Applied Geography 87 (2017) 207–221

Contents lists available at ScienceDirect

Applied Geography

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

Socioeconomic characteristics and crash injury exposure: A case study in Florida using two-step floating catchment area method

Ayberk Kocatepe [a, *], Mehmet Baran Ulak [a], Eren Erman Ozguven [a], Mark W. Horner [b], Reza Arghandeh [c]

[a] *Department of Civil and Environmental Engineering, FAMU–FSU College of Engineering, Tallahassee, FL 32310, United States*
[b] *Department of Geography, Florida State University, Tallahassee, FL 32306, United States*
[c] *Department of Electrical and Computer Engineering, FAMU–FSU College of Engineering, Tallahassee, FL 32310, United States*

A R T I C L E I N F O

Article history:
Received 6 February 2017
Received in revised form
3 July 2017
Accepted 8 August 2017

Keywords:
Socioeconomic analysis
Crash hotspot exposure
Public safety
Empirical-Gaussian approach
Spatial analysis

A B S T R A C T

The objective of this study is to investigate the exposure of different population groups to severe injury crash hotspots using an empirical-Gaussian two-step floating catchment area (EG-2SFCA) method based on roadway network distances and a socioeconomic-based weighting approach. This is performed by developing a special form of a crash-to-population ratio index that incorporates the severe crash hotspots relative to the locations of populations they might impact. While identifying these hotspots, four different age groups are considered: 17 and younger, 18 to 21, 22 to 64 and 65 and older. For each age group, severe crash hotspots are identified based on the roadway network and the number of severely injured crash occupants that belong to the specific age group. Using these age-specific crash hotspots and the EG-2SFCA method, communities that were exposed to elevated crash injury risk (crash injury exposure) have been identified. Furthermore, from a residential perspective, a socioeconomic analysis is conducted in order to develop a socioeconomics-based crash injury exposure measure. This measure assesses the exposure of different socioeconomic groups to the risk of being injured. Results demonstrated by applying this measure in the Tampa Bay region, FL show that different population groups are under varying risk of being injured depending on their residential location. The developed approach has the potential to be a social fairness measure able to be applied by agencies, which could enhance the well-being of communities that are subject to elevated injury risk.

1. Introduction

The Tampa Bay region in Florida, identified as part of District 7 by the Florida Department of Transportation (FDOT), is an important development zone, where significant growth is planned and expected in the future (District 7 Office, 2016). Population growth will lead to more traffic, and therefore related traffic safety problems associated with the roadways such as crashes. As such, the identification of crash hotspots (high crash risk locations) becomes a critical issue for the populations living near locations where severe injury crashes (involving injuries and/or fatalities) are clustered. Understanding the variable exposure of different socioeconomic groups to these severe injury clusters can help us manage this growth, and develop better transportation plans and policies. Therefore, there is a need to evaluate crash injury exposure, especially those associated with fatalities and severe injuries, accounting for the socioeconomics of the population and available transportation network in the region.

Many previous studies have focused on problems associated with roadway crashes. Among those, several studies looked at the relationships between crash frequency and people's age and demographics given a geo-specific location of the crash involvement (Abdel-Aty, Chen, & Radwan, 1999; Boyce & Geller, 2002; Krahe & Fenske, 2002). These studies confirm that driving behavior can be substantially different between age groups. This behavior is also relevant with respect to different geospatial considerations. Geospatial models have also been used to analyze and visually assess spatial roadway crash data. For example, spatial characteristics and the distribution of crashes on roadway networks have been examined via various methods such as hotspot detection analysis

* Corresponding author. 2525 Pottsdamer Street, A-219, Tallahassee, FL 32310, United States.
E-mail address: ak13y@my.fsu.edu (A. Kocatepe).

http://dx.doi.org/10.1016/j.apgeog.2017.08.005
0143-6228/

and regression models (Dai, 2012; Gundogdu, 2010; Ulak, Ozguven, Spainhour, & Vanli, 2017). Additionally, crashes will have an affect not only on the drivers but also on vehicle occupants. Moreover, Blatt and Furman (1998), Ulak et al. (2017) affirm a common notion that people are usually involved in crashes on roadways where they travel the most. This also implies that people might be involved in more crashes on roadways closer to their homes where they access easily (Burdett, Starkey, & Charlton, 2017). Clearly roadways closer to a given residence would be more likely to be used for trip purposes by the people living in that residence.

Transportation accessibility has also been frequently studied in the literature. Numerous studies focus on defining and measuring people's accessibility to various facilities such as supermarkets (Widener, Farber, Neutens, & Horner, 2015), libraries (Horner, Duncan, Wood, Valdez-Torres, & Stansbury, 2015), nursing facilities (Saliba, Buchanan, & Kington, 2004), multimodal facilities, and regular or special needs shelters. Recently, researchers have also investigated freight accessibility and logistic employment in the U.S. (Van den Heuvel et al., 2014), and accessibility to freight terminals (Cartenì, 2014; Thomas, Hermia, Vanelslander, & Verhetsel, 2003). There also several pioneering studies that focused on a variety of mathematical models. These methods consists of the gravity model (Joseph & Bantock, 1982), regional availability model (Khan, 1992), kernel density models (Guagliardo, 2004), and floating catchment methods (Luo & Wang, 2003; Radke & Mu, 2000). In the literature, accessibility is defined as the volume and proximity of services provided to the population of interest or the services that are available to a certain region or population given the prevailing transportation system. In this paper, a "crash injury exposure" measure is developed, which measures how a region or neighborhood is more or less exposed to crash related injuries compared to others. For this purpose, a measure of accessibility is translated into one that captures crash injury exposure (or 'accessibility', but in a negative connotation since it is about proximity to crash hotspots).

Among the existing methods, the two step floating catchment area method (2SFCA) is very promising in terms of applicability to "crash vs. population" studies. The pioneering 2SFCA studies were conducted by Luo and Wang (2003) and Radke and Mu (2000). Over the last years, several studies have developed methodological improvements to the traditional 2SFCA approach. For example, residential segregation in spatial access to healthcare facilities was investigated in the in metropolitan Detroit area using the catchment method integrating a Gaussian function to continuously discount access within the catchment areas (Dai, 2010). Another study proposed to incorporate a kernel function as part of the 2SFCA method in order to capture the variation in each catchment area for accessibility to food stores in the southwest Mississippi (Dai & Wang, 2011). On the other hand, 3SFCA (three-step floating catchment area method) was proposed to account for a reasonable model of healthcare supply-demand in the Austin-San Antonio area (Wan, Zou, & Sternberg, 2012). The aim was to reduce the overestimation of healthcare demand problem and address potential competition among suppliers. A modified 2FSCA method has been published to account for public transport opportunities using continuous decay functions with a case study in Wales (Langford, Fry, & Higgs, 2012). An early application of variable catchment areas proposed a smoother and continuous distance decay function in Victoria, Australia (McGrail, 2012). The Gaussian function was also used to account for the continuous distance decay, with a focus on the accessibility to vaccine sites for rabies in Sao Paolo (Polo, Acosta, & Dias, 2013). Language barriers and ability of physicians to accept new patients were also evaluated with this approach in Ontario, Canada (Bell, Wilson, Bissonnette, & Shah, 2013). In 2015, a variable catchment method (VFCA) was proposed to conceptualize the accessibility to parks in multi-modal cities using the attractiveness of a park as a measure with a case study in Mecklenburg County of North Carolina (Dony, Delmelle, & Delmelle, 2015). In another study, Luo (2014) introduced the Huff model into the catchment method in order to resolve the effect of distance impedance and supply capacity on spatial accessibility, which was enhanced by Lin et al. (2016). Clearly there are a wide range of applications and extensions to the 2SFCA model, as it can be modified to adapt to a range of scenarios.

Literature suggests that the 2SFCA method generally applies to situations where there are supply-demand interactions such as patients' seeking primary care services. In our study, however, the 2SFCA was applied taking a different approach. To elaborate, the hotspots used in the study, on the one hand, are considered as a hazard that threatens the public health, which is analogous to the "supply" in traditional 2SFCA studies. Population, on the other hand, simply represents the "demand". Therefore, as long as a population group has access to this "supply" (i.e. crash hotspots), that group is exposed to 'danger', and hence has a risk of being injured. Therefore, the 2SFCA approach is utilized to assess this "crash hotspots-population" interaction, and this type of use of 2SFCA method is novel in transportation planning and transportation geography fields. Note that these hotspots could potentially be replaced with any other hazard that threatens public health such as crime, pollutants, or any other technological hazard hotspots (Malleson & Andresen, 2015). The modified model is applied in the Tampa Bay region of FL.

More broadly, the purpose of this study is to investigate the proximity of residents living in neighborhoods to severity-weighted crash hotspots (regardless of the prevailing traffic conditions) rather than identifying roadway sections posing a high relative crash risk or having unexpectedly high numbers of crashes with respect to their overall traffic volume. It is important to emphasize that this is not a crash frequency or crash rate study, in which adopting traffic volume-normalized crash frequencies would be clearly more favorable. Additionally, the study does not focus on the number of trips generated from one point to another, or stated more broadly, the origin or destination of particular trips. In that sense, this study defines "crash injury exposure" as the exposure of the population groups in census units (thanks to the catchment method employed) to the presence of severe crash hotspots that are identified based on severely injured occupants involved in accidents. The socioeconomic-based crash injury exposure measure leads to a weighted "fairness measure" controlling for socioeconomic groups depending on their residential location and, finally the total number of people in each of these groups.

2. Methodology

Three main steps are identified as part of the approach: (a) roadway network-based crash hotspot identification for specific age groups, (b) application of the empirical Gaussian two step floating catchment area method (EG-2SFCA), and (c) evaluation of socioeconomic-based crash injury exposure measure (SECIE). Crash injury exposure is measured by the roadway network distance between the crash hotspot locations and the geometric centroids of the U.S. Census blocks. These values are aggregated to higher spatial scales and used to find a weighed metric for U.S. Census tracts and counties using the socioeconomic data associated with the census block groups. Socioeconomic data are based on the American Community Surveys (ACS) data component and is attached to each block group (ACS, 2010), which is obtained from the Florida Geographical Data Library. Variables related to ethnicity, education level, and the poverty level are collected. In order to visually illustrate the results, crash injury exposure maps are created in

ArcGIS 10.2 (ESRI, 2014).

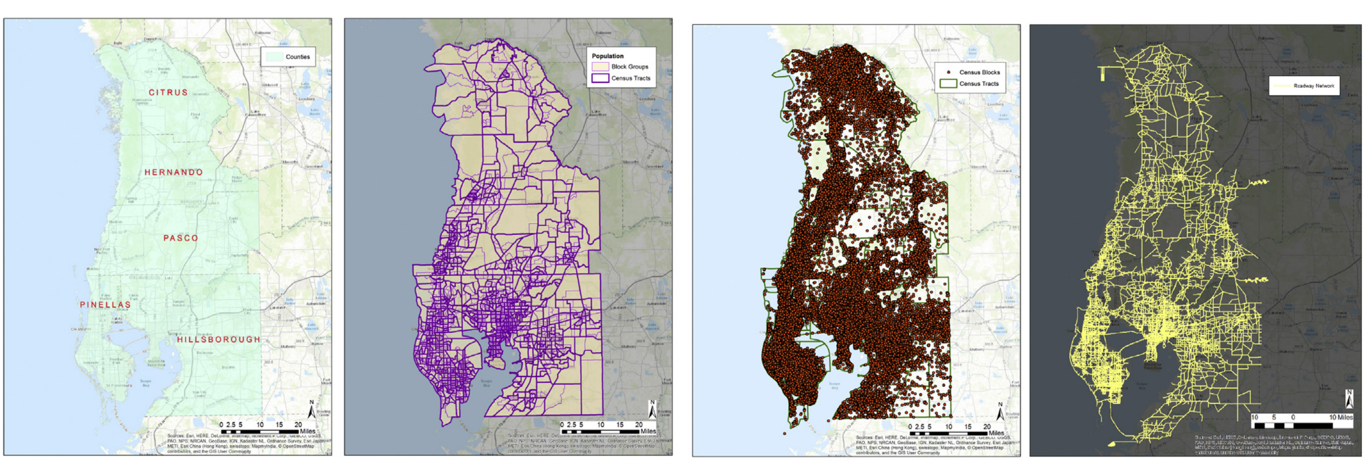
2.1. Study area, population data, and roadway network

The study focuses on the Tampa Bay region in District 7 of Florida, as identified by the Florida Department of Transportation (Fig. 1). District 7 has five counties, and two of them are heavily urbanized: namely Pinellas and Hillsborough counties with 944,971 and 1,325,683 residents, respectively. The other three counties, Citrus, Hernando and Pasco, are more rural counties with around 800,000 residents in total. Demographic data are obtained from the U.S. Census Bureau (U.S. Census Bureau, 2010) and the Florida Geographical Data Library (FGDL, 2016), which represent the population counts for 2010. This data contains census tracts and census block groups, and the latest American Community Survey (ACS) data attached to each census block group. This ACS data contains various socioeconomic information, from ethnicity to income, and from vehicle ownership to education level. Note that District 7 is divided in to 766 census tracts and 2098 census block groups, 62,939 census blocks (Fig. 1). Moreover, Fig. 2 illustrates locations where each age group are concentrated in the region. Four different age groups are used in order to identify the severe crash hotspots for each age group separately: 17 and younger, 18 to 21, 22 to 64 and 65 and older. Note that the U.S. Census data do not include socioeconomic characteristics pertaining to each age group at the block group level. For example, it is unknown how many older (65+) or younger (17-) adults live below poverty or belong to a certain ethnic group based on the U.S. Census. Additionally, roadway network data is obtained from the Tampa Bay Regional Model (District 7 Office, 2015), which contains the Tampa Bay region roadway network (Fig. 1-d) and the related traffic information, through CUBE Software (Citilabs, 2016). This network is utilized to calculate the network distances between census blocks centroids and crash hotpots.

2.2. Crash data and hotspot identification

Crash data are acquired from Florida Department of Transportation (FDOT), which covers the period between 2013 and 2014 (FDOT, 2015). The crash data include all crashes that occurred on all types of roadways (including local and state roads, U.S. highways, and interstates) within the study area (District 7 of FDOT). To be clear, all occupants (drivers, passengers, and non-motorists) who were involved in these crashes are included in the network-based severe crash hotspot analysis. Severe crash hotspots, therefore, are identified based on the number of severely injured occupants in that crash and the roadway network distance between crash points. Note that severe crashes include those that cause non-incapacitating and incapacitating injuries, and the fatalities. The FDOT data for District 7 include 99,432 crashes with 231,834 occupants involved in those crashes, of which 8969 of them are severely injured (including fatality). Furthermore, the ZIP code locations of crash occupants are provided in the data, which is used to determine the decay function used in 2SFCA method (please see section 2.3.). The urban and rural crashes were distinguished in order to increase the accuracy of the hotspot identification. In addition, this study not only focuses on severe crashes but also different age groups involved in those crashes. The crash hotspots of different age groups: 17 and younger, 18 to 21, 22 to 64 and 65 and older, are obtained rather than those hotspots for the whole population. Unfortunately the available US census data (census blocks) classifications are such that population is broken down into age groups at certain cut points (0–5, 5–17, 18–21, 22–29, 30–39, 40–49, 50–64 and 65+). Therefore, we had to work with this age group breakdown in our study. We merged the 22–29, 30–39, 40–49, and 50–64 age groups into one cohort, and considered this group as the 'working age' class, which spans from post-graduation from universities to pre-retirement. Similarly, the 0–17 age group constitutes minors whereas the 18–22 age group includes roughly college students. Finally, 65+ age group covers the population who may be beginning and already in retirement. These population breakdowns into such age cohorts is consistent with previous studies (Horner et al., 2015). Note that, FDOT crash data do not include any ethnic or income related information of individuals for the selected years, and this study does not focus on the "who was involved in crashes where" question. This is strictly a study of exposure, where a region's residents risk of being injured in a crash, based on their proximity (accessibility but in a negative connotation) to severity-weighted crash hotspots, is assessed. Nonetheless, the use of any socioeconomic characteristics of individuals involved in crashes is not possible since this type of data is not readily available.

After processing the crash data, Getis-Ord Gi* (Getis & Ord, 1992; Khan, Qin, & Noyce, 2008) method is utilized in order to identify the severity-weighted crash hotspots. The number of severely injured occupants involved in each crash are used as the



(a) **(b)** **(c)** **(d)**

Fig. 1. District 7 model components (a) counties (b) census block groups and census tracts, (c) geographical centroids of census blocks, (d) roadway network.

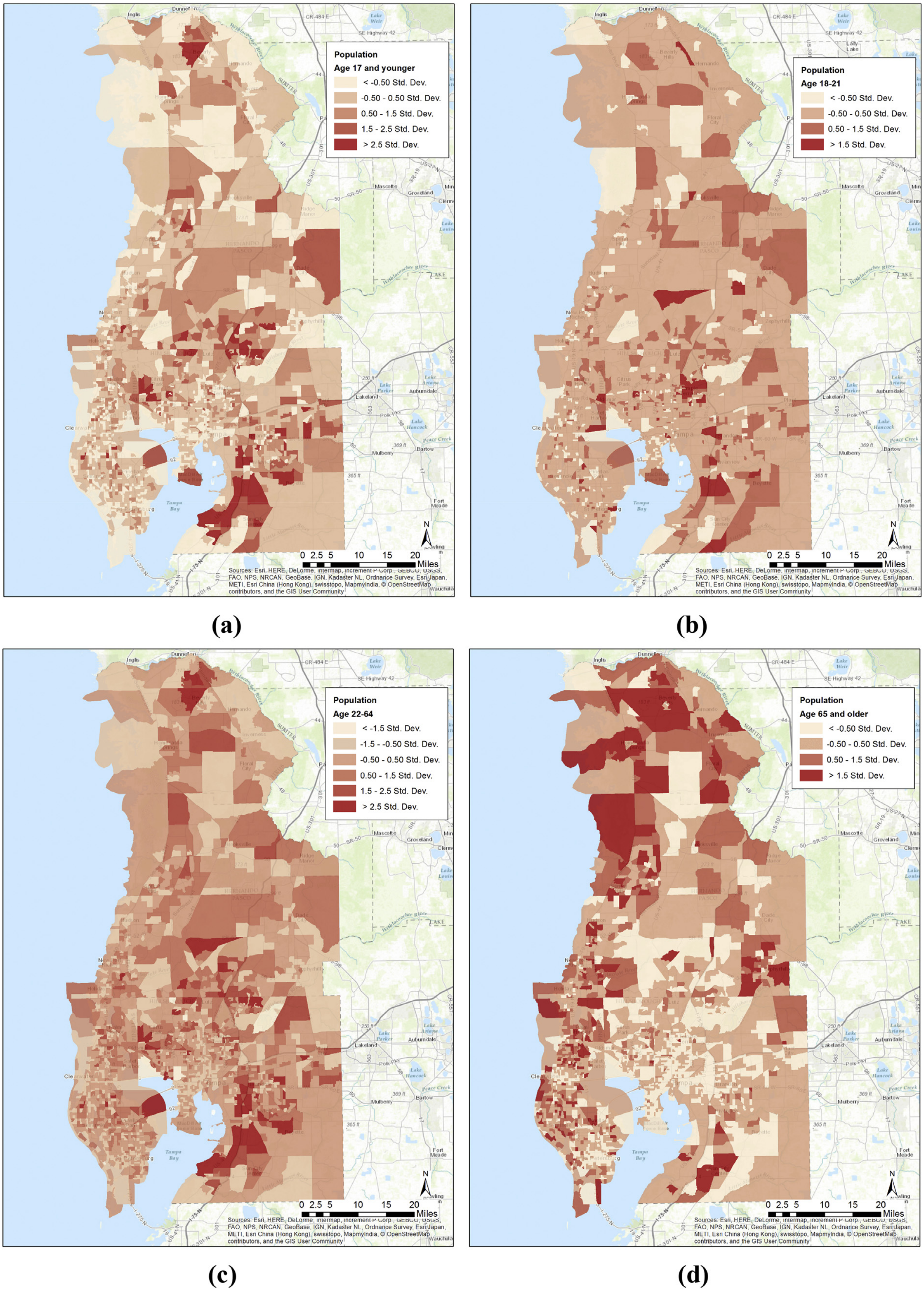


Fig. 2. Overview of District 7 demographic characteristics at census block group level (a) 17- population (b) 18–21 population (c) 22–64 population (d) 65+ population.

weight in the statistic. The ArcGIS 10.2 geoprocessing tool, contains an implementation of the Gi* statistic which allows the user to input a weight from a field in their shapefiles, and is used to identify hotspots that are statistically significant (ArcGIS, 2016). In the analysis, the same distance band value is not used for urban and rural crashes since the spatial relationships between urban crashes (more clustered, distance band: 150 m) are different than rural crashes (less clustered, distance band: 1000 m) (Blazquez & Celis, 2013; Steenberghen, Aerts, & Thomas, 2010; Xie & Yan, 2013). This differentiation may lead to more accurate 2SFCA results since the distance decay parameters likely vary over the study area. For the Getis-Ord Gi* method, roadway network-based distances are used, which is convenient since crash points are distributed along roadway networks.

2.3. Two-step floating catchment area (2SFCA) method

The basis of this effort is the two-step floating catchment area (2SFCA) method conducted using the roadway network distances. 2SFCA has been employed in a number of studies previously to assess spatial access to health services. By utilizing a special form of crash-to-population ratio through a modified 2SFCA method, namely the empirical-Gaussian 2SFCA, which uses a continuous distance decay function, a crash injury exposure measure is defined. This measure is calculated based on the network distance (proximity) between severity weighted crash hotspots and census block centroids. The crash injury exposure measure can be viewed as 'accessibility' but in a negative connotation, unlike traditional approaches where accessibility is defined as the volume and proximity of services provided to the population of interest or the services that are available to a certain region or population given the prevailing transportation system. This negative connotation is that crash hotspots are considered as a hazard that threatens the public health, which is analogous to the "supply" in traditional 2SFCA studies. Population, on the other hand, simply represents the "demand". Therefore, the 2SFCA approach is utilized to assess this "crash hotspots-population" interaction, and this type of use of 2SFCA method is novel in transportation geography field.

The 2FSCA method as applied here has three main components: (a) Population data for each age group (centroids of each U.S. Census blocks), (b) crash hotspot locations for each age group, and (c) an empirical-Gaussian decay function. The case specific decay function was determined based on available crash data and residences of crash occupants. As such, the distribution of network distances between the residence (identified by ZIP codes) and crash locations of every individual involved in crashes were identified. This provided the following two components/insights: 1) It was determined that more than 90% of crash occupants were involved in crashes within 20 network miles from their residential ZIP code 2) an empirical cumulative distribution function (CDF) was utilized to model a decay function for the proposed approach since it reflects how crashes decrease at an increasing the distance from residences (Fig. 3). As a result, a Gaussian function was fitted to the obtained complementary CDF, also called a survival function, in order to obtain the case specific decay function. The 20 mile distance, calculated using network distances, was adopted as the threshold for the catchment network of each census block. That is, severity-weighted crash hotspots within 20 miles of a census block centroid, together with the decay function, were used to calculate the crash injury exposure. The reason behind employing a distance decay is that crashes should not impose the same risk at all distances, as populations are likely to be more exposed to crashes nearby, which is consistent with the literature (Burdett et al., 2017).

The approach taken here utilizes the following empirical-Gaussian form of the model previously developed, and employed in past studies (Dai, 2010; Kwan, 1998; Luo & Qi, 2009; McGrail, 2012) as follows:

$$
\begin{aligned}
A_i &= \sum_{l \in \{d_{ij} \le 20\}} R_l G(d_{ij}, 20) \\
&= \sum_{l \in \{d_{ij} \le 20\}} \frac{S_l G(d_{ij}, 20)}{\sum_{k \in \{d_{kj} < 20\}} D_k \, G(d_{ij}, 20)}
\end{aligned}
\tag{1}
$$

where R_l is the population-to-crash at crash hotspot l falling within the catchment centered at census block centroid i (i.e., $d_{ij} < 20$), 20 miles is the catchment threshold network distance, d_{ij} is the distance between i and j, S_l corresponds to the crash hotspot, and D_k is the total population at location k falling in the catchment (i.e., $d_k < 20$). G is the decay weight based on the continuous decay function, which is a Gaussian function fitted to the obtained complementary CDF as follows:

$$
G(d_{ij}, 20) \begin{cases} a_1 * e^{-\left(\frac{d_{ij}-b_1}{c_1}\right)^2} + a_2 * e^{-\left(\frac{d_{ij}-b_2}{c_2}\right)^2} - k * \frac{1}{e^{d_{ij}}} & \text{if } d_{ij} < 20 \text{ miles} \\ 0 & \text{if } d_{ij} \ge 20 \text{ miles} \end{cases}
\tag{2}
$$

where the parameter values of decay function are given in Table 1.

Note that all these steps are conducted for each age group, and all the distances are based on roadway network. . It is worth mentioning that Wang and Langford (Langford et al., 2012; Wang, 2012) suggest that it is unclear which function is the most appropriate to use as a distance-decay function without any empirical evidence, therefore it remains a matter of choice. Wang (Wang, 2012) defines six different distance decay functions. However, in this study, locations of crashes and crash occupants' residences are utilized to estimate a case-specific function. As seen from Equation (2), this study employs a continuous distance decay function to define G (decay weights). This type of approach can help us differentiate the crash injury exposure inside the catchment threshold network distance. Fig. 3 shows the empirical CDF, complementary CDF (survival function) and the decay weights used to calculate the crash injury exposure for different age groups based on census blocks and severe crash hotspots.

Note that the entire analysis is executed for different age groups at census block levels. However, all the socioeconomic information used in this paper (ACS, 2010; FGDL, 2016) is provided at a census block group level. For that reason, crash injury exposure for each census block group is calculated based on different age groups by using the population-weighted averages. The following section will continue with the explanation of the proposed Socioeconomic-based Crash Injury Exposure Measure (SECIE).

2.4. Socioeconomic-based crash injury exposure measure

In this paper, a socioeconomic-based crash injury exposure measure (SECIE) is derived for each age group by weighting the crash injury exposure (A_i) measures obtained from the EG-2SFCA method, and the specific socioeconomic data associated with each census block group. This socioeconomic data, based on American Community Surveys (ACS) of the U.S. Census (ACS, 2010; FGDL, 2016), is attached to census block groups, and includes information on the ethnicity, poverty, vehicle ownership and education levels of residents. Using aggregation, the SECIE measure will help to compare census tracts and counties in terms of their crash injury exposure. SECIE uses the concept of the weighted average as

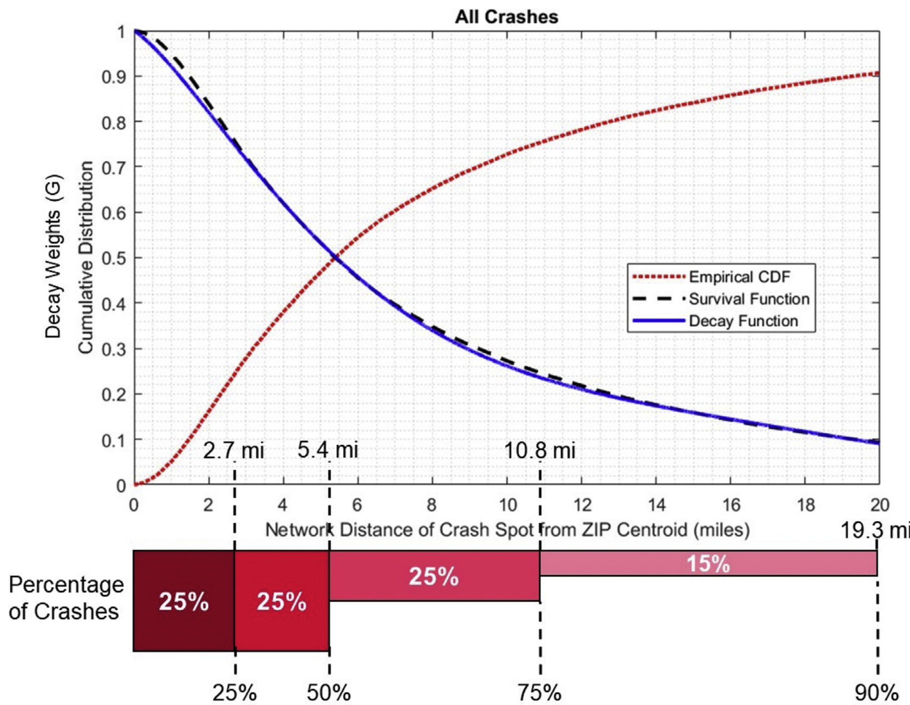


Fig. 3. The empirical CDF, complementary CDF (survival function), and the decay function.

Table 1
Decay function parameter values.

Parameter Values	$a_1 = -36.05$	$b_1 = -0.1759$	$c_1 = 12.71$	$k = 0.063$
	$a_2 = 37.15$	$b_2 = -0.443$	$c_2 = 12.88$	

follows:

$$SECIE_z = \frac{\sum_{i=1}^{n}\left(A_{a,i,z}*Pop_{t,i,z}\right)}{\sum_{i=1}^{n}\left(Pop_{t,i,z}\right)} \tag{3}$$

where $A_{a,i,z}$ is the crash injury exposure of specific age group, "*a*" living at census block group "*i*" pertaining to census tract "*z*", and $Pop_{t,i,z}$ is the number of people living in the census block group "*i*" pertaining to census tract "*z*", "*t*" is the socio economic attribute that SECIE is executed for. For instance, $Pop_{poverty,i,z}$ represents the number of people living below the poverty level in census block group "*i*" pertaining to census tract "*z*". It is important to mention that US census provides the total number of 65+ (65 and older people) but not the African-American age 65+ population in a census block group. SECIE provides a way to incorporate those two attributes. Note that the census tract denoted by the subscript "z", can simply be thought of as a more aggregate spatial unit than the underlying calculations are based (i.e. block groups). In this analysis, "z" can also be used as the designation for counties to provide a SECIE index for that level of aggregation.

The final product gives a crash injury exposure score for census tracts or counties based on each socioeconomic-age group rather than simply different age groups. Note that SECIE is not an algorithm but rather a weighted mean. In a weighted average, some of the data points contribute more than others to the final average. Note that this calculation is performed for each type of age group separately. The socioeconomic-based crash injury exposure measure (SECIE) leads to a weighted "fairness measure" controlling for socioeconomic groups depending on their residential location, and finally the total number of people in each of these groups. This metric also acts as an indicator to quantify and compare social sustainability between different regions (Vallance, Perkins, & Dixon, 2011) That is, if a census tract or county possesses an elevated crash injury exposure, this can be considered as socially unsustainable for residents. These results are potentially beneficial for agencies and stakeholders seeking to effectively allocate resources to people facing elevated risks.

3. Results

In this section, results of our application to District 7 of Florida are presented. First, severe crash hotspots were identified for different age groups. After that, the crash injury exposure of the regions was analyzed, and GIS-based maps were produced. This type of visualization approach can offer a better understanding of areas subject to greater risk. Finally, socioeconomic data were used to create the weighed crash injury exposure maps for each age group.

3.1. Hotspot identification

The hotspot analysis was conducted for 2013–2014 crash data using roadway network distances. Results of the hotspot analysis illustrate that even though hotspots of different age groups have slightly similar spatial patterns, there is still noteworthy spatial variation between severe crashes of different age groups (Fig. 4). That is, a strong visual spatial correlation between hotspot distributions of the age groups under consideration cannot be observed. Nevertheless, it is obvious that all age groups have a number of hotspots in the western Pasco County, parallel to the coast where US-19 lies alongside. Besides that, the hotspots associated with the 22–64 age group are more concentrated around the City of St. Petersburg as well as eastern section of Pasco County than any other age group. However, fewer number of 22–64 age group hotspots were identified in the northern of District 7. Hotspots of

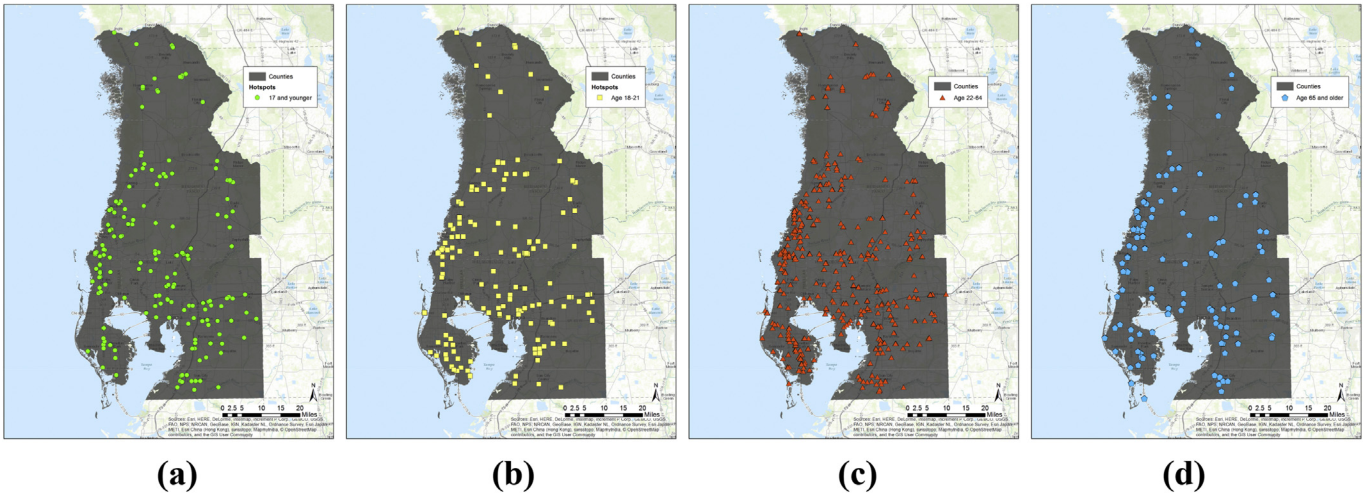


Fig. 4. Severe crash hotspots for different age groups, (a) 17-, (b) 18–21, (c) 22–64, (d) 65+.

the 18–21 age group, on the other hand, are clustered in St. Petersburg in southern Pinellas County, and particularly on the middle sections of Hernando County. Likewise, different patterns are observed for hotspots of 17 and younger and the 65 and older age groups as well. For example, four distinct locations appear to contain areas where age 65+ hotspots are concentrated: South of Hillsborough County, coastal parts of Hernando County, Largo City in Pinellas County, and Citrus County in general. Note that there is an obvious cluster consisting of the 17- hotspots in the northeast of Tampa. This is important because there is no other such high concentration of hotspots in proximity of that region. It can be argued that special attention should be given to those crashes at that location in order to identify the probable reasons for the severe injuries to younger roadway users. This clear difference in the locations of crash hotspots implies that there is a need to focus on individual age groups rather than the overall population.

3.2. Two step floating catchment area method

When 2FSCA method is applied to crashes, the emphasis is not on the accessibility but crash injury exposure. The exposure measure for census blocks (centroids) are illustrated in Fig. 5 whereas Fig. 6 shows the population weighted exposure measure for census block groups. Note that there are highly similar spatial patterns in Figs. 5 and 6. Thus, the discussion is only made for census block groups (note that census block groups are actually groups of census blocks). According to Fig. 6, for the 17- populations, the west and east sections of mid-District 7 and around northeastern parts of Tampa as well as southernmost Hillsborough County are shown to be exposed to elevated injury risk based on the standard deviation on the mean classification approach. Since this is an occupant-based study, families with children seem to be experiencing high crash injury risk around those locations where generally 17- groups populate densely such as southernmost Hillsborough and eastern Pasco (Fig. 2). For the 18–21 years old residents, who are often times college students, crash injury exposure is experienced more on the northwest section of District 7. Moreover, an elevated crash injury risk exists in the whole middle section of District 7, where Pasco County is, for this young age group as evidenced by the observed high deviation from the mean exposure value. One reason for this finding might be the presence of several colleges such as the Pasco-Hernando State College, which in turn exposes college age group occupants to severe injury risk.

The crash injury exposure map of working class (22–64), on the other hand, shows that the residents of this group experience elevated crash injury exposure at western Pasco County. These areas are usually the locations where businesses and work offices are commonly present, which attracts working class people to reside in these regions. This finding, combined with higher number of severe crashes involving working class occupants, indicates that these regions are prone to higher crash injury exposure. For the 65 and older people (65+), there are three areas with high crash injury exposure: the southeast of District 7, and the northwest section along the US-19 and US-98 corridor, where a substantial number of 65+ residents is located. Note that, in Fig. 2, the number of 65+ residents living in each census block group is given. Fig. 2 shows that the census block groups with higher number of aging residents correspond to the findings shown in Fig. 6-d with slight differences (e.g. northernmost Citrus County). The reason of this slight difference is because Fig. 6 illustrates the integrated effect of both population and the severe crash risk (hotspots). Therefore, while there is a high number of 65+ residents in the northernmost part of District 7, that region is not the one of the higher exposed areas based on the mapping classification approach. It is worth mentioning that the most problematic region in terms of crash injury exposure, for all age groups, is the western Pasco County where one finds US-19 and US-98. This information might inform transportation agencies in their efforts to enhance public safety in that region since it is critical for all age groups.

3.3. Application of the socioeconomic-based crash injury exposure measure

In this section, socioeconomic data were integrated into the analysis using the 'SECIE' index, in order to create age-specific weighted crash injury exposure maps primarily for each U.S. census tract (Figs. 7 and 8 show example maps), and then for each county in the study area. Using this approach, for example, the counties and census tracts can be compared to each other in District 7 based on their associated crash injury exposure and socioeconomic variables.

Several important outcomes related to the socioeconomic characteristics of people in the region are observed. For instance, for 17 and younger African American populations living in the eastern and western section of mid-District 7, a high deviation from mean in crash injury exposure is observed, which is similar to all 17

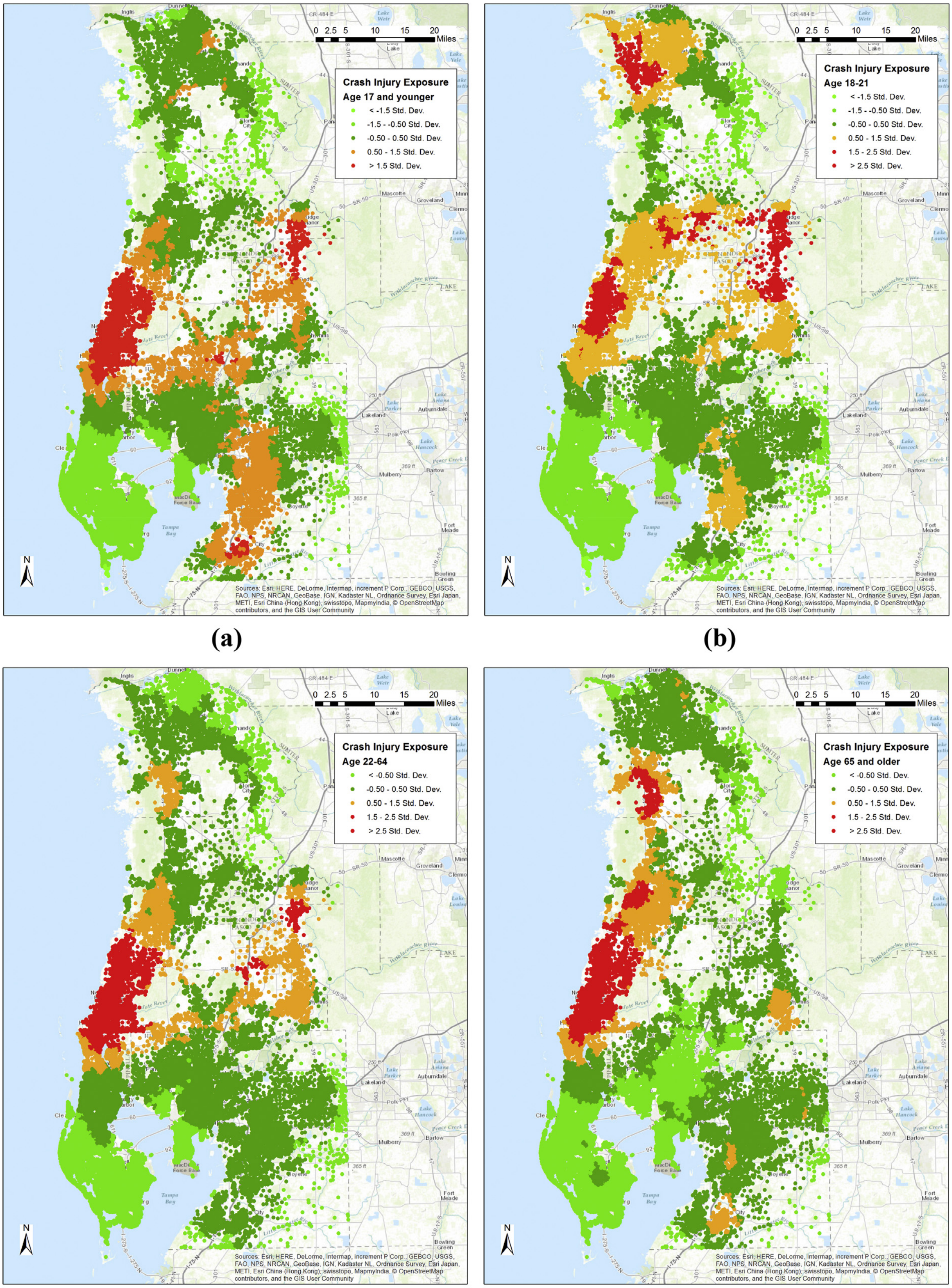


Fig. 5. Crash injury exposure maps for different age groups at census block (centroids) level a) 17- b) 18–21 c) 22–64 and d) 65+.

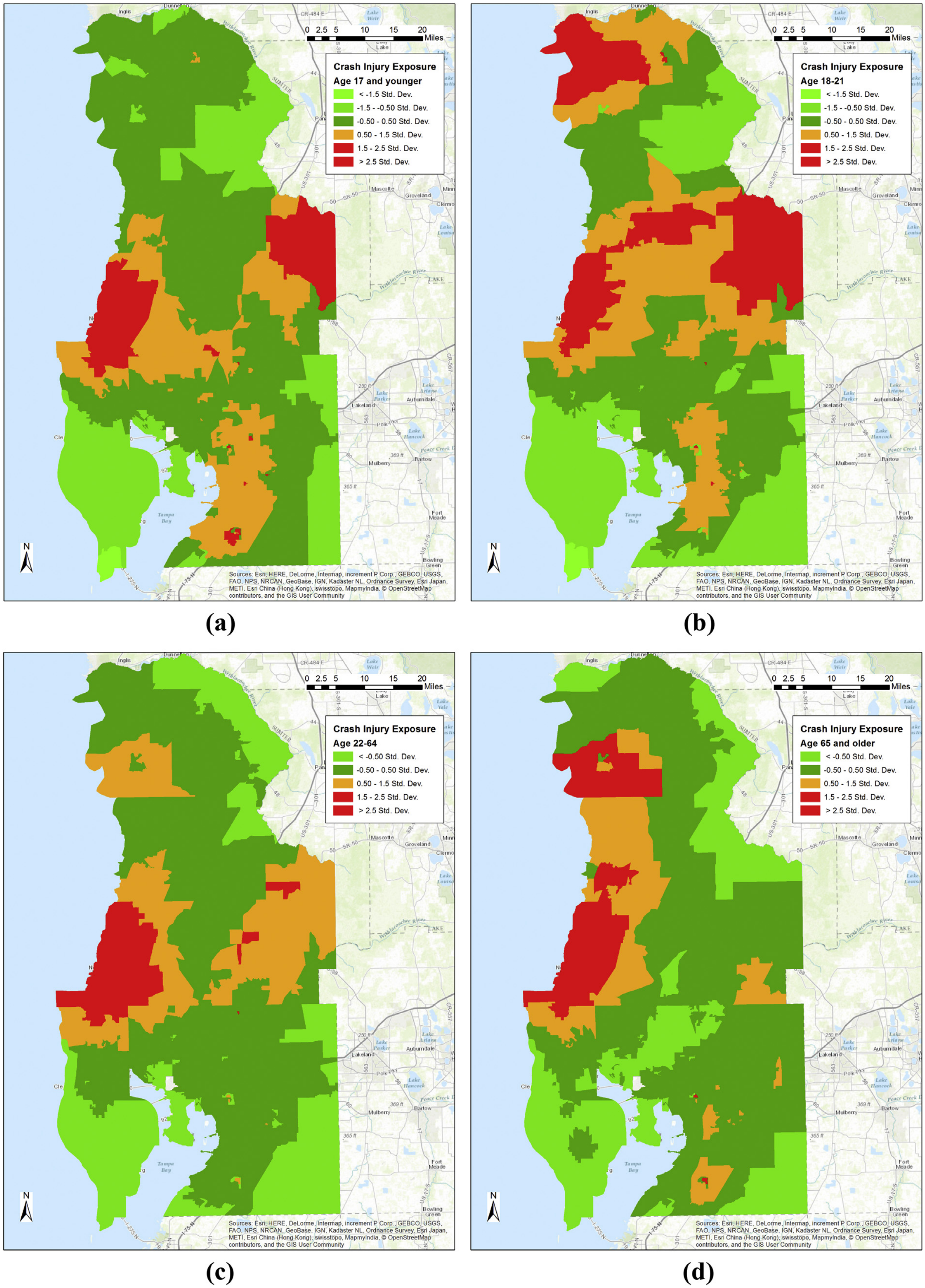


Fig. 6. Crash injury exposure maps for different age groups at census block group level a) 17- b) 18–21 c) 22–64 and d) 65+.

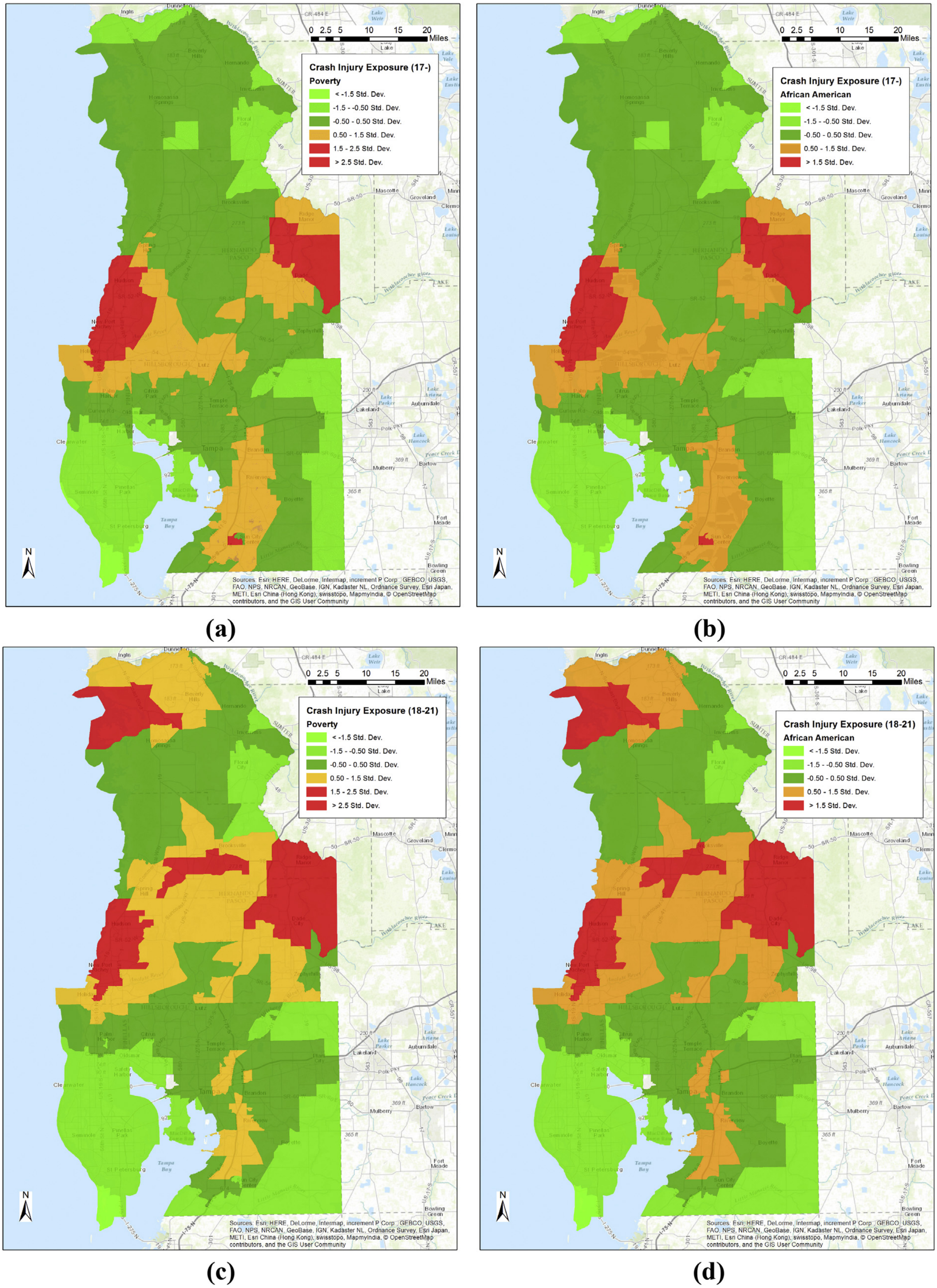


Fig. 7. Socioeconomic characteristics-based crash injury exposure maps for age groups at census tract level (a) 17 and younger for below poverty level, (b) 17 and younger for African American populations, (c) 18–21 for poverty level, and (d) 18–21 for African American populations.

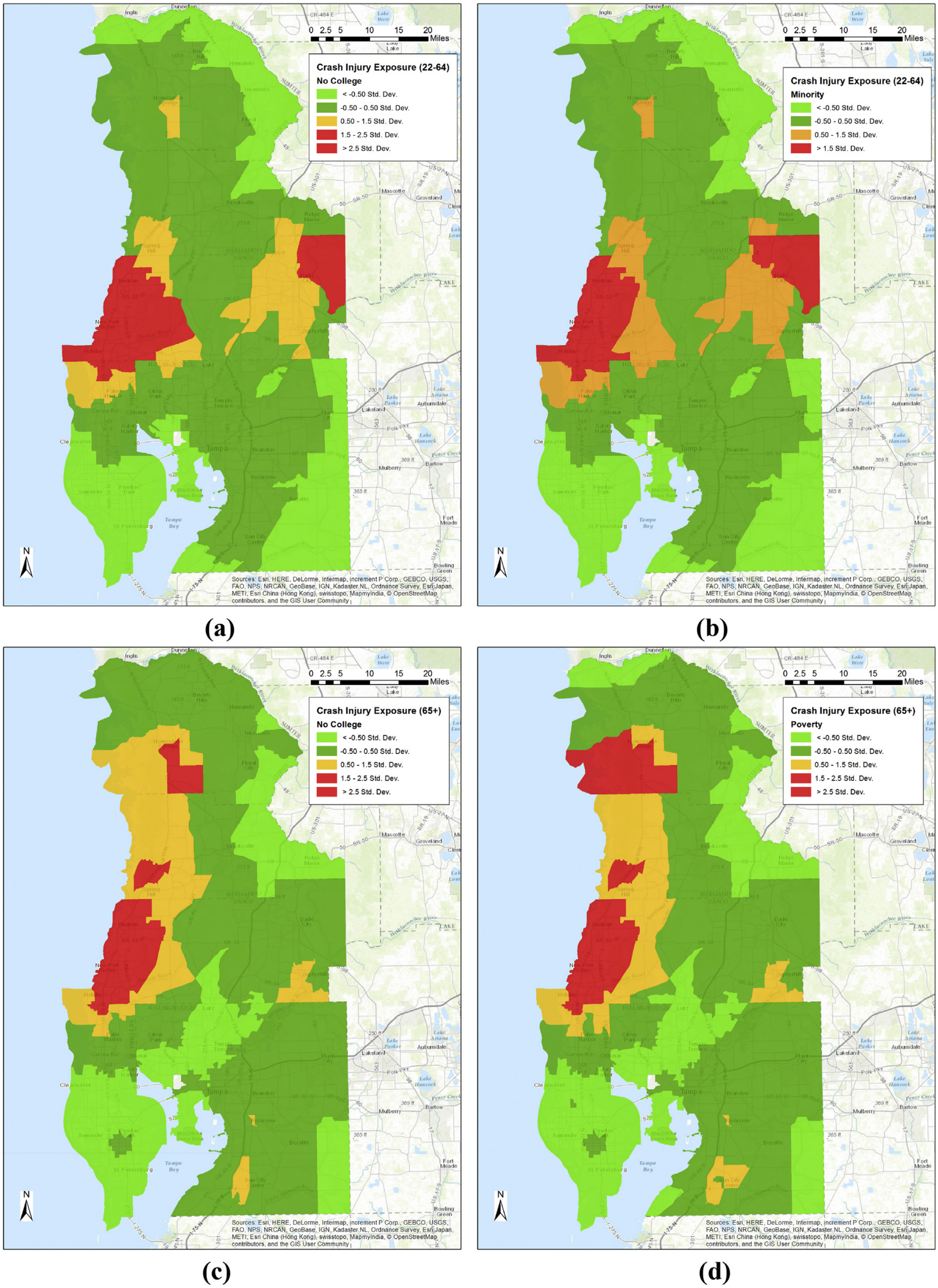


Fig. 8. Socioeconomic characteristics-based crash injury exposure maps for age groups at census tract level (a) 22–64 with education level, (b) 22–64 for minority populations, (c) 65+ with education level, and (d) 65+ with poverty level.

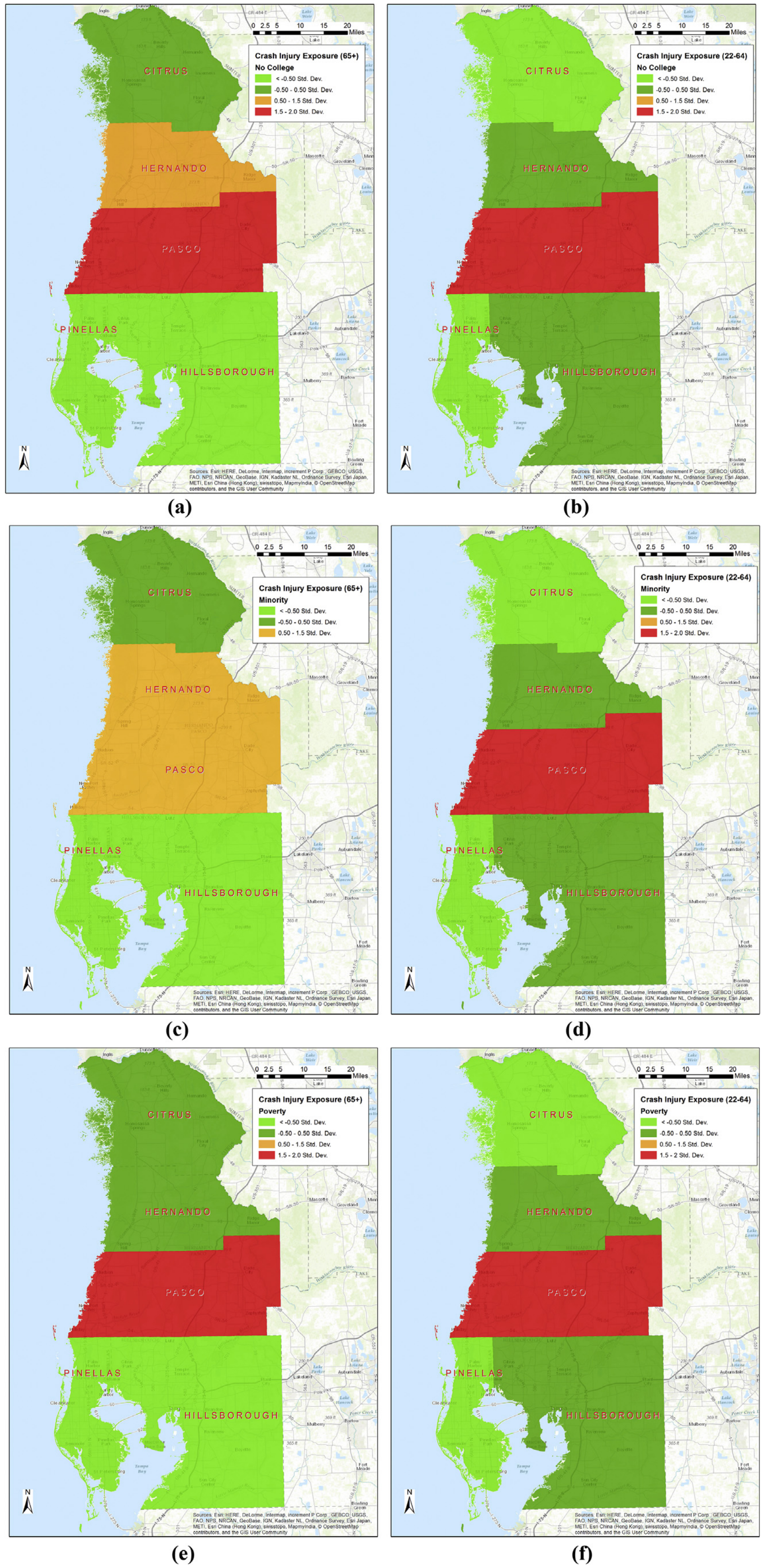


Fig. 9. County-based crash injury exposure ranking maps for the age groups of 22–64 and 65+, respectively: a) – b) education level, c) – d) minority, e) – f) poverty.

and younger populations below poverty (Fig. 7-a). The college age group (18–21) that is below poverty level and from African-American decent, on the other hand, are more exposed to crash injury in the middle as well as northwestern sections of District 7 (Fig. 7c and d). When this finding is incorporated with the findings of Males (2009), who showed that joint effect of age and income-related attributes is the main contributor of young driver fatalities, these identified regions can be considered as highly critical for preventing younger population fatalities. Note that, spatial differences between different locations where these population groups live is the main reason behind this elevated exposure. This knowledge is important since agencies could take countermeasures specific to that location to mitigate injury risk exposed on that socioeconomic group. For the working age group (22–64), a substantial difference is observed when the weight of education and ethnicity (Fig. 8a and b) is evaluated compared to the overall crash injury exposure findings for this group (Fig. 6-c). The identified spatial differences (such as eastern and middle Pasco County) are critical since fatalities were found to be increasing by decreasing education levels (Stamatiadis & Puccini, 2000). Moreover, there is a considerable difference in the effect of being a minority population on crash injury exposure compared to the effect of education level. That is, there is a shift in being exposed to crash injury along the east-west direction (in Pasco County) when the minority dimension is considered instead of that of education level. Additionally, for the 65+ age group, the effect of education and poverty seem to provide substantially different crash injury exposure maps. For instance, crash injury exposure of the northwest section of District 7 increases with poverty. This shows that people belonging to this age group, particularly ones that live below the poverty level, might be exposed to an elevated crash injury risk around the northwestern section of District 7. It is also worth mentioning that, crash injury exposure maps for the census block groups are quite similar to SECIE maps (poverty and African-American) for 17- (Figs. 6a and 7a-b) as well as those for the 18–21 age groups (Figs. 6b and 7c-d). Crash injury exposure maps and the SECIE maps for populations 22–64 (Figs. 6c and 8a-b) and 65+ (Figs. 6d and 8c-d), on the other hand, are quite different from each other.

Note that, socioeconomic measures were carried out in order to identify which groups are unfairly exposed to crash injury risk. That is, for instance, if there is a highly populated African-American community in proximity to a severe crash hotspot, this is an environmental justice issue, which is different than claiming that members of this community are involved in more crashes compared to others. Groups higher in number for a specific block group, due to the weighting function inherent to the model, ultimately will be exposed more risk of being injured than other socioeconomic groups, geographically speaking. In that sense, the crash injury exposure measure is particularly a “fairness measure”. For any geographical unit at any level studied in the paper, the approach estimates an exposure measure geographically based on socioeconomics and accessibility.

When the same methodology was applied to calculate the weighted crash injury exposure for each county, maps like those in Fig. 9 were obtained. In this section, comparison maps for the 22–64 and 65+ age groups are presented. The SECIE approach is further extended to derive a crash injury exposure and compare counties, accounting for the socioeconomic characteristics and crash hotspots contained therein. Results indicate that there is a spatial variation in the “county with highest exposure to crash injury risk” for people attended high school at most (no college) according to the previously selected standard deviation from mean classification approach. For instance, Hernando and Pasco counties are the ones posing the highest crash injury risk for the 65+ age group with no college education by deviating significantly from mean compared to other counties in the region. On the other hand, in Pasco County, the crash injury exposure for the age 22–64 class with no college education is significantly higher with values 0.5 to 1 standard deviation above the mean compared to other counties. For the 65+ age group including the minority population, like the no college education case, Hernando and Pasco County stands out among others with a 0.5 and higher standard deviation from mean. For the age 22–64 age group with including the minority population, Pasco County shows the highest deviation from the mean compared to others, which indicates that the post-graduation and pre-retirement age group, who are also minorities, are exposed to the highest crash risk in that region. When investigating the data based on poverty levels, Pasco County has the highest deviation from mean in terms of crash injury exposure for both the 65+, and the 22–64 age groups. Note that different age groups are exposed to different level of crash exposure depending on the same socioeconomic variable investigated. This indicates that there is spatial variation based on crash exposure of different age groups and socio economics. To conclude, for all three classes (no college education, poverty, and minority) and for both 65+ and 22–64 age groups, Pasco County as appears to be a region of interest for transportation agencies in order to address and investigate the crash exposure risks for those populations.

4. Conclusions and future work

The objective of this study is to investigate the exposure of different population groups to severe injury crash hotspots using an empirical-Gaussian two-step floating catchment area (EG-2SFCA) method based on roadway network distances and a socioeconomic-based weighting approach. In other words, the study examines the proximity of residents living in neighborhoods to severity-weighted crash hotspots (regardless of the prevailing traffic conditions) rather than identifying roadway sections posing a high relative crash risk or having unexpectedly high numbers of crashes with respect to their overall traffic volume. This is performed by developing a special form of crash-to-population ratio that incorporates the severe crash hotspots and population. Results indicate that crash injury exposure changes with respect to age, ethnicity, education and poverty level, and vehicle ownership in the study area. Identifying regions with elevated crash injury exposure can allow transportation agencies to more effectively pinpoint locations to be addressed, and allocate necessary resources to reduce risks to the public. This type of methodology can be easily implemented by transportation planning agencies elsewhere to identify problematic areas. For instance, the most problematic subarea, for all age groups, in terms of crash injury exposure is the western Pasco County where US-19 and US-98 lie along. This information might be very important for transportation agencies in order to enhance the safety of public in that region since it is critical for all age groups.

There are several limitations associated with this study that should be noted. While the empirical-Gaussian 2SFCA implementation provides a mathematical result for any demographic and geographic unit like a census tract or census block group, it produces outputs which are best utilized in a comparative fashion among various spatial units, rather than a unit-specific index which is substantively meaningful on its own. The purpose of this study is to develop a new approach in order to quantify possible crash exposure on the population from a residential perspective, which is a novel perspective in the transportation safety field. The work truly is on development of an index; however, a quantitative validation of findings herein is a natural next step for future research. Validation might entail regressing the metric to determine if the residential areas identified as being high risk do contain people who have been in crashes, after controlling for other factors. Another future direction can be a more thorough categorization of roadways while identifying crash hotspots, such as differentiating between freeways, arterials and other roadways in addition to the urban versus rural designation. Lastly, given the nature of the work, it would be potentially fruitful to interact with governmental agencies and experts such as city DOTs, Council of Governments (COGs), or Metropolitan Planning Organizations (MPOs) regarding the results of this research. This will facilitate the discussion of possible remedial measures that can be used to assess the various challenges presented in this paper under the varying risk of being injured based on statistically significant clusters of injured vehicle occupants in crashes.

Acknowledgments

This research was partly supported by U.S. NSF award 1640587 and United States Department of Transportation grant DTRT13-G-UTC42, and administered by the Center for Accessibility and Safety for an Aging Population (ASAP) at the Florida State University (FSU), Florida A&M University (FAMU), and University of North Florida (UNF). The opinions, results, and findings expressed in this manuscript are those of the authors and do not necessarily represent the views of aforementioned institutions.

References

Abdel-Aty, M., Chen, C. L., & Radwan, E. (1999). Using conditional probability to find driver age effect in crashes. *Journal of Transportation Engineering, 125*(6), 502–507. https://doi.org/10.1061/(ASCE)0733-947X(1999)125:6(502).

ACS. (2010). *American community Survey by U.S. Census*.

ArcGIS. (2016). *Geoprocessing tools*.

Bell, S., Wilson, K., Bissonnette, L., & Shah, T. (2013). Access to primary health Care: Does neighborhood of residence matter? *Annals of the Association of American Geographers, 103*(1), 85–105. https://doi.org/10.1080/00045608.2012.685050.

Blatt, J., & Furman, S. M. (1998). Residence location of drivers involved in fatal crashes. *Accident Analysis and Prevention, 30*(6), 705–711. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/9805513.

Blazquez, C. A., & Celis, M. S. (2013). A spatial and temporal analysis of child pedestrian crashes in Santiago, Chile. *Accident Analysis and Prevention, 50*, 304–311. https://doi.org/10.1016/j.aap.2012.05.001.

Boyce, T. E., & Geller, E. S. (2002). An instrumented vehicle assessment of problem behavior and driving style: Do younger males really take more risks. *Accident Analysis and Prevention, 34*(1), 51–64. https://doi.org/10.1016/S0001-4575(00)00102-0.

Burdett, B. R. D., Starkey, N. J., & Charlton, S. G. (2017). The close to home effect in road crashes. *Safety Science, 98*, 1–8. https://doi.org/10.1016/j.ssci.2017.04.009.

Cartenì, A. (2014). Accessibility indicators for freight transport terminals. *Arabian Journal for Science and Engineering, 39*(11), 7647–7660. https://doi.org/10.1007/s13369-014-1333-y.

Citilabs. (2016). *Cube software*.

Dai, D. (2010). Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. *Health & Place, 16*(5), 1038–1052. https://doi.org/10.1016/j.healthplace.2010.06.012.

Dai, D. (2012). Identifying clusters and risk factors of injuries in pedestrian–vehicle crashes in a GIS environment. *Journal of Transport Geography, 24*, 206–214. https://doi.org/10.1016/j.jtrangeo.2012.02.005.

Dai, D., & Wang, F. (2011). Geographic disparities in accessibility to food stores in southwest Mississippi. *Environment and Planning B: Planning and Design, 38*(4), 659–677. https://doi.org/10.1068/b36149.

District 7 Office. (2015). *Tampa Bay regional planning model*.

District 7 Office. (2016). *Tampa Bay strategic freight plan*.

Dony, C. C., Delmelle, E. M., & Delmelle, E. C. (2015). Re-conceptualizing accessibility to parks in multi-modal cities: A variable-width floating catchment area (VFCA) method. *Landscape and Urban Planning, 143*, 90–99. https://doi.org/10.1016/j.landurbplan.2015.06.011.

ESRI. (2014). *ArcGIS Desktop: Release 10.2*. Redlands, CA: Environmental Systems Research Institute.

FDOT. (2015). *Traffic data*.

FGDL. (2016). *Florida geographic data library*.

Getis, A., & Ord, J. K. (1992). The analysis of spatial associatiion by use of distance statistics. *Geographical Analysis, 24*(3), 189–206. https://doi.org/10.1073/pnas.0703993104.

Guagliardo, M. F. (2004). *Spatial accessibility of primary care;: Concepts, methods and challenges* (Vol. 13, pp. 1–13).

Gundogdu, I. B. (2010). Applying linear analysis methods to GIS-supported procedures for preventing traffic accidents: Case study of Konya. *Safety Science, 48*(6), 763–769. https://doi.org/10.1016/j.ssci.2010.02.016.

Horner, M. W., Duncan, M. D., Wood, B. S., Valdez-Torres, Y., & Stansbury, C. (2015). Do aging populations have differential accessibility to activities? Analyzing the spatial structure of social, professional, and business opportunities. *Travel Behaviour and Society, 2*(3), 182–191. https://doi.org/10.1016/j.tbs.2015.03.002.

Joseph, A. E., & Bantock, P. R. (1982). Measuring potential physical accessibility to general practitioners in rural areas: A method and case study. *Social Science and Medicine, 16*(1), 85–90. https://doi.org/10.1016/0277-9536(82)90428-2.

Khan, A. A. (1992). An integrated approach to measuring potential spatial access to health care services. *Socio-economic Planning Sciences, 26*(4), 275–287. https://doi.org/10.1016/0038-0121(92)90004-O.

Khan, G., Qin, X., & Noyce, D. a. (2008). Spatial analysis of weather crash patterns. *Journal of Transportation Engineering, 134*(5), 191–202. https://doi.org/10.1061/(ASCE)0733-947X(2008)134:5(191).

Krahe, B., & Fenske, I. (2002). Predicting Aggressive Driving Behavior: The Role of Macho Personality, Age, and Power of Car. *Aggressive Behavior, 28*(1), 21–29. https://doi.org/10.1002/ab.90003.

Kwan, M.-P. (1998). Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis, 30*(3), 191–216. https://doi.org/10.1111/j.1538-4632.1998.tb00396.x.

Langford, M., Fry, R., & Higgs, G. (2012). Measuring transit system accessibility using a modified two-step floating catchment technique. *International Journal of Geographical Information Science, 26*(2), 193–214. https://doi.org/10.1080/13658816.2011.574140.

Lin, T. (Grace), Xia, J. (Cecilia), Robinson, T. P., Olaru, D., Smith, B., Taplin, J., et al. (2016). Enhanced Huff model for estimating park and ride (PnR) catchment areas in perth, WA. *Journal of Transport Geography, 54*, 336–348. https://doi.org/10.1016/j.jtrangeo.2016.06.011.

Luo, J. (2014). Integrating the Huff model and floating catchment area methods to analyze spatial access to healthcare services. *Transactions in GIS, 18*(3), 436–448. https://doi.org/10.1111/tgis.12096.

Luo, W., & Qi, Y. (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health and Place, 15*(4), 1100–1107. https://doi.org/10.1016/j.healthplace.2009.06.002.

Luo, W., & Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: Synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design, 30*, 865–884. https://doi.org/10.1068/b29120.

Males, M. A. (2009). Poverty as a determinant of young drivers' fatal crash risks. *Journal of Safety Research, 40*(6), 443–448. https://doi.org/10.1016/j.jsr.2009.10.001.

Malleson, N., & Andresen, M. A. (2015). Spatio-temporal crime hotspots and the ambient population. *Crime Science, 4*(1), 10. https://doi.org/10.1186/s40163-015-0023-8.

McGrail, M. R. (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: An assessment of recent improvements. *International Journal of Health Geographics, 11*(1), 50. https://doi.org/10.1186/1476-072X-11-50.

Polo, G., Acosta, C. M., & Dias, R. A. (2013). Spatial accessibility to vaccination sites in a campaign against rabies in S??o Paulo city, Brazil. *Preventive Veterinary Medicine, 111*(1–2), 10–16. https://doi.org/10.1016/j.prevetmed.2013.03.010.

Radke, J., & Mu, L. (2000). Spatial decompositions, modeling and mapping service regions to predict access to social programs. *Geographic Information Sciences, 6*(2), 105–112.

Saliba, D., Buchanan, J., & Kington, R. S. (2004). Function and response of nursing facilities during community disaster. *American Journal of Public Health*. https://doi.org/10.2105/AJPH.94.8.1436.

Stamatiadis, N., & Puccini, G. (2000). Socioeconomic descriptors of fatal crash rates in the Southeast USA. *Injury Control and Safety Promotion, 7*(3), 165–173. https://doi.org/10.1076/1566-0974(200009)7:3;1-N;FT165.

Steenberghen, T., Aerts, K., & Thomas, I. (2010). Spatial clustering of events on a network. *Journal of Transport Geography, 18*(3), 411–418. https://doi.org/10.1016/j.jtrangeo.2009.08.005.

Thomas, I. B. C., Hermia, J.-P., Vanelslander, T., & Verhetsel, A. (2003). Accessibility to freight transport networks in Belgium: A geographical approach. *Tijdschrift Voor Economische En Sociale Geografie, 94*(4), 424–438. https://doi.org/10.1111/1467-9663.00271.

U.S. Census Bureau. (2010). *2010 TIGER/Line shapefile*.

Ulak, M. B., Ozguven, E. E., Spainhour, L., & Vanli, O. A. (2017). Spatial investigation of aging-involved crashes: A GIS-based case study in northwest Florida. *Journal of Transport Geography, 58*, 71–91. https://doi.org/10.1016/j.jtrangeo.2016.11.011.

Vallance, S., Perkins, H. C., & Dixon, J. E. (2011). What is social sustainability? A clarification of concepts. *Geoforum, 42*(3), 342–348. https://doi.org/10.1016/j.geoforum.2011.01.002.

Van den Heuvel, F. P., Rivera, L., van Donselaar, K. H., de Jong, A., Sheffi, Y., et al. (2014). Relationship between freight accessibility and logistics employment in US counties. *Transportation Research A: Policy and Practice, 59*, 91–105. https://doi.org/10.1016/j.tra.2013.11.002.

Wang, F. (2012). Measurement, optimization, and impact of health care accessibility: A methodological review. *Annals of the Association of American Geographers. Association of American Geographers, 102*(5), 1104–1112. https://doi.org/10.1080/00045608.2012.657146.

Wan, N., Zou, B., & Sternberg, T. (2012). A three-step floating catchment area

method for analyzing spatial access to health services. *International Journal of Geographical Information Science, 26*(6), 1073–1089. https://doi.org/10.1080/13658816.2011.624987.
Widener, M. J., Farber, S., Neutens, T., & Horner, M. (2015). Spatiotemporal accessibility to supermarkets using public transit: An interaction potential approach in Cincinnati, Ohio. *Journal of Transport Geography, 42*, 72–83. https://doi.org/10.1016/j.jtrangeo.2014.11.004.
Xie, Z., & Yan, J. (2013). Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: An integrated approach. *Journal of Transport Geography, 31*, 64–71. https://doi.org/10.1016/j.jtrangeo.2013.05.009.