EXAMINING TORNADO VULNERABILITY AND
INDEX VALIDATION IN OKLAHOMA
USING COMPONENTS OF THE SOCIAL VULNERABILITY INDEX

By

LINDSAY KING

Bachelor of Arts in Geography

Southern Connecticut State University

New Haven, Connecticut

2017

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2019

EXAMINING TORNADO VULNERABILITY AND

INDEX VALIDATION IN OKLAHOMA

USING COMPONENTS OF THE SOCIAL VULNERABILITY INDEX

Thesis Approved:

Dr. Jonathan C. Comer

Committee Chair

Dr. Peter J. Kedron

Thesis Adviser

Dr. Stephen J. Stadler

ACKNOWLEDGEMENTS

Thank you to my research adviser, Dr. Peter Kedron, for your patience, guidance, support, and constructive feedback. I appreciate all your efforts and know that I am a better student and researcher because of it.

Thank you to my committee members, Dr. Jonathan Comer and Dr. Stephen Stadler, for your constructive feedback, advice, and open-door policy.

Additional thanks to Dr. Peter Kedron and the staff and student workers at OSU Cartography Services for their assistance with the creation of some datasets.

To my friends, Tessa Cook, Christina Woehrle, Maria Sol Ramirez Dhillon, and Nicole Pauley, thank you for your endless support, encouragement, love, and advice.

Graduate school would not have been possible without the love and support of my family. Thank you to my parents and sisters for your endless patience and continued support. My education would not have been possible without my godfather, Jim Balitsos, to whom I am forever grateful.

Name: LINDSAY KING

Date of Degree: DECEMBER, 2019

Title of Study: EXAMINING TORNADO VULNERABILITY AND INDEX VALIDATION IN OKLAHOMA USING COMPONENTS OF THE SOCIAL VULNERABILITY INDEX

Major Field: GEOGRAPHY

Abstract: Identifying socially vulnerable groups is an important step toward creating resilient communities and reducing future losses of property and human life. A population's vulnerability to a hazard is not based solely on its proximity to a dangerous event. Instead, vulnerability to a hazard is the product of a complex combination of the socioeconomic, institutional, and environmental systems that affect a group of people, and the disruption of those systems by a hazardous event. Measurement of social vulnerability is already a focus within the hazard's literature. One area of particularly intensive research attention has been the development and application of indices of social vulnerability, which are constructed from a range of measures meant to serve as proxies of aspects of vulnerability. There is an ongoing need to create reliable, useful, and accurate indexes that can inform policymakers and natural hazards scientists for better decision making at various stages of the disaster cycle. However, less attention has been devoted to the validation of these indexes, which is critical to their practical use. The purpose of this work is to validate two alternative social vulnerability indexes within the state of Oklahoma. The indexes included the well-established Social Vulnerability Index (SoVI) and a tornado-specific social vulnerability index (TSVI). A first objective was to examine the spatial distribution of social vulnerability in Oklahoma as defined by the SoVI and TSVI. The indexes identified different areas of the state as more socially vulnerable. A second objective was to externally validate the SoVI and TSVI against a second, independent dataset that measures actual damages and loss from tornado events. Using 4 case studies from the study period and correlation analysis, we found that the SoVI and TSVI were not externally valid. The indexes did not display expected relationships and high damages and losses did not necessarily occur in areas of high social vulnerability. These findings reinforce prior findings that the relationship between social vulnerability and loss is complex, and that further revision of indexes and more validation studies are needed to fully understand their value in hazard planning.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1
1.1 Study Context 2
1.1.2 Where and Why 2
1.1.3 Scale and Study Period 6

II. LITERATURE REVIEW 8
2.1 Defining Vulnerability and Social Vulnerability 8
2.2 Tornado Risk 9
2.3 Tornadoes and Social Vulnerability 11
2.4 Existing Approaches to the Measurement of Social Vulnerability 15
2.5 Measuring Social Vulnerability 16
2.5.1 Index Construction Procedures 19
2.5.2 Index Validation 22
2.5.3 Advantages and Drawbacks of Social Vulnerability Indexes 24

III. DATA 26
3.1 Data Used in the Construction of Social Vulnerability Indices 27
3.2 Data Used to Assess Tornado-Related Damage and Loss 30
3.3 Tornado Frequency 31

IV. METHODS 34
4.1 Social Vulnerability Index Construction 34
4.2 Tornado Social Vulnerability Index Construction 36
4.3 Comparison of Indexes 37
4.3.1 Rank Order Comparison 37
4.3.2 Spatial Analysis 38
4.4 Observed Losses 39
4.5 External Validation 41

V. RESULTS 43
5.1 Indexes 43
5.1.1 Factor Analysis 43
5.1.2 County SoVI 46
5.1.3 Tract SoVI 47
5.1.4 TSVI 48
5.2 Comparison of Indexes 49

Chapter Page

5.2.1 Rank Order Comparison 49
5.2.2 Spatial Analysis 49
5.2.2.1 County SoVI 49
5.2.2.2 Tract SoVI...51
5.2.2.3 TSVI..53
5.2.3 Visual Comparison of the Indexes...55
5.3 External Validation 56
5.3.1 County Losses...56
5.3.2 Tract Losses 57
5.3.3 Tornado Events 58
5.3.3.1 Case 1..58
5.3.3.2 Case 2..60
5.3.3.3 Case 3..62
5.3.3.4 Case 4..64

VI. DISCUSSION...67
6.1 Objective 1 68
6.2 Objective 2 70
6.3 Limitations, Shortcomings, and Future Work..72
6.4 Conclusion 75

REFERENCES 77

APPENDICES 82
APPENDIX A: Terms...82
APPENDIX B: Tornado Classification Scheme...83
APPENDIX C: Tracts Removed from the Indexes...84
APPENDIX D: Variables Included in the SoVI...85
APPENDIX E: Variables Included in the TSVI...87
APPENDIX F: County SoVI Reference Map...88

LIST OF TABLES

Table Page

1. Variables Included in the SoVI..27
2. Variables Included in the TSVI 29
3. Tornado Events and Associated Losses From 2010-2014..31
4. Tornado Activity in Oklahoma for 25 Years 33
5. Factor Interpretations and Variance Explained..45
6. Spearman's *r* Values and *p* Values for County Losses......................................57
7. Spearman's *r* Values and *p* Values for Tract Losses 58

LIST OF FIGURES

Figure Page

1. Tornado Damage in Woodward, OK 1947 4
2. Tornado Damage in Woodward, OK 1947 5
3. Tornado Damage in Woodward, OK 1947 5
4. Tornado Damage in Woodward, OK 1947 6
5. Map of Census Tracts in Oklahoma 2014....7
6. Stages of Index Construction 19
7. Oklahoma Tornado Events From 2010-2014 32
8. Methodology Flowchart 36
9. County SoVI Scree Plot 44
10. Tract SoVI Scree Plot 44
11. County SoVI 2010-2014 46
12. Tract SoVI 2010-2014 47
13. TSVI 2010-2014 48
14. County SoVI – Hot Spot Analysis 50
15. County SoVI – Cluster and Outlier Analysis 51
16. Tract SoVI – Hot Spot Analysis 52
17. Tract SoVI – Cluster and Outlier Analysis 53
18. TSVI – Hot Spot Analysis 54
19. TSVI – Cluster and Outlier Analysis 55
20. Case 1 – EF3A 59
21. Case 1 – EF3B 60
22. Case 2 – EF1A 61
23. Case 2 – EF1B 62
24. Case 3 – EF4 63
25. Case 4 – EF5A 65
26. Case 4 – EF5B 66

CHAPTER I

INTRODUCTION

Vulnerability is a key component of hazards research and identifying socially vulnerable groups is a first step in creating resilient communities and reducing future losses of property and human life (Cutter 1996; Cutter et al. 2000; Flanagan et al. 2011). Natural hazards pose challenges to society, the built environment, and the natural landscape, and also stress interconnections among those systems. Developing resilient communities to decrease hazard-related losses is a top priority for local governments, policymakers, and researchers (McBean and Ajibade 2009; Murphy et al. 2015). Climate change will increase the frequency and intensity of extreme weather events and may exacerbate these stressors, increasing losses significantly (O'Brien et al. 2006). Lessons from past hazardous events such as Hurricane Katrina in 2005 and the 2010 earthquake in Haiti are prime examples of the need and importance of identifying socially vulnerable groups before the onset of extreme events (Tate 2012).

Measurement of social vulnerability is already a focus within the hazard's literature. One area of particularly intensive research attention has been the development and application of indices of social vulnerability. Indices of social vulnerability produce single metrics used to measure social vulnerability (Tate 2013). There is an ongoing need to create reliable, useful, and accurate indexes that can inform policymakers and natural hazard's scientists for better decision making (Eriksen and Kelly 2007; Schmidtlein et al. 2008; Tate 2012; Rufat et al. 2015) at various stages of the disaster cycle (Flanagan et al. 2011). However, while much research attention has gone into developing indices, less attention has been given to their validation. This lack of attention is surprising given the fact that validation of these indices is critical to their practical use (Fekete 2009; Rufat et al. 2015).

The purpose of this study is to assess the validity of two different social vulnerability indicators in the state of Oklahoma. This work has two primary research objectives. **Objective 1** is to replicate a commonly used general index of social vulnerability, develop a tornado-specific social vulnerability index, and compare the spatial pattern of vulnerability identified by each index in the state of Oklahoma. The goal of a general index is to identify the social vulnerability of a population across a range of environmental hazards. In contrast, a specific index focuses on characteristics that would make a population vulnerable to a specific hazard. In this case, the tornado-specific index will try to highlight areas of the state that are especially vulnerable to tornado events. Comparing the spatial pattern of social vulnerability identified by each index can contribute to the body of literature of place-specific case studies, index construction methods, and adaptability of pre-existing indexes. **Objective 2** is to externally validate these indexes against a second, independent dataset measuring observed losses from tornadoes. Examining the relationship between indices and actual losses might provide information on whether social vulnerability indices of different types can inform policy decisions.

1.1 Study Context

1.1.1 Where and Why

The study area for this work focused on the state of Oklahoma. Oklahoma is prone to a range of environmental hazards including drought, severe thunderstorms, and tornadoes due to its geographic location. Situated in Tornado Alley, Oklahoma and other states in this region experience higher concentrations of strong and violent tornadoes (F2-F5 on the Fujita Scale) compared to other regions in the United States (U.S.) (Daley et al. 2005, Romanic et al. 2016). The climate and meteorological conditions in this area are favorable for the construction of supercell thunderstorms that have the capacity to produce tornadoes (Lim et al. 2017). Due to its susceptibility to frequent and extreme tornado events and losses, Oklahoma has been the focus of many tornado-related studies including the works of Brooks and Doswell (2002), Daley et al. (2005), Hout et al. (2010), and Romanic et al. (2016). During the 2010-2014 study period alone, Oklahoma endured 449 tornado events (NOAA, NCEI n.d.). In addition to the state's inherent, physical risk, Oklahoma is home to a variety of social groups known to be especially vulnerable to environmental hazards and tornado events.

There is extensive history of catastrophic tornado events in Oklahoma that resulted in substantial economic and human losses. The deadliest tornado in Oklahoma occurred in April of 1947 in the city of Woodward (Romanic et al. 2016; NOAA, NWS n.d.). The F5 tornado destroyed over 1,000 homes and businesses, killed at least 116 people in and around Woodward, and caused nearly 1,000 additional injuries; some individuals were never found or identified (NOAA, NWS n.d.). See Figures 1-4 to observe some of the damage in Woodward, Oklahoma provided by the National Oceanic and Atmospheric Administration's (NOAA) National Weather Service (NWS) (n.d.). In May of 1999, over 70 tornadoes struck Oklahoma, accounting for the largest recorded tornado outbreak in the state (Romanic et al. 2016). During this event, a tornado with an F5 rating struck counties with densely populated communities in the counties of Grady, McClain, Cleveland, and Oklahoma (Daley et al. 2005). The most financially damaging tornado event recorded in the state transpired in May of 2013 in Moore, Oklahoma. A powerful tornado with a rating of EF5 caused extreme damage in Moore and the Oklahoma City metropolitan area, resulting in many injuries, causalities, and more than $2 billion in damages (Romanic et al. 2016; Lim et al. 2017).

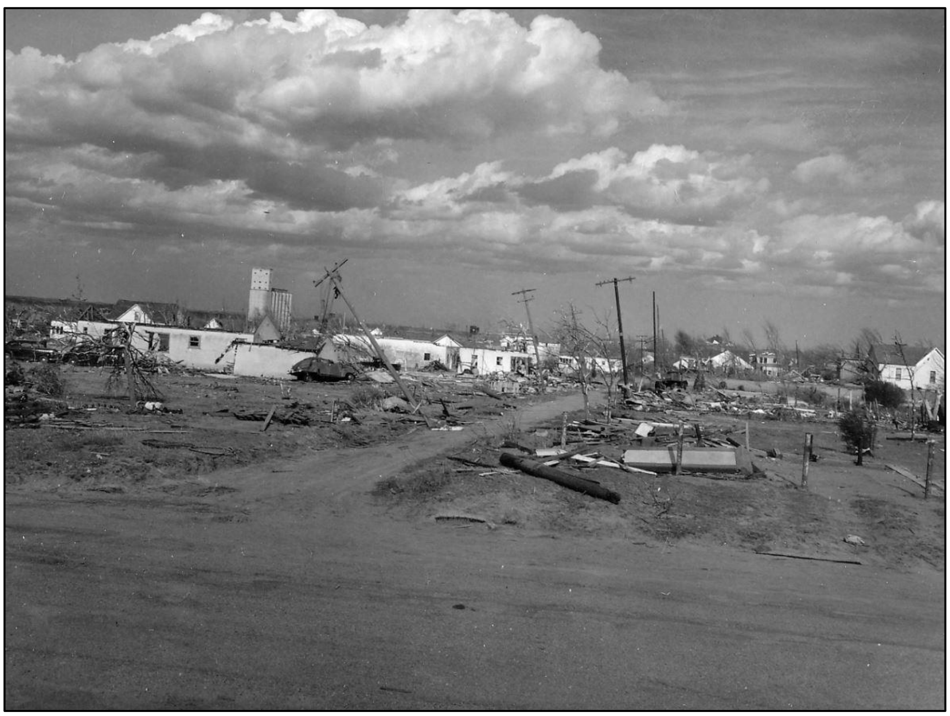


Figure 1: Tornado Damage in Woodward, OK 1947 (NWS n.d.)

Figure 2: Tornado Damage in Woodward, OK 1947 (NWS n.d.)

Figure 3: Tornado Damage in Woodward, OK 1947 (NWS n.d.)

Figure 4: Tornado Damage in Woodward, OK 1947 (NWS n.d.)

1.1.2 Scale and Study Period

This work was conducted primarily at the Census tract level because tornadoes are relatively short, localized events. This scale allows for better representation of the distribution of Oklahoma's population and can help identify socially vulnerable areas. Census tracts are small subdivisions of a county that typically represent an average population size of 4,000 but can range between 1,200 to 8,000 people (U.S. Census Bureau 2012). See Figure 5 for a map of Oklahoma's Census tracts. The map also includes the Oklahoma City and Tulsa metropolitan statistical areas (MSA), two major cities in the state with relatively high population densities. Since the observed losses are recorded at the county scale, this work also includes some county level analyses when necessary.

The study period covers the 5-year period from 2010-2014 and incorporated American Community Survey (ACS) data. The ACS provides up-to-date community estimates in between the full Census counts conducted every ten years (U.S. Census Bureau 2018). This study period was chosen because this work is an extension of a larger project focused on social vulnerability in Oklahoma. The project was supported by an Oklahoma Established Program to Stimulate Competitive Research (EPSCoR) grant (OIA-1301789) through the National Science Foundation.

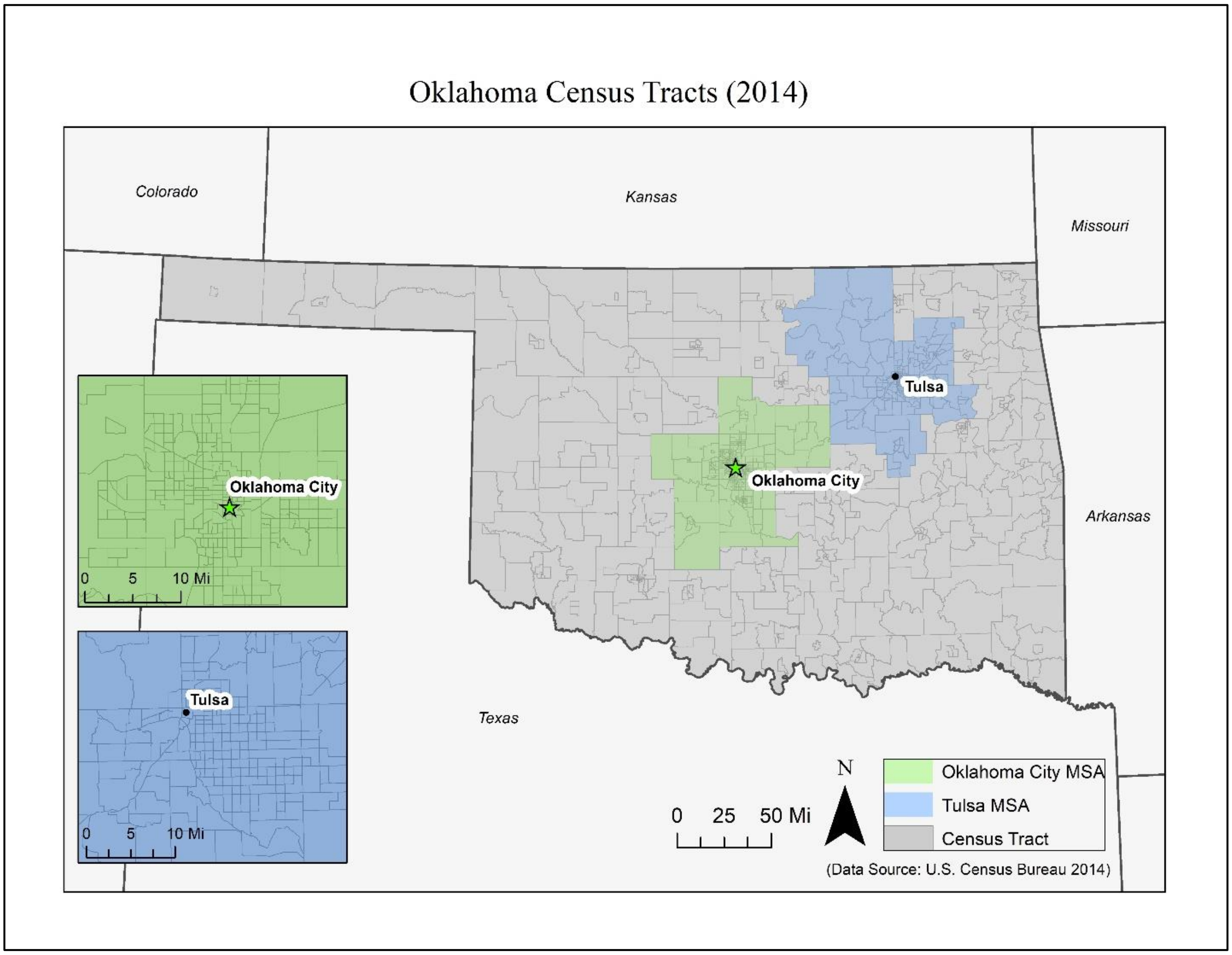


Figure 5: Map of Census Tracts in Oklahoma 2014

CHAPTER II

LITERATURE REVIEW

2.1 Defining Vulnerability and Social Vulnerability

Many authors have stressed the importance of establishing a conceptual framework prior to conducting studies of vulnerability because social vulnerability has different meanings throughout the natural hazards and social sciences literature. Fluctuations in the definition of social vulnerability are related to variation in what related concepts are prioritized (e.g., adaptive capacity, resilience, etc.). These decisions are not trivial. Changing how social vulnerability is measured can alter the outputs of a study and recommendations for policy (Ciurean et al. 2013; Murphy et al. 2015). Researchers commonly adjust the definition used to the purpose of a study (Cutter et al. 2003; Ciurean et al. 2013). Some scholars have defined vulnerability as the "potential for loss" (Cutter 1996, Cutter et al. 2000, 2003; Ciurean et al. 2013), the "likelihood to experience harm" (Boruff et al. 2003; Turner et al. 2003), and "the exposure and sensitivity of a system" (Cutter et al. 2008). What constitutes "loss" is rarely explicitly declared in definitions, but typically refers to a loss of property or life (Cutter et al. 2000).

Across definitions, the concept of social vulnerability incorporates characteristics of both societies and the built environment. Flanagan et al. (2011) explain social vulnerability as, "socioeconomic and demographic factors that affect the resilience of communities." Social vulnerability therefore describes the susceptibility of social groups to potential losses from extreme natural events due to innate characteristics (e.g., ethnicity) or acquired characteristics (e.g., beliefs, customs) (Cutter et al. 1996). Social vulnerability is a result of various conditions including risk, exposure, adaptive capacity, sensitivity, and hazard mitigation measures (Cutter 1996; Cutter et al. 2008). A summation of terms associated or substituted with social vulnerability can be found in Appendix A.

Across definitions, research frameworks integrating social components and vulnerability recognize the interconnectedness of the natural systems, social systems, and the built environment. Each of these frameworks stress that biophysical risk (the likelihood of an event occurring at a defined location or the proximity to an event) and societal risk to environmental hazards cannot be separated (Cutter 1996; Cutter et al. 2008; Ciurean et al. 2013). Social vulnerability therefore combines these two ideas and focuses on examining vulnerability from a socioecological perspective. It also includes other factors that influence vulnerability such as economic, social, environmental, institutional, and political characteristics (Ciurean et al. 2013). This approach to vulnerability science also highlights equity and human rights issues because people are not affected equally by extreme weather events (Cutter et al. 2003, 2008; Flanagan et al. 2011). Socially vulnerable groups are more likely to experience greater losses from a hazardous event and are less likely to recover afterwards (Flanagan et al. 2011).

2.2 Tornado Risk

In addition to constructing a general social vulnerability index, a tornado-specific social vulnerability index will be produced (**Objective 1**). Tornadoes are localized, violent hazards that have the capacity to cause detrimental effects on communities (Simmons and Sutter 2011; Dixon and Moore 2012; Widen 2016). They can, and have, caused considerable economic damage and loss of human life (Romanic et al. 2016; Lim et al. 2017). In the U.S., tornadoes are most commonly reported in a region known colloquially as Tornado Alley. The historic Tornado Alley, which encompasses most of the Great Plains, includes Texas, Oklahoma, Kansas, Nebraska, and Iowa (Coleman and Dixon 2013). However, tornadoes can affect all communities in the U.S.; therefore, all people have some level of tornado vulnerability. The amount of tornado vulnerability varies across space and time due to the different degrees of physical exposure, societal risk (Pielke and Pielke 1997; Dixon and Moore 2012; Romanic et al. 2016), and adaptive capacity (Widen 2016). Understanding the characteristics of society and the built environment that leave groups vulnerable to these hazards can assist in the effort to reduce future losses.

Many studies have explored tornado risk from a biophysical perspective while incorporating tornado-related fatalities such as the works of Boruff et al. (2003), Ashley (2007), Ashley et al. (2008), Coleman and Dixon (2013), and Shen and Hwang (2015). As discussed by Simmons and Sutter (2011), tornado-related causality data are used far more frequently than property damage and economic loss. While examining tornado-related fatalities, Ashley (2007) found the American South to have the greatest fatality rate over any other region in the U.S. from 1985-2005. The author attributed the elevated fatality count to the onset time of the tornado events because this region has a greater occurrence of nocturnal events. These findings, further explored by Ashley et al. (2008), suggest communities in the southern U.S. are more vulnerable to nocturnal tornadoes because they are less likely to seek shelter when woken up during the night (if woken up at all) by a tornado siren system. In addition, tornado siren systems are not uniformly effective (see Mathews et al. 2017). Complementary to these works, Coleman and Dixon (2013) sought to identify regions of the U.S. with the greatest amount of tornado risk from

1973-2011. They used the pathlength of tornadoes rather than the number of events or fatalities, which allows for a more comprehensive understanding of the areas impacted by an event. They found the areas of greatest risk during this study period stretch from Oklahoma to Alabama, with a significant maximum extending from central Mississippi into northern Alabama (Coleman and Dixon 2013).

2.3 Tornadoes and Social Vulnerability

Due to geographic and climatic differences, certain areas are more prone to violent tornado events. Over the past 50 years, tornado-related causalities and injuries have decreased due to the improvements in radar and warning technologies (Boruff et al. 2003) and social media engagement (Ripberger et al. 2014), but socially vulnerable groups are still disproportionately affected by hazardous events (Flanagan et al. 2011; Lim et al. 2017). In addition, areas with high tornado-related deaths do not necessarily mirror the areas exposed to tornadoes with greater Enhanced Fujita (EF) scale ratings (Lim et al. 2017). A brief description of the tornado classification scheme can be found in Appendix B. To account for these discrepancies, some works have identified socioeconomic factors that may increase tornado vulnerability. Many of these characteristics overlap with those known to influence social vulnerability to a range of environmental hazards. Some of these drivers include lack of access to resources, rural populations, physically limited individuals, and non-English speaking minorities (Chaney et al. 2013; Lim et al. 2017).

Income can serve as an indicator for tornado vulnerability. Dixon and Moore (2012) found income tightly linked to people’s ability to recover from hazard-related losses, suggesting people below the poverty line are socially vulnerable. In contrast, Chaney et al. (2013) found individuals with higher incomes tend to exercise enhanced preparedness measures, reducing losses from extreme events. Households with an average annual income less than $40,000 are significantly less likely to have a plan for seeking shelter and their homes are less likely to have a tornado-resistant space (e.g., basement or storm shelter) compared to those with higher incomes (Chaney et al. 2013). Lim et al. (2017) also found per capita income (PCI) to be indicative of tornado impacts; counties with higher levels of PCI had fewer tornado-related fatalities, whereas counties with greater poverty rates and more income disparities had significantly more tornado-related damages. Generally, as discussed by Fothergill and Peek (2004), individuals below the poverty line are significantly more vulnerable to natural hazards due to their, "type of residence, building construction, access to information, low quality infrastructure, and social exclusion" (Lim et al. 2017, 6).

Residents living in mobile homes are especially vulnerable to tornado events. Mobile home residents have an extremely high tornado-related fatality rate and are 20 times more likely to endure a tornado-related casualty than residents in other structures (Brooks and Doswell 2002; Ashley 2007; Chaney and Weaver 2010; Dixon and Moore 2012; Chaney et al. 2013; Lim et al. 2017). NOAA and the U.S. Census Bureau note that despite mobile homes only making up 8% of housing in the U.S. from 1996-2000, mobile home residents made up almost half of all tornado-related deaths during this period (Lim et al. 2017). Similarly, Ashley (2007) found mobile homes residents accounted for 44% of all tornado related fatalities in the U.S. during their study period (1985-2005), and this number has only continued to rise. Mobile homes provide poor physical protection and are extremely vulnerable to the elements, particularly the intensity of tornado events. Residents in these structures are also less likely to devise a plan for seeking proper, safe shelter during tornadoes compared to people living in permanent, traditional style homes (brick or wood-frame) (Chaney and Weaver 2010; Chaney et al. 2013). They often seek shelter within their homes, which may contribute to their high tornado-related fatality rate (Chaney and Weaver 2010).

People living in mobile homes are also perceived to have lower incomes and lower education levels, increasing their level of tornado vulnerability (Chaney and Weaver 2010; Chaney et al. 2013). These factors can pose challenges for residents because they may not receive tornado warnings, and/or may be unsure how to properly respond to warnings if they do not have pre-existing risk-reduction plans in place (Chaney and Weaver 2010). Lim et al. (2017) also found educational attainment (population 25 years or older with a Bachelor, or higher degree) to decrease peoples' vulnerability to tornadoes because it is associated with efficient emergency decision-making and likelihood to perform proper evacuation measures when necessary (e.g., seeking shelter). This group is also found to be more likely to have access to recovery information and resources after an event (Lim et al. 2017). In addition, Chaney et al. (2013) found that increases in community involvement and bonding of citizens in tornado-prone areas were less socially vulnerable and more likely to participate in tornado drills. However, mobile homes residents often rent and are not homeowners, reducing the likelihood of having tight ties to the community and less likely to have participated in a tornado drill prior to an event (Chaney and Weaver 2010; Chaney et al. 2013).

Elderly populations are vulnerable to extreme weather events (Ashley 2007; Dixon and Moore 2012; Chaney et al. 2013). Ashley (2007) found this group (65 and older) to be statistically significant among vulnerable groups and Chaney et al. (2013) suggest they are less likely to be prepared for a tornado event. Elderly groups are also less likely to have participated in a tornado drill compared to younger populations, increasing their susceptibility to harm

(Chaney et al. 2013). This difference may also be attributed to the increased access that younger populations (aged 20-39, 40-59) have to tornado preparedness exercises at work or school, while older populations likely have fewer opportunities and exposure. In addition, elderly populations can be less mobile and may have pre-existing health conditions or challenges that interfere with their ability to seek shelter during tornado events that often require quick reaction times (Dixon and Moore 2012).

Family structure and household dynamics are also indicators of tornado vulnerability. While evaluating household preparedness, Chaney et al. (2013) found no difference between homes with or without children regarding the presence of a plan to seek adequate shelter during a tornado event. However, homes with children are more likely to have participated in tornado drills and practiced their plans compared to homes without children; they were better prepared. This suggests that homes without children are more socially vulnerable and more likely to suffer tornado-related losses (Chaney et al. 2013). In addition, Blaikie et al. (1994) and Lim et al. (2017) argue that women-headed households are among the most socially vulnerable. Wisner et al. (2004) note that it is not women individually who are more socially vulnerable due to their gender, but rather their heightened vulnerability stems from the situations they are more susceptible to (compared to men). For example, women often face challenges during the recovery stage of an event due to their employment status, lower wages, and family care responsibilities (Blaikie et al. 1994; Lim et al. 2017). Further support of this was found by Lim et al. (2017). Results from their study convey a positive correlation of women-headed households and tornado-related causalities. This suggests that women and women-headed households are disproportionately affected by tornadoes because they have limited access to resources (Lim et al. 2017).

Few studies have attempted to quantify social vulnerability to tornadoes. Adapting methods from Cutter et al. (2003) SoVI, Dixon and Moore (2012) explored the spatial distribution of tornado vulnerability in Texas using the framework presented by Pielke and Pielke (1997). This framework defines tornado vulnerability as a sum of the incidence of tornadoes and societal exposure to the hazardous event. Dixon and Moore (2012) presented three different methods to assess tornado vulnerability (using significant and violent events EF2-EF5 on the Enhanced Fujita scale), by assigning scores for each county in Texas. They considered the most vulnerable counties to be those with both high incidences of events and high societal exposure. After comparing three different methods to quantify tornado vulnerability, Dixon and Moore (2012) found the spatial distribution of vulnerability is reliant on and sensitive to the choice of method employed.

2.4 Existing Approaches to the Measurement of Social Vulnerability

Social vulnerability cannot be directly measured because it is not an observed phenomenon (Fekete 2009; Hinkel 2011; Tate 2012). Efforts to analyze vulnerability and the complex human-environment system that generates it are grounded in different conceptual frameworks and models. Some noteworthy models include the pressure-and-release (PAR) model (Turner et al. 2003; Cutter et al. 2008; Ciurean et al. 2013), the vulnerability/sustainability framework (Turner et al. 2003; Cutter et al. 2008), the hazards-of-place model (Cutter 1996; Cutter et al. 2008), and the disaster-resilience-of-place model (Cutter et al. 2008). The PAR model views risk as a function of the various stressors on a system (from a hazard) and inherent vulnerability of communities (Turner et al. 2003; Ciurean et al. 2013). Criticism of the PAR model suggests it does not incorporate the vulnerability of the biophysical stressors on a system and exclusively focuses on social vulnerability, missing the other half of the human-environment system (Turner et al. 2003; Cutter et al. 2008). The vulnerability/sustainability framework by Turner et al. (2003) focuses on place-based, local vulnerability but lacks a temporal component and does not differentiate between exposure and sensitivity (Cutter et al. 2008). The hazards-of-place model analyzes vulnerability as both a biophysical risk and social response within a defined geographic area. This model provides the foundation for the social vulnerability index, SoVI (Cutter et al. 2003). More recently, the disaster-resilience-of-place (DROP) model was introduced to attempt to measure resilience and highlight the relationship between resilience and vulnerability with respect to the built environment and the natural and social systems. The DROP model focuses on the social resilience of places to natural hazards (Cutter et al. 2008).

2.5 Measuring Social Vulnerability

Objective 1 of this study is to replicate a commonly used general index of social vulnerability (SoVI), develop a tornado-specific social vulnerability index (TSVI), and compare the spatial pattern of vulnerability identified by each in the state of Oklahoma. Adopting the hazards-of-place framework, this work will measure social vulnerability using a set of variables to create indexes that work as a proxy. The variables reflect the drivers known to influence social vulnerability such as characteristics of social groups, places, and the built environment (Cutter et al. 2003; Ciurean et al. 2013; Tate 2013). To attempt to ensure the indexes are meaningful and valid, individual variable selection should reflect the adopted conceptual framework (Tate 2012, 2013) and be supported by existing literature and case studies (Cutter et al. 2003; Eriksen and Kelly 2007; Hinkel 2011; Murphy et al. 2015). Throughout the vulnerability and natural hazards literature, there is general agreement about the overarching factors that influence social vulnerability. However, differences arise in the selection of specific variables to reflect these broader concepts (Cutter et al. 2003). Indicators attempt to quantify social vulnerability to natural hazards and aid in simplifying the complex reality of the interconnected systems (Eriksen and Kelly 2007; Hinkel 2011; Tate 2012; Murphy et al. 2015; Rufat et al. 2015). Quantifying social vulnerability can help identify areas that are most susceptible to loss of property and life during hazardous events and allow for objective comparisons of levels of vulnerability across scales and boundaries (Cutter 1996; Cutter et al. 2003; Eriksen and Kelly 2007; Tate 2012; Rufat et al. 2015).

Although several alternative social vulnerability indexes exist, those by Cutter et al. (2003) SoVI and Flanagan et al. (2011) SVI are the most commonly used. The SoVI was initially constructed from 42 U.S. Census Bureau variables to represent the major components known to influence social vulnerability. These components are generally well understood throughout the hazards and vulnerability literature; however, there is variation in the individual variables used to represent them. The major themes influencing social vulnerability include, "lack of access to resources, limited access to political power and representation, social capital, beliefs and customs, building stock and age, frail and physically limited individuals, and the type and density of infrastructure and lifelines" (Cutter et al. 2003, 245). To reduce data using this index, principal component analysis (PCA) condenses the variables into factors that represent the larger dataset and provide SoVI scores for each county (or geographic area being examined).

The SoVI has been used in numerous studies including work led by Armaș and Gavriș (2013), Letsie and Grab (2015), and Zebardast (2013). Most frequently, scores are mapped to display the most and least socially vulnerable areas based on standard deviations from the mean (Cutter et al. 2003). The SoVI and related work at the University of South Carolina's (USC)

Hazards & Vulnerability Research Institute (HVRI) has received funding from a range of government entities including the Federal Emergency Management Agency (FEMA), the National Oceanic and Atmospheric Administration (NOAA), the National Science Foundation (NSF), and the National Aeronautics and Space Administration (NASA) (FEMA 2018).

While the SoVI is designed as a general index, it is not the only one. One alternative to the SoVI is the SVI established by the U.S. Centers for Disease Control and Prevention (CDC) through a collaboration of the Office of Terrorism Preparedness and Emergency Response (OTPER) and the Agency for Toxic Substances and Disease Registry's Geospatial Research, Analysis, and Services Program (ATSDR). Four "domains" form the base of the SVI and within each domain are U.S. Census Bureau variables that influence social vulnerability. The four domains include socioeconomic status (e.g., poverty, education), household composition/disability (e.g., age, disability), minority status/language (e.g., race, ethnicity), and housing/transportation (e.g., housing structure, vehicle access). Together, 15 indicator variables describe the four domains and are assigned SVI values, followed by SVI values applied to each of the four domains. The geographic areas also receive SVI scores in addition to percentile ranks for the 15 variables (Flanagan et al. 2011). Most importantly, the SVI construction procedures can form the basis of the creation of specific indices. Specific indices are needed because they can be more descriptive and informative than general indices and potentially identify groups that were not detected in the general index. These indexes can also inform policy and decisions surrounding specific hazardous events, like tornadoes. Various agencies have adapted the CDC's SVI in practice including the state of Vermont's Department of Public Health, New Hampshire's Department of Health and Human Services, and the U.S. Climate Resilience Toolkit (CDC, ATSDR 2013).

2.5.1 Index Construction Procedures

The construction of a social vulnerability index and the subjective decisions required to develop it are rooted in the chosen conceptual framework and, ultimately, what is being measured (the purpose of the study) (Schmidtlein et al. 2008; Tate 2012, 2013; Yoon 2012). Following the stages of index construction discussed by Tate (2012) and outlined in Figure 6, after formulating a conceptual framework and identifying what the index will measure, the index structure (deductive, hierarchical, or inductive) must be chosen. After identifying the variables that will serve as indicators, the geographic scale at which it will be applied, and a validation medium must be chosen. In addition to these steps, Rufat et al. (2015) expresses the importance of specifying the phase of an event being studied because social groups may be vulnerable at different stages of a hazardous event (preparedness, response, and recovery).

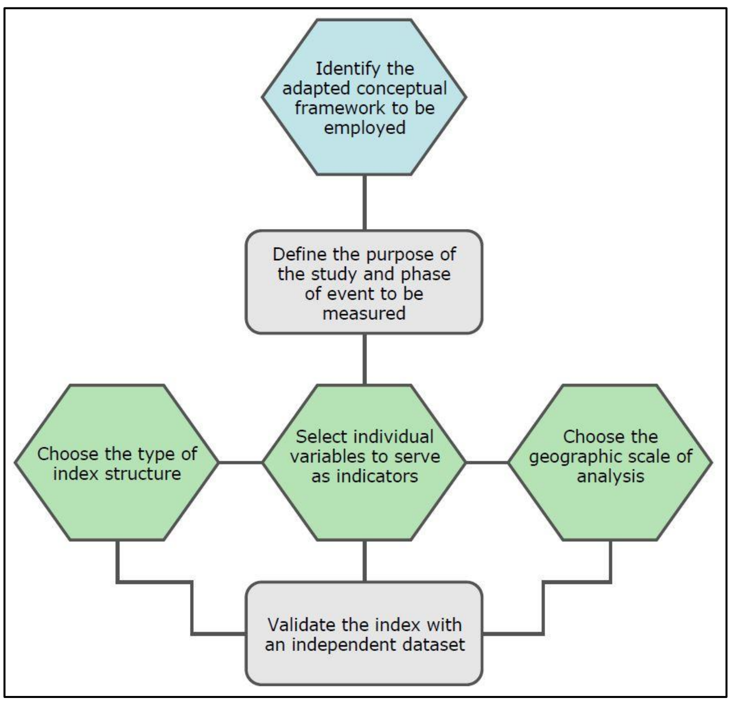


Figure 6: Stages of Index Construction (Adapted from Tate (2012))

An index can be constructed using deductive, hierarchical, or inductive designs. Deductive designs are constructed with normalized variables that are aggregated to an index. This arrangement was most dominant in early social vulnerability indexes and typically have up to 10 indicator variables that are individually selected. Hierarchical designs, such as the CDC's SVI, consist of 10-20 indicator variables that are organized into thematic groups to represent the various factors known to influence social vulnerability. These index designs are the foundation for the TSVI. Lastly, indexes constructed inductively are characterized by larger datasets that are aggregated to a few representative components. These designs, which include the SoVI, typically employ 20 or more variables (Eriksen and Kelly 2007; Tate 2012, 2013; Yoon 2012). These variables are reduced to smaller groups, usually through factor analysis, most notably done by Cutter et al. (2003). Inductive indexes frequently incorporate z-score standardization and principal components analysis (PCA). With PCA, factor selections are commonly chosen based on the Kaiser criterion, supporting the inclusion of all factors with an eigenvalue greater than 1 (Rogerson 2015). Work led by O'Connor (2000) and Patil et al. (2008) suggest the Kaiser criterion may overvalue the number of "important" factors (those with an eigenvalue greater than 1), leading to an increase in the components (factors) that describe social vulnerability (Tate 2012).

One possible solution to these concerns, as discussed by Schmidtlein et al. (2008), is to incorporate local experts from the study area to help interpret and name the factors presented by the PCA analysis. The guidance and incorporation of their expert geographic knowledge of the study area and vulnerable groups can aid in understanding the factor loadings and may even assist in validation of the results (Schmidtlein et al.

2008). Authors have also discussed the option to weight variables during this stage (Schmidtlein et al. 2008; Hinkel 2011; Tate 2012, 2013; Yoon 2012; Rufat et al. 2015). Greater weights can be assigned to variables known to have a larger influence on social vulnerability. Equal weights suggest all variables contribute equally to identify socially vulnerable groups. The decision to weight variables versus the decision to use equal weights is highly subjective (Brooks et al. 2005; Tate 2012, 2013).

Another important step in index construction is designing the index best suited for the scale at which it will be applied. Relationships between variables change at different scales, so tailoring index construction to the anticipated scale is crucial (Tate 2012, 2013; Yoon 2012). As discussed by Fekete et al. (2010), the chosen scale of analysis is most influenced by data availability, policy demand, and the adopted conceptual framework. Social vulnerability is not yet fully understood at regional, national, or global scales, so focusing on the local scale first may assist in measuring and understanding social vulnerability at larger scales (Eriksen and Kelly 2007; Schmidtlein et al. 2008; Fekete et al. 2010; Rufat et al. 2015). Local-level studies are also a good starting point because one can observe the interaction of different systems and components influencing social vulnerability best at this scale (Eriksen and Kelly 2007; Fekete et al. 2010). Hinkel (2011) elaborates, suggesting indicator studies work best when applied locally, and when constructed mindfully can successfully identify vulnerable people, regions, or sectors with a narrowly defined index.

2.5.2 Index Validation

Objective 2 is to externally validate the constructed indexes (SoVI and TSVI) against actual tornado-related damages and loss. Social vulnerability indexes are largely descriptive measures and report the characteristics of a community (Eriksen and Kelly 2007; Murphy et al. 2015; Rufat et al. 2015). They may suggest relationships between variables that are solely hypothetical and might not exist in the real world. Indexes also tend to be outcome-oriented (focused on populations and losses rather than the broader system) (Murphy et al. 2015), linear (have a narrow and simplistic view of vulnerability as physical risk and exposure) (Hinkel 2011; Murphy et al. 2015), and used to inform decision makers (to communicate areas of greatest social vulnerability), which is why validation of an index is essential (Eriksen and Kelly 2007; Hinkel 2011; Tate 2012; Ciurean et al. 2013). Validation of indexes provide insight into whether or not an index measures what it is intended to and can be performed externally or internally. External validation can be achieved using a second, independent dataset and internal validation is typically done through sensitivity analysis and changes to the construction of an index (Fekete 2009; Tate 2012).

Just as there is no best index (Cutter et al. 2003), there is no universal validation medium for social vulnerability indexes (Schmidtlein et al. 2008). Attempts to validate these indexes externally have resulted in varied degrees of success using presidential disaster declarations (Cutter et al. 2003), hazard zone delineation (Cutter et al. 2000), disaster property damages (Yoon 2012), and built environment damage (Burton 2010), among others (see Brooks et al. 2005; Myers et al. 2008; Finch et al. 2010; Flanagan et al. 2011). External validation helps to uncover if the index is representative of the areas and social groups suffering losses compared to other data. For example, in their study examining the validity of a social vulnerability index (SVI) against extreme river-floods in Germany, Fekete (2009) created an independent second dataset consisting of three binary dependent variables. This allowed for the index to be compared against external data. Independent second datasets can be scarce, so Fekete (2009) generated the dependent variables from a series of interviews of households affected by the floods during the study period. With a regression model, Fekete (2009) identified the independent variables that are significant and was able to isolate the ones (9 out of the original 41) that described vulnerability for this study area.

Alternatively, validation of indicators can also be conducted internally through sensitivity analyses. This is done through alterations to the construction of an index or via simulations to compare any differences or similarities in output results and reveal what the index is sensitive to at different stages (Schmidtlein et al. 2008; Fekete 2009; Tate 2012; Ciurean et al. 2013). As discussed by Tate (2012), sensitivity analysis can be achieved using local analysis or global analysis tests. Local sensitivity analysis is simple and usually involves correlation and analyses of variance. The local sensitivity tests can only change one aspect of the index at a time, such as variable selection or the scale to which the index is implemented (Tate 2012).

Global sensitivity analysis is more complex and allows for multiple stages of an index to be assessed simultaneously (Fekete 2009; Tate 2012). There are four crucial steps to global sensitivity analysis including sample selection, Monte Carlo simulation, uncertainty analysis, and sensitivity analysis (Tate 2012). The global sensitivity tests explore model structure, indicator set, analysis scale, weighting, and other components.

Tate (2012) also suggests no index structure (deductive, inductive, or hierarchical) is better or worse for measuring social vulnerability. Rather, the construction of the index and its sensitivities should influence the type of index structure adopted. Ultimately, Eriksen and Kelly (2007) and Tate (2012) stress the need to justify and explain the decisions underlying the construction of a social vulnerability index. Since there are different ways to construct an index, which requires subjective input, (Schmidtlein et al. 2008; Fekete 2009; Tate 2013), it is vital to communicate to the reader why it is designed a certain way (Tate 2012, 2013; Ciurean et al. 2013).

2.5.3 Advantages and Drawbacks of Social Vulnerability Indexes

Based on the discussion above, there are several benefits and shortcomings to indicator studies. They are typically cost effective because data can be obtained from readily available and accessible sources such as the U.S. Census Bureau. Indexes are also inherently spatial and can be tailored to localized areas. In addition, the results of vulnerability indicators can be mapped at various scales that allow for the examination of patterns or hot spots to assess the spatial distribution of social vulnerability. With constant changes in societies due to population growth, migration, aging, and other influences, understanding the relationship between scale and social vulnerability through index implementation can assist in future mitigation measures, as climate change will increase the frequency and intensity of hazardous events (O'Brien et al. 2006; Fekete et al. 2010).

However, social vulnerability indicators may not always measure what they are designed to and can oversimplify complex relationships between variables (Murphy et al.

2015). The indicating variables included in an index are just as important as those that are not. Variables known to increase social vulnerability to natural hazards but are more challenging to quantify are the perceptions and attitudes of citizens, especially in areas with higher incidences of tornadoes. These characteristics, as discussed by Cohen and Nisbett (1998), Ashley (2007), and Ashley et al. (2008), are typically not included in an index. Various studies (see Sims and Baumann 1972; Biddle 1994; Cohen and Nisbett 1998) have tried to quantify these perceptions and behavioral trends and connect them to the spatial distribution of tornado events and fatality reports in the U.S. (Ashley 2007). The attitudes and perceptions of citizens directly influence their will to prepare for extreme weather events, and a lack of preparedness increases risk and will leave individuals more vulnerable and susceptible to losses (Chaney et al. 2013). Rufat et al. (2015) challenge readers to consider the validity and application of indicators when characteristics that greatly influence social vulnerability (such as fatalistic attitudes) are not included. To account for the shortcomings of indexes and fill the gaps of the characteristics left out, qualitative analysis techniques can help provide a holistic understanding of social vulnerability (Ashley et al. 2008; Schmidtlein et al. 2008).

CHAPTER III

DATA

This work has two primary objectives. **Objective 1** is to replicate a commonly used general index of social vulnerability, develop a tornado-specific social vulnerability index, and compare the spatial pattern of vulnerability identified by each in the state of Oklahoma. A general index of social vulnerability was constructed following Cutter et al. (2003), while a tornado-specific index was constructed deductively based on tornado-focused studies (see Blaikie et al. (1994), Brooks and Doswell (2002), Ashley (2007), Mulilis et al. (2000), Chaney and Weaver (2010), Dixon and Moore (2012), Chaney et al. (2013), Widen (2016), and Lim et al. (2017). The construction of each index required the measurement of several socioeconomic variables at the county and Census tract levels. Necessary data were collected from the U.S. Census Bureau (2014). **Objective 2** is to externally validate the two social vulnerability indexes against a secondary, independent dataset measuring observed losses from tornadoes. The data used in this comparison was collected from NOAA and included information on tornado events and observed losses related to those events.

3.1 Data Used in the Construction of Social Vulnerability Indices

This work incorporated a general social vulnerability index (SoVI), designed to measure social vulnerability to a range of environmental hazards and a tornado-specific social vulnerability index (TSVI), intended to measure social vulnerability to tornado events. Data used in the SoVI and TSVI were gathered from the U.S. Census Bureau's American Community Survey (ACS) 5-year estimates for 2010-2014. Data measuring total population was gathered from the 2010 Census summary, from which one table was utilized. The variables were collected and analyzed at the county and tract scale. The county and tract boundaries for Oklahoma were also obtained from the U.S. Census Bureau. These boundary files were sourced from the Topologically Integrated Geographic Encoding and Referencing (TIGER) products and reflect the 2014 data, the last year of the study period.

Due to missing data, 1 Census tract was eliminated from the TSVI and 24 tracts were removed from the SoVI. These tracts, and the variable data they are missing (why they were withheld from index construction), are outlined in Appendix C. Some of these tracts lack data because they represent airports, college campuses, or reservations.

The construction of the general Social Vulnerability Index, SoVI, followed the work of Cutter et al. (2003) and Tate (2012). Cutter et al. (2003) identifies the 7 major themes known to influence social vulnerability and selected variables to represent these drivers in the index. These themes include lack of access to resources, limited access to political power and representation, social capital, beliefs and customs, building stock and age, frail and physically limited individuals, and type and density of infrastructure and lifelines. Many of the variables in the index represent more than one theme. These variables can be found in Table 1 and more detailed information about these variables and their data source can be found in Appendix D.

Table 1: Variables Included in the SoVI

VARIABLE	DESCRIPTION		VARIABLE	DESCRIPTION
QFHH	Percent, families with female-headed households, no spouse present		QRENTER	Percent, renter-occupied housing units
PPUNIT	Estimate, households by type, average number of people per household		QNOAUTO	Percent, housing units with no car available
QED12LES	Percent, educational attainment, population over 25 years old, no high school diploma		MHSEVAL	Estimate, median dollar value of owner-occupied housing units, home value
QESL	Percent, population speaking English as a second language, limited English proficiency		MDGRENT	Estimate, median gross rent for renter-occupied housing units
QCVLUN	Percent, civilian labor force unemployed		QFEMALE	Percent, female population
QFEMLBR	Percent, female participation in the labor force		QBLACK	Percent, Black population
QSERV	Percent, population in service occupations		QNATAM	Percent, Native American population
QEXTRCT	Percent, employment in extractive industries (fishing, farming, mining, etc.)		QASIAN	Percent, Asian population
QRICH200K	Percent, income and benefits, families earning more than $200,000 per year		QHISP	Percent, Hispanic population
QSSBEN	Percent, households receiving Social Security benefits		MEDAGE	Median age
PERCAP	Estimate, income and benefits, per capita income (dollars)		QAGEDEP	Percent, population under 5 years of age or 65 and over, the sum of six categories, divided by TOTPOP from 2010 Census
QNOHLTH	Percent, population without health insurance		QFAM	Percent, children living in married couple families
QPOVTY	Percent, persons living in poverty		QMOHO	Percent, population living in mobile homes, sum of ownership types (owner, renter) divided by total MOHO
QUNOCCHU	Percent, unoccupied housing units, vacancy		QNRES	Percent, population living in nursing facilities, sum and then divided by TOTPOP from 2010 Census

To represent lack of access to resources in the SoVI, the variables QFHH, QED12LES, QESL, QCVLUN, QNOHLTH, QPOVTY, QNOAUTO, QFEMALE, QBLACK, QNATAM, QASIAN, QHISP, and QFAM were included; limited access to political power and representation was represented with QFHH, QED12LES, QESL, QCVLUN, QFEMLBR, QSERV, QEXTRCT, QPOVTY, QRENTER, QFEMALE, QBLACK, QNATAM, QASIAN, and QHISP; and the variables for social capital included PPUNIT, QRICH200K, QSSBEN, MHSEVAL, MDGRENT, and MEDAGE. Also, in the SoVI, beliefs and customs were included using QFHH, PPUNIT, QRICH200K, QSSBEN, QNATAM, QASIAN, and QHISP; and the variables QUNOCCHU, QRENTER, MHSEVAL, MDGRENT, and QMOHO represented building stock and age. To include frail and physically limited individuals the variables QNRES and QAGEDEP were used; and the variables PPUNIT, QRENTER, QMOHO, and QNRES were added to characterize the type and density of infrastructure and lifelines.

The tornado-specific social vulnerability index, TSVI, was constructed using variables following a review of related tornado-focused studies (see Blaikie et al. (1994), Brooks and Doswell (2002), Ashley (2007), Mulilis et al. (2000), Chaney and Weaver (2010), Dixon and Moore (2012), Chaney et al. (2013), Widen (2016), and Lim et al. (2017)). Building on this prior work, 10 variables were selected for the TSVI including QPOVTY, QRICH200K, QMOHO, QED12LES, QELDERLY, QESL, QCVLUN, QFHH, QFAM, and QRENTER. These variables can be found in Table 2 and more detailed information about these variables and their data source can be found in Appendix E.

Table 2: Variables Included in the TSVI

VARIABLE	DESCRIPTION	LITERATURE SOURCE
QPOVTY	Percent, persons living in poverty	Dixon and Moore 2012; Widen 2016; Lim et al. 2017
QRICH200K	Percent, income and benefits, families earning more than $200,000 per year	Chaney et al. 2013; Lim et al. 2017
QMOHO	Percent, population living in mobile homes, sum of ownership types (owner, renter) divided by total MOHO	Brooks and Doswell 2002; Ashley 2007; Chaney and Weaver 2010; Dixon and Moore 2012; Chaney et al. 2013; Widen 2016; Lim et al. 2017
QED12LES	Percent, educational attainment, population over 25 years old, no high school diploma	Chaney and Weaver 2010; Widen 2016
QELDERLY	Population 65 and over, the sum of 5 categories	Ashley 2007; Dixon and Moore 2012; Chaney et al. 2013; Widen 2016
QESL	Percent, population speaking English as a second language, limited English proficiency	Chaney et al. 2013; Widen 2016; Lim et al. 2017
QCVLUN	Percent, civilian labor force unemployed	Widen 2016
QFHH	Percent, families with female-headed households, no spouse present	Blaikie et al. 1994; Lim et al. 2017
QFAM	Percent, children living in married couple families	Chaney et al. 2013
QRENTER	Percent, renter-occupied housing units	Mulilis, Duval, and Bovalino 2000

3.2 Data Used to Assess Tornado-Related Damage and Loss

Objective 2 required the allocation of damages and losses related to tornadoes to the Census geographies used in index construction. Observed tornado event data were obtained from NOAA’s NWS Storm Prediction Center (SPC) database (www.spc.noaa.gov/gis/svrgis/) as a Geographic Information System (GIS) shapefile. The tornadoes were mapped as straight-line paths for each event with several characteristics including the date, time, EF scale rating (see Appendix B), and start and end locations. This work recognizes that tornadoes do not always occur in perfectly straight lines and do not necessarily remain in contact with the ground along the entire path (Widen 2016). In addition, the intensity of an event often changes throughout a tornado path, but the EF ratings included in this dataset catalog the largest EF rating reached for the entire path.

Observed losses were collected from the NOAA Storm Events Database (https://www.ncdc.noaa.gov/stormevents/) and the National Centers for Environmental Information (NCEI). These databases contain counts of observed losses including injuries, deaths, and property damages in dollars at the county level. Only the losses that occurred during the study period (2010-2014) and classified as “Tornado” under the “Event Type” category (losses related to a tornado event) were included in this work. It is possible that other documented losses were related to these tornado events, such as those categorized as “High Wind” or “Flood,” but were not included.

A total of 449 events were reported for the study period. Loss information was matched to event path information giving 381 tornadoes with losses. Further data processing eliminated tornado paths with a shapelength of 0. This included the removal of 102 tornado events; these were a majority of EF0 (92 records) and EF1 (10 records) ratings. These events were likely touchdown points or events without tornado path information. Finally, 3 more tornado events were removed and will be discussed further in the Methods section. For the 2010-2014 study period, there were 276 usable tornado paths.

3.3 Tornado Frequency

Tornado paths were used to identify the tracts and associated losses related to tornado events (**Objective 2**). For the 2010-2014 study period there were 276 tornado events after removing the necessary records as discussed in the Data and Methods sections. The number of events, EF ratings, total injuries, deaths, and property damages for the study period can be found in Table 3. The largest number of injuries, deaths, and property damages were caused by tornadoes with EF5 ratings. There were 0 recorded injuries and deaths for EF0 events and $0 recorded property damages for EF4 events. Tornadoes with an EF4 rating are very strong and often cause damages to property. Despite there being $0 recorded for property damages for the study period from EF4 events, it is possible there is information missing from the dataset. Spatially, tornadoes with EF0-2 ratings were relatively dispersed throughout the state. Tornadoes with EF ratings of 4 and 5 were primarily found in central Oklahoma. Those with an EF3 rating were also concentrated in the central part of the state, in addition to northwest and northeast Oklahoma. A map of the tornado events included in the study period can be found in Figure 7.

Table 3: Tornado Events and Associated Losses From 2010-2014

	NUMBER OF EVENTS	TOTAL INJURIES	TOTAL DEATHS	TOTAL PROPERTY DAMAGES
EF0	99	0	0	$425,500
EF1	124	42	0	$4,720,000
EF2	34	88	2	$16,730,000
EF3	12	135	17	$1,375,000
EF4	5	200	6	$0
EF5	2	393	33	$2,000,000,000

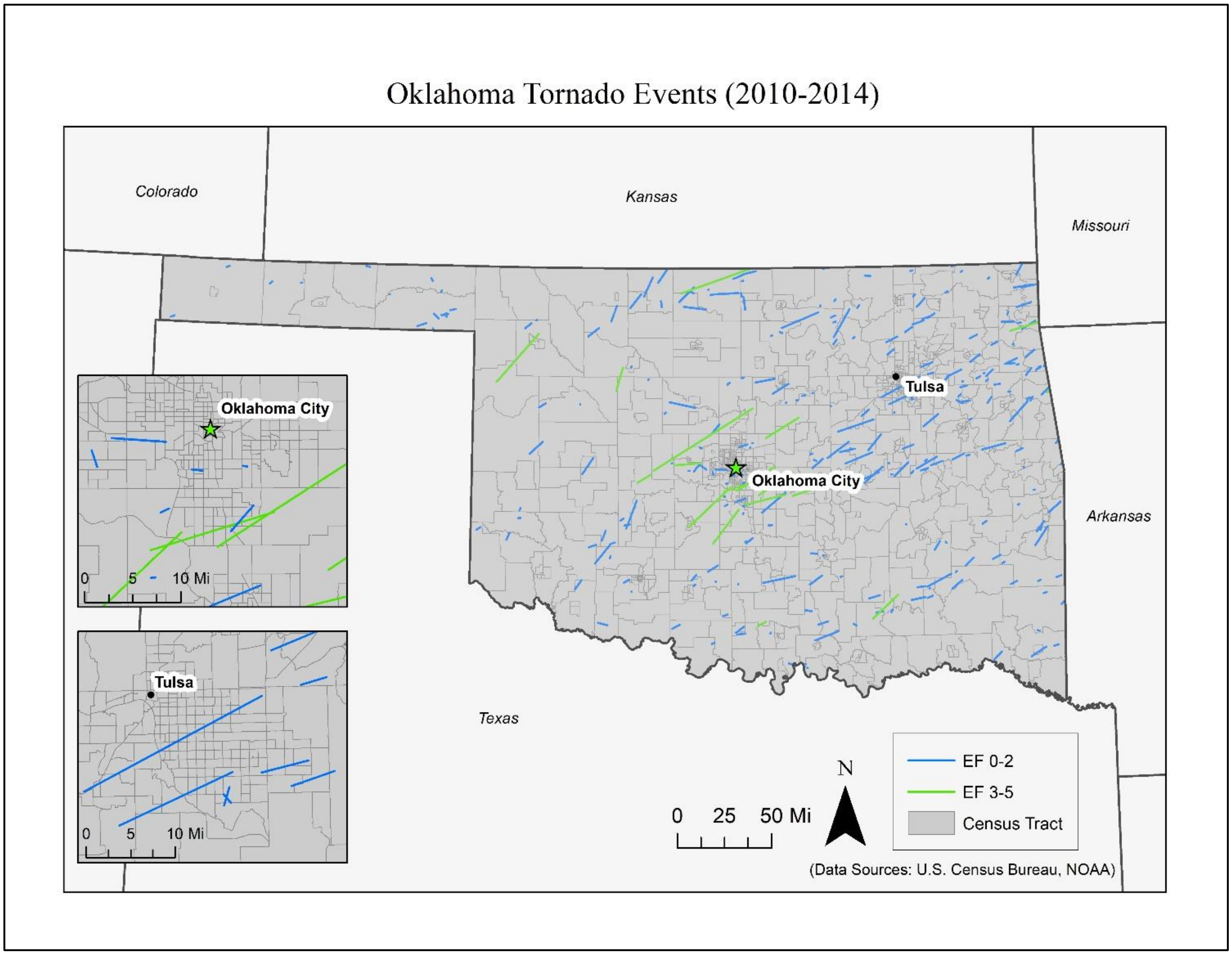


Figure 7: Oklahoma Tornado Events From 2010-2014

To understand how the study period compared to other years of tornado activity in Oklahoma, Table 4 provides a breakdown of the number events over 5-year periods for almost 25 years from NOAA's Storm Events Database (https://www.ncdc.noaa.gov/stormevents/). The number of tornado events for the study period, broken down by EF ratings, are relatively similar to the other years in terms of proportioned frequency. One notable difference for the study period is that there are less EF0 than EF1 events, whereas the opposite is true for the 5-year periods that had a greater number of EF0 than EF1. However, recent improvements and changes in technology and radar have increased the quantity and quality for detecting tornado events. These technologies include but are not limited to social media, storm chasers, and the Oklahoma Mesonet (http://www.mesonet.org/). The Oklahoma Mesonet, founded in 1994, hosts a collection of 120 environmental monitoring stations throughout the 77 counties in Oklahoma. It is important to note how changes in detection and environmental monitoring influences years of comparison.

Table 4: Tornado Activity in Oklahoma for 25 Years

5-YEAR PERIOD	TOTAL EVENTS	F0/EF0	F1/EF1	F2/EF2	F3/EF3	F4/EF4	F5/EF5
1995-1999	484	304	114	45	15	4	2
2000-2004	292	173	87	23	8	1	0
2005-2009	238	140	67	25	4	2	0
2010-2014 (study period)	449	219	164	36	20	8	2
2010-2014 (included in this work)	276	99	124	34	12	5	2
2015-April 2019	351	171	141	31	7	1	0

CHAPTER IV

METHODS

This work builds on Cutter et al. (2003) Social Vulnerability Index (SoVI) and the works of Chaney and Weaver (2010), Dixon and Moore (2012), Widen (2016), Lim et al. (2017), and others, to construct a tornado-specific social vulnerability index (TSVI) for Oklahoma. The SoVI and TSVI are alternative means of identifying socially vulnerable groups in Oklahoma. Though it is understood that different social groups may be vulnerable at different stages of a hazardous event (Rufat et al. 2015), these indexes seek to measure general social vulnerability and tornado-related social vulnerability (**Objective 1**). The purpose of this study is to externally validate these indexes against observed tornado-related losses in Oklahoma (**Objective 2**).

4.1 Social Vulnerability Index Construction

To examine the distribution of social vulnerability throughout Oklahoma (**Objective 1**) the commonly applied SoVI was built at the county and tract scale following the inductive procedure of Cutter et al. (2003) and Tate (2012) using ACS variables. The first step to building this index was to standardize each variable and eliminate outliers' values. Directionality adjustments were performed on four variables including QRICH200K, PERCAP, MHSEVAL, and MDRENT. As discussed by Tate (2011), these directionality changes were needed because for variables representing income and wealth, high values are associated with low social vulnerability.

To reduce data and gather the variable weights, principal components analysis (PCA) or factor analysis was implemented. Factor analysis groups correlating variables into groups known as "factors" (Rogerson 2015). Factor analysis was conducted with a varimax rotation in accordance with Cutter et al. (2003). A varimax rotation, developed by Kaiser, is the most common rotation method and maximizes the variance of the factor loadings. This helps simplify the interpretation of factors to identify the strongest loading variables (Abdi 2003). The factors, representing the main drivers of social vulnerability in Oklahoma, provided the different weights of the variables. These weights demonstrated how much the variables were contributing to each factor. The factors were chosen using the Kaiser criterion. All factors with an eigenvalue greater than 1 were included. With this method, 7 factors had an eigenvalue greater than one.

After obtaining the factors, the z-score of each variable was multiplied by the contributing weight of the 7 factors for all counties and tracts. The social vulnerability scores were created by adding all 7 factors together. Using the social vulnerability scores, the counties and tracts were ranked from most vulnerable (smallest negative value) to least vulnerable (greatest positive value). These ranks were used to assign percentiles so the SoVIs could be mapped using quintiles, or five equal classes ranging from low social vulnerability to high social vulnerability. These steps are outlined in Figure 8.

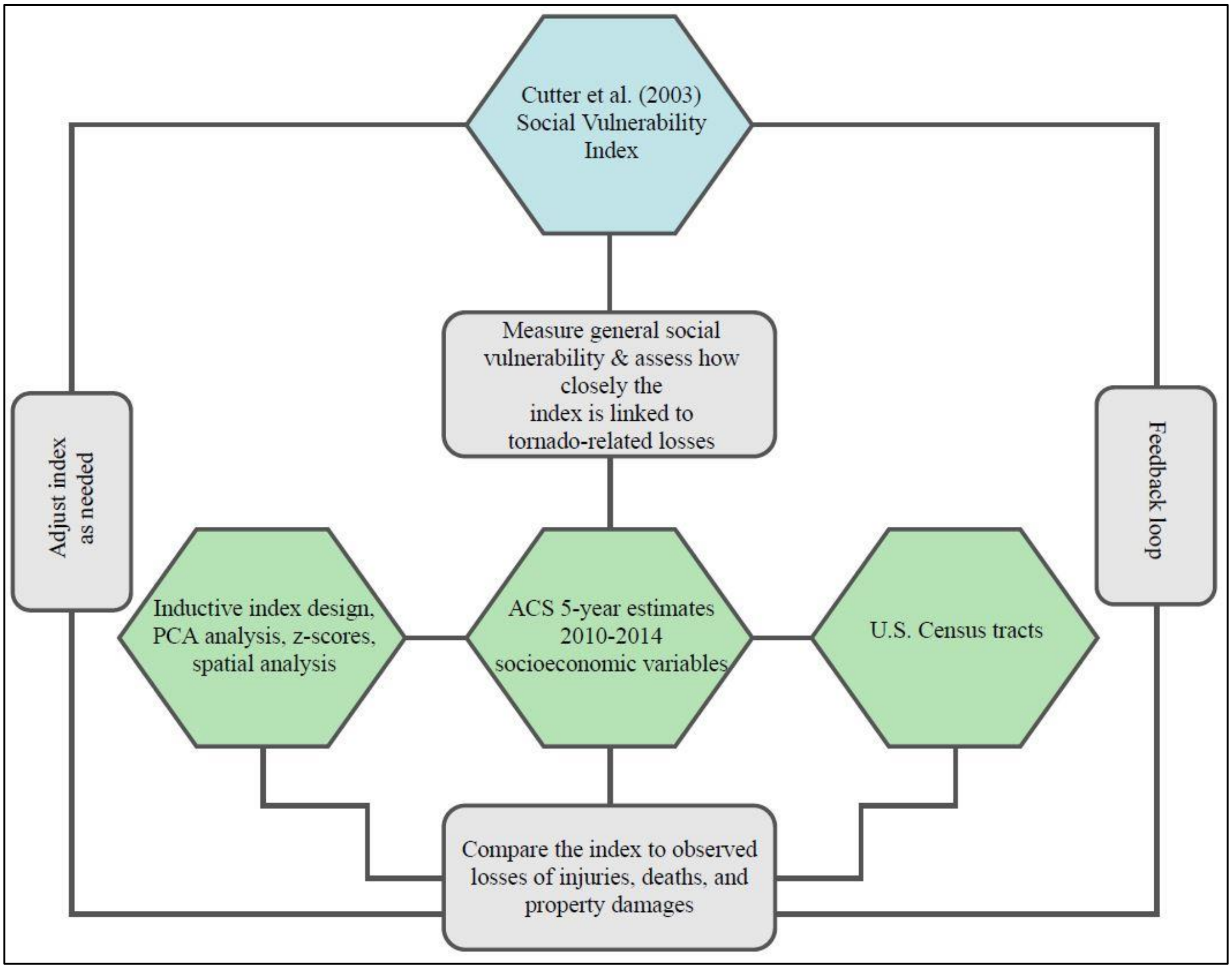


Figure 8: Methodology Flowchart (Adapted from Cutter et al. (2003) and Tate (2012))

4.2 Tornado Social Vulnerability Index Construction

The tornado social vulnerability index, TSVI, was constructed as an alternative measure of tornado-specific social vulnerability and used to identify areas of Oklahoma that were more socially vulnerable to tornadoes (**Objective 1**). Unlike the SoVI, the TSVI was constructed following a deductive design using ACS variables. A deductive design does not incorporate factor analysis since there are not as many variables included in the index. The variables that served as indicators were based on existing literature (see Blaikie et al. (1994), Brooks and

Doswell (2002), Ashley (2007), Mulilis et al. (2000), Chaney and Weaver (2010), Dixon and Moore (2012), Chaney et al. (2013), Widen (2016), and Lim et al. (2017). Like the construction of the SoVI, z-scores were used to standardize the data and eliminate any influence of larger values and extreme outliers. Directionality adjustment was performed on one variable, QRICH200K, since high values are associated with low social vulnerability (Tate 2011). The social vulnerability scores were created by adding together the z-scores for all 10 variables. Using the social vulnerability scores, the tracts were ranked from most vulnerable (smallest negative value) to least vulnerable (greatest positive value). These ranks were used to assign percentiles so the TSVI could be mapped using quintiles, or five equal classes ranging from low social vulnerability to high social vulnerability.

4.3 Comparison of Indexes

4.3.1 Rank Order Comparison

After the tract SoVI and TSVI were created, Spearman's rank correlation coefficient was used to compare the indexes at the tract scale (**Objective 1**). This comparison included the tract SoVI and TSVI. The county SoVI could not be included because it was at a different spatial scale. Spearman's rank correlation coefficient is a nonparametric test that measures the strength of a relationship and direction of association of 2 ranked variables. A nonparametric test does not require any assumptions about the data to be held other than the minimum requirements to use the statistic. Relationships are represented with Spearman's *r*. An *r* value of 0 suggests there is no relationship and an *r* value of +/- 1 indicates a perfectly correlated relationship (Rogerson 2015).

4.3.2 Spatial Analysis

Further comparison of the SoVIs and TSVI incorporated spatial analysis (**Objective 1**). The spatial distribution of social vulnerability scores from all indexes were analyzed using Hot Spot Analysis (Getis-Ord Gi*) and Cluster and Outlier Analysis (Anselin Local Moran's I). Spatial analysis was conducted in ArcMap 10.6.1 to detect statistically significant areas of social vulnerability scores. Based on the Getis-Ord Gi* statistic, Hot Spot Analysis identifies spatial clusters of high values (hot spots) and low values (cold spots). Alternatively, based on the Anselin Local Moran's I statistic, Cluster and Outlier Analysis identifies statistically significant hot spots, cold spots, and outliers. Outliers identify areas where high values are near low values and vice versa, something that Hot Spot Analysis does not highlight. With Getis-Ord Gi*, the value of each feature is included in determining hot and cold spots, whereas Anselin Local Moran's I only include the neighboring features to produce outputs (Rogerson 2015). Both analyses were performed to observe if the minor differences in statistical foundations would result in different outputs. Additionally, Cluster and Outlier Analysis includes an extra measure compared to Hot Spot Analysis and provides more details of statistically significant clusters.

The Hot Spot Analysis and Cluster and Outlier Analysis tools used a weights matrix file to determine results. The weights matrix files were made for each index using the Incremental Spatial Autocorrelation tool to measure the spatial autocorrelation for the county SoVI, tract SoVI, and TSVI. Spatial autocorrelation describes how objects in space are or are not like nearby objects (Rogerson 2015). The Incremental Spatial Autocorrelation tool provides a line graph output that describes the strength of relationships at different distances (in feet). The distance with the strongest relationship, the greatest instance of spatial autocorrelation, was used to generate a weights matrix for each index using the Generate Spatial Weights Matrix tool. The statistically significant areas were used to visually compare the indexes for any major similarities and/or differences.

4.4 Observed Losses

To assess how closely the SoVI and TSVI are linked to tornado-related losses, the spatial distribution of social vulnerability was compared to observed losses from tornadoes (**Objective 2**). From the observed losses dataset, only those that matched the date, time, and locations of the tornado events from the SPC database were included in this work. This allowed for the observed losses to be linked to individual tornado events. However, because losses were recorded at the county scale, they needed to be allocated to the Census tract scale for finer spatial comparison. Census tracts were the more desirable scale of comparison since tornadoes are relatively short, localized events. Other works, such as Shen and Hwang (2015), have used loss data for comparison and validation but only analyzed relationships at the county and state scale.

When a tornado event occurred within a single tract regardless of the method, the losses were assigned to that location. When a tornado spanned multiple tracts, losses were distributed using 3 different methods: (i) averages, (ii) impervious surface area, and (iii) randomization. Three distribution methods were chosen to help identify and account for sensitivities to allocation method.

(i) Averages: When tornado paths overlapped with multiple tracts within a county the losses (e.g., injuries) were divided among those tracts. If the number of losses did not divide evenly with the number of engaged tracts, greater values (e.g., 33, 33, 34) were placed in the tracts with the greatest proportion of tornado path. This approach assumed the losses were evenly distributed throughout an event and tracts with a greater proportion of tornado path had a higher probability of experiencing a tornado-related loss.

(ii) Impervious Surface Area: Impervious surface areas are areas of highly developed land where people reside or work and often include artificial structures. The impervious surface data was gathered from the National Land Cover Database (NLCD) 2011 Land Cover published by the Multi-Resolution Land Cover Characteristics (MRLC) Consortium (https://www.mrlc.gov/). Buffers were created around the tornado paths based on the width reported and the losses were assigned to tracts with the greatest percent of impervious surface area. This approach assumes the injuries, deaths, and/or property damages were more likely to have occurred in areas of dense human activity and development.

(iii) Randomization: When tornado paths overlapped with multiple tracts within a county the losses were distributed to the tracts randomly. This method was used as a baseline to compare with the other distribution methods. Random numbers for the tract placement and amount of losses assigned were generated in Microsoft Office Excel 2016 using the RANDBETWEEN function. When given a range of numbers (e.g., 1 to 6 tracts and $0 to $5,000 damages), this functions output provides a random integer between the values. To randomly assign losses to the tracts, first, the random tract number was generated, followed by the amount of losses to delegate to that tract. Random numbers were generated until the number of losses were reached.

Three tornadoes from the SPC database did not match any date or time stamps of the losses in the Storm Events Database and were removed from analysis. Two of the tornadoes had an EF rating of 0 (4/14/2011, 4/13/2012) and the third had an EF rating of 1 (5/13/2010). After removing these events, a total of 276 tornadoes were included in analyses for the study period.

4.5 External Validation

To assess if the SoVIs and TSVI demonstrated expected relationships with tornado-related losses, the indices were compared to observed loss data (**Objective 2**). Since the injuries, deaths, and property damages were collected and distributed by county, the county data were used as a static reference. Spearman's rank correlation coefficient was used to examine how the different methods of loss distribution were related to the social vulnerability rankings. The Spearman's rank correlation coefficient analysis measured how strongly the SoVIs and TSVI social vulnerability scores related to the different methods of loss distribution of averages, impervious surface area, and randomization. Spearman's rank correlation coefficient is a nonparametric test that measures the strength of a relationship and direction of association of 2 variables. A nonparametric test does not require any assumptions about the data to be held other than the minimum requirements to use the statistic. Relationships are represented with Spearman's *r*. An *r* value of 0 suggests there is no relationship and an *r* value of +/- 1 indicates a perfectly correlated relationship (Rogerson 2015). Spearman's *r* is used as an alternative to Pearson's *r* because the data did not meet the assumptions to use that measure (the data were not normally distributed).

To further examine the relationship between the tract SoVI, TSVI, and observed losses, select tornado events were discussed. The 4 case studies presented examples including how tornadoes of the same EF rating caused different losses across space, how losses varied across different areas of impervious surface areas, and how well the indexes described the tracts that experienced the most damages during the study period.

CHAPTER V

RESULTS

5.1 Indexes

5.1.1 Factor Analysis

Factor analyses were performed at the county and tract levels to create the general social vulnerability index, SoVI. The number of factors were chosen using the Kaiser criterion. This method retains all factors with an eigenvalue greater than 1. Scree plots display eigenvalues and are used to identify the number of factors to be used. As shown in Figures 9 and 10, the county and tract analyses produced 7 factors with eigenvalues greater than 1. Indexes were constructed using all 7 factors explaining 75.5% (county) and 69.9% (tract) of the variance. The strongest contributing factors are those with the greatest number of variances explained. Factor 1 contributed the most to the total variance explained for the county and tract SoVI. For the county SoVI, factor 1 contributed 28.04% of the variance and for the tract SoVI, factor 1 contributed 26.21%. The interpretations of the factors and the percent of variance explained can be found in Table 5.

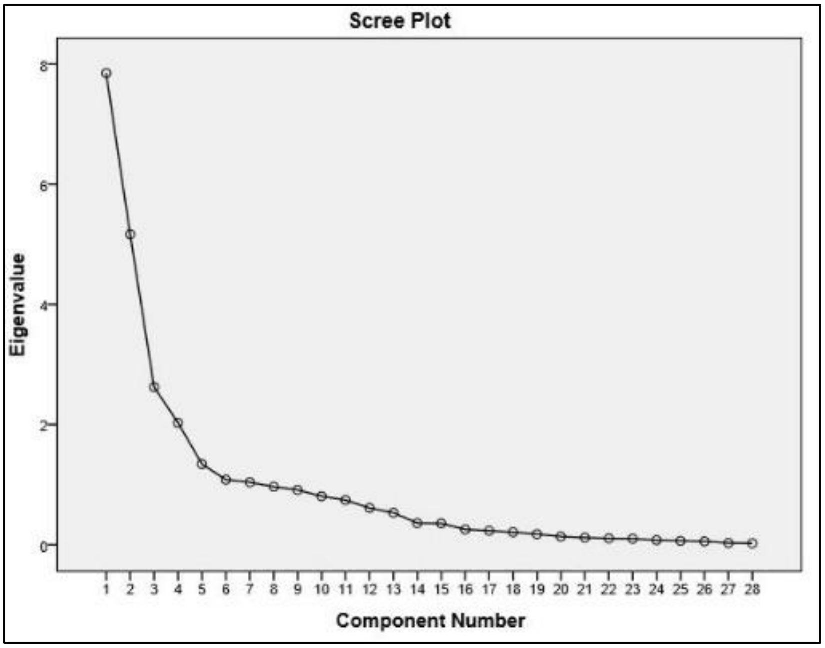


Figure 9: County SoVI Scree Plot

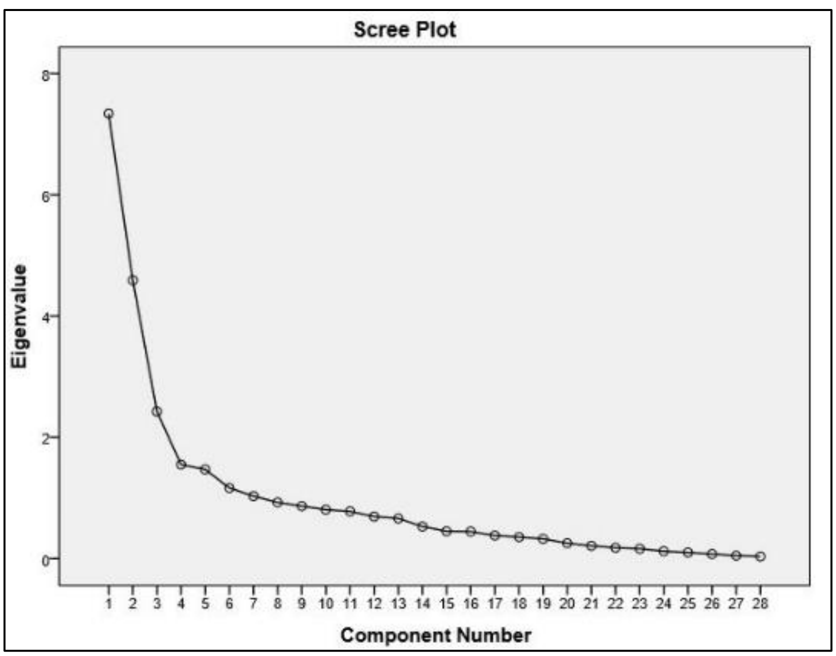


Figure 10: Tract SoVI Scree Plot

Table 5: Factor Interpretations and Variance Explained

	COUNTY INTERPRETATIONS	COUNTY VARIANCE EXPLAINED (CUMULATIVE)	TRACT INTERPRETATIONS	TRACT VARIANCE EXPLAINED (CUMULATIVE)
Factor 1	Income, Home Value	28.038	Income, Home Value	26.208
Factor 2	Age, Elderly	18.459 (46.497)	Elderly, Unemployment	16.391 (42.599)
Factor 3	Hispanic, Non-English Speaking	9.375 (55.872)	Hispanic, Non-English Speaking	8.662 (51.261)
Factor 4	Black, Race	7.250 (63.122)	Race, Gender	5.538 (56.799)
Factor 5	Gender, Family	4.794 (67.916)	House Type, Status (renter)	5.251 (62.050)
Factor 6	Assisted Living, Elderly	3.867 (71.783)	Assisted Living, Elderly	4.164 (66.213)
Factor 7	Mobile Homes	3.713 (75.496)	Native American, Extractive Employment	3.689 (69.896)

The variables for factor 1 were related to income and home value (QRICH200K, QPOVTY, PERCAP, MHSEVAL, and QNOHLTH) and all had loadings above .8 for the county and tract levels. For factor 2, age and elderly variables loaded strongly for the county SoVI, including QSSBEN and MEDAGE. For the tract SoVI, in addition to the age-related variables, QCVLUN, describing civilian unemployment, had a strong loading above .8. For both SoVIs, the variables that best described factor 3 were QHISP and QESL and all had values greater than .9. For factor 4, variables relating to race and gender (QBLACK and QFHH) loaded highly and described this factor. Factor 5 had different loadings for the county and tract. At the county level, this factor was strongly driven by the QFAM and QFEMALE variables, whereas the factor for the tract level had high loadings for QRENTER and QPPUNIT. For the county, factor 5 was driven by gender and family variables and factor 5 for the tract was described relating to house type and status (renter). The variable with the strongest loading for factor 6 was QNRES. For the county and tract, this factor was described as loading high for elderly individuals and assisted living residents (related to elderly). Finally, for factor 7, the strongest contributing variables were different for the county and tract. For the county level QMOHO described the factor best and for the tract, the strongest loading variables were QNATAM and QEXTRCT.

5.1.2 County SoVI

The results of the SoVIs and TSVI reveal the spatial distribution of social vulnerability throughout the state of Oklahoma (**Objective 1**). The county SoVI found the most socially vulnerable groups in southeastern and southwestern Oklahoma. This index also classified central and northeastern Oklahoma as having medium to low social vulnerability as shown in Figure 11.

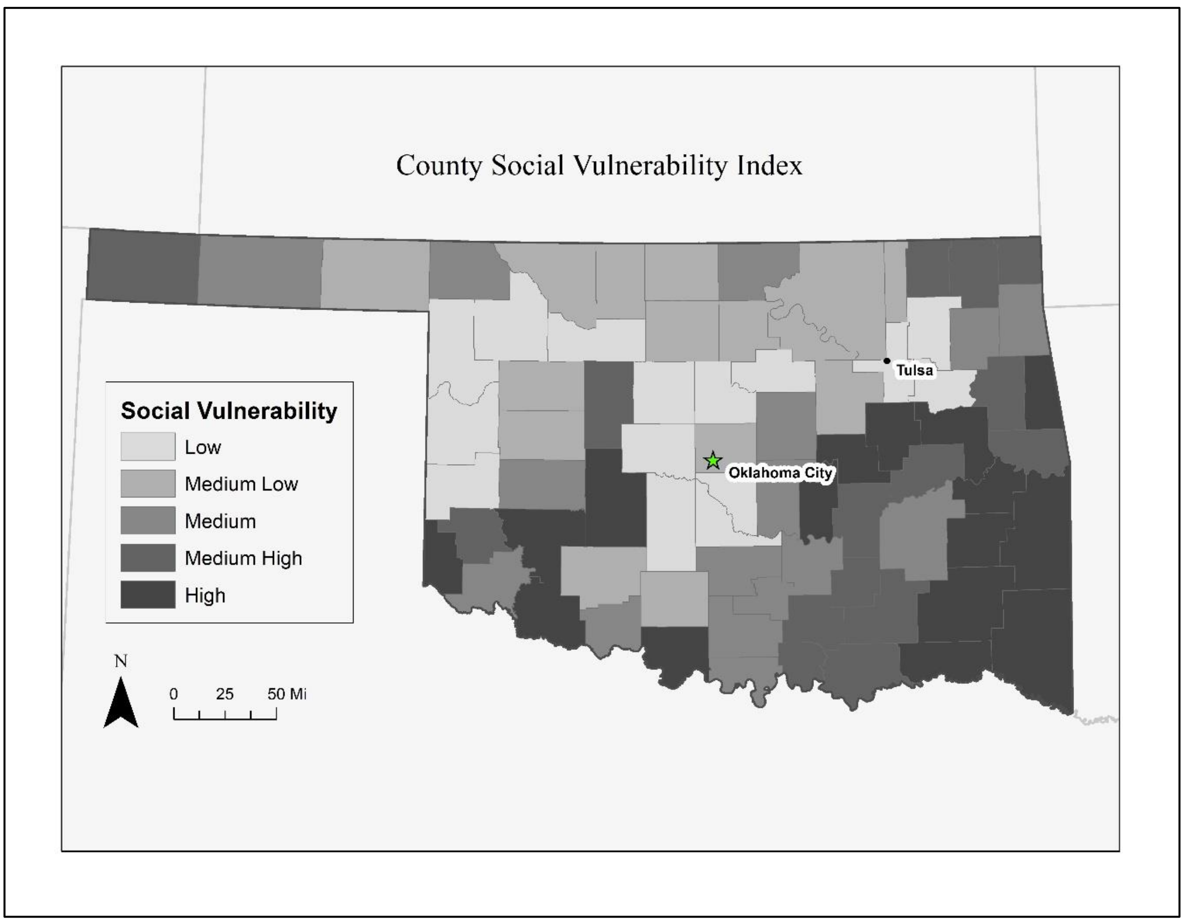


Figure 11: County SoVI 2010-2014

5.1.3 Tract SoVI

The distribution of social vulnerability scores for the Oklahoma tract SoVI can be found in Figure 12. The tract SoVI suggested the most socially vulnerable groups were in the panhandle, central, and northeastern Oklahoma. The Oklahoma City and Tulsa metropolitan areas showed great variability in the assigned social vulnerability scores.

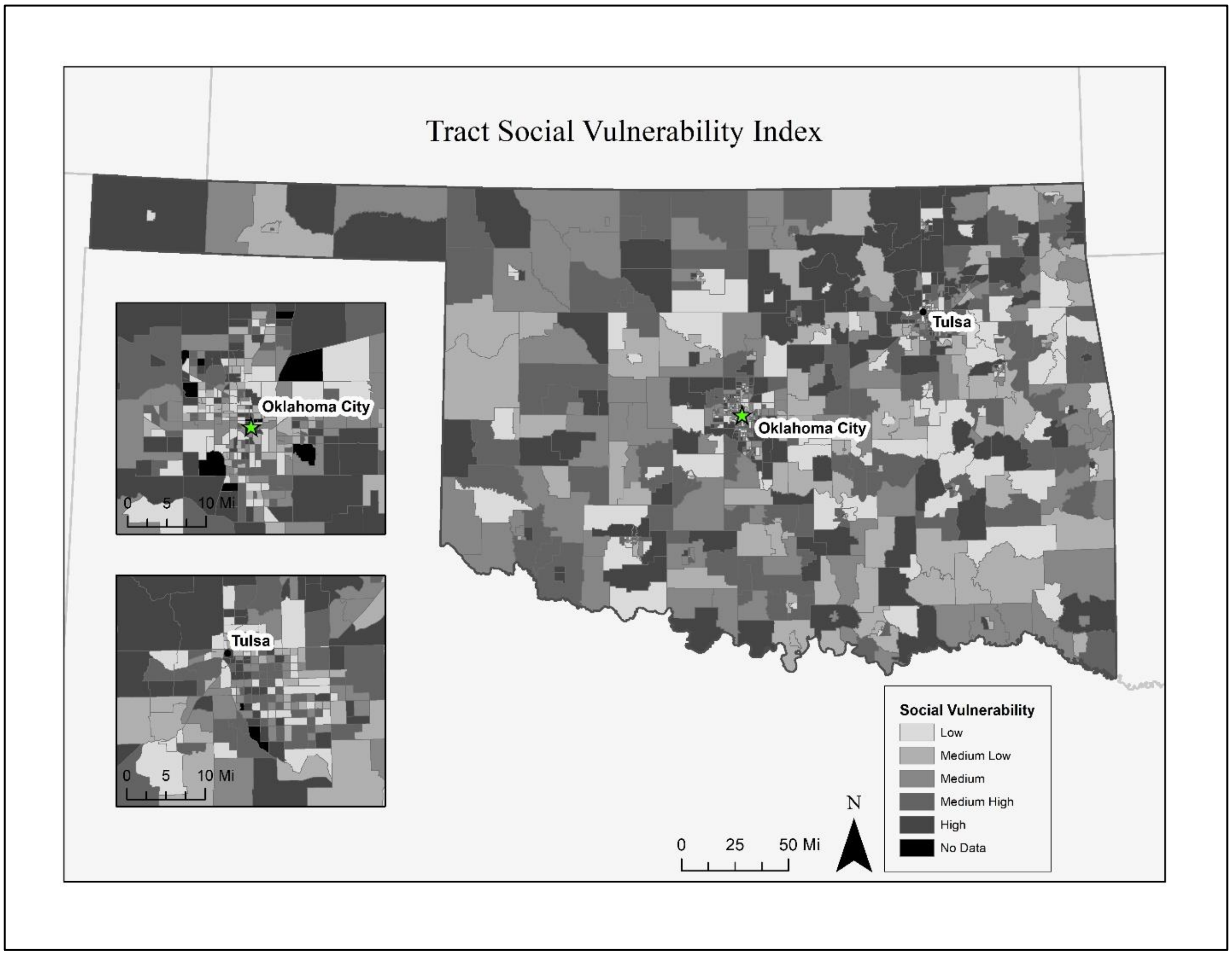


Figure 12: Tract SoVI 2010-2014

5.1.4 TSVI

The TSVI, found in Figure 13, classified parts of the panhandle, southern, and western Oklahoma as the most socially vulnerable. Northwestern Oklahoma was described with low to medium low social vulnerability. The Oklahoma City and Tulsa metropolitan areas showed variability in the assigned social vulnerability scores with distinct concentrations of areas of high to low social vulnerability.

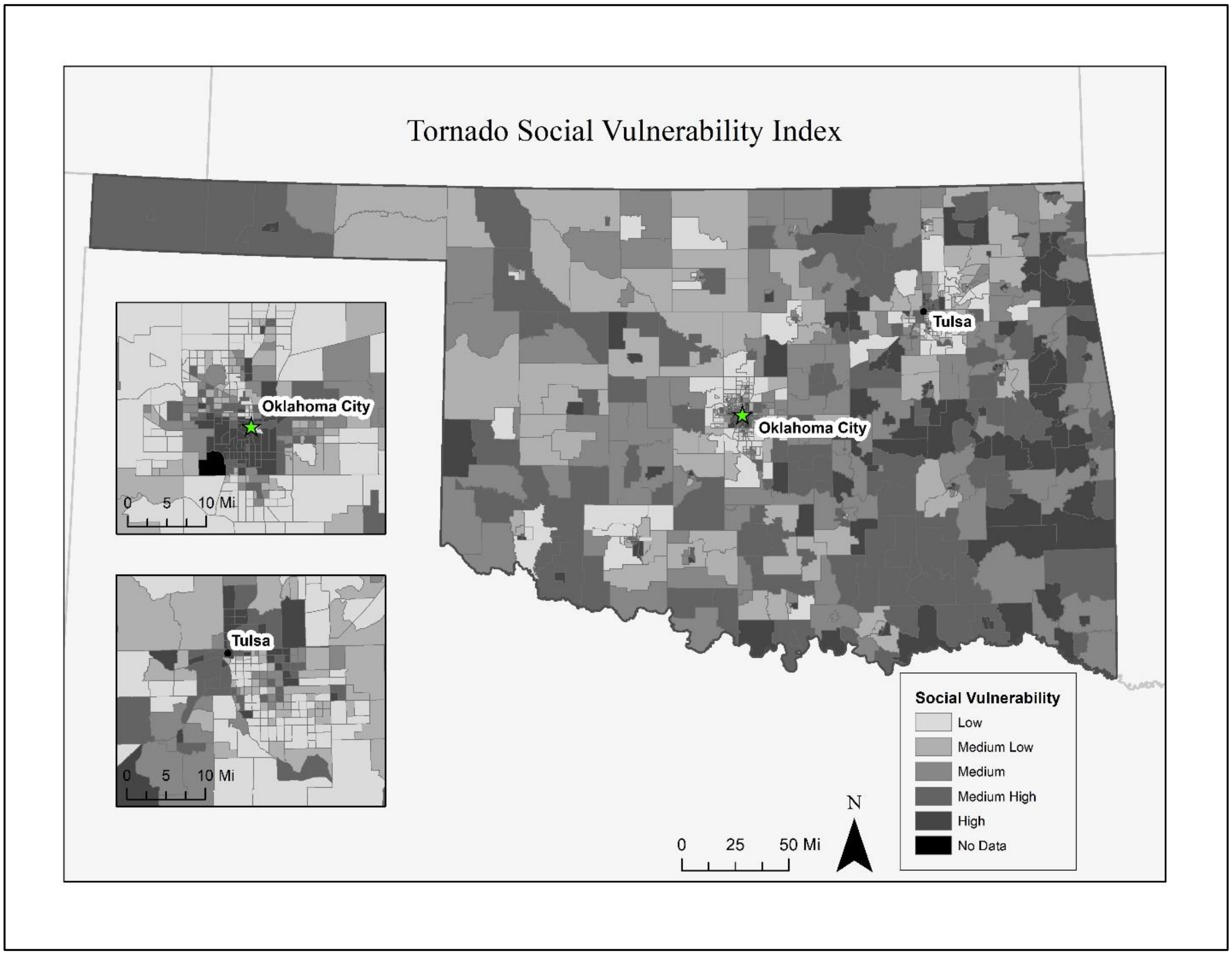


Figure 13: TSVI 2010-2014

5.2 Comparison of Indexes

5.2.1 Rank Order Comparison

The tract SoVI and TSVI were compared to identify any major similarities or differences between the indexes (**Objective 1**). The results of the Spearman's rank correlation coefficient produced an *r* value of -.293. This suggests a weak relationship between the two indexes and demonstrates that the tract SoVI and TSVI did not rank the social vulnerability of Census tracts similarly.

5.2.2 Spatial Analysis

Hot Spot Analysis (Getis-Ord Gi*) and Cluster and Outlier Analysis (Anselin Local Moran's I) were used to identify statically significant areas of the SoVIs and TSVI. The SoVIs and TSVI were then compared visually to identify any major similarities or differences between the indexes (**Objective 1**).

5.2.2.1 County SoVI

Statistically significant areas of the county SoVI were found in the northwestern and southeastern regions of Oklahoma using the Hot Spot Analysis tool. The hot and cold spots generated can be found in Figure 14. The northwestern part of the state was designated as a cold spot because this area and the surrounding region had lower than average values of low social vulnerability. This part of Oklahoma has more open space and fewer people. Southeastern Oklahoma was recognized as a hot spot due to the higher than average concentration of high values or high social vulnerability scores. This region has less tornadoes and a more socially vulnerable population. The outputs of the Cluster and Outlier Analysis tool also identified northeastern Oklahoma with significant low-low clustering (low values near other low values) and southeastern Oklahoma as an area of high-high clustering. An instance of high-low outliers was detected in northeastern Oklahoma, indicating that high social vulnerability scores were closely related in space to an area of low social vulnerability. Similar relationships were found in southeastern Oklahoma. These clusters can be observed in Figure 15.

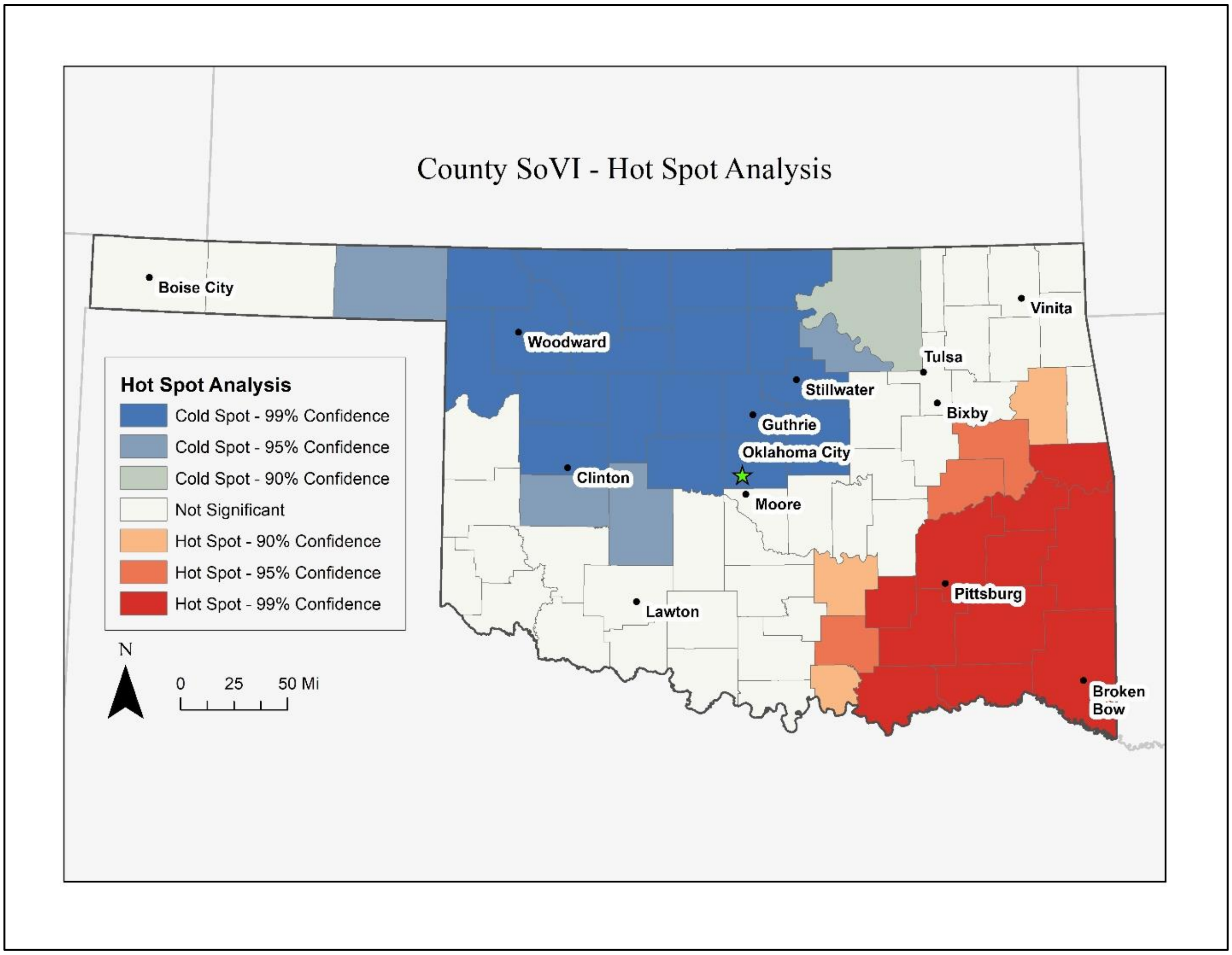


Figure 14: County SoVI - Hot Spot Analysis

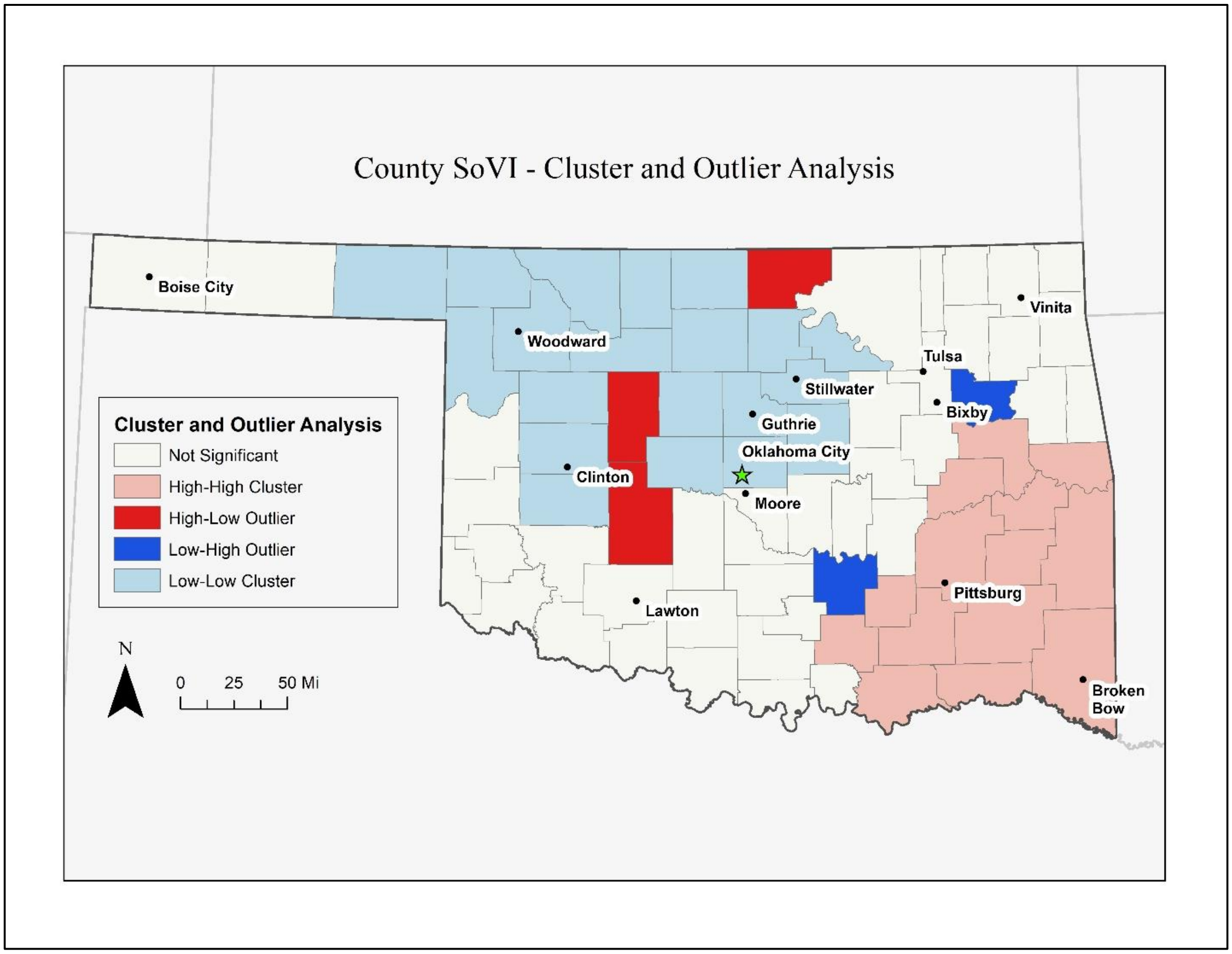


Figure 15: County SoVI - Cluster and Outlier Analysis

5.2.2.2 Tract SoVI

Statistically significant areas of the tract SoVI were found in the panhandle, north-central, and east-central regions of Oklahoma. The hot and cold spots generated from the Hot Spot Analysis tool can be found in Figure 16. The panhandle and north-central part of the state were designated with areas of hot spots because this area had higher than average values of high social vulnerability for tract SoVI. East-central Oklahoma was recognized as a cold spot due to the higher than average concentrations of low values (low social vulnerability scores).

The Cluster and Outlier Analysis tool also identified similar areas of Oklahoma of notable clustering. Low-high outliers were found in north-central and northeastern Oklahoma, whereas high-low outliers were detected in east-central Oklahoma. These clusters suggested high social vulnerability scores were closely related in space to low social vulnerability scores. High-high clustering was found in northeastern Oklahoma and low-low clustering was found in the east-central region. These relationships can be observed in Figure 17.

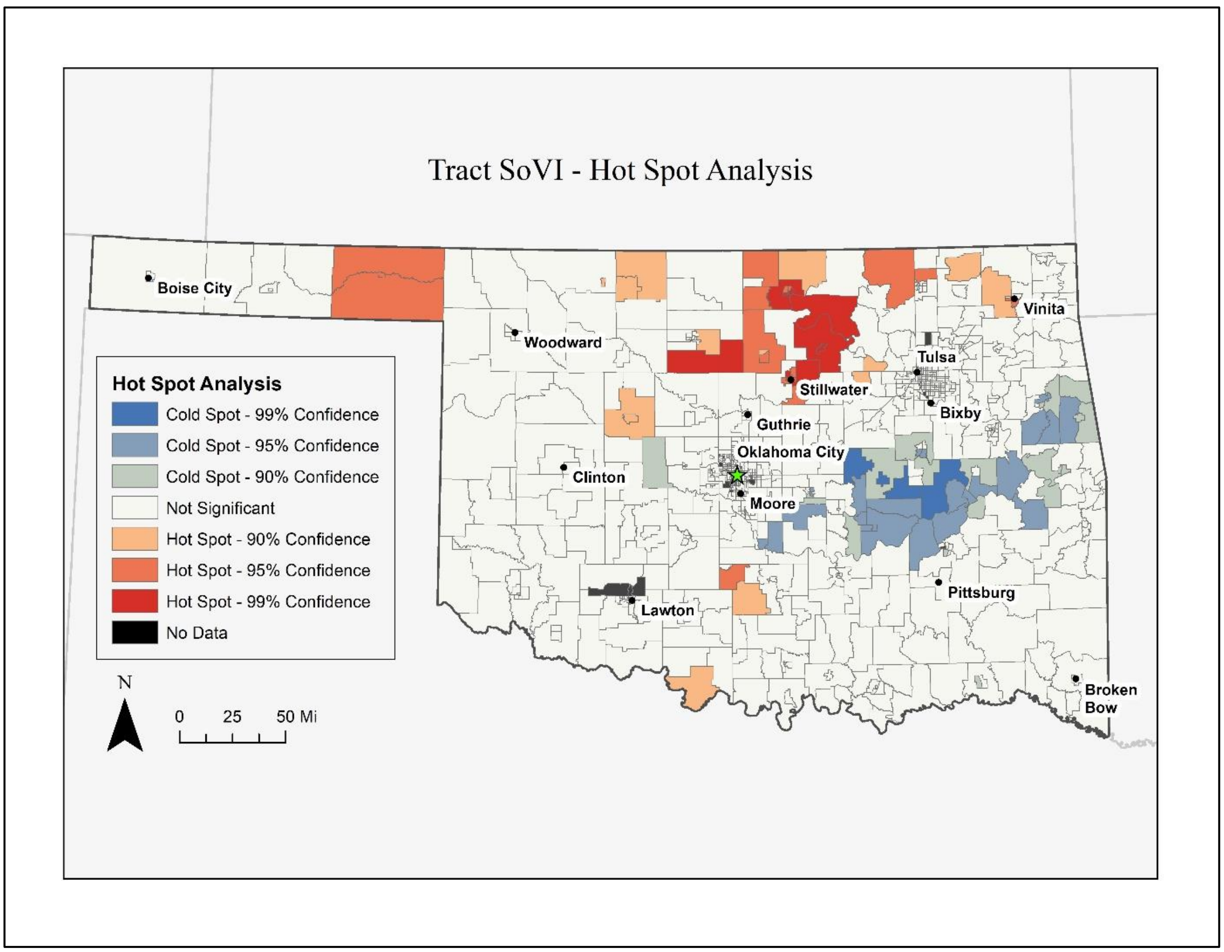


Figure 16: Tract SoVI - Hot Spot Analysis

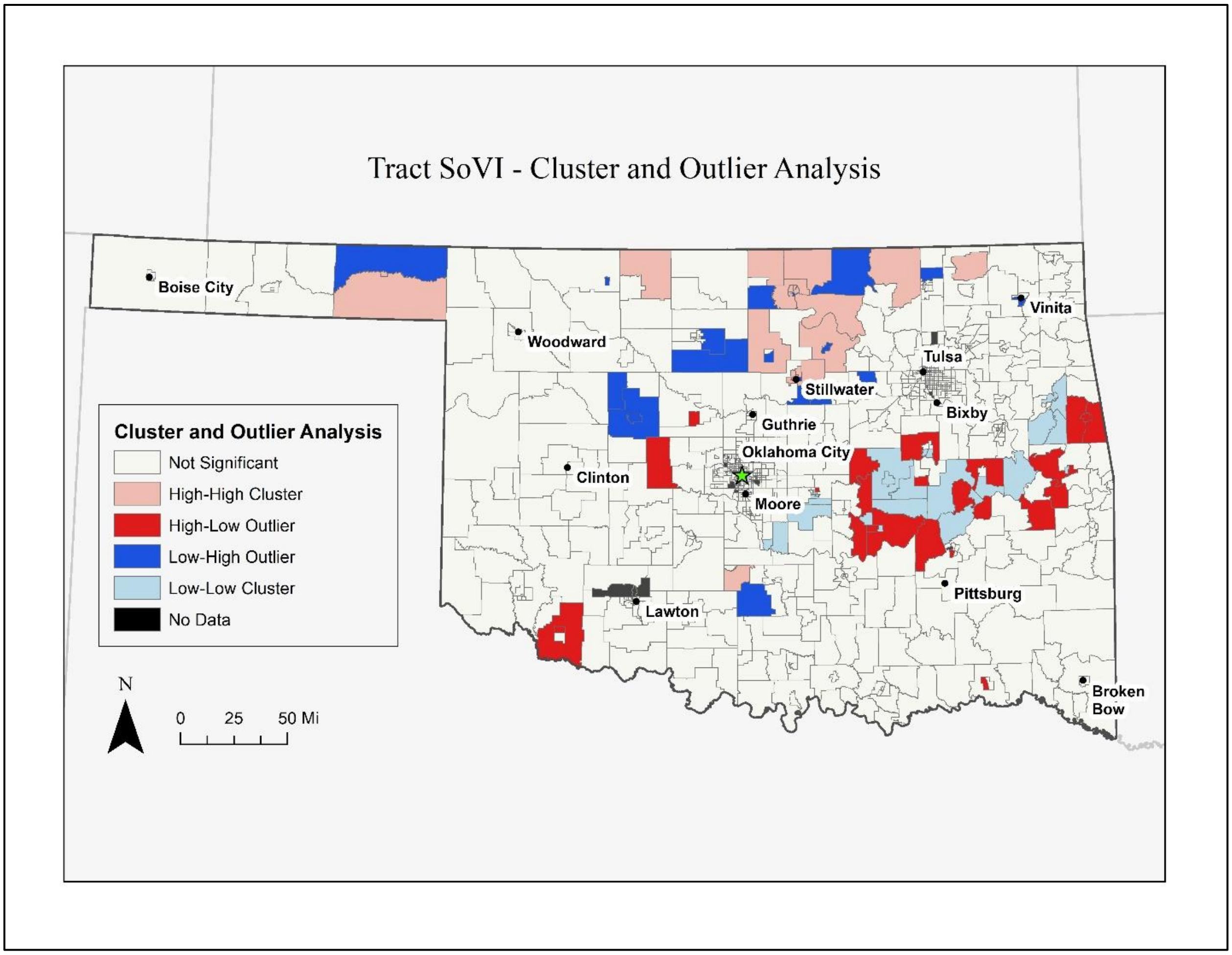


Figure 17: Tract SoVI - Cluster and Outlier Analysis

5.2.2.3 TSVI

The Hot Spot Analysis tool found significant hot spots in southeastern Oklahoma, along the eastern border, and in parts of the panhandle. The hot spots are areas of higher than average values of high social vulnerability. The cold spots were detected around the Oklahoma City and Tulsa metropolitan regions with higher than average concentrations of low social vulnerability scores. These statistically significant areas are shown in Figure 18. The Cluster and Outlier Analysis tool also identified the same general areas of Oklahoma of notable clustering.

Southeastern Oklahoma and along the eastern border had high-high clustering, meaning the TSVI assigned high values for this region (high social vulnerability scores). This area of the state also had low-high outliers meaning that extremely low values were found near extremely high ones. Low-low clustering, concentrations of low values, were assigned to Oklahoma City, Tulsa, and the surrounding region. These regions also had clusters of high-low outliers. These relationships can be observed in Figure 19.

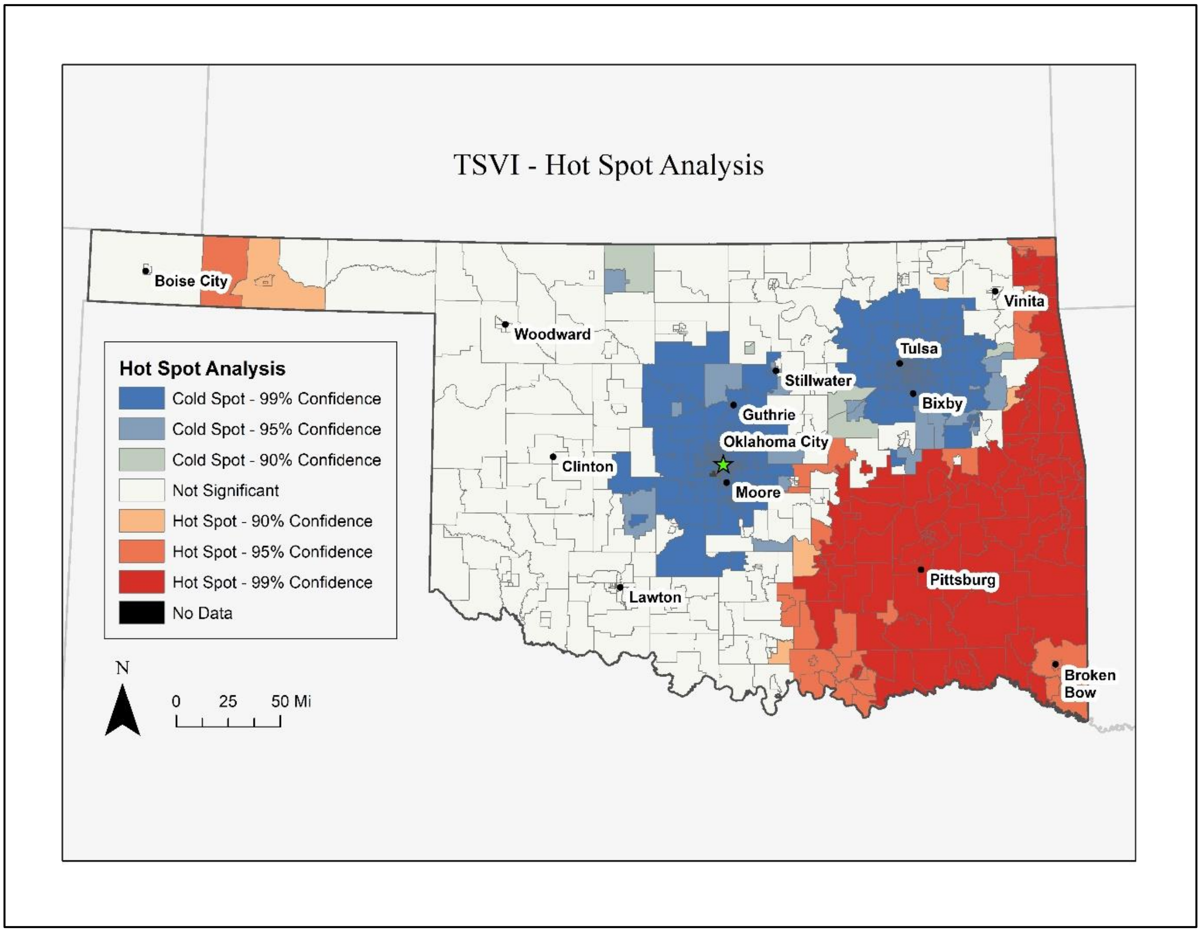


Figure 18: TSVI - Hot Spot Analysis

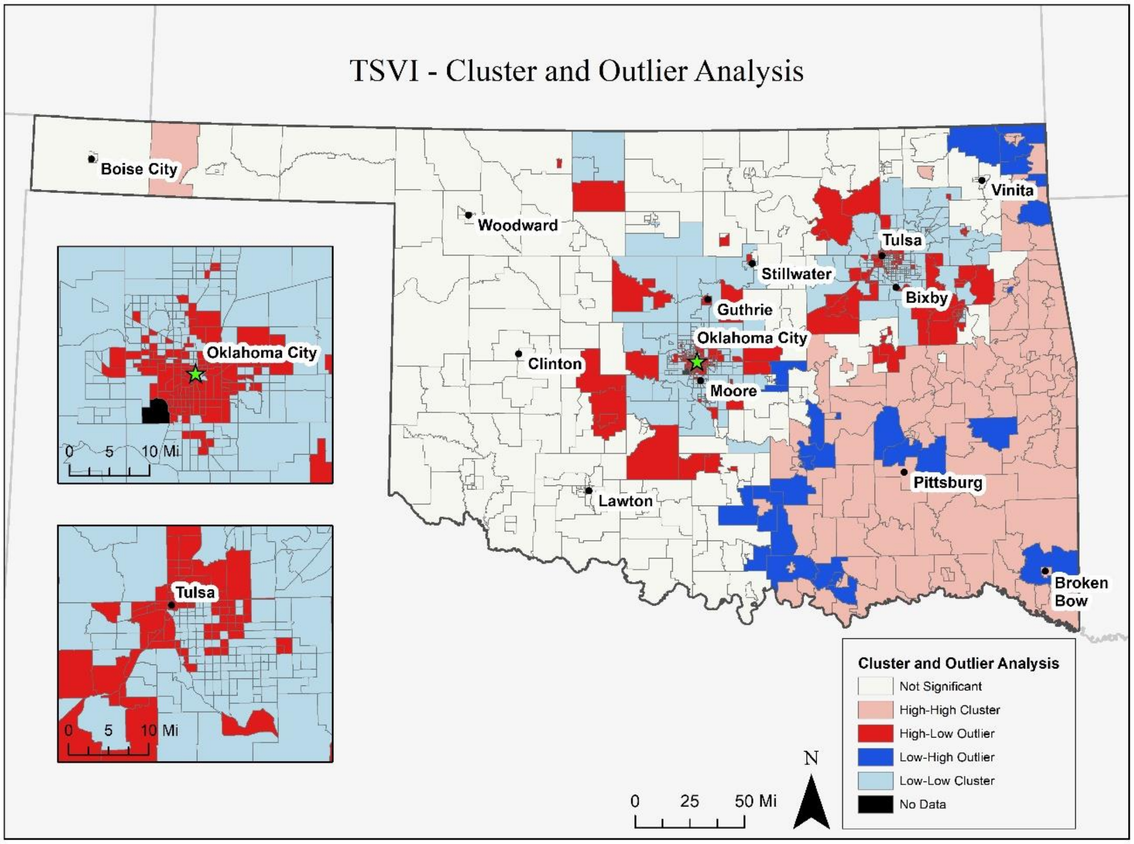


Figure 19: TSVI - Cluster and Outlier Analysis

5.2.3 Visual Comparison of the Indexes

The county SoVI identified northwestern and southeastern Oklahoma as statistically significant regions. Northwestern Oklahoma was designated a cold spot and had a higher concentration of low values (less socially vulnerable). The TSVI did not mark this a region of interest, but the tract SoVI found significant hot spots (high values, high social vulnerability). This demonstrated that the county and tract SoVIs assigned different social vulnerability scores for this part of the state.

The county SoVI described the southeastern region as a hot spot due to the large concentration of high values (more socially vulnerable). The TSVI also identified southeastern Oklahoma as a hot spot in addition to the eastern border. This is a major similarity between the two indexes. The tract SoVI highlighted some geographies in this region with high values but not to the extent of the county SoVI and TSVI. The TSVI was the only index to recognize the greater Oklahoma City and Tulsa metropolitan areas as statistically significant.

5.3 External Validation

5.3.1 County Losses

To assess if the SoVI demonstrated expected relationships with tornado-related losses the county SoVI was compared to observed losses as a static reference since this data was distributed at this scale (**Objective 2**). The results of the Spearman's rank correlation coefficient are shown in Table 6. All *r* values are negative, implying that as one value increased (number of losses) the other value decreased (social vulnerability score). The injuries and deaths had a moderately strong relationship with the county SoVI with *r* values of -.491 and -.553, and the property damages had a weak relationship with an *r* value of -.293. The *p* values represent the statistical significances of the correlations. A *p* value of .1 demonstrates there is a 90% likelihood that the results are not due to chance. Variables with a *p* value of .1 or below are desired. The *p* values of the Spearman's rank correlation coefficient are all $<.1$, suggesting they are statistically significant.

Table 6: Spearman's r Values and p Values for County Losses

	COUNTY SOVI	*p* VALUE
INJURIES	-.491	.013
DEATHS	-.553	.078
PROPERTY DAMAGES	-.293	.083

5.3.2 Tract Losses

To assess if the tract SoVI and TSVI demonstrated expected relationships with tornado-related losses, both indices were compared to observed losses (**Objective 2**). The observed tornado-related losses were allocated to the Census tracts using averages, impervious surface area, and randomization since the original data were at the county scale. To assess the relationship between the indexes and loss allocation methods Spearman's rank correlation coefficient was used; these results can be found in Table 7. All *r* values indicated weak relationships with the SoVI and TSVI. The strongest relationship with the SoVI was an *r* value of .151 with the injuries allocated with averages (AVG_INJ). This value described a weak relationship between the two. The positive loading showed that as the SoVI values increased (more socially vulnerable) the number of observed deaths increased, which is expected. The strongest relationship presented was with the TSVI and deaths allocated by averages (AVG_DTH) with an *r* value of -.313. This value suggests that as social vulnerability scores increased (more socially vulnerable) the observed deaths decreased. The only positive relationship with the TSVI was with the deaths allocated by impervious surface area (IMPV_DTH). This relationship was described with an *r* value of .178, which is very weak.

The *p* values represent the statistical significances of the correlations. A *p* value of .1 demonstrates there is a 90% likelihood that the results are not due to chance. Variables with a *p* value of .1 or below are desired. The correlations between the tract SoVI and losses are all >.1, indicating they are not statistically significant. The only significant correlations occurred between the TSVI and AVG_INJ (.004), AVG_DTH (.098), IMPV_INJ (.002), and IMPV_PROPERTYD (.004).

Table 7: Spearman's r Values and p Values for Tract Losses

	TRACT SOVI	*p* VALUE	TSVI	*p* VALUE
AVERAGES				
AVG_INJ	.151	.146	-.295	.004
AVG_DTH	.135	.484	-.313	.098
AVG_PROPERTYD	-.071	.377	-.080	.317
IMPERVIOUS SURFACE AREA				
IMPV_INJ	-.004	.966	-.308	.002
IMPV_DTH	-.139	.449	-.178	.330
IMPV_PROPERTYD	-.122	.146	-.243	.004
RANDOMIZATION				
RAND_INJ	.010	.936	-.093	.439
RAND_DTH	.022	.926	-.191	.420
RAND_PROPERTYD	-.033	.698	-.066	.435

5.3.3 Tornado Events

To further examine the relationship between the tract SoVI, TSVI, and observed losses from tornado events, select tornado events were used (**Objective 2**). The 4 case studies below present examples including how tornadoes of the same EF rating caused different losses across space, how losses varied across different areas of impervious surface areas, and how well the indexes described the tracts that experienced the most damages during the study period. See Appendix F for a county SoVI reference map.

5.3.3.1 Case 1

Two EF3 tornadoes (EF3A and EF3B) produced different losses while interacting with similar tracts of social vulnerability. Tornado EF3A was smaller in length, width, and duration than tornado EF3B but had the same level of intensity. EF3A did not produce any tornado-related losses in Carter County on 5/10/2010 and passed through 3 Census tracts. Of the area covered, EF3A encountered 4.08% of impervious surface area in Carter County. The SoVI identified these tracts with medium and medium high (2 tracts) social vulnerability ratings. The TSVI found these tracts to have medium low (2 tracts) and medium high levels of social vulnerability as shown in Figure 20.

Tornado EF3B was connected to 40 injuries and 2 deaths in Atoka County on 4/14/2011. EF3B also passed through 3 tracts and engaged with 10.28% of impervious surface area. These tracts were described as having high, low, and medium low levels of social vulnerability by the SoVI. The TSVI identified all 3 tracts with medium social vulnerability as displayed in Figure 21.

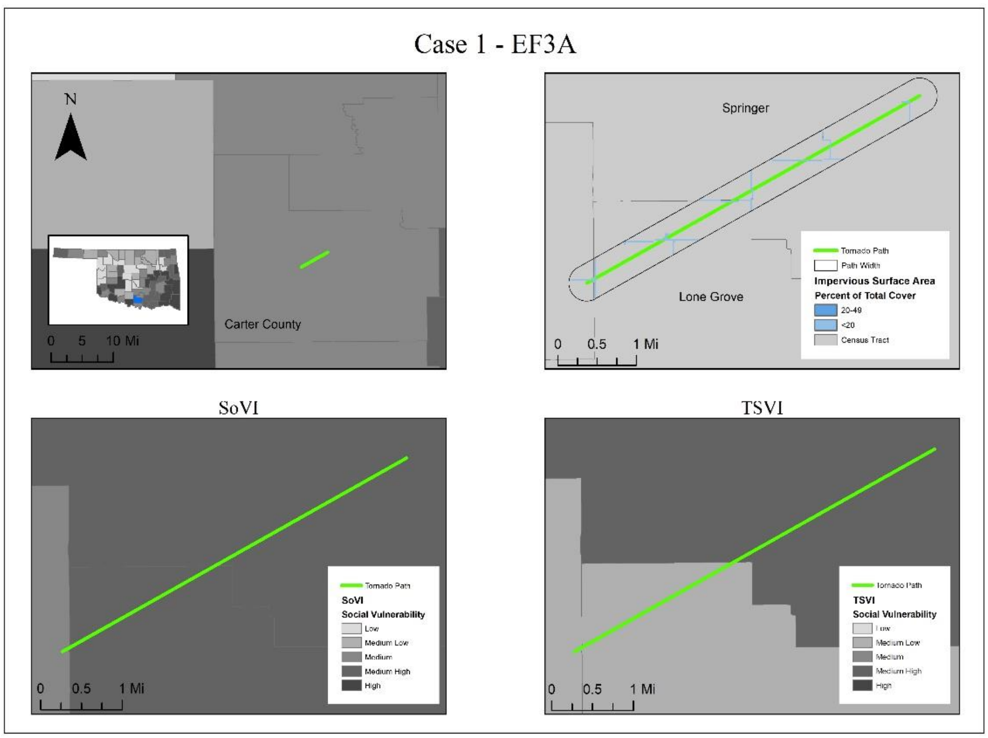


Figure 20: Case 1 – EF3A

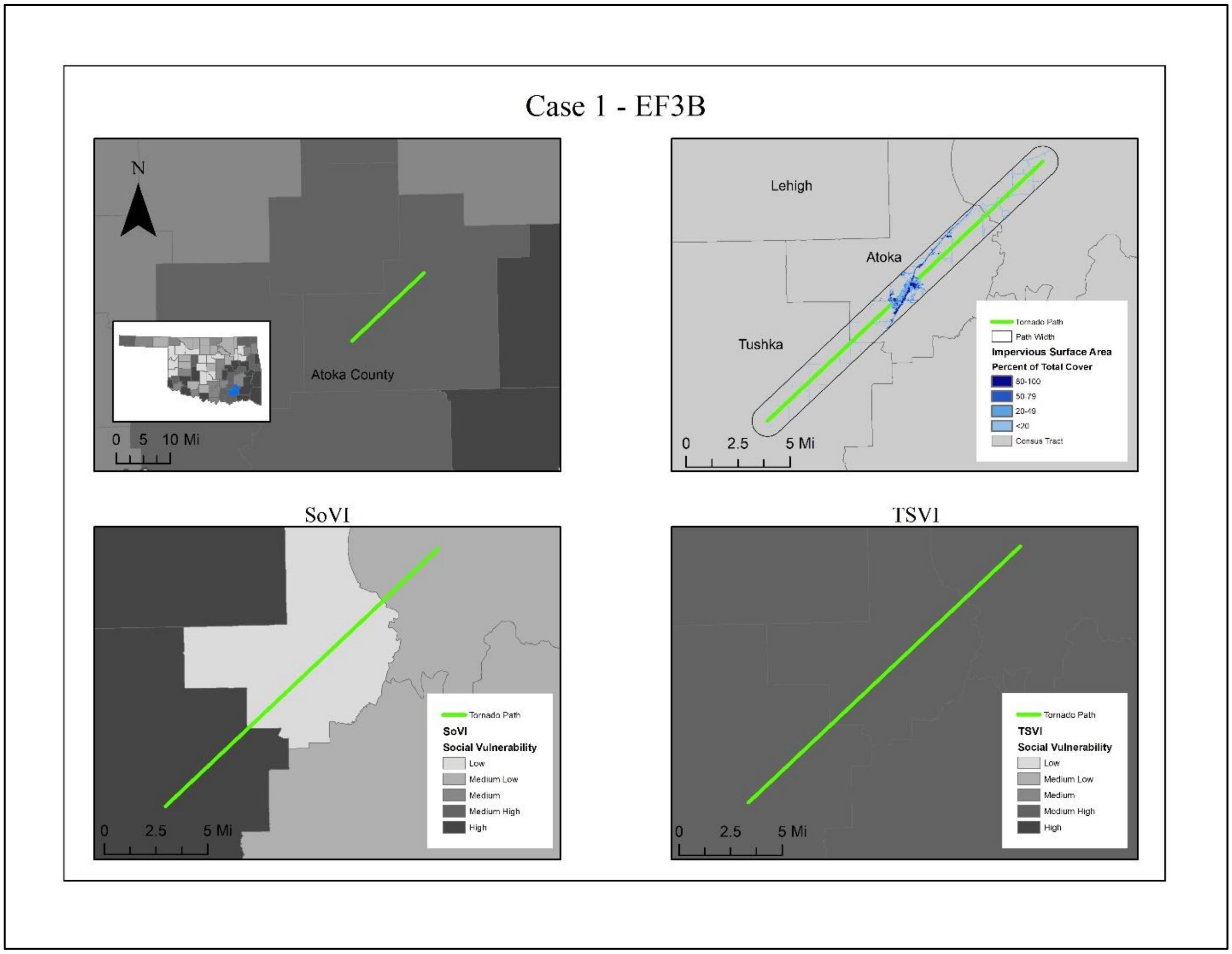


Figure 21: Case 1 – EF3B

5.3.3.2 Case 2

Despite their lower rating on the scale, tornadoes with EF1 ratings have the capacity to cause damages to property and human life. During the study period, two EF1 tornadoes engaged with tracts with considerable amounts of impervious surface area and were credited with different losses. These EF1 tornadoes, EF1A and EF1B, both happened in Oklahoma County. Tornado EF1A, found in Figure 22, occurred on 5/19/2013 and was not credited with any losses. It covered an area with 60.57% of impervious surface area, lasted 8 minutes, was 2,700 feet wide and more than 4 miles long. The SoVI described the affected tracts with medium low, medium high, and high social vulnerability. The TSVI labeled them as having medium low and low (2 tracts) social vulnerability.

Tornado EF1B, shown in Figure 23, occurred on 5/31/2013 and caused 8 injuries and $80,000 worth of property damages. It passed through an area of 100% impervious surface area, lasted 2 minutes, was 900 feet wide, and about 1 mile long. The SoVI described the disturbed tracts as having medium high social vulnerability. The TSVI suggested they had medium high and high social vulnerability. EF1B was a shorter and smaller event than EF1A but caused more damages despite having the same relative intensity.

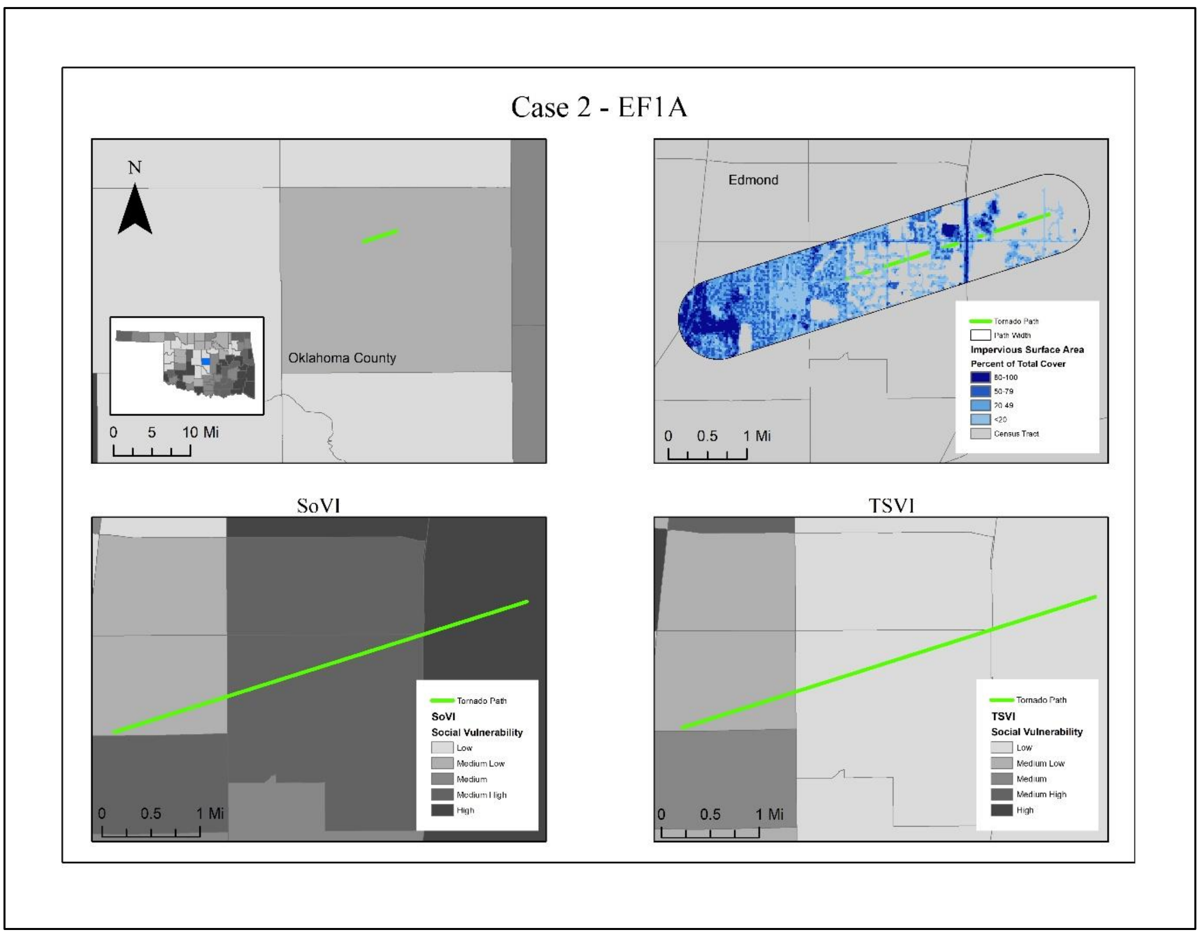


Figure 22: Case 2 – EF1A

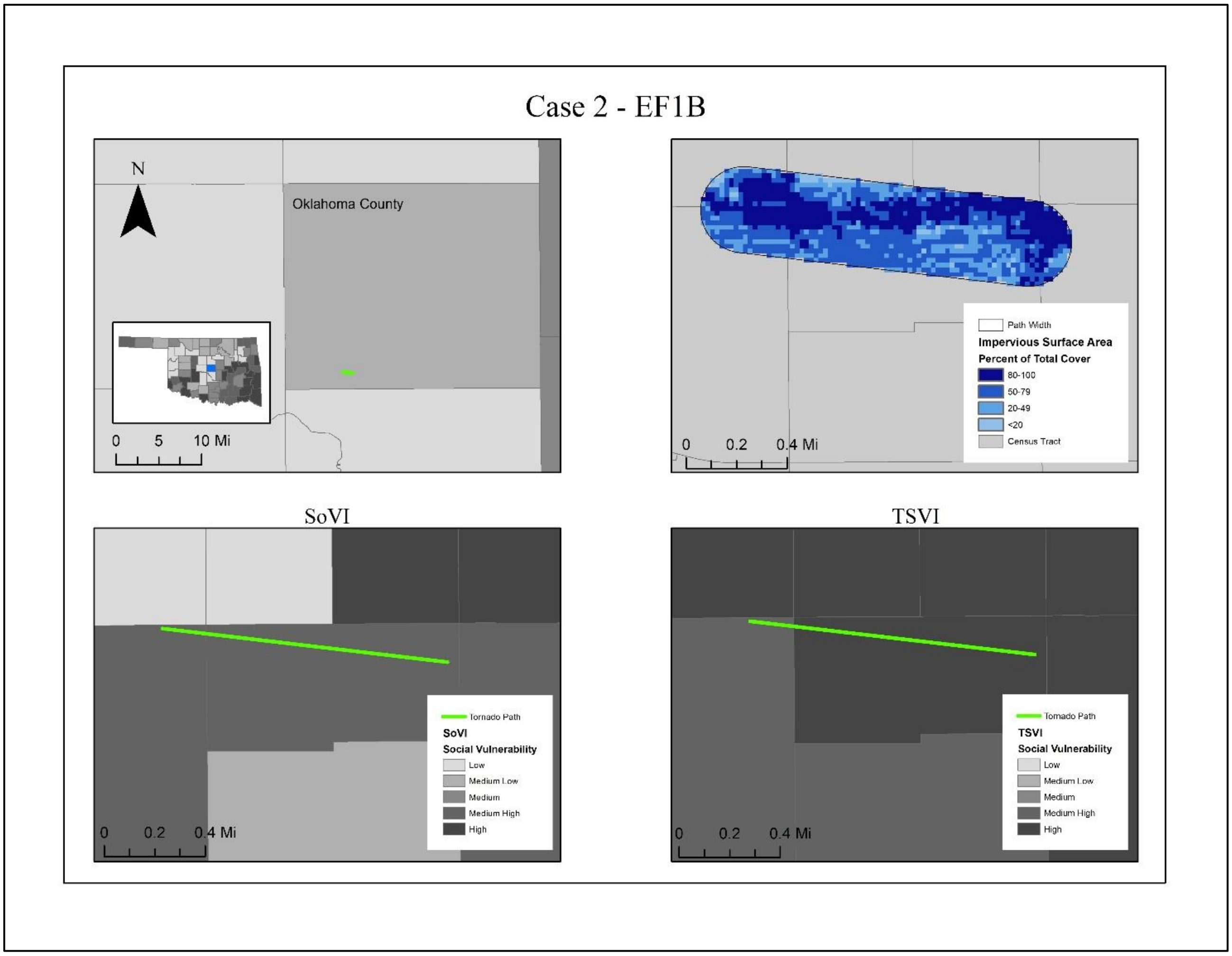


Figure 23: Case 2 – EF1B

5.3.3.3 Case 3

Tornadoes also caused losses in places with minimal areas of impervious surfaces. All EF4 events from the study period caused injuries and/or deaths but did not encounter substantial areas of impervious surfaces. The EF4 tornado event with the greatest number of injuries is explained here. An EF4 tornado that passed through Grady County and McClain County on 5/24/2011 was responsible for 61 injuries in McClain County. This EF4 event was more than 23 miles long, 2,640 feet wide, and lasted 9 minutes. Despite the amount of area this event covered, it only encountered 3.55% of impervious surface area in McClain. This event engaged with 3 census tracts that were described as having medium low, medium, and medium high social vulnerability by the SoVI. The TSVI identified these tracts as having medium, low, and medium low scores of social vulnerability. This case is displayed in Figure 24.

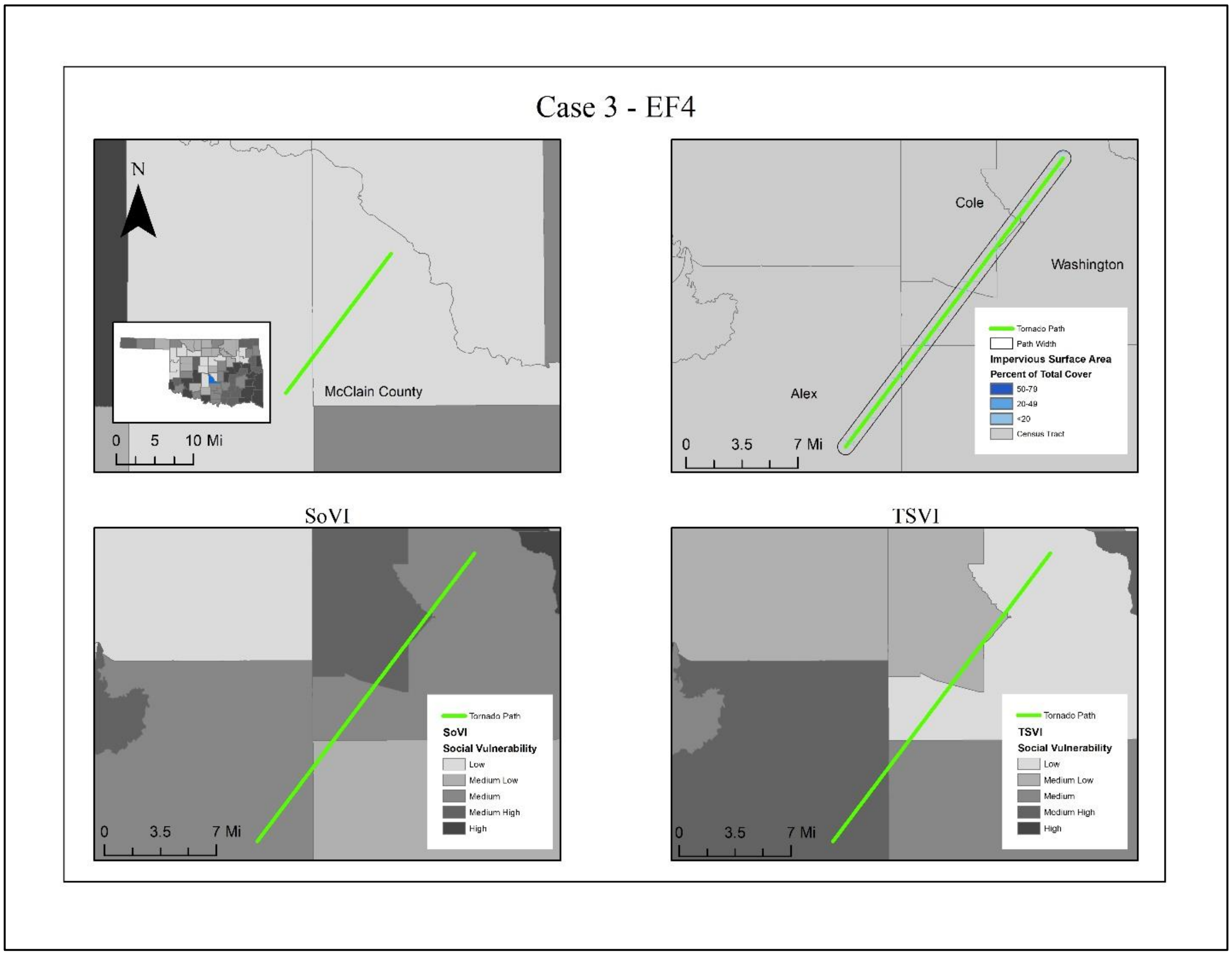


Figure 24: Case 3 – EF4

5.3.3.4 Case 4

To see how well the indexes predicted the areas with the greatest losses during the study period, two of the most damaging events are discussed. The most damaging tornado event during the study period, tornado EF5A, was an EF5 tornado that occurred in Cleveland County on 5/20/2013. The aftermath of this tornado included 212 injuries, 24 deaths, and $2 billion in property damages. EF5A tornado lasted about 5 minutes, was more than 13 miles long, and 5,700 feet wide. Passing through 20 Census tracts, of the area covered, the tornado was in contact with 32.57% of impervious surface area. The SoVI classified the 20 tracts with all classes of social vulnerability including low (3 tracts), medium low (2 tracts), medium (4 tracts), medium high (4 tracts), and high (7 tracts). The TSVI identified these tracts differently assigning social vulnerability ratings of low (17 tracts), medium (2), and medium high (1). These details can be observed in Figure 25.

The tornado that caused the second greatest number of damages, tornado EF5B, occurred on 5/24/2011 and passed through 3 counties. The EF5B tornado mostly occurred in Canadian County and Logan County, but all three counties, including Kingfisher County, experienced losses. This event lasted over 1 hour, was more than 63 miles long, and 5,280 feet wide. In total, this event was credited with causing 181 injures and 9 deaths. Canadian County suffered 112 injuries and 7 deaths. The EF5B event passed through 3 tracts in this county and encountered 4.44% of impervious surface area. The SoVI identified these tracts as having medium high and high (2 tracts) levels of social vulnerability. The TSVI described these tracts to have medium (2 tracts) and low social vulnerability.

The tornado also passed through 1 tract in Kingfisher County that resulted in 46 injuries and encountered 5% of impervious surface area. The SoVI designated this tract as having medium social vulnerability and the TSVI classified it as having medium low social vulnerability. Lastly, the EF5B event met 5 tracts in Logan County resulting in 23 injuries and 2 deaths. Of the area covered, the tornado passed through 11.49% of impervious surface area. The SoVI marked these tracts of social vulnerability as medium high (2 tracts), medium low (2 tracts), and medium. The TSVI described these tracts has having low, medium high (3 tracts), and medium social vulnerability. Tornado EF5B is shown in Figure 26.

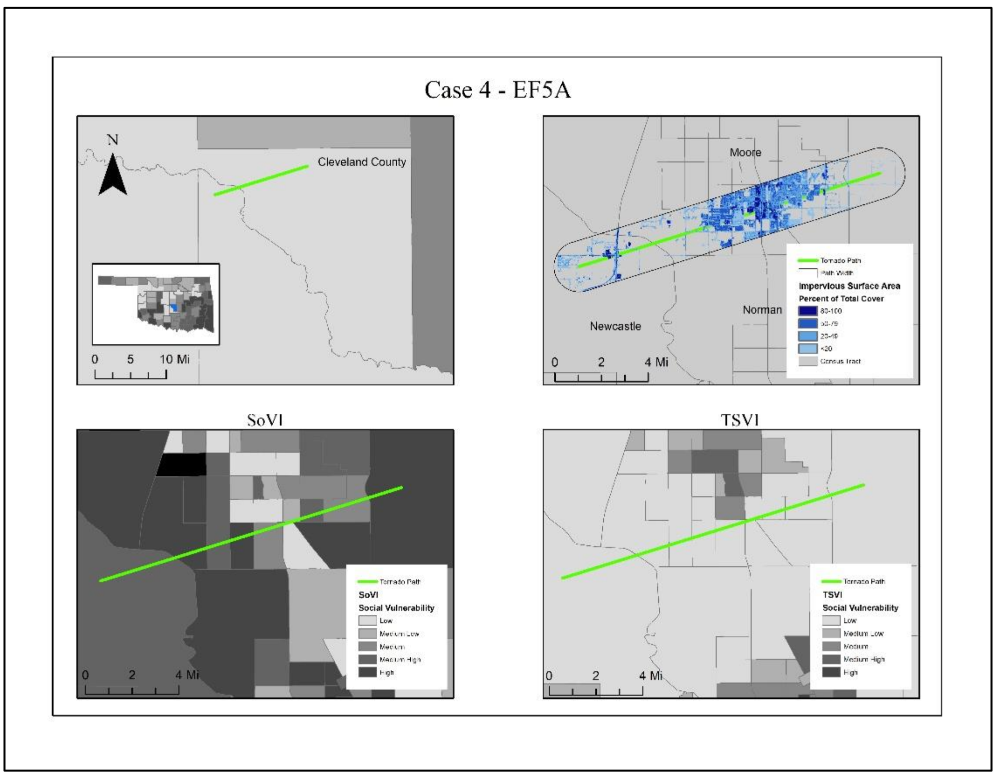


Figure 25: Case 4 – EF5A

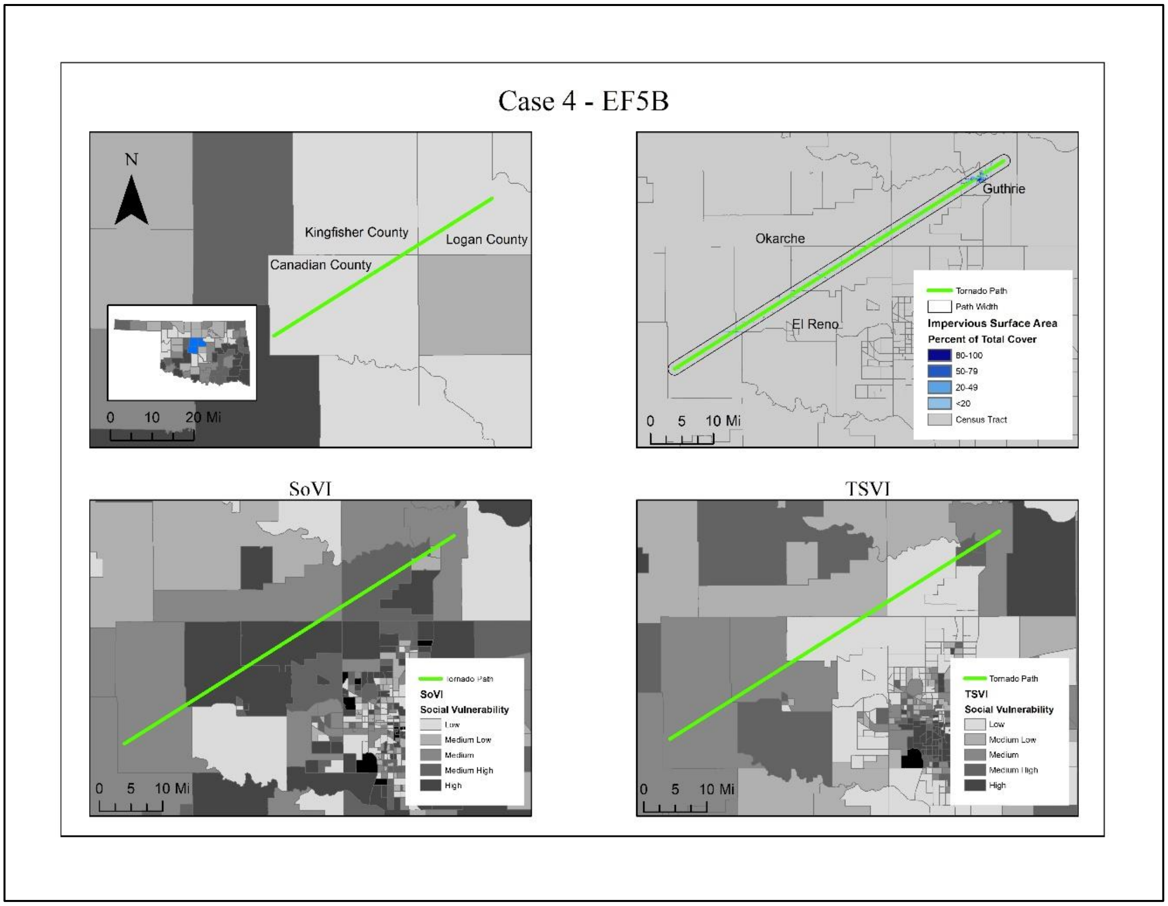


Figure 26: Case 4 – EF5B

CHAPTER VI

DISCUSSION

Identifying socially vulnerable groups is an important step toward creating resilient communities and reducing future losses of property and human life (Cutter 1996; Cutter et al. 2000; Flanagan et al. 2011). Measurement of social vulnerability is typically achieved using social vulnerability indexes, such as the Social Vulnerability Index, SoVI, introduced by Cutter et al. (2003) and the SVI by Flanagan et al. (2011). Despite the popularity and flexibility of incorporating these social vulnerability indexes to research, decision making, and policy, less attention has been given to their validation. Validation of social vulnerability indices is critical to their practical use (Fekete 2009; Rufat et al. 2015). In response, this work had two primary research objectives. The goal of **Objective 1** was to replicate a commonly used general index of social vulnerability, develop a tornado-specific social vulnerability index, and compare the spatial pattern of vulnerability identified by each in the state of Oklahoma. A general index of social vulnerability was constructed following Cutter et al. (2003) SoVI and a tornado-specific social vulnerability index, TSVI, was constructed deductively based on tornado-focused studies (see Blaikie et al. (1994), Brooks and Doswell (2002), Ashley (2007), Mulilis et al. (2000), Chaney and Weaver (2010), Dixon and Moore (2012), Chaney et al. (2013), Widen (2016), and Lim et al. (2017). The goal of **Objective 2** was to examine the relationship between social vulnerability and tornado-related losses (injuries, deaths, property damages) in Oklahoma and observe how closely related they were during the study period.

6.1 Objective 1

To examine the spatial distribution of social vulnerability throughout Oklahoma two social vulnerability indexes, the SoVI and TSVI, were constructed. The SoVI was constructed at the county and tract scale, and the TSVI was only implemented for the tracts. The construction of the SoVI used factor analysis, which produced many of the same factors for the county and tract SoVI. Factor analysis found income and home value, age and elderly (in addition to unemployment for the tract SoVI), and the Hispanic population and non-English speaking minorities to be the main drivers of social vulnerability in Oklahoma. The amount that each of these factors, in addition to the remaining 4, contributed to social vulnerability varied across space.

Based on the factor loadings and literature about social vulnerability to tornadoes, the most important factors for focusing on tornado-related losses include factors 1 (income, home value), 2 (age, elderly), and 7 (mobile homes) for the county SoVI, and factors 1 (income, home value), 2 (elderly, unemployment), and 3 (race, gender) for the tract SoVI. These factors have been found to be indicators of one's access to information, and one's ability to prepare for and recover from tornado events.

The county SoVI, tract SoVI, and TSVI classified different areas of the state as more or less socially vulnerable. The county SoVI and TSVI identified similar areas of Oklahoma with higher social vulnerability scores compared to the tract SoVI. The county SoVI and TSVI classified southeastern Oklahoma with a large concentration of high values, suggesting this area of the state was more socially vulnerable during the study period. This part of Oklahoma has more Native American reservations and a large Native American population. Literature suggests this group, and other minority groups, are more socially vulnerable and more likely to endure losses (see Cutter 1996; Cutter et al. 2003). Regions of Oklahoma that were highlighted as more socially vulnerable by the county SoVI and TSVI were generally labeled as less socially vulnerable by the tract SoVI.

There was a notable change in the locations of the high and low social vulnerability scores when shifting from the county SoVI to the tract SoVI. This may have occurred due to the influence of different tracts when they are aggregated to describe the county. It also shows how social vulnerability changes (drastically or minimally) when focusing on different spatial scales. As discussed in the literature, the outputs from a social vulnerability index are very sensitive to the construction methods used. Depending on the type of index implemented, it can produce different results. Since social vulnerability indexes are used by decisionmakers and vulnerability scientists, this could lead to poor policy and decision making.

Hot Spot Analysis and Cluster and Outlier Analysis identified significant hot and cold spots of the assigned social vulnerability scores of the tract SoVI and TSVI. The results of the two different tools were relatively consistent and did not produce vastly different outputs. The TSVI was the only index that identified the Oklahoma City and Tulsa metropolitan areas as statistically significant. The Oklahoma City and Tulsa metropolitan areas are more urban and diverse and were identified as having a mix of extreme high (more socially vulnerable) and extreme low values (less socially vulnerable). Literature explains that minority groups are more socially vulnerable, so the high social vulnerability rankings seem appropriate for these regions. The low social vulnerability rankings may be sensitive to the composition of the tracts and how the tract boundaries were selected.

Further comparison of the indexes was performed using Spearman's rank correlation coefficient. The results of this found the tract SoVI and TSVI to have a weak relationship. This demonstrated that the tract SoVI and TSVI did not rank the social vulnerability of the Census tracts similarly. This may be attributed to the different weights of the variables in the indexes. The TSVI was constructed with 10 variables whereas the tract SoVI had 28, and then later reduced to 7 factors. One of the 10 variables included in the TSVI included the percent of the population living in mobile homes (QMOHO). Literature suggests this is a crucial indicator for identifying socially vulnerable groups, especially to tornado events (see Brooks and Doswell 2002; Ashley 2007; Chaney and Weaver 2010; Dixon and Moore 2012; Chaney et al. 2013; and Lim et al. 2017). Results of the factor analysis for the tract SoVI did not find mobile homes to be a significant or strongly loading variable for this study period. Of the remaining 9 variables used to construct the TSVI (QED12LES, QESL, QRENTER, QRICH200K, QFAM, QFHH, QCVLUN, QELDERLY, and QPOVTY) 5 variables had strong loadings in the results of the tract SoVI factor analysis. These included QESL, QRENTER, QRICH200K, QFHH, and QCVLUN. Other variables not included in the TSVI had stronger loadings than these 5 for the tract SoVI. The influence of these variables in the construction of the tract SoVI could explain why the indexes had a weak relationship and assigned different social vulnerability scores to the Census tracts.

6.2 Objective 2

To examine the relationship between social vulnerability and observed losses in Oklahoma, statistical analyses and select tornado events were used. The results of the Spearman's rank correlation coefficient did not find any strong relationships between the injuries, deaths, and property damages and the county SoVI. In addition, all the relationships were negative. Negative relationships suggest that as the number of losses increased, the social vulnerability scores decreased (less socially vulnerable). This showed that higher losses occurred in areas described as being less socially vulnerable, which is the opposite of what was expected or intended of the indices. The significance values found these results to be statistically significant.

The observed losses of injuries, deaths, and property damages were allocated from the county level to the tract level using averages, impervious surface area, and randomization. The results of the Spearman's rank correlation coefficient did not find any strong relationships between the different methods of loss allocation with the tract SoVI and TSVI. For the tract SoVI, the directionalities of the relationships were both positive and negative. For the TSVI, all methods, apart from one, were negative relationships. Negative relationships imply that as values for one of the variables (e.g., losses) increased, the values for the other decreased (e.g., social vulnerability scores). The weak relationships suggested the indexes did not assign a proper social vulnerability ranking to the areas experiencing losses. Additionally, the significance values only found 4 of the correlations to be statistically significant. These findings did not display expected relationships between the observed losses and indexes. If the indexes worked as they were designed to, higher losses would have been detected in areas with higher social vulnerability scores.

The select tornado events presented in cases 1-4 provided insight into the relationship between the tract SoVI, TSVI, and observed losses. The 4 case studies included examples of instances when tornadoes with the same EF rating caused different losses across space, how losses varied across different areas of impervious surface areas, and how well the indexes described the tracts that experienced the most damages during the study period. These cases showed how complex social vulnerability and observed losses are. It also suggests there is more happening at the hyper-local level than can by described solely with Census variables. As discussed by Schmidtlein et al. (2008), it would be wise to incorporate local experts to more refined study areas to understand what is going on and possibly understand why the indexes did not work well.

The case studies illustrated that the SoVI and TSVI were not externally valid and did not accurately identify where the greatest number of losses were likely to have occurred. The SoVI and TSVI did not display the expected relationships; we expected the greatest number of losses in areas with high social vulnerability scores. Although neither index was externally valid, the SoVI was a better predictor of identifying where the losses occurred compared to the TSVI.

Externally validating indexes requires more data and information about places and tornado-related losses. With the data used in this work, we do not know the specifics of those that endured the actual damages and losses. Understanding exactly who and/or what experienced the tornado-related losses in Oklahoma could introduce ways the indexes could be adjusted. Having this information could also help explain why the indexes were not great predictors of identifying the tracts most likely to endure losses. The losses during the study period may have been experienced by social groups and/or properties not included in the SoVIs and TSVI.

6.3 Limitations, Shortcomings, and Future Work

This work addressed a gap in the social vulnerability and natural hazards literature concerning index validation. It proposed an external validation source and method to examine if social vulnerability indexes measure what they are intended to. However, this work was restricted by some limitations and shortcomings of the data and decisions utilized. Some limitations and shortcomings of this work include the study period, tornado data, observed loss data, and the methods of loss allocation to the tracts.

The findings from this work are strongly dependent on the 5-year study period chosen. The results and patterns identified may be specific to this particular time period (2010-2014) and may not be generalizable across other years. The study period also limited the number of tornado-related losses. A longer study period would have supplied more data and more observations. More data would provide more information and ultimately, more descriptive, comprehensive results. This would improve future works because sample size was a concern for some of the analyses undertaken here. At the county scale, there are 77 counties in Oklahoma and the number of observations included 26 (injuries), 12 (deaths), and 37 (property damages), which is not desirable, especially for statistical analyses. For the observed losses at the tract scale, there were more injuries and property damages than deaths. The smallest sample size of the different losses were the deaths distributed randomly (RAND_DTH) with 21 observations and the largest dataset was property damages allocated with averages (AVG_PROPERTYD) with 158 observations. A longer study period would increase the number of observations, likely yielding more robust results.

The tornado data from the SPC database is known to have some discrepancies as discussed by Widen (2016). Researchers in the tornado research community acknowledges the database's inconsistencies in rating assessments and data collection procedures of tornado damage, pathlength, width, and other attributes. The data are also known to vary from one NWS Office to another (Widen 2016). This dataset maps the tornadoes as straight-line paths with several attributes including the EF ratings. This work recognized that tornadoes do not always occur in perfectly straight lines and do not necessarily remain in contact with the ground along the entire path. In this work, no attempt was made to identify the breaks in the tornado paths. In addition, the intensity of an event and the width of the tornadoes often change throughout a tornado path, but the EF ratings and widths included in this dataset catalog the largest ratings reached for the entire path.

If the SPC or alternative databases incorporated more detailed information, future works could incorporate more accurate tornado path data. This information could provide researchers with a better understanding of how tornado events behave across space and identify more precise areas affected. They would be able to identify the people and places exposed to these hazards more precisely.

The SoVI and TSVI were externally validated using observed losses from the Storm Events Database. The database of observed losses collects injuries, deaths, and property damages (in dollars) at the county level. In this work, only the losses that occurred during the study period (2010-2014) and classified as “Tornado” under the “Event Type” category (losses related to tornado events) were included. It is possible that other documented losses were related to the tornado events from the study period but were not included. It would be useful to compare loss data across sources and various databases to obtain this information. Depending on the size of the study area, future works could contact local authorities for more precise information of observed losses.

Additionally, this work required the loss data at the tract scale. The losses were distributed to the tracts by averages, impervious surface area, and randomization. It is likely that losses were assigned to tracts that did not actually experience a tornado-related loss or that the number of losses assigned were incorrect. Three different methods were used to try and account for this. Future works could incorporate newspaper reports and or other records to help identify where these losses occurred and delegate the losses to the tracts accordingly. They could also focus on the onset time of tornadoes in relation to loss records to identify if the greatest number of losses occurred at nighttime when people are more socially vulnerable (see Ashley 2007; Ashley et al. 2008).

The SoVI and TSVI were constructed and applied to the study area without any internal adjustments. Future works should incorporate methods used by Tate (2012) and test for any internal sensitivities of the social vulnerability indexes. Sensitivities can be identified, and the indexes can be adjusted as needed to best serve the study area. A better understanding of those who experienced tornado-related losses in Oklahoma could provide insight to how these indexes could be improved upon.

6.4 Conclusion

The SoVI and TSVI did not perform as expected. This is likely due to the variable selections and data availability of the tornado-related losses. Moving forward, the SoVI and TSVI could benefit from alterations to the index construction. One index is not necessarily better than the other based solely on the methods used. Instead, careful consideration of the variables included in an index are needed to identify whether they are representative of those who experienced losses in the study area. More insight into this could supply better indexes for describing social vulnerability to tornadoes in Oklahoma. Additionally, works could incorporate on-site validation using qualitative methods and collaboration with local experts or agencies. Qualitative investigations may help better identify the specific groups that experienced tornado-related losses in Oklahoma. Index validation in this work required fine spatial data at the tract scale, which was not directly available. There is a need for data at this scale, especially for focusing on very localized hazards like tornadoes.

The findings from this work reinforce prior findings that the relationship between social vulnerability and loss is complex, and that further revision of indexes and more validation studies are needed to fully understand their value in hazard planning and decision making. This need also includes more empirical validation studies in different places at various scales, and with different hazards and temporal components (Rufat et al. 2019). As discussed by Rufat et al. (2019), indexes should not be used for informing policy or decision making until they consistently explain loss outcomes. The indexes in this work did not consistently identify the areas that experienced the greatest number of losses as more socially vulnerable, so they should not be considered for informing policy and/or decision making in Oklahoma as they currently are.

REFERENCES

Abdi, H. 2003. Factor rotations in factor analyses. *Encyclopedia for Research Methods for the Social Sciences*: 792-795.

Armaș, I., and A. Gavriș. 2013. Social vulnerability assessment using spatial multi-criteria analysis (SEVI model) and the Social Vulnerability Index (SoVI model) – a case study for Bucharest, Romania. *Natural Hazards and Earth System Sciences* 13 (6): 1481-1499.

Ashley, W. S. 2007. Spatial and temporal analysis of tornado fatalities in the United States: 1880–2005. *Weather and Forecasting* 22 (6): 1214-1228.

Ashley, W., A. Krmenec, and R. Schwantes. 2008. Vulnerability due to nocturnal tornadoes. *Bulletin of the American Meteorological Society* 89 (10): 795-807.

Biddle, M. D. 1994. Tornado hazards, coping styles, and modernized warning systems. M.A. thesis, Dept. of Geography, University of Oklahoma.

Blaikie, P., T. Cannon, I. Davis., and B. Wisner. 1994. *At Risk: Natural hazards, people's vulnerability and disasters* (1st ed.). New York, NY: Routledge.

Boruff, B. J., J. A. Easoz, S. D. Jones, H. R. Landry, J. D. Mitchem, and S. L. Cutter. 2003. Tornado hazards in the United States. *Climate Research* 24 (2): 103-117.

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

Brooks, H. E., and C. A. Doswell III. 2002. Deaths in the 3 May 1999 Oklahoma City tornado from a historical perspective. *Weather and Forecasting* 17 (3): 354-361.

Brooks, N., W. N. Adger, and P. M. Kelly. 2005. The determinants of vulnerability and adaptive capacity at the national level and the implications for adaptation. *Global Environmental Change* 15 (2): 151-163.

Centers for Disease Control and Prevention (CDC), Agency for Toxic Substances and Disease Registry (ATSDR). 2013. *ATSDR – The Social Vulnerability Index (SVI).* (https://svi.cdc.gov/SVIPublicationsPostersPresentations.html). (last accessed 21 August 2018).

Chaney, P. L., and G. S. Weaver. 2010. The vulnerability of mobile home residents in tornado disasters: The 2008 Super Tuesday tornado in Macon County, Tennessee. *Weather, Climate, and Society* 2 (3): 190-199.

Chaney, P. L., G. S. Weaver, S. A. Youngblood, and K. Pitts. 2013. Household preparedness for tornado hazards: The 2011 disaster in DeKalb County, Alabama. *Weather, Climate, and Society* 5 (4): 345-358.

Ciurean, R. L., D. Schröter, and T. Glade. 2013 Conceptual frameworks of vulnerability assessments for natural disasters reduction. In *Approaches to Disaster Management-Examining the Implications of Hazards, Emergencies and Disasters*. InTech. 3-31.

Cohen, D., and R. E. Nisbett. 1998. Are there differences in fatalism between rural Southerners and Midwesterners? *Journal of Applied Social Psychology* 28 (23): 2181-2195.

Coleman, T. A., and P. G. Dixon. 2013. An objective analysis of tornado risk in the United States. *Weather and Forecasting* 29 (2): 366-376.
Cutter, S. L. 1996. Vulnerability to environmental hazards. *Progress in human geography* 20 (4): 529-539.
Cutter, S. L., J. T. Mitchell, and M. S. Scott. 2000. Revealing the vulnerability of people and places: a case study of Georgetown County, South Carolina. *Annals of the Association of American Geographers* 90 (4): 712-737.
Cutter, S. L., B. J. Boruff, and W. L. Shirley. 2003. Social vulnerability to environmental hazards. *Social Science Quarterly* 84 (2): 242-261.
Cutter, S. L., L. Barnes, M. Berry, C. Burton, E. Evans, E. Tate, J. Webb. 2008. A place-based model for understanding community resilience to natural disasters. Global Environmental Change (18??): 598-606.
Daley, W. R., S. Brown, P. Archer, E. Kruger, F. Jordan, D. Batts, and S. Mallonee. 2005. Risk of tornado-related death and injury in Oklahoma, May 3, 1999. *American Journal of Epidemiology* 161 (12): 1144-1150.
Dixon, R. W., and T. W. Moore. 2012. Tornado vulnerability in Texas. *Weather, Climate, and Society* 4 (1): 59-68.
Edwards, R., J. G. LaDue, J. T. Ferree, K. Scharfenberg, C. Maier, and W. L. Coulbourne. 2013. Tornado intensity estimation: Past, present, and future. *Bulletin of the American Meteorological Society* 94 (5): 641-653.
Eriksen, S., and P. Kelly. 2007. Developing credible vulnerability indicators for climate adaptation policy assessment. *Mitigation and Adaptation Strategies for Global Change* 12: 495-524.
Federal Emergency Management Agency (FEMA). 2018. Mitigation Best Practices, Mapping Risks and Vulnerabilities to Increase Resilience Planning. (https://www.fema.gov/media-library-data/1461933984986-bbb94b669d354f1bff4420a8272f46df/20-Mapping-Risks-and-Vulnerabilities_web2.pdf). (last accessed 21 August 2018).
Fekete, A. 2009. Validation of a social vulnerability index in context to river-floods in Germany. *Natural Hazards and Earth System Sciences* 9 (2): 393-403.
Fekete, A., M. Damm, and J. Birkmann. 2010. Scales as a challenge for vulnerability assessment. *Natural Hazards* 55 (3): 729-747.
Finch, C., C. T. Emrich, and S. L. Cutter. 2010. Disaster disparities and differential recovery in New Orleans. *Population and Environment* 31 (4): 179-202.
Flanagan, B. E., E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis. 2011. A social vulnerability index for disaster management. *Journal of Homeland Security and Emergency Management* 8 (1, 3): 1-22.
Fothergill, A., and L. A. Peek. 2004. Poverty and disasters in the United States: A review of recent sociological findings. *Natural Hazards* 32 (1): 89-110.
Hinkel, J. 2011. "Indicators of vulnerability and adaptive capacity": Towards a clarification of the science-policy interface. *Global Environmental Change* 21 (1): 198-208.
Hout, E. M., M. Yuan, J. McIntosh, and C. Weaver. 2010. Spatial analysis of tornado vulnerability trends in Oklahoma and northern Texas. *Proceedings of the 25th Conference on Severe Local Storms*, Denver, CO, American Meteorological Society (2): 1-13.
Letsie, M. M., and S. W. Grab. 2015. Assessment of social vulnerability to natural hazards in the Mountain Kingdom of Lesotho. *Mountain Research and Development* 35 (2): 115-125.

Lim, J., S. Loveridge, R. Shupp, and M. Skidmore. 2017. Double danger in the double wide: Dimensions of poverty, housing quality and tornado impacts. *Regional Science and Urban Economics* 65: 1-15.

Mathews, A. J., M. Haffner, and E. A. Ellis. 2017. GIS-based modeling of tornado siren sound propagation: refining spatial extent and coverage estimations. *International Journal of Disaster Risk Reduction* 23: 36-44.

McBean, G., and I. Ajibade. 2009. Climate change, related hazards and human settlements. *Current Opinion in Environmental Sustainability* 1 (2): 179-186.

Mileti, D.S., 1999. *Disasters by design: a reassessment of natural hazards in the United States, Natural hazards and disasters*. Joseph Henry Press, Washington, DC. 5.

Mulilis, J.-P., T. S. Duval, and K. Bovalino. 2000. Tornado preparedness of students, nonstudent renters, and nonstudent owners: Issues of PrE Theory 1. *Journal of Applied Social Psychology* 30 (6): 1310-1329.

Multi-Resolution Land Characteristics (MRLC) Consortium. 2019. 2011 Landcover, *National Land Cover Database* (NLCD) (https://www.mrlc.gov/). (last accessed June 2019).

Murphy, D. J., C. Wyborn, L. Yung, and D. R. Williams. 2015. Key concepts and methods in social vulnerability and adaptive capacity. United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-GTR-328. Fort Collins, CO: 24.

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

National Oceanic and Atmospheric Administration (NOAA), National Centers for Environmental Information (NCEI). 2019. *Storm Events Database* (https://www.ncdc.noaa.gov/stormevents/). (last accessed June 2019).

National Oceanic and Atmospheric Administration (NOAA), National Weather Service (NWS). 2019. *Storm Prediction Center (SPC) Database* (www.spc.noaa.gov/gis/svrgis/). (last accessed June 2019).

National Oceanic and Atmospheric Administration (NOAA), National Weather Service (NWS). n.d. *The Woodward Tornado of 9 April 1947*. (https://www.weather.gov/oun/events-19470409). (last accessed 15 September 2018).

O'Brien, G., P. O'Keefe, J. Rose, and B. Wisner. 2006. Climate change and disaster management. *Disasters* 30 (1): 64-80.

O'Connor, B. P. 2000. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers* 32, no. 3 (2000): 396-402.

Oklahoma Mesonet. 2019. (http://www.mesonet.org/). (last accessed June 2019).

Patil, V. H., N. S. N. Singh, S. Mishra, and D. T. Donavan. 2008. Efficient theory development and factor retention criteria: Abandon the 'eigenvalue greater than one' criterion. *Journal of Business Research* 61 (2): 162-170.

Pielke R. A. Jr., and R. A. Pielke Sr. 1997. *Hurricanes: Their nature and impacts on society.* England: John Wiley and Sons.

Ripberger, J. T., H. C. Jenkins-Smith, C. L. Silva, D. E. Carlson, and M. Henderson. 2014. Social media and severe weather: Do tweets provide a valid indicator of public attention to severe weather risk communication? *Weather, Climate, and Society* 6 (4): 520-530.

Romanic, D., M. Refan, C-H. Wu, and G. Michel. 2016. Oklahoma tornado risk and variability: a statistical model. *International Journal of Disaster Risk Reduction* 16: 19-32.
Rogerson, P. A. 2015. *Statistical methods for geography: a student's guide* (4th ed.). Los Angeles, CA: Sage.
Rufat, S., E. Tate, C. G. Burton, and A. S. Maroof. (2015). Social vulnerability to floods: Review of case studies and implications for measurement. *International Journal of Disaster Risk Reduction* 14: 470-486.
Rufat, S., E. Tate, C. T. Emrich, and F. Antolini. 2019. How valid are social vulnerability models?. *Annals of the American Association of Geographers* 109 (4): 1131-1153.
Schmidtlein, M. C., R. C. Deutsch, W. W. Piegorsch, and S. L. Cutter. 2008. A sensitivity analysis of the social vulnerability index. *Risk Analysis: An International Journal* 28 (4): 1099-1114.
Shen, G., and S. N. Hwang. 2015. A spatial risk analysis of tornado-induced human injuries and fatalities in the USA. *Natural Hazards* 77 (2): 1223-1242.
Sims, J. H., and D. D. Baumann. 1972. The tornado threat: Coping styles of the North and South. *Science* 176 (4042): 1386-1392.
Simmons, K., and D. Sutter. 2011. *Economic and societal impacts of tornadoes*. Boston, MA: American Meteorological Society.
Tate, E. C. 2011. Indices of social vulnerability to hazards: Model uncertainty and sensitivity. PhD dissertation, University of South Carolina. 386.
Tate, E. 2012. Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards* 63 (2): 325-347.
Tate, E. 2013. Uncertainty analysis for a social vulnerability index. *Annals of the Association of American Geographers* 103 (3): 526-543.
Turner, B. L., R. E. Kasperson, P. A. Matson, J. J. McCarthy, R. W. Corell, L. Christensen, N. Eckley, J. X. Kasperson, A. Luers, M. L. Martello, C. Polsky, A. Pulsipher, and A. Schiller. 2003. A framework for vulnerability analysis in sustainability science. *Proceedings of the National Academy of Sciences* 100 (14): 8074-8079.
United States Census Bureau. 2012. *Geographic Terms and Concepts – Census Tract.* (https://www.census.gov/geo/reference/gtc/gtc_ct.html). (last accessed 23 August 2018).
United States Census Bureau, 2010-2014 American Community Survey 5-Year Estimates, Tables B09002, B25033, DP02, DP03, DP04, DP05, S0101, S0601, SF1_P42 using American FactFinder. (http://factfinder.census.gov). (last accessed November 2018).
United States Census Bureau, 2010 Census Summary, Table SF1DP1 using American FactFinder. (http://factfinder.census.gov). (last accessed November 2018).
United States Census Bureau, 2014 TIGER/Line Shapefiles, Shapefiles county, tract, stateline, metropolitan statistical areas using TIGER/Line Shapefiles. (https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2014.html). (last accessed November 2018).
Widen, H. M. 2016. New methods in tornado risk and vulnerability assessments. PhD Dissertation, Dept. of Geography, Florida State University. 159.
Wisner, B., P. Blaikie, T. Cannon, and I. Davis. 2004. *At Risk: Natural hazards, people's vulnerability and disasters* (2nd ed.). New York, NY: Routledge.
Yoon, D. K. 2012. Assessment of social vulnerability to natural disasters: a comparative study. *Natural Hazards* 63 (2): 823-843.

Zebardast, E. 2013. Constructing a social vulnerability index to earthquake hazards using a hybrid factor analysis and analytic network process (F'ANP) model. *Natural Hazards* 65 (3): 1331-1359.

APPENDICES

APPENDIX A: Terms

TERM	MEANING
Adaptive Capacity	A key element of vulnerability (Murphy et al. 2015), it is "the ability of a system to adjust to change, moderate the effects, and cope with a disturbance" (Cutter et al. 2008, 600). Also, "a combination of characteristics that are internal to an individual, community or organization and external factors that are beyond their control that either enable or constrain their ability to respond to change" (Murphy et al. 2015, 3).
Coping	Actions used, "to describe shorter-term adjustments made to simply survive a disturbance" (Murphy et al. 2015, 5). The way, "people act within the limits of existing resources and range of expectations to achieve various ends" (Wisner et al. 2004, 113).
Disaster	"The result of the impact of hazards on vulnerable people" (Wisner et al. 2004; 87).
Exposure	"A measure of the people or property that are subject to a given risk" (Boruff et al. 2003, 104). "The proximity of units or systems to disturbances" (Murphy et al. 2015, 3).
Hazard/Natural Hazard	"A dangerous phenomenon…that may cause loss of life, injury, or other health impacts, property damage, loss of livelihoods and services, social and economic disruption, or environmental damage" (Ciurean et al. 2013, 5). They can be large scale, such as forest fires, or relatively local, like tornadoes (Flanagan et al. 2011). "The interaction between physical systems and human-use systems that produce a 'loss" (Boruff et al. 2003, 104).
Hazard Mitigation	"Any action taken to reduce or avoid risk or damage from hazard events…the use of mitigation techniques and planning can increase a system's or society's resilience to hazards" (Cutter et al. 2008, 600).
Resilience	"The ability of a social system to respond and recover from disasters and includes those inherent conditions that allow the system to absorb impacts and cope with an event, as well as post-event, adaptive processes that facilitate the ability of the social system to re-organize, change, and learn in response to a threat." This includes "a system's capacity to absorb disturbance and re-organize into a fully functioning system" (Cutter et al. 2008, 599).
Resources	"The physical and social means of gaining a livelihood and access to safety" (Wisner et al. 2004, 113).
Risk	"Risk is the likelihood of occurrence (or probability) of the hazard. Risk has two domains: it includes the potential sources of risk (industrial, flooding, transportation) and the contextual nature of the risk itself (high consequence, low consequence)" (Cutter 1996, 536; Cutter et al. 2008). Ciurean et al. (2013, 5) describes risk as, "the combination of the probability of an event and its negative consequences."
Sustainability	The capacity to, "tolerate – and overcome – damage, diminished productivity, and reduced quality of life from an extreme event without significant outside assistance" (Mileti 1999, 4).

APPENDIX B: Tornado Classification Scheme

Tornado events have traditionally been evaluated using the Fujita scale (F scale). The F scale was introduced in 1971 by T. Theodore Fujita to assign ratings to tornado intensity based on observed damages. This rating system was adopted by organizations including the NWS and later utilized in the U.S. in 1973. The F scale assigns ratings based on levels of destruction and classifies tornadoes from F0 (minimal damages) to F5 (catastrophic damages) (Edwards et al. 2013). Tornado events prior to 1973 were evaluated using photographs and newspaper articles to assess damages and later included in the SPC database, an extension of NOAA's NWS (Coleman and Dixon 2013). Since the implementation of the F scale, there has been an increase in annual reported tornado events due to improved spotting techniques and radar, and NWS warning verification procedures (Coleman and Dixon 2013). Due to concerns regarding the consistency and accuracy of the F scale, a new tornado rating system was introduced shortly after; the Enhanced Fujita scale (EF scale). The EF scale is based on wind speeds, providing a clearer, objective approach to classifying these extreme hazard events. The EF scale also organizes tornadoes into categories ranging from EF0 to EF5, increasing in intensity the larger the rating. These classes also align with the original F scale (Edwards et al. 2013).

APPENDIX C: Tracts Removed from the Indexes

EXCLUDED INDEX	TRACT NUMBER	COUNTY	MISSING VARIABLE(S)
SoVI and TSVI	107.01	Oklahoma	ALL
TSVI	2007	Cleveland	PPUNIT, MHSEVAL
TSVI	24.01	Comanche	MHSEVAL, MDGRENT
TSVI	24.03	Comanche	MHSEVAL
TSVI	24.04	Comanche	MHSEVAL
TSVI	1075	Oklahoma	MHSEVAL
TSVI	1026	Oklahoma	MHSEVAL
TSVI	76.41	Tulsa	MHSEVAL
TSVI	104	Payne	MHSEVAL
TSVI	4863	Pittsburg	MHSEVAL, MDGRENT
TSVI	1027	Oklahoma	MHSEVAL
TSVI	1036.01	Oklahoma	MHSEVAL
TSVI	1025	Oklahoma	MHSEVAL
TSVI	1036.02	Oklahoma	MHSEVAL
TSVI	1091	Oklahoma	MHSEVAL
TSVI	1081.01	Oklahoma	MDGRENT
TSVI	76.38	Tulsa	MDGRENT
TSVI	1081.07	Oklahoma	MDGRENT
TSVI	2018.01	Cleveland	MDGRENT
TSVI	1085.29	Oklahoma	MDGRENT
TSVI	1067.08	Oklahoma	MDGRENT
TSVI	1085.24	Oklahoma	MDGRENT
TSVI	54.01	Tulsa	MDGRENT
TSVI	1037	Oklahoma	MDGRENT

APPENDIX D: Variables Included in the SoVI

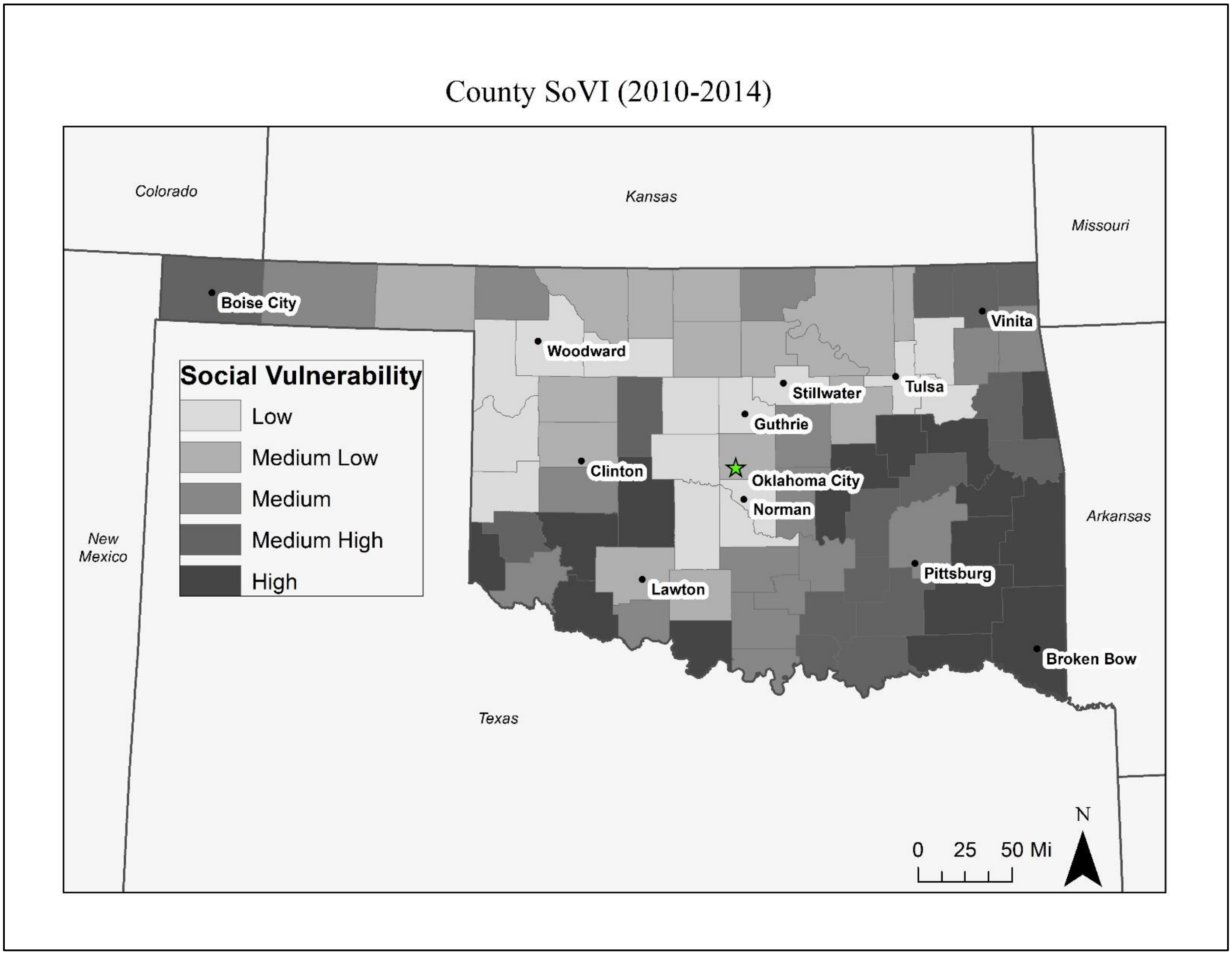
TABLE	SOURCE	ACS/CENSUS ID	VARIABLE	DESCRIPTION	SV THEME
DP02	ACS	HC03_VC10	QFHH	Percent, families with female-headed households, no spouse present	Lack of access to resources, limited access to political power and representation, beliefs and customs
DP02	ACS	HC01_VC21	PPUNIT	Estimate, households by type, average number of people per household	Social capital, beliefs and customs, type and density of infrastructure and lifelines
DP02	ACS	HC03_VC87	QED12LES	Percent, educational attainment, population over 25 years old, no high school diploma	Lack of access to resources, limited access to political power and representation
DP02	ACS	HC03_VC173	QESL	Percent, population speaking English as a second language, limited English proficiency	Lack of access to resources, limited access to political power and representation
DP03	ACS	HC03_VC09	QCVLUN	Percent, civilian labor force unemployed	Lack of access to resources, limited access to political power and representation
DP03	ACS	HC03_VC15	QFEMLBR	Percent, female participation in the labor force	Limited access to political power and representation
DP03	ACS	HC03_VC42	QSERV	Percent, population in service occupations	Limited access to political power and representation
DP03	ACS	HC03_VC50	QEXTRCT	Percent, employment in extractive industries (fishing, farming, mining, etc.)	Limited access to political power and representation
DP03	ACS	HC03_VC84	QRICH200K	Percent, income and benefits, families earning more than $200,000 per year	Social capital, beliefs and customs
DP03	ACS	HC03_VC91	QSSBEN	Percent, households receiving Social Security benefits	Social capital, beliefs and customs
DP03	ACS	HC01_VC118	PERCAP	Estimate, income and benefits, per capita income (dollars)	Social capital
DP03	ACS	HC03_VC134	QNOHLTH	Percent, population without health insurance	Lack of access to resources
DP04	ACS	HC03_VC05	QUNOCCHU	Percent, unoccupied housing units, vacancy	Building stock and age
DP04	ACS	HC03_VC65	QRENTER	Percent, renter-occupied housing units	Limited access to political power and representation, building stock and age, type and density of infrastructure and lifelines
DP04	ACS	HC03_VC84	QNOAUTO	Percent, housing units with no car available	Lack of access to resources

DP04	ACS	HC01_VC127	MHSEVAL	Estimate, median dollar value of owner-occupied housing units, home value	Social capital, building stock and age
DP04	ACS	HC01_VC189	MDGRENT	Estimate, median gross rent for renter-occupied housing units	Social capital, building stock and age
DP05	ACS	HC03_VC05	QFEMALE	Percent, female population	Lack of access to resources, limited access to political power and representation
DP05	ACS	HC03_VC50	QBLACK	Percent, Black population	Lack of access to resources, limited access to political power and representation
DP05	ACS	HC03_VC51	QNATAM	Percent, Native American population	Lack of access to resources, limited access to political power and representation, beliefs and customs
DP05	ACS	HC03_VC56	QASIAN	Percent, Asian population	Lack of access to resources, limited access to political power and representation, beliefs and customs
DP05	ACS	HC03_VC88	QHISP	Percent, Hispanic population	Lack of access to resources, limited access to political power and representation, beliefs and customs
B09002	ACS	HD01_VD010 HD01_VD02	QFAM	Percent, children living in married couple families	Lack of access to resources
B25033	ACS	HD01_VD06 HD01_VD12 HD01_VD01	QMOHO	Percent, population living in mobile homes, sum of ownership types (owner, renter) divided by total MOHO	Building stock and age, type and density of infrastructure and lifelines
S0101	ACS	HC01_EST_VC35	MEDAGE	Median age	Social capital
S0101	ACS, CENSUS	HC01_EST_VC03 HC01_EST_VC20 HC01_EST_VC19 HC01_EST_VC18 HC01_EST_VC17 HC01_EST_VC16	QAGEDEP	Percent, population under 5 years of age or 65 and over, the sum of six categories, divided by TOTPOP from 2010 Census	Frail and physically limited individuals
S0601	ACS	HC01_EST_VC67	QPOVTY	Percent, persons living in poverty	Lack of access to resources, limited access to political power and representation
SF1_P42 SF1DP1	ACS, CENSUS	D005 HD01_S001	QNRES	Percent, population living in nursing facilities, sum and then divided by TOTPOP from 2010 Census	Frail and physically limited individuals, type and density of infrastructure and lifelines

APPENDIX E: Variables Included in the TSVI

TABLE	SOURCE	ACS/CENSUS ID	VARIABLE	DESCRIPTION	LITERATURE SOURCE
S0601	ACS	HC01_EST_VC67	QPOVTY	Percent, persons living in poverty	(Dixon and Moore 2012; Widen 2016; Lim et al. 2017)
DP03	ACS	HC03_VC84	QRICH200K	Percent, income and benefits, families earning more than $200,000 per year	(Chaney et al. 2013; Lim et al. 2017)
B25033	ACS	HD01_VD06 HD01_VD12 HD01_VD01	QMOHO	Percent, population living in mobile homes, sum of ownership types (owner, renter) divided by total MOHO	(Brooks and Doswell 2002; Ashley 2007; Chaney and Weaver 2010; Dixon and Moore 2012; Chaney et al. 2013; Widen 2016; Lim et al. 2017
DP02	ACS	HC03_VC87	QED12LES	Percent, educational attainment, population over 25 years old, no high school diploma	(Chaney and Weaver 2010; Widen 2016)
S0101	ACS	HC01_EST_VC16 HC01_EST_VC17 HC01_EST_VC18 HC01_EST_VC19 HC01_EST_VC20	QELDERLY	Population 65 and over, the sum of 5 categories	(Ashley 2007; Dixon and Moore 2012; Chaney et al. 2013; Widen 2016)
DP02	ACS	HC03_VC173	QESL	Percent, population speaking English as a second language, limited English proficiency	(Chaney et al. 2013; Widen 2016; Lim et al. 2017)
DP03	ACS	HC03_VC09	QCVLUN	Percent, civilian labor force unemployed	(Widen 2016)
DP02	ACS	HC03_VC10	QFHH	Percent, families with female-headed households, no spouse present	(Blaikie et al. 1994; Lim et al. 2017)
B09002	ACS	HD01_VD010 HD01_VD02	QFAM	Percent, children living in married couple families	(Chaney et al. 2013)
DP04	ACS	HC03_VC65	QRENTER	Percent, renter-occupied housing units	(Mulilis, Duval, and Bovalino 2000)

APPENDIX F: County SoVI Reference Map



VITA

Lindsay King

Candidate for the Degree of

Master of Science

Thesis: EXAMINING TORNADO VULNERABILITY AND INDEX VALIDATION IN OKLAHOMA USING COMPONENTS OF THE SOCIAL VULNERABILITY INDEX

Major Field: Geography

Biographical:

Education:

Completed the requirements for the Master of Science in Geography at Oklahoma State University, Stillwater, Oklahoma in December, 2019.

Completed the requirements for the Bachelor of Arts in Geography at Southern Connecticut State University, New Haven, Connecticut in 2017.

Professional Memberships:

American Association of Geographers
Connecticut GIS User Network