UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SPATIOTEMPORAL ANALYSES OF TORNADO RISK AND EXPOSURE IN THE

CONTIGUOUS UNITED STATES: A MODELING STUDY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

JOSHUA JOHN HATZIS
Norman, Oklahoma
2019

SPATOTEMPORAL ANALYSES OF TORNADO RISK AND EXPOSURE IN THE
CONTIGUOUS UNITED STATES: A MODELING STUDY

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL SUSTAINABILITY

BY THE COMMITTEE CONSISTING OF

Dr. Jennifer Koch, Chair

Dr. Harold Brooks, Co-Chair

Dr. Thomas Neeson

Dr. Mark Shafer

Dr. Elinor Martin

Acknowledgments

I would first like to acknowledge the contributions of my co-authors for the studies presented in this manuscript: Dr. Jennifer Koch, Dr. Harold Brooks, and Dr. Naci Dilekli. The writing and analyses were primarily mine; however, they provided invaluable data, ideas and feedback. Without their contributions this project would not have been a success.

I would like to thank my advisor Dr. Jennifer Koch for her academic, financial, and emotional support, throughout my time as a doctoral student. When I was plagued by doubt and considering dropping out, she gave me the strength and support to push through to the end. Without her, I never would have made it this far. I also would like to thank Dr. Harold Brooks for his support in picking a topic for my dissertation and for his advice on how to make my model better. I would also like to thank the remainder of my committee, Dr. Elinor Martin, Dr. Mark Shafer, and Dr. Thomas Neeson, for their comments and advice throughout this process.

I would like to thank the staff within the University of Oklahoma Department of Geography and Environmental Sustainability for their support and help in navigating the hurdles of completing a Ph.D. I would also like to thank the Department of Geography and Environmental Sustainability as well as the Graduate College at the University of Oklahoma for funding opportunities to present early versions of my dissertation chapters at professional conferences. The thoughtful feedback and experience I gained through those presentations helped improve this dissertation immensely.

Finally, I would like to thank my friends and family for their continued love and support throughout this process. You've made me what I am today and I will be forever in your debt for everything you've done for me.

Abstract

Tornadoes represent a significant threat to both life and property across the United States. It is unclear exactly how climate change may influence the occurrence and intensity of small-scale phenomenon (below the resolution of current climate models), such as tornadoes. However, changes have already been observed in the spatial and temporal variability of tornadoes. Regardless of whether climate change results in substantial changes in tornado frequency or severity, continued population growth and the expansion of urban areas will likely lead to an increasing amount of exposure of people and buildings to tornadoes. Potential future changes in tornado risk and exposure require new methods for studying tornado impacts through simulation. This dissertation discusses the development and application of a new tornado impacts model, the Tornado Daily Impacts Simulator (TorDIS). First, we conducted a spatiotemporal analysis of near-miss violent tornadoes as a justification for the use of spatial models in tornado impact analysis. Second, we discussed the development of TorDIS and showcased its utility via comparisons of annual tornado exposures between six metropolitan areas. We also show an example of using TorDIS to assess potential tornado impacts on individual high-risk days. Finally, we describe a case study, using TorDIS, over the Oklahoma City Metropolitan Area of the combined and isolated effects of climate change and urban development on tornado impacts. Models such as TorDIS can be used by emergency managers to pre-allocate resources to the areas of greatest risk/potential impact, by city planners to assess how changes in land use could affect the potential tornado risk and impacts, and as a justification for the placement of public storm shelters in the areas of highest potential tornado risk and impacts.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Tornadoes are an extremely powerful force of nature capable of posing a significant threat to both life and property. Based on the 1981 – 2010 climate norms, an average of 1,100 tornadoes are reported in the United States each year (SPC, 2017) and, while most tornadoes cause relatively little damage and few casualties (Ashley, 2007), a single tornado can cause a large number of fatalities (Grazulis, 1993; Simmons and Sutter, 2012) and result in billions of dollars in damages (Brooks and Doswell, 2001a; Paul and Stimers, 2012).

While major improvements in tornado detection, tornado warning dissemination systems, and greater public awareness have led to a reduction in tornado casualties since the early 20th century (Ashley, 2007; Doswell *et al.*, 2012; Brotzge and Donner, 2013) the damages caused by the strongest tornadoes have increased due to population growth and the expansion of urban area and more expensive built structures in the damage path of tornadoes (Ashley and Strader, 2016; Strader *et al.*, 2017a). As the climate continues to change, it is possible that tornado frequency and intensity may also change (Trapp *et al.*, 2007; Gensini *et al.*, 2014b). Concurrently, the population within the United States is becoming increasingly urban, with projections suggesting that 89% of the population will be living in urban areas by 2050 (United Nations, 2018). To better understand the effects climate change and urbanization may have on future tornado impacts it is important to understand the risk and vulnerability as well as how models can be used to improve understanding of the processes and their interlinkages associated with tornadoes.

1.1) Risk and Vulnerability

The hazards literature has many definitions of risk and vulnerability (Paul, 2011). In this study, risk is defined as the likelihood that a hazard will occur and impact an area at a given time

(Ashley and Strader, 2016). Vulnerability is defined as the potential for loss (Cutter *et al.*, 2003). This potential itself is a function of exposure (the placement of a system in position to experience a hazard), sensitivity (the degree to which a system is affected by the hazard), and adaptive capacity (the ability or potential of a system to modify itself to better cope/adapt to current and future hazards; Yohe and Tol, 2002; Eakin and Luers, 2006; Morss *et al.*, 2011). Specifically, the hazard of interest here – a tornado – is defined as a violently rotating column of air which descends to the ground from a cumuliform cloud and is usually visible as a funnel cloud (Federal Emergency Management Agency, 2018). With respect to tornadoes, exposure, for this study, is defined as the number of persons residing in the damage path of the tornado (Wurman *et al.*, 2007; Ashley *et al.*, 2014; Rosencrants and Ashley, 2015). In this research, exposure is the primary metric of tornado impact.

Tornadoes are typically generated from thunderstorms with the majority being produced by a specific type of thunderstorm with a rotating updraft (supercell; Moller *et al.*, 1994; Shafer *et al.*, 2012). These types of thunderstorms and the accompanying tornadoes require several key ingredients to form: (1) warm, moist air at the surface, (2) cold, dry air aloft, (3) steep vertical lapse rates (instability), (4) strong vertical wind shear (change in wind speed and/or direction with height), and (5) a source of lift (*e.g.* convection or frontal boundaries; Schultz *et al.*, 2014; Sherburn and Parker, 2014). These ingredients are most prevalent in North America with mountains to the west and warm oceans to the southeast, but tornadoes have been reported on every continent except Antarctica (Goliger and Milford, 1998; Brooks and Doswell, 2001b; Groenemeijer and Kühne, 2014). Within the United States, these ingredients are present most frequently in the central and southern states, where dry air aloft, from the Rocky Mountains, meets moist southerly flow from the Gulf of Mexico in close proximity to the jet stream (Ashley,

2007; Dixon *et al.*, 2011; Schultz *et al.*, 2014). Peaks in the significant (rated two or higher on the Enhanced Fujita scale [EF2+]) and violent (EF4+) tornadoes occur along an "L"-shaped corridor extending from Iowa in the north down through Oklahoma to Alabama in the east (Concannon *et al.*, 2000; Doswell *et al.*, 2012).

Tornado risk peaks, in the contiguous U. S., between March and June when temperature differences between north and south are greatest, but tornadoes can occur at any time of year (Brooks *et al.*, 2003a; Krocak and Brooks, 2018). In the contiguous U.S., risk peaks earliest in the south-central states and shifts northward with the jet stream and advected moisture from the Gulf of Mexico (Concannon *et al.*, 2000; Schultz *et al.*, 2014). States in the southeastern U.S. have more complex seasonality with respect to tornadoes, with a primary peak during spring, a secondary peak in fall, and an occasional hurricane-spawned tornado in the summer (Dixon *et al.*, 2011; Chaney *et al.*, 2013a). Tornadoes can occur at any time of day but are most likely to occur during late afternoon and early evening due to maxima in surface heating during those times (Davies and Fischer, 2009; Krocak and Brooks, 2018). Nocturnal tornadoes (*i.e.*, occurring between sunset and sunrise) represent only about a quarter of all observed events but cause 40% of tornado casualties (Ashley, 2007; Kis and Straka, 2010). While nocturnal tornadoes do occur anywhere daytime tornadoes occur, they represent a larger relative proportion of tornadoes in the southeastern states, with the greatest risk of nocturnal tornadoes in Arkansas and Tennessee (Ashley *et al.*, 2008; Simmons and Sutter, 2011).

Tornado casualties occur where tornado risk meets social and/or biophysical vulnerability (Wurman *et al.*, 2007; Rosencrants and Ashley, 2015; Ashley and Strader, 2016). Tornado vulnerabilities are usually related to one of three factors: situational awareness, risk perception, and ability to seek shelter (Cutter *et al.*, 2003; Merrell *et al.*, 2005; Dixon and Moore, 2012).

Since the 1990s, advanced radar and dense storm spotter networks have ensured that most tornadoes are warned. When a warning is issued it is thoroughly disseminated through many sources, including NOAA weather radio, broadcast and social media, and mobile and internet-based applications, enabling most of the population to receive the warning (Coleman *et al.*, 2011; Brotzge and Donner, 2013). In spite of these advances, many people still may remain unaware of an impending tornado due to lack of understanding (*e.g.*, language barriers), lack of reception of a warning (lack of a television, radio, computer; power outage; civil defense siren failure; being asleep), or lack of issuance of a warning (*e.g.*, failure to detect due to a lack of visibility; Paul and Stimers, 2012; Uccellini, 2014).

To act to reduce vulnerability, a person first has to perceive that they are at risk. Tornadoes are low-frequency events and are thus often treated as no risk events leading to cavaliere attitudes towards tornado warnings (Doswell *et al.*, 1999; Simmons and Sutter, 2007). Frequent false alarms (due to radar-indicated tornadoes that never touchdown or tornado warning sirens that are used too frequently) can erode public confidence in the warning dissemination systems making people less likely to seek shelter (Simmons and Sutter, 2009; Brotzge *et al.*, 2011; Paul and Stimers, 2012). For example, a survey conducted by Paul and Stimers (2012), about warning response for the EF5 tornado that hit Joplin, Missouri on May 22, 2011, found that approximately 28% of the respondants ignored the tornado warnings due to frequent false alarms. Additionally, tornado myths and folklore (*e.g.*, hills act as protective barriers to tornadoes) can lead people to assume they are safe and thus not seek shelter or seek shelter in the wrong location (Hoekstra *et al.*, 2011; Hoffman, 2013; Klockow *et al.*, 2014).

Once a person has recognized the need to seek shelter, they have to have access to shelter and the ability to seek it (be mobile). The safest place to seek shelter during a tornado is a

Federal Emergency Management Agency (FEMA) rated storm shelter as these provide near-total protection; however, these are only available to homeowners with sufficient disposable income (Merrell *et al.*, 2002; Simmons and Sutter, 2006), and communities that are able and willing to invest in public storm shelters (Buckley, 2002). Even without a FEMA-rated shelter, the National Weather Service (NWS) recommends sheltering in place in an interior room without windows or in the nearest sturdy structure as any building will provide protection from flying debris (Schmidlin *et al.*, 2009; Chaney *et al.*, 2013b; Lindell *et al.*, 2016). Mobile and manufactured homes are the one exception to the shelter-in-place recommendation as they can easily be severely damaged and/or flipped in tornado winds (Brooks and Doswell, 2002; Simmons and Sutter, 2007; Lim *et al.*, 2017). In fact, a study by Brooks and Doswell (Brooks and Doswell, 2002) found the risk of death in tornado in a mobile home to be about 10 – 15 times higher than the risk of death in a single-family home.

A person who is mobile will be able to either move to a safe location inside or travel to the nearest private/public shelter. Physical mobility can be impaired for the elderly, infirm and disabled (Ashley, 2007; Simmons and Sutter, 2011; Kuligowski *et al.*, 2014). A lack of access to transport can further reduce mobility for those who are physically impaired as well as those living in conditions of poverty (Zoraster, 2010; Bowser and Cutter, 2015; Strader *et al.*, 2019).

1.2) Future Tornado Impacts

Climate change and urbanization will both likely play a role in shaping future tornado impacts. Modeling studies indicate that climate change likely will have a much smaller impact than urbanization (Ashley and Strader, 2016; Jones, 2017; Strader *et al.*, 2017a). However, as increases in moisture and instability, due to warming temperatures, contrast with decreases in vertical wind shear, due to decreased latitudinal temperature gradients, the favorability of the

atmosphere towards tornado development may change in the future (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2008; Gensini *et al.*, 2014b). There have already been changes in the variability of tornado occurrence with the number of tornado outbreak days increasing while the overall number of tornado days decreases (Brooks *et al.*, 2014; Elsner *et al.*, 2014).

Whether or not the number or intensity of tornadoes increases with climate change, the landscape that experiences tornado impacts is continuing to change. The amount of urban area is increasing and at the same time population densities in existing urban areas are also increasing (Seto *et al.*, 2011; United Nations, 2018). As more people and properties are placed in areas of high tornado risk, the potential impacts continue to rise (Ashley *et al.*, 2014). A study by Ashley and Strader (2016) found that, east of the Continental Divide in the United States, the number of housing units impacted by EF1+ tornadoes per square kilometer of tornado path area nearly tripled from 1954 to 2014. The character of the development also influences the nature of the tornado impacts. Two common development types are sprawl (expansion of a city's boundaries through the addition of low density housing) and infill (development of vacant space within an urban area or an increase in population density for existing developments; Hamidi and Ewing, 2014; Laidley, 2016). When the development is low density (*i.e.*, sprawl) the median impacts tend to be larger, while when the development is high density (*i.e.*, infill) the maximum impacts appears to be larger (Morss *et al.*, 2011; Rosencrants and Ashley, 2015; Strader *et al.*, 2018). With the US population projected to increase to nearly 390 million by 2050, 89% of that total residing in urban areas (United Nations, 2018), the number of people impacted by tornadoes is likely to continue to increase (Strader *et al.*, 2017a, 2018).

1.3) Modeling Tornado Impacts

Since tornadoes are rare events and a consistent, long-term tornado record does not exist (Doswell and Burgess, 1988; Brooks *et al.*, 2003a; Verbout *et al.*, 2006), it is not possible to develop a reasonable assessment of tornado impacts from the historical record alone (Doswell, 2007). One way around this problem is to create models to simulate or project potential tornado impacts. Modeling studies have the advantage of being able to simulate thousands of years' worth of tornadoes, as opposed to the ~67-year record available from the Storm Prediction Center (SPC; SPC, 2017). While these simulations are subject to the same biases as the historical record (*e.g.*, inaccuracies due to amateur observations and underestimation of tornado intensity due to a lack of damage to sturdy structures; Verbout *et al.*, 2006; Edwards *et al.*, 2013), they present a better snapshot of overall tornado risk due to the increased number of outcomes being assessed (Strader *et al.*, 2016).

Most tornado simulation studies have involved transposing historical (Rae and Stefkovich, 2000) or synthetic (with path lengths, widths, and intensities derived from historical averages; Ashley *et al.*, 2014; Wurman *et al.*, 2007) tornado paths over urban areas. These studies have mostly focused on the impacts of worst-case scenario events where EF4+ tornadoes track through the central business district or urban core of a major city without regards to the likelihood of such an event. Some studies have focused on specific external influences on these impacts including urban growth and expansion (Ashley *et al.*, 2014) and daily mobility patterns (Paulikas, 2015). While major cities, such as St. Louis (Missouri), have been hit by tornadoes in the past (Galway, 1981), the odds of a rare, violent tornado hitting the urban core of a major city (covering a relatively small surface area) are minimal (Wurman *et al.*, 2007) and, thus, worst-case scenario events are extremely unlikely to occur. A more likely scenario is that a strong to

violent tornado (EF3+) will affect a high-density development in the suburbs or outskirts of a city, causing moderate to high fatalities.

One way around this problem is to create large scale models that can simulate tornado impacts across the United States. One modeling approach is to use spatial statistical models to project tornado risk as a function of climatology, population density or teleconnections (*e.g.*, El Niño Southern Oscillation (ENSO); Elsner *et al.*, 2013b, 2016). This methodology allows for the assessment of tornado risk but does not address tornado impacts. Other statistical models have linked tornado casualties to factors such as tornado kinetic energy, population density, number of mobile homes, and time of day (Elsner *et al.*, 2018; Fricker and Elsner, 2019). However, these models are incapable of directly addressing future changes in tornado impacts.

Another approach is to directly simulate tornado footprints (rectangular shaped polygons with length and width corresponding to that of the tornado) and intersect these footprints with population or housing data to assess potential impacts (Meyer *et al.*, 2002; Daneshvaran and Morden, 2007). This approach requires sampling from historical data or theoretical distributions to determine various characteristics of a given simulated tornado (*e.g.*, tornado path width and intensity). Sampling is typically done via the Monte Carlo method (Meyer *et al.*, 2002; Strader *et al.*, 2016), a technique using repeated random sampling to determine the probability distribution of some unknown quantity (Mooney, 1997). When theoretical distributions are used it is still possible to get record characteristics through random sampling. This is important given that we do not know if we have yet seen the full range of possible tornado characteristics (*e.g.*, the widest tornado on record (4.2 km) occurred in El Reno, Oklahoma on May 31, 2013; Bluestein *et al.*, 2015). These characteristics can be used to create realistic tornado footprints (rectangular shaped polygons with length and width corresponding to that of the tornado). These footprints can in

turn be intersected with population or housing unit data to determine the number of people or houses potentially impacted by tornadoes (Strader *et al.*, 2016).

1.4) Research Overview

Past spatial modeling efforts for tornado impacts have focused on the impacts at an annual time scale. Emergency managers have expressed interest in projecting tornado impacts at a daily time step to better prepare for potential high-impact days. The current methodology employed in models such as Strader *et al.*'s (2016) Tornado Impact Monte Carlo (TorMC) model is to use historical climatology with an annual time step and would not be suitable for projecting daily tornado impacts. One methodology for projecting tornado impacts at a daily time scale would be to combine the stochastic methods of the TorMC model (Strader *et al.*, 2016) with daily atmospheric environmental data (*e.g.*, reanalysis data; Trapp *et al.*, 2007; Gensini and Ashley, 2011). By using daily atmospheric data, it is possible to spatially distribute tornadoes within favorable environments (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2008), on a given day, and to link tornado parameters to the environment (Colquhoun and Riley, 1996; Naylor and Gilmore, 2012). The research described here begins with a study on the spatiotemporal patterns of near-miss violent tornadoes in the U.S. using observed data as a justification for future modeling studies. The dissertation then proceeds to describe the development of a new spatially explicit, environmentally driven, tornado impacts model and assess its utility for the spatial analysis of potential tornado risk and exposure. It concludes by discussing a case study using the model to assess the combined and isolated effects of climate change and urbanization on future tornado impacts in the Oklahoma City metropolitan area. The individual chapters are described hereafter.

Chapter 2 examines the spatiotemporal distribution of near-miss violent tornadoes. The hazards literature defines a near-miss as an event that narrowly avoids occurrence by random

chance (Dillon *et al.*, 2011; Tinsley *et al.*, 2012). Examples of near-miss tornadoes include

tornadoes passing just outside of a major city (Prosser, 1976) and supercells, with a history of

producing tornadoes, passing over an urban area without one (Chaney and Weaver, 2010;

Sherman-Morris, 2010). Near-misses are important because they can lead to two diametrically

opposed responses. If an individual treats the near-miss as a nonevent they could underestimate

tornado risk and be less likely to respond in the future (Tinsley *et al.*, 2012). Conversely, a

person could also interpret the near-miss as a lucky break and actually be more likely to respond

next time (Dillon *et al.*, 2014). Most studies on near-miss severe weather events have focused on

the impacts of a single event without explaining what defines a near-miss (Prosser, 1976;

Sherman-Morris, 2010). The only large-scale study with a clear definition of a near-miss that the

author is aware of was for tropical cyclones that were forecast to make landfall in the United

States but did not (Powell and Aberson, 2001). The study presented in Chapter 2 sought to fill in

the literature gap by clearly defining near-miss violent tornadoes and study their spatiotemporal

distribution from 1995 to 2016. This study also updated the violent tornado climatology of

Concannon *et al.* (2000). This chapter was co-authored by myself, Dr. Jennifer Koch and Dr.

Harold Brooks and published in the January 2019 issue of *Weather, Climate and Society*.

Chapter 3 introduces the Tornado Daily Impacts Simulator (TorDIS) - a spatially explicit,

Monte Carlo-based, tornado impacts model which operates at a daily time scale and links

tornado occurrence to the atmospheric environment. Tornadoes are rare events and have a short

period of historical record. Within the U. S., it is likely that the true risk posed by tornadoes is

underestimated (Brooks *et al.*, 2003b; Verbout *et al.*, 2006; Doswell, 2007). One method to

overcome this problem is to simulate thousands of years of tornadoes (including tornadoes with

potential record-breaking tornado path lengths and widths) to get a broader picture of tornado

risk (Meyer *et al.*, 2002; Daneshvaran and Morden, 2007; Strader *et al.*, 2016). Previous models have used annual time scales for their simulations. However, from an operational standpoint it would be useful to be able to simulate tornadoes at a daily time step (Karstens *et al.*, 2015; Smith *et al.*, 2015). The study displayed in Chapter 3 describes the development of TorDIS which uses atmospheric reanalysis data to determine the favorability of the environment for tornadoes on a given day (Sobash *et al.*, 2011; Nowotarski and Jensen, 2013; Karstens *et al.*, 2015). Tornadoes are only simulated on days with favorable conditions. Their spatial distributions, intensities, path lengths, and path widths are all linked to the environmental conditions on that day (Colquhoun and Riley, 1996; Thompson and Edwards, 2000; Corfidi *et al.*, 2010). The tornadoes are spatially distributed using a weighted random placement with the weighting associated with a tornado probability raster. The probability rasters are created using a series of logistic regression equations which incorporate common severe weather diagnostic parameters (*e.g.*, convective available potential energy (CAPE) and vertical wind shear (VWS); Brooks *et al.*, 2003b; Gensini *et al.*, 2014a). The selection of the variables included in these equation was determined through a global sensitivity analysis using SimLab software (Giglioli and Saltelli, 2008). The study showcases the model's utility through 1,000-year simulations over select metropolitan areas as well as comparisons between modeled and observed tornado impacts on several high-impact tornado days. This chapter was co-authored by myself, Dr. Jennifer Koch and Dr. Harold Brooks and has been submitted to the journal *Meteorological Applications* and is currently accepted with major revisions.

Chapter 4 is a simulation study of the isolated and combined effects of climate change and urban development on tornado impacts in the Oklahoma City metropolitan area using TorDIS. It is unclear how climate change will influence tornado risk across the U. S.; however,

studies indicate an increase in the number of days favorable for the development of severe weather (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2013; Gensini *et al.*, 2014b). At the same time, urban areas are expected to continue to expand (Seto *et al.*, 2011; Strader *et al.*, 2017a) with the urban population of the U. S. expected to exceed 89% by 2050 (United Nations, 2018). Oklahoma City, the state capitol and largest city in Oklahoma (U.S. Census Bureau, 2018), is an example of a sprawling urban area that is expected to continue to expand into the late 21st century (Barker, 2012). The Oklahoma City metropolitan area is also in an area of high tornado risk (Concannon *et al.*, 2000; Doswell *et al.*, 2012) having experienced two particularly deadly tornadoes in the last 20 years (Brooks and Doswell, 2002; Atkins *et al.*, 2014). This study uses TorDIS to run four 1,000-year simulations over the three primary counties in the Oklahoma City Metropolitan Area (Canadian, Cleveland and Oklahoma). The four simulations represent the following climate change and urbanization scenarios: (1) no climate change and no urban development (baseline); (2) climate change only, (3) urban development only; and (4) climate change and urban development. The simulations are compared to test how both climate change and urban development can shape future tornado impacts. The effect that county-level differences in land use had on tornado impacts between and within simulations was also assessed. This chapter was co-authored by myself, Dr. Jennifer Koch, Dr. Naci Dilekli and Dr. Harold Brooks and will be submitted to *The Southwestern Geographer*.

1.5) Broader Impacts and Intellectual Merit

This dissertation describes a first attempt to fill a gap in the field of tornado impacts research with an environmentally driven tornado impacts model. The proposed model, TorDIS, is capable of assessing tornado impacts on a daily time step and thus can be used to project tornado impacts on any day with anticipated tornado risk. Foreknowledge of the potential

severity of a high-risk day could allow emergency managers to allocate their resources ahead of

time potentially saving lives (Brooks *et al.*, 2008). Tornado impact models, such as TorDIS, can

be used by city planners to assess how changes in land use will affect potential tornado exposure

and vulnerability (Stone *et al.*, 2010; Ashley *et al.*, 2014; Strader *et al.*, 2018). This knowledge

can help justify changes to a city's land-use plan to reduce tornado risk or justify the addition of

public storm shelters in potential high-risk/high-impact areas (Liu *et al.*, 1996; Balluz *et al.*,

2000). The use of atmospheric environmental data also allows for the testing of how climate

change may influence tornado risk through a potential changes in the number of tornado

favorable days (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2013), the number of tornadoes occurring

on favorable days (Brooks *et al.*, 2014; Elsner *et al.*, 2014), and the interannual variability in

tornado occurrence (Strader *et al.*, 2017a).

Chapter 2 Spatiotemporal Analysis of Near-Miss Violent Tornadoes in the United States

Abstract

In the hazards literature, a near-miss is defined as an event which had a nontrivial probability of causing loss of life or property but did not due to chance. Frequent near-misses can desensitize the public to tornado risk and reduce responses to warnings. Violent tornadoes rarely hit densely populated areas, but when they do they can cause substantial loss of life. It is unknown how frequently violent tornadoes narrowly miss a populated area. To address this question, this study looks at the spatial distribution of possible exposures of people to violent tornadoes in the US. We collected and replicated tornado footprints for all reported US violent tornadoes between 1995 and 2016, across a uniform circular grid, with a radius of 40 km and a resolution of 0.5 km, surrounding the centroid of the original footprint. We then estimated the number of people exposed to each tornado footprint using proportional allocation. We found that violent tornadoes tended to touch down in less populated areas with only 33.1% potentially impacting 5,000 persons or more. Hits and near-misses were most common in the Southern Plains and Southeastern US with the highest risk in central Oklahoma and northern Alabama. Knowledge about the location of frequent near-misses can help emergency managers and risk communicators target communities that might be more vulnerable, due to an underestimation of tornado risk, for educational campaigns. By increasing educational efforts in these high-risk areas, it might be possible to improve local knowledge and reduce casualties when violent tornadoes do hit.

2.1) Introduction

Tornadoes are one of the most destructive forces on Earth and present a substantial threat to both life and property. Each year approximately 1,200 tornadoes are reported in the United States and while the majority (~98.0%) cause no fatalities (SPC, 2017), a single tornado can result in a large number of fatalities (*e.g.*, Joplin, Missouri on 22 May 2011; Paul and Stimers, 2012). High-fatality tornadoes (hereafter any tornado causing 100 or more fatalities) are extremely rare with only 14 occurring since 1880 (Grazulis, 1993, 1997), and of those only three occurring after the advent of the first tornado forecasts in 1948 (Doswell *et al.*, 1999). While major improvements in tornado detection and warning dissemination systems, building technology, and general public awareness have dramatically reduced the likelihood of the occurrence of a high-fatality tornado, the 2011 tornado season proved that they can and still do occur (Paul and Stimers, 2012; Simmons and Sutter, 2012).

High-fatality tornadoes tend to occur when a violent tornado (rated [E]F4 or higher on the [Enhanced] Fujita Scale) hits a densely populated area, but this is a rare occurrence given that between 1995 and 2016 there were, on average, only seven violent tornadoes per year (SPC, 2017), and densely populated areas are relatively small targets (Ashley and Strader, 2016). Many studies have investigated these "worst-case scenarios," where violent tornadoes track through the central business district or urban core of a major city by transposing historical (Rae and Stefkovich, 2000; Hall and Ashley, 2008; Ashley *et al.*, 2014) or synthetic (Wurman *et al.*, 2007; Ashley *et al.*, 2014) tornado footprints over urban areas. Some studies have focused on specific external influences on these impacts including urban growth and expansion (Ashley *et al.*, 2014), and daily mobility patterns (Paulikas, 2015). While tornadoes have hit major cities in the past (*e.g.*, St. Louis, Missouri; Waco, Texas; Nashville, Tennessee; Galway, 1981; Grazulis, 1993;

Edwards and Schaefer, 2012), the odds of a rare, violent tornado hitting the urban core of a major city are minimal (due to its relatively small surface area) and, thus, "worst-case scenario" events are extremely unlikely to occur (Doswell *et al.*, 2012).

While direct hits have been historically rare, there appears to be an increase in the potential for high-impact tornadoes with the expansion of the urban environment (Ashley *et al.*, 2014; Ashley and Strader, 2016; Strader *et al.*, 2018), increased traffic along the interstate (Blair and Lunde, 2010), and the influence of climate change (Trapp *et al.*, 2007; Gensini and Ashley, 2011; Gensini *et al.*, 2014b; Gensini and Mote, 2015; Strader *et al.*, 2017a). The exact role that climate change may have on tornado risk (here defined as the probability of the occurrence of a tornado) is unclear. However, there have already been changes in the interannual variability of tornadoes (Brooks *et al.*, 2014; Elsner *et al.*, 2015) and many studies have projected that there will be increases in the number of days with environments favorable for tornado development (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2008; Gensini and Ashley, 2011; Gensini *et al.*, 2014b; Gensini and Mote, 2015) .

True tornado risk includes both direct hits and near-misses. The latter are tornadoes that come close to hitting densely populated areas or potentially tornadic storms that pass over a densely populated area without producing a tornado. The hazards literature defines a near-miss as an event that had some nontrivial probability of causing a disaster, but by chance, it did not (Dillon *et al.*, 2011). Examples would include a tornado that ends just before entering a populated area or a hurricane that suddenly curves away from the coast (Tinsley *et al.*, 2012). Near-misses are important because they could have caused a disaster and they can influence risk perception and future disaster preparedness (Dillon *et al.*, 2014).

Risk perception is a key component of public safety during a tornado. If a person does not believe themselves to be at risk, they are unlikely to seek shelter if a tornado warning is issued (Biddle, 1994; Ashley, 2007). Frequent false alarms due to near-misses can erode public confidence in the warning dissemination systems making people less likely to seek shelter (Barnes *et al.*, 2007; Brotzge *et al.*, 2011; Simmons and Sutter, 2011). Simmons and Sutter (2009) showed a direct causal link between the average tornado false-alarm rate of an area and the average rate of casualties. Paul (2012) found that 27% of respondents to a survey, given to tornado survivors after the 2011 Joplin, Missouri tornado, received a warning about the tornado but did not act because of how frequently the city sounded the tornado sirens. Frequent near-misses can also prompt the development of tornado myths and folklore that can lead people to assume they are safe and thus not seek shelter or seek shelter in the wrong location (Hoekstra *et al.*, 2011). Some examples of these myths include: when on the road, it is safest to seek shelter under an overpass (Hoffman, 2013); tornadoes will not cross rivers; tornadoes will not stay on the ground for many miles (Klockow *et al.*, 2014), and tornadoes will not hit large cities (Hoekstra *et al.*, 2011; Hoffman, 2013; Klockow *et al.*, 2014). Near-misses can influence disaster preparedness in two ways (1) if a person interprets a near-miss as a non-event they may underestimate their risk; (2) if a person interprets it as a lucky break they may prepare more for future events as if they had been hit (Tinsley *et al.*, 2012; Dillon *et al.*, 2014).

Studies on near-miss severe weather events often focus on the impacts of a singular event without formally defining the parameters of a near-miss. Prosser (1976) studied the unusual characteristics of a tornado which nearly hit Denver, Colorado on 18 May 1975 and Sherman-Morris (2010) studied sheltering behavior during a near-miss tornado at Mississippi State University on 10 January 2008. However, the authors are aware of no large-scale study on near-

miss tornadoes which provides a clear definition of a near-miss. One such study on *hurricanes* defined near-misses as hurricanes which were forecast to make landfall but did not, and used this definition as a part of an analysis on the accuracy of tropical cyclone forecasts in the Atlantic Ocean between 1976 and 2000 (Powell and Aberson, 2001). The authors propose to fill in this gap in the literature by developing a methodology to define near-miss violent tornadoes as a function of the population surrounding the tornado footprint, and to apply this definition to all violent tornadoes in the US between 1995 and 2016.

The purpose of this study is to determine the frequency with which violent tornadoes in the US are near-misses and to assess any spatiotemporal patterns that exist for near-miss violent tornadoes in the US. To answer these questions, we update the violent tornado climatology of Concannon *et al.* (2000) to extend from 1880 – 2016 and use replicates of historical violent tornado footprints and gridded representations of historical census data to determine possible exposure (here defined as the number of people residing in the footprint of the tornado) scenarios for violent tornadoes. We conclude this study by mapping out the high risk area for near-miss violent tornadoes.

2.2) Data and Methods

The track that a tornado takes is primarily dependent upon the atmospheric environment during its life cycle with small changes potentially shifting a track towards or away from a populated area (Kurdzo *et al.*, 2015). While the environment typically dictates tornado tracks, there is a certain level of randomness to when and where a tornado forms (Klockow *et al.*, 2014). As such, there are many potential tracks a tornado could take and since risk includes all events that did or could have happened (Kaplan and Garrick, 1981) it is of interest to know the potential distribution of tornado exposure if a tornado took a different, yet reasonable, track through the

area. To determine this potential distribution, we replicated historical tornado footprints throughout their respective surrounding areas and estimated the number of people exposed for each replicate.

The bounds of historical violent tornado activity between 1995 and 2016 closely match the four study regions used by Gensini and Ashley (2011) in their study on severe convective environment climatology. Since the regions are representative of environments favorable for violent tornado development, they were replicated for this study (Figure 2.1). In this study, we define a tornado track as a line from the point of touchdown for the tornado to the point of dissipation and a footprint as the total area experiencing tornadic winds. We started our analysis by updating the violent tornado climatology of Concannon *et al.* (2000) and conducting a sensitivity test to determine whether the error in the exposure estimates, introduced by using linear tornado footprints, could be minimized by careful selection of the census resolution.

2.2.1) Data and Data Accuracy

We collected population data from the US Census Bureau at multiple levels between 1880 and 2010 from the University of Minnesota's National Historical Geographic Information System (NHGIS; Manson *et al.*, 2017). We used county-level data from 1880 to 2010 for the violent tornado climatology and sensitivity test, tract and block group level data from 2000 for the sensitivity test, and block-level data from 1990 to 2010 for the sensitivity test as well as for our primary analysis.

We obtained historical violent tornado data for 1880 to 2016 from two sources: tornado reports from a long-term study of US tornadoes by Tom Grazulis, hereafter referred to as the Grazulis dataset (1880 – 1994; Grazulis, 1993, 1997), and tornado tracks from the US Storm Prediction Center's (SPC) SVRGIS database, hereafter referred to as the SPC dataset (1995 –

2016; SPC, 2017). The starting year was selected as 1880 as this is the period after which John Park Finley started collecting regular tornado reports (Brooks and Doswell, 2002; Ashley, 2007). These reports and tracks were collected from both amateur and professional observers and included information on the time and location, size and intensity, and the damage and casualties caused by each tornado. Many articles discuss the quality problems inherent in this data including errors in report accuracy and consistency (Doswell and Burgess, 1988; Grazulis, 1993), limits of using damage to assess tornado intensity (Verbout *et al.*, 2006), and changes in reporting methodology (Agee and Childs, 2014; Strader *et al.*, 2016).

Two of these quality problems are particularly relevant for this analysis: changes in width reporting and use of damage-based intensity ratings. In 1995, the Storm Prediction Center switched from reporting tornado widths as the mean width of the path to the maximum width (Brooks, 2004; Ashley *et al.*, 2014). This switch resulted in significantly larger widths after 1994 (Agee and Childs, 2014; Strader *et al.*, 2015) and introduced a source of error in estimating tornado exposure. Previous studies have resolved this inconsistency by using only tornadoes after 1995 (Ashley *et al.*, 2014; Strader *et al.*, 2015) or adjusting the widths before 1995 by adding the difference in the mean width between the two periods (Agee and Childs, 2014) or defining a standard width to use for all tornadoes based on the intensity (Ashley and Strader, 2016). We chose to employ the prior method and started our analysis in 1995 since the required block-level population data was only available nationally after 1990 (Ashley *et al.*, 2014).

Tornado intensity was estimated using the [Enhanced] Fujita ([E]F) scale, a scale that measures the degree of damage (DoD; *i.e.*, magnitude of damage to an object) a tornado causes to various damage indicators (DI; *e.g.*, mobile homes, trees) and relates it, empirically, to a wind speed. The scale was originally devised by T. Theodore Fujita and colleagues in the 1970s, after

the completion of many detailed tornado damage surveys, in reference to the damage caused to "well built" one- and two-family homes (Fujita, 1971; Fujita and Pearson, 1973; Abbey, 1976). Many flaws were noticed in the F scale including too few DIs, especially in rural areas (Doswell and Burgess, 1988; Doswell *et al.*, 2009), the maximum intensity rating was limited by the presence of DIs and the wind speed required to cause the maximum DoD (*e.g.*, double-wide manufactured homes have their maximum DoD at 134 mph while one- and two-family residents have their maximum DoD at 200 mph; McDonald and Mehta, 2006a; Edwards *et al.*, 2013), and according to wind engineers the maximum wind speeds for each intensity category were generally too high (McDonald *et al.*, 2003). These concerns led to the development and adoption of the Enhanced Fujita (EF) scale which provided 28 DIs a number of which are common in rural areas (*e.g.*, small barns, hardwood and softwood trees), as well as a more engineering-based understanding of the wind speeds required to cause damage/failure to various DIs (Edwards *et al.*, 2013). In the EF scale each DI has a maximum possible DoD, associated with an expected wind speed, so each DI has a maximum intensity rating associated with that wind speed that it can indicate (*e.g.*, double-wide manufactured homes are expected to be destroyed in 134 mph winds (EF2 range), so they cannot be used to indicate intensities over EF2).

The adoption of the EF scale helped alleviate some of the concerns with using the F scale to rate tornado intensity; however, for tornadoes that hit in rural areas, it is still very easy to either not hit any damage indicators or to not hit damage indicators with DoDs allowing for a rating of EF4-5. This implies that, historically, especially before the adoption of the EF scale in 2007, rural tornadoes may have been underrated and there may have been more violent tornadoes than the records currently indicate (Strader *et al.*, 2015). A good example of a tornado that was likely underrated was the tornado which hit El Reno, Oklahoma on 31 May 2013. This tornado

tracked over mostly rural areas and did not hit many DIs, resulting in a rating of only EF3. However, a mobile Doppler unit (RaXPol) recorded multiple wind speed measurements exceeding 100 m s^{-1} (maximum recorded wind speed was 135 m s^{-1}) which is in the EF5 range (Snyder and Bluestein, 2014). Despite these quality issues, this database is the best record of tornado occurrence currently available and is considered suitable for climatological studies (Verbout *et al.*, 2006; Ashley, 2007; Brooks *et al.*, 2014).

We obtained official tornado footprints from the US National Weather Service (NWS) offices in Norman (Oklahoma), Birmingham (Alabama), and Springfield (Missouri) for select tornadoes during the Great Plains Outbreak of 3 May 1999 (NWS, 1999) and the Super Outbreak of 27 April 27 2011 (NWS, 2011) as well as for the tornado that hit Joplin, Missouri on 22 May 2011 (NWS, 2017). The Great Plains Outbreak tornadoes were used in the sensitivity test while the others were used to compare the accuracy of synthetic versus observed tornado footprints.

2.2.2) Violent Tornado Climatology

Concannon *et al.* (2000) produced a spatial and temporal climatology of violent tornadoes for the period from 1921 to 1995 using the database of Grazulis (1993). We propose an update to the climatology to include the period from 1880 to 2016. To extend our period of record from 1880 to 2016 we combined the Grazulis and SPC datasets. Since the Grazulis dataset only had the name of the counties impacted by each tornado, we chose to follow the method of Concannon *et al.* (2000) and assigned each tornado (for both datasets) a location based on the coordinates of the centroid of the county where it touched down. Once each tornado has coordinates we create an 80 km by 80 km grid (corresponding to the Storm Prediction Center's practice of using 40 km as a proximity distance for severe weather events; Hitchens *et al.*, 2013) over the continental US and calculated the number of days where a violent tornado touched

down in each grid cell ("violent tornado days"). We used "tornado days" instead of actual

tornado reports to reduce the influence of changes in reporting frequency (Concannon *et al.*,

2000; Brooks *et al.*, 2003a). We are interested in changes in the violent tornado climatology over

time, so we break the record into three 30-year periods: 1880 - 1909, 1930 - 1959, 1987 - 2016

and calculated the mean number of violent tornado days per millennium. Since we are only

interested in large-scale patterns, we smooth the data using a Gaussian lowpass filter with a 120

km standard deviation following Concannon *et al.* (2000). We test for trends in both the number

of violent tornadoes and violent tornado days during this period using the nonparametric Mann-

Kendall test ($\alpha = 0.05$; Mann, 1945; Kendall, 1975). We chose Mann-Kendall as our data was

non-normal and Mann-Kendall is commonly used to test for trends in climate research (Fraedrich

et al., 2001; Yue *et al.*, 2002; Westra *et al.*, 2013; Sayemuzzaman and Jha, 2014). All subsequent

trend testing was done using the same Mann-Kendall test.

2.2.3) Sensitivity Test for Grid Resolution

Three main methodologies exist for creating tornado footprint polygons for exposure

analysis: (1) digitizing post-event damage surveys by government agencies, consulting

meteorologists and others (Ashley *et al.*, 2014), (2) digitizing observed (radar generated) or

modeled wind field data for a tornado (Wurman *et al.*, 2007; Strader *et al.*, 2015), and (3)

synthesizing tornado footprints by buffering a tornado track, in a GIS, to a distance

corresponding to the reported tornado width (Strader *et al.*, 2015). Of these three methods, the

most accurate is the first as observations constrain the footprints. However, post-event surveys

are mostly only available for recent high-end events (Ashley *et al.*, 2014; NWS, 2018). Radar-

based wind fields can be unrealistically large (Wurman *et al.*, 2007) due to the angle and height

at which the radar samples the tornado and can lead to overestimations of exposure (Ashley *et*

al., 2014; Strader *et al.*, 2015) and additionally, tornadoes rarely pass close enough to radar to obtain wind field data (Wurman *et al.*, 2007; Simmons and Sutter, 2011). While synthetic footprints are typically (unrealistically) linear and constant width they can be produced for most tornadoes in the Storm Events Database, and thus are frequently used in "worst-case scenario" work for tornado hazards (Wurman *et al.*, 2007; Ashley *et al.*, 2014; Strader *et al.*, 2016). Hence, we chose to buffer the collected tornado tracks in a GIS using the reported width to create synthetic footprints for all violent tornadoes between 1995 and 2016.

Since tornadoes frequently have curved tracks and their widths vary over the lifetime of the storm, it is of interest to note how accurately tornado exposure can be estimated using a synthetic footprint. It is also of interest to analyze the error in exposure (here defined as the difference between the exposures calculated for an observed and synthetic footprint) to determine if it can be minimized by carefully choosing the resolution of the census data to use in the exposure calculations. We required the finest resolution census data available (blocks) for our analysis because we were interested in the influence of small-scale variations in population on tornado exposure. However, it was unclear if census blocks provided the most accurate estimate of the population in a synthetic tornado footprint. To validate our choice, we conducted a sensitivity test to determine which census resolution yielded the smallest error in exposure.

In this study, we define tornado exposure as the number of persons residing in the footprint of the tornado based on data from the US Census Bureau. The census data is effectively a nighttime estimate of population, in a given area, as many people leave home for certain periods during the day to work or run errands (Paulikas, 2015; Ashley and Strader, 2016). We chose this measure instead of the more conservative housing units used by many studies (Ashley *et al.*, 2014; Strader *et al.*, 2015, 2016; Ashley and Strader, 2016) because we were interested in

relating the people potentially exposed to a tornado with potential fatalities (Merrell *et al.*, 2005; Simmons and Sutter, 2011). Tornado exposure was estimated using proportional allocation based on the intersection between a tornado footprint and the census population data. We allocated each section of the tornado a population, based on the proportion of the area of each census unit that is in that section of the tornado. For example, if the tornado footprint covered 70% of a census unit, we would assign 70% of that census unit's population to that segment of the tornado. The total population exposure for the tornado footprint was then the sum of the population in each segment (Ashley *et al.*, 2014).

To determine how changes in population distribution might influence the error in exposure between observed and synthetic tornado footprints, we replicate selected tornadoes from the 3 May 1999 Great Plains Outbreak over the entire study area. We paired each of the 28 selected observed tornado footprints with the corresponding synthetic footprints generated from the SPC SVRGIS database (see Section 2.2.1). We then created a 10 km resolution replication grid over the entire study area and replicated and shifted the paired footprints to the center of each grid cell maintaining the size and distribution of the tornadoes within the original outbreak (Figure 2.2). The tornado footprints were intersected in a GIS with the population data (2000 census) at each census level (county, tract, block group, and block) for each region in the study area. During the intersection procedure, we split the tornado footprint into many small segments, one for each census unit it crossed. We assigned each segment a population, based on the proportion of the original census unit it covered (proportional allocation; Ashley *et al.*, 2014). For each cell in the replication grid, we calculated the root mean square error (RMSE) of the exposure. For our RMSE calculation, we compare the exposure for the synthetic footprint, intersected with the population data for each census level, to the observed footprint, intersected

with only the block-level population data. We only use the block-level population data for the observed footprint since we want an exact measure of the people residing in the actual footprint. We also calculated the RMSE at the region and study area level. To determine which resolution performed best we selected the resolution that yielded the lowest RMSE value at each level (cell, region, and study area).

2.2.4) Exposure Distributions

To simulate many possible tornado footprints, we first created a uniform circular replication grid with a 40 km radius and a resolution of 0.5 km surrounding the centroid of the synthetic footprint. We define the extent of the area occupied by all tornado replicates as the potential impact zone. We then created replicates of the synthetic footprint (maintaining the size and orientation of the original) and shifted each one to the centroid of each of the 20,140 grid cells in the replication grid (Figure 2.2). We chose a 40 km radius following the Storm Prediction Center's practice of using 40 km as a proximity distance for severe weather events (Hitchens *et al.*, 2013). We used block-level population data in our analysis since we aim to assess small-scale changes in exposure and those required the finest resolution possible. Since census blocks vary in size and shape both in space and time, we followed the method of Ashley *et al.* (2014) to interpolate census blocks onto a fixed grid using proportional allocation. As in Ashley *et al.* (2014), we used a grid resolution which corresponded to the mean resolution of the census blocks located within the grid. We then linearly interpolated the population grids to the year of the tornado using the preceding and succeeding census data. Finally, we calculated exposure for each replicate as above by intersecting the tornado footprint and the population grid.

We subset our replicate exposures into those within 10 km (the upper range for mesocyclone size; Adlerman *et al.*, 1999) and 40 km (SPC's proximity radius; Hitchens *et al.*,

2013) of the original footprint and calculated summary statistics for the following variables:

footprint area (*AREA*), observed tornado exposure (*OBS*), median (*PPEMED*) and maximum

(*PPEMAX*) number of persons potentially exposed to the tornado (from the distribution of all

replicate exposures), and the probability that a tornado hitting within 10 km (40 km) of the

original footprint would exceed the observed exposure (*EPOBS*), 5,000 persons (*EP5K*), or

20,000 persons (*EP20K*). The probability of exceedance (*EP*) was calculated as follows:

$$EP = 1 - F(t),$$

where *t* is the threshold (*e.g.*, 5,000 persons), *F(t)* is the empirical cumulative distribution

function derived from the set of all replicate exposures within the specified radius

(10 km or 40 km), and is calculated as:

$$F(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[x_i \leq t]},$$

where *n* is the number of exposures within the specified radius and *1* is the indicator

function (1 if $x_i \leq t$ and 0 otherwise). The likelihood related to the *EP* is defined as very likely

($EP > 75\%$), likely ($50\% < EP \leq 75\%$), unlikely ($25\% < EP \leq 50\%$), or very unlikely

($EP \leq 25\%$).

Using the exposure distributions, we defined a hit as a tornado where $OBS \geq$

$5,000\ persons$, a near-miss where $OBS < 5,000\ persons$ and $PPEMAX \geq 5,000\ persons$

within 10 km, and a far-miss where $OBS < 5,000\ persons$, $PPEMAX < 5,000\ persons$ within

10 km and $PPEMAX \geq 5,000\ persons$ within 40 km. We chose 5,000 persons as the threshold

as Brooks *et al.* (2008) report a 0.1 to 1.9% fatality rate for select violent tornadoes and this

range was supported by fatality estimates from the 27 April 2011 Tuscaloosa-Birmingham,

Alabama (1.9%) and the 22 May 2011 Joplin, Missouri tornadoes (0.9%). These estimates were

derived by following the methodology of Brooks *et al.* (2008), using the number of homes

destroyed as reported by Prevatt *et al.* (2012) and assuming the national average of 2.64 persons

per home (U.S. Census Bureau, 2015). A threshold of 5,000 persons would yield an expected

fatality total of 5 to ~100 and only 10.4% of all violent tornadoes occurring between 1995 and

2016 had fatality totals exceeding five (SPC, 2017). We used a similar procedure as that used to

produce the violent tornado climatology to determine the high-risk areas for violent tornadoes

and for all tornadoes with observed or potential exposures exceeding 5,000 persons (hits or near-

misses) during the period of 1995 to 2016. We used the reported coordinates of tornado

touchdown to place each tornado on an 80 km by 80 km grid. We then proceeded to calculate the

mean number of days where a violent tornado touched down in each grid cell during the study

period. As in the violent tornado climatology we were only interested in large-scale trends so we

smoothed the data using a Gaussian lowpass filter with a standard deviation of 120 km

(Concannon *et al.*, 2000). We arbitrarily defined high-risk areas as all areas estimated to have

had at least two violent tornado days per century as the one day per century area covered most of

the middle part of the country. We tested for trends in the number of hits, near-misses, far-

misses, observed exposure, median and maximum potential exposure and the probability that a

tornado would impact at least 5,000 persons. We used a Quadrat Analysis with a Chi-Squared

Test ($\alpha = 0.05$; Griffith and Haining, 2006; Arnold *et al.*, 2017) to determine which regions had

more hits, near-misses and far-misses. We also used global (Moran, 1950; Legendre and

Legendre, 2012) and local Moran's I tests (Anselin, 1995) to determine the degree of spatial

autocorrelation in the tornado locations and the locations of clusters with reference to large

metropolitan statistical areas. To perform the Moran's I tests we first counted the number of

tornadoes in each county during the whole study period and then defined our neighborhood using the county boundaries with the standard "queen's case" contiguity rule (Greenbaum, 2002).

2.2.5) Comparing Exposures Using Synthetic and Observed Damage Paths

We selected ten violent historical tornadoes with both synthetic and observed footprints available (Table 2.1) and re-ran the replication analysis to calculate the observed exposure (*OBS*) and maximum replicate exposure (*PPEMAX*) within 10 km and 40 km for each footprint type for each tornado. We used these results to classify each footprint as a hit, near-miss, or far-miss to determine if the use of synthetic footprints affects the classification of the tornado.

2.3) Results

2.3.1) Violent Tornado Climatology

Risk occurs when a hazard (*i.e.*, tornado) and a vulnerable population/structure are collocated. An understanding of the violent tornado climatology is thus key in determining violent tornado risk (Dixon *et al.*, 2011; Coleman and Dixon, 2014; Ashley and Strader, 2016). A large body of literature has discussed tornado climatology (Abbey, 1976; Doswell and Burgess, 1988; Grazulis, 1993; Concannon *et al.*, 2000; Ashley, 2007; Doswell *et al.*, 2012; Brooks *et al.*, 2014; Strader *et al.*, 2015). We update the violent tornado climatologies of Concannon *et al.* (2000) and Doswell *et al.* (2012) to include the period from 1880 to 2016 during which there were 1,255 violent tornadoes reported in the US, occurring over 786 days (Grazulis, 1993, 1997; SPC, 2017). During this period violent tornadoes were most common in an 'L' shaped pattern between Iowa, Oklahoma, and Alabama, similar to the patterns found by Concannon *et al.* (2000) and Doswell *et al.* (2012) for the period of 1921 to 1995 (2010). The greatest number of mean violent tornado days (35 to 39 days per millennium) was found in central Oklahoma. The

1880 to 1909 period saw peak activity (35 to 39 days per millennium) in the central Great Plains while the 1930 to 1959 and 1987 to 2016 periods saw peak activity shift to the southern Great Plains (50+ days per millennium) and Southeast (35 to 39 days per millennium) respectively (Figure 2.3). While there was no significant temporal trend nationally for either annual violent tornado counts ($p = 0.57$) or tornado days ($p = 0.63$) between 1880 and 2016, we found a small but significant decrease in annual counts the Northern Plains ($p = 0.02$) and a corresponding increase in the Southeast ($p = 0.05$; Figure 2.4). With regards to tornado days only the increase in the Southeast was significant ($p = 0.04$). As in Concannon *et al.* (2000) and Ashley and Strader (2016) we found a high degree of interannual variability in violent tornado occurrence and location resulting in some periods where certain regions are more active than others. It is unclear if the long-term significant increasing (decreasing) trend for violent tornado counts in the Southeast (Northern Plains) is due to overall changes in tornado favorable environments or due to other nonmeteorological factors such as a short period of record or changes in reporting frequency over time (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2008, 2013; Ashley and Strader, 2016).

2.3.2) Population Distribution in Potential Impact Zones

The population distribution over the study area showed significant changes during the period of our study. The regions with the highest (Great Lakes) and lowest (Northern Plains) population densities remained the same. However, the mean population density (per census block) decreased over time as more people moved from rural to urban areas, as found by Ashley and Strader (2016).

We defined the potential impact zone for each violent tornado as the extent of the area surrounding all of its replicates. We found that the potential impact zones were largest in the

Southeast (Table 2.2), due to larger tornado footprints (Ashley, 2007; Strader *et al.*, 2015; Ashley and Strader, 2016). The greatest risk for violent tornadoes (defined here as the shortest recurrence interval) was in the Southern Plains and Southeast (Ashley, 2007; Ashley *et al.*, 2008; Dixon *et al.*, 2011; Strader *et al.*, 2015; Ashley and Strader, 2016), while the greatest population density was in the Great Lakes (Table 2.2). Relative to the mean population density in each region, the population density in each potential impact zone was very small, indicating that most violent tornadoes were hitting in less populated areas. The percent of the areas that were "urban" (here defined following one of the US Census Bureau's criteria for urban area classification; population density exceeds 386 persons km^{-2} (Ratcliffe *et al.*, 2016) or populated were generally greater in the potential impact zones than in the regions themselves. This is an interesting finding, but given that most of the developable area in the study area is populated (Table 2.2), it is not surprising that populated areas are more frequently hit (Ashley and Strader, 2016). The probable explanation is a reduction in the number of reported violent tornadoes in unpopulated areas due to a combination of the underreporting of tornadoes in unpopulated regions (Brooks *et al.*, 2003a; Simmons and Sutter, 2011; Elsner *et al.*, 2013a; Strader *et al.*, 2015) and the underrating of tornadoes in rural areas because of a lack of people/structures to impact (Doswell and Burgess, 1988; Doswell *et al.*, 2009; Strader *et al.*, 2015).

2.3.3) Sensitivity Test

At the grid cell-level, we found that no one census resolution significantly outperformed the others in terms of providing the lowest RMSE. The block-level resolution yielded the lowest RMSE in the most grid cells (greatest area) for the study area as a whole as well as for each individual region. However, the remaining census levels only performed marginally worse (Figure 2.5). In regions with higher population densities (Great Lakes and Southeast; Table 2.2),

31

the difference in performance was very low, while in areas with lower population densities (Northern and Southern Plains) the differences were greater but still not significant (Table 2.3). At the regional level, we found that the RMSE did not significantly vary by census level for each region, but the lowest RMSE was produced using the county-level resolution for all regions except the Northern Plains. Similar results were found at the level of the study area with the county-level resolution outperforming the others (Table 2.3).

The difference in the performance of the block-level resolution between the grid cell-level and the regional and study area levels is likely due to the mismatch between the observed and synthetic footprints. When the observed footprint closely matches the synthetic footprint the finest resolution (block-level) typically yields the lowest RMSE. This is not always the case when the footprints are mismatched. When the observed footprint curves significantly it can hit a nearby population center that the synthetic footprint missed yielding a large RMSE. These large differences can be most pronounced when block-level population data is used since coarser resolutions (census level) typically result in fewer variations in population over the length of the footprint. In aggregate, these large differences appear to result in larger RMSEs for the block-level than the county-level. Based on these results it is evident that tornado exposure is only marginally sensitive to the selection of the census resolution. Each resolution has its weaknesses and no one resolution works significantly better than the others.

2.3.4) Potential Exposure to Violent Tornadoes

Due to the lack of block-level census data prior to 1990 (Ashley *et al.*, 2014) and a shift in tornado width reporting in 1994 (Brooks, 2004; Agee and Childs, 2014; Strader *et al.*, 2015), our analysis was limited to the period of 1995 to 2016. This 22-year period is very small and thus is likely influenced by small-sample bias due to the rarity of violent tornadoes (Doswell, 2007;

Ashley and Strader, 2016). While the characteristics of individual tornadoes during this period might be biased (Doswell, 2007), the climatology during this period is in general agreement with the long-term climatology from 1880 to 2016 (Figure 2.3; Ashley, 2007; Doswell *et al.*, 2012; Ashley and Strader, 2016) implying that the spatial trends may be reliable. In spite of the potential bias, the data can still provide some information about violent tornadoes.

Between 1995 and 2016, there were 154 violent tornadoes in the United States with most of them occurring between the Rocky and Appalachian Mountains. The greatest number occurred in the Southern Plains, but the Northern Plains and Southeast also had significant numbers of violent tornadoes (differences between regions were significant based on a Pearson Chi-Square test; $p = 0.002$; Figure 2.6a). The high-risk area for violent tornadoes (area expected to have at least two violent tornado days per century) covered approximately 483,788 km^2 and extended in a broken 'L' shape between Iowa, Oklahoma and Alabama (Figure 2.7) following the pattern found by Concannon *et al.* (2000) and Doswell *et al.* (2012). These violent tornadoes had a median area, observed exposure, potential exposure (within 10 km), and probability of impacting 5,000 persons or more (within 10 km) of 22.9 km^2, 564 persons, 311 persons, and 30.3% respectively. As found by Ashley (2007) and Ashley and Strader (2016), the tornadoes were largest and had the greatest observed and potential exposure in the Southeast (Table 2.4), due to greater rural population densities (Table 2.2) and larger damage areas. Observed and potential exposure were lower in the Northern Plains due to lower rural population densities.

Within 10 km (40 km) of the original footprint 33.1% (57.8%) of all violent tornadoes between 1995 and 2016 had a potential exposure of at least 5,000 persons while only 8.4% (24.7%) had an exposure of at least 20,000 persons. The high-risk area for violent tornadoes with observed or potential exposures of at least 5,000 persons (within 10 km) covered approximately

35,663 km² and was located primarily in central Oklahoma and northern Alabama (Figure 2.7). This area matches where major metropolitan areas (Oklahoma City, OK and Birmingham, AL) meet the areas with the greatest risk of violent tornadoes (Concannon *et al.*, 2000; Ashley, 2007; Doswell *et al.*, 2012; Ashley and Strader, 2016).

Approximately 10.4% of the violent tornadoes were likely to very likely to be hits $EP5K > 50\%$ within 10 km) with all of these occurring in the Southern Plains and Southeast (Figure 2.8). There were no significant temporal trends in median annual observed ($p = 0.61$), maximum ($p = 1.00$) or median ($p = 1.00$) potential exposure (within 10 km) nor in the median annual probability of impacting 5,000 persons or more (within 10 km; $p = 0.54$) between 1995 and 2016.

2.3.5) Near-Misses and Far-Misses

Near-misses (far-misses) were defined as violent tornadoes where $OBS < 5,000\ persons$ and $PPEMAX \geq 5,000\ persons$ within 10 km (where $OBS < 5,000\ persons$, $PPEMAX < 5,000\ persons$ within 10 km and $PPEMAX \geq 5,000\ persons$ within 40 km). Figure 2.9 shows an example of a near-miss to the City of Norman, Oklahoma on 10 May 2010 (b) and a far-miss for the Cities of Canton and Pekin in Illinois on 13 May 1995 (c). Between 1995 and 2016 there were 30 near-misses and 38 far-misses. Near-misses and far-misses occurred in all regions with their distribution being similar to the distribution of violent tornadoes in general. Near-misses (far-misses) occurred most frequently in the Southern Plains (Southeast) but the regional differences were not significant ($p = 0.09$ and $p = 0.13$ respectively; Figure 2.6c-d). The median maximum potential exposure was similar between near-misses and far-misses (9,415 and 9,234 persons respectively) with the highest value for near-misses (far-misses) in the Southern Plains (Great Lakes; Table 2.5). The median probability of

being a hit was 8.1% (1.9%) indicating that it is very unlikely that any near-misses or far-misses could have been hits. In fact, no near-miss was likely to very likely to have been a hit (Table 2.6). However, the median probability of impacting more persons than was observed was 47.8% indicating it was nearly likely that near-misses could have impacted more persons (Table 2.5). There was no significant temporal trend found for either near-misses ($p = 0.38$) or far-misses ($p = 0.60$) between 1995 and 2016 for the entire US or for any subregion.

2.3.6) Hits

Hits were defined as violent tornadoes where $OBS \geq 5,000\ persons$. There were 21 reported hits, between 1995 and 2016, occurring in all regions, except the Great Lakes. They were most common in the Southern Plains and Southeast and rare in the Northern Plains (regional differences were significant; $p = 0.008$; Figure 2.6b). These locations match the findings of Ashley and Strader (2016) who found the high rural population density in the Southeast combined with the high risk of tornadoes resulted in significant potential impacts. The Northern Plains have high tornado risk but low population density while the Great Lakes have high population density but low tornado risk (Ashley and Strader, 2016). Figure 2.9d shows an example of a hit on the City of Cleveland, Tennessee on 27 April 2011. Hits had a median observed exposure of 8,848 persons with the highest value in the Southern Plains and lowest in the Northern Plains (Table 2.5). They had a median maximum potential exposure (within 10 km) of 19,534 persons with the greatest maximum potential exposure in the Southern Plains. Hits that occurred in the Northern Plains were unlikely to be hits while hits in the Southern Plains and Southeast were very likely to be hits. All but five of the hits were likely to very likely to be hits (Table 2.7). Hits in the Southeast were likely to have impacted more persons than they did while they were very unlikely to do so in the Northern Plains (Table 2.5). Similarly to the near-misses

and far-misses, we found no significant temporal trend for hits ($p = 0.92$) between 1995 and 2016 for the entire US or for any subregion.

2.3.7) Characteristics of Select Violent Tornadoes

A total of 13 violent tornadoes had a maximum potential exposure (within 10 km) exceeding 20,000 persons; of these ten were hits and three were near-misses. Three of these were likely to very likely to have had an exposure of at least 20,000 persons and another seven were very unlikely to have had such an exposure. Only five were likely to impact more people than was observed (Tables 2.6 and 2.7).

2.3.8) Comparison of Synthetic and Observed Damage Paths

Tornado footprints come in all shapes and sizes (Wurman *et al.*, 2007; Ashley *et al.*, 2014; Strader *et al.*, 2015) with some taking a relatively straight track (Ashley *et al.*, 2014), others curving significantly (Paul and Stimers, 2012), some even moving in a loop (Wurman *et al.*, 2014). The width of the footprint can also change significantly throughout the life of the tornado as it weakens or strengthens (Burgess *et al.*, 2014). Wind speed also varies throughout the tornado footprint with only a small fraction of the total area experiencing EF4-5 wind speeds and damage (Wurman *et al.*, 2007; Ashley *et al.*, 2014; Strader *et al.*, 2015). These variations in footprint shape and area can result in large departures from the linear synthetic footprint produced by buffering the SPC tornado tracks (SPC, 2017). To test if these variations affected the classification of a violent tornado as a hit, near-miss or far-miss, we ran our replication analysis on the observed footprints of ten violent tornadoes and compared the exposures to those from the matching synthetic footprints. We found substantial differences between synthetic and observed tornado footprints in terms of both observed exposure and maximum potential exposure (within 10 km). The largest difference in observed exposure was 16,062 persons for Hackleburg-

Phil Campbell, AL while the smallest difference was 167 persons for Dover, OK. The largest difference in maximum potential exposure was 65,042 persons for Tuscaloosa-Birmingham while the smallest difference was 167 persons for Cimarron City-Mulhall, OK. These differences in exposure arose from both curvature in the observed footprints and differences in footprint area between observed and synthetic footprints. The latter finding was in agreement with a study by Strader *et al.* (2015) that found a mean overestimation of significant tornado footprint area of 39% for synthetic footprints between 1995 and 2013. While most of the synthetic and observed tornado footprints were classified the same (*e.g.*, near-miss), three (Joplin, MO; Cordova, AL; and Shoal Creek-Ohatchee-Argo, AL) had large enough differences to result in different classifications. All three of the misclassifications were the result of tornado footprints with significant curvature. The Joplin tornado curved into a densely populated area resulting in a hit while the others curved away from densely populated areas resulting in near-misses (Table 2.1).

2.3.9) Spatial Autocorrelation of Violent Tornadoes

Violent tornadoes typically form as a result of the presence of key ingredients in the atmosphere: low-level moisture, increases in wind speed with height, rapid change in temperature with height and the presence of a thunderstorm. It is rare to get all of these ingredients together to produce violent tornadoes (Doswell *et al.*, 2012); however, there are certain regions where these ingredients are more common (*e.g.*, "Tornado Alley" and "Dixie Alley"; Dixon *et al.*, 2011; Gensini and Ashley, 2011; Ashley and Strader, 2016). Due to the spatial dependence of tornado-favorable environments, there is spatial clustering (spatial autocorrelation) in the tornado climatology. Spatial clustering also exists in population data with a significant proportions of the population living in clustered urban areas (Ashley *et al.*, 2014; Ashley and Strader, 2016). We tested for spatial autocorrelation in all tornadoes, near-misses,

far-misses, and hits using a global and local Moran's I test. For the global Moran's test we found positive spatial autocorrelation (clustering) in all cases ($p \leq 0.05$) as expected. At the local level, for all violent tornadoes, we found significant clustering scattered throughout the traditional "Tornado Alley" and "Dixie Alley" (Dixon *et al.*, 2011; Gensini and Ashley, 2011). For near-misses, far-misses, and hits we found significant clustering in or near large metropolitan statistical areas (population of 500,000 persons or more; Figure 2.10). This was also expected as, by definition, near-misses, far-misses, and hits require significant populations living within the potential impact zone.

2.4) Discussion

The risk of tornado exposure is typically measured as a function of the number of tornadoes hitting an area during a specified time period (Boruff *et al.*, 2003; Ashley *et al.*, 2014; Strader *et al.*, 2016). These risk assessments rarely include tallies of tornadoes which narrowly missed a populated area, however, near-misses are equally likely events and thus are an important part of the true exposure risk (Dillon *et al.*, 2011, 2014; Tinsley *et al.*, 2012). In addition to impacting exposure risk, near-misses can also influence vulnerability via their effect on risk perception and shelter-seeking behavior (Dillon *et al.*, 2014). This study represents a first attempt to determine the frequency of near-misses for violent tornadoes by replicating and translating each original tornado footprint across the area surrounding the potential impact zone. This method allows us to consider scenarios where a tornado struck a more or less populated area nearby the original footprint. We chose this methodology since it enables us to test many possible exposure scenarios throughout the area surrounding the potential impact zone and also because it has the advantage of ease of use.

We first updated the US violent tornado climatology of Concannon *et al.* (2000) to include the period of 1880 to 2016. We found that the general pattern over the 137 year period was the same (Figure 2.3) as has been found by others (Concannon *et al.*, 2000; Ashley, 2007; Doswell *et al.*, 2012; Ashley and Strader, 2016). However, there was a small but statistically significant increase (decrease) in the number of violent tornadoes in the Southeast (Northern Plains; Figure 2.4) over the period as found by Ashley and Strader (2016). The climatology shows that the general pattern has not changed (with regular peaks in the Southern Plains and Southeast), however, during some periods certain regions are more active than others (Concannon *et al.*, 2000; Doswell *et al.*, 2012; Ashley and Strader, 2016). It is unclear if the increasing trend for violent tornadoes in the Southeast is due to nonmeteorological factors, such as small-sample bias (Doswell, 2007; Ashley and Strader, 2016) or population bias (Brooks *et al.*, 2003a; Simmons and Sutter, 2011; Elsner *et al.*, 2013a; Strader *et al.*, 2015) or climate change (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2008; Gensini and Ashley, 2011; Gensini *et al.*, 2014a; Gensini and Mote, 2015). If this increase in violent tornado activity in the Southeast is due to climate change it is a major concern given that the population growth in the Southeast has been rapid (Ashley, 2007; Ashley and Strader, 2016) and the highly-vulnerable mobile/manufactured home market continues to grow there (Merrell *et al.*, 2002; Ashley, 2007).

We tested the sensitivity of the error in exposure estimates, between synthetic and observed tornado footprints, to the selection of census level. Most studies which estimate tornado exposure tend to ignore the error generated by using synthetic tornado footprints since the focus is on scenario work and exact historical exposure values are unnecessary (Wurman *et al.*, 2007; Ashley *et al.*, 2014; Ashley and Strader, 2016). This error can become important for historical estimates of tornado exposure, however, since it is possible for exposure estimates to be highly

inaccurate if the observed footprint curves into or away from a densely populated area (*e.g.*, Joplin, Missouri 22 May 2011; Paul and Stimers, 2012). Our findings that the error between synthetic and observed footprints is large, is not surprising as tornado widths change over the lifetime of the tornado and tracks frequently curve (Strader *et al.*, 2015). It is noteworthy that the differences in error between the different census levels for each region, as well as for the whole study area, are relatively small indicating a lack of sensitivity to the selection. The finding that the block-level census data performs best at the replication level while the county-level performs best in aggregate is interesting as it shows that the error generated when a curved tornado footprint travels through a densely populated area is large enough to overcome the local inaccuracies in county-level data. It is also noteworthy that the error is sensitive to the regional population with more densely populated areas (*e.g.*, Great Lakes) having a much higher error than more sparsely populated areas (*e.g.*, Northern Plains). The overall findings of the sensitivity test indicate that the error in exposure (relative to a nonlinear, width-changing footprint) cannot be minimized through the selection of the census level. This implies that if a study doesn't require fine-resolution census data (Ashley *et al.*, 2014; Ashley and Strader, 2016; Strader *et al.*, 2016, 2017b), it is reasonable to use county-level data (Boruff *et al.*, 2003; Merrell *et al.*, 2005; Simmons and Sutter, 2011, 2012). This is highly relevant since it allows the use of county-level census data to study historical tornado exposure going back to the beginning of the tornado record in the 1880s (Grazulis, 1993; Ashley, 2007).

Many studies have shown that tornadoes rarely hit densely populated areas due to both the rarity of violent tornadoes and the small amount of developed area that exists in the US (Rae and Stefkovich, 2000; Wurman *et al.*, 2007; Ashley and Strader, 2016; Strader *et al.*, 2016, 2017b, 2018). Our findings that violent tornadoes typically hit sparsely populated areas with

40

median observed and potential exposures under 1,000 persons (in all regions except the Southeast; Table 2.6) were thus unsurprising. Most tornado-prone regions (Concannon *et al.*, 2000; Doswell *et al.*, 2012) have low population densities (Table 2.2), and thus low exposure, while the Southeast is the exception with greater rural population densities, longer tornado tracks and more fatalities (Ashley, 2007; Strader *et al.*, 2015; Ashley and Strader, 2016). While densely populated areas are rarely hit, contrary to folklore, cities are no safer from tornadoes than rural areas (Hoekstra *et al.*, 2011; Klockow *et al.*, 2014). In fact, there have been many tornadoes that have even hit the downtown areas (central business districts) of major cities (Edwards and Schaefer, 2012).

We found that only 33.1% of all violent tornadoes had observed or potential exposures of at least 5,000 persons (within 10 km; were near-misses or hits) and the high-risk area for such tornadoes only covered 7.4% of the total high-risk area (area expected to have at least two violent tornado days per century). The high-risk area was primarily located in central Oklahoma and northern Alabama matching the areas of peak violent tornado activity (Concannon *et al.*, 2000; Doswell *et al.*, 2012) and killer tornado activity (Ashley, 2007), respectively. It was interesting that no near-miss was likely to be a hit (Table 2.6) while several hits were very unlikely to have been hits (Table 2.7). This is likely because the near-miss definition only required the maximum potential exposure to be 5,000 persons. We chose to use a maximum potential exposure due to the small sample size of violent tornadoes, but future work looking at near-misses for all tornadoes could include a more stringent definition by, for example, using a median potential exposure of 5,000 persons. If we had used such a measure for our small sample we would have found no near-misses.

The primary goal of this study was to understand the spatiotemporal patterns of near-miss violent tornadoes and their relation to population distributions in the area surrounding the potential impact zone. We found that the likelihood of hits, near-misses and far-misses were a function of the underlying population distribution (Table 2.2) as well as the climatology of violent tornadoes (Figure 2.6). The spatial distribution of near-misses, far-misses and hits was similar to the general distribution of all violent tornadoes with peak occurrence in the Southern Plains and Southeast (Concannon *et al.*, 2000; Ashley, 2007; Doswell *et al.*, 2012; Ashley and Strader, 2016). The differences in the spatial patterns between the three types of violent tornadoes is primarily evident in the differences in population density. The results indicate that hits were relatively more common in the Southeast where the rural population density is higher and less common in the less populated areas (*e.g.*, Northern Plains; Table 2.2; Ashley and Strader, 2016). The results also indicate that near and far-misses were relatively highest in the Great Lakes, likely due to the expansion of the urban areas into the countryside as well as the rarity of violent tornadoes in this region (Ashley and Strader, 2016). We also found significant spatial autocorrelation (clustering) for counties with hits, near-misses and far-misses near large metropolitan areas (Figure 2.10). The clustering implies that the risk for near-misses is low outside large urban areas. As urban expansion continues, the vulnerable areas will likely see an increase in population making it more likely that tornadoes will pass close by or hit densely populated areas (Ashley *et al.*, 2014; Ashley and Strader, 2016; Strader *et al.*, 2017b, 2017a).

The methodology used in this study is a first attempt at determining the frequency of near-miss violent tornadoes and as such we employed a definition that made use of existing tornado footprint data. The definition appears reasonable as small shifts in environmental conditions have been known to shift the footprints of violent tornadoes (Bluestein *et al.*, 2015).

Our study was limited to a 22-year period due to a lack of block-level census data prior to 1990 and a change in tornado width reporting in 1994. This short period is likely subject to small-sample bias due to the extreme rarity of violent tornadoes (Doswell, 2007; Ashley and Strader, 2016). However, the distribution of violent tornadoes was found to match the long-term climatology (Figure 2.3; Concannon *et al.*, 2000; Doswell *et al.*, 2012) implying the distribution observed was reasonable. Likewise, the high-risk area for near-misses and hits (Figure 2.7) was located where the highest risk for violent tornadoes met large metropolitan areas in central Oklahoma (Oklahoma City) and northern Alabama (Birmingham-Hoover). Given that hits and near-misses require large populations, by definition, this distribution also makes sense (Ashley and Strader, 2016; Strader *et al.*, 2017b).

Near-misses do not only result from tornadoes which dissipated before, shifted away from or just narrowly missed a densely populated area. Some cyclic tornadic supercells pass over populated areas without producing a tornado (*e.g.*, Nashville, Tennessee on 5 February 2008; Murphy and Knupp, 2013) resulting in a near-miss. Future work on this topic could include creating synthetic tornado tracks and translating them along and about the track of tornadic supercells (identified via radar; Trapp *et al.*, 2005) to determine the likelihood of near-misses for tornadoes that did not happen.

The use of synthetic tornado footprints to estimate exposure also creates a source of error for this analysis as true tornado footprints often curve and change strengths and widths over the life of the tornado (Paul and Stimers, 2012; Strader *et al.*, 2015). The error introduced by using synthetic footprints can be significant. As an example, the official footprint from the NWS for the EF5 tornado which hit Joplin, MO on 22 May 2011 had an exposure of 17,292 persons while the synthetic footprint from the SPC only had an exposure of 4,474 persons. The reason for the

difference is that the SPC used a linear footprint which passed south of the highest population areas in Joplin. Conversely the NWS footprint for the EF4 tornado which hit Tuscaloosa-Birmingham, AL on 27 April 2011 had an exposure of only 18,690 persons while the SPC footprint had an exposure of 39,231 persons. As a result of errors such as these three of the ten tornadoes tested (Joplin, MO; Cordova, AL; and Shoal Creek-Ohatchee-Argo, AL) were misclassified as near-miss, hit and hit respectively (Table 2.3). Since the sensitivity test showed that changing the census resolution used in the analysis had minimal effect on the accuracy of the exposures using the SPC footprints this represents a distinct limitation of this methodology. However, we believe our methodology justified given that 70% of the tested tornadoes were classified correctly (Table 2.3).

Since we found no temporal trend in near-misses, far-misses or hits, since 1995, it is unclear if either climate change or urban expansion have influenced their occurrence but it seems likely that urban expansion has had at least some effect since various scenario studies have showed exposures increasing over time (Ashley *et al.*, 2014; Ashley and Strader, 2016). We limited our analysis to violent tornadoes because they cause the most fatalities (Ashley, 2007), but it would be of interest to re-run this analysis on all historical US tornadoes since 1995 to see if a temporal trend could be found with a larger sample size. Additionally, by limiting our analysis to only those tornadoes that were classified as violent we could have missed many tornadoes that might have had winds on the level of a violent tornado but didn't hit any damage indicators capable of receiving [E]F4/5-level damage (*e.g.*, El Reno, OK on 31 May 2013; Snyder and Bluestein, 2014). Our findings that the percentage of urban and populated areas were generally higher, in the potential impact zone than in the surrounding region, could be an indicator of an underrating problem for tornadoes in unpopulated areas (Table 2.2). While it is

not possible to know exactly how many violent tornadoes were underrated due to a lack of damage indicators (Doswell and Burgess, 1988; Doswell *et al.*, 2009; Edwards *et al.*, 2013; Strader *et al.*, 2015), including all, or at least all significant ([E]F2+-level), tornadoes would allow for a much better picture of the risk strong tornadoes pose to populated areas (Ashley *et al.*, 2014; Ashley and Strader, 2016; Strader *et al.*, 2017b, 2018).

2.5) Conclusions

This study has represented a first large-scale attempt to determine the frequency of near-miss violent tornadoes by replicating and translating each original tornado footprint across the area surrounding the potential impact zone. Not surprisingly, we found that tornadoes tended to hit in less populated areas and that hits, near-misses, and far-misses had spatial distributions that matched the violent tornado climatology. The primary difference we noted in the distributions was related to rural population density with locations with higher rural population densities (*e.g.*, Southeast) favoring hits and areas with lower rural population densities (*e.g.*, Southern Plains) favoring near-misses. Our analysis also found that the error introduced by using synthetic tornado footprints is not sensitive to the selection of the level of census data used. This finding is important because (1) it allows a user to select the census level that best fits their tornado hazard assessment and (2) it enables analysis of tornadoes going back to the late 1800s (Grazulis, 1993, 1997) when county-level census data was the only census data available.

Emergency managers require an in-depth understanding of tornado hazard risk to help mitigate tornado disasters, but they often ignore near-miss tornadoes when assessing risk due to a lack of direct impacts. Near-misses are important because in addition to representing realistic outcomes they can also influence risk perception and sheltering behavior (Dillon *et al.*, 2011, 2014; Tinsley *et al.*, 2012). Tornado warnings do not always reach the entire population at risk,

due to factors such as language barriers or lack of television/radio/tv/internet access (Brotzge and Donner, 2013). However, often the warning is received and not heeded (Sherman-Morris, 2010; Paul and Stimers, 2012) due to a lack of personalization of the risk. Studies have shown that frequent false alarms, due to near-misses, can desensitize the public to tornado risk and reduce the likelihood of response to a warning (Barnes *et al.*, 2007; Simmons and Sutter, 2009, 2011; Brotzge *et al.*, 2011; Paul and Stimers, 2012). Frequent near-misses can also prompt the development of tornado folklore that can lead people to assume they are safe and thus not seek shelter or seek shelter in the wrong location (Hoekstra *et al.*, 2011; Klockow *et al.*, 2014). Knowledge about the location of frequent near-misses can help emergency managers and risk communicators target communities that might be more vulnerable, due to an underestimation of tornado risk, for educational campaigns (Brotzge and Donner, 2013). Expert-led town halls could be conducted to combat tornado myths and better explain the true nature of local tornado risk (Stewart *et al.*, 2018). By increasing educational efforts in these high-risk areas, it might be possible to improve local knowledge and reduce casualties when violent tornadoes do hit.

Table 2.1. Comparison of exposure, potential exposure, and type between synthetic and observed tornado footprints for select violent tornadoes between 1995 and 2016. Variables include the date (DATE) and location (LOCATION) of the tornado, its exposure (OBS), and maximum potential exposure within the specified radius (PPEMAX) and whether the tornado is a near-miss, far-miss, or hit (TYPE).

| | | Synthetic Footprint | | | | Observed Footprint | | | |
| | | | *r=10 km* | *r=40 km* | | | *r=10 km* | *r=40 km* | |
DATE	LOCATION	OBS	PPEMAX	PPEMAX	TYPE	OBS	PPEMAX	PPEMAX	TYPE
3 May 1999	Bridge Creek-Moore, OK	23,649	37,617	47,503	hit	10,057	16,810	42,282	hit
3 May 1999	Cimarron City-Mulhall, OK	1,209	3,749	38,478	far-miss	440	3,582	39,048	far-miss
3 May 1999	Dover, OK	62	1,783	6,112	far-miss	229	1,026	5,832	far-miss
27 Apr 2011	Cullman, AL	4,162	4,825	16,255	far-miss	2,662	3,666	11,313	far-miss
27 Apr 2011	Hackleburg-Phil Camp., AL	22,784	63,522	63,645	hit	6,722	20,303	33,835	hit
27 Apr 2011	Cordova, AL	5,712	10,413	40,943	hit	2,445	5,032	46,434	near-miss
27 Apr 2011	Tuscaloosa-Birmingham, AL	39,231	102,942	114,368	hit	18,690	37,900	46,454	hit
27 Apr 2011	Shoal Cr.-Ohatchee-Argo, AL	7,887	17,661	57,974	hit	2,278	7,766	12,546	near-miss
27 Apr 2011	Lake Martin, AL	1,199	3,205	17,455	far-miss	500	1,392	9,662	far-miss
22 May 2011	Joplin, MO	4,474	14,449	14,579	near-miss	17,292	18,119	18,119	hit

47

Table 2.2. Population distributions and violent tornado recurrence intervals for each region in the study area. For the potential impact zones the data was interpolated to a grid with a resolution equivalent to the mean census block size in each impact zone while for the other tables the data is in the original census blocks for each decennial census (1990 - 2010). Variables include: region name (REGION), area of impact zone or total land area for census blocks (km^2; AREA), percent of blocks or grid cells with a population density exceeding 386 persons km^{-2} (PCTURB; one definition of an urban area; Ratcliffe *et al.*, 2016), percent of blocks or grid cells which are populated (PCTPOP), mean grid cell or census block population density (persons km^{-2}; POPDEN), and number of violent tornadoes (NTOR) and recurrence interval for the period 1995 to 2016 (months; RECUR).

Potential Impact Zones						
REGION	**AREA**	**PCTURB**	**PCTPOP**	**POPDEN**	**NTOR**	**RECUR**
GL	14,697.30	7.5	96.6	125.3	11	24
NP	15,743.40	1.2	96.5	19.3	41	6.4
SE	20,750.90	2.2	96.3	41.2	44	6
SP	18,046.80	2.6	94.9	43.8	58	4.6
1990 Census Blocks						
REGION	**AREA**	**PCTURB**	**PCTPOP**	**POPDEN**		
GL	774,146.90	4.1	87	2,014.00		
NP	1,324,138.80	0.5	76.2	504.4		
SE	1,002,375.10	2	82.8	699.6		
SP	1,390,897.10	1.2	78	769.1		
2000 Census Blocks						
REGION	**AREA**	**PCTURB**	**PCTPOP**	**POPDEN**		
GL	773,830.40	4.5	87.5	1,793.90		
NP	1,322,642.40	0.6	74.1	523.2		
SE	1,001,832.10	2.4	84.1	613.9		
SP	1,385,149.30	1.3	70.1	740.4		
2010 Census Blocks						
REGION	**AREA**	**PCTURB**	**PCTPOP**	**POPDEN**		
GL	773,680.30	4.9	86	1,442.50		
NP	1,322,228.10	0.6	71.1	453.5		
SE	1,001,024.50	2.9	82.4	510.4		
SP	1,385,004.20	1.5	67.5	623.6		

Table 2.3. Results of sensitivity test for selection of the best census level. At the grid cell-level, each grid cell is assigned a best census level which corresponds to the census level yielding the lowest root mean square error (RMSE) between the synthetic and observed population exposures. The overall best census level to use for each region (and for the while study area) corresponds to the census level which is best in the greatest number of grid cells. At the regional level, the RMSE is calculated for each census level for the entire region (study area) and the best census level corresponds to the census level with the lowest RMSE for the region.

Grid Cell Level					
Region	**Grid Cell Count**				**Best Level**
	County	Tract	Block Group	Block	
GL	104	95	79	110	Block
NP	90	138	135	197	Block
SE	117	114	106	127	Block
SP	107	150	136	167	Block
Study Area	*418*	*497*	*456*	*601*	*Block*

Regional Level					
Region	**Grid Cell Count**				**Best Level**
	County	Tract	Block Group	Block	
GL	2,006.40	2,158.30	2,194.50	2,230.30	County
NP	687.6	596.7	636.9	650.5	Tract
SE	987.7	1,130.40	1,163.90	1,239.60	County
SP	870.4	916.5	968.8	1,011.50	County
Study Area	*1,170.10*	*1,246.70*	*1,282.90*	*1,322.40*	*County*

Table 2.4. Characteristics of violent tornadoes and the regions they strike in the United States between 1995 and 2016. Tornado footprint area (AREA) refers to the area of the tornado footprint polygon (km^2). Persons Exposed (OBS) refers to the residential population (based on the US Census data at the time of each tornado) within the tornado footprint. Median Persons Potentially Exposed (PPEMED) refers to the median value of the residential population within the footprint of all replicate tornadoes within the specified distance radius. Probability that a replicate tornado within the specified distance from the original footprint will have a potential exposure of at least 5,000 persons (EP5K) is also shown. Statistics included are minimum (MIN), median (MED), interquartile range (IQR), and maximum (MAX) values for all tornadoes within each region.

	AREA				OBS			
REGION	**MIN**	**MED**	**IQR**	**MAX**	**MIN**	**MED**	**IQR**	**MAX**
GL	1.1	7.8	19.4	68.4	41	993	1,375	3,907
NP	0.4	16.1	36.7	349.6	0	112	364	5,815
SE	0.1	39.8	49.4	427.3	4	1,232	3,760	39,231
SP	0.9	28.2	39.2	675.2	0	782	2,955	24,130
US	*0.1*	*22.9*	*48.9*	*675.2*	*0*	*564*	*2,441*	*39,231*

	PPEMED							
	r = 10 km				r = 40 km			
REGION	**MIN**	**MED**	**IQR**	**MAX**	**MIN**	**MED**	**IQR**	**MAX**
GL	59	554	1,068	3,166	39	404	821	3,027
NP	0	40	261	4,408	0	52	231	3,116
SE	2	967	3,258	41,949	1	659	2,110	17,807
SP	1	346	2,496	23,508	0	215	1,547	8,415
US	*0*	*311*	*1,891*	*41,949*	*0*	*251*	*1,206*	*17,807*

	EP5K							
	r = 10 km				r = 40 km			
REGION	**MIN**	**MED**	**IQR**	**MAX**	**MIN**	**MED**	**IQR**	**MAX**
GL	0.2	2.9	7.5	17.4	0.1	6.2	4.7	34.1
NP	2.3	6	17.1	42	0	2	10.7	36.7
SE	4.1	49.5	51.8	100	0.4	12.4	25.2	100
SP	1.1	39.8	60.8	100	0	14.3	34.9	62.8
US	*0.2*	*30.3*	*57.2*	*100*	*0*	*9.2*	*26.2*	*100*

Table 2.5. Characteristics of violent tornadoes by type and region. Characteristics include the median value for each region of persons exposed (OBS), tornado footprint area (AREA) in km^2, maximum persons potentially exposed (PPEMAX), probability (%) that the persons potentially exposed exceeds the persons exposed (EPOBS), 5,000 persons (EP5K), or 20,000 persons (EP20K) within the specified distance radius (r). * denotes sample size for probability calculation was less than 3.

Far-Misses

				r = 10 km				r = 40 km			
REGION	**n**	**OBS**	**AREA**	**PPEMAX**	**EPOBS**	**EP5K**	**EP20K**	**PPEMAX**	**EPOBS**	**EP5K**	**EP20K**
GL	4	1,100	10.3	4,226	38	-	-	15,488	43.3	5.9	1.4*
NP	12	206	17.7	1,433	27	-	-	8,872	37.3	1.2	1.4*
SE	14	1,198	55.1	3,565	38.9	-	-	13,337	43.7	8.4	1.3
SP	8	492	26.3	3,386	30.9	-	-	7,677	30.1	0.5	1.1*
US	38	647	37.5	3,142	31.8	-	-	9,234	40	1.9	1.3

Near-Misses

				r = 10 km				r = 40 km			
REGION	**n**	**OBS**	**AREA**	**PPEMAX**	**EPOBS**	**EP5K**	**EP20K**	**PPEMAX**	**EPOBS**	**EP5K**	**EP20K**
GL	5	1,969	7.8	10,012	31.1	2.9	-	14,123	33.6	7.2	11.7*
NP	5	813	103	6,394	69.6	5.1	7.0*	21,130	58.2	8.2	6.4
SE	6	2,699	38.1	8,915	40.8	19.8	9.2*	11,190	26.2	10.1	5.6*
SP	14	2,412	45.5	11,144	51.1	17.5	16.6*	23,617	40.5	13.6	2.1
US	30	2,182	49.2	9,415	47.8	8.1	9.2	19,611	36.3	12.3	2.6

Hits

				r = 10 km				r = 40 km			
REGION	**n**	**OBS**	**AREA**	**PPEMAX**	**EPOBS**	**EP5K**	**EP20K**	**PPEMAX**	**EPOBS**	**EP5K**	**EP20K**
GL	0	-	-	-	-	-	-	-	-	-	-
NP	2	5,631	204.8	12,327	22.7	30.3*	-	59,941	26.8	29.6*	18.0*
SE	9	7,887	87.1	18,144	52.2	82.5	63.3	40,943	27.6	42.5	15.2
SP	10	10,873	34.9	23,698	42.1	82.6	30.1	41,431	18.5	43.7	10.7
US	21	8,848	70.6	19,534	43	82.5	34.6	40,943	27.3	42.5	16.8

Table 2.6. Characteristics of all near-misses from 1995 to 2016. Characteristics include the date (DATE) and location (LOCATION) of the tornado, its magnitude on the EF scale (MAG), tornado footprint area (AREA) in km^2, fatalities (FAT), persons exposed (OBS), median (PPEMED) and maximum (PPEMAX) persons potentially exposed, and the probability that the persons potentially exposed exceeds the persons exposed (EPOBS), or 5,000 persons (EP5K) within 10 km. Table is sorted by the probability that the near-miss would have been a hit. Bolded rows have a maximum potential exposure (within 10 km) of 20,000 persons or more.

DATE	LOCATION	MAG	AREA	FAT	OBS	PPEMED	PPEMAX	EPOBS	EP5K	EP20K
10 May 2008	Picher, OK	4	195.6	21	3,076	4,651	10,888	87.8	44.9	-
1 Mar 1997	**Little Rock, AR**	**4**	**22.8**	**5**	**3,269**	**3,420**	**34,315**	**52.5**	**44.5**	**16.6**
22 May 2011	Joplin, MO	5	50.9	158	4,474	3,851	14,449	46.7	44.5	-
8 Apr 1998	**Oak Grove-Rock Creek, AL**	**5**	**58.9**	**32**	**4,251**	**4,227**	**32,927**	**49**	**40.3**	**9.2**
1 Mar 1997	Vimy Ridge-Shannon Hills, AR	4	29	10	3,625	2,806	17,352	42.8	35.2	-
10 Apr 2009	Murfreesboro, TN	4	25.7	2	4,996	4,592	8,168	32.7	32.6	-
16 Dec 2000	Tuscaloosa, AL	4	19.9	11	2,558	2,423	9,662	49.6	30.3	-
27 Apr 2014	Mayflower-Vilonia, AR	4	79.8	16	2,697	2,923	12,313	67.4	23.7	-
25 May 2008	**Parkersburg-New Hartford, IA**	**5**	**126.6**	**9**	**1,665**	**2,498**	**25,321**	**85.4**	**23.1**	**7**
10 May 2010	Norman-Little Axe, OK	4	28.8	1	2,985	1,522	14,480	27.2	18.4	-
24 May 2011	Chickasha-Oklahoma City, OK	4	43.1	1	2,564	2,057	11,401	40	18.1	-
2 Jun 1998	Frostburg, MD	4	64.5	0	2,095	3,166	13,766	70.5	17.4	-
24 Nov 2001	Madison, MS	4	14.9	2	2,259	2,173	15,674	46.5	16.9	-
4 May 2003	Jackson, TN	4	50.5	11	2,780	1,846	6,988	30.5	9.3	-
5 Feb 2008	Atkins-Clinton, AR	4	236.7	13	2,184	2,456	5,874	56.1	8.2	-
28 Apr 2002	La Plata, MD	4	36.3	3	3,907	2,348	10,633	25.2	8	-
16 Apr 1998	Lawrence County, TN	5	179.7	3	2,619	1,941	11,639	32.1	7.1	-
8 Apr 1999	Creston-Granger, IA	4	103	0	813	1,103	6,395	69.6	6	-
4 Oct 2013	Wayne, NE	4	65.6	0	624	341	5,401	38.8	5.1	-

Table 2.6. Continued.

DATE	LOCATION	MAG	AREA	FAT	OBS	PPEMED	PPEMAX	EPOBS	EP5K	EP20K
19 May 2013	Lake Thunderbird-Shawnee, OK	4	50.8	2	2,179	2,128	6,988	44.8	4.5	-
6 Feb 2008	Moulton-Decatur, AL	4	21.6	4	944	949	7,377	50.4	4.1	-
24 Nov 2001	Winterville, MS	4	39.2	0	349	379	8,241	57.8	3.8	-
5 Jun 2010	Millbury, OH	4	5.2	7	937	498	10,012	31.1	2.9	-
13 May 1995	Niota, IL	4	73.6	0	195	447	7,391	98.4	2.4	-
29 Mar 1998	Comfrey, MN	4	162.2	1	1,021	783	5,337	29.5	2.3	-
24 May 2011	Washington-Goldsby, OK	4	29.9	0	284	474	9,168	82.6	1.8	-
10 Feb 2009	Lone Grove, OK	4	47.9	8	791	788	5,071	49.6	1.2	-
24 May 2011	Etna-Denning, AR	4	148	4	1,660	1,914	6,408	69.8	1.1	-
9 Apr 1999	Blue Ash, OH	4	4	4	1,969	1,499	6,143	28	0.5	-
11 Jun 1998	Greenfield-Maxwell, IN	4	7.8	0	993	579	5,449	35.3	0.2	-

Table 2.7. Same as Table 2.6, but for hits.

DATE	LOCATION	MAG	AREA	FAT	OBS	PPEMED	PPEMAX	EPOBS	EP5K	EP20K
22 Apr 2011	**St. Louis, MO**	**4**	**27.6**	**0**	**24,130**	**23,508**	**38,957**	**45.5**	**100**	**71.5**
27 Apr 2011	**Hackleburg-Phil Camp., AL**	**5**	**427.3**	**72**	**22,784**	**22,678**	**63,522**	**43**	**100**	**63.3**
27 Apr 2011	**Tuscaloosa-Birmingham, AL**	**4**	**308.7**	**64**	**39,231**	**41,949**	**102,942**	**52.2**	**100**	**80.8**
27 Apr 2011	Shoal Cr.-Ohatchee-Argo, AL	4	252.1	22	7,887	10,969	17,661	88.9	100	-
3 May 1999	**Bridge Creek-Moore, OK**	**5**	**77.9**	**36**	**23,649**	**18,992**	**37,617**	**30.6**	**94.7**	**45.5**
10 May 2010	**Moore-Choctaw, OK**	**4**	**70.6**	**2**	**10,582**	**11,710**	**55,020**	**53**	**90.5**	**30.1**
4 May 2003	Franklin, KS	4	15.5	2	8,848	8,286	14,277	41.2	89.6	-
27 Apr 2011	Cordova, AL	4	264.8	13	5,712	5,961	10,414	58.5	83.9	-
24 Apr 2010	Tallulah-Yazoo City-Durant, MS	4	675.2	10	5,628	7,337	13,533	74.6	82.6	-
20 May 2013	**Moore, OK**	**5**	**38.7**	**24**	**19,181**	**17,548**	**42,244**	**43**	**82.6**	**39**
28 May 1996	**Pioneer Village-Hillview, KY**	**4**	**37.5**	**0**	**7,877**	**8,340**	**32,553**	**54.4**	**82.5**	**9.3**
3 May 1999	**Wichita-Haysville, KS**	**4**	**31.1**	**6**	**15,361**	**9,074**	**21,860**	**21**	**73.2**	**5.5**
26 Dec 2015	**Sunnyvale-Garland, TX**	**4**	**10.6**	**10**	**7,399**	**7,197**	**25,536**	**49.2**	**63.7**	**4**
18 May 1995	Deerfield-Campbellsville, AL	4	74.6	1	6,173	6,383	19,534	52.6	63.6	-
27 Apr 2011	Apison-Cleveland, TN	4	56.5	20	9,833	8,680	13,696	33.9	63.6	-
8 May 2003	**Moore-Oklahoma City, OK**	**4**	**17.8**	**0**	**11,163**	**7,586**	**20,620**	**32.6**	**59.9**	**1.7**
10 Feb 2013	Hattiesburg, MS	4	42.1	0	16,491	4,947	18,144	5.6	49.5	-
27 Apr 2011	Pisgah-Trenton, AL-GA	4	87.1	14	5,293	4,968	17,380	43.4	49.3	-
17 Nov 2013	Washington, IL	4	60	3	5,815	4,408	18,304	30.6	42	-
24 May 2011	El Reno-Piedmont, OK	5	163.4	9	5,123	2,566	14,352	24.1	26.1	-
22 May 2004	Hallam, NE	4	349.6	1	5,446	2,658	6,350	14.8	18.5	-
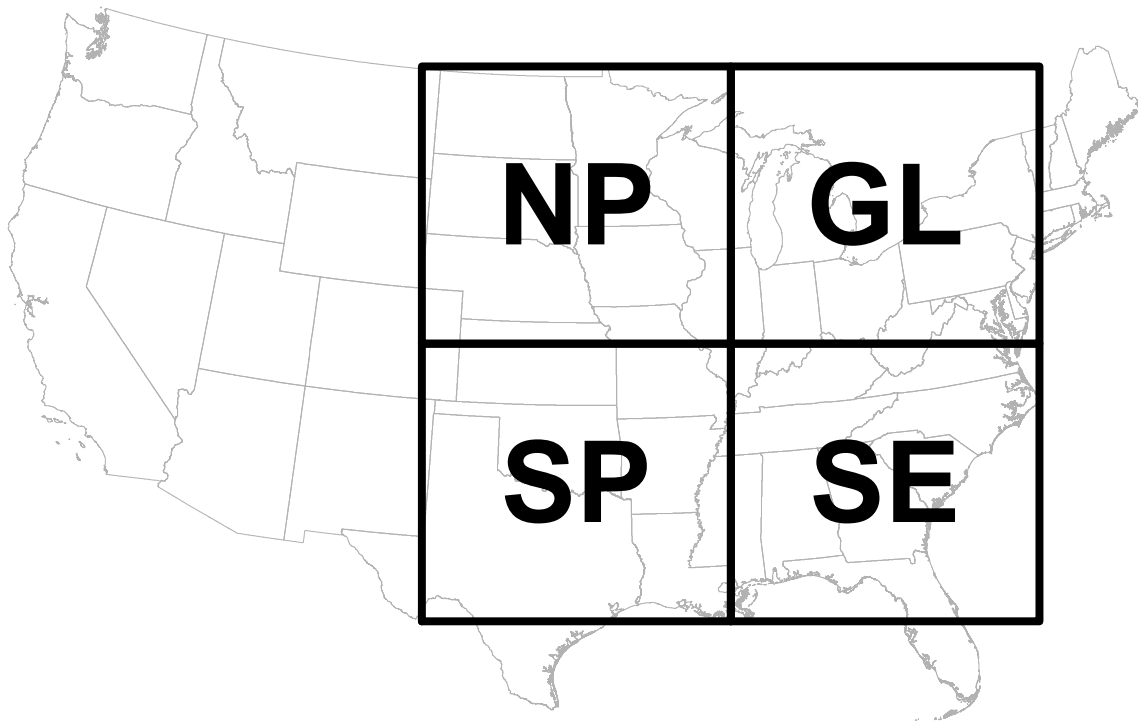
Figure 2.1. Four regions used in this study: Northern Plains (NP); Southern Plains (SP); Great Lakes(GL); Southeast (SE).
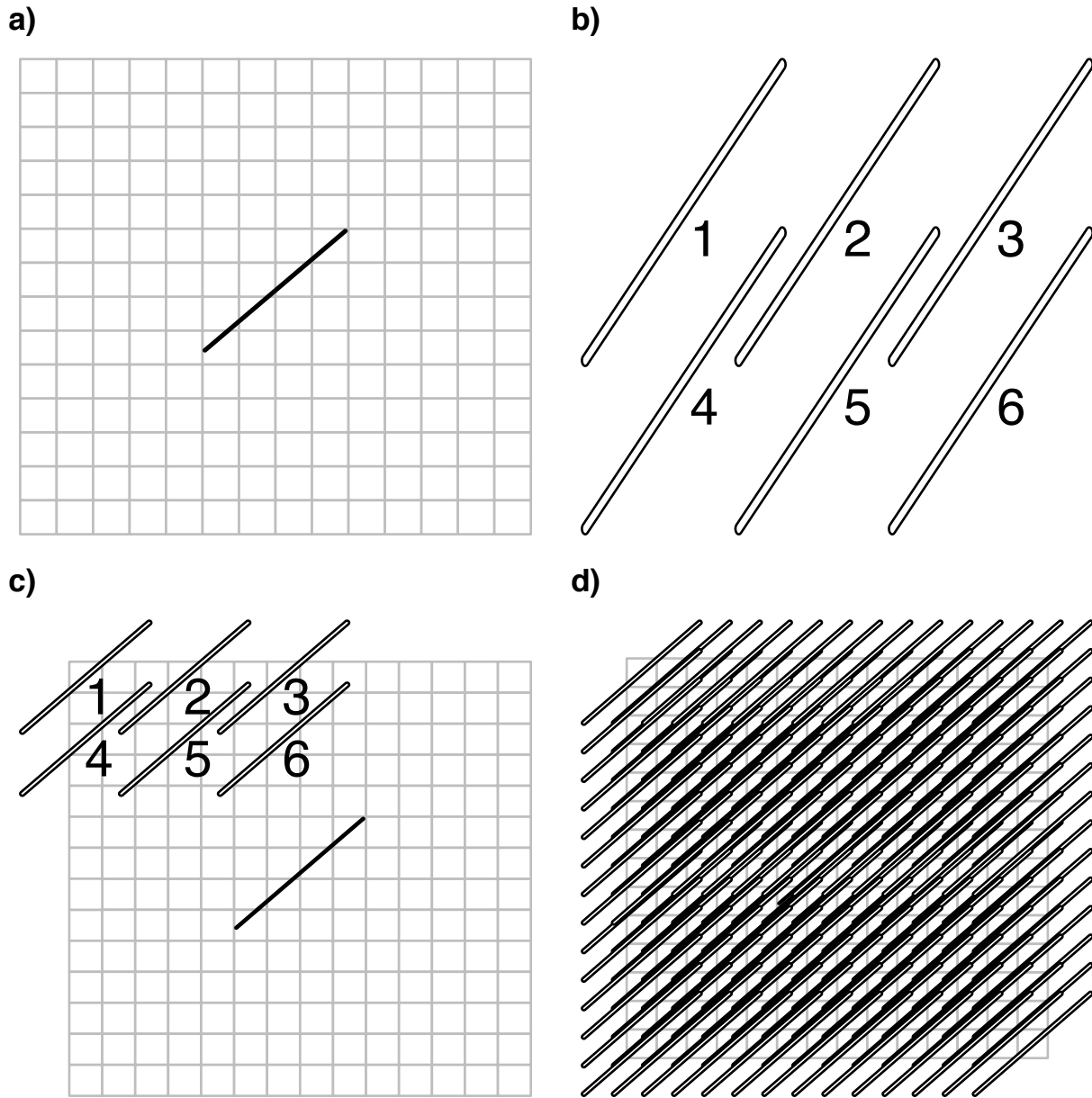
Figure 2.2. Simplified workflow for the replication procedure: a) Create a uniform grid surrounding the tornado footprint to be replicated. b) Make copies of the tornado footprint maintaining the size and orientation of the original footprint. c) Move each copy to the center of its corresponding grid cell. d) Repeat b and c until all grid cells have a copy of the original footprint. This example is not to scale; the actual grid is a uniform circular grid with a radius of 40 km and a resolution of 0.5 km for a total of 20,140 grid cells.
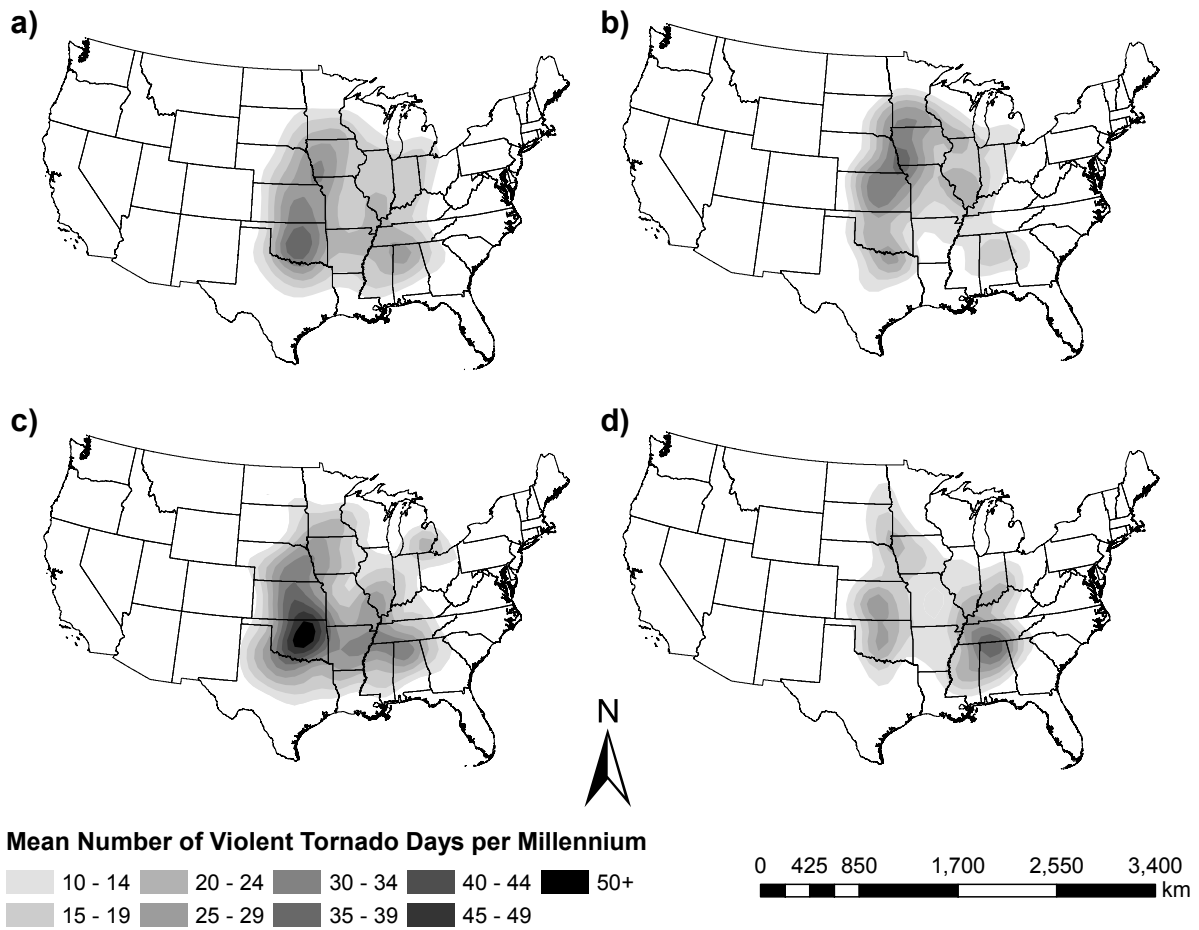
Figure 2.3. Mean number of violent tornado days in the US per millennium for the period of a) 1880 - 2016 and select 30 year periods (b) 1880 - 1909, c) 1930 - 1959, d) 1987 - 2016. Figure created following the methods of (Concannon *et al.*, 2000).

a)

b)
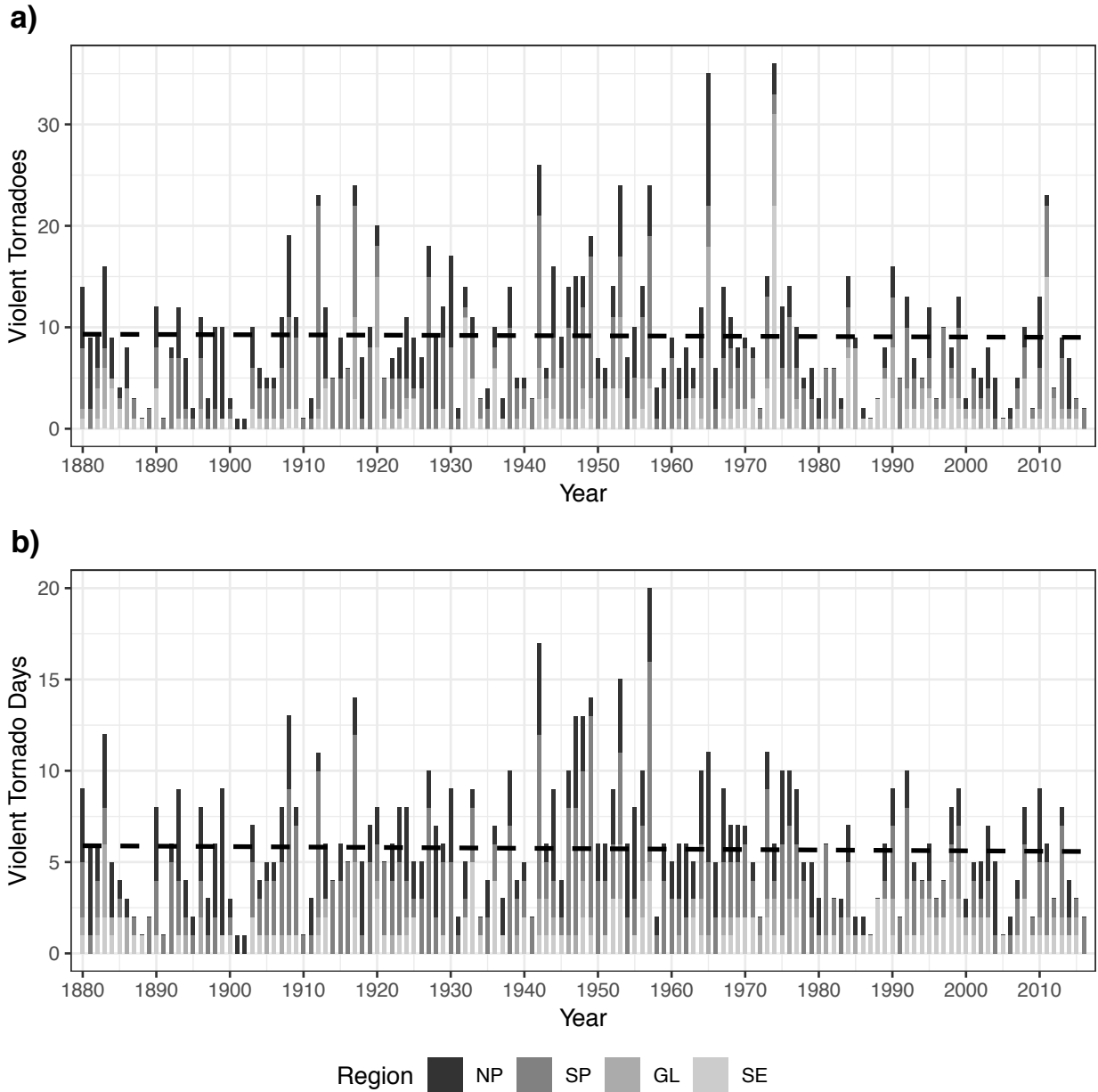
Figure 2.4. Violent tornado counts (a) and days with a violent tornado (b) per year from 1880 to 2016. Stacked bars are colored by region. Dashed line is the best fit using a linear model.
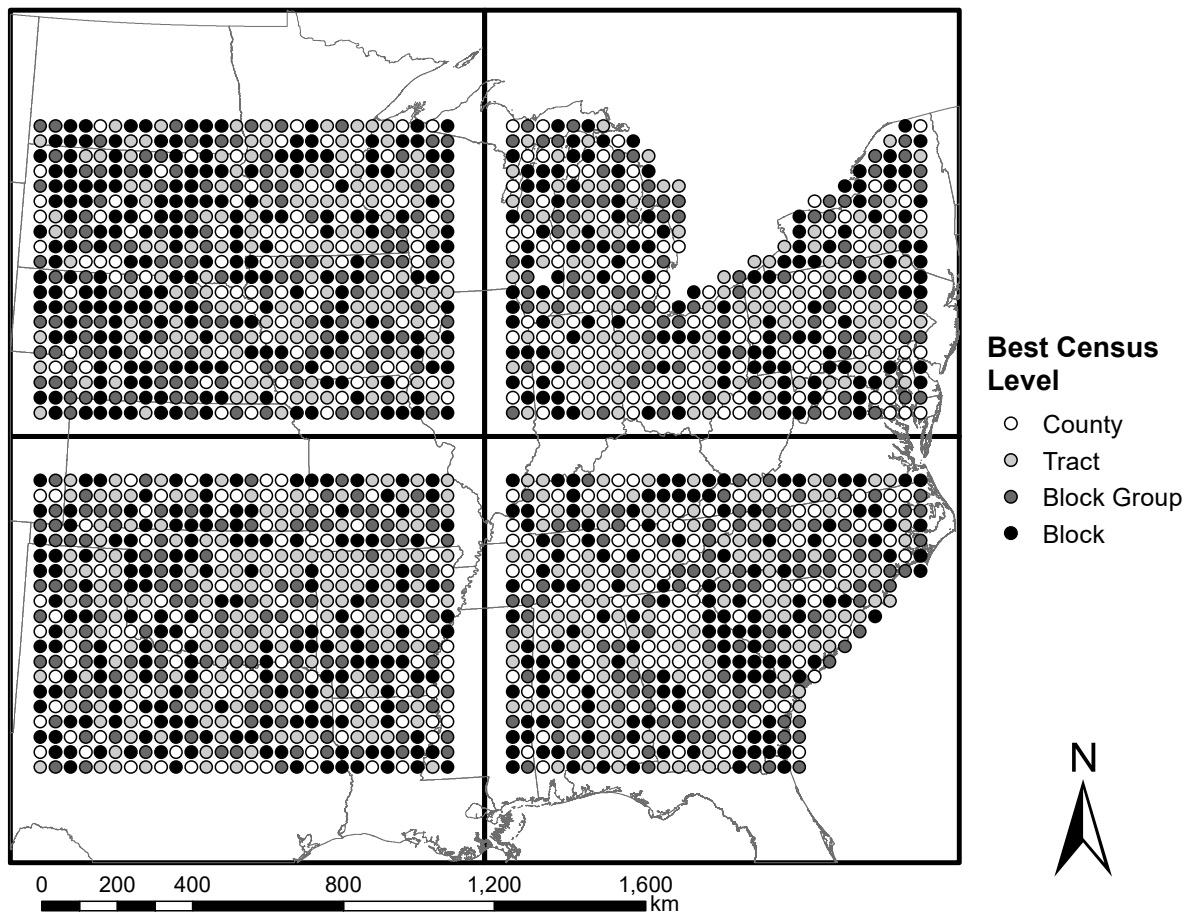
Figure 2.5. Results of the sensitivity test for best census resolution at the grid cell-level over the study area. The best census level was the one that yielded the lowest root mean square error for the estimation error between the observed and synthetic footprint exposures. The sensitivity test involved replicating the 3 May 1999 Oklahoma tornado outbreak in each grid cell over a uniform rectangular grid with a resolution of 10 km spanning the entire study area.
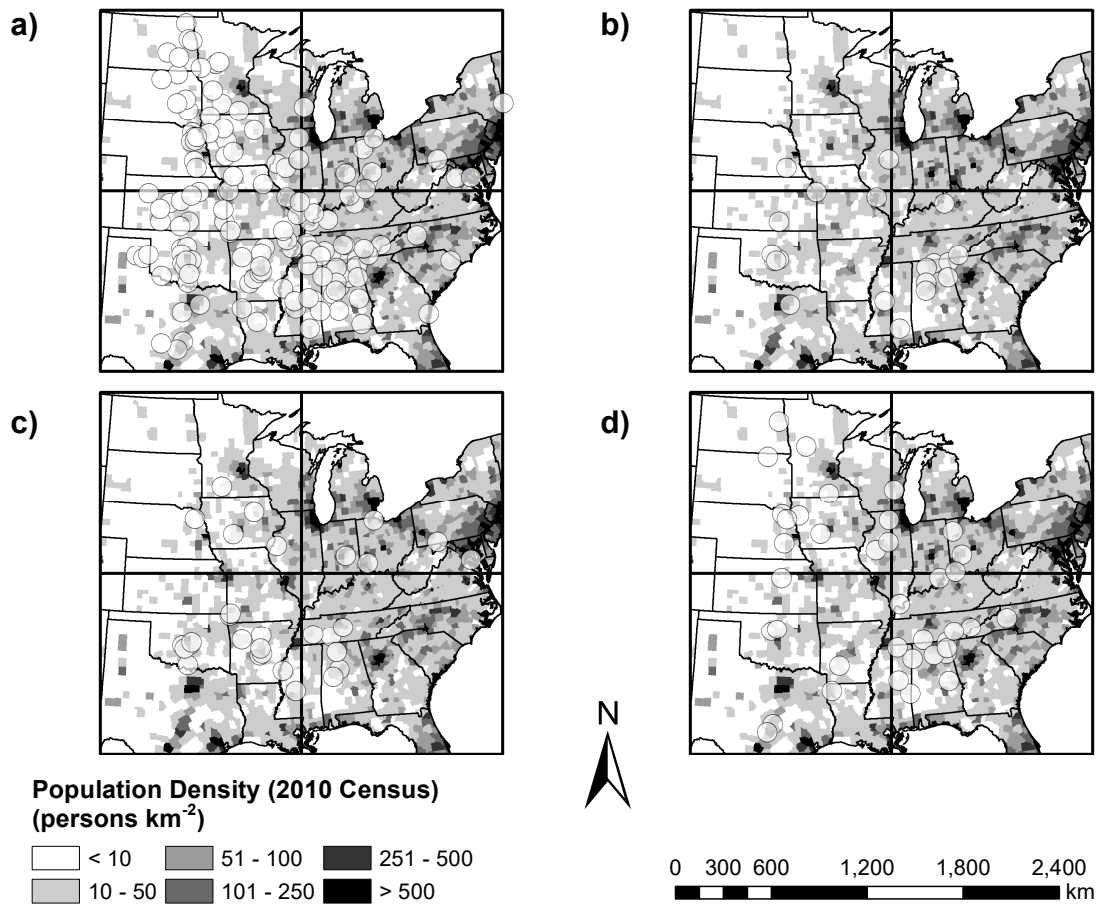
Figure 2.6. Centroids of (a) violent tornadoes, (b) hits, (c) near-misses, and (d) far-misses between 1995 and 2016. Population density, at the county level, from the 2010 census is included for reference.
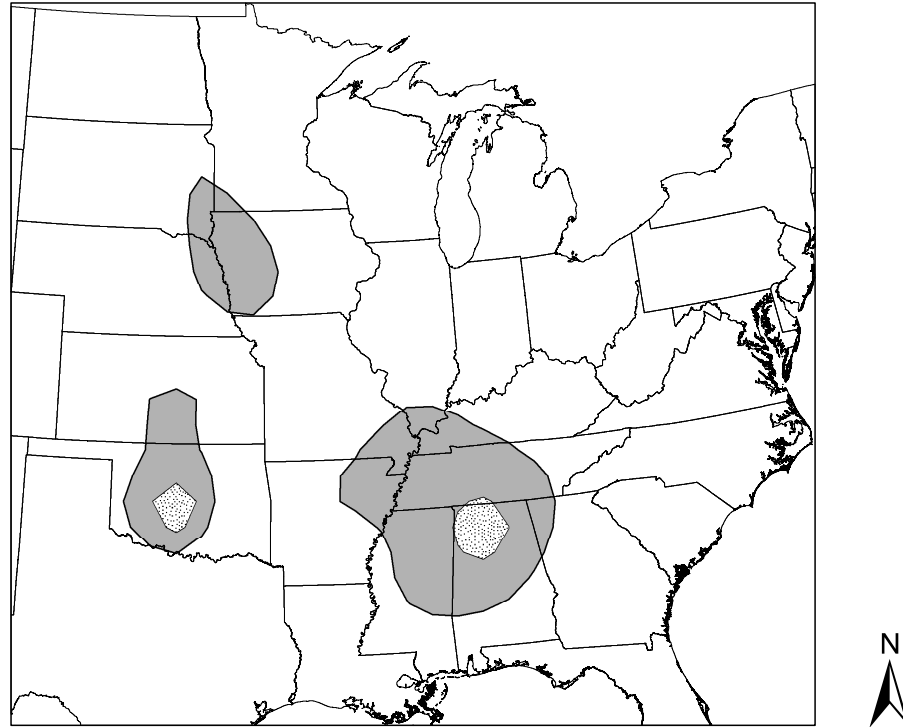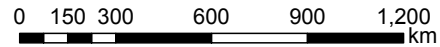
Figure 2.7. High-risk area for all violent tornadoes (grey) and near-misses or hits (stippled) between 1995 and 2016. High-risk area is defined as the area expected to have had at least two violent tornado days per century. High-risk areas were calculated after the data had been smoothed with a Gaussian lowpass filter with a standard deviation of 120 km (Concannon *et al.*, 2000).
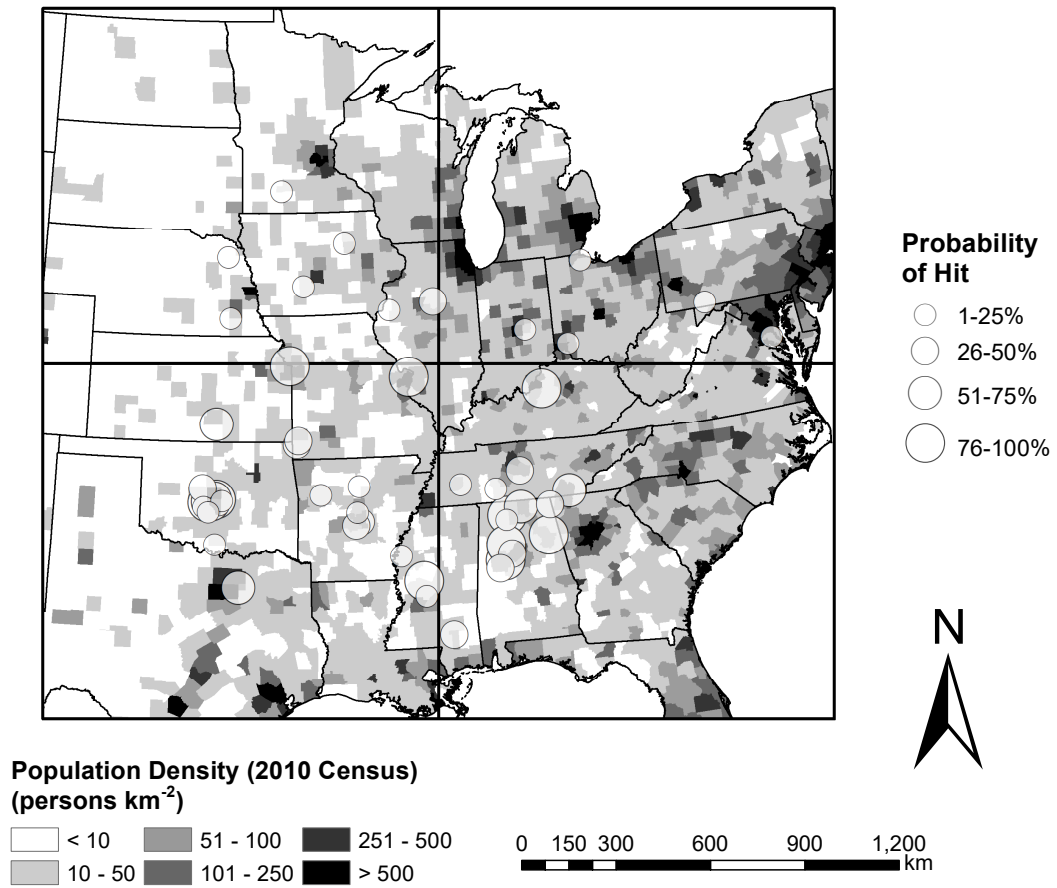
Figure 2.8. Probability that a violent tornado hitting within 10 km of its original footprint would have resulted in a hit between 1995 and 2016. Population density, at the county level, from the 2010 census is included for reference. The size of the dots relate to the probability category with higher probabilities represented by larger circles.
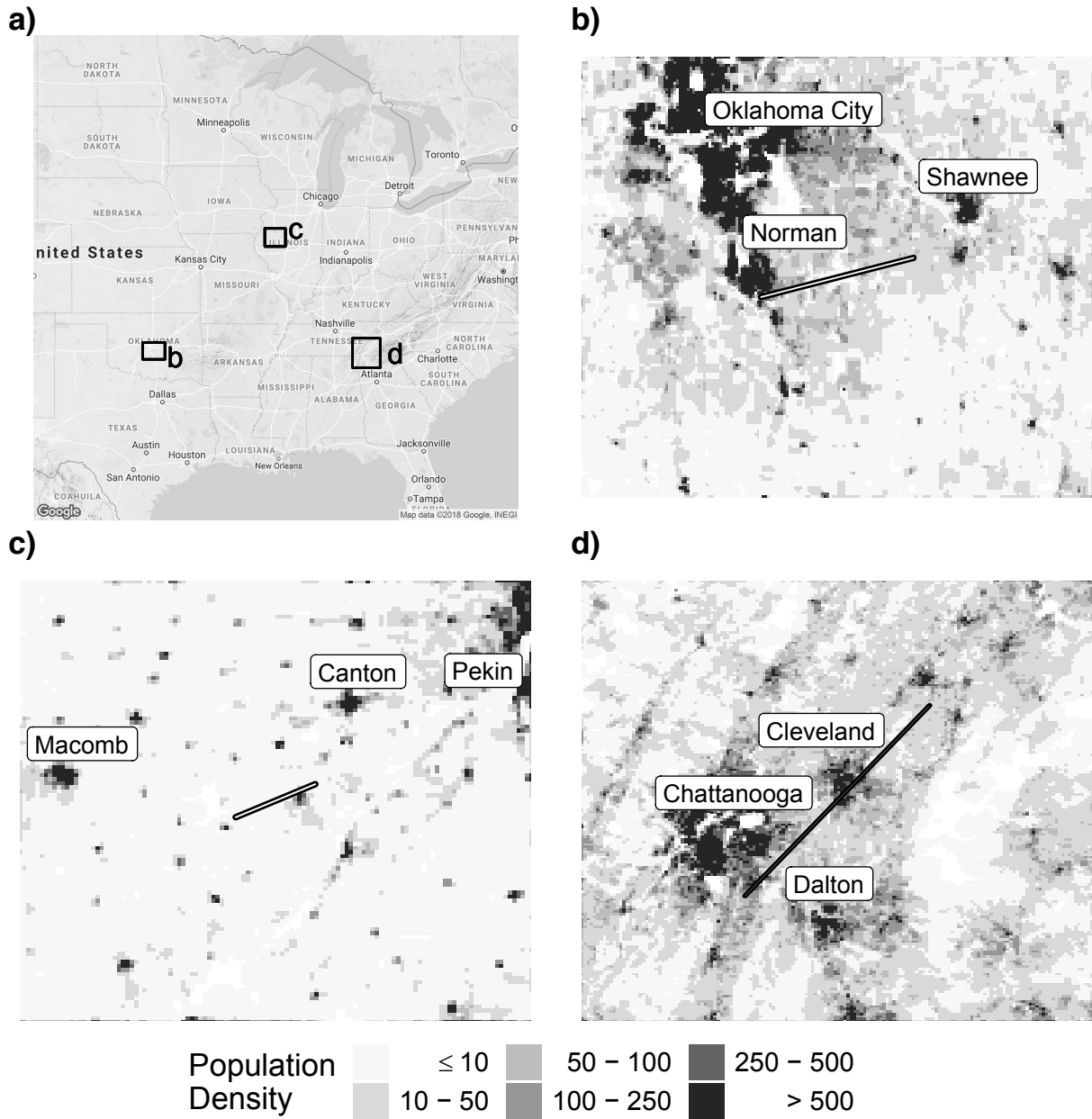
Figure 2.9. Examples of the tornado footprints for a near-miss, far-miss, and hit. a) A reference map of the US displaying the extent for each example. b) A near-miss on the City of Norman, Oklahoma on 10 May 2010. c) A far-miss for the Cities of Canton and Pekin in Illinois on 13 May 1995. d) A hit on the City of Cleveland, Tennessee on 27 April 2011. The population grids (for each potential impact zone) were generated following the method of (Ashley *et al.*, 2014) and use the block-level census data from the 1990 and 2010 censuses and a grid cell resolution matching the mean block size in the potential impact zone. The population was linearly interpolated to the year of the tornado.
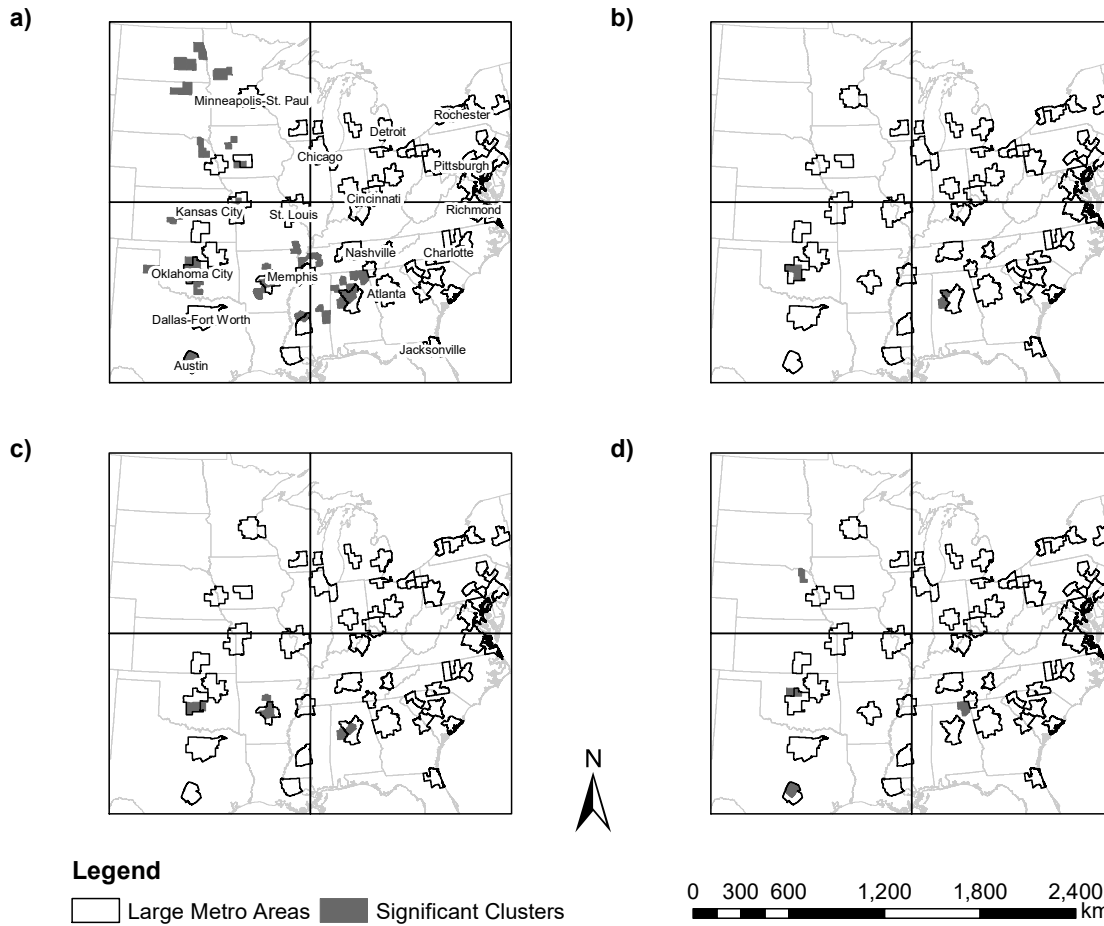
Figure 2.10. Counties with clusters of (a) violent tornadoes, (b) hits, (c) near-misses, and (d) far-misses. Counties are identified as clusters based on a local Moran's I test with $\alpha = 0.05$. Metropolitan Statistical Areas with populations greater than or equal to 500,000 persons are also plotted for reference.

Chapter 3 A Tornado Daily Impacts Simulator for the Central and Southern United States

Abstract:

In an average year (1979 – 2016), the US experiences nearly 1,100 tornadoes, which cause a total of 68 fatalities. Annual fatality rates have decreased since the peak in the 1920s, but there is a concern that they could start to rise again with increases in vulnerable populations and the impacts of climate change. It is possible to assess the risk of tornado fatalities using the historical record. However, the rarity of tornadoes and the short period of record may not capture the true risk. One way around this problem is to simulate thousands of years' worth of tornadoes to get a broader picture of risk. Previous tornado risk models have distributed tornadoes randomly or used climatology to generate realistic tornado patterns on an annual (or longer) time scale. From an operational standpoint, it would be useful to have a model that distributes tornadoes on a daily time step to enable the forecasting of potential tornado impacts on a given day. This study introduces one such model that distributes tornadoes using information about the favorability of the atmospheric environment for tornado development: the Tornado Daily Impacts Simulator (TorDIS). We demonstrate model utility through 1,000-year simulations over several metropolitan areas and through comparison between modeled and observed impacts for several high impact tornado days. Forecasting potential tornado impacts on a daily time step could allow emergency managers to plan ahead for high-risk days to prioritize their resources and save lives.

3.1) Introduction

Since the early 1950s, the U.S. has experienced a steady increase in economic losses due to weather-related hazards (Bouwer, 2011; Field *et al.*, 2012; Smith and Katz, 2013). The exact cause of this increase is unclear; however, it is likely caused by some combination of changing patterns in urban development (Ashley *et al.*, 2014; Ashley and Strader, 2016) and changes in the environmental conditions that favor the formation of weather-related hazards (Strader *et al.*, 2017a). These changes are likely to continue into the future as the U.S. continues to become increasingly urban (Alig *et al.*, 2004; EPA, 2009; Bierwagen *et al.*, 2010) and as climate change causes an increase in the number of days favoring severe weather events (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2013; Gensini and Mote, 2015).

Here, we focus on one particular hazard, the tornado. We define tornado risk as the probability of occurrence of a tornado at a given time and place. We define tornado exposure as the number of persons or housing units residing in the direct path of the tornado (Strader *et al.*, 2016). The exact effect that climate change has on tornadoes is unknown, but there have already been changes in interannual variability and the spatiotemporal clustering of tornadoes (Brooks *et al.*, 2014; Elsner *et al.*, 2015) as well as shifts in regional climatology (Gensini and Brooks, 2018).

Advances in computational resources and software have enabled researchers to model the impact of hazards on people and property over large spatial domains and time periods (Burton, 2010; Ashley *et al.*, 2014). Much of the modeling research has focused on the impacts of hurricanes (Pinelli *et al.*, 2004; Peduzzi *et al.*, 2012), floods (Remo *et al.*, 2012) and earthquakes (Remo and Pinter, 2012; Dell'Acqua *et al.*, 2013). Research on tornado impacts has mostly focused on scenario work involving transposing historical or synthetic tornado footprints

over heavily populated areas to identify worst-case scenario events (Rae and Stefkovich, 2000; Wurman *et al.*, 2007; Ashley *et al.*, 2014). For example, a study by Hatzis *et al.* (2019) focused on replicating historical violent tornadoes to determine the likelihood of a violent tornado hitting or narrowly missing a heavily populated area. Other recent tornado modeling research has employed Monte Carlo methods, a technique which uses repeated random sampling to ascertain the probability distribution for some unknown quantity (Mooney, 1997), to better understand probabilistic impacts of tornadoes (Meyer *et al.*, 2002; Daneshvaran and Morden, 2007). One model, in particular, the Tornado Impact Monte Carlo (TorMC) model, developed by Strader *et al.* (2016), is unique in its examination of the dynamic interaction between tornado risk, tornado severity, and exposure. This model was also used to study the impact of urbanization and climate change on future tornado exposure (Strader *et al.*, 2017a).

The TorMC model represents an important step in the simulation of the dynamic relationship between tornado risk, tornado severity, and vulnerability over time and space by assessing tornado impacts on an annual time scale. However, from an operational standpoint, it would be useful to have a model capable of simulating tornado impacts on shorter time scales (Karstens *et al.*, 2015; Smith *et al.*, 2015). A daily impacts model would allow a user to determine how the impact of an observed tornado footprint compares to the potential impacts that could have occurred under the same atmospheric environment (*i.e.*, how many more people could have potentially been exposed to the tornado). Such a model would rely on information regarding the favorability of the atmospheric environment for the production of tornadoes (Sobash *et al.*, 2011; Nowotarski and Jensen, 2013; Karstens *et al.*, 2015). Sources of such information could be atmospheric reanalysis data (Brooks *et al.*, 2003b; Trapp *et al.*, 2007; Gensini and Ashley, 2011) or probabilistic outlooks from the Storm Prediction Center (SPC; Kay

and Brooks, 2000; Stensrud *et al.*, 2009). Environmental data could also be used to constrain tornado parameters such as magnitude (Colquhoun and Riley, 1996; Naylor and Gilmore, 2012) or even the number of tornadoes occurring on a given day (Thompson and Edwards, 2000; Corfidi *et al.*, 2010). Additionally, climate projections for variables related to the atmospheric environment (*e.g.*, shear and convective available potential energy) allow for a more rigorous approach to estimating future tornado impacts under climate change (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2013; Gensini *et al.*, 2014b). By using projected atmospheric environments, it may be possible to estimate the location of future changes in tornado occurrence and severity as well as the magnitude of those changes (Diffenbaugh *et al.*, 2008; Tippett *et al.*, 2015).

No current published modeling approach simulates tornado impacts at the daily time scale. Hence, the objective of this research is to present a proof of concept for one such daily impacts model: The Tornado Daily Impacts Simulator (TorDIS). Like the TorMC model, TorDIS overlays tornado paths upon cost surfaces such as population or housing units. However, in TorDIS, all aspects of the tornado path (*i.e.*, location, size, intensity, direction) are constrained by the environment in which it forms. In the following sections, we describe the TorDIS approach, validate the model's deterministic components through a global sensitivity analysis, and showcase the model's utility through applications at the daily and annual time scales over the central and southern United States.

3.2) TorDIS Development

TorDIS was developed as an extension of the work of Strader *et al.* (2016) to link tornado distribution and characteristics to daily atmospheric environments instead of basing them on climatology. Like TorMC, TorDIS is limited by the accuracy and scope of the historical data used by the model (Brooks *et al.*, 2003a; Verbout *et al.*, 2006; Doswell, 2007). The primary

objective of this study is to expand upon the methods of Strader *et al.* (2016) to work towards simulating tornado impacts at time scale suitable for operational use.

TorDIS has five process steps (Figure 3.1): (1) parameterization and study area selection; (2) production of a tornado probability raster and determination of whether the day is favorable for tornado development; (3) selection of tornado parameters and creation of footprints; (4) extraction of cost information across the footprint; and (5) production of model output. Like the TorMC model, TorDIS is modular in nature with many user-defined parameters to allow the user to control model output.

3.2.1) Model Input

a.) Tornado Record

TorDIS simulates tornadoes by sampling various tornado parameters from the historical tornado distribution or theoretical distributions based on the historical record. The model requires a tornado database, in shapefile format, containing tornado locations, path lengths, widths, intensities, and dates, as lines. By default, TorDIS uses the SVRGIS tornado database (US) from the SPC for the period 1979 to 2016 (https://www.spc.noaa.gov/gis/svrgis/). However, the model can also be adjusted to use other databases such as the U.S. Significant Tornado dataset (Grazulis, 1993, 1997) or the European Severe Weather Database (https://www.essl.org/cms/european-severe-weather-database/). The SVRGIS tornado database has many known flaws that have been reported elsewhere, including, but not limited to: population bias, reporting frequency (fewer weak tornadoes were observed early in the record), questions of accuracy (due to amateur reports), changes in reporting methodologies (*e.g.*, reporting mean path width versus maximum path width), and concerns regarding using damage assessments to determine tornado intensity (Verbout *et al.*, 2006; Ashley *et al.*, 2014; Strader *et*

al., 2015). However, we selected it for use in the model since it is still the best record available (Strader *et al.*, 2016).

b.) Atmospheric Environmental Data

An important aspect of TorDIS is the linkage between tornado occurrence and intensity, and the atmospheric environment in which the tornado forms. Tornadoes require certain key ingredients to form, including a moist, unstable atmosphere, a source of lift, rotation, and low cloud bases (Brooks *et al.*, 2003b; Ashley and Strader, 2016). These ingredients can be represented by certain severe weather diagnostic parameters, such as Convective Available Potential Energy (CAPE), Storm Relative Helicity (SRH), and Vertical Wind Shear (VWS). These parameters can be derived from reanalysis data (Brooks *et al.*, 2003b; Gensini and Ashley, 2011; Gensini *et al.*, 2014b). TorDIS uses both gridded diagnostic parameters and proximity soundings (atmospheric profiles at the closest points in space and time to the occurrence of tornadoes) to determine the favorability of the environment for tornado formation, and the intensity of a tornado should one occur, respectively.

By default, TorDIS uses reanalysis data from 1979 to 2016 from the North American Regional Reanalysis (NARR) project for the central and southern United States (Figure 3.2). The NARR is an extension of the National Centers for Environmental Prediction (NCEP) Global Reanalysis project and uses the high-resolution Eta model (with 32 km horizontal resolution and 45 sigma levels) combined with the Regional Data Assimilation System (RDAS) to generate high-resolution atmospheric field variables (Gensini and Ashley, 2011). The fields are available eight times daily from 1979 to 2016, and all vertical fields are available at 29 pressure levels. Mesinger *et al.* (2006) evaluated the quality of the NARR data and found it to be a major improvement in both accuracy and resolution over previous global reanalysis efforts. We chose

this data set because of its high resolution over North America. There are caveats with the data: small biases have been documented in temperature and precipitation fields (Gensini and Ashley, 2011) and a lower vertical resolution in the lowest layers can cause biases for some thermodynamic parameters which require vertical integration (*e.g.*, CAPE) when compared to collocated observations (Gensini *et al.*, 2014a). We collected all gridded reanalysis data only once per day (at 0000 UTC, the peak time for severe weather activity in the central United States; Brooks *et al.*, 2003b; Gensini and Ashley, 2011) while we collected proximity sounding data at the time closest to the tornado report (Lee, 2002). The NARR data is processed using the Sounding/Hodograph Analysis Research Program in Python (Halbert *et al.*, 2015) to obtain the severe weather diagnostic parameters used by the model.

TorDIS can also be calibrated to use other reanalysis data sets, such as the NCEP Reanalysis-2 (Kanamitsu *et al.*, 2002) as well as to use future atmospheric environments from climate models, such as the North American Regional Climate Change Assessment Program (Mearns *et al.*, 2013) as long as they are available at a daily time scale.

3.2.2) Model Mode and Study Area Selection

TorDIS has two modes in which it can be run: daily or annually. Like TorMC, TorDIS is a stochastic model which relies on Monte Carlo simulations to estimate tornado impacts. In daily mode, the user specifies the number of Monte Carlo simulations to run (*n*) and tornado production on the specified day (*e.g.*, 27 April 2011) is simulated *n* times to assess the potential impacts over the spatial domain on the day. This mode can be used to assess the likelihood of population exposure exceeding that of a particularly impactful tornado on the day (*e.g.*, the EF4 tornado that hit Tuscaloosa and Birmingham, AL on 27 April 2011; Doswell *et al.*, 2012). In annual mode, the user specifies the number of Monte Carlo simulations to run and for each

simulation, a year of atmospheric environment data is randomly selected from the available data and tornadoes are simulated for each favorable day over that year. This mode is used to assess the long-term risk of tornado impacts, such as the risk of population exposure exceeding 5,000 persons in a given area (Hatzis *et al.*, 2019).

In addition to selecting a mode and number of simulations, the user also must select a study area over which to run the model. The study area must be at least 5,000 km^2 (corresponding to the SPC's practice of forecasting tornadoes within 40 km of a given location; Hitchens *et al.*, 2013) and be within the model's spatial domain (Figure 3.2). The maximum study area size is limited by both the size of the domain and the buffer size used to reduce edge effects (Strader *et al.*, 2016). Figure 3.2 shows the maximum study area size for a 50 km buffer as a dashed line. The model is first run over the entire domain, and then the footprints are clipped to the specified study area. The procedure is done this way to ensure that sufficient tornado footprints are generated over the study area for each day since daily tornado probabilities are low for smaller areas (Brooks *et al.*, 2003a).

3.2.3) Tornado Probability Surface and Daily Tornado Count

a.) Tornado Probability Regression Model Specification and Sensitivity Analysis

Tornado distribution in the model begins with a decision of whether or not the atmospheric environment is favorable for tornado development on a given day of the simulation. A number of methods have been suggested for defining environmental favorability, within both atmospheric and health sciences, including linear discriminant analysis (Lee, 2002; Brooks *et al.*, 2003b; Gensini and Ashley, 2011), geographic weighted regression (GWR; Nakaya *et al.*, 2005; Ivajnšič *et al.*, 2014), and logistic regression (Billet *et al.*, 1997). TorDIS makes use of logistic regression as most atmospheric variables do not follow a normal distribution (days with high

CAPE and/or high VWS are rare; Brooks *et al.*, 2003b; Gensini and Ashley, 2011), one of the

assumptions of linear discriminant analysis (Pohar *et al.*, 2004). Additionally, geographically

weighted regression has a high computational cost (Ali *et al.*, 2007) and so it is not uncommon

for large-scale studies on severe weather favorability to opt for non-geographically weighted

methods (*e.g.*, Brooks, Lee and Craven, 2003; Gensini and Ashley, 2011; Diffenbaugh, Scherer

and Trapp, 2013).

TorDIS uses four logistic regression equations (Table 3.1) to determine tornado

favorability: severe weather probability in a low CAPE environment (svr_{low}), tornado probability

given severe weather in a low CAPE environment ($torn_{low}$), severe weather probability in a

moderate to high CAPE environment (svr_{mod}), and tornado probability given severe weather in a

moderate to high CAPE environment ($torn_{mod}$). We used these equations since studies have

shown that predicting tornadoes is more difficult in low CAPE environments (Dean and

Schneider, 2008; Sherburn and Parker, 2014) and Brooks *et al.* (2003b) suggest predicting the

occurrence of tornadoes given that severe weather is already occurring. An environment was

considered to be low CAPE where CAPE ≤ 500 J kg^{-1}, following Sherburn and Parker (2014).

We considered many common severe weather diagnostic parameters for inclusion in the

regression equations (*e.g.*, CAPE for a surface-based parcel, Freezing Layer; Lifted

Condensation Level, 0 to 1 km vertical wind shear; vertical velocities at 850 mb, 500 mb, and

200 mb; Brooks *et al.*, 2003b; Lock, 2012; Gensini *et al.*, 2014a). However, after testing for

multicollinearity (Zuur *et al.*, 2010) and an application of the least absolute shrinkage and

selection operator (Thompson, 2009), only ten were selected for the initial model specification:

CAPE, Convective Inhibition (CIN), and Lifted Condensation Level (LCL) for a mixed-layer

(lowest 100 mb) parcel; surface height (ELEV); 2 – 4 km lapse rate (LR24); 0 – 1 (VWS1) and 0

– 6 (VWS6) km vertical wind shear; vertical velocity at 850 mb (ω_{850}) and 500 mb (ω_{500}), and 0 – 3 km Storm Relative Helicity (SRH3).

After the initial model specification, a global sensitivity analysis was conducted on all regression equations to determine which variables explained most of the variability in model output. The global sensitivity analysis method used was the method of Sobol (1993) as improved by Saltelli (2002) and implemented using SimLab software (Giglioli and Saltelli, 2008). In this method, model output variances are decomposed into partial variances representing the contribution of model parameters (individually and together through interactions) to the overall uncertainty of the model output. The measure of this uncertainty, used in this study, was a normalized version of the Total Order Index (S_t) which represents the contribution of each parameter (and its interactions) to the overall uncertainty (Confalonieri, 2010) of the model output. Following Confalonieri (2010), a model parameter was considered important enough to be included if its S_t was at least 5%. Table 3.1 shows the final list of parameters selected for each model. The Tornado Forecast Model Selection and Validation section of Appendix A shows the results of the regression model validation and sensitivity analysis.

b.) Daily Tornado Production

The process of determining whether a simulated day will have tornadoes starts with the creation of a tornado probability surface for that day. The tornado probability surfaces are created by applying the regression equations to the gridded environmental data. First, the CAPE status (low or moderate to high) is assessed for each grid cell to determine which set of regression equations to use on that grid cell. The severe weather probability is then calculated using the relevant equation. Following SPC's practice of not including severe weather probabilities below 2% (Hitchens *et al.*, 2013), if the cell has a severe weather probability of less

than 2% the tornado probability is set to 0. Otherwise, the tornado probability is calculated using the relevant equation. Finally, a Gaussian low-pass filter with a 120 km standard deviation is applied to the tornado probability surface to smooth the data (Gensini and Ashley, 2011) and all grid cells with tornado probabilities below 2% are set to 0 as well (Figure 3.3). Smoothing is performed to help spread favorable environments over a larger area since environments change over a day (Brooks *et al.*, 2007).

Days are considered to be favorable for tornado production when the favourable area (area with 2% or higher probability) covers at least 5,000 km^2 (Hitchens *et al.*, 2013). Since daily tornado production is highly variable (Elsner *et al.*, 2014; Tippett and Cohen, 2016) and not all days that are considered favorable produce tornadoes (Trapp *et al.*, 2007; Lock, 2012), daily tornado counts are randomly drawn from the historical record. Tornado production is linked to LCL with higher LCLs corresponding to increased outflow, and a reduced likelihood of tornado formation (Rasmussen and Blanchard, 1998). Tornado outbreaks also tend to be the largest in the spring and fall (Brooks *et al.*, 2003a; Doswell *et al.*, 2006). To account for these patterns, daily tornado counts are broken down into eight categories based on season, and the 5th percentile daily LCL height over the domain and a daily tornado count is randomly drawn from the relevant category based on the atmospheric conditions of the day for each day that is considered favorable for tornado production. Once the tornado probability surface has been created, and a daily tornado count has been selected, tornado touchdown points are distributed across the model domain using a weighted random distribution based on the tornado probability surface.

3.2.4) Tornado Intensity

Tornado intensity is related to SRH3 with more intense tornadoes more common in high helicity environments more conducive of supercell production (Colquhoun and Riley, 1996;

Rasmussen and Blanchard, 1998). The literature suggests multiple SRH3 thresholds for supercell production, and thus stronger tornadoes, including 150 m^2 s^{-2} and 250 m^2 s^{-2} (Droegemeier *et al.*, 1993; Moller *et al.*, 1994; Colquhoun and Riley, 1996). A preliminary investigation of proximity sounding data for tornadoes in the SVRGIS database showed a clear distinction between tornado intensities for tornadoes with SRH3 in the following categories: low ($SRH3 < 150 \ m^2s^{-2}$), moderate ($150 \ m^2s^{-2} \leq SRH3 < 250 \ m^2s^{-2}$), and high ($SRH3 \geq 250 \ m^2s^{-2}$). All the tornado intensities in the SVRGIS database are separated into groups based on the SRH3 category. Once a tornado is placed, the SRH3 value for that location is extracted and a tornado intensity is randomly drawn from the relevant group.

3.2.5) Path Length and Width

Brooks (2004) showed that, when separated by intensity class, both tornado path lengths and widths were well approximated by Weibull distributions due to their non-negative and positively skewed nature. Tornado path lengths and widths are randomly drawn from intensity-specific Weibull distributions fitted to the observed path width and length data. Before fitting the Weibull distributions, both path length and width values in the SVRGIS database were adjusted to account for non-meteorological trends in their values (Strader *et al.*, 2015, 2016). Jumps in the mean annual path width were found in 1995 (when path width reporting switched from mean width to maximum width; Agee and Childs, 2014; Ashley *et al.*, 2014) and 2007 (with the switch from the F to EF-scale for measuring tornado intensity; Strader *et al.*, 2016), while a jump in the mean annual path length was only evident in 2007. Following the method of Agee and Childs (2014), path widths are detrended by determining the difference between the lower threshold of the mean annual path widths during 1979 – 1994 (1995 – 2006) and 2007 – 2016, and adding that difference (61.2 m and 51.5 m, respectively) to the path widths during the

76

earlier two periods. Path lengths are similarly detrended by adding the difference between the lower threshold of the mean annual path length during 1979 – 2006 and 2007 – 2016 (1.04 km) to each of the path lengths during the earlier period. The exact reason for the increase in path widths and lengths after 2006 is unknown (Strader *et al.*, 2016). However, possible reasons include improvements to damage assessment methods (Agee and Childs, 2014) and the addition of non-structural damage indicators to the EF scale that allowed previously undetectable wind damage to be identified (Doswell *et al.*, 2009; Edwards *et al.*, 2013; Hatzis *et al.*, 2019). Due to the improvements in damage assessment introduced in the EF scale, TorDIS considers the 2007 – 2016 length and width data the most accurate. Once a tornado's intensity is set, path width and length are randomly selected from the appropriate Weibull distribution. Alternatively, if the user specifies, the path width and length can be randomly selected from the adjusted historical data based on the intensity.

3.2.6) Tornado Direction

Tornado direction is closely linked to the 500 mb wind direction, which acts to steer the storm systems that generate tornadoes (Notis and Stanford, 1973; Suckling and Ashley, 2006). For each tornado touchdown point, the 500 mb wind direction is extracted (as a bearing). This bearing is assigned as the tornado's direction. After the direction has been assigned, a tornado footprint (area covered by tornadic winds) is created by using the direction and path length and width. The results of the validation of this method can be found in the Predicting Tornado Direction section of Appendix A.

3.2.7) Cost Extraction

Cost extraction, the determination of the impact of a given tornado, for TorDIS begins, as with the TorMC model, with the clipping of all tornado footprints to the specified study area.

Once the clipping is complete an area-weighted sum (Ashley *et al.*, 2014; Strader *et al.*, 2016) is performed on all cost units intersected by the clipped footprints to determine the total cost over each footprint. By default, the cost surface used by TorDIS is the U.S. Population Grids for 2010 (Summary File 1) from the Socioeconomic Data and Applications Center (SEDAC; Center for International Earth Science Information Network - CIESIN - Columbia University, 2017). However, any raster (*e.g.*, gridded housing unit data) or polygon-based cost surface (*e.g.*, percent population in poverty at the census block level in Oklahoma) can be used by the model.

3.2.8) Model Output

The final output of TorDIS includes shapefiles (containing the simulated tornado footprints with cost data, as well as tracks, and touchdown points) and comma-separated value files containing the attribute data of the simulated tornadoes. The model also produces probability of exceedance curves for the tornado costs.

3.3) Model Application

3.3.1) Model Performance

To demonstrate TorDIS's ability to estimate probabilitiy distributions for tornado exposures given atmospheric environmental information, a simulation of 1,000 years (in annual mode) was run over the maximum study area size assuming a 50 km buffer. The number of simulations was limited to 1,000 compared to the 10,000 recommended by Strader *et al.* (2016) due to computational restraints. We then clipped these runs to the Oklahoma City Metropolitan Statistical Area (MSA) in the state of Oklahoma (hereafter Oklahoma City refers to the Oklahoma City MSA and not the city itself). Oklahoma City was chosen as an urban area with a high risk of significant and violent tornadoes (Concannon *et al.*, 2000; Ashley, 2007;

Rosencrants and Ashley, 2015). We randomly selected tornado path lengths and widths from Weibull distributions. The environmental data were derived from NARR for 2005 to 2016 while the tornado database used was from SVRGIS for 1979 to 2016 (the period when the NARR was available for proximity soundings). The cost surface used was the default 2010 population counts on a 1 km resolution grid from SEDAC. For comparability, the following section emulates the structure used in Strader *et al.* (2016).

Over a 1,000-year simulation, TorDIS generated 7,408 tornadoes over Oklahoma City with 44.2% EF0, 35.0% EF1, 13.8% EF2, 5.6% EF3, 1.3% EF4, and 0.1% EF5 (Table 3.2). The model shows a slight bias towards lower intensity tornadoes when compared to the observed record over Oklahoma City for 1979 to 2016 (*i.e.*, 31.4% EF0, 41.9% EF1, 16.9% EF2, 5.1% EF3, 3.4% EF4, and 1.3% EF5). This bias could be due to 0 to 3 km Storm Relative Helicities being higher at the time of tornado formation than at 0000 UTC when the atmospheric environment is valid. Since tornadoes are more likely to have higher intensity in higher helicity environments, lower helicities would explain a bias towards lower intensities (Colquhoun and Riley, 1996).

Overall the model showed a slight bias towards the overproduction of all tornadoes, while simultaneously underproducing significant and violent tornadoes. The mean annual simulated tornado count was 7.4, 1.5, and 0.1 for all tornadoes, significant tornadoes, and violent tornadoes respectively compared to 6.3, 1.7, and 0.3 for the observed tornado counts. The mean annual simulated tornado day count was 3.0 compared to a mean of 2.8 days per year observed over Oklahoma City. The model's seasonal tornado production mirrors the observed production with a maximum in spring and a minimum in winter. However, tornado production is slightly underestimated in spring and overestimated in all other seasons. The general overproduction of

tornadoes is due to the model's high false alarm rate at predicting tornado days (Table A.1). For further comparison with the observed record over Oklahoma City, a random sample of 38 years was chosen from the 1,000-year simulation to match the 1979 to 2016 observed period. The sample shows a similar distribution across Oklahoma City to what was observed (Figure 3.4), but a greater number of tornadoes (261 versus 236). The random sample also revealed a higher mean (6.9) and median (5) tornado count when compared to the observed record (6.2 and 4, respectively). The breakdown by magnitude for the random sample (46.4% EF0, 34.1% EF1, 14.2% EF2, 3.8% EF3, 1.1% EF4, 0.4% EF5) showed the same slight bias towards lower intensity tornadoes present in the full 1,000-year simulation with the random sample containing no EF5 tornadoes.

The 1,000-year simulation over Oklahoma City yielded a mean tornado length and width of 9.2 km and 264.3 m, respectively (Table 3.2). The modeled lengths and widths increased as the tornado intensity increased similar to the findings of Strader *et al.* (2016), with lengths ranging from 3.1 km (EF0) to 58.6 km (EF5) and widths ranging from 90.8 m (EF0) to 1,459.4 m (EF5). Other than for EF5 tornadoes (which the model underproduces) the mean lengths for each intensity class are within 3 km of the mean for the respective Weibull distribution. For EF5 tornadoes the mean length is overestimated by more than 14 km, however, with only 4 EF5 tornadoes simulated over Oklahoma City it is likely, this is due to random chance. The mean widths are within 35 m of the respective Weibull distributions, except for EF5 tornadoes which are off by over 105 m. Maximum modeled widths and lengths were mostly overestimated when compared to the observed maxima over Oklahoma City (*e.g.*, maximum width (length) for simulated EF2 tornadoes was 2,748.5 (107.5) compared to an observed maximum of 1,258 m (47.9 km)). This is likely due to the tendency of the Weibull distribution to overestimate or

underestimate the tail ends of the distribution (Brooks, 2004). The mean modeled direction for all tornadoes over Oklahoma City was 77.4° (roughly west southwest to east northeast) which matches the findings of Suckling and Ashley (2006) for the South Central region.

Based on the 2010 SEDAC population grids, the mean (median) number of persons impacted by a tornado footprint was 212 (3; Table 3.3). Mean impacts (persons exposed) per footprint increase with increasing tornado magnitude from 15 (EF0) to 2,756 (EF4). While the simulated EF5 tornadoes did not have the maximum mean impacts, they did have the maximum median impacts. Mean annual numbers of persons impacted by tornadoes peaked at the EF2-3 magnitude range due to the small number of higher intensity tornadoes (only 1.5% (Table 3.2) of the simulated tornadoes had magnitudes of EF4+). One interesting finding was that among the top ten most impactful (in terms of persons exposed) simulated tornadoes over Oklahoma City, one was an EF2. Lower intensity tornadoes tend to have smaller footprints and would thus be expected to be less impactful (Brooks, 2004; Strader *et al.*, 2015). However, the placement of this simulation over a metropolitan area coupled with the tendency for maximum tornado lengths and widths to be overestimated lead to the possibility of a high impact EF2 tornado.

A significant difference between the TorMC model and TorDIS is in the ability to assess daily and seasonal differences in tornado impacts. From the 1,000-year simulation over Oklahoma City, the most active months for significant tornadoes were April and May with a secondary peak in October and November (Figure 3.5). This mirrors the observed record, which indicates peak significant tornado activity in spring and a smaller peak in autumn (SPC, 2017). Mean annual impacts (persons exposed) for significant tornadoes over each month of the simulation show a pattern that matches that of the significant tornado counts with peaks in April and October (Figure 3.5). The correlation between the mean annual impacts and the tornado

counts results from the fact that tornado impacts are a function of both tornado footprint area and the environment in which they hit (Ashley *et al.*, 2014; Ashley and Strader, 2016; Strader *et al.*, 2018).

3.3.2) Comparison of Tornado Risk across Major Metropolitan Statistical Areas

To show how TorDIS can be used to study spatial variations in tornado impacts, the 1,000-year annual simulation was also clipped to five other metropolitan statistical areas: Dallas/Fort Worth (Texas), Chicago (Illinois), Birmingham (Alabama), Omaha (Nebraska), and St. Louis (Missouri), hereafter references to the city name refer to the MSA. These MSAs were chosen for their size (from less than 1 million persons in Omaha to over 9 million persons in Chicago; U.S. Census Bureau, 2010), distribution throughout the model domain, and varying levels of tornado risk (risk is highest in an 'L' shaped pattern between Nebraska, Texas, and Alabama and drops off further towards the northeast; Concannon *et al.*, 2000; Ashley, 2007; Hatzis *et al.*, 2019). All model parameters and decisions were the same as for the 1,000-year simulation over Oklahoma City. The simulation produced tornadoes of all magnitudes, but this analysis focuses on significant tornadoes.

The normalized significant tornado count was highest over Omaha and Dallas/Fort Worth (103) and lowest over St. Louis (69; Table 3.4). Despite having one of the lowest normalized significant tornado counts (75), Chicago had the highest mean, median, and maximum annual impact per 1,000 km^2 (381, 49, 11,367 persons exposed respectively). Saint Louis had the lowest mean (107) persons exposed per 1,000 km^2 while Omaha and Oklahoma City tied for the lowest median (8) and Birmingham had the lowest maximum (2,844) annual persons exposed per 1,000 km^2 (Table 3.4). A probability of exceedance curve for annual persons exposed (per 1,000 km^2) to significant (Figure 3.6a) and violent (Figure 3.6b) tornado footprints shows how the combined

factors of footprint area and population density affect the annual impact. The greatest impacts are in the MSAs with the greatest population density (Chicago and Dallas/Fort Worth; U.S. Census Bureau, 2010).

Meanwhile, MSA's like Oklahoma City and St. Louis tend to experience reduced impacts due to the sprawling nature of their population distributions (Rosencrants and Ashley, 2015). The annual impacts are greater in Birmingham at lower thresholds due to the higher rural density (Ashley, 2007) while the annual impacts are greater in Omaha at slightly higher thresholds because most of the population in Omaha is confined to the center (U.S. Census Bureau, 2010). This implies that tornadoes are unlikely to impact many people, but when they do, the totals will be higher (Hatzis *et al.*, 2019). In fact, despite being the least populous MSA in this study, significant tornadoes impacting Omaha nearly had the same probability of impacting 20,000 persons as in St. Louis which had over twice the population in 2010 (Table 3.4).

The most impactful tornado from any of the simulations was a 2.5 km wide and 102.5 km long EF3 tornado which was simulated to hit Chicago and impact 165,142 persons. For comparison, from 1979 to 2016, the widest tornado, which impacted Chicago was 1.2 km, and the longest tornado was 49.3 km (SPC, 2017). There were six tornadoes which impacted more than 100,000 persons, all of which impacted either Chicago or Dallas/Fort Worth. Among the 25 most impactful tornadoes all hit either Chicago or Dallas/Fort Worth. More than half (14) were EF4 or EF5 tornadoes with the rest mostly being EF3 and one anomalously wide (1.7 km) EF2 tornado that hit Dallas/Fort Worth.

3.3.3) Tornado Impacts on a Daily Time Scale

A unique use of TorDIS's daily time step is to determine how the impact of an observed tornado footprint compares to the potential impacts that could have occurred under that same

atmospheric environment. As an example of this sort of analysis, a 10,000-year simulation was run, in daily mode, for four significant tornado outbreak days: 5 February 2008, 27 April 2011, 22 May 2011, and 20 May 2013. These runs were conducted over the maximum study area size assuming a 50 km buffer (Figure 3.2). All other model settings and data sets were the same as for the annual runs. Additionally, observed tornado footprints (damage path polygons) were collected from the National Weather Service for three high impact violent tornadoes (an EF4 at Tuscaloosa-Birmingham, Alabama (27 April 2011; Doswell *et al.*, 2012), an EF5 at Joplin, Missouri (22 May 2011; Paul and Stimers, 2012), and an EF5 at Moore, Oklahoma (20 May 2013; Kurdzo *et al.*, 2015) and the total number of persons exposed (using the 2010 SEDAC population grids) to each footprint was calculated.

On 5 February 2008, a large weather system moved over the southeastern United States, producing 87 tornadoes and causing 57 fatalities (Chaney and Weaver, 2010). During this event, one tornadic supercell passed over Nashville, Tennessee producing tornadoes to the southwest and northeast of the city but not downtown (Hatzis *et al.*, 2019). The 10,000-year simulation of this environment yielded a mean of 3 significant tornadoes and a maximum of 27. These significant tornadoes impacted a mean of 1,391 persons per day and a maximum of 222,690 (Table 3.5) with the most impactful footprint exposing 152,967 persons to tornado winds. For comparison, the observed impact over the day was 9,096 persons with the most impactful tornado footprint having an exposure of 2,327 persons. The observed impact was high with only a 2.7% chance that it would have been exceeded on that day.

The probability that the exposure of a simulated violent tornado would be higher than observed, on the date of the event, was 1.6%, 2.2% and 3.3% for Tuscaloosa-Birmingham,

Joplin, and Moore respectively. Given that all three tornadoes impacted more than 5,000 persons, it is not surprising that these exposures are not very likely to be exceeded (Hatzis *et al.*, 2019).

3.3.4) Limitations and Future Development of TorDIS

TorDIS performs well at simulating the spatial distribution of tornadoes. However, its performance is weaker at predicting whether or not a day is a tornado day. Storm initiation is often one of the biggest uncertainties for tornado forecasting in environments with sufficient instability and vertical wind shear (Lock, 2012; Schultz *et al.*, 2014). The current version of the model tends to overpredict storm initiation resulting in nearly double the observed number of tornado days and a resultant bias towards more tornadoes overall. Another model limitation is the bias towards lower intensity tornadoes due to a higher frequency of low helicity environments (Colquhoun and Riley, 1996). This could be because of differences in storm relative helicity between 0000 UTC and the actual time of tornado initiation. Future model runs will increase the temporal resolution for the gridded environmental data (beyond one time step per day) to ensure that the environment varies throughout the day to more accurately represent real tornado production.

The next step in model development is implementing multiple footprints for each tornado to represent the wind field. Tornadoes are given intensity ratings based on the maximum amount of damage over their footprints with much of the footprint area covered by less intense winds (Fricker and Elsner, 2015; Strader *et al.*, 2015). Currently, TorDIS creates only one footprint for each tornado (representing the EF0 wind field), however, future versions will add footprints for each intensity class so that it is possible to know how many persons or buildings are actually experiencing significant (49.2 m s^{-1}) or violent (74.2 m s^{-1}) tornado winds (McDonald and Mehta, 2006b). Intersecting these wind fields over building data, it would be possible to know

whether a particular type of building in the footprint might be destroyed increasing the likelihood of fatalities for anyone inside (Wurman *et al.*, 2007; Brooks *et al.*, 2008). Future work using TorDIS will also seek to estimate potential fatalities in the path of individual tornadoes using such building-based estimates of fatalities and/or regressions taking into account components of vulnerability such as awareness (*e.g.*, off-season or nocturnal tornado), access to shelter, mobility and risk perception (Simmons and Sutter, 2011; Klockow *et al.*, 2014; Paul *et al.*, 2015).

3.4) Conclusion

This study introduces the Tornado Daily Impacts Simulator (TorDIS), a Monte Carlo simulation-based tornado impacts model which distributes tornadoes based on the favorability of the atmospheric environment on a given day. The daily time step is useful as it allows for the prediction of the potential tornado exposures on any day as well as the analysis of the severity of a historical day's tornado exposures (*i.e.*, how many more people might have been impacted by tornadoes on that day). TorDIS builds on the work of Strader *et al.* (2016) by linking tornado distribution and parameters to the environment in which they form, allowing for daily and annual assessments of tornado impacts. Stochastic models such as TorDIS and TorMC (Strader *et al.*, 2016), enable users to understand better the true risk posed by tornadoes through the use of repetition.

To the authors' knowledge, TorDIS is the first spatial tornado impact model to link tornado distribution and parameters to the atmospheric environment in order to enable daily tornado impact analysis. The authors hope that this study can be used as a first step towards research-to-operations for daily impact analysis. The future addition of a fatality estimation module will also hopefully aid in the Federal Emergency Management Agency's goal to project casualties on high-risk severe weather days. Knowledge of potential casualty estimates could

allow emergency managers to plan ahead for these high-risk days in order to prioritize their resources and save lives.

Supporting Information

A supplement to this article is available containing the validation and sensitivity analysis for the regression model (Appendix A). Appendix A has two sections: Tornado Forecast Model Selection and Validation and Predicting Tornado Direction as well as one table (Table A.1: Verification Metrics) and one figure (Figure A.1: Global Sensitivity Analysis of Regression Equations.)

Table 3.1. Parameters and coefficients for logistic regression equations for the tornado probability forecasts. Parameters which were not included in an equation are marked with a dash. Parameters include Convective Available Potential Energy (CAPE), Convective Inhibition (CIN), and Lifted Condensation Level (LCL) for a mixed-layer (lowest 100 mb) parcel; surface height (ELEV); 2 – 4 km lapse rate (LR24); 0 – 1 (VWS1) and 0 – 6 (VWS6) km vertical wind shear; and vertical velocity at 850 mb (ω_{850}) and 500 mb (ω_{500}), and 0 – 3 km Storm Relative Helicity (SRH3). Predictors are the probability of severe weather (svr$_{low}$; svr$_{mod}$) or tornadoes given severe weather (torn$_{low}$; torn$_{mod}$) in low or moderate to high CAPE environments. Here, low CAPE refers to values at or below 500 J kg^{-1} while moderate to high refers to values above 500 J kg^{-1}.

Predictor	Coefficients										
	Intercept	CAPE	CIN	LCL	ELEV	LR24	VWS6	VWS1	ω_{850}	ω_{500}	SRH3
svr$_{low}$	-6.7131	0.0051	-	-	-	-	-	-	-	-1.1814	-
torn$_{low}$	-0.9920	0.0004	-	-0.0004	-	-0.1270	-	0.0793	-0.3829	-	-0.0024
svr$_{mod}$	-10.3173		0.0077	-	-	0.7604	0.0996	-	-	-0.9747	-
torn$_{mod}$	-3.3608	-	0.0020	-0.0007	0.0016	-	0.0321	0.1369	-0.5515	-	-

Table 3.2. Tornado attributes from a 1,000-year TorDIS simulation over the Oklahoma City Metropolitan Statistical Area. Attributes presented include tornado magnitude (EF scale), count, mean length, mean width, and mean direction.

Magnitude	Count	Mean Length (km)	Mean Width (m)	Mean Direction (∘)
EF0	3,279	3.1	90.8	78.0
EF1	2,594	8.5	240.5	77.6
EF2	1,021	17.5	483.3	75.4
EF3	414	34.4	993.3	77.7
EF4	96	41.6	1,312.1	75.2
EF5	4	58.6	1,459.4	101.8
All	7,408	9.2	264.3	77.4
Significant (EF2+)	1,535	23.7	675.2	76.1
Violent (EF4+)	100	42.3	1,318.0	76.2

Table 3.3. Impact analysis from a 1,000-year TorDIS simulation over the Oklahoma City Metropolitan Statistical Area. Results are organised by tornado magnitude and include annual occurrence probability, return period, mean number of persons impacted by an individual tornado, and the mean number of persons impacted per year.

Magnitude	Annual Occurrence Probability	Return Period (years)	Mean Tornado Impact (Persons)	Mean Annual Impact (Persons)
EF0	3.279	0.305	15	48
EF1	2.594	0.386	113	294
EF2	1.021	0.979	410	419
EF3	0.414	2.415	1,312	543
EF4	0.096	10.417	2,756	265
EF5	0.004	250.000	1,375	6
All	7.408	0.135	212	1,574
Significant (EF2+)	1.535	0.651	802	1,232
Violent (EF4+)	0.100	10.000	2,701	270

Table 3.4. The annual number of persons impacted by significant tornado winds per 1000 km^2 over select metropolitan statistical areas during a 1000 year TorDIS simulation using the 2010 SEDAC gridded population counts. Summary statistics include the normalised count of significant tornadoes, mean, median, maximum, and standard deviation of annual persons impacted, and the probability that a tornado hitting each metro area would impact more than 5000 (EP$_{5000}$) or 20000 persons (EP$_{20000}$).

Metropolitan Statistical Area	Normalized Count	Mean	Median	Maximum	Standard Deviation
Oklahoma City, OK	102	127	8	3,125	348
Dallas/Fort Worth, TX	103	331	45	6,467	736
Chicago, IL	75	381	49	11,367	959
Birmingham, AL	92	115	24	2,844	272
Omaha, NE	103	109	8	5,030	381
St. Louis, MO	69	107	12	3,190	277

Table 3.5. TorDIS model results from a 10,000-year simulation of all tornadoes over the domain on the specified days. Summary statistics include the mean, median and maximum values for both tornado count (by magnitude) and persons exposed per tornado.

Date	Magnitude	Tornado Count			Persons Exposed		
		Median	Mean	Maximum	Median	Mean	Maximum
5 February 2008	*All*	4	10	72	46	1,020	244,878
	Significant (EF2+)	2	3	27	151	1,391	222,690
	Violent (EF4+)	1	1	5	321	1,738	93,619
27 April 2011	*All*	4	9	72	70	817	79,897
	Significant (EF2+)	2	3	24	221	1,102	68,045
	Violent (EF4+)	1	1	5	472	1,569	37,201
22 May 2011	*All*	2	4	67	11	292	55,320
	Significant (EF2+)	1	2	26	65	592	55,247
	Violent (EF4+)	1	1	4	315	1,369	55,247
20 May 2013	*All*	4	10	72	34	652	77,381
	Significant (EF2+)	2	3	20	106	889	77,367
	Violent (EF4+)	1	1	3	272	1,355	77,367
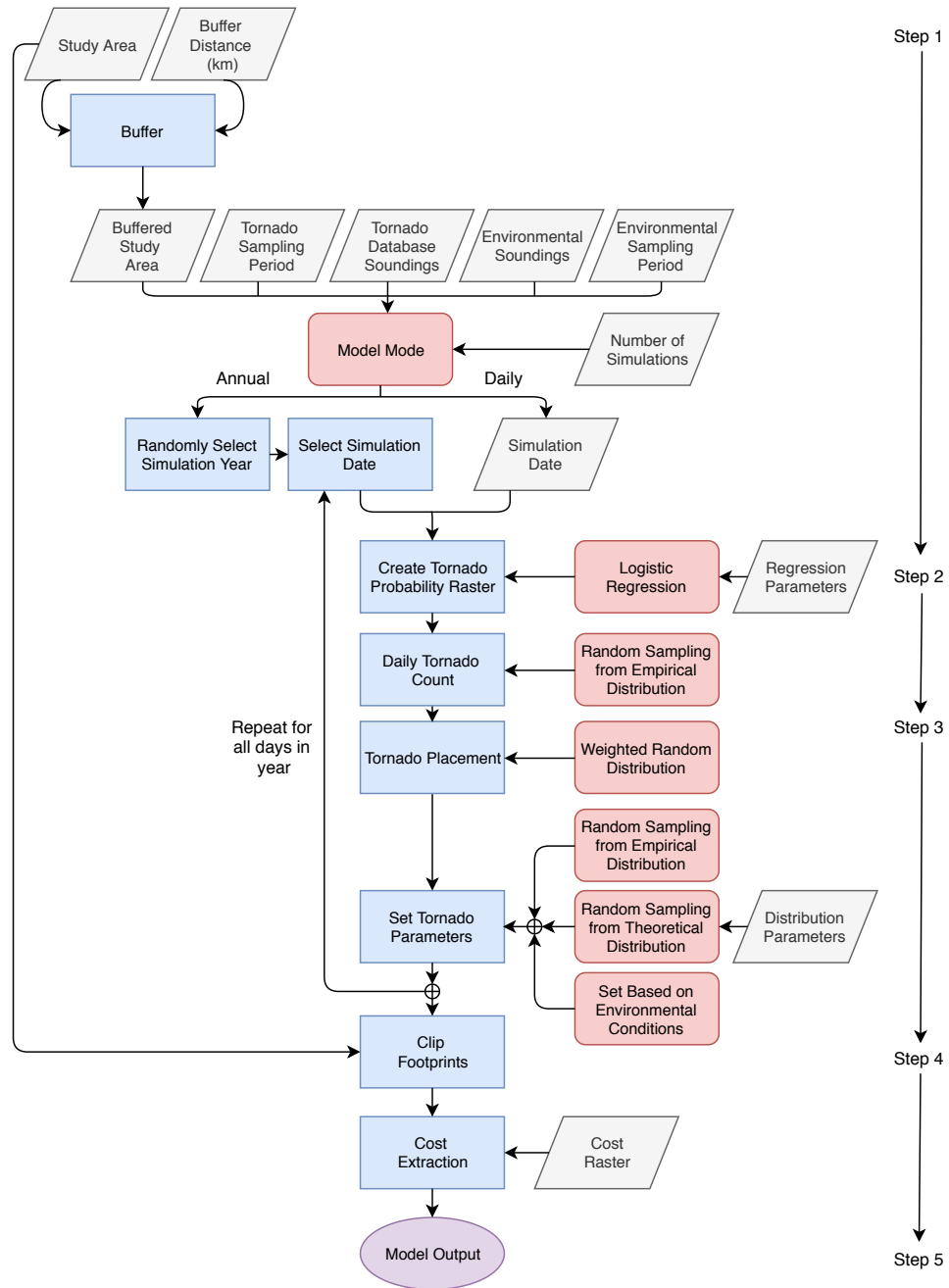
Figure 3.1. Tornado Daily Impacts Simulator (TorDIS) model flow chart. Rhombus shapes represent model input, squares represent model processes, rounded rectangles represent model decisions, and the oval represents the model output. Modelled after Figure 1 in Strader *et al.* (2016) for comparison.
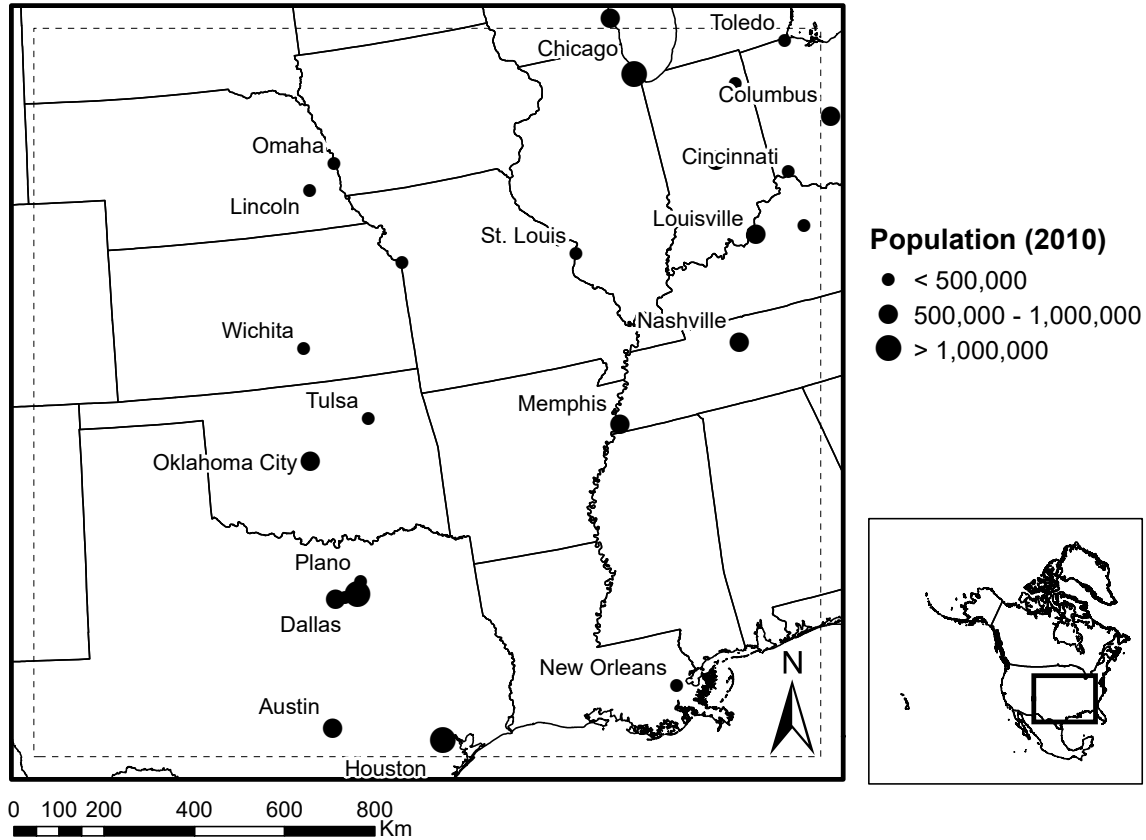
Figure 3.2. Population for major cities within the model domain. The dashed line refers to the maximum study area size for a model run assuming a 50 km buffer to reduce edge effects.

a.)

b.)
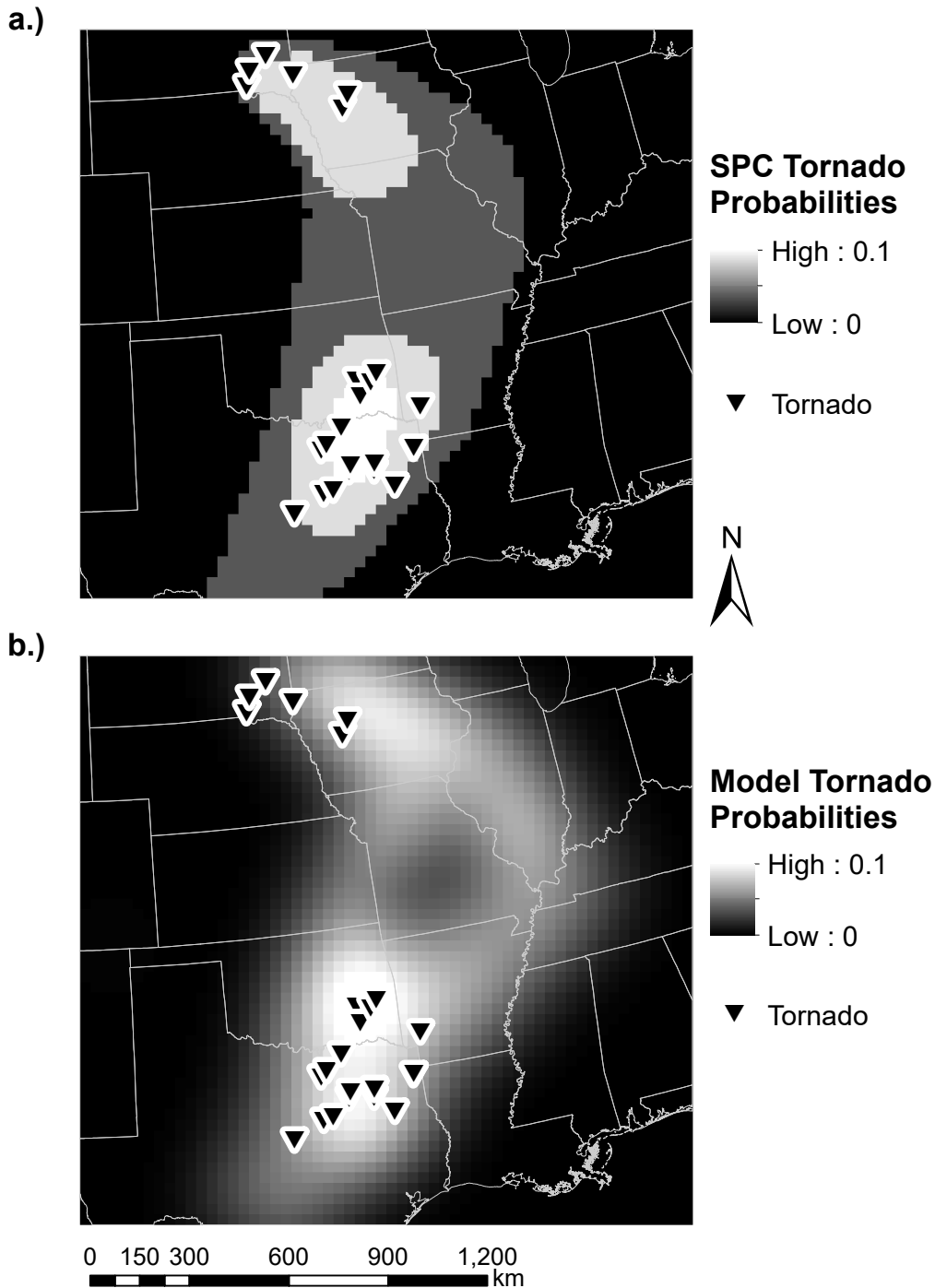
Figure 3.3. Comparison between probabilistic tornado forecasts for 2015 May 10 generated by the SPC (a) and TorDIS (b). Tornado reports for 2015 May 10 are plotted as upside-down triangles to show the accuracy of the forecast areas. Probabilistic forecasts are smoothed using a Gaussian low-pass filter with a 120 km standard deviation. All smoothed forecasts with a probability of less than 2% are set to 0 to match the SPC's practice of not showing probabilities below 2% (Hitchens *et al.*, 2013).

SVRGIS Footprints 1979-2016

a)

TorDIS Random 38 Years Footprints

b)

TorDIS 38 Most Impactful Years Footprints

c)

2010 Population

0
1 - 5
6 - 10
11 - 50
51 - 100
101 - 250
251 - 500
> 500

N

0 20 40   80   120   160
Km
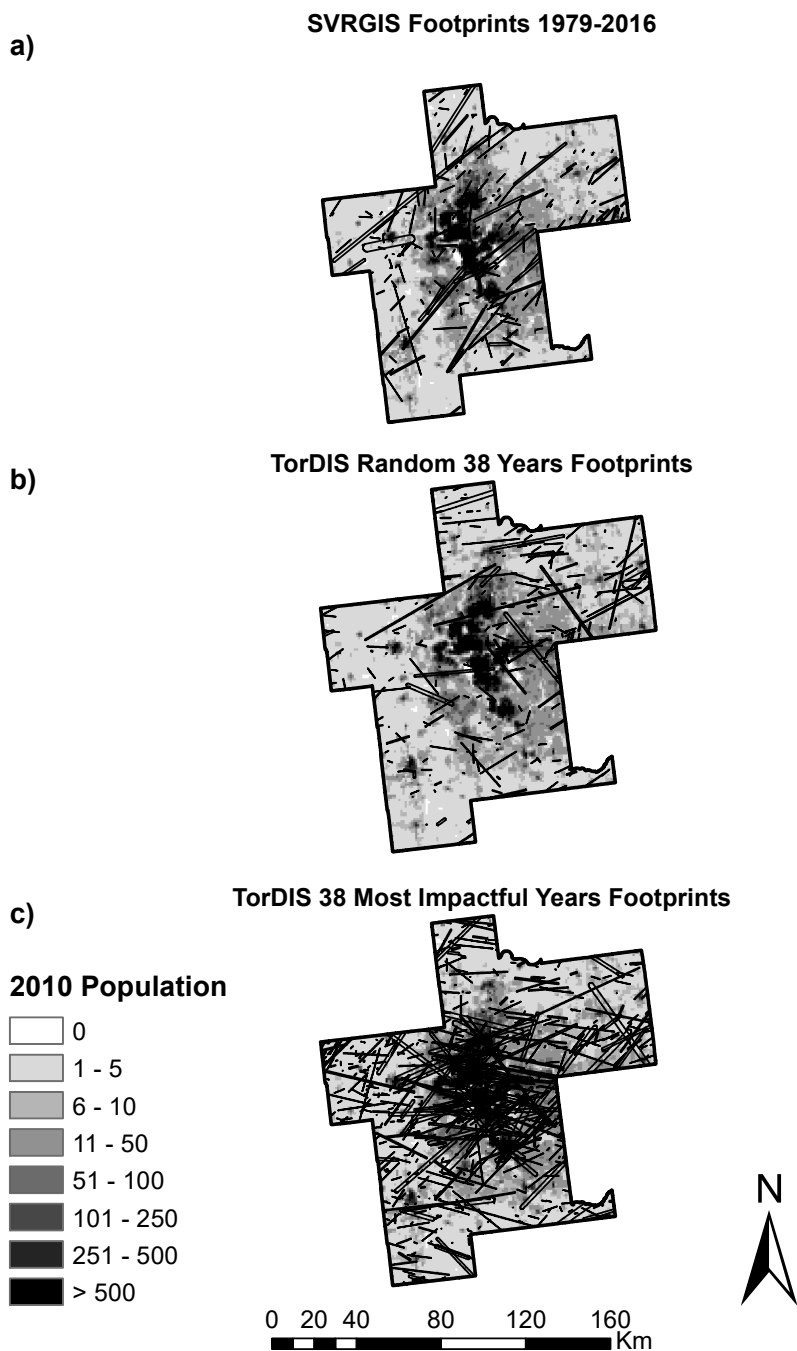
Figure 3.4. (a) Observed tornado footprints for 1979 to 2016 from SVRGIS over gridded 2010 population counts from SEDAC for the Oklahoma City Metropolitan Area. (b) Same as (a) but for a random sample of 38 years of simulated tornado footprints. (c) Same as (a) but for all tornado footprints from the 38 most impactful years over the Oklahoma City Metropolitan Area (in terms of total persons exposed to tornado winds per year). Sample sizes were chosen to be over 38 years since it matches the length of 1979 to 2016 observed period. The figure is modelled after Figure 4 from Strader *et al.* (2016).
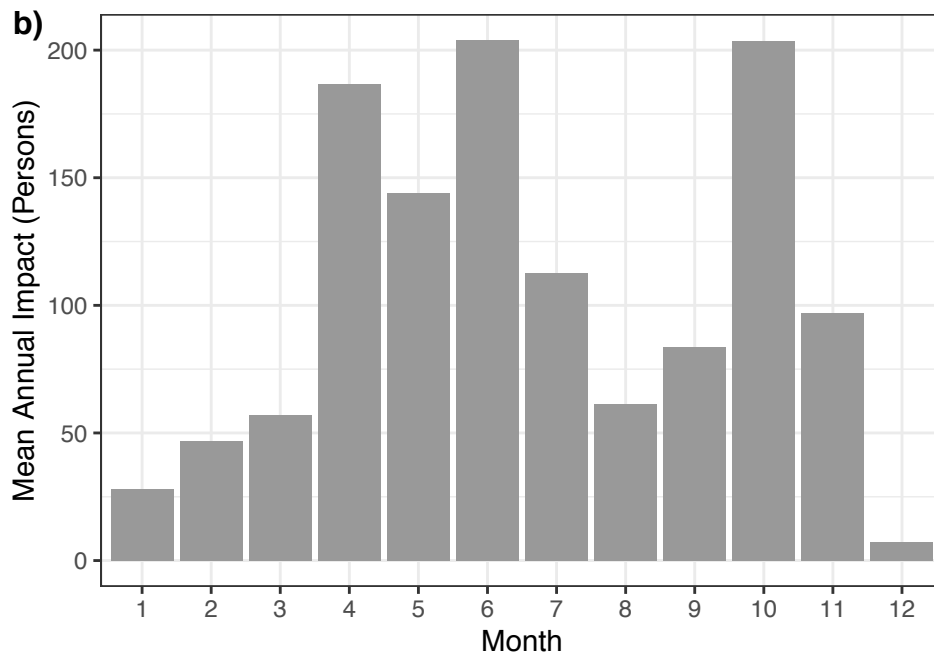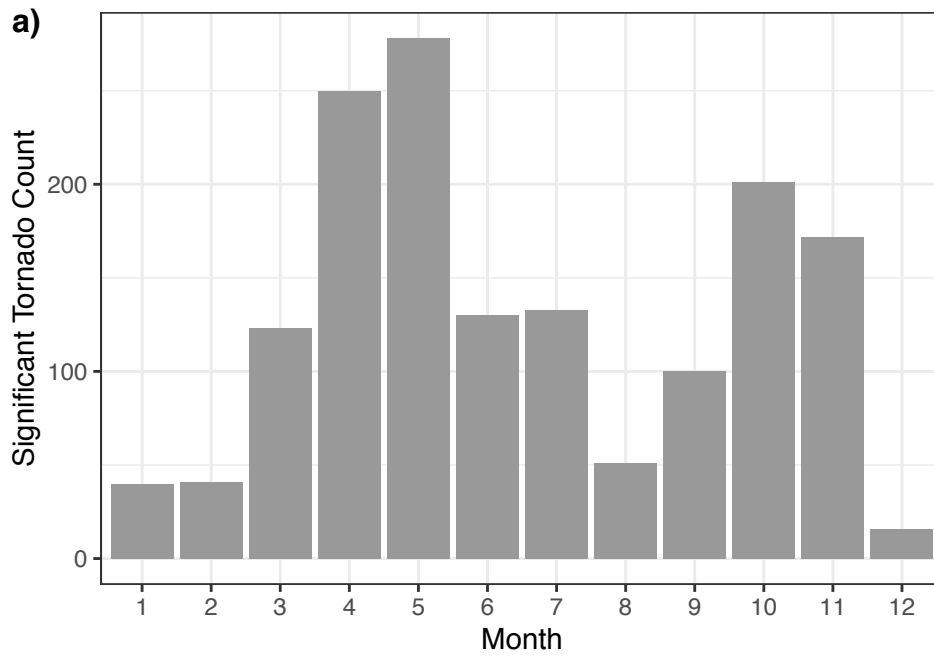
Figure 3.5. Monthly variation in significant tornado counts and mean annual persons exposed to significant tornado winds over the Oklahoma City Metropolitan Statistical Area.
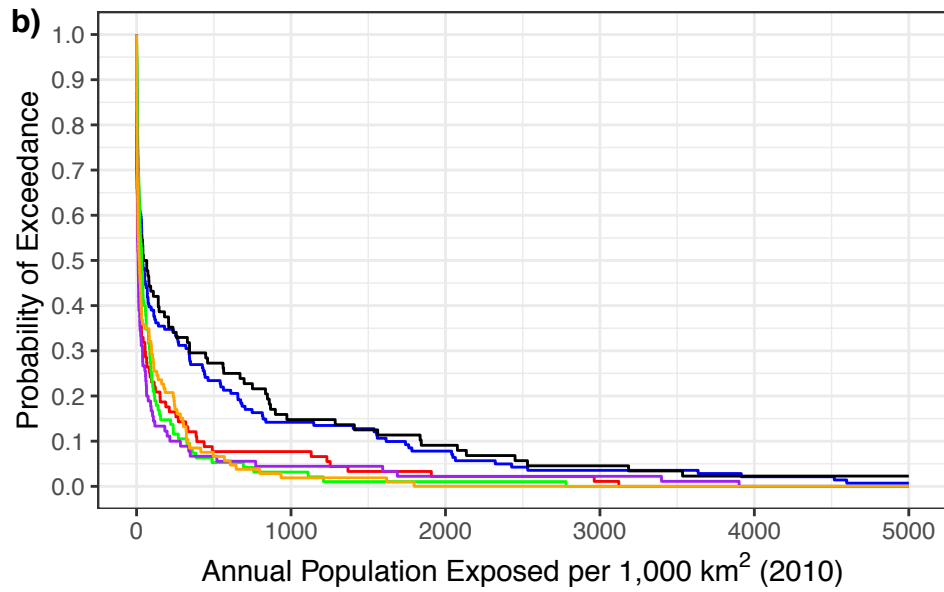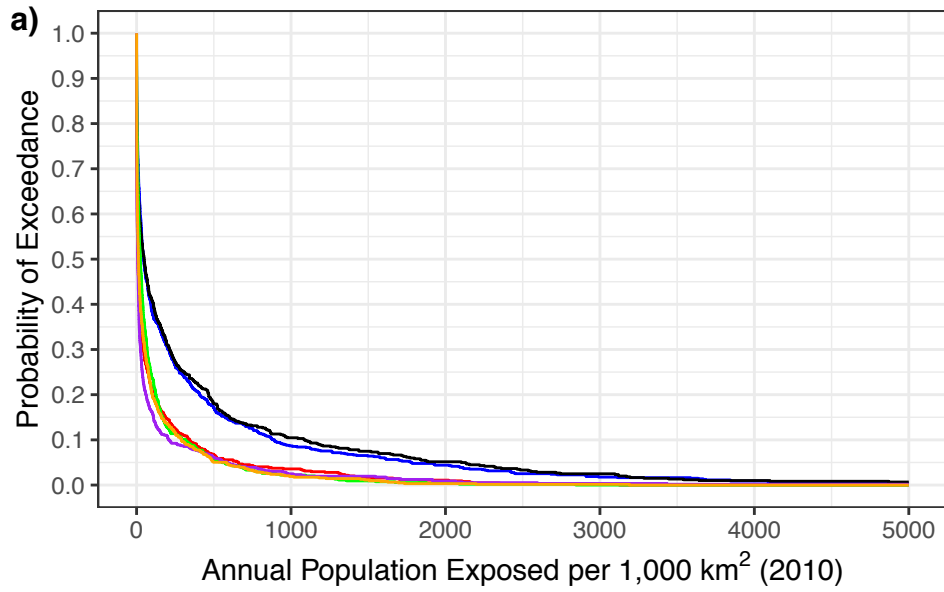
Figure 3.6. The probability that a significant (a) or violent (b) tornado would impact more persons than a set threshold for a 1,000-year simulation over the Birmingham, Dallas/Fort Worth, Omaha, Chicago, Oklahoma City and St. Louis Metropolitan Statistical Areas.

Chapter 4 Analyzing the Combined Effects of Urbanization and a Changing Climate on Tornado Exposure: A Case Study for the Oklahoma City Metropolitan Area

Abstract

In recent years the U.S. has seen a significant increase in the frequency of weather-related disasters and their attendant consequences. The exact role climate change will play on the frequency and intensity of severe weather is widely debated, but it will likely become more common. At the same time, urban areas continue to expand in size and population, putting more people at risk of severe weather events. One example of this is Oklahoma City, the fourth largest city in the U.S. (by area) which has seen a population growth of over 39% since 2000 and frequently experiences tornadoes. Population growth within Oklahoma City is expected to continue into the year 2050 and beyond. We tested how tornado exposure over the Oklahoma City Metropolitan Area may change by 2050 using a business-as-usual projection of population and projected climate data from the North American Regional Climate Change Assessment Program (NARCCAP). To test the relative contribution of climate change and urban development on tornado impacts, we simulated 1,000 years of significant tornadoes using the Tornado Daily Impacts Simulator under four different scenarios: control, climate change only, urban development only and urban development and climate change. We found no significant change in either tornado risk, or impacts, due to climate change, over the Oklahoma City Metropolitan Area while tornado impacts were likely to increase due to the effects of urban development. Our findings are in agreement with other studies suggesting that urban development plays a larger role in changing tornado impacts than does climate change.

4.1) Introduction

Following World War II, the global population began to shift from mostly rural (30% urban in 1950) to more urban (55% urban in 2018), and this trend is likely to continue. The projections from the United Nations suggest that the global urban population will reach 68% by 2050 (United Nations, 2018). The development in population distributions has changed and affected land use and land cover change considerably. Over the last 30 years global urban areas have expanded at a rate, on average, twice that of the population growth (Seto *et al.*, 2011), showing an increase in area consumption per capita (Scheidel and Sorman, 2012; Sharma *et al.*, 2012). In 2018, 82% of the North American population was living in urban areas (United Nations, 2018) and projections indicate that urban land cover is likely to nearly double by 2030 (Seto *et al.*, 2012). Urban expansion in the United States has been characterized by mostly sprawl (as defined by Hamidi and Ewing (2014) where the predominant development character is low-density housing spread over a large land area) since the mid-1940s with populations moving from the urban core to more suburban locations (Danielsen *et al.*, 1999; Handy, 2005; Radeloff *et al.*, 2005).

Land use and land cover change due to urban expansion have considerable effects on both people and the environment. Urban expansion is responsible for increases in habitat loss and species extinction (Foley *et al.*, 2005; Seto *et al.*, 2012; IPCC, 2019), loss of agricultural land (Seto and Ramankutty, 2016; d'Amour *et al.*, 2017), pollution (Frumkin, 2016), and income disparity (Powell, 1998). Additionally, urban expansion has implications for the impacts of hazards. As population numbers increase, more land is developed, and more expensive real estate is built, the potential for high-impact hazard events increases (Ashley and Strader, 2016; Strader *et al.*, 2017a, 2018). Studies have shown an increase in economic losses from weather-related

100

disasters, including tornadoes in recent decades (Simmons *et al.*, 2013; Smith and Katz, 2013). While increasing population numbers alone can influence hazard impacts (Ashley *et al.*, 2014; Ashley and Strader, 2016), the distribution of this additional population is an important driver as well. Sprawl-type development covers more of the landscape in low-density housing, while infill-type development builds up existing developments (*i.e.*, taller buildings) or fills in vacant plots within an existing urban development (Hamidi and Ewing, 2014; Laidley, 2016). Sprawl-type development tends to increase the overall risk posed by hazards as the urban area expands placing more people in potential impact zones (Stone *et al.*, 2010; Strader *et al.*, 2018). Infill-type development does not expand the perimeter of the developed area. However, it does increase the population density within the urban area. The net result of infill is that the overall risk is lower than for sprawl; however, the maximum potential impacts increase when hazards occur in densely populated areas due to increases in population density (Wurman *et al.*, 2007; Ashley *et al.*, 2014).

In addition to urban expansion, another process is likely to affect the potential for hazardous conditions in urban environments, namely climate change. There is much uncertainty in how climate change may impact tornado frequency or intensity. However, studies have already shown changes in tornado outbreak frequency and interannual variability (Brooks *et al.*, 2014; Elsner *et al.*, 2014) and climate projections suggest an increase in the number of days favorable for severe weather development (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2013; Gensini *et al.*, 2014b). In addition to changes in frequency, a study by Gensini and Brooks (2018) suggests that the spatial distribution of tornadoes has also changed with tornadoes becoming more frequent in the southeastern United States and less frequent in the traditional Tornado Alley (including Oklahoma). The projected continuing growth of the Oklahoma City metropolitan area

(Barker, 2012) will likely place an increasing number of people and homes in the path of tornadoes (Ashley and Strader, 2016; Strader *et al.*, 2017a).

Oklahoma City is a prime example of a sprawling city with an extensive area covered by low-density housing, and it is located in a high-risk area for tornadoes, which makes it an excellent location to study the combined and isolated effects of urbanization and climate change on tornado exposure. Oklahoma City experiences an average (over the period 1950 – 2017) of seven tornadoes per year within the metropolitan statistical area and a violent tornado (rated four or higher on the Enhanced Fujita Scale) occurring once every four years (SPC, 2017). Oklahoma City is the state capital and largest city in the state of Oklahoma. The Oklahoma City Metropolitan Statistical Area encompasses the seven counties surrounding the city and, in 2010, had a population of 1,252,987 (U.S. Census Bureau, 2010). Several high fatality tornadoes have impacted the Oklahoma City Metropolitan Area including the May 3, 1999, EF5 tornado that killed 36 people (Brooks and Doswell, 2002) and the May 20, 2013 EF5 tornado that killed 24 including seven children at an elementary school (Atkins *et al.*, 2014).

The purpose of this case study is to use a newly developed, environmentally driven tornado impacts model to assess how the isolated and combined effects of urbanization and climate change may alter tornado exposure over the Oklahoma City Metropolitan Area in the mid-21st century. In a similar study by Strader *et al.* (2017a), the authors assumed tornado risk would increase at a flat rate across the central US, while studies have already suggested that tornado risk is shifting to the east as the west becomes drier (Gensini and Brooks, 2018; Seager *et al.*, 2018). Hence, while other studies have looked at the impacts of climate change on tornado risk (Diffenbaugh *et al.*, 2008; Strader *et al.*, 2017a), the novelty of this study is that it uses future atmospheric environmental data, from a climate model, to directly determine how climate

change may influence tornado risk. This distinction is crucial as it allows us to determine how tornado risk may vary spatially in the future.

4.2) Data and Methods

4.2.1) Study Area

The study area covers the three counties (Oklahoma, Cleveland, and Canadian) containing the majority of the City of Oklahoma City in the State of Oklahoma (Figure 4.1). Hereafter, OKCMA is used to refer to the three-county metro area of Oklahoma City. The three counties have a total area of 5,554 km^2, with an estimated 2010 population of 1,171,185 people and a population density of 197 persons km^{-2}. The most densely populated county is Oklahoma County (391 persons km^{-2}) while the least populated is Canadian County (57 persons km^{-2}). As of 2010, the OKCMA was mostly characterized by the low-density development typical of a sprawling city (Hamidi and Ewing, 2014; Laidley, 2016; Figure 4.2a). Here, we classify rural parcels as any parcel with a population density less than 518 persons km^{-2} with all other parcels classified as urban. Urban parcels are further separated into low-density urban (urban-low; 518 – 9,064 persons km^{-2}) and high-density urban (urban-high; 9,065 persons km^{-2} or greater). These definitions correspond with those of Laidley (2016) with the urban-high classification, including Laidley's moderate and high-density classes. Most of the urban area was located in Oklahoma and Cleveland Counties (70% and 60% urban respectively) with Canadian County being mostly of a rural character (44% urban; Table 4.1). All three counties have experienced population growth since 2010 and are projected to continue to grow into the mid-21st century (Barker, 2012).

The OKCMA is situated in the heart of the traditional Tornado Alley, a high-risk area for significant tornadoes (Concannon *et al.*, 2000; Dixon *et al.*, 2011; Doswell *et al.*, 2012). During the 1990 – 1999 period, the OKCMA experienced ten significant tornadoes over four days. Cleveland County experienced the most significant tornadoes during this period (5) with Canadian and Oklahoma Counties experiencing three each (SPC, 2017). One tornado, the Bridge Creek-Moore EF5 on 3 May 1999 (Brooks and Doswell, 2002), impacted both Cleveland and Oklahoma Counties. The average number of days during 1990 – 1999 with atmospheric environmental conditions favorable for tornado development was high (~170 days) over OKCMA with little variation (~3 days) over the metropolitan area (Figure 4.2c).

4.2.2) Potential Future Urbanization Patterns in the Oklahoma City Metropolitan Area

We used the TARGET plug-in within the ENVISION framework (Bolte *et al.*, 2007; Spies *et al.*, 2017) to project population densities over OKCMA for the period 2016 to 2050. The TARGET plug-in uses a baseline distribution of population densities specified on the parcel level and works by increasing the population density of this density-surface towards a maximum capacity using a fixed growth rate on an annual time step. We used an annual population density growth rate of 2.5% and a further set spatial allocation rules to distribute population growth preferentially to areas proximal to certain desirable spatial features such as the city center or bodies of water, based on established procedures in land change models (*e.g.*, Dorning *et al.*, 2015; Koch, 2010). Simulations were carried out on an annual basis between 2011 and 2050 with the original 2010 parcel-level population densities derived from the 2010 census block data (U.S. Census Bureau, 2010). This scenario describes a Business as Usual (BAU) pattern (*i.e.*, a continuation of the historically observed spatial sprawl patterns) in the OKCMA. Table 4.2 provides the allocation rules used in the population density projections. Dilekli *et al.* (in

104

preparation) discuss model validation and caveats. Figure 4.2 displays the simulation results resulting from the TARGET application.

4.2.3) North American Regional Climate Change Assessment Program Current and Future Climate Scenarios

We used current (1990 – 1999) and future (2045 – 2054) gridded atmospheric environmental data from the North American Regional Climate Change Assessment Program (NARCCAP) project (Mearns *et al.*, 2014). NARCCAP is an international effort to create downscaled, high-resolution simulations of various atmospheric environmental variables for use in climate change research. The NARCCAP gridded data sets were produced by running several Regional Climate Models (RCMs) forced by a suite of atmosphere-ocean general circulation models (AOGCMs). The AOGCMs in the future climate scenarios were forced with the Special Report on Emission Scenarios (SRES) A2 emissions scenario, which is one of the higher Representative Concentration Pathway (RCP) emissions scenarios (Nakicenovic *et al.*, 2000), but still one that can be adapted to by humans (Gensini *et al.*, 2014b) The spatial resolution of the data is 50 km, and the domain is the entire conterminous United States and parts of Canada and Mexico. Following a study on future convective environments in North America by Gensini *et al.* (2014b), we use the Weather Research and Forecasting Model (WRF-G) RCM forced with the Community Climate System Model (CCSM3) version 3.0 for our data. We collected the following variables at one daily time step (0000 UTC): specific humidity, temperature, vertical velocity, and u- and v-component winds. These variables were pre-processed in Python using the SHARPpy library (Halbert *et al.*, 2015) to obtain the severe weather diagnostic parameters required by the Tornado Daily Impacts Simulator (Hatzis *et al.*, in review).

4.2.4) Tornado Daily Impacts Simulator

The Tornado Daily Impacts Simulator (TorDIS) is a spatially explicit and

environmentally driven model that simulates the impacts of many tornadoes on a specified cost

surface at a daily time step (Hatzis *et al.*, in review). Hatzis *et al.* (in review) describe TorDIS, in

detail, along with use examples and model limitations. Here, we used TorDIS to simulate 1,000

years' worth of tornado footprints over the state of Oklahoma and surrounding areas, assuming a

100 km buffer (Figure 4.1). We limited the number of runs to 1,000 years due to computational

constraints. We then clipped the runs to the OKCMA. The simulation produced tornadoes of all

magnitudes, but this study focuses on significant (rated two or higher on the Enhanced Fujita

(EF) scale) tornadoes as they cause the most fatalities (Ashley, 2007; Simmons and Sutter,

2011). We randomly selected tornado path lengths and widths from Weibull distributions. We

derived the environmental data from NARCAPP and used the parcel-based census population

data for the OKCMA as the cost surface. The tornado database used in this study is the SVRGIS

database from the Storm Prediction Center for 1979 to 2016, the same as used in Hatzis *et al.* (in

review).

Following Hatzis *et al.* (in review), tornado favorability and spatial distribution were

determined using logistic regressions. We used the same regression variables as in Hatzis *et al.*

(in review) but we specified the logistic regressions using the 1990 – 1999 NARCCAP and

SVRGIS data over the central and southern United States (see Figure 2; Hatzis *et al.*, in review),

instead of the North American Regional Reanalysis (NARR) data. The regressions were trained

on the 1987 – 1997 observations and validated against the 1998 and 1999 observations to

confirm that the regressions performed acceptably well. We tested multiple thresholds, to

determine whether a grid cell was considered favorable for severe weather, to find the thresholds

that maximized the probability of detection (POD; and minimized false alarm rates (FAR) for

both tornado days and tornado distributions. The best threshold proved to be 0.00509 for the

severe weather regressions and 0.009 for the tornado regressions. The POD for the spatial

distribution of tornadoes was 62.5% with a FAR of 99.9%. The POD for tornado days was

81.9% with an FAR of 53.5%. The PODs were lower compared to the findings of Hatzis *et al.*

(in review) when using the North American Regional Reanalysis data, but still high enough to be

acceptable for our purposes. All other model settings were as in Hatzis *et al.* (in review).

4.2.5) Study Design

In this spatially explicit, scenario study we combined an urbanization scenario and

climate change scenarios with a recently-developed model, TorDIS, to test how tornado exposure

varies between 2010 and 2050 due to the combined and isolated effects of urbanization and

climate change. The model output consists of spatial vector (polygon) data, representing the

footprints of all simulated tornadoes over the OKCMA. We simulated tornado footprints from 10

years of atmospheric environmental data with multiple simulations using each year of data. Each

tornado has a population exposure representing the number of people who reside in the footprint

during the decade in which the tornado occurs (*i.e.*, 2010 or 2050). For this purpose, we applied a

6-step procedure (Figure 4.3).

Figure 4.3 shows the steps performed at each model time step. Each time step begins with

a selection of the input data. All simulations use the 100 km buffer around the study area (Figure

4.1) as well as the tornado data with environmental information (SVRGIS SND; Hatzis *et al.*, in

review). The atmospheric data used in the model run depends upon the scenario: scenarios using

current climate use the NARCCAP 1990 – 1999 data while scenarios using future climate

projections use the NARCCAP 2045 – 2054 data. The population data (cost surface) used in the

model run are also selected based on the scenario: we used the 2010 population data for the no urbanization scenarios while we used the 2050 population projections for the urbanization scenarios. We used the 1990 – 1999 atmospheric environment data from NARCCAP as it was the most recent period available. We use this in conjunction with 2010 population counts instead of counts from the 1990 or 2000 census because 2010 was the starting date for the urbanization scenario (Dilekli *et al.*, in preparation). Additionally, the tornado climatology indicates no significant change in significant tornado occurrence between 1990 and 2010 (Brooks *et al.*, 2014; Elsner *et al.*, 2014).

Once the atmospheric data and population data are selected, we followed the standard TorDIS model procedures (Hatzis *et al.*, in review): (1) we use logistic regressions to create tornado probability rasters; (2) if the day is determined to be favorable for tornado development, there is at least 5,000 km^2 of favorable area (grid cells with a tornado probability of 0.009 or greater; Hatzis *et al.*, in review), we place a randomly drawn number of tornadoes over the buffered study area using a random distribution weighed by the tornado probability raster; (3) we select tornado intensity, path length, width, and direction, and create a footprint for each simulated tornado; (4) we clip the footprints to the study area; (5) we intersect the footprints with the population data and sum up the total exposed population across the footprint (Figure 4.3). While the model operates at a daily time step, each simulation begins with a year randomly selected from the atmospheric data (*e.g.*, between 1990 and 1999) and each day during that year is simulated. We run the simulation for 1,000 years.

We tested four scenarios to improve our understanding of the isolated and combined effects of urbanization and climate change on tornado impacts. First, we established a baseline for tornado impacts over the OKCMA. This baseline tested the impacts under current (1990 –

1999) climate and current population (2010) over the OKCMA. The second scenario tested the impacts under future (2045 – 2054) climate and current population to isolate the effects of climate change alone. The third scenario used current climate and future population (2050) to isolate the effects of urbanization. The final scenario tested the combined effects of urbanization and climate change by using future climate and future population.

We assessed the tornado impacts both for individual significant tornadoes as well as for annual total impacts of all significant tornadoes. We assessed the impacts at the metropolitan area level and the county level. Annual impacts were normalized by county (or total) area to aid in the comparison between counties. Tornado impact statistics included mean, median, 95th and 99th percentiles and maximum number of persons exposed. Additionally, we calculated the probabilities that at least 5,000 or 20,000 persons would be impacted by individual tornadoes. We chose these thresholds as they represent tornadoes that are likely to have multiple to many fatalities. A study of select violent (EF4+) tornadoes by Brooks at others (2008) showed fatality rates of 0.1% to 1.9% of those exposed. An exposure of 5,000 (20,000) persons would suggest at least 5 (20) fatalities for a given violent tornado. While fatality rates for weaker significant tornadoes (EF2-3) are lower (Ashley, 2007), these thresholds are still applicable for stronger significant tornadoes, so we chose to use them. Comparisons between counties and scenarios were conducted visually using probability of exceedance curves (Strader *et al.*, 2016; Hatzis *et al.*, in review). Direct comparisons of the summary statistics were also used to determine which scenarios led to the greatest impacts, following Strader *et al.* (2017a). We used the non-parametric Mann-Whitney test, with a 95% confidence value for statistical comparisons (Hollander *et al.*, 2013).

We assigned each simulated tornado a type (land use class) based on the most densely populated parcel it impacted and not the dominant land use type (*i.e.*, a tornado impacting mostly rural parcels and only a single urban-high (densely populated) parcel would be classified as urban-high). This was done given that most tornado impacts occur in the more densely populated areas (Edwards *et al.*, 2013). The counts of tornadoes (by type) were tallied for each county and for the OKCMA as a whole and normalized by county area (per 2,500 km^2 to approximate the size of the largest county, Canadian County) for each scenario. We averaged the counts between scenarios using the same population estimates (baseline with CC and UD with CCUD) to show only the influence of urban development on changes in land use status of significant tornadoes in the OKCMA (Table 4.3). We did this because climate change did not significantly alter the tornado counts.

4.3) Results

4.3.1) Projections of Future Tornado Risk and Urban Development over OKCMA

Our projections indicate that the population of the OKCMA will nearly double to 2,252,564 persons by 2050 (Table 4.1). The BAU scenario settings used for the 2050 population projections continues the sprawling trend with low-density urban development expanding in all counties at the expense of rural area (Figure 4.2b). The simulation results indicate the greatest amount of urban development for Cleveland County with a 10.3% loss of rural area while the least development is projected to occur in Canadian County (8.6% loss; Table 4.1). The number of tornado favorable days is projected to increase to ~180 days by the 2045 – 2054 period; however, the variation across OKCMA is projected to decrease to ~1 day (Figure 4.2d). In spite of the slight increase in the number of tornado favorable days over the OKCMA, we found that the mean annual number of simulated significant tornadoes and significant tornado days

110

decreased slightly in all three counties with the greatest decrease in both significant tornadoes

(0.6) and significant tornado days (0.6) in Cleveland County (Table 4.4). This decrease was not

statistically significant and may have been the result of random effects and a greater tornado

favorability to the west of the study area; however, it is in line with findings by Gensini and

Brooks (2018), which suggest a reduction in tornado activity over Oklahoma that may continue

into the future.

4.3.2) Baseline Simulation

For the baseline simulation, we used atmospheric environments from 1990 – 1999 with

population estimates from the 2010 census. Figure 4.4a shows an example of the ten most

impactful years' (highest annual exposures) worth of significant tornadoes over the OKCMA for

the baseline simulation. Hereafter, we estimate all annual exposure totals per 2,500 km^2 to enable

comparison between different sized counties. During an average baseline year over OKCMA, a

median (mean) of 1,328 (4,541) persons were impacted by significant tornadoes with the most

impactful years (95th and 99th percentile) affecting 21,284 and 33,322 persons respectively.

Exposure was greatest (least) in the most (least) populated counties with a median exposure of

3,929 persons in Oklahoma County compared to 345 persons in Canadian County. The

simulation results indicated that the probability that the annual exposure over the OKCMA

would exceed 5,000 or 20,000 persons was 28.0% and 5.8% respectively with the highest

(lowest) probabilities in Oklahoma County (Canadian County; Table 4.5). For individual

significant tornadoes, the simulation results indicated a median (mean) exposure of 172 (2,523)

persons with a maximum exposure of 93,448 persons for the OKCMA. The maximum impact for

an individual significant tornado was highest in Oklahoma County (88,059 persons) and lowest

in Canadian County (20,135 persons). The probability that an individual significant tornado over the OKCMA would impact more than 5,000 (20,000) persons was 12.2% (3.1%; Table 4.6).

4.3.3) Climate Change Scenario

The climate change (CC) scenario simulations combined the 2045 – 2054 projected atmospheric environments with the baseline population estimates from the 2010 census to analyze the isolated effects of a changing climate on tornado exposure. Figure 4.4b shows an example of the ten most impactful years' for the CC scenario over the OKCMA. The slight decrease in the number of significant tornadoes under the CC scenario led to a reduction in both individual and annual significant tornado impacts over OKCMA. The median (mean) annual impact decreased to 1,120 (3,984) persons while the most impactful years decreased to 17,757 and 28,305 persons, respectively (Table 4.5). As for the baseline simulations, the annual impacts were highest in Oklahoma County and lowest in Canadian County. For impacts of individual significant tornadoes, the median (mean) exposure over OKCMA decreased to 145 (2,187) persons while the maximum impact decreased to 63,173 persons. The simulations for the CC scenario indicated that the probability that an individual significant tornado over the OKCMA would impact more than 5,000 (20,000) persons decreased to 11.3% (2.7%) as compared to 12.2% (3.1%; Table 4.6) for the baseline simulation (see section 3.2). The difference between the simulations for the baseline and CC scenario exposures were very small (Figure 4.5) and not statistically significant. While none of the differences in exposures were significant, the individual significant tornado exposures in Canadian County were nearly significant ($p = 0.06$).

4.3.4) Urban Development Scenario

The urban development (UD) scenario combined the 1990 – 1999 atmospheric environments for the baseline period with the projected population density estimates for 2050 to

simulate and analyze the isolated effects of population growth and spatial distribution on tornado exposure. Figure 4.4c shows an example of the ten most impactful years' for the UD scenario over the OKCMA. The UD scenario simulations indicated near doubling of the population in OKCMA by 2050 (Figure 4.2). This had a significant effect on tornado impacts; the median (mean) annual impact over OKCMA increased to 4,175 (8,576) persons with the most impactful years increased to 34,688 and 57,559 persons, respectively. Mean and median annual impacts were still highest in Oklahoma County (7,266 and 19,387 persons respectively) and lowest in Canadian County (3,889 and 6,465 persons respectively; Table 4.5). For individual significant tornadoes over the OKCMA, the mean (median) impact increased to 1,147 (4,693) persons with a maximum impact of 153,522 persons. According to the UD scenario simulations, the probability that an individual significant tornado over the OKCMA would impact more than 5,000 (20,000) persons increased to 20.8% (5.3%; Table 4.6). The UD scenario showed a statistically significant increase in exposures over the baseline, as can be observed from the large differences between the POE curves (Figure 4.5).

4.3.5) Climate Change and Urban Development Scenario

The Climate Change and Urban Development (CCUD) scenario used the 2045 – 2054 projected atmospheric environments combined with the projected population density estimates for 2050. Figure 4.4d shows an example of the ten most impactful years' for the CCUD scenario over the OKCMA. This scenario enables the analysis of the combined effects of climate change and population growth on tornado exposure. According to the CCUD simulation results, the median (mean) annual impacts were 3,820 (8,396) persons over the OKCMA with the most impactful years at 31,979 and 46,747 persons, respectively. These values were higher than both the results for the baseline simulation and the CC scenario; however, they were lower than the

values for the UD scenario under some thresholds (Figure 4.5). We attribute this to the slight

decrease in significant tornado activity over the OKCMA under the CC scenario. Again, the

results showed the highest median annual impact (8,040 persons) for Oklahoma County, while

Canadian County had the lowest (3,103 persons; Table 4.5). For individual significant tornadoes,

the mean and median impact of 1,201 and 4,812 persons, respectively, was higher than for the

UD scenario. However, the maximum impact of 102,430 persons was lower as compared to

153,522 persons for the UD scenario. According to the CCUD scenario simulations, the

probability that an individual significant tornado over the OKCMA would impact more than

5,000 (20,000) persons was also higher than for the UD scenario at 22.2% (6.1%; Table 4.6).

Like the UD scenario, the CCUD scenario showed a statistically significant increase in both

annual total exposure and individual exposure over the baseline simulation. However,

differences between the UD and CCUD scenario exposures were very small (Figure 4.5) and not

statistically significant, except for Cleveland County for individual significant tornado

exposures. Population change is the primary driver of the change in normalized annual impacts

with a 75-times greater median change in value and an opposite direction of change (although

the decrease due to the climate change scenario was not statistically significant).

4.3.6) Exposure of Urban Versus Rural Population

For the OKCMA, under the no urban development scenario, most (51%) significant

tornadoes only impacted rural parcels while that number dropped to 42% under the urban

development scenario (Table 4.3). Most of the change was because of an increase in the number

tornadoes impacting low-density parcels (~9%). However, tornadoes impacting moderate to

high-density parcels also increased (~1%). Each county saw a similar increase in urban tornadoes

(~10%) with most of that increase in urban-low tornadoes. The simulation showed the greatest

increase in urban-high tornadoes (1.5%) in Oklahoma County (Table 4.3). Overall the tornado type distribution matches the land use type distribution (Table 4.1) with the greatest number of rural (urban) tornadoes in Canadian County (Oklahoma County).

4.4) Discussion

Here, we conducted an integrated scenario analysis to analyze and quantify the isolated and combined effects of climate change and change in population density for the Oklahoma City Metropolitan Area. For this, we combined spatially explicit population density simulations, carried out with ENVISION (Bolte *et al.*, 2007), with spatial tornado simulations produced with the environmentally driven daily impacts simulation model TorDIS (Hatzis *et al.*, in review). One novel component of our study was the use of climate projections to assess the influence of climate change on tornado risk. For this purpose, we used the newly developed TorDIS model, which uses data on the atmospheric environment to spatially distribute tornadoes on a daily time step and to constrain the intensity and size of the simulated tornado footprints. This approach allowed us to use projected atmospheric environments (from NARCCAP; Mearns *et al.*, 2014) to simulate tornadoes under potential future environments and to compare and contrast these projections with simulations for the baseline (using historical atmospheric environments). This approach differs from that similar studies (*e.g.*, that of Strader *et al.* (2017a) who compare simulations using historical tornado climatology (Strader *et al.*, 2016) with simulations using a manual increase, from the climatology, in tornado frequency and interannual variability. The advantage of our novel approach is that it avoids the use of a flat percent change in tornado frequency across the US and instead allows the change to vary according to the projected changes in the atmospheric environment (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2013; Gensini and Mote, 2015).

While TorDIS provides improved functionalities by using environmental data, improvements in the logistic regressions that control the spatial and temporal distribution of the tornadoes are required. The current regressions led to a significant overestimation of the number of tornado favorable days (as evidenced by the high false alarm rate (53.5%) for the tornado day forecasts). While the number of tornado favorable days is not the same as the number of days with tornadoes (*i.e.*, favorable days still need storms to initiate in order for tornadoes to occur; Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2013; Gensini *et al.*, 2014b), overestimations of the number of days will still lead to overestimation of the total number of tornadoes. The ability to capture the spatial distribution of tornadoes could also be improved. This study used atmospheric environments at one time step daily (0000 UTC; Gensini and Ashley, 2011) due to time constraints. Future studies will use atmospheric environmental data at multiple time steps daily to get atmospheric environments closer in time the actual occurrence of tornadoes (Lee, 2002; Brooks *et al.*, 2003b). This may result in more realistic relationships between the tornadoes and their environments, thus improving the accuracy of the logistic regressions.

One key result of this scenario study is that population growth is the main driver of increased tornado exposure. This is in line with earlier findings from studies analyzing observed population distributions (Ashley and Strader, 2016; Hatzis *et al.*, 2019) as well as spatially explicit simulation studies (Strader *et al.*, 2017a, 2017b, 2018). The OKCMA is projected to experience significant population growth by 2050 (Barker, 2012) and during this time atmospheric conditions are also projected to become more conducive for severe weather and potentially tornadoes (Diffenbaugh *et al.*, 2008, 2013; Gensini *et al.*, 2014b). There are some indications that tornado risk is shifting to the east, as the western states become drier (Gensini and Brooks, 2018; Seager *et al.*, 2018). This eastward shift could lead to a reduction in tornado

occurrence over the OKCMA in the future. However, we did not observe a significant downward trend over OKCMA in this study. Future studies using an ensemble of climate models and emissions scenarios might confirm such a trend (IPCC, 2013; Pryor *et al.*, 2014). While the tornado risk may decrease, the increasing population of the OKCMA will likely lead to greater tornado impacts (Ashley *et al.*, 2014; Ashley and Strader, 2016). Our findings showed that, despite a slight reduction (not significant) in significant tornado risk, the annual number of persons exposed to significant tornadoes increased over time (Figure 4.5). As the population within the US continues to become increasing urban (Seto *et al.*, 2012; United Nations, 2018) more people will reside in the paths of hazards such as tornadoes (Ashley *et al.*, 2014; Ashley and Strader, 2016). The risk posed by significant tornadoes also varies depending upon the development paradigm employed (*e.g.*, sprawl or infill; Strader *et al.*, 2017b, 2018). Studies have shown that sprawl-type development leads to greater median tornado impacts with more people potentially in harm's way (Ashley and Strader, 2016). Infill-type (concentrated) development leads to greater high-end tornado impacts with people more densely packed into the biggest tornadoes' paths (Rosencrants and Ashley, 2015; Strader *et al.*, 2018; Hatzis *et al.*, 2019). For city planning purposes, TorDIS could be run with multiple development paradigms such as these to see which mode of development reduces the risk the most. Future improvements to TorDIS will also add the capability to assess social vulnerability (Cutter *et al.*, 2003; Simmons and Sutter, 2011) in the path of tornadoes. Such improvements would allow us to assess where to establish public storm shelters (Merrell *et al.*, 2002, 2005) to protect the most vulnerable at-risk populations.

The Urban Development scenario presented in this study represents one (out of many possible) potential future pathway(s) for population development in the OKCMA - a BAU

scenario simulating a continuation of the observed historical development pattern (Dilekli *et al.*, in preparation). The OKCMA has historically grown outward by increasing low-density housing on the outskirts of the city, a sprawl-type development (Hamidi and Ewing, 2014; Laidley, 2016). Sprawl-type development has led the OKCMA to be dominated by low-density housing (Table 4.1). This development paradigm may change in the future with the current comprehensive plan for Oklahoma City encouraging/incentivizing infill-type development (increasing density in existing developments or filling in unoccupied areas within the city limits; Oklahoma City Planning Commission, 2017). Future studies on the influence of urban development on tornado impacts in the OKCMA will incorporate urban development scenarios for other possible development paradigms, including increasing sprawl and increasing infill.

In summary, this study evaluated the effects of climate change and urban development on tornado impacts in the Oklahoma City Metropolitan Area. Similar to previous studies (Bouwer, 2013; Rosencrants and Ashley, 2015; Strader *et al.*, 2017a), we found that population growth plays a much larger role in changes to the tornado disaster landscape than climate change. We found that the OKCMA is expected to see an increase in annual tornado impacts by 2050, primarily due to the projected near doubling of the population (Barker, 2012). While not statistically significant, the downward trend in significant tornado risk found during this time period was in line with the suggestions that tornado risk in the OKCMA may be decreasing (and shifting eastward) as western states, like Oklahoma, become drier (Gensini and Brooks, 2018; Seager *et al.*, 2018). Future sensitivity analyses will be required to confirm if the downward trend was due to random effects or climate change. Furthermore, this study showcased the use of the TorDIS model (Hatzis *et al.*, in review) to assess the effect of climate change on tornado impacts. As the population of the US continues to grow (Seto *et al.*, 2011; United Nations, 2018)

and increasing greenhouse gas emissions further alter the climate (Trapp *et al.*, 2007; Gensini and Ashley, 2011; Diffenbaugh *et al.*, 2013), the impact of tornadoes is likely to keep changing. To reduce the risk of casualties in the future, it is critical to understand how these changes will occur. Such knowledge can reduce risk by encouraging the placement of more storm shelters and increasing general awareness.

4.5) Conclusion

The objective of this research was to use an environmentally driven tornado impacts model to analyze how the isolated and combined effects of urbanization and climate change may alter tornado exposure over the Oklahoma City Metropolitan Area in the mid-21st century. The novelty of this study was that it was the first to use the atmospheric environment to project future tornado impacts. While we were unable to find a statistically significant change in tornado risk over the Oklahoma City Metropolitan Area, this was only a limited case study. Future modeling studies using other emissions scenarios and an ensemble of climate models with TorDIS may be able to investigate further how projected future atmospheric environments may influence tornado exposure.

Table 4.1. Parcel-based population and land use statistics for the Oklahoma City Metropolitan Area by county from the 2010 census and 2050 population projections. Variables include county area (km^2), population count (persons), population density (persons km^{-2}), and percentage of parcels under each of three land use classes. Land use classes are based on the definitions from Laidley (2016) with parcels classified as rural if they have a population density of under 518 persons km^{-2}, low density if the population density is between 518 and 9,064 persons km^{-2}, and high density if the population density is 9,065 persons km^{-2} or greater (high density here corresponds to Laidley's moderate to high density definition).

County	Area	2010					2050				
		Count	Density	Land Use by Parcel			Count	Density	Land Use by Parcel		
	km^2		persons km^{-2}	% Rural	% Urban-Low	% Urban-High		persons km^{-2}	% Rural	% Urban-Low	% Urban-High
Canadian	2,439	138,029	57	55.91	44.07	0.02	498,926	205	47.29	52.69	0.02
Cleveland	1,506	280,521	186	39.91	60.07	0.02	548,686	364	29.65	70.33	0.02
Oklahoma	1,935	756,598	391	29.91	70.04	0.05	1,212,823	627	20.75	79.20	0.05
OKC Metro	5,944	1,171,185	197	35.95	64.02	0.04	2,252,564	379	26.62	73.34	0.04

Table 4.2. Rules for allocation of population density at each time step for the urbanization scenarios

Rule	Percent Allocation
Nearest population center with a density of greater than 2000 persons km^{-2} is within 500 m	20
Nearest population center with a density of greater than 1000 persons km^{-2} is within 1000 m	20
Nearest population center is within 100 m	25
Nearest population center is within 101 to 500 m	10
Nearest population center is within 501 to 1000 m	5
Parcel density is no greater than 1,000 persons km^{-2}	10

Table 4.3. Number of simulated significant tornadoes impacting each county by land use class. The land use class represents the land use of the most densely populated parcel the simulated tornado impacted (*i.e.*, a tornado hitting mostly rural parcels but a single urban-high parcel would be classified as urban-high). Counts are normalized by county area (per 2,500 km^2). Counts are averaged between scenarios using the same population estimates (baseline with CC and UD with CCUD) as climate change did not significantly alter tornado counts.

			Rural		Urban - Low		Urban - High	
Scenario	**County**	**Total**	**Count**	**Percent**	**Count**	**Percent**	**Count**	**Percent**
No Urban Development	Canadian	1,857	1,197	64.4	633	34.1	28	1.5
	Cleveland	2,083	1,044	50.1	977	46.9	62	3.0
	Oklahoma	1,748	549	31.4	928	53.1	271	15.5
	OKC Metro	*1,618*	*829*	*51.2*	*675*	*41.7*	*114*	*7.0*
Urban Development	Canadian	1,837	998	54.3	808	44.0	31	1.7
	Cleveland	2,047	816	39.9	1,164	56.9	67	3.3
	Oklahoma	1,710	369	21.6	1,051	61.5	290	17.0
	OKC Metro	*1,596*	*668*	*41.8*	*807*	*50.6*	*121*	*7.6*

Table 4.4. Normalized (per 2,500 km^2) number of simulated significant (EF2+) tornadoes and significant tornado days using the 1990 – 1999 atmospheric environment and the projected 2045 – 2054 environment.

County	# of Tornadoes		# of Tornado Days	
	1990 - 1999	2045 - 2054	1990 - 1999	2045 - 2054
Canadian	1.8	1.7	1.4	1.3
Cleveland	2.0	1.9	1.4	1.3
Oklahoma	1.7	1.6	1.2	1.1
OKC Metro	*1.5*	*1.5*	*1.3*	*1.3*

Table 4.5. Summary statistics for the normalized (by county per 2,500 km^2) annual number of persons exposed to significant tornado damage paths for each scenario by county and for the entire Oklahoma City Metropolitan Area. Statistics include median, mean, 95th and 99th percentiles, and maximum. Also included are the probability that the normalized annual exposure total will exceed 5,000 or 20,000 persons.

Scenario	County	Population Exposed					POE for Threshold (%)	
		Median	Mean	95%	99%	Max	5,000	20,000
Baseline	Canadian	345	1,688	8,138	17,305	23,607	10.36	0.59
	Cleveland	1,403	6,010	26,835	54,201	84,235	30.67	7.98
	Oklahoma	3,929	12,506	55,730	98,130	117,080	46.07	21.03
	OKC Metro	*1,328*	*4,541*	*21,284*	*33,322*	*42,014*	*28.02*	*5.81*
Climate Change Only	Canadian	243	1,519	8,128	14,634	22,112	9.70	0.30
	Cleveland	1,290	5,681	31,117	50,713	95,078	25.64	7.35
	Oklahoma	3,658	10,857	46,618	78,042	105,061	45.10	18.01
	OKC Metro	*1,120*	*3,984*	*17,757*	*28,305*	*43,871*	*25.39*	*3.73*
Urban Development Only	Canadian	3,889	6,465	21,790	36,120	58,141	42.21	6.91
	Cleveland	4,561	11,344	49,013	87,636	304,060	48.13	16.58
	Oklahoma	7,266	19,387	83,019	139,980	198,126	58.89	28.19
	OKC Metro	*4,175*	*8,576*	*34,688*	*57,559*	*105,968*	*44.44*	*12.06*
Urban Development and Climate Change	Canadian	3,103	6,013	21,124	39,268	66,460	38.75	5.78
	Cleveland	4,613	11,443	48,435	72,677	104,758	48.65	18.56
	Oklahoma	8,040	19,351	80,243	107,700	187,483	58.24	31.11
	OKC Metro	*3,820*	*8,396*	*31,979*	*46,747*	*68,203*	*43.55*	*13.14*

Table 4.6. Same as for Table 4.5 but for persons exposed to individual significant tornado damage paths

Scenario	County	Population Exposed					POE for Threshold (%)	
		Median	Mean	95%	99%	Max	5,000	20,000
Baseline	Canadian	55	642	3,598	9,543	20,135	2.71	0.06
	Cleveland	236	1,729	9,338	20,559	48,503	9.06	1.02
	Oklahoma	627	4,547	24,049	47,347	88,059	23.08	6.18
	OKC Metro	*172*	*2,523*	*14,083*	*39,083*	*93,448*	*12.17*	*3.08*
Climate Change Only	Canadian	48	576	3,153	8,793	19,049	3.06	0.00
	Cleveland	212	1,681	8,709	22,133	52,402	8.31	1.68
	Oklahoma	628	3,891	18,661	39,269	61,562	21.05	4.62
	OKC Metro	*145*	*2,187*	*12,410*	*31,761*	*63,173*	*11.29*	*2.70*
Urban Development Only	Canadian	902	2,427	9,698	19,145	42,639	14.66	0.85
	Cleveland	898	3,196	14,273	36,320	65,874	16.50	3.17
	Oklahoma	1,784	7,153	34,931	64,532	153,315	30.00	9.52
	OKC Metro	*1,147*	*4,693*	*20,768*	*62,837*	*153,522*	*20.79*	*5.28*
Urban Development and Climate Change	Canadian	883	2,374	9,757	19,331	45,275	13.59	0.98
	Cleveland	1,098	3,384	15,691	28,968	56,137	18.83	3.09
	Oklahoma	1,889	7,165	34,566	62,771	94,162	32.64	9.95
	OKC Metro	*1,201*	*4,812*	*22,819*	*53,812*	*102,430*	*22.15*	*6.06*
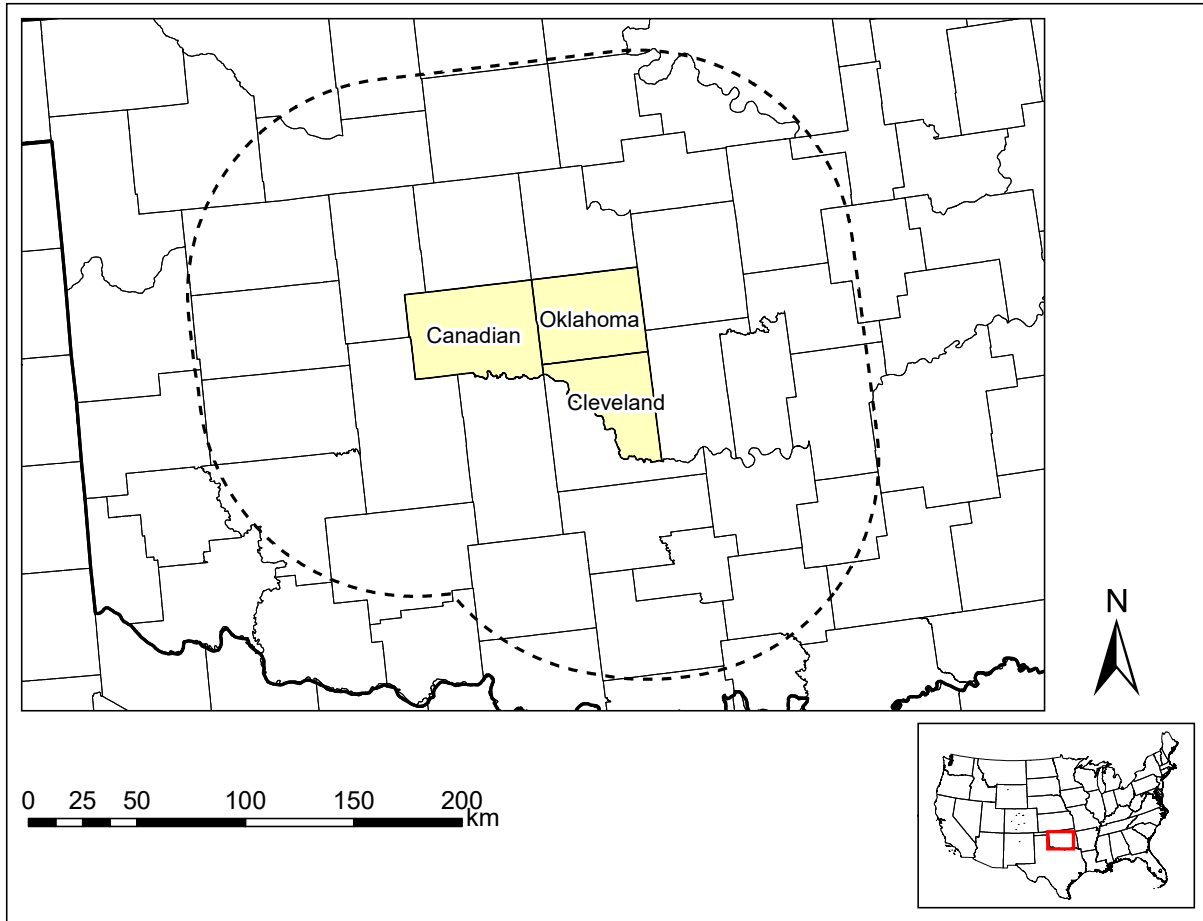
Figure 4.1. Study area representing the three primary counties of the Oklahoma City Metropolitan Area (Canadian, Cleveland, and Oklahoma) in the state of Oklahoma. The dashed line represents the 100 kilometer buffer used by the TorDIS model (Hatzis *et al.*, in review) to reduce edge effects.
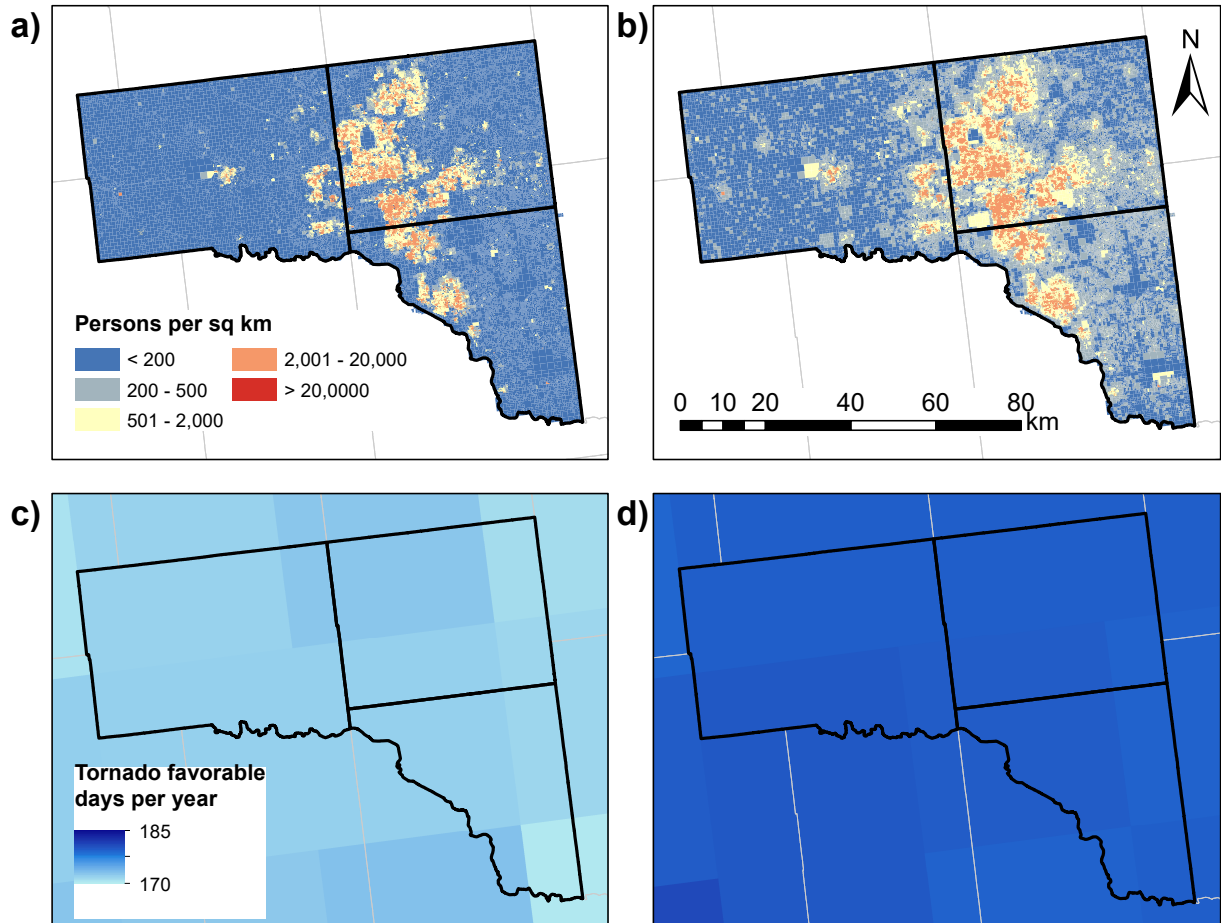
Figure 4.2. Population and tornado climate change over the Oklahoma City Metropolitan Area in the 21st century. a.) 2010 parcel-based estimates of population density (persons km^{-2}). (b) Same as a) but for 2050 parcel-based projections of population density. c.) Mean number of days with tornado favorable conditions (based on the severe weather regression equations) during the 1990 – 1999 period on a 50 km resolution grid. d.) Same as c) but for the projected period of 2045 – 2054, based on the NARCCAP climate data.
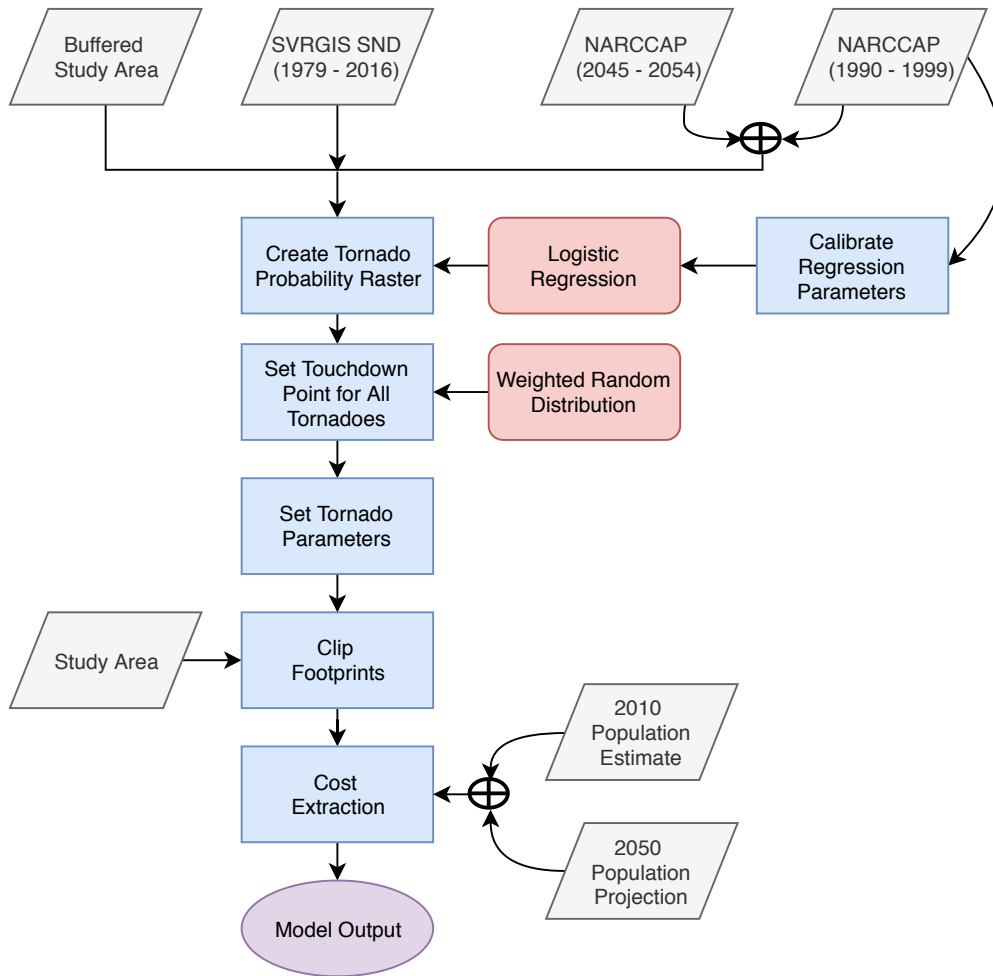
Figure 4.3. Flowchart for each daily simulation using TorDIS. Rhombus shapes represent model input, squares represent model processes, rounded rectangles represent model decisions, cross in circle shapes represents a choice between two inputs (two climate scenarios or two urban development scenarios) and the oval represents the model output. Adapted from Figure 1 in Hatzis *et al.* (in review).
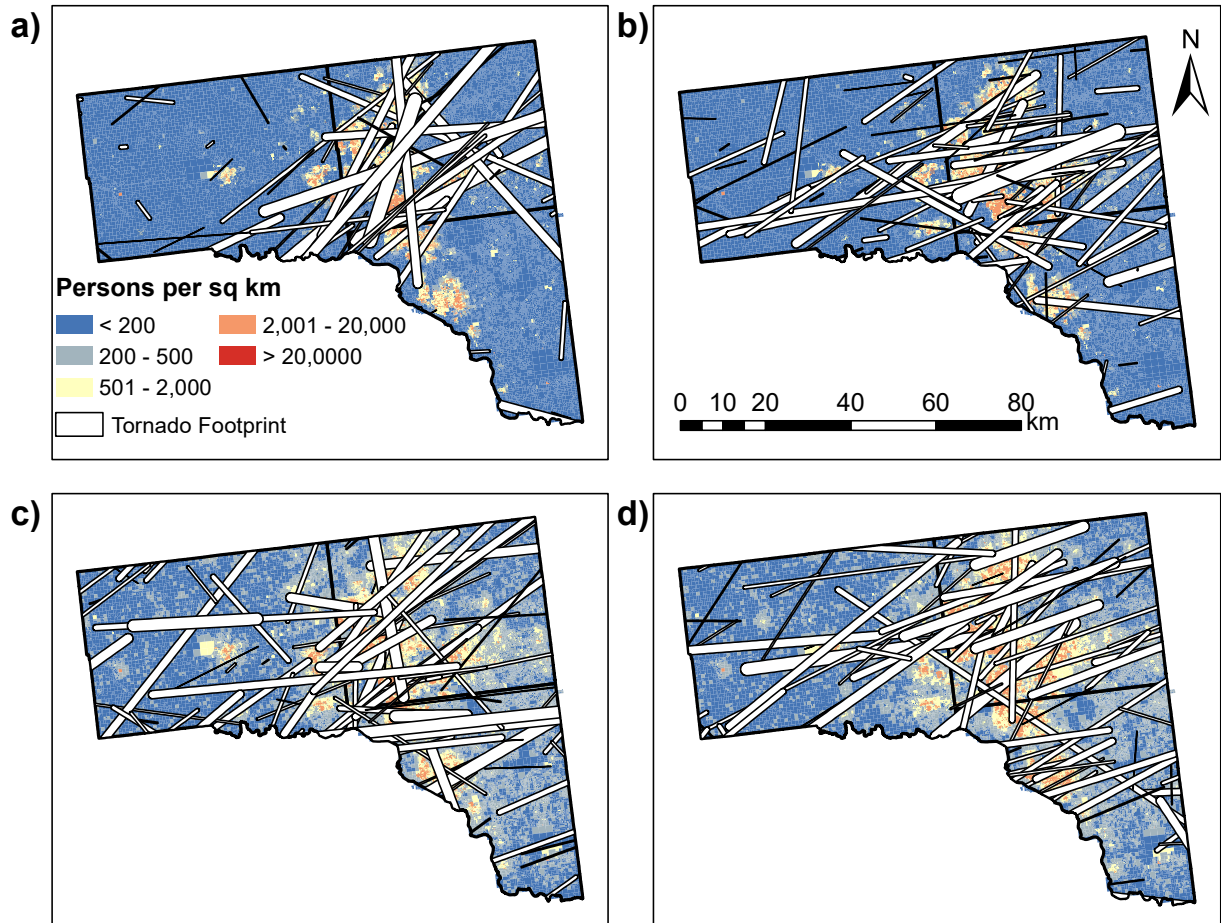
Figure 4.4. Significant tornado footprints from the ten most impactful years of each simulation. Footprints correspond to the following scenarios: baseline control a), climate change only b), urban development only c), and urban development and climate change d). Footprints are overlaid upon the population densities used in each scenario: 2010 estimates (a and b) and 2050 projections (c and d).
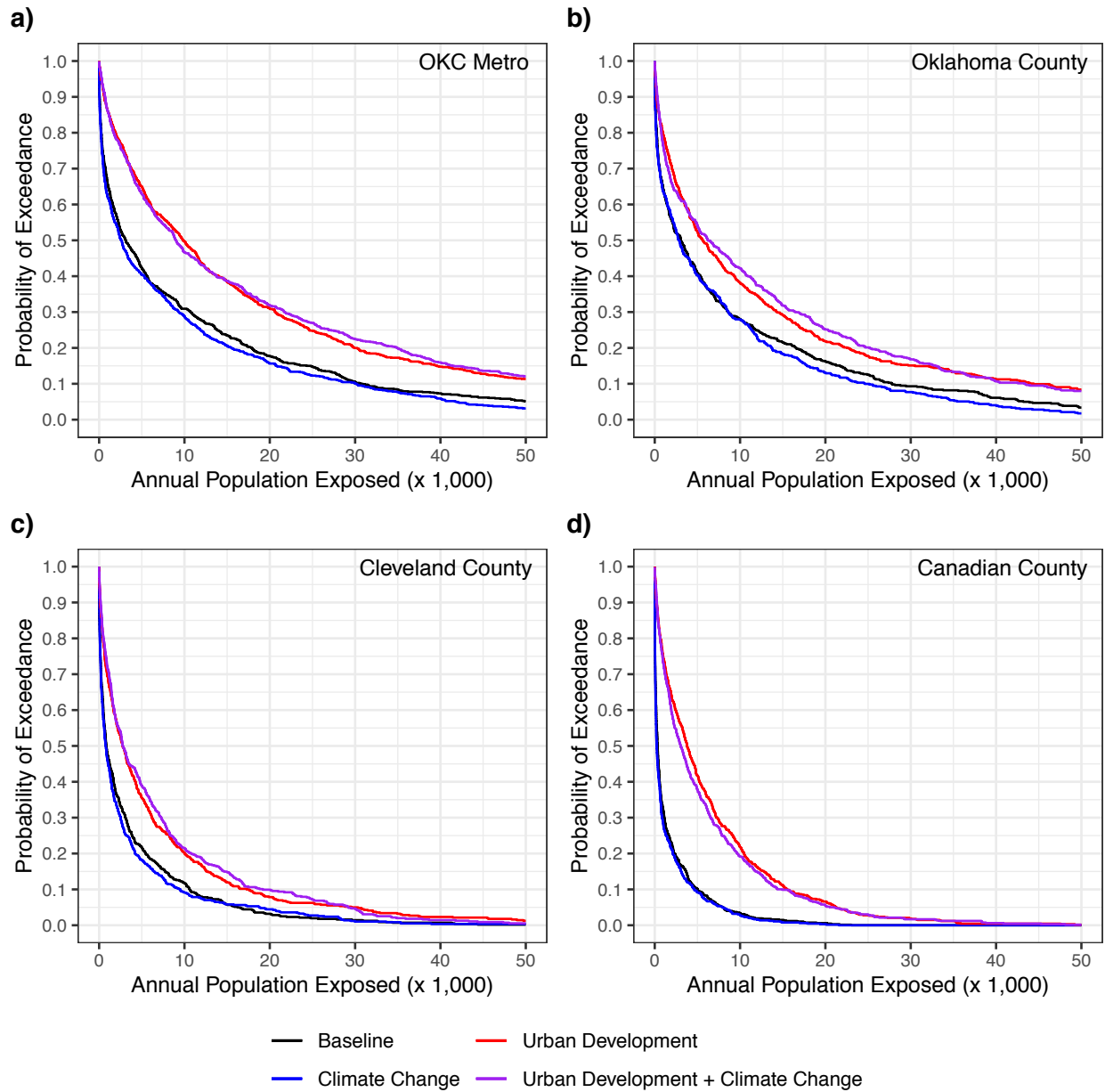
Figure 4.5. Probability of exceedance curves for the annual total population exposed to tornado winds for (a) the Oklahoma City Metropolitan Area, (b) Oklahoma County, (c) Cleveland County, and (d) Canadian County. Black line represents the baseline scenario, blue represents the climate change only scenario, red represents the population change only scenario, and purple represents the climate and population change scenario.

Chapter 5 Conclusions and Future Work

As we continue to alter the world around us, our actions will have consequences for the risk posed by severe convective hazards such as tornadoes. Greenhouse gas emissions continue to rise with no prospect of a resolution to mitigate climate change, causing weather patterns to keep changing (IPCC, 2013). While it is unclear if tornado activity will increase in a warmer world (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2008), it is clear that the spatial distribution (Gensini and Brooks, 2018) and the frequency of tornado outbreaks is changing and will likely continue to do so into the future (Brooks *et al.*, 2014; Elsner *et al.*, 2015). With continued population growth and urban development causing more people to reside in the paths of potential tornadoes, it is critical that we better understand tornado risk. The three studies presented in this dissertation attempt to further our knowledge of the spatial patterns of tornado risk and impacts in a changing world.

Chapter 2 is the first known large-scale spatiotemporal analysis of near-miss violent tornadoes in the United States. In this study, I developed and applied a new methodology for assessing near-misses. First, I replicated historical violent tornado footprints across a uniform circular grid with a radius of 40 km and a resolution of 0.5 km. Then, I used these replicated footprints to assess the number of persons residing in each footprint via proportional allocation. Using these replicates, this study provided clear definitions for near-misses and hits (I defined a hit as a tornado that impacted at least 5,000 persons and a near-miss as a tornado that was not a hit but had at least one replicate within 10 km that impacted at least 5,000 persons). This analysis revealed that only a third of all violent tornadoes hit close enough to population centers to potentially impact 5,000 persons or more. Hits and near-misses both had similar spatial distributions as other violent tornadoes with the majority located in the Southern Plains and the

Southeast. I found the highest risk of near-misses and hits in central Oklahoma and northern Alabama, where a high violent tornado risk met large population centers. None of the near-misses were likely to have been hits, and nearly a quarter of the hits were unlikely to have impacted 5,000 persons or more. The implication here is that hits tended to occur where violent tornadoes impacted the edge or middle of a large population center, while near-misses occurred on the outside of, but nearby to, a large population center where the tornadoes were unlikely to impact 5,000 people or more. Furthermore, this study was an update of the violent tornado climatologies of Concannon *et al.* (2000) and Doswell *et al.* (2012) to extend from 1880 to 2016. The climatology showed that the risk for violent tornadoes has shifted to the east as the climate in the western states began to get drier (Gensini and Brooks, 2018; Seager *et al.*, 2018). It also showed that the use of fine-resolution population data did not improve the accuracy of population exposure estimates for synthetic tornado footprints (linear paths with a constant width), compared to estimates using official damage paths from the National Weather Service (which may curve and have variable widths).

However, the study described in Chapter 2 was subject to small sample bias (Doswell, 2007) and, as such, it is unclear if the true risk for near-misses and hits for violent tornadoes was assessed. Future work on this topic should extend the analysis to all significant (EF2+) tornadoes, as significant tornadoes are responsible for most of the fatalities and damages caused by tornadoes (Ashley, 2007; Simmons and Sutter, 2011; Strader *et al.*, 2017a). Assessing near-misses for all significant tornadoes could help confirm the pattern identified in this study. However, even when including all significant tornadoes, the sample size is still small due to the rarity of significant tornadoes (Verbout *et al.*, 2006; Doswell, 2007; SPC, 2017). One way to address this limitation is to use a spatial impacts model as it enables the user to create a much

larger sample size and can draw from the full envelope of historical tornado sizes and intensities and even simulate record-breaking tornadoes.

The implications of the study conducted in Chapter 2 are twofold. First, the finding that population exposure estimates for synthetic tornado footprints cannot be improved by using finer resolution population data is relevant because it justifies the use of coarser population data in the exposure estimates. This is important because it allows for the study of tornado exposures going back to the early part of the US tornado record (*e.g.*, the 1950s when the modern record first started or the late 1800s when the first routine collection of tornado reports began; Ashley, 2007) when only county-level data was available from the US Census Bureau (Simmons and Sutter, 2011). Second, since near-misses can affect a person's perception of tornado risk, by either making a person extra cautious because they felt lucky to have been spared or extra careless because frequent near-misses have made them underestimate the true risk (Dillon *et al.*, 2011, 2014; Dillon and Tinsley, 2016), understanding the spatial distribution of high-risk areas for near-misses is important in casualty mitigation. By understanding the location of these high-risk areas, it is possible for emergency managers to target these areas for educational town hall meetings to improve the locals understanding of the true tornado risk and potentially save lives (Stewart *et al.*, 2018).

Chapter 3 introduced a new spatial tornado impacts model, the Tornado Daily Impacts Simulator (TorDIS), which I developed by building on the work of Strader *et al.* (2016) and their Tornado Impact Monte Carlo model (TorMC). The novelty of TorDIS lies in the introduction of daily environmental atmospheric data, such as that from atmospheric reanalyses or climate model outputs, as a way of constraining the size/intensity and spatial distribution/daily count of the simulated tornadoes. Like TorMC, TorDIS uses thousands of Monte Carlo simulations to

stochastically place and create tornado footprints to replicate the apparent randomness associated with tornado occurrence (Klockow *et al.*, 2014). In a TorDIS simulation, the number of tornadoes simulated on a given day is determined by a random draw from historical daily tornado counts over the study area, limited by season (Brooks *et al.*, 2003a; Doswell *et al.*, 2006) and lifting condensation level (Rasmussen and Blanchard, 1998). The model randomly places each tornado footprint within the study area based on the weighted probability of tornado occurrence. This probability is based on the favorability of the atmospheric environment for the production of tornadoes (Brooks *et al.*, 2003b; Gensini and Ashley, 2011; Diffenbaugh *et al.*, 2013) according to a series of logistic regression equations. These equations are derived using historical reanalysis data from the North American Regional Reanalysis (NARR; Mesinger *et al.*, 2006) along with historical tornado report data from SVRGIS (SPC, 2017). The variables chosen for inclusion in these logistic regressions are selected through a global sensitivity analysis using SimLab software (Giglioli and Saltelli, 2008). The intensity of each tornado is randomly drawn from a sample of historical intensities limited by storm relative helicity within three storm relative helicity categories (Droegemeier *et al.*, 1993; Moller *et al.*, 1994; Colquhoun and Riley, 1996). The length and width of each tornado footprint are randomly drawn from an intensity-specific Weibull distribution, calibrated to the historical data (Brooks, 2004; Strader *et al.*, 2016). The direction of the tornado is determined by the steering winds (500 mb wind direction) at the point of touchdown (Notis and Stanford, 1973; Suckling and Ashley, 2006). The generated footprints can then be intersected with any potential cost surface (*e.g.*, population, housing units) and a total impact can be assessed via an area-weighted sum across the footprint (Ashley *et al.*, 2014). This study showcased the reliability and utility of the model in four ways. First, it showed the reliability of TorDIS by yielding tornadoes with similar distributions of intensity, path length

and width, and direction when compared with historical data. Second, it showed how TorDIS could be used for comparisons of annual tornado risk between multiple metropolitan areas. Third, the daily time step, as compared to the annual time step used by TorMC (Strader *et al.*, 2016), employed by TorDIS allowed for analysis of the seasonal time scale. By using this capability, I was able to show that simulated tornado frequency and tornado impacts followed the same seasonal pattern with the greatest peak in April and a lesser peak in October. Last, the study displayed how daily simulations could be used to assess the likelihood that a given high-impact tornado (*e.g.*, Bridge Creek-Moore, OK on May 3, 1999 or Tuscaloosa-Birmingham, AL on April 27, 2011; Brooks and Doswell, 2002; Doswell *et al.*, 2012) would have impacted as many people as it did. The model was also able to assess the likelihood of a high-impact tornado on a near-miss day such as February 5, 2008 (Chaney and Weaver, 2010).

The major limitation of this study is in the accuracy of the spatial and temporal distribution of tornadoes. The study had a very low skill score (albeit a very high detection rate) in predicting the spatial pattern of tornado probability and a very high false alarm rate in determining whether a day should be tornadic. To address this limitation, I plan to expand the number of years of reanalysis data used (I only used the period from 2005 – 2016 while the data is available since 1979) as well as use more daily time steps (reanalysis data was only used for 0000 UTC while it is available eight times daily; Mesinger *et al.*, 2006). The range of data used was limited due to computational restraints with 0000 UTC chosen as the time step closest to the peak in tornado activity in the central and southern US (Brooks *et al.*, 2003b; Gensini and Ashley, 2011). The use of more data will enable the development of proximity soundings with atmospheric environments more representative of tornadoes (Lee, 2002; Brooks *et al.*, 2003b).

The proposed improvements in the proximity soundings are expected to increase the accuracy of the logistic regressions and allow better tornado and tornado day forecasts.

The research described in Chapter 3 displays a first attempt at linking daily atmospheric environments with tornado impacts by using the newly developed model TorDIS. There are three advantages to this new approach. (1) By using a model with a daily time step, it is possible to project future tornado impacts on a given high-risk day. While it was previously possible to do such projections using the Storm Prediction Center's Convective Outlooks (Hitchens and Brooks, 2014), it was only possible to use probabilistic forecasts for tornadoes for up to 24 hours out (Grams *et al.*, 2014). (2) Using information provided by the atmospheric environment to constrain the simulation of tornado intensity and path direction may allow for a more realistically drawn sample from the observed data. (3) By using atmospheric environmental data to constrain and distribute simulated tornadoes, it is possible to test the potential effects of climate change on tornado impacts (see Chapter 4).

Chapter 4 describes an application example of using TorDIS to assess the isolated and combined impacts of climate change and urban development in the three-county Oklahoma City Metropolitan Area (Canadian, Cleveland, and Oklahoma Counties). This study showed how the logistic regressions used by TorDIS to distribute tornado footprints could be calibrated against climate model output as well as reanalysis data. In this case, I calibrated the logistic regressions against the current (1990 – 1999) atmospheric environmental data from the North American Regional Climate Change Assessment (NARCCAP). I also collected projected future (2045 – 2054) atmospheric environmental data from NARCCAP. The cost surface data used by TorDIS for this study were the current (derived from the 2010 US census) population estimates distributed to the parcel level in the Oklahoma City Metropolitan Area and future population

simulations (projected to 2050 assuming a business-as-usual development paradigm), developed with the ENVISION framework (Bolte *et al.*, 2007).

The study tested and quantified the relative importance of climate change and urban development to annual tornado impacts by running four simulations: (1) a baseline simulation with the current climate and population distribution; (2) a climate change only scenario using the future climate with current population distribution; (3) an urban development only scenario using current climate with the future population distribution; and (4) an urban development and climate change scenario using the future climate and population distributions. I found that despite the fact that climate change was leading to a slight reduction of tornado risk over the Oklahoma City Metropolitan Area, the combined effects of climate change and urban development led to a significant increase in annual tornado impacts by 2050, under the conditions tested. The simulation results indicated that urban development was the dominant force in future changes in annual tornado impacts. This is line with findings from other studies (Ashley and Strader, 2016; Strader *et al.*, 2017a).

Chapter 4 showcased the ability of TorDIS to use future projections of the atmospheric environment to assess changes in tornado risk due to a changing climate (Trapp *et al.*, 2007; Diffenbaugh *et al.*, 2008; Gensini and Mote, 2015). A previous study by Strader *et al.* (2017a) simulated the effects of climate change by employing a flat percent increase in tornado frequency and another one in the interannual variability of tornadoes. However, the advantage of the approach in Chapter 4 is that changes in tornado frequency are not likely to be uniform across the US. There are currently indications that tornado risk is shifting to the east away from the traditional Tornado Alley as the western states become drier (Gensini and Brooks, 2018; Seager

et al., 2018). If this trend holds true in the future, then a flat change in tornado climatology would yield unrealistic future tornado risk maps.

The primary limitation of this study was in the use of only one data set each to represent possible future tornado environments and population growth. Climate projections tend to use multiple climate models forced with multiple emissions scenarios to determine how climate change may influence future atmospheric hazards (IPCC, 2013; Pryor *et al.*, 2014). Urban development can also take many different paths forward, resulting in multiple possible population growth scenarios (Verburg *et al.*, 2002; Dorning *et al.*, 2015). Future studies will add multiple climate change and urban development scenarios (*e.g.*, sprawl versus infill) to assess a more complete picture of potential future tornado impacts over the Oklahoma City Metropolitan Area.

This study also looked at how projected urban development patterns of a sprawling city, such as Oklahoma City, could affect tornado impacts. Previous studies have shown that median impacts from hazards, such as tornadoes, are higher in sprawling cities because there are more people and buildings in the path of any potential tornado. Development using an infill-type paradigm (building up or filling in instead of building out) can lead to reduced median impacts (Etkin, 1999). However, the maximum impacts tend to increase as the number of people in the path of, the rare large tornado that hits an urban core, increases impacts (Rosencrants and Ashley, 2015; Strader *et al.*, 2018). The knowledge that median tornado impacts increase as cities become more sprawled can act as an incentive to promote an infill-type development paradigm in cities in high-risk areas such as Oklahoma and Alabama. While infill increases the impacts for potential urban tornadoes, this fact could also incentivize the investment in public

138

storm shelters in heavily populated areas where more people are at risk (Liu *et al.*, 1996; Balluz *et al.*, 2000).

Overall the research presented in this dissertation improved the understanding of the spatial impacts of tornadoes in the following ways. (1) It created a clear definition of near-misses for tornadoes as well as a transparent methodology for studying them. This methodology could be employed to further study the spatial distribution of near-misses for all intensity classes of tornadoes and not just violent ones. (2) It introduced a new spatially explicit and environmentally driven tornado impacts model that operates at a daily time step, TorDIS. I showed how this model could be used to assess the potential impacts for tornadoes on a given high-risk day. The identification of potential future impacts could allow emergency managers to be aware of areas that may have increased risk and allow them to prepare their resources in advance and perhaps save lives. (3) I showed how by using atmospheric environmental data from climate models, it is possible to assess how climate change may influence tornado risk and where these influences might be positive or negative. (4) Finally, I showed how business-as-usual development scenario in the Oklahoma City Metropolitan Area could lead to a significant increase in tornado impacts as sprawl dominates the development. Oklahoma City has already begun incentivizing infill-type development in its latest comprehensive development plan (Oklahoma City Planning Commission, 2017). This research could add further incentive for infill-type development in order to reduce the risk of tornado exposure.

There are a few lines of future research suggested by my findings. First, I plan to make improvements to the accuracy of the tornado day forecasts and probabilistic tornado forecast maps used by TorDIS. The model currently has a known bias towards the overproduction of tornadoes and the overestimation of the number of tornado days. This study used only one time

step a day for the atmospheric environmental data, which means that many tornado reports were associated with environmental conditions occurring up to 12 hours before or after the tornado occurred. Since data is available up to eight times daily for many reanalysis data sets and climate models, future work will attempt to link tornadoes to atmospheric conditions that are within 1.5 – 3 hours of the event. This will ensure that the atmospheric environments are more representative of the tornadoes being forecast and hopefully will improve the accuracy. Eventually, I hope that once the accuracy of the tornado probability forecasts can be improved, the model can be tested for operational use.

I plan to use atmospheric environments from an ensemble of downscaled climate models with TorDIS to get a broader picture of the potential impacts of climate change on tornado risk. Future work will also employ multiple developmental paradigms for different metropolitan areas to see specifically how realistic development changes the tornado impacts as opposed to theoretical development paradigms as employed in Strader *et al.* (2018). Furthermore, I plan to look at Oklahoma City, and other cities, to see how tornado impacts from business-as-usual scenarios compare to scenarios where sprawl or infill are promoted. Such studies could be presented to city planners to help encourage infill as a way of reducing tornado impacts.

As a long term goal for the simulation model introduced in this dissertation, I hope that improvements of TorDIS will lead to a sufficient level of model accuracy to make the simulation tool useable for operational forecasts of potential tornado impacts on high-risk days. With the future addition of measures of vulnerability (*e.g.*, poverty, visibility, language barriers; Cutter, 1996; Simmons and Sutter, 2011), we could project casualty estimates as well so that emergency managers will be made aware, in advance, of what resources they may need on a given day. Future modeling studies using TorDIS could be conducted by city planners to test the effect of

local changes on land use in tornado impacts. Finally, TorDIS could be used by emergency managers to promote the establishment of public storm shelters to offset the increased risk of high impact events when infill-type development is encouraged by city planners, as a way to reduce the overall risk of tornado impacts.

References

Abbey, R.F., 1976. Risk probabilities associated with tornado wind speeds, in: Abbey, R.F., Mehta, K.C. (Eds.), Proceedings of the Symposium on Tornadoes: Assessment of Knowledge and Implications for Man. Lubbock, TX, pp. 177–236.

Adlerman, E.J., Droegemeier, K.K., Davies-Jones, R., 1999. A numerical simulation of cyclic mesocyclogenesis. J. Atmos. Sci. 56, 2045–2069.

Agee, E., Childs, S., 2014. Adjustments in tornado counts, F-scale intensity, and path width for assessing significant tornado destruction. J. Appl. Meteorol. Climatol. 53, 1494–1505. https://doi.org/10.1175/JAMC-D-13-0235.1

Ali, K., Partridge, M.D., Olfert, M.R., 2007. Can geographically weighted regressions improve regional analysis and policy making? Int. Reg. Sci. Rev. 30, 300–329.

Alig, R.J., Kline, J.D., Lichtenstein, M., 2004. Urbanization on the US landscape: looking ahead in the 21st century. Landsc. Urban Plan. 69, 219–234.

Anselin, L., 1995. Local indicators of spatial association-LISA. Geogr. Anal. 27, 93–115.

Arnold, K., Golino, D.A., McRoberts, N., 2017. A synoptic analysis of the temporal and spatial aspects of grapevine leafroll disease in a historic Napa vineyard and experimental vine blocks. Phytopathology 107, 418–426.

Ashley, W.S., 2007. Spatial and Temporal Analysis of Tornado Fatalities in the United States: 1880–2005. Weather Forecast. 22, 1214–1228. https://doi.org/10.1175/2007WAF2007004.1

Ashley, W.S., Krmenec, A.J., Schwantes, R., 2008. Vulnerability due to Nocturnal Tornadoes. Weather Forecast. 23, 795–807. https://doi.org/10.1175/2008WAF2222132.1

Ashley, W.S., Strader, S.M., 2016. Recipe for disaster: how the dynamic ingredients of risk and exposure are changing the tornado disaster landscape. Bull. Am. Meteorol. Soc. 97, 767–786.

Ashley, W.S., Strader, S.M., Rosencrants, T., Krmenec, A.J., 2014. Spatiotemporal changes in tornado hazard exposure: The case of the expanding bull's-eye effect in Chicago, illinois. Weather. Clim. Soc. 6, 175–193.

Atkins, N.T., Butler, K.M., Flynn, K.R., Wakimoto, R.M., 2014. An integrated damage, visual, and radar analysis of the 2013 Moore, Oklahoma, EF5 Tornado. Bull. Am. Meteorol. Soc. 95, 1549–1561. https://doi.org/10.1175/BAMS-D-14-00033.1

Balluz, L., Schieve, L., Holmes, T., Malilay, J., 2000. Predictors for People' s Response to a Tornado Warning : Arkansas , 1 March 1997. Disasters 24, 71–77. https://doi.org/10.1111/1467-7717.00132

Barker, S., 2012. 2012 Demographic State of the State Report, Oklahoma City.

Barnes, L.R., Gruntfest, E.C., Hayden, M.H., Schultz, D.M., Benight, C., 2007. False alarms and close calls: A conceptual model of warning accuracy. Weather Forecast. 22, 1140–1147.

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

Bierwagen, B.G., Theobald, D.M., Pyke, C.R., Choate, A., Groth, P., Thomas, J. V, Morefield, P., 2010. National housing and impervious surface scenarios for integrated climate impact assessments. Proc. Natl. Acad. Sci. 107, 20887–20892.

Billet, J., DeLisi, M., Smith, B.G., Gates, C., 1997. Use of regression techniques to predict hail size and the probability of large hail. Weather Forecast. 12, 154–164.

Blair, S.F., Lunde, E.P.K., 2010. Tornadoes impacting interstates: Service and societal considerations. E-Journal Sev. Storms Meteorol. 5.

Bluestein, H.B., Snyder, J.C., Houser, J.B., 2015. A multiscale overview of the El Reno, Oklahoma, tornadic supercell of 31 May 2013. Weather Forecast. 30, 525–552. https://doi.org/10.1175/WAF-D-14-00152.1

Bolte, J.P., Hulse, D.W., Gregory, S. V., Smith, C., 2007. Modeling biocomplexity - actors, landscapes and alternative futures. Environ. Model. Softw. 22, 570–579. https://doi.org/10.1016/j.envsoft.2005.12.033

Boruff, B.J., Easoz, J.A., Jones, S.D., Landry, H.R., Mitchem, J.D., Cutter, S.L., 2003. Tornado hazards in the United States. Clim. Res. 24, 103–117.

Bouwer, L.M., 2011. Have disaster losses increased due to anthropogenic climate change? Bull. Am. Meteorol. Soc. 92, 39–46.

Bouwer, L.M., 2013. Projections of future extreme weather losses under changes in climate and exposure. Risk Anal. 33, 915–930.

Bowser, G.C., Cutter, S.L., 2015. Stay or go? Examining decision making and behavior in hurricane evacuations. Environ. Sci. Policy Sustain. Dev. 57, 28–41.

Brooks, H.E., 2004. On the relationship of tornado path length and width to intensity. Weather Forecast. 19, 310–319.

Brooks, H.E., Anderson, A.R., Riemann, K., Ebbers, I., Flachs, H., 2007. Climatological aspects of convective parameters from the NCAR/NCEP reanalysis. Atmos. Res. 83, 294–305.

Brooks, H.E., Carbin, G.W., Marsh, P.T., 2014. Increased variability of tornado occurrence in the United States. Science (80-.). 346, 349–352. https://doi.org/10.1126/science.1257460

Brooks, H.E., Doswell, C.A. III., 2001a. Normalized Damage from Major Tornadoes in the United States: 1890–1999. Weather Forecast. 16, 168–176. https://doi.org/10.1175/1520-0434(2001)016<0168:NDFMTI>2.0.CO;2

Brooks, H.E., Doswell, C.A. III., 2001b. Some aspects of the international climatology of tornadoes by damage classification. Atmos. Res. 56, 191–201.

Brooks, H.E., Doswell, C.A. III., 2002. Deaths in the 3 May 1999 Oklahoma City Tornado from a Historical Perspective. Weather Forecast. 17, 354–361. https://doi.org/10.1175/1520-0434(2002)017<0354:DITMOC>2.0.CO;2

Brooks, H.E., Doswell, C.A. III., Kay, M.P., 2003a. Climatological estimates of local daily tornado probability for the United States. Weather Forecast. 18, 626–640. https://doi.org/10.1175/1520-0434(2003)018<0626:CEOLDT>2.0.CO;2

Brooks, H.E., Doswell, C.A. III., Sutter, D., 2008. Low-level winds in tornadoes and potential catastrophic tornado impacts in urban areas. Bull. Am. Meteorol. Soc. 89, 87–90. https://doi.org/10.1175/BAMS-89-1-87

Brooks, H.E., Lee, J.W., Craven, J.P., 2003b. The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. Atmos. Res. 67, 73–94.

Brotzge, J., Donner, W., 2013. The tornado warning process: A review of current research, challenges, and opportunities. Bull. Am. Meteorol. Soc. 94, 1715–1733. https://doi.org/10.1175/BAMS-D-12-00147.1

Brotzge, J., Erickson, S., Brooks, H.E., 2011. A 5-yr climatology of tornado false alarms. Weather Forecast. 26, 534–544. https://doi.org/10.1175/WAF-D-10-05004.1

Buckley, J.W., 2002. After the storm: Building a safe shelter for the school children of Mulhall, Oklahoma. Weather Forecast. 17, 626–634.

Bunkers, M.J., Klimowski, B.A., Zeitler, J.W., Thompson, R.L., Weisman, M.L., 2000. Predicting supercell motion using a new hodograph technique. Weather Forecast. 15, 61–79. https://doi.org/10.1175/1520-0434(2000)015<0061:PSMUAN>2.0.CO;2

Bunkers, M.J., Wetenkamp Jr, J.R., Schild, J.J., Fischer, A., 2010. Observations of the relationship between 700-mb temperatures and severe weather reports across the contiguous United States. Weather Forecast. 25, 799–814.

Burgess, D., Ortega, K., Stumpf, G., Garfield, G., Karstens, C., Meyer, T., Smith, B., Speheger, D., Ladue, J., Smith, R., Marshall, T., 2014. 20 May 2013 Moore, Oklahoma, Tornado: Damage Survey and Analysis. Weather Forecast. 29, 1229–1237. https://doi.org/10.1175/WAF-D-14-00039.1

Burton, C.G., 2010. Social vulnerability and hurricane impact modeling. Nat. Hazards Rev. 11, 58–68.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2017. U.S. Census Grids (Summary File 1), 2010.

Chaney, P.L., Herbert, J., Curtis, A., 2013a. A Climatological Perspective on the 2011 Alabama Tornado Outbreak. J. Oper. Meteorol. 1.

Chaney, P.L., Weaver, G.S., 2010. The Vulnerability of Mobile Home Residents in Tornado Disasters: The 2008 Super Tuesday Tornado in Macon County, Tennessee. Weather. Clim. Soc. 2, 190–199. https://doi.org/10.1175/2010wcas1042.1

Chaney, P.L., Weaver, G.S., Youngblood, S.A., Pitts, K., 2013b. Household Preparedness for Tornado Hazards: The 2011 Disaster in DeKalb County, Alabama. Weather. Clim. Soc. 5, 345–358. https://doi.org/10.1175/WCAS-D-12-00046.1

Coleman, T.A., Dixon, P.G., 2014. An objective analysis of tornado risk in the United States. Weather Forecast. 29, 366–376. https://doi.org/10.1175/WAF-D-13-00057.1

Coleman, T.A., Knupp, K.R., Spann, J., Elliott, J.B., Peters, B.E., 2011. The history (and future) of tornado warning dissemination in the United States. Bull. Am. Meteorol. Soc. 92, 567.

Colquhoun, J.R., Riley, P.A., 1996. Relationships between tornado intensity and various wind and thermodynamic variables. Weather Forecast. 11, 360–371.

Concannon, P.R., Brooks, H.E., Doswell, C.A. III., 2000. Climatological risk of strong and violent tornadoes in the United States, in: Preprints, 2nd Conf. on Environmental Applications. American Meteorological Society, Long Beach, CA, pp. 212–219.

Confalonieri, R., 2010. Monte Carlo based sensitivity analysis of two crop simulators and considerations on model balance. Eur. J. Agron. 33, 89–93.

Corfidi, S.F., Weiss, S.J., Kain, J.S., Corfidi, S.J., Rabin, R.M., Levit, J.J., 2010. Revisiting the 3–4 April 1974 super outbreak of tornadoes. Weather Forecast. 25, 465–510. https://doi.org/10.1175/2009WAF2222297.1

Cutter, S.L., 1996. Vulnerability to hazards. Prog. Hum. Geogr. 20, 529–539. https://doi.org/10.1177/030913259602000407

Cutter, S.L., Boruff, B.J., Shirley, W.L., 2003. Social vulnerability to environmental hazards. Soc. Sci. Q. 84, 242–261. https://doi.org/10.1111/1540-6237.8402002

d'Amour, C.B., Reitsma, F., Baiocchi, G., Barthel, S., Güneralp, B., Erb, K.-H., Haberl, H., Creutzig, F., Seto, K.C., 2017. Future urban land expansion and implications for global croplands. Proc. Natl. Acad. Sci. 114, 8939–8944.

Daneshvaran, S., Morden, R.E., 2007. Tornado risk analysis in the United States. J. Risk Financ. 8, 97–111.

Danielsen, K.A., Lang, R.E., Fulton, W., 1999. Retracting suburbia: Smart growth and the future of housing.

Davies, J.M., Fischer, A., 2009. Environmental characteristics associated with nighttime tornadoes. Electron. J. Oper. Meteor 10, 1–29.

Dean, A.R., Schneider, R.S., 2008. 9A. 2 Forecast Challenges at the NWS Storm Prediction Center Relating to the Frequency of Favorable Severe Storm Environments.

Dell'Acqua, F., Gamba, P., Jaiswal, K., 2013. Spatial aspects of building and population exposure data and their implications for global earthquake exposure modeling. Nat. hazards 68, 1291–1309.

Diffenbaugh, N.S., Brooks, H.E., Trapp, R.J., 2008. Does global warming influence tornado activity? Eos, Trans. Am. Geophys. Union 89, 553–554. https://doi.org/10.1029/2008EO530001

Diffenbaugh, N.S., Scherer, M., Trapp, R.J., 2013. Robust increases in severe thunderstorm environments in response to greenhouse forcing. Proc. Natl. Acad. Sci. 201307758.

Dileckli, N., Koch, J., Jawarneh, R.N., n.d. A Spatial Analysis of Historical and Potential Future Urbanization Patterns in the Oklahoma City Metropolitan Area. Unpubl. Work.

Dillon, R.L., Tinsley, C.H., 2016. Near-miss events, risk messages, and decision making. Environ. Syst. Decis. 36, 34–44. https://doi.org/10.1007/s10669-015-9578-x

Dillon, R.L., Tinsley, C.H., Burns, W.J., 2014. Near-misses and future disaster preparedness. Risk Anal. 34, 1907–1922. https://doi.org/10.1111/risa.12209

Dillon, R.L., Tinsley, C.H., Cronin, M., 2011. Why near-miss events can decrease an individual's protective response to hurricanes. Risk Anal. 31, 440–449. https://doi.org/10.1111/j.1539-6924.2010.01506.x

Dixon, P.G., Mercer, A.E., Choi, J., Allen, J.S., 2011. Tornado risk analysis: is Dixie Alley an extension of Tornado Alley? Bull. Am. Meteorol. Soc. 92, 433. https://doi.org/10.1175/2010BAMS3102.1

Dixon, R.W., Moore, T.W., 2012. Tornado Vulnerability in Texas. Weather. Clim. Soc. 4, 59–68. https://doi.org/10.1175/WCAS-D-11-00004.1

Dorning, M.A., Koch, J., Shoemaker, D.A., Meentemeyer, R.K., 2015. Simulating urbanization scenarios reveals tradeoffs between conservation planning strategies. Landsc. Urban Plan. 136, 28–39. https://doi.org/10.1016/j.landurbplan.2014.11.011

Doswell, C.A. III., 2007. Small sample size and data quality issues illustrated using tornado occurrence data. E-Journal Sev. Storms Meteorol. 2, 1–16.

Doswell, C.A. III., Brooks, H.E., Dotzek, N., 2009. On the implementation of the enhanced Fujita scale in the USA. Atmos. Res. 93, 554–563. https://doi.org/10.1016/j.atmosres.2008.11.003

Doswell, C.A. III., Burgess, D.W., 1988. On some issues of United States tornado climatology. Mon. Weather Rev. 116, 495–501. https://doi.org/10.1175/1520-0493(1988)116<0495:OSIOUS>2.0.CO;2

Doswell, C.A. III., Carbin, G.W., Brooks, H.E., 2012. The tornadoes of spring 2011 in the USA: An historical perspective. Weather 67, 88–94.

Doswell, C.A. III., Edwards, R., Thompson, R.L., Hart, J.A., Crosbie, K.C., 2006. A Simple and Flexible Method for Ranking Severe Weather Events. Weather Forecast. 21, 939–951. https://doi.org/10.1175/WAF959.1

Doswell, C.A. III., Moller, A.R., Brooks, H.E., 1999. Storm Spotting and Public Awareness since the First Tornado Forecasts of 1948. Weather Forecast. 14, 544–557. https://doi.org/10.1175/1520-0434(1999)014<0544:SSAPAS>2.0.CO;2

Droegemeier, K.K., Lazarus, S.M., Davies-Jones, R., 1993. The influence of helicity on numerically simulated convective storms. Mon. Weather Rev. 121, 2005–2029.

Eakin, H., Luers, A.L., 2006. Assessing the vulnerability of social-environmental systems. Annu. Rev. Environ. Resour. 31, 365–394.

Edwards, R., La Due, J.G., Ferree, J.T., Scharfenberg, K., Maier, C., Coulbourne, W.L., LaDue, J.G., Ferree, J.T., Scharfenberg, K., Maier, C., Coulbourne, W.L., 2013. Tornado intensity estimation: Past, present, and future. Bull. Am. Meteorol. Soc. 94, 641–653. https://doi.org/10.1175/BAMS-D-11-00006.1

Edwards, R., Schaefer, J., 2012. Downtown tornadoes. Accessed through NOAA 13.

Elsner, J.B., Elsner, S.C., Jagger, T.H., 2015. The increasing efficiency of tornado days in the United States. Clim. Dyn. 45, 651–659. https://doi.org/10.1007/s00382-014-2277-3

Elsner, J.B., Fricker, T., Berry, W.D., 2018. A model for US tornado casualties involving interaction between damage path estimates of population density and energy dissipation. J. Appl. Meteorol. Climatol. 57, 2035–2046.

Elsner, J.B., Jagger, T.H., Fricker, T., 2016. Statistical models for tornado climatology: long and short-term views. PLoS One 11, e0166895.

Elsner, J.B., Jagger, T.H., Widen, H.M., Chavas, D.R., 2014. Daily tornado frequency distributions in the United States. Environ. Res. Lett. 9, 24018.

Elsner, J.B., Michaels, L.E., Scheitlin, K.N., Elsner, I.J., 2013a. The decreasing population bias in tornado reports across the central Plains. Weather. Clim. Soc. 5, 221–232.

Elsner, J.B., Murnane, R.J., Jagger, T.H., Widen, H.M., 2013b. A Spatial Point Process Model for Violent Tornado Occurrence in the US Great Plains. Math. Geosci. 45, 667–679. https://doi.org/10.1007/s11004-013-9458-1

EPA, 2009. Land-use scenarios: National-scale housing-density scenarios consistent with climate change storylines.

Etkin, D., 1999. Risk transference and related trends: driving forces towards more mega-disasters. Glob. Environ. Chang. Part B Environ. Hazards 1, 69–75.

Federal Emergency Management Agency, 2018. Be prepared for a tornado [WWW Document]. FEMA V-1010.

Field, C.B., Barros, V., Stocker, T.F., Dahe, Q., 2012. Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change. Cambridge University Press.

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., others, 2005. Global consequences of land use. Science (80-.). 309, 570–574.

Fraedrich, K., Gerstengarbe, F.-W., Werner, P.C., 2001. Climate shifts during the last century. Clim. Change 50, 405–417.

Fricker, T., Elsner, J.B., 2015. Kinetic energy of tornadoes in the United States. PLoS One 10, e0131090.

Fricker, T., Elsner, J.B., 2019. Unusually devastating tornadoes in the United States: 1995--2016. Ann. Am. Assoc. Geogr. 1–15.

Frumkin, H., 2016. Urban sprawl and public health. Public Health Rep.

Fujita, T.T., 1971. Proposed characterization of tornadoes and hurricanes by area and intensity (No. 91), Satellite & Mesometeorology Research Project. Chicago, IL.

Fujita, T.T., Pearson, A.D., 1973. Results of FPP classification of 1971 and 1972 tornadoes, in: Bulletin of the American Meteorological Society.

Galway, J.G., 1981. Ten famous tornado outbreaks. Weatherwise 34, 100–109. https://doi.org/10.1080/00431672.1981.9931955 To

Gensini, V.A., Ashley, W.S., 2011. Climatology of potentially severe convective environments from the North American Regional Reanalysis. E-Journal Sev. Storms Meteorol. 6, 1–40.

Gensini, V.A., Brooks, H.E., 2018. Spatial trends in United States tornado frequency. npj Clim. Atmos. Sci. 1, 38.

Gensini, V.A., Mote, T.L., 2015. Downscaled estimates of late 21st century severe weather from CCSM3. Clim. Change 129, 307–321.

Gensini, V.A., Mote, T.L., Brooks, H.E., 2014a. Severe-Thunderstorm Reanalysis Environments and Collocated Radiosonde Observations. J. Appl. Meteorol. Climatol. 53, 742–751. https://doi.org/10.1175/jamc-d-13-0263.1

Gensini, V.A., Ramseyer, C., Mote, T.L., 2014b. Future convective environments using NARCCAP. Int. J. Climatol. 34, 1699–1705. https://doi.org/10.1002/joc.3769

Giglioli, N., Saltelli, A., 2008. Simlab 2.2 Reference Manual. Ispra, Italy Inst. Syst. Informatics Saf. (Joint Res. Centre, Eur. Comm.

Goliger, A.M., Milford, R. V., 1998. A review of worldwide occurrence of tornadoes. J. Wind Eng. Ind. Aerodyn. 74–76, 111–121. https://doi.org/10.1016/S0167-6105(98)00009-9

Grams, J., Bunting, B., Weiss, S., 2014. SPC Convective Outlooks [WWW Document].

Grazulis, T.P., 1993. Significant Tornadoes: 1680-1991. Enviromental Films.

Grazulis, T.P., 1997. Significant Tornadoes (Update): 1992-1995. Environmental Films.

Greenbaum, R.T., 2002. A spatial study of teachers' salaries in Pennsylvania school districts. J. Labor Res. 23, 69–86.

Griffith, D.A., Haining, R., 2006. Beyond mule kicks: The Poisson distribution in geographical analysis. Geogr. Anal. 38, 123–139.

Groenemeijer, P., Kühne, T., 2014. A climatology of tornadoes in Europe: Results from the European Severe Weather Database. Mon. Weather Rev. 142, 4775–4790.

Halbert, K.T., Blumberg, W.G., Marsh, P.T., 2015. SHARPpy: Fueling the Python Cult, in: Preprints, 5th Symposium on Advances in Modeling and Analysis Using Python, Phoenix AZ.

Hall, S.G., Ashley, W.S., 2008. Effects of Urban Sprawl on the Vulnerability to a Significant Tornado Impact in Northeastern Illinois. Nat. Hazards Rev. 9, 209–219. https://doi.org/10.1061/ASCE1527-698820089:4209

Hamidi, S., Ewing, R., 2014. A longitudinal study of changes in urban sprawl between 2000 and 2010 in the United States. Landsc. Urban Plan. 128, 72–82.

Handy, S., 2005. Smart growth and the transportation-land use connection: What does the research tell us? Int. Reg. Sci. Rev. 28, 146–167.

Hatzis, J.J., Koch, J., Brooks, H.E., 2019. Spatiotemporal Analysis of Near-Miss Violent Tornadoes in the United States. Weather. Clim. Soc. 11, 159–182.

Hatzis, J.J., Koch, J., Brooks, H.E., n.d. A Tornado Daily Impacts Simulator for the Central and Southern United States. Unpubl. Work.

Hitchens, N.M., Brooks, H.E., 2014. Evaluation of the Storm Prediction Center's Convective Outlooks from Day 3 through Day 1. Weather Forecast. 29, 1134–1142. https://doi.org/10.1175/WAF-D-13-00132.1

Hitchens, N.M., Brooks, H.E., Kay, M.P., 2013. Objective limits on forecasting skill of rare events. Wea.\ Forecast. 28, 525–534. https://doi.org/10.1175/WAF-D-12-00113.1

Hoekstra, S., Klockow, K., Riley, R., Brotzge, J., Brooks, H.E., Erickson, S., 2011. A Preliminary Look at the Social Perspective of Warn-on-Forecast: Preferred Tornado Warning Lead Time and the General Public's Perceptions of Weather Risks. Weather. Clim. Soc. 3, 128–140. https://doi.org/10.1175/2011WCAS1076.1

Hoffman, R.L., 2013. To the Southwest Corner: Tornado Myths and Socio-Demographic Vulnerability. Kent State University.

Hollander, M., Wolfe, D.A., Chicken, E., 2013. Nonparametric statistical methods. John Wiley & Sons.

IPCC, 2013. Summary for Policymakers, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom. https://doi.org/10.1017/CBO9781107415324.004

IPCC, 2019. Special Report on Climate Change and Land.

Ivajnšič, D., Kaligarič, M., Žiberna, I., 2014. Geographically weighted regression of the urban heat island of a small city. Appl. Geogr. 53, 341–353.

Jones, B., 2017. Natural disasters: Cities build their vulnerability. Nat. Clim. Chang. 7, 237–238.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J.J., Fiorino, M., Potter, G.L., 2002. Ncep--doe amip-ii reanalysis (r-2). Bull. Am. Meteorol. Soc. 83, 1631–1644.

Kaplan, S., Garrick, B.J., 1981. On the quantitative definition of risk. Risk Anal. 1, 11–27. https://doi.org/10.1111/j.1539-6924.1981.tb01350.x

Karstens, C.D., Stumpf, G., Ling, C., Hua, L., Kingfield, D., Smith, T.M., Correia Jr, J., Calhoun, K., Ortega, K., Melick, C., others, 2015. Evaluation of a probabilistic forecasting methodology for severe convective weather in the 2014 Hazardous Weather Testbed. Weather Forecast. 30, 1551–1570.

Kay, M.P., Brooks, H.E., 2000. Verification of probabilistic severe storm forecasts at the SPC, in: Preprints, 20th Conf. on Severe Local Storms, Orlando, FL, Amer. Meteor. Soc.

Kendall, M.G., 1975. Rank Correlation Methods 4th edition. Charles Griffin, London.

Kis, A.K., Straka, J.M., 2010. Nocturnal Tornado Climatology. Weather Forecast. 25, 545–561. https://doi.org/10.1175/2009waf2222294.1

Klockow, K.E., Peppler, R.A., McPherson, R.A., 2014. Tornado folk science in Alabama and Mississippi in the 27 April 2011 tornado outbreak. GeoJournal 79, 791–804. https://doi.org/10.1007/s10708-013-9518-6

Knupp, K.R., Murphy, T.A., Coleman, T.A., Wade, R.A., Mullins, S.A., Schultz, C.J., Schultz, E. V., Carey, L., Sherrer, A., McCaul, E.W., Carcione, B., Latimer, S., Kula, A., Laws, K., Marsh, P.T., Klockow, K., 2014. Meteorological overview of the devastating 27 april 2011 tornado outbreak. Bull. Am. Meteorol. Soc. 95, 1041–1062. https://doi.org/10.1175/BAMS-D-11-00229.1

Koch, J., 2010. Modeling the impacts of land-use change on ecosystems at the regional and continental scale. kassel university press GmbH.

Krocak, M.J., Brooks, H.E., 2018. Climatological estimates of hourly tornado probability for the United States. Weather Forecast. 33, 59–69.

Kuligowski, E.D., Lombardo, F.T., Phan, L.T., Levitan, M.L., Jorgensen, D.P., 2014. Technical investigation of the May 22, 2011 tornado in Joplin, Missouri. Gaithersburg, Maryland.

Kurdzo, J.M., Bodine, D.J., Cheong, B.L., Palmer, R.D., 2015. High-Temporal Resolution Polarimetric X-Band Doppler Radar Observations of the 20 May 2013 Moore, Oklahoma, Tornado. Mon. Weather Rev. 143, 2711–2735. https://doi.org/10.1175/MWR-D-14-00357.1

Laidley, T., 2016. Measuring sprawl: a new index, recent trends, and future research. Urban Aff. Rev. 52, 66–97.

Lee, J.W., 2002. Tornado proximity soundings from the NCEP/NCAR reanalysis data. University of Oklahoma.

Legendre, P., Legendre, L., 2012. Numerical ecology, 3rd English edition. Elsevier, Amsterdam.

Lim, J., Loveridge, S., Shupp, R., Skidmore, M., 2017. Double danger in the double wide: Dimensions of poverty, housing quality and tornado impacts. Reg. Sci. Urban Econ. 65, 1–15. https://doi.org/10.1016/j.regsciurbeco.2017.04.003

Lindell, M.K., Huang, S.K., Wei, H.L., Samuelson, C.D., 2016. Perceptions and expected immediate reactions to tornado warning polygons. Nat. Hazards 80, 683–707. https://doi.org/10.1007/s11069-015-1990-5

Liu, S., Quenemoen, L.E., Malilay, J., Noji, E., Sinks, T., Mendlein, J., 1996. Assessment of a severe-weather warning system and disaster preparedness, Calhoun County, Alabama, 1994. Am. J. Public Health 86, 87–89.

Lock, N., 2012. An Analysis of Deep Convection Initiation Environments. M.S. Thesis. University of Nebraska.

Mann, H.B., 1945. Nonparametric tests against trend. Econom. J. Econom. Soc. 245–259.

Manson, S., Schroeder, J., Van Riper, D., Ruggles, S., 2017. IPUMS National Historical Geographic Information System: Version 12.0. https://doi.org/10.18128/D050.V12.0

McDonald, J.R., Mehta, K.C., 2006a. A recommendation for an Enhanced Fujita scale (EF-Scale). Wind Science and Engineering Center, Texas Tech University.

McDonald, J.R., Mehta, K.C., 2006b. A Recommendation for an Enhanced Fujita Scale (EF-Scale) 1–95.

McDonald, J.R., Mehta, K.C., Mani, S., 2003. F-scale modification process and proposed revisions. Preprints, in: Symp. on the F-Scale and Severe-Weather Damage Assessment.

Mearns, L., McGinnis, S., Arritt, R., Biner, S., Duffy, P., Gutowski, W., Held, I., Jones, R., Leung, R., Nunes, A., Snyder, M., Caya, D., Correia, J., Flory, D., Herzmann, D., Laprise, R., Moufouma-Okia, W., Takle, G., Teng, H., Thompson, J., Tucker, S., Wyman, B., Anitha, A., Buja, L., Macintosh, C., McDaniel, L., O'Brien, T., Qian, Y., Sloan, L., Strand, G., Zoellick, C., 2014. The North American Regional Climate Change Assessment Program dataset. https://doi.org/10.5065/D6RN35ST

Mearns, L.O., Sain, S., Leung, L.R., Bukovsky, M.S., McGinnis, S., Biner, S., Caya, D., Arritt, R.W., Gutowski, W., Takle, E., others, 2013. Climate change projections of the North American regional climate change assessment program (NARCCAP). Clim. Change 120, 965–975.

Merrell, D., Simmons, K., Sutter, D., 2002. Taking Shelter : Estimating the Safety Benefits of Tornado Safe Rooms. Am. Meteorol. Soc. 17, 619–625. https://doi.org/10.1175/1520-0434(2002)017<0619:TSETSB>2.0.CO;2

Merrell, D., Simmons, K.M., Sutter, D., 2005. The determinants of tornado casualties and the benefits of tornado shelters. Land Econ. 81, 87–99.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E., Berbery, E.H., 2006. North American regional reanalysis. Bull. Am. Meteorol. Soc. 87, 343–360.

Meyer, C.L., Brooks, H.E., Kay, M.P., 2002. A hazard model for tornado occurrence in the United States, in: Preprints, 16th Conf. on Probability and Statistics, Orlando, FL, Amer. Meteor. Soc., J88--J95.

Moller, A.R., Doswell, C.A. III., Foster, M.P., Woodall, G.R., 1994. The operational recognition of supercell thunderstorm environments and storm structures. Weather Forecast. 9, 327–347.

Mooney, C.Z., 1997. Monte carlo simulation. Sage Publications.

Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17–23.

Morss, R.E., Wilhelmi, O. V, Meehl, G.A., Dilling, L., 2011. Improving societal outcomes of extreme weather in a changing climate: an integrated perspective. Annu. Rev. Environ. Resour. 36, 1–25.

Murphy, T.A., Knupp, K.R., 2013. An analysis of cold season supercell storms using the synthetic dual-Doppler technique. Mon. Wea. Rev. 141, 602–624. https://doi.org/10.1175/MWR-D-12-00035.1

Nakaya, T., Fotheringham, A.S., Brunsdon, C., Charlton, M., 2005. Geographically weighted Poisson regression for disease association mapping. Stat. Med. 24, 2695–2717.

Nakicenovic, N., Alcamo, J., Grubler, A., Riahi, K., Roehrl, R.A., Rogner, H.-H., Victor, N., 2000. Special report on emissions scenarios (SRES), a special report of Working Group III of the intergovernmental panel on climate change. Cambridge University Press.

Naylor, J., Gilmore, M.S., 2012. Environmental factors influential to the duration and intensity of tornadoes in simulated supercells. Geophys. Res. Lett. 39.

Notis, C., Stanford, J.L., 1973. The contrasting synoptic and physical character of northeast and southeast advancing tornadoes in Iowa. J. Appl. Meteorol. 12, 1163–1173.

Nowotarski, C.J., Jensen, A.A., 2013. Classifying Proximity Soundings with Self-Organizing Maps toward Improving Supercell and Tornado Forecasting. Weather Forecast. 28, 783–801. https://doi.org/10.1175/WAF-D-12-00125.1

NWS, 1999. The Great Plains tornado outbreak of May 3-4, 1999.

NWS, 2011. Historic Outbreak of April 27, 2011.

NWS, 2017. 6th anniversary of the Joplin tornado - May 22nd, 2011.

NWS, 2018. NWS Damage Assessment Toolkit.

Oklahoma City Planning Commission, 2017. PlanOKC.

Paul, B.K., 2011. Environmental hazards and disasters: contexts, perspectives and management. John Wiley & Sons.

Paul, B.K., Stimers, M., 2012. Exploring probable reasons for record fatalities: the case of 2011 Joplin, Missouri, tornado. Nat. Hazards 64, 1511–1526. https://doi.org/10.1007/s11069-012-0313-3

Paul, B.K., Stimers, M., Caldas, M., 2015. Predictors of compliance with tornado warnings issued in Joplin, Missouri, in 2011. Disasters 39, 108–124.

Paulikas, M.J., 2015. Potential tornado vulnerability variance over a 24-hour cycle for an urban metropolitan region. Dissertation. Kent State University.

Peduzzi, P., Chatenoux, B., Dao, H., De Bono, A., Herold, C., Kossin, J., Mouton, F., Nordbeck, O., 2012. Global trends in tropical cyclone risk. Nat. Clim. Chang. 2, 289.

Pinelli, J.-P., Simiu, E., Gurley, K., Subramanian, C., Zhang, L., Cope, A., Filliben, J.J., Hamid, S., 2004. Hurricane damage prediction model for residential structures. J. Struct. Eng. 130, 1685–1691.

Pohar, M., Blas, M., Turk, S., 2004. Comparison of logistic regression and linear discriminant analysis: a simulation study. Metod. Zv. 1, 143.

Powell, J.A., 1998. Race and space: What really drives metropolitan growth. Brookings Rev. 16, 20.

Powell, M.D., Aberson, S.D., 2001. Accuracy of United States tropical cyclone landfall forecasts in the Atlantic basin (1976-2000). Bull. Am. Meteorol. Soc. 82, 2749–2767.

Prevatt, D.O., van de Lindt, J.W., Back, E.W., Graettinger, A.J., Pei, S., Coulbourne, W., Gupta, R., James, D., Agdas, D., 2012. Making the case for improved structural design: Tornado outbreaks of 2011. Leadersh. Manag. Eng. 12, 254–270.

Prosser, N.E., 1976. Denver's near-miss tornado of 1975. Mon. Weather Rev. 104, 800–803.

Pryor, S.C., Scavia, D., Downer, C., Gaden, M., Iverson, L.R., Nordstrom, R., Patz, J., Robertson, G.P., 2014. Midwest, Climate Change Impacts in the United States: The Third National Climate Assessment. https://doi.org/10.7930/J0J1012N

Radeloff, V.C., Hammer, R.B., Stewart, S.I., 2005. Rural and suburban sprawl in the US Midwest from 1940 to 2000 and its relation to forest fragmentation. Conserv. Biol. 19, 793–805.

Rae, S., Stefkovich, J., 2000. The tornado damage risk assessment predicting the impact of a big outbreak in Dallas-Fort Worth, Texas. Ext. Abstr. 20th Conf. Sev. Local Storms 9.6.

Rasmussen, E.N., Blanchard, D.O., 1998. A Baseline Climatology of Sounding-Derived Supercell andTornado Forecast Parameters. Weather Forecast. 13, 1148–1164. https://doi.org/10.1175/1520-0434(1998)013<1148:ABCOSD>2.0.CO;2

Ratcliffe, M., Burd, C., Holder, K., Fields, A., 2016. Defining rural at the US Census Bureau. Am. community Surv. Geogr. Br. 1–8.

Remo, J.W.F., Carlson, M., Pinter, N., 2012. Hydraulic and flood-loss modeling of levee, floodplain, and river management strategies, Middle Mississippi River, USA. Nat. hazards 61, 551–575.

Remo, J.W.F., Pinter, N., 2012. Hazus-MH earthquake modeling in the central USA. Nat. hazards 63, 1055–1081.

Rosencrants, T.D., Ashley, W.S., 2015. Spatiotemporal analysis of tornado exposure in five US metropolitan areas. Nat. Hazards 78, 121–140. https://doi.org/10.1007/s11069-015-1704-z

Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Commun. 145, 280–297.

Sayemuzzaman, M., Jha, M.K., 2014. Seasonal and annual precipitation time series trend analysis in North Carolina, United States. Atmos. Res. 137, 183–194.

Scheidel, A., Sorman, A.H., 2012. Energy transitions and the global land rush: Ultimate drivers and persistent consequences. Glob. Environ. Chang. 22, 588–595.

Schmidlin, T.W., Hammer, B.O., Ono, Y., King, P.S., 2009. Tornado shelter-seeking behavior and tornado shelter options among mobile home residents in the United States. Nat. Hazards 48, 191–201. https://doi.org/10.1007/s11069-008-9257-z

Schultz, D.M., Richardson, Y.P., Markowski, P.M., Doswell, C.A. III., 2014. Tornadoes in the central United States and the "Clash of Air Masses." Bull. Am. Meteorol. Soc. 95, 1704–1712. https://doi.org/10.1175/BAMS-D-13-00252.1

Seager, R., Lis, N., Feldman, J., Ting, M., Williams, A.P., Nakamura, J., Liu, H., Henderson, N., 2018. Whither the 100th Meridian? The once and future physical and human geography of America's arid--humid divide. Part I: The story so far. Earth Interact. 22, 1–22.

Seto, K.C., Fragkias, M., Güneralp, B., Reilly, M.K., 2011. A meta-analysis of global urban land expansion. PLoS One 6, e23777.

Seto, K.C., Güneralp, B., Hutyra, L.R., 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proc. Natl. Acad. Sci. 109, 16083–16088.

Seto, K.C., Ramankutty, N., 2016. Hidden linkages between urbanization and food systems. Science (80-.). 352, 943–945.

Shafer, C.M., Mercer, A.E., Richman, M.B., Leslie, L.M., Doswell, C.A. III., 2012. An Assessment of Areal Coverage of Severe Weather Parameters for Severe Weather Outbreak Diagnosis. Weather Forecast. 27, 809–831. https://doi.org/10.1175/WAF-D-11-00142.1

Sharma, L., Pandey, P.C., Nathawat, M.S., 2012. Assessment of land consumption rate with urban dynamics change using geospatial techniques. J. Land Use Sci. 7, 135–148.

Sherburn, K.D., Parker, M.D., 2014. Climatology and ingredients of significant severe convection in high-shear, low-CAPE environments. Weather Forecast. 29, 854–877.

Sherman-Morris, K., 2010. Tornado warning dissemination and response at a university campus. Nat. Hazards 52, 623–638. https://doi.org/10.1007/s11069-009-9405-0

Simmons, K.M., Sutter, D., 2006. Direct estimation of the cost effectiveness of tornado shelters. Risk Anal. 26, 945–954. https://doi.org/10.1111/j.1539-6924.2006.00790.x

Simmons, K.M., Sutter, D., 2007. Tornado shelters and the manufactured home parks market. Nat. Hazards 43, 365–378. https://doi.org/10.1007/s11069-007-9123-4

Simmons, K.M., Sutter, D., 2009. False alarms, tornado warnings, and tornado casualties. Weather. Clim. Soc. 1, 38–53. https://doi.org/10.1175/2009WCAS1005.1

Simmons, K.M., Sutter, D., 2011. Economic and societal impacts of tornadoes, 1st ed. American Meteorological Society, Boston, MA.

Simmons, K.M., Sutter, D., 2012. The 2011 tornadoes and the future of tornado research. Bull. Am. Meteorol. Soc. 93, 959–961. https://doi.org/10.1175/BAMS-D-11-00126.1

Simmons, K.M., Sutter, D., Pielke, R., 2013. Normalized tornado damage in the United States: 1950--2011. Environ. Hazards 12, 132–147.

Smith, A.B., Katz, R.W., 2013. US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. Nat. Hazards 67, 387–410. https://doi.org/10.1007/s11069-013-0566-5

Smith, B.T., Thompson, R.L., Dean, A.R., Marsh, P.T., 2015. Diagnosing the conditional probability of tornado damage rating using environmental and radar attributes. Weather Forecast. 30, 914–932.

Snyder, J.C., Bluestein, H.B., 2014. Some considerations for the use of high-resolution mobile radar data in tornado intensity determination. Weather Forecast. 29, 799–827.

Sobash, R.A., Kain, J.S., Bright, D.R., Dean, A.R., Coniglio, M.C., Weiss, S.J., 2011. Probabilistic Forecast Guidance for Severe Thunderstorms Based on the Identification of Extreme Phenomena in Convection-Allowing Model Forecasts. Weather Forecast. 26, 714–728. https://doi.org/10.1175/WAF-D-10-05046.1

Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comput. Exp. 1, 407–414.

SPC, 2017. SVRGIS.

Spies, T.A., White, E., Ager, A., Kline, J.D., Bolte, J.P., Platt, E.K., Olsen, K.A., Pabst, R.J., Barros, A.M.G., Bailey, J.D., others, 2017. Using an agent-based model to examine forest management outcomes in a fire-prone landscape in Oregon, USA. Ecol. Soc. 22 25. 22, 25.

Stensrud, D.J., Xue, M., Wicker, L.J., Kelleher, K.E., Foster, M.P., Schaefer, J.T., Schneider, R.S., Benjamin, S.G., Weygandt, S.S., Ferree, J.T., others, 2009. Convective-scale warn-on-forecast system: A vision for 2020. Bull. Am. Meteorol. Soc. 90, 1487–1500.

Stewart, A.E., Knox, J.A., Schneider, P., 2018. Reaching students and parents through weather science and safety workshops for teachers. Bull. Am. Meteorol. Soc. 99, 1545–1555.

Stone, B., Hess, J.J., Frumkin, H., 2010. Urban form and extreme heat events: are sprawling cities more vulnerable to climate change than compact cities? Environ. Health Perspect. 118, 1425–1428.

Strader, S.M., Ash, K., Wagner, E., Sherrod, C., 2019. Mobile home resident evacuation vulnerability and emergency medical service access during tornado events in the Southeast United States. Int. J. Disaster Risk Reduct. 101210.

Strader, S.M., Ashley, W.S., Irizarry, A., Hall, S., 2015. A climatology of tornado intensity assessments. Meteorol. Appl. 22, 513–524. https://doi.org/10.1002/met.1482

Strader, S.M., Ashley, W.S., Pingel, T.J., Krmenec, A.J., 2017a. Projected 21st century changes in tornado exposure, risk, and disaster potential. Clim. Change 141, 301–313. https://doi.org/10.1007/s10584-017-1905-4

Strader, S.M., Ashley, W.S., Pingel, T.J., Krmenec, A.J., 2017b. Observed and projected changes in United States tornado exposure. Weather. Clim. Soc. 9, 109–123.

Strader, S.M., Ashley, W.S., Pingel, T.J., Krmenec, A.J., 2018. How land use alters the tornado disaster landscape. Appl. Geogr. 94, 18–29. https://doi.org/10.1016/j.apgeog.2018.03.005

Strader, S.M., Pingel, T.J., Ashley, W.S., 2016. A Monte Carlo model for estimating tornado impacts. Meteorol. Appl. 23, 269–281. https://doi.org/10.1002/met.1552

Suckling, P.W., Ashley, W.S., 2006. Spatial and temporal characteristics of tornado path direction. Prof. Geogr. 58, 20–38. https://doi.org/10.1111/j.1467-9272.2006.00509.x

Thompson, R.L., Edwards, R., 2000. An Overview of Environmental Conditions and Forecast Implications of the 3 May 1999 Tornado Outbreak. Weather Forecast. 15, 682–699. https://doi.org/10.1175/1520-0434(2000)015<0682:AOOECA>2.0.CO;2

Thompson, W.R., 2009. Variable selection of correlated predictors in logistic regression: investigating the diet-heart hypothesis. Florida State University.

Tinsley, C.H., Dillon, R.L., Cronin, M.A., 2012. How near-miss events amplify or attenuate risky decision making. Manage. Sci. 58, 1596–1613. https://doi.org/10.1287/mnsc.1120.1517

Tippett, M.K., Allen, J.T., Gensini, V.A., Brooks, H.E., 2015. Climate and Hazardous Convective Weather. Curr. Clim. Chang. Reports 1, 60–73. https://doi.org/10.1007/s40641-015-0006-6

Tippett, M.K., Cohen, J.E., 2016. Tornado outbreak variability follows Taylor's power law of fluctuation scaling and increases dramatically with severity. Nat. Commun. 7, 10668.

Trapp, R.J., Diffenbaugh, N.S., Brooks, H.E., Baldwin, M.E., Robinson, E.D., Pal, J.S., 2007. Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. Proc. Natl. Acad. Sci. 104, 19719–19723. https://doi.org/10.1073/pnas.0705494104

Trapp, R.J., Tessendorf, S.A., Godfrey, E.S., Brooks, H.E., 2005. Tornadoes from squall lines and bow echoes. Part I: Climatological distribution. Wea. Forecast. 20, 23–34. https://doi.org/10.1175/WAF-835.1

U.S. Census Bureau, 2010. American Fact Finder [WWW Document]. URL https://factfinder.census.gov (accessed 7.18.19).

U.S. Census Bureau, 2015. 2011-2015 American Community Survey 5-Year Estimates.

U.S. Census Bureau, 2018. QuickFacts: Oklahoma City city, Oklahoma [WWW Document].

Uccellini, L.W., 2014. May 2013 Oklahoma Tornadoes and Flash Flooding National Oceanic and Atmospheric Administration 63.

United Nations, 2018. 2018 revision of world urbanization prospects.

Verbout, S.M., Brooks, H.E., Leslie, L.M., Schultz, D.M., 2006. Evolution of the U.S. tornado database: 1954-2003. Weather Forecast. 21, 86–93. https://doi.org/10.1175/WAF910.1

Verburg, P.H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., Mastura, S.S.A., 2002. Modeling the spatial dynamics of regional land use: The CLUE-S model. Environ. Manage. 30, 391–405. https://doi.org/10.1007/s00267-002-2630-x

Westra, S., Alexander, L. V, Zwiers, F.W., 2013. Global increasing trends in annual maximum daily precipitation. J. Clim. 26, 3904–3918.

Wurman, J., Alexander, C., Robinson, P., Richardson, Y., 2007. Low-level winds in tornadoes and potential catastrophic tornado impacts in urban areas. Bull. Am. Meteorol. Soc. 88, 31–46. https://doi.org/10.1175/BAMS-88-1-31

Wurman, J., Kosiba, K., Robinson, P., Marshall, T., 2014. The role of multiple-vortex tornado structure in causing storm researcher fatalities. Bull. Am. Meteorol. Soc. 95, 31–45. https://doi.org/10.1175/BAMS-D-13-00221.1

Yohe, G., Tol, R.S.J., 2002. Indicators for social and economic coping capacity—moving toward a working definition of adaptive capacity. Glob. Environ. Chang. 12, 25–40.

Yue, S., Pilon, P., Cavadias, G., 2002. Power of the Mann--Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. J. Hydrol. 259, 254–271.

Zoraster, R.M., 2010. Vulnerable populations: Hurricane Katrina as a case study. Prehosp. Disaster Med. 25, 74–78.

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Appendix A - Supplementary Material for Tordis Model Validation

Model validation was conducted on the deterministic components of TorDIS, with comparisons between the stochastic components and the observed record created during the model application examples. For the regression equations and tornado day forecasts the validation was done on the 2015 – 2016 testing data, while the tornado direction was validated against the entire 1979 – 2016 tornado record as none of the data was used to specify the relationship between tornado direction and 500 mb wind direction.

A.1) Tornado Forecast Model Selection and Validation

The initial tornado forecast model was validated against observed tornado reports from 2015 – 2016 for both, its ability to capture the spatial distribution of tornadoes and its ability to forecast whether a day would be tornadic (at least one tornado reported in the model domain). Three verification metrics were used in the model validation (Table A.1): probability of detection (POD), false alarm rate (FAR), and Heidke Skill Score (HSS). The initial model had a low HSS (0.9%) for the tornado distribution forecasts and moderate HSS (39.1%) for the tornado day forecasts. To determine whether each variable was necessary to the model, a global sensitivity analysis was conducted, and the Sobol Total Order Index (S_t) was calculated for each model parameter. All variables with an S_t of 5% or greater were kept for the final model. At least half of the model output variance was explained by the top two variables for each of the four regressions (svr$_{low}$, torn$_{low}$, svr$_{mod}$, and torn$_{mod}$) with more than 90% explained by all variables with an S_t greater than 5% (Figure A.1). The variables to which the regression equations were most sensitive (CAPE for svr$_{low}$, VWS1 for torn$_{low}$, CIN for svr$_{mod}$, and ELEV for torn$_{mod}$) were all reasonable. Severe thunderstorms require sufficient instability in order to promote the required

157

vertical motion (Schultz *et al.*, 2014). Tornadoes which occur in low CAPE environments typically occur overnight or during the cool season when vertical wind shear is higher (Sherburn and Parker, 2014). As a result, small changes in CAPE can push the instability to a sufficient value to promote tornadoes. In environments with sufficient CAPE, strong CIN can hamper the development of severe thunderstorms with small reductions in CIN making initiation possible in some circumstances (Bunkers *et al.*, 2010). Tornadic thunderstorms require high low-level wind shear in order to have enough rotation to tilt into the updraft (Knupp *et al.*, 2014). It is unclear why ELEV explained most of the variance for tornadoes in moderate CAPE environments. However, tornadoes are rare at high elevations (Brooks *et al.*, 2003b) and VWS1 was only slightly less important. Initially, ω_{850} and ω_{500} were dropped due to low S_t (Figure A.1). Additionally, further testing showed that these variables, while not important overall, significantly improved the probability of detecting tornadoes (by over 30%) in winter. This could be because instability tends to be lower in the cold season making the presence of broad-scale lifting (*e.g.*, air-mass boundaries) more important (Schultz *et al.*, 2014; Sherburn and Parker, 2014).

The final model showed a low HSS (0.9%) for the tornado distribution forecasts (0.9%); however, it showed a high POD (76.7%). This was slightly below the HSS (2.0%) for the probabilistic outlooks and the POD for the outlooks was slightly lower (72.6%). Seasonally, the model had the greatest HSS in winter (1.3%) and the lowest HSS in summer (0.4%), similar to the outlooks (Table A.1). The final model had moderate HSS (43.8%) for the tornado day forecasts with a high POD (94.2%). The probabilistic outlooks again had a greater HSS (55.8%) and lower POD (82.7%). The model showed the greatest HSS for tornado day forecasts in spring (47.3%) and lowest HSS in summer (12.4%) with the seasonal patterns in HSS differing slightly

from those of the probabilistic outlooks (highest HSS was in winter for the outlooks). Overall, the final model had a slightly higher HSS for both tornado day and distribution forecasts and a slightly lower POD over the initial model (Table A.1). Due to the small differences the final model was accepted for simplicity.

A.2) Predicting Tornado Direction

Tornado direction was simulated using the 500 mb wind direction at the location of tornado touchdown. This method was chosen over the more complex method of Bunkers *et al.* (2000) for simplicity given a lack of data on parent storm direction for tornadoes in the historical record. When testing for the quadrant in which the tornadoes move, it was found that most tornadoes (89%) during 1979 – 2016 moved to the northeast with the second most (7%) moving to the southeast. Likewise, 500 mb winds, at the point of touchdown during 1979 – 2016, tended to be towards the northeast (79%) and southeast (16%). A comparison of tornado direction against 500 mb wind direction showed that they were the same 75% of the time. This matches findings by Suckling and Ashley (2006), which suggested that 500 mb flow is closely linked to tornado direction due to the steering effects of upper-level winds on storms.

Table A.1. Verification metrics (in per cent) for the spatial distribution of tornadoes as well as for the ability to forecast a tornado day (a day with at least one reported tornado in the model domain) by season and annually. Metrics include Probability of Detection (POD), False Alarm Rate (FAR), and Heidke Skill Score (HSS) and were calculated before and after the model parameter screening procedure.

Before Screening

Verification of Spatial Distribution										
	Winter		Spring		Summer		Autumn		All Months	
Metric	Model	SPC	Model	SPC	Model	SPC	Model	SPC	Model	SPC
POD	86.8	90.7	84.2	77.1	58.1	49.5	78.6	76.7	78.1	72.6
FAR	99.4	98.3	99.3	98.8	99.7	99.4	99.6	98.8	99.5	98.9
HSS	1.2	3.2	1.1	2.2	0.5	1.1	0.7	2.3	0.9	2

Verification of Daily Forecasts										
	Winter		Spring		Summer		Autumn		All Months	
Metric	Model	SPC	Model	SPC	Model	SPC	Model	SPC	Model	SPC
POD	94.4	94.4	100	92.2	88.6	72.7	93.3	76.7	94.2	82.7
FAR	64.6	43.3	35.7	23.9	49.7	41.3	71.1	54	51.6	37.2
HSS	43.3	66.7	46.3	64.3	8.2	25.6	25.3	46.5	39.1	55.8

After Screening

Verification of Spatial Distribution										
	Winter		Spring		Summer		Autumn		All Months	
Metric	Model	SPC	Model	SPC	Model	SPC	Model	SPC	Model	SPC
POD	80.1	90.7	83.7	77.1	55.2	49.5	73	76.7	75.7	72.6
FAR	99.2	98.3	99.3	98.8	99.7	99.4	99.6	98.8	99.4	98.9
HSS	1.5	3.2	1.1	2.2	0.6	1.1	0.8	2.3	1	2

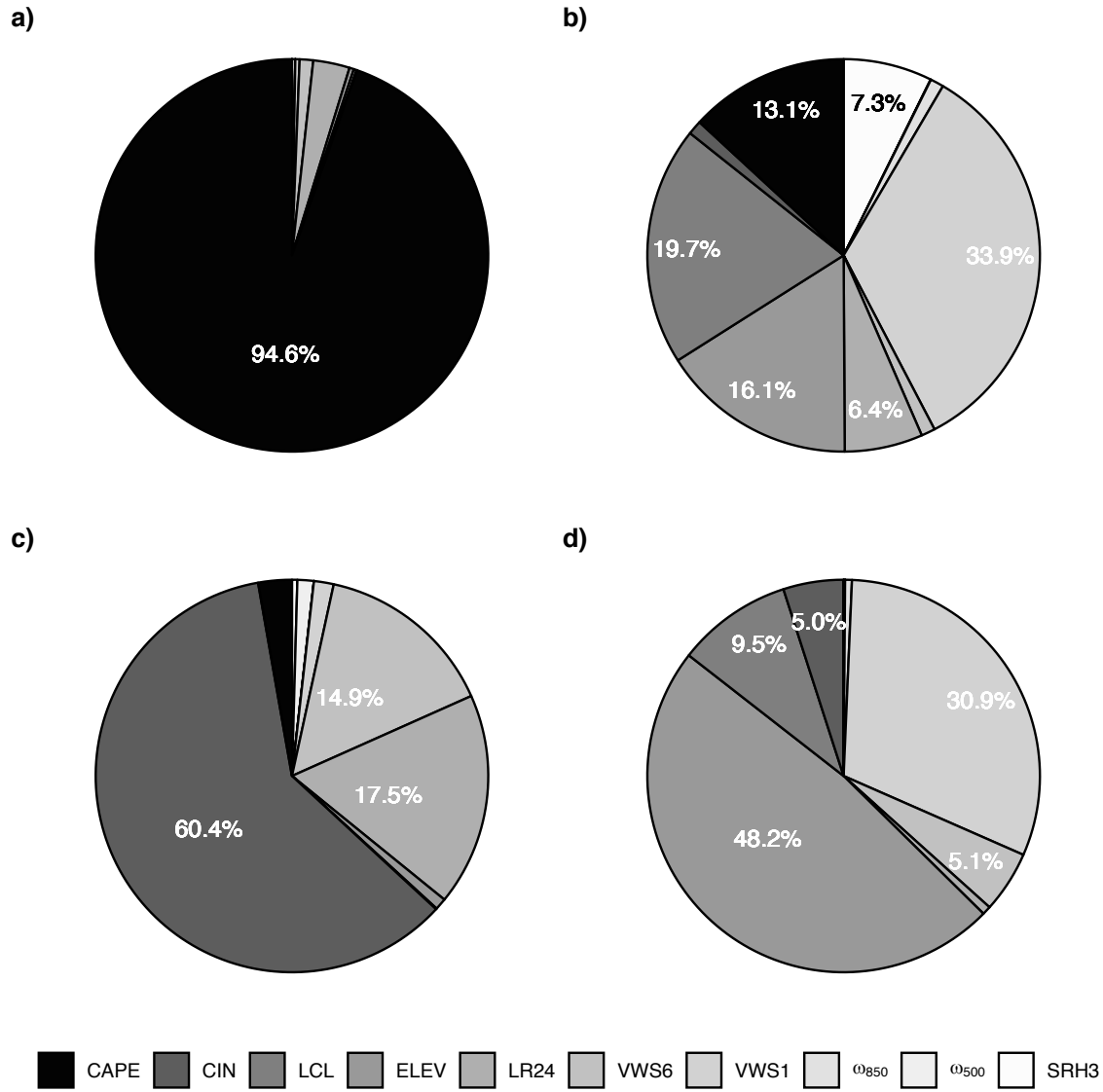| Verification of Daily Forecasts										
	Winter		Spring		Summer		Autumn		All Months	
Metric	Model	SPC	Model	SPC	Model	SPC	Model	SPC	Model	SPC
POD	88.9	94.4	100	92.2	88.6	72.7	90	76.7	93.4	82.7
FAR	59	43.3	35.7	23.9	46.9	41.3	68.2	54	48.7	37.2
HSS	49.2	66.7	46.3	64.3	16.3	25.6	29.9	46.5	43.8	55.8

Figure A.1. Total order global sensitivity indices for all logistic regression parameters following the method of Sobol (1993). The following equations are included: the probability of severe weather in a low or moderate to high CAPE environment (a, c) and the probability of tornadoes given severe weather in a low or moderate to high CAPE environment (b, d). Parameter indices with values above 5% are labelled.