



Investigating surrogate safety measures at midblock pedestrian crossings using multivariate models with roadside camera data

Nafis Anwari^{*}, Mohamed Abdel-Aty, Amrita Goswamy, Ou Zheng

Department of Civil, Environment & Construction Engineering, University of Central Florida (UCF), Orlando, FL 32816, USA



ARTICLE INFO

Keywords:

Pedestrian Safety
Surrogate Safety Measures
Rectangular Rapid Flashing Beacon
Pedestrian Hybrid Beacon
Multivariate Linear and Non-linear Regression

ABSTRACT

This study aims to evaluate and compare Surrogate Safety Measures (SSMs) at five midblock Rectangular Rapid Flashing Beacons (RRFB) and two midblock Pedestrian Hybrid Beacons (PHB) sites in Florida using extensive video data collected over the study period of July to November 2021. Computer vision and data processing resulted in four pedestrian SSMs, namely spatial gap, temporal gap, relative time to collision (RTTC) and Post Encroachment Time (PET). An initial investigation of the SSMs using Mann-Whitney-Wilcoxon tests revealed significant differences in the SSM values across different treatment types and hours of the day. Additionally, univariate regression of spatial gap, and multivariate regression of temporal gap, RTTC and PET revealed significant differences of SSMs across RRFB and PHB sites. The study considered both linear and non-linear (gamma, inverse Gaussian and lognormal) regression models. After considering various traffic and operational parameters, the data were aggregated for each pedestrian-vehicle interaction on each lane to create a total of 395 observations. The SSMs included average spatial gap, temporal gap, RTTC and PET for each interaction of pedestrian and vehicle on each lane. The results indicated that non-linear models performed better than the linear models. Moreover, the presence of the PHB, weekday, signal activation, lane count, pedestrian speed, vehicle speed, land use mix, morning period and pedestrian starting position from the sidewalk have been found to be significant determinants of the SSMs. Results also suggest temporal SSMs increase at the PHB sites compared to the RRFB sites, indicating an improvement of traffic safety at PHB sites. However, the spatial gap decreased for PHB sites compared to the RRFB sites, which suggests that pedestrians tend to start to cross the RRFB sites when they perceive vehicles to be further away than at the PHB sites.

1. Introduction

Pedestrian safety is one of the most serious concerns in Florida. In 2021, Florida ranked as the 2nd state nationally with the highest number of pedestrian fatalities (i.e., estimated 899) as well as 2nd for pedestrian fatality rate (i.e., 4.13 pedestrian fatalities per 100,000 population) according to the State Highway Safety Offices and U.S. Census Bureau. Additionally, pedestrian fatalities in Florida have increased almost every year since 2010 (486 fatalities) to 2021 (899 fatalities) (Macek, 2022; National Highway Traffic Safety Administration, 2020).

More pedestrian crashes were observed at non-intersection locations compared to intersections. Crosswalks are usually placed at midblock for pedestrians to walk across the road. However, it would be difficult for pedestrians to navigate if crosswalks are unsignalized or at locations

without traffic control devices (uncontrolled crosswalks), which can pose significant risks. Drivers often fail to stop or yield to pedestrians in uncontrolled marked crosswalks. The main reasons could be that drivers fail to recognize crosswalks and abide by traffic laws, which leads to crossing point issues and crashes involving pedestrians in uncontrolled crosswalks.

Pedestrian crossing enhancement treatments such as Pedestrian Hybrid Beacon (PHB) and Rectangular Rapid Flash Beacon (RRFB) have been developed and implemented to improve safety at uncontrolled crosswalks. The PHB offers a unique beacon configuration with two red lights and a single yellow light. The light cycle is activated by a pedestrian, through automated (i.e., via detection) or manual means. A flashing yellow indication is first lit on, followed by a solid yellow indication, which informs drivers to prepare to stop. Then, a solid red indication is on during the walk period and vehicles should stop before

* Corresponding author.

E-mail addresses: nafis.anwari@ucf.edu (N. Anwari), M.Aty@ucf.edu (M. Abdel-Aty), amrita.goswamy@ucf.edu (A. Goswamy), ouzheng1993@knights.ucf.edu (O. Zheng).

the stop line. Finally, the signal finishes with a flashing red indication, which encourages the drivers to stop and then proceed with caution if there is no pedestrian crossing. The PHB is mounted overhead like a traffic signal. The RRFB is another safety enhancement tool, which consists of two rectangular organized amber light-emitting diodes (LEDs) flashing adjacent to one another. The lights are mounted just below a rectangular pedestrian sign and are installed on the side of the road and in the center median. Previous studies have reported a statistically significant reduction in crashes, and an increase in drivers' compliance of yielding pedestrians for PHB and RRFB. For example, Zegeer et al. (2017) showed that PHBs decreased total crashes and pedestrian crashes by 18% and 75.6%, respectively compared to crashes at untreated sites (Zegeer et al. 2017). Monsere et al. (2017) concluded that RRFBs reduced pedestrian crashes by 16% compared to the time before installation (Monsere et al. 2017). Driver yielding was found to be 96% and 86% at PHB and RRFB sites, respectively (Fitzpatrick et al. 2014, 2016).

Traffic conflicts have been defined in this study as an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged (Arun et al., 2021). Previous studies assessed the safety aspect of pedestrian crossings using only crash data, (Goswamy et al., 2023; Monsere et al., 2017; Zegeer et al., 2017). To assess safety microscopically using real-time data, Surrogate Safety Measures (SSMs) serve as near-crash indicators to measure the severity and frequency of traffic conflict events. Thus, SSMs can be used to assess the safety aspect of a treatment in real time (Chen et al., 2017).

2. Literature review

The literature review was conducted to summarize the existing Surrogate Safety Measures (SSMs) for pedestrian and vehicular interactions. Previous efforts on pedestrians' crossing and vehicles' yielding behaviors and traffic efficiency measures were also investigated. As an alternative to crash risk estimation based on historical crash data that depends on police-reported data that can have huge variability in coding between officers, SSMs serve as near-crash indicators to measure the severity and frequency of traffic conflict events (Chen et al., 2017). Previous pedestrian safety studies involving SSMs have explored three types of interactions, namely, pedestrian-vehicle, pedestrian-pedestrian and pedestrian-bicycle interactions (e.g. (Beitel et al., 2018; Gruden et al., 2019; Li, 2013)). This study focuses on evaluating SSMs for pedestrian-vehicle interactions at signalized midblock crossings using the RRFB and PHB.

2.1. Pedestrian crossing behavior

The bulk of the literature on pedestrian road-crossing behavior has used the gap acceptance to gauge the safety of a particular pedestrian crossing facility. Most of the studies on gap acceptance have defined the gap in traffic as the time duration between two consecutive vehicles arriving at the same exact reference point (on the crossing facility in this case). The shortest gap that the pedestrian would accept for crossing the facility is termed as the critical gap (Li, 2013, 2014; Saleh and Lashin, 2020). A slight variation of this approach is to determine the accepted gap of each pedestrian by measuring the time interval between the two successive vehicles that arrive at the crosswalk before and after the pedestrian started crossing, respectively. The smallest gap is taken as the critical gap (Tezcan et al., 2019). Another variation of the temporal gap has been studied as the time difference between the instant the vehicle arrives at the crossing path (latter time point) and the pedestrian is just ready to set foot on the street (earlier time point) (Yannis et al., 2013). Some studies have explored the temporal gaps further by investigating rolling gaps. When a pedestrian has to consider more than one vehicle for making a gap acceptance decision, the pedestrian will consider the gaps with reference to the vehicles travelling in different lanes and

accept rolling gaps (Pawar and Patil, 2015). As opposed to the temporal nature of the above-mentioned gap acceptance, a few studies have focused on the spatial gap acceptance. During the spatial gap determination, the gap is measured in terms of distance between pedestrian and vehicle. The safety distance is considered as the distance between the pedestrian and the vehicle in a specific lane at the time of crossing this lane. The critical distance is the minimum safety distance between the lanes in one direction. The concept of using the distance between the pedestrians and the vehicles to represent the gap has been utilized in (Oxley et al., 1997, 2005; Pawar and Yadav, 2022; Shaaban et al., 2018). A study by Theofilatos et al. (2021) assessed the effect of vehicle speed, gap size, waiting time and frequency of attempts on pedestrian gap acceptance probabilities through meta-analyses of 14 studies. The study found that frequency of attempts had the greatest impact on pedestrian gap acceptance, followed by gap size, waiting time then speed. For each crossing attempt, the odds of crossing the road (accepting the incoming time gap) was 16.6 times larger for pedestrians, compared to the previous unsuccessful crossing attempts (Theofilatos et al. 2021).

2.2. Application of SSMs in pedestrian-vehicle interactions

In addition to the gap acceptance, driver yielding behavior has been investigated in the literature to assess the effectiveness of given crossing facilities. The literature has considered vehicle yielding through two actions: vehicular speed reduction and vehicular lane change. Usually, such studies only investigate the proportion of vehicles yielding to the pedestrians (Avinash et al., 2019; Kadali and Vedagiri, 2019; Shaaban et al., 2018). A notable exception to this is Katz et al. (1975), who defined an index of relative speed reduction as the crossing velocity minus approach velocity, divided by approach velocity (Katz et al., 1975). The study found significant differences in drivers' crossing velocities at crosswalks as compared to the no-crosswalk condition. The study concluded that lower crossing velocities can be expected when the approach speed of a vehicle is low, when a group of pedestrians' attempts crossing on a marked crosswalk, when there is a relatively long-distance gap between the pedestrians and the vehicles, and when pedestrians do not look at the approaching vehicle. When segregating the studied drivers, the study found that female drivers and older drivers have a higher yielding tendencies than others (Katz et al., 1975). Kadali and Perumal (2016) used Driver Yield Behavior (DYB) to calculate critical gap accepted by pedestrians using a logit model. However, the exact method to determine DYB has not been explicitly mentioned in this study (Kadali and Perumal 2016). Literature review reveals that the driver's yield behavior affects the pedestrian's gap acceptance size. DYB can vary between developing and developed countries, because developing countries have a higher pedestrian density, leading to higher pedestrian-vehicle interaction. In mixed traffic conditions, different vehicles have different speeds, and may not maintain lane-based driving. The vehicle drivers may exhibit a path change to avoid interaction with pedestrians, which can influence pedestrians' critical gap estimation (Kadali and Perumal, 2016).

Other SSMs that were studied in the literature include Time to Collision (TTC) and Post Encroachment Time (PET), focusing on uncontrolled midblock crossings and signalized intersection crossings (Avinash et al. 2019; Ismail et al. 2009, 2010; Jiang et al. 2015; Kadali and Vedagiri 2019; Kaparias et al. 2013; Layegh et al. 2020; Marisamynathan and Vedagiri 2020; Ni et al. 2016; Onelcin and Alver 2015; Ren et al. 2012; Várhelyi 1998; Zhang et al. 2017). Cafiso et al. (2011) and Chaudhari et al. (2020) have used TTC and Stopping Time to develop Pedestrian Risk Index (PRI). The studies calculated both Time to Collision for pedestrian (TTCp) and Time to Collision for vehicle (TTCv), but only uses TTCv for direct calculations of PRI. While the PRI itself seems like a promising parameter to study, it is dependent on the stopping time, which depends partially on the reaction time. The reaction time can vary from driver to driver and even from situation to situation (light intensity, weather, mental and physical condition of the

driver, etc.), and using a single value of reaction time (usually 2.5 s) does not fit all situations in our study. The considered reaction times were different even in the two above-mentioned studies (1.07 s and 0.7 s respectively) (Cafiso et al. 2011; Chaudhari et al. 2020). A naturalistic driving study by Arbabzadeh et al. (2019) found that drivers' reaction times vary between 0.5 and 7.5 s. Another shortcoming of the PRI is that it only considers a case as a conflict only when $TTC_p < TTC_v$, meaning only when the pedestrian is projected to cross first. Setting aside the eventual outcome of which road-user crosses the conflict point first, many interactions show that pedestrians are projected to cross first during some timeframes ($TTC_p < TTC_v$) and vehicles are projected to cross during other timeframes ($TTC_p > TTC_v$). This means that in the same interaction, sometimes the vehicle is projected to cross first, while the pedestrian is expected to cross first in other times. Both scenarios create confusion in the minds of both road-users (pedestrians and drivers), but PRI only considers one scenario, making its risk assessment incomplete.

2.3. Pedestrian safety at RRFB and PHB

Studies assessing effectiveness of Pedestrian Hybrid Beacon (PHB) have typically used driver yield rates (Anderson, 2019; Bennett et al., 2014; Brewer et al., 2015; Fitzpatrick et al., 2020). Rista and Fitzpatrick evaluated and compared the driver yielding behavior among different traffic safety infrastructures, including LED pedestrian crossing signs, RRFB and PHB (Rista and Fitzpatrick 2020). Mendez assessed the RRFB performance using PET and yielding behavior (Méndez 2020). Ugan et al. (2022) compared driver yielding behavior based on the speed reductions among the RRFB and PHB sites, and concluded that PHB sites had a more sustained effect downstream on speed reductions (Ugan et al. 2022). Fitzpatrick et al. (2014) and Fitzpatrick et al. (2016) found driver yielding and pedestrian push-button compliance to be better at PHB sites compared to RRFB sites (Fitzpatrick et al., 2014, 2016).

The above literature review reveals that most of the pedestrian-vehicle conflict studies have focused on both signalized and unsignalized intersections and mid-block crossings. However, there is a gap in the literature in evaluating and comparing among RRFB and PHB sites using well-established SSMs. Moreover, there is a gap in the literature in using Computer Vision (CV) technology in pedestrian surrogate safety studies. Given the identified literature gaps, this study has contributed to the literature by being the first to evaluate SSMs at PHB sites. Moreover, this study has compared the safety effectiveness of RRFB and PHB sites based on four SSMs, namely spatial gap, temporal gap, RTTC and PET. This study has investigated the effect of a unique factor, namely, land use mix, on SSMs. Besides, the results of this study can suggest whether RRFB or PHB is suitable for a site based on the land use mix, the traffic characteristics and the roadway characteristics. To achieve these objectives, this study has utilized extensive roadway video data through CV processing. These are described further in the following sections.

3. Data collection and extraction

Seven locations in the state of Florida have been selected for this study to compare effectiveness of Rectangular Rapid Flashing Beacons (RRFB) and including Pedestrian Hybrid Beacons (PHB), which are

described in Table 1.

The raw data were collected using several video cameras set up at suitable angles, direction and heights at all study sites to capture pedestrian-vehicle interactions as well as vehicular trajectories upstream and downstream of the pedestrian crosswalk. An example of camera positions for location VI is provided in Fig. 1. Depending on the site requirements, two to four cameras were placed at the roadside existing poles to synchronously record the oncoming vehicles of two directions, pedestrians crossing the roads, and the signal status. The heights of the cameras were adjusted to capture more upstream portions (when cameras were installed higher) or entire crosswalk portion (when cameras were installed lower). Sample screenshots from cameras in Location VI are provided in Fig. 2.

3.1. Data extraction by computer vision (CV) processing

The Near Miss Event Detection System (NMEDS) was developed by the research team for the purpose of data extraction from videos. The NMEDS framework utilizes a combination of the Mask-RCNN bounding box and Occlusion-Net detection algorithms to reconstruct the key points of road users in a 3D representation (Abdel-Aty et al. 2022). The system offers various possibilities for traffic analysis, including:

1. Grid-based spatial autocorrelation analytics: This technique was employed to identify locations that exhibit significant levels of danger. By analyzing the spatial patterns of near-miss events, the system can pinpoint areas with a higher likelihood of accidents or risky situations.
2. Prediction of road users' intentions: The NMEDS system has the capability to predict the intentions of road users, such as pedestrians crossing the road. By analyzing their movements and trajectories, the system can provide insights into potential crossing points and help improve safety measures.
3. Assessment of road safety conditions using surrogate safety measures: The system enables the estimation of road safety conditions by utilizing surrogate safety measures. By analyzing various factors such as traffic volume, vehicle speed, and other relevant variables, NMEDS can provide valuable information regarding the safety levels of specific road segments.

By employing the NMEDS system, researchers and traffic analysts can leverage these capabilities to enhance their understanding of traffic dynamics, identify dangerous locations, predict road user behavior, and evaluate road safety conditions using surrogate measures. The NMEDS framework leverages a hybrid approach, incorporating both the Mask-RCNN bounding box and Occlusion Net detection algorithms. The Mask R-CNN model was adapted from Faster-RCNN (He et al. 2017; Ren et al. 2015). This integration allows for the reconstruction of crucial landmarks or key points associated with road users within a three-dimensional (3D) representation. Moreover, CSRT was used for tracking the objects that were detected by Mask R-CNN. In case of a missed detection by the Mask-RCNN algorithm, the CSRT (DCF-CSR, Discriminative Correlation Filter with Channel and Spatial Reliability) can provide an interpolated tracking location of the road user in question. By combining the strengths of these algorithms, NMEDS can

Table 1
Description of Study sites.

Location number	County	Treatment type	Latitude	Longitude	Lane count	Presence of median	One-way/two way
I	Aalachua	PHB	29.63709	-82.31455	4	Yes	Two-way
II	St. Lucie	PHB	27.51173	-80.30718	2	No	Two-way
III	St. Lucie	PHB	27.52405	-80.31224	2	No	Two-way
IV	St. Lucie	PHB	27.5261	-80.31283	2	No	Two-way
V	St. Lucie	PHB	27.53148	-80.31435	2	No	Two-way
VI	Marion	RRFB	29.04662	-82.46389	4	Yes	Two-way
VII	Brevard	RRFB	28.32562	-80.60923	2	No	One-way

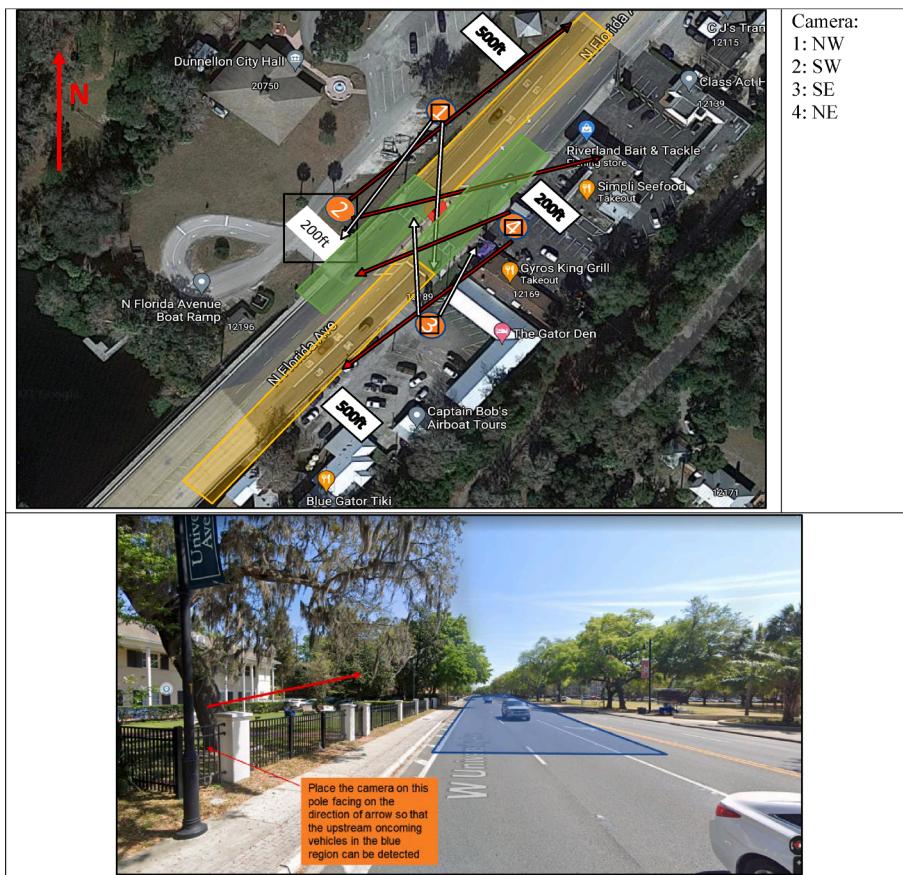


Fig. 1. Video Camera Positions at an RRFB site.

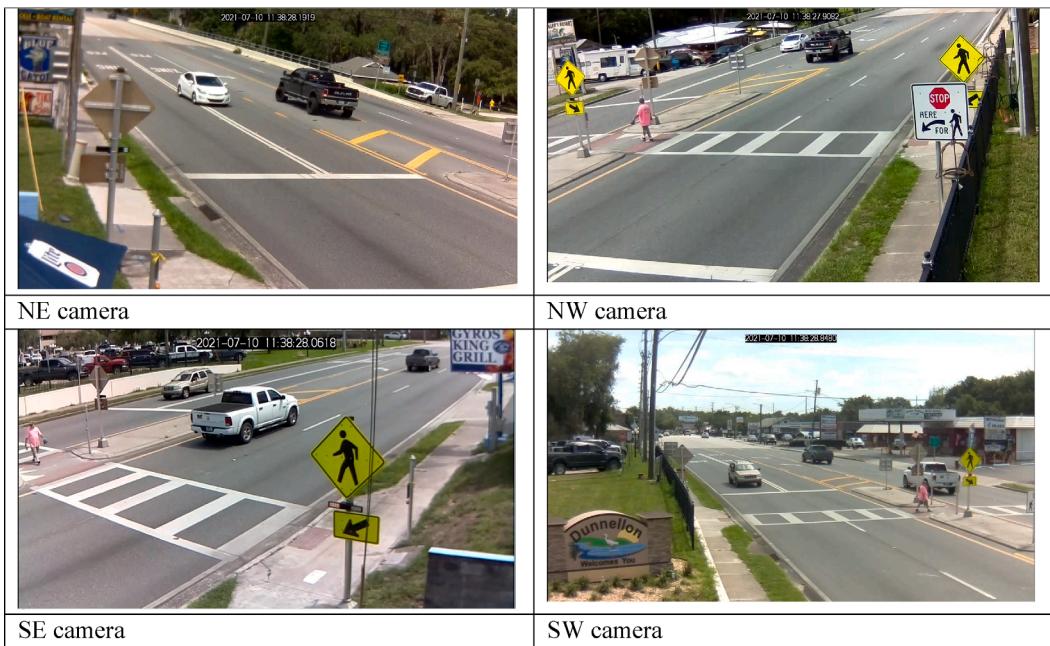


Fig. 2. Sample Screenshots from Cameras.

accurately capture and visualize the spatial positions and orientations of road users in the analyzed video data (Abdel-Aty et al. 2022). Trajectories extracted using CV were validated by comparing with human labeled ground truth. Intersection of Union (IOU) are calculated based on the following equation (Zheng 2019):

$$IOU_{gt} = \frac{DetectionResult \cap GroundTruth}{DetectionResult \cup GroundTruth} \quad (i)$$

The tracking accuracy has been validated in previous papers (Abdel-Aty et al. 2022; Mahmoud et al. 2021; Zheng 2019). In total, 1,588 road

user were detected and tracked based on the UAV video. The accuracy of the proposed algorithm was evaluated based on IOU values. For each type of movement (i.e. left turn, right turn, straight) IOU values were calculated at the movement duration for each vehicle based on the outputs and the ground truths that were collected manually. Totally 3,541 video images were collected to calculate the IOU for the selected vehicles. Higher IOU values indicate higher accuracy of the detection and tracking results. Based on previous studies, the IOU values for left turn, right turn and straight trajectories were 0.78, 0.70 and 0.85, respectively. This creates an average IOU of 0.76. It outperforms the average performance of general computer vision models, whose IOU is around 0.5–0.6 (Abdel-Aty et al. 2022; Mahmoud et al. 2021; Zheng 2019).

Fig. 3 shows an example of the output videos. The road users were first detected with the classification result (pedestrian or vehicle). Then, the object tracking model was used to follow the movements of each road user, and trackers' ID numbers were displayed at the top left corner of the bounding box. The moving trajectory of the road user (the latest ten movements) were also plotted. The trajectories of the road users in the study area (within the yellow boundary) were generated. The road users such as parked vehicles outside the studies area were only detected but not tracked, which is shown by masks instead of the bounding boxes. The resulting pedestrian (red) and vehicle (blue) trajectories shown in Fig. 3 for the given camera angle (SW camera) show that the employed CV algorithm was able to accurately track pedestrians.

Videos were collected for 2–3 days with 8 h per day for each location. To capture sufficient pedestrian crossing activities and different driving behavior, both weekdays and weekends were considered. The 8-hour video on each day considered the following time: morning: 8–10 am; afternoon: 11:30 – 2 pm; and evening: 5:30—9 pm. The time slots were chosen with an aim to capture the maximum pedestrian-vehicle interactions. Thus, videos were able to capture morning and evening peak traffic during weekdays. Surprisingly, during the weekends, pedestrian activity peaked during noon in many locations. The data collection period covering all sites was from July to November 2021. A total of 144 hours of video data trajectory were extracted for analysis of SSMs at midblock crossings, which has never been attempted in the previous literature. Thus, even though this study only covered seven sites, the rich video dataset provides adequate opportunities for analyzing SSMs at signalized midblock crossings.

3.2. Vehicle trajectory data processing

Time-stamped vehicle trajectory files were extracted from the video cameras. The literature has commonly used 0.1 s as a time-step for trajectory extraction (for example, (Punzo et al. 2011; Wang et al. 2010)). Smaller time steps can capture more fluctuations in vehicle

trajectories. Thus, the current study used trajectory data based on 26 frames/second video data (i.e. time-step of 0.038 s). Trajectory files from multiple camera angles were merged, if required, to get the full vehicle and pedestrian trajectory. The trajectories were then processed in python to identify pedestrian-vehicle (conflicts) interactions and calculate surrogate safety measures, namely, spatial gap, temporal gap, RTTC and PET for each conflict for the relevant time frames.

One important step during data processing is ensuring that the data represents the real-world situation as close as possible. Some of the time-step observations from raw vehicle trajectory data can lead to outlier estimates of distances and instantaneous speeds and accelerations. These arise because of local errors in vehicle tracking. Thus smoothening techniques are needed to reduce outliers. The literature has used a variety of smoothening techniques, including moving average method, local regression method, Savitzky-Golay filtering technique, and Kalman Filtering technique. Moving average is a calculation to analyze data points by creating a series of averages of different selections of the full data set. It is also called a moving mean or rolling mean and is a type of finite impulse response filter. With the increase in the degree of moving average (number of observations used for averaging), the degree of smoothening increases. However, this reduces the chance of observing important fluctuations in speed and acceleration. In local weighted regression, position measurements are smoothed by fitting a local curve at the points of interest. In addition to the regression weight function, different robust weight functions can be used to make the data resistant to outliers. Various sub-methods of local regression include Lowess and Loess methods. Savitzky-Golay filtering technique fits successive subsets of data points with a low-degree polynomial function using linear least square method in a process called convolution. When the data points are equally spaced, an analytical solution to the least-squares equations can be found in the form of a single set of "convolution coefficients". This can be applied to all data sub-sets to determine the estimates of the smoothed data points (Raju et al. 2017). Kalman filtering produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone, by estimating a joint probability distribution over the variables for each timeframe. The Kalman filter produces estimates of the current state variables, along with their uncertainties. These estimates are updated using weighted average when the outcome of the next measurement is observed. Greater weight is given to estimates with greater certainty. The algorithm is recursive. Kalman filtering is a forward pass through the data. Kalman smoothing adds a backward pass through the data (Punzo et al. 2011).

Although each smoothening technique has its strengths and benefits, previous research has shown that moving average method can achieve good results while saving processing time (Raju et al. 2017). Thus, the current study has used the moving average method to smoothen distances of road-users from reference points, which in turn produces good



Fig. 3. Automated Object Detection and Tracking Models and Resultant Road-user Trajectories.

results for speed and acceleration. For example, the results of speeds obtained from raw trajectories and smoothed trajectories for four random vehicles in the study dataset is provided in Fig. 4.

The obtained data were checked against videos manually for each observed pedestrian-vehicle interaction to validate the obtained output. In addition, the pedestrian speed, traffic flow (veh/h), timestamps for the SSMs, flashing status of the pedestrian signal was extracted for each pedestrian-vehicle interaction for each lane of the road. The study considered only the lanes that the pedestrian would have to cross without taking refuge. Therefore, in the sites without medians, all lanes were considered. However, in the sites having medians, only the lanes before reaching the median were considered for each pedestrian. Pedestrians who walked on the sidewalk and did not cross the road were eliminated. Trajectory merging of pedestrians in python solved the issue of disjointed identification of pedestrians in Computer Vision (CV) processing. The same was done for the vehicles interacting with pedestrians. The aggregated dataset resulted in over 50,000 observations, where each observation represented a timeframe and the resultant SSM values occurring in that time frame at a particular lane. This was finally compressed to obtain spatial gap, average temporal gap, average RTTC, and PET for each pedestrian-vehicle interaction at each lane. This created a total of 395 observations for the final dataset. If a fleet of vehicles interacted with pedestrians, only the leading vehicle in each lane were considered.

3.3. Description of performance measures

Different performance measures were computed based on the trajectories of vehicles and pedestrians, signal activation time, speeds of vehicles and pedestrians using computer vision and validated by manual inspection. This study has considered the following SSMs in evaluating the effectiveness of RRFB and PHB sites:

3.3.1. Post Encroachment Time (PET)

PET is the difference between the time frames in the video for a pair of pedestrian/s and vehicle when they are likely to cross each other at the conflict/intersection point in their line of path. The vehicle and pedestrian trajectory are obtained from the videos, from which the coordinates of the conflict point were determined (the point where the

pedestrian and the vehicle are likely to collide, which can be determined using their trajectories). The PET can be calculated by taking the difference of time when a pedestrian reaches the crossing point and the time when vehicle reaches the crossing point (refer to Fig. 5).

PET = Time difference between when pedestrian and vehicle reach the conflict point

PET provides the ultimate outcome of a pedestrian-vehicle conflict. By definition, a PET value of zero implies the occurrence of a crash. The risk of conflict increases as the value of PET approaches zero. PET takes into account only the actual trajectory (as opposed to projected trajectories). However, the conflict severity indicated by PET is only relevant to the last time frame of the pedestrian-vehicle interaction. PET does not indicate the change in conflict severity over the entire time period of pedestrian-vehicle interaction. Moreover, it describes neither the initial stage of the conflict nor the action taken by the road users (both driver and pedestrian) involved. Studies have also indicated that critical PET value changes with the speed of interacting vehicle (Chaudhari et al. 2021). Thus, interactions having low PET may be less severe if the approaching vehicle speed is low. Nevertheless, PET is still beneficial as a SSM since it provides a definite boundary to differentiate a crash from a non-crash event (Peesapati et al. 2018).

3.3.2. Spatial gap for pedestrians

Pedestrian and vehicle GPS coordinates were used to geolocate them. The distance between the pedestrian crossing a particular lane and the nearest vehicle in that lane at a particular timeframe was defined as the spatial gap in that frame (refer to Fig. 5). Spatial gap gives a continuous value of the conflict severity at each timeframe of the conflict. Spatial gap of zero indicates occurrence of a crash. The risk of conflict increases as the value of PET approaches zero. Spatial gap can be used to assess the extent to which pedestrians are comfortable to be in the spatial proximity of the approaching vehicle. Thus, spatial gap is often used in gap acceptance studies. Similar to PET, spatial gap takes into account the actual trajectories of road users. However, spatial gap does not take into account the speed of the approaching vehicle. Thus, spatial gap does not give an accurate representation of the severity of the conflict.

3.3.3. Temporal gap for pedestrians

Temporal gap calculation used vehicle speed for the frame where a

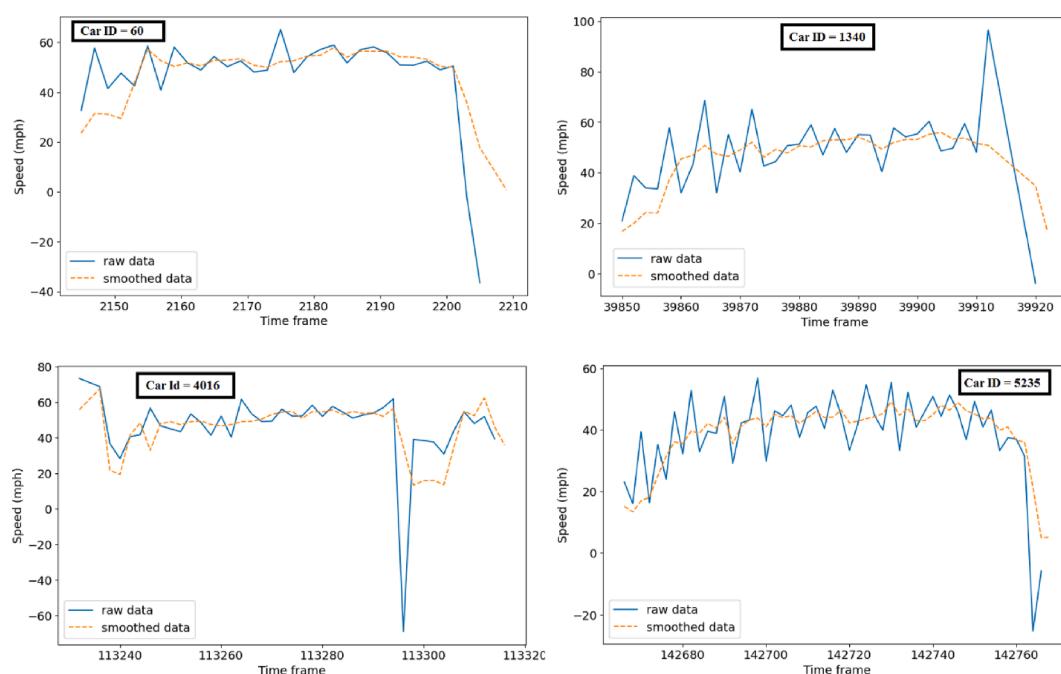


Fig. 4. Examples of Comparison of Speeds Obtained from Raw Data and Smoothed Data.

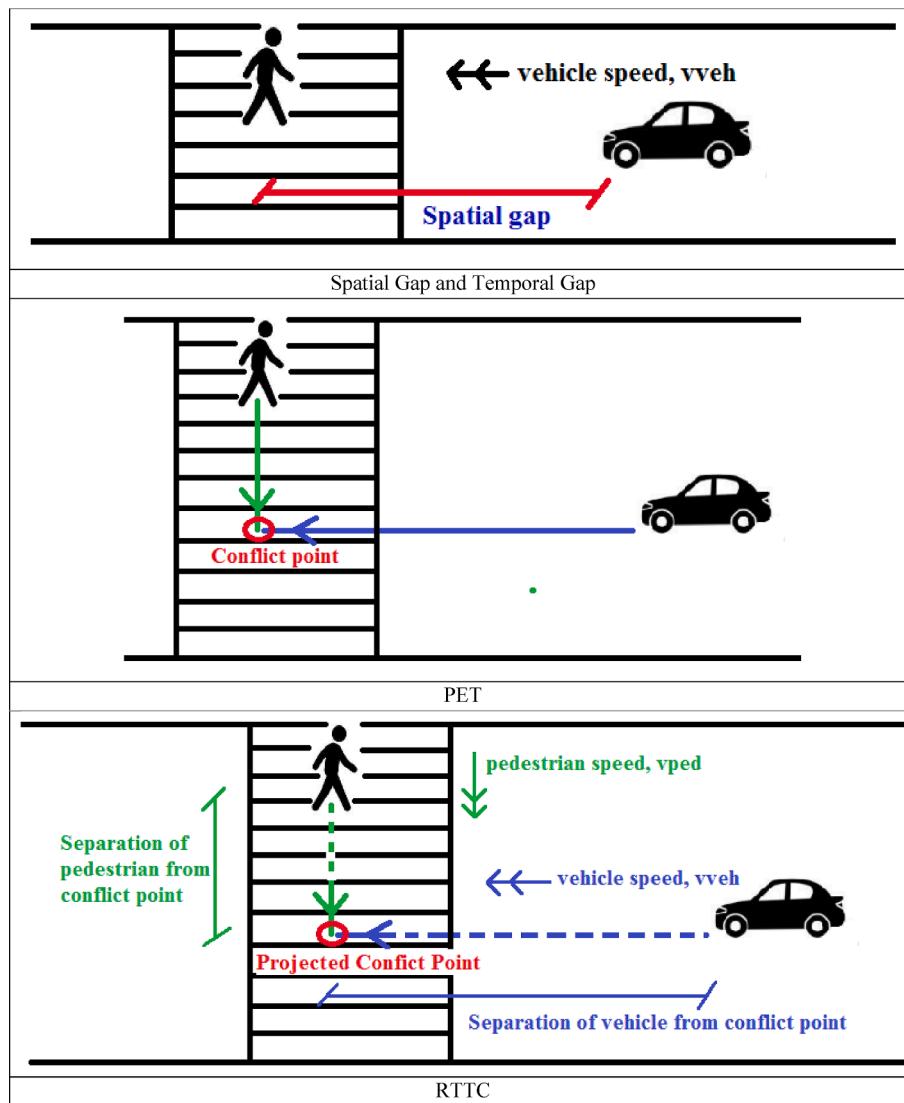


Fig. 5. Different Surrogate Safety Measures used in the Study. Post Encroachment Time (PET).

spatial gap was calculated. The spatial gap was divided by the instantaneous vehicle speed to obtain the temporal gap (in seconds) for a particular frame. Thus temporal gap is the time taken by vehicle to collide with the pedestrian if the pedestrian stopped moving and the vehicle continued its trajectory to the pedestrian at the same speed. Like the spatial gap, temporal gap gives continuous values of traffic conflict severity. In many cases, the temporal proximity is better than spatial proximity, because temporal gap integrates both space and speed simultaneously. For example, a vehicle that is coming to a rolling stop as it approaches a pedestrian may have decreasing spatial proximity. However, the decreasing speed of the vehicle will increase temporal gap and reduce the severity of the conflict. On the other hand, a vehicle moving at a high speed may be at a high spatial proximity but low temporal proximity (and thereby increasing the conflict severity) (Pawar and Patil, 2019). Temporal gap can indicate actions taken by the driver in response to a conflict. For example, a decrease in temporal gap followed by a rise in temporal gap can indicate that the driver is braking. A drastic increase in temporal gap can indicate an event of hard braking. However, temporal gap does not give the full picture of conflict severity because it does not take into account the reaction of the pedestrian.

3.3.4. Relative time to Collision (RTTC)

In each frame, the headway of the vehicles and crossing pedestrians

were extrapolated to identify their theoretical/ projected conflict point. The separation of the pedestrian and vehicle was calculated independently from that conflict point. RTTC was then calculated by the following equation, which is also depicted in the RTTC section of Fig. 5. TTC of the pedestrian and vehicle are calculated separately in each timeframe through the following equations:

$$TTC_p = \frac{\text{separation between pedestrian and conflict point}}{\text{instantaneous pedestrian speed}} \quad (2)$$

$$TTC_v = \frac{\text{separation between vehicle and conflict point}}{\text{instantaneous vehicle speed}} \quad (3)$$

RTTC is the absolute difference between the TTC_p and TTC_v :

$$RTTC = abs(TTC_v - TTC_p) \quad (4)$$

The concept of TTC arose from vehicle-vehicle conflicts. It was originally defined as the time that remains for the paired vehicles before they collide, if both continue at their present speeds along their respective trajectories (Tarko 2018). TTC can be easily detected in the rear-end conflict events because the trajectories of the paired vehicles are assumed to be overlapped. However, it cannot be detected (or does not exist) in most of the interactions if the trajectories of the paired users intersect, for example, the pedestrian-vehicle conflict and the conflict

between left-turn and opposing through vehicles. In the rear-end conflict, the following vehicle will definitely collide with the leader vehicle if the speed of the follower is higher. However, for the pedestrian-vehicle conflict, the cases that the pedestrian and the vehicle occupy the trajectory intersection point at the same moment are rare. Thus, since pedestrians and vehicles do not have continuously overlapping trajectories, the formula of TTC used in vehicle-vehicle conflicts is not applicable for pedestrian-vehicle conflicts (Chen et al. 2017).

Consequently, the literature has used several variations of the TTC to depict pedestrian-vehicle conflict severity. Most of the literature has calculated TTCP and TTCv separately for each timeframe. The larger of the values from TTCP and TTCv is taken as the TTC for that particular timeframe. The argument for taking the larger value is that the road user with the larger TTC will arrive with the conflict point later and collide with the first road user. For example, if $TTCv < TTCP$, then the vehicle arrives first at the conflict point, and the latter user (pedestrian) will collide with vehicle, and vice-versa (Cafiso et al. 2011; Jiang et al. 2015; Kathuria and Vedagiri 2020; Layegh et al. 2020; Ni et al. 2016). However, the problem with this method is that it cannot fully capture the conflict severity. For example, two events having TTCP of 1 s and TTCv of 5 s (Event A) and TTCP of 4.5 s and TTCv of 5 s (Event B) would have the same value of TTC (i.e. 5 s). However, Event B is more critical because the two road users would meet within 0.5 s of each other at the conflict point, if they continue at the same trajectory and speed. This would prompt at least one of the road users to take an evasive action (yielding, lane changing, etc.). The smaller the time difference between TTCP and TTCv, the greater the anticipation of the conflict severity in the minds of the road users. This would prompt them to take more drastic evasive actions. This time difference between the TTC of the two road users has been called different names, such as, Time Advantage (Hansson 1975; Laureshyn et al. 2010), Time Difference Time to Collision (Zhang et al. 2011, 2012, 2014), and Relative Time to Collision (RTTC) (Chen et al. 2017). Therefore, because of the above-mentioned advantage of RTTC, this study has opted to use RTTC as a measure for conflict severity. The values of TTCP and TTCv can change with time for a particular interaction. If absolute values were not used for RTTC calculation, the sign of RTTC would change each timeframe depending on the values of TTCv and TTCP. This would have created problems when trying to find a representative RTTC for each pedestrian-vehicle interaction. The later sections of the paper reveal that the average value of RTTC were taken for modelling purposes. Many studies in the literature take the minimum TTC value for each pedestrian-vehicle interaction (e.g. (Kathuria and Vedagiri 2020)). However, following this method in the current study would mean taking negative RTTC for each pedestrian-vehicle interaction (if absolute value were not used). On the other hand, if the average RTTC were used, the values would be close to zero (because average of a large number of positive and negative values can be close to zero). If the value of average RTTC is close to zero for all interactions, it would be difficult to gauge which interactions are relatively dangerous and which are relatively safer. Thus, this study decided to take the average of absolute RTTC values in the model.

The four SSM investigated in this study have unique advantages, which, when studied together can offer the full picture of conflict severity over the entire period of conflict. For example, before crossing the road, the pedestrian is mentally evaluating the spatial and temporal gaps and looking for an acceptable gap to begin crossing. Once the pedestrian starts crossing, both road users (driver and pedestrian) are evaluating the chances and severity of conflict based on RTTC, spatial gap and temporal gap. RTTC can be calculated only when both road users are moving simultaneously towards the conflict point. After the first road user has crossed the conflict point, the conflict severity for the remaining conflict period is indicated by the PET. Therefore, to get a holistic picture of conflict severity, this study has investigated the aforementioned SSMs.

Analysis of pedestrian-vehicle interactions in this study created more than 50,000 observations, where each observation represented a

timeframe, the corresponding pedestrian-vehicle interaction identities, lane during pedestrian and vehicle interaction, spatial gap, temporal gap and RTTC at the timeframe. These time frames covered a total of 395 pedestrian-vehicle lane interactions. The data from each of the 395 pedestrian-vehicle lane interactions were then averaged over their corresponding timeframes of that particular interaction to create a dataset containing 395 observations. Each of the 395 observations in the final dataset represented a pedestrian-vehicle interaction in a particular lane, and each observation in the final dataset contained PET, average spatial gap, average temporal gap, and average RTTC for that specific pedestrian-vehicle interaction in that lane. In addition, other geometric and operational attributes were considered to create multivariate models to compare surrogate safety features between RRFB and PHB sites. Both linear and non-linear univariate (for spatial gap) and multivariate models (for temporal gap, TTC, PET) were considered in this study. All models were estimated in R using the Vector Generalized Linear and Additive Models (VGAM) package.

3.4. Data Description

The data description is provided in Table 2. All the variable values are within reasonable limits. The land use mix was calculated from land use data extracted from the FDOT land using GIS shape files. Land-use data considered in this study include residential, commercial, industrial, agricultural, institutional, government and miscellaneous areas. The data was extracted using a 1000ft buffer around the study sites, which is comparable to those used in the literature (McConville et al. 2011; Sliva and Williams 2001). The proportions of several types of land uses like residential, commercial, industrial, agricultural, institutional, governmental, etc. within a one-mile buffer around each treatment and control study site were calculated. Besides, the land use mix was computed by:

$$\text{Land use mix} = \frac{-\sum_k^N (p_k(\ln p_k))}{\ln N} \quad (5)$$

where k is the category of land use, p is the proportion of land use category k, N is the number of land use categories. Land use mix can indicate the attractiveness of the site for trips. A greater land use mix is expected to provide a greater range of purposes to visit the site, and consequently generate more pedestrian and traffic flow (Goswamy and Abdel-Aty, 2023).

The concept of pedestrian starting position (before crossing) used in this study is illustrated in Fig. 6. The starting position of each pedestrian before crossing varies from near end to far end with respect to the direction of flow of the interacting vehicle. The red arrows and blue arrows indicate the trajectory path of pedestrians and vehicles respectively. The period type was divided into three dummy variables, namely, Morning dummy, Afternoon dummy and Evening dummy.

The 15-minute time blocks considered in this study are the four 15-minute slices in a particular hour. For example, 8:00–8:15 am, 8:15–8:30 am, 8:30–8:45 am, 8:45–9:00 am, etc. The speed of all pedestrians crossing during each of the 15-minute time block is calculated. The speeds of all vehicles were calculated from vehicle trajectories. Pedestrian-vehicle interactions are identified based on cases when pedestrians and vehicles are on the road segment at the same time. During such interactions, the speed of all interacting vehicles during that 15-minute time block are considered in the model.

In this study, pedestrian crossing behavior was divided into three types, namely, proper crossing, unsafe crossing and jaywalking. A pedestrian follows a proper crossing if he/she crosses using the crosswalk at the pedestrian signal phase. A pedestrian is considered as unsafe if he/she crosses using the crosswalk, but not during the pedestrian signal phase. All other pedestrians are considered as jaywalkers. This study did not have sufficient number of jaywalkers to put them in a separate category. Hence, unsafe pedestrians and jaywalkers were

Table 2
Descriptive Statistics of Variables used in Modelling.

Variable	Description	Value	count	mean	std	min	max
Dependent variables:							
Spatial Gap	Average Spatial Gap (ft) for each pedestrian-vehicle interaction per lane	Continuous variable	395	116.38	62	6.31	344.88
Temporal Gap	Average Temporal Gap (s) for each pedestrian-vehicle interaction per lane	Continuous variable	395	11.79	10.2	1.19	59.93
RTTC	Average RTTC (s) for each pedestrian-vehicle interaction per lane	Continuous variable	395	6.46	4.89	1.2	33.28
PET	PET (s) for each pedestrian-vehicle interaction per lane	Continuous variable	395	8.48	7.2	0.72	47.83
Independent variables:							
PHB/ RRFB presence	Presence of PHB or RRFB	0: RRFB 1: PHB	395	–	–	0	1
Weekend dummy	Presence of weekday or weekend	0: Weekday 1: Weekend	395	–	–	0	1
Period type dummies:							
Morning dummy	Period of the day Dummy variable for morning period (8–10 am)	0: Otherwise 1: Morning	395	–	–	0	1
Afternoon dummy	Dummy variable for afternoon period (11:30–2 pm)	0: Otherwise 1: Afternoon	395	–	–	0	1
Evening dummy	Dummy variable for evening period (5:30–9 pm)	0: Otherwise 1: Evening	395	–	–	0	1
Signal activation	Activation of pedestrian signal by pedestrian	0: Not activated 1: Activated	395	–	–	0	1
Total lane count	Total number of lanes on road considering both directions	Discrete variable	395	3.12	0.99	2	4
Pedestrian crossing speed	Average pedestrian crossing speed (mph) during the 15-minute time block covering the interaction	Continuous variable	395	4.11	1.27	2.15	9.84
Average vehicle lane interaction speed	Average vehicle travel speed (mph) while interacting with pedestrians in that particular lane during the 15-minute time block covering the interaction	Continuous variable	395	18.25	9.48	0.72	62.31
Log (Average vehicle lane interaction speed)	Natural logarithm of average vehicle lane interaction speed	Continuous variable	395	2.90	0.64	-0.32	4.13
Land use mix	Land use mix	Continuous variable	395	0.4	0.17	0	0.56
Pedestrian starting position	Position of the crossing pedestrian with respect to the direction of movement of interaction vehicle	0: Near End 1: Far End	395	–	–	0	1

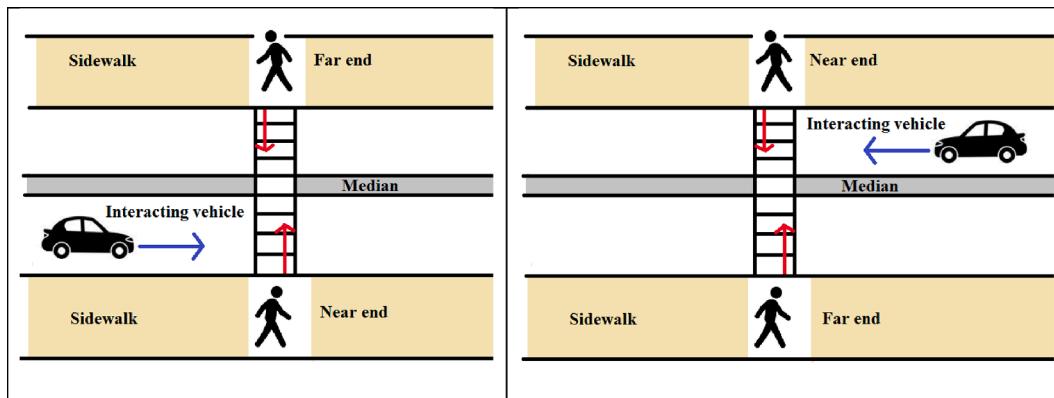


Fig. 6. Position of near end and far end pedestrians with respect to interacting vehicle.

considered together in this study. Unsafe pedestrians and jaywalkers crossed illegally or unexpectedly. Neither the unsafe pedestrian nor the jaywalker pressed the button to start pedestrian signal phase. So, for those pedestrians, the *Signal Activation* variable would have a value of zero (Table 2). Hence, this study evaluated illegally crossing pedestrians using the *Signal Activation* variable.

Histograms of the considered SSMS are presented in Fig. 7 and indicate that the data points are positively skewed. Previous literature has indicated that the positively skewed distributions can be fit well by the gamma, inverse Gaussian and the lognormal regression models. Moreover, these models have been specifically chosen in this study because they can only handle positive values, which fit the characteristics of the considered dependent variables in this study (Jorgensen, 2012; Pawar and Patil, 2016; Vilca et al., 2014). Hence, the non-linear

regression models are explored and are compared with linear regression models in this study.

The correlation matrix of the considered variables is shown in Fig. 8. The results reveal that there were no multi-collinearity among the considered variables (Yoo et al. 2014). Further tests for multicollinearity are presented in the Appendix.

3.5. Statistical methodologies

3.5.1. Mann-Whitney-Wilcoxon test

The Mann-Whitney-Wilcoxon Test (also known as Mann-Whitney *U* test; or the Wilcoxon rank-sum test) is an analog to the two-sample *t*-test for independent samples, in which the actual values are replaced by rank scores. Under the Mann-Whitney *U* test, here, the responses are assumed

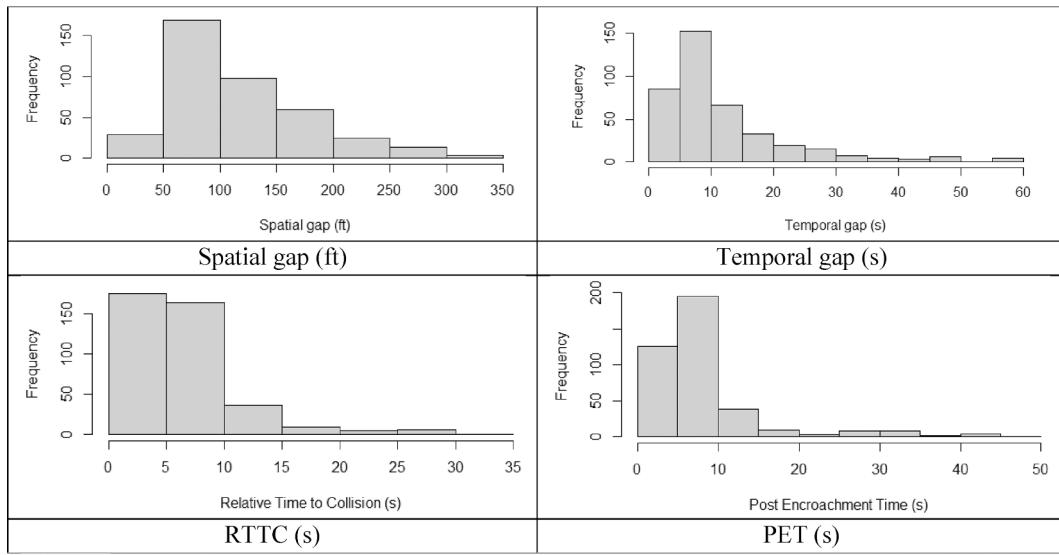


Fig. 7. Distribution of Considered SSMs.

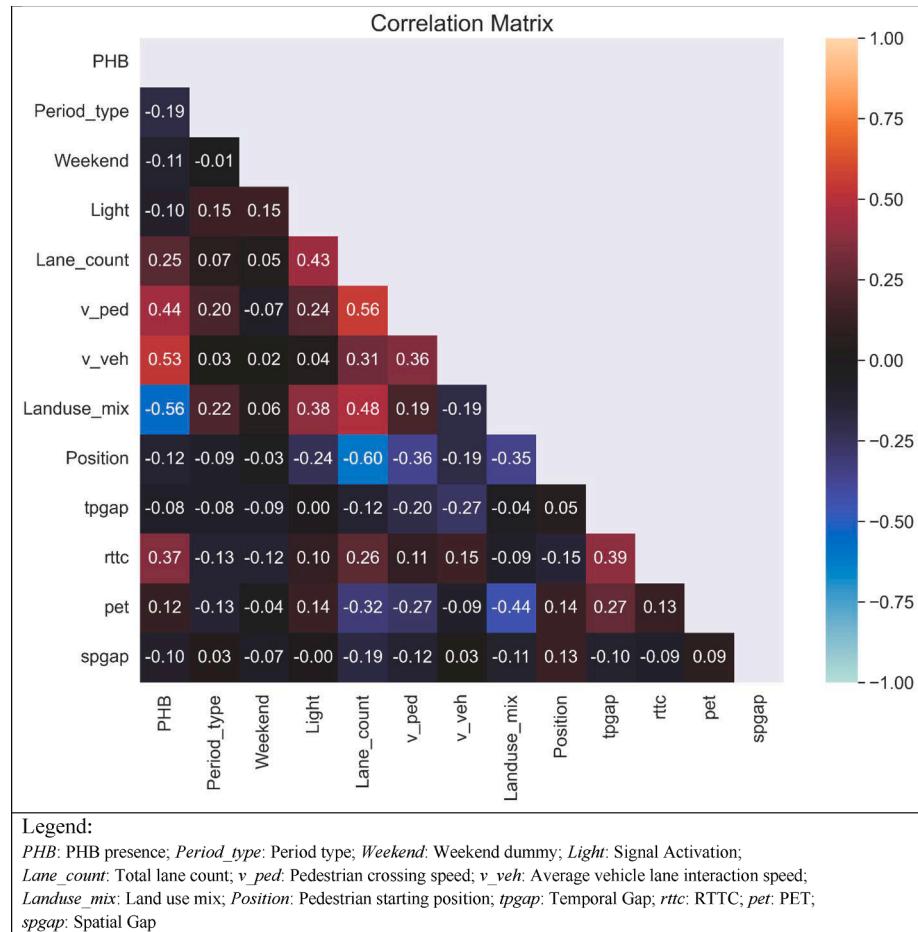


Fig. 8. Correlation Matrix of Considered Variables.

to be continuous. The null hypothesis assumes that the medians from both data are the same. If the null hypothesis is rejected, i.e., p-value < 0.05, then the medians and the relevant samples are considered to be significantly different (Rosner, 2015). The study has not used t-tests because t-tests are more relevant to normally distributed data.

3.5.2. Multivariate linear regression

The multivariate regression has the form

$$y_{ik} = b_{0k} + \sum_{j=1}^{\rho} b_{jk} x_{ij} + e_{ik} \quad (5)$$

for $i \in \{1, 2, \dots, n\}$, and $k \in \{1, 2, \dots, m\}$

where,

1. $y_{jk} \in R$ is the k-th real valued response for the i-th observation.
2. $b_{0k} \in R$ is the regression intercept for the k-th response.
3. $b_{jk} \in R$ is the j-th predictor's regression slope for i-th observation
4. $x_{ij} \in R$ is the j-th predictor for the i-th observation
5. $(e_{11}, \dots, e_{im}) N(0_m, \Sigma)$ is the multivariate Gaussian error vector.

Some measures of the goodness-of-fit including Log-Likelihood, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to find the best fit for the data and determine the best model (Timm, 2002).

3.5.3. Multivariate gamma regression

A random variable Y follows a gamma distribution if its probability density function is given by:

$$f(\alpha, \lambda) = \frac{\lambda^\alpha y^{\alpha-1} e^{-\lambda y}}{\Gamma(\alpha)} I(0, \infty)(y), \quad (6)$$

where $\alpha, \lambda > 0$, $\Gamma(\cdot)$ denotes the gamma function, and $I(\cdot)$ is the indicator function. Given that $\lambda = \alpha/\mu$, the gamma distribution function can be re-parameterized as a function of the mean (μ) and shape (α) parameters and written in the following form:

$$f(y|\mu, \alpha) = \frac{1}{y\Gamma(\alpha)} \left(\frac{\alpha y}{\mu}\right)^\alpha e^{-\frac{\alpha y}{\mu}} I_{(0,\infty)}(y) \quad (7)$$

where $\mu, \alpha > 0$, $\Gamma(\cdot)$ denotes the gamma function, and $I(\cdot)$ is the indicator function. Gamma regression models can be determined by optimizing the relevant likelihood function (Bhaumik et al. 2009). The multivariate concept of the gamma regression is similar to that of the linear regression.

3.5.4. Multivariate inverse Gaussian regression

The inverse Gaussian distribution (also known as the Wald distribution) is a two-parameter family of continuous probability distributions with support on $(0, \infty)$. Its probability density function is given by:

$$f(y; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi y^3}} \exp\left(-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right) \quad (8)$$

for $y > 0$, where $\mu > 0$ is the mean and $\lambda > 0$ is the shape parameter.

The parameters μ and λ can be estimated using the relevant log-likelihood function (Cheng and Amin, 1981). The multivariate concept of inverse Gaussian regression is similar to that of linear regression. Similar link functions can also be applied to the inverse Gaussian regression. This study used the log link function for both gamma and inverse Gaussian regressions to facilitate the interpretation of coefficient estimates.

3.5.5. Multivariate lognormal regression

A lognormal distribution (also known as Galton's distribution) is a continuous probability distribution of a random variable whose logarithm is normally distributed. A random variable Y follows a lognormal distribution if its probability density function is given by (Johnson et al., 1995):

$$f(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{\ln(y)-\mu}{\sigma}\right)^2\right\} \quad (9)$$

where μ is the location parameter and σ is the scale parameter of the distribution, which can be estimated by optimizing the relevant likelihood functions. The multivariate concept of a lognormal regression is similar to that of linear regression.

4. Results

4.1. Mann-Whitney-Wilcoxon test

Since the data are non-normal, the study used Mann-Whitney U test to assess the difference in medians of SSMs, namely spatial gap, temporal gap, RTTC and PET among RRFB and PHB sites for various conditions as depicted in Table 3. The difference in median is obtained by subtracting the median in the PHB case from that of the RRFB case. So, a negative median difference means that the PHB median value of the SSM is higher and vice-versa. Mann-Whitney tests give an introductory diagnosis to justify the investigation of differences in SSMs among the RRFB and PHB groups. The test by itself is not sufficient to draw the final conclusions, however, it sets the premise for investigating the differences in SSMs among the RRFB and PHB sites. Based on the significant cases presented in Table 3, larger median values for the spatial gap, temporal gap, RTTC and PET have been observed in 15 out of 28, 15 out of 32, 30 out of 33 and 2 out of 8 cases respectively. The low number of significant cases for PET can be attributed to the fact that only one PET value occurs per interaction per lane, leading to a small sample size for PET. Based on the above numbers it is ambiguous whether RRFB or PHB sites offer better safety. However, this may be attributed to the confounding factors, such as traffic volume, speed, and other operational and geometric features. Thus, to accurately assess the difference in SSM values in RRFB and PHB sites, this study has used multivariate modeling to consider a variety of factors, which are elaborated in the following sections.

4.2. Regression analysis results

4.2.1. Analysis of spatial gaps

Spatial gaps (ft) were modeled separately from the other three time-based SSMs (i.e. temporal gap, RTTC, PET), because their dimension was different from those of other SSMs. The results of the linear model and generalized linear models are presented in Table 4:

Based on the goodness of fit measures shown in Table 4, lognormal regression performed the best, followed by the inverse Gaussian regression, linear regression and gamma regression. Even though the Gamma regression model has the largest number of significant variables, it cannot be chosen as the optimum model because of its poor goodness of fit. Thus subsequent discussion on spatial gap will be based on the lognormal regression model results. The period type was divided into three dummy variables, namely, Morning dummy, Afternoon dummy and Evening dummy. The only dummy found to have significant impact in the model was the Morning dummy, hence the other two dummies were removed from the final model. As per Table 4, for spatial gaps, the significant factors were the presence of PHB, and the variables indicating weekend, signal activation, vehicle interaction speed, morning period, and land use mix. Moreover, the interaction between the PHB sites and the morning period was found to be statistically significant. Spatial gap was higher during weekdays, at signal activation, larger vehicle interaction speeds and during the morning period. On the contrary, spatial gap was lower at PHB sites and at a greater variety of land use mix. Moreover, the interaction of the PHB with the morning period decreased spatial gaps.

4.2.2. Multivariate analysis of temporal gap, RTTC and PET

Temporal gap (s), RTTC (s) and PET (s) were modeled together in a multivariate model. Multivariate analysis of variance (MANOVA) was done to assess whether the considered explanatory variables jointly affected the three temporal SSMs. The results of MANOVA are provided in Table 5. As per Table 5, all the considered factors have significantly affected the temporal SSMs, which justifies the usage of a multivariate analysis for the temporal SSMs (Timm, 2002).

Results of the multivariate modeling are provided in Table 6:

Table 3

Results of Mann-Whitney-Wilcoxon test on Comparing SSMs among RRFB and PHB sites.

Pair	Spatial Gap (ft)		Temporal Gap (s)		RTTC (s)		PET (s)	
	Test statistic (p-value)	Difference in Median	Test statistic (p-value)	Difference in Median	Test statistic (p-value)	Difference in Median	Test statistic (p-value)	Difference in Median
1	RRFB vs PHB	1.75 (0.19)	-4.19	420.49 (*)	-2.64	3695.21 (*)	-3.77	0.003 (0.96)
2	RRFB weekday vs PHB	564.65 (*)	27.16	20.14 (*)	-1.2	281.58 (*)	-1.59	0.31 (0.58)
3	RRFB weekend vs PHB	230.28 (*)	-17.26	271.4 (*)	-2.84	3007.46 (*)	-3.8	0.08 (0.77)
4	RRFB Before sunset vs PHB	525.51 (*)	-48.52	900.1 (*)	6.3	591.08 (*)	-4.09	1.47 (0.22)
	Before sunset							
5	RRFB After sunset vs PHB	27.27 (*)	21.25	55.6 (*)	1.68	43.78 (*)	-2.17	696 (*)
	After sunset							
6	RRFB No Median vs PHB No	15.57 (*)	-1.82	649.69 (*)	-3.91	4347.24 (*)	-4.21	19.81 (*)
	Median							
7	RRFB Median vs PHB Median	0.04 (0.85)	-1.48	226.15 (*)	4.41	280.18 (*)	-2.19	1.26 (0.26)
8	RRFB Morning vs PHB	1433.02 (*)	61.7	280.15 (*)	-6.58	240.2 (*)	-3.78	0.04 (0.85)
	Morning							
9	RRFB Afternoon vs PHB	250.16 (*)	-23.63	42.12 (*)	-0.41	778.06 (*)	-3.91	2.28 (0.13)
	Afternoon							
10	RRFB Day vs PHB Day	486.52 (*)	16.61	1075.94 (*)	-5.56	3224.7 (*)	-3.92	3.43 (0.064)
11	RRFB Evening vs PHB	436.66 (*)	-35.69	782.18 (*)	4.6	464.7 (*)	-3.25	6.91 (*)
	Evening							
12	RRFB Weekday morning vs	1474.1 (*)	64.73	291.8 (*)	-6.66	661.11 (*)	-4.24	0.09 (0.77)
	PHB Weekday morning							
13	RRFB Weekday afternoon vs	73.42 (*)	-18.88	15.61 (*)	-0.24	352.91 (*)	-3.34	1.72 (0.18)
	PHB Weekday afternoon							
14	RRFB Weekday Day vs	494.87 (*)	27.9	207.49 (*)	-4.91	862 (*)	-3.71	1.01 (0.31)
	Weekday Day							
15	RRFB Weekday Evening vs	0.33 (*)	-1.69	987.36 (*)	6.5	97.39 (*)	1.78	4.97 (*)
	PHB Weekday Evening							
16	RRFB Weekend Day vs PHB	30.55 (*)	6.02	460.11 (*)	-4.09	1949.87 (*)	-3.76	2.37 (0.12)
	Weekend Day							
17	RRFB Weekend Evening vs	210.02 (*)	-32.99	30.45 (*)	0.88	671.54 (*)	-4.96	4.83 (*)
	PHB Weekend Evening							
18	RRFB Weekday Before Sunset	15.07 (*)	27.13	858.38 (*)	7.45	93.49 (*)	2.58	0.48 (0.49)
	vs PHB Weekday Before							
	Sunset							
19	RRFB Weekday After Sunset	1.12 (0.29)	-10.88	70.93 (*)	4.34	10.85 (*)	1.28	4.55 (*)
	vs PHB Weekday After Sunset							
20	RRFB Weekend Before	397.4 (*)	-54.28	15.76 (*)	0.53	605.69 (*)	-5.84	1.87 (0.17)
	Sunset vs PHB Weekend							
	Before Sunset							
21	RRFB Weekend After Sunset	45.79 (*)	26.92	24.98 (*)	1.46	93.53 (*)	-1.44	2.68 (0.1)
	vs PHB Weekend After Sunset							
22	RRFB No Median, Day vs	709.98 (*)	24.63	1435.95 (*)	-6.7	3614.25 (*)	-4.4	21.76 (*)
	PHB No Median, Day							
23	RRFB No Median, Evening vs	458.35 (*)	-39.62	722.25 (*)	4.94	357.57 (*)	-3.12	0.42 (0.52)
	PHB No Median, Evening							
24	RRFB Median, Day vs PHB	1.67 (0.2)	4.01	317.03 (*)	6.18	222.45 (*)	-2.46	0.87 (0.35)
	Median, Day							
25	RRFB Median, Evening vs	6.17 (*)	-17.66	1.018 (0.31)	0.46	66.27 (*)	-1.59	0.13 (0.72)
	PHB Median, Evening							
26	RRFB No Median, before	460.06 (*)	-48.96	795.78 (*)	6.45	392.21 (*)	-4.3	0.65 (0.42)
	sunset vs No Median, before							
	sunset							
27	RRFB No Median, after	26.42 (*)	22.58	53.6 (*)	2.09	9.47 (*)	-0.56	0.04 (0.84)
	sunset vs PHB No Median,							
	after sunset							
28	RRFB Median, before sunset	84.99 (*)	-60.62	71.95 (*)	5.48	106.72 (*)	-2.87	0.06 (0.8)
	vs Median, before sunset							
29	RRFB Median, after sunset vs	23.77 (*)	40.6	15.16 (*)	-2.45	17.28 (*)	-0.86	0.74 (0.39)
	PHB Median, after sunset							
30	RRFB Weekday, No median	728.93 (*)	39.42	77.68 (*)	-2.61	267.93 (*)	-1.74	1.6 (0.21)
	vs PHB Weekday, No median							
31	RRFB Weekday, median vs	8.88 (*)	-20.37	282.53 (*)	9.29	16.48 (*)	-0.68	0.18 (0.67)
	PHB Weekday, median							
32	RRFB Weekend, No median	216 (*)	-17.17	415.06 (*)	-4.56	3147.39 (*)	-3.94	21.04 (*)
	vs PHB Weekend, No median							
33	RRFB Weekend, median vs	0.66 (0.42)	-0.23	64.02 (*)	2.56	325.21 (*)	-2.87	1.33 (0.25)
	PHB Weekend, median							

*: significant at 95% confidence level.

Table 4
Results of Univariate Modelling on Spatial Gap.

Variable	Estimate (p-value) for Different Regression Models			
	Linear (Gaussian)	Gamma ^a	Inverse Gaussian ^a	Lognormal
Intercept	142.523 (<0.001)*	5.007 (<0.001)*	4.956 (<0.001)*	5.008 (<0.001)*
PHB dummy	-44.220 (<0.001)*	-0.470 (<0.001)*	-0.413 (<0.001)*	-0.442 (<0.001)*
Weekend dummy	-14.280 (0.025)*	-0.174 (<0.001)*	-0.114 (0.025)*	-0.177 (0.001)*
Signal activation	15.482 (0.057)*	0.138 (<0.001)*	0.095 (0.165)	0.144 (0.037)*
log (Average vehicle lane interaction speed)	17.870 (0.002)*	0.151 (<0.001)*	0.169 (<0.001)*	0.136 (0.005)*
Land use mix	-123.658 (<0.001)*	-1.232 (<0.001)*	-1.010 (<0.001)*	-1.192 (<0.001)*
Dummies for Period type:				
Morning dummy	26.052 (0.050)*	0.213 (<0.001)*	0.183 (0.076)	0.209 (0.064)
(PHB * Morning dummy) interaction	-31.158 (0.042)*	-0.225 (<0.001)*	-0.228 (0.062)	-0.234 (0.071)
Goodness of Fit				
AIC	4359.006	12759.920	4234.377	589.685
BIC	4394.816	12791.750	4298.039	625.495

a: log link .. significant at 90% confidence level *: significant at 95% confidence level.

Table 5
Results of Pillai test Statistic from Type II MANOVA Tests.

Explanatory variable	test stat	approximate F	P-value [Pr
PHB/ RRFB presence	0.256	44.03	<0.001
Weekend Dummy	0.019	2.505	0.059
Signal activation	0.243	41.088	<0.001
Total lane count	0.200	31.902	<0.001
Pedestrian crossing speed	0.045	5.949	0.001
log (Average vehicle lane interaction speed)	0.149	22.39	<0.001
Land use mix	0.043	5.689	0.001
Pedestrian starting position	0.025	3.308	0.020

4.3. Model interpretation

4.3.1. Findings based on temporal SSMs

Based on the goodness of fit measures shown in Tables 6, lognormal regression performed the best, followed by the inverse Gaussian, linear, then gamma regression. Hence, the following analysis and discussions are based on the lognormal regression model results.

As per Table 6, for temporal gap, the significant factors were the presence of PHB, and variables indicating signal activation, total lane count, pedestrian crossing speed, vehicle interaction speed, land use mix, and pedestrian starting position. The presence of the PHB, activation of pedestrian signal, and a greater variety of land use mix increased the temporal gap, and hence improve the overall safety for pedestrians. On the other hand, the temporal gap decreased when pedestrian and vehicle speeds increased, at roads having more lanes, and when pedestrians crossed starting from the end relatively further from the direction of the approaching vehicle. The coefficient estimates of the lognormal regression model can be interpreted using the following examples. Compared to RRFB sites, PHB sites have increased average temporal gap, average RTTC, and PET by a factor of $\exp(0.649)$, i.e. 1.914 times, 2.298 times and 1.218 times respectively.

For RTTC models, the significant factors were the presence of PHB, and variables indicating weekday presence, signal activation, total lane

Table 6
Results of Multivariate Modelling on Temporal Gap, RTTC, PET.

Variable	Estimate (p-value) for Different Regression Models			
	Linear (Gaussian)	Gamma ^a	Inverse Gaussian ^a	Lognormal
Temporal Gap (s)				
Intercept	31.434 (<0.001)*	3.583 (<0.001)*	4.067 (<0.001)*	3.533 (<0.001)*
PHB dummy	5.878 (0.005)*	0.653 (<0.001)*	0.550 (0.001)*	0.649 (<0.001)*
Weekend dummy	-1.727 (0.088)*	-0.146 (<0.001)*	-0.040 (0.579)	-0.112 (0.111)
Signal activation	2.103 (0.118)	0.200 (<0.001)*	0.155 (0.168)	0.252 (0.007)*
Total lane count	-0.628 (0.472)	-0.101 (<0.001)*	-0.090 (0.225)	-0.155 (0.011)*
Pedestrian crossing speed	-1.631 (0.002)*	-0.132 (<0.001)*	-0.199 (<0.001)*	-0.140 (<0.001)*
log (Average vehicle lane interaction speed)	-6.328 (<0.001)*	-0.513 (<0.001)*	-0.437 (<0.001)*	-0.455 (<0.001)*
Land use mix	5.762 (0.305)	1.156 (<0.001)*	0.595 (0.242)	1.151 (0.003)*
Pedestrian starting position	-1.736 (0.223)	-0.158 (0.001)*	-0.205 (0.084)	-0.235 (0.017)*
RTTC (s)				
Intercept	4.177 (0.003)*	1.543 (<0.001)*	1.438 (<0.001)*	1.449 (<0.001)*
PHB dummy	5.781 (<0.001)*	0.832 (<0.001)*	0.875 (<0.001)*	0.832 (<0.001)*
Weekend dummy	-1.154 (0.013)*	-0.202 (<0.001)*	-0.547 (<0.001)*	-0.289 (<0.001)*
Signal activation	1.261 (0.042)*	0.157 (0.007)*	0.094 (0.119)	0.126 (0.099).
Total lane count	0.870 (0.031)*	0.144 (<0.001)*	0.322 (<0.001)*	0.211 (<0.001)*
Pedestrian crossing speed	-1.041 (<0.001)*	-0.161 (<0.001)*	-0.164 (<0.001)*	-0.168 (<0.001)*
log (Average vehicle lane interaction speed)	-0.638 (0.126)	-0.088 (0.021)*	-0.054 (0.266)	-0.082 (0.111)
Land use mix	3.991 (0.124)	0.293 (0.227)	-0.201 (0.522)	-0.005 (0.986)
Pedestrian starting position	-0.266 (0.686)	-0.053 (0.409)	-0.065 (0.269)	-0.073 (0.363)
PET (s)				
Intercept	24.480 (<0.001)*	3.202 (<0.001)*	3.203 (<0.001)*	3.190 (<0.001)*
PHB dummy	1.639 (0.188)	0.253 (0.001)*	0.114 (0.282)	0.197 (0.068).
Weekend dummy	-1.040 (0.082)	-0.063 (0.088)	-0.116 (0.033)*	-0.072 (0.168)
Signal activation	8.152 (<0.001)*	0.727 (<0.001)*	0.535 (<0.001)*	0.580 (<0.001)*
Total lane count	-1.834 (<0.001)*	-0.204 (<0.001)*	-0.123 (0.006)*	-0.155 (0.001)*
Pedestrian crossing speed	-1.148 (<0.001)*	-0.114 (<0.001)*	-0.131 (<0.001)*	-0.137 (<0.001)*
log (Average vehicle lane interaction speed)	-1.547 (0.004)*	-0.104 (<0.001)*	-0.086 (0.076)	-0.097 (0.036)*
Land use mix	-19.148 (<0.001)*	-1.198 (<0.001)*	-1.045 (0.001)*	-1.145 (<0.001)*
Pedestrian starting position	-2.717 (0.001)*	-0.231 (<0.001)*	-0.103 (0.15)	-0.174 (0.018)*
Goodness of Fit				
AIC	7686.996	7845.372	6516.867	1991.357
BIC	7806.363	7952.802	6731.727	2110.724

a: log link.: significant at 90% confidence level *: significant at 95% confidence level.

count, and pedestrian crossing speed. The presence of PHB, signal activation, and greater lane count increased RTTC. RTTC was higher during weekdays. However, RTTC was lower at greater pedestrian speeds.

For PET, the significant factors were the presence of PHB, signal activation, total lane count, pedestrian crossing speed, vehicle interaction speed, land use mix, and pedestrian starting position. The presence of PHB, activation of pedestrian signal increased PET, while PET was lower at larger pedestrian crossing speeds, larger vehicle speeds, greater land use mix and when pedestrians crossed starting from the end relatively further from the direction of the approaching vehicle.

4.3.2. Comparison of influence of various predictor variables on different SSMs

Based on the goodness of fit measures, both the finally selected spatial gap and temporal gap models performed similarly. However, spatial gap model had fewer significant variables. Moreover, the interpretation of the regression coefficient estimates for modelling spatial gap is not straightforward. This may be because of the variable distances, at which vehicles yielded to allow the pedestrians to pass. Vehicles stopping were observed to usually have lower spatial gaps as they stopped at the stop line, while vehicles slowing down were further away. Based on the results presented in [Tables 4–6](#), the larger temporal SSM values, but smaller spatial SSM values, were found in PHB sites compared to RRFB sites. Moreover, SSM values increased during weekdays, indicating more pedestrian safety during that time. This may be because of higher vehicle flows during the weekdays, which reduced vehicle travel speeds and interaction speeds. The activation of the PHB signal increased pedestrian safety. Pedestrian crossing speed was found to increase in times of low SSM values, indicating that pedestrians crossed at faster speeds when they perceived greater danger during crossing. In fact, pedestrians were observed to be running in several cases. On the other hand, it may be argued that pedestrians who cross by running tend to accept gaps more easily, which is also confirmed by ([Pawar and Yadav, 2022](#)). Nevertheless, running poses a distraction as this may injure pedestrians ([Li et al., 2013](#)). Total lane count, land use mix, and vehicle interaction speed have been found to have ambiguous effects on the SSMs. This may be because the SSMs are affected by both pedestrians and drivers, whose characteristics vary from person to person. [Table 6](#) reveals larger temporal gaps but smaller PETs at sites having greater land use variety. One possible explanation of this phenomenon may be that the temporal gaps take into account the interactions as long as the pedestrian has not reached the conflict point. However, PET also takes into account the interaction after the pedestrian crosses the conflict point and until the following vehicle reaches the conflict point. This may indicate that vehicles decelerate more before pedestrians reach conflict point, then accelerate faster after pedestrian has crossed the conflict point in areas having greater variety of land use. The interpretation about deceleration is supported by [Table 4](#) results where vehicles in areas having greater land use mix have smaller spatial gaps. Thus areas having greater variety of land use may indicate greater urgency for vehicles to travel.

4.3.3. Interesting observations based on the spatial gap model

As per [Table 4](#), spatial gaps have been found to increase in cases having larger interaction speeds. To address this seemingly contradictory result, this study has looked at the interaction of PHB sites with vehicle interaction speed. Spatial gap seems to increase at PHB sites when vehicle interaction speed is large. However, the coefficient estimate of the interaction term was not found to be significant, and was hence dropped from the final model. The above results lead to the possible emergence of two types of spatial gaps. If the perceived spatial gap is too low, pedestrians either wait for the vehicles to pass the crossing, or drivers decide to pass through the crossing without stopping seeing they cannot yield in time. Such vehicles passing during the signal activation stage are known as violators. Another category of spatial gap

occurs when the spatial gap is high. During that time, if the vehicle interaction speed is high, pedestrians may cross the road only if they feel the perceived spatial gap is acceptable (even during signal activation phase). Such findings are in line with ([Pawar and Patil, 2015](#)), who concluded that the probability of pedestrians accepting a spatial gap for crossing decreases with the increase in the vehicle speed. This indicates that as vehicle interaction speed increases, pedestrians will only be comfortable with higher spatial gaps. These effects are more pronounced at RRFB sites, possibly because vehicles have a greater probability of failing to register the activation of pedestrian signals in RRFB sites ([Ugan et al., 2022](#)). However, such actions vary from person to person (both drivers and pedestrians). Another possible explanation for the low spatial gap in the PHB sites may be that pedestrians feel more comfortable to cross PHB sites than RRFB sites when vehicles come too close to the crosswalk. This may be because PHB signal is more visible than an RRFB one, and hence drivers may be alerted to decelerate quickly at PHB sites. Thus, vehicles may be more likely to slow or come to a rolling stop as they get closer to the pedestrians. This is supported by the larger temporal gaps and RTTC at PHB sites as shown in [Table 6](#).

The significant interaction of term of PHB with the morning period in [Table 4](#) indicates PHB sites have lower spatial gaps during the evening period compared to the morning period. This is completely opposite to the overall direct effect of the morning period type has on spatial gaps. Again here, the interpretation is not straightforward. Morning rush hours have higher traffic flow and lower interaction speeds. This may indicate that pedestrians are more comfortable at PHB sites than at RRFB sites to cross the road as vehicles come closer to them at lower interaction speeds. This may indicate that during morning rush hours, PHB provides more safety benefits than RRFB sites do.

4.4. Practical applications

The methodology followed in this study can be used to evaluate and compare among multiple pedestrian treatment facilities using SSMs. Moreover, the model coefficients developed in this study can be used to calculate SSMs and compare the relative safety among different pedestrian-vehicle interactions. If reliable thresholds for SSMs are used, the interactions can be classified under different near-crash severities.

5. Conclusions

This study evaluated and compared four types of SSMs, namely, spatial gap, temporal gap, RTTC and PET between two advanced pedestrian mid-block crossing facilities, namely RRFB and PHB. The study used CV processing and data processing of video camera recordings of five selected PHB and two selected RRFB crossings to calculate SSMs, which were aggregated at the timeframe level to create a robust dataset, where each observation represented a timeframe, the corresponding pedestrian-vehicle interaction ID, lane, spatial gap, temporal gap and RTTC at the timeframe. Initial Mann-Whitney U-tests performed in this dataset revealed significant differences in SSMs among the RRFB and PHB sites. Multivariate models (temporal gap, RTTC, PET) and separate regression model (spatial gap) considering both linear regression and non-linear models (gamma regression and inverse Gaussian regression) were estimated after compressing the dataset to obtain the PET and average spatial gap, temporal gap, and RTTC for each pedestrian-vehicle interaction per lane. The dataset was further consolidated by adding various roadway geometric, operational, and land use factors. The models revealed that the non-linear regression performed better than the linear regression, because the SSMs followed a skewed distribution instead of a normal distribution. In fact, the lognormal regression performed the best among all considered models, which is also in line with results in ([Pawar and Patil, 2016](#)), who visualized the distribution of temporal gap and spatial gap. The models revealed that the presence of PHB increased temporal gap, RTTC, PET, but reduced spatial gap. Thus, the results provide definite proof that

PHB offers more safety in terms of the temporal SSMs. The results of this study confirm conclusions drawn from various Crash Modification Factor (CMF) studies where CMF for pedestrian crashes varied from 0.244 to 0.630 and from 0.640 to 0.840 for PHB and RRFB sites, respectively. Thus CMF studies indicated higher crash reductions at PHB sites (Fitzpatrick et al. 2021; Monsere et al. 2017, 2020; Zegeer et al. 2017). Goswamy et al. (2023), calculated a CMF of 0.31 for pedestrian crashes at RRFB sites, compared to unsignalized zebra crossings (Goswamy et al. 2023b).

However, the results from analysis of spatial gaps is not so straightforward. The results of this study suggest heterogeneous behavior among pedestrians when perceiving the spatial gap. If the spatial gap is too low, pedestrians do not tend to cross the road. On the other hand, during crossing, if the perceived temporal gaps are too low, pedestrians may run (possibly in interactions involving low temporal gaps but high spatial gaps). These effects are more pronounced in RRFB sites. The study reveals that pedestrians accepted larger spatial gaps when crossing RRFB rather than when crossing PHB. On the other hand, the results can also be interpreted as pedestrians feeling more comfortable in accepting smaller spatial gaps at PHB sites, because PHB sites offer more incentives for vehicles to decelerate quickly. Future studies can explore in more detail pedestrian gap acceptance at the RRFB and PHB sites.

This study investigated a variety of factors (including a unique factor landuse) affecting SSMs at signalized midblock crossings. The study has demonstrated that sites having a larger range of land-uses offer less pedestrian safety. Moreover, results from the study confirm that PHB facilities offer better temporal safety compared to what RRFB sites. In addition, the study has used vehicle trajectories based on the image processing of video recordings to increase the accuracy of the obtained SSMs and relevant roadway and traffic characteristics. The findings from this study could be useful to transportation planners in understanding the safety effectiveness of the RRFB and PHB treatments. The study reveals that signalized midblock crossings offer more safety during the weekdays, as compared to during the weekends. Since areas having greater variety of land use reduces traffic safety, transportation planners can consider installing PHB treatment facilities in those areas. This study also provides possible proof that vehicles decelerate more at PHB sites. This justifies the preferential deployment of PHB over RRFB in high-speed areas, which is also supported by our previous study (Ugan et al. 2022).

One avenue that the authors plan to explore in the future is incorporating Stopping Time in assessing conflict severity. Although several studies have tried this in the past using PRI, it has been argued in this

paper that the risk assessments used in Cafiso et al. (2011) and Chaudhari et al. (2020) were incomplete (Cafiso et al. 2011; Chaudhari et al. 2020). Nevertheless, these two studies still hold merit as they were able to show that comparison between TTCv and Stopping time can indicate conflict severity. This has inspired the authors to investigate the relation between RTTC and Stopping time in the future. Future research can also look at developing accurate thresholds for SSMs to classify conflict severities.

6. Data availability statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The data used in this study were collected using funding from the FDOT under the project titled “Evaluation of Midblock Pedestrian Signals (MPS)”. The authors are also grateful to the anonymous reviewers for their constructive suggestions, which have greatly improved the quality of the manuscript.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Nafis Anwari, Mohamed Abdel-Aty, Amrita Goswamy, Ou Zheng; data collection: Nafis Anwari, Amrita Goswamy, Ou Zheng; analysis and interpretation of results: Nafis Anwari, Mohamed Abdel-Aty, Amrita Goswamy, Ou Zheng; draft manuscript preparation: Nafis Anwari, Mohamed Abdel-Aty, Amrita Goswamy, Ou Zheng. All authors reviewed the results and approved the final version of the manuscript.

Appendix

Results from variance inflation factor (VIF) tests on temporal SSMs

VIF can be used to check for multi-collinearity. A variable having VIF greater than 10 shows multi-collinearity, and should be removed from the final analysis (Xie et al. 2011).

Results from variance inflation factor (VIF) tests on temporal SSMs

Variable	Temporal gap (s)		RTTC (s)		PET (s)	
	Gaussian Regression	Gamma Regression	Gaussian Regression	Gamma Regression	Gaussian Regression	Gamma Regression
PHB dummy	4.530	22.080	4.530	17.577	4.530	18.890
Weekend dummy	1.056	5.217	1.056	3.802	1.056	4.381
Signal activation	1.342	1.056	1.342	1.060	1.342	1.045
Total lane count	3.416	1.234	3.416	1.363	3.416	1.212
Pedestrian crossing speed	1.948	3.107	1.948	3.920	1.948	3.529
log (Average vehicle lane interaction speed)	1.516	2.031	1.516	1.973	1.516	2.030
Land use mix	4.233	1.646	4.233	1.506	4.233	1.385
Pedestrian starting position	1.600	4.915	1.600	4.417	1.600	5.128

References

- Abdel-Aty, M., Wu, Y., Zheng, O.u., Yuan, J., 2022. Using closed-circuit television cameras to analyze traffic safety at intersections based on vehicle key points detection. *Accid. Anal. Prev.* 176, 106794.
- Anderson, I., 2019. Assessment of the HAWK Crosswalk Traffic Signal.
- Arbabzadeh, N., Jafari, M., Jalayer, M., Jiang, S., Kharbeche, M., 2019. A hybrid approach for identifying factors affecting driver reaction time using naturalistic driving data. *Transp. Res. Part C: Emerg. Technol.* 100, 107–124.
- Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. *Accid. Anal. Prev.* 153, 106016.
- Avinash, C., Jiten, S., Arkatkar, S., Gaurang, J., Manoranjan, P., 2019. Evaluation of pedestrian safety margin at mid-block crosswalks in India. *Saf. Sci.* 119, 188–198.
- Beitel, D., Stipancic, J., Manaugh, K., Miranda-Moreno, L., 2018. Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis. *Transp. Res. Part D: Transp. Environ.* 65, 710–724.
- Bennett, M.K., Manal, H., Van Houten, R., 2014. A comparison of gateway in-street sign configuration to other driver prompts to increase yielding to pedestrians at crosswalks. *J. Appl. Behav. Anal.* 47 (1), 3–15. <https://doi.org/10.1002/jaba.103>.
- Bhaumik, D.K., Kapur, K., Gibbons, R.D., 2009. Testing Parameters of a Gamma Distribution for Small Samples. *Technometrics* 51 (3), 326–334.
- Brewer, M.A., Fitzpatrick, K., Avelar, R., 2015. Rectangular rapid flashing beacons and pedestrian hybrid beacons: Pedestrian and driver behavior before and after installation. *Transp. Res. Rec.* 2519, 1–9. <https://doi.org/10.3141/2519-01>.
- Cafiso, S., Garcia, A., Cavarra, R., Rojas, M.A.R., 2011. "Crosswalk Safety evaluation using a Pedestrian Risk Index as Traffic Conflict Measure." 3rd Int. Conf. Road Saf. Simul., 15.
- Chaudhari, A.R., Gore, N., Arkatkar, S., Joshi, G., Pulugurtha, S.S., 2020. Deriving pedestrian risk index by vehicle type and road geometry at midblock crosswalks under heterogeneous traffic conditions. *J. Transp. Eng. Part A Syst.* 146 (10) <https://doi.org/10.1061/jtpebs.0000421>.
- Chaudhari, A., Gore, N., Arkatkar, S., Joshi, G., Pulugurtha, S., 2021. Exploring pedestrian surrogate safety measures by road geometry at midblock crosswalks: a perspective under mixed traffic conditions. *IATSS Research* 45 (1), 87–101.
- Chen, P., Zeng, W., Yu, G., Wang, Y., 2017. Surrogate safety analysis of pedestrian-vehicle conflict at intersections using unmanned aerial vehicle videos. *J. Adv. Transp.* 2017, 1–12.
- Cheng, R.C.H., Amin, N.A.K., 1981. Maximum likelihood estimation of parameters in the inverse gaussian distribution, with unknown origin. *Technometrics* 23 (3), 257–263. <https://doi.org/10.1080/00401706.1981.10486295>.
- Fitzpatrick, K., Brewer, M.A., Avelar, R., 2014. Driver yielding at traffic control signals, pedestrian hybrid beacons, and rectangular rapid-flashing beacons in Texas. *Transp. Res. Rec.* 2463 (2463), 46–54. <https://doi.org/10.3141/2463-06>.
- Fitzpatrick, K., Cynecki, M.J., Pratt, M.P., Beckley, M., 2020. Analysis of pedestrian hybrid beacon operation on higher-speed roadways. *Transp. Res. Rec.* 2674 (5), 22–32. <https://doi.org/10.1177/036198120913558>.
- Fitzpatrick, K., Avelar, R., Pratt, M., Brewer, M., Robertson, J., Lindheimer, T., Miles, J., T. T. Institute, and F. H. Administration. 2016. Evaluation of Pedestrian Hybrid Beacons and Rapid Flashing Beacons.
- Fitzpatrick, K., Park, E.S., Cynecki, M.J., Pratt, M.P., Beckley, M., 2021. Investigation of crashes at pedestrian hybrid beacons: Results of a large-scale study in Arizona. *J. Saf. Res.* 78, 59–68.
- Goswamy, A., Abdel-Aty, M., 2023. Safety Effectiveness of Rectangular Rapid Flashing Beacons Pedestrian Enhancement. *J. Transp. Eng. Part A: Syst.* 149 (8), 04023070.
- Goswamy, A., Abdel-Aty, M., Islam, Z., 2023a. Factors affecting injury severity at pedestrian crossing locations with Rectangular RAPID Flashing Beacons (RRFB) using XGBoost and random parameters discrete outcome models. *Accid. Anal. Prev.* 181, 106937.
- Goswamy, A., Ph, D., Abdel-Aty, M., Ph, D., Asce, F., 2023b. Safety effectiveness of rectangular rapid flashing beacons pedestrian enhancement. *J. Transp. Eng. Part A Syst.* 149 (8), 1–10. <https://doi.org/10.1061/JTPEBS.TEENG-7507>.
- Gruden, C., Campisi, T., Canale, A., Tesoriere, G., Srimal, M., 2019. A cross-study on video data gathering and microsimulation techniques to estimate pedestrian safety level in a confined space. *IOP Conf. Ser. Mater. Sci. Eng.*, 42008.
- Hansson, A., 1975. *Studies in Driver Behaviour, With Applications in Traffic Design and Planning: Two Examples*. Lund, Sweden.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn. *Proc. IEEE Int. Conf. Comput. Vis.* 2961–2969.
- Ismail, K., Sayed, T., Saunier, N., Lim, C., 2009. Automated analysis of pedestrian-vehicle conflicts using video data. *Transp. Res. Rec.* 2140, 44–54. <https://doi.org/10.3141/2140-05>.
- Ismail, K., Sayed, T., Saunier, N., 2010. Automated analysis of pedestrian-vehicle: context for before-and-after studies. *Transp. Res. Rec.* 2198, 52–64. <https://doi.org/10.3141/2198-07>.
- Jiang, X., Wang, W., Bengler, K., 2015. "Intercultural Analyses of Time-to-Collision in Vehicle-Pedestrian Conflict on an Urban Midblock Crosswalk." *IEEE Trans. Intell. Transp. Syst.*, 16 (2): 1048–1053. IEEE. doi: 10.1109/TITS.2014.2345555.
- Johnson, N.L., Kotz, S., Balakrishnan, N., 1995. *Continuous Univariate Distributions, volume 2*. John wiley & sons.
- Jorgensen, B., 2012. *Statistical Properties of the Generalized Inverse Gaussian Distribution*. Springer Science & Business Media.
- Kadali, B.R., Perumal, V., 2016. Critical gap analysis based on pedestrian behavior during peak-hour traffic at unprotected midblock crosswalks. *Asian Transp. Stud.* 4 (1), 261–277. <https://doi.org/10.11175/eastsats.4.261>.
- Kadali, B.R., Vedagiri, P., 2019. Evaluation of pedestrian accepted vehicle gaps with varied roadway width under mixed traffic conditions. *Transportation Letters* 11 (9), 527–534.
- Kaparias, I., Bell, M., Dong, W., Sastrawinata, A., Singh, A., Wang, X., Mount, B., 2013. Analysis of pedestrian-vehicle traffic conflicts in street designs with elements of shared space. *Transp. Res. Rec.* 2393, 21–30. <https://doi.org/10.3141/2393-03>.
- Kathuria, A., Vedagiri, P., 2020. Evaluating pedestrian vehicle interaction dynamics at un-signaled intersections: a proactive approach for safety analysis. *Accid. Anal. Prev.* 134, 105316.
- Katz, A., Zaidel, D., Elgrishi, A., 1975. An experimental study of driver and pedestrian interaction during the crossing conflict. *Hum. Factors* 17 (5), 514–527. <https://doi.org/10.1177/00182087501700510>.
- Laureshyn, A., Å. Svensson, and C. Hydén. 2010. "Evaluation of traffic safety , based on micro-level behavioural data : Theoretical framework and first implementation." *Accid. Anal. Prev.*, 42 (6): 1637–1646. Elsevier Ltd. doi: 10.1016/j.aap.2010.03.021.
- Layegh, M., Mirbaha, B., Rassafi, A.A., 2020. Modeling the pedestrian behavior at conflicts with vehicles in multi-lane roundabouts (a cellular automata approach). *Phys. A* 556, 124843.
- Li, B., 2013. A model of pedestrians' intended waiting times for street crossings at signalized intersections. *Transp. Res. B Methodol.* 51, 17–28.
- Li, B., 2014. A bilevel model for multivariate risk analysis of pedestrians' crossing behavior at signalized intersections. *Transp. Res. B Methodol.* 65, 18–30.
- Li, P., Bian, Y., Rong, J., Zhao, L., Shu, S., 2013. Pedestrian crossing behavior at unsignalized mid-block crosswalks around the primary school. *Procedia Soc. Behav. Sci.* 96, 442–450.
- Macek, K., 2022. Pedestrian Traffic Fatalities by State: 2021 Preliminary Data.
- Mahmoud, N., Abdel-Aty, M., Cai, Q., Zheng, O.u., 2021. Vulnerable road users' crash hotspot identification on multi-lane arterial roads using estimated exposure and considering context classification. *Accid. Anal. Prev.* 159, 106294.
- Marisamynathan, S., Vedagiri, P., 2020. Pedestrian safety evaluation of signalized intersections using surrogate safety measures. *Transport* 35 (1), 48–56. <https://doi.org/10.3846/transport.2020.12157>.
- Méndez, J.G.S., 2020. Evaluating the effectiveness of Rectangular Rapid Flashing Beacons (RRFB) implementation using a before-after video-based surrogate safety approach. McGill University, Montreal.
- Monsere, C., S. Kothuri, and J. Anderson. 2020. Best Practices for Installation of Rectangular Rapid Flashing Beacons with and without Median Refuge Islands. Salem.
- Monsere, C., Figliozzi, M., Kothuri, S., Razmpa, A., Hazel, D., et al., 2017. Safety effectiveness of pedestrian crossing enhancements. Oregon. Dept. of Transportation. Research Section, Salem.
- National Highway Traffic Safety Administration. 2020. "2019 Ranking of STATE Pedestrian Fatality Rates - State : USA." Accessed November 27, 2021. <https://www-fars.nhtsa.dot.gov/States/StatesPedestrians.aspx>.
- Ni, Y., Wang, M., Sun, J., Li, K., 2016. Evaluation of pedestrian safety at intersections: a theoretical framework based on pedestrian-vehicle interaction patterns. *Accid. Anal. Prev.* 96, 118–129.
- Onelcin, P., Alver, Y., 2015. Illegal crossing behavior of pedestrians at signalized intersections: factors affecting the gap acceptance. *Transport. Res. F: Traffic Psychol. Behav.* 31, 124–132.
- Oxley, J., Fildes, B., Ihnsen, E., Charlton, J., Day, R., 1997. Differences in traffic judgements between young and old adult pedestrians. *Accid. Anal. Prev.* 29 (6), 839–847.
- Oxley, J.A., Ihnsen, E., Fildes, B.N., Charlton, J.L., Day, R.H., 2005. Crossing roads safely: an experimental study of age differences in gap selection by pedestrians. *Accid. Anal. Prev.* 37 (5), 962–971.
- Pawar, D.S., Patil, G.R., 2015. Pedestrian temporal and spatial gap acceptance at mid-block street crossing in developing world. *J. Saf. Res.* 52, 39–46.
- Pawar, D.S., Patil, G.R., 2016. Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials. *Saf. Sci.* 86, 295–303.
- Pawar, D.S., Patil, G.R., 2021. Analyzing variations in spatial critical gaps at two-way stop controlled intersections using parametric and non-parametric techniques. *J. Traffic Transp. Eng. (Engl. Ed.)* 8 (1), 129–138.
- Pawar, D.S., Yadav, A.K., 2022. Modelling the pedestrian dilemma zone at uncontrolled midblock sections. *J. Saf. Res.* 80, 87–96.
- Peesapati, L.N., Hunter, M.P., Rodgers, M.O., 2018. Can post encroachment time substitute intersection characteristics in crash prediction models? *J. Saf. Res.* 66, 205–211.
- Punzo, V., Borzacchiello, M.T., Ciuffo, B., 2011. On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data. *Transp. Res. Part C: Emerg. Technol.* 19 (6), 1243–1262.
- Raju, N., P. Kumar, C. Reddy, S. Arkatkar, G. Joshi. 2017. Examining Smoothing Techniques for Developing Vehicular Trajectory Data under Heterogeneous Conditions. *J. East. Asia Soc. Transp. Stud.*, 12 (2017): 1549–1568. doi: 10.11175/easts.12.1549.
- Ren, S., He, K., Girshick, R., Sun, J., 2015. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Adv. Neural Inf. Process. Syst.*, 28.

- Ren, G., Gu, C., Lu, L., Zhou, Z., Ding, C., 2012. Modeling risk degree of conflicts between crossing pedestrians and vehicles at signalized intersections. *J. Transp. Syst. Eng. Inf. Technol.* 12 (5), 76–81.
- Rista, E., Fitzpatrick, K., 2020. Comparison of LED-embedded pedestrian crossing signs with rectangular rapid flashing beacons and pedestrian hybrid beacons. *Transp. Res. Rec.* 2674 (11), 856–866. <https://doi.org/10.1177/0361198120941849>.
- Rosner, B., 2015. *Fundamentals of Biostatistics*. Cengage learning.
- Saleh, W.S., Lashin, M.M.A., 2020. Investigation of critical gap for pedestrian crossing using fuzzy logic system. *Appl. Sci.* 10 (10), 1–12. <https://doi.org/10.3390/app10103653>.
- Shaaban, K., Muley, D., Mohammed, A., 2018. Analysis of illegal pedestrian crossing behavior on a major divided arterial road. *Transport. Res. F: Traffic Psychol. Behav.* 54, 124–137.
- Tarko, A.P., 2018. "Surrogate measures of safety." *Safe Mobil. Challenges, Methodol. Solut.*, 383–405. Emerald Publishing Limited.
- Tezcan, H.O., Elmorssy, M., Aksoy, G., 2019. Pedestrian crossing behavior at midblock crosswalks. *J. Safety Res.* 71, 49–57. <https://doi.org/10.1016/j.jsr.2019.09.014>.
- Theofilatos, A., Ziakopoulos, A., Oviedo-Trespalacios, O., Timmis, A., 2021. To cross or not to cross? Review and meta-analysis of pedestrian gap acceptance decisions at midblock street crossings. *J. Transp. Health* 22, 101108.
- Timm, N.H., 2002. *Applied Multivariate Analysis*. Springer.
- Ugan, J., Abdel-Aty, M., Cai, Q., 2022. Estimating effectiveness of speed reduction measures for pedestrian crossing treatments using an empirically supported speed choice modeling framework. *Transport. Res. F: Traffic Psychol. Behav.* 89, 276–288.
- Várhelyi, A., 1998. Drivers' speed behaviour at a zebra crossing: A case study. *Accid. Anal. Prev.* 30 (6), 731–743. [https://doi.org/10.1016/S0001-4575\(98\)00026-8](https://doi.org/10.1016/S0001-4575(98)00026-8).
- Vilca, F., Balakrishnan, N., Zeller, C.B., 2014. Multivariate Skew-Normal Generalized Hyperbolic distribution and its properties. *J. Multivar. Anal.* 128, 73–85.
- Wang, H., Wang, W., Chen, J., Jing, M., 2010. Using trajectory data to analyze intradriver heterogeneity in car-following. *Transp. Res. Rec.* 2188, 85–95. <https://doi.org/10.3141/2188-10>.
- Xie, F., Gladhill, K., Dixon, K., Monsere, C., 2011. Calibration of highway safety manual predictive models for oregon state highways. *Transp. Res. Rec.* 2241, 19–28. <https://doi.org/10.3141/2241-03>.
- Yannis, G., Papadimitriou, E., Theofilatos, A., 2013. Pedestrian gap acceptance for mid-block street crossing. *Transp. Plan. Technol.* 36 (5), 450–462. <https://doi.org/10.1080/03081060.2013.818274>.
- Yoo, W., M. Robert, B. Sejong, S. Karan, (Peter) He Qinghua, and W. L. J. James. 2014. A study of effects of multicollinearity in the multivariable analysis. *Int. J. Appl. Sci. Technol.*, 6 (8): 9–19.
- Zegeer, C., Lyon, C., Srinivasan, R., Persaud, B., Lan, B., Smith, S., Carter, D., Thirsk, N.J., Zegeer, J., Ferguson, E., Van Houten, R., Sundstrom, C., 2017. Development of crash modification factors for uncontrolled pedestrian crossing treatments. *Transp. Res. Rec.* 2636, 1–8. <https://doi.org/10.3141/2636-01>.
- Zhang, Y., D. Yao, T. Z. Qiu, and L. Peng. 2011. "Vehicle-pedestrian interaction analysis in mixed traffic condition." *ICTIS 2011 Multimodal Approach to Sustain. Transp. Syst. Dev. - Information, Technol. Implement. - Proc. 1st Int. Conf. Transp. Inf. Saf.*, 552–559. Wuhan, China.
- Zhang, Y., Yao, D., Qiu, T.Z., 2012. Pedestrian safety analysis in mixed traffic conditions using video data. *IEEE Trans. Intell. Transp. Syst.* 13 (4), 1832–1844.
- Zhang, Y., Yao, D., Qiu, T.Z., Peng, L., 2014. Scene-based pedestrian safety performance model in mixed traffic situation. *IET Intell. Transp. Syst.* 8 (3), 209–218. <https://doi.org/10.1049/iet-its.2013.0012>.
- Zhang, C., Zhou, B., Chen, G., Chen, F., 2017. Quantitative analysis of pedestrian safety at uncontrolled multi-lane mid-block crosswalks in China. *Accid. Anal. Prev.* 108, 19–26.
- Zheng, O., 2019. Developing a Traffic Safety Diagnostics System for Unmanned Aerial Vehicles Using Deep Learning Algorithms. University of Central Florida.