Accident Analysis and Prevention 131 (2019) 15–24

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

A geographically weighted regression to estimate the comprehensive cost of traffic crashes at a zonal level

Amin Mohamadi Hezaveh, Ramin Arvin, Christopher R. Cherry*

Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, United States

ARTICLE INFO

Keywords:
Comprehensive crash cost
Home-based approach
Geographically weighted regression
Road safety

ABSTRACT

Global road safety records demonstrate spatial variation of comprehensive cost of traffic crashes across countries. To the best of our knowledge, no study has explored the variation of this matter at a local geographical level. This study proposes a method to estimate the comprehensive crash cost at the zonal level by using person-injury cost. The current metric of road safety attributes safety to the location of the crash, which makes it challenging to assign the crash cost to home-location of the individuals who were involved in traffic crashes. To overcome this limitation, we defined Home-Based Approach crash frequency as the expected number of crashes by severity that road users who live in a certain geographic area have during a specified period. Using crash data from Tennessee, we assign those involved in traffic crashes to the census tract corresponding to their home address. The average Comprehensive Crash Cost at the Zonal Level (CCCAZ) for the period of the study was $18.2 million (2018 dollars). Poisson and Geographically Weighted Poisson Regression (GWPR) models were used to analyzing the data. The GWPR model was more suitable compared to the global model to address spatial heterogeneity. Findings indicate population of people over 60-years-old, the proportion of residents that use non-motorized transportation, household income, population density, household size, and metropolitan indicator have a negative association with CCCAZ. Alternatively, VMT, vehicle per capita, percent educated over 25-year-old, population under 16-year-old, and proportion of non-white races and individuals who use a motorcycle as their commute mode have a positive association with CCCAZ. Findings are discussed in line with road safety literature.

1. Introduction

Annually more than 1.2 million individuals lose their lives on roads globally. Likewise, between 20 to 50 million individuals are impacted by serious and sometimes permanent injuries in traffic crashes. World Health Organization reports on road safety indicate road fatality rate (death per 100,000 population) varies across countries (WHO, 2015b). The traffic fatality rate is approximately two times higher in low and middle-income countries compared to high-income countries (21.5, 19.5, and 10.3 per 100,000 respectively). Spatial variation in road safety performance indicators and social costs of traffic crashes on roads could be attributed to several external factors; namely safety measures and programs, traffic structure, and culture of a country (Koornstra et al., 2002). These factors also reflect study area characteristics such as demographics, weather, reporting practices, and the economy (Koornstra et al., 2002).

Social cost or Comprehensive crash cost of traffic crashes consists of several components including medical cost, loss of production capacity, costs of property damage, administrative costs, and economic valuation of lost quality of life (Elvik et al., 2009; Harmon et al., 2018). These components could be categorized into two main categories; tangible and intangible cost. Tangible costs reports the economic costs of traffic crashes that can be directly measured such as medical bills and lost wages. The intangible costs comprise the other impacts of crashes and can be monetized as quality-adjusted life years (QALY) (Harmon et al., 2018).

Road safety literature provides several instances of studies regarding the comprehensive crash cost at country level (e.g., Mohan, 2002; García-Altés and Pérez, 2007; Wegman and Oppe, 2010; Ahadi and Razi-Ardakani, 2015; Blincoe et al., 2015). Findings indicate that traffic crashes cost 1–2% of Gross Domestic Product (GDP) of high-income countries and 3% of GDP in low and middle-income countries. (Jacobs et al., 2000. In the United States, societal harm from traffic crashes in 2010 was estimated to be over $836 billion. the economic cost of traffic crashes in the US was estimated to be over $242 billion, which is equal to 1.6% of the US GDP (Blincoe et al., 2015). This trend

* Corresponding author.
E-mail address: Cherry@utk.edu (C.R. Cherry).

https://doi.org/10.1016/j.aap.2019.05.028
Received 3 October 2018; Received in revised form 21 February 2019; Accepted 29 May 2019
Available online 21 June 2019

also holds within a country; for example, several studies in the United States showed that in rural areas the fatality rate is several times higher than the majority of urban areas (Marshall and Ferenchak, 2017). In addition, some ethnicities such as Hispanic, African-American, and Native American are more prone to traffic crashes –i.e., they have higher crash rates (Mayrose and Jehle, 2002; Braver, 2003; Campos-Outcalt et al., 2003; McAndrews et al., 2013); this is also the case for the fatality rate (Schiff and Becker, 1996; Baker et al., 1998; Harper et al., 2000). Furthermore, vulnerable road users (i.e., pedestrians and bicyclists) and lower-income neighborhoods have higher fatality rates compared to motorized road users and wealthier neighborhoods (Marshall and Ferenchak, 2017).

To the best of our knowledge, the road safety literature has abundant examples of estimating the societal outcome of traffic crashes at aggregate level (i.e., country); however, there are no studies that investigated comprehensive crash cost at a fine geographical level (e.g., traffic analysis zone, census tract) and explores the factors correlating with it by using a Macroscopic Crash Prediction Models (MCPM). Furthermore, we may expect that traffic crashes do not impact geographic areas in equitable ways. Therefore, we expect the comprehensive cost of traffic crashes has spatial variation within a country, city or finer geographic unit. In addition, learning about the areas where their residents are more prone to the burden of traffic crashes would help safety practitioners and researchers to allocate proper countermeasures to reduce the burden of traffic crashes or providing resources to reduce the burden of traffic crashes.

1.1. Macroscopic crash prediction models

In order to study the societal cost of traffic crashes, we will measure societal cost at the zonal level by using MCPM. MCPM is a set of methods that are used to investigate the relationship between safety at zonal level and socioeconomic factors (Aguero-Valverde and Jovanis, 2006; Hadayeghi et al., 2010b; Huang and Abdel-Aty, 2010; Pulugurtha et al., 2013), travel behavior (Naderan and Shahi, 2010), road infrastructure and traffic flow (Quddus, 2008; Huang and Abdel-Aty, 2010; Abdel-Aty et al., 2011; Pirdavani et al., 2012; Xu and Huang, 2015), and environment condition (Aguero-Valverde and Jovanis, 2006). Various type of spatial units have been used by researchers; from fine level such as traffic analysis zones (Hadayeghi et al., 2010b; Pirdavani et al., 2012; Pirdavani et al., 2013; Pulugurtha et al., 2013; Xu and Huang, 2015, Dong et al. 2016, Gomes et al., 2017) and census tracts (Ukkusuri et al., 2011; Wang and Kockelman, 2013; Hezaveh and Cherry, 2019a), block groups (Levine et al., 1995) to coarser levels such as zip codes (Girasek and Taylor, 2010), districts (Haynes et al., 2007), counties (Miaou et al., 2003; Aguero-Valverde and Jovanis, 2006; Huang and Abdel-Aty, 2010), and regions (Washington et al., 1999).

In MCPM, safety usually is measured with different indices, namely number of all traffic crashes (Miaou et al., 2003; Naderan and Shahi, 2010; Pirdavani et al., 2012; Pirdavani et al., 2013, Huang et al., 2016; Cai et al., 2017; Hezaveh and Cherry, 2019a), number of property damage only crashes (Naderan and Shahi, 2010; Aguero-Valverde, 2013), frequency of injury/severe crashes (Aguero-Valverde, 2013; Xu and Huang, 2015; Cai et al., 2017), and crashes of specific road users (e.g., non-motorized, bicyclists) (Lee et al., 2015; Cai et al., 2017; Cheng et al., 2018; Saha et al., 2018). Although using different dependent variables enable researchers to investigate the correlation of exogenous variables and traffic crash outcomes, it does not provide information about the association of exogenous variable and ultimate outcome – societal cost (i.e., comprehensive cost of traffic crashes) - of road safety. For example, the crash rates in an urban area are higher than rural area; but the fatal injury rate and crash injury rate in the urban area is lower than the rural area (Zwerling et al., 2005; NHTSA, 2013).

Learning about the relationship between exogenous variables and comprehensive cost of traffic crashes would help safety practitioners and researchers prioritize their countermeasures based on the monetary values of traffic crashes regardless of road user type, the location of the crash (e.g., rural vs. urban), and crash severity. The current practices of assessing road safety relies on MCPM based on the location of the crash. This metric was best described by Hauer (1997) "*the number of accidents (crashes) by kind and severity, expected to occur on the entity during a specified period.*" This definition attributes road safety to the location of the crashes rather than individuals who were involved in traffic crashes (i.e., pedestrians, bicyclists, motorcyclists, vehicle occupants, and drivers). As a result, it is challenging to attribute crash burden to the location where individuals reside by using this definition.

Another concern in MCPM models is spatial heterogeneity or spatial non-stationary (Fotheringham et al., 2002; LeSage and Pace, 2009). Spatial heterogeneity exists when exogenous variables do not vary identically across space (Xu et al., 2017). One reason for the presence of unobserved heterogeneity in the data is the presence of unknown or known factors that are unlikely to be available for the analysis (Mannering et al., 2016). This phenomenon influences the association among exogenous variables and dependent variables; as a result, this relationship may not be constant across the observation. Failing to consider the unobserved heterogeneity in count data analysis would lead to overdispersion; hence, the variance of the exogenous variable is larger than the mean (Cameron and Trivedi, 1986; Gourieroux and Visser, 1997). Likewise, If unobserved factors correlate with known exogenous factors, the estimates would yield biased parameters, which eventually lead to drawing incorrect inferences (Mannering et al., 2016).

There are different methods to address the heterogeneity in count models. Random parameters count data, and geographically weighted poison regression (GWPR) are two common methods to address this issue (Xu and Huang, 2015; Arvin et al., 2019). Random parameters models are drawn from some random distribution and are assumed to vary randomly over observations (Xu and Huang, 2015). One of the shortcomings of the random parameter model is that it usually fails to reflect the location of observation. Alternatively, spatial models consider the location of the observations to capture spatially structured variability in the effect of contributing factors (Xu and Huang, 2015; Xu et al., 2017). Several studies showed the advantage of GWPR models with regards to improvement in model goodness of fit and capability to explore the spatially varying association among dependent variables and contributing factors (Hadayeghi et al., 2010b; Pirdavani et al., 2014; Xu and Huang, 2015; Xu et al., 2017). In addition, estimated parameters of the GWPR reflects local characteristics by enabling coefficients to vary across the study area; therefore, GWPR results could be used as a reference for transportation agencies focusing on geographic differences (Chiou et al., 2015).

Regarding the current state of practice in MCPM, this study has three aims. First, it aims to measure comprehensive crash cost at the zonal level by using the home address of the road users (i.e., Home-Based Approach). Secondly, we aim to display the geographical distribution of the comprehensive cost of traffic crash across the study area. Third, we will use a geographically weighted regression to account for spatial heterogeneity and explore the relationship between sociodemographic factors and comprehensive crash cost at the zonal level. The result of this study could help safety practitioners and researchers to identify the neighborhoods where their residents are more prone to the burden of traffic crashes and target them with proper countermeasures or interventions to reduce their risk.

2. Methodology

2.1. Data and geocoding process

To achieve the aims of this study, we need to attribute the safety outcome of traffic crashes to the home address of the individuals who were involved in traffic crashes. Therefore, we defined the HBA to measure the expected number of crashes by severity that road users who live in a certain geographic area have during a specified period. This definition attributes traffic crashes to individuals and their residential addresses rather than the location of traffic crashes. We use crash frequency by severity to calculate the comprehensive crash cost at each zone.

The data in this study was provided by the Tennessee Integrated Traffic Analysis Network (TITAN), a portal provided by Tennessee Highway Patrol (THP). The records include 694,276 crashes and information on 2,026,666 individuals who were involved in traffic crashes between 2014–16. Each record includes information about road user type (e.g., driver, motorcyclist, passenger, pedestrian, bicyclist), coordinates of the crashes and addresses of all individuals who were involved in traffic crashes. It is also worthy to mention that the TITAN database does not identify at-fault road users in traffic crashes so we could not conduct any analysis on the role of each person in the traffic crash.

After obtaining the address of the pedestrians, bicyclists, motorcyclists, drivers, and vehicle occupants (n = 1,615,374), we used the Bing and Google application program interface services to geocode the addresses. The quality of the geocoding was checked by controlling for the locality of the addresses. Only those records that had an accuracy level of premises (e.g., property name, building name), address level accuracy, or intersection level accuracy was used for the analysis (Hezaveh and Cherry, 2019a,b, Merlin et al., 2019). After controlling for the address quality, 1,521,583 (94.1%) of the records met the minimum address quality filter. Of those, 1,358,117 had a Tennessee home-address (89.3% of geocoded addresses); the number out of state individuals was 163,466 (10.7% of geocoded addresses).

In this study, one goal was to investigate the relationship between sociodemographic variables and crash frequency at the zonal level. For that reason, we used the census tract as the geographic unit. Census data from the US survey in 2010 was also used to obtain sociodemographic data elements in each census tract in Tennessee. Table 1 presents the descriptive statistics of the sociodemographic variables obtained from the US census in 2010. Fig. 1 presents the histogram of HBA crash frequency by severity at the zonal level in this study.

Furthermore, we used highway performance monitoring system data for Tennessee in 2015 to obtain Average Annual Daily Traffic for each road segment and calculate total Vehicle Miles Travelled (VMT) at the census tract level.

Table 1
Descriptive statistics of the variables.
Source: United States Census and HPMS.

Variable	Mean	SD	Range
Total Population	1526	789	[0, 9281]
Population Density (Person per square km)	625	979	[0, 32989]
Average Household Size	2.72	5.3	[0, 243.18]
Race Proportion			
White	0.77	0.3	[0, 1]
Non-White	0.22	0.28	[0, 1]
Means Of Transportation To Work Proportion			
Personal Vehicle	0.92	0.11	[0, 1]
Carpool	0.1	0.08	[0, 0.82]
Bus	0.01	0.04	[0, 0.62]
Motorcycle	0	0.01	[0, 0.17]
Bicycle	0	0.01	[0, 0.18]
Walk	0.02	0.05	[0, 1]
Other Means	0.01	0.03	[0, 0.6]
Age Cohort Proportion			
16 Years And Younger	0.23	0.08	[0, 0.71]
16-42 Years Old	0.32	0.11	[0, 1]
43-59 Years Old	0.25	0.08	[0, 1]
60 Years Old And More	0.2	0.1	[0, 1]
Vehicles' Ownership Per Capita	0.69	0.16	[0, 1.2]
% Of Educated People Over 25 Years Old	67.67	10.37	[0,99.93]
Housing Unit			
Percent Of Vacant Housing Unit	0.12	0.1	[0, 1]
VMT (1,000,000)	0.57	0.69	[0, .74]
Average Travel Time To Work (Minutes)	25.1	6.6	[0, 65.85]
Median Household Income ($1,000)	45.7	25.09	[0, 249.3]

2.2. Comprehensive crash cost at the zonal level

The injury severity in this database followed the KABCO scale for Tennessee. In the KABCO scale K, A, B, C, and O respectively stand for a crash with fatal, Incapacitating, Non-Incapacitating Evident, Possible Injury, and No Injury (FHWA, 2017). The comprehensive crash cost consisted of two elements: economic person-injury unit costs and quality-adjusted life years (QALY), which respectively take account for tangible and intangible consequences of traffic crashes. In order to convert the injury severities to crash cost, we used number presented in Table 2, which are based on FHWA recommendation for the person-injury unit cost (Harmon et al., 2018). It is also worthy to mention that, in their report, Harmon et al. (2018) considered injury misclassification that controlled for more accurate injury accounting at emergency departments. For more details about this issue please see Gomes et al. (2017). Furthermore, based on the Harmon et al. (2018) study recommendation, we updated the person injury unit cost to reflect 2018 US Dollars. For more details, please see Harmon et al. (2018).

In this study, comprehensive cost of traffic crashes at the zonal level consists of two parts. The first part reflects the injury severity cost and the second part reflects the vehicle damage cost. By using numbers presented in Table 2 and counting crash frequencies by severity at each census tract, Comprehensive Crash Cost at Zonal Level (CCCAZ) at census tract was calculated using the following equation:

$$CCCAZ_i = PCI*\left(\sum_{\alpha=\{K,A,B,C,O\}} N_{\alpha,i}*Cost_{\alpha} + (N_{v,i}*Cost_{PDO})\right) \quad (1)$$

where $N_{\alpha,i}$ represents the number of individual who live in zone i with the level of injury α, and $Cost_{\alpha}$ presents the traffic injury cost per injury presented in Table 2. $N_{v,i}$ presents the number of vehicles with a registered address in zone i which were involved in traffic crashes, $Cost_{PDO}$ also presents the vehicle unit damage cost. PCI also represents the Tennessee crash cost Per Capita Income (PCI) ratio adjustment factor, the PCI of the state of Tennessee for year 2018 was 0.855 (Bureau of Economic Analysis, 2018). Fig. 2 presents the histogram of the Comprehensive Crash Cost at the zonal level for the period of the study.

2.3. Modeling approach

To evaluate safety at the zonal level, traditionally, count data models such as Poisson, negative binomial and zero-inflated models are commonly utilized owing to the nature of traffic crashes that are usually measured as non-negative integers in a specific period of time (Anastasopoulos and Mannering, 2009). Similar to crash frequency, the comprehensive cost of the traffic crash is a non-negative integer. Hence the models that would be used to evaluate CCCAZ must follow the nature of counts model. In addition, variations in the relationship over space also could be present in the data (i.e., spatial heterogeneity) (LeSage and Pace, 2009). The stationary relationship may hide some spatial factors affecting the safety at the zonal level, which may eventually affect the accuracy of models only use one constant coefficient for the study area (i.e., global model). Using the analogy between crash frequency and CCCAZ, in this study, we will use the Poisson Regression model and the Geographically Weighted Poisson Regression Model (GWPR). To directly model the CCCAZ as an integer variable, we also used the population of each census tract as the generalized offset variable. We used the GWPR model to reflect the role of spatial heterogeneity in the modeling process by using the coordinates of the center of census tracts.

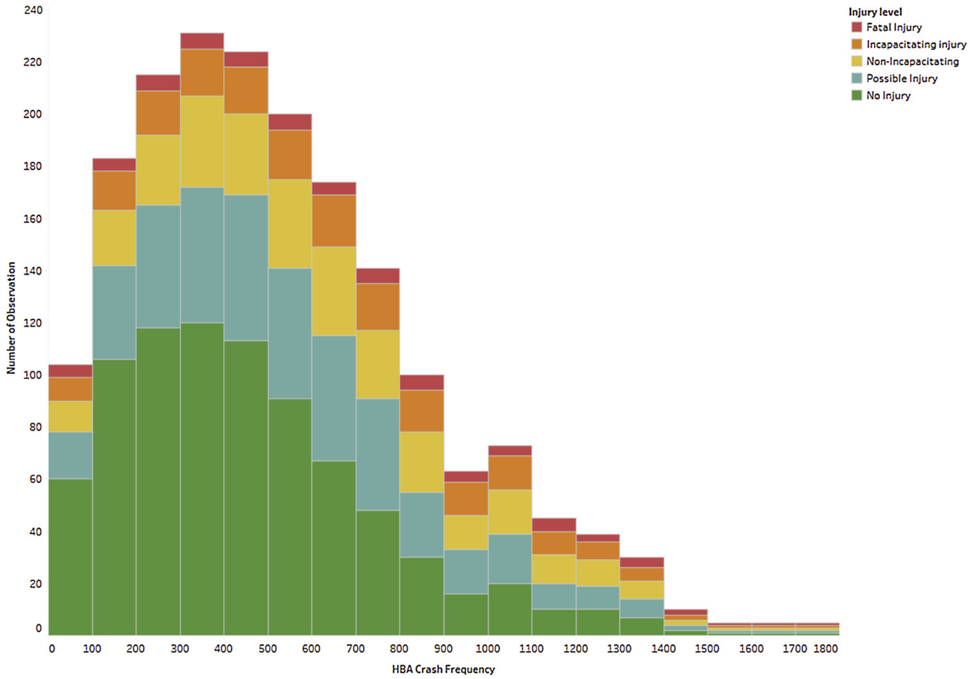

Fig. 1. Histogram of HBA Crash frequency by severity type at the census tract level.

Table 2
National KABCO person-injury unit costs in 2018 dollars.

Injury Type	Crash Cost Per Injury		
	Economic person- Injury Unit Costs	QALY Person-Injury Unit Costs	Comprehensive Crash Cost (2018 Dollars)
No Injury[†]	6,553 (5,717[*])	2,938 (2,563[*])	9,491 (8,280[*])
Possible Injury	24,930 (21,749[*])	57,227 (49,926[*])	82,157 (71,675[*])
Non-Incapacitating Injury	36,800 (32,105[*])	112,302 (97,974[*])	149,102 (130,079[*])
Incapacitating Injury	96,866 (84,507[*])	41,6459 (363,324[*])	513,325 (447,832[*])
Fatal Injury	1,603,502 (1,398,916[*])	8,880,060 (7,747,082[*])	10,483,562 (9,145,998[*])
Unknown			
Vehicle unit cost	6,965 (6,076[*])		6,965 (6,076[*])

† The cost reflects the cases where injury severity was falsely assigned.
* Source: adjusted person-injury cost based on 2010 US Dollar based on Harmon et al. (2018).

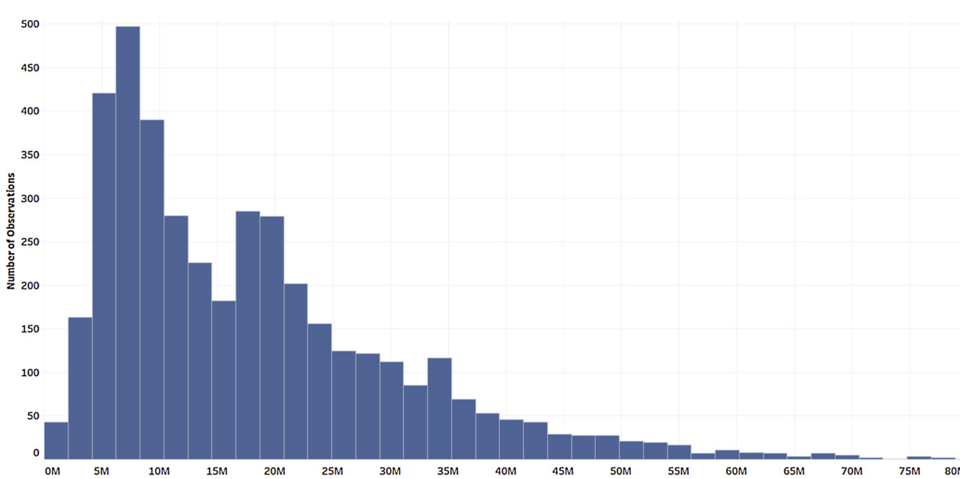

Fig. 2. Histogram of CCCAZ (2018 Dollars).

2.3.1. Poisson model

In the Poisson regression, the probability that comprehensive cost of the crash at zone i equal to n could be written as (Greene, 2003):

$$P(n_i) = \frac{\lambda_i^{n_i}\exp(-\lambda_i)}{n_i!} \tag{2}$$

where λ_i (Poisson parameter) is the expected CCCAZ for zone i in a three year period, $E(n_i)$. In order to fit the regression model, the Poisson parameter, λ_i, can be written in a logarithm format (Greene, 2003):

$$\ln(\lambda_i) = \beta X_i \tag{3}$$

where X_i is the vector of the sociodemographic data element extracted from the census tract and β is a vector of the estimated coefficients. To consider the population variable as an offset variable, we constrained the value of the population's (in logarithm scale) coefficient equal to one (Pérez-Marín and Guillen, 2019).

Furthermore, in cases where the mean and the variance are not equal, applying the Poisson regression might lead to inappropriate results. in order to statistically test the existence of over-dispersion in the Poisson model, the Lagrange multiplier method was performed (Greene, 2003):

$$LL = \left(\frac{\sum_{i=1}^{N}((y_i-\mu_i)^2-y_i)}{2\sum_{i=1}^{N}\mu_i^2}\right)^2 \tag{4}$$

where y_i is the observed CCCAZ at zone i, μ_i is the predicted CCCAZ at zone i, and N is the number of zones.

2.3.2. Geographically Weighted Poisson Regression Model

GWPR can be used to examine whether the association between exogenous variables and CCCAZ substantially varies across space (Fotheringham et al., 2003). The model can be written as:

$$\ln(\lambda_i) = \beta_0(u_i,\ v_i) + \beta_1(u_i,\ v_i)\ln(E_{vi}) + \sum_{k=1}^{K}\beta_k(u_i,\ v_i)x_{ij} \tag{5}$$

where $(u_i,\ v_i)$ denotes the coordinates of zone i. It should be noted that in the GWPR, $\beta_k(u_i,\ v_i)$ is a function of the coordinates of the center of census tract i. The following equation can be used to estimate $\beta_k(u_i,\ v_i)$:

$$\hat{\beta}(u_i,\ v_i) = (X^TW(u_i,\ v_i)X)^{-1}X^TW(u_i,v_i)Y \tag{6}$$

where $\hat{\beta}(u_i,\ v_i)$ is the vector of estimated coefficients at zone i, X is the matrix of exogenous variables, Y is the $n \times 1$ vector of the dependent variable (CCCAZ), and $W(u_i, v_i)$ is $n \times n$ spatial weight matrix:

$$W(u_i,\ v_i) = \begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & w_{in} \end{bmatrix} \tag{7}$$

where w_{ij} is the weight of variable j at location i. In this approach, a regression equation is estimated for each location based on observations at nearby areas. Based on the distance from the regression point each area is weighted (where areas that are closer have a higher weight and vice versa). The (W) matrix can be estimated using an adaptive bi-square kernel, which can be written as:

$$w_{ij}\begin{cases}\left(1-\left({}^{d_{ij}}/_{d_{iN}}\right)^2\right)^2 & if\ d_{ij} < d_{iN} \\ 0\ otherwise\end{cases} \tag{8}$$

where d_{iN} denotes the distance to the N^{th} nearest zone of zone i. Compared to the fixed bandwidth kernels, the adaptive bi-square bandwidth varies based on the data's sparsity. To determine the bandwidth of the adaptive kernel, the corrected Akaike Information Criteria (AICc) (Hurvich et al., 1998) was used. The best model is the one with the lowest AICc score (Fotheringham et al., 2003; Hadayeghi et al., 2010a).

The non-stationarity test was used to evaluate the existence of variation in the estimated coefficients across space (Liu and Khattak, 2017; Arvin et al., 2019). Substantial variations among the estimated coefficients across space exist if the difference between upper and lower quartile ($\delta = \beta_{upper} - \beta_{lower}$) of the estimated coefficients from the GWPR model meets both of the following conditions:

$$\begin{cases}\delta > 1.96{*}SE \\ and \\ 1.96 < \max(|z_i|)\end{cases} \tag{9}$$

where SE is the standard error of the coefficient in the global Poisson model and $|z_i|$ is the absolute value of the significance z-score of the GWPR model at census tract i. Otherwise, the coefficient is considered as the global coefficient, which does not have a substantial spatial variation. In order to estimate the GWPR model, GWR4.0 software which is developed by Nakaya et al. (2012) was used.

2.4. Variable selection

A combination of intuition and stepwise regression modeling was used to select the best subset of the predictors with an exclusion criterion of p-values greater than 0.20. Moreover, Variance Inflation Factors (VIF) was used to control for the multicollinearity in each step. Curious readers could refer to O'brien (2007) for more details about the VIF.

2.5. Measures of goodness of fit

To evaluate and compare the performance of traditional Poisson regression, and GWPR, three statistics were utilized to measure estimation accuracy. First, we used AIC, a lower value of AIC (Bozdogan, 1987) represents the better goodness of fit. We can measure AIC as following:

$$AIC = D + 2k \tag{10}$$

where D denotes the model deviance, and k is the number of parameters. In the GWPR, due to the non-parametric framework of the model, the number of parameters is meaningless. Therefore, an effective number of parameters should be considered, which can be written as (Nakaya et al., 2005)

$$K = trace(S) \tag{11}$$

where S is the hat matrix. In addition to the AIC, we will also use Mean Absolute Error (MAE), Root Mean Square Error (RMSE) to compare the model performances. The lower value of MAE and RMSE indicates a better performace.

3. Results and discussion

Between 2014–2016, 215,481 individuals were injured in traffic crashes and 3082 died on Tennessee's road. The total compressive crash cost of Tennessean in a three-year period was $75.0 Billion (2018 dollars). Table 3 shows the cost of traffic crashes based on road users types. Overall, drivers and passenger have the biggest share of Comprehensive Cost of traffic crashes (94%). Yet, pedestrian and bicyclists crashes average comprehensive cost of traffic crashes was 15.7 and 6.7 times higher than average drivers' comprehensive crash cost, respectively. This value reflects the vulnerability of pedestrians and bicyclist compared to motorized road users.

On average, the comprehensive cost of a traffic crash in Tennessee was $92,374. Table 4 presents the average number crashes by severity at the zonal level. As crash severity increases, the frequency of individuals who suffered decreases. On average, on each census tract 28, 14, and 5 individuals received possible injury, non-incapacitating injury and incapacitating injury in a three-year period, respectively. On

Table 3
Crash cost by road user type (2018 Dollars).

		Total Cost ($ million)			Average Cost ($ 1,000)		
User Type	Number of users	Compreh. Cost	Economic Costs	QALY Costs	Compreh. Cost	Economic Costs	QALY Costs
Driver	995,670	$41,282	$11,617	$29,815	$41.5	$11.7	$29.9
Passenger	352,389	$12,644	$3,828	$8,850	$35.9	$10.9	$25.1
Pedestrian	5,262	$3,431	$579	$2,836	$652.0	$110.1	$539.0
Cyclists	1,387	$390	$74	$314	$280.9	$53.2	$226.4
Grand Total	**1,354,708**	**$57,746**	**$16,098**	**$41,815**	**$42.6**	**$11.9**	**$30.9**

average, 0.7 individuals in each census tract were fatally injured in a traffic crash over the period. The mean comprehensive crash cost at census tract level (CCCAZ) for the study period was $18.2 million (SD = $13.9 million; Median = $15 million). Fig. 3 exhibits a geographical distribution of CCCAZ in Tennessee.

3.1. Model comparison

Table 5 presents the result of the Poisson model and Geographically Weighted Poisson Regression with the adaptive bi-square kernel for predicting CCCAZ. We also tested the sensitivity of the model with fixed Gaussian, fixed bi-square and adaptive Gaussian kernels for model estimation. The adaptive bi-square kernel had the lowest value of the AICc. Therefore, to maintain concision, we only included the adaptive bi-square kernel results. Notably, we found that the results were largely consistent with the adaptive bi-square kernel.

The value of the Lagrange multiplier for the GWPR model and global model are 0.10, 0.13, respectively, which is less than the critical value of Lagrange multiplier ($\chi^2_{(1)} = 3.84$). Therefore, we can conclude the overdispersion is not an issue in this study. Moreover, the comparison of the AIC, AICc, deviance, MAE, and RMSE presented in Table 4 also indicate that the GWPR model is more suitable compared to the global model. The value of the Moran's I (Moran's I = 0.009) indicates that in the GWPR model the residuals are not spatially correlated. Furthermore, the VIF values (average = 1.6, max = 2.9) also indicate that the multicollinearity in not an issue.

3.2. Parameter estimation

Results of the stationary test indicated in the GWPR model, all the covariates in the GWPR model have a local effect. Fig. 4 presents the spatial effect of the estimated parameter on CCCAZ. Only those coefficients that have a significant effect are presented in Fig. 4; the insignificant coefficients are presented with a white color. It is worthy to mention that the estimated coefficients in the traditional fixed models fall into the range of correspondence counterparts in the spatial models (Xu and Huang, 2015), indicating that the estimated parameters in the global models (i.e., fixed models) characterize the average effects of the factors on the dependent variable. The sign of the median of all the variables in the GWPR model (except mean of the proportion of road users who use bus and proportion of road users 16 years old and younger) is consistent with the Poisson model which attest the aforementioned.

Analysis of the local distribution of the estimated coefficients indicates that in most variables (except population density, median family income, and household size) the sign of the variables vary from negative to positive, which is some cases are unexpected. The counterintuitive sign is not an uncommon issue considering the geographically weighted regression models and has been reported in previous studies (Chow et al., 2006; Hadayeghi et al., 2010b; Pirdavani et al., 2013; Xu and Huang, 2015). Some studies attributed this issue to the local multicollinearity (Hadayeghi et al., 2010b); however, this was not the case for this study. Results of the VIF test on areas where signs were counterintuitive did not raise the local multicollinearity issue; VIF values ranged between 1.01 to 3.05. Another issue could be the presence of over-dispersion in the dependent variables (Xu and Huang, 2015). As a result, the Poisson model could produce more significant variables compared to the negative binomial model (Lord and Mannering, 2010). This issue needs to be investigated in future studies.

Means of travel to work have a significant impact on CCCAZ. The proportion of road users who use motorcycle has a positive sign in most of Tennessee; however, the sign of the estimated coefficients has a negative association with CCCAZ in the Knoxville metropolitan area. Moreover, means of estimated coefficients of non-motorized modes (i.e., walk and bicycle) and bus in the local model have a significant negative association with the CCCAZ. Local model estimation indicates that non-motorized modes of transportation has a negative association with CCCAZ in Memphis and Nashville metropolitan areas, whereas, in Knoxville and Chattanooga metropolitan areas this variable has a positive association with CCCAZ. In the Knoxville, Nashville and Chattanooga metropolitan areas, the estimated coefficients for the proportion of road users who use public transit has a negative association with CCCAZ. In the Memphis metropolitan area, this variable has a significant positive association with CCCAZ; however, the magnitude of the estimated coefficients is close to zero. One explanation for the negative sign of both bus and non-motorized road users could be a

Table 4
National KABCO person-injury unit costs and number of crashes in Tennessee 2014-16 based on 2018 Dollar.

Injury Type	Zonal Level Crash Frequency				Total Cost Of Crashes Between 2014-16 (2010 Dollars)**
	Mean	Std. Dev.	Max	Total Number Of Observed Crashes	
No Injury	267.6	179.1	2472	1,099,523	$9,041,888,254
Possible Injury	28.4	20.8	275	116,652	$8,441,878,352
Non-Incapacitating Injury	13.5	9.7	114	55,330	$7,284,213,036
Incapacitating Injury	4.9	3.9	40	20,287	$9,217,372,280
Fatal Injury	0.7	0.9	7	2,725	$25,323,439,334
Unknown	1.7	1.9	27	7,071	$0
PDO Vehicle*	197.1	134.9	19971	810,055	$4,823,938,279
Total					$64,132,729,534

* Unit: Crash cost per vehicle.

** Only Reflects Tennesseans' Part.

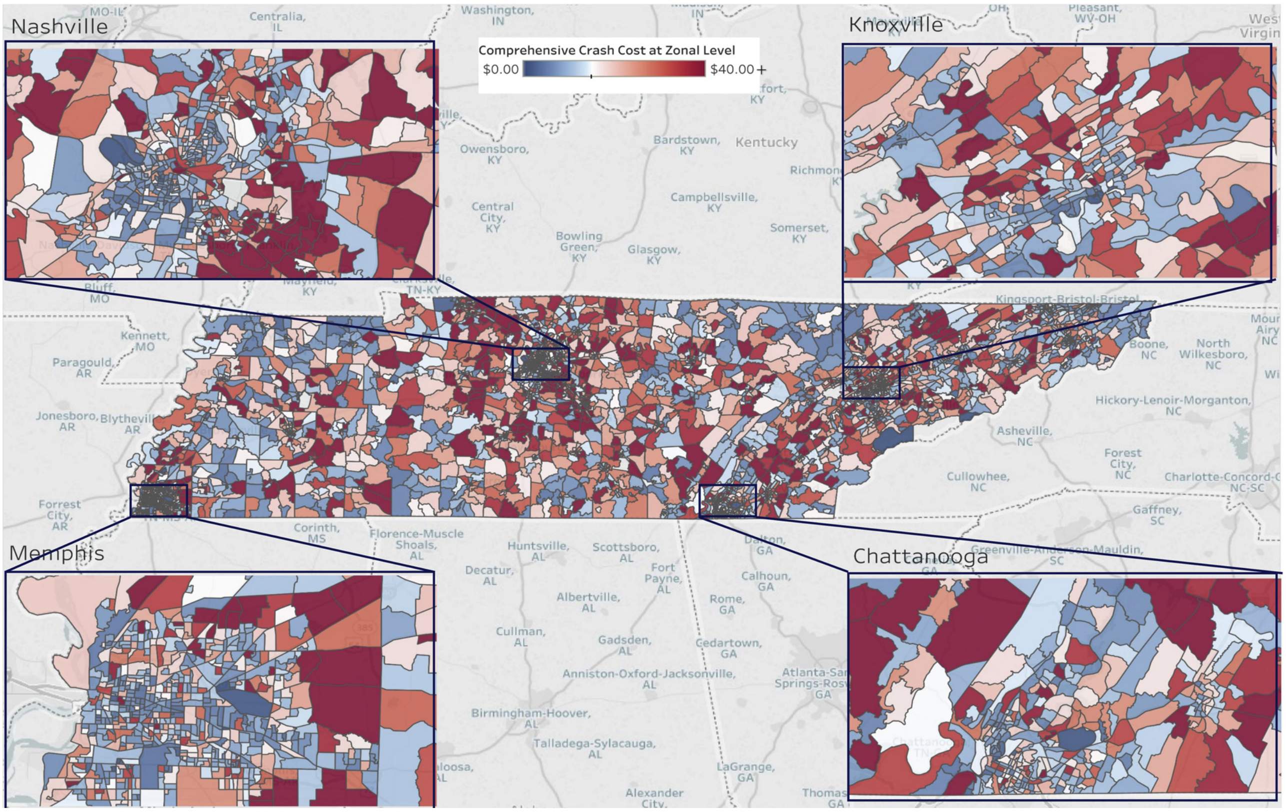

Fig. 3. Geographical distribution of the comprehensive crash cost at the zonal level ($ million).

Table 5
Results of the Poisson model for predicting CCCAZ ($100,000).

Variable	Estimate	Stand. Error	z(Est/SE)	Mean	STD	Min	lwr Quartile	Median	Upr Quartile	Max	Local
Constant	−3.2939	0.0233	−141.23	−3.1656	0.4972	−4.4814	−3.5978	−3.2531	−2.7313	−1.9068	yes
Age Cohorts											
16 Years And Younger	−0.0004	0.0003	−1.55	0.0017	0.0053	−0.0147	−0.0013	0.0007	0.0062	0.0140	yes
60 Years Old and More	−0.0070	0.0002	−35.42	−0.0044	0.0037	−0.0110	−0.0077	−0.0049	−0.0013	0.0031	yes
Proportion of Minor Race	0.0032	0.0001	47.10	0.0033	0.0020	−0.0056	0.0026	0.0034	0.0045	0.0080	yes
Travel to Work Mode											
Motorcycle	0.0442	0.0017	26.20	0.0351	0.0392	−0.0638	0.0094	0.0364	0.0635	0.1522	yes
Bus	0.0019	0.0004	4.28	−0.0033	0.0133	−0.0365	−0.0143	−0.0034	0.0038	0.0481	yes
Non-Motorized Modes	−0.0043	0.0004	−11.06	−0.0056	0.0094	−0.0366	−0.0117	−0.0079	0.0022	0.0292	yes
Average Travel Time To Work	0.0141	0.0002	63.57	0.0101	0.0045	−0.0031	0.0065	0.0093	0.0137	0.0251	yes
Household Size	−0.0223	0.0014	−15.75	−0.0571	0.0519	−0.2342	−0.0953	−0.0316	−0.0180	0.0544	yes
% Educated Over 25 Years Old	0.0079	0.0003	30.90	0.0092	0.0067	−0.0121	0.0077	0.0096	0.0137	0.0198	yes
Median Family Income ($1000)	−0.0044	0.0001	−73.89	−0.0048	0.0013	−0.0080	−0.0056	−0.0051	−0.0039	−0.0008	yes
% Vacant Household	0.0057	0.0001	40.06	0.0089	0.0050	−0.0015	0.0059	0.0084	0.0112	0.0216	yes
Vehicle Per Capita	0.6573	0.0152	43.21	0.5253	0.4432	−0.8005	0.1758	0.3955	0.8985	1.9799	yes
Vehicle Miles Traveled (1000,000)	0.2708	0.0165	16.38	0.0061	0.4124	−1.3157	−0.2697	0.0232	0.3177	0.7178	yes
Population Density (Population Per Square Kilometer)	−9.00E-05	2.0E-06	−39.80	−1.3E-04	6.1E-05	−3.7E-04	−1.6E-04	−1.5E-04	−8.4E-05	−1.9E-05	yes
Metropolitan Indicator	−0.0059	0.0032	−1.85	0.0005	0.1083	−0.2598	−0.0631	0.0285	0.0721	0.2895	yes
AIC	184877			163309							
AICc	184877			163324							
Deviance	184845			162961							
Percent Deviance Explained	0.145			0.246							
Lagrange Multiplier	0.125			0.106							
RMSE	7743.11			6861.31							
Mean Absolute Deviation	66.59			62.63							
R^2 Poisson	0.41			0.49							
Moran's I	0.065			0.009							

reduction in motorized vehicle volume in the surrounding of the residential area, which reduces traffic conflicts and eventually exposure to motorized traffic to other residents of the census tract. On the other hand, the poor design of a multimodal network could adversely impact the safety of non-motorized and public transit users. The difference between signs of the estimated coefficients in the different metropolitan areas need to be investigated in more details in future studies.

Population density also has a negative association CCCAZ; the

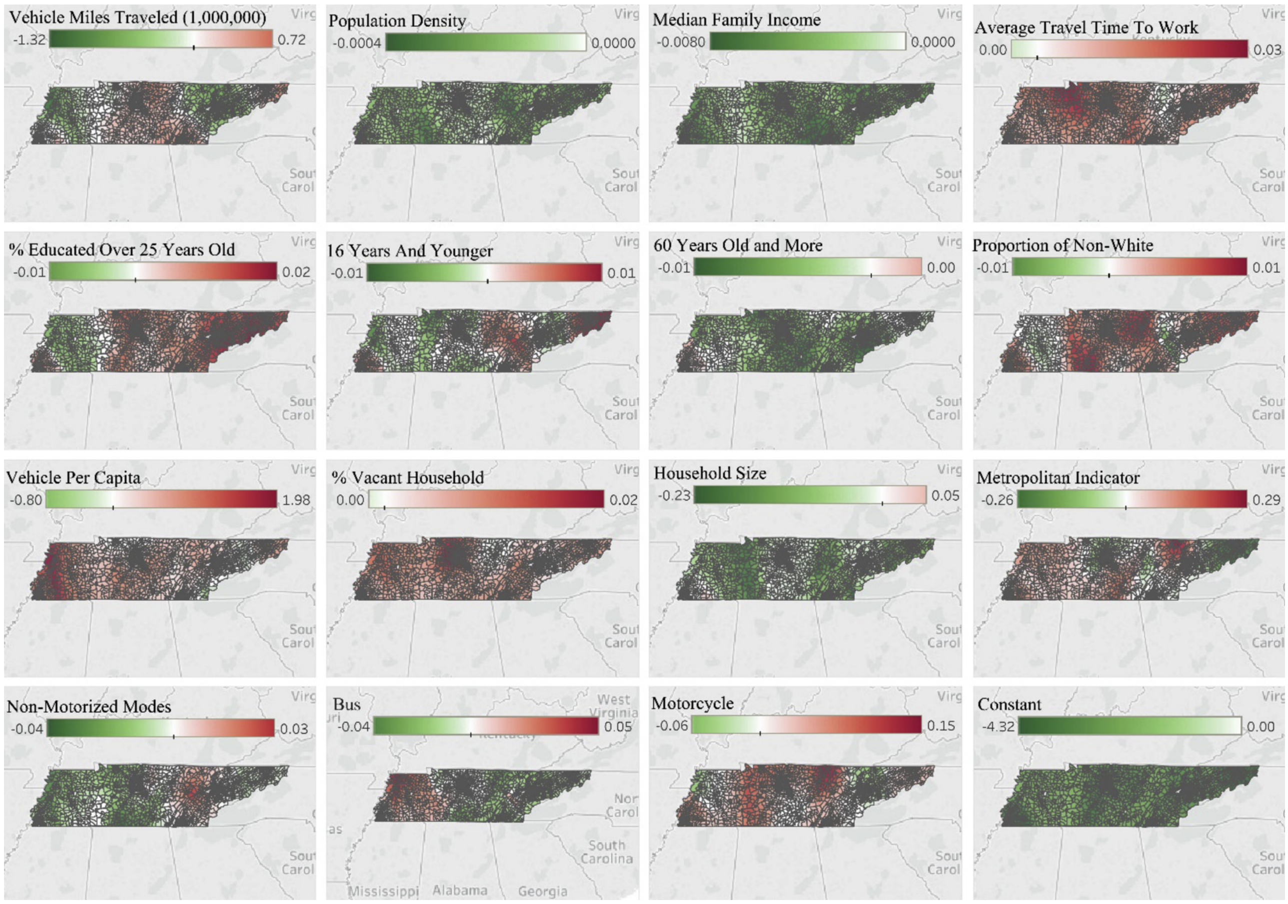

Fig. 4. Local effect of the estimated coefficients in the GWPR model.

model predicts that as density increases the CCCAZ decreases. The sign of the population density is intuitive and is in agreement with the previous studies. For example, population density is associated with higher crash frequency for all traffic crashes, and vulnerable road user crashes (Zwerling et al., 2005; Marshall and Ferenchak, 2017). Crash frequency in high-density areas such as urban areas or metropolitan areas is usually higher than rural or non-metropolitan areas; but, the crash severity is relatively lower (Clark, 2003; Zwerling et al., 2005; Dumbaugh and Rae, 2009). As a result, the overall effect of the population density is constructive and reduces the comprehensive cost of traffic crashes. The Metropolitan indicator also is a proxy for urban areas. Interestingly, Knoxville and Memphis Metropolitan areas coefficients have a negative association with CCCAZ, which is different from the corresponding signs in the Nashville Metropolitan.

Along with previous literature, findings indicate that age cohorts have a significant relationship with the crash outcome (e.g., Wier et al., 2009; Dong et al., 2016; Gomes et al., 2017). The proportion of population 16 years and younger has a varying sign across the state. While the percentage of individuals over 60-years-old has a significant negative association with CCCAZ (except in the Memphis metropolitan area). One may expect the senior population due to their vulnerability will suffer from higher injury severity (Yee et al., 2006); conversely, senior population have a lower trip rate (e.g., exposure to traffic) compared to other groups (Williams and Carsten, 1989; Massie et al., 1995; KRTPO, 2008). As a result, in this study percent of the senior population has a negative effect on CCCAZ (with the exception of the Memphis metropolitan area) compared to other age cohorts.

Considering racial distribution, the estimated model indicates that the population of non-white residents has a significant association with increasing CCCAZ. This finding agrees with previous research (McAndrews et al., 2013; Marshall and Ferenchak, 2017). The percentage of the population educated over 25 years old (except in some rural areas in West-Tennessee), and the percentage of a vacant houses in a census tract also has a significant positive impact on CCCAZ. Although one may expect safer behavior from educated people, it was surprising that this variable's estimated coefficients' sign is counter-intuitive. The negative sign could be attributed to a higher trip rate of this group. This issue needs further analysis. Household size also has a negative association with CCCAZ, which indicates as average household size increases the CCCAZ decreases. One explanation could be the lower per-capita trip rate of individuals in families with bigger household size compared to smaller households in the study area (KRTPO, 2008).

Variables that explain the economic status of each census tract are also associated with the CCCAZ. Median family income is a significant predictor of the CCCAZ; a negative sign of the variable suggests that as family income increases the CCCAZ decreases. The sign agrees with previous studies that show road users with lower income are more prone to traffic crashes (Males, 2009; Lee et al., 2014; WHO, 2015; Lee and Abdel-Aty, 2018). Furthermore, lower-income families' vehicles usually have fewer safety features which may increase the likelihood of severe injuries (Girasek and Taylor, 2010). In contrast, vehicles per capita has a significant impact on the CCCAZ; the positive sign indicates that as this variable increases, the social outcome of traffic crashes gets worse. Vehicles per capita also could be used as a proxy for activity (i.e., amount of vehicle traveled) or lack of multi-modality; we expect a higher trip rate in areas with higher vehicle ownership (e.g., Khattak and Rodriguez, 2005; KRTPO, 2008). These findings are also in agreement with studies that focused on human factors that show that some groups (e.g., lower income, lower education, young road users) are more prone to aberrant behaviors (Özkan et al., 2006; Davey et al., 2007; Elliott et al., 2007; Nordfjærn et al., 2015; Hezaveh et al., 2017, 2018).

VMT and average travel time to work could be interpreted as proxies to exposure. The expectation was to see higher crash cost as these two variables value increase. The models indicate that VMT in the surrounding area of residents has a positive association with CCCAZ. However, considering the local effect and geographical distribution of this variable in Fig. 4, we noticed that in the Knoxville metropolitan area and some rural area, VMT has a negative association with CCCAZ. Analyzing the local coefficient indicate that the multicollinearity was not an issue in the areas with counterintuitive signs. This issue needs to be investigated in future studies.

Average travel time to work, which represent the amount of time that individuals spend in traffic on their work trips also has a significant positive association with CCCAZ. The positive sign in the model indicates that as travel time increases the crash cost at zonal level increases. Travel time could be interpreted as an indicator of accessibility (Marshall and Ferenchak, 2017; Merlin et al., 2019). Increase in accessibility would decrease the travel time (by reducing trip length), VMT and eventually would reduce the comprehensive cost of traffic crashes.

4. Conclusion

In this study, we used the Home-Based Approach crash frequency at the zonal level to calculate the comprehensive crash cost at the zonal level. Unlike traditional road safety analysis that aggregates crashes at the location of the crash, the HBA attributes road safety to the home-address of individuals in a traffic crash. Consequently, we measured the comprehensive cost of traffic crashes at the zonal level by using person-injury crash cost.

Findings indicate that the burden of traffic crashes does not affect the study area in equitable ways. Moreover, over-dispersion is not an issue regarding CCCAZ analysis in this study, hence the Poisson model is suitable for evaluation of the relationship between sociodemographic variables and CCCAZ at the zonal level. Comparison of the performance of the GWPR and Poisson models shows the substantial existence of spatial heterogeneity in the analysis.

This study's findings are broadly in agreement with road safety literature. We find that an increase in population density reduces the societal cost of traffic crashes at the zonal level; increase in residential density, particularly in the urban areas is correlated with the reduction in speeds. On the other hand, an increase in travel time and consequently higher traffic exposure adversely affect the social cost of traffic crashes.

Comprehensive crash cost at the zonal level could be used as a tool for assigning proper countermeasures or interventions to the areas where the disproportionate economic burden of traffic crashes exists or to promote vertical equity in the distribution of road safety countermeasures. Moreover, the HBA could be an advantageous element for developing policies that support groups that are more prone to burden from road traffic crashes.

There are several possible extensions for this study; first we can learn to reduce the injury misclassification error by linking police crash reports to health-oriented databases (Cherry et al., 2018) to get a better understanding of injury outcome and subsequently a more accurate measurement of the injury. Second, the variables that we used in this study was mostly limited to the demographics of residents extracted from the US Census. Adding extra variables regarding transportation network and travel behavior would help us understand the relationship between travel behavior and the comprehensive cost of traffic crashes. Third, based on our findings, we are recommending the use of the home address of the road users to target the areas that are more prone to the burden of traffic crashes by proper education and enforcement countermeasures.

Acknowledgments

The authors would like to thank the Tennessee Department of Safety and Homeland Security for providing the data for this study. This project was supported by the Collaborative Sciences Center for Road Safety, www.roadsafety.unc.edu, a U.S. Department of Transportation National University Transportation Center promoting safety. The study design was reviewed and approved by the University of Tennessee Institutional Review Board.

References

Abdel-Aty, M., Siddiqui, C., Huang, H., Wang, X., 2011. Integrating trip and roadway characteristics to manage safety in traffic analysis zones. Transp. Res. Rec.: J. Transp. Res. Board 2213, 20–28.
Aguero-Valverde, J., 2013. Multivariate spatial models of excess crash frequency at area level: case of Costa Rica. Accid. Anal. Prev. 59, 365–373.
Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accid. Anal. Prev. 38 (3), 618–625.
Ahadi, M.R., Razi-Ardakani, H., 2015. Estimating the cost of road traffic accidents in Iran using human capital method. Int. J. Transp. Eng. 2 (3), 163–178.
Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41 (1), 153–159.
Arvin, R., Kamrani, M., Khattak, A.J., 2019. How instantaneous driving behavior contributes to crashes at intersections: extracting useful information from connected vehicle message data. Accid. Anal. Prev. 127, 118–133.
Baker, S.P., Braver, E.R., Chen, L.-H., Pantula, J.F., Massie, D., 1998. Motor vehicle occupant deaths among hispanic and black children and teenagers. Arch. Pediatr. Adolesc. Med. 152 (12), 1209–1212.
Blincoe, L., Miller, T., Zaloshnja, E., Lawrence, B.A., 2015. The Economic and Societal Impact of Motor Vehicle Crashes, 2010. (revised)(report No. Dot Hs 812 013). National Highway Traffic Safety Administration., Washington, DC.
Bozdogan, H., 1987. Model selection and akaike's information criterion (aic): the general theory and its analytical extensions. Psychometrika 52 (3), 345–370.
Braver, E.R., 2003. Race, hispanic origin, and socioeconomic status in relation to motor vehicle occupant death rates and risk factors among adults. Accid. Anal. Prev. 35 (3), 295–309.
Bureau of Economic Analysis, 2018. Per Capita Real Gdp by State (chained 2009 Dollars).
Cai, Q., Abdel-Aty, M., Lee, J., Eluru, N., 2017. Comparative analysis of zonal systems for macro-level crash modeling. J. Safety Res. 61, 157–166.
Cameron, A.C., Trivedi, P.K., 1986. Econometric models based on count data. Comparisons and applications of some estimators and tests. J. Appl. Econom. 1 (1), 29–53.
Campos-Outcalt, D., Bay, C., Dellapena, A., Cota, M.K., 2003. Motor vehicle crash fatalities by race/ethnicity in Arizona, 1990–96. Inj. Prev. 9 (3), 251–256.
Cheng, W., Gill, G.S., Ensch, J.L., Kwong, J., Jia, X., 2018. Multimodal crash frequency modeling: multivariate space-time models with alternate spatiotemporal interactions. Accid. Anal. Prev. 113, 159–170.
Cherry, C., Hezaveh, A.M., Noltenius, M., Khattak, A., Merlin, L., Dumbaugh, E., Ragland, D., Sandt, L., 2018. Completing the Picture of Traffic Injuries: Understanding Data Needs and Opportunities for Road Safety.
Chiou, Y.-C., Jou, R.-C., Yang, C.-H., 2015. Factors affecting public transportation usage rate: geographically weighted regression. Transp. Res. Part A Policy Pract. 78, 161–177.
Chow, L.-F., Zhao, F., Liu, X., Li, M.-T., Ubaka, I., 2006. Transit ridership model based on geographically weighted regression. Transp. Res. Rec.: J. Transp. Res. Board 1972, 105–114.
Clark, D.E., 2003. Effect of population density on mortality after motor vehicle collisions. Accid. Anal. Prev. 35 (6), 965–971.
Davey, J., Wishart, D., Freeman, J., Watson, B., 2007. An application of the driver behaviour questionnaire in an australian organisational fleet setting. Transp. Res. Part F Traffic Psychol. Behav. 10 (1), 11–21.
Dong, N., Huang, H., Lee, J., Gao, M., Abdel-Aty, M., 2016. Macroscopic hotspots identification: a Bayesian spatio-temporal interaction approach. Accid. Anal. Prev. 92, 256–264.
Dumbaugh, E., Rae, R., 2009. Safe urban form: revisiting the relationship between community design and traffic safety. J. Am. Plan. Assoc. 75 (3), 309–329.
Elliott, M.A., Baughan, C.J., Sexton, B.F., 2007. Errors and violations in relation to motorcyclists' crash risk. Accid. Anal. Prev. 39 (3), 491–499.
Elvik, R., Vaa, T., Erke, A., Sorensen, M., 2009. . The Handbook of Road Safety Measures. Emerald Group Publishing.
Fhwa, 2017. Kabco injury classification scale and injury Classification Scale and definations.
Fotheringham, A., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression-The Analysis of Spatially Varying Relationships. John Wiley & Sons, Chichester, UK.
Fotheringham, A.S., Brunsdon, C., Charlton, M., 2003. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. John Wiley & Sons.
García-Altés, A., Pérez, K., 2007. The economic cost of road traffic crashes in an urban setting. Inj. Prev. 13 (1), 65–68.
Girasek, D.C., Taylor, B., 2010. An exploratory study of the relationship between socioeconomic status and motor vehicle safety features. Traffic Inj. Prev. 11 (2), 151–155.

Gomes, M.J.T.L., Cunto, F., Da Silva, A.R., 2017. Geographically weighted negative binomial regression applied to zonal level safety performance models. Accid. Anal. Prev. 106, 254–261.
Gourieroux, C., Visser, M., 1997. A count data model with unobserved heterogeneity. J. Econom. 79 (2), 247–268.
Greene, W.H., 2003. Econometric Analysis. Pearson Education India.
Hadayeghi, A., Shalaby, A., Persaud, B., 2010a. Development of planning-level transportation safety models using full Bayesian semiparametric additive techniques. J. Transp. Saf. Secur. 2 (1), 45–68.
Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010b. Development of planning level transportation safety tools using geographically weighted poisson regression. Accid. Anal. Prev. 42 (2), 676–688.
Harmon, T., Bahar, G., Gross, F., 2018. Crash Costs for Highway Safety Analysis. USDOT FHWA Office, Washington, DC.
Harper, J.S., Marine, W.M., Garrett, C.J., Lezotte, D., Lowenstein, S.R., 2000. Motor vehicle crash fatalities: a comparison of hispanic and non-hispanic motorists in colorado. Ann. Emerg. Med. 36 (6), 589–596.
Hauer, E., 1997. Observational Before-After Studies in Road Safety–Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety.
Haynes, R., Jones, A., Kennedy, V., Harvey, I., Jewell, T., 2007. District variations in road curvature in England and Wales and their association with road-traffic crashes. Environ. Plan. A 39 (5), 1222–1237.
Hezaveh, A.M., Cherry, C.R., 2019a. Individuals' likelihood of involvement in traffic crashes: introduction of home-based approach –a complementary definition of road safety. Proceedings of the Transportation Research Board 98th Annual Meeting.
Hezaveh, A.M., Cherry, C.R., 2019b. Neighborhood-level factors affecting seat belt use. Accid. Anal. Prev. 122, 153–161.
Hezaveh, A.M., Nordfjærn, T., Mamdoohi, A.R., Şimşekoğlu, Ö., 2018. Predictors of self-reported crashes among iranian drivers: exploratory analysis of an extended driver behavior questionnaire. PROMET-Traffic Transp. 30 (1), 35–43.
Hezaveh, A.M., Zavareh, M.F., Cherry, C.R., Nordfjærn, T., 2017. Errors and violations in relation to bicyclists' crash risks: development of the bicycle rider behavior questionnaire (brbq). J. Transp. Health 8, 289–298.
Huang, H., Abdel-Aty, M., 2010. Multilevel data and Bayesian analysis in traffic safety. Accid. Anal. Prev. 42 (6), 1556–1565.
Huang, H., Song, B., Xu, P., Zeng, Q., Lee, J., Abdel-Aty, M., 2016. Macro and micro models for zonal crash prediction with application in hot zones identification. J. Transp. Geogr. 54, 248–256.
Hurvich, C.M., Simonoff, J.S., Tsai, C.L., 1998. Smoothing parameter selection in nonparametric regression using an improved akaike information criterion. J. R. Stat. Soc. Series B Stat. Methodol. 60 (2), 271–293.
Jacobs, G., Aeron-Thomas, A., Astrop, A., 2000. Estimating Global Road Fatalities.
Khattak, A.J., Rodriguez, D., 2005. Travel behavior in neo-traditional neighborhood developments: a case study in USA. Transp. Res. Part A Policy Pract. 39 (6), 481–500.
Koornstra, M., Lynam, D., Nilsson, G., 2002. Sunflower: A Comparative Study of the Development of Road. SWOV Institute for Road Safety Research, Leidschendam.
Krtpo, 2008. 2008 East Tennessee Household Travel Survey Final Report. Knoxville Regional Transportation Planning Organization. Knoxville Regional Transportation Planning Organization, 206 Wild Basin Rd., Suite A-300 Austin, Texas 78746.
Lee, J., Abdel-Aty, M., 2018. Macro-level analysis of bicycle safety: focusing on the characteristics of both crash location and residence. Int. J. Sustain. Transp. 12 (8), 553–560.
Lee, J., Abdel-Aty, M., Choi, K., 2014. Analysis of residence characteristics of at-fault drivers in traffic crashes. Saf. Sci. 68, 6–13.
Lee, J., Abdel-Aty, M., Jiang, X., 2015. Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. Accid. Anal. Prev. 78, 146–154.
Lesage, J., Pace, R.K., 2009. Introduction to Spatial Econometrics. Chapman and Hall/CRC.
Levine, N., Kim, K.E., Nitz, L.H., 1995. Spatial analysis of honolulu motor vehicle crashes: ii. Zonal generators. Accid. Anal. Prev. 27 (5), 675–685.
Liu, J., Khattak, A.J., 2017. Gate-violation behavior at highway-rail grade crossings and the consequences: using geo-spatial modeling integrated with path analysis. Accid. Anal. Prev. 109, 99–112.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp. Res. Part A Policy Pract. 44 (5), 291–305.
Males, M.A., 2009. Poverty as a determinant of young drivers' fatal crash risks. J. Safety Res. 40 (6), 443–448.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.
Marshall, W.E., Ferenchak, N.N., 2017. Assessing equity and urban/rural road safety disparities in the US. J. Urban. Int. Res. Placemaking Urban Sustain. 10 (4), 422–441.
Massie, D.L., Campbell, K.L., Williams, A.F., 1995. Traffic accident involvement rates by driver age and gender. Accid. Anal. Prev. 27 (1), 73–87.
Mayrose, J., Jehle, D.V., 2002. An analysis of race and demographic factors among motor vehicle fatalities. J. Trauma Acute Care Surg. 52 (4), 752–755.
Mcandrews, C., Beyer, K., Guse, C.E., Layde, P., 2013. Revisiting exposure: fatal and non-fatal traffic injury risk across different populations of travelers in wisconsin, 2001–2009. Accid. Anal. Prev. 60, 103–112.
Merlin, L.A., Cherry, C.R., Hezaveh, A.M., Dumbaugh, E., 2019. Residential accessibility's relationships with crash rates per capita. Year. Proceedings of the Transportation Research Board 98th Annual Meeting.
Miaou, S.-P., Song, J.J., Mallick, B.K., 2003. Roadway traffic crash mapping: a space-time modeling approach. J. Transp. Stat. 6 (1).
Mohan, D., 2002. Social cost of road traffic crashes in india. Proceedings of the Proceedings First Safe Community Conference on Cost of Injury 33–38.
Naderan, A., Shahi, J., 2010. Aggregate crash prediction models: introducing crash generation concept. Accid. Anal. Prev. 42 (1), 339–346.
Nakaya, T., Charlton, M., Lewis, P., Fortheringham, S., Brunsdon, C., 2012. Windows Application for Geographically Weighted Regression Modeling. Ritsumeikan University, Kyoto, Japan.
Nakaya, T., Fotheringham, A.S., Brunsdon, C., Charlton, M., 2005. Geographically weighted poisson regression for disease association mapping. Stat. Med. 24 (17), 2695–2717.
Nhtsa, 2013. Traffic Safety Facts 2013. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC.
Nordfjærn, T., Hezaveh, A.M., Mamdoohi, A.R., 2015. An analysis of reported driver behaviour in samples of domestic and expatriate Iranians. J. Risk Res. 18 (5), 566–580.
O'brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41 (5), 673–690.
Özkan, T., Lajunen, T., Chliaoutakis, J.E., Parker, D., Summala, H., 2006. Cross-cultural differences in driving skills: a comparison of six countries. Accid. Anal. Prev. 38 (5), 1011–1018.
Pérez-Marín, A.M., Guillen, M., 2019. Semi-autonomous vehicles: usage-based data evidences of what could be expected from eliminating speed limit violations. Accid. Anal. Prev. 123, 99–106.
Pirdavani, A., Bellemans, T., Brijs, T., Kochan, B., Wets, G., 2014. Assessing the road safety impacts of a teleworking policy by means of geographically weighted regression method. J. Transp. Geogr. 39, 96–110.
Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., 2012. Developing zonal crash prediction models with a focus on application of different exposure measures.
Pirdavani, A., Brijs, T., Bellemans, T., Wets, G., 2013. Spatial Analysis of Fatal and Injury Crashes in Flanders, Belgium: Application of Geographically Weighted Regression Technique.
Pulugurtha, S.S., Duddu, V.R., Kotagiri, Y., 2013. Traffic analysis zone level crash estimation models based on land use characteristics. Accid. Anal. Prev. 50, 678–687.
Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accid. Anal. Prev. 40 (4), 1486–1497.
Saha, D., Alluri, P., Gan, A., Wu, W., 2018. Spatial analysis of macro-level bicycle crashes using the class of conditional autoregressive models. Accid. Anal. Prev. 118, 166–177.
Schiff, M., Becker, T., 1996. Trends in motor vehicle traffic fatalities among hispanics, non-hispanic whites and american indians in new mexico, 1958–1990. Ethn. Health 1 (3), 283–291.
Ukkusuri, S., Hasan, S., Aziz, H., 2011. Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. Transp. Res. Rec.: J. Transp. Res. Board 2237, 98–106.
Wang, Y., Kockelman, K.M., 2013. A poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. Accid. Anal. Prev. 60, 71–84.
Washington, S., Metarko, J., Fomunung, I., Ross, R., Julian, F., Moran, E., 1999. An inter-regional comparison: fatal crashes in the southeastern and non-southeastern United States: preliminary findings. Accid. Anal. Prev. 31 (1), 135–146.
Wegman, F., Oppe, S., 2010. Benchmarking road safety performances of countries. Saf. Sci. 48 (9), 1203–1211.
WHO, 2015. Global Status Report on Road Safety 2015. World Health Organization. Violence Injury Prevention World Health Organization.
Wier, M., Weintraub, J., Humphreys, E.H., Seto, E., Bhatia, R., 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accid. Anal. Prev. 41 (1), 137–145.
Williams, A.F., Carsten, O., 1989. Driver age and crash involvement. Am. J. Public Health 79 (3), 326–327.
Xu, P., Huang, H., 2015. Modeling crash spatial heterogeneity: random parameter versus geographically weighting. Accid. Anal. Prev. 75, 16–25.
Xu, P., Huang, H., Dong, N., Wong, S., 2017. Revisiting crash spatial heterogeneity: a Bayesian spatially varying coefficients approach. Accid. Anal. Prev. 98, 330–337.
Yee, W., Cameron, P.A., Bailey, M., 2006. Road traffic injuries in the elderly. Emergy Med. J. 23 (1), 42–46.
Zwerling, C., Peek-Asa, C., Whitten, P., Choi, S.-W., Sprince, N., Jones, M.P., 2005. Fatal motor vehicle crashes in rural and urban areas: decomposing rates into contributing factors. Inj. Prev. 11 (1), 24–28.