A COMPARATIVE ANALYSIS OF SOCIAL VULNERABILITY MAPPING TECHNIQUES FOR TRAVIS COUNTY, TEXAS

by

Mogahid Adam Ahmed Hussein, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Geography
May 2018

Committee Members:

Ronald R. Hagelman III, Chair

Tzee-Kiu Edwin Chow

Russell Weaver

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGEMENTS

I am grateful to all of those whom I have had the pleasure to work with during this and other related projects. Each of the members of my thesis committee has provided me extensive personal and professional guidance and taught me a great deal about the scientific research and life in general. I would like especially to thank Dr. Ron Hagelman, the chair of my committee. As my teacher and mentor, he has taught me more than I could ever give him credit for here. He has shown me, by his example, what a good scientist (and person) should be.

Nobody has been more important to me in this project than the members of my family. I would like to thank my parents; whose love and guidance are with me in whatever I pursue. They are the ultimate role models.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

LIST OF TABLES vi

LIST OF FIGURES vii

LIST OF ABBREVIATIONS viii

CHAPTER

1. INTRODUCTION AND PROBLEM STATEMENT 1
2. PURPOSE OF RESEARCH 7
3. LITERATURE REVIEW 10
4. DATA AND METHODS 15
5. RESULTS 17
6. LIMITATIONS OF STUDY 31
7. CONCLUSION 33

REFERENCES 35

LIST OF TABLES

Table **Page**

1. Selected social vulnerability indicators/variables 3

2. Social vulnerability index construction stages and options 6

3. Floodplain overlay analysis for additive aggregation 26

4. Floodplain overlay analysis for Pareto ranking aggregation 27

LIST OF FIGURES

Figure **Page**

1. A conceptual framework that explains the steps of this research 14

2. Overall social vulnerability in Travis County, Texas calculated using additive aggregation of the eight standardized social vulnerability variables 18

3. Overall social vulnerability in Travis County, Texas calculated using Pareto ranking of the eight standardized social vulnerability variables 20

4. Different classification schemes using the same data 23

5. Additive aggregation Vulnerability Index map with 25-years, 100-years, and 500-years floodplains 24

6. Pareto ranking Vulnerability Index map with 25-years, 100-years, and 500-years floodplains..25

7. Overall social vulnerability of block groups in Travis County, TX: top image represents additive aggregation of indicators; bottom image represents Pareto ranking 28

8. Overall social vulnerability of block groups in Travis County, TX: top image represents additive aggregation of indicators; bottom image represents Pareto ranking 29

9. Overall social vulnerability of block groups in Travis County, TX: top image represents additive aggregation of indicators; bottom image represents Pareto ranking 30

LIST OF ABBREVIATIONS

Abbreviation	**Description**
SOVI	Social Vulnerability Index
FEMA	Federal Emergency Management Administration

1. INTRODUCTION AND PROBLEM STATEMENT

According to the Intergovernmental Panel on Climate Change (IPCC), human action, primarily the burning of fossil fuels, is causing the earth's climate to change. It is a virtual certainty that the frequency and intensity of extreme weather events will increase with warming atmosphere. In a working group report for the IPCC's 5th Assessment, the group reported to policymakers that, "continued emission of greenhouse gases will cause further warming and long-lasting changes in all components of the climate system, increasing the likelihood of severe, pervasive and irreversible impacts on people and ecosystems…" (IPCC 2014). These changes, though global in nature, will manifest at regional, continental and local scales. It follows that successful adaptation and mitigation will need to be effectuated at a variety of spatial scales as well.

Although it is difficult to precisely predict these changes at the geographic scale of a city or U.S. county, recent analysis contracted by the City of Austin from Katharine Hayhoe at Texas Tech University projected increasing variability in annual and seasonal average temperatures, more frequent high-temperature extremes, more frequent extreme precipitation, and more frequent drought conditions due to hotter summer weather (Hayhoe 2014). These results indicated a need for further climate adaption and extreme weather mitigation by city and county managers.

Extreme weather events that lead to flooding like the Memorial Day Flood in 2015 and wildfire like the Bastrop Fire of 2011 due to climate variability will pose a challenge to the city of Austin and the surrounding communities. Previous research indicates that all populations are affected by the impacts of climate change, but some communities bear a greater burden than others do (Cutter and Emrich 2006; Zahran, S. et al. 2008; IPCC 2014). To aid hazards researchers, urban managers and other decision makers, communities can be chracterized by their demographic or *social* vulnerabilities to environmental hazards. Social vulnerability refers to how likely it is that a population will be adversely affected by an extreme event (Wisner et al. 1994; Cutter and Emrich 2006; Flanagan et al. 2011). Researchers have recognized that vulnerability involves demographic and socio-economic factors that affect a community's resilience or ability to bounce back from a climate weather event like a flood or wildfire. (Wisner et al. 1994). Despite the perceived utility and popularity of social vulnerability assessments, there are no standard practices for the methods employed to create them. Recent research has focused attention on the choice of variables, degree of practitioner involvement, and indexing methods (Rygel, L. et al. 2006; Tate 2012; Oulahen, G. et al. 2015; Araya-Munoz et al. 2017).

First, this research examines the literature on social vulnerability to identify the key demographic and socio-economic variables that contribute to social vulnerability. A list of variables was developed drawing on the multidimensional characteristics that adversely affect the ability of communities to recover and bounce back in pre-disaster and post-disaster events. Second, this list was presented to urban managers in the study area at the Office of Sustainability at the City of Austin. Through consultation with city staff, eight socioeconomic U.S. Census block group-scale variables were selected as being the most relevant to Austin's current climate management goals (Table 1). In previous research, these variables are often selected through statistical analysis, rather than leveraging expert opinion. Although not the primary intent of this research, using expert opinion to focus the list of indictors is an emerging approach in social vulnerability scholarship that warrants further exploration on future research (Oulahen, G. et al. 2015). This research will use two of the most prevalent methods of aggregation for eight standardized variables of social vulnerability. These social vulnerability indices (SOVI) will be mapped for Travis County using U.S. Census block groups as the unit of analysis. The goal of this research is to discover the visual and quantitative differences between the two SOVI methods. This analysis has the potential to aid in the further refinement of SOVI scholarship and to assist City of Austin and Travis County managers in their efforts to improve social resilience to extreme events in their respective jurisdictions.

Table 1. Selected social vulnerability indicators/variables

Demographic and Socio-Economic Indicators	**References from Previous Studies**
Percent of population five years and under	Cutter 2003, 2008; Chakraborty 2005; Flanagan 2011; Holand 2013; Lixin 2014
Percent of population 65 years and older	Cutter 2003, 2008; Chakraborty 2005; Flanagan 2011; Holand 2013; Lixin 2014
Percent of single-parent households	Cutter 2003, 2008; Chakraborty 2005; Flanagan 2011; Holand 2013; Lixin 2014
Median Household Income	Cutter 2003, 2008; Li 2010; Wood 2010; Ghadiri 2013; Armas 2013; Chen 2013; Ge 2013
Percent of population with less than a high school education	Cutter 2003; Chakraborty 2005; Flanagan 2011; Holand 2013; Martins 2012; Chen 2013; Lixin 2014
Percent of population that speaks Spanish as a first language at home	Flanagan 2011; Holand 2011; Chen 2013; Lixin 2014
Percent of households with no vehicle	Flanagan 2011
Percent of housing that is trailers or mobile homes	Armas 2013; Holand 2013

According to a review of previous SOVI research, there are eleven stages that a researcher should consider when constructing an index of social vulnerability. Tate (2012) summarizes the stages (Table 2) and provides a comparative assessment of the various choices within each stage. However, that study did not include an assessment of the aggregation stage in its assessment, and the study recommended that future analysis should investigate alternative aggregation schemes as is proposed here. The *additive aggregation approach*, which has been used in much of the SOVI research to date, assumes that all indicators of vulnerability are equally important. Thus, it sums up the indicators to compute the arithmetic mean. This approach of aggregation has been criticized due to its potential to obscure block groups that score highly in one indicator of vulnerability (Rygel, L. et al. 2006). In addition, indicators of vulnerability are not believed to be equally important in the realm of policymakers due to the variability of geographic spaces (Oulahen, G. et al. 2015). For example, for emergency managers a variable such as households without a vehicle might be of a great concern when evacuation is advised. In previous research, decision-makers found it useful to have individual maps of variables or customized combinations of variables (Oulahen, G. et al. 2015). The second method of aggregation this research will use is called *Pareto ranking algorithm.* Rygel et al. (2006) used this method in Hampton Roads, Virginia to rank block groups in order from high to low based on their *Pareto ranking.* This approach leverages a multi-objective optimization technique based on genetic algorithm in order to increase the dimensionality of the SOVI assessments that assumes that all indicators of vulnerability contribute equally to the construction of an index. Rygel et al. (2006) found that block groups that were moderately vulnerable on the scale of the *additive* *aggregation approach* fell into the highest *Pareto ranks*, indicating that the method selected for a SOVI assessment can have a substantive influence on the outcome of the assessment. This study employs a comparative analysis of SOVI methods in an environmentally and socially dynamic city in order to advance our understanding of how the choice of SOVI ranking methods can influence the outcomes of SOVI mapping exercises and to better inform urban managers who may want to use SOVI mapping as a tool for decision-making. Although multiple approaches have been introduced to construct an index, they do not offer a solution to the criticism of using additive aggregation to assess social vulnerability as proposed by Rygel et al. (2006). For example, Araya-Munoz et al. (2017) used an approach called fuzzy logic aggregation to examine the impact of multiple hazards in a metropolitan area in Chile. The study assessed the impacts of different types of hazards that were analyzed using components that do not share the same attribute range and scale of analysis. However, most research of vulnerability uses indicators that share attributes and have the same scale of analysis, but the challenges are which indicator is more important? Should researcher apply weight to the indicators? Alternatively, should they be treated equally? Most studies assume that the indicators are equally important when constructing an index. Rygel et al. (2006) proposes a multi-objective optimization technique, called Pareto ranking, to avoid such an assumption. This approach works well when several criteria are present simultaneously, and it is not possible or wise to combine these into a single number. When this is the case, the problem is said to be a multi-objective optimization problem (Goldberg 2006). Thus, this study aims at comparing the *additive aggregation* that assumes all indicators are of equal importance and the *Pareto ranking*

Table 2. Social vulnerability index construction stages and options (adopted from Tate 2012)

Stage	**Description**	**Example options**
Conceptual framework	Vulnerability dimensions to include	Access to resources, demographic structure, evacuation, institutional
Structural design	Organization of indicators within the index	Deductive, hierarchical, inductive
Analysis scale	Geographic aggregation level of indicators	US county, census enumeration unit, neighborhood, raster cell size
Indicator selection	Proxy variables for dimensions	Income, education, age, ethnicity, gender, occupation, disability
Measurement error	Accuracy and precision of the demographic data	Census undercounts, reported margin of error
Transformation	Indicator representation	Counts, proportions, density
Normalization	Standardization to common measurement units	Ordinal, linear scaling (min–max, maximum value), z-scores
Data reduction	Reduction of large correlated indicator set to a smaller set	Factor analysis
Factor retention	How many principal components to retain?	Scree plot, Kaiser criterion, parallel analysis
Weighting	Relative degree of indicator importance	Equal, expert, data envelopment analysis, budget allocation, analytic hierarchy process
Aggregation	Combination of normalized indicators to the final index	Additive, geometric, multi-criteria analysis

2. PURPOSE OF RESEARCH

The effect of extreme weather events enhanced by climate change can have a tremendous economic cost in the form of damaged infrastructure and the loss of homes. There may also be a severe social cost in the form of home displacement, injury, and even the loss of life. Previous research has explored a variety of methods and data set aimed at identifying and spatially delineating people most likely to suffer the greatest losses in extreme events. Hazards managers and practitioners have also sought to identify vulnerable populations in order to develop and implement strategies, policies, and programs to make vulnerable communities more resilient (Flanagan et al. 2011; Ge 2017). The purpose of this research is to explore the visual and quantitative differences in two of the more prevalent methods for assessing spatial patterns of social vulnerability, and to implement both using expert input for Travis County and the City of Austin. In order to achieve this purpose, two primarily research questions will be answered. First, of the U.S. Census-based social indicators most often utilized in social vulnerability mapping, which ones are considered most germane to city managers currently working on climate resiliency planning for the City of Austin? This approach to identifying social indicators stands in contrast to much of the social vulnerability research that starts with multiple indicators and then employs statistical techniques, such as principal component analysis, to reduce the number of indicators used in mapping. Using any method, the selection of SOVI indicators is somewhat subjective. This approach allows for a focused list of indicators that most closely match the goals of practitioners working in the study area. Second, what are the visual and quantitative differences between traditional additive approaches to social vulnerability index mapping and a multi-objective optimization approach, specifically the Pareto ranking method for Travis County, Texas?

This research compares the two methods of aggregation visually using ESRI's ArcMap, and applies Spearman's rank correlation in order to verify whether there is a monotonic correlation between the additive aggregation and the Pareto ranking. The Spearman's rank correlation is a nonparametric statistical test of rank that serves as a measure of the strength between two variables and ranges from -1.00 to 1.00. This statistical measure does not require that the data is normally distributed, and it assesses how well a monotonic function between ordered sets can describe the relationship between two variables (Hauke and Kossowski 2011). The null and alternative hypothesis tested here are as follow:

- Ho: There's no significant difference between *additive aggregation* and the *Pareto algorithm aggregation (*i.e., there's a monotonic correlation)
- H1: *Additive aggregation* and the *Pareto algorithm aggregation* are significantly different (i.e., there's no monotonic correlation)

Previous research tends to either weight all variables considered to create the SOVI equally or have an expert opinion to assign a different weight of importance to each one of the indicators (Cutter et al. 2000; Rygel, L. et al. 2006; Oulahen, G. et al. 2015). Rygel, L. et al. (2006) argue that the Pareto ranking algorithm aggregation approach creates a SOVI that does not suffer from the problem of considering that all variables that affect a community's ability to respond and recover from a flood event are equally important. This research adds to the emerging body of scholarship on the differences in SOVI techniques by testing two of the more prevalent approaches in a rapidly urbanizing, hazard-prone city. Previous studies have showed that practitioners prefer individual maps instead of a composite index (Oulahen, G. et al. 2015). They argue that indices have the potential to obscure indicators that are of great importance to their goal management when averaged with low score of vulnerability indicators. Practitioners consider assessing the vulnerability of population to extreme events as a multi-objective optimization problem where optimizing one attribute means worsening another attribute. Therefore, the Pareto ranking approach is worth comparing to the most prevalent additive aggregation approach. Pareto ranking is an approach that has the potential to deal with problems that are considered as multi-objective optimization by ranking block groups based on non-domination in the complete dataset (Rygel et al. 2006).

3. LITERATURE REVIEW

There is extensive evidence that the climate is changing and that adaptation is unavoidable (Berrang-Ford et al. 2011; IPCC 2014). In recent years, there have been many devastating events around the world that adversely affected cities; and caused loss of life. For example, Hurricane Katrina in 2005, the Memorial Day weekend floods in Central Texas in 2015, and Hurricane Harvey in 2017. Those events flooded roads and houses leaving several people dead, missing, injured or displaced. Although it is difficult to ascribe such events to climate change, they have reminded the world that many cities in developed and developing countries are vulnerable to climate variability (Funfgeld 2010). The concept of social vulnerability identifies sensitive populations that may be less likely to respond to, cope with, and recover from a natural disaster. (Cutter 2000; Johnson et al. 2016). As a result of such devastating events, there is an emerging approach to emergency management systems that emphasize strategies to reduce losses through effective mitigation, preparedness, and recovery programs instead of the simple traditional approach of post-event response (Cutter et al. 2000; Hallegate 2009; The City of Austin Resilience Plan 2014).

Previous research indicates that climate change management in cities yield several institutional obstacles that confine the ability of city managers to address climate change risks appropriately. These barriers include the limitation associated with the understanding of rising scientific information about climate change hazards and their impact on cities, in addition to the limited understanding of the complex socio-economic processes that influence urban vulnerabilities to determine the best urban climate adaptation (Funfgeld 2010). Moreover, studies have found that there is a lack of integrating effective integration of information related to hazard exposure and vulnerability to local planning processes and development programs (Fothergill, A. et al. 1999; Oulahen, G. et al. 2015). Traditionally the view of climate governance at the local level is influenced by international agreements and national policies, the preferences of funders, ideas inspired by nongovernmental organizations and transnational networks (Anguelovski and Carmin 2011). However, increasingly most cities have shifted from this traditional view to climate governance strategies that are motivated by internal goals and actions that will advance their climate agendas. For example, the City of Austin is working on developing a resilience plan that will help build a community that is resilient to extreme weather and hazards (The City of Austin Resilience Plan 2014).

To accomplish such a plan, proactive efforts to identify where Austin citizens most vulnerable are located is the first step in developing effective adaptation strategies and programs. This research will assess the social vulnerability in Travis County using the two proposed methods of aggregation. Furthermore, assessing the social vulnerability of populations is challenging in many cases because it requires considering the complex interrelations between socio-economic indicators that significantly influence how cities operate and grow (Funfgeld 2010).

Social vulnerability is regarded as the product of social stratification and social inequalities that exist among different groups of people in each location (Cutter, Boruff and Shirley 2003; Yenneti et al. 2016). Previous research has found that the characteristics that influence social vulnerability include such factors as age, gender, race, employment, level of income, transportation and housing conditions (Cutter, Boruff and Shirley 2003; Lee 2014; Flanagan et al. 2011; Nguyen et al. 2017). In addition, it is important to integrate vulnerable biophysical places with social indicators that define vulnerable population when assessing vulnerability. People within a location bear the burden of a disaster disproportionately because not in all cases do vulnerable biophysical places intersect with the most vulnerable population (Cutter and Emrich 2006; Zahran, S. et al. 2008; IPCC 2014). In some places with high-risk of exposure, the residence will experience a great loss economically but have sufficient resources to overcome a disaster due to insurance or other financial resources which will help to absorb and recover quickly (Cutter et al. 2000). Cutter (1996) developed a hazard-of-place model of vulnerability, in which she argues that the intersection between biophysical vulnerability and social vulnerability is what creates the vulnerability of a place. Government and agencies have utilized the model to visualize the spatial distribution of vulnerable populations at different scales of U.S Census data (Zandt et al. 2012). Building on this model, there have been numerous studies that adopted the approach and created an index that is place-specific (Cutter et al. 2000; Boruff et al. 2005; Oulahen, G. et al. 2015; Araya-Munoz et al. 2017). This research follows the same approach by intersecting the result of the two methods of aggregation with FEMA and Austin watershed department floodplains.

Previous studies have attempted to compare the most common approaches in the literature to construct an index of vulnerability (Tate 2012; Yoon 2012). Findings from such studies show that the most prominent difference of vulnerability indices is the structural design, which includes three approaches: deductive, inductive, and hieratical. The deductive design selects a limited number of variables deductively based on a priori theory and knowledge from the literature review. Cutter et al. (2000) selected eight variables to examine the social vulnerability of populations living in hazard zones in Georgetown County, South Carolina. Wu et al. (2002) used nine variables to assess the social vulnerability of Cape May County, New Jersey. Another study selected only three variables as a proxy to assess social vulnerability (Zahran et al. 2008). This research follows the same approach and deductively chooses eight variables (Table 1) as a proxy to assess social vulnerability. The second design, the inductive approach, aims to create a systematic social vulnerability index using a large set of all possible variables that influence social vulnerability (Boruff et al. 2005; Rygel et al. 2006; Myers et al. 2008; Azar and Rain 2007; Fekete 2009). Finally, the hierarchical designs have selected roughly ten to twenty indicators that are separated into sub-indices that share a common theme of vulnerability (Chakraborty et al. 2005; Flanagan et al. 2011). A detailed assessment of the different approaches that researchers have to select from within each stage (Table 2) of indices construction can be found in (Tate 2012; Yoon 2012). Tate (2012) offers recommendations for each stage of index construction and provides insights for future research in social vulnerability indices so that they can be developed with more robustness and reliability. The following section explains the data and methods used in this research, and (Figure 1) provides an overview of the conceptual framework of the proposed research.

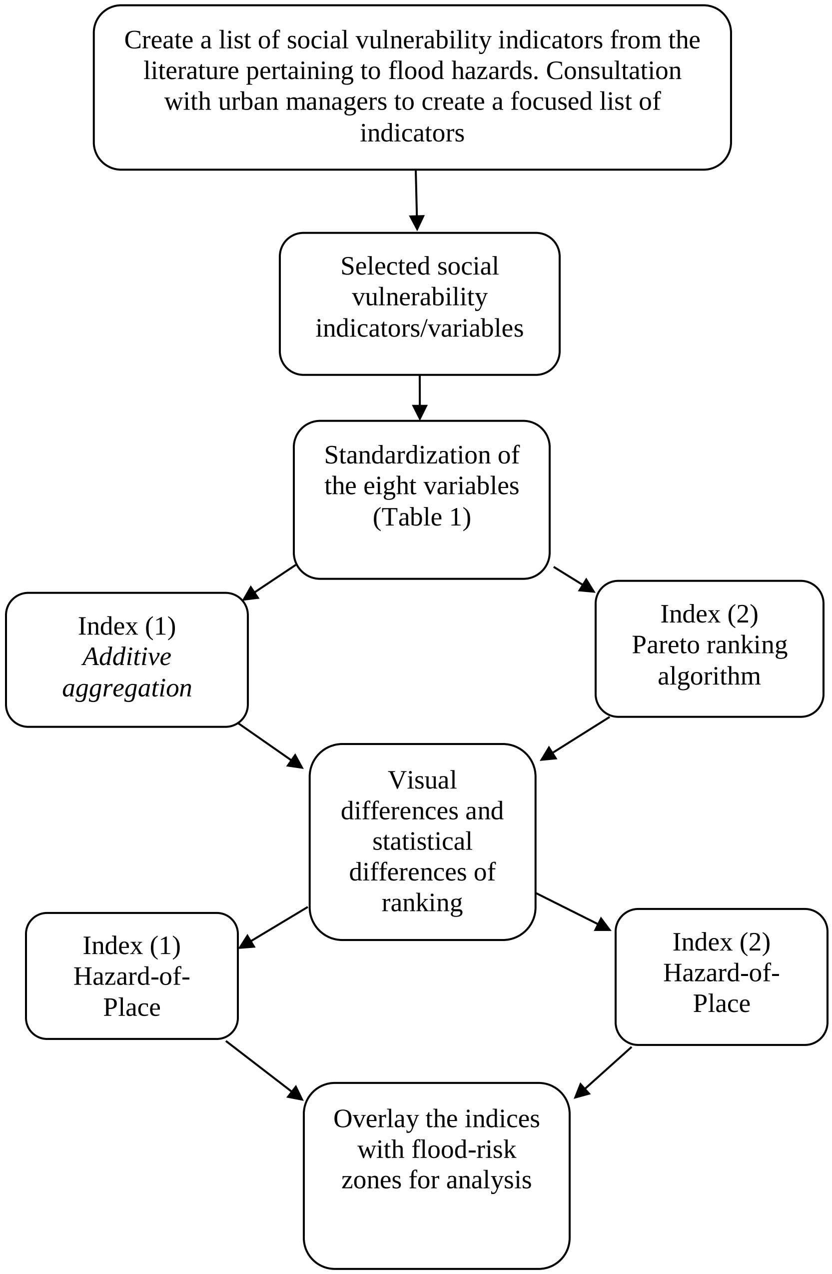

Figure 1. A concptual framework that explains the steps of this research.

4. DATA AND METHODS

This research utilizes Simply Analytics, computer software that allows users to search for census data and then plot the data on a choropleth map using an online version of ESRI's ArcMap. The data is collected at the block group level for Travis County to analyze the finest scale that the U.S. Census is captured at, and to provide sufficiently fine resolution that planners and emergency managers might be able to use to identify and target socially vulnerable populations easily. This research uses census data from the most recent 2016 estimates because Travis County is part of the Austin-Round Rock Metropolitan Statistical Area, which is a rapidly growing part of the county. The City of Austin has released a report on their official website titled: *Austin Area Population Histories and Forecast* that shows an increase of about 159,197 people between the years 2010 and 2017. Thus, using the most recent 2016 estimate will give a better understanding of the distribution of the landscape in the present time. In addition, this research uses data that were obtained from the Federal Emergency Management Agency (FEMA) and the Watershed Department in Austin. These data contain the Federal FEMA's 100-year and 500-year floodplains layers in addition to 25-year floodplain layer that was created by the Watershed department in Austin to capture more localized flooding events in Austin. The following section explains the second approach used in this research to construct an index of social vulnerability.

4.1. PARETO RANKING OPTIMIZATION TECHNIQUE

Pareto ranking is a technique that originates from the context of the genetic algorithm, and it ranks cases on multiple criteria based on a fitness function (Rygel, L. et al. 2006). The search method applied is based on the concept of Pareto *non-dominant* in evaluating fitness or assigning selection probability to solutions (Konak et al. 2006). To understand the proposed multi-objective optimization technique fully, two concepts must be introduced: the concept of *non-domination* and the *Pareto-optimal front* (Rygel, L. et al. 2006; Konak et al. 2006). In any complete dataset, the *non-dominated* cases are the ones that have no other cases in the dataset that are more vulnerable than them, and this is because of their scoring at least as high or higher on all variables. The process of selecting the non-dominated solutions is an iterative one to allow researchers to assign a vulnerability ranking to every block group in the dataset. Each time a set of *non-dominated* solutions or block groups are selected, they are removed from the dataset in the next iteration so a new set can be selected.

The *non-dominated* set that is being selected in each iteration is what is called the *Pareto-optimal front* (Rygel, L. et al. 2006). This gives practitioners an opportunity to examine the Pareto-optimal fronts, and ultimately make a value judgment among the alternatives to arrive at a particular decision (Goldberg 2006). The ranking will give city officials an opportunity to consider optimizing the block groups in order of vulnerability. This research utilized the R environment to accomplish its goal. R is an integrated suite of software facilities for data manipulation, calculation and graphical display among other things. *Emoa* is the packaged this research used to apply the function *nds_rank (data),* which ranks block groups in order of vulnerability.

5. RESULTS

5.1. SOCIAL VULNERABILITY INDEX USING ADDITIVE AGGREGATION

To create the first index proposed in this research, the eight social variables (Table 1) were normalized to keep the range between zero and one by using the minimum-maximum stretching (Equation 1):

$$\text{Equation 1: } z = \frac{x - x_{min}}{x_{max} - x_{min}}$$

All variables were normalized using the above equation except the variable median household income was normalized using the following equation:

$$\text{Equation 2: } z = \left| \frac{x - x_{min}}{x_{max} - x_{min}} - 1 \right|$$

This was done because all the variables (Table 1), not including median household income, have direct or positive relationship to social vulnerability. However, the median household income variable has an inverse relationship to social vulnerability. In other words, higher percentages in all of the indicators, except for median household income, indicate higher vulnerabilities. Once all the variables were normalized, they were added together to produce the first social vulnerability index for each block group using the following equation:

$$\text{Equation 3: } Index = z_1 + z_2 + z_3 \ldots z_n/\text{N}$$

All of the above calculations were done using Excel and the results were exported to ArcMap, where characteristics of the variables were analyzed and discussed. In ArcMap, the variables were projected to NAD_1983_2011_StatePlane_Texas_Central_ FIPS_4203_ft_US. Several classification methods were considered (equal interval, standard deviation, quantile), however, the natural breaks (Jenks) method was chosen and used for this research. The natural breaks (Jenks) method finds natural clusters and creates classes around those clusters. Finally, a social vulnerability map was created in ArcMap, and the block groups were classified into low, medium low, medium high and high vulnerability (Figure 2).

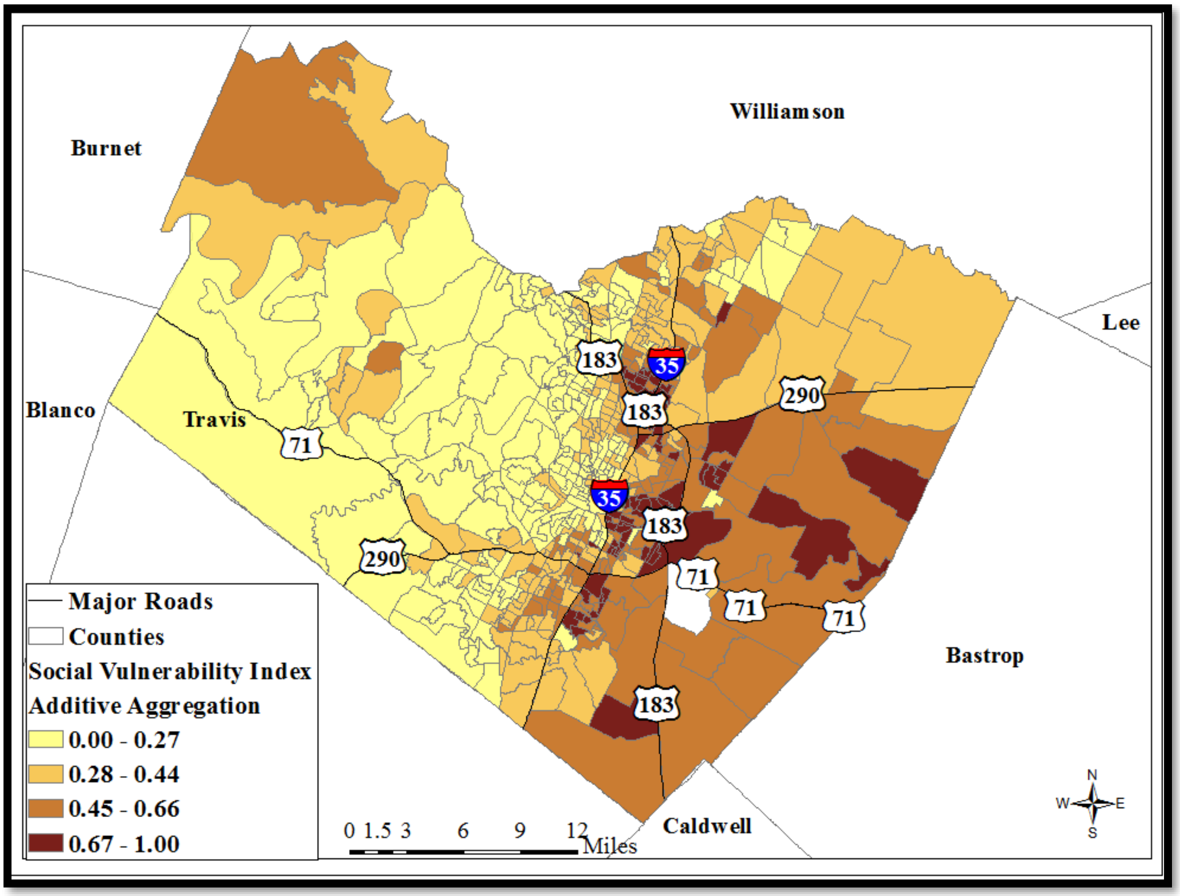

Figure 2. Overall social vulnerability in Travis County, Texas calculated using additive aggregation of the eight standardized social vulnerability variables

5.2. SOCIAL VULNERABILITY INDEX USING PARETO ALGORITHM

To create the second index, the Excel file that contains the eight standardized social variables was exported to the statistical software packages R. The package that was used to implement the *Pareto ranking algorithm* is called *emoa,* and the function that was applied to rank the *non-dominated* block groups is called *nds_rank*. In R, a matrix from the data was created because the function *nds_rank* works on matrices, not on data frames. Columns in a matrix must have the same data type and length compare to a data frame. After that, the function was applied to the matrix and a new column named *Rank* was created to store the ranking of each block group. In this study, with 579 block groups and 8 variables, block groups were sorted into six ranks. To assess overall social vulnerability, the six Pareto ranks were reassigned such that the most vulnerable block groups had a score of 1 and the least vulnerable block groups had a score of 6. The social vulnerability of each block group was then defined as its Pareto rank. Finally, the results were rescaled from 0 to 1 to increase interpretability, and overall vulnerability zones were established by sorting the scores into four classes (i.e., low, medium low, medium high and high vulnerability) using the natural breaks (Jenks) method of classification in ArcMap (Figure 3).

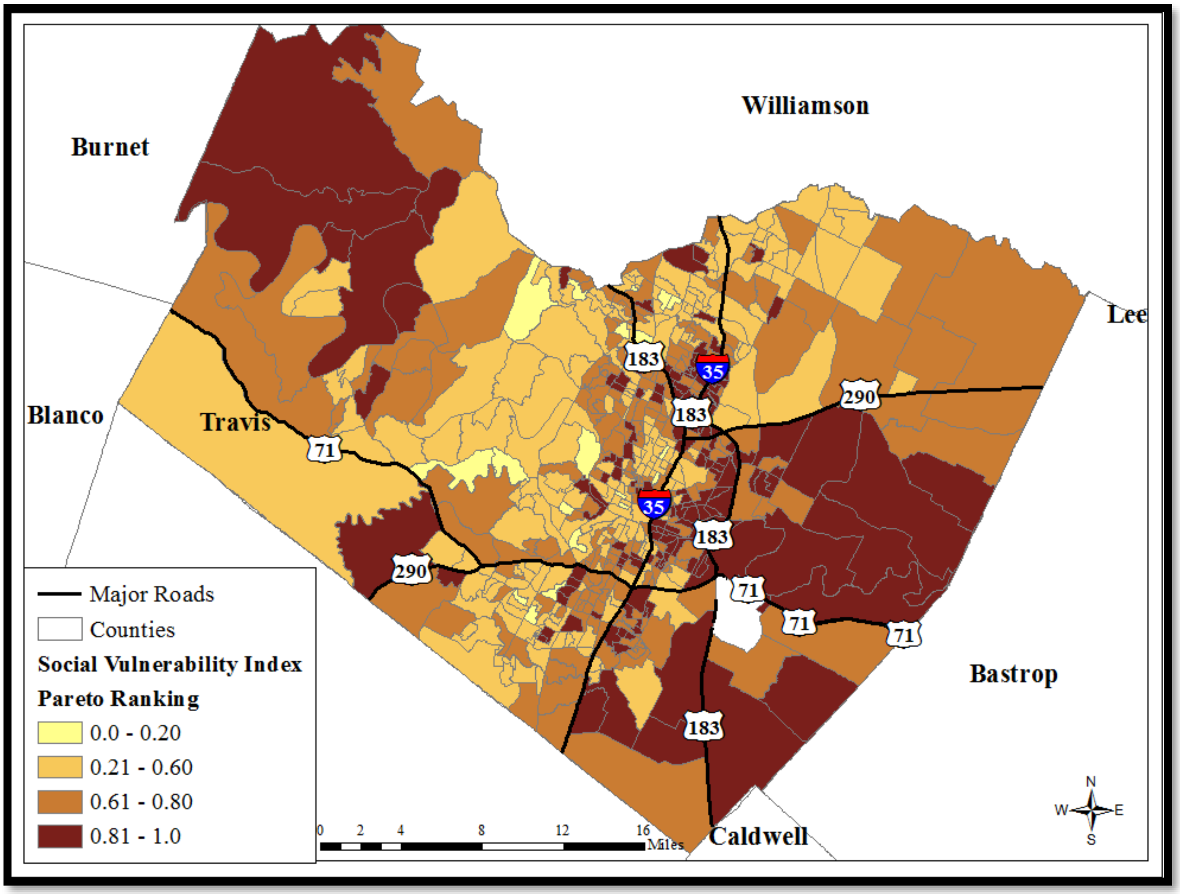

Figure 3. Overall social vulnerability in Travis County, Texas calculated using Pareto ranking of the eight standardized social vulnerability variables

5.3. COMPARISON OF ADDITIVE AGGREGATION AND PARETO RANKING

The additive and Pareto ranking methods of aggregation scores were compared using Spearman's rank correlation. In order to compare their social vulnerability scores, block groups are rank-ordered based on their overall composite social vulnerability scores. Spearman's rank correlation coefficient is a nonparametric rank statistic that measures the strength of the association between two variables and ranges from a -1.00 to 1.00 (Hauke and Kossowski 2011). The null and alternative hypothesis tested here are as follow:

- Ho: There's no significant difference between *additive aggregation* and the *Pareto algorithm aggregation (*i.e., there's a monotonic correlation)
- H1: *Additive aggregation* and the *Pareto algorithm aggregation* are significantly different (i.e., there's no monotonic correlation)

The Spearman correlation shows that the overall social vulnerability scores using the two different aggregation methods is 0.659, which indicates a moderate, positive monotonic correlation between *additive aggregation* and *Pareto ranking*. The Spearman's rank-order correlation results indicate that the alternative hypothesis can be rejected for this test (i.e., there is a monotonic correlation). The results from the Pareto ranking show more vulnerable block groups in the western portion of the county than what the results from the additive aggregation indicate. Pareto ranking will help direct the attention of City of Austin and Travis County managers toward the western part of the county than what the results from the additive aggregation indicate, which is that only the part of the county with a high concentration of low-income households is vulnerable. Because the additive aggregation calculates the mean arithmetic of the standardized eight variables (Table 1), most of the block groups in the western portion of the county appear less vulnerable than what they are when considering individual indicators more important than their arithmetic mean. Oulahen, G. et al. (2015) study showed that practitioners in municipalities in Canada found that individual maps are more helpful to guide their effort to manage flood hazards because in some situations an indicator such as households with no vehicle to escape the risk is more important than a composite of indicators, which has the potential to obscure such information. Since Pareto ranking highlights block groups that score high in one or two of the indicators, it could provide more variability for Austin and Travis County managers in their effort to identify vulnerable populations.

5.4. MAP CLASSIFICATIONS

Maps and other data graphics are one of the ways that play a role in generating ideas and communicating model results. Researchers have used maps to represent data to different audiences such as decision makers or a concerned public. However, a single map could reveal different patterns based on factors such as symbol types, color choices, and data classing. The different choices in each of the factors mentioned above depend on the type of data and analysis. This study considered different options to classify the results in ArcMap such as equal interval, natural breaks (Jenks), quantile, and standard deviation. By examining the data, the histogram shows that the data is not normally distributed which eliminated the choices of the equal interval, quantile, and standard deviation classification methods. Equal interval divides the attribute range into equally sized classes, and it is best used to emphasize the relative amount of attribute values compared to others. The quantile classification method will contain an equal number of features and is well suited for linearly distributed data. The standard deviation emphasizes how much feature values vary from the mean and is best used on normally distributed data because outliers might skew the distance from the mean. The final method of classification is natural breaks (Jenks) which this study used for data classification. Jenks arrange data values in order and the class breaks are determined statistically to find a relatively large difference between adjacent classes. This study classified the final social vulnerability indices using natural breaks with four classes, which are low, medium low, medium high and high vulnerability. (Figure 4) illustrates the difference between the four methods of classification in ArcMap using the additive aggregation. Moreover, considering the different ways to classify the data as shown in

(Figure 4), the Pareto ranking could offer an alternative to solve this issue of classification if the mapping of the block groups is based on their Pareto ranking which eliminates the need to create arbitrary zones of vulnerability to classify the data.

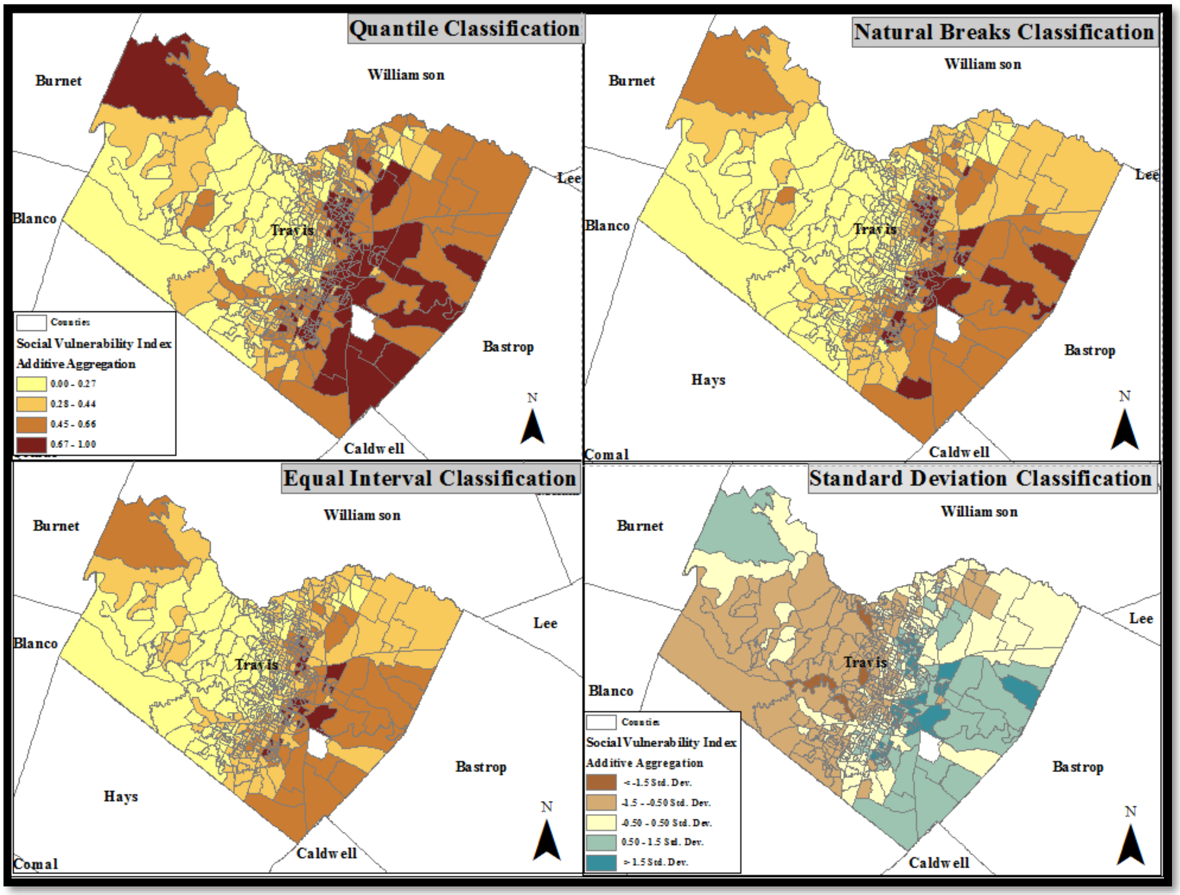

Figure 4. Different classification schemes using the same data

5.5. FLOODPLAIN OVERLAY

Following the conceptual model created by Cutter (1996), which proposes that the intersection between the biophysical vulnerability and social vulnerability creates the hazard-of-place, this research overlaid FEMA floodplains and Austin watershed department fully developed floodplain with the social vulnerability indices. By doing so, this study aims to help further guide the City of Austin and Travis County managers in their goal to identify vulnerable populations to extreme weather events such as flooding. The floodplain layers contain the 25-year, 100-year and 500-year floodplains. (Figure 5) and (Figure 6) show the results of the additive aggregation and Pareto ranking, respectively, overlaid with the floodplains layer obtained from FEMA and Austin watershed department.

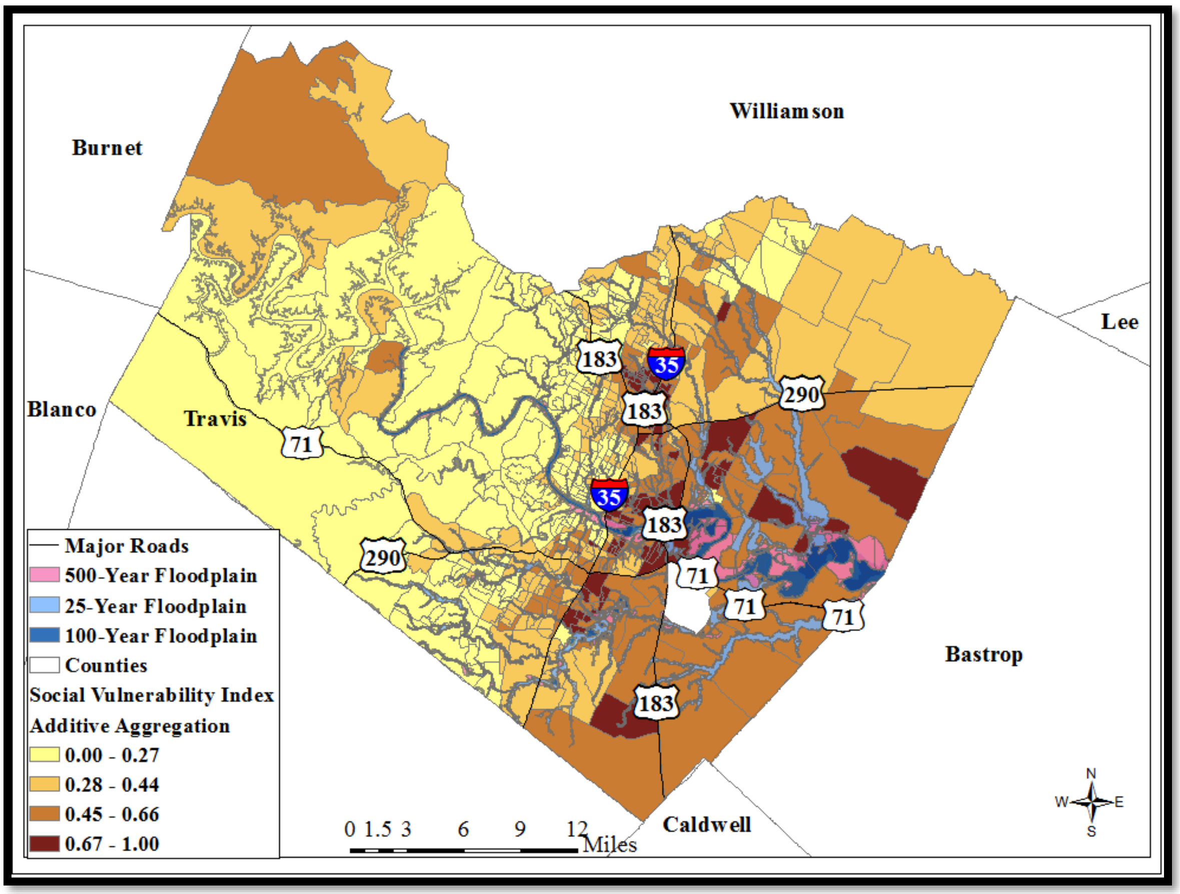

Figure 5. Additive aggregation Vulnerability Index map with 25-years, 100-years, and 500-years floodplains

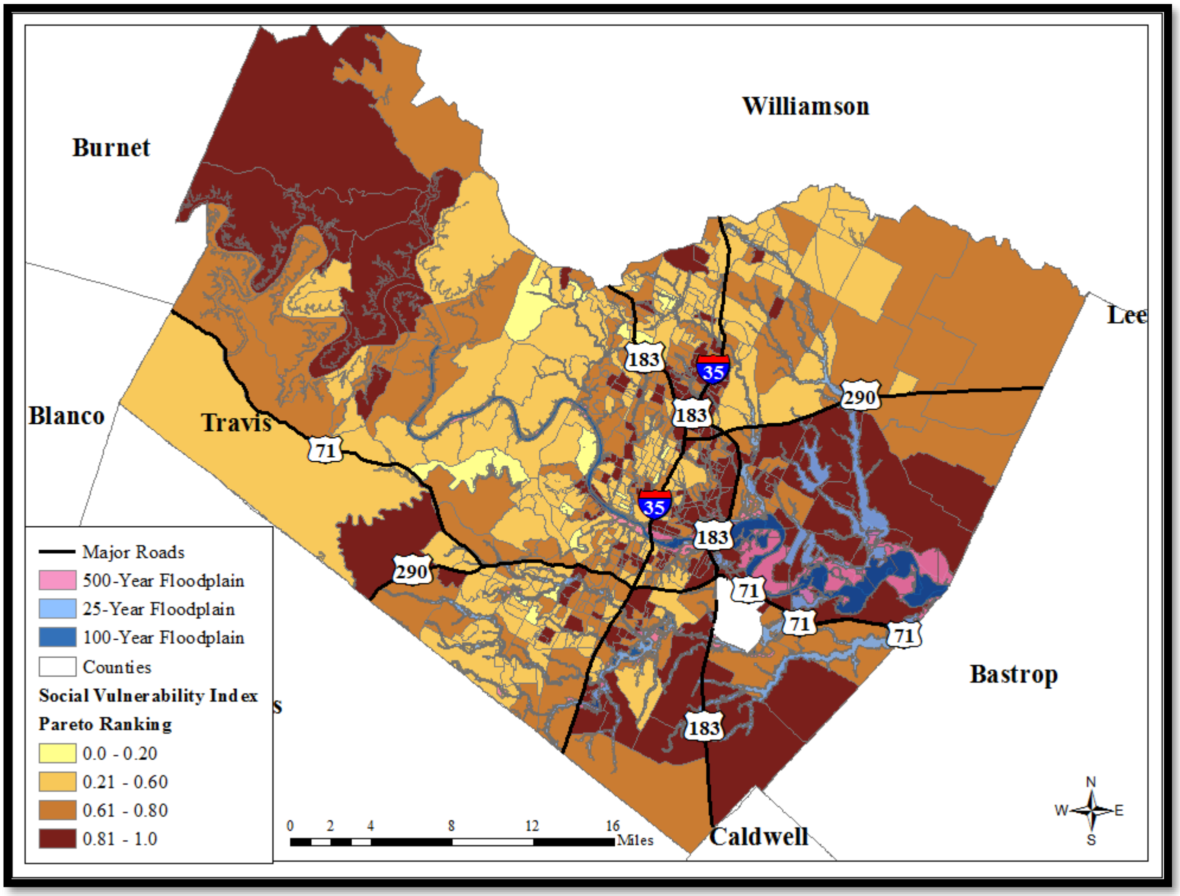

Figure 6. Pareto ranking Vulnerability Index map with 25-years, 100-years, and 500-years floodplains

To increase interpretability, (Table 3) and (Table 4) present the numbers and percentages of block groups in each of the four vulnerability zones of the two indices of social vulnerability. The four zones are the sorting of block groups into low, medium low, medium high and high social vulnerability. In addition, the tables show the numbers of block groups from each of the four zones that intersect with the 25-year, 100-year and 500-year floodplains. As the results show, most of the high vulnerability block groups, from both methods, intersect with the 25-year and 100-year floodplains, which have a higher chance of being flooded than block groups that intersect only with the 500-year floodplain. The study concludes that drill-down maps that focus on specific areas in

Travis County with high vulnerability are necessary to understand the intersection between floodplains and socially vulnerable block groups.

Moreover, the differences between the two methods show that the additive aggregation indicates that there are less block groups that are classified as high vulnerability and intersect with floodplains than what the Pareto ranking show. Therefore, using only the additive approach to assess vulnerability could potentially mislead City of Austin and Travis County managers in their effort to implement effective strategies and programs to mitigate the effect of flood hazards in their respective jurisdictions. Thus, future studies should consider that additive aggregation has the potential to obscure block groups that score highly in one or two of the indicators, and that potentially could lead to incorrect interpretation of the spatial distribution of social vulnerability.

Table 3. Floodplain overlay analysis for additive aggregation

Additive aggregation	**Classes of vulnerability**	**Number of block groups**	**Percentage of block groups**	**Number of block groups intersect with 25-yr & 100-yr floodplain**	**Number of block groups intersect with the 500-yr floodplain**
	Low	209	≈ 36 %	172	209
	Medium low	186	≈ 32 %	134	186
	Medium high	110	≈ 19 %	93	118
	High	74	≈ 13 %	52	74
	Total	579			

Table 4. Floodplain overlay analysis for Pareto ranking aggregation

	Classes of vulnerability	Number of block groups	Percentage of block groups	Number of block groups intersect with 25-yr & 100-yr floodplain	Number of block groups intersect with the 500-yr floodplain
Pareto ranking aggregation	Low	19	≈ 3 %	15	15
	Medium low	213	≈ 37 %	168	213
	Medium high	196	≈ 34 %	156	196
	High	151	≈ 26 %	105	151
	Total	579			

5.6. DRILL-DOWN COMPARATIVE ANALYSIS OF BLOCK GROUPS

Although the Spearman's correlation indicated a positive monotonic relationship between the additive aggregation and the Pareto ranking, some block groups that were classified as low vulnerability using the additive aggregation have a Pareto ranking of one, which indicates high vulnerability. This difference shows that the Pareto ranking highlights block groups that score high in one or more indicators, and that result in more variability in the distribution of block groups that require attention. For example, block group BG0017681_Travis_County_Tx (Figure 7) has a low vulnerability score in the additive aggregation, but in the Pareto ranking it has a high ranking of one. This block group has a low score in all indicators of vulnerability except for the indicator that represents the percentage of populations 65 years and older which further proves that the additive aggregation has the potential to obscure block groups that score high in one indicator.

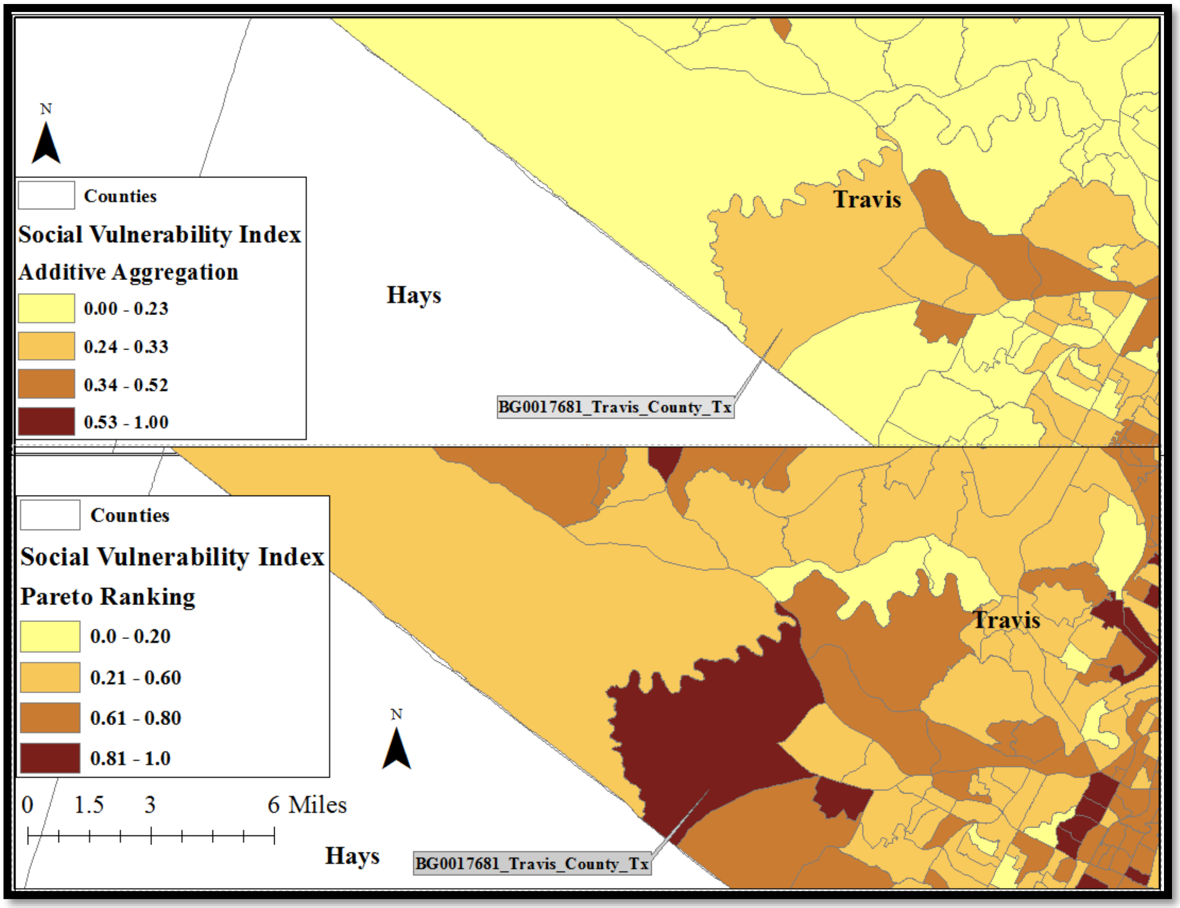

Figure 7. Overall social vulnerability of block groups in Travis County, TX: top image represents additive aggregation of indicators; bottom image represents Pareto ranking.

Many other block groups move up in the Pareto ranking aggregation because of scoring high in one or two of the indicators of social vulnerability. For instance, block group BG0024211_Travis_County_Tx (Figure 8) has a medium-low score using the additive aggregation but moves up to medium-high using the Pareto ranking. This block group has a low score in all indicators except for the indicator that represents single-parent household, and it has a relatively low level of income, which increases the vulnerability of the block group.

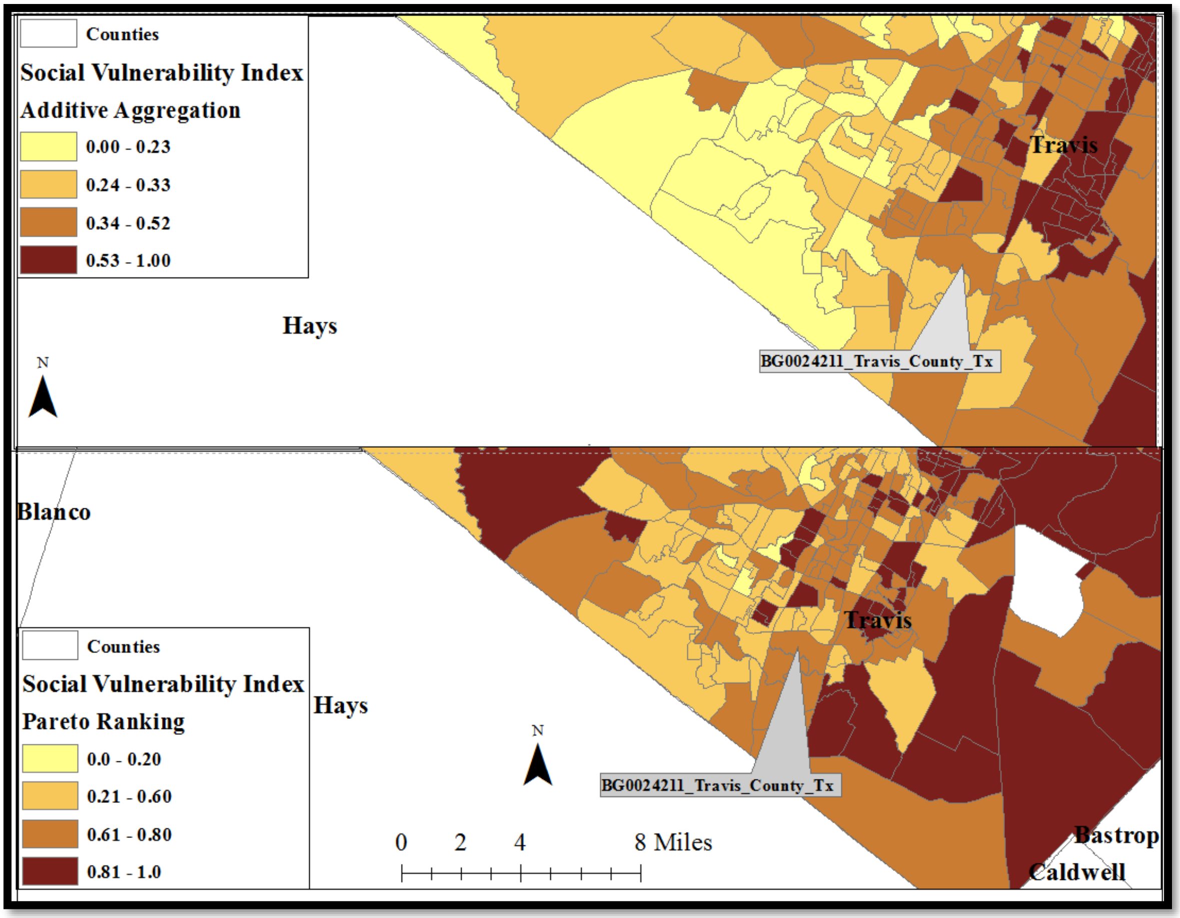

Figure 8. Overall social vulnerability of block groups in Travis County, TX: top image represents additive aggregation of indicators; bottom image represents Pareto ranking

Another example, block group BG0016023_Travis_County_Tx (Figure 9) has a medium-low score using the additive aggregation and a high vulnerability score using the Pareto ranking. The reason for this difference is due to the indicator that represents populations of age 5 and under, for which this block group scores high.

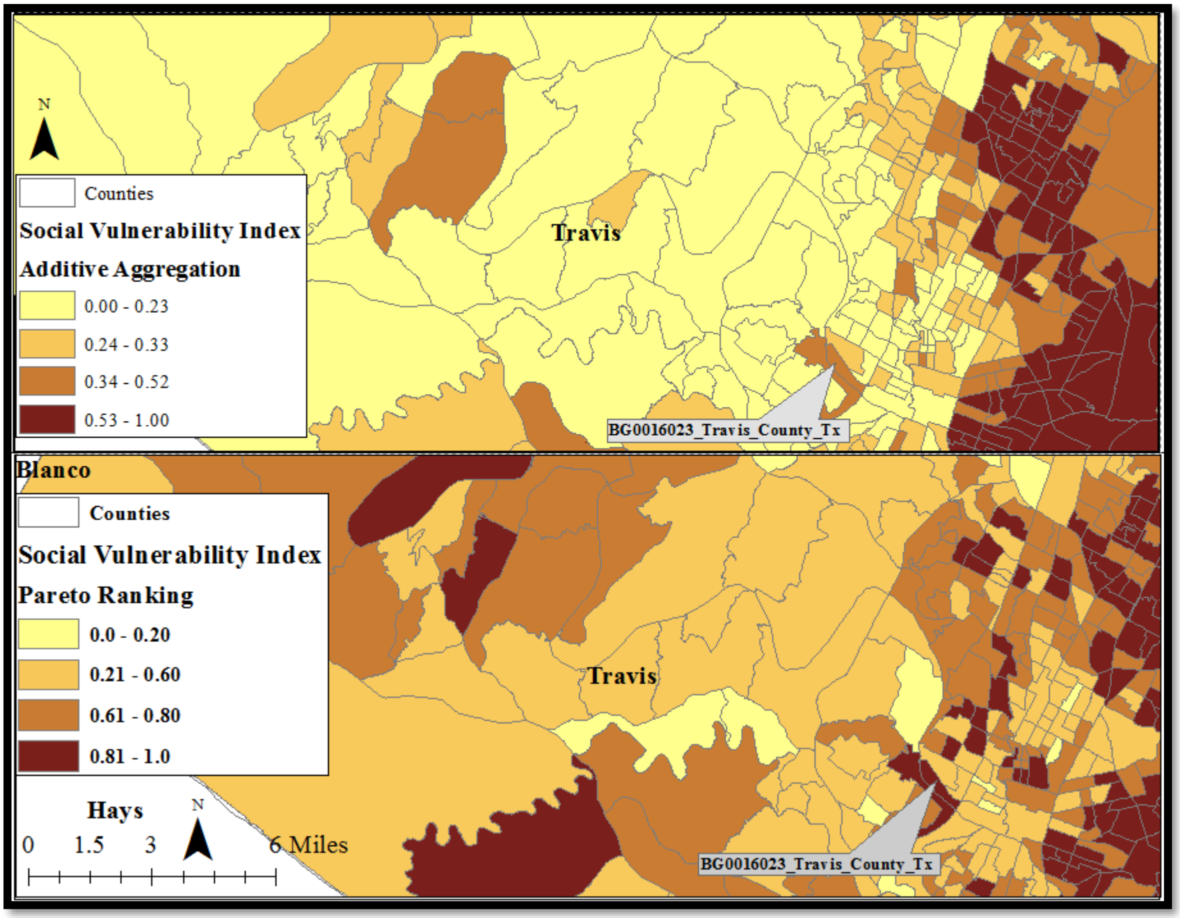

Figure 9. Overall social vulnerability of block groups in Travis County, TX: top image represents additive aggregation of indicators; bottom image represents Pareto ranking.

6. LIMITATIONS OF STUDY

The research of vulnerability to environmental hazards has been a challenge to researchers and policymakers. Often, practitioners who are involved in making decisions criticized research that is done without their inputs when assessing the vulnerability of their communities (Oulahen, G. et al. 2015). This research has incorporated expert's knowledge from urban managers at the Office of Sustainability in Austin to focus the list of indicators (Table 1). However, there is a general agreement that the list is not final, but has the potential to start conversations between different concerned parties within the county. The staff from the Office of Sustainability expressed that other departments within the county will have different indicators of vulnerability that will be of more importance to their management goals than what is in (Table 1). Therefore, users of the results of this research should consider it only within the context of flood hazards.

For Austin and Travis county managers to gain a better understanding of the spatial distribution of populations with higher vulnerability to flood hazards, they can use the SOVI maps created in this study. Indexes are means of understanding the spatial distribution of a statistical unit such as block groups. They help researchers or policymakers to visualize the social vulnerability distribution of the landscape, but those indexes do not reveal the underlying causes of any of the indicators such as why a certain block group has more low-income populations than the adjacent block group?

Quantifying social phenomenon in a context of a polygon representation in software, such as ArcMap, does not allow the researcher to understand the root of the phenomena. Therefore, indexes should be thought of as a starting point of a complex process to alleviate and understand the vulnerability of populations. In addition, another limitation of this research is considering that the socio-economic variables selected here are independent of influences from surrounding counties. People commute between counties daily, and the effect of that on Travis County social vulnerability is a factor that warrants further investigation.

Moreover, the number of processes to construct indices of vulnerability such as weighting, standardizing, and aggregating indicators are numerous. The availability of numerous approaches make it difficult to have a baseline of constructing an index, therefore, determining the social vulnerability of people should be a mixed approach rather than only using quantitative measures. Finally, future research should combine the land use/land cover variable to the analysis to help further city staff to focus on residential areas of the county.

7. CONCLUSION

Human actions, primarily the burning of fossil fuels, are causing the earth's climate to change. Extreme weather events that lead to flooding like the Memorial Day Flood in 2015 and wildfires like the Bastrop Fire of 2011 due to climate variability will pose a challenge to the city of Austin and the surrounding communities. Thus, determining where the vulnerable populations to such events are located is important in order to implement effective strategies and programs for mitigation. This research compared two of the more prevalent methods of aggregation to assess the vulnerability of block groups in Travis County. A quantitative and visual comparison were done using the Spearman's correlation and ESRI's ArcMap, respectively. The two methods are different in that the additive aggregation uses the arithmetic mean to compute the index, on the other hand, Pareto ranking is a multi-objective optimization technique that ranks block groups based on *non-domination* in the complete data set.

The additive aggregation has been criticized because it has the potential to obscure block groups that score high in one of the indicators. However, Pareto ranking has been proposed as an alternative to the additive aggregation approach because it avoids assuming that all the indicators that contribute to social vulnerability are equally important. The *non-dominated* set that is being selected in each iteration is what is called the *Pareto-optimal front* (Rygel, L. et al. 2006). This gives practitioners an opportunity to examine the Pareto-optimal fronts, and ultimately make a value judgment among the alternatives to arrive at a decision (Goldberg 2006).

Although the analysis showed that the two methods have a moderate positive monotonic correlation, some block groups that were classified as low vulnerability using the additive aggregation fell into the high vulnerability classification using the Pareto ranking. This proves that the additive aggregation has the potential to obscure block groups that score high in one or two of the indicator, therefore, researchers and urban managers should be aware of this drawback when assessing social vulnerability.

In addition, the Pareto ranking offered more variability in the spatial distribution of vulnerability due to its algorithm that treats block groups as a multi-objective optimization problem where several indicators of vulnerability are present simultaneously, and it is not possible or wise to combine the indicators into a single number. Therefore, Pareto ranking has the potential to give city officials an opportunity to consider optimizing the block groups in order of vulnerability.

Finally, this research concludes that SOVI maps created in this comparison between the two aggregation methods are a useful tool to visualize the spatial distribution of vulnerable populations to flood hazards. However, validating the results through qualitative approaches that aim at understanding the historical and structural factors that constrain the adaptive capacity of vulnerable populations will help Austin and Travis managers in their effort to implement strategies and allocate resources that will serve the need of Travis County citizens (Singh et al. 2016). This could be accomplished through semi-structured interviews with local practitioners from a broader array of people before using the SOVI maps to disburse funds or allocate resources.

REFERENCES

Adger, W. Neil. 2003. Social Capital, Collective Action, and Adaptation to Climate Change. *Economic Geography* 79 (4): 387-404.

ArcMap, Ver. 10.5.1, ESRI, Redlands, CA, USA.

Armas, I., A. Gavris. 2013. Social Vulnerability Assessments Using Spatial Multi-Criteria Analysis (SEVI model) and the Social Vulnerability Index (SoVI Model) - A Case Study of Bucharest, Romania. *Natural Hazards and Earth System Science* 13 (6): 1481-1499.

Araya-Muñoz, D., M. J. Metzger, N. Stuart, A. M. W. Wilson, and D. Carvajal. 2017. A spatial fuzzy logic approach to urban multi-hazard impact assessment in Concepción, Chile. *Science of the Total Environment* 576:508–519.

Azar, D., and D. Rain. 2007. Identifying population vulnerable to hydrological hazards in San Juan, Puerto Rico. *GeoJournal* 69 (1-2):23–43.

Boruff, B. J., C. Emrich, and S. L. Cutter. 2005. Erosion Hazard Vulnerability of US Coastal Counties. *Journal of Coastal Research* 215:932–942.

Chakraborty, Jayajit, G. A. Tobin, and B. E. Montz. 2005. Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards. *Natural Hazards Review* 6 (1): 23-33.

Chen, W., P. Shi, S. L. Cutter and C. T. Emrich. 2013. Measuring Social Vulnerability to Natural Hazards in Yangtze River Delta Region, China. International Journal of Disaster Risk Science 4 (4):169-181.

The city of Austin. 2014. *Net-Zero: Austin Community Climate Plan.*

City of Portland. 2014. *Climate Change Preparation Strategy-Risk and Vulnerabilities Assessment: Preparing for Local Impacts in Portland and Multnomah County 2014.*

Cutter, Susan L., J. T. Mitchell, and M. S. Scott. 2000. Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina. *Annals of the Association of American Geographers* 90 (4):713–737.

Cutter, Susan L., Bryan J. Boruff. and W. Lynn Shirley. 2003. Social Vulnerability to Environmental Hazards. *Social Science Quarterly* 84 (2): 242-261.

Cutter, Susan L. and Bryan J. Boruff. 2007. The Environmental Vulnerability of Caribbean Island Nations. Geographical Review. 97(1):24-45

Cutter, Susan L. and C. Finch. 2008. Temporal and Spatial Changes in Social Vulnerability to Natural Hazards. *Proceedings of the National Academy of Sciences of the United States of America* 105 (7):2301-2306.

Chakraborty, J., G. A. Tobin, and B. E. Montz. 2005. Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards. *Natural Hazards Review* 6 (1):23–33.

Dwyer, A., Zoppou, C., Nielsen, O., Day, S. Roberts, S., 2004. Quantifying Social Vulnerability: A methodology for identifying those at risk to natural hazards. *Geoscience Australia* Record 2004 (14)

Fünfgeld, H. 2010. Institutional challenges to climate risk management in cities. *Current Opinion in Environmental Sustainability* 2 (3):156–160.

Flanagan, Barry E., Edward W. Gregory, Elaine J. Hallisey, Janet L. Heitgerd, and Brian Lewis 2011. A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management* 8 (1)

Fothergill, A., E. G. M. Maestas, and J. D. Darlington. 1999. Race, Ethnicity and Disasters in the United States: A Review of the Literature. *Disasters* 23 (2):156–173.

Fekete, A. 2009. Validation of a social vulnerability index in context to river-floods in Germany. *Natural Hazards and Earth System Science* 9 (2):393–403.

Goldberg, D. E. 2012. *Genetic algorithms in search, optimization, and machine learning*. Boston: Addison-Wesley.

Ge, Yi, X. Quian, X. Zhou, Z. Gu, J. Wang, W. Xu, P. Shi, X. Ming and Y. Chen. 2013. Assessment of Social Vulnerability to Natural Hazards in the Yangtze River Delta, China. *Stochastic Environmental Research and Risk Assessment* 27 (8):1899-1908.

Ge, Yi, Wen Dou and Ning Liu. 2017. Planning Resilient and Sustainable Cities: Identifying and Targeting Social Vulnerability to Climate Change. *Sustainability* 9 (8): 1394.

Ghadiri, M and A. R. Roknodineftekhari. 2013. The Relation between Social Structure of Cities and Earthquake Vulnerability: A Case Study Tehran City's Neighborhoods. Geography & Environmental Planning 50 (2):37-41.

Krishnan, V. 2010. Constructing a Multidimensional Socioeconomic Index and the Validation of It with Early Child Developmental Outcomes. *Advances in Data Mining and Database Management Emerging Trends in the Development and Application of Composite Indicators*: 163–199.

Hallegatte, S. C. A. 2009. Strategies to adapt to an uncertain climate change. *Global Environmental Change* 19 (2):240–247.

Hayhoe, Katharine. 2014. Climate Change Projections for the City of Austin: Draft Report April 2014. Atmos Research and Consulting

Holand, Ivar S., P. Lujala, J. K. Rod. 2011. Social Vulnerability Assessment for Norway: A Quantitative Approach. Norsk Geografisk Tidsskrift 65 (1):1-17.

Holand, Ivar S. and P. Lujala. 2013. Replicating and Adapting an Index of Social Vulnerability to a New Context: A Comparison Study for Norway. Professional Geographer 65 (2): 312-328.

Hauke, J., and T. Kossowski. 2011. Comparison of Values of Pearsons and Spearman's Correlation Coefficients on the Same Sets of Data. *Quaestiones Geographicae* 30 (2).

Johnson, J. E., D. J. Welch, J. A. Maynard, J. D. Bell, G. Pecl, J. Robins, and T. Saunders. 2016. Assessing and reducing vulnerability to climate change: Moving from theory to practical decision-support. *Marine Policy* 74:220–229.

Konak, A., D. W. Coit, and A. E. Smith. 2006. Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering & System Safety* 91 (9):992–1007.

Intergovernmental Panel on Climate Change (IPCC). 2014. Climate Change 2014 - Impacts, Adaptation, and Vulnerability: Summary for Policymakers.

Lee, Yung-Jaan. 2014. Social vulnerability indicators as a sustainable planning tool. *Environmental Impact Assessment Review* 44 (2014): 31-42.

Li F., J. Bi, L. Huang, C. Qu, J. Yang and Q. Bu. 2010. Mapping Human Vulnerability to Chemical Accidents in the Vicinity of Chemical Industry Parks. *Journal of Hazardous Materials* 179 (1):500-506.

Lixin, Y., Z. Xi, G. Lingling, and Z. Dong. 2014. Analysis of Social Vulnerability to Hazards in China. *Natural Hazards* 87 (2): 1223-1243.

Martins, V. N., D. S. e Silva and P. Cabral. 2012. Social Vulnerability Assessment to Seismic Risk Using Multicriteria Analysis: The Case Study of Vila Franca do Campo. Natural Hazards 62 (2):385-404.

Melillo, Jerry M., Terese (T. C.) Richmond, and Gary W. Yohe, Eds., 2014: Climate Change Impacts in the United States: The Third National Climate Assessment. U.S. Global Change Research Program, 841 pp. doi: 10.7930/J0Z31WJ2.

Myers, C. A., T. Slack, and J. Singelmann. 2008. Social vulnerability and migration in the wake of disaster: the case of Hurricanes Katrina and Rita. *Population and Environment* 29 (6):271–291.

Nguyen, Cuong, Ralph Horne, John Fien, and France Cheong. 2017. Assessment of social vulnerability to climate change at the local scale: development and application of a Social Vulnerability Index." *Climatic Change* 143 (¾): 355-370.

Oulahen, G., L. Mortsch, K. Tang, and D. Harford. 2015. Unequal Vulnerability to Flood Hazards: “Ground Truthing” a Social Vulnerability Index of Five Municipalities in Metro Vancouver, Canada. *Annals of the Association of American Geographers* 105 (3):473– 495.

Parry M. L., O. F. Canziani, J. P. Palutikof, P. J. van der Linden and C. E. Hanson. 2007. “Working Group II Report “Impacts, Adaptation and Vulnerability" Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change 2007.

Rygel, L., D. O'Sullivan, and B. Yarnal. 2006. A Method for Constructing a Social Vulnerability Index: An Application to Hurricane Storm Surges in a Developed Country. *Mitigation and Adaptation Strategies for Global Change* 11 (3):741–764

Simply Analytics, Simply Analytics Inc., New York, NY, USA.

Singh, A., and Z. Zommers. 2016. *REDUCING DISASTER: early warning systems for climate change*. *SPRINGER*.

Tate, E. 2012. Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards* 63 (2):325–347.

Wood, N. J., C. G. Burton and S. L. Cutter. 2010. Community Variations in Social Vulnerability in Cascadia-Related Tsunamis in the U.S. Pacific Northwest. *Natural Hazards 52:369-389.*

Wisner, B., P. Blaikie, T. Cannon, and I. Davis. 1994. at Risk**.**

Yenneti, Komali, Sabyasachi Tripathi, Yehua Dennis Wei, Wen Chen, and Gaurav Joshi. 2016. The truly disadvantaged? Assessing social vulnerability to climate change in urban India. *Habitat International* 56: 124-135.

Yoon, D. K. 2012. Assessment of social vulnerability to natural disasters: a comparative study. *Natural Hazards* 63 (2):823–843.

Zahran, S., S. D. Brody, W. G. Peacock, A. Vedlitz, and H. Grover. 2008. Social vulnerability and the natural and built environment: a model of flood casualties in Texas. *Disasters* 32 (4):537–560.