



# Center for Advanced Multimodal Mobility Solutions and Education

**Project ID: 2020 Project 12**

## **ESTIMATION OF PEDESTRIAN VOLUME USING GEOSPATIAL AND TRAFFIC CONFLICT DATA**

### **Final Report**

by

John Ivan (ORCID ID: <https://orcid.org/0000-0002-8517-4354>)  
Professor, Department of Civil and Environmental Engineering  
University of Connecticut  
261 Glenbrook Road Unit, 3037, Storrs, CT 06269-3037  
Phone: ; Email: [john.ivan@uconn.edu](mailto:john.ivan@uconn.edu)

Amy Burnicki (ORCID ID: <https://orcid.org/0000-0002-5232-4555>)  
Associate Professor in Residence, Department of Geography  
University of Connecticut  
Phone: ; Email: [amy.burnicki@uconn.edu](mailto:amy.burnicki@uconn.edu)

Pankaj Joshi (ORCID ID: <https://orcid.org/0000-0002-5068-7969>)  
Associate Regional Planner, Fresno Council of Governments  
2035 Tulare Street Suite 201, Fresno, CA 93721  
Phone: 1-959-929-2925; Email: [pankaj.joshi@uconn.edu](mailto:pankaj.joshi@uconn.edu)

for

Center for Advanced Multimodal Mobility Solutions and Education  
(Cammse @ UNC Charlotte)  
The University of North Carolina at Charlotte  
9201 University City Blvd  
Charlotte, NC 28223

**September 2021**

## **ACKNOWLEDGMENTS**

This project was funded by the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE @ UNC Charlotte), one of the Tier I University Transportation Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT), under the FAST Act. The authors are grateful to Vincent Huang for his help in the GIS data preparation.

## **DISCLAIMER**

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation University Transportation Centers Program in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the U.S. Government. This report does not constitute a standard, specification, or regulation.

# TABLE OF CONTENTS

ABSTRACT.....	4
1. INTRODUCTION .....	6
2. LITERATURE REVIEW .....	8
3. DATA PREPARATION.....	10
3.1.Traffic Conflict Data.....	10
3.2.Spatial Data.....	11
4. METHODOLOGY .....	17
4.1.Conflict Ratio.....	17
4.2.Model Formulation .....	17
4.3.Spatial Autocorrelation .....	19
5. RESULTS .....	24
5.1.Models Using Variables from the One-Eighth Mile Buffer .....	24
5.2.Models Using Variables from the Quarter Mile Buffer.....	25
5.3.Models Using Variables from the Half Mile Buffer.....	26
6. DISCUSSION.....	29
7. CONCLUSION.....	31
REFERENCES .....	33

## LIST OF TABLES

Table 1 Summary statistics from the Traffic Conflict Observations.....	11
Table 2 Definitions of the Land Use Categories.....	14
Table 3 Summary statistics for variables in Half-mile Polygon.....	15
Table 4 Models using variables from the One-Eighth mile buffer.....	25
Table 5 Models using variables from the Quarter Mile Buffer.....	26
Table 6 Models using variables from the Half-Mile Buffer.....	27

## LIST OF FIGURES

Figure 1 Total obligations to Pedestrians and Bicycle Facilities & Programs by year.....	6
Figure 2 Study Sites – Towns.....	10
Figure 3 Example Crosswalk Locations in Hartford.....	12
Figure 4 Buffer Polygons at four example Locations in Hartford.....	13
Figure 5 Situation where parcel whose access road is not within buffer polygon.....	16
Figure 6 Density plot of the Logarithm of Pedestrian Volume per Hour.....	18
Figure 7 QQ plot of the Logarithm of Pedestrian Volume per Hour.....	18
Figure 8 Spatially plotted Residuals from equation (2) for Hartford Locations.....	20
Figure 9 Spatially plotted Residuals from equation (2) for West Hartford Locations.....	21
Figure 10 Spatially plotted Residuals from equation (3) for Hartford Locations.....	22
Figure 11 Spatially plotted Residuals from equation (3) for West Hartford Locations.....	23

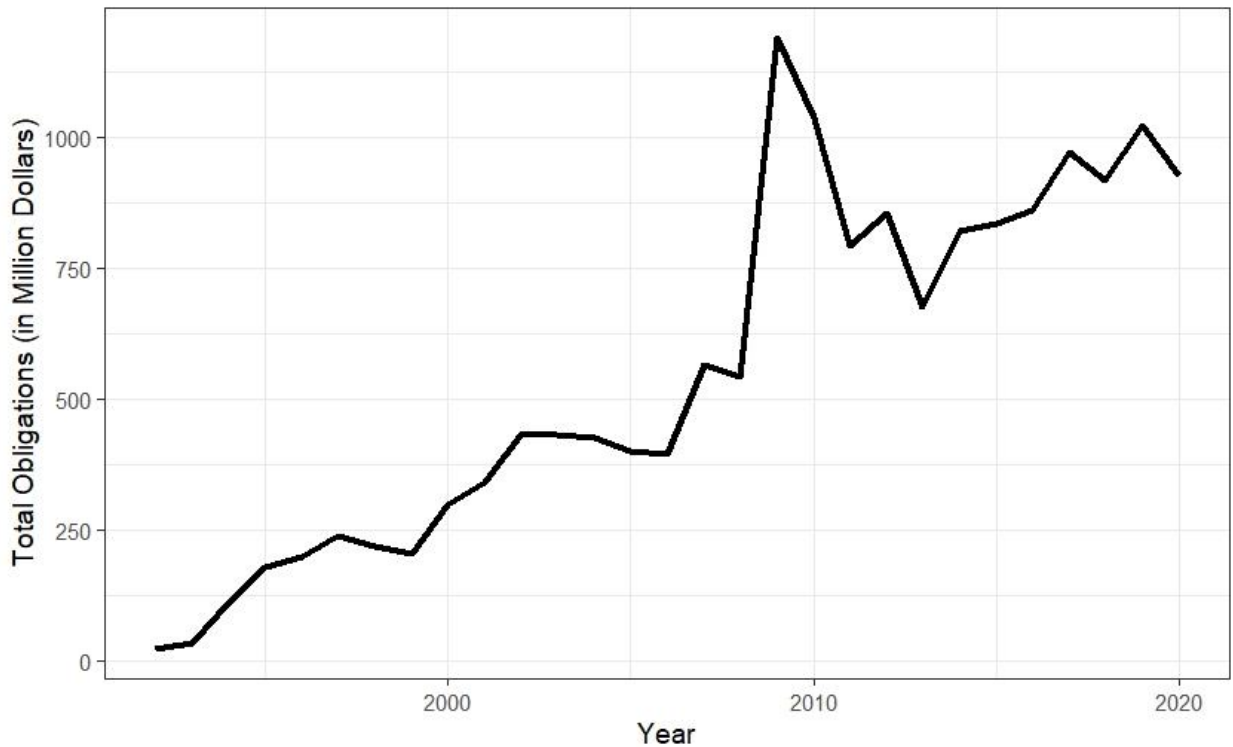
## **ABSTRACT**

Walking is widely promoted in the U.S. with known benefits for personal health, the environment, and economic activity. This project investigated the association between pedestrian volume and pedestrian-vehicle interactions on crosswalks. Pedestrian-vehicle interactions were classified based on their severity following a modified version of the Swedish Traffic Conflict Technique. The severity of interactions on crosswalks was categorized using conflict ratio, the ratio of highly severe interactions to less severe interactions. A random-effects log-linear regression model was developed using conflict ratio, the ratio of highly severe interactions to less severe interactions, as a covariate along with other geospatial variables such as land use and network connectivity. The results indicate that the conflict ratio significantly explains variation in pedestrian volume at the studied crosswalk locations. Locations with higher conflict ratios have lower pedestrian volumes than locations with lower conflict ratios. Facilities that reduce the severity of pedestrian interaction more comfortable for pedestrians and can help planners to achieve their environmental, societal, and economical goals by promoting walking.

# 1. INTRODUCTION

Walking is widely recognized as a sustainable mode of transportation. Increasing the share of walking as a travel mode is essential for urban areas to reduce auto dependency and promote green and environment-friendly development (Giles-Corti et al., 2016). Studies have also found a significant association between health and walking (Murphy et al., 2007; Hanson and Jones, 2015). Further, there is evidence that consumers who arrive at businesses by modes other than automobile, such as walking, make more frequent visits and are “competitive consumers” impacting businesses in cities (Clifton et al., 2013).

With known benefits of walking on personal health, the environment, and economic activity, walking is widely promoted in the U.S. Urban agencies have developed strategies, educational campaigns, and prepared strategic plans focused on increasing the mode share of walking. The California Department of Transportation published a strategic plan describing their Active Transportation Program to encourage increased trips via walking in an effort to double the walking mode share (Brown, Jr. et al., 2015). Likewise, health professionals have promoted walking to increase physical activity in the U.S. (U.S. Department of Health and Human Services, 2015). The Partnership for Active Transportation has policies to increase federal investments and to prioritize creating networks for walking (Rails-to-Trails Conservancy, 2018). United States Department of Transportation (USDOT) also announced their approach to incorporate safe and convenient walking facilities into transportation projects (LaHood, R., 2010). Moreover, the Federal Highway Administration shows increasing obligations to projects focusing on pedestrian and bicycle facilities (Figure 1).



**Figure 1 Total obligations to Pedestrians and Bicycle Facilities & Programs by year**

(data from Federal-Aid Highway Program Funding for Pedestrian and Bicycle Facilities and Programs, [https://www.fhwa.dot.gov/environment/bicycle\\_pedestrian/funding/bipedfund.cfm](https://www.fhwa.dot.gov/environment/bicycle_pedestrian/funding/bipedfund.cfm))

When factors that affect walking are known, focused efforts and investments on promoting walking and increasing share of walking as a transportation mode can succeed. Researchers have studied the association between the characteristics of the built-environment and pedestrian activity (Pushkarev and Zupan, 1971; Behnam and Patel, 1977; Schneider et al., 2009; Pulugurtha and Repaka, 2008; Miranda-Moreno et al., 2011). Studies have also examined the relationship between street network connectivity and pedestrian activity (Miranda-Moreno et al., 2011; Lu et al., 2012; Do et al., 2013).

Studies on the association between behavioral and psychological factors and walking as a mode of transportation have found that fear of crashes impacts mode adaptation significantly for walking when compared with users of other modes of transportation such as bus, taxi, metro, etc. (Backer-Grøndahl et al., 2009; Fyhri and Backer-Grøndahl, 2012). Pedestrians and parents of young children avoid high-risk areas in their route while walking (Quistberg et al., 2017).

The Theory of Planned Behavior says that the intention to perform different kinds of behavior is formed and predicted from the attitude towards the behavior, subjective norm, and perceived behavioral control (Ajzen 1991). The perceived behavioral control is based on a set of control beliefs that may be based—in part—on past experiences (Ajzen 1991). Hence, it is important to make walking experiences comfortable and safe to promote walking. Previous studies have found a correlation between pedestrian-vehicle conflicts and the pedestrian's perception of safety (Muraleetharan et al., 2005; Ying Ni et al., 2017). Based on these ideas, a need to explore the effects of interactions between pedestrians and vehicles on pedestrian volumes is deemed necessary. Findings on the influences of pedestrian-vehicle interactions would help in planning and designing aspects of crosswalk locations for promoting walking.

Hence, the primary hypothesis of this project is that the number of conflicts between pedestrians and vehicles contributes to variation in pedestrian volume on crosswalks. This is based on the assumption that conflicts with motor vehicles influence pedestrians' perception of safety. It is expected that if the hypothesis is correct, and conflict frequency is found to be significant in explaining variation in pedestrian volume, any location where interactions between pedestrians and vehicles are more severe, lower pedestrian volumes will be observed. To test the hypothesis, regression analysis is carried out using traffic conflict data while controlling for geospatial and network variables already known to be related to pedestrian activity. The model interpretation will focus on testing this hypothesis, meaning that the models are not developed with the intention of being used to predict pedestrian volume. As a secondary goal, this project will also confirm how other geospatial land use variables and network characteristics influence pedestrian activity.

The report is organized into seven sections. Section 1 contains the motivation and objective of the study. Section 2 presents the literature review of previous studies that have explored the association of pedestrian-vehicle interactions with pedestrian perception of safety. Section 2 also provides a review of research into estimating pedestrian volume using geospatial and network characteristics variables. Section 3 includes a detailed description of the data collection process and definition of variables used in the study. Details about the methodology used in the study are explained in Section 4. Model estimation results are presented in Section 5, followed by a discussion on the findings in Section 6. Lastly, Section 7 concludes the results of the research and offers ideas for future research.

## 2. LITERATURE REVIEW

Researchers in the past have explored the effect of interaction of pedestrians with vehicles on pedestrian's perception of safety. Muraleetharan et al. (2005) identified factors affecting the pedestrian Level of Service. They showed that pedestrian level of service decreases with higher pedestrian-vehicle interactions, and pedestrians feel discomfort due to conflicts with vehicles. Likewise, Bian et al. (2009) studied pedestrian Level of Service through a questionnaire. They concluded that conflicts with turning vehicles at intersections have a negative effect on pedestrians' sense of comfort and safety, ultimately hurting the pedestrian level of service. Ying Ni et al. (2017) used ordered logit modeling to find an association between pedestrian perception of safety at a signalized intersection and conflicts between pedestrians and vehicles. The perception of safety was collected through a survey. It was found that conflicts with vehicles have significant impacts on the pedestrian perception of safety, with results showing that locations with less conflicts perceived as safer. In a study of decision strategy at marked crossings, Sucha et al. (2017) observed a correlation between the distribution of conflicts between the site under observation and the pedestrian reported subjective safety of the site. In other words, pedestrians reported feeling safe where the fewest conflicts with vehicles were observed.

Another stream of research has studied the behavioral and psychological factors that deter people from walking. A questionnaire-based study examined the effects of fear and worry on travel behavior (Amundsen et al., 2009). The participants showed worry about crashes due to prior crash experiences. The study concluded that behavior adaptation in pedestrians is a consequence of concern about crashes and unpleasant incidents. In the study, 53.8% of people sometimes or often underwent a travel behavior adaptation by choosing another mode of transportation due to worry about crashes and unpleasant experiences. Quistberg et al. (2017) found that participants and parents of young children avoided areas that were categorized as riskier based on their knowledge of past collision history.

Past research has modeled pedestrian volumes using land use and network variables. Pushkarev and Zupan (1971) generated equations for estimating pedestrian volumes using linear regression and found a positive association between retail and office land use and restaurants in the vicinity. Behnam and Patel (1977) generated log-linear regression models for estimating pedestrian volume for a central business district (CBD). Positive coefficients for commercial space, office space, and residential space in the CBD were found. A synthesis of travel and the built environment (Ewing and Cervero, 2001) found a negative association between street block length and pedestrian volume. Pulugurtha and Repaka (2008) generated linear regression models to measure pedestrian activity at signalized intersections. Pedestrian activities at different time intervals were assessed using geospatial data from multiple buffer distances around the observation intersection. The models developed in the study found a positive association between vehicular volume and pedestrian activity. Single-family residential areas and commercial centers in the vicinity of the intersection had a negative impact on pedestrian activity. Distance to the nearest transit stop was also found to be significant in modeling pedestrian activity at intersections. The study also pointed out that different variables were influential when modeled for one mile, half mile and quarter mile buffer sizes. Another pilot model (Schneider et al. 2009) for pedestrian intersection crossing volumes showed that more commercial land use resulted in more pedestrian activity. A case study of San Francisco (Liu and Griswold, 2009) conducted regression analysis at multiple geographical scales using socioeconomic and built-environment variables. Land use mix, transit access, and job density



were found to have strong explanatory power on the variation in pedestrian volumes. Miranda-Morena et al. (2011) studied the link between the built environment and pedestrian activity using linear regression. The results showed a significant association between built environment variables such as commercial buildings in the vicinity of the crossing location and pedestrian activity. Network variables such as average street length, number of bus stops, and percentage of arterial streets were also significantly associated with pedestrian activity.

The relationship between the built environment and pedestrian activity has been studied through spatial analysis as well. Lu et al. (2012) carried out geographically weighted regression to analyze non-motorized traffic volumes and ambient built environment elements. The number of all buildings and number of intersections within a buffer area around the intersection of interest were significantly correlated with the pedestrian count in the study. Yang et al. (2017) used geographically weighted regression to model walking travel demands at intersections. They found that the number of buildings, distance to downtown, number of transit stops, among other variables, to be significantly associated with pedestrian travel.

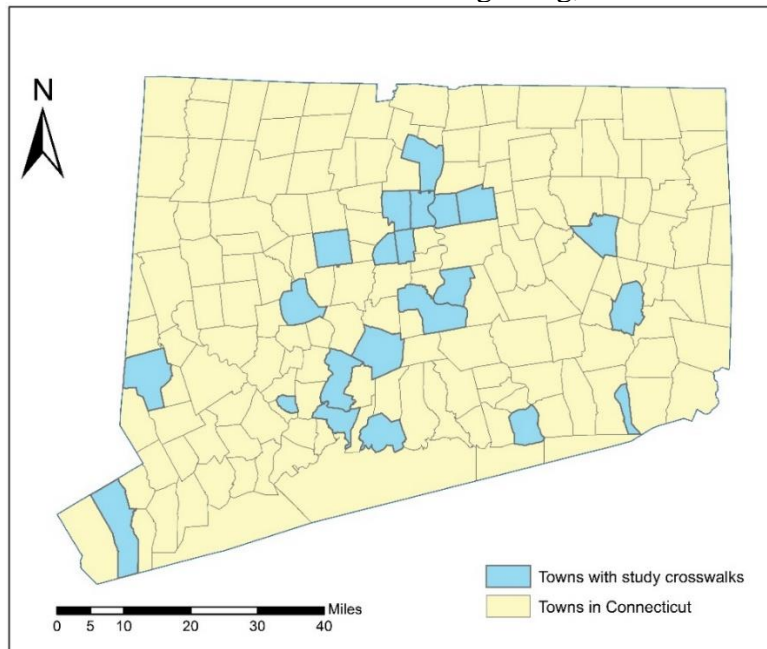
In summary, previous studies have investigated the impacts of fear, worry, unpleasant experiences, and conflicts on the perception of safety and have found that these can be discouraging factors for walking. However, whether that impact brings variation in pedestrian volume is still a question. This project examined the association between pedestrian-vehicle interactions and variation in pedestrian volume on crosswalks while controlling for other factors such as land use and network characteristics that have been known to affect pedestrian volumes. The results will have the potential to be helpful in the safety analysis of crossing locations and provide support to planners and engineers for making informed decisions.

### 3. DATA PREPARATION

The study aims to test the hypothesis that pedestrian perception of safety, as measured by observed conflicts between pedestrians and vehicles, is significant for explaining pedestrian volume on a crosswalk while also controlling for surrounding land use and network characteristics. Therefore, there are two types of data needed in this study: traffic conflict data and spatial data.

#### 3.1 Traffic Conflict Data

Three previous studies (Ivan et al. 2012, Ivan et al. 2015a, Ivan et al. 2015b) collected information related to pedestrian-vehicle conflicts for a total of 316 crosswalks located at 206 different intersections and midblock locations in 25 different towns in Connecticut, USA (Figure 2). The selection of the crosswalks was based on the needs of the previous studies that required mid-block locations, painted crosswalks, crosswalks with signs, textured crosswalks, crosswalks at an intersection with exclusive or concurrent traffic signaling, and other conditions.



**Figure 2 Study Sites – Towns**

The conflict types were categorized based on a variation of the Swedish Traffic Conflict Technique (Hyden and Linderholm, 1984). All the observers in these projects were appropriately trained, and test runs were conducted to ensure consistency in observations among observers in categorizing the interactions. The observation period varied from one to six hours. The shorter observation durations were due to rain, unusual and unexpected local events, or negligible pedestrian volume. The summary statistics of the variables collected from the traffic conflict observation are presented in Table 1. The definitions used in the classification of vehicle-pedestrian interactions in these previous studies are as follows:

**Table 1 Summary statistics from the Traffic Conflict Observations**

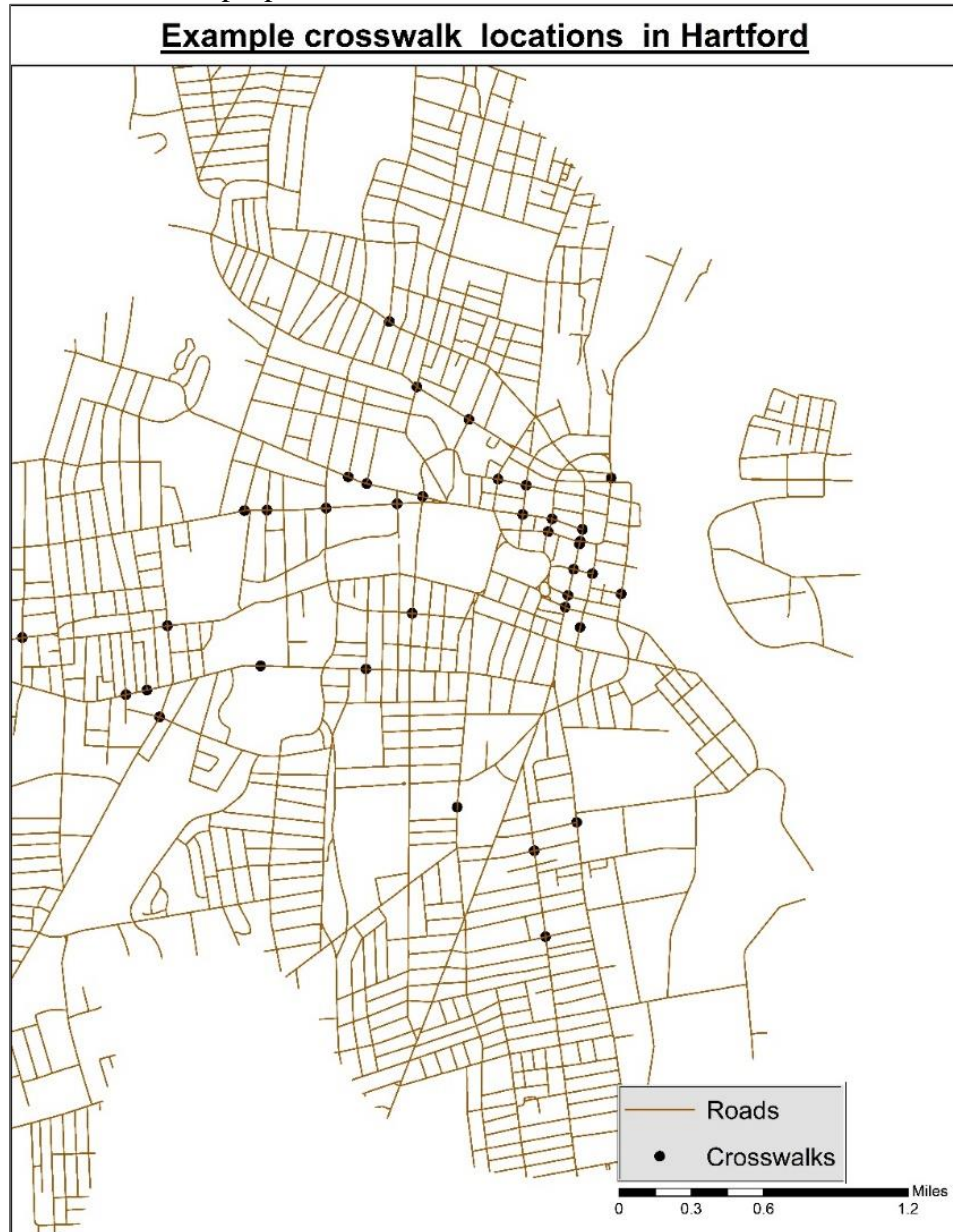
Variable	Minimum	Mean	Maximum
Undisturbed Conflicts (U)	0	253.53	3504
Potential Conflicts (P)	0	21.6	345
Minor Conflicts (M)	0	4.67	48
Severe Conflicts (S)	0	0.73	20
Total Pedestrian Count	1	280.5	3576
Vehicular Volume	103	3074	14864
Observed Hours (hr)	1	4.476	6
Pedestrian Volume per Hour	0.17	70.05	715.2
Vehicular Volume per Hour	32.49	711.41	2812.98

- **Undisturbed crossing:** When a pedestrian crosses the road when there are no vehicles in the vicinity, or all vehicles are already completely stopped before the pedestrian begins to cross, the interaction is classified as an “undisturbed crossing”. These interactions do not present any kind of threat to pedestrians.
- **Potential conflict:** This level of interaction means there is no likelihood of a collision, but there are simple interactions such as eye contact between the pedestrian and vehicle driver as the vehicle decelerates to a stop. These interactions also do not present any threat to pedestrians.
- **Minor conflict:** This level of interaction has a low possibility of a collision. However, the speed of the vehicle is low, and it makes a quick stop a few feet away from the pedestrian at a crosswalk to avoid hitting the pedestrian in the crosswalk. Since the speed is low in these interactions, it is unlikely to result in fatality or serious injury if it were to become a collision.
- **Serious conflict:** When a very late evasive action to avoid a collision is taken by a vehicle driver (such as unusual deceleration or abrupt stopping) or pedestrian (such as jumping back onto the sidewalk or running out of the vehicle's path) is observed, the interaction is classified as a “serious conflict”. These interactions are likely to result in a fatality or serious injury in the absence of evasive actions.

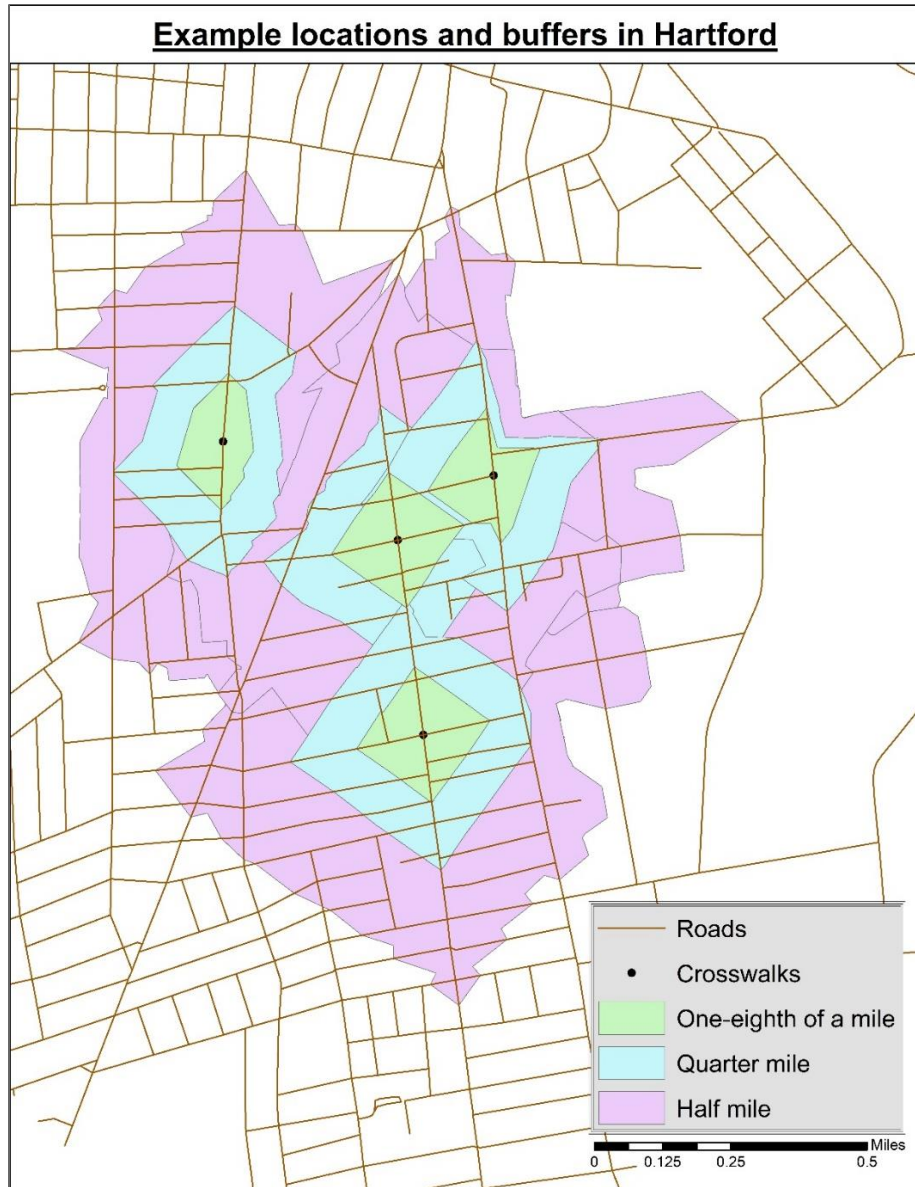
### 3.2 Spatial Data

In addition to traffic conflict data, geospatial variables are used as covariates in the models. Two types of spatial data are used: road network data and parcel land use data. These variables are expected to influence pedestrian volume at a location with different magnitude and significance based on their vicinity to the location (Liu et al. 2009). For this reason, spatial data were generated and compiled at three threshold distances from the observation location: half-mile, quarter-mile, and one-eighth of a mile. These distances are selected based on previous research (Pulugurtha and Repaka 2008; Miranda-Morena et al. 2011) and the assumption that

people usually walk for trips that are within these distances. Hence, to extract geospatial data that are within the threshold walking distances from the study locations, network service areas were generated for three threshold walking distances using Network Analyst Extension in ArcGIS Desktop (ESRI). A network service area is a region or a polygon that contains all accessible street locations that are within a specified impedance—which is the distance one-eighth mile, quarter-mile, and half-mile in this case. To generate these service area buffers, a road network from OpenStreetMap (OpenStreetMap contributors) was used. The network file was converted into a network dataset to generate service area buffers. As a result of this process, three polygons around each observation location were generated. Figure 3 shows a sample of crosswalk locations in Hartford; Figure 4 shows four selected crosswalk locations in Hartford and their three buffers for demonstration purposes.



**Figure 3 Example Crosswalk Locations in Hartford**



**Figure 4 Buffer Polygons at four example Locations in Hartford**

Previous studies have found a positive association between network connectivity and pedestrian activity (Miranda-Morena et al. 2011; Lu et al. 2012; Do et al. 2013). Higher pedestrian volume is expected at locations with better network connectivity in the vicinity. Hence, two network connectivity characteristics were determined—link-to-node ratio and average block length. The link-to-node ratio was calculated by dividing the number of links in the polygon by the number of nodes and cul-de-sacs. The average block length for each polygon was calculated by dividing the sum of the length of links in the polygon by the number of links.

Digitized parcel data containing land-use attributes were obtained for all towns in the study area. Some towns had readily available digital parcel-level data. However, the land use classification system of the parcel varied by the town. Therefore, a revised classification of land use type was constructed for simplification and consistency across all towns for this study as

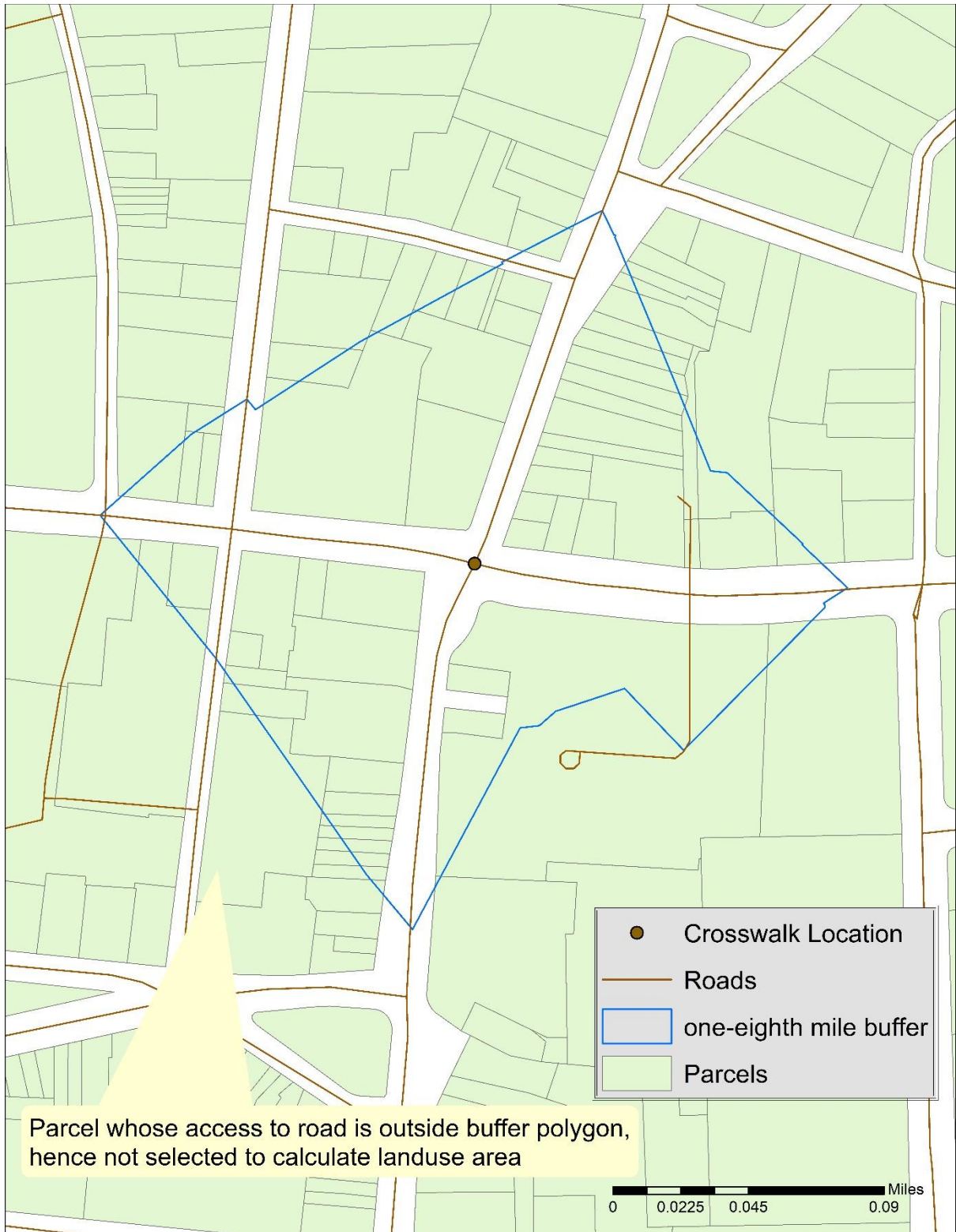
defined in Table 2. For towns that did not have parcel-level data available in digital form, parcels were digitized and land use types were assigned manually in ArcGIS Desktop (ESRI). The land use attributes were taken from the town’s parcel data viewer on the web and categorized according to the classification for this study. Then, buffer polygons of each size for each location were laid over the parcel data. At first, it was anticipated that parcels intersected by the buffer polygons at each location would be summarized to generate the area of each land-use class. However, we found that there were some situations where a parcel whose access to the road was not within walking distance from the respective study crosswalk was being selected, as shown in Figure 5. Therefore, parcels were checked manually whether they are accessible within the threshold walking distance and selected carefully for each buffer polygon of every location. Then, the area of the selected parcels was summarized to get the total square footage of different land use classes within one-eighth of a mile, quarter-mile, and half-mile walking distance from the study crosswalks. The number of parcels corresponding to each land-use class was also counted. Hence, counts and areas of different land use classes were generated for each crosswalk location at three different scales. The summary of the extracted variables for the half-mile buffer is presented in Table 3.

**Table 2 Definitions of the Land Use Categories**

<b>Land-use Category</b>	<b>Definition</b>
Single-Family Residential	Single family detached housing
Multi-Family Residential	Multi-family housing such as condominiums, apartments or townhouses
Major Commercial	Commercial locations that require a longer presence of customers, such as automobile dealerships, major department stores, and hotels
Industrial	Warehouses and manufacturing sites
Office and minor commercial	Offices that provide services, such as government buildings, banks, and commercial places that require a relatively shorter stay of customers, such as fast food franchises, restaurants, convenience stores, etc.
Institutional	Churches, religious institutions, hospitals, etc.
Parking	Parcels and structures used for parking
Parks and Recreation	Public parks and recreational places
Higher Educational	Colleges and universities

**Table 3 Summary statistics for variables in Half-mile Polygon**

	<b>Variable</b>	<b>Minimum</b>	<b>Mean</b>	<b>Maximum</b>
<b>Parcel Area (10,000 sq ft)</b>	Single-Family Residential	0.000	16.529	77.252
	Multi-Family Residential	0.000	24.600	129.930
	Major Commercial	0.000	3.043	32.133
	Industrial	0.000	1.292	30.174
	Office and minor commercial	0.000	19.658	64.041
	Institutional	0.000	8.506	33.236
	Parking	0.000	4.787	22.209
	Parks and Recreation	0.000	3.966	18.921
	Higher educational	0.000	9.246	75.960
<b>Parcel Count</b>	Single-Family Residential	0.000	166.200	775.000
	Multi-Family Residential	0.000	211.860	741.000
	Major Commercial	0.000	13.070	144.000
	Industrial	0.000	1.766	30.000
	Office and minor commercial	0.000	76.590	263.000
	Institutional	0.000	15.300	71.000
	Parking	0.000	22.600	90.000
	Parks and Recreation	0.000	3.399	14.000
	Higher educational	0.000	6.079	145.000
<b>Network Variable</b>	Average block length	94.930	351.950	1529.050



**Figure 5 Situation where parcel whose access to a road is not within the buffer polygon**



## 4. METHODOLOGY

### 3.3 Conflict Ratio

Interactions between vehicles and pedestrians at crosswalks have the potential to be used in assessing the pedestrian perception of safety. The association between vehicle-pedestrian conflicts with the pedestrian-vehicle crashes has been studied (Islam et al., 2014; Zhang et al., 2015; Zhang et al., 2019) and their potential use as a proactive approach to safety analysis kindles curiosity to explore their significance and prospective use in transportation planning. Past research (Bian et al., 2009; Amundsen et al., 2009; Suchaa et al., 2017) have studied the perception of safety through subjective questionnaires. However, the interactions between vehicles and pedestrians possess the potential to represent—how pedestrians perceive safety—objectively. Hence, conflict-ratio is used in the model, defined as the ratio of highly severe interactions to less severe interactions as shown in Equation (1):

$$\text{Conflict ratio} = \frac{\text{Number of highly severe interactions}}{\text{Number of less severe interactions}} \quad (1)$$

Among the conflict observations in the data, fortunately, there were very few serious conflicts. Accordingly, using serious conflicts alone in the numerator of the conflict ratio would not have been very useful. Therefore, minor conflicts and serious conflicts were combined to constitute “highly severe interactions”, whereas undisturbed crossings and potential conflicts were totaled to form “less severe interactions”. Evasive actions or certain communicative gestures can be seen in highly severe interactions that conceivably prevented an impending collision. Hence, it can be assumed that these interactions have some impact on the pedestrian perception of safety at the crosswalk. On this assumption, the conflict ratio has the potential for quantitatively representing the perception of safety at a crosswalk.

### 3.4 Model formulation

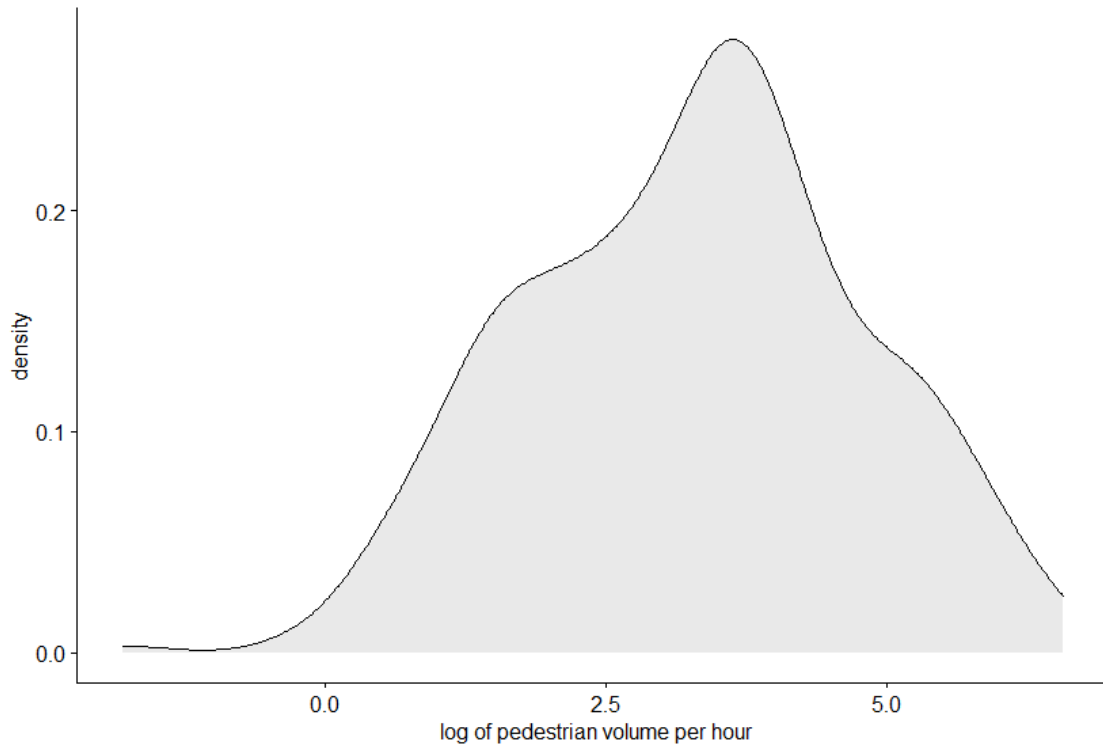
A multiple regression by Ordinary Least Squares (OLS) was carried out to test the hypothesis that traffic conflict frequency can significantly explain variation in pedestrian volume at crosswalks. Before deciding on the form of the regression equation, the normality of the independent variable, the pedestrian volume per hour, was examined. A visual inspection of the data histogram as well as the Shapiro-Wilks normality test (Shapiro and Wilk, 1965) was carried out. The visual inspection showed that the pedestrian volume per hour did not follow a normal distribution; however, the logarithm of this variable cleared the normality check. Visual inspection included a density plot shown in Figure 6 with an approximate bell curve and a QQ plot shown in Figure 7 with the observations plotted in a straight line. Both plots indicated the visual confirmation of normality of logarithm of pedestrian volume. The Shapiro-Wilks normality test parameter was calculated to be 0.991 with a p-value of 0.06, which significantly failed to reject that the sample represents a normal distribution.

Hence, a log-linear model was decided to be appropriate for the analysis. The basic model form used is given in Equation (2):

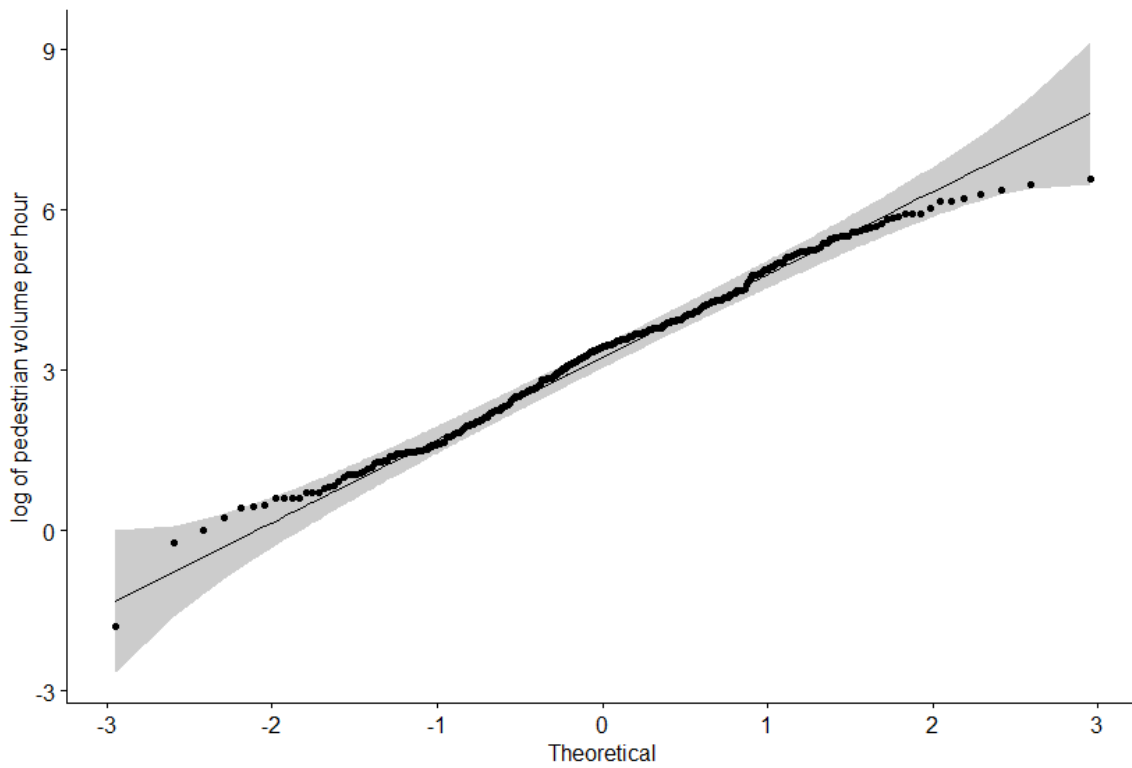
$$\ln P_i = \alpha + \beta_1 x_i + \beta_2 w_i + \varepsilon_i \quad (2)$$

Where  $P_i$  is the observed pedestrian crossing volume per hour at location  $i$ ,  $x_i$  is the conflict ratio,  $w_i$  is a vector of geospatial variables at location  $i$ , including those described earlier,

along with vehicular volume,  $\alpha$ ,  $\beta_1$  and  $\beta_2$  are parameters to be estimated, and  $\varepsilon_i$  is a random disturbance for the location  $i$ .



**Figure 6 Density plot of the Logarithm of Pedestrian Volume per Hour**



**Figure 7 QQ plot of the Logarithm of Pedestrian Volume per Hour**

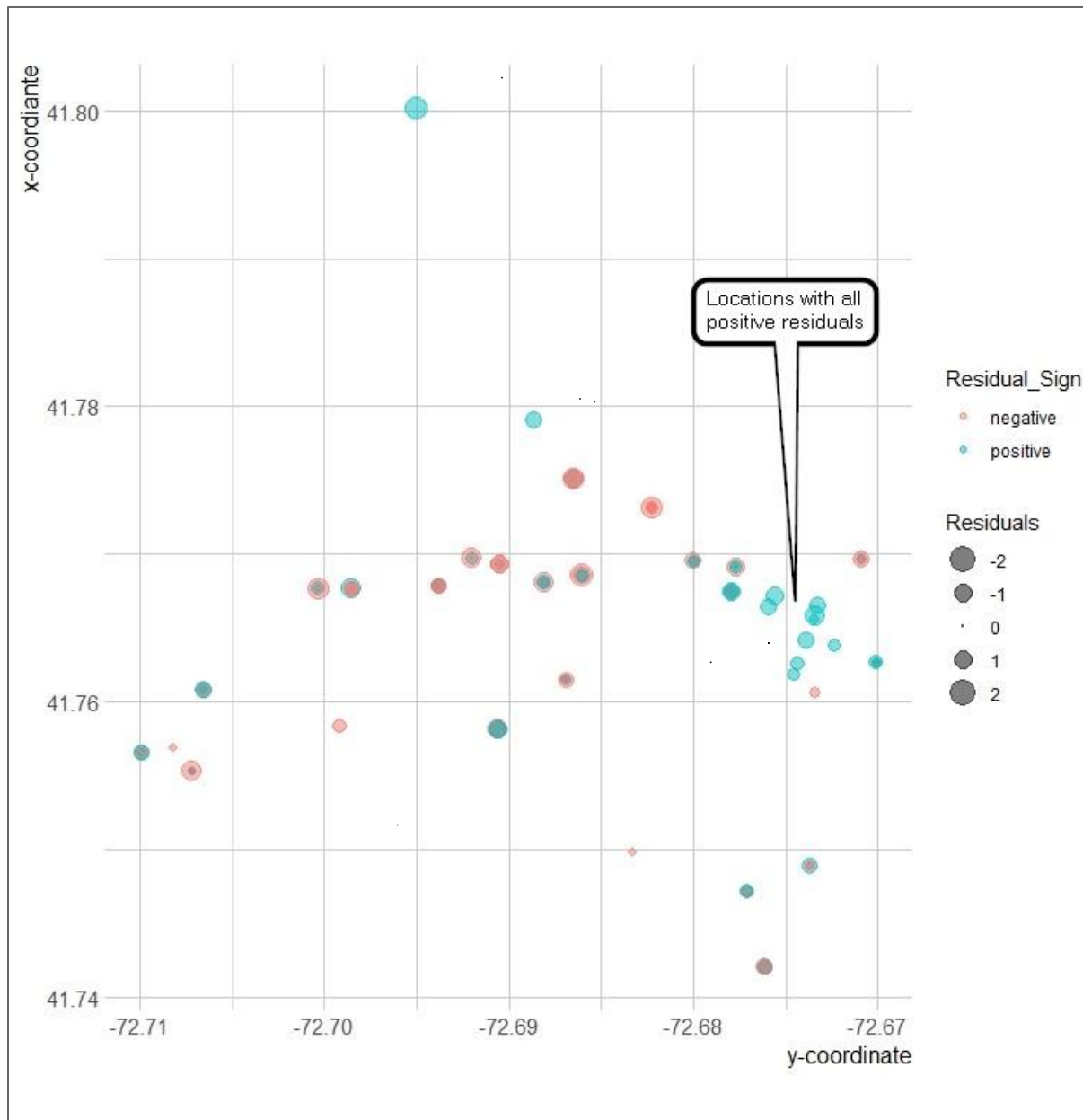
The crosswalks were located at intersections or at mid-block. Some crosswalks were located in the same intersection, crossing up to all three or four entering roadways. There were 206 unique locations each with 1 to 4 crosswalks observed. These crosswalks at the same intersections might have some inherent shared quality or property that influences the pedestrian volume. Such characteristics can be the crime history of the location, ethnic and demographic features of the people around the location, etc., that are not included in the model. Also, the observations at different crosswalks at the same intersection can act like multiple or repeated observations of the same intersection since they will have the same geospatial variables. To address this concern, a random effect model was adopted. In other words, an intersection-specific random term was added to capture the unique characteristics of locations that could not be observed. The model form after introducing the intersection-specific random intercept is given in Equation (3):

$$\ln P_{ji} = \alpha + \beta_1 x_{ji} + \beta_2 w_{ji} + \varepsilon_{ji} + \delta_j \quad (3)$$

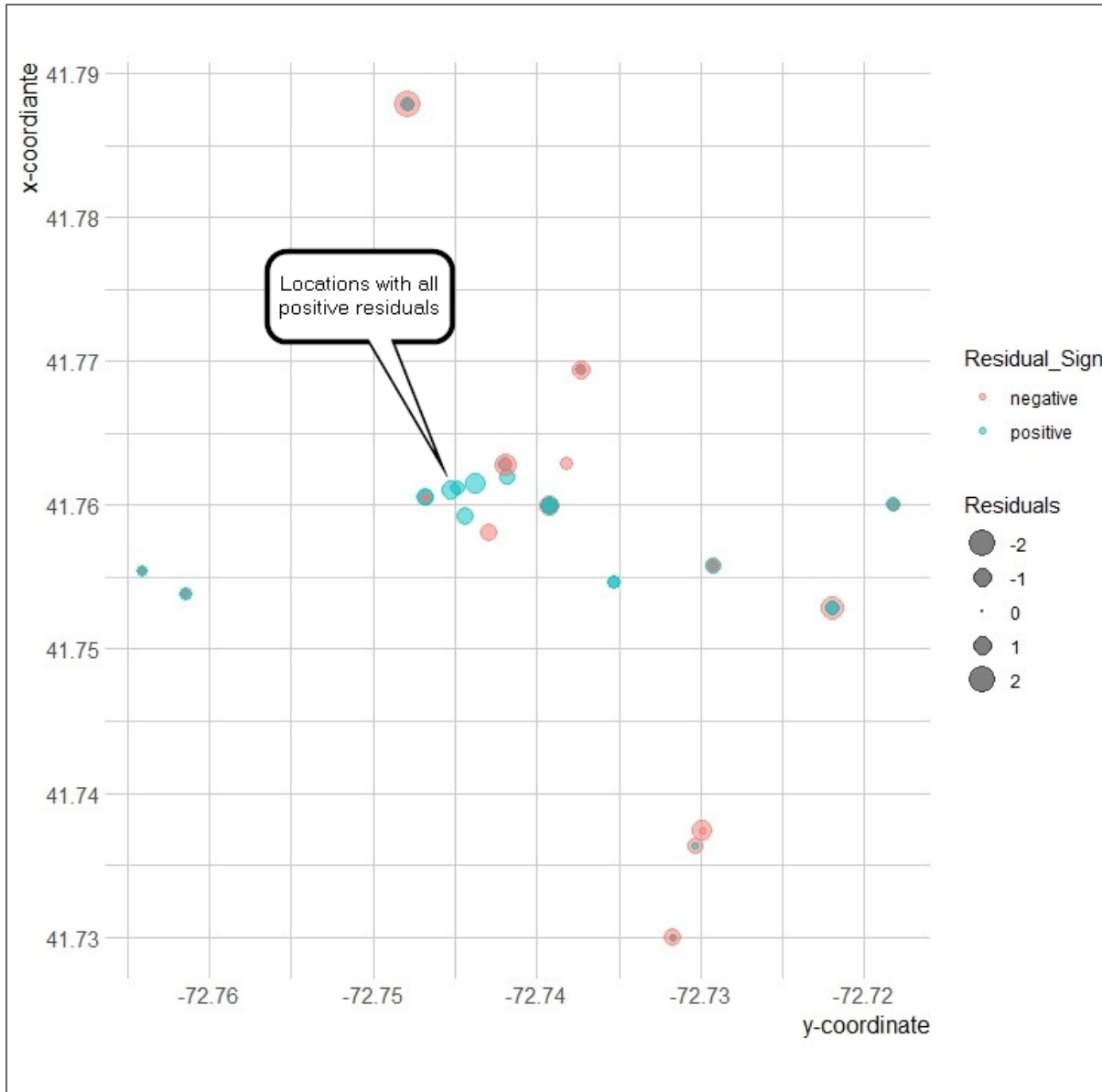
Where  $\delta_j$  is the random effect for intersection  $j$  where crosswalk  $i$  is located, and all other symbols represent variables as explained before.

### 3.5 Spatial Autocorrelation

Tobler's first law of geography states that everything is related to everything else, but near things are more related than distant things (Tobler, 1970). Recognizing this law, we considered the spatial relationship between the pedestrian volumes at nearby locations. To explore the presence of spatial autocorrelation, Moran's I test was carried out on residuals from the modeling results of equation (2). It was found that the Moran's I statistic was close to 0 and statistically insignificant at a 10% significance level. The residuals were spatially plotted by town as well to confirm that there were no visible clusters. The plot of residuals from equation (2) showed possible clusters in two towns: Hartford (Figure 8) and West Hartford (Figure 9). No other towns exhibited any signs of clustering.



**Figure 8 Spatially plotted Residuals from equation (2) for Hartford Locations**



**Figure 9 Spatially plotted Residuals from equation (2) for West Hartford Locations**

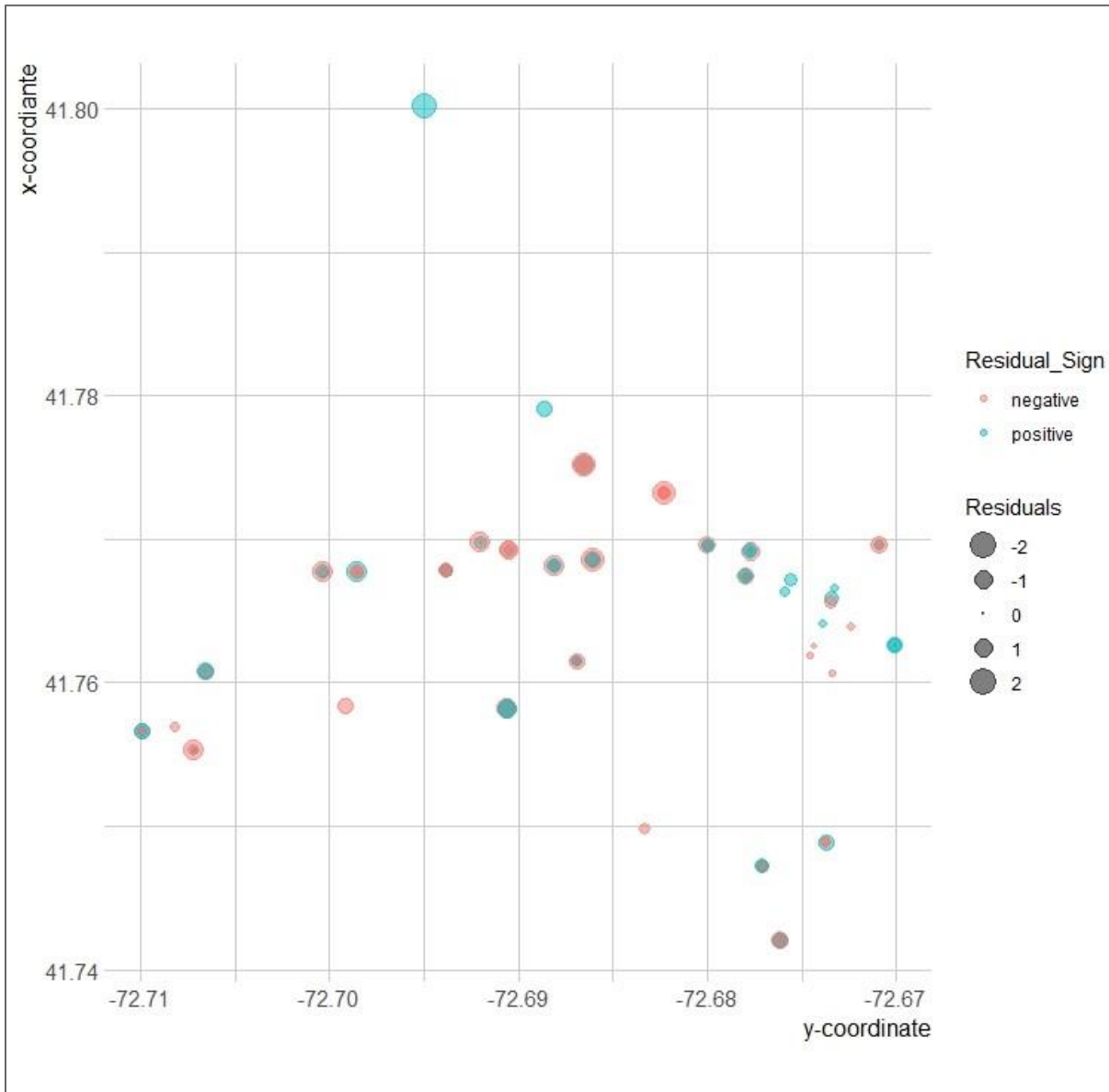
After reexamining the locations in Hartford and West Harford, it was found that the locations that showed a possible clustering were in downtown areas of the towns, and the residuals were positive, indicating underestimation of pedestrian volume at these locations. Therefore, two dichotomous variables representing whether the location is in these two downtown areas or not was included in the model. The inclusion of the new downtown variables was to eliminate the problem which underestimated the pedestrian volume in these areas, as shown in equation (4):

$$\ln P_{ji} = \alpha + \beta_1 x_{ji} + \beta_2 w_{ji} + \beta_3 u_j + \beta_4 v_j + \varepsilon_{ji} + \delta_j \quad (4)$$

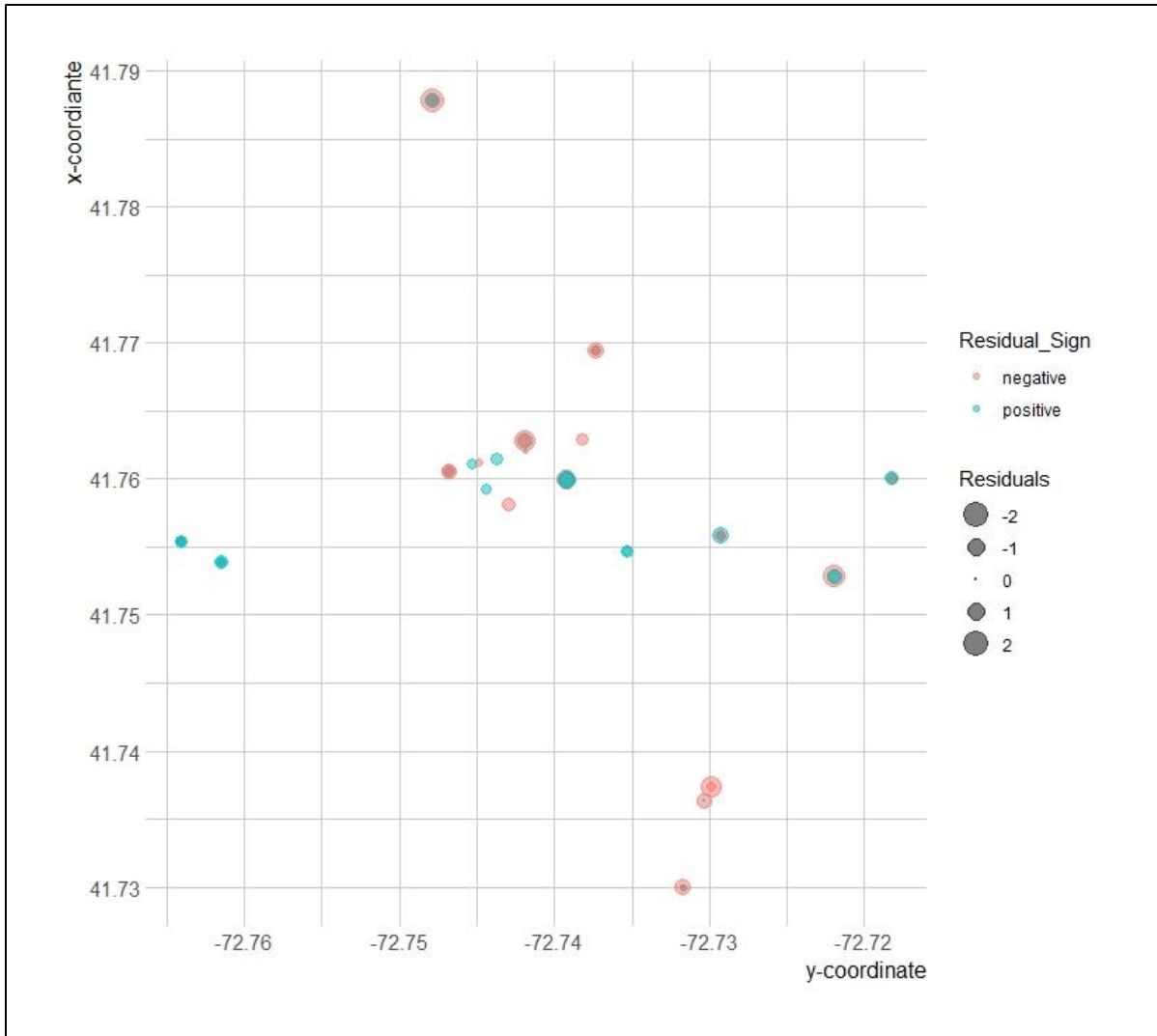
Where  $u_j$  and  $v_j$  are the variables indicating whether the intersection  $j$  is located in downtown Hartford and West Hartford, respectively, and  $\beta_3$  and  $\beta_4$  are the parameters to be determined for these variables.

After including the new downtown variables in the model, the resulting residuals were again checked for any spatial autocorrelation and possible clustering. The results showed that Moran's I statistic was again close to 0 and was not significant at a 10% significance level. The spatial plotting of the residuals in Hartford and West Hartford, shown in Figure 10 and Figure 11 respectively, also did not show any visible clustering like before.

Hence, equation (3) was decided to be the final form of the model.



**Figure 10 Spatially plotted Residuals from equation (3) for Hartford Locations**



**Figure 11 Spatially plotted Residuals from equation (3) for West Hartford Locations**

## 5. RESULTS

The main goal of this project was to find the significance of vehicle-pedestrian conflicts in explaining the variation in pedestrian volumes at crosswalks. Conflict ratio, the ratio of severe conflicts to the less severe interactions, is used as a representation of the severity level of interactions at the crosswalks. Random-effect multiple log-linear regression analysis was carried out using conflict ratio as one of the covariates along with geospatial variables and vehicular volumes. The modeling process was carried out in RStudio (version 1.2.5019) using package “lme4” (version 1.1-23). Three sets of multiple linear regression models were estimated for the three threshold buffers representing walking distance of half-mile, quarter mile, and one-eighth of a mile from the crosswalk. A backward elimination process was used to determine the best model for each buffer size. To start the modeling process, variables were also checked for collinearity between variables and correlation with the dependent variable. Initially, only land use and network variables were used to develop the best model using the geospatial variables to determine the significant geospatial variables explaining variation in pedestrian volume at crosswalks. Then, vehicular volume was introduced in the model to test its significance and improvement in the model. The third model was generated by using conflict ratio and geospatial variables that were found to be significant in the first step—not including vehicular volume. Ultimately, a fourth model using all the significant geospatial variables, vehicular volume, and conflict ratio was developed. AIC and log-likelihood were used to see relative model fit.

### 5.1 Models Using Variables from the One-Eighth Mile Buffer

First, models for the one-eighth mile buffer were generated. The coefficients and standard deviations of all four models are shown in Table 4. Model 1 used the land-use and network characteristics variables. Only four variables were found to be significant at a 5% significance level. A negative coefficient was estimated for the single-family parcel areas indicating that single-family parcels are negatively associated with the pedestrian volume. Higher educational land use has a positive estimate. This means the greater the area of higher educational parcels in the one-eighth mile walking distance from the crosswalk, the higher will be the pedestrian volume. The positive estimates for the dummy variables representing whether the crosswalk is in the downtown area in Hartford and West Hartford show that if the crosswalk is located in the downtown areas of these towns, the pedestrian volume is expected to be higher. Once the best model using geospatial variables was estimated, vehicular volume was introduced to Model 1. The geospatial variables from the initial model remained significant and had the same sign. The estimate for the vehicular volume was computed to be significant and positive. This outcome implies that pedestrian volume is higher at crosswalks where vehicular volume is also higher.

Then, conflict ratio was introduced to Model 1 and Model 2 to develop Model 3 and Model 4, respectively. In both models, the conflict ratio was statistically significant at a 5% level. The results suggested that the conflict ratio is negatively associated with the pedestrian volume. The AIC of Model 4 is the lowest and has the best log-likelihood. Hence, Model 4 is decided to be the best among the four models.

Since this is a log-linear model, an increase in variable impacts the logarithm of the pedestrian volume. Hence, a 1% increase in the conflict ratio multiplies the expected value of pedestrian volume at the crosswalk by  $e^{-4.3491*0.01}$ . As a result, the expected decrease in the pedestrian volume will be  $1 - e^{-4.3491*0.01} = 0.0425$  times the pedestrian volume, which corresponds to a 4.25% decrease.



**Table 4 Models using variables from the One-Eighth mile buffer**

<b>Variable</b>	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
<b>Constant</b>	3.5935 (0.088)	3.2912 (0.128)	3.7417 (0.089)	3.3695 (0.122)
<b>Conflict Ratio</b>			-3.7782 (0.733)	-4.3491 (0.739)
<b>Traffic Volume (100 veh/hr)</b>		0.0334 (0.011)		0.0444 (0.011)
<b>Single Family (10,000 sq. ft.)</b>	-0.6366 (0.064)	-0.6001 (0.061)	-0.6108 (0.061)	-0.5612 (0.057)
<b>Higher Educational (10,000 sq. ft.)</b>	0.3345 (0.045)	0.3581 (0.044)	0.3292 (0.043)	0.3582 (0.042)
<b>Hartford Downtown<sup>a</sup></b>	1.709 (0.317)	1.6822 (0.306)	1.6133 (0.304)	1.5618 (0.289)
<b>West Hartford Downtown<sup>a</sup></b>	0.9228 (0.387)	0.8121 (0.375)	0.8366 (0.37)	0.6773 (0.354)
<b>AIC</b>	894.9	889.5	871.5	858.8
<b>Log-Likelihood</b>	-440.4	-436.7	-427.7	-420.4

Note: All of the variables are significant at a 5% significance level. Numbers in parentheses are the standard deviations of the estimates.

a = dummy variable indicating the crosswalk is in a downtown location

## 5.2 Models Using Variables from the Quarter Mile Buffer

Four models with geospatial variables generated from quarter-mile buffer were also estimated in the same sequence as described in Section 5.1 and are presented in Table 5. Like model 1 in Section 5.1, higher educational land use and location of a crosswalk in the downtown area were positively related to pedestrian volume in model 5. Similarly, single-family land use had a negative and significant coefficient. Two new variables were found to be significant. Office and minor commercial land use were significant with a positive estimate while industrial land use was significant with a negative coefficient. This means a crosswalk with a greater area of the office and minor commercial land use in the quarter-mile walking area will have higher pedestrian volume while those with higher industrial land use will have lower pedestrian volume. In the same manner, model 6 had the same directionality and significance of variables as in model 5, and additionally, vehicular volume had a positive estimate suggesting the same relationship as shown by models in Section 5.1.

**Table 5 Models using variables from the Quarter Mile Buffer**

<b>Variable</b>	<b>Model 5</b>	<b>Model 6</b>	<b>Model 7</b>	<b>Model 8</b>
<b>Constant</b>	3.2966 (0.152)	3.103 (0.161)	3.4943 (0.15)	3.2632 (0.154)
<b>Conflict Ratio</b>			-3.5779 (0.694)	-4.0822 (0.699)
<b>Traffic Volume (100 veh/hr)</b>		0.0291 (0.011)		0.0398 (0.01)
<b>Single Family (10,000 sq. ft.)</b>	-0.1482 (0.016)	-0.1437 (0.016)	-0.1461 (0.016)	-0.1403 (0.015)
<b>Office and Minor Commercial (10,000 sq. ft.)</b>	0.0427 (0.013)	0.0341 (0.013)	0.0368 (0.013)	0.0242 (0.013)
<b>Industrial (10,000 sq. ft.)</b>	-0.2111 (0.09)	-0.2008 (0.088)	-0.2046 (0.086)	-0.1902 (0.083)
<b>Higher Educational (10,000 sq. ft.)</b>	0.1347 (0.015)	0.1365 (0.015)	0.1285 (0.014)	0.1297 (0.014)
<b>Hartford Downtown<sup>a</sup></b>	1.4861 (0.292)	1.4969 (0.283)	1.4088 (0.28)	1.41 (0.268)
<b>West Hartford Downtown<sup>a</sup></b>	1.2314 (0.363)	1.1649 (0.353)	1.1675 (0.347)	1.0689 (0.333)
<b>AIC</b>	851	846.7	827.5	816.4
<b>Log-Likelihood</b>	-416.5	-413.3	-403.8	-397.2

Note: All of the variables are significant at a 5% significance level. Numbers in parentheses are the standard deviations of the estimates.

a = dummy variable for a downtown location

Following these results, the conflict ratio was introduced to both model 5 and model 6 to arrive at model 7 and model 8. The results of these models show a similar relationship between pedestrian volume and conflict ratio. The estimate of conflict ratio was found to be significant and negatively associated with the pedestrian volume. AIC and log-likelihood of these models showed that model 8 consisting of both conflict ratio and vehicular volume along with geospatial variables was the best fit model. From this model, it was determined that a 1% increase in conflict ratio corresponds to a decrease of  $1 - e^{-4.0822*0.01} = 0.04$  times the pedestrian volume, which is 4% less.

### 5.3 Models Using Variables from the Half-Mile Buffer

Four models were developed for the half-mile buffer in a similar way described in Section 5.1 and the results are presented in Table 6. The first model, model 9, determined the significant geospatial variables generated using parcels within a half-mile walking distance from crosswalks for explaining variation in pedestrian volume. Single-family, office and minor commercial, and higher educational parcels in the half-mile buffer were also significant and

positively associated with the pedestrian volume as in models in Section 5.2. Being in the downtown area of Hartford and West Hartford still meant a significant factor in determining pedestrian volume. However, industrial land use in a half-mile buffer was not significant anymore. But a new variable, major commercial land use was found to be significant with a positive estimate. This implies that greater square footage of major commercial land use in the half-mile walking distance from a crosswalk will result in higher pedestrian volume at the crosswalk. No network characteristics variables were found to be significant in previous models from one-eighth and quarter-mile buffers, however, the estimate of average block length generated from half-mile buffer was significant here. The estimate for average block length had a negative sign, therefore, indicating that longer average block lengths will result in lower pedestrian volume. Addition of vehicular volume to model 9 to estimate model 10 was found beneficial as observed in previous models in Section 5.1 and Section 5.2. The positive estimate implies that pedestrian volume is higher at locations with higher vehicular volume.

**Table 6 Models using variables from the Half-Mile Buffer**

<b>Variable</b>	<b>Model 9</b>	<b>Model 10</b>	<b>Model 11</b>	<b>Model 12</b>
<b>Constant</b>	2.645 (0.317)	2.513 (0.306)	2.93 (0.308)	2.805 (0.291)
<b>Conflict Ratio</b>			-3.58 (0.695)	-4.24 (0.7)
<b>Traffic Volume (100 veh/hr)</b>		0.0346 (0.011)		0.0472 (0.01)
<b>Single Family (10,000 sq. ft.)</b>	-0.0235 (0.005)	-0.0246 (0.005)	-0.0249 (0.005)	-0.0267 (0.004)
<b>Multi-Family (10,000 sq. ft.)</b>	0.0097 (0.004)	0.0092 (0.004)	0.0095 (0.004)	0.0087 (0.003)
<b>Major Commercial (10,000 sq. ft.)</b>	0.0373 (0.014)	0.0349 (0.014)	0.0357 (0.013)	0.0325 (0.013)
<b>Office and Minor Commercial (10,000 sq. ft.)</b>	0.029 (0.007)	0.023 (0.007)	0.0246 (0.007)	0.0156 (0.006)
<b>Higher Educational (10,000 sq. ft.)</b>	0.0468 (0.005)	0.046 (0.005)	0.0438 (0.005)	0.042 (0.004)
<b>Average block Length</b>	-0.0008 (0.000)	-0.0008 (0.000)	-0.0008 (0.000)	-0.0007 (0.000)
<b>Hartford Downtown<sup>a</sup></b>	1.462 (0.31)	1.49 (0.298)	1.396 (0.297)	1.418 (0.279)
<b>West Hartford Downtown<sup>a</sup></b>	1.684 (0.391)	1.637 (0.375)	1.659 (0.374)	1.586 (0.35)
<b>AIC</b>	851.9	845.4	828.4	812.9
<b>Log-Likelihood</b>	-414.9	-410.7	-402.2	-393.5

Note: All of the variables are significant at a 5% significance level. Numbers in parentheses are the standard deviations of the estimates.

a = dummy variable indicating for a downtown location

Model 11 and model 12 were estimated introducing conflict ratio to model 9 and model 10, respectively. The variables that were significant in previous models were still significant and the conflict ratio was also significant. The estimate of the conflict ratio had a negative sign as observed in previous models in Section 5.1 and Section 5.2. Based on AIC and log-likelihood model 12 was decided to be the best among the four. The estimate of the conflict ratio in this model indicates that a crosswalk with all other variables same but 1% higher conflict ratio will have  $1 - e^{-4.24*0.01} = 0.0415$  times the initial pedestrian volume, which is 4.15% fewer pedestrians.

## 6. DISCUSSION

Multiple linear regression results presented in the previous section clearly show that the conflict ratio significantly explains pedestrian volumes on the studied crosswalk locations. A negative coefficient was estimated, meaning that a crosswalk with a relatively higher number of severe interactions has fewer pedestrians. Generally speaking, this implies people use crosswalks less where more severe interactions with vehicles are more frequent. This matches the expectations that we had when the research hypothesis was formulated. It is logical that people would avoid routes or locations that have a higher percentage of uncomfortable and dangerous interactions with vehicles. In other words, when all other variables that affect pedestrian volume are the same for two locations, the location where crossing the street involves more severe interactions with vehicles—indicated by a higher conflict ratio—would observe the lower pedestrian volume.

One interesting finding is the association between vehicular volume and pedestrian volume at crosswalks. The estimate of the vehicular volume in models for all three buffers is positive, which implies that the pedestrian volume is higher at locations where the vehicular volume is also higher. This might be because the study locations are mostly located in urban areas in Connecticut, where land use is mostly mixed and higher vehicle and pedestrian activities are expected. Since the study areas have many destinations, the result is higher vehicular volumes along with generating higher pedestrian volumes.

The results of the analysis also found geospatial variables to be significant in explaining pedestrian volume. It is found that square footage of single-family residential parcels in models for all three buffers is negatively associated with pedestrian volume. In contrast, higher square footage of multi-family parcels within half-mile walking distance from the crosswalk is found to be significant. This means that having more multi-family parcels than single family parcels in the vicinity generates higher pedestrian volume. This is likely because multi-family parcels increase population density in an area. Another variable that is found to be significant in models for all three buffers is square footage of institutions of higher education. The positive association of higher educational land use and pedestrian volume means that having greater square footage of higher educational parcels increases the pedestrian volume. This may be because such institutions attract a higher density of road users who usually walk to and from such institutions, such as students.

Moreover, office and minor commercial land uses were found to be positively associated in models generated for the quarter mile and half mile walking distance. The definition of office and minor commercial land-use type consisted of government buildings, banks, retails, food franchises and restaurants, and others. The positive association between pedestrian volume and office and minor commercial is intuitive as well because having facilities like offices and restaurants nearby might generate trips from offices to restaurants and vice versa by foot resulting in greater pedestrian activity.

In addition to these, the square footage of industrial land use in quarter-mile walking distance from crosswalks is found to be related to lower pedestrian volume. This is expected as industrial destinations are usually likely to attract trips by cars and other modes of transportation. Another variable that has negative estimate is the average block length. The relation is significant in the model used with other variables generated using half-mile buffers. This is instinctive because higher average block length indicates lower connectivity due to longer distances between intersections which can be discouraging for walking.

It is also noticed that the dummy variable indicating the location of a crosswalk in the downtown area of Hartford and West Hartford is significant in all models. The estimates for these variables are positive indicating that without these variables the model would underestimate pedestrian volumes at these locations. The reason behind this might be that the pedestrian volumes at these locations are affected by other variables that are inherent to these locations which are not observed in the model by any other variables. Such variables can be cultural, social, or localized street design factor which were not included in the model.

Ultimately, when all these models are observed and compared it can be seen that the variables are sensitive to the considered buffer distance. Not all variables are significant in all three models from three different buffers.

## 7. CONCLUSION

The main goal of this project was to examine the association between pedestrian volume and pedestrian-vehicle interactions on crosswalks. This was done by estimating random effect log-linear regression models. Pedestrian-vehicle interactions were classified by their severity level. The conflict ratio is defined as the ratio of highly severe interactions to less severe interactions was used as the representation of the level of severity of interactions on the crosswalk. Land use variables, network connectivity, and vehicular volume were also used as covariates in the models. Models for three buffer distances, half-mile, quarter mile, and one-eighth of a mile, were estimated. The buffer distances were selected based on past research results and with the assumption that people usually walk for trips that are within this distance.

The results showed a significant negative association between pedestrian volume and conflict ratio indicating that 1% increase in conflict ratio can result into about 4% reduction in pedestrian volume. Lower pedestrian volumes are observed at the locations where conflict ratio is higher. Conflict ratio represented the discomfort while crossing road sections. This finding suggests that people walk less on crosswalks that have higher number of unpleasant and uncomfortable interactions with vehicles. This result can be helpful to guide planners and designers on decision making for design consideration of a crossing location. For example, one of the reasons for severe interactions might be that drivers do not become aware of the crosswalk location due to its low visibility. Improving visibility with reflective crosswalk signs and flashing beacons can help drivers to see the crosswalk and have enough time to react to pedestrians and make the interactions safer while crossing. Bulb-outs to move pedestrians from behind the parked vehicles at intersections can also alert drivers to the possible presence of pedestrians. Raised crosswalks and textured crosswalks are other methods to warn drivers about the presence of the crossing location. Another reason for severe interactions might be higher speed of the vehicles at the crossing locations. This can be tackled by introducing speed humps several feet before to decrease the speed of the vehicle and make sure that drivers are paying enough attention to their surroundings. Refuge islands can also help to ease the crossings (Federal Highway Administration, 2013). Whenever pedestrians feel unsafe to continue crossing, they can wait at the islands and continue crossing when it is safer. Right-turn-on-red movements can cause frequent conflicts with pedestrians (Furth et al., 2014). Limiting such movements can help pedestrians to feel safe to cross streets. With these solutions that make the road crossing comfortable and easier, people may perceive the locations as safer and walk often.

The results of the analysis also found other land use variables that are significant in explaining variation in pedestrian volume. Having higher square footage of single-family parcels, industrial parcels, and greater average block length are found to have negative association with pedestrian volume. Higher square footage of multi-family parcels, major commercial, and office and minor commercials is found to have positive impact on pedestrian volumes. These results can be used for land use planning to increase the share of walking.

It is recommended that the accommodation of pedestrians at crossing locations be given proper consideration while designing. Planning land use and designing pedestrian friendly street networks with adequate solutions at road crossing location to make the interactions with vehicles safe and comfortable can help planners to achieve their environmental, societal, and economical goals by promoting walking.

This study found that the pedestrian-vehicle conflict plays significant role in explaining variation in pedestrian volume at crossing locations. There are a few limitations of the findings. The crosswalk locations are all located in Connecticut. The cultural and behavioral differences in

other places might play a role in the relationship. The conflict ratio is determined by observing the conflicts using a modified version of Swedish Traffic Conflict Technique. Using a different procedure and method might also result in different findings. The conflict observation time varied for different locations due to adverse weather conditions, absence of pedestrian, etc. The conflicts were all observed during the daytime as well.

Based on the past research on perception of safety, this study opened a potential use of interaction of pedestrian and vehicle at crossing locations to quantify pedestrians' perception of safety. A future study that surveys the pedestrians' interaction experience with vehicles at crosswalks validating the relationship between severity of interactions and perception of safety would be a contribution to behavioral research in transportation. Another potential research idea would be to see the relationship between presence of different signage and infrastructure at crosswalk and conflict ratio. If such a relationship is found and can be clustered, an index based on the conflict ratio can then be used as a screening tool for evaluating safety at the crosswalks. The difference in effects of the pedestrian vehicle interaction across various age groups and genders can also be studied as a future research.



## REFERENCES

1. Ajzen, I., 1991. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), pp.179-211.
2. Backer-Grøndahl, A., Fyhri, A., Ulleberg, P. and Amundsen, A.H., 2009. Accidents and unpleasant incidents: worry in transport and prediction of travel behavior. *Risk Analysis: An International Journal*, 29(9), pp.1217-1226.
3. Behnam, J. and Patel, B.G., 1977. A method for estimating pedestrian volume in a central business district. *Transportation Research Record*, 629, pp.22-26.
4. Bian, Y., Ma, J., Rong, J., Wang, W. and Lu, J., 2009. Pedestrians' level of service at signalized intersections in China. *Transportation research record*, 2114(1), pp.83-89.
5. Brown Jr, E.G., Kelly, B.P., Dougherty, M. and Ajise, K., 2015. Caltrans Strategic Management Plan 2015–2020. California Department of Transportation.
6. Clifton, K.J., Muhs, C.D., Morrissey, S., Morrissey, T., Currans, K.M. and Ritter, C., 2013. Examining consumer behavior and travel choices. OTREC-RR-12-15. Portland, OR: Transportation Research and Education Center (TREC)
7. Do, M.T., Grembek, O., Ragland, D. and Chan, C.Y., 2013, January. Weighting integration by block heterogeneity to evaluate pedestrian activity. In *TRB 92nd Annual Meeting (Transportation Research Board)* (pp. 14p-schémas).
8. ESRI 2019. ArcGIS Desktop 10.7.1. Environmental Systems Research Institute.
9. Ewing, R. and Cervero, R., 2001. Travel and the built environment: a synthesis. *Transportation research record*, 1780(1), pp.87-114.
10. Federal Highway Administration, 2013. “Crossing Islands” in PEDSAFE: Pedestrian Safety Guide and Countermeasure Selection System. Available: [http://www.pedbikesafe.org/PEDSAFE/countermeasures\\_detail.cfm?CM\\_NUM=6](http://www.pedbikesafe.org/PEDSAFE/countermeasures_detail.cfm?CM_NUM=6)
11. Furth, P.G., Koonce, P.J., Miao, Y., Peng, F. and Littman, M., 2014. Mitigating Right-Turn Conflict with Protected Yet Concurrent Phasing for Cycle Track and Pedestrian Crossings. *Transportation Research Record*, 2438(1), pp.81-88.
12. Fyhri, A. and Backer-Grøndahl, A., 2012. Personality and risk perception in transport. *Accident Analysis & Prevention*, 49, pp.470-475.
13. Giles-Corti, B., Vernez-Moudon, A., Reis, R., Turrell, G., Dannenberg, A.L., Badland, H., Foster, S., Lowe, M., Sallis, J.F., Stevenson, M. and Owen, N., 2016. City planning and population health: a global challenge. *The lancet*, 388(10062), pp.2912-2924.
14. Hanson, S. and Jones, A., 2015. Is there evidence that walking groups have health benefits? A systematic review and meta-analysis. *British journal of sports medicine*, 49(11), pp.710-715.
15. Hyden, C. and Linderholm, L., 1984. The Swedish traffic-conflicts technique. In *International Calibration Study of Traffic Conflict Techniques* (pp. 133-139). Springer, Berlin, Heidelberg.
16. Islam, M.S., Serhiyenko, V., Ivan, J.N., Ravishanker, N. and Garder, P.E., 2014. Explaining pedestrian safety experience at urban and suburban street crossings considering observed conflicts and pedestrian counts. *Journal of Transportation Safety & Security*, 6(4), pp.335-355.

17. Ivan, J., Ravishanker, N., 2015a. Investigation of Road and Roadside Design Elements Associated with Elderly Pedestrian Safety. Region 1 University Transportation Center, Final Report, Project UCNR24-30A.
18. Ivan, J., Ravishanker, N., 2015b. Effectiveness of Interventions at Midblock Crossings for Improving Senior and Other Pedestrian Safety. New England University Transportation Center, Final Report, Project UCNR24-30A.
19. Ivan, J.N., Ravishanker, N. and Islam, M.S., 2012. *Evaluation of Surrogate Measures for Pedestrian Safety in Various Road and Roadside Environments* (No. 11-04). Center for Transportation and Livable Systems.
20. LaHood, R., 2010. United States Department of Transportation Policy Statement on Bicycle and Pedestrian Accommodation Regulations and Recommendations. *Signed on March, 11*.
21. Liu, X. and Griswold, J., 2009. Pedestrian volume modeling: a case study of San Francisco. *Yearbook of the Association of Pacific Coast Geographers*, pp.164-181.
22. Lu, G.X., Sullivan, J. and Troy, A., 2012. Impact of ambient built-environment attributes on sustainable travel modes: a spatial analysis in Chittenden county, Vermont. In *91st Annual Meeting of the Transportation Research Board. Washington, DC* (Vol. 21).
23. Miranda-Moreno, L.F., Morency, P. and El-Geneidy, A.M., 2011. The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. *Accident Analysis & Prevention*, 43(5), pp.1624-1634.
24. Muraleetharan, T., Adachi, T., Hagiwara, T. and Kagaya, S., 2005. Method to determine pedestrian level-of-service for crosswalks at urban intersections. *Journal of the Eastern Asia Society for Transportation Studies*, 6, pp.127-136.
25. Murphy, M.H., Nevill, A.M., Murtagh, E.M. and Holder, R.L., 2007. The effect of walking on fitness, fatness and resting blood pressure: a meta-analysis of randomised, controlled trials. *Preventive medicine*, 44(5), pp.377-385.
26. Ni, Y., Cao, Y. and Li, K., 2017. Pedestrians' safety perception at signalized intersections in Shanghai. *Transportation research procedia*, 25, pp.1955-1963.
27. OpenStreetMap Contributors. <https://www.openstreetmap.org/copyright>
28. Pulugurtha, S.S. and Repaka, S.R., 2008. Assessment of models to measure pedestrian activity at signalized intersections. *Transportation Research Record*, 2073(1), pp.39-48.
29. Quistberg, D.A., Howard, E.J., Hurvitz, P.M., Moudon, A.V., Ebel, B.E., Rivara, F.P. and Saelens, B.E., 2017. The relationship between objectively measured walking and risk of pedestrian–motor vehicle collision. *American journal of epidemiology*, 185(9), pp.810-821.
30. Rails-to-Trails Conservancy, 2018. About the Partnership. <https://www.railstotrails.org/partnership-for-active-transportation/about/>
31. RStudio Team, 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. <http://www.rstudio.com/>
32. Schneider, R.J., Arnold, L.S. and Ragland, D.R., 2009. Pilot model for estimating pedestrian intersection crossing volumes. *Transportation research record*, 2140(1), pp.13-26.

33. Shapiro, S.S. & Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3–4), pp.591–611. Available at: <https://doi.org/10.1093/biomet/52.3-4.591>.
34. Sucha, M., Dostal, D. and Risser, R., 2017. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention*, 102, pp.41-50.
35. Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1), pp.234-240.
36. US Department of Health and Human Services, 2015. Step it up! the surgeon general’s call to action to promote walking and walkable communities. *Washington, DC: US Dept of Health and Human Services, Office of the Surgeon General*. <http://www.surgeongeneral.gov/library/calls/walking-and-walkable-communities>.
37. Yang, H., Lu, X., Cherry, C., Liu, X. and Li, Y., 2017. Spatial variations in active mode trip volume at intersections: a local analysis utilizing geographically weighted regression. *Journal of Transport Geography*, 64, pp.184-194.
38. Zhang, Y., Mamun, S.A., Ivan, J.N., Ravishanker, N. and Haque, K., 2015. Safety effects of exclusive and concurrent signal phasing for pedestrian crossing. *Accident Analysis & Prevention*, 83, pp.26-36.
39. Zhang, Y., Ravishanker, N., Ivan, J.N. and Mamun, S.A., 2019. An Application of the Tau-Path Method in Highway Safety. *Journal of the Indian Society for Probability and Statistics*, 20(1), pp.117-139.
40. Zupan, J.M. and Pushkarev, B., 1971. Pedestrian Travel Demand. *Highway Research Record*, 355, pp.37-53.