Report No. K-TRAN: KU-21-4 ▪ FINAL REPORT ▪ November 2023

Evaluation of Near-Miss Crashes Using a Video-Based Tool

Alexandra Kondyli, Ph.D.
Steven D. Schrock, Ph.D., P.E., F. ITE
Fahmid Tousif

The University of Kansas

Kansas Department of Transportation

1 Report No. K-TRAN: KU-21-4	**2 Government Accession No.**	**3 Recipient Catalog No.**
4 Title and Subtitle Evaluation of Near-Miss Crashes Using a Video-Based Tool		**5 Report Date** November 2023 **6 Performing Organization Code**
7 Author(s) Alexandra Kondyli, Ph.D. Steven D. Schrock, Ph.D., P.E., F. ITE Fahmid Tousif		**8 Performing Organization Report No.**
9 Performing Organization Name and Address The University of Kansas Department of Civil, Environmental & Architectural Engineering 1530 West 15th St. Lawrence, Kansas 66045-7609		**10 Work Unit No. (TRAIS)** **11 Contract or Grant No.** C2171
12 Sponsoring Agency Name and Address Kansas Department of Transportation Bureau of Research 2300 SW Van Buren Topeka, Kansas 66611-1195		**13 Type of Report and Period Covered** Final Report July 2020 — February 2022 **14 Sponsoring Agency Code** RE-0816-01

15 Supplementary Notes
For more information write to address in block 9.

16 Abstract

This research investigated the prediction accuracy of a video-based tool developed by Transoft Solutions for predicting near-miss crashes at signalized intersections. The research team selected two signalized intersections in Overland Park, Kansas, and collected two weeks of video data from both locations. Only weekday data were collected for the two weeks in February and March 2021. The data were provided to Transoft Solutions, and analyzed results were accessed from the vendor website. Approximately 10% of the data were sampled for manual validation, which included drawing vehicle trajectories and conflict spots on top of the computer screen and measuring time manually in milliseconds. Both post-encroachment time (PET) and time-to-collision (TTC) data were validated based on three conflict categories (critical conflicts, minor conflicts, and potential conflicts) and three weather and traffic conditions (rainy peak condition, sunny peak condition, and sunny off-peak condition). Four performance measures (mean absolute deviation - MAD, root mean squared error - RMSE, mean absolute percentage error - MAPE, and root mean squared log error - RMSLE) were selected, and an ANOVA with Tukey's post hoc test was carried out for each analysis. Both the PET and the TTC data showed that sunny weather had better predictability than rainy weather. Statistical analysis revealed a significant difference between means from observed and predicted values for the PET data. Overall, the video-based tool by Transoft Solutions demonstrated moderate predictability and overestimated the conflict measures for both the PET and the TTC data.

17 Key Words Accuracy, Machine Vision, Mathematical Prediction, Near Crashes, Signalized Intersections, Software	**18 Distribution Statement** No restrictions. This document is available to the public through the National Technical Information Service www.ntis.gov.

19 Security Classification (of this report) Unclassified	**20 Security Classification (of this page)** Unclassified	**21 No. of pages** 71	**22 Price**

Form DOT F 1700.7 (8-72)

This page intentionally left blank.

Evaluation of Near-Miss Crashes Using a Video-Based Tool

Final Report

Prepared by

Alexandra Kondyli, Ph.D.
Steven D. Schrock, Ph.D., P.E., F. ITE
Fahmid Tousif

The University of Kansas

A Report on Research Sponsored by

THE KANSAS DEPARTMENT OF TRANSPORTATION
TOPEKA, KANSAS

and

THE UNIVERSITY OF KANSAS
LAWRENCE, KANSAS

November 2023

PREFACE

The Kansas Department of Transportation's (KDOT) Kansas Transportation Research and New-Developments (K-TRAN) Research Program funded this research project. It is an ongoing, cooperative and comprehensive research program addressing transportation needs of the state of Kansas utilizing academic and research resources from KDOT, Kansas State University and the University of Kansas. Transportation professionals in KDOT and the universities jointly develop the projects included in the research program.

NOTICE

The authors and the state of Kansas do not endorse products or manufacturers. Trade and manufacturers names appear herein solely because they are considered essential to the object of this report.

This information is available in alternative accessible formats. To obtain an alternative format, contact the Office of Public Affairs, Kansas Department of Transportation, 700 SW Harrison, 2nd Floor – West Wing, Topeka, Kansas 66603-3745 or phone (785) 296-3585 (Voice) (TDD).

DISCLAIMER

The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the views or the policies of the state of Kansas. This report does not constitute a standard, specification or regulation.

Abstract

This research investigated the prediction accuracy of a video-based tool developed by Transoft Solutions for predicting near-miss crashes at signalized intersections. The research team selected two signalized intersections in Overland Park, Kansas, and collected two weeks of video data from both locations. Only weekday data were collected for the two weeks in February and March 2021. The data were provided to Transoft Solutions, and analyzed results were accessed from the vendor website. Approximately 10 percent of the data were sampled for manual validation, which included drawing vehicle trajectories and conflict spots on top of the computer screen and measuring time manually in milliseconds. Both post-encroachment time (PET) and time-to-collision (TTC) data were validated based on three conflict categories (critical conflicts, minor conflicts, and potential conflicts) and three weather and traffic conditions (rainy peak condition, sunny peak condition, and sunny off-peak condition). Four performance measures (mean absolute deviation - MAD, root mean squared error - RMSE, mean absolute percentage error - MAPE, and root mean squared log error - RMSLE) were selected, and a one-way analysis of variance (ANOVA) with Tukey's post hoc test was carried out for each analysis. Both the PET and the TTC data showed that sunny weather had better predictability than rainy weather. Statistical analysis revealed a significant difference between means from observed and predicted values for the PET data. Overall, the video-based tool by Transoft Solutions demonstrated moderate predictability and overestimated the conflict measures for both the PET and the TTC data.

Acknowledgments

The authors of this report thank project monitors Alan Spicer, Steven Buckley, and William Hughes for their guidance and suggestions. The authors also thank the supervisory engineer from the city of Overland Park, Mr. Shawn Gotfredson, for assisting with the traffic video data.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background

Vision-based trajectory data provide useful information for analyzing roadway safety and interactions of roadway users (e.g., drivers, pedestrians, bicyclists). Several companies offer vision-based software that promises to identify near-miss crashes between vehicles, pedestrians and vehicles, or bicycles and vehicles by estimating surrogate measures for these crashes using trajectory data. However, the accuracy of these tools for predicting near-miss crashes has not been evaluated. Transoft Solutions offers a popular vision-based tool that has been deployed in several cities throughout the United States and Canada. This research project evaluated the Transoft Solutions tool to assist local agencies, such as the Kansas Department of Transportation (KDOT), in identifying the benefits of deploying such technology in their network.

1.2 Objectives

The main objective of this research was to evaluate a video-based tool offered by Transoft Solutions for its prediction accuracy of near-miss crashes at signalized intersections. This project specifically evaluated two commonly used near-miss safety measures: time to collision (TTC) and post-encroachment time (PET). Manual validation of these two measures was conducted and compared to tool-based validation measures. The findings of this study will help KDOT and local transportation agencies determine whether or not to invest in the Transoft Solutions video-based tool at signalized intersections.

Chapter 2: Literature Review

Vehicle conflicts are often used as safety measurement indicators of road safety. Various traffic conflict techniques (TCTs) have been used to study intersection safety. Chin et al. (1991) defined TCT as a procedure to observe and infer crash potential in any physical location; while Amundsen and Hydén (1977) defined vehicle conflict as a probable collision course involving two or more road users, if no evasive action is taken. Near-miss traffic conflict can be defined as a situation when two or more vehicles evade an actual collision by a margin of time span. Hourdos et al. (2006) defined near-misses as when one or more vehicles deviate from the original lane and end up being on the shoulder to avoid rear-end collisions.

2.1 Surrogate Safety Measures Definitions

Previous studies have utilized the following surrogate safety measures to analyze near-miss conflicts: post-encroachment time (PET), time to collision (TTC), minimum time to collision (MTTC), time to intersection (TTI), gap time (GT), time headway, deceleration to safety time (DST), deceleration rate to avoid crash (DRAC), and proportion of stopping distance (PSD). PET, the time lapse of two road users in a conflict zone (Allen et al., 1978), has been used by several researchers in the literature (Peesapati et al., 2013; Zangenehpour et al., 2017). Hayward (1972) introduced the concept of TTC as the time for two consecutive vehicles to be in a collision if they continue at the same speed on the same path; while MTTC considers vehicle acceleration and nullifies the assumption that the lag vehicle is traveling at a constant speed (Ozbay et al., 2008). TTI is the expected time of a vehicle to enter the intersection at a constant speed from the start of braking (Van der Horst, 1990). GT is the time interval between the rear bumper of the leading vehicle and the front bumper of the following vehicle (FHWA, 1976); and time headway refers to the time that elapses between the front of the leading vehicle when it passes a point on the roadway and the front of the following vehicle as it passes the same point (Evans, 1991). DST measures the deceleration required to achieve a non-negative PET value while the conflicting movements of road users stay unchanged (Hupfer, 1997). DRAC, which is the minimum required deceleration rate to avoid a crash with the leading vehicle, has also been used as a safety indicator in several studies (Cooper & Ferguson, 1976; Gettman & Head, 2003; Guido et al., 2011; Astarita et al., 2012; Fazekas et al., 2017). Allen et al. (1978) defined PSD as the ratio of available distance for a driver to maneuver and the projected distance to the collision point. According to Gettman et al. (2008), PET and TTC are the most effective surrogate safety measures for analyzing near-miss conflicts at intersections. In general, PET provides a discrete value; while TTC provides a set of values continuously calculated over time (Kathuria & Vedagiri, 2020).

2.1.1 Post-Encroachment Time

PET, as defined above, is a popular surrogate safety measure used in many investigations. Encroachment time (ET) is the time lapse between when the front bumper of a vehicle reaches the conflict spot and when the rear bumper of the same vehicle leaves the conflict spot. The conflict spot is defined as the intersecting point of two crossing vehicles (Allen et al., 1978; Songchitruksa & Tarko, 2004). PET includes the time lapse between two vehicles when the rear bumper of the first vehicle leaves the conflict spot, and the front bumper of the second vehicle reaches the same spot. Therefore, GT is the summation of ET and PET. In Figure 2.1, t_1 is the time when the front bumper of the first vehicle reaches the conflict spot, t_2 is the time when the rear bumper of the first vehicle leaves the conflict spot, and t_3 is the time when the front bumper of the second vehicle reaches the same conflict spot. Therefore, ET is defined as (t_2-t_1), PET is defined as (t_3-t_2), and GT is defined as (t_3-t_1).

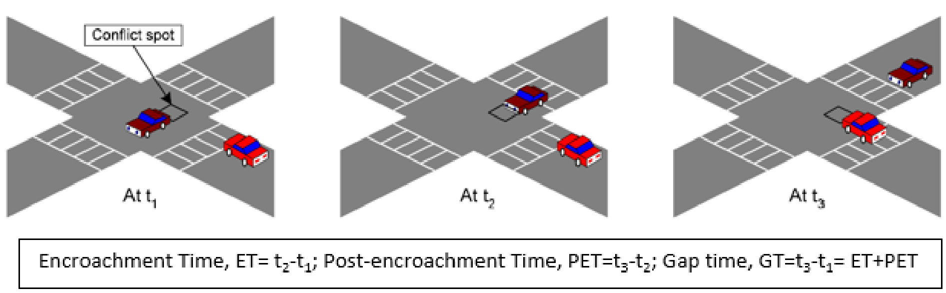

Figure 2.1: Illustrations of ET, PET, and GT
Source: Songchitruksa & Tarko (2004)

Laureshyn et al. (2010) defined PET as the minimum delay between trajectories of two road users. Researchers have used various PET threshold values to define whether a vehicle conflict is considered critical or not; however, these values were generally not consistent in the literature. For example, Vogel (2003) stated that a PET value up to 6 seconds is dangerous, while Archer (2000) considered any PET value less than 1 second to be very unsafe. According to Tang and Kuawahara (2011), a PET value of less than 2 seconds implies probable collision potential. Caliendo and Guglielmo (2013) used 5 seconds as a PET threshold in their study. Songchitruksa and Tarko (2004) and Zhang et al. (2020) used 6.5 and 6 seconds, respectively; and Zangenehpour et al. (2017) used 3 seconds as the PET threshold value.

Songchitruksa and Tarko (2004) used manual PET measurements, an automated commercial video detection system, and semi-automated proprietary image processing software to evaluate right-angle collisions at signalized intersections. Historical crash data were collected from Indiana state police for right-angle collisions from 1997 to 2000, and eight hours of video data were collected (7:00–9:30 a.m., 11:00 a.m. to 2:00 p.m., and 3:30–6:00 p.m.). For the manual detection procedure, PET values were calculated via frame-by-frame analysis using Adobe premiere software. For the automated method, Autoscope, a virtual tripwire image-processing unit was used to detect and obtain PET values. Autoscope was also used for the semi-automated method to collect data and verify a subset of data using frame-by-frame manual analysis to extract PET values. All PET data from these three processes were then compared to the historical crash data. Using a PET threshold of 6.5 seconds, Poisson and negative binomial regression revealed a significant correlation between manual PET count and observed crash data. Overall, the manual process was found to be the most effective method, as the automated and semi-automated methods produced a significant amount of false and missed detections. In a follow-up study using the same data, Songchitruksa and Tarko (2006) used extreme value theory (EVT) and a PET threshold of 4.5 seconds, resulting in a positive correlation between right-angle collisions and PET.

Pirdavani et al. (2010) used the S-Paramics (Sykes, 2010) microsimulation tool to obtain PET values for various traffic volumes and speeds at an unsignalized intersection. The study assumed four conflict zones. As a result, four loop detectors in four approaches were placed to collect the speed and position of each vehicle. Major road traffic volume was 500–650 vph, while minor road volume was 150–250 vph. Major road speed ranged from 28 mph to 47 mph, and minor road speed varied from 22 mph to 31 mph. The PET threshold used in this study was 3 seconds.

Results showed that increased traffic volume and posted speed limits caused PET values to decrease; thereby, decreasing intersection safety.

Peesapati et al. (2011) proposed a semi-automated data collection method to extract surrogate safety measures. Their study considered three surrogate safety measures (acceleration and deceleration profiles, PET, and intersection approach speed). The researchers used Java script to collect speed, acceleration, and deceleration data from vehicle trajectories at a signalized intersection in Atlanta, Georgia, and PET was calculated manually from timestamps of the extracted data. The methodology was validated using data from vehicles instrumented with global positioning systems (GPS). Results showed that increasing speed led to increasing error and noise in the collected data. Therefore, the authors concluded that their data collection method would be more accurate for low-speed roads and arterials.

In a similar study, Peesapati et al. (2013) investigated the usability of PET as a surrogate safety measure for left turning and opposing through vehicles in four-leg signalized intersections in Atlanta, Georgia. Crash data were collected from 2006 to 2009 for 18 intersections. Video data from one day (2:00–7:00 p.m.) were collected for each intersection, and the study used a semiautomatic video processing software to analyze the data to obtain PET. The researchers experimented with 10 PET threshold values ranging from 1 second to 10 seconds to count conflicts and compare them to the collected crash data. Results showed that a PET threshold of 1 second had the highest correlation with the collected crash data.

Zangenehpour, Strauss, et al. (2015) examined the safety of cycle tracks in 23 intersections in Montreal, Canada. The study focused on a right-hook scenario, where a right-turning vehicle interacts with a through bicycle. Out of 23 intersections, eight intersections had cycle tracks on the right side, seven had cycle tracks on the left side, and eight had no cycle tracks. The open-source software Traffic Intelligence (Saunier & Sayed, 2006) was used to extract PET values from 90 hours of video data. A logistic random effects model was developed for each type of intersection. Results showed that intersections with cycle tracks on the right were safer than intersections with cycle tracks on the left or intersections without cycle tracks.

Razmpa (2016) used field data, driving simulator data, and microsimulation data to compare the PET values of bicycle-vehicle interactions. The study identified 52 right-hook conflicts from 135 hours of video footage collected from a signalized intersection in Portland, Oregon. The SMplayer program was used to analyze the field data frame-by-frame, and Fisher's Exact Test was used to compare the frequency distribution of PET in all three data groups. Simulation results differed significantly from the field data, but PET values from the driving simulator and field data sets did not differ significantly.

Zangenehpour et al. (2017) utilized 72 hours of video data in a before-after study to investigate the effect of curb radius adjustment on pedestrian safety. They used Traffic Intelligence software to collect PET data as the conflict measure, and they categorized the conflict measure as high ($PET \leq 1s$), medium ($1s < PET \leq 3s$), and low ($3s < PET \leq 5s$). Results showed that a curb radius reduction helped reduce the occurrence of high-risk conflicts and vehicle speeds.

Shekhar Babu and Vedagiri (2018) used PET and vehicle speed to evaluate the safety of unsignalized intersections. AutoCAD, Corel Video Studio Pro X6 software, and AVS video editor software were used to extract the PET values. The PET threshold for critical conflicts and for each PET value was 6 seconds. The critical speed of the conflicting through vehicle was calculated as twice the product of gravitational acceleration, coefficient of friction between tire and road surface and the recorded PET, using the concept of braking distance. The study compared critical conflicts among two wheelers (motorbikes), light motor vehicles (cars and minivans), and heavy vehicles (buses and trucks). Results showed that light motor vehicles had higher occurrences of critical conflicts than heavy vehicles and two wheelers.

2.1.2 Time to Collision

Hayward (1972) defined TTC as the time to get in a collision with the leading vehicle if the path and speed of both vehicles remain unchanged. A higher TTC value indicates increased safety conditions, as shown in Figure 2.2 (Minderhoud & Bovy, 2001).

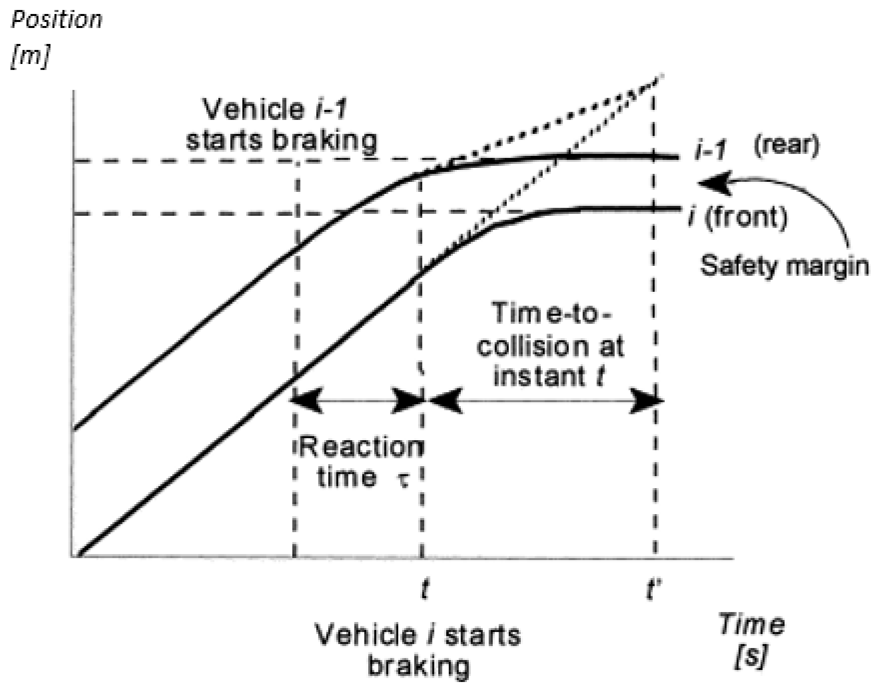

Figure 2.2: Illustration of TTC
Source: Minderhoud & Bovy (2001)

As shown in the figure, *i* and *i*-1 represent vehicle trajectories of the subject vehicle and lead vehicle, respectively. At time *t*, the lead vehicle brakes, and then the subject vehicle begins to brake. If the subject vehicle maintains a constant speed differential, it will collide with the lead vehicle at *t*` time, as shown by the dashed lines. Then the TTC value at time *t* can be calculated as the ratio of distance between the two vehicles and the speed between the two vehicles. Minderhoud and Bovy (2001) also proposed two new indicators based on TTC: time-exposed time (TET) to collision and time-integrated time (TIT) to collision. TET is defined as the summation of all instances when the driver of the lag vehicle tends to enclose with the front vehicle with a TTC value less than the threshold. On the other hand, TIT utilizes the integral of the TTC profile of the vehicle, thereby depicting the safety level with respect to the TTC threshold (Figure 2.3).

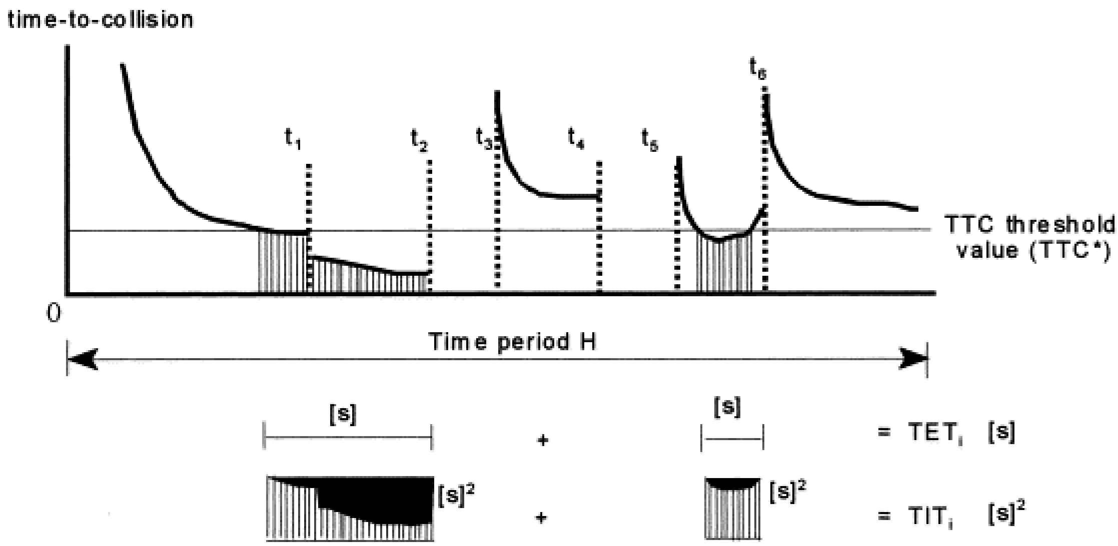

Figure 2.3: TET to Collision and TIT to Collision
Source: Minderhoud & Bovy (2001)

Sayed et al. (2013) used computer vision techniques to investigate traffic conflict in a major signalized intersection in Vancouver, Canada. The authors collected TTC data as a surrogate safety measure with a threshold of 3 seconds to categorize a conflict as critical. The researchers conducted a before-after safety evaluation for right-turn smart channels that were designed to have short distance exposure and short signal cycles. They collected TTC data in a signalized intersection with no treatment and three other locations with smart channels. Results showed a significant reduction in conflict frequency (conflicts/hour) with this treatment.

In a similar study, Zaki et al. (2013) used computer vision to investigate intersection safety for vehicle-pedestrian conflicts. The study was conducted at an intersection in downtown Vancouver, Canada. Data were collected from a single camera on four weekdays from 9:00 a.m. to 5:00 p.m., and a different camera view was taken each day. A TTC threshold of 3 seconds was used to define a critical conflict. Results showed that pedestrians crossing in one of the approaches of the selected intersection were most vulnerable for vehicular conflict because pedestrians had the highest number of critical conflicts.

Jackson et al. (2013) proposed a video camera-based system that utilizes Traffic Intelligence software to investigate microscopic traffic data. The study collected video data to conduct a before-after study for a special lane-change ban on urban highway segments near exit and entrance ramps in Montreal, Canada. Individual pixels were detected frame-by-frame, and feature trajectories were obtained using a tracking algorithm and then grouped based on common motion patterns. The trajectories were smoothed using the moving average technique. Based on these trajectories, various road safety measures were then computed. Overall, the results revealed various TTC distributions due to the implementation of the treatment, proving the effectiveness of the proposed system.

Bai et al. (2015) identified factors that affect conflicts between vehicles and bicycles. The study collected 735 hours (approximately one month) of video data from 20 four-legged signalized intersections in Kunming, China. TTC data were manually extracted from the video data. The researchers developed conflict models to observe how variables such as peak period, channelization, cross-street width, bicycle lane width, barrier type, median type, two-wheeled vehicles, through vehicles, and left-turning motorized vehicles impact traffic conflicts. Results showed that motorized vehicles contribute more towards traffic conflicts than two-wheeled vehicles at signalized intersections.

St-Aubin et al. (2015) analyzed roundabout safety using 473 hours (approximately three weeks) of video data from 41 roundabouts in Montreal, Canada. Traffic Intelligence software was used to analyze the video data. All 41 sites were clustered into six groups, and TTC values less than 1.5 seconds were considered the threshold for critical conflicts. Based on TTC distribution for the six groups, the roundabouts converted from the traffic circles were least safe.

Tageldin et al. (2018) studied safety benefits associated with the extension of left-turn length at three intersections in Surrey, Canada. Video data were collected for two days before and after the treatment. The researchers used the longest common subsequence (LCSS) algorithm to calculate TTC. A TTC threshold of 3 seconds was considered, and the average hourly conflict was compared before and after the treatment. Significant safety improvements due to the extended length of the left turn were observed.

Ke et al. (2017) studied the cost effectiveness of an onboard monocular camera to detect vehicle-pedestrian near-miss crashes. The camera was placed on a metro transit bus in Seattle, Washington, and more than 30 hours of video data were collected. A histogram of oriented gradients pedestrian detector and the Kanade-Lucas-Tomasi (KLT) feature tracker algorithm were used to detect and track pedestrian movements. The study also collected the comparison dataset organized by the Rosco/MobilEye Shield+ system, which utilizes multiple camera sensors. The researchers used TTC values ranging from 1 to 4 seconds to compare the number of different detections from both systems. Results showed that their camera achieved more than a 90% overlap rate with a TTC threshold of 2 seconds.

Guo et al. (2019) investigated the correlation between field-measured and simulation conflicts. TTC was the safety indicator. Seven hours of video data were collected from two signalized intersections in two cities in Australia. The study utilized the KLT feature tracker algorithm and the LCSS algorithm for automatic conflict data extraction. A total of 21 TTC thresholds ranging from 1 second to 3 seconds were used at an increment of 0.1 seconds. The TTC from the VISSIM simulation were extracted using the surrogate safety assessment model (SSAM) tool. The results showed a higher correlation between field measured and simulated TTC with increasing TTC threshold value.

2.1.3 Multiple Conflict Measures

Ismail et al. (2009) used video data to analyze pedestrian-vehicle conflict. Twenty hours of video data were collected over two days from a busy intersection in downtown Vancouver, Canada. The authors used the KLT feature tracker algorithm to extract TTC, PET, DST, and GT data. The thresholds were 1.5 seconds, 3 m/s^2, 1 second, and 1 second for TTC, DST, PET, and GT, respectively. Validation of all four surrogate safety measures was conducted according to the Observers Manual from the Federal Highway Administration (FHWA) (Parker & Zegeer, 1989). A comparison of the system-generated values and manually calculated data for the four safety indicators showed that no surrogate safety measure was individually capable of capturing all possible dangerous interactions among road users.

Caliendo and Guglielmo (2013) used the SSAM tool (Pu et al., 2008) with AIMSUN traffic simulation software to identify the number of critical conflicts at nine unsignalized urban intersections in Salerno, Italy. A TTC threshold of 1.5 seconds and a PET threshold of 5 seconds were used to define critical conflicts, and hourly volume data were collected from a video camera placed at the intersections. The study also collected crash data for five years from the selected intersections. Results showed a good match between recorded crashes and computed conflicts.

Zangenehpour, Miranda-Moreno, et al. (2015) conducted a safety analysis between two intersections in Montreal, Canada. The two intersections had similar traffic and geometric conditions, with the exception of the presence of a separate bicycle facility in one intersection. Traffic Intelligence software was used to extract PET and TTC values from seven hours of video data. A threshold of 5 seconds was used to compute the conflict rate, and a threshold of 1.5 seconds was used to compute the dangerous conflict rate. Results showed that the intersection with a separate bicycle facility was safer than the intersection without a bicycle facility.

Zheng et al. (2019) developed bivariate extreme value models to integrate various safety indicators for road safety estimation. Their study utilized four signalized intersections in two Canadian cities (Alberta and Surrey), as well as computer vision techniques with the KLT feature tracking algorithm to infer rear-end traffic conflicts. Four safety indicators (TTC, MTTC, PET, and DRAC) were used to develop the models in a combination of two indicators each time. TTC, MTTC, and PET values of less than 4 seconds and DRAC greater than 0 m/s^2 were set as thresholds for collecting conflict data. The numbers of estimated crashes from the models and observed crashes were then compared. Results showed that most of the estimated crashes were within 95% Poisson interval of the observed crashes; the combination of TTC and PET had the most accurate estimates.

Zhang et al. (2020) proposed a gated recurrent unit (GRU) neural network to predict pedestrian near-miss conflicts at signalized intersections. A total of 80 hours of video data were collected from two signalized intersections located in Seminole County, Florida. Computer vision techniques were used to extract the PET and TTC data. EVT was used to obtain the threshold values for both indicators. The threshold for PET and minimum TTC were 6 seconds and 3 seconds, respectively. A GRU model was then used to predict near-miss conflicts. Results showed a high accuracy for the newly proposed model, making it a recommendation for future development.

Stipanic et al. (2021) utilized vehicle speed, TTC, and PET to investigate the safety impact of stop-controlled intersections. Video data were collected from 97 intersections in Montreal, Canada, and were processed by an open-source computer vision platform. Results showed a significant decrease in vehicle speed in stop-controlled intersections and a significant decrease in TTCs between vehicles in partially stop-controlled intersections. The study did not find any decisive effect of stop-controlled intersections on PET.

2.2 Transoft Solutions Tool

TrafxSAFE, formerly known as BriskLUMINA, is a safety assessment tool developed by Transoft Solutions, formerly Brisk Synergies. The tool utilizes a web platform to provide insights regarding high-risk, crash-prone sites, contributing factors, and effects of countermeasures. It utilizes various complex deep-learning algorithms to analyze traffic video data and infer temporal and spatial information about road users (e.g., vehicles, bikes, pedestrians, etc.), position, speed, motion trajectories, and safety measures, such as TTC and PET. High resolution video data are first collected for the selected intersections, and then the video data are uploaded to the TrafxSAFE web platform, followed immediately by camera calibration to match the camera projection (image space) with the ground projection (world space). The tool then extracts road users' trajectories for each site using complex deep-learning algorithms and identifies, classifies, and labels each road user as cars, bikes, pedestrians, etc. (Figure 2.4). The algorithm analyzes the interaction among all road users frame-by-frame and generates information on the coordinates, speed, approximate location of near-miss collisions, and values of safety measures (TTC and PET) to download from the web platform. The tool also generates statistics, graphs, and maps from the generated information.

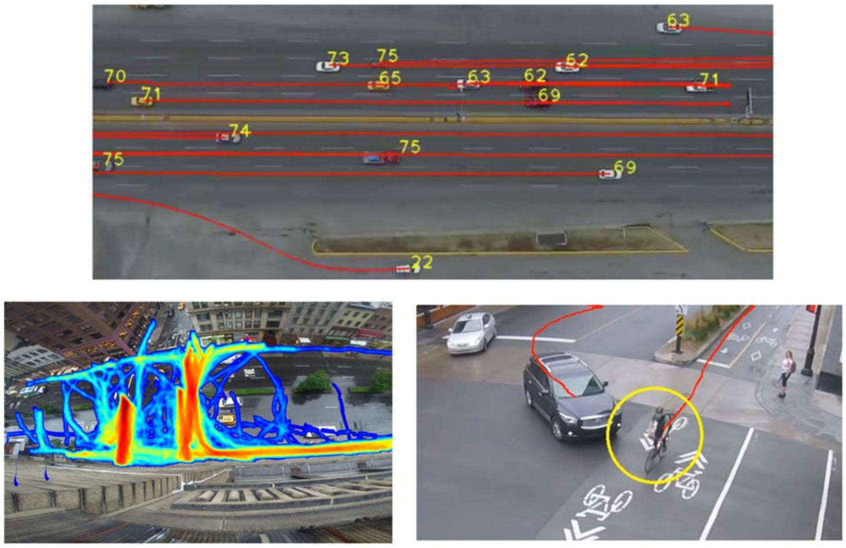

Figure 2.4: Road Users' Trajectories and Patterns
Source: (Brisk Synergies, 2018)

A report published by Samara et al. (2020) describes the analysis of approximately 4,500 hours of traffic video data collected from 40 intersections across the city of Bellevue, WA. A PET threshold of 2 seconds was defined as the critical conflict. The authors obtained data on different traffic safety metrics (speed, PET, etc), frequency of speeding events, and conflict rates. The authors observed the majority of critical traffic conflicts between through and left-turning vehicles.

2.3 Summary of Literature Review

In summary, previous studies have used various surrogate safety measures to quantify near-miss crashes, with PET and TTC being two of the most common safety measures. Many researchers developed their own procedure to manually calculate the PET and TTC values via frame-by-frame video data analysis (Songchitruksa & Tarko, 2004, 2006; Bai et al., 2015; Shekhar Babu & Vedagiri, 2018). Previous research utilized video data (Ismail et al., 2009; Laureshyn et al., 2010), in-vehicle sensors (Matsui et al., 2013), and simulation models (Gettman & Head, 2003) to investigate near-miss conflicts. Many studies also used computer vision techniques to develop their own algorithms to calculate PET and TTC (Ismail et al., 2009; Jackson et al., 2013; Sayed et al., 2013; Zaki et al., 2013; St-Aubin et al., 2015; Ke et al., 2017; Zangenehpour et al., 2017; Tageldin et al., 2018; Zhang et al., 2020; Stipanic et al., 2021). In addition, some researchers experimented with PET and TTC values obtained from microsimulation tools and driving simulators (Pirdavani et al., 2010; Caliendo & Guglielmo, 2013; Razmpa, 2016; Guo et al., 2019). Although vendors such as Transoft Solutions have produced their own near-miss crash detection algorithms that are deployed in cities throughout the United States (Samara et al., 2020), an insufficient amount of literature has evaluated the accuracy of these algorithms. Therefore, the objective of this study was to evaluate the accuracy of the safety measures provided by a vendor named Transoft Solutions.

Chapter 3: Data Collection

3.1 Location Selection

The research team initially reached out to the Multimodal Transportation Commission in Lawrence, Kansas, to request traffic video data, but the quality of obtained videos did not meet the resolution requirement of Transoft Solutions (1280 x 720 or higher) for video data analysis. Appendix A.1 contains detailed camera specification requirements by Transoft Solutions. After consulting with KDOT, the Public Works Department of Overland Park, Kansas, was contacted because video camera resolution of traffic cameras deployed throughout Overland Park met the resolution requirement of Transoft Solutions.

3.2 Intersection Selection and Data Collection Periods

Overland Park has high resolution traffic cameras in 105 locations throughout the city (Figure 3.1), and the city website (https://www.opkansas.org/city-services/traffic-roads-transportation/traffic-roads/traffic-cameras/) contains real-time traffic video feeds of all its traffic camera locations (Figure 3.2).

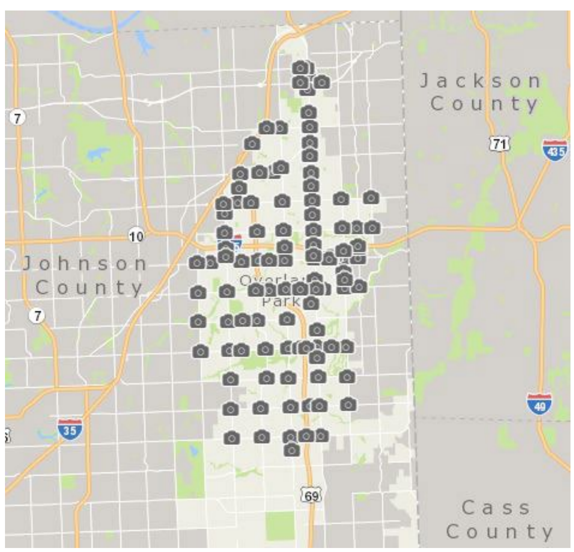

Figure 3.1: Overland Park Traffic Camera Locations

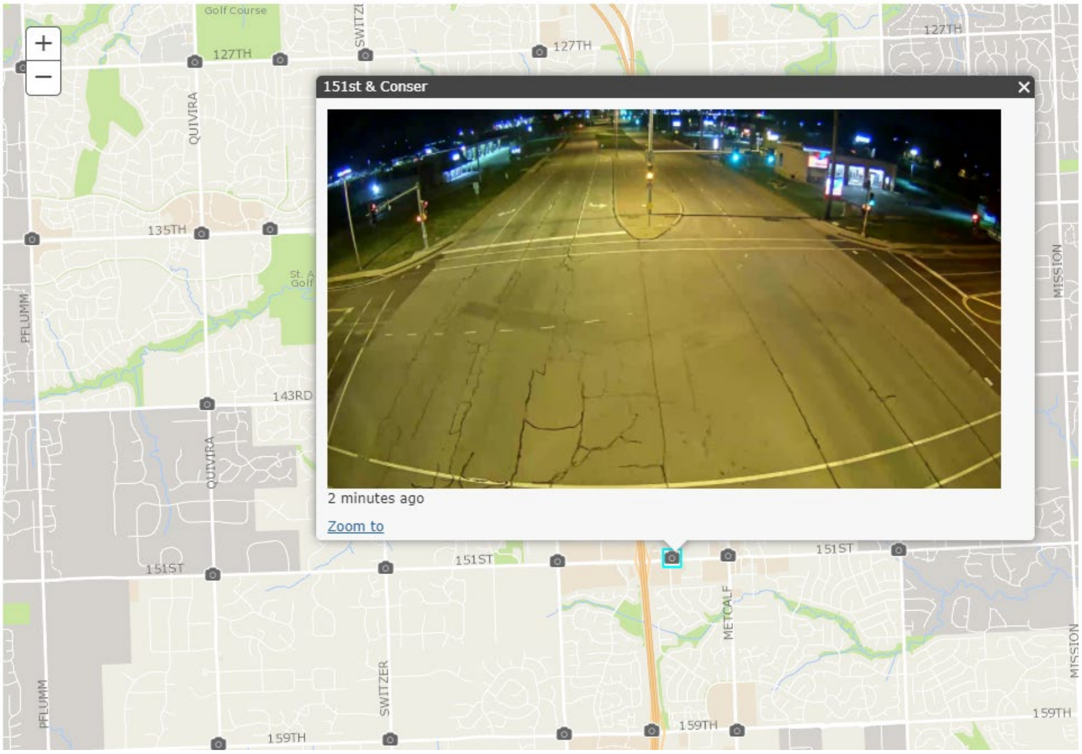

Figure 3.2: Overland Park Traffic Camera Locations with Real-Time Update

Six intersections were initially considered for monitoring video data. Table 3.1 presents each location's traffic and pedestrian volumes and crash intensity.

Table 3.1: Initial Study Locations

Intersection Name	Existing Condition
80th/Metcalf/E	high traffic, high pedestrian, low crash volumes
87th/Farley	medium traffic, high pedestrian, low crash volumes
95th/Nall/N	medium traffic, low pedestrian, medium crash volumes
110th/Quivira/N	high traffic, low pedestrian, medium crash volumes
151st/Conser/W	high traffic, low pedestrian, high crash volumes
College/Roe/S	medium traffic, low pedestrian, medium crash volumes

After consulting with KDOT, one study location with high pedestrian traffic volume and one location with high crash frequency were selected. Location 1 was 80th St. & Metcalf (Figure

3.3), and location 2 was 151st St. & Conser St. (Figure 3.4). For location 1, Metcalf Street has a posted speed limit of 35 mph, the east leg of 80th Street speed limit is 20 mph, and the west leg of 80th Street speed limit is 25 mph. For location 2, 151st Street posted speed limit is 35 mph, and Conser Street posted speed limit is 25 mph.

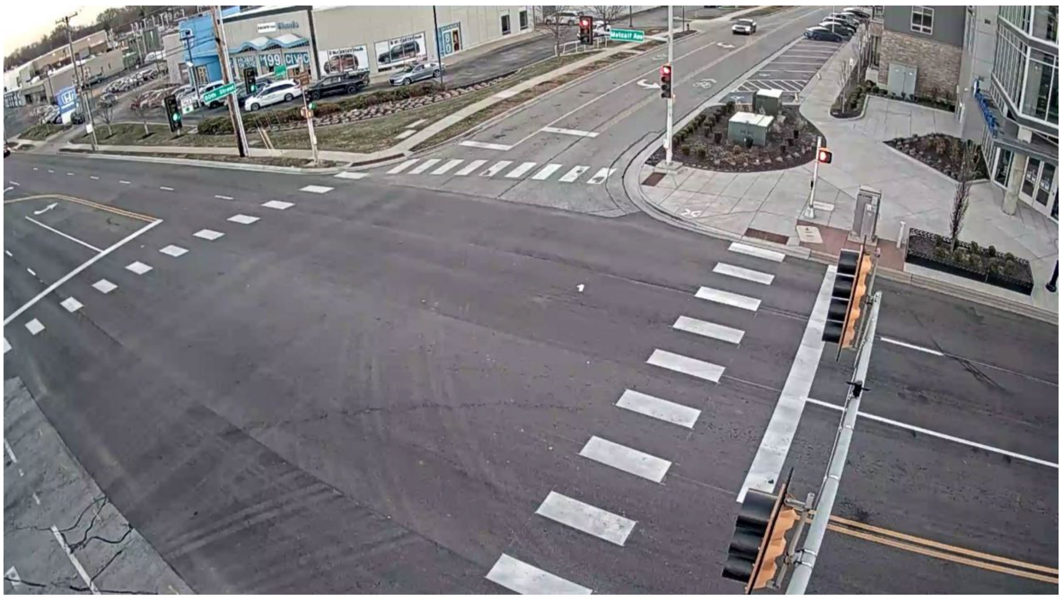

Figure 3.3: Location 1 - Intersection of 80th St. & Metcalf

Figure 3.4: Location 2 - Intersection of 151st St. & Conser St.

Video data were collected for two weeks for both locations. The Public Works Department of Overland Park provided video feeds from February 1–5, 2021 (7:00 a.m. to 7:00 p.m.), while additional data were collected at locations 1 and 2 on March 1–5, 2021, and March 15–19, 2021, respectively. Only weekday data were taken for this study. A thorough intersection study was carried out to determine the AM peak, off-peak (midday) and PM peak for both locations. Table 3.2 presents the data collection time for each location. Notably, the COVID-19 pandemic and resulting work-from-home mandates significantly decreased pedestrian and traffic volume.

Table 3.2: Data Collection Periods

Location 1 - 80th St. & Metcalf		Location 2 - 151st & Conser St.	
AM	8:00–9:00 a.m.	AM	8:00–10:00 a.m.
Off-Peak	11:00 a.m.–2:00 p.m.	Off-Peak	11:00–1:00 p.m.
PM	4:00–6:00 p.m.	PM	4:00–6:00 p.m.

3.3 Transoft Solutions Video Tool

This study utilized TrafxSAFE, the Transoft Solutions safety assessment tool, to analyze the collected video feeds. As shown in Table 3.2, six hours of daily traffic videos were selected for each location, totaling 120 hours of data. The video data were uploaded to the designated portal of the Transoft Solutions website, and then the analyzed data were available to download as a CSV file according to location and date range (Figure 3.5).

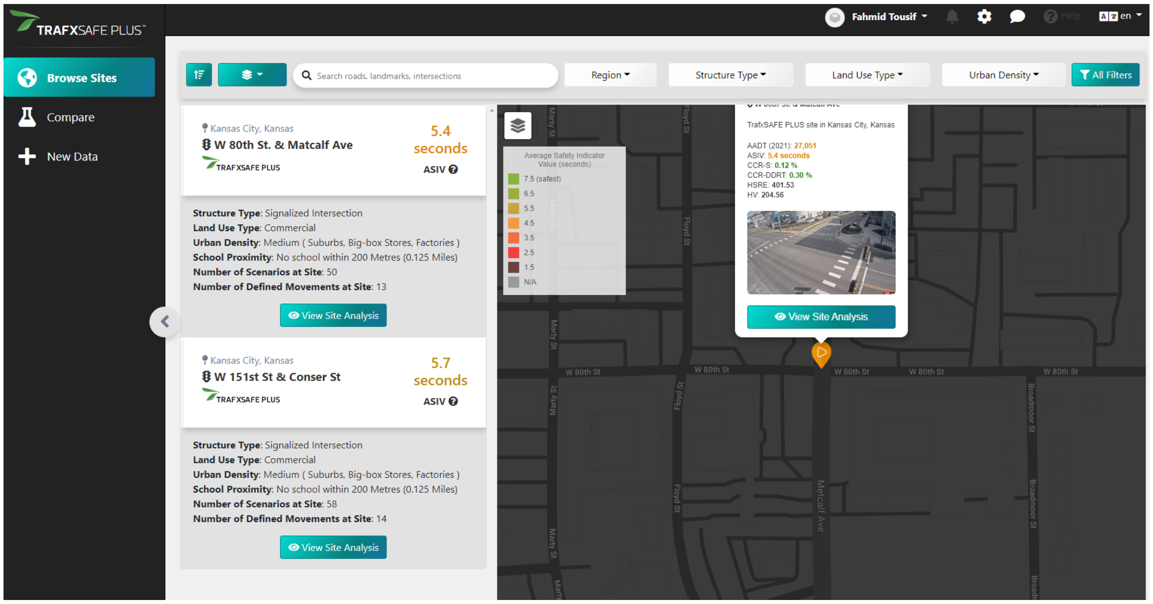

Figure 3.5: Dashboard for Analyzed Data Download Module in Transoft Solutions' Portal

The raw data showed PET and TTC values for all possible conflicts up to 10 seconds, and the CSV file provided date, time, scenario type (e.g., rear-end-following-through versus through or crossing-through versus through-adjacent, etc.), vehicle type, movement, conflict speed, and median speed of road users (Figure 3.6).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Safety Indi	Safety Indi	Date	Time	Arrived Firs	Road User	Road User	Road User	Road User	Road User	Road User	Road User	Road User	Road User	Road User	Scenario Ty	URL
2	2.32	TTC	2/1/2021	16:47:20	-	2.12E+08	Eastbound	Car	16.65274	12.31557	2.12E+08	Eastbound	Car	17.60344	17.88306	Rear-end -	N/A
3	1.62	TTC	2/1/2021	16:00:39	-	2.12E+08	Southboun	Car	22.5247	13.87521	2.12E+08	Southboun	Car	17.01935	16.57818	Rear-end -	https:
4	2.18	TTC	2/1/2021	12:13:36	-	2.12E+08	Southboun	Bus	32.29886	30.6398	2.12E+08	Southboun	Car	26.68167	30.62738	Rear-end -	N/A
5	2.21	TTC	2/1/2021	16:55:09	-	2.12E+08	Southboun	Work Van	26.44555	5.909238	2.12E+08	Southboun	Car	19.39299	11.66313	Rear-end -	N/A
6	2.18	TTC	2/1/2021	16:10:48	-	2.12E+08	Southboun	Car	25.78068	9.289496	2.12E+08	Southboun	Car	17.03799	14.79484	Rear-end -	N/A
7	2.07	TTC	2/1/2021	16:08:51	-	2.12E+08	Southboun	Car	13.79444	16.94479	2.12E+08	Southboun	Car	15.83253	10.07864	Rear-end -	N/A
8	2.2	TTC	2/1/2021	12:25:54	-	2.12E+08	Southboun	Car	16.28613	13.96221	2.12E+08	Southboun	Car	11.21575	17.23062	Rear-end -	N/A
9	2.51	TTC	2/1/2021	12:34:31	-	2.12E+08	Southboun	Articulated	14.08648	15.22359	2.12E+08	Southboun	Car	23.59967	30.42232	Rear-end -	N/A
10	2.28	TTC	2/1/2021	16:25:12	-	2.12E+08	Southboun	Car	5.412141	11.18468	2.12E+08	Southboun	Car	13.74473	11.16604	Rear-end -	N/A
11	2.39	TTC	2/1/2021	16:25:12	-	2.12E+08	Southboun	Car	2.118875	11.96761	2.12E+08	Southboun	Car	8.400936	11.18468	Rear-end -	N/A
12	3.58	TTC	2/1/2021	12:55:08	-	2.12E+08	Southboun	Car	18.83997	11.70663	2.12E+08	Southboun	Car	10.29612	14.61465	Rear-end -	N/A
13	4.22	TTC	2/1/2021	8:22:11	-	2.12E+08	Southboun	Car	18.3118	7.599367	2.12E+08	Southboun	Car	8.183456	6.872363	Rear-end -	N/A
14	2.24	TTC	2/1/2021	12:06:21	-	2.12E+08	Westbounc	Car	26.95507	22.53091	2.12E+08	Southboun	Car	4.33717	8.581134	Crossing - 1	N/A
15	2.28	TTC	2/1/2021	13:57:42	-	2.12E+08	Westbounc	Car	25.48242	24.76785	2.12E+08	Southboun	Car	0	8.040541	Crossing - 1	N/A
16	2.36	TTC	2/1/2021	17:38:44	-	2.12E+08	Westbounc	Car	20.85321	17.57859	2.12E+08	Southboun	Car	1.75848	8.456859	Crossing - 1	N/A
17	2.56	TTC	2/1/2021	13:57:34	-	2.12E+08	Westbounc	Pickup Truc	10.83671	7.03392	2.12E+08	Southboun	Car	2.808597	8.711621	Crossing - 1	N/A
18	3.05	TTC	2/1/2021	8:19:18	-	2.12E+08	Westbounc	Car	20.59845	20.25048	2.12E+08	Southboun	Car	0.018641	9.357847	Crossing - 1	N/A
19	3.04	TTC	2/1/2021	16:33:17	-	2.12E+08	Westbounc	Pickup Truc	15.5902	15.85739	2.12E+08	Southboun	Car	0.031069	9.941936	Crossing - 1	N/A
20	3.15	TTC	2/1/2021	13:00:59	-	2.12E+08	Westbounc	Car	18.25588	18.74676	2.12E+08	Southboun	Car	3.324335	7.785779	Crossing - 1	N/A
21	4.84	TTC	2/1/2021	12:25:14	-	2.12E+08	Westbounc	Car	13.37812	15.78282	2.12E+08	Southboun	Car	1.454008	8.127533	Crossing - 1	N/A
22	2.25	TTC	2/1/2021	8:27:36	-	2.12E+08	Northboun	Car	24.34532	28.1916	2.12E+08	Westbounc	Car	8.605988	9.109299	Crossing - 1	N/A
23	2.72	TTC	2/1/2021	13:56:24	-	2.12E+08	Northboun	Pickup Truc	20.23805	19.72853	2.12E+08	Westbounc	Pickup Truc	14.1859	7.03392	Crossing - 1	N/A
24	2.85	TTC	2/1/2021	17:13:34	-	2.12E+08	Northboun	Car	18.86482	17.75257	2.12E+08	Westbounc	Car	11.20953	8.767545	Crossing - 1	N/A
25	3.25	TTC	2/1/2021	17:34:33	-	2.12E+08	Northboun	Car	18.43608	17.23683	2.12E+08	Westbounc	Car	9.886013	9.886013	Crossing - 1	N/A

Figure 3.6: CSV Raw Data from Transoft Solutions

3.4 Data Sampling

Approximately 2,000 data points, including PET and TTC, were recorded for up to 10 seconds each day for each location. Transoft Solutions classified the PET values into four categories, as shown in Figure 3.7. The first category, Critical Conflicts, included PET values up to 1.99 seconds, the second category, Minor Conflicts, ranged from 2 seconds to 2.99 seconds, and Potential Conflicts, the third category, included PET values from 3 seconds to 4.99 seconds. Any values between 5 seconds and 10 seconds were labeled as Interactions. This study validated Critical, Minor, and Potential Conflicts for PET and TTC safety measures. A maximum of 5 seconds were considered for data validation.

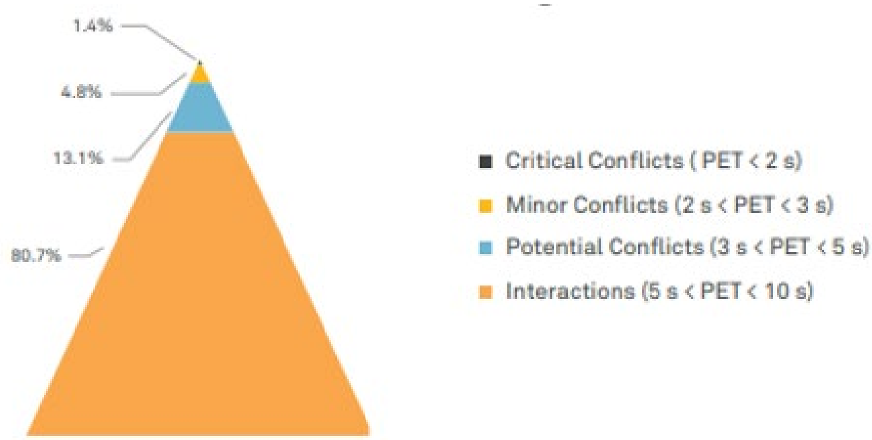

Figure 3.7: PET Conflict Categories
Source: Samara et al. (2020)

Approximately 1,000 data points were recorded each day for location 1 and 600 data points were recorded for location 2, including PET and TTC up to 5 seconds, resulting in approximately 10,000 data points for 10 days at location 1 and 6,000 data points for 10 days at location 2. The selected data were sorted into the three conflict categories discussed previously. Due to the large volume of data, 10% of data from each category were sampled randomly and selected for validation.

3.5 Ground Truth Data

To analyze the data provided by Transoft Solutions, this study collected PET and TTC ground truth data using detailed information provided by Transoft Solutions, including the video URL for measured conflicts up to 2 seconds. Vehicle trajectories and conflict spots were drawn on the computer screen by Transoft Solutions (Figures 3.8 and 3.9). However, due to the large volume of data, the vendor did not provide trajectory information from 2 to 5 seconds. Therefore, vehicle trajectories and conflict spots were manually drawn on the computer screen based on information provided in the CSV file. Manual data were obtained via frame-by-frame analysis of the video file using a media player. Because TTC calculations do not include crossing trajectories, TTC values exceeding 2 seconds were determined by assuming the conflict spots, or presumed locations of the near-miss crashes, to be on the edge of the intersection, as circled in Figures 3.8 and 3.9. The detailed procedure for ground truth extraction is described in Appendix A.2.

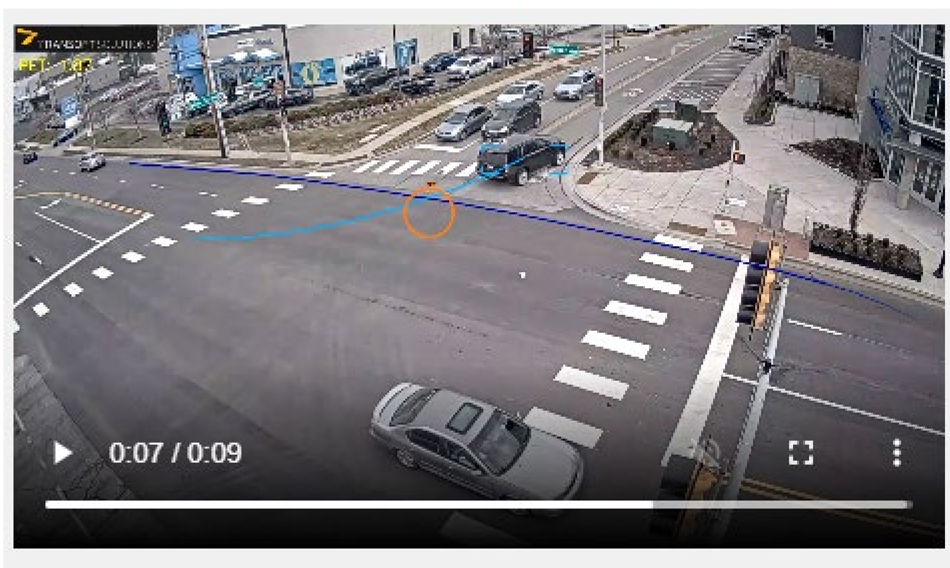

Figure 3.8: Trajectory Provided by Transoft Solutions for Calculating PET

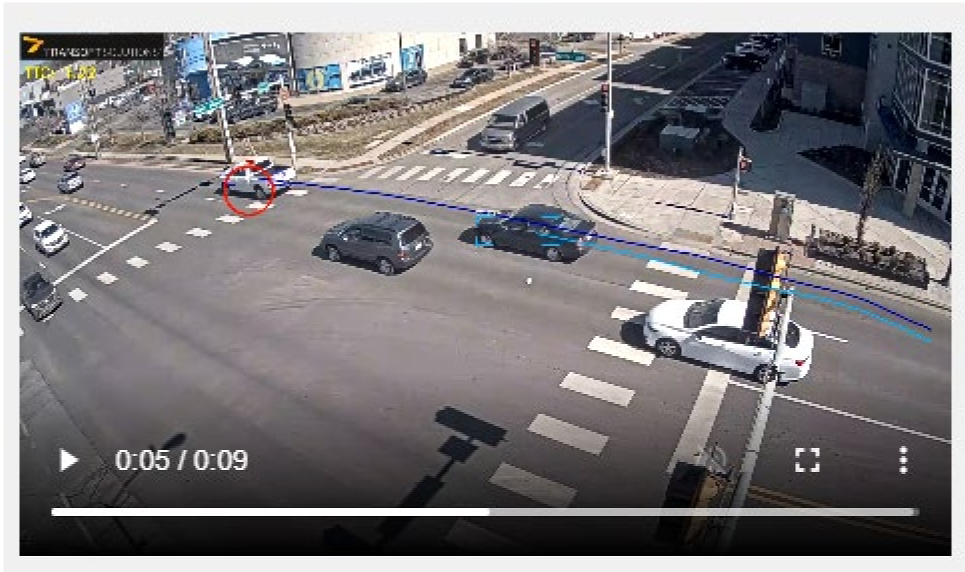

Figure 3.9: Trajectory Provided by Transoft Solutions for Calculating TTC

Chapter 4: Data Analysis and Statistical Comparison

4.1 Comparison Framework

The collected data included observations under various weather conditions (e.g., rain/light snow, cloudy/foggy, and sunny) and traffic conditions (e.g., peak and off-peak periods). Tables 4.1 and 4.2 present the three weather conditions by peak or off-peak period for both locations.

Table 4.1: Weather Conditions in Location 1 (80th St. & Metcalf)

Date	AM Peak	Off-Peak	PM Peak
1-Feb	Sunny	Sunny	Sunny
2-Feb	Sunny	Sunny	Sunny
3-Feb	Sunny	Sunny	Sunny
4-Feb	Rainy	Sunny	Sunny
5-Feb	Sunny	Sunny	Sunny
1-Mar	Sunny	Sunny	Sunny
2-Mar	Sunny	Sunny	Sunny
3-Mar	Sunny	Sunny	Sunny
4-Mar	Sunny	Sunny	Sunny
5-Mar	Sunny	Sunny	Sunny

Table 4.2: Weather Conditions in Location 2 (151st St. & Conser St.)

Date	AM Peak	Off-Peak	PM Peak
1-Feb	Sunny	Sunny	Sunny
2-Feb	Sunny	Sunny	Sunny
3-Feb	Sunny	Sunny	Sunny
4-Feb	Rainy/light snow	Sunny	Sunny
5-Feb	Sunny	Sunny	Sunny
15-Mar	Sunny	Sunny	Rainy
16-Mar	Sunny	Sunny	Sunny
17-Mar	Rainy	Sunny	Sunny
18-Mar	Sunny	Sunny	Sunny
19-Mar	Sunny	Sunny	Sunny

Data analysis focused on predication accuracy for PET and TTC under various weather and traffic conditions at both locations. To ensure a sufficient number of observations was included in all conflict intervals (described in Section 3.4), the data were aggregated to account for only three traffic and weather conditions: rainy peak hour, sunny peak hour, and sunny off-peak hour. Table 4.3 shows the analysis framework for both locations.

Table 4.3: Analysis Outline

Data Categories	Safety Measures	Locations	Weather and Traffic Conditions	No. of Observations
Critical Conflicts (0–2 sec)	PET	Location 1	Rainy Peak Hour	6
			Sunny Peak Hour	35
			Sunny Off-Peak Hour	23
		Location 2	Rainy Peak Hour	10
			Sunny Peak Hour	16
			Sunny Off-Peak Hour	10
	TTC	Location 1	Rainy Peak Hour	0
			Sunny Peak Hour	11
			Sunny Off-Peak Hour	12
		Location 2	Rainy Peak Hour	12
			Sunny Peak Hour	28
			Sunny Off-Peak Hour	11
Minor Conflicts (2–3 sec)	PET	Location 1	Rainy Peak Hour	10
			Sunny Peak Hour	142
			Sunny Off-Peak Hour	119
		Location 2	Rainy Peak Hour	12
			Sunny Peak Hour	72
			Sunny Off-Peak Hour	44
	TTC	Location 1	Rainy Peak Hour	4
			Sunny Peak Hour	30
			Sunny Off-Peak Hour	33
		Location 2	Rainy Peak Hour	10
			Sunny Peak Hour	28
			Sunny Off-Peak Hour	21
Potential Conflicts (3–5 sec)	PET	Location 1	Rainy Peak Hour	11
			Sunny Peak Hour	219
			Sunny Off-Peak Hour	186
		Location 2	Rainy Peak Hour	23
			Sunny Peak Hour	154
			Sunny Off-Peak Hour	74
	TTC	Location 1	Rainy Peak Hour	4
			Sunny Peak Hour	35
			Sunny Off-Peak Hour	30
		Location 2	Rainy Peak Hour	10
			Sunny Peak Hour	19
			Sunny Off-Peak Hour	22

4.2 Comparison Methodology

The predicted value (provided by Transoft Solutions) was compared to the observed value (ground truth) using mean absolute deviation (MAD), root mean squared error (RMSE), mean absolute percentage error (MAPE), and root mean squared log error (RMSLE).

The MAD can be calculated by taking the summation of the absolute differences between the observed and predicted values and dividing it by the number of observations, as shown by:

$$\boldsymbol{MAD} = \frac{\boldsymbol{\Sigma} \ \mid \boldsymbol{Observed\,Value} - \boldsymbol{Predicted\,Value} \mid}{\boldsymbol{Number\ of\ Observations}}$$

Equation 4.1

The RMSE shows the deviation of residual errors and is calculated using the root of the mean squared error (MSE), as shown in Equation 4.2. The MSE is calculated by taking the summation of squared differences between observed value and predicted value and dividing it by the number of observations.

$$\boldsymbol{RMSE} = \sqrt{\frac{\boldsymbol{\Sigma}\ (\boldsymbol{Observed\,Value} - \boldsymbol{Predicted\,Value})\text{^}\mathbf{2}}{\boldsymbol{Number\ of\ Observations}}}$$

Equation 4.2

The MAPE is calculated by taking the summation of absolute percent error for each instance between observed value and predicted value and dividing it by the number of observations, as shown by:

$$\boldsymbol{MAPE}\ (\%) = \frac{\boldsymbol{\Sigma}\ \mid \dfrac{\boldsymbol{Observed\,Value} - \boldsymbol{Predicted\,Value}}{\boldsymbol{Observed\,Value}} \mid}{\boldsymbol{Number\ of\ Observations}} * \mathbf{100}\%$$

Equation 4.3

Finally, although the RMSLE is the RMSE in logarithmic scale (Equation 4.4), it does not explode in magnitude as RMSE does when including an outlier. A common concern with RMSE is its sensitivity towards an outlier (Chai & Draxler, 2014).

$$\boldsymbol{RMSLE} = \sqrt{\frac{\boldsymbol{\Sigma}\ [\boldsymbol{log}(\boldsymbol{Observed\,Value} + \mathbf{1}) - \boldsymbol{log}(\boldsymbol{Predicted\,Value} + \mathbf{1})]\text{^}\mathbf{2}}{\boldsymbol{Number\ of\ Observations}}}$$

Equation 4.4

4.3 Summary of Results by Location

4.3.1 Critical Conflicts (0–2 sec)

As shown in Table 4.4, the rainy peak hours for PET MAD show a high value of 0.49 and low values for sunny peak hours and sunny off-peak hours for location 1, indicating that, in rainy weather, the PET values predicted by Transoft Solutions differ more significantly from the ground truth than in sunny weather. As shown in the table, the MAPE values are 26.24% and 19.59%, respectively, for sunny peak and off-peak hours, and 38.28% for rainy peak hours. This observation is consistent with the RMSE and RMSLE values in location 1, proving that Transoft Solutions more accurately predicts PET values in sunny weather, potentially due to camera vision obstruction as a result of raindrops. Although the trend for location 2 was similar to location 1, all four statistical measures showed lower values than location 1 in all three conditions, indicating that the algorithm predicted PET values in location 2 more accurately than location 1.

Table 4.4: Comparison of Statistical Measures for Critical Conflicts (PET)

Statistical Measures	PET		
	Loc 1: Rainy Peak	Loc 1: Sunny Peak	Loc 1: Sunny Off-Peak
MAD	0.49	0.36	0.28
RMSE	0.50	0.39	0.32
MAPE (%)	38.28	26.24	19.59
RMSLE	0.09	0.06	0.05
	Loc 2: Rainy Peak	Loc 2: Sunny Peak	Loc 2: Sunny Off-Peak
MAD	0.30	0.26	0.18
RMSE	0.32	0.28	0.21
MAPE (%)	22.48	17.73	12.22
RMSLE	0.06	0.05	0.04

Table 4.5 shows the statistical measures for critical conflicts for TTC. In the table, not applicable (N/A) is used to report that this study did not observe any rain under peak hours at location 1. As shown, the MAD at location 1 is 0.36 and 0.28 for sunny peak and off-peak hours, respectively, and the RMSE values are 0.4 and 0.32 for those same weather conditions. RMSLE

showed a similar trend. The MAPE values indicate that the average difference between the forecasted value and the actual value in sunny peak hours is higher (31.35%) than in sunny off-peak hours (19.59%). Overall, all four metrics showed lower values for sunny off-peak hours than sunny peak hours potentially due to a decreased number of vehicles in off-peak hours. For location 2, the data showed a similar trend for TTC as was shown for PET. Overall, all four statistical measures had lower values for location 2. In terms of weather conditions, MAD decreased from 0.23 to 0.19 when comparing sunny and rainy weather, respectively. This trend was also observed for RMSE, MAPE, and RMSLE. Overall, sunny peak hours showed increased accuracy over rainy peak hours, and sunny off-peak hours showed a slightly increased accuracy over sunny peak hours for all four statistical measures. As with PET, Transoft Solutions most accurately predicted TTC at location 2 in sunny weather conditions.

Table 4.5: Comparison of Statistical Measures for Critical Conflicts (TTC)

Statistical Measures	TTC		
	Loc 1: Rainy Peak	Loc 1: Sunny Peak	Loc 1: Sunny Off-Peak
MAD	N/A	0.36	0.28
RMSE	N/A	0.40	0.32
MAPE (%)	N/A	31.35	19.59
RMSLE	N/A	0.07	0.05
	Loc 2: Rainy Peak	Loc 2: Sunny Peak	Loc 1: Sunny Off-Peak
MAD	0.23	0.19	0.16
RMSE	0.27	0.25	0.20
MAPE (%)	15.33	13.58	10.73
RMSLE	0.04	0.04	0.03

4.3.2 Minor Conflicts (2–3 sec)

For minor conflicts, all four statistical measures showed smaller differences in sunny conditions than in rainy weather conditions (Table 4.6). The MAD and RMSE values were approximately 0.4 for location 1 and approximately 0.3 for location 2, indicating that Transoft

Solutions more accurately predicted PET values at location 2. This observation is consistent with the MAPE and RMSLE values.

Table 4.6: Comparison of Statistical Measures for Minor Conflicts (PET)

Statistical Measures	PET		
	Loc 1: Rainy Peak	Loc 1: Sunny Peak	Loc 1: Sunny Off-Peak
MAD	0.41	0.39	0.35
RMSE	0.47	0.43	0.39
MAPE (%)	21.54	19.91	17.23
RMSLE	0.06	0.06	0.05
	Loc 2: Rainy Peak	Loc 2: Sunny Peak	Loc 2: Sunny Off-Peak
MAD	0.31	0.25	0.23
RMSE	0.32	0.28	0.26
MAPE (%)	14.37	11.87	10.48
RMSLE	0.04	0.04	0.03

For TTC, sunny weather conditions (peak and off-peak) had lower statistics than rainy conditions. The MAD value in location 1 was 0.47 for rainy peak hours and then decreased to 0.39 for sunny peak hours and 0.37 for sunny off-peak hours. For location 2, MAD values decreased to 0.37, 0.21, and 0.20 for each weather condition, respectively, demonstrating increased prediction accuracy in location 2. The MAPE values for the three weather conditions show that the average difference between the forecasted value and the actual value was approximately 20% for location 1 and 15% for location 2.

Table 4.7: Comparison of Statistical Measures for Minor Conflicts (TTC)

Statistical Measures	TTC		
	Loc 1: Rainy Peak	Loc 1: Sunny Peak	Loc 1: Sunny Off-Peak
MAD	0.47	0.39	0.37
RMSE	0.47	0.47	0.45
MAPE (%)	21.86	18.88	18.18
RMSLE	0.06	0.06	0.06
	Loc 2: Rainy Peak	Loc 2: Sunny Peak	Loc 2: Sunny Off-Peak
MAD	0.37	0.21	0.20
RMSE	0.43	0.27	0.25
MAPE (%)	19.32	9.76	8.36
RMSLE	0.06	0.04	0.03

4.3.3 Potential Conflicts (3–5 sec)

Table 4.8 shows the PET-related statistical measures for potential conflicts. As shown, MAD has similar results across all three conditions, while the MAPE demonstrates slightly increased accuracy during sunny weather conditions. However, in terms of RMSE and RMSLE, rainy conditions were predicted with slightly higher accuracy, potentially due to the smaller sample size for rainy weather conditions compared to the sample size for sunny weather conditions.

Table 4.8: Comparison of Statistical Measures for Potential Conflicts (PET)

Statistical Measures	PET		
	Loc 1: Rainy Peak	Loc 1: Sunny Peak	Loc 1: Sunny Off-Peak
MAD	0.29	0.29	0.29
RMSE	0.33	0.36	0.35
MAPE (%)	8.57	8.04	7.97
RMSLE	0.03	0.03	0.03
	Loc 2: Rainy Peak	Loc 2: Sunny Peak	Loc 2: Sunny Off-Peak
MAD	0.26	0.24	0.23
RMSE	0.29	0.29	0.26
MAPE (%)	7.28	6.63	6.30
RMSLE	0.03	0.03	0.02

The results for TTC were nearly similar to PET (Table 4.9). In location 1, sunny weather conditions showed slightly increased accuracy over rainy weather conditions in terms of MAD, RMSE, and RMSLE. However, MAPE in sunny peak conditions was slightly higher (7.45%) than in rainy peak condition (7.27%). For location 2, rainy weather conditions showed slightly increased accuracy than sunny weather conditions in terms of MAD, RMSE, and RMSLE. Unlike location 1, MAPE was slightly higher in rainy weather (3.62%) than in sunny peak conditions (3.55%). However, MAPE in sunny off-peak conditions (3.55%) was slightly higher than in sunny peak conditions (3.73%).

Table 4.9: Comparison of Statistical Measures for Potential Conflicts (TTC)

Statistical Measures	TTC		
	Loc 1: Rainy Peak	Loc 1: Sunny Peak	Loc 1: Sunny Off-Peak
MAD	0.27	0.25	0.20
RMSE	0.31	0.30	0.25
MAPE (%)	7.27	7.45	5.44
RMSLE	0.03	0.03	0.02
	Loc 2: Rainy Peak	Loc 2: Sunny Peak	Loc 2: Sunny Off-Peak
MAD	0.13	0.14	0.13
RMSE	0.15	0.15	0.15
MAPE (%)	3.62	3.55	3.73
RMSLE	0.01	0.01	0.01

4.4 Summary of Results for Both Locations

4.4.1 Critical Conflicts (0–2 sec)

Because there were a limited number of observations during rainy peak hours for location 1, this study merged and analyzed observations from both locations. PET was more accurately predicted during sunny weather (peak and off-peak) than rainy weather based on all four statistical measures (Table 4.10). In addition, PET was more accurately predicted during sunny off-peak conditions than sunny peak conditions. However, for TTC, rainy weather demonstrated slightly increased accuracy across all four metrics, although when analyzed separately for each location, sunny weather showed smaller differences between measured and observed values.

Table 4.10: Comparison of Statistical Measures for Critical Conflicts (Both Locations: PET & TTC)

Statistical Measures	PET		
	Rainy Peak	Sunny Peak	Sunny Off-Peak
MAD	0.37	0.33	0.25
RMSE	0.40	0.36	0.29
MAPE (%)	28.40	23.57	17.35
RMSLE	0.07	0.06	0.05
	TTC		
MAD	0.23	0.24	0.25
RMSE	0.27	0.30	0.29
MAPE (%)	15.33	18.59	17.35
RMSLE	0.04	0.05	0.05

4.4.2 Minor Conflicts (2–3 sec)

Table 4.11 shows the results for minor conflicts at locations 1 and 2. For PET, sunny weather showed slightly increased accuracy over rainy weather based on all four statistical measures. As with critical conflicts, minor conflicts also demonstrated increased accuracy during sunny off-peak hours compared to sunny peak hours. The trend for TTC was identical to PET, with low values estimated for sunny weather.

Table 4.11: Comparison of Statistical Measures for Minor Conflicts (Both Locations: PET & TTC)

Statistical Measures	PET		
	Rainy Peak	Sunny Peak	Sunny Off-Peak
MAD	0.36	0.35	0.32
RMSE	0.39	0.39	0.36
MAPE (%)	17.75	17.21	15.41
RMSLE	0.05	0.05	0.05
	TTC		
MAD	0.40	0.31	0.31
RMSE	0.44	0.38	0.39
MAPE (%)	20.03	14.48	14.36
RMSLE	0.06	0.05	0.05

4.4.3 Potential Conflicts (3–5 sec)

For potential conflicts (Table 4.12), the MAD and RMSLE resulted in values similar to PET. Although RMSE during rainy weather was slightly lower than sunny weather, MAPE was higher for rainy weather. For TTC, rainy peak conditions and sunny off-peak conditions showed higher prediction accuracy than sunny peak conditions.

Table 4.12: Comparison of Statistical Measures for Potential Conflicts (Both Locations: PET & TTC)

Statistical Measures	PET		
	Rainy Peak	Sunny Peak	Sunny Off-Peak
MAD	0.27	0.27	0.27
RMSE	0.30	0.33	0.33
MAPE (%)	7.70	7.46	7.50
RMSLE	0.03	0.03	0.03
	TTC		
MAD	0.17	0.21	0.17
RMSE	0.21	0.26	0.21
MAPE (%)	4.66	6.08	4.72
RMSLE	0.02	0.02	0.02

4.5 Statistical Analysis: Analysis of Variance

Statistical analysis was carried out to investigate differences in the means for PET and TTC values with respect to location and weather conditions. A one-way analysis of variance (ANOVA) was selected for this purpose. The null hypothesis was that the means of different groups are equal, suggesting no significant difference among selected groups. The alternative hypothesis was that at least two group means are statistically significantly different.

Null Hypothesis: H_0: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_N$

Alternative Hypothesis: H_1: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \ldots \neq \mu_N$

The desired confidence level was 95%, meaning the significance level α was 0.05. The null hypothesis is rejected when the test statistic is less than the significance level. ANOVA only shows significant differences among means, and it does not identify means differences. Therefore, Tukey's Honest Significant Difference (HSD) post-hoc test was conducted with ANOVA for this study. Tukey's HSD identifies individual differences between means by assigning an English letter to each group. Means that do not share the same letter are significantly different.

This study organized the data by conflict category (critical conflicts, minor conflicts, and potential conflicts), location (1 or 2), and weather and traffic conditions (sunny peak, sunny off-peak, rainy peak). Abbreviations used for individual locations are presented in Table 4.13. In addition, this study combined the data from both locations to infer any statistical observations in terms of various weather and traffic conditions. Relevant abbreviations are presented in Table 4.14.

Table 4.13: Abbreviations Used in ANOVA and Tukey's HSD for Each Location

Abbreviations	Full Form
L1_RP_TS	Location 1: Rainy Peak measured by Transoft Solutions
L1_RP_OB	Location 1: Rainy Peak measured from observation (Ground Truth)
L1_SP_TS	Location 1: Sunny Peak measured by Transoft Solutions
L1_SP_OB	Location 1: Sunny Peak measured from observation (Ground Truth)
L1_SOP_TS	Location 1: Sunny Off-Peak measured by Transoft Solutions
L1_SOP_OB	Location 1: Sunny Off-Peak measured from observation (Ground Truth)
L2_RP_TS	Location 2: Rainy Peak measured by Transoft Solutions
L2_RP_OB	Location 2: Rainy Peak measured from observation (Ground Truth)
L2_SP_TS	Location 2: Sunny Peak measured by Transoft Solutions
L2_SP_OB	Location 2: Sunny Peak measured from observation (Ground Truth)
L2_SOP_TS	Location 2: Sunny Off-Peak measured by Transoft Solutions
L2_SOP_OB	Location 2: Sunny Off-Peak measured from observation (Ground Truth)

Table 4.14: Abbreviations Used in ANOVA and Tukey's HSD for Both Locations

Abbreviations	Full Form
RP_TS	Rainy Peak measured by Transoft Solutions
RP_OB	Rainy Peak measured from observation (Ground Truth)
SP_TS	Sunny Peak measured by Transoft Solutions
SP_OB	Sunny Peak measured from observation (Ground Truth)
SOP_TS	Sunny Off-Peak measured by Transoft Solutions
SOP_OB	Sunny Off-Peak measured from observation (Ground Truth)

4.5.1 Critical Conflicts (0–2 sec) for Individual Locations

Table 4.15 shows results from the ANOVA and post-hoc Tukey's HSD test for critical conflicts (0–2 sec). No significant difference was observed between all mean PET values provided by Transoft Solutions (predicted values) in both locations and for all three weather and traffic conditions (rainy peak, sunny off-peak, and sunny peak), as denoted by the letter A. Similarly, the means of observed values (ground truths) for the groups did not vary significantly, as designated with the letter C in the table. However, the observed values of location 1 (80th St. & Metcalf) differed significantly from the predicted values for all three weather and traffic conditions, and the observed values of location 2 (151st St. & Conser St.) differed significantly from the predicted values for rainy peak and sunny peak conditions, although they share a common letter B, indicating no significant difference between these two groups for sunny off-peak conditions.

Table 4.15: ANOVA of Critical Conflicts for Individual Locations (PET)

Factor	N	Mean	Grouping		
L1_RP_TS	6	1.802	A		
L1_SP_TS	35	1.791	A		
L1_SOP_TS	23	1.782	A		
L2_SP_TS	16	1.764	A		
L2_RP_TS	10	1.683	A	B	
L2_SOP_TS	10	1.660	A	B	
L2_SP_OB	16	1.503		B	C
L1_SOP_OB	23	1.497		B	C
L2_SOP_OB	10	1.495		B	C
L1_SP_OB	35	1.429			C
L2_RP_OB	10	1.384			C
L1_RP_OB	6	1.310			C

No data were available for rainy peak conditions at location 1 for TTC, so the overall conclusions for TTC were consistent with PET (Table 4.16). No significant difference was observed among the means of TTC measured by Transoft Solutions for all locations and weather and traffic conditions. In the table, all observed values, regardless of location, are marked with the letter C, indicating no difference, except for sunny peak hour. The observed mean of sunny peak hours for location 1 varied significantly from the observed mean of sunny peak hours for location 2. In location 1, the observed values differed significantly from the predicted values for sunny peak conditions, but no significant difference was observed between the observed and predicted values for off-peak conditions. In location 2, no significant difference in TTC values was found between the observed and predicted values for all three weather conditions.

Table 4.16: ANOVA of Critical Conflicts for Individual Locations (TTC)

Factor	N	Mean	Grouping			
L2_RP_TS	12	1.740	A			
L2_SP_TS	28	1.544	A	B		
L2_SOP_TS	11	1.542	A	B		
L1_SP_TS	11	1.536	A	B		
L2_RP_OB	12	1.520	A	B	C	
L1_SOP_TS	12	1.509	A	B	C	
L2_SOP_OB	11	1.503	A	B	C	
L2_SP_OB	28	1.449		B	C	
L1_SOP_OB	12	1.234			C	D
L1_SP_OB	11	1.179				D

4.5.2 Minor Conflicts (2–3 sec) for Individual Locations

For minor conflicts (2–3 sec), no significant difference was observed among all the predicted values (Table 4.17). For location 1, the observed values differed significantly from the predicted values for rainy peak, sunny peak, and sunny off-peak conditions. Although no significant difference was found for rainy peak conditions, significant differences were found for sunny peak and sunny off-peak conditions for location 2.

Table 4.17: ANOVA of Minor Conflicts for Individual Locations (PET)

Factor	N	Mean	Grouping					
L1_SOP_TS	119	2.486	A					
L1_SP_TS	142	2.478	A					
L2_RP_TS	12	2.477	A	B	C	D		
L2_SOP_TS	44	2.475	A					
L2_SP_TS	72	2.460	A	B				
L1_RP_TS	10	2.454	A	B	C	D		
L2_SOP_OB	44	2.260				D	E	
L2_SP_OB	72	2.212			C	D	E	F
L2_RP_OB	12	2.172		B	C	D	E	F
L1_SOP_OB	119	2.145					E	F
L1_SP_OB	142	2.083						F
L1_RP_OB	10	2.026					E	F

The observed and predicted values for TTC are presented in Table 4.18. For location 1, significant difference was found between observed values and Transoft Solutions values for sunny peak and sunny off-peak conditions, while no significant difference was found for observed and predicted values for all three weather conditions at location 2.

Table 4.18: ANOVA of Minor Conflicts for Individual Locations (TTC)

Factor	N	Mean	Grouping										
L1_RP_TS	4	2.625	A	B	C	D	E	F	G	H	I	J	K
L1_SOP_TS	33	2.572	A										
L2_SOP_TS	21	2.487	A	B		D		F			I		
L1_SP_TS	30	2.476	A	B		D		F			I		
L2_RP_TS	10	2.368	A	B	C	D	E	F	G	H	I	J	K
L2_SOP_OB	21	2.358	A	B	C	D	E	F	G	H	I	J	K
L2_SP_TS	28	2.309									I	J	K
L2_SP_OB	28	2.243						F	G	H	I	J	K
L1_SOP_OB	33	2.222				D	E	F		H	I		K
L1_SP_OB	30	2.193			C		E		G	H		J	K
L1_RP_OB	4	2.159	A	B	C	D	E	F	G	H	I	J	K
L2_RP_OB	10	2.135		B	C	D	E	F	G	H	I	J	K

4.5.3 Potential Conflicts (3–5 sec) for Individual Locations

As shown in Table 4.19, the PET values for potential conflicts (3–5 sec) differed significantly in sunny peak and off-peak conditions (location 1), but no difference was observed for rainy peak conditions due to the small sample size. For location 2, significant differences in predicted and observed values were found only in sunny peak conditions, while no significant differences were found for sunny off-peak and rainy peak conditions. However, no significant difference was observed for TTC across all weather and traffic conditions (Table 4.20).

Table 4.19: ANOVA of Potential Conflicts for Individual Locations (PET)

Factor	N	Mean	Grouping		
L2_SOP_TS	74	4.035	A	B	
L2_SP_TS	154	4.029	A		
L1_SP_TS	219	4.026	A		
L1_SOP_TS	186	4.018	A		
L1_RP_TS	11	3.892	A	B	C
L2_RP_TS	23	3.868	A	B	C
L2_SOP_OB	74	3.824	A	B	C
L2_SP_OB	154	3.805		B	C
L1_SP_OB	219	3.765			C
L1_SOP_OB	186	3.745			C
L2_RP_OB	23	3.633	A	B	C
L1_RP_OB	11	3.613	A	B	C

Table 4.20: ANOVA of Potential Conflicts for Individual Locations (TTC)

Factor	N	Mean	Grouping
L1_RP_OB	4	3.818	A
L2_SP_TS	19	3.811	A
L2_SP_OB	19	3.785	A
L1_SOP_TS	30	3.774	A
L2_RP_OB	10	3.750	A
L1_SP_TS	35	3.748	A
L2_RP_TS	10	3.735	A
L1_RP_TS	4	3.713	A
L2_SOP_TS	22	3.646	A
L1_SOP_OB	30	3.623	A
L2_SOP_OB	22	3.598	A
L1_SP_OB	35	3.550	A

4.5.4 Critical Conflicts (0–2 sec) for Both Locations

Due to the decreased number of critical conflicts in rainy peak hour observations in location 1, data from both locations were merged and analyzed. When merged, all the predicted values (measured by Transoft Solutions) were designated with the letter A, while all the observed values were given the letter B (Table 4.21). This indicates that the observed values differed significantly from the predicted values for all three weather and traffic conditions. As shown in Table 4.22, significant differences for TTC were found only in sunny peak conditions, but no significant difference was observed for sunny off-peak and rainy peak conditions.

Table 4.21: ANOVA of Critical Conflicts for Both Locations (PET)

Factor	N	Mean	Grouping	
SP_TS	51	1.782	A	
SOP_TS	33	1.745	A	
RP_TS	16	1.728	A	
SOP_OB	33	1.496		B
SP_OB	51	1.452		B
RP_OB	16	1.356		B

Table 4.22: ANOVA of Critical Conflicts for Both Locations (TTC)

Factor	N	Mean	Grouping	
RP_TS	12	1.740	A	
SP_TS	39	1.542	A	
SOP_TS	23	1.525	A	B
RP_OB	12	1.520	A	B
SP_OB	39	1.373		B
SOP_OB	23	1.362		B

4.5.5 Minor Conflicts (2–3 sec) for Both Locations

Similar results to critical conflicts were observed for minor conflicts when calculating PET values (Table 4.23). Although the observed values differed significantly from the predicted values for all three weather and traffic conditions, significant differences were found in sunny peak and off-peak conditions for TTC (Table 4.24). No significant difference was observed for rainy peak conditions.

Table 4.23: ANOVA of Minor Conflicts for Both Locations (PET)

Factor	N	Mean	Grouping	
SOP_TS	163	2.483	A	
SP_TS	214	2.472	A	
RP_TS	22	2.466	A	
SOP_OB	163	2.176		B
SP_OB	214	2.126		B
RP_OB	22	2.106		B

Table 4.24: ANOVA of Minor Conflicts for Both Locations (TTC)

Factor	N	Mean	Grouping		
SOP_TS	54	2.539	A		
RP_TS	14	2.441	A	B	C
SP_TS	58	2.395	A	B	
SOP_OB	54	2.275		B	C
SP_OB	58	2.217			C
RP_OB	14	2.142		B	C

4.5.6 Potential Conflicts (3–5 sec) for Both Locations

Regarding potential conflicts, significant differences in PET values were found between sunny peak and off-peak conditions (Table 4.25), but no significant difference was observed for rainy peak conditions. For TTC, no significant difference was found between observed and predicted values for all three weather and traffic conditions (Table 4.26).

Table 4.25: ANOVA of Potential Conflicts for Both Locations (PET)

Factor	N	Mean	Grouping	
SP_TS	373	4.027	A	
SOP_TS	260	4.023	A	
RP_TS	34	3.876	A	B
SP_OB	373	3.780		B
SOP_OB	260	3.767		B
RP_OB	34	3.627		B

Table 4.26: ANOVA of Potential Conflicts for Both Locations (TTC)

Factor	N	Mean	Grouping
SP_TS	54	3.770	A
RP_OB	14	3.769	A
RP_TS	14	3.729	A
SOP_TS	52	3.720	A
SP_OB	54	3.633	A
SOP_OB	52	3.613	A

4.5.7 Summary of Statistical Analysis

In summary, sunny weather conditions (peak and off-peak) resulted in more accurate predictions of PET and TTC values than rainy weather conditions for all performance measures, potentially due to vision obstruction as a result of raindrops on the camera lenses. Sunny off-peak hours generally demonstrated better accuracy than sunny peak hours. When traffic is low, the algorithm seems to work better. In addition, location 2 (151st St. & Conser St.) had better prediction accuracy than location 1 (80th St. & Metcalf) due to the camera field of view and roadway geometry. Overall, the RMSE values were within 0.1–0.5, indicating moderate predictability.

Based on the statistical analysis, however, the predicted values differed significantly from the observed values of PET in most cases, although the predicted and observed values of TTC did not differ significantly. When the data were analyzed separately for both locations, observed PET values for critical conflicts (0–2 sec) differed significantly from predicted PET values at location 1 for all three weather and traffic conditions, but the sunny off-peak condition did not show any significant difference for TTC. For location 2, the sunny off-peak condition did not show any significant difference between the observed PET values and predicted PET values, and no significant difference was observed for TTC for all three weather and traffic conditions. When the data were merged for both locations, significant differences were noticed for PET values for all three weather and traffic conditions, while a significant difference was noticed for TTC values only in sunny peak conditions. For potential conflict, observed PET values differed significantly from the predicted PET values for all three weather and traffic conditions. However, significant differences between observed and predicted TTC values were only found for sunny peak and off-peak conditions.

Chapter 5: Conclusions and Recommendations

Near-miss crash detection technology has immense potential to contribute significantly to improved traffic safety. Transoft Solutions is one of the leading vendors of safety assessment tools for near-miss crash prediction. This research project attempted to validate the Transoft Solutions algorithm and predict values of two surrogate conflict measures (TTC and PET) with manually measured observed values. Two weeks of video data were collected and analyzed from two intersections in Overland Park, Kansas, for a total of 60 hours of weekday (Monday–Friday) traffic data at each location.

The Transoft Solutions tool provided PET and TTC data for each location up to 10 seconds. Data points up to 5 seconds were sorted for analysis, meaning approximately 1,000 data points were recorded daily for location 1 and 600 data points were recorded for location 2, including PET and TTC values. Approximately 10 percent of the data were sampled for manual validation. Data were sorted based on conflict measures (PET and TTC) and conflict category (critical conflicts, minor conflicts, and potential conflicts). Manual observations (ground truth) were made by drawing estimated vehicle trajectories on the computer screen and measuring the time lapse in milliseconds in frame-by-frame analysis. The sorted data were further categorized based on three weather and traffic conditions (rainy peak hours, sunny peak hours, and sunny off-peak hours) during the two weeks of data collection.

Two types of analysis were conducted: a statistical analysis that considered four measures (MAD, RMSE, MAPE, and RMSLE) and an ANOVA. The following conclusions were drawn after the statistical analysis with the four performance measures:

- Transoft Solutions' predictions of PET and TTC values were more accurate in sunny weather than in rainy weather conditions. For critical, minor, and potential conflicts, errors observed for all four performance measures (MAD, RMSE, MAPE, and RMSLE) were higher for rainy peak conditions than sunny peak conditions, potentially due to vision obstruction on the camera's field of view. These observations were consistent for both locations.

- Transoft Solutions' predictions of PET and TTC values were slightly better in sunny off-peak conditions than sunny peak conditions. For critical, minor, and potential conflicts, errors observed for all four performance measures (MAD, RMSE, MAPE, and RMSLE) were slightly higher for sunny peak conditions than sunny off-peak conditions. Traffic volume could have affected the accuracy of the PET and TTC predictions. These observations were consistent for both locations.
- Transoft Solutions' predictions of PET and TTC values were slightly better in location 2 (151st St. & Conser St.) than in location 1 (80th St. & Metcalf). For critical, minor, and potential conflicts, errors observed for location 2 were slightly less than errors observed for location 1 for all four performance measures. This difference could be attributed to the roadway geometry or camera position since location 2 has wider road geometry (more lanes to facilitate left-turning and right-turning vehicles) and cameras with wider fields of view.
- Overall, the TTC and PET of potential conflicts (3–5 sec) showed a lower error rate than for minor (2–3 sec) and critical conflicts (0–2 sec), potentially due to increased precision with increasing PET or TTC time interval for detecting near-miss conflicts.

The goal of the ANOVA and Tukey's post-hoc test was to identify statistically significant differences between the means of the observed and predicted values. The following conclusions were drawn from this analysis:

- Observed values for PET differed significantly from predicted values in location 1 (80th St. & Metcalf) for all three conflict categories (critical conflicts, minor conflicts, and potential conflicts) except for rainy peak conditions of potential conflicts.
- In location 2 (151st St. & Conser St.), for PET, significant differences were found except for sunny peak condition for critical conflicts, rainy

peak condition for minor conflicts and sunny off-peak and rainy peak conditions for potential conflicts.

- When the PET data were merged for both locations, the observed values differed significantly from the predicted values for critical and minor conflicts. For potential conflicts, significant differences between means of observed and predicted values were found for sunny peak and off-peak conditions. No statistically significant difference was found for potential conflicts in rainy peak conditions.
- For TTC, location 1 (80th St. & Metcalf) observed and predicted values differed significantly in sunny peak conditions for critical conflicts and sunny peak and off-peak conditions for minor conflicts. No significant differences were found in sunny off-peak conditions for critical conflicts, rainy peak conditions for minor conflicts, and all three conditions (rainy peak, sunny peak, and sunny off-peak) for potential conflicts.
- In location 2 (151st St. & Conser St.), no significant difference among the means was found for TTC in all three weather and traffic conditions and for all three conflict categories.
- When the data were merged, significant differences were found in sunny peak conditions for critical conflicts, as well as sunny peak and off-peak conditions for minor conflicts. However, no significant differences were found in rainy peak and sunny off-peak conditions for critical conflicts, rainy peak conditions for minor conflicts, and all three conditions (rainy peak, sunny peak, and sunny off-peak) for potential conflicts.

In general, the RMSE ranged from 0.15 to 0.5 for both PET and TTC. The statistical analysis measure of MAD showed a deviation from 0.15 to 0.5 for both PET and TTC, the MAPE measure showed a range of 3.55–38.28%, and the RMSLE demonstrated a range from 0.01 to 0.09. Although high values were notably observed in rainy peak conditions in location 1, the very small sample size (less than 30 observations) in rainy peak conditions may have inaccurately reflected representative conditions, meaning additional data during rainy peak should be collected in future research. Overall, the conclusion was made that the Transoft Solutions tool offers moderate prediction accuracy since the ground truth was generally lower than the predicted values. Statistical analysis showed that the means of observed values were significantly different from the predicted values in most cases for PET, but many instances showed no significant difference between observed and predicted values for TTC, potentially due to relatively fewer TTC observations.

General findings of this study resulted in the following recommendations:

- This research should be expanded with more data from different weather conditions. Although the data were collected for this research in February and March 2021, additional data should be collected during precipitation seasons, preferably late fall and late spring. The weather conditions need to be expanded to include snow and rainy and dry conditions.
- This research took place during the COVID-19 pandemic, when traffic conditions had not returned to pre-pandemic levels. Therefore, comparatively less traffic was observed. The pedestrian volume was also very low due to work-from-home mandates.
- The manual observation showed comparatively lower values than the predicted values due to a known limitation in the reliable, consistent estimation of the 3D volume of objects (finding the front and rear vehicle bumper) across a wide array of cameras and camera angles and road user types. This limitation is a systematic error that affected all results equally and consistently across deployments. Transoft Solutions' improved algorithm to address this overfitting issue should be made available in their next update.
- Though the manually measured values were less than the predicted values in most cases, KDOT may still benefit from the near-miss crash probabilities using the current algorithm of Transoft Solutions.

- Overall, the current algorithm of Transoft Solutions detects near-miss incidents with moderate accuracy but overestimates the conflict measures (PET and TTC). However, KDOT can explore with the same vendor once Transoft Solutions has updated their algorithm.

References

Allen, B. L., Shin, B. T., & Cooper, P. J. (1978). Analysis of traffic conflicts and collisions. *Transportation Research Record 667*, pp. 67-74. http://onlinepubs.trb.org/Onlinepubs/trr/1978/667/667-009.pdf

Amundsen, F. H. E., & Hydén, C. H. (Eds.) (1977). *Proceedings of the first workshop on traffic conflicts*. Institute of Transport Economics/Lund Institute of Technology, Oslo/Lund. https://www.ictct.net/wp-content/uploads/XX-Oslo-1977/1977-Proceedings.pdf

Archer, J. (2000, October). Developing the potential of micro-simulation modelling for traffic safety assessment. In *Proceedings of the 13th ICTCT Workshop*, pp. 233-246. https://www.ictct.net/wp-content/uploads/13-Corfu-2000/ictct_document_nr_268_Archer.pdf

Astarita, V., Guido, G., Vitale, A., & Giofré, V. (2012). A new microsimulation model for the evaluation of traffic safety performances. *European Transport / Trasporti Europei*, Issue 51/XVII. https://www.openstarts.units.it/handle/10077/6117

Bai, L., Liu, P., Guo, Y., & Xu, C. (2015). Understanding factors affecting frequency of traffic conflicts between electric bicycles and motorized vehicles at signalized intersections. *Transportation Research Record, 2514,* pp. 68-78. https://journals.sagepub.com/doi/pdf/10.3141/2514-08

Brisk Synergies. (2018). *BriskVANTAGE platform data sheet*. https://safety.transoftsolutions.com/wpcontent/uploads/2019/06/BriskVANTAGE_Platform_Data_Sheet.pdf

Caliendo, C., & De Guglielmo, M. L. (2013). Road transition zones between the rural and urban environment: Evaluation of speed and traffic performance using a microsimulation approach. *Journal of Transportation Engineering*, *139*(3), pp. 295–305. https://ascelibrary.org/doi/10.1061/%28ASCE%29TE.1943-5436.0000495

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), pp. 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chin, H. C., Quek, S. T., & Cheu, R. L. (1991). Traffic conflicts in expressway merging. *Journal of Transportation Engineering*, *117*(6), pp. 633–643. https://doi.org/10.1061/(ASCE)0733-947X(1991)117:6(633)

Cooper, D. F., & Ferguson, N. (1976). Traffic studies at T-Junctions. 2. A conflict simulation record. *Traffic Engineering & Control*, *17*(7), pp. 306-309.

Evans, L. (1991). *Traffic safety and the driver*. Van Nostrand Reinhold. https://books.google.com/books?id=OyuGJJ6rKQ0C

Fazekas, A., Hennecke, F., Kalló, E., & Oeser, M. (2017). A novel surrogate safety indicator based on constant initial acceleration and reaction time assumption. *Journal of Advanced Transportation*, *2017*. https://doi.org/10.1155/2017/8376572

Federal Highway Administration (FHWA). (1976). *Traffic control systems handbook* (Implementation Package 76-10). https://books.google.com/books?id=FbEhwQEACAAJ

Fu, T., Stipancic, J., Miranda-Moreno, L., Zangenehpour, S., & Saunier, N. (2016). Traffic data collection using thermal camera under varying lighting and temperature conditions in multimodal environments. In *TAC - Efficient transportation-managing the demand-2016 conference and exhibition of the Transportation Association of Canada*. https://tac-atc.ca/sites/tac-atc.ca/files/conf_papers/fu.pdf

Gettman, D., & Head, L. (2003). Surrogate safety measures from traffic simulation models. *Transportation Research Record*, *1840*, pp. 104–115. https://doi.org/10.3141/1840-12

Gettman, D., Pu, L., Sayed, T., & Shelby, S. G. (2008). *Surrogate safety assessment model and validation final report* (Report No. FHWA-HRT-08-051). Federal Highway Administration Research Center. https://rosap.ntl.bts.gov/view/dot/39210

Goyani, J., Pawar, N., Gore, N., Jain, M., & Arkatkar, S. (2019). Investigation of traffic conflicts at unsignalized intersection for reckoning crash probability under mixed traffic conditions. *Journal of the Eastern Asia Society for Transportation Studies*, *13*, pp. 2091–2110. https://doi.org/10.11175/easts.13.2091

Guido, G., Saccomanno, F., Vitale, A., Astarita, V., & Festa, D. (2011). Comparing safety performance measures obtained from video capture data. *Journal of Transportation Engineering*, *137*(7), pp. 481–491. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000230

Guo, Y., Essa, M., Sayed, T., Haque, M. M., & Washington, S. (2019). A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. *Transportation Research Part C: Emerging Technologies*, *101*, pp. 96–110. https://doi.org/10.1016/j.trc.2019.02.009

Hayward, J. C. (1972). Near-miss determination through use of a scale of danger. *Highway Research Record, 384*, pp. 24-34. https://onlinepubs.trb.org/Onlinepubs/hrr/1972/384/384-004.pdf

Hourdos, J. N., Garg, V., Michalopoulos, P. G., & Davis, G. A. (2006). Real-time detection of crash-prone conditions at freeway high-crash locations. *Transportation Research Record*, *1968*, pp. 83–91. https://doi.org/10.1177/0361198106196800110

Hupfer, C. (1997, November). Deceleration to safety time (DST)-a useful figure to evaluate traffic safety? In *ICTCT Conference Proceedings of Seminar* (Vol. 3). https://www.ictct.net/wp-content/uploads/10-Lund-1997/10-Hupfer.pdf

Ismail, K., Sayed, T., Saunier, N., & Lim, C. (2009). Automated analysis of pedestrian–vehicle conflicts using video data. *Transportation Research Record*, *2140*, pp. 44–54. https://doi.org/10.3141/2140-05

Jackson, S., Miranda-Moreno, L. F., St-Aubin, P., & Saunier, N. (2013). Flexible, mobile video camera system and open source video analysis software for road safety and behavioral analysis. *Transportation Research Record*, *2365*, pp. 90–98. https://doi.org/10.3141/2365-12

Kathuria, A., & Vedagiri, P. (2020). Evaluating pedestrian vehicle interaction dynamics at un-signalized intersections: A proactive approach for safety analysis. *Accident Analysis & Prevention*, *134*, Article 105316. https://doi.org/10.1016/j.aap.2019.105316

Ke, R., Lutin, J., Spears, J., & Wang, Y. (2017). A cost-effective framework for automated vehicle-pedestrian near-miss detection through onboard monocular vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 25–32. https://openaccess.thecvf.com/content_cvpr_2017_workshops/w9/papers/Ke_A_Cost-Effective_Framework_CVPR_2017_paper.pdf

Laureshyn, A., Svensson, Å., & Hydén, C. (2010). Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. *Accident Analysis & Prevention*, *42*(6), pp. 1637–1646. https://doi.org/10.1016/j.aap.2010.03.021

Matsui, Y., Hitosugi, M., Doi, T., Oikawa, S., Takahashi, K., & Ando, K. (2013). Features of pedestrian behavior in car-to-pedestrian contact situations in near-miss incidents in Japan. *Traffic Injury Prevention*, *14*(sup1), pp. S58–S63. https://doi.org/10.1080/15389588.2013.796372

Minderhoud, M. M., & Bovy, P. H. (2001). Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention*, *33*(1), pp. 89–97. https://doi.org/10.1016/S0001-4575(00)00019-1

Ozbay, K., Yang, H., Bartin, B., & Mudigonda, S. (2008). Derivation and validation of new simulation-based surrogate safety measure. *Transportation Research Record*, *2083*, pp. 105–113. https://journals.sagepub.com/doi/pdf/10.3141/2083-12

Parker, M. R., Jr., & Zegeer, C. V. (1989). *Traffic conflict techniques for safety and operations: Observers manual* (Report No. FHWA-IP-88-027). Federal Highway Administration. https://rosap.ntl.bts.gov/view/dot/14308

Peesapati, L. N., Hunter, M., Rodgers, M., & Guin, A. (2011). A profiling-based approach to safety surrogate data collection. In *3rd International Conference on Road Safety and Simulation*. https://onlinepubs.trb.org/onlinepubs/conferences/2011/RSS/1/Peesapati,L.pdf

Peesapati, L. N., Hunter, M. P., & Rodgers, M. O. (2013). Evaluation of postencroachment time as surrogate for opposing left-turn crashes. *Transportation Research Record*, *2386*, pp. 42–51. https://journals.sagepub.com/doi/pdf/10.3141/2386-06

Pirdavani, A., Brijs, T., Bellemans, T., & Wets, G. (2010). Evaluation of traffic safety at un-signalized intersections using microsimulation: a utilization of proximal safety indicators. *Advances in Transportation Studies*, *22*, pp. 43–50. https://www.researchgate.net/publication/228660289

Pu, L., & Joshi, R. (2008). *Surrogate safety assessment model (SSAM)--Software user manual* (Report No. FHWA-HRT-08-050). Federal Highway Administration Research Center. https://rosap.ntl.bts.gov/view/dot/1008

Razmpa, A. (2016). *An assessment of post-encroachment times for bicycle-vehicle interactions observed in the field, a driving simulator, and in traffic simulation models* (Portland State University). https://www.proquest.com/docview/1873476981

Samara, L., St-Aubin, P., Loewenherz, F., Budnick, N., & Miranda-Moreno, L. (2021, January). *Video-based network-wide surrogate safety analysis to support a proactive network screening using connected cameras: Case study in the City of Bellevue (WA) United States*. Presented at 100th Annual Meeting of the Transportation Research Board. https://dev.bellevuewa.gov/sites/default/files/media/pdf_document/2020/TRB-Accepted-Paper-2021-Video-Surrogate-Safety-Analysis-Bellevue.pdf

Saunier, N., & Sayed, T. (2006, June). A feature-based tracking algorithm for vehicles in intersections. In *3rd Canadian Conference on Computer and Robot Vision (CRV'06)*. IEEE. https://doi.org/10.1109/CRV.2006.3

Sayed, T., Zaki, M. H., & Autey, J. (2013). A novel approach for diagnosing road safety issues using automated computer vision techniques. In *16th International Conference Road Safety on Four Continents. Beijing, China (RS4C 2013)*. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A759740&dswid=2816

Shekhar Babu, S., & Vedagiri, P. (2018). Proactive safety evaluation of a multilane unsignalized intersection using surrogate measures. *Transportation Letters*, *10*(2), pp. 104–112. https://doi.org/10.1080/19427867.2016.1230172

Songchitruksa, P., & Tarko, A. (2004). *Using imaging technology to evaluate highway safety* (Report No. FHWA/IN/JTRP-2004/27). Purdue University. https://doi.org/10.5703/1288284313343

Songchitruksa, P., & Tarko, A. P. (2006). The extreme value theory approach to safety estimation. *Accident Analysis & Prevention*, *38*(4), pp. 811–822. https://doi.org/10.1016/j.aap.2006.02.003

St-Aubin, P., Saunier, N., & Miranda-Moreno, L. (2015). Large-scale automated proactive road safety analysis using video data. *Transportation Research Part C: Emerging Technologies*, *58*, pp. 363–379. https://doi.org/10.1016/j.trc.2015.04.007

Stipancic, J., St-Aubin, P. G., Ledezma-Navarro, B., Labbe, A., Saunier, N., & Miranda-Moreno, L. (2021). Evaluating safety-influencing factors at stop-controlled intersections using automated video analysis. *Journal of Safety Research*, *77*, pp. 311-323. https://doi.org/10.1016/j.jsr.2021.03.006

Sykes, P. (2010). Traffic simulation with paramics. In J. Barceló (Ed.), *Fundamentals of traffic simulation*, pp. 131–171. Springer. https://link.springer.com/chapter/10.1007/978-1-4419-6142-6_4

Tageldin, A., Sayed, T., & Ismail, K. (2018). Evaluating the safety and operational impacts of left-turn bay extension at signalized intersections using automated video analysis. *Accident Analysis & Prevention*, *120*, pp. 13–27. https://doi.org/10.1016/j.aap.2018.07.029

Tang, K., & Kuwahara, M. (2011). *Implementing the concept of critical post-encroachment time for all-red clearance interval design at signalized intersections*. Presented at The 9th International Conference of Eastern Asia Society for Transportation Studies. Eastern Asia Society for Transportation Studies. https://doi.org/10.11175/eastpro.2011.0.299.0

Thallinger, G., Krebs, F., Kolla, E., Vertal, P., Kasanický, G., Neuschmied, H., & Ambrosch, K. E. (2018). Near-miss accidents – Classification and automatic detection. In T. Kováčiková, Ľ. Buzna, G. Pourhashem, G. Lugano, Y. Cornet, N. Lugano (Eds.), *Intelligent transport systems – From research and development to the market uptake*: *Proceedings first international conference*, *INTSYS 2017*, *Hyvinkää, Finland, November 29-30, 2017* (pp. 144-152). Springer. https://books.google.com/books?id=tkpjDwAAQBAJ

Van der Horst, R. (1990). A time-based analysis of road user behaviour at intersections. In Draskóczy, M. (Ed.), *Proceedings of the Third Workshop of International Cooperation on Theories and Concepts in Traffic Safety*, pp. 91–107. https://www.ictct.net/wp-content/uploads/03-Krakow-1990/03-Proceedings.pdf

Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. *Accident Analysis & Prevention*, *35*(3), pp. 427–433. https://doi.org/10.1016/S0001-4575(02)00022-2

Zaki, M. H., Sayed, T., Tageldin, A., & Hussein, M. (2013). Application of computer vision to diagnosis of pedestrian safety issues. *Transportation Research Record*, *2393*, pp. 75–84. https://journals.sagepub.com/doi/pdf/10.3141/2393-09

Zangenehpour, S., Chung, C., & Saneinejad, S. (2017). Impact of curb radius reduction on pedestrian safety: A before-after surrogate safety study in Toronto. In *Proceedings of 27th CARSP conference, Toronto, ON, June 18-21, 2017.* Canadian Association of Road Safety Professionals. https://carsp.ca/en/presentations-and-papers/carsp-conference-acpser-toronto-2017/automated-evaluation-of-the-effectiveness-of-curb-radii-adjustments-on-turning-vehicle-speed-a-case-study-in-toronto/

Zangenehpour, S., Miranda-Moreno, L. F., & Saunier, N. (2015). Automated classification based on video data at intersections with heavy pedestrian and bicycle traffic: Methodology and application. *Transportation Research Part C: Emerging Technologies*, *56*, pp. 161–176. https://doi.org/10.1016/j.trc.2015.04.003

Zangenehpour, S., Strauss, J., Miranda-Moreno, L. F., & Saunier, N. (2015). *Are intersections with cycle tracks safer? A control-case study based on automated surrogate safety analysis using video data*. Presented at 94th Annual Meeting of the Transportation Research Board. https://citeseerx.ist.psu.edu/pdf/9d0ec7f8b6610099389fac2d3e2637023fbc2a1b

Zhang, S., Abdel-Aty, M., Wu, Y., & Zheng, O. (2020). Modeling pedestrians' near-accident events at signalized intersections using gated recurrent unit (GRU). *Accident Analysis & Prevention*, *148*, Article No 105844. https://doi.org/10.1016/j.aap.2020.105844

Zheng, L., Sayed, T., & Essa, M. (2019). Validating the bivariate extreme value modeling approach for road safety estimation with different traffic conflict indicators. *Accident Analysis & Prevention*, *123*, pp. 314–323. https://doi.org/10.1016/j.aap.2018.12.007

Appendix A.1: Transoft Solutions' Camera Specifications

Transoft Solutions requires a minimum resolution of 1280 x 720 (HD resolution), and the camera field (angle) of view must be wide enough to capture the scenario to be analyzed. The camera must be installed at a height of at least 6 m and a maximum of 10 m (20–32 ft) to capture a top view of the intersection. The camera must be secured so it does not move during data collection.

Figure A.1: Wide Field of View of the Installed Camera

Figure A.2: Correct Height and Angle of Camera

Appendix A.2: Ground Truth Manual Extraction Process

The detailed manual extraction process to obtain ground truth can be summarized as follows:

1. The data were first sorted by PET and TTC using Microsoft Excel.
2. Then PET and TTC data were further sorted in 0-1, 1-2, 2-3, 3-4, and 4-5s groups.
3. Approximately 10 percent of the data from each tab were then sorted, which was final data.
4. The final data from each tab was highlighted in green for each tab.
5. Log in to TrafxSAFE website using login credentials provided by the vendor.
6. In the highlighted area look for the URL in the very right column of the Excel and copy/paste that link into any browser.
7. A video clip will pop up and run (approximately 10 seconds, the length may vary).
8. Some straight solid lines and a circle will appear.
9. Enlarge to full screen. Pause the video and draw the same circle on the monitor. Take another marker and point to the approximate center of that circle as a conflict point (for both PET and TTC) on the monitor. For PET, the conflict point will be the point where the trajectories intersect.
10. Navigate to the folder where all the video files were saved.
11. Select the respective video from the folder and play in VLC media player.
12. Enlarge to full screen and, based on the timestamp and other information provided in the Excel file such as first road user type and movement, second road user type and movement, and scenario type, locate the same

event. Note: the event in VLC media player will appear approximately four seconds earlier than the time reported in the Excel file.

13. Pause the video once the same event (as observed in step 8) has been observed. Then press the letter E on the keyboard to advance frame by frame in VLC media player. The milliseconds will appear at the top of the screen.
14. When the back of the first vehicle leaves the marked conflict point, pause the video and record timestamp 1. When the front of the second road user touches the same conflict point, record timestamp 2.
15. Then calculate the numerical difference in milliseconds for these two timestamps and divide it by 1,000 to obtain the time in seconds, which is the value of PET or TTC.
16. Repeat this process for all other events.

K-TRAN

KANSAS TRANSPORTATION RESEARCH AND NEW-DEVELOPMENT PROGRAM

AD ASTRA PER ASPERA
Kansas
Department of Transportation