vnet.Nodes.GetMultiAttValues('Name'); %listing for user info
[NoNodes,n]=size(NameOfNodes);
[No_TMCs,n]=size(Obs_TMCs);

k=1;
vnet.Nodes.ItemByKey(1).Movements.GetMultiAttValues('Vehs(1,1,All)'); %(Simulation
Run,TimeInterval,VehType)
vnet.Nodes.ItemByKey(1).Movements.GetMultiAttValues('Vehs(Avg,1,All)');
TempResults=[];
for i=1:NoNodes

TempResults=[TempResults;vnet.Nodes.ItemByKey(i).Movements.GetMultiAttValues('Vehs(Avg,1,All)')];
    TempResults(end,:)=[]; %I delete the last row imported because Vissim outputs an extra Avg of
all movements for the node, which is not needed
    [m,n]=size(TempResults);
    %for j=1:m-1
     %   Sim_TMCs(k,1)=cell2mat(TempResults(j,2));
      %  k=k+1;
    %end
end
TempResults=TempResults(:,2);
for GG=1:m
    Sim_TMCs(GG,1)=cell2mat(TempResults(GG,1));
end
Sim_TMCs=double(Sim_TMCs);
% ******* Calculating the MOEs (Mean Square of errors, and GEHs)
GEHs=zeros(m,1);
MSE=0;

for i=1:No_TMCs
    Temp=(Obs_TMCs(i,1)-Sim_TMCs(i,1))^2;
    MSE=MSE+Temp;
    GEH(i,1)=sqrt(2*Temp/(Obs_TMCs(i,1)+Sim_TMCs(i,1)));
end

MSE=MSE/m;
GEH0to5=sum(GEHs<=5)/m*100;
GEH5to10=(sum(GEHs<=10)-sum(GEHs<=5))/m*100;
GEHover10=sum(GEHs>10)/m*100;
```

```
if GEH0to5+GEH5to10+GEHover10==100
    disp('All Good with GEHs'); %all tests for GEH passed.
    %pause;
else
    disp('sum of GEHs are',num2str(GEH0to5+GEH5to10+GEHover10),'!!!!!!'); %Just to make sure!
    pause;
end

end
```

Function +++++++++++++++ Vis_Run_GA +++++++++++++++++

```
function [Selected_parents,New_generation,best]=Vis_Run_GA_2(Parents_mat,dim,Steps,ULs,LLs)

New_pop=[];
temp=[];
temp2=[];
Mutate_pop=[];
[m,n]=size(Parents_mat);
MSEsum=0;

ProbOfSelec_mat=zeros(m-1,1);
Com_ProbOfSelec_mat=zeros(m-1,1);

% ******* Identify and Keep the best parent in the next iteration (Pass over the best
% combination to the next generation)
[I,c] = (min(Parents_mat));% 'I' min of each column and 'c' position of that minimum (which row)
C=c(dim);
New_pop=[New_pop,Parents_mat(C,:)];

% ******* Remove the best answer parent to do the rest of GA
temp=[temp;Parents_mat(1:C-1,:)];
temp=[temp;Parents_mat(C+1:m,:)];
temp2=temp;

%Probabilty of Selcetion (pre-step to roulette wheel.
tempSum=sum(temp);
MSEsum=MSEsum+tempSum(dim);

for j=1:m-1
   ProbOfSelec_mat(j,1)= temp(j,dim)/MSEsum; %probofselection is based on MSE of each chromosme
divided to summary fo all.
   if j==1
       Com_ProbOfSelec_mat(j,1)=ProbOfSelec_mat(j,1);
   else
       Com_ProbOfSelec_mat(j,1)=ProbOfSelec_mat(j,1)+Com_ProbOfSelec_mat(j-1,1);
   end
end

% ******* generate roulette wheel for this generation
Roulette_mat=rand(m-3,1); % excluding the best and 2 that will be mutated, 6 parents will be
selected
% ******* Find the indices for selected parents matrix
Index_parent=Return_Index_2(Roulette_mat,Com_ProbOfSelec_mat);
New_pop=[New_pop;temp(Index_parent,:)]; % New_pop is the selected parents that will produce new
children through CrossOver
% remove the Index_parent row from temp2 (to have the non-selected parents
% at the end) [deleteing a row afects the row numbers-> therfore, change its MSE to
temp2(Index_parent,dim)=inf;% infinity; so that it won't be a "best unselected solution"
```

```
% Crossover : (for m-2 chromosomes: All except the ones that will be mutated)
Selected_parents=New_pop(:,1:dim-1); % only the chromosomes of the New-pop

% *** find Mutate_pop
for b=1:2 %we take two but remove one of the chromosomes
    [~,cc] = (min(temp2));
    CC=cc(dim);
    Mutate_pop=[Mutate_pop;temp2(CC,1:dim-1)];
    temp2(CC,dim)=inf;
end

best=C;

New_generation_crossover=Vis_new_generation_2(Selected_parents); %use the function of
'Vis_new_generation_2" for crossover from 6 parents to 6 childrens
New_generation_mutated=Vis_Mutation_2(Mutate_pop,Steps,ULs,LLs); %use the functiuon of
'Vis_Mutation_2' for mutation  from 1 parent to 1 children
New_generation=[New_generation_crossover;New_generation_mutated];
```

Function +++++++++++++++ Vis_new_generation +++++++++++++++++

```
% ******* This function does the crossover
function New_Generation=Vis_new_generation_2(SelectedParents)
temp_pop=[];
% ******** The best solution is passed over without any changes
temp_pop=[temp_pop;SelectedParents(1,:)];

[m,n]=size(SelectedParents);
num=(m-1)/2;
% Use the two-point crossover method to produce off-springs
for i=1:num
    CrossOverPoints=sort(floor(1+(n-1)*rand(2,1)));


child1=[SelectedParents(2*i,1:CrossOverPoints(1)),SelectedParents(2*i+1,CrossOverPoints(1)+1:CrossOv
erPoints(2)),SelectedParents(2*i,CrossOverPoints(2)+1:n)];

child2=[SelectedParents(2*i+1,1:CrossOverPoints(1)),SelectedParents(2*i,CrossOverPoints(1)+1:CrossOv
erPoints(2)),SelectedParents(2*i+1,CrossOverPoints(2)+1:n)];

    temp_pop=[temp_pop;child1;child2];
end
New_Generation=temp_pop;
```

Function +++++++++++++++ Vis_Mutation +++++++++++++++++

```
function [New_generation_mutated]=Vis_Mutation_2(Mutate_pop,Steps,ULs,LLs)

% Random decision to mutate (if not mutate, it will replace with a NEW
% chromosome all together;
% If mutate, the get the direction of mutation, and increase it by 3 steps
% (NOTE that this number was sleected based on the fact that the search
% areas have a range between 10 to 26 potential values. This number can be
% modified in the future)
```

```
[h,hh]=size(Mutate_pop); %hh is the counter for the gene (same as it would be in Steps, ULs, and
LLs)
New_generation_mutated=[];

for j=1:h-1
    Gene=Mutate_pop(j,:);
    for i=1:hh
        if rand(1)>=0.5   % whether or not a gene will be mutated (probability of mutation)
            if rand(1)>=0.5 %direction of adding: increase
                if Gene(i)+(5*Steps(i))<=ULs(i)
                    New_Gene(i)=Gene(i)+(3*Steps(i));
                else
                    New_Gene(i)=ULs(i);
                end
            elseif Gene(i)-(3*Steps(i))>=LLs(i) %decrease
                New_Gene(i)=Gene(i)-(5*Steps(i));
            else
                New_Gene(i)=LLs(i);
            end
        else
            New_Gene(i)=Gene(i);
        end
    end
    New_generation_mutated=[New_generation_mutated;New_Gene];
end

DCC_Range=[-4:0.5:-1];
DCC_Dist_Range=[50:2:100];
LH_Dist_Range=[100:8:300];
SS_Dis_Range=[0:0.5:5];
H_Range=[0.1:0.1:5];
Safety_Dis_Range=[1:0.1:3];
[Max_Dcc_O,Acc_Dcc_O,DecelRedDistOwn,DecelRedDistTrail,Max_LAH_Dis,Avg_SS_Dis,Min_head,Add_SS_Dis,Mu
lti_SS_Dis]=Vis_CreatChromosome(DCC_Range,DCC_Dist_Range,LH_Dist_Range,SS_Dis_Range,H_Range,Safety_D
is_Range);

New_Gene=[Max_Dcc_O,Acc_Dcc_O,DecelRedDistOwn,DecelRedDistTrail,Max_LAH_Dis,Avg_SS_Dis,Min_head,Add_
SS_Dis,Multi_SS_Dis];
New_generation_mutated=[New_generation_mutated;New_Gene];
```