

SPATIAL ASSESSMENT OF ROAD TRAFFIC INJURIES IN THE GREATER TORONTO AREA (GTA): SPATIAL ANALYSIS FRAMEWORK

*Eric Vaz*¹

*Sina Tehranchi*²

*Michael Cusimano*³

ABSTRACT

This research presents a Geographic Information Systems (GIS) and spatial analysis approach based on the global spatial autocorrelation of road traffic injuries for identifying spatial patterns. A locational spatial autocorrelation was also used for identifying traffic injury at spatial level. Data for this research study were acquired from Canadian Institute for Health Information (CIHI) based on 2004 and 2011. Moran's I statistics were used to examine spatial patterns of road traffic injuries in the Greater Toronto Area (GTA). An assessment of Getis-Ord G_i^* statistic was followed as to identify hot spots and cold spots within the study area. The results revealed that Peel and Durham have the highest collision rate for other motor vehicle with motor vehicle. Geographic weighted regression (GWR) technique was conducted to test the relationships between the dependent variable, number of road traffic injury incidents and independent variables such as number of seniors, low education, unemployed, vulnerable groups, people smoking and drinking, urban density and average median income. The result of this model suggested that number of seniors and low education have a very strong correlation with the number of road traffic injury incidents.

Keywords: Spatial Analysis, Geographic Information Systems, Injury Analytics, Traffic Injuries, Geographically Weighted Regression.

JEL Classification: I10, I18, C31

1. INTRODUCTION

Road traffic accidents are the leading and the most frequent cause of death and injury worldwide (Morency *et al.*, 2012). Injuries sustained in road traffic accidents are a major burden on healthcare system in terms of emergency treatment, chronic care, and rehabilitation (Ramage-Morin, 2008). In 2009, road traffic accidents accounted for about 2300 fatalities and 11450 serious injuries which required hospitalization (Transport Canada, 2011). According to World Health Organization (WHO) report in 2005, currently road traffic injuries are the leading cause of death and injuries, the 10th leading cause of all deaths and 9th leading contributor to the burden of disease worldwide. It has been predicted that by 2030, road traffic injuries will become the third largest contributor to the burden of disease worldwide (Chisholm *et al.*, 2012).

¹ Department of Geography and Environmental Studies, Ryerson University, Toronto, ON, Canada. Research Centre for Spatial and Organizational Dynamics, University of the Algarve, Faro Portugal. evaz@ryerson.ca

² Department of Geography and Environmental Studies, Ryerson University, Toronto, ON, Canada. sina.tehranchi@ryerson.ca

³ Department of Neurosurgery University of Toronto and St. Michael's Hospital, Toronto, ON, Canada. cusimanom@smh.ca

Moreover, WHO also reports that majority of road traffic injuries are among “*pedestrians, cyclists, and motorcyclist who are the most vulnerable road users*”. For instance, according to Transport Canada (2011), in total about 25% of all road traffic fatalities are among vulnerable road users, of which 13% of fatalities are among pedestrian alone. According to Schuurman and others (2009) pedestrian safety is often overlooked, primarily because smooth movement of motorized vehicle remains a priority for engineers in road design (Schuurman *et al.*, 2009).

Consequently, one of the major contributing causes of road traffic accidents is urban sprawl (Ewing, Schieber & Zegeer, 2003). The reason being is because since cities have spread outward into rural areas (de Noronha & Vaz, 2015), residents of sprawled cities tend to commute by a vehicle to their job in the city. This as a result increase the use of motor vehicles, which ultimately leads to an increase in the road traffic accidents.

Road traffic accident, is a global occurrence, which results in increased injury and mortality rates (World Health Organization, 2005). Increase in the number of motorized vehicle over the past few decades has been accompanied by an increasing number of road traffic injuries (Ameratunga, Hajar & Norton, 2006). According to Ashokbhai (2014) road traffic injuries have become a threat to the public health in many countries as it contributes to “*poverty by causing death, injuries, disabilities, grief and loss of productivity*”. The reason being is because the victim may take time off from work which would result in loss of income due to their injuries.

In general, it is important to study road traffic injuries as it is a major problem and it is not well recognized in many countries. For the purpose of this study, road traffic injury data from 2004 and 2011 was acquired from the Canadian Institute for Health Information (CIHI). The purpose of this study is as follow,

- To investigate and determine the spatial pattern of road traffic injuries throughout the study area.
- To identify areas of hotspots with high frequency of road traffic injury
- To develop a spatial explanatory model for underlying causes of road traffic injury.
- And lastly, to understand and determine contributing factors behind the spatial pattern using geographic weighted regression (GWR).

2. INTRODUCTION

2.1 Socioeconomic Factors

Road traffic accidents have been recognized as one of the major causes of disability and death among people every day (Scheidt *et al.*, 1995). There have been several literatures in the field of road traffic injury, however majority of research studies have been focused on the vulnerable road users, as they consider to have the highest number of injury and mortality (Pratte, 1998). Studies have also focused on the factors that contribute to the increase in the number of road traffic injuries and mortality.

According to the study by Link and Phelan (1995) and Camilloni and others (2013), socioeconomic status (SES) can play a major role as a fundamental determinants of injuries and deaths. For instance, study by Camilloni and others (2013) suggested that people with lower income are more likely to get involve in a road traffic injury, primarily due to their lack of access to new and safer mean of transportation. Similarly, low socioeconomic status is highly correlated with each of the “*14 major causes of injury and death in the International Classification of Diseases (ICD)*” (Link & Phelan, 1995). Study by Cubbin, LeClere and Smith (2000) discovered that risks for motor vehicle related fatalities can vary by occupation. People with lower status job are more prone to motor vehicle related fatalities (Cubbin, LeClere & Smith, 2000). The high risk of motor vehicle related fatalities among people with

lower job status, may possibly reflect the fact that due to lower income, they are more likely to live in outer city-core (rural areas), which would result in greater distance commute by car, and thus be exposed to higher risk of road traffic injury. Similarly, it is possible that due to lower income, they are less likely to be able to afford newer and safer vehicle, which would put them at a great risk in the event of an accident.

In addition to socioeconomic factors, some studies have suggested that both urban and rural environment can have an impact on the road traffic injuries as well. Afukarr, Antwi and Ofosu-Amaah (2003) explain that majority of road traffic injuries occur on rural areas, where roads are deteriorated and driving rules and regulations are frequently tend to be ignored by drivers (Jacob & Sayer, 1983). As a result the chances of getting into an accidents are much higher than in urban areas. On the contrary, other studies have suggested otherwise. For instance, study by Al-Omari and Obaidat (2013) showed that the rates of casualties occurring in urban areas among vulnerable road users are relatively higher compare to rural areas. In fact, pedestrians constitute a higher proportion of road traffic injuries in urban area (Jacob & Sayer, 1983) among vulnerable road users, followed by cyclist (WHO, 2005). Urban areas tend to have a higher risk of road traffic injuries due to several reasons. Concentration of vehicles and population in city center which causes congestion and heavy traffic, lack of attention to traffic rules by both drivers and pedestrians (jaywalking), as wells as lack of segregated pedestrian and bicycle facility (bike lane) on the road network (Pratte, 1998) are the main reason of high rate of road traffic injury in urban setting.

2.2 Global and Local Spatial Analysis

Road traffic accidents causing injury have increased throughout the years (Odero, Garner & Zwi, 1997), as a result it has become the main motivation behind the analysis of motor vehicle accident pattern. Due to the importance of road traffic accidents, number of studies have focused their attention to road traffic accident hotspot identification (Flahaut *et al.*, 2003; Geurts & Wets, 2003; Manepalli, Bham & Kandada, 2011) in order to reduce and prevent the increase in road traffic injuries. Martin, Crandall and Pilkey (2000) explained that road traffic studies have focused on identifying past, present and future injury patterns from motor vehicle accidents, in order to help researchers establish injury prevention priorities. However, to address these priorities, spatial analysis could be joined with road traffic injury data for the analysis.

The growth in spatially references datasets, advance visualization, rapid data retrieval and ability to manipulate data in geographic information system (GIS) (Vaz & Khaper, 2016), have allowed new techniques to thrive for spatial data analysis (Anselin, 1995). Spatial autocorrelation is one of those methods that has been widely used to evaluate the interconnectedness of values in a geographic area to those nearby (Jackson & Waller, 2005). Spatial autocorrelation is important because the outcome in one area can greatly influence the outcome of its neighboring areas (Anselin, 1995). According to the Tobler's first law of geography, "*everything is related to everything else, but near things are more related than distance things*" (Tobler, 1970). Number of studies in the field of road traffic injury (LaScala, Gerber & Gruenewald, 1999; Aguero-Valverde & Jovanis, 2006) have implemented this method to examine the distribution of incidents in space, and conduct analysis accordingly.

Study of hotspots which also is referred to as "*black spots or high risk locations*" (Geurts & Wets, 2003) are sites with significantly higher frequency of injuries or accidents compare to neighboring locations (Hakkert & Mahalel, 1978). However, to reduce the number of road traffic injuries, a feasible solution is to identify hotspots (Geurts & Wets, 2003; Montella, 2010). Some of the commonly used hotspot identification methods are local spatial autocorrelation, known as Getis-Ord G_i^* Statistics (Ord & Getis, 1995), and Kernel

density estimator (Flahaut *et al.*, 2003). Among the existing methods, Getis-Ord G_i^* Statistics is preferred to be used for hotspot identification (Manepalli, Bham & Kandada, 2011). Though, both Getis-Ord G_i^* Statistics and Kernel density estimator have different conceptualization, both produce similar results under specific conditions of the selected parameters (Manepalli, Bham & Kandada, 2011).

Number of researchers have made the link between road traffic injury and spatial analysis using hot spot analysis. Erdogan (2009) conducted a study in Turkey aiming at road traffic accidents using global and local spatial autocorrelation analysis, to examine whether provinces in Turkey with high rate of road traffic accident are clustered or are located close to each other randomly. Additionally, G_i^* Statistics and Z-score of statistics were used to identify hotspots that may not be visible with global spatial autocorrelation (Getis and Ord, 1992). More recently, Truong and Somenahalli (2011) used Moran's I statistics to examine spatial patterns of pedestrian-vehicle crash data. Getis-Ord G_i^* Statistics was also used to identify the clustering of low and high values, and also to generate a pedestrian vehicle crash hotspot map. It is apparent that global and local spatial autocorrelation is well known method and is used by many researchers in different fields, especially in the field of road traffic accidents.

2.3 Geographic Weighted Regression (GWR)

One of the main objectives in spatial analysis is the identification of relationships that exist among different variables (Helbich *et al.*, 2013). In order to analyze the relationship among variables, it is important to conduct regression analysis. Regression analysis is a popular method of analysis among researchers when it comes to estimating effects of explanatory variables on the dependent variable (Charlton, Fotheringham & Brunson, 2006). Moreover, regression analysis allows researchers to study and explore spatial relationships, to understand the influencing factors that may affect the spatial patterns, and also to predict outcomes based on that understanding (Moutinho & Hutcheson, 2011: 225). For that matter, several techniques have already been implemented. For instance, ordinary least square (OLS) which is one of the most basic and most commonly used techniques, is a generalized linear modelling technique which is used to estimate regression models at a global level (Moutinho & Hutcheson, 2011: 224). Geographic weighted regression (GWR), however, is another technique to estimate regression with spatially varying relationships at a local level (Brunson, Fotheringham & Charlton, 1996). According to Lu, Charlton and Fotheringham (2011), in the last few years, there have been an increasing interest in the local forms of spatial analysis methods that produces results locally rather than globally. As a result, GWR has been identified to be the most popular, reliable and accurate method for exploring the spatial relationships among variables (Zhang *et al.*, 2004).

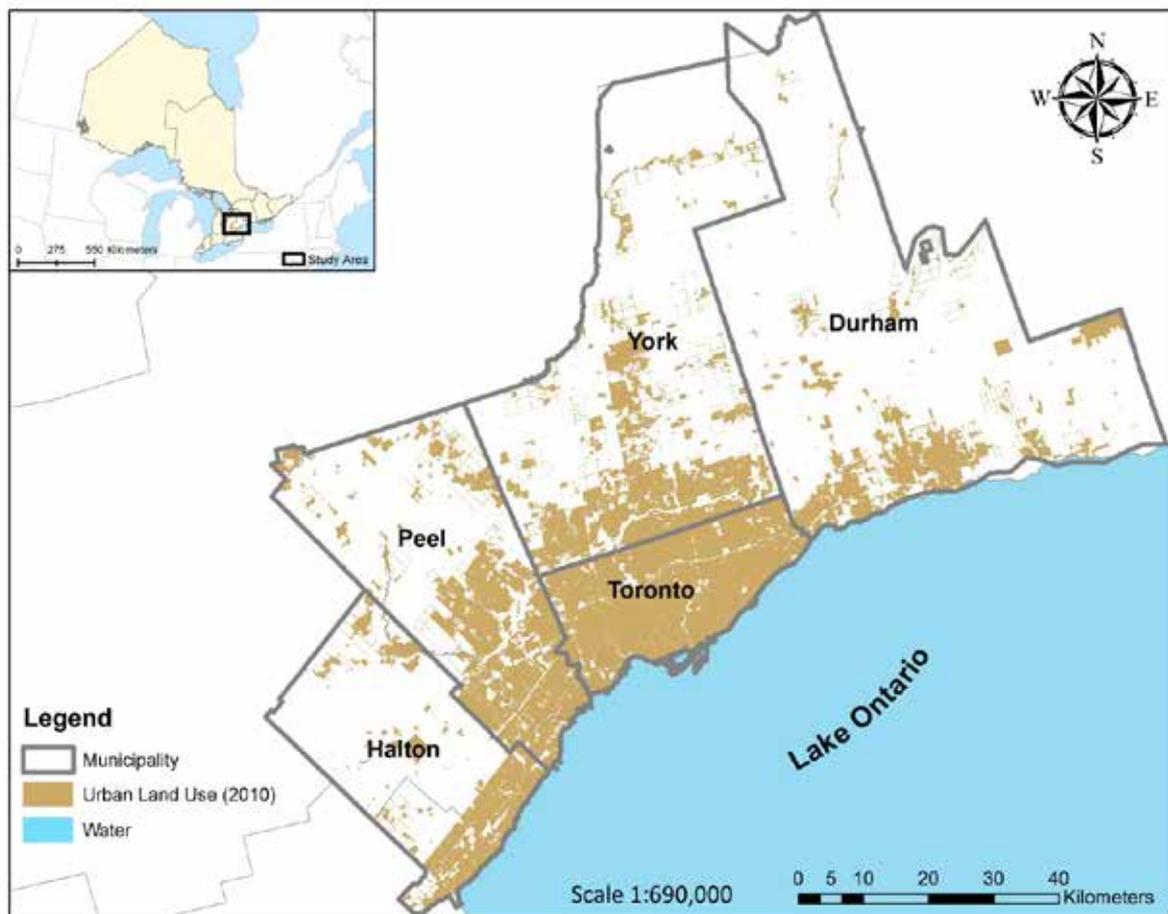
In addition, geographic weighted regression technique has been widely used in different research fields such as economics (Huang, Wu & Barry, 2010), ecology and environment (Zhang *et al.*, 2004), urban analysis (Gao & Li, 2011) and health analysis (Nakaya, 2005; Vaz, Cusimano & Hernandez, 2015). According to Zhang and others (2004), one of the advantages of using GWR is that it significantly improves the model fitting over ordinary least square (OLS). Also, since GWR produces results at a local level, the result are more accurate and reliable compare to global regressions.

3. STUDY AREA

The Greater Toronto Area (GTA), which is the focus for this study, is comprised of the City of Toronto and the four surrounding regional municipalities of Durham, Halton, Peel and

York, which are located along the northern shore of Lake Ontario (Figure 1). The GTA has a population of over six million people, of which more than half of that population live in the surrounding regional municipalities (GTA Alliance, 2011). The GTA is considered to be the fourth largest populated area in North America (GreaterToronto, 2013). Furthermore, the GTA is bounded by Lake Simcoe to the north, Lake Ontario to the south, Kawartha Lake to the east and Niagara Escarpment to the west. The region covers an area of over 7000 km² with the newly amalgamated City of Toronto at its urban core. This urban core includes the former cities of Toronto, North York, Etobicoke, Scarborough, and the District of East York. The GTA is the largest urban area in Canada and one of the largest in North America. The GTA continues to experience rapid growth with the region adding nearly 100,000 new residents each year. This growth is expected to continue over the coming decades with the GTA becoming home to nearly 9.4 million people by 2041 (MetropolisIQ, 2012).

Figure 1. Urban land use cover in the Greater Toronto Area (GTA), located in southern Ontario, Canada



Source: Own Elaboration

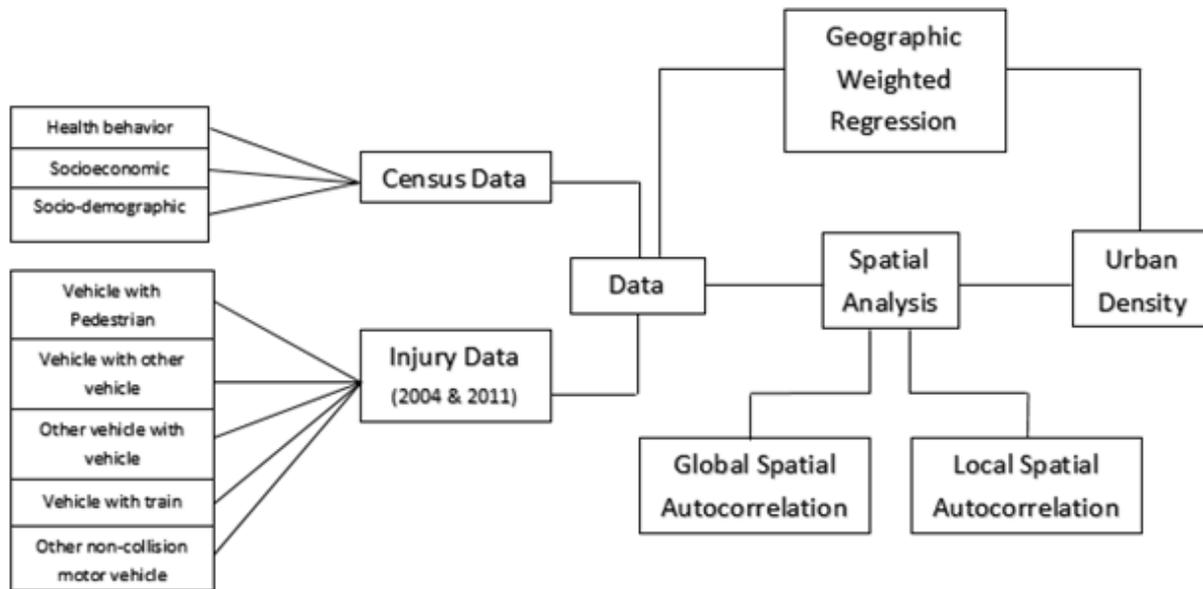
4. METHODOLOGY

4.1 Data

In terms of the data analysis, number of census variables and injury data types were aggregated to conduct the analysis (Figure 2). These data variables were then used to perform global and local spatial autocorrelation and perform geographic weighted regression along with

urban density data to identify factors that may contribute to the increase of road traffic injury in the Greater Toronto Area.

Figure 2. Construction of road traffic injury spatial analysis



Source: Own Elaboration

Road traffic injury data were acquired from The Canadian Institute for Health Information (CIHI). The CIHI which was founded in 1994, is an independent and non for profit organization created by Canada’s federal, provincial and territorial governments (Statistics Canada, 2011). Their goal is to coordinate, develop and improve Canadian health and health care system by providing and distributing health information from hospitals and other organizations all over the country. The CIHI also provides information that are required for effective healthcare management, enable sound policy and raise public awareness of the factors that could contribute to affect health (Statistics Canada, 2011; CIHI, 2014). The CIHI manages and holds about 27 different databases of health information. Some of these databases are on health personnel, health spending and health services such as ambulatory care and discharge abstract.

To extract motor vehicle injury data from a larger dataset, a tabular search of International Classification of Diseases (ICD) codes was conducted to search for data of interest. International Classification of Diseases (ICD) codes are the “standard diagnostic tool for epidemiology, health management and clinical purposes” (WHO, 2014). ICD codes are used to classify diseases and other health problems, based on their type and severity. Motor vehicle injury data related to traffic were extracted from the 2004 and 2011 National Ambulatory Care⁴ (NAC) dataset using set of SQL queries in Microsoft Access. After retrieving the desired data, they were categorized into five primary injury types, such as motor vehicle collision with pedestrian, motor vehicle collision with other vehicle, other motor vehicle collision with motor vehicle, motor vehicle collision with train and finally, other non-collision motor vehicle accidents (Figure 2). However, since the extracted data were not spatially defined, the data sets were geocoded⁵ using postal codes available in the dataset. The geocoded

⁴ The National Ambulatory Care (NAC) contains data for all hospital-based and community-based ambulatory care on day surgery, outpatient clinics, and emergency departments.

⁵ Process of transforming non spatial data into spatial data that can be displayed as features on a map

incidents, which now contained x and y coordinates were then georeferenced⁶ using ArcGIS built-in tools to display the incident in each category on a map.

Following that, statistical analysis CHASS (Computing in the Humanities and Social Sciences) from the University of Toronto was used to provide census information for the Greater Toronto Area. Number of socioeconomic and socio-demographic factors that affect road traffic injury according to previous studies were selected from the census. Variables such as low education, number of unemployed, average median income, mobility, and number of seniors are used in statistical analysis to quantify their significance to the number of road traffic injury incidents (Figure 2). For effective statistical analysis, all variables were transformed to the square root, in order to reduce skewness that existed within the data.

Last but not least, for health behavior analysis, 2010 Ontario health survey from Canadian Institute Health Research (CIHR) were retrieved for statistical analysis. Data such as number of people smoking and drinking alcohol were collected for the Greater Toronto Area (Figure 2). Similar to the socioeconomic and socio-demographic data, health behavior data were used in statistical analysis to measure their significance to the number of road traffic injury incidents. They were also transformed to the square root for an effective and accurate outcomes.

4.2 Global Spatial Autocorrelation

Spatial patterns is defined by the arrangement and distribution of features in space and the geographic relationship among them (Gatrell *et al.*, 1996). Accordingly, in order to be able to measure the relationship among features according to their spatial arrangement, it is necessary to conduct spatial autocorrelation (Cliff & Ord, 1973). There are two types of spatial autocorrelation. Global spatial autocorrelation identifies whether the values of a variable show a significant pattern of regional clustering (Vaz, 2016). In contrast, local spatial autocorrelation identifies the location of high and low value clusters, usually known as hotspots and cold spots (Ord & Getis, 1995). In global measures, the relationship among features may be described as positive correlation if similar values are spatially close to each other (clustered), negative correlation if different values are located near one another (dispersed) and random if spatial pattern cannot be distinguished from the arrangement of values (Getis, 2008). In order to understand and measure such pattern, statistical tools such as Moran's I (Anselin, 1995; Ord-Getis, 1995) can be used to measure the dispersal or clustering of the features in space. As a result, Moran's I value can be computed as,

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (1)$$

where z_i is the deviation of an attribute for feature i from its mean ($x_i - X$), $w_{i,j}$ is the spatial weight between features i and j , and n is the total number of features (Anselin, 1995). S_0 is the aggregate of all the spatial weights, which is calculated as,

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{i,j} \quad (2)$$

Moran's I value are ranged from -1 to 1. A Moran's I value of +1 indicates that features are clustered and there is a positive autocorrelation; conversely, a Moran's I value of -1 indicates that features are dispersed and there is a negative spatial autocorrelation. Moran's I value

⁶ Process of aligning geographic data to a known coordinate system so it can be viewed, queried, and analyzed with other geographic data.

of 0 indicates that features are distributed randomly and therefore, spatial autocorrelation would not exist (Slocum *et al.*, 2009: 52). However, before calculating Moran's I, a spatial weight matrix was defined using queen contiguity, which outlines a locations neighbor with either a shared border or vertex (Anselin, 1995).

Moreover, the statistical significance for Moran's I can be calculated using z-score method (Erdogan, 2009). This is primarily used to see whether the outcome is the result of random distribution or not. Z-score greater than 1.96 or less than -1.96 indicates that there is a spatial autocorrelation at a 5% significant level; thus the outcome is not as a result of random distribution.

4.3 Local Spatial Autocorrelation

In order to evaluate the spatial variation and the spatial association, it is necessary to conduct local measures such as local Moran's I (Anselin, 1995) or Getis-Ord G_i^* statistics (Getis & Ord, 1992). This is carried out by counting the number of road traffic injury incidents in each dissemination area within the study area. This will allow us to determine whether the features contain high or low clustering value. Unlike measures of global spatial autocorrelation, which gives only one value indicating the degree of regional clustering across the whole study area, Getis-Ord G_i^* statistics gives a value for each location which indicates the degree of high or low value clusters (Getis & Ord, 1992). Also, Getis-Ord G_i^* statistics is used to measure the degree of association from a concentration of weighted point (Ord & Getis, 1995). For the purpose of this study, the local statistics were carried out by using queen contiguity weight matrix in relation to the number of road traffic injury incidents which was calculated earlier. Higher Getis-Ord G_i^* statistics value indicates that spatial clustering among features is significant (Vaz, 2013). This value is highly correlated with the z-score or significance level. For instance, higher value of z-score indicates the intensity of clustering among higher value known as hot spot, and low z-score value indicates the intensity of clustering among lower values known as cold spot (Ord & Getis, 1995). The Getis-Ord G_i^* statistics is computed as,

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{x} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{[n \sum_{j=1}^n w_{i,j}^2 x_j^2 - (\sum_{j=1}^n w_{i,j} x_j)^2]}{n-1}}} \quad (3)$$

where x_j is the attribute value for feature j , $w_{i,j}$ is the spatial weight between features i and j and n is the total number of features (Ord & Getis, 1995).

4.4 Level of Urbanization

According to studies, one of the main contributing factors influencing road traffic injury is urban density. It has been found that many of the road traffic injuries occur on urban roads where population is higher and traffic congestion is at its peak. This could be due to the lack of attention on the road or the way roads have been designed (Pratte, 1998). GTA has seen a significant growth in urban development in the past 30 years. As a result of urban growth (Vaz *et al.*, 2012; Vaz & Nijkamp, 2015), population has increased which has led to higher rate of road traffic injuries (Atubi, 2012). Urban density is defined as areas that are continuously being or have been developed with a population of 50,000 or more. These areas are usually considered to be developed and densely settled (Carlino, Chatterjee & Hunt, 2007). To evaluate the influence of urban density on road traffic injuries, a model was

carried out by measuring the urban size and the geographic size of each census tract. As a result, urban density in each census tract was computed as,

$$Urban\ Density = \frac{\sum_{i=1}^n U_i}{A} \quad (4)$$

where U_i is the size of each urban area and A is the geographic size of each census tract. However, urban areas were developed by combining urban categories such as commercial, residential, industrial, parks/recreation and government/institutions. The importance of evaluating the urban density in this research, is to explain if the increase in urbanization has any influence on the road traffic injuries (Vaz *et al.*, 2017).

4.5 Geographic Weighted Regression

According to number of studies (Camilloni *et al.*, 2013; Pratte, 1998; Link & Phelan, 1995), several factors have contributed in the increase of road traffic injuries worldwide. In order to investigate the relationship between the number of road traffic injuries and other related factors such as income, education, unemployment and etc., geographic weighted regression (GWR) method was used. GWR is an extension of classical standard regression (OLS) with a difference that GWR allows local parameters to be estimated rather than global parameters (Gao & Li, 2011). The main reason that GWR is used in the study is because results are more reliable and more accurate since they are at a local level. The results in GWR is similar to the classical standard regression (OLS) in which R^2 represents the goodness of fit. Similar to OLS, R^2 value ranges from 0 to 1, and the higher the R^2 value is, the better the model fits the hypothesis. Geographic weighted regression can be expressed as,

$$y_i = \beta_0(\mu_i, v_i) + \sum_{k=1}^p \beta_k(\mu_i, v_i)x_{ik} + \varepsilon_i \quad (5)$$

where y_i is the dependent variable at location i , $\beta_0(\mu_i, v_i)$ is the intercept parameter at location i , (μ_i, v_i) is the coordinate of location i , $\beta_k(\mu_i, v_i)$ is the local regression coefficient for the k -th explanatory variable at location i , X_{ik} is the value of the k -th explanatory variable at location i and lastly, ε_i are random errors at point i terms (Lu, 2011).

In addition, in order to be able to estimate the parameters in GWR, all observations must be weighted. Observations closer to point i have higher impact on the local parameter and therefore are weighted more than data that are located further away (Gao & Li, 2011). As a result the parameters are computed as,

$$\hat{\beta}_{(\mu,v)} = (X^T W_{(\mu,v)} X)^{-1} X^T W_{(\mu,v)} y \quad (6)$$

where $\hat{\beta}_{(\mu,v)}$ represents the unbiased estimate of β , $W_{(\mu,v)}$ is the weighting matrix which acts to ensure that observations near to the specific point have bigger weight value (Lu, Charlton & Fotheringham, 2011). The weighting functions which are also called kernel are being implemented using the Gaussian distance decay which is computed as

$$W_{ij} = \exp\left(-\frac{d_{ij}^2}{h^2}\right) \quad (7)$$

where W_{ij} represents the weight of observation j for location i , h is a non-negative parameter known as bandwidth which produces a decay of influence with distance and d_{ij} is the measure of Euclidean distance between location of observation i and j and is computed as (Brunsdon, Fotheringham & Charlton, 1998),

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{8}$$

where (x_i, y_i) and (x_j, y_j) are point coordinates. As a result, if observation j coincide with i , the weight value is one, and if the distance is greater than the kernel bandwidth then the weight will be set to zero. Moreover, as Huang, Wu and Barry (2010) explain, there are two weighting methods that can be used in GWR, fixed and adaptive kernel. In fixed kernel, distance is constant, though the number of adjacent neighbors vary. In adaptive kernel, distance among features vary but the number of neighbors remains constant. For instance, in adaptive kernel if feature contain dense distribution, then the spatial context is smaller, and if the feature have sparse distribution then the spatial context is larger. For the purpose of this study, adaptive kernel was chosen for the analysis.

5. RESULTS AND DISCUSSION

5.1 Road Traffic Injury Incidents Statistics

The result of this research study follows the three stream of analysis as defined earlier, defining spatial pattern of road traffic injuries, identify areas of hotspots and cold spots, and determine variables that contribute and influence the increase in road traffic injuries. Table 1 is a summary statistics comparing the number of road traffic injury incident between Ontario and GTA between 2004 and 2011.

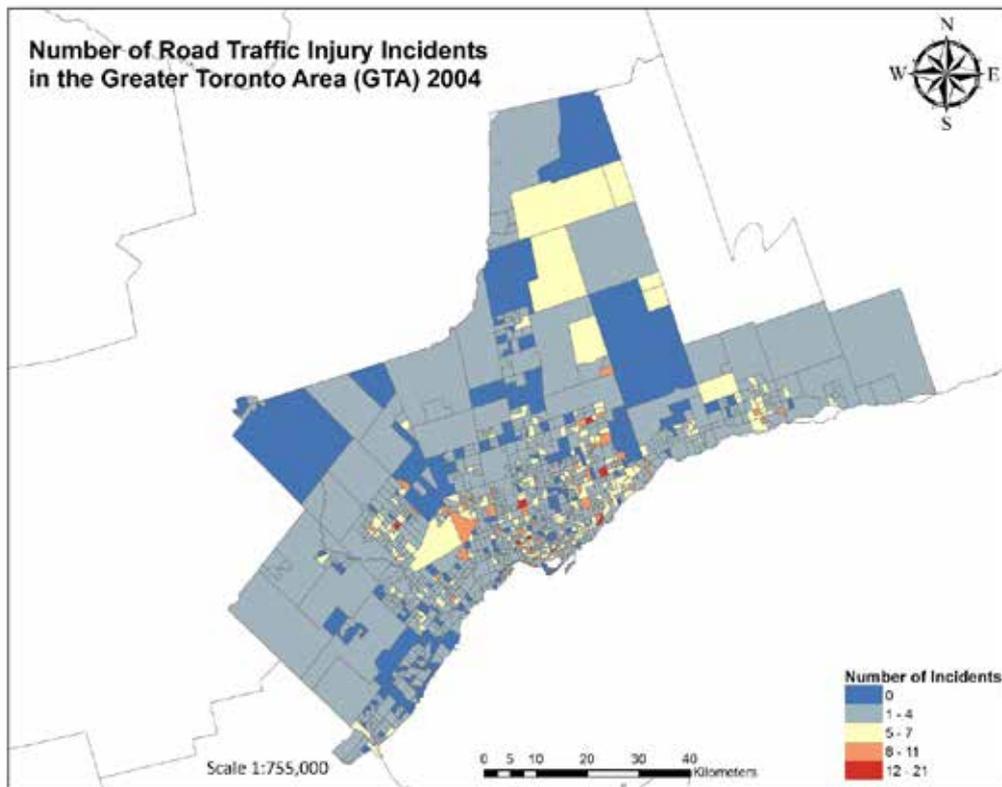
Table 1. Summary statistics of road traffic injury incidents statistics

	2004		2011		Change (%)	
	Ontario	GTA	Ontario	GTA	Ontario	GTA
Motor vehicle collision with pedestrian	3916	1837	4132	1892	6	3
Motor vehicle collision with other vehicle	137	53	139	61	1	15
Motor vehicle collision with train	196	24	76	14	-61	-42
Other motor vehicle collision with motor vehicle	3522	1621	4488	2224	27	37
Other non-collision vehicle accidents	102	25	61	16	-40	-36
TOTAL	7873	3563	8896	4207	13	18

Source: Own Elaboration

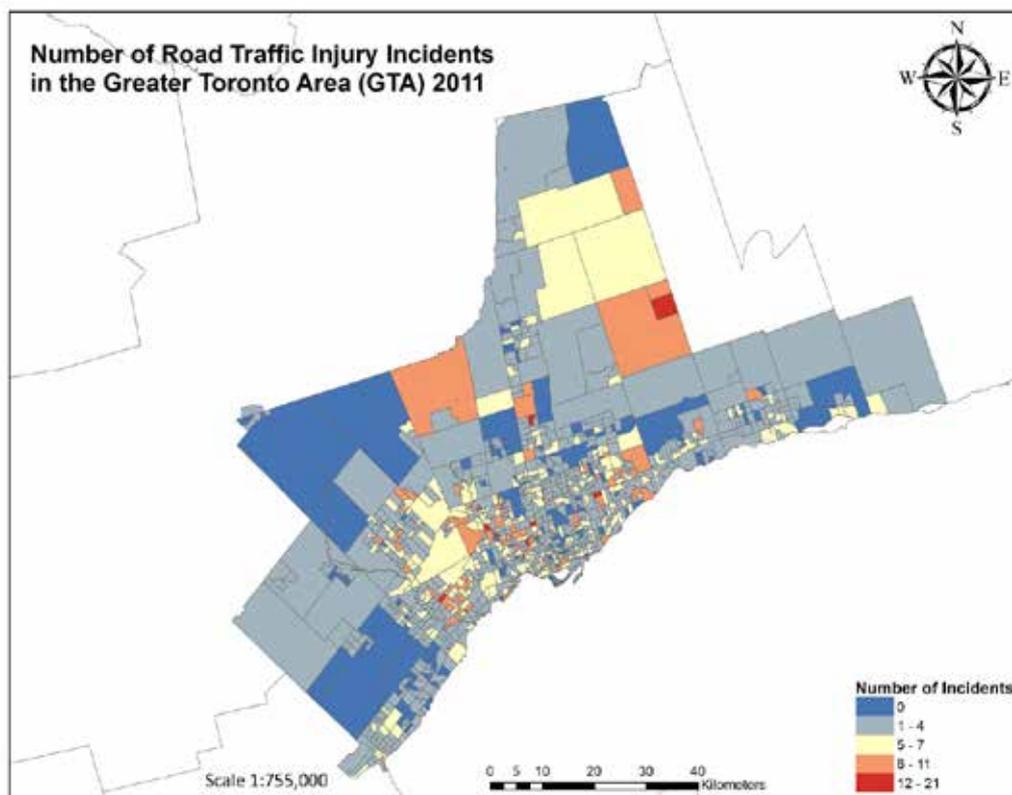
As Table 1 shows, there has been an 18% increase in the total number of road traffic injuries just within GTA between 2004 and 2011. The table also illustrates that nearly 50% of all traffic injury incidents in Ontario occur within Greater Toronto Area (GTA), and therefore this establishes a huge threat to the public. The change and increase in the number of road traffic injury incidents between 2004 and 2011 can clearly be seen in Figures 3 and 4 below.

Figure 3. Number of road traffic injury incidents in the Greater Toronto Area (GTA) 2004



Source: Own Elaboration

Figure 4. Number of road traffic injury incidents in the Greater Toronto Area (GTA) 2011



Source: Own Elaboration

This significant growth in the total number of road traffic injury in the GTA from 2004 to 2011 is primarily due to the population increase that had occurred in the GTA. However after normalizing the data by total population, it was observed that there was about 17% increase in the terms of the total number of road traffic injuries in both Ontario and the GTA between 2004 and 2011. Moreover, increase in the use of motor vehicle and vehicle dependability as a result of urban sprawl can be one of the factors of this significant change in the total number of road traffic injury in the GTA.

5.2 Global Pattern of Road Traffic Injury in the Greater Toronto Area

The Moran’s I analysis indicates a statistically significant clustering pattern of vehicle collision with pedestrian, and other motor vehicle collision with motor vehicle in the Greater Toronto Area (Table 2). With z-scores above 1.96, we know that there is a positive spatial autocorrelation with 95 percent confidence, and therefore the outcomes are not as a result of random distribution.

Table 2. Summary table of global spatial autocorrelation

	2011		
	Moran’s I	Z-score	Pattern
Motor vehicle collision with pedestrian	0.24	10.88	Cluster
Motor vehicle collision with other vehicle	0.08	0.66	Random
Motor vehicle collision with train	0.69	1.78	Cluster
Other motor vehicle collision with motor vehicle	0.20	10.45	Cluster
Other non-collision vehicle accidents	-0.15	-0.20	Random

Source: Own Elaboration

One of the reasons that could explain the clustering pattern of pedestrian collision with a motor vehicle is population density. According to the study by Clark (2003), there is a positive relationship when it comes to traffic fatalities and population density. As for the GTA, majority of Ontario’s population is concentrated in the GTA which increases the risk of traffic injury and fatalities among people. In addition, pedestrian safety is often overlooked in the main cities, because smooth movements of motor vehicles has become a priority of engineers in road design (Schuurman *et al.*, 2009).

5.3 Greater Toronto Area Hotspots

Local approaches were also used for analyzing spatial association to identify locations with highest number of incidents compare to other locations which is also known as hot spots. The results revealed that majority of motor vehicle collisions with pedestrian occur in urban areas of GTA where population density is higher. In contrast, motor vehicle collision with other vehicle hotspots are mostly concentrated on the southern part of GTA, with majority of hotspots in the Toronto region. On the contrary, Peel region and western part of Durham region have seen a significant existence of hotspots when it comes to other motor vehicle collision with motor vehicle. The result indicates that these two regions have the highest rate of injury as a result of vehicle collision compare to other regions in the study area. Compare to other three collision types, motor vehicle collision with train had the lowest number of road traffic injury incidents. However, according to the hotspot analysis results, majority of this types of incidents happen on the northern part of Toronto region where there is a rail

track that pass by. In addition to all these, other non-collision vehicle accidents hotspots have occurred on the outskirts of the city where areas are rural and population is lower.

5.4 Contributing Factors to Road Traffic Injury in the Greater Toronto Area

Though the road traffic injury contributing factors do not affect the identification of hotspots, some of the key factors can be considered. For this research, eight geographic weighted regression models have been developed based on their category (Table 3). For socio-demographic factors, variables such as age, mobility and employment status were selected. For socioeconomic factors, variables such as income and education were selected, and for health behavior factors, variables such as number of smokers and alcohol drinkers were selected for the analysis. In addition, the urban size in each census tract was also included in the analysis to see whether level of urbanization has any effect on the number of road traffic injuries. The strength of each model is evaluated by considering the R^2 value which is created after the model is executed.

Table 3. Summary statistics of geographic weighted regression

	Variables	R ²	Interpretation
Socio-Demographic	Seniors	0.93	Very strong
	Vulnerable Groups	0.60	Moderately strong
	Unemployed	0.67	Moderately strong
Socio-Economic	Average Median Income	0.58	Moderate
	Low Education	0.80	Very strong
Health Behavior	Smoke	0.53	Moderate
	Drink	0.57	Moderate
Land use cover	Urban	0.61	Moderately strong

Source: Own Elaboration

Based on the results presented in Table 3, both seniors and low education have a very strong correlation with the dependent variable, number of road traffic injury incidents. This result suggests that areas with higher number of seniors and higher number of population with lower education leads to higher number of road traffic injury incidents. Following that, number of unemployed, urban land use (Vaz & Arsanjani, 2015) cover and vulnerable groups explained 67, 61 and 60 percent of the model respectively. Surprisingly, according to these results, health behavior factors such as smoking and drinking does not have a strong correlation with number of road traffic injuries, which conveys that there should be other health behavior factors that should correlate with the number of road traffic injuries.

With seniors and low education being identified as main factors that contribute to the increase of road traffic injuries, some policies should be implemented in order to alleviate road traffic injury accidents in major cities with high population density such as Greater

Toronto Area. As an example, in London UK a congestion pricing project was implemented in the highest vehicular congested area which would force people to pay a fee in order to be able to drive in those congested areas (Mackie, 2005). This has reduced congestion significantly, allowing people to walk, shop and live in a more pleasant and safer environment. Moreover, according to the study by Mackie (2005), this policy has also reduced personal injury accidents in central London by 5%.

6. CONCLUSION

6.1 Findings

This research presented an approach based on the global spatial autocorrelation of road traffic injury for identifying the spatial pattern, and local spatial autocorrelation for identifying traffic injury hotspots. Global spatial autocorrelation allowed us to examine the pattern of injury incidents and determine whether they are clustered or dispersed in the study area. This was measured by using Moran's I value. On the other hand, local spatial autocorrelation or Getis-Ord G_i^* statistics, was used to identify hot spots in the study area. The hot spots were measured by using z-score value. Identification of hot spots are important because the results reveal areas where governments should pay more attention when it comes to road injury preventions. In addition to these two methods, geographic weighted regression (GWR) was also used in determining factors that would influence the increase in road traffic injury. One of the benefits of using GWR was that it allowed locational information to be included in the analysis, and thus the results became more accurate and reliable.

6.2 Limitations

Furthermore, different geographic units were used in the analysis. For instance, in both global and local spatial autocorrelation, dissemination areas (DA) were used, primarily due to the accuracy and reliability of the results in smaller areas. In contrast, census tracts (CT) geographic units were used for statistical analysis, because since DA geographic units are small in size, they tend to create more variance and as a result skew the results in regression analysis. Therefore, a larger geographic unit such as CT was applied in the statistical analysis to avoid skewness in the result. However, using CT geographic unit had its own limitation. As an example, the north western municipalities of Durham region (Scugog and Brock) which are considered as secluded and rural area, were excluded from the analysis, due to the lack of CT geographic units, and thus they were not included in the regression analysis. In addition to this, there were some issues within the census data as well. For example, 17 instances of null data or no data values were present within the GTA census tracts, and as a result they were also not included in the regression analysis.

6.3 Future Research

Identification of road traffic injury hotspots and determining influencing factors in road traffic injury can definitely be beneficial in injury prevention measures. The findings of this research supports those of other studies which associate low education and urban density with road traffic injury. In addition, number of seniors was also found to be a factors in road traffic injury. Consequently, these results provide a foundation for extending road traffic injury research. Future studies should definitely analyze and study the road design and visibility in urban areas. Road design is widely acknowledged in several studies. For instance studies have found that motor vehicle speed and fatality frequencies are higher on wider major arteries, and as a result there are more crashes and injuries at main intersections. Also,

study by Morency (2012) examined how road design can increase the rate of pedestrian and cyclist injuries within an urban areas. The results of this study suggested that arterial roads with increased turning range and free-flow right turn lanes may threaten pedestrian and cyclist. Therefore, to better orient street level environment and prevent such incidents to occur, further studies should include more detailed measures of street design and visibility. Furthermore, since health behaviors factors such as smoking and drinking demonstrated a moderate relationship with road traffic injury incidents, future studies should examine physical factor such as time and the distance travelled in relation to road traffic injury incidents. Some studies have found that distance travelled can pose a great risk to pedestrian and other drivers (Chapman, 1973; Lassarre *et al.*, 2007). Study by Lassarre and others, examined the relationship between road accidents and distance travelled. The results explained that drivers that travel a greater distance, are more likely to driver at a higher speed and be prone to exhaustion and thus threaten the safety of other motorists and pedestrians. Study by Chapman (1973) also found that driving time can also be an exposure to road traffic injury. For instance, driving between day and night can greatly impact the risk of road injuries. Drivers with poor visions are more prone to get in an accident since they would not be able to see and distinguish the objects clearly at night time. Having said this, to prevent and alleviate the rate of road traffic injury, further studies should include more detailed examination and analysis on time and distance travelled.

REFERENCES

- Afukarr, F.K., Antwi, P. and Ofosu-Amaah, S. (2003). Pattern of Road Traffic Injuries in Ghana: Implications for Control. *Injury Control and Safety Promotion*. 10(1-2): 69-76.
- Aguero-Valverde, J. and Jovanis, P.P. (2006). Spatial Analysis of Fatal and Injury Crashes in Pennsylvania. *Accident Analysis and Prevention*. 38: 618-625.
- Al-Omari, B.H. and Obaidat, E.S. (2013). Analysis of Pedestrian Accidents in Irbid City, Jordan. *The Open Transportation Journal*. 7: 1-6.
- Ameratunga, S., Hajar, M. and Norton, R. (2006). Road Traffic Injuries: Confronting Disparities to Address a Global-Health Problem. *The Lancet*. 367: 1533-1540.
- Anselin, L. (1995). Local Indicators of Spatial Association – LISA. *Geographical Analysis*. 27(2): 93-115.
- Ashokbhai, P.V. and Jain, A.M. (2014). Road Safety Audit: An Identification of Black Spots on Busy Corridor Between Narol-Naroda of Ahmedabad City. *International Journal of Engineering and Technical Research*. 2(3): 86-89.
- Atubi, A.O. (2012). Determinants of Road Traffic Accident Occurrences in Lagos State: Some Lessons for Nigeria. *International Journal of Humanities and Social Science*. 2(6): 252-259.
- Brunsdon, C., Fotheringham, A.S. and Charlton, M.E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*. 28(4): 281-298.
- Brunsdon, C., Fotheringham, S. and Charlton, M. (1998). Geographically Weighted Regression-Modelling Spatial Non-Stationarity. *Journal of the Royal Statistical Society*. 47(3): 431-443.
- Camilloni, L., Farchi, S., Chini, F., Rossi, P.G., Borgia, P. and Guasticchi, G. (2013). How Socioeconomic Status Influences Road Traffic Injuries and Home Injuries in Rome. *International Journal of Injury Control and Safety Promotion*. 20(2): 134-143.

- Carlino, G.A., Chatterjee, S. and Hunt, R.M. (2007). Urban Density and the Rate of Invention. *Journal of Urban Economics*. 61: 389-419.
- Chapman, R. (1973). The concept of exposure. *Accident Analysis and Prevention*. 5: 95-110.
- Charlton, M., Fotheringham, S. and Brunson, C. (2006). Geographically Weighted Regression. Discussion Paper. Unpublished. ESRC National Center for Research Methods.
- Chisholm, D., Naci, H., Hyder, A.A., Tran, N.T. and Peden, M. (2012). Cost effectiveness of strategies for reducing road traffic injuries in sub-Saharan Africa and South East Asia: mathematical modelling study. *BMJ*. 344(612): 1-9.
- CIHI (2014). *Vision and Mandate*. Retrieved May 15, 2014, from <http://www.cihi.ca/CIHI-ext-portal/internet/EN/SubTheme/about+cihi/vision+and+mandate/cihi010703>
- Clark, D.E. and Cushing, B.M. (2003). Rural and urban traffic fatalities, vehicle miles, and population density. *Accident Analysis and Prevention*. 36: 967-972.
- Cliff, A.D. and Ord, J.K. (1973). *Spatial Autocorrelation*. Pion:London.
- Cubbin, C., LeClere, F.B. and Smith, G.S. (2000). Socioeconomic Status and the Occurrence of Fatal and Nonfatal Injury in the United States. *American Journal of Public Health*. 90(1): 70-77.
- Erdogan, S. (2009). Explorative Spatial Analysis of Traffic Accident Statistics and Road Mortality among the Provinces of Turkey. *Journal of Safety Research*. 40: 341-351.
- Ewing, R., Schieber, R.A. and Zegeer, C.V. (2003). Urban Sprawl as a Risk Factor in Motor Vehicle Occupant and Pedestrian Fatalities. *American Journal of Public Health*. 93(9): 1541-1545.
- Flahaut, B., Mouchart, M., Martin, E.S. and Thomas, I. (2003). The Local Spatial Autocorrelation and The Kernel Method for Identifying Black Zones: A Comparative Approach. *Accident Analysis and Prevention*. 35: 991-1004.
- Gao, J. and Li, S. (2011). Detecting Spatially Non-Stationary and Scale-Dependent Relationships Between Urban Landscape Fragmentation and Related Factors Using Geographically Weighted Regression. *Applied Geography*. 31: 292-302.
- Gatrell, A.C., Bailey, T.C., Diggle, P.J. and Rowlingson, B.S. (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Transaction of the Institute of British Geographers*. 21(1): 526-274.
- Getis, A. and Ord, J.K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*. 24(3): 189-206.
- Getis, A. (2008). A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective. *Geographical Analysis*. 40: 297-309.
- Geurts, K. and Wets, G. (2003). *Black Spot Analysis Methods: Literature Review*. Flemish Research Center for Traffic Safety. Diepenbeek, Belgium.
- Greater Toronto Marketing Alliance (2011). *Population Forecast and Growth Rate*. Retrieved April 27, 2014, from <http://www.greatertoronto.org/why-greater-toronto/economic-overview/population>
- Hakkert, A.S. and Mahalel, D. (1978). Estimating the Number of Accidents at Intersections from a Knowledge of the Traffic Flows on the Approaches. *Accident Analysis and Prevention*. 10: 69-79.
- Huang, B., Wu, B. and Barry, M. (2010). Geographically and Temporally Weighted Regression for Modeling Spatio-Temporal Variation in House Prices. *International Journal of Geographical Information Science*. 24(3): 383-401.

- Jackson, M.C. and Waller, L.A. (2005). Exploring Goodness-of-Fit and Spatial Correlation Using Components of Tango's Index of Spatial Clustering. *Geographical Analysis*. 37: 371-382.
- Jacob, G.D. and Sayer, I. (1983). Road Accidents in Developing Countries. *Accident Analysis and Prevention*. 15(5): 337-353.
- LaScala, E.A., Gerber, D. and Gruenewald, P.J. (1999). Demographic and Environmental Correlates of Pedestrian Injury Collision: A Spatial Analysis. *Accident Analysis and Prevention*. 32: 651-658.
- Lassarre, S., Papadimitriou, E., Yannis, G. and Golias, J. (2007). Measuring accident risk exposure for pedestrians in different micro-environments. *Accident Analysis and Prevention*. 39: 1226-1238.
- Link, B.G. and Phelan, J. (1995). Social Conditions as Fundamental Causes of Disease. *Journal of Health and Social Behavior*. 35: 80-9.
- Lu, B., Charlton, M. and Fotheringham, A.S. (2011). Geographically Weighted Regression Using a Non-Euclidean Distance Metric with a Study on London House Price Data. *Procedia Environmental Science*. 7: 92-97.
- Mackie, P. (2005). The London congestion charge: A tentative economic appraisal. *Transport Policy*. 12: 288-290.
- Manepalli, U.R., Bham, G.H. and Kandada, S. (2011). Evaluation of Hotspots Identification Using Kernel Density Estimation (k) and Getis-Ord (G_i^*) on I-630. In: *International Conference on Road Safety and Simulation*.
- Martin, P.G., Crandall, J.R. and Pilkey, W.D. (2000). Injury Trends of Passenger Car Drivers in Frontal Crashes in the USA. *Accident Analysis and Prevention*. 32: 541-557.
- MetropolisIQ (2012). *GTA Population, Employment and Household Projections to 2041*. Retrieved April 27, 2014, from <https://metropolisiq.com/blog/entry/updated-gta-population-employment-and-household-projections-to-2041>
- Montella, A. (2010). A Comparative Analysis of Hotspots Identification Method. *Accident Analysis and Prevention*. 42: 571-581.
- Morency, P., Gauvin, L., Plante, C., Fournier, M. and Morency, C. (2012). Neighborhood Social Inequalities in Road Traffic Injuries: The Influence of Traffic Volume and Road Design. *American Journal of Public Health*. 102(6): 1112-1119.
- Moutinho, L.A.M. and Hutcheson, G.D. (2011). *The SAGE Dictionary of Quantitative Management Research*. London: SAGE Publications Ltd.
- Nakaya, T., Fotheringham, A.S., Brunson, C. and Charlton, M. (2005). Geographically Weighted Poisson Regression for Disease Association Mapping. *Statistics in Medicine*. 24: 2695-2717.
- Odero, W., Garner, P. and Zwi, A. (1997). Road Traffic Injuries in Developing Countries: A Comprehensive Review of Epidemiological Studies. *Tropical Medicine and International Health*. 2(5): 445-460.
- Ord, J.K. and Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*. 27(4): 286-306.
- Pratte, D. (1998). Road to Ruin: Road Traffic Accidents in the Developing World. *NEXUS*. 13: 46-62.
- Ramage-Morin, P.L. (2008). Motor Vehicle Accident Death, 1979 to 2004. *Statistics Canada Health Reports*. 19(3) :1-7.

- Scheidt, P.C., Harel, Y., Trumble, A.C., Jones, D.H., Overpeck, M.D. and Bijur, P.E. (1995). The Epidemiology of Nonfatal Injuries among US Children and Youth. *American Journal of Public Health*. 85(7): 932-938.
- Schuurman, N., Cinnamon, J., Crooks, V.A. and Hameed, S.M. (2009). Pedestrian Injury and the Built Environment: An Environmental scan of hotspots. *BMC Public Health*. 9(233): 1-10.
- Slocum, T.A., McMaster, R.B., Kessler, F.C. and Howard, H.H. (2009). *Thematic Cartography and Geovisualization*. Upper Saddle River, N.J.: Pearson Prentice Hall.
- Statistics Canada (2011). *About the Canadian Institute for Health Information (CIHI)*. Retrieved May 7, 2014, from <http://www.statcan.gc.ca/pub/82-221-x/2011002/abc-eng.htm>
- Tobler, W.R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*. 46(2): 234-240.
- Transport Canada (2011). *Road Safety in Canada*. Retrieved May 22, 2014, from <http://www.tc.gc.ca/eng/motorvehiclesafety/tp-tp15145-1201.htm>
- Truong, L.T. and Somenahalli, S.V. (2011). Using GIS to Identify Pedestrian-Vehicle Crash Hot Spots and Unsafe Bus Stops. *Journal of Public Transportation*. 14(1): 99-114.
- Vaz, E., Nijkamp, P., Painho, M. and Caetano, M. (2012). A multi-scenario forecast of urban change: A study on urban growth in the Algarve. *Landscape and Urban Planning*. 104(2): 201-211.
- Vaz, E. and Jokar Arsanjani, J. (2015). Crowdsourced mapping of land use in urban dense environments: An assessment of Toronto. *The Canadian Geographer/Le Géographe canadien*. 59(2): 246-255.
- Helbich, M., Brunauer, W., Vaz, E. and Nijkamp, P. (2013). Spatial heterogeneity in hedonic house price models: the case of Austria. *Urban Studies*. 0042098013492234.
- Vaz, E. (2016). The future of landscapes and habitats: The regional science contribution to the understanding of geographical space. *Habitat International*. 51: 70-78.
- Vaz, E., Lee, K., Moonilal, V. and Pereira, K. (2017). Potential of Geographic Information Systems for Refugee Crisis: Syrian Refugee Relocation in Urban Habitats. *Habitat International*.
- Vaz, E. and Khaper, M. (2016). New Resources For Smart Food Retail Mapping. A Gis And The Open Source Perspective. *Journal of Spatial and Organizational Dynamics*. 4(4): 305-313.
- Vaz, E., Cusimano, M. and Hernandez, T. (2015). Land use perception of self-reported health: Exploratory analysis of anthropogenic land use phenotypes. *Land Use Policy*. 46: 232-240.
- de Noronha, T. and Vaz, E. (2015). Framing urban habitats: the small and medium towns in the peripheries. *Habitat International*. 45: 147-155.
- Vaz, E. (2013). The spatial business landscape of India. *Journal of Spatial and Organizational Dynamics*. 1(4): 241-253.
- Vaz, E. and Nijkamp, P. (2015). Gravitational forces in the spatial impacts of urban sprawl: An investigation of the region of Veneto, Italy. *Habitat International*. 45: 99-105.
- World Health Organization (2005). *World Report on Road Traffic Injury Prevention*. Retrieved May 3, 2014, from http://siteresources.worldbank.org/EXTTOPGLOASAF/Resources/WHO_full_report_en.pdf

- World Health Organization (2014). *International Classification of Diseases (ICD)*. Retrieved May 16, 2014, from <http://www.who.int/classifications/icd/en/>
- World Web (2013). *Greater Toronto Area*. Retrieved April 27, 2104, from <http://www.greatertoronto.worldweb.com/>
- Zhang, L., Bi, H., Cheng, P. and Davis, C.J. (2004). Modeling Spatial Variation in Tree Diameter-Height Relationships. *Forest Ecology and Management*. 189: 317-329.