Fall 11-29-2016

An Assessment of Post-Encroachment Times for Bicycle-Vehicle Interactions Observed in the Field, a Driving Simulator, and in Traffic Simulation Models

Ali Razmpa
Portland State University

10.15760/etd.5270

An Assessment of Post-Encroachment Times for Bicycle-Vehicle Interactions Observed in the Field, a Driving Simulator, and in Traffic Simulation Models

by

Ali Razmpa

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Civil and Environmental Engineering

Thesis Committee:
Christopher M. Monsere, Chair
Miguel Figliozzi
Avinash Unnikrishnan

Portland State University
2016

ABSTRACT

Most safety analysis is conducted using crash data. Surrogate safety measures, such as various time-based measures of time-to-collision can be related to crash potential and used to gain insight into the frequency and severity of crashes at a specific location. One of the most common and acknowledged measures is post-encroachment time (PET) which defines the time between vehicles occupying a conflicting space. While commonly used in studies of motor vehicle interactions, studies of PET for bicycle-vehicle interactions are few. In this research, the PET of bicycle-vehicle interactions measured in the field, a driving simulator, and in a micro-simulation are compared. A total of 52 right-hook conflicts were identified in 135 hours of video footage over 14 days at a signalized intersection in Portland, OR (SW Taylor and SW Naito Pkwy). The results showed that 4 of 17 high-risk conflicts could not be identified by the conventional definition of PET and PET values of some conflicts did not reflect true risk of collision. Therefore, right-hook conflicts were categorized into two types and a modified measure of PET was proposed so that their frequency and severity were properly measured. PETs from the field were then compared to those measures in the Oregon State University driving simulator during research conducted by Dr. Hurwitz et al. (2015) studying the right-hook conflicts. Statistical and graphical methods were used to compare field PETs to those in the simulator. The results suggest that the relative validity of the OSU driving simulator was good but not conclusive due to differences in traffic conditions and intersections. To further explore the field-observed PET values, traffic simulation models of the field intersection were

developed and calibrated. Right-hook conflicts were extracted from the simulation files and conflicts observed in PM-peak hours over 6 days in the field were compared to those obtained from 24 traffic simulation runs. The field-observed PET values did not match the values from the simulation values very well. However, the approach does show promise. Further calibration of driving and bicycling behaviors would likely improve the result.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to the following people, who have provided guidance and support in pursuing my graduate studies.

To Dr. Christopher Monsere, my major advisor, whose encouragement, patience, and knowledge gave me great confidence for both my personal life and academic success at Portland State University.

To my parents, who supported and encouraged my decision to begin the long journey of pursuing my graduate degree in the US.

To my classmates and officemates in the ITS lab, whose friendship has enriched my time here, and also for their help in this research, especially Daniel Hazel for his contribution in collecting data.

To my committee members and faculty in Civil Engineering Department, Dr. Christopher Monsere, Dr. Robert Lawrence Bertini, Dr. Miguel Andres Figliozzi, Dr. Kelly Clifton, and Dr. Franz Rad for their insight and support. I am honored to have had such great faculty.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

The use of advanced driving simulators as a tool for studying driving behaviors and developing countermeasures, particularly in high-risk situations or crashes, is growing substantially. Their usefulness as such a tool is dependent on their ability to reflect actual driver behavior. In order to use them for research, a process is required to validate a driving simulator for each research project. A portion of this thesis describes such an effort that was part of a research project conducted by Hurwitz et al. (2015) and funded by the Oregon Department of Transportation (ODOT) using the Oregon State University (OSU) driving simulator. Although the OSU driving simulator validation was part of that research, the validation was not conclusive due to a lack of observed conflicts in the field. This research aims to enhance the validation of the OSU driving simulator for the right-turn vehicle-bicycle interaction by observing more conflicts in another intersection. In addition, traffic simulation models were developed in VISSIM and evaluated for the effect of geometric designs and other conditions that differ between the intersections designed in the OSU driving simulator and those in the field to analyze the models ability to reflect driver behavior on the road. Furthermore, the measurement method of Post-Encroachment Times (PETs) in the field is revised and a new approach for measuring this surrogate safety measure for right-hook conflicts is proposed. Finally, the relative validity of the OSU driving simulator is determined through graphical and statistical validation approaches.

1.1 Research Objectives

The primary purpose of this research is to validate the ability of the Oregon State University driving simulator to reflect driving behavior in the field where a right-turning motorist fails to notice the bicyclist approaching the intersection, leading to a conflict and possible crash. This kind of crash is termed a "right-hook (RH) crash." Hurwitz et al. (2015) used the OSU driving simulator to investigate safety countermeasures for right-hook crashes. The research objectives, research questions, data sources, and analysis methods of this work are described below.

Objectives

- Validation of the Oregon State University Driving Simulator for right-hook conflicts.
- Identifying and measuring post-encroachment time for right-hook conflicts in the field.
- Assessment of post-encroachment time in the VISSIM traffic simulation models.

Research Questions

- Does driver behavior in the OSU driving simulator differ significantly from that in the field?
- Does driver behavior in VISSIM simulation models differ significantly from that in the field?
- Is the conventional definition of post-encroachment time appropriate for identifying and measuring right-hook conflicts?

Data Sources

- Frequency of PETs observed in the Oregon State University driving simulator for right-hook conflicts.
- Frequency of PETs observed in the intersection of SW Taylor and SW Naito Pkwy in Portland, Oregon for right-hook conflicts.
- Frequency of PETs observed in traffic simulation models in VISSIM for right-hook conflicts.

To validate the OSU driving simulator, graphical and statistical validation approaches are employed. The graphical validation approach compares the distributions of the frequencies of PET data from the OSU driving simulator to those in the intersection of SW Taylor St and SW Naito Pkwy graphically, grouping the frequency of PETs data by low, moderate, and high risk of collision. Fisher's Exact test is employed to test whether the distribution of frequencies of PETs by group in OSU driving simulator is equal to that of the field. The Exact Multinomial test of goodness of fit analyzes whether the distribution of frequencies of PETs observed in the OSU driving simulator fits the model given by field data. Traffic simulation models in VISSIM are also analyzed using these methods to check its relative validity.

1.2 Organization

This thesis research is organized into six chapters. The first chapter introduces the research and its objectives. The second chapter is a literature review of driving simulator validation, traffic simulation models, and the use of surrogate safety measures. The third chapter describes the Oregon State University driving simulator and Chapter four describes the method, dates, and duration of data collection in the field. Chapter five describes extracting post-encroachment times from video records and the OSU driving simulator and proposes new measures of PET for right-hook conflicts in the field. Chapter six describes graphical and statistical validation approaches for the OSU driving simulator through analysis of frequencies of PETs observed in the OSU driving simulator and the field. Chapter seven describes the application of traffic simulation models in VISSIM for safety research assessment and Chapter eight describes the method of data collection and calibration in VISSIM. Chapter nine summarizes and explains data reduction methods for VISSIM. Chapter ten describes statistical and graphical validation approaches for VISSIM traffic models through analysis of frequencies of PETs between traffic simulation models and the field; finally, Chapter eleven summarizes findings and considers areas of potential research.

2.0 LITERATURE REVIEW

2.1 Driving Simulator

Driving simulators provide a virtual environment that allows for the study of crashes while avoiding damages and costs. However, the challenge with driving simulators is their reflection of actual driving performance on the road. Although driving simulators are not able to replicate all the complexity of the real world, they can be validated as a useful tool if they actually represent the main information of interest in any particular research. Two types of validations are often discussed in the literature.

2.1.1 Physical Fidelity and Behavioral Validity

Physical fidelity measures how similar driving a simulator is to driving an actual vehicle on the road and behavioral validity measures how well a driving simulator reflects actual driving behavior observed in the real world. Physical fidelity is determined by the physical properties of a simulator such as motion, steering control, audio and visual systems. The physical fidelity of the simulator depends on the level of systems applied in the simulator. The low fidelity simulator has basic systems and lower simulator costs, and the high fidelity simulator has advanced systems and higher simulator costs. The behavioral validity is determined by measuring an appropriate driving performance metric under investigation and comparing measured data in the simulator to that observed on the road (Blaauw 1982).

The relationship between the physical fidelity and the behavioral validity of a simulator has been discussed in the literature.

In a study sponsored by FHWA, four simulators with different levels of fidelity and costs were used to investigate how well they contributed to engineers. One hundred sixty-seven subjects, ages 25 to 45 years old, participated in driving simulators. Two road segments and identical scenarios were replicated in four simulators. The spot-speed data of those road segments was collected from both published reports and driving simulators. Comparing mean speeds between simulators and the field data indicated that high physical fidelity simulators had better behavioral validity. (Philips and Morton 2015).

The high fidelity simulator is more likely to demonstrate high behavioral fidelity, however according to Godley et al (2001) the level of physical fidelity does not matter if behavioral validity is not established. A low fidelity simulator can have the same level of behavioral validity as a high fidelity simulator does for a research question under investigation (Godley, Triggs and Fildes 2001).

Behavioral validity may be absolute or relative. Absolute validity is established if a simulator produces the same numerical values of driving performance as those observed in the real world, and relative validity is claimed if numerical values between a simulator and field data are different, but they are in the same direction and have similar magnitude (Godley, Triggs and Fildes 2001). Since most research questions investigate whether causal factors affect results significantly, it is not essential to determine numerical values (absolute validity) but it is necessary to determine the magnitude and direction of values (relative validity) (Tornros 1998).

2.1.2 Driving Simulator Validation

One advantage of simulators is that a vast variety of driving performance measures such as speed, acceleration and deceleration rates, lane position, braking reaction time, headway, and so on that are difficult to gather real-time data on can be easily collected in the simulator. To establish the behavioral validity of a simulator, both research objective and limitations should be considered in selecting an appropriate driving performance measure to be used in validation. Simulator validation research for safety assessment in the literature is summarized below.

Chilakapati (2006) conducted research to investigate the behavioral validity of the University of Central Florida driving simulator in speed and identifying safety countermeasures at high-risk locations. Crash reports from 1999 to 2002 at the intersection of Alafaya Trail (SR-434) and Colonial Drive (SR-50) were used. Free flow speeds of vehicles were recorded using a radar gun during the green phase, around 50 m downstream of each approach at the intersection. The intersection was replicated in the simulator and eight scenarios were designed. Sixty-one objects, aged 16 to over 45, participated while the position and speed of vehicle were recorded in the simulator. Statistical results indicated that speed data followed the normal distribution, and the mean speeds of the simulator and the field data were equal. Overall, the simulator was validated for speed as a traffic parameter at the intersection.

Chilakapati (2006) also validated the UCF simulator for safety assessment. Surrogate safety measures such as average and maximum deceleration, speed at stop line, and following distance were measured in the simulator as safety parameters. The subjects' levels of risky behavior were determined based on those parameters for two approaches. Results indicated that subjects who drove in the approach with high rear-end crash records in the field showed higher risky behavior in the simulator. Hence, the level of risky behaviors corresponded to rear-end crash history records in the field. Therefore, the UCF driving simulator was validated for traffic safety, and it was concluded that the UCF simulator is a useful tool to test high risk locations at intersections.

McGehee et al. (2000) conducted research to validate the Iowa Driving Simulator (IDS) for studying driver performance and the effect of ABS on avoiding a collision in a crash scenario. Sixty men and sixty women, ages 25 to 55, participated in the IDS study, and 129 subjects with the same age range participated in a test site. Experiments lasted 15 minutes in both environments and ended with a crash scenario at an intersection. The test site included 3.5 laps with three intersections. Real vehicles and drivers were used at intersections, except in the last lap a mock-up vehicle with regular car dimensions was used for a crash. Several measures were compared between simulator and test site experiments. Brake reaction time (2.2 sec vs 2.3 sec) and time-to-initial steering (1.64 sec vs 1.67 sec) were equivalent using a 95th percentile confidence interval, but time to throttle (0.96 sec vs 1.28 sec) was not statistically equivalent due to some methodological differences. Overall, IDS was validated for the safety assessment.

Engen (2008) conducted a series of experiments to validate the NTNU/SINTEF driving simulator for reaction time, speed and lateral position, and time gap. For reaction time, each subject was exposed to eight crash scenarios during which driving and reaction times in the simulator were measured. The reaction times were also measured in the field at three sites in six hours of video records. Simulator results complied with results from the literature review and field studies at three sites. Hence, the NTNU/SINTEF driving simulator was validated for studying driver reaction time.

Engen (2008) studied the effect of road markings on speed and lateral position through the NTNU/SINTEF driving simulator. A real road (E6 Støren-Soknedal) with two different road widths was replicated in the simulator, and driving speeds and lateral positions were collected. “Fifteen test subjects drove the 10 metre road and 14 others drove the 8.5 metre road. Each subject drove approximately 8 minutes on the 10 to 11 km test road. The speed limit was 80 km/h” (Engen 2008). Simulator data were compared with three sources of real data and results indicated that absolute and relative mean speeds and mean lateral positions were of equal size. Smaller variation in the simulator was explained by the experimental situation in the simulator and other confounding variables in the real road.

Engen (2008) also examined the ability of the NTNU/SINTEF driving simulator to study the interaction between adjacent vehicles. Time gaps less than 5 seconds were measured and analyzed as an appropriate parameter to validate the simulator. The simulator data was compared with the data measured in the field through an instrumented vehicle as well as roadside measurement data at six sites obtained from other research. Results indicated that the mean time gap between the simulator and the instrumented vehicle differed significantly, and the mean time gap in six sites was between the mean time gap in the simulator and the instrumented vehicle. However, differences among the driving simulator, instrumented vehicle, and roadside measurements were explained through different situations, so in light of these differences, all measurements were found acceptable.

Brown (2012) conducted research to validate the Oregon State University driving simulator for speed, acceleration and deceleration rate data. A total of 10 subjects drove two segments of a road. Those two segments were replicated in the simulator environment and participated in the simulator experiment. An actual vehicle was equipped with a "CarChip E/X" device in the road test to collect speed and travel time data, calculating acceleration and deceleration rates. Simulator data was compared with data measured in the road. Statistical results indicated that mean, maximum, and the 85th percentile of speed, acceleration, and deceleration differed significantly between the two environments. However, in practice, differences fell in an acceptable range (speed $\leq$ 5 mph, acceleration and deceleration rates $\leq$ 1.6 ft/sec^2). As a result, relative validity was established for the OSU driving simulator.

2.2 Traffic Simulation Models

Traffic simulation models have been mostly used to evaluate traffic efficiency in areas such as operation, planning, and ITS technologies because of their ability to study traffic operations in a network (Byungkyu and Jongsun 2006). Transportation professionals have found potential application of traffic simulation models for traffic safety assessment as well. Their use is attractive due to their ability to study driving behavior without risk of casualties and damages. Both driving simulators and simulation models are useful tools for research, so long as they provide reliable information that reasonably reflects real road driving performance. For simulation models, validity is established when traffic parameters generated by the simulation model and those observed in the field are reasonably similar. Previous research on the ability of simulation models to study traffic safety is summarized below.

The Federal Highway Administration (FHWA) sponsored a research project to develop the Surrogate Safety Assessment Model (SSAM). SSAM can extract data from existing simulation models, identify conflict events, compute surrogate safety measures and classify conflicts into three maneuver types including crossing, lane change, and rear-end conflicts. A rear-end conflict results if the conflict angle is less than 2 degrees. If conflict angle is larger than 45 degrees, it is a lane change conflict, and if conflict angle is between 2 and 45 degrees, it is a crossing conflict. (Gettman, et al. 2008). In right-hook conflicts, conflict angle between right-turning motorist and through bicyclist can range from 10 degrees to 90 degrees, and therefore they may be classified as either crossing or lane change conflicts in the SSAM approach. Gettman et al. (2008) conducted research to validate the SSAM approach. A total of 83 real signalized intersections were simulated in traffic simulation models in VISSIM, and SSAM identified traffic conflicts, computed surrogate safety measures, and classified conflicts into three maneuver types. SSAM results were compared to the actual crash history records. Results showed that the frequency of conflicts by type in SSAM significantly differed from the frequency of historical crashes by type. The ratio of conflicts-per-hour to crashes-per-year for crossing conflicts was close to zero, equal to 0.01 indicating that frequency of crossing conflicts identified in SSAM did not reflect frequency of crossing crashes in the real world. This ratio was 0.65 for lane change conflicts. Also, the average of hourly conflicts for crossing and lane change conflicts and percentage of crossing and lane change conflicts in SSAM substantially differed from the average yearly crashes for crossing and lane change crashes and percentage of crossing and lane change crashes in the real world. These differences revealed that either SSAM is not able to correctly identify all crossing and lane change types of simulated conflicts in VISSIM or VISSIM traffic simulation models are not able to represent a true frequency of crossing or lane change types of conflicts, or both. Furthermore, traditional volume-based prediction models were still a better representative of crash records than the SSAM approach. However, SSAM can help analyze traffic facilities and control policies before they are implemented. Overall, the conclusion was that the validation of SSAM approach was promising, but not definitive.

Archer (2004) investigated how well safety indicators such as TTC, PET, and their severity, defined by required braking rate severity (RBR-severity) in VISSIM simulation models matched those observed in the real road. Safety indicators were measured for five different types of conflicts during three time periods including morning-peak, off-peak, and afternoon-peak hours at three T-junctions. Three T-junctions were replicated and calibrated in the simulation models. Safety indicators in the field were compared with those in the calibrated simulation models. The average TTC frequencies and their RBR severity measures over three time-periods in the simulation showed very high consistency with those observed in the field. The average PET frequencies and their required braking rate (RBR) severity in the simulation showed little consistency with those observed in the field. Because PET was not a useful indicator for road users traveling in the same direction, PET values were not measured for two conflicts. However, results showed a very consistent pattern of order for the other three conflicts within three time periods between simulated PET frequencies and PET frequencies observed in the field (Archer 2004).

Sayed et al. (1994) applied a traffic computer simulation model, called General Purpose Simulation System (GPSS/H) to study traffic conflicts. Three types of conflicts were identified at unsignalized intersections and simulated TTC values were measured as a severity measure with the threshold value of 1.5 seconds. A total of four intersections and 32 hours of video records for each intersection were used to identify traffic conflicts. The severity of conflicts were measured based on a combination of TTC, and risk of collision (ROC) scores ranging between 2 (a low-risk conflict) and 6 (a high-risk conflict). The number of conflicts was compared between the simulated conflicts and the observed conflicts. Results indicated that the distribution of conflicts was very close between the two environments. They also investigated the effect of volume and speed on the frequency and severity of conflicts. Overall, the simulation models were validated for safety assessment at four unsignalized intersections under investigation.

Huang et al. (2013) conducted research to investigate the validity of VISSIM simulation models and the SSAM approach in identifying traffic conflicts at signalized intersections. Traffic conflicts were identified from 80 hours of video records at 10 intersections in the field, and volume and geometric configuration information was collected and applied in the VISSIM simulation models. In total, 1774 rear-end, 551 lane change and 572 crossing conflicts were observed in the field. Traffic volumes were extracted from 32 hours of video at eight intersections and 32 separate models were created, each corresponding to an hour of video in the relevant intersection. A two-stage calibration procedure compared traffic conflicts and TTC values between simulated models and the field observations. The minimum gap time parameter was also calibrated from 3 seconds to 2 seconds in the VISSIM simulation models. Linear regression models and the Spearman rank correlation were analyzed to study the relationship between the calibrated simulated conflicts and the observed conflicts. The results indicated that calibrated simulated conflicts were reasonable indicators for rear-end conflicts and the total number of observed conflicts in the field but they were only moderate indicators for lane change and crossing conflicts in the field.

Zhou and Huang (2013) used the simulation model in VISSIM to evaluate safety performance at a signalized intersection. Traffic conflicts were measured from video records in the field. Using a radar gun and a measuring wheel, geometric characteristics and volume data of the intersection were replicated in the simulation model. After calibration, the trajectory files were extracted from the VISSIM outputs and applied in SSAM to identify simulated conflicts. Simulated conflicts were compared with conflicts observed in the field to validate the simulation model. After validation, simulation conflicts were compared under different speed limits. Results indicated that reducing the speed limit would improve the safety performance of the intersection.

2.2.1 Calibration

Driving simulators are able to directly study driver's behavior because an actual subject drives the simulator, but predefined parameters and default values for a driver's behavior are used in traffic simulation models. In order to validate simulation models, it is essential to calibrate these parameters for the specific segment under investigation because drivers show different behaviors in different segments of the road. Otherwise, the simulation results will be significantly different from the field data, and the validation of the simulation model will be rejected. While there are many proposed calibration procedures such as linear regression and genetic algorithms in the literature (see e.g. Miller, 2009; Archer 2004; Park & Qi 2005), procedures applicable to this research are outlined below.

The Oregon Department of Transportation (ODOT) (2011) developed a protocol for VISSIM simulation models and a calibration procedure in which a universal measure, GEH, was used to compare the observed volumes in the field with the volumes in the simulation output. If the differences calculated by the GEH formula result in a value less than 5, the calibration is appropriate. Another important calibration criterion is the minimum number of simulation runs, N. Both the GEH and N equations are shown below.

$$GEH = \sqrt{\frac{2(m-c)^2}{m+c}}$$

$m = output\ traffic\ volume\ from\ the\ simulation\ model\ (vph)$

$c = input\ traffic\ volume\ (vph)$

$$N = \left(2 * t_{0.025,N-1}\frac{S}{R}\right)^2$$

$R = 95 - percent\ confidence\ interval\ for\ the\ true\ mean$

$t_{0.025,N-1} = Student's\ t\_statistic\ for\ two\_sided\ error\ of\ 2.5\ percent\ (total\ 5\ percent)$

$with\ N-1\ degree\ of\ freedom$

$S = standard\ deviation\ of\ about\ the\ mean\ for\ selected\ MOE$

$N = number\ of\ required\ simulation\ runs$

Huang et al. (2013) used the mean absolute percent error (MAPE) to measure differences between observed and simulated conflicts. A small MAPE value indicates little difference between simulated conflicts and observed conflicts, suggesting that the calibration of the simulation model is appropriate. The Spearman rank correlation coefficient (ρ_S), a non-parametric statistical test, is used to evaluate the correlation between simulated conflicts and observed conflicts based on their safety ranking. A coefficient of one represents a perfect correlation and a coefficient of zero represents no correlation between observed and simulated traffic conflicts. They also use a simple linear regression model to determine the percentage of the variation in the observed data explained by the simulation model. Both MAPE and ρ_s equations are given below.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{c_m^i - c_f^i}{c_f^i}\right|$$

$n = the\ number\ of\ observations$

$c_m^i = the\ number\ of\ the\ simulated\ conflicts\ for\ time\ interval\ i$

$c_f^i = the\ number\ of\ the\ observed\ conflicts\ in\ the\ field\ for\ time\ interval\ i$

$$\rho_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

$d_i = the\ difference\ between\ ranks\ for\ observation\ i$

$n = the\ number\ of\ observations$

2.3 Post-Encroahment Time (PET), A Surrogate Safety Measure

Allen et al. (1977) defined the post-encroachment time as "*the time from the end of encroachment to the time that the through vehicle actually arrives at the potential point of collision*" (Allen, Shin and Cooper 1977). This concept is illustrated in Figure 2-1.

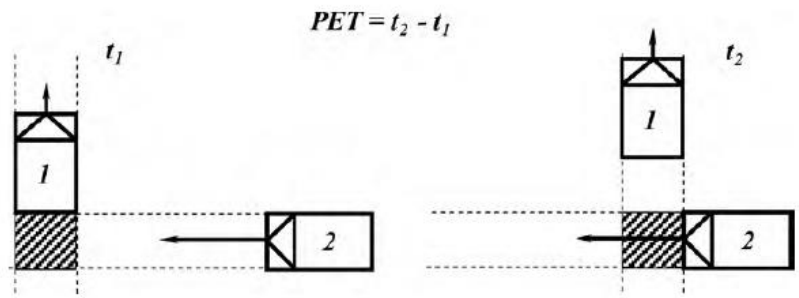


Figure 2-1: Conventional definition of PET (Allen et al., 1977).

Killi, and Vedagiri (2014) proposed grids as conflict areas and a rhombus of grids is used to differentiate between a close conflict and a far conflict (Figure **2-2**) (Killi and Vedagiri 2014).

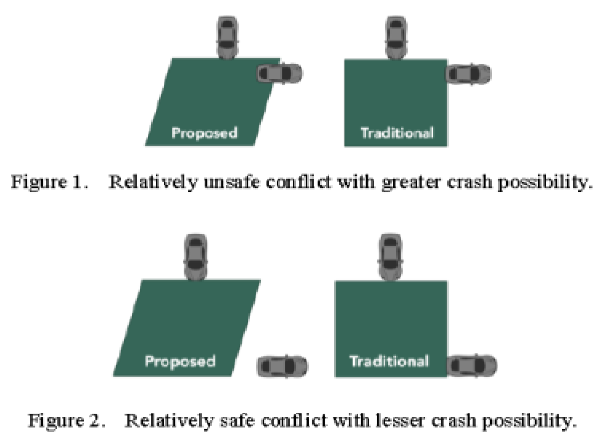


Figure 2-2: Relatively unsafe and safe conflicts (Kili and Vedagiri, 2014).

Laureshyn et al. (2010) mapped the trajectory of road users and defined PET as the minimal delay between their trajectories (Figure **2-3**), (Laureshyn, Svensson and Hyden 2010).

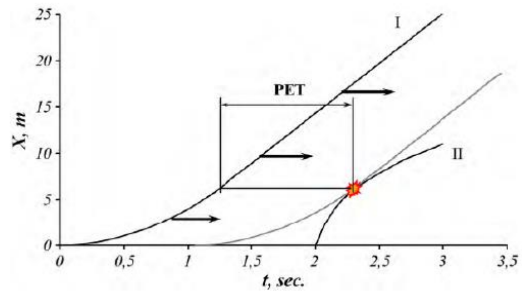


Figure 2-3: “Delay”-based definition of PET (Laureshyn et al., 2010).

2.3.1 Application of PET to Angle Collisions

The literature review indicates that PET is an appropriate indicator of right-angle collisions such as right-hook crashes. Songchitruksa and Tarko (2006) found a positive correlation between the risk of right-angle collisions, PETs, and traffic volumes. They concluded that PET can represent traffic interactions well for right-angle collisions (Songchitruksa and Tarko 2006). Alhajyaseen (2014) found that PET is the best surrogate measure in application for angle conflicts. Furthermore, measuring PET is more convenient than other indicators as it does not require measuring relative speed and distance (Songchitruksa and Tarko 2006)

2.3.2 Critical PET

In order to study traffic safety performance related to crashes, it is important to identify conflict events that represent an interaction and a risk of collision between two road users. Because PET is an indicator of traffic conflicts, a threshold value of PET that accounts for the lowest level of collision must be determined. Pessapati et al. (2013) found that a PET threshold value of one second provided the best correlation with opposing left-turn crashes (Peesapati, Hunter and Rodgers 2013). Research by Tang and Kuwahra (2011) concluded that the application of a minimum value of PET of approximately two seconds in design of all-red time at signalized intersections could achieve significant operational and safety benefits (Tang and Kuwahara 2011). In other words, a PET of less than two seconds would result in an interaction and a sufficiently high risk of a collision between road users.

2.4 Application of Surrogate Safety Measures

The traditional approach to road safety analysis studies the frequency and severity of crashes. However, crashes are rare events and it takes at least two to three years to collect data on a large enough number of crashes. Small sample size in crash records makes it difficult and unreliable to study traffic safety performance. An alternative, innovative approach to investigate safety performance is called traffic conflict technique (TCT). TCT determines the risk of collision by identifying conflict events and near-miss crashes instead of crashes. Conflict events can be identified and measured by safety indicators which are called surrogate safety measures. According to Tarko et al. (2009), road safety analysis can benefit from surrogate safety measures instead of the accumulation of crash records. Tarko et al (2009) defines "*A traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their* *movements remain unchanged*." Laureshyn (2010) defines a near-miss as "*a situation when two road users unintentionally pass each other with a very small margin, so that the general feeling is that a collision was "near."* The purpose of studying surrogate safety measures is to gain insight into the frequency and severity of crashes at a specific location. The most acknowledged surrogate safety measures include time to collision (TTC), post-encroachment time (PET), gap time (GT), deceleration rate (DR), maximum speed, and speed deferential (SD) (Douglas and Head 2003). Although the relationship between traffic conflicts and crashes has been subject to a great deal of controversy over the last several years, literature supporting the use of traffic conflicts instead of historical crash data follows.

Glauz et al. (1985) conducted research to establish a relationship between traffic conflicts and crashes. Data was collected at 46 intersections in the greater Kansas City area in 1982, and the ratio of accident per conflict was determined for each type of collision and each type of intersection. Using these ratios, the expected number of each type of collisions was determined. Results suggested that traffic conflicts are good surrogates for accidents and they can estimate the average accident rate as accurately as historical accident data.

Sayed and Zein (1998) used collected data from 52 signalized and 42 unsignalized intersections across British Columbia, and established traffic conflict standards. They found strong correlation between accidents and traffic conflicts for signalized intersection models and very weak correlation between those across unsignalized intersection models.

Sayed and Zein (1998) also conducted a traffic conflict survey to identify the causes of six crashes within three years at the unsignalized intersection of Highway 97 and Oyama Road in Okanagan region of British Columbia. A total of 56 conflicts were collected in two days. Results showed that interactions between drivers crossing Highway 97 from Oyama Road and pedestrians crossing the crosswalk on Highway 97 resulted in greater than average conflicts compared to other similar intersections. These results revealed that drivers on Highway 97 heading south failed to see the intersection or crosswalk. As a solution, they recommended improving traffic safety by providing a traffic signal with warning flashers and pedestrian push-button activation.

Bai et al. (2015) conducted research to identify factors that affect the frequency of traffic conflicts between motorized vehicles and electric bicycles, including e-bikes and e-scooters. The frequency of traffic conflicts was observed for three types of conflicts including right-hook conflicts at 14 intersections. There were 1472 right-hook conflicts observed during 162 hours of video records. Time to collision (TTC) was measured to identify traffic conflicts. Results indicated that the frequency of right-hook conflicts in 30-minute intervals followed a negative binomial distribution. Thus, a generalized linear regression model was developed to identify determining factors of right-hook conflicts. Results indicated that the increase in e-bikes, e-scooters, and the volume of right-turning vehicles result in an increase in the frequency of right-hook conflicts, on average, by 0.49%, 1.62%, 0.40%, respectively. The frequency of right-hook conflicts during peak periods was on average 29% greater than those in non-peak periods. Finally, the presence of traffic channelization reduced the average frequency of right-hook conflicts by 72%.

Zangenehpour et al. (2014) conducted research to investigate the safety of cycle tracks at intersections where a right-turning vehicle interacts with a through movement bicyclist (right-hook scenario). PETs were collected and measured by using a tracking tool and video records at 23 intersections, including eight intersections with a cycle track on the right side, seven intersections with a cycle track on the left side, and eight intersections without a cycle track. Random effects ordered logit models were developed for each type of intersection. These models were compared across three types of intersections. Results indicated that intersections with a cycle track on the right side are safer than intersections with a cycle track on the left side and intersections without a cycle track. However, intersections with a cycle track on the left side did not improve the safety of intersections without a cycle track.

Songchitruksa and Tarko (2006) developed regression models to estimate the expected frequency of right-angle collisions by using the frequency of PETs at 16 signalized intersections. Eight hours of videos were recorded for each intersection and traffic volumes and PETs were collected at intersections. Results of regression models indicated that the frequency of PETs is a key factor in determining the expected number of right-angle collisions and different safety levels across locations.

2.5 Summary

The literature review supports this research in several aspects. High fidelity simulators are more likely to establish behavioral validity, and behavioral validity is determined by comparing an appropriate performance metric between simulator data and field observations. A performance metric measures a key action of the driving performance under investigation. If relative validity is established, a driving simulator is sufficient for studying most research questions. Surrogate safety measures such as TTC and PET are appropriate performance metrics for traffic safety assessment. They can be applied to validate driving simulators. Traffic simulation models are also used to study traffic safety performance. However, calibration is critical for the validation of traffic simulation models. PET is an appropriate performance metric of driving performance for studying angle collisions such as right-hook crashes. An added benefit of PET is that it can be measured more easily than other surrogate safety measures. A PET of less than two seconds suggests an interaction and a risk of collision between two road users approaching one another.

3.0 THE OREGON STATE UNIVERSITY DRIVING SIMULATOR

According to Oregon State University,

> The Oregon State driving simulator is a high-fidelity moving-base simulator. The simulator consists of a full size 2009 Ford Fusion cab mounted on top of a high performance electric pitch motion system. […] The pitch motion system allows for on set cues during acceleration and braking events. The motion base moves +/- 4 degrees with the center of rotation around the driver head position. […] Three LCOS projectors with a resolution of 1400 x 1050 are used to project a 180 degrees by 40 degrees front view, these front screens measure to 11 feet by 7.5 feet. A DLP projector is used to display a rear image for the driver's center mirror. The two side mirrors have embedded LCD displays. Sound is provided by surround sound speakers capable of 500 watts. […] The vehicle cab instruments are fully functional and include a steering control loading system to accurately represent steering torques based on vehicle speed and steering angle. […] The computer system consists of a quad core host running Realtime Technologies SimCreator Software. The vehicle model is a 15-dof multi-body chassis model with a combined Pacejka tire model. The visual system is comprised of dual core computers each running a Nvidia 280 graphics card. The update rate for the graphics is 60 Hz.

Figure 3-1 and Figure 3-2 display the Oregon State University Driving Simulator and the simulated environment.

Figure 3-1: The mounted rear projector and rear projection screen (Right Panel) and the driver's view looking over the right shoulder out the rear vehicle window (Left Panel) (OSU).

Figure 3-2: Simulated Environment in OSU driving simulator (OSU).

3.1 Safety Parameter of Driving Performance

Hurwitz et al. (2015) conducted research using the OSU driving simulator to investigate safety performance of alternative traffic control strategies that reduce crashes between right-turning vehicles and through bicyclists at signalized intersections. Although the OSU driving simulator has already been validated for speed, acceleration and deceleration rates for traffic parameters (Brown 2012), it is necessary to validate the OSU driving simulator for the safety assessment in the right-hook project. Hurwitz et al. (2015) considered post-encroachment time an appropriate safety parameter of driving performance because PET represents the risk of collision and interaction between right-turning vehicles and through bicyclists at intersections.

4.0 DATA ASSEMBLY

To validate the OSU driving simulator, it is important to study an intersection that closely resembles the intersections in the simulator. The main characteristics of the intersections in the OSU driving simulator include a signalized intersection, an approaching single through lane, two opposite through lanes, no right-turn or left-turn lanes, and a striped bicycle lane with no bike box (Hurwitz, et al. 2015). The intersection of SW Taylor St and SW Naito Pkwy in Portland, Oregon was selected as the study location.

4.1 Study Location

Unfortunately, no one intersection was found that would match with the simulated intersections exactly. As a result, the intersection differed from the designed intersections in the OSU driving simulator in terms of geometric designs, traffic conditions, phasing and traffic signal plan. The intersection had oncoming protected left-turn lanes that eliminated the conflict between approaching right-turning and oncoming left-turning vehicles. The designed intersections in the OSU driving simulator had different numbers of lanes and widths from the intersection in the field. The traffic signal in the intersection also differed from the designed intersections in the OSU driving simulator and the intersection in the field. The intersection approach in the OSU driving simulator had a speed limit of 35 mph, but the intersection of SW Taylor and SW Naito Pkwy in the field had a speed limit of 30 mph. The speed of bicyclists was constant in designed scenarios in the OSU driving simulator, while bicyclists had variable speeds in the field. It will be shown later that these different conditions have significant effect on the validation of the OSU driving simulator. Figure 4-1 through Figure 4-4 depict the environment and the geometric design of each intersection. Table 3-1 shows a summary of critical parameters.

Figure 4-1: Simulated intersection designed in the OSU driving simulator (OSU).

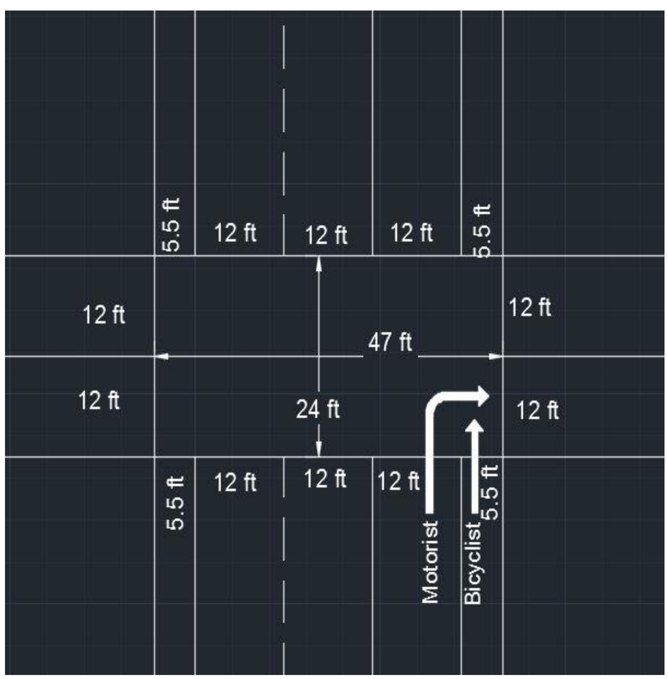


Figure 4-2: The dimensions of the simulated intersection designed in the OSU driving simulator.

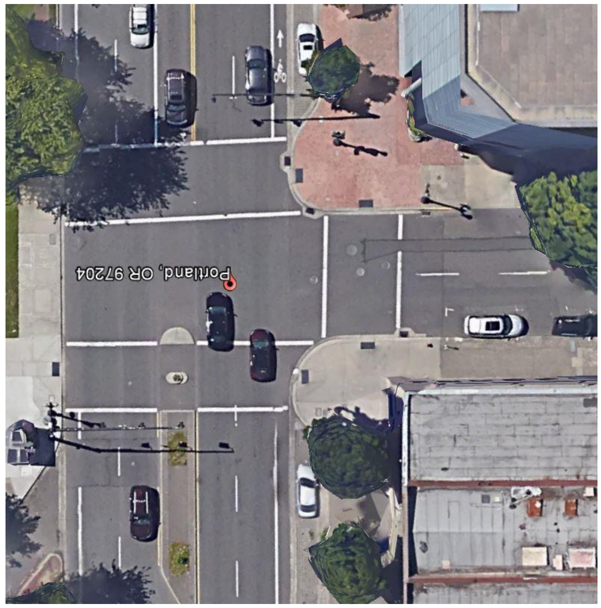


Figure 4-3: SW Taylor St & SW Naito Pkwy (Google Map).

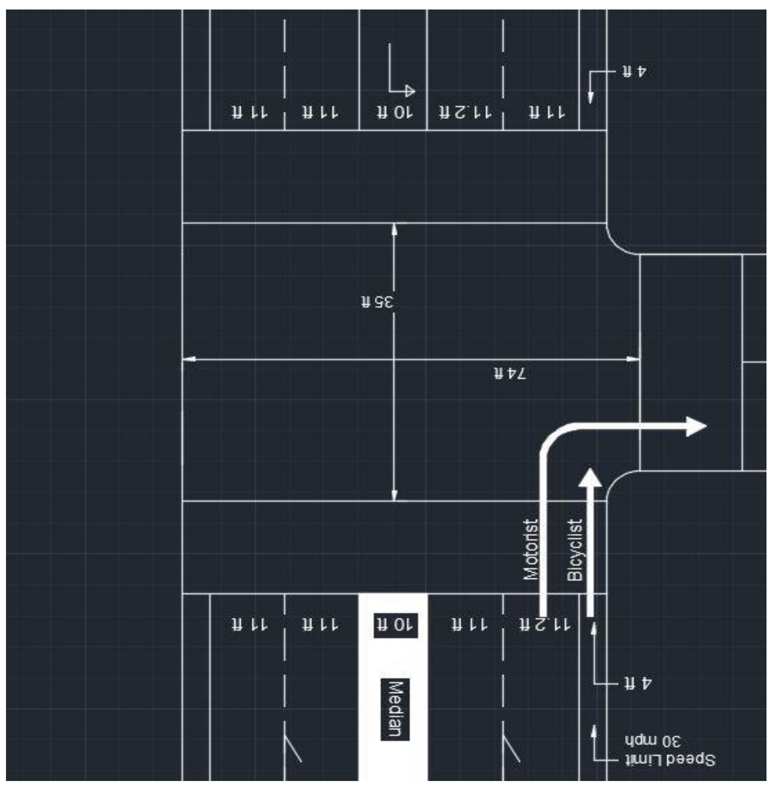


Figure 4-4: The dimensions of SW Taylor St & SW Naito Pkwy.

Table 4-1: Summary of Critical Parameters

Intersections	Dimension	# Approaching Through Lanes	Width of Bike Lane	Speed Limit	Protected on-coming left turn	Area Type
Simulated Intersections	47 × 24 ft.	1	5.5 ft.	35 mph	No	Suburban
SW Taylor & SW Naito Pkwy	74 × 35 ft.	2	4 ft.	30 mph	Yes	Urban

4.1.1 Field Setup

Two video cameras were connected to the top of a 10-foot bar and attached to a light pole on the northwestern corner of the intersection of SW Taylor and SW Naito Pkwy. As is shown in Figure 4-5, one camera showed the crosswalk and the bike lane in SW Naito Pkwy. Another camera showed the crosswalk in SW Taylor St and the oncoming vehicles in SW Naito Pkwy. The views of the two cameras overlapped each other providing a continuous observation of bicyclists and motorists. A box, attached and secured to the light pole, supplied electricity to cameras and video was recorded by memory card between 8:00 AM and 7:00 PM.

Figure 4-5: Camera view of the SW Naito Pkwy and the SW Taylor St.

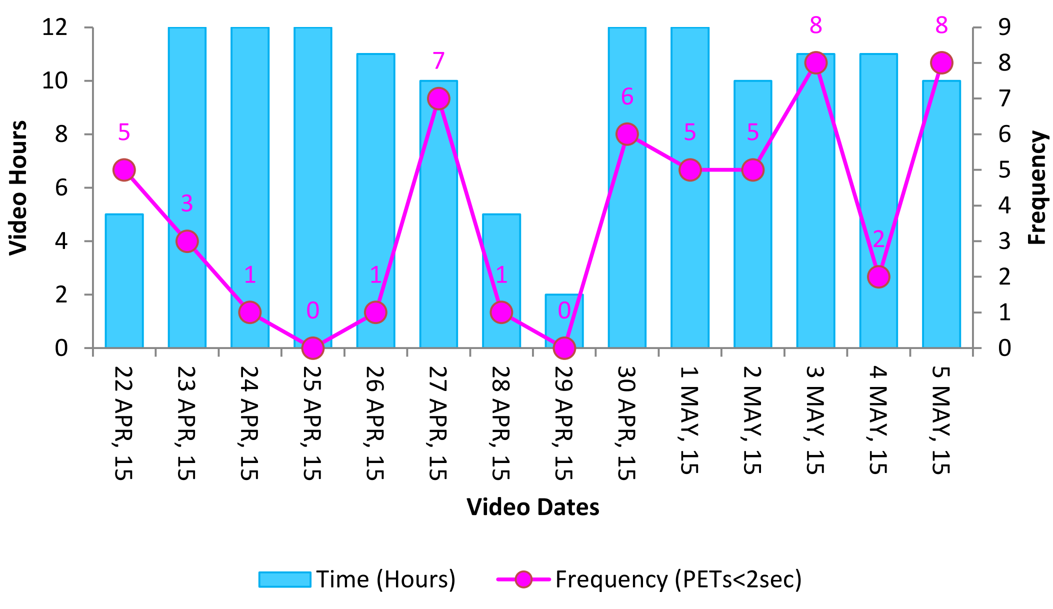
4.1.2 Field Observations

Cameras recorded a total of 135 hours of video between April 22, 2015 and May 5, 2015 at the intersection of SW Taylor Street and SW Naito Pkwy.

Dates, hours of video records, and frequency of conflicts corresponding to the intersection in the field and those in the OSU driving simulator are summarized in Table 4-2 and shown in Figure 4-6.

Table 4-2: Summary of data collection in the OSU driving simulator and the intersection of SW Taylor and SW Naito Pkwy in the field.

Study Area	Days	Hours	Freq. PETs < 5_{sec}	Freq. PETs < 2_{sec}
OSU Driving Simulator	---	22	153	50
SW Taylor & SW Naito	14	135	159	52



SW Taylor St & SW Naito Pkwy

Figure 4-6: The total number of dates, hours of video records, and the frequency of PETs at SW Taylor St & SW Naito Pkwy.

5.0 DATA REDUCTION

After detecting conflicts by analyzing video records, the SMPlayer program was used to measure PETs less than 5 seconds, frame by frame, where a second comprises 20 frames. The speed of bicyclists (ft./sec) was determined by dividing the width of the crosswalk (20 ft.) by the time a bicyclist passed the crosswalk at the upstream intersection in the SW Naito Pkwy. Finally, PETs less than 2 seconds were checked and edited by Kinovea, a video editor program. It may be noted that the process of identifying conflicts and measuring PETs from video records motivated the development of the conventional definition of PET proposed in this work.

5.1 Extracting PETs from Video Records

The risk of collision is emphasized in the definition of the traffic conflict and the near-miss in the literature (Tarko, et al. 2009), (Laureshyn, Svensson and Hyden 2010). In fact, traffic conflicts are used as alternative data for crashes, and surrogate safety measures such as PET and TTC are indicators of traffic conflicts, their values reflecting a risk of collision. According to the literature, a PET less than two seconds represents conflicts with a risk of collision and an interaction between road users (Tang and Kuwahara 2011). However, it takes considerable time to collect a small sample size of PETs less than two seconds. In this research, collecting 52 PETs less than two seconds took 135 hours of video records over 14 days. Additionally, small sample size makes the analysis sensitive to any changes, especially if frequencies of

PETs are analyzed instead of raw PET values. Hence, it is vital to capture and measure all severe conflicts (PET < 2sec) correctly.

The conventional definition of PET by Allen et al (1978) only measures the time from the end of the first vehicle encroachment in the potential area of collision. In fact, it assumes that road users are two homogeneous vehicles, so a collision is inevitable if the second vehicle arrives at potential area of collision before the end of the first vehicle encroachment. However, a vehicle and a bicyclist are our road users in a right-hook conflict and field observations showed that some bicyclists arrived in the potential area of collision before the end of vehicle encroachment, yet the two road users were able to avoid collision. The choice of potential area of collision is also important; it depends on the type of conflict and road users. In right-hook conflicts, vehicles sometimes create high-risk conflicts with bicyclists before they even enter the bike line. Therefore, the area of potential collision was determined to be the bike lane plus an additional one-foot buffer. Finally, some interactions were observed in which the PET value was less than two seconds with no true risk of collision and thus these observations should not be recorded as conflicts.

5.2 Proposed Measures of PET for Right-Hook Conflicts

In order to identify all right-hook conflicts and measure them correctly, conflicts were grouped into two types. In the first type of right-hook conflict, the conventional definition of PET was modified so that observed conflicts in which the motorist did not end the encroachment but the bicyclist arrived at potential area of collision were measured and recorded as right-hook conflicts. For understanding the second type of right-hook conflict, we note that encroachment time (ET) is another surrogate safety measure defined by Allen et al. (1978). The conventional definition of ET is the "*time duration during which the turning vehicle infringes upon the right-of-way of through vehicle.*" This definition is not a time interval between two road users but measures the encroachment duration of the turning vehicle in a potential area of collision. For the second type of right-hook conflicts, we use a definition related to both PET and ET concepts. The time interval of encroachment between the turning vehicle and through bicyclist is measured, the conventional PET concept, but it begins at the encroachment of the turning vehicle into the potential area of collision, an ET concept. In the conventional definition, it begins at the end of the vehicle encroachment into the area of collision. Modified definitions of PET for right-hook conflicts are described and shown graphically in Figure 5-1 through Figure 5-3.

5.2.1 Type I: After Vehicle Occupation

PET is the interval of time from vehicle occupation of a potential area (line or point) of collision to the time the bicyclist arrives at the potential area of collision (Figure 5-1).

Vehicle occupation occurs when the center of the vehicle is located in the center of the potential area of collision. Potential area of collision was determined to be the bike lane plus a one-foot buffer for a right-hook conflict after careful analysis of video records at the intersection of SW Taylor St and SW Naito Pkwy.

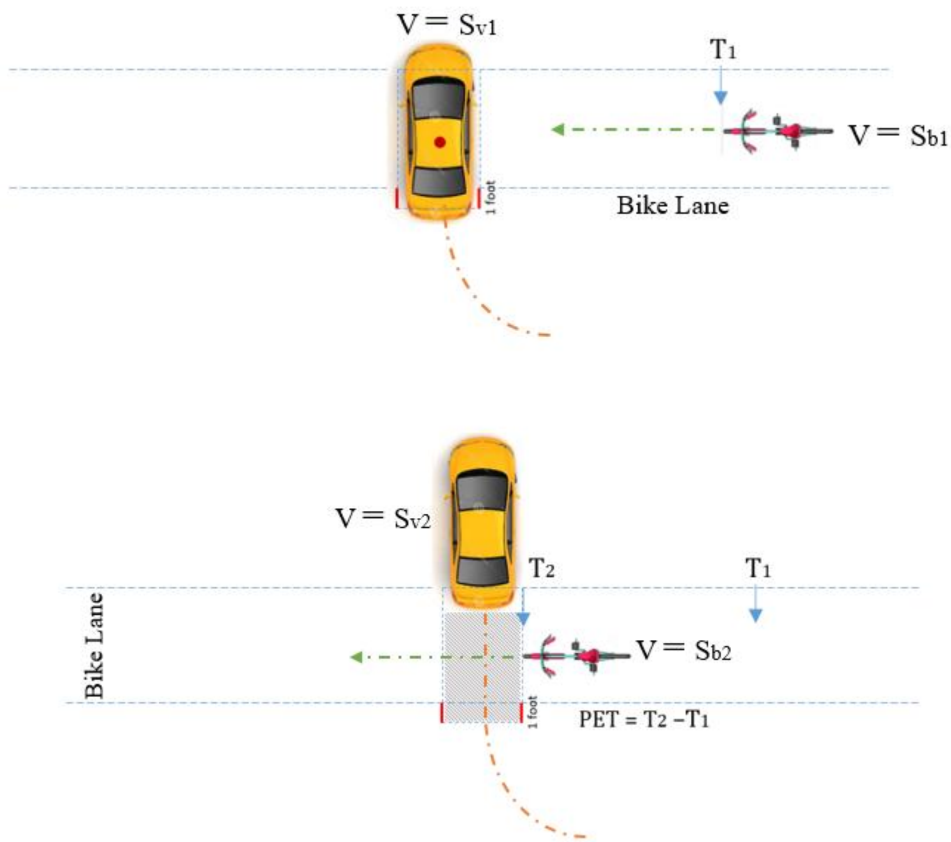


Figure 5-1: Bicyclist arrives at conflict area after vehicle occupation of conflict area (Type I).

5.2.2 Type II: Before Vehicle Occupation

PET is the time from vehicle encroachment into the potential area (line, point) of collision to the time the bicyclist and motorist take evasive actions (Figure 5-2).

An evasive action consists of the activation of brakes (V→ 0) or a noted change in direction. PET should not be measured if the bicyclist waits behind the potential area of collision for the vehicle to cross the bike lane as the vehicle and bicyclist are not on a collision course (Figure 5-3).

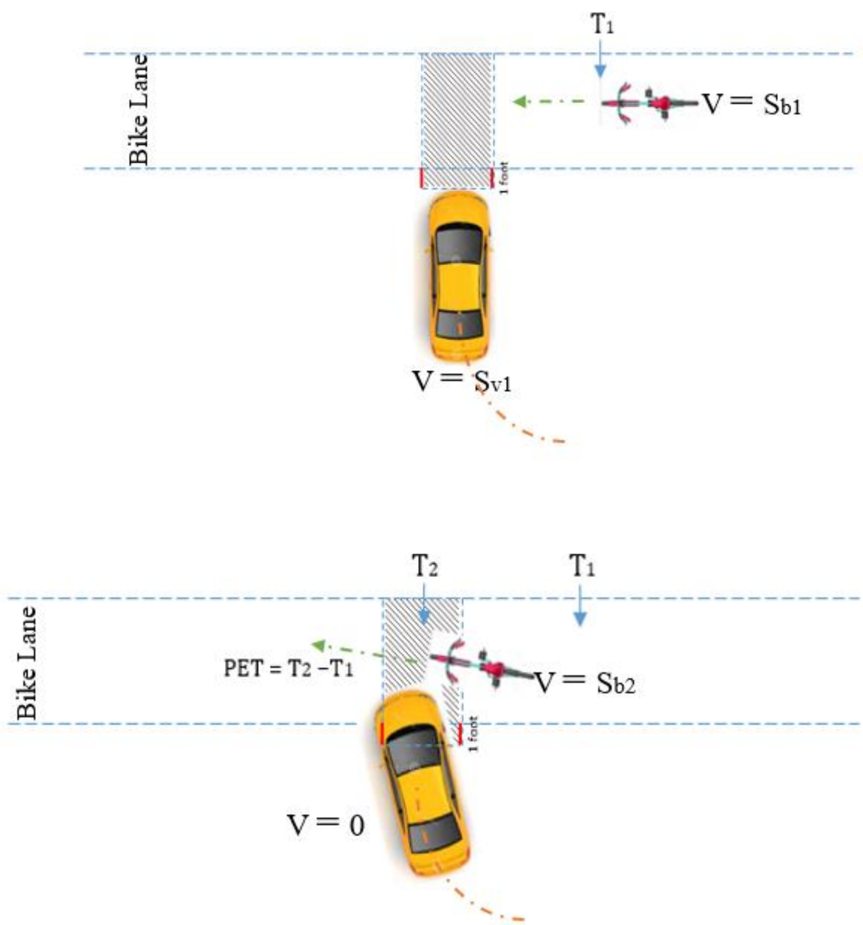


Figure 5-2: Motorist stops and bicyclist changes his direction (Type II).

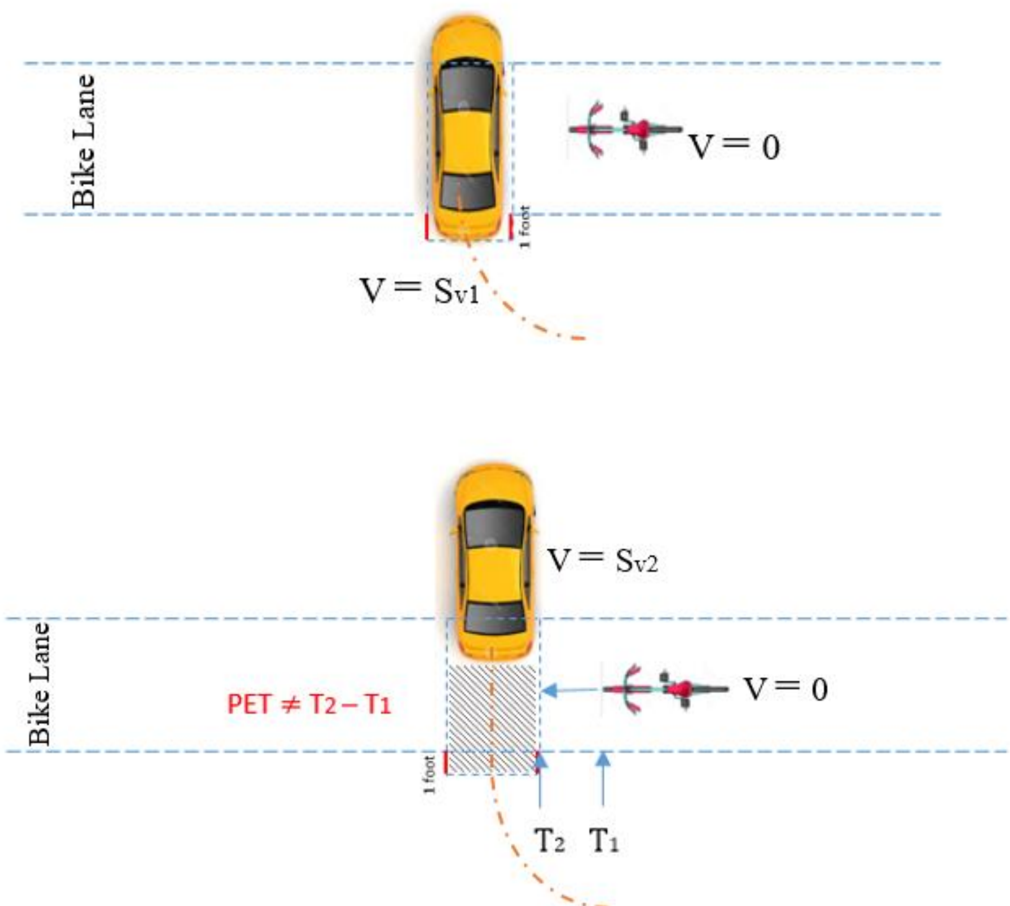


Figure 5-3: Bicyclist waits behind conflict area occupied by motorist.

5.3 Extraction of PETs from the Driving Simulator

In the research project conducted by Hurwitz et al. (2015), PET was calculated by recording the location of the vehicle and bicycle centroids, and the constant speed of bicyclist in the OSU driving simulator. Figure 5-4 displays the right-hook conflict and PET calculation.

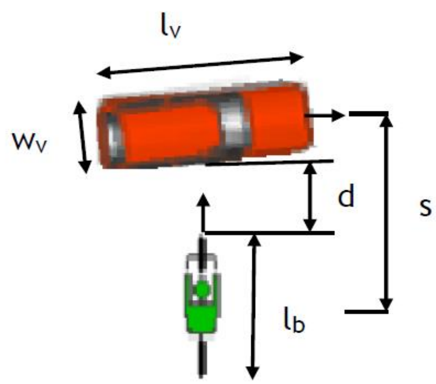


Figure 5-4: PET Calculation for a RH Crash Scenario (Hurwitz et al., 2015).

$$\boldsymbol{PET = \frac{d}{v_b}} \qquad \boldsymbol{d = s - \frac{w_v}{2} - \frac{l_b}{2}}$$

w_v = width of vehicle (i.e., car)
l_b and l_v = length of bicycle and car, respectively
v_v = velocity of car
v_b = velocity of bicycle (Constant)
d = distance from middle point of the side of the car and front of the bicycle
s = center to center distance between bicycle and car

5.4 Summary

PETs measured in the OSU driving simulator fit the proposed measure of PET for type I right-hook conflicts because they are measured from the time that the centroid of the vehicle is in the middle of the bike lane to the time that the bicyclist arrives at that location. Type II conflicts are also measured in the field as described above. Figure 5-5 shows how a PET for a type I conflict is measured from video recorded in the field.

Figure 5-5: Measuring the post-encroachment time for a type I conflict from video recorded in SW Taylor St and SW Naito Pkwy.

6.0 ANALYSIS OF OSU DRIVING SIMULATOR DATA

This chapter investigates the degree to which the OSU driving simulator reflects actual driving behavior in right-hook conflicts at intersections. The PET values were used to identify conflicts and the severity level of risk of collision. Conflicts were also grouped by speed into high-speed bicyclists whose speed was greater than the average bicycle speed in the field (≥13.6 mph) and low-speed bicyclists whose speed was less than the average bicycle speed in the field (<13.6 mph) so that the conflicts are comparable to scenarios designed in the simulator.

6.1 Summary of Data Collection

As discussed in the literature review, PET is an appropriate representative measure of angle collisions. A threshold value of PET between one and two seconds represents a risk of collision (Pessapati et al 2013 & Tang and Kuwahra 2011). PETs were grouped into three time intervals based on their risk of collision including high risk ($0 < PET < 1$ sec), moderate risk ($1 \leq PET < 1.5$ sec), and low risk ($1.5 \leq PET < 2.0$ sec). PETs larger than two seconds were removed from analysis as they do not represent an interaction and a risk of collision between road users. The frequency of each group of PETs in the OSU driving simulator and the intersection of SW Taylor and SW Naito in the field is shown in Table 6-1.

Table 6-1: Frequency of each group of PETs

	Frequency of PETs			
	High Risk (0-1)	**Medium Risk [1-1.5)**	**Low Risk [1.5-2)**	**Total**
SW Taylor & SW Naito	17	18	17	52
OSU Driving Simulator	8	18	24	50

Descriptive Data Analysis Table 6-2 summarizes a total of 52 PET values of observed conflicts less than two seconds and speeds of bicyclists at the intersection of SW Taylor & SW Naito. PETs were measured through video, mostly recorded between 8:00 AM and 7:00 PM. Average PET and PET standard deviation were 1.2 seconds and 0.46 seconds, respectively. Average bicyclist speed and the speed standard deviation were 13.6 mph and 3.5 mph, respectively. Pedestrians were present in five conflicts. Sixty percent of vehicles were SUV, pick-up, van, or truck. Minimum PET was 0.3 seconds, representing the highest risk of collision. Maximum bicyclist speed was 22.7 mph. A total of four type II conflicts with high risk of collision, including the minimum PET, were observed in the intersection of SW Taylor and SW Naito Pkwy. These conflicts are starred in Table 6-2.

Table 6-2: Summary of observed PETs (≤ 2sec) (SW Taylor St & SW Naito Pkwy)

No.	Date	Time	Vehicle Type	PET (Sec)	Bicyclist Speed (mph)	Crossing Pedestrian
1	04/30/15	13:42:20	Van	0.8	15.1	None
2	04/30/15	17:49:23	Car	0.9	15.1	None
3	05/01/15	18:09:20	Car	0.8	17	None
4	05/02/15	15:54:11	SUV	0.6	19.5	None
5	04/27/15	17:21:54	SUV	0.8	12.4	None
6	05/02/15	13:14:17	SUV	0.4	8	None
7	05/02/15	16:31:33	Car	0.8	9.7	None
8	05/03/15	12:55:51	Car	0.6	11.4	None
9	05/03/15	14:11:03	Car	0.8	10.5	None
10	05/03/15	18:04:20	SUV	0.5	12.4	None
11	05/05/15	16:29:29	SUV	0.6	8.9	None
12	05/05/15	17:04:34	Car	0.6	12.4	None
13	05/05/15	16:55:33	SUV	0.7	9.75	4 Ped
14	04/23/15	18:16:03	Pick-up	1.4	22.7	None
15	05/01/15	17:25:59	SUV	1.3	22.7	None
16	05/01/15	18:10:29	Car	1	15.1	None
17	05/02/15	17:18:23	Car	1.3	15.1	None

No.	Date	Time	Vehicle Type	PET (Sec)	Bicyclist Speed (mph)	Crossing Pedestrian
18	05/03/15	13:36:07	Car	1.4	17	None
19	05/04/15	17:11:37	SUV	1.2	15.1	None
20	05/05/15	17:10:24	SUV	1	17	None
21	04/23/15	11:03:05	Car	1.2	13.3	None
22	04/23/15	17:59:56	Van	1.3	13.6	None
23	04/27/15	15:38:07	Car	1	13.6	None
24	04/27/15	18:54:48	Truck	1.3	10.5	None
25	05/02/15	15:23:32	Van	1.2	12.4	None
26	05/03/15	18:02:44	SUV	1	12.4	None
27	05/03/15	11:18:41	SUV	1.2	9	None
28	05/05/15	14:53:24	SUV	1.4	9.7	None
29	05/05/15	17:28:15	SUV	1.3	8.5	None
30	05/03/15	11:41:17	Van	1.3	8.5	2 Ped
31	05/05/15	16:34:00	SUV	1.8	10.5	None
32	04/22/15	17:07:40	Car	1.7	17	None
33	04/26/15	10:58:38	SUV	1.7	17	1 Ped
34	04/27/15	14:13:14	Pick-up	1.9	14.3	2 Ped
35	04/27/15	17:17:50	SUV	1.9	17	None
36	04/28/15	17:20:27	Car	1.6	15.1	None
37	04/30/15	12:55:10	Pick-up	1.7	15.2	None
38	04/30/15	15:47:56	SUV	1.6	15.1	None
39	05/01/15	15:09:17	Car	1.6	17	None
40	04/24/15	11:53:38	Car	1.7	12.4	None
41	04/27/15	17:31:14	Van	1.8	5.5	None
42	04/27/15	08:51:29	Car	1.5	13.6	None
43	04/30/15	14:31:12	SUV	1.7	12.4	None
44	04/30/15	17:22:37	Car	1.8	13.6	None
45	05/01/15	17:12:29	SUV	1.6	13.6	None
46	05/05/15	14:49:12	Pick-up	1.6	13.6	None
47	04/22/15	16:57:01	Car	1.1	15.1	None
48	05/03/15	14:05:13	Car	1.7	10.5	1 Ped
49*	05/04/15	16:11:42	SUV	0.3	15.2	None
50*	04/22/15	18:29:30	SUV	0.6	13.6	None
51*	04/22/15	17:52:24	SUV	0.5	19.5	None
52*	04/22/15	14:41:40	SUV	0.4	12.4	None
	Average	1.2	13.6			
	Standard Deviation	0.46	3.5			
	Maximum	1.9	22.7			
	Minimum	0.2	5.5			

6.2 Conflict Scenarios

PETs were measured for eight scenarios designed in the OSU driving simulator by Hurwitz et al. (2015). These scenarios consisted of the combination of three factors (1) "the presence of oncoming left-turning vehicular traffic" (2) "the presence of a conflicting pedestrian in the crosswalk" and (3) "bicyclist speed" (Hurwitz, et al. 2015). Table 6-3 and Table 6-4 show the different combinations of these variables in each given scenario and the frequency of PETs corresponding to each scenario in the OSU driving simulator and the intersection of SW Taylor St and SW Naito Pkwy. Bicyclist speed in the OSU driving simulator was categorized as low-speed (12 mph) and high-speed (16 mph). Average bicyclist speed, 13.6 mph was chosen to separate high-speed bicyclists (V_{bike} ≥13.6 mph) from low-speed bicyclists (V_{bike}<13.6 mph) in the intersection of SW Taylor and SW Naito Pkwy.

Table 6-3: Frequency of PETs in the simulator

Simulator	Bicyclist	Pedestrian	Oncoming Traffic	Bicyclist Speed	Frequency of PETs < 2sec	Total
Scenario 1	x			High (16mph)	16	26
Scenario 2	x			Low (12mph)	10	
Scenario 3	x	x		High (16mph)	0	
Scenario 4	x	x		Low (12mph)	0	
Scenario 5	x	x	x	High (16mph)	4	24
Scenario 6	x	x	x	Low (12mph)	1	
Scenario 7	x		x	High (16mph)	14	
Scenario 8	x		x	Low (12mph)	5	
Total						50

Table 6-4: Frequency of PETs in the intersection of SW Taylor and SW Naito

Field	Bicyclist	Pedestrian	Oncoming Traffic	Bicyclist Speed	Frequency of PETs < 2sec	Total
Scenario 1	x			High (≥ 13.5 mph)	27	52
Scenario 2	x			Low (< 13.5 mph)	20	
Scenario 3	x	x		High (≥ 13.5 mph)	2	
Scenario 4	x	x		Low (< 13.5 mph)	3	
Scenario 5	x	x	x	High (≥ 13.5 mph)	0	0
Scenario 6	x	x	x	Low (< 13.5 mph)	0	
Scenario 7	x		x	High (≥ 13.5 mph)	0	
Scenario 8	x		x	Low (< 13.5 mph)	0	
Total						52

Oncoming left-turning vehicles were involved in right-hook conflicts in scenarios four through eight, but there was a protected left-turn for oncoming vehicles at the intersection of SW Taylor and SW Naito Pkwy so such conflicts were not observed in the field (Table 6-4). Therefore, the validation of the OSU driving simulator was inconclusive for these four scenarios, and scenarios one through four were left for comparison. The frequency of each group of PETs from scenarios one through four in the OSU driving simulator and the intersection of SW Taylor and SW Naito Pkwy is compared in Table 6-5 and Table 6-6.

Table 6-5: Frequency of each group of PETs in the OSU driving simulator

Simulator	Bicyclist	Bicyclist speed (mph)	Crossing Pedestrian	Frequency of PETs (sec)			
				High Risk [0-1)	Moderate Risk [1.0-1.5)	Low Risk [1.5-2.0)	Total
Scenario 1	x	High (16mph)		2	5	9	16
Scenario 2	x	Low (12mph)		2	4	4	10
Scenario 3	x	High (16mph)	x	0	0	0	0
Scenario 4	x	Low (12mph)	x	0	0	0	0
Total				4	9	13	26

Table 6-6: Frequency of each group of PETs in the field

Field	Bicyclist	Bicyclist speed (mph)	Crossing Pedestrian	Frequency of PETs (sec)			
				High Risk [0-1)	Moderate Risk [1.0-1.5)	Low Risk [1.5-2.0)	Total
Scenario 1	x	High (≥ 13.5)		7	10	10	27
Scenario 2	x	Low (< 13.5)		9	7	4	20
Scenario 3	x	High (≥13.5)	x	0	0	2	2
Scenario 4	x	Low (<13.5)	x	1	1	1	3
Total				15	18	17	52

For scenarios three and four, pedestrians were involved in the conflict. No conflict was observed for scenario three and four in the OSU driving simulator (Table 6-5), while two conflicts for scenario three, and three conflicts for scenario four were observed in the intersection of SW Taylor St and SW Naito Pkwy (Table 6-6). The validation of the OSU driving simulator was inconclusive for these two scenarios, and thus only scenarios one and two are used for comparison in the rest of this analysis.

Scenarios one and two were analyzed through two validation approaches between the OSU driving simulator and the intersection of SW Taylor St & SW Naito Pkwy in the field.

6.3 Graphical Validation Approach

The comparison of the frequency of PETs between the OSU driving simulator and the field are displayed below. The cumulative percent of the frequency of each group of PETs (line) along with their frequencies (bars) are given in Figure 6-1. The percentage of frequencies in each group of observations within each respective bar is given in Figure 6-2. These plots help visualize differences between PETs observed in the OSU driving simulator and those observed in the field. The following figures depict a hypothetical example of a perfect match between the frequency of each group of PETs observed in the OSU driving simulator and those in the field.

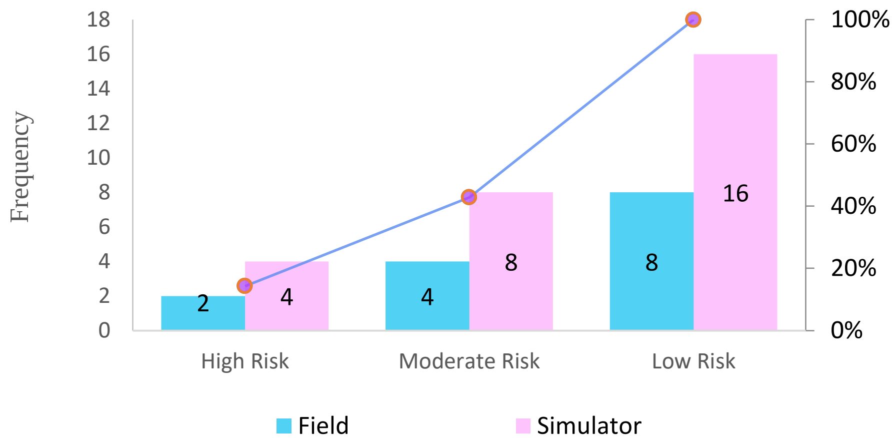


Figure 6-1: The cumulative percent of the frequency of each group of PETs (line) along with their frequencies (bars). A hypothetical illustration of perfect match b/t simulator and field PET frequencies.

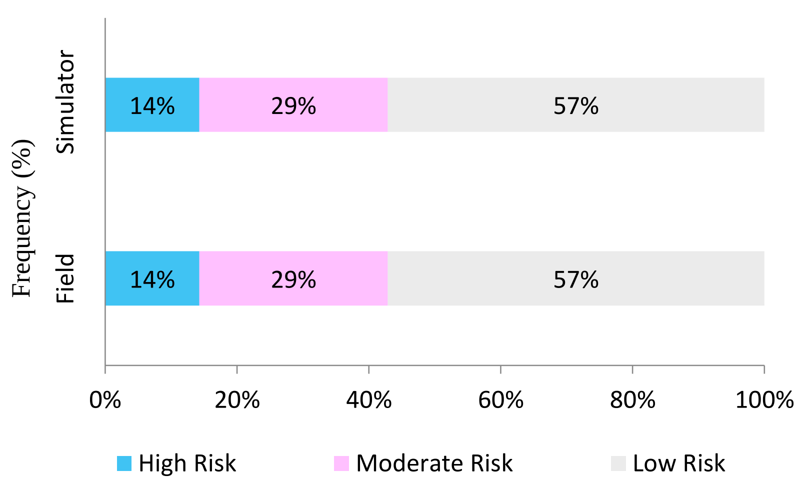


Figure 6-2: The percentage of frequencies in each group of observations within each respective bar. A hypothetical illustration of perfect match between simulator and field PET frequencies.

6.4 Statistical Validation Approach

Fisher's Exact Test was computed to determine whether the distribution of the frequency of PETs differs significantly between the OSU driving simulator and the field (Fisher's Exact Test for Count Data n.d.).

The null hypothesis of Fisher's Exact Test states that the proportion of the frequency of PETs for each group in the OSU driving simulator is identical to the proportion of the frequency of PETs for each corresponding group in the field. Thus,

$$H_0: P_{High\ risk\ of\ collision\ in\ the\ OSU\ driving\ simulator} = P_{High\ risk\ of\ collision\ in\ the\ field}$$

$$H_0: P_{Moderate\ risk\ of\ collision\ in\ the\ OSU\ driving\ simulator} = P_{Moderate\ risk\ of\ collision\ in\ the\ field}$$

$$H_0: P_{Low\ risk\ of\ colision\ in\ the\ OSU\ driving\ simulator} = P_{Low\ risk\ of\ collision\ in\ the\ field},$$

Where $\mathrm{P}_{\text{High risk of collision in the OSU driving simulator}}$ is the relative frequency of PETs with high risk of collision in the OSU driving simulator.

The alternative hypothesis states that at least one of the null hypotheses is false.

The Exact Multinomial Test of goodness of fit was computed to test how well the frequency of each group of PETs observed in the OSU driving simulator fits the distribution observed in the field (Engles n.d.). This distribution is our model and consists of the proportion of each group of PETs observed in the field for each scenario. The likelihood-ratio test statistic was applied to the hypothesis to determine the associated p-value (Engles n.d.). The p-value of the test indicates the probability of observing the frequency of each group of PETs obtained in the OSU driving simulator given our model, i.e. Model 1 for scenario one and Model 2 for scenario two. Because the sample size of frequencies observed in the OSU driving simulator remains unchanged, p-values are comparable for the given models (ReliaSoft Corporation 2007).

Because only type I conflicts were measured in the OSU driving simulator, the frequencies of PETs observed in the OSU driving simulator are only comparable to models with only type I conflicts (Table 6-7). Therefore, four type II conflicts were excluded from the frequency of PETs comprising Model 1 and Model 2. The null hypothesis of the Exact Multinomial Test is specified below.

Table 6-7: Model 1 (Scenario 1) and Model 2 (Scenario 2)

SW Taylor & SW Naito	Unit	High Risk (0-1)	Medium Risk [1-1.5)	Low Risk [1.5-2)	Total
Scenario 1	Frequency	4	10	10	24
Model 1	**Probability**	**4/24**	**10/24**	**10/24**	**1**
Scenario 2	Frequency	8	7	4	19
Model 2	**Probability**	**8/19**	**7/19**	**4/19**	**1**

Model 1 for Scenario 1:

$$H_0: P_{High\ risk\ of\ collision\ in\ the\ OSU\ driving\ simulator} = 4/24$$

$$H_0: P_{Moderate\ risk\ of\ collision\ in\ the\ OSU\ driving\ simulator} = 10/24$$

$$H_0: P_{Low\ risk\ of\ colision\ in\ the\ OSU\ driving\ simulator} = 10/24$$

Model 2 for Scenario 2:

$$H_0: P_{High\ risk\ of\ collision\ in\ the\ OSU\ driving\ simulator} = 8/19$$

$$H_0: P_{Moderate\ risk\ of\ collision\ in\ the\ OSU\ driving\ simulator} = 7/19$$

$$H_0: P_{Low\ risk\ of\ colision\ in\ the\ OSU\ driving\ simulator} = 4/19$$

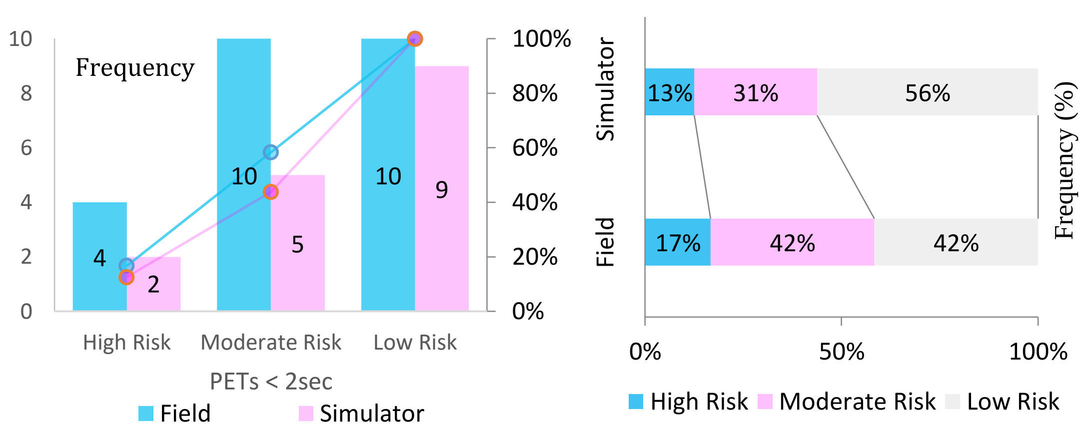
Where $\mathrm{P}_{High\ risk\ of\ collision\ in\ theOSU\ driving\ simulator}$ is the relative frequency of PETs with high risk of collision observed in the OSU driving simulator.

The alternative hypothesis states that at least one of the null hypotheses is false.

6.5 Model 1 & Model 2 (Type I Conflicts)

The result of Fisher's Exact Test indicates that the frequency of PETs observed in the field did not significantly differ from those observed in the OSU driving simulator for both scenarios one and two. The p-value for scenario one was 0.68, and the p-value for scenario two was 0.43. Thus, there is not sufficient evidence to reject the null hypothesis of equal frequencies of PETs in the OSU driving simulator and the field. The Exact Multinomial Test results in 58% and 27% probabilities that frequencies of PETs observed in the OSU driving simulator fit Model 1 for scenario one and Model 2 for scenario two, respectively (Table 6-8). P-values equal to 0.58 and 0.27, for scenario one and two respectively, give the probability of observing the frequencies in the OSU driving simulator given the model. Plots of the graphical validation approach are shown in Figure 6-3.

- Model 1 for PETs with Type I conflicts (Scenario 1)



- Model 2 for PETs with Type I conflicts (Scenario 2)

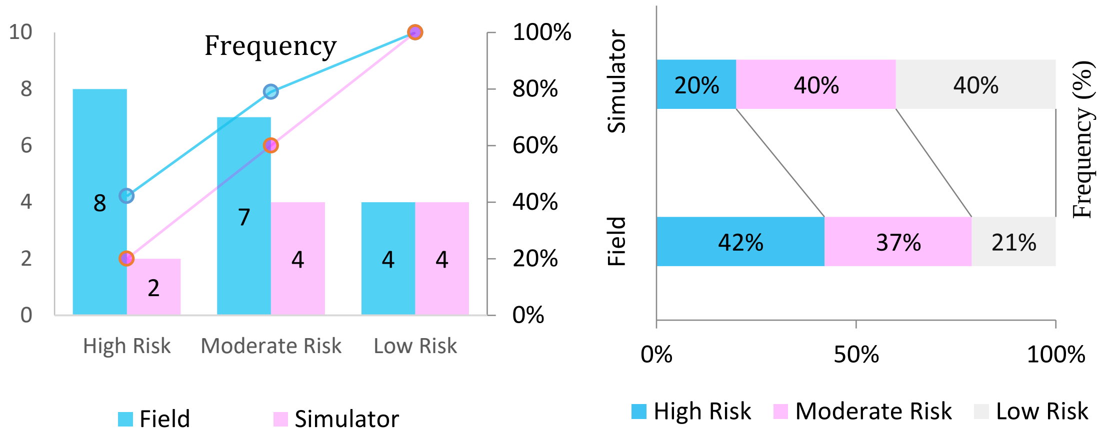


Figure 6-3: The frequency of each group of PETs (bars) along with their cumulative percent (line) for scenario one (Left), the percentage of frequencies in each group of observations within each respective bar (Right).

Table 6-8: The results of Fisher's Exact and Exact Multinomial Tests

SW Taylor & SW Naito	PETs with Type I conflicts	
	Exact Multinomial Test	Fisher's Exact Test
Model 1	**0.58**	**0.68**
Model 2	**0.27**	**0.43**

Figure 6-4 shows the summary of results of the Exact Multinomial Test.

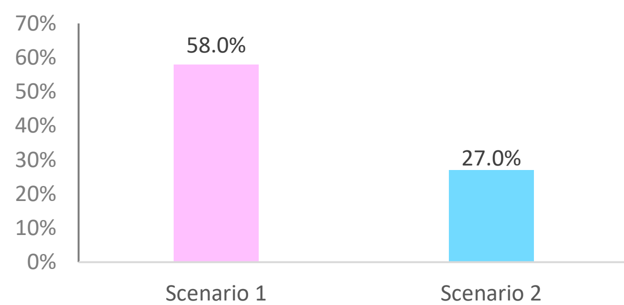


Figure 6-4: Probabilities that frequencies of PETs observed in the OSU driving simulator fit model 1 for scenario 1 and model 2 for scenario 2.

Now, the important question is whether relative validity can be claimed for the OSU driving simulator by these tests of goodness of fit. To answer this question, two considerations are noted. (1) We are comparing frequencies of PETs observed in the OSU driving simulator to those observed in the field to determine if driving behaviors are similar in both environments. (2) There were other differences aside from driving behaviors between the OSU driving simulator environment and field environment. These differences included method of measuring PETs, bicyclist speed, geometric designs, traffic conditions, traffic signals and signal timing. To validate the OSU driving simulator for driving behaviors, these other differences need to be accounted for. The method of measuring PETs is controlled for in Model 1 and Model 2. These other differences cannot be controlled for, but their effect on the goodness of fit may be estimated. To estimate the effect of these other conditions, various models are computed in which these conditions differ and the difference in goodness of fit is analyzed. These PET frequencies include (1) PETs of both type I and type II conflicts (2) PETs of type I conflicts during PM-Peak hours (3) PETs of type I conflicts with an alternate bicyclist speed threshold. If these different models can explain the change in the probability of goodness of fit from prior probabilities, then the magnitude and direction of a change in the probability may be attributed to the effect of a differing condition between the OSU driving simulator environment and the intersection of SW Taylor and SW Naito Pkwy, and thereby relative validity for the OSU driving simulator may be established.

6.6 Model 3 & Model 4 (Type I & Type II Conflicts)

We expect to see a decrease in the probability of goodness of fit if four type II conflicts are added into Model 1 and Model 2 as the frequency of PETs were measured only for type I conflicts in the OSU driving simulator. The frequency of each group of PETs for scenario one and two and their corresponding Models 3 and 4 are shown in Table 6-9.

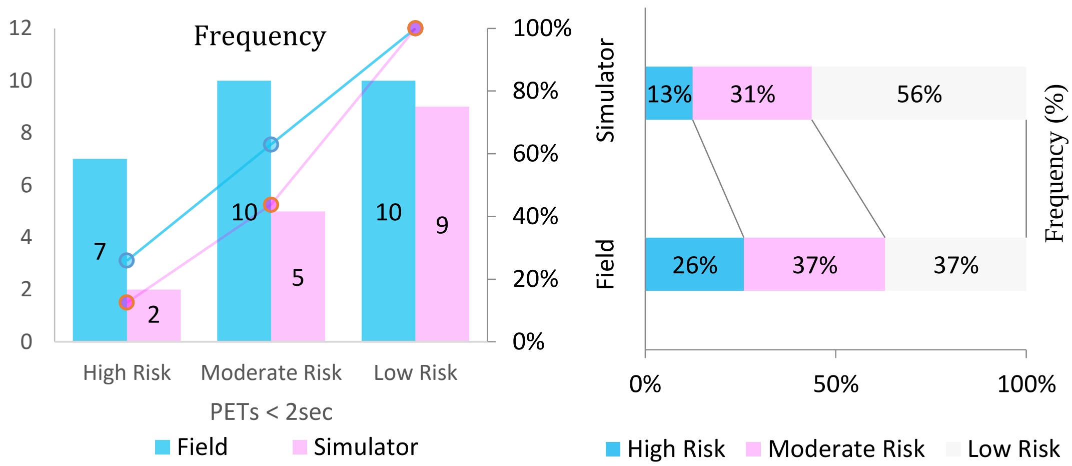
Table 6-9: Frequency of each group of PETs at the intersection of SW Taylor & SW Naito

SW Taylor & SW Naito	The frequency of PET < 2sec		
	High Risk [0-1)	Medium Risk [1-1.5)	Low Risk [1.5-2)
Scenario 1	4+3*=7	10	10
Model 3	P1=7/27	P2=10/27	P3=10/27
Scenario 2	8+1*=9	7	4
Model 4	P1=9/20	P2=7/20	P3=4/20

*Three type II conflicts in scenario 1 and one type II conflict in scenario 2

The results of Fisher's Exact test and the Exact Multinomial test of goodness of fit are shown in Table 6-10 and plots of the graphical validation approach are shown in Figure 6-5.

- Model 3 for PETs with Type I and Type II conflicts (Scenario 1)



- Model 4 for PETs with Type I and Type II conflicts (Scenario 2)

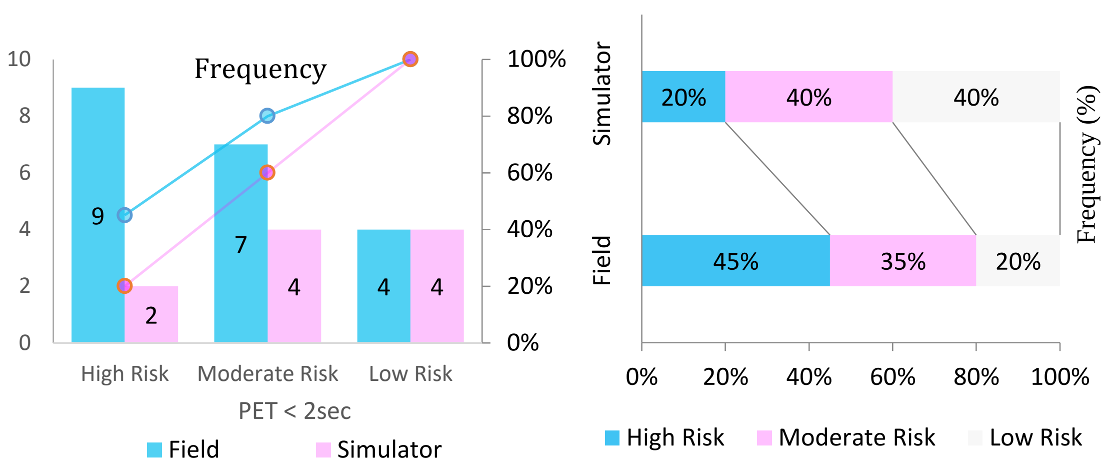


Figure 6-5: The frequency of each group of PETs (bars) along with their cumulative percent (line) (Left), the percentage of frequencies in each group of observations within each respective bar (Right).

Table 6-10: The results of Fisher's Exact and Exact Multinomial Tests

SW Taylor & SW Naito	PETs with Type I and Type II conflicts	
	Exact Multinomial Test	Fisher's Exact Test
Model 3	**0.25**	**0.43**
Model 4	**0.26**	**0.39**

6.7 Model 5 & Model 6 (Type I Conflicts during PM-Peak Hours)

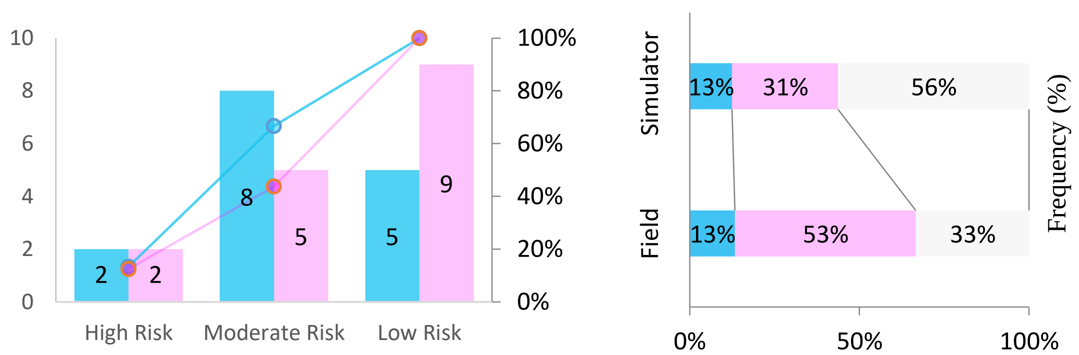
We expect to see a decrease in probability of goodness of fit when PETs corresponding to non-PM-peak hours are excluded from Model 1 and Model 2 because traffic conditions in the OSU driving simulator represent a suburban area, but the intersection of SW Taylor St and SW Naito Pkwy is located in an urban area. Therefore, traffic conditions during PM-peak hours, as opposed to the entire day, should make the difference between the two environments more marked. The frequency of each group of PETs observed in the field for scenarios one and two during PM-peak hours and their corresponding models, Model 5 and Model 6, are shown in Table 6-11.

Table 6-11: Frequency of each group of PETs at the intersection of SW Taylor & SW Naito

SW Taylor & SW Naito	The frequency of PET < 2sec		
	High Risk [0-1)	Medium Risk [1-1.5)	Low Risk [1.5-2)
Scenario 1	2	8	5
Model 5	$P1 = 2/15$	$P2 = 8/15$	$P3 = 5/15$
Scenario 2	5	3	2
Model 6	$P1 = 5/10$	$P2 = 3/10$	$P3 = 2/10$

The results of Fisher's Exact test and the Exact Multinomial test of goodness of fit are shown in Table 6-12 and plots of the graphical validation approach are in Figure 6-6.

- Model 5 for PETs with Type I conflict during PM-peak hours (Scenario 1)



- Model 6 for PETs with Type I during PM-peak hours (Scenario 2)

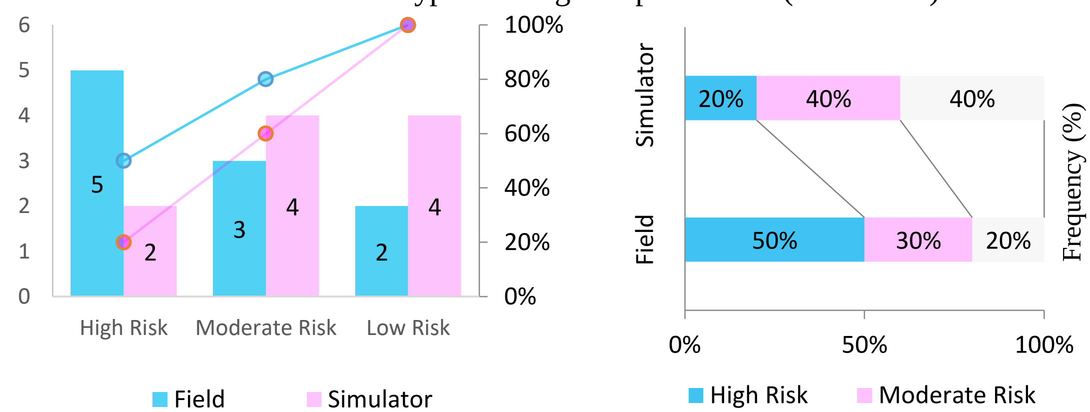


Figure 6-6: The frequency of each group of PETs (bars) along with their cumulative percent (line) (Left), the percentage of frequencies in each group of observations within each respective bar (Right).

Table 6-12: The results of Fisher's Exact and Exact Multinomial Tests

SW Taylor & SW Naito	PETs with Type I conflicts during PM-Peak Hours	
	Exact Multinomial Test	Fisher's Exact Test
Model 5	**0.20**	**0.42**
Model 6	**0.19**	**0.54**

6.8 Model 7 & Model 8 (Type I Conflicts with 14 mph Bicyclist Speed Threshold)

We estimate the effect of bicyclist speed on the goodness of fit by changing the threshold from average bicyclist speed in the field (13.6mph) to average bicyclist speed in the OSU driving simulator (14 mph). Although we cannot explain the direction and magnitude of changes in the probabilities from Model 1 and Model 2, we can attribute their changes to bicyclist speed only because all frequencies of PETs included in Model 1 and Model 2 remain in the new models. PETs with bicyclist speed between 13.6 mph and 14 mph are simply transferred between Model 1 (high-speed bicyclists in scenario one) and Model 2 (low-speed bicyclists in scenario two) based on the new speed threshold. In other words, by changing the bicyclist speed threshold, there are no PETs added or eliminated from the total number of PETs comprising Model 1 and Model 2. The change in frequency of each group of PETs between scenario one and scenario two when the threshold is changed from 13.6 mph to 14 mph can be seen in Table 6-13 and Table 6-14. Models 7 and 8 correspond to the new threshold of 14 mph (Table 6-15).

Table 6-13: Frequency of PETs for type I conflicts when bicyclist speed threshold is 13.6 mph

SW Taylor & SW Naito	The frequency of PET < 2sec			Sum
	High Risk [0-1)	Medium Risk [1-1.5)	Low Risk [1.5-2)	
Scenario 1	4	10	10	24
Scenario 2	8	7	4	19
Total	12	17	14	43

Table 6-14: Frequency of PETs for type I conflicts when bicyclist speed threshold is 14 mph

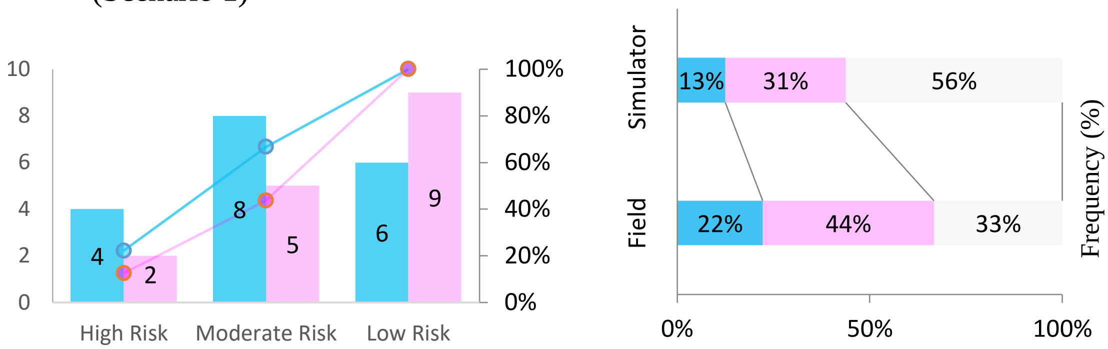
SW Taylor & SW Naito	The frequency of PET < 2sec			Sum
	High Risk [0-1)	Medium Risk [1-1.5)	Low Risk [1.5-2)	
Scenario 1	4	8	6	24
Scenario 2	8	9	8	19
Total	12	17	14	43

Table 6-15: Model 7 & Model 8 (Type I conflicts when bicyclist speed threshold is 14mph)

SW Taylor & SW Naito	High Risk [0-1)	Medium Risk [1-1.5)	Low Risk [1.5-2)
Model 7	P1 = 4/18	P2 = 8/18	P3 = 6/18
Model 8	P1 = 8/25	P2 = 9/25	P3 = 8/25

Results of Fisher's Exact test and the Exact Multinomial test of goodness of fit are shown in Table 6-16 and plots of the graphical validation approach are shown in Figure 6-7.

- Model 7 for PETs with Type I conflicts and 14 mph bicyclist speed threshold (Scenario 1)



- Model 8 for PETs with Type I conflicts and 14 mph bicyclist speed threshold (Scenario 2)

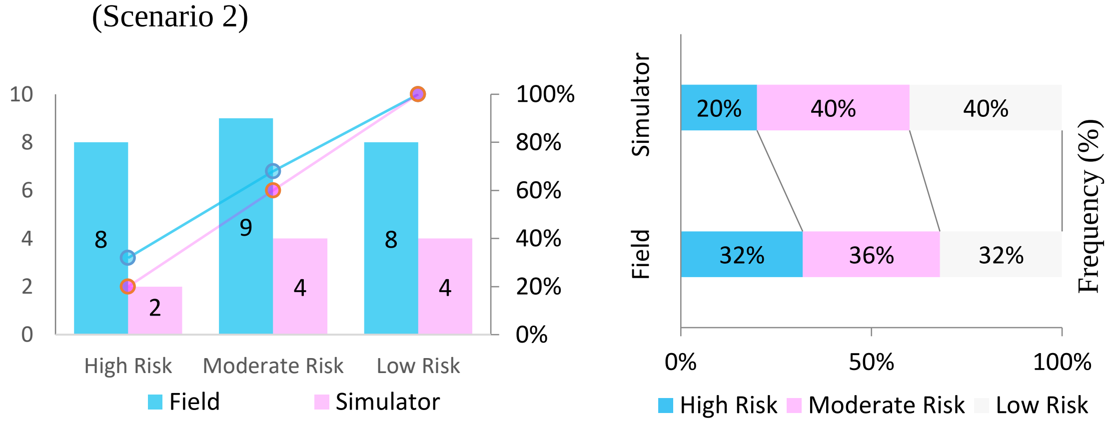


Figure 6-7: The frequency of each group of PETs (bars) along with their cumulative percent (line) (Left), the percentage of frequencies in each group of observations within each respective bar (Right).

Table 6-16: The results of Fisher's Exact and Exact Multinomial Tests

SW Taylor & SW Naito	PETs with type I conflicts when bicyclist speed threshold is 14mph	
	Exact Multinomial Test	Fisher's Exact Test
Model 7	**0.22**	**0.48**
Model 8	**0.79**	**0.8**

6.9 Results

Probability of goodness of fit for each model and the change in the probability of goodness of fit from Model 1 and Model 2 are shown in Figure 6-8 and Figure 6-9.

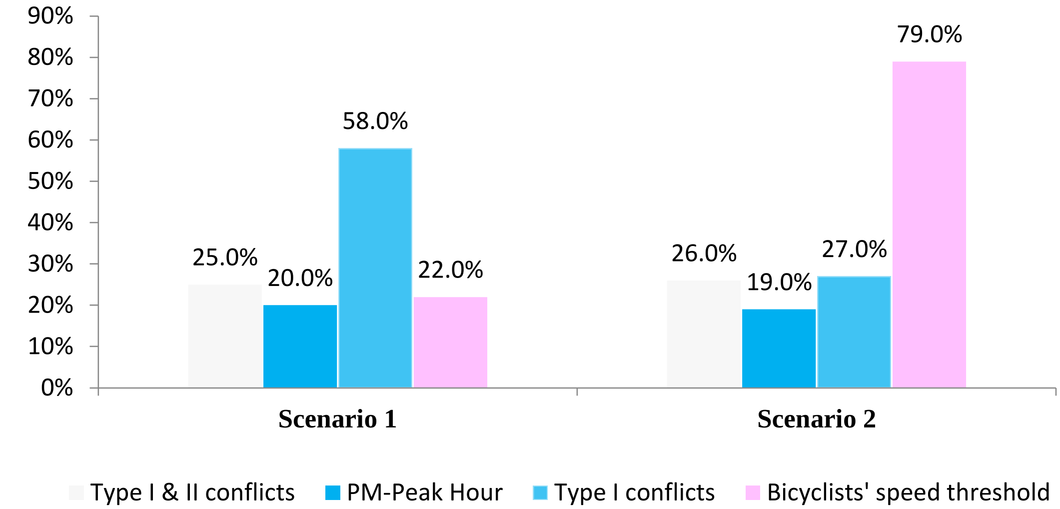


Figure 6-8: Probability of goodness of fit for each model.

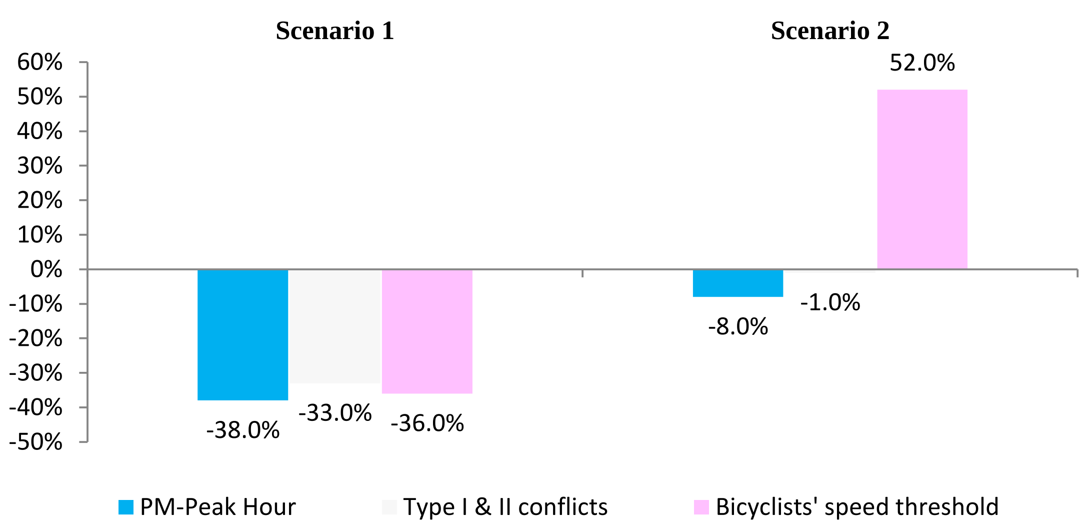


Figure 6-9: Change in the probability of goodness of fit from Model 1 and Model 2.

Results of Fisher's Exact test indicated that the distribution of frequency of PETs did not significantly differ at the 5% significance level between the OSU driving simulator and the intersection of SW Taylor and SW Naito ($p > 0.05$). The effects of traffic condition, method of measuring PETs, and bicyclist speed on the probability of goodness of fit were examined. The effect of traffic condition after removing non-PM-peak hour's frequencies of PETs from Model 1 for scenario one was negative and significant. The probability of goodness of fit in this model changed (decreased) by 38%, this decrease in probability was about 8% in Model 2 for scenario two. This negative effect was expected because the differences in the traffic condition between the two environments were greater. The effect of measuring PETs with both type I and type II conflicts was also negative in both scenarios. The probability of goodness of fit in Model 1 decreased significantly by 33%, but this decrease was only 1% in Model 2 for scenario two. This negative effect was also to be expected as PETs for type II conflicts were not measured in the OSU driving simulator. The effect of bicyclist speed after changing the threshold of bicyclist speed from 13.6 mph to 14 mph was significantly negative on Model 1 for scenario one and significantly positive on Model 2 for scenario two. The probability of goodness of fit in Model 1 decreased by 36% and increased in Model 2 by 52%. Although the change in the probability could not be explained, this change could be attributed to the change in the bicyclist speed threshold only because all frequencies of PETs in Model 1 and Model 2 remained in the new models. The magnitude of changes in the probability of goodness of fit indicated the significant effect of bicyclist speed threshold on the OSU driving simulator validation.

6.10 Conclusion

To answer the question of whether relative validity can be claimed for the OSU driving simulator with 58% and 27% probability of goodness of fit for scenarios one and two, we note that this analysis showed that differences between the OSU driving simulator and the field have significant effect on the probability of goodness of fit. In other words, if these differences between the OSU driving simulator environment and the intersection of SW Taylor and SW Naito Pkwy could be controlled for, we would expect that the models would improve significantly. Fisher's Exact test results indicated that distribution of frequencies of PETs did not significantly differ between those observed in the OSU driving simulator and those comprising our eight models observed in the intersection of SW Taylor and SW Naito Pkwy. Overall, although it is not possible to claim a definite behavioral validity for the OSU driving simulator, the results showed that the validation of the OSU driving simulator for right-hook conflicts is promising.

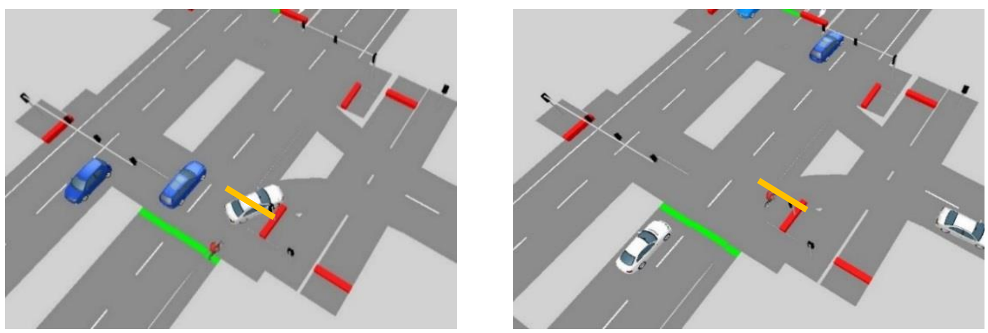
The following chapter will investigate the probability that the frequencies of PETs observed in traffic simulation models in VISSIM fit frequencies of PETs in the intersection of SW Taylor and SW Naito Pkwy after replicating the geometric design of the intersection of SW Taylor and SW Naito Pkwy, the traffic conditions, and signal timing in the simulation models.

7.0 TRAFFIC SIMULATION MODELS (VISSIM)

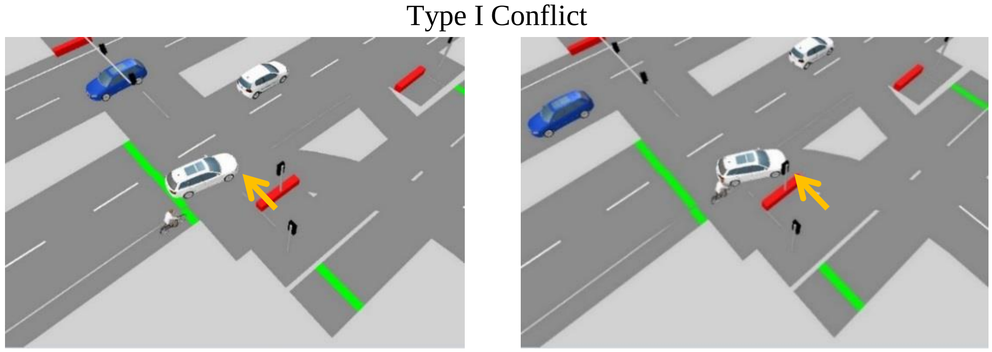
Although traffic simulation models are widely used to study traffic operations in a network, many professionals have recently investigated the potential of traffic simulation models for safety assessment. Accessibility, the ability to assess traffic operations in a network, and the ease of replicating real road designs are the main advantages of using traffic simulation models. The simulation road users instead of analyzing actual road users is the main weak point of using simulation models in evaluating driving behaviors and therefore safety assessment. Nevertheless, traffic simulation models have been developed to account for driving behaviors by incorporating driving behavior parameters. This chapter investigates how similarly traffic simulation models in VISSIM generate frequencies for each group of PETs to those observed in the field. The intersection of SW Taylor St and SW Naito Pkwy was replicated in traffic simulation models in VISSIM. In other words, all the differences between the OSU driving simulator and the intersection of SW Taylor and SW Naito Pkwy, including variable bicyclist speed, traffic volumes, traffic signals, and geometric designs are addressed in the traffic simulation models in VISSIM. The goal is to see if addressing these issues at the expense of observing actual driver behavior can reasonably represent driver behavior in the field important to safety assessment.

8.0 DATA ASSEMBLY

SSAM software was used to collect PETs from the output of simulation models in VISSIM. However, the SSAM could not identify most right-hook conflicts (crossing and lane change conflicts defined in SSAM) from simulation outputs. Therefore, all conflicts were identified manually by recording, replaying, and analyzing simulation runs. Both type I and type II conflicts were identified, and their PETs were measured in the simulation models. Figure 8-1 illustrates measuring a type I conflict (top) and type II conflict (bottom) in the VISSIM simulation models.



Type I Conflict



Type II Conflict

Figure 8-1: Measuring the post-encroachment time in the simulation model (VISSIM).

Figure 8-2 shows that PET should not be measured when a bicyclist waits behind the conflict area for the vehicle to cross the bike lane because the road users are not on a collision course.

Figure 8-2: A bicyclist is waiting for a vehicle to cross the bike lane.

8.1 Calibration

Priority rules were modeled and driving behavior parameters were defined to calibrate simulation models (PTV AG 2014). The Urban (motorized) parameter set was selected and Car Following model was set to Wiedemann 74, which is suitable for urban traffic. In order to simulate more aggressive driving behavior, minimum gap time was reduced to 0.5 second and 1 second for car and bicyclist, respectively. However, minimum look ahead distance was increased to 65.6 ft. to reduce run-over scenarios. Figure 8-3 shows priority rules, including conflict markers and stop lines, in VISSIM. Table 8-1 shows the main driving behavior parameter sets in the simulation model.

Table 8-1: Driving behavior parameter sets in VISSIM

Car following model	Wiedemann 74
Look ahead distance	[65.6, 820] ft
Minimum lateral distance	3.28 ft
Accepted deceleration	-3.28 ft/sec^2
Safety distance reduction factor	0.6
Average standstill distance	3.28 ft.
Minimum longitudinal speed	2.24 mph
Temporary lack of attention	0 sec

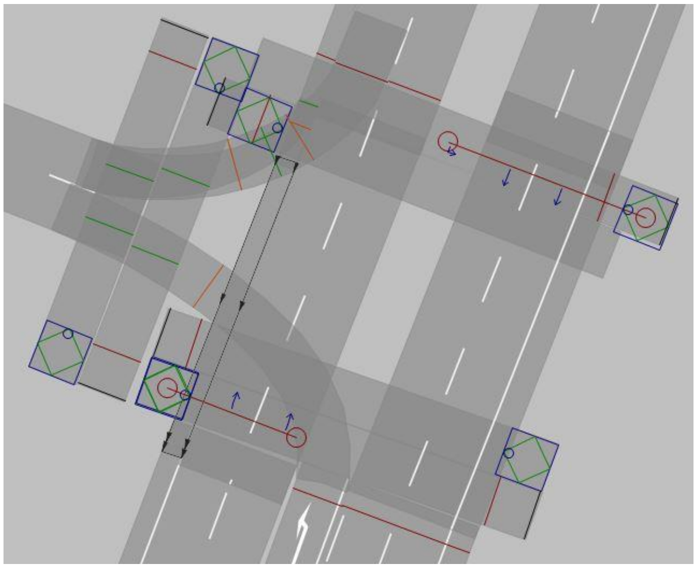


Figure 8-3: Priority rules include Conflict marker (Green lines), and Stop line (Orange lines).

After calibration, three criteria were calculated to determine if the calibration was appropriate:

1. A Universal Measure: the GEH Statistic

GEH was calculated to compare the observed volumes in the field with the volumes in the simulation output. A GEH value of less than 5 indicates an appropriate calibration. As is shown in Table 8-2, all GEH values were less than 3 (Oregon Department of Transportation 2011).

Table 8-2: Volume and GEH values

Observed volumes in the field			
Right Turn Veh	Through Veh	Through bike	Total
85	1337	40	1462
79	1196	58	1333
100	1188	88	1376
91	1208	91	1390
100	1080	78	1258
105	1131	50	1286
Volumes in the simulation outputs			
Right Turn Veh	Through Veh	Through bike	Total
72	1399	44	1515
88	1183	50	1321
108	1185	94	1387
88	1184	121	1393
91	1082	84	1257
108	1130	46	1284
GEH Values			
Right Turn Veh	Through Veh	Through bike	Total
1.46	1.67	0.61	1.37
0.98	0.37	1.08	0.32
0.78	0.08	0.62	0.29
0.31	0.69	2.91	0.08
0.92	0.06	0.66	0.02
0.29	0.02	0.57	0.05

2. The Mean Absolute Percent Error (MAPE)

The MAPE showed a 22% differences between the frequency of simulated conflicts and observed conflicts. This percentage indicates an acceptable calibration of the simulation model. Dates, the frequency of observed conflicts, the average frequency of simulated conflicts, and MAPE value calculations are shown in Table 8-3.

Table 8-3: The computation of the Mean Absolute Percent Error

The frequency of PETs < 2 sec					
No.	Time	Date	Obs.(C_f)	Sim. (C_m)	$(C_m - C_f)/C_f$
1	4-5 PM	4/22/2015	1	1	0
2	5-6 PM	4/23/2015	1	1	0
3	5-6 PM	4/27/2015	3	3	0
4	5-6 PM	4/30/2015	2	3	0.5
5	5-6 PM	5/1/2015	2	3	0.5
6	4-5 PM	5/5/2015	3	2	0.333
			MAPE		0.222

3. The Spearman Rank Correlation Coefficient

The Spearman rank correlation coefficient showed a moderate correlation, 68.6% between the frequency of PETs measured in the field and the average frequency of PETs measured in the simulation model. It indicates an acceptable calibration of the simulation model. The rank of observed conflicts, simulated conflicts, and the Spearman rank coefficient are shown in Table 8-4.

Table 8-4: The Spearman rank correlation coefficient

PET < 2 sec		Spearman's rank correlation coefficient (ρ_s)			
No.	Time	Date	Rank Obs.	Rank Sim.	Diff^2
1	4-5 PM	4/22/2015	1.5	1.5	0
2	5-6 PM	4/23/2015	1.5	1.5	0
3	5-6 PM	4/30/2015	3.5	5	2.25
4	5-6 PM	5/1/2015	3.5	5	2.25
5	4-5 PM	5/5/2015	5.5	3	6.25
6	5-6 PM	4/27/2015	5.5	5	0.25
			ρ_s		0.6857

Although the calibration satisfies all three criteria for appropriateness, it should be noted that it is possible to reach a better calibration by further adjusting driving behavior parameters, particularly in regard to bicyclists because most of the default parameters were unchanged in this research.

As the calibration was found to be satisfactory, the frequency of PETs derived from the simulation models may be analyzed to determine how similarly the simulation models generate frequencies of PETs compared to those observed in the field. To do so, Fisher’s Exact Test and the Exact Multinomial Test are employed. Plots of the cumulative percent of each group of the frequency of PETs, as well as the proportion of the frequency of PETs for each group, are displayed in the following sections.

9.0 DATA REDUCTION

The geometric design, phasing and signal timing plan at the intersection of SW Taylor St & SW Naito Pkwy were replicated in the simulation models. Traffic volumes of vehicles, pedestrians, and bicyclists were extracted from video cameras over the course of six days, with data from one PM-peak hour corresponding to each day, and imported into the traffic simulation models (Table 9-1). It may be noted that the VISSIM traffic simulation models analyzes PM-peak hour data only, our primary interest in right-hook conflicts.

Table 9-1: Dates and traffic volumes of each PM-peak hour in the field

SW Taylor & SW Naito			Pedestrian Vol			Vehicle Vol		Bicycle Vol
Dates of video records	Days of video records	Time of video records (PM)	Up-Stream	crossing	Down-Stream	Right Turn	Through	Through
4/22/2015	Wed	4 – 5	29	50	14	85	1337	40
4/23/2015	Thu	5 – 6	22	40	28	79	1196	58
4/27/2015	Mon	5 – 6	46	52	26	100	1188	88
4/30/2015	Thu	5 – 6	18	70	5	91	1208	91
5/1/2015	Fri	5 – 6	40	80	10	100	1080	78
5/5/2015	Tue	4 – 5	86	66	60	105	1131	50
Average			40	60	24	93	1190	68

A total of 12 PETs were recorded in the intersection of SW Taylor St and SW Naito Pkwy. Minimum and maximum PETs were 0.7 seconds, and 1.9 seconds, respectively (Table 9-2). A total of 53 PETs were recorded within 24 simulation runs, consisting of 4 simulation runs per each PM-peak hour. Minimum PET was 0.1 seconds and maximum PET was 1.8 seconds. Table 9-3 shows PETs measured in 24 simulation runs.

Table 9-2: Dates, Times and PETs measured in the field

SW Taylor & SW Naito			
date	**Day**	**Time**	**PET (sec)**
4/22/2015	Wednesday	4 – 5 PM	1.1
4/23/2015	Thursday	5 – 6 PM	1.3
4/27/2015	Monday	5 – 6 PM	1.85
4/27/2015	Monday	5 – 6 PM	0.9
4/27/2015	Monday	5 – 6 PM	1.9
4/30/2015	Thursday	5 – 6 PM	1.75
4/30/2015	Thursday	5 – 6 PM	0.9
5/1/2015	Friday	5 – 6 PM	1.75
5/1/2015	Friday	5 – 6 PM	1.45
5/5/2015	Tuesday	4 – 5 PM	0.55
5/5/2015	Tuesday	4 – 5 PM	1.6
5/5/2015	Tuesday	4 – 5 PM	0.7

Table 9-3: PETs measured within 24 simulation runs

	Simulation PETs < 2sec			
Dates	**Seed**	**Time 1**	**Time 2**	**PET**
4/22/2015	35	1644.8	1645.3	0.5
	35	3067.6	3068.9	1.3
	38	2874.9	2876.6	1.7
	42	2928.9	2930.3	1.4
4/23/2015	40	2710.7	2712.1	1.4
	40	3524.9	3526.7	1.8
	54	545.8	546.3	0.5
4/27/2015	24	550.5	551.7	1.2
	24	1110.9	1112.3	1.4
	24	1860.7	1860.9	0.2
	24	2272.3	2272.5	0.2
	24	2924.5	2924.6	0.1
	24	2962.7	2963	0.3
	29	343.6	345.1	1.5
	29	1252.3	1253.9	1.6
	29	3774.3	3774.6	0.3
	83	791.5	791.8	0.3
	1	336.3	337.9	1.6
	1	2804	2805.5	1.5
	1	3673.6	3675	1.4

	Simulation PETs < 2sec			
Dates	**Seed**	**Dates**	**Seed**	**Dates**
4/30/2015	50	368.1	368.4	0.3
	50	453.2	453.7	0.5
	50	1056.2	1056.5	0.3
	50	2715.2	2715.3	0.1
	50	2895.1	2896.7	1.6
	50	3369.3	3370.5	1.2
	55	3286.5	3286.8	0.3
	73	1378.8	1379.1	0.3
	73	1238.2	1239.4	1.2
	73	1781.6	1781.8	0.2
	73	3502.4	3503.7	1.3
	73	3860.4	3861.1	0.7
	86	646.9	648.1	1.2
	86	1009.4	1010	0.6
5/1/2015	60	1601.6	1602.4	0.8
	60	2091.6	2093.2	1.6
	60	2399.9	2401.5	1.6
	60	2497.6	2499.2	1.6
	60	2846.5	2846.9	0.4
	65	1538.1	1539.4	1.3
	65	1866.8	1868	1.2
	91	1983.9	1984.4	0.5
	100	462.3	463.5	1.2
	100	1221.5	1221.9	0.4
	100	1310.7	1312	1.3
	100	1515.9	1516.1	0.2
5/5/2015	75	583.2	583.8	0.6
	75	1516.8	1517.2	0.4
	75	3634.1	3634.7	0.6
	9	3396.3	3396.4	0.1
	18	324.2	324.8	0.6
	18	388.2	389.5	1.3
	18	3662.1	3663.7	1.6

The frequency of each group of PETs was averaged across the simulation runs for each PM-peak hour and rounded (Table 9-4). Simulation models were calibrated by adjusting appropriate priority rules and driving behavior parameters.

Table 9-4: Average frequency of each group of PETs within 24 simulation runs

Simulation			The Average Frequency of PETs < 2 sec		
date	**Day**	**Time (PM)**	**High Risk (0-1)**	**Medium Risk [1-1.5)**	**Low Risk [1.5-2)**
4/22/2015	Wed	4 – 5	1/4≈0	2/4≈1	1/4≈0
4/23/2015	Thu	5 – 6	1/4≈0	1/4≈0	1/4≈0
4/27/2015	Mon	5 – 6	6/4≈2	3/4≈1	4/4=1
4/30/2015	Thu	5 – 6	9/4≈2	4/4=1	1/4≈0
5/1/2015	Fri	5 – 6	5/4≈1	4/4=1	3/4≈1
5/5/2015	Tue	4 – 5	5/4≈1	1/4≈0	1/4≈0
		Total	**6**	**4**	**2**

The frequency of each group of PETs corresponding to each PM-peak hour in the field and the average frequency of PETs corresponding to each PM-peak hour in the traffic simulation models are shown in Table 9-5.

Table 9-5: Frequency of PETs between the field and the simulation model

	Frequency of PETs < 2sec (one PM-peak hour)		
SW Taylor & SW Naito	**High Risk (0-1)**	**Medium Risk [1-1.5)**	**Low Risk [1.5-2)**
	4	3	5
Simulation	6	4	2

10.0 ANALYSIS OF TRAFFIC SIMULATION MODELS DATA

The Exact Multinomial Test was used to determine the probability that the frequencies of PETs observed in the VISSIM simulation models fits the distribution observed in the field (Engles n.d.). This distribution is our model and consists of the relative frequency of each group of PETs observed in the intersection of SW Taylor St and SW Naito Pkwy. These model proportions are displayed in Table 10-1.

Table 10-1: Frequency of each group of PETs at the intersection of SW Taylor & SW Naito

SW Taylor & SW Naito	The frequency of PET during 6 hours		
	High Risk [0-1)	Medium Risk [1-1.5)	Low Risk [1.5-2)
	4	3	5
Model 9	P1=4/12	P2=3/12	P3=5/12

The null hypothesis:

$$H_0: P_{High\ risk\ of\ collision\ in\ the\ VISSIM\ simulation\ models} = 4/12$$

$$H_0: P_{Moderate\ risk\ of\ collision\ in\ the\ VISSIM\ simulation\ models} = 3/12$$

$$H_0: P_{Low\ risk\ of\ colision\ in\ the\ VISSIM\ simulation\ models} = 5/12$$

Where $\mathrm{P}_{\text{High risk of collision in the VISSIM simulation models}}$ is the relative frequency of PETs with high risk of collision in the VISSIM simulation models.

The alternative hypothesis states that at least one of the null hypotheses is false.

10.1 Model 9 (PETs for Type I and Type II Conflicts within 6 PM-Peak Hours)

The result of Fisher's Exact Test indicates that the frequency of PETs observed in the field did not significantly differ from those observed in the simulation models, with a p-value of 0.58. The Exact Multinomial Test indicated 22% probability that the frequency of each group of PETs observed in the VISSIM simulation models fit Model 9. P-values equal to 0.22 gives the probability of observing the frequencies in the traffic simulation models given the model 9. Results of Fisher's Exact test and the Exact Multinomial test of goodness of fit are shown in Table 10-2 and plots of the graphical validation approach are shown in Figure 10-1.

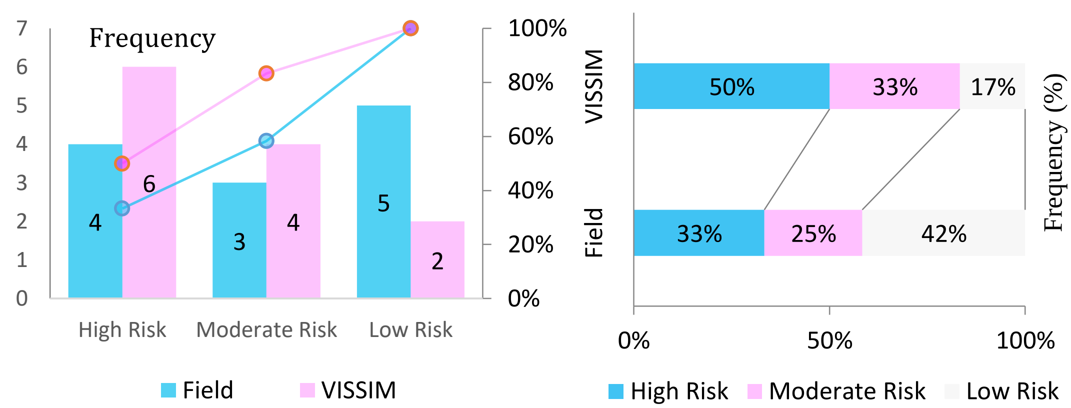


Figure 10-1: The frequency of each group of PETs (bars) along with their cumulative percent (line) (Left), the percentage of frequencies in each group of observations within each respective bar (Right).

Table 10-2: **The results of Fisher's Exact and Exact Multinomial Tests**

SW Taylor & SW Naito	Conventional Method of Measuring PET	
	Exact Multinomial Test	Fisher's Exact Test
Model 9	**0.22**	**0.58**

10.2 Conclusion

We tested the frequency of PETs during PM-peak-hours for type I and type II conflicts in the traffic simulation models in VISSIM and the intersection of SW Taylor St and SW Naito Pkwy. The probability that the frequency of PETs observed in the VISSIM simulation models fit Model 9 was 22%. As all differences between the VISSIM simulation models and the field are controlled for, 22% probability of goodness of fit is rather low. This result indicates that the VISSIM traffic simulation software may not reflect actual driving behavior in bicycle–vehicle conflict events and may not be a valid tool for traffic safety assessment for bicycle-vehicle interactions. However, as discussed before, further adjustment in travel behavior parameters, especially for bicyclists, may improve the validation result.

One potential reason for the poor result may be that actual drivers have unusual behaviors in conflict events with other road users, and traffic simulation models may not be truly able to reflect these behaviors during conflict events. For an example, run-over crashes usually occur in simulation models even after adjusted calibration of driving behavior parameters. This is not reflective of behavior in the field. It may be surmised that the simulation approach may be more applicable in comparing relative scenarios.

11.0 SUMMARY OF FINDINGS

The conventional definition of post-encroachment time (PET) as a surrogate safety measure was found to be unable to identify all right-hook conflicts between right-turning motorists and through bicyclists at intersections, and its measure could not truly represent the severity of risk of collision at times. Hence, the conventional definition of PET was extended using insights from the analysis of video records in the field. The proposed measures of PET helped to identify four out of seventeen right-hook conflicts with high risk of collision and eliminate conflicts without a risk of collision. The validation of the OSU driving simulator was inconclusive where pedestrians and oncoming left turning vehicles were involved in right-hook conflicts due to the lack of observations and different lane configurations between intersections designed in the OSU driving simulator environment and the intersection of SW Taylor St and SW Naito Pkwy in the field. The validation of the OSU driving simulator between only right-turning motorists and through bicyclists was promising but not definite because of differences between two environments. The statistical validation approach revealed that simulated bicyclists with constant speed in the OSU driving simulator environment versus actual bicyclists with variable speeds in the field had considerable negative effect on the validation of the OSU driving simulator. Likewise, the effect of traffic conditions and measuring method of PETs had considerable effect on the validation of the OSU driving simulator where high-speed bicyclists were involved in right-hook conflicts. Traffic simulation models in VISSIM were found to poorly reflect actual driving behaviors in the field for bicycle-vehicle interactions.

However, the validation of traffic simulation models may be improved by further adjustments of travel behavior parameters in simulation models, particularly for bicyclists. In sum, these results suggested that the performance of the OSU driving simulator was better than traffic simulation models in VISSIM for traffic safety assessment.

11.1 Limitations

The main limitation of this research was that a perfectly matched intersection in the field for the intersections designed in the OSU driving simulator was not identified in Portland, Oregon, and therefore could not be used as a study location. Another limitation was that the true population distribution of relative frequencies of PETs is unknown, and thus the relative frequency of each group of PETs observed in the field was used as the best guess to estimate the true distribution in the Exact Multinomial Test of Goodness of Fit. Finally, more research needs to be done in order to determine the best calibration of travel driving behavior parameters, particularly for bicyclists, at conflict events in traffic simulation models such as VISSIM.

11.2 Future Work

Surrogate safety measures should be analyzed by viewing video records to gain insight into road user dynamics. In addition, their definitions should take into account conflict and road user type so that conflicts are properly measured. In order to obtain the true population proportion of post-encroachment time for each risk-level group, more intersections need to be studied. For stronger validation of the OSU driving simulator for right-hook conflicts, a bicycling simulator should be synchronized with the OSU driving simulator and intersections designed in the OSU driving simulator should reflect actual intersections in the field.

12.0 REFRENCES

Alhajyaseen, Wael K. M. 2014. "The Development of Conflict Index for the Safety Assessment of Intersections Considering Crash Probability and Severity." *Procedia Computer Science* 364 – 371.

Allen, Brian L., B. Tom Shin, and Peter J. Cooper. 1977. *Analysis of Traffic Conflicts and Collisions.* Hamilton: Department of Civil Engineering, Mc Master University.

Archer, Jeffery. 2004. *Methods for the Assessment and Prediction of Traffic Safety at Urban Intersections and their Application in Micro-simulation Modeling* . Stockholm, Sweden: Department of Infrastructor, Royal Institute of Technology .

Bai, Lu, Pan Liu, Yanyong Guo, and Chengcheng Xu. 2015. *Understanding Factors Affecting Frequency of Traffic conflicts between Electeric Bicycles and Motorized Vehicles at Signalized Intersections.* Nanjing: the 94nd Annual Meeting of the Transportation Research Board.

Blaauw, J. G. 1982. "Driving experience and task demands in simulator and instrumented car: a validation study." *Human Factors 24 (4)* 473–486.

Brown, Lacy S. 2012. *A Validation of the Oregon State University Driving Simulator*. Oregon State University.

Byungkyu, Park, and Won Jongsun. 2006. *Microscopic Simulation Model Calibration and Validation Handbook.* Virginia Transportation Research Council.

CHILAKAPATI, PRAVEEN KUMAR. 2006. *DRIVING SIMULATOR VALIDATION AND REAR-END CRASH RISK ANALYSIS AT A SIGNALISED INTERSECTION.* Orlando: the University of Central Florida.

Douglas, Gettman, and Larry Head. 2003. *Surrogate Safety Measures from Traffic Simulation Models.* Tucson: Transportation Research Board.

Engen, Thomas. 2008. *Thomas Engen Use and validation of driving simulators.* Trondheim: Norwegian University of Science and Technology.

Engles, William R. n.d. *XNomial – Exact Test For Multinomial.* Accessed May 05, 2016. https://cran.r-project.org/web/packages/XNomial/vignettes/XNomial.html.

n.d. *Fisher's Exact Test for Count Data.* Accessed October 1st, 2016. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/fisher.test.html.

Gettman, Douglas, Lili Pu, Tarek Sayed, and Steve Shelby. 2008. *Surrogate Safety Assessment Model and Validation: Final Report* . FHWA-HRT-08-051.

Glauz, W D, K M Bauer, and D J Migletz. 1985. *EXPECTED TRAFFIC CONFLICT RATES AND THEIR USE IN PREDICTING ACCIDENTS.* Transportation Research Record.

Godley, Stuart T., Thomas J. Triggs, and Brian N. Fildes. 2001. "Driving simulator validation for speed research." *Accident Analysis and Prevention 34* 589–600.

Huang, Fei, Pan Liu, Hao Yu, and Wei Wang. 2013. "Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections." *Accident Analysis and Prevention 50* 1014-1024.

Hurwitz, David, Christopher Monsere, Jannat Mafruhatul, Jennifer Warner, and Ali Razmpa. 2015. *TOWARDS EFFECTIVE DESIGN TREATMENT FOR RIGHT-TURNS AT INTERSECTIONS WITH BICYCLE TRAFFIC.* Salem: Oregon Department of Transportation.

Killi, Deepak V., and P. Vedagiri. 2014. "Proactive Evaluation of Traffic Safety at An Unsignalized Intersection Using MicroSimulation." *Traffic and Logistics Engineering* Vol. 2, No. 2.

Laureshyn, Aliaksei, Ase Svensson, and Christer Hyden. 2010. "Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation." *Accident Analysis and Prevention* 1637-1646.

McGehee, Daniel V, G.H. Scott Baldwin, and Elizabeth N. Mazzae. 2000. "Driver Reaction Time in Crash Avoidance Research: Validation of a Driving Simulator Study on a Test Track." *Human Factors and Ergonomics Society Annual Meeting Proceedings* .

Miller, David Michael. 2009. *DEVELOPING A PROCEDURE TO IDENTIFY PARAMETERS FOR CALIBRATION OF A VISSIM MODEL.* Georgia Institute of Technology.

Oregon Department of Transportation. 2011. *PROTOCOL FOR VISSIM SIMULATION.* Salem: Oregon Department of Transportation.

Park, Byungkyu (Brian), and Hongtu (Maggie) Qi. 2005. "Development and Evaluation of a Procedure for the Calibration of Simulation Models." *Transportation Research Record: Journal of the Transportation Research Board* 208-217.

Peesapati, Lakshmi N., Michael P. Hunter, and Michael O. Rodgers. 2013. "Evaluation of Post-encroachment Time as Surrogate for Opposing Left-Turn Crashes." *Transportation Research Board* 42-51.

Philips, Brian, and Tom Morton. 2015. *Making Driving Simulators More Useful for Behavioral Research-Simulator Characteristics Comparison and Model-Based Transformation.* Iowa: U.S. Department of Transportation Fedral Highway Administration.

PTV AG. 2014. *PTV VISSIM USER MANUAL.* Germany: PTV GROUP.

ReliaSoft Corporation. 2007. *Reliability Hot Wire.* January 17. Accessed June 24, 2016. http://www.weibull.com/hotwire/issue71/relbasics71.htm.

Sayed, Tarek, and Sany Zein. 1998. *TRAFFIC CONFLICT STANDARDS FOR INTERSECTIONS.* Vancouver: Department of Civil Engineering, University of British Columbia.

Sayed, Tarek, Gerald Brown, and Francis Navin. 1994. "SIMULATION OF TRAFFIC CONFLICTS AT UNSIGNALIZED INTERSECTIONS WITH TSC-Sim." *Accid.Anal. and Prev* 593-607.

Songchitruksa, Praprut, and Andrew P. Tarko. 2006. "The extreme value theory approach to safety estimation." *Accident Analysis and Prevention* 811–822.

Tang, Keshuang, and Masao Kuwahara. 2011. "Implementing the Concept of Critical Post-Encroachment Time for All-Red Clearance Interval Design at Signalized Intersections." *the Eastern Asia Society for Transportation Studies.*

Tarko, Andrew, Gary Davis, Nicolas Saunier, Tarek Sayed, and Simon Washington. 2009. *SURROGATE MEASURES OF SAFETY White Paper.* Subcommittee on Surrogate Measures of Safety and Committee on Safety Data Evaluation and Analysis.

Tornros, J. 1998. "Driving behaviour in a real and a simulated roadtunnel--a validation study." *Accident Analysis and Prevention 30 (4)* 497-503.

Zangenehpour, Sohail, Jillian Strauss, Luis F. Miranda-Moreno, and Nicolas Saunier. 2014. *Are Intersections With Cycle Tracks Safer? A Control-Case Study Based On Automated Surrogate Safety Analysis Using Video Data.* Montréal: Transportation Research Board 2015 Annual Meeting.

Zhou, Huanyun, and Fei Huang. 2013. "Development of traffic safety evaluation method based on simulated conflicts at signalized intersections." *Procedia-Social and Behavioral Sciences 96* 881-885.

APPENDIX

Video Files

Video records of some Type I and four Type II right-hook conflicts observed in the intersection of SW Taylor and SW Naito, along with their measure of PET are attached to this paper for illustration. Video records of some right-hook conflicts and their PET measures in traffic simulation runs are also attached for illustration. The open source VLC media player and Kinovea video editor were used to play footages and measure PETs. File names, types, and sizes are listed in the table below.

Name	Types	Size (KB)
Type I	.avi	74,165
Type I_VISSIM	.avi	59,928
Type II	.avi	29,151
Type II_VISSIM	.avi	2,731
Waiting	.avi	29,453
Waiting_VISSIM	.avi	12,043