Social Vulnerability and Earthquake Impact Modeling in

Federal Emergency Management Agency (FEMA) Region IV (Southeast of the U.S.)

by

Sumaiya Tul Siddique

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 4, 2018

Keywords: social vulnerability, natural hazard,
HAZUS-MH, estimated damage

Approved by

Christopher Burton, Chair, Assistant Professor of Geosciences
Chandana Mitra, Associate Professor of Geosciences
Philip Chaney, Associate Professor of Geosciences

Abstract

At the forefront of the natural hazard risk assessment sciences is the assessment of earthquake risk. Countries such as the United States have been targets of extensive earthquake risk assessments to communicate potential damages and loss of life. Few studies, however, have gone beyond the estimation of direct earthquake impacts such as damages to buildings by integrating estimates the socio-economic characteristics of populations. This research seeks to address this missing societal component in earthquake loss modeling using Federal Emergency Management Agency Region IV (Alabama, Florida, South Carolina, North Carolina, Mississippi, Kentucky, Georgia, and Tennessee) as a case study. Social vulnerability and economic losses are modeled respectively by developing a Social Vulnerability Index (SoVI) and by integrating the SoVI with probabilistic impact estimates from potential earthquakes within the region. The results of this research highlight the areas of management concern in which high earthquake losses may be coupled with populations that are unlikely to be able to prepare for, respond to, and recover from damaging earthquake events when they occur.

Acknowledgements

For me, the Master’s was very challenging in the United States, and the completion of this thesis was an arduous journey. Throughout the period of research and writing, I had to overcome adversity and a number of hurdles. Without several people and their contributions in different ways, my academic status and thesis completion would not have been possible. Because of this, I would like to express my gratitude.

First, I would like to thank my parents, Mr. Md. Siddiqur Rahman and Mrs. Delowara Akter Jahan. Without your cordial support, my higher education may not have been possible. You are always beside me when I need your help. Second, I’d like to thank my husband, Mohammed Humed Yusuf. You are always with me as my closest friend and available when I need your suggestions. It would have been almost impossible to continue my higher education without your support and sacrifice.

To Dr. Christopher Burton, I would also like to express my special thanks and gratitude. You are a fine fellow, researcher and mentor. I thank you for your instructions, expert advice and draft corrections in time. You were generous to advise me again and again without being tired when sometimes I failed to understand certain concepts. I appreciate your patience.

To my committee members, Dr. Chandana Mitra and Dr. Philip Chaney, I would like to express my thanks. You two are the educators who helped to guide me and to gain the background that allowed me to develop my research. Taking your courses helped me a lot. To

my entire committee, I thank you for attending my thesis proposal presentation and advising me at the primary stages of my thesis.

And finally, this is time to thank all faculty, staff, my classmates and colleagues in the Geosciences Department at Auburn University. Thank you all for making it feel easier to adjust to a new department and a new environment during my first semester. I will never forget you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

Earthquakes are a natural phenomenon. There are hardly any early warning systems to notify populations of a potential earthquake event, and earthquakes can cause serious damage when they impact the areas in which people live. It is within this context that damages to infrastructure and livelihoods could result in a disaster where a serious disruption of the functioning of a community or society occurs (United Nations Office for Disaster Risk Reduction (UNISDR) 2017). Over-populated areas are among the most vulnerable to earthquakes and their associated impacts as a result. Highly developed urban areas, for instance, are at more risk from earthquakes due to their large number of infrastructure, high population densities and a lack of environmentally sustainable development activities in some places. From 2000 to 2015, earthquakes caused 801,629 deaths worldwide (USGS Earthquake Hazards Program 2017). Recent billion dollar loss causing earthquake events include the 2008 Wenchuan Earthquake in China (US$ 150 billion), the 2010 Christchurch Earthquake in New Zealand (NZ$ 40 billion), and the 2010 earthquake in Chile (US$ 30 billion) (Araneda et al. 2010, Sun and Xu 2010).

Although often overlooked, the Southeastern part of the U.S. is situated within a vulnerable zone to earthquake events. The result is that both the region’s infrastructure and social systems are at risk from potential adverse impacts, and there has been considerable seismic activity in this area. In the past, there were damaging earthquakes centered in Western North Carolina (i.e. Wilkes County experienced an earthquake of 5.1 magnitude in 1861). The Skyland

Earthquake in western North Carolina occurred on February 21, 1916 which was 5.5 magnitude event where shaking was felt for approximately 200,000 square miles. Mitchell County, Georgia experienced a 5.2 magnitude earthquake in 1926 in which shaking was felt over an area of about 40,000 square miles. This earthquake damaged chimneys, cracked house foundations, broke water pipes, and displaced windowpanes. Other damaging regional earthquakes, in 2003 and 2011 respectively, occurred Richmond (4.5 magnitude) and Mineral (5.8 magnitude), Virginia (NCDCCPS 2016). The Richmond earthquake was felt as far south as Raleigh, North Carolina, and the Mineral Earthquake was felt throughout most of western North Carolina.

In addition to moderate earthquakes, two very large earthquakes in 1811 and 1812 (magnitude 7.3 and 8.3 respectively) have occurred within the region. These were centered in the Mississippi Valley. Situated within close proximity to the New Madrid seismic zone, the effects of these events were severe and included damage to brick-made structures all over the state. In 1886, a major earthquake also occurred in Charleston South Carolina (magnitude of 6.7). The shaking affected windows, collapsed brick-made structures, and damaged internal contents of buildings. Here, approximately seven thousand buildings were damaged by this event, and the building materials were found as a major single factor causing the damage. Total damages were tallied at US$ 5 million (currency not adjusted to today's value) (Robinson and Talwani 1983). The New Madrid seismic zone is one of the most active faults in the Southeast of the U.S., and previous research has found that large earthquakes recur on average 450 years apart (Tuttle et al. 1999). These histories of events recognized Southeastern U.S. as an earthquake prone zone.

1.1 Research Perspective

Damage estimation and the ability to map potential causalities from earthquake events is becoming an increasingly important aspect of earthquake disaster risk reduction. Damage estimates from earthquakes include factors such as building materials, replacement cost, floor replacement and many other quantifiable losses. There are many earthquake loss estimation models to estimate these costs. To better understand risk from earthquakes in the U.S., the Federal Emergency Management Agency (FEMA) released the HAZUS-MH hazard and loss modeling tool which estimates dollar losses due to natural hazards like floods, hurricanes, earthquakes, storm surges and tsunamis. HAZUS-MH is a GIS-based tool which works as an extension to the ArcGIS software. HAZUS displays loss information through maps and spreadsheets. Data such as the location and intensity of past earthquakes, proximity to faults, slip rates, and location and cost of infrastructure drive the analysis of a HAZUS loss estimation.

The primary component of an earthquake impact model, such as those developed in HAZUS, is the measurement of potential ground shaking. HAZUS, for instance, uses several quantitative parameters such as peak ground acceleration and spectral acceleration for expressing ground shaking in a hazard analysis. A hazard analysis can be accomplished deterministically by replicating the ground shaking from a past earthquake. Conversely, a hazard analysis can be accomplished probabilistically by simulating thousands of potential earthquakes. A secondary component of a hazard analysis is the estimation of losses. Estimated losses are referred to as a physical risk assessment (Burton and Silva 2016). In HAZUS, a physical risk assessment can be conducted using three levels of data. The first level is based on default inventory data that includes building square footage, building value, occupancy information, and repair cost estimations. The second level of analysis utilizes user-supplied data that is often made publicly available through city, county, or state data sharing initiatives. Default inventory data is improved when high resolution local-level data can be used, resulting in a more robust analysis of economic loss. The third level of analysis requires experts such as public workers, planners, engineers, information technology specialists, GIS specialists, land recorders and natural resources management officials (FEMA 2003). Here, primary source data is collected exclusively for a given risk assessment project and could include a building by building inventory within a city.

Losses from earthquakes are not explicitly the result of ground shaking and building characteristics, however. Rather, losses are the result of interactions between the hazard, the engineered environment (e.g. residential and commercial infrastructure) and social characteristics of populations (Burton and Silva 2016). Social characteristics that make people more susceptible to hazard events can be explained by the concept of social vulnerability. Social vulnerability is defined as characteristics within social systems that create the potential for harm or loss (Cutter, Boruff, and Shirley 2003). Social vulnerability results from those social factors that make the society susceptible to harm such as lack of access to education, politics, technology, culture, religion, population and building density (Cutter, Boruff, and Shirley 2003). The social vulnerability of populations is largely ignored in earthquake risk assessments, however, even though social is an integral factor determining a community's earthquake risk.

Few studies and tools have been developed to demonstrate where highest losses from earthquakes are likely to occur combined with the likelihood of exposed populations to be able to prepare for, respond to, and recover from damaging earthquakes when they happen (i.e. a population's social vulnerability). More studies are thus needed that assess earthquake risk in a manner that is more holistic and integrtaed. In other words, more studies are needed that combine earthquake loss estimation with measures of social vulnerability.

This research seeks to address the integration of this missing social component in earthquake loss modeling in the southeastern United States. This research utilizes all counties within FEMA Region IV (Alabama, Florida, South Carolina, North Carolina, Mississippi, Kentucky, Georgia, and Tennessee) as a case study. The impacts due to earthquakes on communities are expressed in terms of both physical damages (i.e. physical risk) and socio-economic factors which affect the capacity of populations to absorb and recover from the events (Burton and Silva 2016, Carreño, Cardona, and Barbat 2007, Davidson 1997, Khazai et al. 2015). Physical damages are demonstrated using modelled losses and fatalities, whereas social characteristics are modelled using a well-established index, the Social Vulnerability Index (SoVI) (Cutter, Boruff, and Shirley 2003). Estimates of potential losses were coupled with the quantified measure of social vulnerability in order to assess earthquake risk in a manner that is more holistic than describing ground shaking and earthquake losses alone. This is to better understand the spatial distribution and drivers of earthquake risk that are a function of not only the vulnerability of the built environment, but also the vulnerability of a place's population that that creates a differential potential for loss and recovery. The overall goal of this study is to outline areas of management concern ─ those areas with the highest risk and social vulnerability. Defining these areas spatially and visualizing contributing factors to the region's risk and social vulnerability may help planners and decision makers to develop equitable public policies to reduce earthquake risk.

1.2 Research Questions

This research seeks to answer the following questions:

Research Q. 1. Which counties within FEMA Region IV have the highest physical risk (i.e. estimated property loss and fatalities) from earthquakes?

Research Q. 2. Which counties within FEMA Region IV have the highest physical risk from earthquakes coupled with a high social vulnerability?

Research Q. 3. What social vulnerability indicators may be best for predicting adverse effects such as losses from earthquakes?

1.3 Organization

This thesis is organized as follows. Chapter 1 provides a brief introduction. Chapter 2 provides a review of background literature. In this section, background literature of social vulnerability and earthquake physical impact modeling both are discussed. Chapter 3 describes the data sources and methodology. Chapter 4 presents the results after applying the methodology. The final chapter discusses all significance of the research and suggestions for the future research in this focus.

Chapter 2: Background Literature

2.1 Modeling Potential Earthquake Impacts from a Physical Perspective

In recent decades, there has been considerable advancements in earthquake impact model development. A variety of earthquake prediction models have been developed to predict the probable losses from the hazard. To estimate earthquake loss, assessment platforms HAZUS-MH (Hazus-MH 2003), the CAPRA Probabilistic Risk Assessment Platform (Anderson 2008), and OpenQuake (Silva et al. 2014b) are very popular. All of these platforms follow comprehensive methodologies for loss estimation (Erdik et al. 2014). These include the modelling of a basic set of components representing hazard, exposure and structural vulnerability. These are briefly described below.

2.1.1 The Earthquake Hazard Component

Earthquake hazard modelling components are created combining history, geophysical and engineering knowledge of earthquakes to provide estimates of future earthquakes and their potential threats (Rao et al. 2017). The hazard component represents the earthquake event (Burton 2010), and it is often modeled using peak ground acceleration (PGA) and spectral acceleration (SA). Peak ground acceleration is the maximum ground acceleration that occurs during earthquake shaking at a location. Spectral acceleration (SA) is a unit that describes the maximum acceleration of an object in an earthquake. Permanent ground deformations (PGD) from earthquake events is another hazard component which was derived to represent liquefaction, land sliding and surface fault rupture potential. Essentially, ground motions, earthquake fault rupture, ground displacement, shaking intensity, topography, nature of the rocks, and surface soil types are considered earthquake hazard components (HAZUS-MH 2004).

2.1.2 The Earthquake Exposure Component

The term exposure indicates those elements that are at-risk from hazards. Generally speaking, elements at-risk can be population, the building stock and essential facilities, and critical infrastructure. It is within this context that the exposure component describes the location, structural attributes, and values of assets that are in harm's way (Burton 2010). Damages are often calculated by accounting for the exposure of different construction types, roof types, number of floors, number of square footage, value of structures, and replacement costs. These are often utilized as GIS layers in tools such as HAZUS (HAZUS-MH 2003) and OpenQuake. Rao et al. (2017) provide an example of an exposure model developed for the state of California that was developed to identify the importance of uncertainty in estimating California's earthquake risk. For loss calculations, a residential exposure model was constructed at the census tract level for California by accounting for the total number of residential, commercial, and industrial structures. Seismic vulnerability models (discussed in the following section) for different building classes were created using the exposure database, and two analyses were undertaken: 1) a scenario-based analysis and 2) a probabilistic risk assessment. The scenario analysis was developed to address the risk associated with a single earthquake, the repetition of the 1906 San Francisco earthquake. The probabilistic analysis was used to address the risk associated with all possible events that could occur in the state (Rao et al. 2017).

2.1.3 The Physical Vulnerability Component

The physical vulnerability component of earthquake risk modelling represents the link between seismic hazard, exposure, and estimated damage (Burton 2010, Davidson, Zhao, and Kumar 2003). This component describes potential damage states such “no damage” and “collapse” associated with a given earthquake magnitude earthquake. Vulnerability is a potential for loss from hazard events (Mileti 1999, Yucel and Arun 2012). To assess physical earthquake risk, the vulnerability component is a key element along with the hazard and exposure components. Vulnerability models are used to express physical damage and economic losses based on an earthquake’s intensity and the amount and type of property exposed (i.e. the exposure component) (Ryu and Edward 2015). In an earthquake risk assessment, physical vulnerability refers to the probabilistic distribution of loss at a certain intensity level. Physical vulnerability is often delineated as functions that can be derived using the losses from past events at certain locations. These functions refer to the probability of exceeding limit states such as a certain damage threshold (e.g., limited, moderate, extensive, catastrophic) or injury levels (Crowley and Silva 2013, Ryu and Edwards 2015).

2.2 Modeling Potential Earthquake Impacts from a Social Perspective

2.2.1 Social Vulnerability as a Concept

Understanding losses and social vulnerability to an earthquake is a key step for earthquake hazard mitigation, planning, and risk reduction. In the recent Haiti earthquake (2010), for instance, economic losses were too vast for the country to recovery effectively. Total economic losses were more than 120% of the nominal GDP of the country. Moreover, approximately 300,000 people were injured (Silva et al. 2014a). Here, the vulnerability of Haiti varied and was dependent on a number of factors where a similar shaking intensity throughout the affected area likely caused different impacts in different places due to difference in the vulnerability of populations (Blaikie et al. 1994). A 1976 earthquake that occurred within Guatemala is a prominent example of how an earthquake impact varies with socio-economic factors. People living in undeveloped and unplanned housing were the victims of the highest residential damage. On the other hand, middle class people that sustained similar damages were able to recover their losses more easily. The Guatemala earthquake explicitly identified the influences of social components on losses, and it was aptly named the “Class-quake” (Blaikie et al. 1994).

In essence, the “Class-quake” indicated that vulnerability to the earthquake hazard varied by social class, and is thus, partially a social factor. In the context of natural hazards and disasters, social vulnerability is the potential within social systems for losses or harm due to natural hazards. Social vulnerability changes with time and space (Cutter, Boruff, and Shirley 2003), and three terms are often used to help describe the concept. These are exposure, resistance and resilience. Exposure relates to the physical location and environmental characteristics of a particular place that makes populations vulnerable, resistance is the protection ability from the hazard and resilience means the capability of recovery after a damaging hazard impact or disaster (Cutter et al. 2006). As the definition posits, social vulnerability influences negative impacts on an individual or group of people due to hazard exposure.

In the context of natural hazards and disasters, social vulnerability increases the potential for losses due to characteristics within social systems. This loss is mainly the outcome of interacting social and biophysical factors (Cutter 1996). Different types of hazards in different places cause variation in vulnerability. There are three classes to categorize vulnerability in the literature. These are: (a) vulnerability as risk/hazard exposure; (b) vulnerability as social response; and (c) vulnerability of places (Cutter 1993). Vulnerability as a pre-existing condition (vulnerability as risk/hazard exposure) includes bio-physical vulnerability and refers to the distribution of natural hazards and human occupancy in hazard zones. Vulnerability as social response includes the social response to hazards. It focuses how various factors of the society such as historical, social, and cultural characteristics can play a role in adaptation to hazard events. Vulnerability of places is the result of both biophysical risk and the social vulnerability of populations. The combination of the two concepts at a particular location is referred to as place-based vulnerability. Measurements of place-based vulnerability can help to identify which places and populations are most vulnerable (Cutter 1996). A number of theoretical frameworks have been developed to explain both vulnerability and place-based vulnerability. These are briefly outlined below.

2.2.2 Conceptual Frameworks on Social Vulnerability to Hazards and Disasters

Vulnerability depends on many factors. The most commonly cited frameworks that are used to explain and assess the concept include the Pressure and Release Model (Blaikie et al. 1994), the Vulnerability/Sustainability Framework (Turner et al. 2003), the Integrated Risk

Assessment (Burton and Silva 2016), the Disaster Resilience of Place (DROP) Model (Cutter et al. 2008) and the Vulnerability of Place Model (Cutter 1996). These models are important within the context of this research because they focus on the coupled human, social, and physical conditions of vulnerability.

2.2.2.1 Disaster Pressure and Release Model (PAR)

The vulnerability of a community depends not only a natural hazard occurrence, but also on the socio-economic processes in a hazard zone. According to Disaster Pressure and Release Model (PAR) (Blaikie et al. 1994), pressure increases with the increased vulnerability of the community and the effects of the hazard. In this model, human conditions like limited access to resources, power, structures and human ideologies are basic reasons behind vulnerability and increased pressures. The term "Release", that is expressed by the reduction of pressure caused by hazards and vulnerability, is expressed by factors that reduce vulnerability and risk. The model (Figure 1) describes how vulnerability arises from root causes and then progression continues to dynamic pressure to unsafe conditions. The PAR model focuses on social vulnerability only and leaves out the impacts of the physical event itself. This is a remarkable disadvantage of this model. According to Cutter et al. (2009), this model is more theory based than useful for empirical analysis. (Cutter et al. 2009).

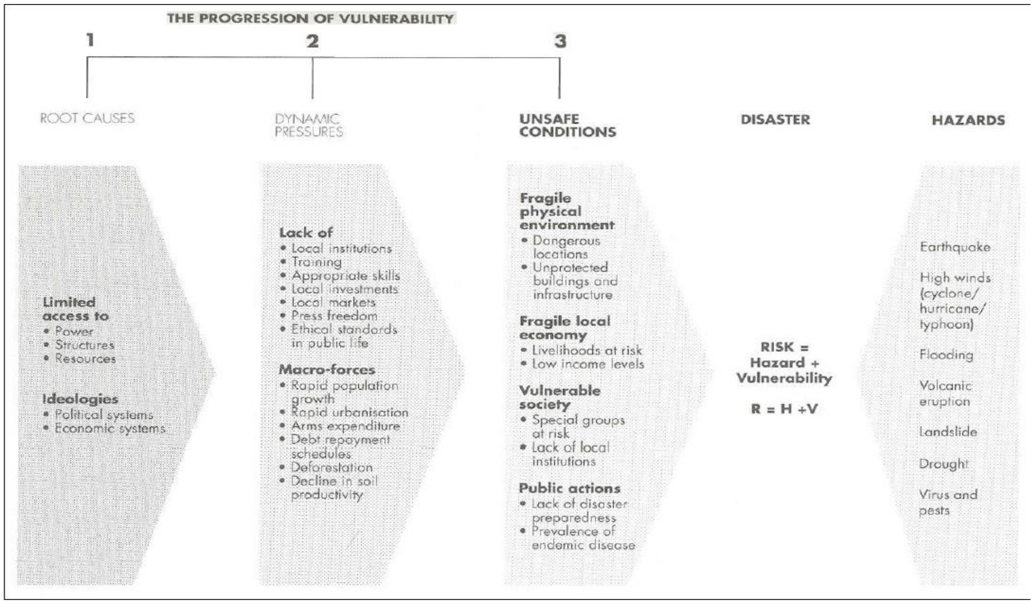


Figure 1: The Pressure and Release model (Blaikie et al. 1994)

2.2.2.2 Vulnerability/Sustainability Framework

The vulnerability concept of Turner et al. (2003a) shows vulnerability to natural hazards and disasters from a global environmental change perspective. This concept measures vulnerability at the global level using local vulnerability measures. This model comes from the combination of two archetypal models regarding vulnerability. They are the Risk the Hazard (RH) and Pressure and Release models (PAR). The Risk Hazard portion of the model reveals how hazard impacts, exposure and vulnerability are interrelated (Figure 2). The Pressure and Release portion of the model exhibits the progression of vulnerability from root causes to unsafe conditions as mentioned in the previous section. The pressure and release portion of the model refers to how vulnerability influences the occurrence of disasters by social vulnerability interacting with natural hazards (Figure 3).

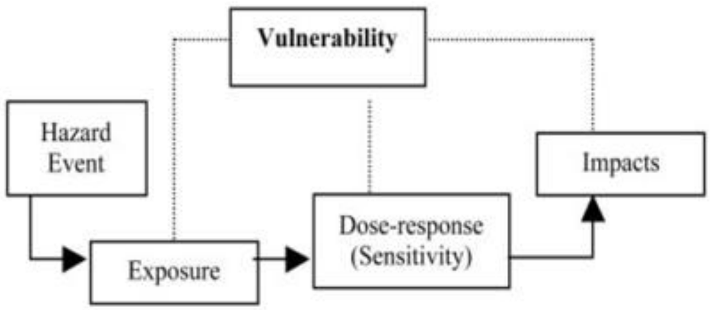


Figure 2: Risk hazard framework; Chain sequence is from hazard to impacts; vulnerability indirectly noted by dotted lines (Turner et al. 2003).

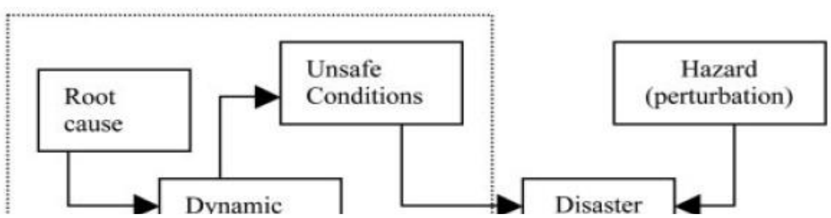


Figure 3: Pressure and release framework (common to risk research) with emphasis focused on "social" conditions of exposure; vulnerability usually marked very clearly (Turner et al. 2003).

The finalized vulnerability framework (Figure 4) is a representation of hazard vulnerability in human-environmental systems (exposure, sensitivity, resilience). According to Cutter et al. (2009), the major disadvantage of this model is that there is no way to identify the start point and end point of the vulnerability of a particular location. Moreover, the model is better suited for qualitative analysis rather than empirical analysis (Cutter et al. 2009).

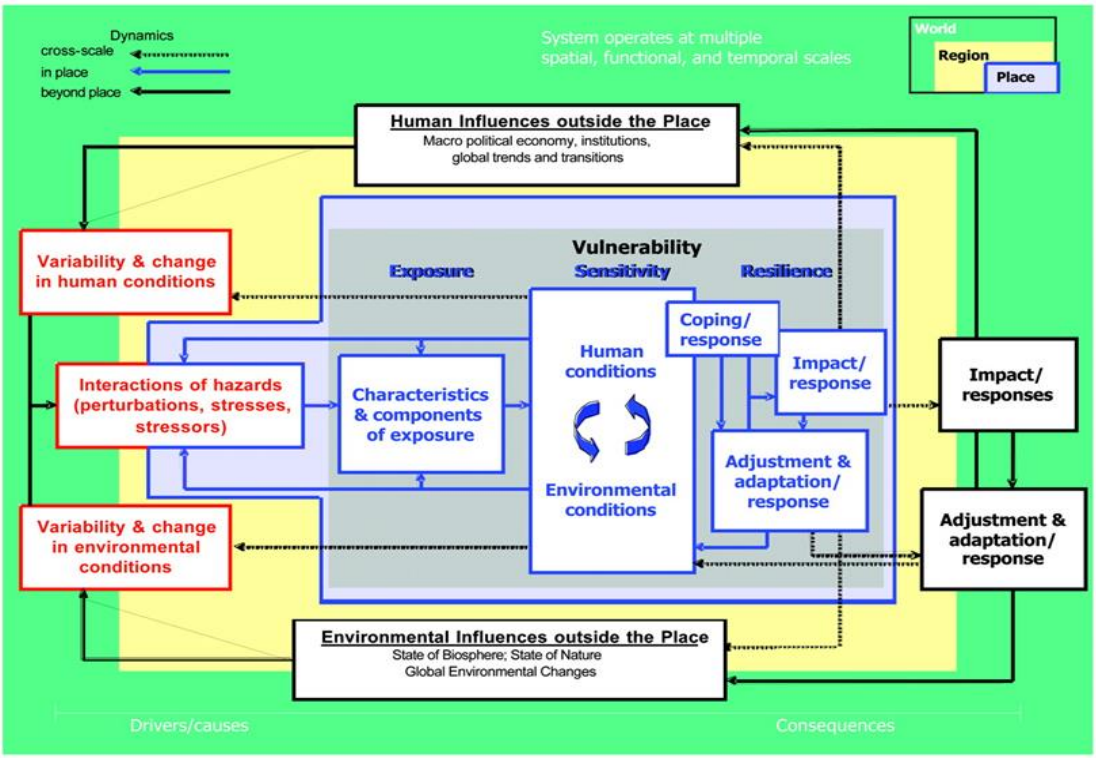


Figure 4: Exposure, sensitivity, and resilience components of the vulnerability framework (Turner et al. 2003)

2.2.2.3 Integrated Risk Assessment Framework

The integrated Risk Assessment framework was developed to better understand earthquake risk as a combination of social and physical risk characteristics. Here, the earthquake impact potential of a place is delineated by integrating a seismic hazard assessment, a physical risk assessment, and human dimensions of a respective hazard zone (Burton and Silva 2016). Burton and Silva (2016) build upon the Vulnerability of Place Model (Cutter 1996) (discussed below) for the assessment of integrated risk in Portugal (Figure 5). The GEM OpenQuake-engine and Integrated Risk Modelling Toolkit (Burton and Tormene 2018) was utilized to calculate risk for the country accounting for: A) seismic hazard potential, B) geographic context in the way of population and infrastructure exposure modeling, C) physical risk through calculations of potential building loss, D) social fabric by the exploration of social characterizes that drive vulnerability, E) social vulnerability using a social vulnerability index, and F) integrated risk by mathematically combining a quantified social vulnerability and risk assessment. In this model, it is unknown whether qualitative (survey-based) data can be incorporated into the framework. Moreover, more research is needed on how to effectively combine physical risk estimates with estimates of social vulnerability in this model.

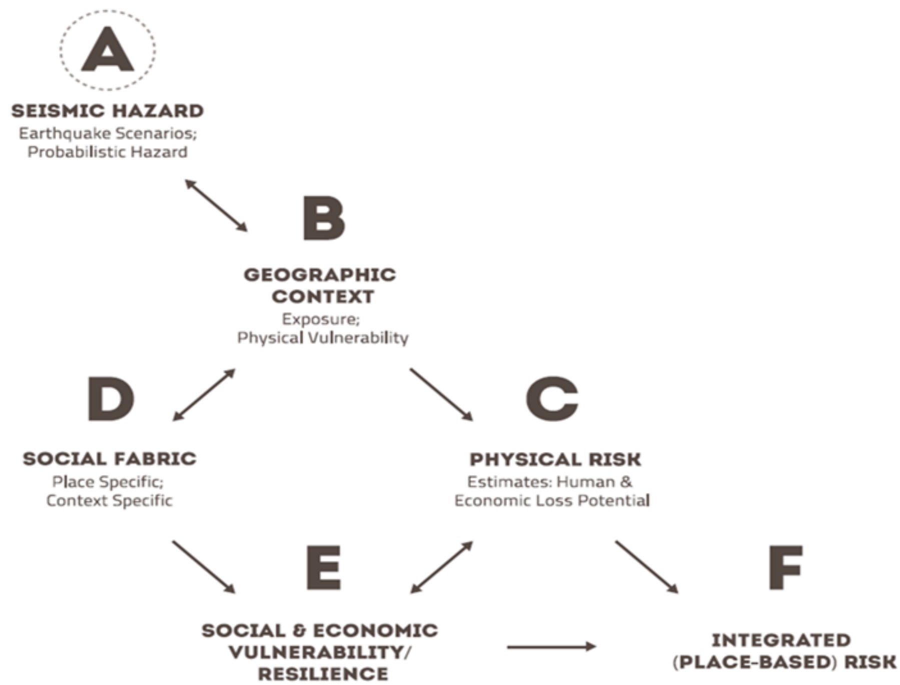


Figure 5: Integrated Risk Assessment Framework (Burton and Silva 2016)

2.2.2.4 The Vulnerability of Place Model

In 1996, social vulnerability scientist Susan L. Cutter first proposed the Vulnerability of Place Model (VPM) (Figure 6) (Cutter 1996). This model posits that the overall vulnerability of a place is based on both the social and biophysical vulnerability of that place. The total vulnerability of a place is therefore a function of biophysical and social factors. Cutter, Mitchell, and Scott (2000) demonstrated the utility of the combination of social and biophysical factors in a GIS to measure the overall place vulnerability of Georgetown County, South Carolina. The vulnerability of census blocks in Georgetown County was modelled considering twelve environmental and eight social factors. The study found the interaction of two types of vulnerability in a common place. Biophysical vulnerability included hazard frequency and the locational impacts of past hazard events. Then to calculate the overall vulnerability of the place, a social vulnerability layer (using a calculated index of social vulnerability) was combined with the estimate of biophysical vulnerability in a GIS. The study showed that in most cases the areas which had high biophysical vulnerability did not intersect with the areas of high social vulnerability. Instead, the total vulnerability of Georgetown County results when medium levels of biophysical vulnerability intersected with medium to high levels of social vulnerability (Cutter, Mitchell, and Scott 2000). The Georgetown study concepts of vulnerability are essential to this thesis.

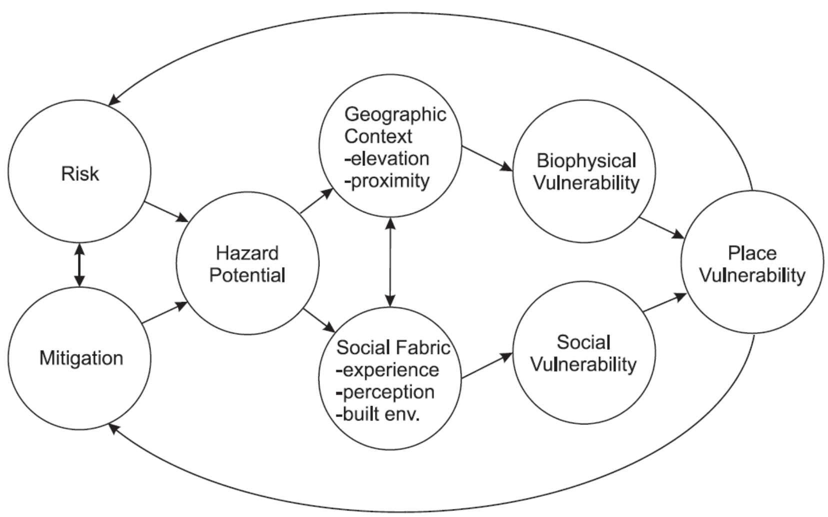


Figure 6: Vulnerability of Place Model (Cutter, Boruff, and Shirley 2003)

2.2.2.5 Disaster Resilience of Place (DROP) Model

Cutter et al. 2008 proposed the Disaster Resilience of Place (DROP) Model (Figure 7) which describes the concept of natural disaster resilience (Cutter et al. 2008). This model is based on three assumptions to represent how social vulnerability and resilience are related. Firstly, the model addresses natural hazards specifically, but also may be adapted to other rapid onset hazards (i.e. terrorism, technological hazards) or slow onset natural hazards (i.e. drought). Secondly, the model focuses on community level resilience. Finally, the main focus of the model is social resilience of places. This means the other forms of resilience cannot be isolated from the social aspects. Thus, resilience is an inherent or antecedent condition or process according to DROP model representation. The starting point of this model is antecedent conditions which are a product of processes occurring within societal, natural and built environment. Antecedent condition refers to both inherent vulnerability and resilience. This concept represents how the inherent process occurs at the local scale and results at the broader scale (Figure 7). Antecedent conditions interact with hazard characteristics such as frequency, duration, intensity, magnitude etc. This interaction produces hazard effects which might be reduced by implementing proper responses. This effects reduction leads to a high recovery. (Cutter et al. 2008).

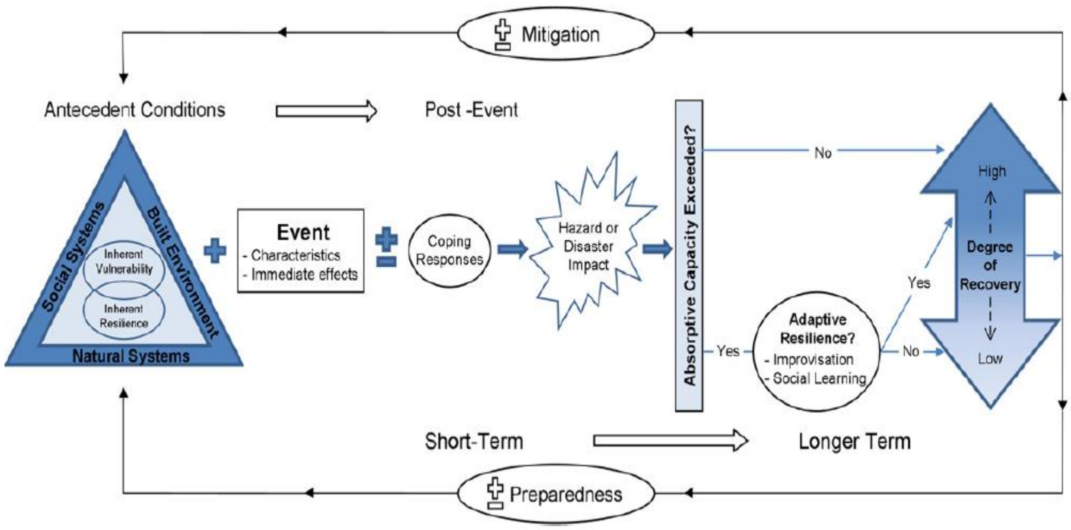


Figure 7: Disaster Resilience of Place (DROP) Model (Cutter et al. 2008)

2.3 GIS and Quantified Social Vulnerability

The Federal Emergency Management Agency (2003) focuses on residential building vulnerability caused by natural hazards such as floods, hurricanes, storm surges, and earthquakes. Social aspects of vulnerability analysis using GIS is often ignored in these modelling efforts because social vulnerability is not observable directly (Burton 2010). Therefore, physical aspects such as hazard severity have been the focus of the majority of GIS-based analysis. In many recent researches, however, social vulnerability has been quantified by using composite indicators (Cutter, Boruff and Shirly 2003, Burton 2010, Boruff, Emrich, and Cutter 2005). Geographic Information Systems (GIS) and geostatistical applications (i.e. GeoDa, SPSS) have made social vulnerability analysis robust and meaningful as differential vulnerabilities can be demonstrated across space. These software and tools are used to analyze and visualize social vulnerability and to focus on changes in the vulnerability of places due to the spatial variability of indicators used in the modelling process (Burton 2010, Chakraborty, Tobin, and Montz 2005).

2.4 Summary

Vulnerability depends on many factors. The most commonly cited frameworks include the Pressure and Release Model (Blaikie et al. 1994), the Vulnerability/Sustainability Framework (Turner et al. 2003), the Integrated Risk Assessment (Burton and Silva 2016), the Disaster Resilience of Place (DROP) Model (Cutter et al. 2008) and the Vulnerability of Place Model (Cutter 1996). These models are important within the context of this research because they focus on the interaction between social and physical components of earthquake risk. Within these models, natural and societal factors contribute to the vulnerability of populations to hazards and their ability to mitigate against them. Some factors increase vulnerability such as lack of access to wealth and resources, poverty, and illiteracy, and some factors like access to wealth, education, high living standard decrease vulnerability. This research sought to adopt an approach for modeling earthquake risk in a manner in which social and physical dimensions are quantifiable. Since the Burton and Silva (2016) framework was developed explicitly for earthquake assessments using quantified indicators, it was selected as the theoretical model that will guide this research.

Chapter 3: Research Design & Methodology

3.1 Study area

The study area of this research project is FEMA Region IV which encompasses the Southeastern U.S. and includes eight states: Alabama, Florida, South Carolina, North Carolina, Mississippi, Kentucky, Georgia, and Tennessee (Figure 8). The work was conducted at the U.S. Census County level of geography since counties are a main administrative unit in the U.S. for the application of major disaster risk reduction policies.

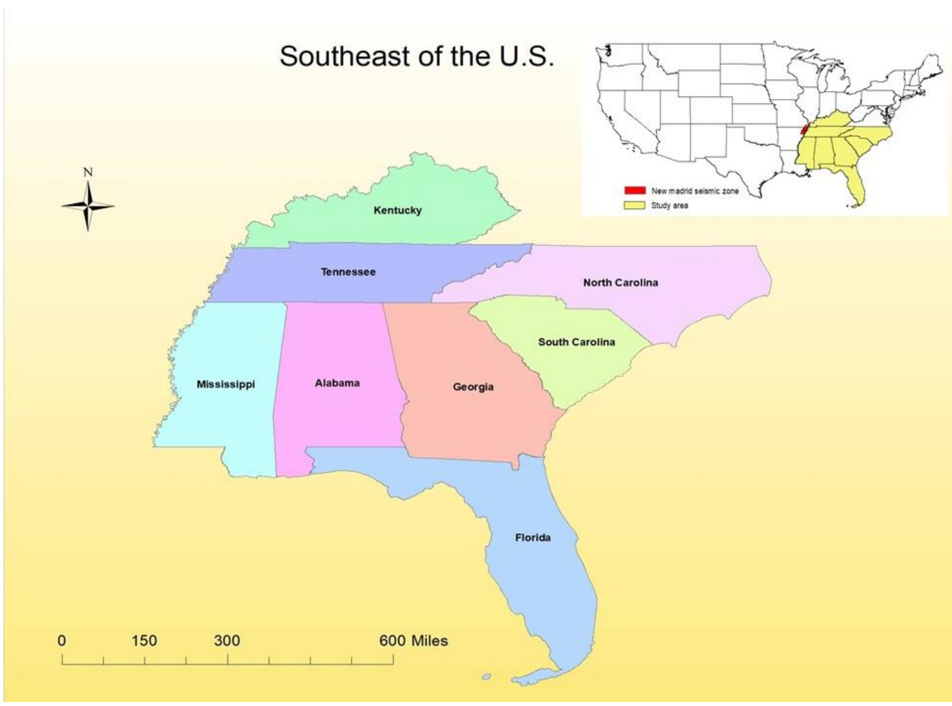


Figure 8: Study Region

Although much of the land area in FEMA Region IV is far from the New Madrid Seismic Zone, there is still considerable seismic activity throughout the region. For instance, two devastating earthquakes were observed in this region, respectively in Charleston, South Carolina (1886) and Mineral, Virginia (2011) (Biryol et al. 2016). The New Madrid seismic zone is one of the most active faults in the Southeast of the U.S., and previous research has found that large earthquakes reoccur on average at 450-year intervals (Tuttle et al. 1999). Another devastating earthquake in the Southeast occurred in 1886 in Charleston, South Carolina. Most of the buildings in the city were damaged during the event, and the overall amount of loss from that seismic hazard was about 24 million dollars (1886 dollar value) (Schmidtlein et al. 2011). Brick and stone made buildings were all effected (Figure 9); whereas wooden structures were subjected to less damage. Roads were also very badly damaged. The location in proximity to the fault was not the main reason behind this devastation. Likely, socio-economic factors were among the primary contributing factors to the loss.

The FEMA Region IV study region contains some of the more vulnerable populations in the entire nation. The demographic characteristics of this region (Table 1) help to describe its vulnerabilities from a social perspective.

Table 1: Demographic characteristics of the region

State	%African American	% Hispanics	Per Capita Income annually	% Age dependency
AL	26.178	3.657	23020.72	20.142
FL	15.956	21.500	26547.01	23.047
GA	30.456	8.239	25184.24	17.742
KY	7.778	2.806	22554.16	19.832
MS	37.016	2.537	20036.72	19.929
NC	21.484	7.862	24801.26	19.570
SC	27.904	4.755	23486.97	20.197
TN	16.661	4.256	23756.27	19.875

Table 1 demonstrates that Florida and Georgia enjoy higher incomes than states such as Mississippi and Kentucky. This differential income could affect the poorer state’s ability to mitigate and respond to an earthquake. The places with more income are likely to have higher losses due to a large number of assets in harm’s way. Age dependency is the highest in Florida which could lead to the state’s residents being more vulnerable to hazards because older age can affect mobility out of harm’s way, income, and the ability to recover. Race is another characteristic which describes the vulnerability of populations. The percentage of Hispanics is higher in Florida due to migration from Cuba. Race is also a contributing factor to social vulnerability to hazards. A Hispanic or African American resident, for instance, may have different education levels (e.g., lower levels of education or poorer access to education), lower incomes, a high degree of marginalization, and a lack of language proficiency which will affect hazard preparedness, response, and recovery.

Figure 9: The post-earthquake damage after 1886, Charleston, South Carolina. Originally photographed by J. K. Hillers in 1886. Source: USGS

3.2 Data and Methods

The modelling and respective methodology for this thesis consists of three major components in which an integrated assessment of earthquake risk was developed for the entirety of the study area (Figure 10). The first is a physical risk model to delineate loss potential to buildings from earthquakes in the study area. A social vulnerability index is the second component. The third component, an integrated risk index, was the result of the convolution of the physical risk estimates and social vulnerability indicators. The combination of these results in what is referred to in the literature as an integrated or holistic risk assessment (Carreño et al. 2007, Burton and Silva 2016). Here, hotspot areas defined where both social and physical vulnerabilities are high, and which could require highest priority in hazard mitigation and disaster risk reduction approaches and policies. Integrated/holistic risk assessments have been conducted for Bogota (Colombia) and Barcelona (Spain) (Carreño et al. 2007) and for mainland Portugal (Burton and Silva 2016), to name just a few places.

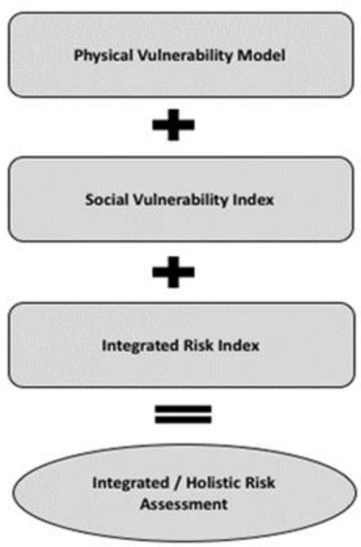


Figure 10: Integrated Risk Methodology Workflow

3.2.1 Modeling Direct Loss

The starting point for the estimation of the region's risk from an integrated perspective is the modelling of earthquake losses (i.e. physical risk). For the modeling of potential losses, the HAZUS risk assessment platform HAZUS-MH-IV (an extension of ArcGIS) was utilized. HAZUS provides the ability to calculate average annualized loss statistics for U.S. states, counties, and census tracts. HAZUS includes a national building inventory that makes up its exposure database that includes: 1) essential facilities (police, fire, emergency operations facilities, school, medical facilities); 2) lifelines (utilities and transportation); 3) the general building stock (residential, commercial and industrial); and 4) some demographic data (FEMA 2018).

In this thesis, property losses to residential and commercial infrastructure and losses of life were estimated for each county in FEMA region IV. Due to time constraints, this research was conducted using a HAZUS Level 1 analysis that is based on data from national databases within the HAZUS platform and includes historic hazard events, demographic data, and estimates of the location and costs of the building stock. Here, losses were estimated probabilistically utilizing existing seismic source models within HAZUS-MH and a set of ground motion prediction equations for the region that were also found in HAZUS-MH. For what concerns the physical vulnerability of the exposed building infrastructure, a set of vulnerability functions were employed for all building typologies relevant to the study area (e.g. wood, masonry, and reinforced concrete). Economic losses were calculated probabilistically as direct losses in the form of average annual loss for a 500-year return period. Direct losses refer to the physical or structural impact caused by a disaster such as the destruction of infrastructure from ground shaking (Erdik et al. 2011). A probabilistic model was favored over a deterministic model because a probabilistic model accounts for all possible magnitudes of events that can occur in the region, rather than simulating a single historic event. Casualties that are as a function of structural and non-structural building damage were also calculated. HAZUS provides a methodology for calculating casualties that based on the correlation between building damage and the number and severity of casualties that could occur in a given area (Jaiswal et al. 2017). Only casualties that are a result of direct damages were accounted for in this research.

3.2.2 Modeling Social Vulnerability

A number of factors that influence the social vulnerability of populations can be measured. This thesis used an indicator-based approach to measure the social vulnerability of FEMA Region IV's counties. An indicator is a quantitative or qualitative measure derived from observed facts that simplify and communicate the reality of a complex situation (Burton and Silva 2016). A composite indicator is the mathematical combination of individual indicators or thematic indicators sets that represent different dimensions of a concept that cannot be fully captured by any individual indicator alone. The Social Vulnerability Index (SoVI) is a composite indicator that was developed to model social vulnerability (Cutter, Boruff and Shirley 2003) within the U.S. This method was adopted for this research due to its ubiquity in the peer reviewed literature and due to the ability of the social vulnerability modeler to better understand potential driving factors of social vulnerability.

The original SoVI utilized 42 variables. These variables refer to the demographic, economic and built environment variables which are contributing factors to a community's ability to respond to and recover from a disaster according to the reviewed literatures (Cutter, Boruff, and Shirley 2003, Molas et al. 2012, Burton 2010). These variables include measures of: 1) age characteristics since extremes of age affect movement out of harm's way and often results in financial constraints (Bolin and Stanford 1991, Cutter et al. 2000, Cutter, Boruff and Shirley 2003, Hewitt 1997, Pulido 2000); 2) socioeconomic status since income and socioeconomic status affects the ability of populations to absorb losses and enhance resilience to hazard impacts (Blaikie et al. 1994; Hewitt 1997; Cutter et al. 2000; Cutter, Boruff and Shirley 2003); 3) gender status because in many countries, including the U.S., woman often have a more difficult time during recovery then men, often due to sector-specific employment, lower wages, and family care responsibilities (Blaikie et al. 1994, Hewitt 1997, Cutter 1996, Cutter, Boruff and Shirley 2003); and race and ethnicity due to the potential of language and cultural barriers that affect access to post-disaster funding and the location of populations in high hazard areas (Bolin and Stanford 1998, Pulido 2000, Cutter, Boruff and Shirley 2003).

These variables were input into a Principal Components Analysis (PCA), more specifically a Factor Analysis (FA) that reduced the original 42 variables into 7 factors used to explain the driving characteristics of social vulnerability within the United States. These factors were summed to derive composite social vulnerability scores for each county in the nation. For this research, 28 out of the original 42 SoVI variables were used. This reduction in indicators was due to data availability within the current U.S. Census as well as improvements to the SoVI indicator selection over time (HVRI 2015). The source of all data to measure social vulnerability was the U. S. Census Bureau American Factfinder (U.S. Census Bureau 2017). These variables are delineated in Table 2.

Table 2: SoVI variables

Median Age	Percent of household collecting social security benefits
Percent Black	Percent speaking English as a second language
Percent Native American	Percent of population without health insurance
Percent Asian and Hawaiian Islanders	Percent civilian unemployment
Percent Hispanic	Percent of households earning $200,000 or more
Percent of population under 5 years or 65 and over	Percent living below poverty level
Median Value of Owner Occupied Housing Units	Percent employment in extractive industries

Percent of children living in married couple families	Percent of housing units that are mobile homes
Percent renter	Percent of population with no high school diploma or less than 12^{th} grade education
Percent residents in nursing homes	Percent of housing units with no car
Percent female population	Percent females participating in the labor force
Percent female headed households	Percent unoccupied housing units
Per capita number of community hospitals	Median Gross Rent
People per unit	Percent Employed in service industry
Per Capita Income (in dollars)	

The SoVI methodology includes standardization of all variables as percentages, per capita or density functions. Once this step was accomplished, the data was converted to Z-scores and input into a Principal Component Analysis (PCA), more specifically a Factor Analysis (FA). An FA is a type of PCA which reduces the total number of variables into a smaller set of common factors that explain the driving components of social vulnerability in the study area. These reduced factors are uncorrelated and suitable for using in various statistical analysis. In this thesis, the number of dimensions in the FA were reduced using the varimax rotation & Kaiser Criterion (Cutter, Boruff and Shirley 2003). Following rotation, each factor was categorized and named based on the contribution of its variables to the vulnerability of populations. This was accomplished by highlighting variables in which correlations between each variable and its respective factor exceeded >=0.500 and <=-0.500. In other words, if correlations between a given factor and its variables include high loadings on percent of the population in poverty, percent African American populations, and the percent of population living in mobile homes, one might name the factor “Race and Socioeconomic Status” and determine a positive directionality for the factor. Here, a positive directionality was assigned to the factors which increase vulnerability and a negative directionality was assigned to the factors which decrease vulnerability by multiplying that factor by -1. All factors were then put into an additive model to sum the factors to derive and map a final SoVI score (Burton 2010, Jackson 2005).

Once derived, SoVI scores were mapped as standard deviations from the mean in order to highlight areas of potential management concern. Positive deviations from the mean were utilized to indicate potentially more vulnerable areas, comparatively, whereas negative deviations were used to illustrate less socially vulnerable populations, comparatively. Final social vulnerability scores were mapped in addition to each factor score. The mapping of the factor scores allowed the spatial representation of potential contributing factors to the social vulnerability of populations in the Southeastern United States to be mapped and analyzed.

3.2.3 Integrated Risk Model

The integrated risk model is the result of the convolution of physical risk estimates from HAZUS-MH with the social vulnerability model. The goal of the integrated risk assessment portion of this thesis is to provide an overview of an end-to-end earthquake risk assessment for FEMA Region IV that is integrated and holistic. Carreño et al. (2007), Carreño et al. (2012), and Burton and Silva (2016) provide the aggregation method that was adjusted for this work due to its mathematical simplicity. In this method, the risk and social vulnerability model outputs are made commensurate through MIN-MAX rescaling or conversion or Z-scores. The direct potential impact of an earthquake (in a general sense) is denoted as $R_T = R_F\left(1+F\right)$ where R_T is a total risk index, R_F is a physical earthquake risk index, and F is the composite social vulnerability index which might be described as an aggravating coefficient of the estimated loss.

This thesis simplifies Carreño's aggregation method where a risk matrix was constructed to calculate a final integrated risk score. In this matrix, the estimated earthquake losses and the SoVI results were normalized based on standard deviations from the mean. Classes that are at the negetive end of the scoring spectrum (< 0.0 standard deviations), moderate (>=0.0 to +0.5 standard deviations), and high end (>+0.5 standard deviation) were assigned values of -1, 0 and +1 respectively. An additive model was then created where the scaled social vulnerability scores and the physical impact scores were summed to create final values ranging from -2 to +2. The combined values were mapped into the five classes (-2, -1, 0, +1, +2). These five classes represent places with the lowest social vulnerability and lowest potential loss (i.e. a -1 in social vulnerability and -1 in hazard loss in a value of -2) to the most vulnerable and most exposed (+2).

3.3 Statistical analysis

In addition to the physical risk, social vulnerability, and integrated risk modelling, a series of multivariate regression models were calibrated in an attempt to find a statistical link between the estimated losses and social vulnerability. It is important to note, however, that all earthquake losses are probabilistic estimations. As such, the regression modelling effort was not conducted to make interferences regarding the social vulnerability concept's association with earthquake damages and loss or to make predictions. Rather, the regression models were calibrated to test the assumption that some social vulnerability variables might be able to be used as proxy variables to delineate potential adverse impacts from earthquakes. Here, Ordinary Least Square (OLS) regression models were calibrated where the factors of the SoVI index were input as independent variables and predicted damages were input into the model as dependent variables using:

$$P_i = x_0 + x_1FAC1_i + x_2FAC2_i + x_3FAC3_i + x_4FAC4_i + x_5FAC5_i + x_6FAC6_i + x_7FAC7_i + \mathcal{E}_i \ldots\ldots(1)$$

where,

$P_{i\,=}$ damaged building structures per county and FAC1- FAC7 represent the SoVI factors derived using the social vulnerability index. To explore the potential contribution of the social vulnerability factors to casualties, a second OLS regression model was calibrated using a similar equation where the social factors are the independent variables and casualties is the dependent variable.

3.4 Summary

The integrated risk modelling approach established for this thesis includes a social vulnerability model and a probabilistic damage assessment for earthquakes in the Southeastern United States. All variables and parameters were calculated at the county level within the study region. In an effort to model how risk from earthquakes varies across space, earthquake risk, social vulnerability, and integrated risk models were developed. These were then input into a series of spatial and statistical analyses. Figure 11 provides a workflow and an overview of the methods conducted for this research.

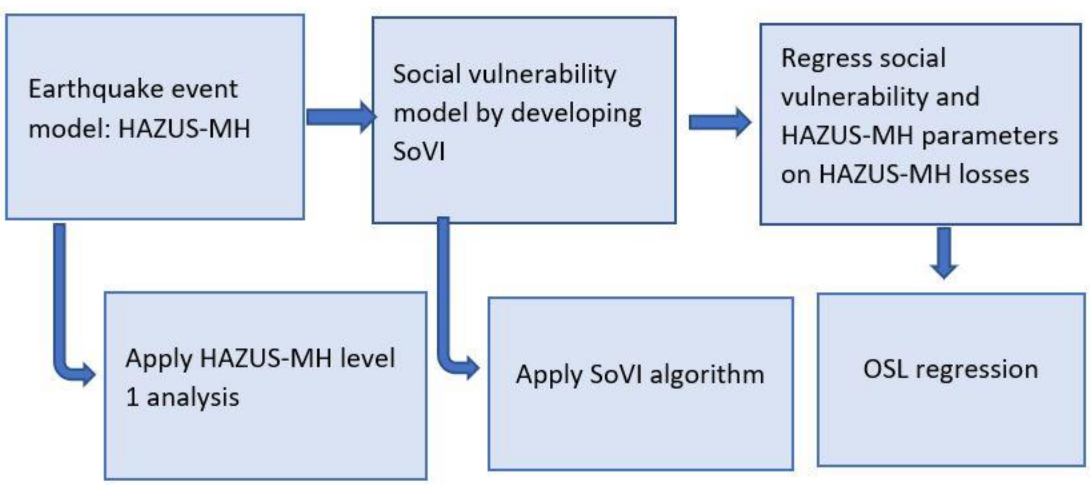


Figure 11: Flowchart of analysis methodology adapted from Schmidtlein et al. 2011

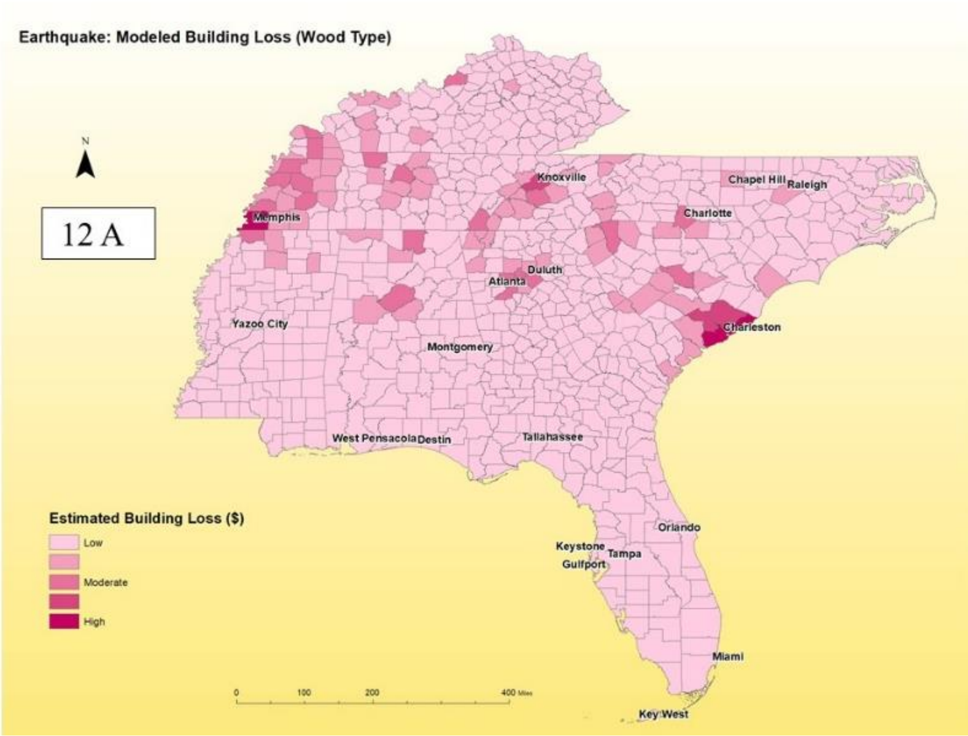
Chapter 4: Results

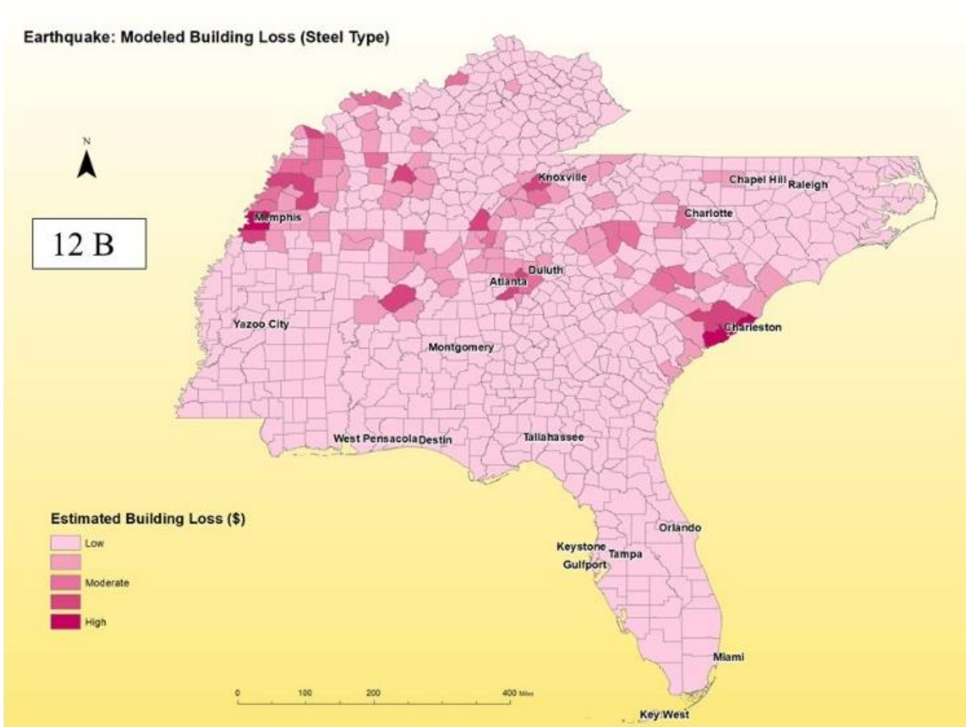
The modelling and associated methodology for this thesis consists of three major components in order to develop an integrated assessment of earthquake risk. The first is a physical risk model to delineate loss potential to buildings from earthquakes for the study area. A social vulnerability index is the second component. The social vulnerability index was constructed to describe characteristics within social systems that create the potential for loss. The third component, an integrated risk index, is the result of the convolution of the physical risk estimates and social vulnerability indicators. The combination of these result in what is referred to as an Integrated Risk Assessment. The results of each of the component parts is explained in detail in the sub-sections below.

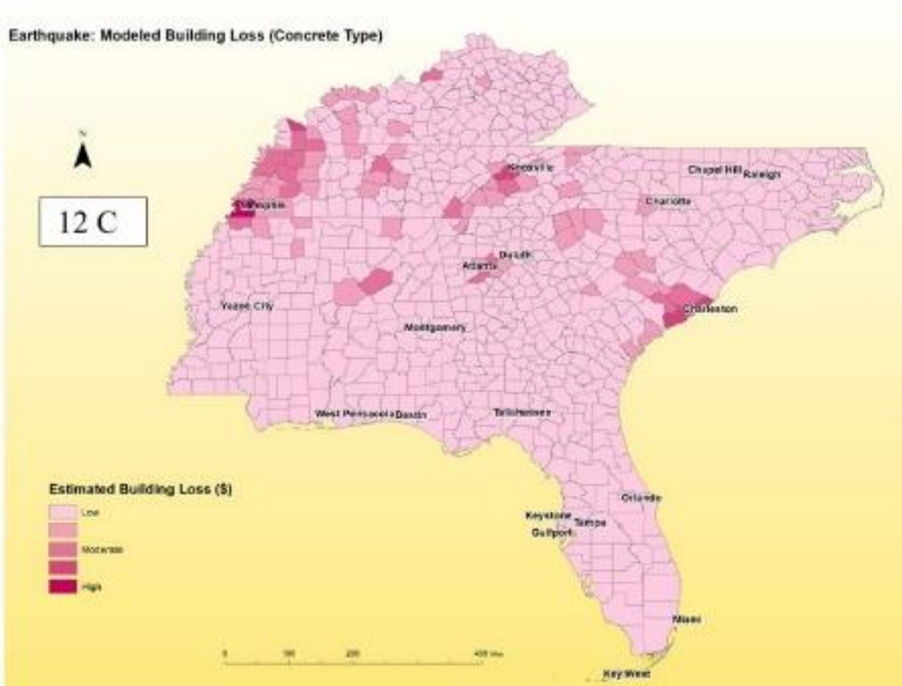
4.1 Modeled Estimated Building Loss

Earthquakes can produce structural damage and cause complete building loss. The direct economic losses (annual average losses) from the HAZUS-MH model for each county range from approximately $6,000 to $2,500,000. Figure 12 A-E demonstrates the spatial distribution of direct economic losses caused by different building types from earthquakes (i.e. wood, steal, concrete, masonry, mobile homes). Figure 12 A is the spatial distribution of direct economic loss of wood type buildings. Charleston county (South Carolina) has the highest amount of estimated economic loss in dollar value based on this building type. Average annual losses for wood buildings is also high in Tennessee (Shelby and Knox counties) and South Carolina (Charleston and Berkeley counties).

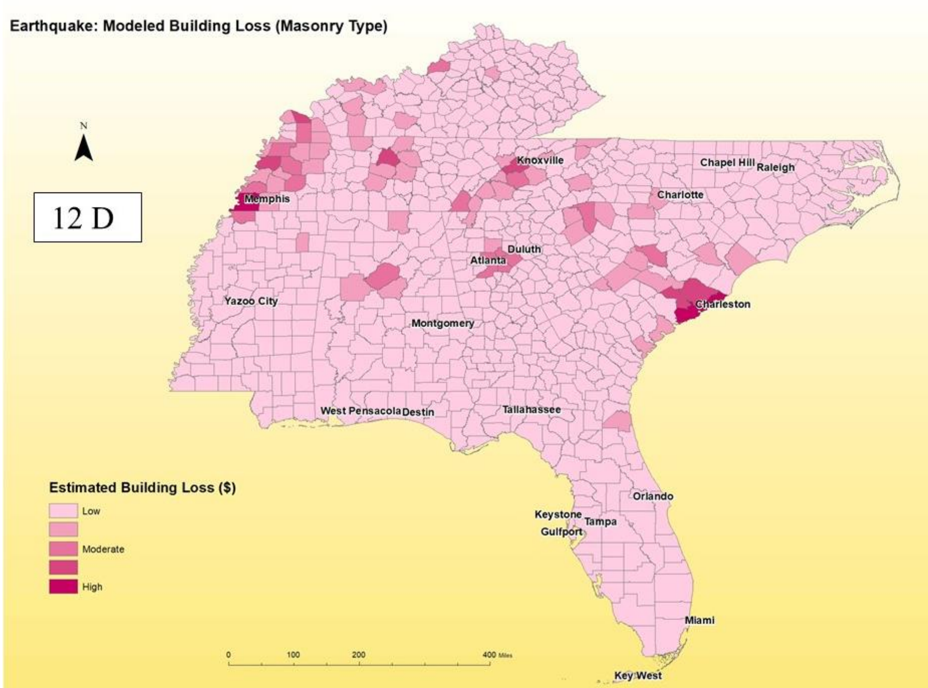
Figure 12 B and 12 C represents estimated building losses for steel and concrete building types. The highest potential losses are observed in McCracken (KY) and Knox (TN) counties along with Shelby County (TN) and Charleston County (SC). Figures 12 D and 12 E represent masonry building types and mobile homes. For masonry, Shelby County (TN), Charleston County (SC), Knox County (TN) and Berkeley County (SC) have a high potential for loss. For mobile homes, there is high potentiality of loss in Tennessee (Obion, Dyer, Tipton and Shelby Counties).




Earthquake: Modeled Building Loss (Steel Type)
N
12 B
Memphis
Knoxville
Chapel Hill
Raleigh
Charlotte
Duluth
Atlanta
Yazoo City
Charleston
Montgomery
West Pensacola
Destin
Tallahassee
Estimated Building Loss ($)
Low
Moderate
High
Orlando
Keystone
Tampa
Gulfport
Miami
0
100
200
400 Miles
Key West



Earthquake: Modeled Building Loss (Concrete Type)
12 C
Knoxville
Chapel Hill
Raleigh
Charlotte
Memphis
Atlanta
Duluth
Yazoo City
Charleston
Montgomery
West Pensacola
Destin
Tallahassee
Estimated Building Loss ($)
Low
Moderate
High
Orlando
Keystone
Tampa
Gulfport
Miami
Key West



Earthquake: Modeled Building Loss (Masonry Type)
N
12 D
Knoxville
Chapel Hill
Raleigh
Charlotte
Memphis
Duluth
Atlanta
Yazoo City
Charleston
Montgomery
West Pensacola
Destin
Tallahassee
Estimated Building Loss ($)
Low
Moderate
High
Orlando
Keystone
Tampa
Gulfport
Miami
0
100
200
400 Miles
Key West

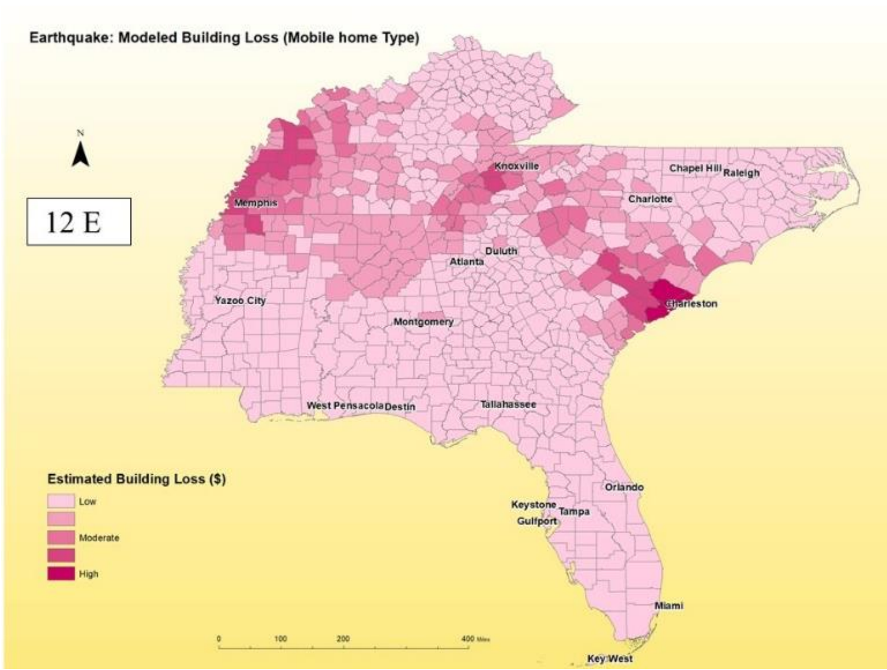


Figure 12 (A-E): Spatial distribution of losses for wood, steel, concrete, masonry, and mobile home building types

4.2 Casualty Modeling

Figure 13 represents the total casualties due to earthquakes in the study region. High casualties were estimated in both western and eastern Tennessee and in eastern South Carolina. The highest estimated causalities for an individual county were in Charleston (SC) and Shelby County (TN), where damaging earthquake events have been experienced before. In Mississippi, central and southern Alabama and Georgia, and in Florida, casualties are negligible. This is because these areas are too far from the earthquake fault to experience ground-shaking strong enough to cause building collapse and other damage that could injure or kill.

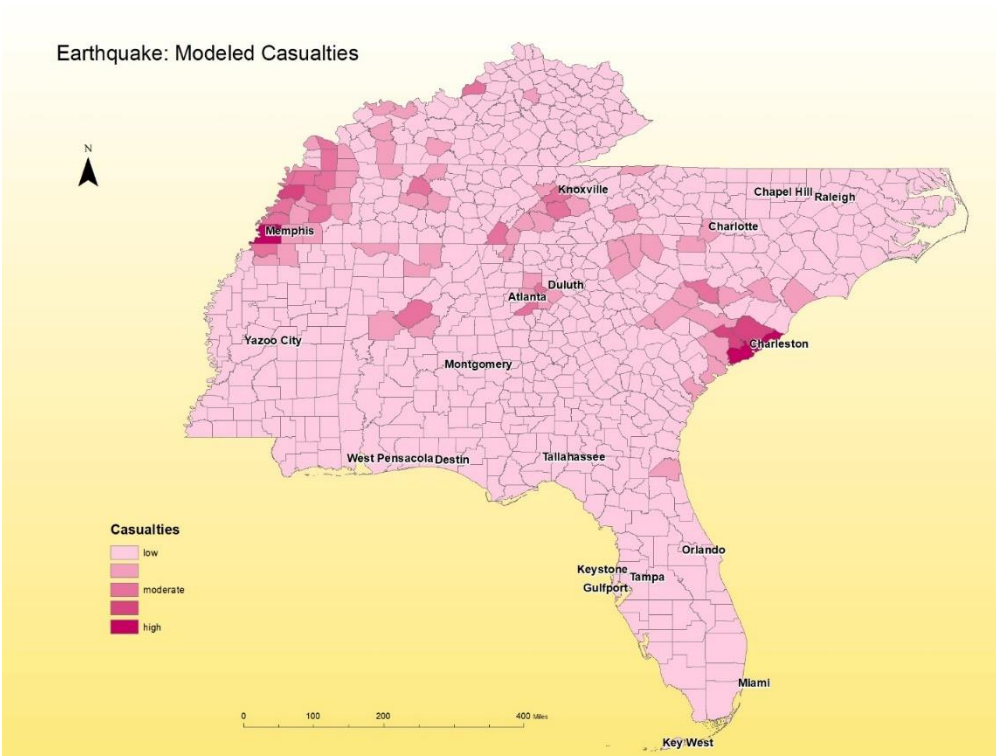


Figure 13: Potential Casualties from earthquakes within the study region

4.3 Social vulnerability in FEMA Region IV

Potential damages from earthquakes as demonstrated in section 4.1 are not random. Rather, there is an association between the study area's social characteristics and the potential for damages at particular places. The social analog to the quantitative physical risk analysis is the social vulnerability index. Social vulnerability helps to explain why some communities will experience damages and recovery from an earthquake differentially, even though they are subject to the same ground-shaking intensity (Burton and Silva 2016). For this research, seven composite factors were found following the SoVI methodology (Table 3) (see Appendix – 1 for factor loadings). These factors explain 76.1 percent of the variance among counties and are ordered in term of their variance explained).

Table 3: SoVI Factors with Percentage of Variance Explained

SoVI factors	Percent variance explained
(1) Wealth (-)	21.84
(2) Age dependence (elderly) (+)	14.72
(3) Race (African American) & poverty (+)	12.91
(4) Race (Hispanics) (+)	8.74
(5) Mobile homes (+)	7.97
(6) Unemployment (+)	5.61
(7) Service industry employment and native Americans (+)	4.31

The first highest factor in term of variance explained pertains to the wealth of the counties. The wealth factor is explained by per capita income, percentage of households earning greater than $200,000 annually, percent poverty, median gross rents, and median house values. The wealth factor explains 21.84 percent of the total variance within the data. Essentially, wealth is an important component of social vulnerability because wealth allows those impacted the financial means to recover, whereas those in poverty will struggle to recover. Wealth is also an important factor due to increased building and critical infrastructure concentrations. The greater Atlanta Metropolitan area provides an example of a large and sprawling city, covering two counties, where the generation of vast amounts of wealth has resulted in the ever-increasing development of the constructed environment. Although wealthy counties may be considered vulnerable due to increased assets in harm's way, there is strong evidence differential access to wealth and/or poverty is a major contributing factor to not only the social vulnerability of populations (Cutter, Boruff, and Shirley 2003), but even hazard losses (Burton 2010).

Figure 14 (A) represents the spatial distribution of the mapped component score for the wealth factor. The distribution of wealth demonstrates a coastal and urban bias where the highest component scores are found along the southern coast and in western portions of Florida. In Georgia, the wealth factor is concentrated in Fulton and Forsyth Counties. When this spatial distribution of the wealth factor is compared to the estimated damage patterns, an obvious association between wealth based on more assets in harm's way can be identified.

Age dependency is described by the second factor. Age dependency was identified by the percent population under five years or over sixty-five years, a high percentage of social security benefits recipients, and a high median age of populations. The age dependency factor explains 14.72 percent of variance. Here, the elderly may have mobility constraints which makes them less resilient (Cutter, Boruff and Shirley 2003; Cutter, Mitchell and Scott 2000). Moreover, the elderly living on limited retirement income or social security may not have the ability or opportunity to perform costly mitigation prior to an event. Figure14 (B) maps the spatial distribution of the age dependency factor. High loadings on the age factor are found in Florida coastal counties including Sumter, Charlotte, Citrus, Sarasota, Highlands, and Collier. The latter may be a direct result of migration of elderly from the colder and costlier portions of the United States.

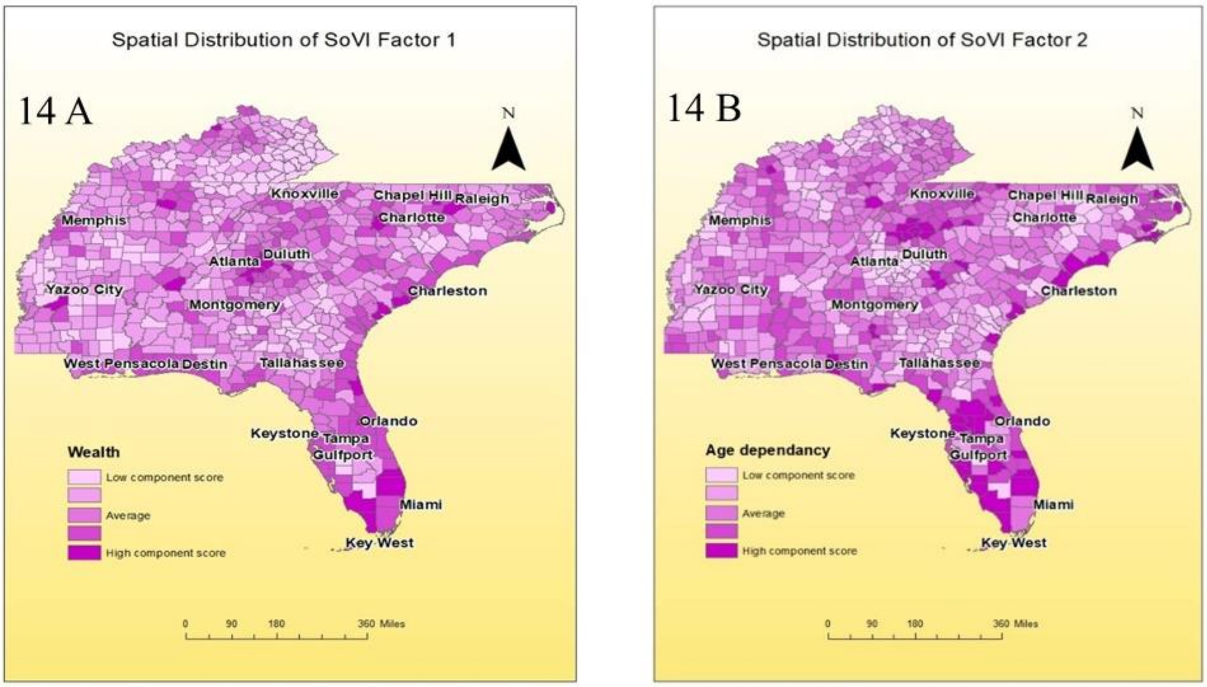
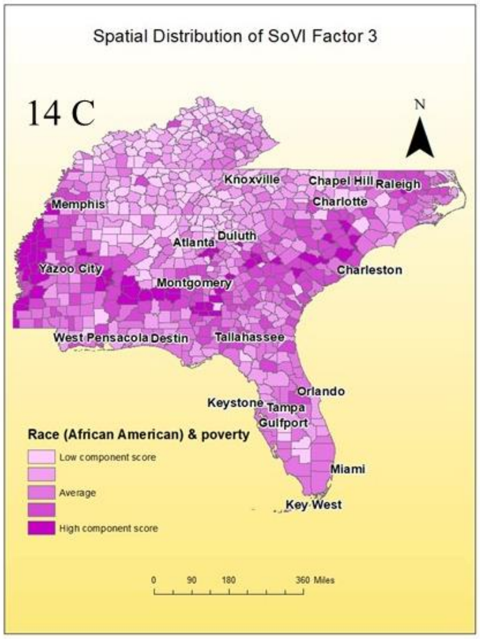
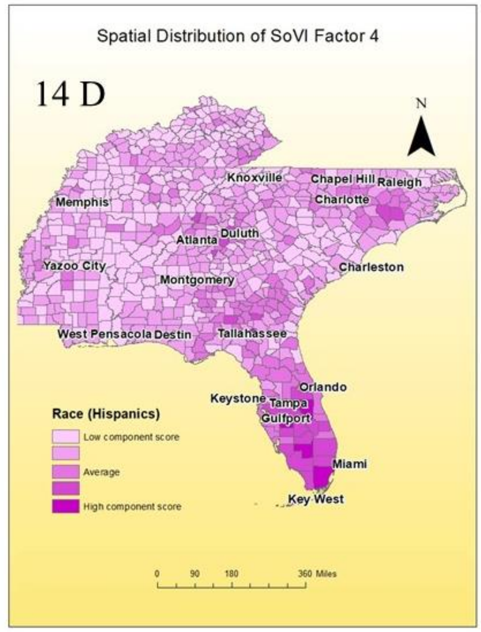


Figure 14: A, B, C, D, E, F & G; Spatial distribution of SoVI factors

The third factor, identified race (African American) and poverty, is an important contributing factor to the social vulnerability of the region. This factor explains 12.91 percent of variation among counties. Race contributes to social vulnerability through the lack of access to resources, cultural differences, and the social, economic, and political marginalization that is often associated with racial disparities (Cutter, Boruff and Shirley 2003, Burton 2010). In factor 3, poverty, female headed households, and percent African American populations load significantly high. Essentially, these variables may contribute to adverse earthquake impacts due to the nature of the housing stock where minorities and the poor often reside. Structural maintenance to the home, costly upgrades to conform to building standards, and the adoption of mitigation strategies are often out of reach and not possible for low income residents (Burton 2010). Figure 14 (C) represents spatial distribution of race (African American) and poverty factor. High factor scores were found in Mississippi (Tunica and Holmes County) and within the black belt of Alabama (Wilcox, Perry, and Macon Counties) where African Americans make up a large percentage of the population.

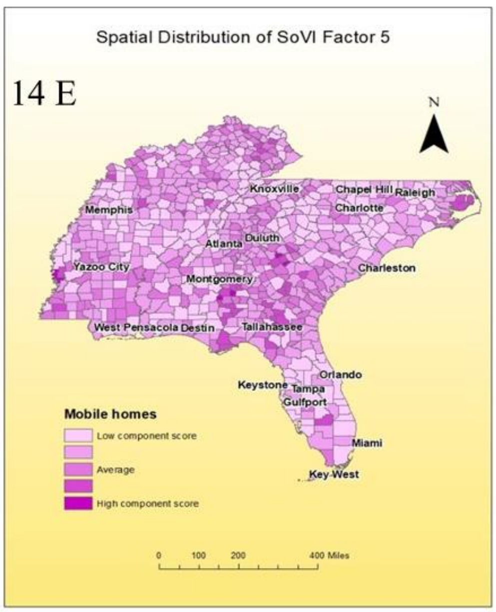
The fourth most significant factor in terms of variance explained is race (Hispanics). The latter explained 8.74 percent of variance within counties. This factor identifies Hispanics as a contributing factor to the social vulnerability of the study region. As previously mentioned, membership in a racial minority often results in marginalization. This influences disaster impact potential, resiliency, and hazard event outcomes (Morrow 1999; Burton 2010). The Race (Hispanics) factor was identified by the percent population that is Hispanic, percentage of non-English speaking populations and the percentage of population without health insurance. Figure 14 (D) represents spatial distribution of the Race (Hispanics) factor. Here, coastal counties of Florida such as Miami-Dade, Hendry, Hardee, Orange, Osceola, and Glades show high factor scores. In addition, Duplin and Sampson Counties in North Carolina have high factor scores for Hispanics. Other parts of the study area show low to moderate component scores.

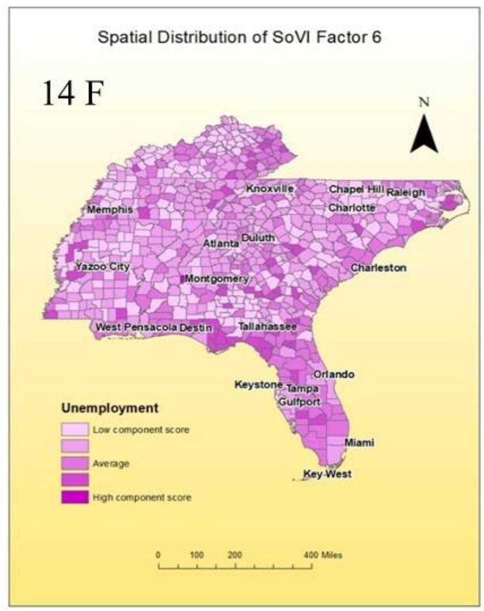




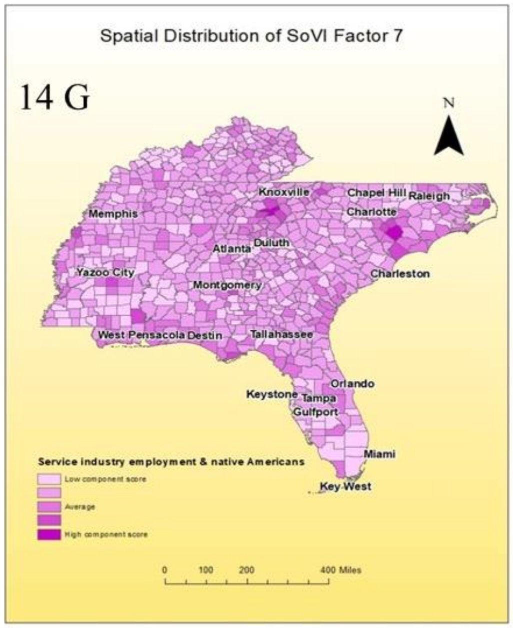
The fifth major factor in term of variance explained is mobile homes (Figure 14E). Mobile homes explained 7.97 percent of variance within counties. This factor was identified by the total percentage of population living in mobile homes. Mobile homes are vulnerable due to their construction type and these are often occupied by populations earning lower incomes. Essentially, this is a contributing factor to vulnerability.

Factor 6 was identified as representing unemployment (Fig. 14 F) which explains 5.61 percentage of variation among counties. The percentage of unemployed civilian populations and female populations are significant variables in this factor identification. Unemployed populations may lack the monetary resources needed to recover, and the potential loss of employment following an event exacerbates the number of unemployed workers in a community contributing to a slower recovery (Cutter, Boruff and Shirley 2003). The spatial distribution of the unemployment factor appears ubiquitous, yet higher areas of unemployment can be found primarily in rural counties that do not contain large metropolitan centers.





Finally, factor 7 (service industry employment and Native Americans) explained 4.31 percent of variation (Fig. 14 G). Some occupations such as service sector occupations may be severely impacted by a hazard event. Here, disposable incomes may fade as the need for services decline following a damaging event. Native Americans may be socially and economically marginalized and require additional support both in pre- and post-disaster periods. The highest scores for Factor 7 occur in North Carolina and in Tennessee, just south of Knoxville.



4.4 Using Spatial Autocorrelation to Identify Interrelationships Between Damage, Casualities, and Social Vulnerability

Spatial autocorrelation defines how one object or phenomenon across space is similar to other objects (Cliff and Ord 1970). Here, the concept was used here to pinpoint counties with high loss potential and high social vulnerability, and also to delineate areas in which counties with high social vulnerability and high loss potential cluster together to form hot-spots. Moran's I is a metric of association which measures the degree to which spatial autocorrelation occurs within a given study area. Randomness or spatial continuity of observed patters can be identified by spatial autocorrelation (Waldhör 1996). Moran's I is a popular and widely used measure of spatial autocorrelation which measures the similarity of values in nearby places from a mean value and visualizes a spatially weighted distribution of the autocorrelation between places (Moran 1950, Jackson et al. 2010).

The section describes the patterns that are the result of the analyses of spatial autocorrelation (Figure 15- 20). The analysis was conducted using a Local Bivariate Moran's I to associate the two SoVI factors explaining the highest percent variance (i.e. wealth and age) with the estimated earthquake losses and casualties. Such an analysis is often described as a Local Indicator of Spatial Association (or LISA). Here, bivariate LISA maps were produced where results were classified into four groups based on the extent of autocorrelation between the physical risk and social variables. Here, positive spatial autocorrelation is used to identify clusters representing high physical risk and a high loading on one of the respective social factors (high-high) or a low loading on physical risk and a low loading on one of the respective social factors (low-low). Conversely, negative spatial autocorrelation identifies relationships where high risk exists coupled with low social vulnerability (high-low) and vice versa (low-high).

Figure 15 A describes the relationship between the physical risk estimates and the wealth factor (Factor 1). Nineteen counties contain a high amount of wealth and also have the highest damage potential from earthquakes in the study area. These are primarily urban counties along the South Carolina coast and in the greater Atlanta Metropolitan area. These are areas of management concern in terms of the enormity of losses that could be sustained. The counties, however, contain wealth that could fund mitigation initiatives, pre-impact planning, response, and recovery actions if a damaging earthquake were to occur. Figure 15 B describes the statistical significance of the relationship between direct damages and wealth. Here, relationships between the two variables are very highly significant at p=0.0001. Five hundred forty-nine counties demonstrate no statistically significant relationship.

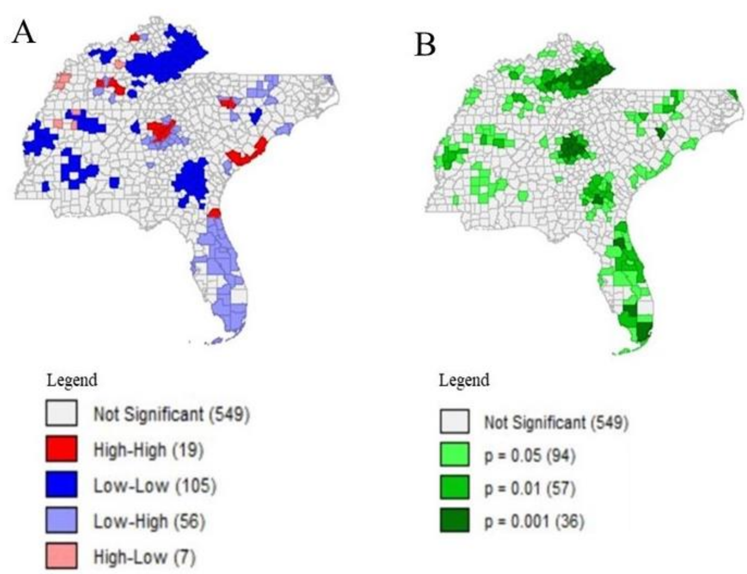


Figure 15: Direct damage vs wealth (Factor 1) (A. Bivariate LISA cluster map & B. Bivariate LISA significance map).

Figure 16 A describes the relationship between physical risk and age dependency (Factor 2) using maps. Plots of these spatial autocorrelations are located in Appendix 2. The results show that both high damage potential and a high age dependency occur in only six counties, mainly in eastern Tennessee and northeastern South Carolina. Positive spatial relationships between low age dependence and low damage potential also exist. These relationships occur in eighty-five counties and appear to be randomly distributed. Figure 16 B describes the significance of the relationship of direct damage and age dependency. The spatial associations between twenty-two counties are significant at p=0.0001. However, a statistical association between the age and damage does not exist in 559 counties.

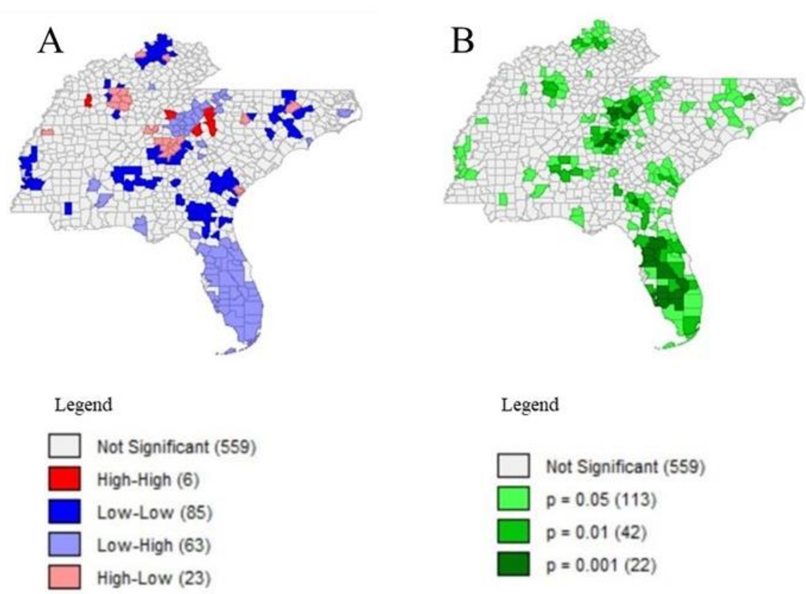


Figure 16: Direct damage vs Age dependency (Factor 2) (A. Bivariate LISA cluster map & B. Bivariate LISA significance map).

Figure 17 A describes the relationship between the modeled physical risk and SoVI scores which is the final composite metric of social vulnerability to define the social vulnerability of the study area. Areas of management concern include counties with high physical risk and high social vulnerability (6 counties) as well as counties low physical risk and high social vulnerability (90 counties). High-High counties occur mainly along the South Carolina coast. Low-high counties occur throughout the study area but mainly in Florida. Due to their high social vulnerability, these counties should still be seen as a management concern. This is primarily because these counties may be more inclined to suffer damages from a low magnitude event due to the potential condition and age of the infrastructure. In addition, costly mitigation may not be performed in these counties. In areas where the physical risk is negligible, such as in Florida, highly socially vulnerable populations may be at considerably high risk to other perils such as tropical cyclones and floods where risk is compounded by social vulnerabilities within the population (Burton and Silva 2016). Figure 17 B describes the significance of the relationship of physical risk and SoVI scores.

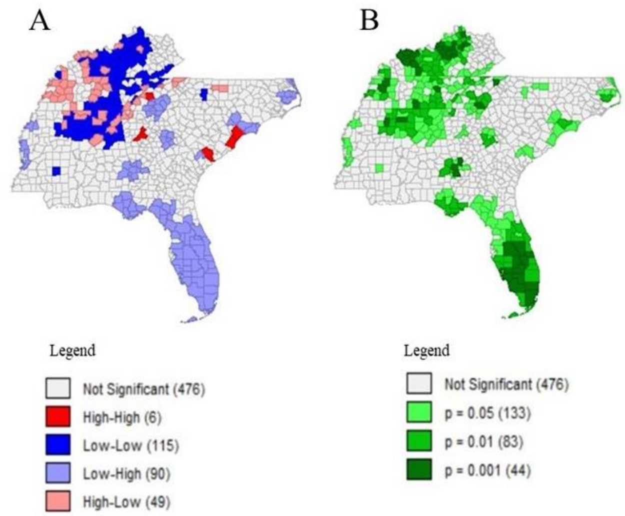


Figure 17: Direct damage vs SoVI score (A. Bivariate LISA cluster map & B. Bivariate LISA significance map).

Figures 18 A-B and 19 A-B describes the relationship between casualties and the wealth factor (Factor 1) and causalities and the age factor (Factor 2), respectively. Casualties and the wealth factor tend to cluster in urban areas along the South Carolina coast and in the Atlanta area. Conversely, high casualties and high age dependence (Figure 19 A) demonstrate a pattern that appears random. Counties with the potential for low casualties and high age dependence, however, are found throughout Florida. Figures 18 B and 19 B are representations of the statistical significance of the autocorrelation between the variables.

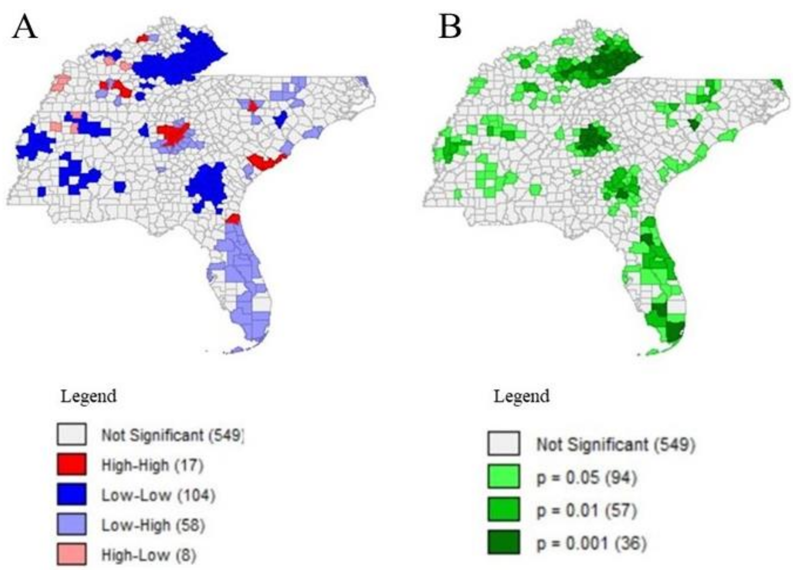


Figure 18: Casualties vs Wealth (Factor 1) (A. Bivariate LISA cluster map & B. Bivariate LISA significance map).

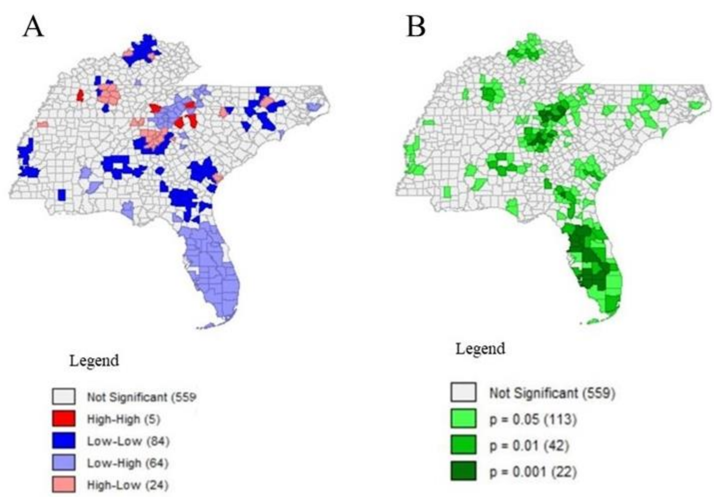


Figure 19: Casualties vs Age dependence (Factor 2) (A. Bivariate LISA cluster map & B. Bivariate LISA significance map).

Finally, figure 20 A-B describes the relationship between casualties and SoVI scores. Figure 20 A represents the areas of concern where both casualties and SoVI scores are high. The High-High counties are found mainly in South Carolina coast and Atlanta Metropolitan area. Figure 20 B represents the statistical significance of the relationship between these two variables.

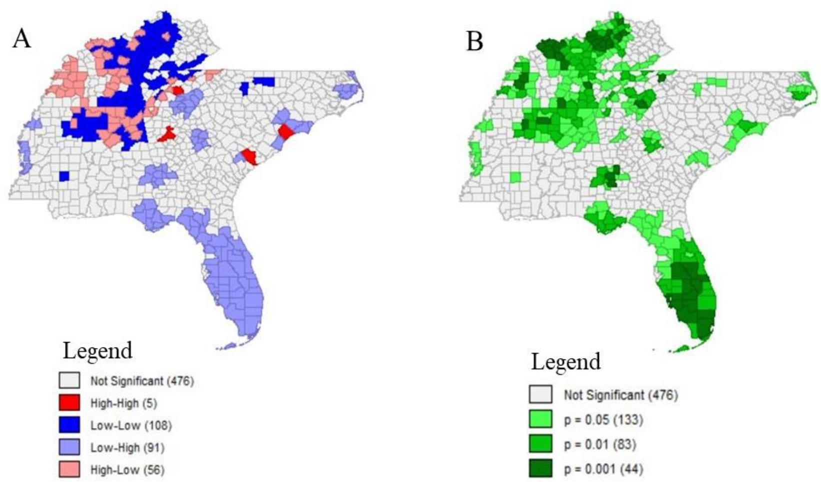


Figure 20: Casualties vs SoVI (A. Bivariate LISA cluster map & B. Bivariate LISA significance map)

4.5 Integrated Risk Assessment

An Integrated Risk Assessment is the integration of the social vulnerability model and the physical risk model. This was accomplished to determine areas of risk and social vulnerability simultaneously where high losses are likely to occur coupled with populations that may not be able to prepare for, respond to, and recover from such losses. The bivariate LISA analysis in section 4.4 was dependent on spatial autocorrelation between the physical risk and social vulnerability. The integrated assessment accompanies the previous method because it provides a straight forward metric to simultaneously measure areas of high risk and social vulnerability regardless of the relationships that are occurring in surrounding polygons. It is a comparative metric rather than a hot-spot analysis.

According to the probabilistic scenario (see Table 4), the Integrated Risk Assessment shows there are eight counties identified to be in highest integrated risk category (+2) and no county is in lowest category (-2). Nearly half (42.93%) of the counties fall into the medium risk category. One hundred and eighty-one counties (24.59%) fall into the high-risk category (+1 score) and two hundred and thirty-one counties (31.39%) fall into low risk category (-1).

Table 4: Number of counties in each risk level for probabilistic scenario.

Integrated Risk Score	Probabilistic scenario	Percent (%) of total counties
Highest (+2)	8	1.09
High (+1)	181	24.59
Medium (0)	316	42.93
Low (-1)	231	31.39
Lowest (-2)	0	0

Figure 21 is a map of the social vulnerability of the region using the SoVI. Counties mapped with a red hue represent the most socially vulnerable based on the index construction method and the three-class mapping scheme demonstrated here. Highly socially vulnerable counties are found throughout Florida, in a multitude of the counties along the southeast coast., and along the gulf coast of Mississippi. Other socially vulnerable areas include a swath of counties in eastern Tennessee and counties in proximity to the Atlanta Metropolitan areas.

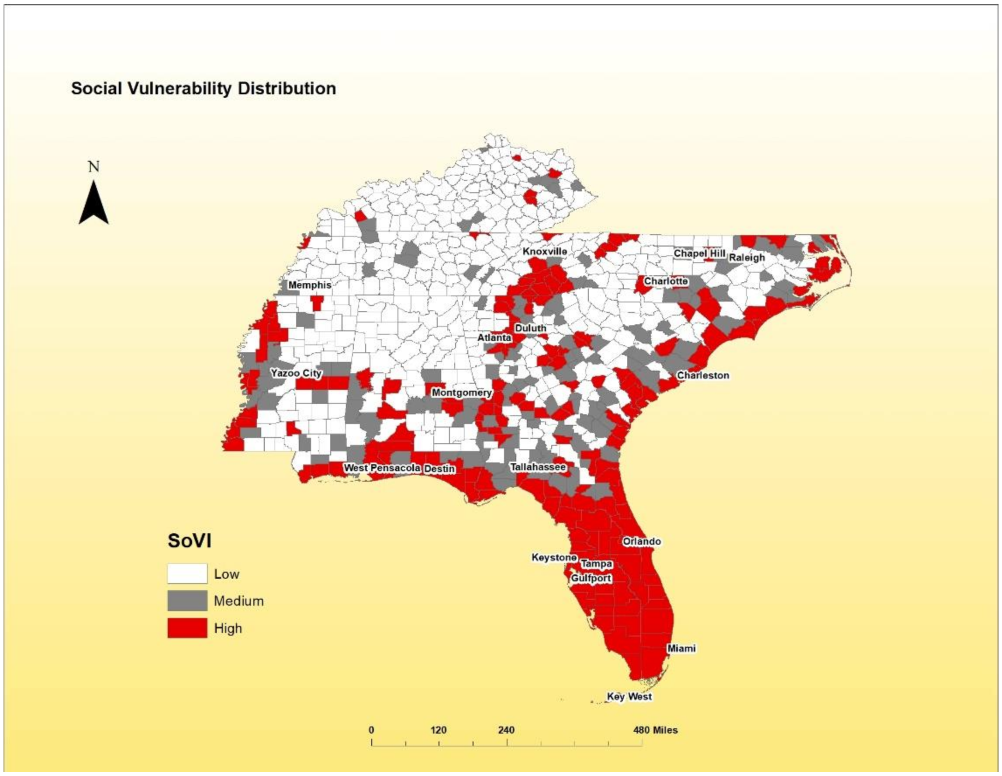


Figure 21: Social vulnerability distribution in the study region.

Figure 22 maps the modeled physical risk within study area. Not surprisingly, the lowest levels of physical risk occur in central and southern Florida which is an area furthest removed from the New Madrid Fault and its ground shaking potential. The highest levels of physical risk are found in the greater Memphis area of Tennessee, north to the Kentucky border. Charleston (Charleston County) and Columbia, South Carolina (Richland and Lexington Counties) also have a considerable amount of seismic risk when compared to the rest of the region. The same applies to Fulton and Dekalb Counties in Georgia where the sprawling, abundant, and high-density infrastructure of Atlanta could be subjected to considerable loss.

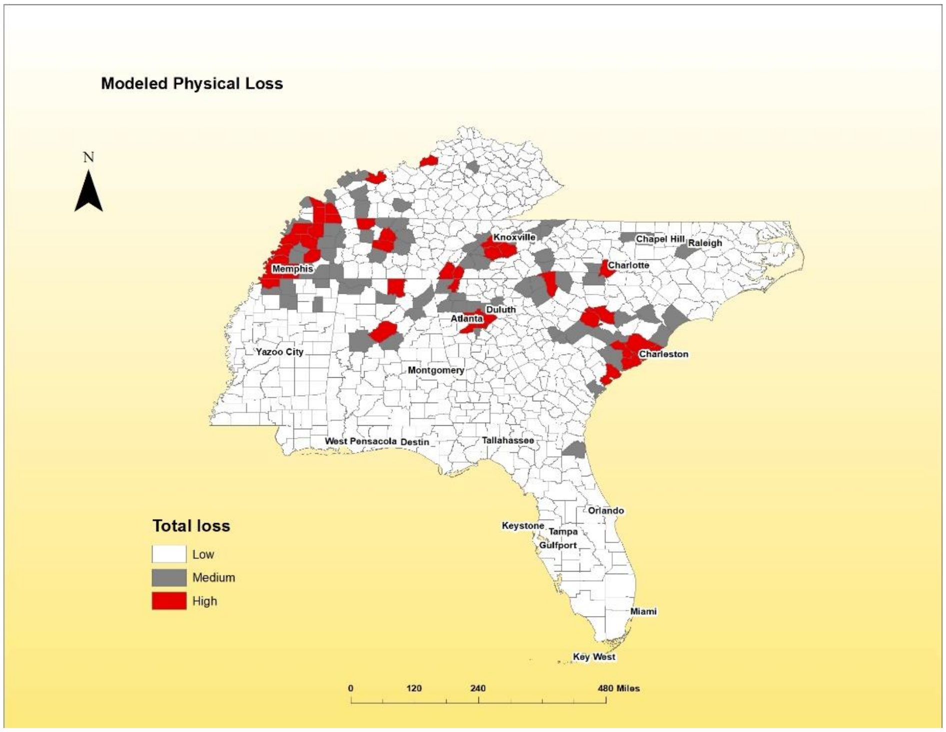


Figure 22: Total loss distribution (in probabilistic scenario) for FEMA Region IV

Figure 23 is a representation of the integrated risk of the region. Here, South Carolina (Beaufort, Charleston, Georgetown, and Horry Counties) are subjected to the highest physical risk coupled with highly socially vulnerable populations, comparably. Counties that encompass Knoxville and the Atlanta Metropolitan area also have a high degree of integrated risk. It's important to note that there is no lowest category of integrated risk because there was no county in which the calculated social vulnerability and physical scores were the lowest value (-1). As a result, the lowest value of -2 for any county could not be achieved.

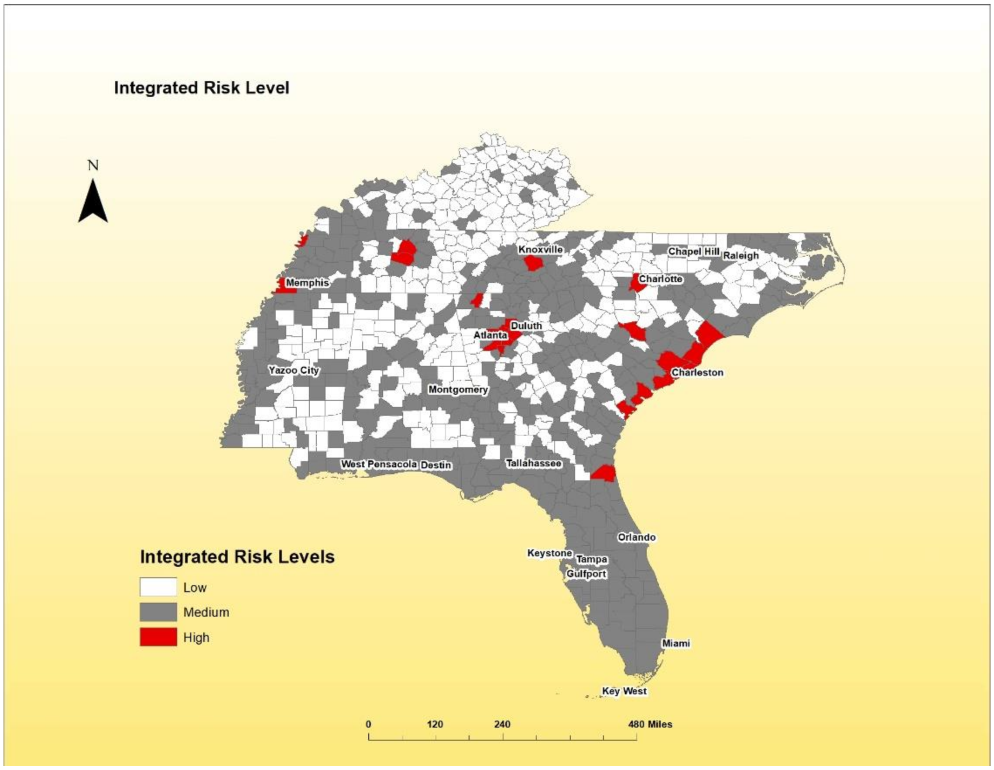


Figure 23: Integrated risk levels for FEMA region IV- Probabilistic earthquake scenario.

4.6 Global Regression Model

Up until this point, this thesis has concentrated mostly on the spatial description of the concepts of physical risk, social vulnerability, and integrated risk. As a supporting and ancillary analysis, OLS regression models were calibrated in order to test the assumption that quantifiable characteristics of socially vulnerability of populations may be used as proxy variables for modeled earthquake loss. This was accomplished first by regressing the SoVI factors against the predicted average annual losses, and second, by regressing the SoVI factors against the predicted casualties. Since the SoVI factors are a linear and parametric combination of the original variables, it was found that regression's basic assumptions were not violated.

The ANOVA results (Table 5) for the SoVI and direct damage model demonstrates a statistically significant model outcome overall. Table 6 presents the results where Factors 1, 2, and 6 where found to have a statistically significant association with the modeled losses. According to the regression beta coefficients which are used to assess the strength of the relationship between the dependent and independent variables, the biggest potential predictor of damage is the wealth factor. The predictive strength of the model based on its coefficient of determination (adjusted R-Square) is extremely low, however. The Adjusted R-Square results demonstrate that approximately 94% of the variance in the model is explained by characteristics not measured here. The assumption that characteristics of social vulnerability may be used as a proxy for earthquake impacts (Direct damage and casualties) was supported by the model. The explanatory power of the model is so low, however, that the results are negligible.

Table 5: ANOVA Table for Global Regression of SoVI factors and Direct Damage

	Sum of squares	df	Mean square	F	Sig.
Regression	48.020	7	6.860	7.259	.000
Residual	687.980	728	0.945		
Total	736.000	735			

Table 6: Global Regression of SoVI factors for Direct Damage using Standardized Regression Coefficients

Variable	**Beta**	**t**	**Sig**
Factor 1_Wealth	**0.194**	**5.420**	**0.000**
Factor 2_ Age dependency	**-0.104**	**-2.907**	**0.004**
Factor 3_Race Black & Poverty	**0.069**	**1.923**	**0.055**
Factor 4_Race Hispanics	**-0.054**	**-1.513**	**0.131**
Factor 5_Mobile homes	**-0.062**	**-1.735**	**0.083**
Factor 6_Unemployment	**-0.071**	**-1.978**	**0.048**
Factor 7_Service Industry Employment and Native Americans	**-0.010**	**-0.281**	**0.779**
R-square: 0.065, Adjusted R-Square 0.058			

For the association between casualties and the SoVI factors, a statistically significant model was also achieved (see Table 7). In addition, several variables (Factor 1, Factor 2, and Factor 5) achieved a statistically significant association with casualties (Table 8). Factor 1 (wealth) likewise was found to be the biggest predictor of casualties based on the regression's beta coefficients. Here, the results are also negligible, however. The predictive strength of the model is extremely low (Adjusted R-Squared = 0.057), leaving the majority of the variance unexplained.

Table 7: ANOVA of Global Regression Parameters with SoVI Factors and Casualties

	Sum of squares	**df**	**Mean square**	**F**	**Sig.**
Regression	**48.375**	**7**	**6.911**	**7.316**	**0.000**
Residual	**687.625**	**728**	**0.945**		
Total	**736.000**	**735**			

Table 8: Global Regression Parameters with SoVI Factors for Casualty using Standardized Regression Coefficients

Variable	**Beta**	**T**	**Sig**
Factor 1_Wealth	**0.184**	**5.138**	**0.000**
Factor 2_ Age dependency	**-0.119**	**-3.308**	**0.001**

Factor 3_Race Black & Poverty	**0.064**	**1.777**	**0.076**
Factor 4_Race Hispanics	**-0.066**	**-1.835**	**0.067**
Factor 5_Mobile homes	**-0.075**	**-2.081**	**0.038**
Factor 6_Unemployment	**-0.061**	**-1.716**	**0.087**
Factor 7_Service Industry Employment and Native Americans	**-0.009**	**-0.253**	**0.800**
R-square: 0.066, Adjusted R-Square 0.057			

4.7 Summary

For the social vulnerability model, seven composite factors were identified as contributors to the spatial variation of social vulnerability within the counties of the study region. These seven factors were derived by Factor Analysis using the SoVI method. The social vulnerability estimates were coupled with probabilistic damage estimates from HAZUS-MH to estimate integrated risk. The integrated risk assessment identified areas of management concern where counties both demonstrated high social vulnerability and high physical risk scores. These include counties along the South Carolina Coast, counties in eastern Tennessee and counties in the greater Atlanta Metropolitan area.

Chapter 5: Conclusions and Directions for Further Research

5.1 Conclusions

Property damages and casualties from earthquakes are an issue of concern for states, counties, cities, and individuals within the United States. It is within this context that this thesis contributes to current research by identifying locations of physical risk from earthquakes coupled with the social vulnerability of populations that will be subjected to the hazard. The objective of this research was not only to identify the social factors that contribute to earthquake losses, but also to demonstrate procedures on how assess earthquake risk spatially and holistically. The procedures and results outlined in this thesis show promise for hazard mitigation, communication, planning, disaster risk reduction initiatives, and research. Mapping earthquake risk and the contributing factors to it can improve resiliency and loss reduction within the region. It is hoped that the successful answers to the research questions outlined in this thesis can contribute to future research within the area of earthquake risk assessment.

5.2 Research Question 1 Summary

The first research question asked which counties of the study region have high physical risk. Here, HAZUS-MH software was utilized to calculate direct damages and casualties from earthquakes. It was found that the spatial distribution of losses was not random but based on a number of factors such as building type and the location of buildings in large urban centers. Very high losses were estimated for Charleston (SC), Berkeley (SC), and Shelby (TN), to name just a few counties. South Carolina and Tennessee have had previous earthquake records. Here, earthquake impact modeling was accomplished considering previous earthquakes and these areas were found to be higher in vulnerable physically. The results of the earthquake impact modeling were found to be aligned with the assumptions of where the majority of high dollar losses might occur (i.e. in large and wealthy cities such as Charleston and Atlanta).

5.3 Research Question 2 Summary

The second research question asked which counties entertain a high physical risk coupled with populations that have a high social vulnerability. The purpose of this question was to delineate the spatial distribution of clusters of counties where both physical risk and social vulnerability are high. Hotspots were found in some parts of South Carolina, Tennessee and Georgia based on the comparative integrated risk levels of each county. While Georgia counties in the Atlanta area have a low probability for the occurrence of damaging earthquakes, this area was found to be vulnerable due to the large number of assets in harm's way coupled with social factors.

5.4 Research Question 3 Summary

The final research question asked which indicators of social vulnerability provide the best predictors of the spatial differences in impacts from earthquakes. This question was asked in order to see if a proxy variable for estimated losses could be derived from socio-economic data, not to make inferences regarding the extent to which social vulnerability increases damages. Two linear regression analyses were conducted to test if social vulnerability has a place in physical risk modelling directly. In both models, wealth was identified as the strongest predictor of damages, mainly because wealthier counties contain more infrastructure that is exposed to potential damage from ground shaking. However, the adjusted R-Squared values were very low. It's arguable that that the results of the regression analysis are negligible, and future studies intending to use social vulnerability characteristics as a proxy for estimated loss should be approached with extreme caution.

5.5 Research Opportunities

This section refers to the problems that came up while working on this thesis. A number of areas of opportunities roused which constitute the need for future research. The first is the study region itself. If the study area were California, for instance, the results would likely be different. The second area of research opportunity results from issues of scale. This research was conducted at the county level which is a very large areal unit. Thus, results at one scale of geography could be quite different from the results derived at other spatial resolutions. The third area of opportunity is the variable selection. This thesis adopted the SoVI method because it is well established and replicable, but there are more variables out there that could have been explored, even qualitative ones. Perhaps with different variables, the regression results would have been improved. Finally, the regression models were failed to explain more than 6% of the variation in the models. Having real earthquake losses at the county level would have been helpful. While there are a number of caveats, it is important to address these areas of opportunity because they provide a wide-open opportunity for future research.

5.6 Research Contributions

In the fields of natural hazards and disasters and disaster risk reduction, researchers and practitioners alike recognize that risk assessments, regardless of the hazard, all-to-often overlook characteristics within social systems that could aggravate earthquake loss. The successful completion of this thesis resulted in several outcomes which are notable. First, this research has provided a mechanism in which social indicators are used in earthquake impact modeling for a very large study region where the earthquake hazard is often overlooked. Secondly, understanding how social indicators contribute to earthquake impact allows researchers to identify why differential impacts exist within built environment. Finally, the integration of social vulnerability with impact prediction provides a comparative metric and benchmarking tool for decision making and planning.

References

Anderson, E. 2008. "Central American Probabilistic Risk Assessment (CAPRA): objectives, applications and potential benefits of an open access architecture." Global Risk Forum, GRF Davos, Switzerland.

Araneda, Juan C, H Rudnick, Sebastian Mocarquer, and Pedro Miquel. 2010. "Lessons from the 2010 Chilean earthquake and its impact on electricity supply." Power System Technology (POWERCON), 2010 International Conference on.

Biryol, C Berk, Lara S Wagner, Karen M Fischer, and Robert B Hawman. 2016. "Relationship between observed upper mantle structures and recent tectonic activity across the Southeastern United States." *Journal of Geophysical Research: Solid Earth* 121 (5):3393-3414.

Blaikie, P., Cannon, T., and Wisner, B., 1994. *At Risk: Natural Hazards, People's Vulnerability, and Disasters*. New York: Routledge.

Bolin, Robert, and Lois Stanford. 1991. "Shelter, housing and recovery: a comparison of US disasters." *Disasters* 15 (1):24-34.

Boruff, Bryan J, Christopher Emrich, and Susan L Cutter. 2005. "Erosion hazard vulnerability of US coastal counties." *Journal of Coastal Research*:932-942.

Burton, Christopher G. 2010. "Social vulnerability and hurricane impact modeling." *Natural Hazards Review* 11 (2):58-68.

Burton, Christopher G, and Vitor Silva. 2016. "Assessing integrated earthquake risk in OpenQuake with an application to Mainland Portugal." *Earthquake Spectra* 32 (3):1383-1403.

Burton, Christopher G, and Tormene, Paolo. 2018. *Integrated Risk Modelling Toolkit – User Manual*, Release 2.9.2, (https://docs.openquake.org/oq-irmt-qgis/v2.9.2/), 55 p.

Carreño, Martha-Liliana, Omar D Cardona, and Alex H Barbat. 2007. "Urban seismic risk evaluation: a holistic approach." *Natural Hazards* 40 (1):137-172.

Chakraborty, Jayajit, Graham A Tobin, and Burrell E Montz. 2005. "Population evacuation: assessing spatial variability in geophysical risk and social vulnerability to natural hazards." *Natural Hazards Review* 6 (1):23-33.

Cliff, Andrew D, and Keith Ord. 1970. "Spatial autocorrelation: a review of existing and new measures with applications." *Economic Geography* 46 (sup1):269-292.

Crowley, H, and V Silva. 2013. "OpenQuake engine book: risk v1. 0.0." *GEM Foundation, Pavia, Italy*.

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. 1993. *Living with risk: the geography of technological hazards*: E. Arnold.

Cutter, Susan L. 1996. "Vulnerability to environmental hazards." *Progress in human geography* 20 (4):529-539.

Cutter, Susan L, Lindsey Barnes, Melissa Berry, Christopher Burton, Elijah Evans, Eric Tate, and Jennifer Webb. 2008. "A place-based model for understanding community resilience to natural disasters." *Global environmental change* 18 (4):598-606.

Cutter, Susan L, Christopher T Emrich, Jerry T Mitchell, Bryan J Boruff, Melanie Gall, Mathew C Schmidtlein, Christopher G Burton, and Ginni Melton. 2006. "The long road home: Race, class, and recovery from Hurricane Katrina." *Environment: Science and Policy for Sustainable Development* 48 (2):8-20.

Cutter, Susan L, Christopher T Emrich, Jennifer J Webb, and Daniel Morath. 2009. "Social vulnerability to climate variability hazards: A review of the literature." *Final Report to Oxfam America* 5:1-44.

Cutter, Susan L, Jerry T Mitchell, and Michael 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.

Davidson, Rachel. 1997. "EERI annual student paper award a multidisciplinary urban earthquake disaster risk index." *Earthquake Spectra* 13 (2):211-223.

Davidson, Rachel A, Huan Zhao, and Vineet Kumar. 2003. "Quantitative model to forecast changes in hurricane vulnerability of regional building inventory." *Journal of Infrastructure Systems* 9 (2):55-64.

Erdik, Mustafa, K Şeşetyan, MB Demircioğlu, C Zülfikar, U Hancılar, C Tüzün, and E Harmandar. 2014. "Rapid earthquake loss assessment after damaging earthquakes." In *Perspectives on European earthquake engineering and seismology,* 53-95. Springer.

Erdik, Mustafa, K. Şeşetyan, M. B. Demircioğlu, U. Hancılar, and C. Zülfikar. 2011. "Rapid earthquake loss assessment after damaging earthquakes." *Soil Dynamics and Earthquake Engineering* 2: 247-266.

Hazus-MH, FEMA. 2003. "2.1 Technical Manual: Multi-hazard Loss Estimation Methodology." *Washington, DC: Federal Emergency Management Agency*.

Hazus Earthquake Model." What Is Disaster Assistance? | FEMA.gov. Accessed May 09, 2018. https://www.fema.gov/hazus-mh-earthquake-model

Hazus-MH, FEMA. 2004. "Using HAZUS-MH for Risk Assessment." *Washington, DC: Federal Emergency Management Agency*.

Hewitt, Kenneth. 1997. *Regions of risk: A geographical introduction to disasters*: Essex, U.K.

Jackson, J Edward. 2005. *A user's guide to principal components*. Vol. 587: John Wiley & Sons.

Jackson, Monica C, Lan Huang, Qian Xie, and Ram C Tiwari. 2010. "A modified version of Moran's I." *International journal of health geographics* 9 (1):33.

Jaiswal, Kishor, Doug Bausch, Jesse Rozelle, John Holub, and Sean McGowan. 2017. Hazus® estimated annualized earthquake losses for the United States. Federal Emergency Management Agency.

Khazai, Bijan, Fouad Bendimerad, Omar D Cardona, Martha-Liliana Carreño, Alex H Barbat, and CG Buton. 2015. "A guide to measuring urban risk resilience: Principles, tools and practice of urban indicators." *Earthquakes and Megacities Initiative (EMI), The Philippines*.

Mileti, Dennis. 1999. *Disasters by design: A reassessment of natural hazards in the United States*: Joseph Henry Press.

Molas, GL, H Aslani, N Shome, C Cabrera, S Taghavi, and J Bryngelson. 2012. "Impacts of Earthquake Hazard Uncertainties On Probabilistic Portfolio Loss Risk Assessment." 15th World Conference on Earthquake Engineering (WCEE), Lisbon, Portugal.

Moran, Patrick AP. 1950. "Notes on continuous stochastic phenomena." *Biometrika* 37 (1/2):17-23.

Morrow, Betty Hearn. 1999. "Identifying and mapping community vulnerability." *Disasters* 23 (1):1-18.

North Carolina Department of Crime Control and Public Safety, Division of Emergency Management. 2016. *Earthquake History, Response and Emergency Kit.*

Pulido, Laura. 2000. "Rethinking environmental racism: White privilege and urban development in Southern California." *Annals of the Association of American Geographers* 90 (1):12-40.

Rao, Anirudh, Graeme Weatherill, Vitor Silva and John Schneider. 2017. "Beyond Button Pushing: Seismic Risk Assessment for California." *CSSC Publication No. 17-05*

Robinson, Andrew, and Pradeep Talwani. 1983. "Building damage at Charleston, South Carolina, associated with the 1886 earthquake." *Bulletin of the Seismological Society of America* 73 (2):633-652.

Ryu, H, and M Edwards. 2015. "A refinement to earthquake vulnerability models implemented in the EQRM: correcting underestimation of the damage probability for non-structural acceleration-sensitive components." *Proceedings of the Tenth Pacific Conference on Earthquake Engineering Building an Earthquake-Resilient Pacific 6- 8 November 2015, Sydney, Australia.*

Schmidtlein, Mathew C, John M Shafer, Melissa Berry, and Susan L Cutter. 2011. "Modeled earthquake losses and social vulnerability in Charleston, South Carolina." *Applied Geography* 31 (1):269-281.

Silva, Vitor, Helen Crowley, Marco Pagani, Damiano Monelli, and Rui Pinho. 2014b. "Development of the OpenQuake engine, the Global Earthquake Model's open-source software for seismic risk assessment." *Natural Hazards* 72 (3):1409-1427.

SoVI® | Hazards & Vulnerability Research Institute | University of South Carolina. "SoVI®."Accessed October 04, 2017. http://artsandsciences.sc.edu/geog/hvri/sovi%C2%AE-0

Sun, Caiyu, and Jiuping Xu. 2010. "Estimation of time for Wenchuan Earthquake reconstruction in China." *Journal of Construction Engineering and Management* 137 (3):179-187.

Turner, Billie L, Roger E Kasperson, Pamela A Matson, James J McCarthy, Robert W Corell, Lindsey Christensen, Noelle Eckley, Jeanne X Kasperson, Amy Luers, and Marybeth L Martello. 2003. "A framework for vulnerability analysis in sustainability science." *Proceedings of the national academy of sciences* 100 (14):8074-8079.

Tuttle, MP, J Collier, LW Wolf, and RH Lafferty. 1999. "New evidence for a large earthquake in the New Madrid seismic zone between AD 1400 and 1670." *Geology* 27 (9):771-774.

Yucel, G, and Görün Arun. 2012. "Earthquake and Physical and Social Vulnerability Assessment for Settlements: Case Study Avcilar District." World Conference on Earthquake Engineering.

United Nations Office for Disaster Risk Reduction (UNISDR) (2017). Terminology on Disaster Risk Reduction. https://www.unisdr.org/we/inform/terminology#letter-d

USGS Earthquake Hazards Program. (2017b). Earthquake Statistics. Retrieved March 25, 2017, from https://earthquake.usgs.gov/earthquakes/browse/stats.php

Appendix-1

SoVI variable names and corresponding codes

Variable Names	**Codes**	**Variable Name**	**Codes**
Median Age	MEDAGE	Percent female population	QFEMALE
Percent Black	QBLACK	Percent female headed households	QFHH
Percent Native American	QNATAM	Percent households receiving social security benefits	QSSBEN
Percent Asian and Hawaiian Islanders	QASIAN	People per unit	PPUNIT
Percent Hispanic	QHISP	Per Capita Income (in dollars)	PERCAP
Percent of population under 5 years or 65 and over	QAGEDEP	Percent employment in extractive industries	QEXTRCT
Median Value of Owner Occupied Housing Units	MDHSEVAL	Percent of housing units that are mobile homes	QMOHO
Percent of children living in married couple families	QFAM	Percent of population with no high school diploma or less	QED12LES

		than 12th grade education	
Percent renter	QRENTERS	Percent of housing units with no car	QNOAUTO
Percent residents in nursing homes	QNRRES	Percent females participating in the labor force	QFEMLBR
Percent speaking English as a second language	QESL	Percent unoccupied housing units	QUNOCCHU
Percent of population without health insurance	QNOHLTH	Median Gross Rent	MDGRENT
Percent civilian unemployment	QCVLUN	Percent Employed in service industry	QSERV
Percent of households earning $200,000 or more	QRICH200K		
Percent living below poverty level	QPOVTY		

SPSS Table of factor analysis output showing %variance explained

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.200	29.287	29.287	8.200	29.287	29.287	6.115	21.840	21.840
2	4.448	15.886	45.173	4.448	15.886	45.173	4.122	14.722	36.562
3	2.751	9.824	54.997	2.751	9.824	54.997	3.615	12.912	49.474
4	2.267	8.098	63.095	2.267	8.098	63.095	2.447	8.738	58.212
5	1.348	4.815	67.909	1.348	4.815	67.909	2.232	7.972	66.183
6	1.236	4.413	72.323	1.236	4.413	72.323	1.570	5.606	71.789
7	1.057	3.775	76.097	1.057	3.775	76.097	1.206	4.309	76.097
8	0.936	3.343	79.440						
9	0.866	3.092	82.532						
10	0.675	2.409	84.942						
11	0.599	2.138	87.080						
12	0.586	2.094	89.174						
13	0.505	1.802	90.976						
14	0.406	1.449	92.425						
15	0.348	1.244	93.668						
16	0.307	1.098	94.766						
17	0.277	0.991	95.757						
18	0.231	0.826	96.583						
19	0.184	0.657	97.240						
20	0.157	0.560	97.799						
21	0.131	0.467	98.267						
22	0.118	0.420	98.686						
23	0.094	0.337	99.023						
24	0.078	0.278	99.301						
25	0.069	0.248	99.549						
26	0.059	0.210	99.759						
27	0.038	0.136	99.895						
28	0.029	0.105	100.000						

Extraction Method: Principal Component Analysis.

SPSS PCA output of rotated component matrix with variables >=0.500 and <= -0.500 highlighted in yellow

Rotated Component Matrix[a]

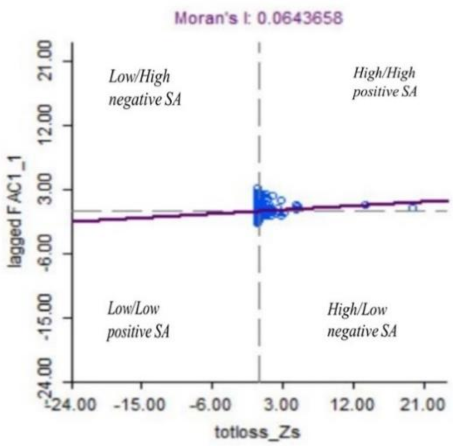
	Component						
	1	2	3	4	5	6	7
Zscore(QNOHLTH)	-0.211	0.090	0.280	0.672	0.264	0.036	0.225
Zscore(QFAM)	0.128	-0.001	-0.905	-0.054	-0.005	0.066	-0.077
Zscore(MDGRENT)	0.844	-0.070	-0.037	0.266	-0.184	0.045	0.069
Zscore(QASIAN)	0.699	-0.301	0.123	0.235	-0.085	-0.065	-0.014
Zscore(QBLACK)	-0.050	-0.106	0.897	-0.019	0.122	-0.008	-0.076
Zscore(QHISP)	0.252	-0.069	-0.062	0.910	-0.075	0.042	-0.007
Zscore(QNATAM)	-0.119	-0.059	-0.047	0.064	-0.011	-0.054	0.755
Zscore(QAGEDEP)	-0.142	0.877	-0.090	0.066	-0.093	-0.196	-0.061
Zscore(MEDAGE)	-0.069	0.859	-0.262	-0.114	0.013	0.011	-0.163
Zscore(QFEMALE)	0.071	0.120	0.159	-0.092	-0.172	-0.900	-0.003
Zscore(QFHH)	-0.237	-0.224	0.884	0.028	0.026	-0.038	0.005
Zscore(QNRRES)	-0.396	0.200	0.062	-0.155	0.043	-0.128	-0.154
Zscore(QMOHO)	-0.283	0.041	0.053	0.030	0.874	0.081	-0.037
Zscore(PPUNIT)	-0.043	-0.741	-0.020	0.032	-0.143	0.471	-0.201
Zscore(QESL)	0.406	-0.084	-0.032	0.859	-0.112	0.013	0.009
Zscore(QPOVTY)	-0.640	0.006	0.569	0.023	0.217	0.053	0.059
Zscore(QRENTERS)	0.126	-0.567	0.477	0.172	-0.274	-0.178	0.096
Zscore(QSSBEN)	-0.394	0.813	0.038	-0.027	-0.010	0.164	-0.056
Zscore(QRICH200K)	0.808	-0.082	-0.079	0.075	-0.055	-0.066	-0.155
Zscore(PERCAP)	0.878	0.014	-0.263	-0.004	-0.192	-0.159	-0.088
Zscore(QED12LES)	-0.730	0.202	0.140	0.193	0.193	0.136	-0.023
Zscore(QUNOCCHU)	0.080	0.710	0.043	0.083	0.337	0.156	0.307
Zscore(QNOAUTO)	-0.475	0.266	0.434	-0.090	0.154	0.169	-0.174

Zscore(QCVLUN)	-0.468	0.398	0.278	-0.058	0.126	0.516	0.055
Zscore(QEXTRCT)	-0.479	0.042	0.017	0.328	0.435	0.066	-0.169
Zscore(QSERV)	0.439	0.148	0.138	0.031	-0.240	0.119	0.516
Zscore(QFEMLBR)	-0.293	0.179	0.150	-0.108	0.805	0.106	-0.044
Zscore(MDHSEVAL)	0.889	0.096	-0.201	0.178	-0.129	0.005	0.030
	Wealth (-)	Age dependence (elderly)(+)	Race (African American) & poverty (+)	Race (Hispanics) (+)	Mobile homes (+)	Unemployment (+)	Service industry employment and Native Americans (+)

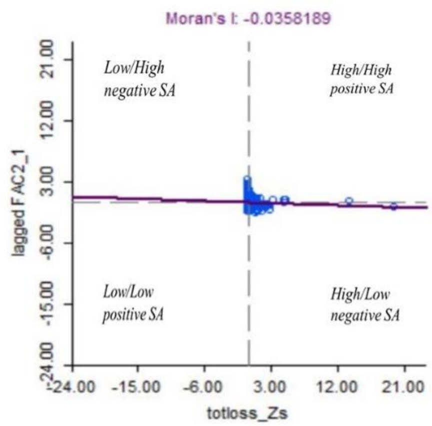
Appendix-2

Moran's I scatter plot

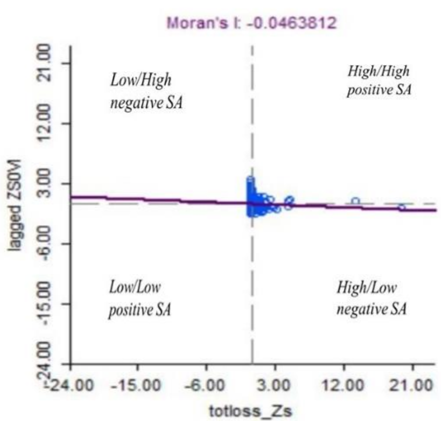
Direct damage vs wealth



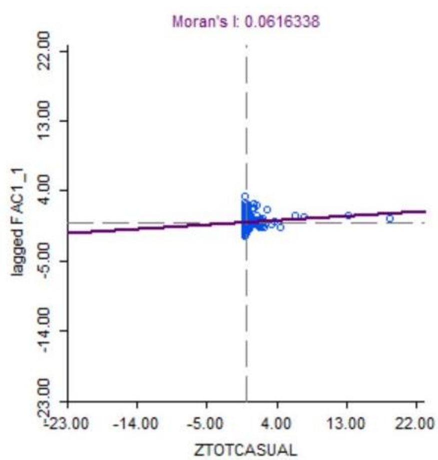
Direct damage vs age dependency



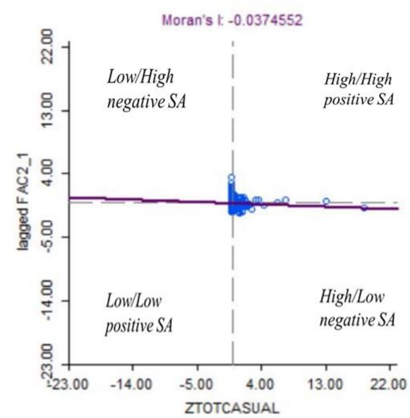
Direct damage vs SoVI score



Casualty vs wealth



Casualty vs age dependency



Casualty vs SoVI