Highlights

Improving tornado casualty predictions in the US with population exposure data and a modified social vulnerability index

Vikalp Mishra, Eric R. Anderson, Skyler Edwards, Robert E. Griffin

- A new spatially-weighted social vulnerability index specific to tornadoes was created.
- Combining exposure data with hazard data more accurately predicts tornado casualties. The addition of social vulnerability information marginally improves casualty predictions.
- Together, models with population exposure and underlying social vulnerability were better able to predict casualties than models with single predictors.

Improving tornado casualty predictions in the US with population exposure data and a modified social vulnerability index

Vikalp Mishra[a,b], Eric R. Anderson[a,c], Skyler Edwards[a,d], Robert E. Griffin[a,d]

[a]NASA SERVIR Science Coordination Office, 320 Sparkman Dr., Huntsville, 35805, AL, USA
[b]Earth System Science Center, The University of Alabama in Huntsville, Huntsville, 35805, AL, USA
[c]Earth Science Branch, NASA Marshall Space Flight Center, Huntsville, 35805, AL, USA
[d]Department of Atmospheric and Earth Science, The University of Alabama in Huntsville, Huntsville, 35805, AL, USA

Abstract

Tornadoes are frequent and widespread events that often account for hundreds of injuries and fatalities, and millions of dollars in damages. Multiple studies have analyzed tornado climatology and associated exposure; however, fewer have focused on predicting fatalities and injuries by coupling geospatial data on tornado characteristics and underlying social and economic vulnerability. In this study, we test the ability of negative binomial regression models to predict tornado-induced injuries and fatalities by coupling data on physical characteristics of tornadoes, exposure of populations, and underlying social vulnerability. We also present a modified spatially weighted social vulnerability index SVI^*_{wt}. We used 10-year (2005-2014) tornado data over the continental United States for this analysis. The results of this study indicate that the tornado length, magnitude and nocturnality seem to be the major hazard-related indicators of fatalities (McFadden's Pseudo-R^2 ranged from 0.01-0.12). Population exposure and SVI^*_{wt} are positive and statistically significant in predicting tornado related fatalities and injuries (Pseudo R^2 of 0.11-0.18). Although combining SVI^*_{wt} with hazard variables does not substantially improve model fit when compared to adding population exposure, combining SVI^*_{wt}, hazard and exposure results in better predictions of both injuries and fatalities (Pseudo R^2 of 0.17) and is also an improvement on previous similar studies.

Keywords: Disaster, Vulnerability, Tornado, Natural Hazard, SVI

1. Introduction

Tornadoes are frequent and widespread hazards in the United States. Nearly 800-1400 tornadoes are reported annually, which resulted in a mean fatality count of more than 350 per year between 1880-2005 (Ashley, 2007). A total of 60,114 tornado events have been officially recorded between 1950 and 2015 by the National Centers for Environmental Information (NCEI, formally known as National Climatic Data Center) in the U.S. This has resulted in more than 5,800 fatalities and nearly 94,000 injuries in addition to billions of dollars of losses in terms of property and crop damage. Over the years, tornado related fatalities and personal injuries have witnessed a decreasing trend (Platt, 1999) which could be attributed to the advancement in detection technologies, systematic warnings, with increased public awareness. Since 1950, overall tornado related fatalities in the U.S. have decreased (1415 in the 1950s compared to 603 in the 2000s). However, such trends may not be consistent across space and time (Ashley & Strader, 2016). For instance, the tornado events of 27 April 2011 over the southeastern U.S. caused more than 500 fatalities and several hundred injuries. Despite advances such as state-of-the-art detection and warning dissemination technology, tornado related casualties cannot be prevented entirely. Recent tornado events also show that not only the intensity of such extreme events varies by region and time (Strader & Ashley, 2018; Nouri et al., 2021) but also by people's responses to the risk (Fricker, 2020).

Significant research on tornadoes has been focused on aspects of climatological and meteorological hazards (Doswell & Burgess, 1988; Bluestein, 1999; Grazulis, 2001; Boruff et al., 2003; Brooks et al., 2003; Feuerstein et al., 2005; Verbout et al., 2006). Further, impact-based research has focused on human casualties [e.g., Ashley (2007); Ashley et al. (2008)] and on combined human and economic losses [e.g., Brooks & Doswell (2001); Boruff et al. (2003); Diaz & Joseph (2019)]. Factors such as tornado magnitude, length, and width, in association with the natural and built environment have been found to be directly related to most of the fatalities and injuries caused by tornadoes (Ashley et al., 2014; Wurman et al., 2007; Simmons & Sutter, 2011; Paul & Stimers, 2014; Elsner et al., 2016; Gensini & Brooks, 2018; Strader & Ashley, 2018; Grieser & Haines, 2020). Studies such as Paul (2003); Simmons & Sutter (2005); Ashley (2007); Ashley et al. (2008) etc. found that nocturnal tornadoes tend to be relatively more lethal to humans, whereas tornado timing has little impact on the extent of economic losses. Mobile and manufactured homes are another factor that is commonly associated with a higher percentage of fatalities and has been extensively investigated over the years (Eidson et al., 1990; Schmidlin & King, 1995; Simmons & Sutter, 2011;

Strader & Ashley, 2018; Ash, 2017). Small differences between mobile and manufactured notwithstanding, much of the tornado literature suggests that the two terms can be used interchangeably and refer to housing structures that tend to be poorly anchored and can be more susceptible to sustaining damage from strong winds Strader & Ashley (2018); Strader et al. (2021); Simmons & Sutter (2011); Fricker (2020); Sutter & Simmons (2010). Ashley (2007) reported that nearly 44% of total tornado fatalities are associated with manufactured homes. Further, Simmons & Sutter (2011) reported that there is almost 15 times greater probability of tornado related fatalities in manufactured homes. More recently, Strader & Ashley (2018) argue that the impact of tornadoes on manufactured homes varies across regions, and other complementary socioeconomic factors also play a significant role. The findings were in line with the observations by Schmidlin & King (1995), which suggest that in addition to tornado physical characteristics and population exposure, several demographic and socioeconomic conditions of communities also influence the likelihood of sustaining loss. Demographic variables such as higher percentages of disabled, elderly, or child populations, contribute towards vulnerability to natural hazards at a location (Kilijanek & Drabek, 1979; Carter et al., 1989; Eidson et al., 1990; Cutter et al., 2003; Flanagan et al., 2011; Cutter et al., 2014; Strader et al., 2021). In the context of this study, vulnerability to tornadoes can be described as intrinsic characteristics of the socioeconomic ecosystem that create the potential for harm during a tornado event (Cutter & Finch, 2008).

Over the years, researchers have tried to assess social vulnerability to external stresses by constructing indices using combinations of socioeconomic variables. Examples of such indices with a limited number of variables include Cutter et al. (2000); Wu et al. (2002); Wood & Good (2004); Rygel et al. (2006); Dixon & Moore (2012) etc. Petit et al. (2012) developed indices to assess the vulnerability of critical infrastructure to disasters. Cutter et al. (2003) and later Flanagan et al. (2011) developed an index of social vulnerability (SoVI) to environmental hazards. SoVI has been replicated and used in multiple geographical settings (Boruff & Cutter, 2007) and at various spatial and temporal scales (Cutter, 2006; Borden et al., 2007; Cutter & Finch, 2008; Gottlieb et al., 2008). Flanagan's vulnerability index has been adopted as the official social vulnerability index (SVI) used by the U.S. Centers for Disease Control and Prevention (CDC). It should be noted that vulnerability indices alone do not depict the potential impact of events such as tornadoes. Disaster risk research points to the need to combine knowledge on hazard, exposure, and vulnerability, when understanding overall risk (Shah et al., 2020; Alexander, 2002; Amendola et al., 2008) and to better prioritize disaster mitigation and prevention strategies (Marin et al., 2021). Targeted reduction of social vulnerability and improved risk coverage in the most exposed regions and census tracts may go further to reduce nationwide tornado-induced casualties, versus allocating disaster relief funds solely based on economic damages. While there are multiple studies that have addressed tornado risk assessments, most go as far as linking hazard and exposure aspects (Shen & Hwang, 2015; Fricker et al., 2017; Elsner et al., 2018; Masoomi & van de Lindt, 2018), leaving comprehensive studies that connect hazard, exposure, and vulnerability to tornadoes relatively unexplored. More recently, Strader et al. (2021) assessed tornado risk, vulnerability and exposure at a county warning area scale, and Fricker (2020) argued for the need of tornado-level social vulnerability to better predict the number of causalities and fill in the gaps in current prediction.

Thus, it is imperative to improve our understanding of the interactions between the physical characteristics of tornadoes and the underlying socioeconomic conditions of exposed populations, and at finer spatial scales commensurate with the small hazard footprints of tornadoes. This could allow for better prediction of related casualties and enhance mitigation, preparedness, response and recovery efforts to ultimately reduce future casualties. Therefore, the aim of this study is to identify and evaluate potential predictors of tornado injuries and fatalities. We approach this by testing the explanatory value of data on population exposure and a modified spatially-weighted social vulnerability index (SVI), versus considering hazard-related variables alone. More broadly, this will allow us to test to what extent the generally-held theory that defining disaster risk by its determinants (hazard, exposure, and vulnerability) holds for the specific context of tornado casualties. This study utilizes multi-year (2005-2014) tornado data over the continental United States (CONUS). We used negative binomial regression models with tornadoes as the unit of analysis, building models to test four hypotheses to estimate tornado related casualties: hazard only (HO); hazard and exposure variables (H1); hazard with vulnerability variables (H2); and finally, a combination of hazard, exposure and vulnerability variables (H3; see the Modeling Strategy section for the framework of statistical models).

2. Data Sources and Pre-Processing

2.1. Tornado hazard variables

The tornado data utilized in this study was obtained from the Storm Prediction Center (SPC) severe weather database files. Despite having certain shortcomings particularly in the estimation of economic losses (Gall et al., 2009), casualty counts in terms of location, timings of the incident and direct vs indirect impact (Ashley, 2007; Ashley et al., 2008) and issues pertaining to biases and trends as studied by Doswell & Burgess (1988); Verbout et al. (2006) etc., the database is the official national tornado archive constructed from National Weather Services (NWS) storm reports since 1950. The database contains several important tornado related information such as the start and end locations, timings, magnitude, losses, etc., and has been the primary source of information. Tornado data from 2005 to 2014 were used in this study for the CONUS. A total of 11,970 tornado events were reported in the database during the study period.

For a comprehensive tornado risk assessment, tornado occurrences with all of the associated physical attributes are considered. Tornado records with both start and end coordinates can be mapped, therefore only 11,940 tornado events can be used in this analysis. These tornado events resulted in 9,940 injuries and 1,103 fatalities for the CONUS. Tornado events are not uniformly distributed either annually or monthly and may play into distinct vulnerabilities during different years and months (Figure 1a and 1b). As expected, low damage rating tornadoes (F/EF0) hardly caused fatalities and injuries, whereas each of the highest damage rating tornadoes (F/EF5) accounted for nearly 90 injuries and 37 fatalities on average during the study period (Figure

1c). Although a recent studies by Edwards et al. (2021) show slight differences in tornado climatology (track width and length in particular) between the F and EF scales, both methods are subjective and susceptible to systematic biases. Additionally, since the scale is derived using damages rather than wind speed, F and EF scale can be considered in the same category. Therefore, in this study F and EF scale ratings are assumed to represent same damage category.

Characteristic tornado attributes such as date and time of occurrence; location; damage rating (F/EF-scale); length and widths were recorded for each tornado event in the database. Information on the location, length, width and damage rating, were directly used to map the spatial extent of the tornado event over census tracts. Tornado frequencies and associated fatalities and injuries are known to vary regionally (Ashley, 2007; Simmons & Sutter, 2012; Long et al., 2018; Ashley & Strader, 2016), resulting in distinct tornado seasons per region. For this study, the CONUS is broadly divided into four major regions, and region-specific seasonality is determined using a long-term (1950-2012) regional monthly tornado frequency distribution (Figure 2). Months with tornado activities above the long-term annual average tornado frequency (shown in bold) are considered to be the regional tornado season. Tornado season in the Southern US lasts from March – June; Eastern tornado season lasts longest from April – September; whereas Central tornado season extends from April – July; and the Western region's season extends from May - August.

Responses and attitudes towards tornadoes differ per season, location, and timing Simmons & Sutter (2008). Studies by Paul (2003); Simmons &

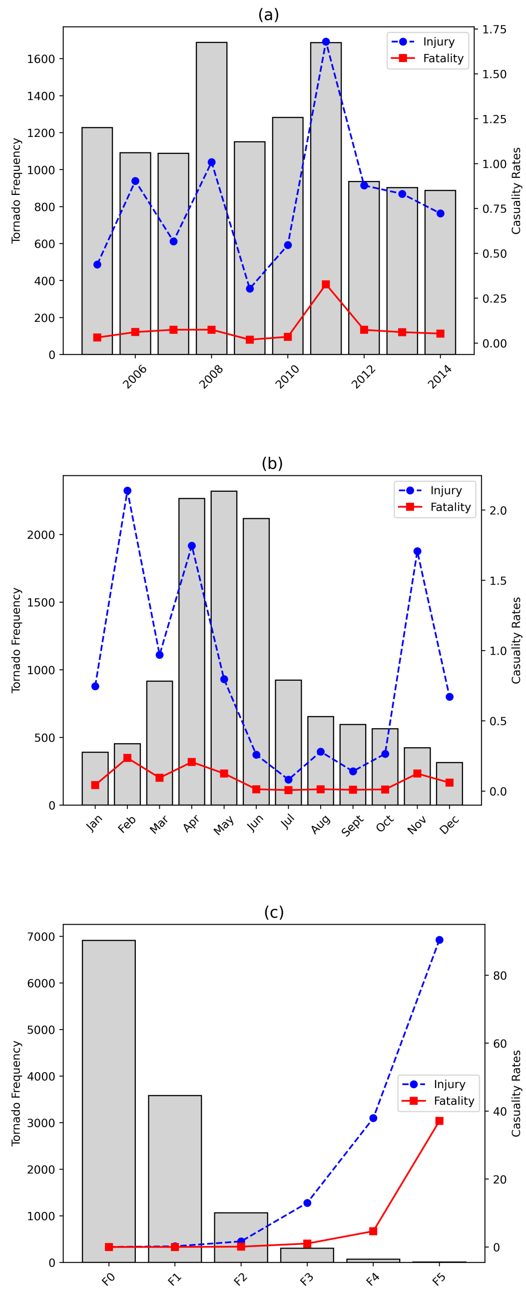

Figure 1: Frequency distribution along with total injury and fatality at (a) Annual, (b) Monthly scales and with respect to magnitude from 2005-2014 over CONUS

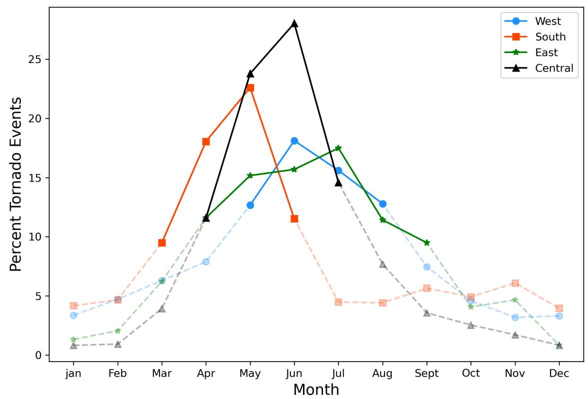

Figure 2: Regional tornado seasonality based on long-term tornado monthly frequency (1950-2013). Bold lines denote the months identified as regional tornado season based on annual regional means.

Sutter (2005); Ashley et al. (2008) showed that tornado timing also affects the rate of potential fatalities and injuries to such events. Nocturnal (between local sunset to sunrise) tornado events were relatively more fatal and accounted for nearly 2.5 times more fatalities than daytime tornadoes (Ashley et al., 2008). It is assumed that during a tornado season, in the daytime or on a workday, people are more likely to be alert and reactive to tornado alerts and recovery actions as opposed to non-seasonal, weekend nighttime. Therefore, tornadoes as a function of factors such as weekend vs. weekdays; nocturnal vs. daytime; and tornado season vs. non-seasonal were evaluated. Nocturnal in this study is defined as tornadoes occurring between local 12:00 am to 6:00 am.

2.2. Socioeconomic and Demographic Factors: a Social Vulnerability Index

The vulnerability index by CDC is created using decadal census data, but our study period of 2005-2014 spans across two decades. To remedy this mismatch, the vulnerability index from the Agency for Toxic Substances and Disease Registry (ATSDR) was recreated using data from the U.S. Census Bureau and American Community Survey (ACS) following the original method by Flanagan et al. (2011). The recreated index is referred to as SVI* to distinguish it from CDC's official SVI. Both 2010 decadal data from the Census Bureau and 2006-2010 survey data from ACS were used in this study. The census data mainly provided the social counts, whereas household and economic counts were available from ACS data. In this study the vulnerability was computed at the census tract level since it is a typical unit to collect and analyze data for policy and management practices. Census tracts are designed to be as demographically homogeneous as possible with populations ranging from 1500 to 8000. The SVI* was constructed with a total of 14 census variables consisting of indicators such as social and economic status; housing and household composition; disability and minority status; language barrier; and access to transportation and basic amenities as a measure of vulnerability. All variables were grouped into four basic SVI themes: (a) socioeconomic factors; (b) household composition; (c) minority/language and (d) housing/transportation (Table 1).

Socioeconomic variables

Economically disadvantaged sections of society are more likely to be the least prepared for, worst hit, and generally lacking in resources necessary to recover from a disaster (Cutter et al., 2003; Schmidtlein et al., 2008).

Further, such sections are also less likely to have access to either the health or property insurances needed to help recover from disasters, increasing their vulnerability to such events (Fothergill & Peek, 2004). Therefore, people with less education, unemployed people, and those living under poverty or with less per capita income are seen as more socioeconomically vulnerable.

Another section consisting of the elderly and young people are known to be vulnerable to natural hazards (Cutter et al., 2003; Ngo, 2001) due to insufficient life experience or knowledge; limited access to resources; having physical, sensory or cognitive challenges; and are relatively more susceptible to injuries and fatalities (Schmidlin & King, 1995; Kar, 2009; Anderson, 2005). Hence, people aged over 65 or less than 17, households with single parents, and children under 18, were considered as vulnerable. Language barrier, poverty and social discrimination also increase vulnerability. This places many racial and ethnic minorities in more vulnerable positions to disasters (Fothergill et al., 1999; Bolin, 2006; Elliott & Pais, 2006).

Manufactured homes have been proven to account for a disproportionately high rate of injuries and fatalities during tornadoes (Eidson et al., 1990; Donner, 2007; Simmons & Sutter, 2009). Similarly, multi-unit housing in densely populated areas, people living in group quarters such as hospitals, dormitories, prisons, etc., also possess additional risks during a disaster (Cutter et al., 2003). Finally, inaccessibility to transportation during a disaster can also increase social vulnerability.

Variables	**SVI Themes**
Proportion of persons aged 65 and above	Household Composition
Proportion of persons aged below 17	
Proportion of single parent household with children<18	
Proportion of people with minority status	Minority/Language
Proportion of persons (5+) who speak English less than well	
Proportion of persons below poverty level	Socioeconomic
Per capita Income	
Proportion of civilians (16+) unemployed	
Persons 25+ with no high school diploma	
Proportion of manufactured homes	Housing/Transportation
Proportion of people living in group quarters	
Proportion of household with no access to vehicle	
Proportion of housing structures with 10 or more units	
Proportion of households with more people than room	

Table 1: Socioeconomic and demographic variables used to construct SVI* using data from Census data 2010 (SF1) (italicized data fields are obtained from ACS 2006-2010)

2.3. Data Processing

2.3.1. SVI^* Calculation

All demographic variables (except for per capita income) were first arranged in decreasing order and ranked from highest to lowest for all census tracts across CONUS. A higher value indicates a higher vulnerability. The per capita income on the other hand was ranked from lowest (most vulnerable) to highest. A percentile ranking was then computed for each variable:

$$PR = \frac{(Rank - 1)}{(N - 1)} \tag{1}$$

Where PR is the percentile rank, and N is the total number of census tracts. The method is consistent with the approach in Flanagan et al. (2011); the variable PR(s) were summed first across themes, and the process was repeated to construct a composite index for all variables. In this study errors in either census or ACS were considered inherent, and the data is used as is. A total of 14 states had 30% or more tracts with $SVI^* > 0.75$, indicating higher vulnerability, with Mississippi and Alabama being the states with the highest percentages (52.8% and 42.44%) of tracts in vulnerable condition. Incidentally, of these 14 states with highly vulnerable census tracts, 9 are in well-known tornado prone locations (such as Texas, Alabama, Mississippi, West Virginia, Kentucky, Louisiana, South Carolina, Georgia and Tennessee). Seven of those are in the southeastern US, making them extremely susceptible to natural hazards and subsequent recovery efforts (Figure 3).

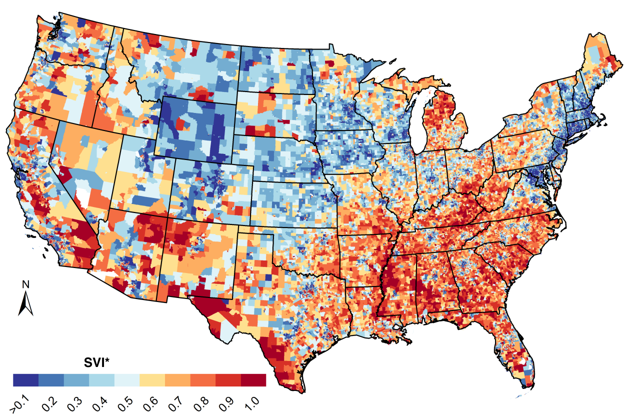

Figure 3: A map showing the computed social vulnerability index (SVI*) for 2010 at the census tract level for the CONUS

2.3.2. *GIS Analysis: a Spatially Modified Social Vulnerability Index*

Spatially collocated tornado tracks and their underlying SVI* data at a census tract level were joined using a geospatial platform. Multiple spatial transformations were performed such as buffering, splitting, spatial joining, summary statistics, etc., to allow for the integration of tornado tracks (lines) with census tract SVI* data (polygons).

Many tornadoes cross several census tracts, each tract having its unique SVI* value. To account for the varying degrees of social vulnerability within each tornado path, the weighted average of the SVI* data were calculated for each tornado. Tornado tracks (Figure 4) completely contained within a census tract boundary were assigned the respective tract's vulnerability indices. Whereas, for tornado tracks across multiple tracts, a weighted average vulnerability index was used to account for each census tract based upon the fraction of area affected by a buffered tornado track. Using the number of tracts through which a particular tornado passes as n and f_{area} as the fraction of total tornado area that falls in a given census tract, then the weighted average social vulnerability index SVI^*_{wt} was computed as:

$$SVI^*_{wt} = \sum f_{area} \sum SVI^* \quad (2)$$

To calculate the population exposed to each tornado event, a similar fractional area-based approach was later applied to estimate the total population expected to be under a tornado path. A uniform population density across the census tract was assumed, then the total population expected to be under the tornado's area of influence was estimated based on the fraction of the tornado path area within a census tract as:

$$POP_{Tor} = \sum f_{area} \sum POP_{den} \tag{3}$$

The spatial distribution of tornadoes by magnitude is uneven across the CONUS (Figure 4). Table 2 summarizes all tornado hazard characteristic variables used in this study: F/EF-scale, seasonality, nocturnality, and weekday. Counts are given per each characteristic along with the total number of associated fatalities and injuries.

It should be noted that the tornado width data is estimated at the widest portion across the tornado path and may contain uncertainties (Masoomi & van de Lindt, 2018). In this study, a constant maximum width is assumed across tornado tracks, which may result in some overestimation of the actual area under tornado path. Moreover, not all tornado paths are straight (such as the 2011 Joplin, Missouri tornado). In the absence any data on the varying widths of tornado paths along their tracks, this is the closest approximation to tornado impact area Fricker (2020).

3. Modeling Strategy

The aim of this study is to assess the how well data on tornado hazard characteristics and underlying socioeconomic conditions can predict injuries and fatalities. Earlier studies on related topics have shown the value of using Poisson and Negative Binomial Regression in predicting dependent variables of a count data (Donner, 2007; Elsner et al., 2018; Fricker et al., 2017; Masoomi & van de Lindt, 2018; Potvin et al., 2019; Refan et al., 2020; Schroder & Elsner, 2021; Simmons & Sutter, 2005, 2009; Zahran et al., 2013; Lim et al., 2017). Simmons & Sutter (2008) suggested that the Poisson model is

Tornado Characteristics	Description	Frequency	Injury	Fatality
Magnitude	F0	6912	130 (0.02)	2 (0.00)
	F1	3584	758 (0.21)	38 (0.01)
	F2	1066	1770 (1.66)	109 (0.10)
	F3	302	3926 (13.0)	308 (1.02)
	F4	67	2542 (37.9)	312 (4.66)
	F5	9	814 (90.4)	334 (37.1)
Seasonality	Seasonal	8148	6895 (0.85)	832 (0.10)
	Non-seasonal	3792	3045 (0.80)	271 (0.07)
Nocturnality	Daytime	11089	8901 (0.80)	1008 (0.09)
	Nocturnal	851	1039 (1.23)	95 (0.11)
Weekday	Weekday	8758	6759 (0.77)	745 (0.09)
	Weekend	3182	3181 (1.00)	358 (0.11)

Table 2: Tornado hazard characteristics with summary statistics used in the model (2005-2014). Values in parentheses denote average fatalities and injuries per tornado event.

preferable for fatalities while the negative binomial is optimum for injuries. However, recent studies by Elsner et al. (2018); Masoomi & van de Lindt (2018) argued that negative binomial regression is more appropriate than Poisson regression when modeling count variables such as tornado related injuries and fatalities that exhibit overdispersion [i.e. the condition where variance is greater than the mean (Long & Freese, 2001)]. If the dispersion parameter (α) is zero (or closer to zero), then the negative binomial distribution approaches the Poisson distribution. Consistent with the earlier studies, the dispersion parameter is found to be both significant and greater than zero (> 0.6) indicating that both injuries and fatalities exhibit overdisper-

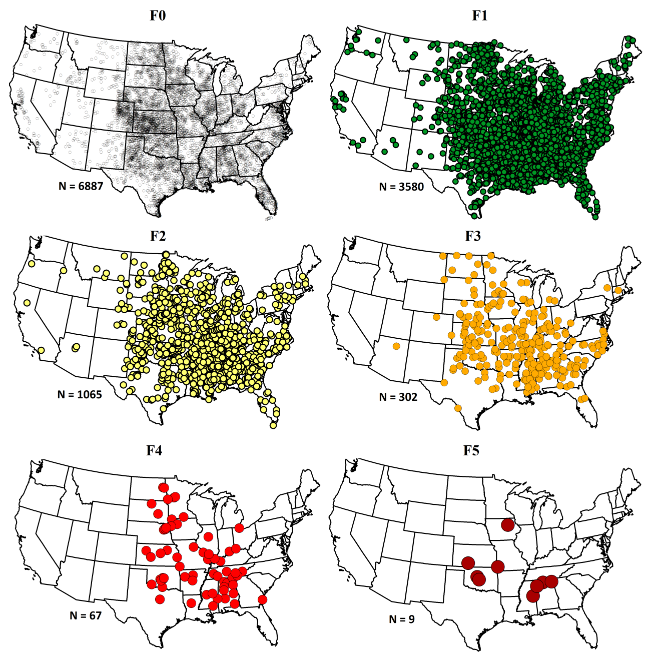

Figure 4: Tornado distribution map by magnitude (2005-2014)

sion. To test for multicollinearity, we created a correlation matrix with all independent variables and excluded any combinations of variables with high correlations ($r^2 > 0.7$) (such as length in combination with width or area) in any model. Moreover, the variable inflation factor (VIF) for all of the variables was less than 1.5, so we expect that multicollinearity is not a factor and will have no negative impact on the model and statistical outcomes. Therefore, negative binomial regression was used in this study to predict the independent variables. The negative binomial regression model for an i^{th} observation can be expressed as:

$$f(Y = y_i|\mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(y_i + 1)}\Big(\frac{1}{1 + \alpha\mu_i}\Big)^{\alpha^{-1}}\Big(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\Big)^{y_i} \quad (4)$$

$$\mu_i = exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} .... + \beta_n x_{ni}) \tag{5}$$

Where y is the dependent variable (fatality or injury); x is the explanatory variable; n refers to the number of explanatory variables in the model, whereas μ is the link parameter and β are the unknown regressor coefficients. β_0 is the intercept.

The Akaike information criteria (AIC) and McFadden's pseudo R^2 were used as metrics to quantify model performances. Models with lower AIC values are preferred. McFadden's pseudo R^2 values range from 0-1, where 0 refers to no model fit, and 1 represents the best fit model. McFadden's pseudo R^2 are not 1:1 mapped to typical coefficient of determination (R^2). In fact, a model with pseudo R^2 in the range of 0.2-0.4 represents an excellent fit (McFadden, 1977).

A framework of statistical models allowed for a comparative assessment of the predictive capability of models that included various combinations of tornado hazard characteristics, exposure, and underlying social vulnerability:

- H0: hazard-only variables (i.e., those corresponding to physical attributes of the tornado)
- H1: H0 with exposure data (total population exposed to the tornado's path)
- H2: H0 with vulnerability data (weighted average SVI*)
- H3: a combination of H0, H1 & H2 (hazard, exposure and vulnerability)

Each set of tests considered both dependent variables (i.e. fatality and injury). The independent variables representing hazard included: the tornado damage rating as indicated by F-scale (prior to 2007) and EF-scale since 2007; tornado length; and Boolean variables to represent weekdays vs weekend, seasonality, and nocturnal vs daytime. The natural log of variables with large magnitudes (such as population and tornado length) was used. The vulnerability variable considered was the spatially weighted social vulnerability index produced in this study SVI^*_{wt}.

For this study, data from all years except 2011 were used to develop the model whereas the year 2011 was used to validate the model performance. The main motivation of selecting 2011 as the validation year was since it was an anomalous year in terms of tornado event frequency and damages in the US, with some extreme tornado events that caused major devastation. This will be critical to test the robustness of the model, especially considering that the model was developed using relatively normal years. Each of these models were tested in multiple combinations of variables, and the top five models for each case with the lowest AIC and highest pseudo R^2 were identified. Since the differences in pseudo R^2 and AIC among the top five set of variables were insignificant, a model consistently present in all the cases was selected to represent all hypothesis. The use of a single set of variables across all models allows for a consistent basis to compare the results. Our analysis revealed that the variable set consisting of tornado length; nocturnality, seasonality; weekend and magnitude consistently appeared in top five sets, although some of the variables may exhibit non-significant contributions in more than one model.

Considering the tornado events of all magnitudes simultaneously, the violent tornado events tend to be underrepresented in the model due to their relatively lower frequency. Therefore, rare but extreme events may not be predicted effectively. To adequately represent and distinguish violent tornadoes from relatively moderate ones, the dataset was sub-divided into three categories: Weak (F/EF: 0 & 1), Moderate (F/EF: 2 & 3) and Violent tornadoes (F/EF: 4 & 5), and separate models were developed and evaluated for each case. Similar to Elsner et al. (2018) we only used a subset of tornado events with non-zero casualties to develop and evaluate the model.

4. Results and Discussion

Given these inputs, we developed a series of models that considered hazard, exposure, social vulnerability, or combinations thereof, and assessed their abilities to predict injuries and fatalities. To test the robustness of the results, we withheld the year 2011 from the training set, allowing for a validation and intercomparison with previous studies.

4.1. Injury Prediction

Considering all magnitude tornado events (N=778) together, the tornado hazard characteristic only model (H0) had an AIC of 4571.1 with Pseudo-R^2 of 0.08. The model was able to predict a total of 68.5% of injuries for year 2011 with 14.6 MAE and an absolute relative error of 227.5%. Combining hazard variables with population exposure (H1) reduced the AIC to 4342.6 and nearly doubled the Pseudo-R^2 to 0.15. The H1 model predicted $\sim$65% of injuries but with relatively less errors (MAE of 12.8 and 164.8%

relative error). The H2 model combined hazard variables with social vulnerability (excluding exposure), caused the AIC to jump to 4573.5 (and dropped Pseudo-R^2 to 0.08). The MAE and relative errors were similar to those of H0 model (Table 3), indicating that the inclusion of SVI^*_{wt} did not improve the model performance compared to the hazard only (H0) model. In fact, in some instances, including SVI^*_{wt} made the model slightly worse with higher relative error. The H3 model that combined hazard, exposure and vulnerability yielded the lowest AIC of all models (4337.1) and same Pseudo-R^2 as H1 (0.15). When compared to 2011 data, the MAE and relative errors of the H3 model were again lowest of all the models (12.7 and 162.9%, respectively). Therefore, although both H1 and H3 models performed very similarly, but the slightly lower AIC and absolute relative errors of the combined hazard, exposure and vulnerability (H3) model makes it a better candidate than the exposure only (H1) model.

In terms of weak tornadoes (N=300), the H0 model had an AIC of 1266.3 with Pseudo-R^2 of 0.01. Compared to 2011 data, the MAE was 2.4 with 95.2% absolute relative error. The inclusion of exposure information in H1 reduced the AIC to 1261.4 but did not improve Pseudo-R^2 or MAE values. In fact, there was a slight increase in absolute relative error (97.6% from 95.2%) compared to H0 model. Replacing exposure with SVI^*_{wt} information in H2 again resulted in a higher AIC, although there was a slight improvement in terms of MAE and relative error (2.3 and 94.3%) compared to both H0 and H1 models. The H3 model reduced AIC back to H1 levels with a slightly better Pseudo-R^2 of all other models (0.012) and produced one of the lowest MAE values for weak tornadoes.

For moderate magnitude tornadoes (N=439), the addition of exposure information in H1 reduced the AIC compared to the H0 model and nearly tripled the Pseudo-R^2 while reducing the MAE and relative error. However, with the H2 model, the AIC did decline compared to H0 but only slightly (2769.4 against 2779.1) and produced higher error statistics, indicating that the inclusion of SVI^*_{wt} made the model even worse than the hazard only model for moderate intensity tornadoes. The H3 model on the other hand, had the lowest AIC (2609.5) of all the models and highest Pseudo-R^2 value (0.13). Although the error statistics are slightly higher than the H1 model, they are significantly lower than the H0 or H2 models. Based on AIC and Pseudo-R^2 values, the H3 model would be preferred over others where the exposure information seems to offer the most significant positive impact on overall model performance. Together, exposure and SVI^*_{wt} were able to further reduce the model AIC and improve the Pseudo-R^2. This shows that both of these datasets convey more information and can improve model performance for moderate intensity tornadoes than any individual dataset can.

Considering only violent tornadoes (N=39), the inclusion of exposure in H1 made the model slightly better compared to hazard only in H0, whereby AIC decreased from 391.1 to 367.6. Furthermore, the H1 model was better able to predict injuries with a lower MAE (41.7 vs 48.2) and lower relative errors (96.0% vs 120.2%). The inclusion of SVI^*_{wt} instead of exposure in H2 seems to have negative impact compared to including exposure data. The AIC was increased to 389.6 (although still slightly lower than the H0 model) and Pseudo-R^2 was reduced to 0.08. The H2 model also saw an increase in MAE (46.5) compared to the H1 model. With the H3 model, the AIC was

369.7 (lower than H0 & H2 but higher than H1) and produced a Pseudo-R^2 of 0.18. The MAE and absolute relative errors were the lowest of all models (40.67 and 95.4%, respectively). Based on AIC and Pseudo-R^2 values, the H1 model seems more appropriate to predict violent tornado injuries; however, validation results show that for the 2011 tornadoes, model H3 was slightly better able to predict injuries than H1. Most of the errors in the H0 and H2 models are due to a single tornado event that resulted in over 300 injuries, which the model was not able to predict. The H0 model predicted only 56 injuries whereas with the inclusion of SVI^*_{wt}, the prediction was further reduced to 53 injuries. On the other hand, including exposure data improved the prediction, although still underestimating, from 50s to 186.

4.2. Fatality Prediction

Taking all fatality-causing tornado events of all magnitudes (N=181), the H0 model resulted in an AIC of 746.7 with a Pseudo-R^2 of 0.1 (Table 4). Except for weekday/weekend variable and nocturnality, all other variables were significant contributors with $P < 0.001$. When compared to the 2011 tornado-induced fatalities, the model underestimated overall fatalities by nearly 54%. The total fatalities caused in 2011 was 553, but the model predicted only 255 fatalities. The MAE of the model was 7.46, whereas the mean relative error was found to be 101.4%

By including the exposure parameter (H1 model), the AIC was reduced to 735.9 and also resulted in a slight increase in Pseudo-R^2 to 0.12. With the addition of the exposure parameter, the total count of fatalities increased to 264. The MAE and mean absolute relative errors were improved slightly with the inclusion of exposure data to 7.41 and 101.3%, respectively. However,

		H0	H1	H2	H3
All	AIC	4571.1	4342.6	4573.5	4337.1
Tornadoes	R^2	0.08	0.15	0.08	0.15
(N=778)	MAE	14.6	12.8	14.6	12.7
	Abs. Rel Error (%)	227.5	164.8	228.1	162.9
Weak	AIC	1266.3	1261.4	1267.8	1261.8
Tornadoes	R^2	0.01	0.01	0.01	0.012
(N=300)	MAE	2.4	2.4	2.3	2.3
	Abs. Rel Error (%)	95.2	97.6	94.3	95.7
Moderate.	AIC	2779.1	2619.9	2769.4	2609.5
Tornadoes	R^2	0.04	0.11	0.05	0.13
(N=439)	MAE	11.8	10.0	11.7	10.2
	Abs. Rel Error (%)	232.1	147.5	225.7	148.9
Violent	AIC	391.1	367.6	389.6	369.7
Tornadoes	R^2	0.07	0.18	0.08	0.18
(N=39)	MAE	48.2	41.7	46.5	40.7
	Abs. Rel Error (%)	120.2	96.0	117.9	95.4

Table 3: Statistical comparison of the injury prediction models. The model parameters (AIC and Pseudo-R^2) and the evaluation matrices (MAE and percent absolute relative error)

the model was still significantly underpredicting the high end of large-fatality events.

Using the underlying social vulnerability instead of the exposure (H2), caused the model AIC to increase relative to H1 to 748.3 and a reduction in Pseudo-R^2 (0.11, but still higher than H0). Furthermore, the MAE and mean absolute relative errors for the H2 model was higher than both H0 and H1 at 7.49 and 102.6%. Although error statistics increased, the total fatality prediction was slightly better at 277 fatalities for the year 2011. The results indicate that, despite improving the overall predictability, the inclusion of SVI^*_{wt} data into the model made it relatively less accurate for individual tornado events. The H3 model produced the lowest AIC (737.1) of all the models, and the total fatality prediction jumped to 283 (403 at the upper limits). The H3 model outperformed any of the individual models with the lowest MAE and mean absolute relative errors (7.23 and 95.6%, respectively). The majority of the errors were due to the underprediction of three major tornado events each causing fatalities in excess of 50 for all the models (the maximum being 158 on 22 May 2011; see table 5). Moreover, the models tend to overpredict fatalities for more common but less devastating tornado events (fatality <= 10 and lower F/EF scales). If we only consider tornado events with less than 50 fatalities (n=56), the H3 model was able to predict 244 out of 259 fatalities (~94%) with an MAE <2.5.

There were a total of 97 fatality-causing tornado events of F/EF scale 0 or 1 during the study period. The H0 model had the lowest AIC (93.9) compared to other models whereas H3 model had the highest AIC of 97.8. Pseudo-R^2 was almost negligible for all the models, ranging from 0.005 – 0.007. When compared to the 2011 tornadoes, all models performed the same with 5 out of 6 fatality predictions producing an MAE of 0.2 and absolute relative error of 10%.

Considering moderate tornado events (N=48), the H0 model had an AIC of 469.3 and predicted 102 fatalities, compared to 92 observed for the year 2011. The mean absolute difference in predicated tornado fatalities was a little over 1 per tornado event with an absolute relative error of 68.6%. The addition of exposure data (H1 model) reduced the AIC to 465.8 (from 469.3 of H0 model), and total fatalities predicated for moderate tornado was 99 (against 94 observed) with a >94% success rate. Using underlying social vulnerability as an independent variable instead of exposure (H2 model), the AIC was further reduced to 465.5; however, total the fatality prediction jumped to 104, indicating overprediction by the model compared to both H0 and H1, despite having the lowest AIC. This also resulted in the highest MAE of 1.22. In line with the aforementioned results, the H3 model produced the lowest AIC of all (461.9) and the lowest absolute relative error of 64.4%. The total fatality prediction was 101 with a Pseudo-R^2 of 0.16. The maximum overprediction was observed for tornado events with less than 4 fatalities. Tornado events with higher fatalities were predicted with relatively greater accuracy, with absolute error of 25.4% against 80.9%.

The H0 model for violent tornadoes (N=36) produced an AIC of 176.3, where only tornado length and magnitude were significant contributors ($P < 0.05$). The model was able to predict only 47.5% of the total fatalities with MAE of 16.8. With the addition of population exposure information (H1), the model AIC was reduced to 172.3, and predictive capabilities increased to 51.2% and reduced the MAE to 15.7. With H2 model, the AIC increased to 177.6, but the total fatality predictability was increased to nearly 56% with 14.7 MAE. Finally, the H3 model had the lowest AIC of 172.2 and the highest Pseudo-R^2 of 0.16 of all models. The model was able to predict more than 68.8% of the total fatality estimation (312 out of total 453) with lowest MAE of 10.9 and absolute relative error of 73.9%.

Similar to all tornado scenarios, most of the underpredictions in the prediction model for violent tornadoes is due to three extreme events with fatalities in excess of 50 (Table 5) that all models were unable to fully capture. Excluding those three events, the H0 model MAE is reduced from 16.8 to 5.4. Similarly, for the H1 and H2 models the MAE was also reduced to 5.8 and 4.6. Likewise, for the H3 model, the MAE was reduced to less than half from 10.9 to 4.5 considering only tornado events causing fewer than 50 fatalities. The results clearly indicate that the model was able to predict fatalities with reasonable success for non-anomalous tornado events. However, a limitation of the method presents itself when attempting to predict fatalities of the highest-impact events.

4.3. Fatality Intercomparisons

We used violent tornado estimates of 2011 from two earlier studies Masoomi & van de Lindt (2018) and Simmons & Sutter (2014) to intercompare the regression results. There were a total of 18 violent tornado events for 2011. The overall absolute relative error for the Simon and Sutter (SS) model was 110 whereas the Masoomi & van de Lindt (MH) model published an error of 78.2%. The results from the best performing model H3 model constrained to violent tornadoes from the current study is used for the intercomparison. Overall, the H3 model in this study was better able to predict tornado fatalities for the year 2011 with an absolute relative error of 73.9% (Figure 5). Interestingly, both the SS and MH models significantly over predicted the 27 Apr 2011 tornado in Alabama (3^{rd} in table 5) at 174 and 185, respectively, compared to the observed 72 fatalities. Conversely, the EF5 tornado in Missouri with 158 fatalities was underpredicted with only 48 (SS) and 12 (MH) fatalities, respectively. Both tornado events are similar in many aspects since

		H0	**H1**	**H2**	**H3**
All	AIC	746.7	735.9	748.3	737.1
Tornadoes	R^2	0.10	0.12	0.11	0.12
(N=181)	MAE	7.46	7.41	7.49	7.23
	Abs. Rel Error (%)	101.4	101.3	102.6	95.6
Weak	AIC	93.9	95.9	95.9	97.8
Tornadoes	R^2	0.005	0.006	0.006	0.007
(N=97)	MAE	0.2	0.2	0.2	0.2
	Abs. Rel Error (%)	10.0	10.0	10.0	10.0
Moderate.	AIC	469.3	465.8	465.5	461.9
Tornadoes	R^2	0.12	0.14	0.13	0.16
(N=48)	MAE	1.16	1.19	1.22	1.16
	Abs. Rel Error (%)	68.6	63.5	68.6	64.6
Violent	AIC	176.3	172.3	177.6	172.2
Tornadoes	R^2	0.09	0.15	0.09	0.16
(N=36)	MAE	16.8	15.7	14.7	10.9
	Abs. Rel Error (%)	102.2	100.0	93.9	73.9

Table 4: Statistical comparison of the fatality prediction models including the model parameters (AIC and Pseudo-R^2) and the evaluation measures (MAE and percent absolute relative error)

both are EF5 tornadoes with paths in the range of 122-132 miles, both occurring during the tornado season and during the day. The populations that were directly in the path of the tornado track were also similar with nearly 3700 in Alabama and ∼4500 in Missouri. The only difference was in weekend vs weekday. The Alabama tornado occurred on a weekday, whereas the Missouri tornado occurred on a weekend. The H1 model predicted only 35 and 25 tornado fatalities for the Alabama and Missouri tornadoes, respectively.

The H1 model results are in line with the SS and MH models at least in terms of trends (35 out of 72 and 14 out of 158). The current H3 model underpredicted both tornado fatalities, but the trends were in the right direction (47 out of 72 and 91 out of 158). The major distinction that the H3 model has compared to the earlier studies (or the H1 model) is the inclusion of vulnerability information. The underlying weighted social vulnerability for the Alabama tornado was 0.63, whereas the MO tornado had 0.88. These values are reflected in the fatality count of the model. Despite an underestimation, the use of social vulnerability information seems to further improve predictive capabilities of a casualty model.

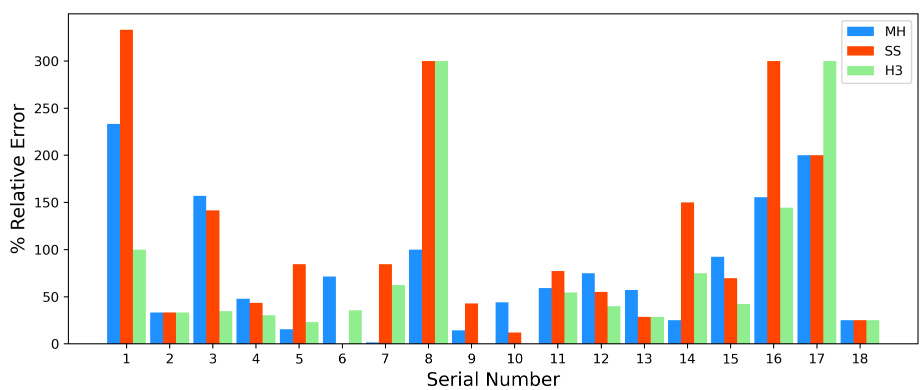

Figure 5: Percent absolute relative errors for each of the models [SS - Simmons and Sutter; MH - Masoomi and van de Lindt; H3 - current study; serial number refers to tornado events described in Table 5]

4.4. Limitations and error characterization

In this study, the tornado paths are assumed to be straight despite the fact that not all tornado paths are straight (Masoomi & van de Lindt, 2018;

S. No.	Date	State	EF	Fatality	Predicted H3	Predicted SS	Predicted MH	Abs. Error H3	Abs. Error SS	Abs. Error MH	H3 Model Info. SE	95% CI LB	95% CI UP
1	4/27/11	MS	5	3	6	13	10	3	10	7	0.7	4	7
2	4/27/11	AL	4	6	8	4	4	2	2	2	1.7	6	13
3	4/27/11	AL	5	72	47	174	185	25	102	113	14.3	14	77
4	4/27/11	MS	5	23	16	13	12	7	10	11	3.5	8	39
5	4/27/11	AL	4	13	10	2	11	23	11	2	7.2	8	24
6	4/27/11	AL	4	14	9	14	4	5	0	10	2.3	7	16
7	4/27/11	AL	4	64	24	10	63	40	54	1	10.2	13	57
8	4/27/11	AL	4	1	4	4	2	3	3	1	1.1	3	7
9	4/27/11	MS	4	7	7	4	8	0	3	1	2.5	5	15
10	4/27/11	AL	5	25	25	28	14	0	3	11	11.8	10	63
11	4/27/11	AL	4	22	10	5	9	12	17	13	3.2	8	21
12	4/27/11	GA*	4	20	12	9	5	8	11	15	7.6	9	42
13	4/27/11	AL	4	7	5	5	3	2	2	4	0.9	4	8
14	4/27/11	TN	4	4	7	10	3	3	6	1	1.8	5	12
15	5/22/11	MO	5	158	91	48	12	67	110	146	41.7	23	224
16	5/24/11	OK	5	9	22	36	23	13	27	14	5.8	10	35
17	5/24/11	OK	4	1	4	3	3	3	2	2	1.4	3	10
18	5/24/11	AR	4	4	5	5	3	1	1	1	0.9	4	8
	Mean							**10.9**	**20.8**	**19.7**			

Table 5: An intercomparison of 2011 violent tornado fatalities as predicted in this study vs. observed data and previous studies. [SS – Simon and Sutter model; MH – Masoomi and van de Lindt model; H3 - current study; * cross state Tornado passing through GA and TN; SE refers to standard error; LB and UB refers to the lower and upper bounds on the 95% confidence interval]

Fricker et al., 2017). The findings from Fricker (2020) show that the differences in actual vs assumed path only resulted in differences of 700 in the total population count. They also argued that only a handful of the tornadoes significantly deviate from the straight-line path and that such instances can be considered outliers. Furthermore, the assumption of the tornado path as a straight line is consistent with recent studies such as Fricker et al. (2017); Masoomi & van de Lindt (2018); Fricker (2020); Elsner et al. (2018).

Furthermore, the modified spatially weighted social vulnerability index SVI^*_{wt} introduced in this study attempts to separate and quantify the populations exposed to tornadoes from those that were not affected within a census tract. The geospatial subsetting approach used to determine SVI^*_{wt} is prone to the modifiable unit problem (MAUP). The limitations due to MAUP on using coarser, county-level demographic data to estimate finer-scale impacts of tornadoes as been discussed earlier by Schlossberg (2003) and Ashley et al. (2014) and partially motivated this study. To mitigate the MAUP challenges and to some extent and to leverage more detailed vulnerability information, we used fine scale census tract data in this study. However, the results still rely on the assumption of demographic homogeneity within census tracts, an assumption that we inherit from the delineation of census tracts and the creation of the SVI themselves (Flanagan et al., 2011).

Finally, for simplicity, we used same set of predictor variables for all of the models in this study, despite the fact that not all of the predicting variables were significant (p-value >0.1) for every model. Inclusion of non-significant variable(s) in the model may reduce model performance in some instances. When compared with models that only including significant variables, the difference in pseudo-R^2 was less than 0.015, and the difference in AIC was $< 1\%$ for all the models. This resulted in a MAE difference of 3% or less when compared to 2011 tornado fatality/injury data. Therefore, although in some instances non-significant variables were included, analysis show that the impact is usually modest.

5. Conclusions

This paper focused on developing a modeling framework to investigate possible relationships between casualties and the physical characteristics of tornadoes in combination with population exposure and/or social vulnerability. The developed model was evaluated against 2011 tornado season. This study presented a method to create a tornado track customized to a spatial weighting of the commonly used social vulnerability index (SVI). Although including population exposure data (H1) resulted in a significant positive impact on the model performance over the basic hazard (H0) model, the results suggest that the SVI^*_{wt} also has a positive effect on predicting tornado fatalities and injuries (H3). Still, the inclusion of SVI^*_{wt} appears to add relatively little value in comparison to exposure data. Models that include both population exposure and underlying SVI^*_{wt} seem to be better at predicting both injuries and fatalities than any other model in terms of MAE and relative errors for 2011 tornado events, despite not producing the best AIC or Pseudo-R^2 values in all instances. Furthermore, developing separate models for different intensities of tornadoes demonstrated better predictive ability, compared to the models that considered scales of tornadoes all together.

These models can inform decisions on prioritizing and targeting disaster risk reduction and resilience building initiatives. Simulation exercises could further explore the operationalization of such modeling frameworks and how they might improve response and relief efforts. Such a setup could provide probabilistic casualty estimates as soon as a tornado damage swath or track (e.g., from the field, airborne, or satellite-based), is available. This type of information could complement field reports during the crisis response phase to help efficient and optimal allocation of relief resources.

The predictive capability of the model can be further improved with better information of the casualty count, accurate tornado path, etc. such as one being provided by the NWS Damage Assessment Toolkit's maps of sub-tornado track level. Further studies could expand the time period selected and test against high casualty causing events. Inclusion of more tornado events of varying magnitudes and characteristics would make the model more robust, especially to extreme events such as those in 2011. Advances in both remote sensing and socioeconomic studies will likely improve estimates of tornado damage areas and the underlying vulnerability of exposed populations, both of which will lead to greater understanding of tornado impacts. We recognize that these models may be sensitive to changes in the populations and social vulnerability. Therefore a thorough analysis of factors involved in the estimation of SVI tailored towards a disaster specific vulnerability index may be needed.

6. Acknowledgement

Funding for this work was provided through a subgrant to the University of Alabama in Huntsville (UAH) from the Alabama Emergency Management

Agency (AEMA), via the Federal Emergency Management Agency (FEMA) Hazard Mitigation Assistance (HMA) Pre-disaster Mitigation (PDM) Planning Grant, PDM-PL-04-AL-2010-002. Support for this work was also provided by the NASA Applied Sciences Capacity Building Program through NASA Cooperative Agreement 80MSFC22M001with UAH.

References

Alexander, D. E. (2002). *Principles of emergency planning and management.* Oxford University Press on Demand.

Amendola, A., Linnerooth-Bayer, J., Okada, N., & Shi, P. (2008). Towards integrated disaster risk management: case studies and trends from asia. *Natural Hazards*, 44(2), 163–168.

Anderson, W. A. (2005). Bringing children into focus on the social science disaster research agenda. *International Journal of Mass Emergencies and Disasters*, 23(3), 159.

Ash, K. D. (2017). A qualitative study of mobile home resident perspectives on tornadoes and tornado protective actions in South Carolina, USA. *GeoJournal*, 82(3), 533–552.

Ashley, W. S. (2007). Spatial and Temporal Analysis of Tornado Fatalities in the United States: 1880–2005. *Weather and Forecasting*, 22(6), 1214–1228.

Ashley, W. S., Krmenec, A. J., & Schwantes, R. (2008). Vulnerability due to Nocturnal Tornadoes. *Weather and Forecasting*, 23(5), 795–807.

Ashley, W. S., Strader, S., 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, Climate, and Society*, 6, 175–193.

Ashley, W. S. & Strader, S. M. (2016). Recipe for disaster: How the dynamic ingredients of risk and exposure are changing the tornado disaster landscape. *Bulletin of the American Meteorological Society*, 97(5), 767 – 786.

Bluestein, H. (1999). Tornado Alley: Monster Storms of the Great Plains.

Bolin, B. (2006). Race, Class, Ethnicity, and Disaster Vulnerability. In *Handbook of Disaster Research* (pp. 113–129). New York, NY: Springer New York.

Borden, K. A., Schmidtlein, M. C., Emrich, C. T., Piegorsch, W. W., & Cutter, S. L. (2007). Vulnerability of U.S. Cities to Environmental Hazards. *Journal of Homeland Security and Emergency Management*, 4(2).

Boruff, B. J. & Cutter, S. L. (2007). The Environmental Vulnerability of Caribbean Island Nations*. *Geographical Review*, 97(1), 24–45.

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. *Climate Research*, 24, 103–117.

Brooks, H. E. & Doswell, C. A. (2001). Normalized Damage from Major Tornadoes in the United States: 1890–1999.

Brooks, H. E., Lee, J. W., & Craven, J. P. (2003). The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmospheric Research*, 67-68, 73–94.

Carter, A. O., Millson, M. E., & Allen, D. E. (1989). Epidemiologic Study of Deaths and Injuries Due to Tornadoes. *Am. J. Epidemiol.*, 130(6), 1209–1218.

Cutter, S. L. (2006). Moral Hazard, Social Catastrophe: The Changing Face of Vulnerability along the Hurricane Coasts. *The ANNALS of the American Academy of Political and Social Science*, 604(1), 102–112.

Cutter, S. L., Ash, K. D., & Emrich, C. T. (2014). The geographies of community disaster resilience. *Global Environmental Change*, 29, 65–77.

Cutter, S. L., Carolina, S., Boruff, B. J., & Shirley, W. L. (2003). Social Vulnerability to Environmental Hazards*. *Social Science Quarterly*, 84(2), 242–261.

Cutter, S. L. & Finch, C. (2008). Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the National Academy of Sciences of the United States of America*, 105(7), 2301–6.

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

Diaz, J. & Joseph, M. B. (2019). Predicting property damage from tornadoes with zero-inflated neural networks. *Weather and Climate Extremes*, 25(25), 100216.

Dixon, R. W. & Moore, T. W. (2012). Tornado Vulnerability in Texas. *Weather, Climate, and Society*, 4(1), 59–68.

Donner, W. R. (2007). The Political Ecology of Disaster: An Analysis of Factors Influencing U . S . Tornado Fatalities And Injuries, 1998-2000*. *Demography*, 44(3), 669–685.

Doswell, C. a. & Burgess, D. W. (1988). On Some Issues of United States Tornado Climatology. *Monthly Weather Review*, 116(2), 495–501.

Edwards, R., Brooks, H. E., & Cohn, H. (2021). Changes in tornado climatology accompanying the enhanced fujita scale. *Journal of Applied Meteorology and Climatology*, 60(10), 1465 – 1482.

Eidson, M., Lybarger, J. A., Parsons, J. E., MacCormack, J. N., & Freeman, J. I. (1990). Risk Factors for Tornado Injuries. *International Journal of Epidemiology*, 19(4), 1051–1056.

Elliott, J. R. & Pais, J. (2006). Race, class, and Hurricane Katrina: Social differences in human responses to disaster. *Social Science Research*, 35(2), 295–321.

Elsner, J. B., Fricker, T., & Berry, W. D. (2018). A Model for U.S. Tornado Casualties Involving Interaction between Damage Path Estimates of Population Density and Energy Dissipation. *Journal of Applied Meteorology and Climatology*, 57(9), 2035–2046.

Elsner, J. B., Jagger, T. H., & Fricker, T. (2016). Statistical Models for Tornado Climatology: Long and Short-Term Views. *PLOS ONE*, 11(11), e0166895.

Feuerstein, B., Dotzek, N., & Grieser, J. (2005). Assessing a Tornado Climatology from Global Tornado Intensity Distributions. *Journal of Climate*, 18(1), 585–596.

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

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

Fothergill, A. & Peek, L. A. (2004). Poverty and Disasters in the United States: A Review of Recent Sociological Findings. *Natural Hazards*, 32(1), 89–110. Publisher: Kluwer Academic Publishers.

Fricker, T. (2020). Tornado-level estimates of socioeconomic and demographic variables. *Natural Hazards Review*, 3(21), 1–11.

Fricker, T., Elsner, J. B., Mesev, V., & Jagger, T. H. (2017). A dasymetric method to spatially apportion tornado casualty counts. *Geomatics, Natural Hazards and Risk*, 8(2), 1768–1782.

Gall, M., Borden, K. A., & Cutter, S. L. (2009). When Do Losses Count? *Bulletin of the American Meteorological Society*, 90(6), 799–809.

Gensini, V. A. & Brooks, H. E. (2018). Spatial trends in United States tornado frequency. *npj Climate and Atmospheric Science*, 1(1), 38.

Gottlieb, R. J., Brooks, H. E., Shafer, M. A., Richman, M. B., et al. (2008). Uusing gis to assess social vulnerability to climate hazards in the southern united states.

Grazulis, T. P. (2001). *The Tornado: Nature's Ultimate Windstorm.* University of Oklahoma Press.

Grieser, J. & Haines, P. (2020). Tornado Risk Climatology in Europe. *Atmosphere*, 11(7), 768.

Kar, N. (2009). Psychological impact of disasters on children: review of assessment and interventions. *World Journal of Pediatrics*, 5(1), 5–11. Publisher: Springer-Verlag.

Kilijanek, T. S. & Drabek, T. E. (1979). Assessing Long-term Impacts of a Natural Disaster: A Focus on the Elderly. *The Gerontologist*, 19(6), 555–566.

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

Long, J. A., Stoy, P. C., & Gerken, T. (2018). Tornado seasonality in the southeastern United States. *Weather and Climate Extremes*, 20(20), 81–91.

Long, J. S. & Freese, J. (2001). Predicted probabilities for count models. *The Stata Journal*, 1(1), 51–57.

Marin, G., Modica, M., Paleari, S., & Zoboli, R. (2021). Assessing disaster risk by integrating natural and socio-economic dimensions: A decision-support tool. *Socio-Economic Planning Sciences*, 77, 101032.

Masoomi, H. & van de Lindt, J. W. (2018). Fatality and Injury Prediction Model for Tornadoes. *Natural Hazards Review*, 19(3), 04018009.

McFadden, D. (1977). *Quantitative methods for analysing travel behaviour of individuals: some recent developments*. Routledge.

Ngo, E. B. (2001). When Disasters and Age Collide: Reviewing Vulnerability of the Elderly. *Natural Hazard Review*, 2(2), 80–89. Publisher: American Society of Civil Engineers.

Nouri, N., Devineni, N., Were, V., & Khanbilvardi, R. (2021). Explaining the trends and variability in the United States tornado records using climate teleconnections and shifts in observational practices. *Scientific Reports*, 11(1), 1741.

Paul, B. K. (2003). Public response to tornado warnings: A comparative study of the May 4, 2003, tornados in Kansas, Missouri, and Tennessee.

Paul, B. K. & Stimers, M. (2014). Spatial Analyses of the 2011 Joplin Tornado Mortality: Deaths by Interpolated Damage Zones and Location of Victims. *Weather, Climate, and Society*, 6(2), 161–174.

Petit, F. D., Eaton, L. K., Fisher, R. E., McAraw, S. F., & Collins III, M. J. (2012). Developing an index to assess the resilience of critical infrastructure. *International Journal of Risk Assessment and Management*, 16(1-3), 28–47.

Platt, R. H. (1999). *Disasters and Democracy: The Politics Of Extreme Natural Events*. Island Press.

Potvin, C. K., Broyles, C., Skinner, P. S., Brooks, H. E., & Rasmussen, E. (2019). A Bayesian Hierarchical Modeling Framework for Correcting Reporting Bias in the U.S. Tornado Database. *Weather and Forecasting*, 34(1), 15–30.

Refan, M., Romanic, D., Parvu, D., & Michel, G. (2020). Tornado loss model of Oklahoma and Kansas, United States, based on the historical tornado data and Monte Carlo simulation. *International Journal of Disaster Risk Reduction*, 43, 101369.

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

Schlossberg, M. (2003). Gis, the us census and neighbourhood scale analysis. *Planning, Practice & Research*, 18(2-3), 213.

Schmidlin, T. W. & King, P. S. (1995). Risk Factors for Death in the 27 March 1994 Georgia and Alabama Tornadoes. *Disasters*, 19(2), 170–177.

Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., & Cutter, S. L. (2008). A sensitivity analysis of the social vulnerability index. *Risk analysis : an official publication of the Society for Risk Analysis*, 28(4), 1099–114.

Schroder, Z. & Elsner, J. B. (2021). Estimating 'outbreak'-level tornado counts and casualties from environmental variables. *Weather, Climate, and Society*.

Shah, M. A. R., Renaud, F. G., Anderson, C. C., Wild, A., Domeneghetti, A., Polderman, A., Votsis, A., Pulvirenti, B., Basu, B., Thomson, C., Panga, D., Pouta, E., Toth, E., Pilla, F., Sahani, J., Ommer, J., El Zohbi, J., Munro, K., Stefanopoulou, M., Loupis, M., Pangas, N., Kumar, P., Debele, S., Preuschmann, S., & Zixuan, W. (2020). A review of hydro-meteorological hazard, vulnerability, and risk assessment frameworks and indicators in the context of nature-based solutions. *International Journal of Disaster Risk Reduction*, 50, 101728.

Shen, G. & Hwang, S. N. (2015). A spatial risk analysis of tornado-induced human injuries and fatalities in the USA. *Natural Hazards*, 77(2), 1223–1242.

Simmons, K. & Sutter, D. (2012). The 2011 Tornado Season in Historical Perspective. In *Deadly Season* chapter 1, (pp. 1–15).

Simmons, K. & Sutter, D. (2014). Fatality prediction for the 2011 tornado season based on historical extreme weather data. *Natural Hazards Review*, 3(15), 1–7.

Simmons, K. M. & Sutter, D. (2005). Protection from Nature's Fury: Analysis of Fatalities and Injuries from F5 Tornadoes. *Natural Hazards Review*, 6(2), 82–87.

Simmons, K. M. & Sutter, D. (2008). Tornado Warnings, Lead Times, and

Tornado Casualties: An Empirical Investigation. *Weather and Forecasting*, 23(2), 246–258.

Simmons, K. M. & Sutter, D. (2009). False Alarms, Tornado Warnings, and Tornado Casualties. *Weather, Climate, and Society*, 1(1), 38–53.

Simmons, K. M. & Sutter, D. (2011). *Economic and Societal Impacts of Tornadoes*. Boston, MA: American Meteorological Society.

Strader, S. M. & Ashley, W. S. (2018). Finescale Assessment of Mobile Home Tornado Vulnerability in the Central and Southeast United States. *Weather, Climate, and Society*, 10(4), 797–812.

Strader, S. M., Haberlie, A. M., & Loitz, A. G. (2021). Assessment of nws county warning area tornado risk, exposure, and vulnerability. *Weather, Climate, and Society*, 13(2), 189 – 209.

Sutter, D. & Simmons, K. M. (2010). Tornado fatalities and mobile homes in the united states. *Natural Hazards*, 53, 125 – 137.

Verbout, S. M., Brooks, H. E., Leslie, L. M., & Schultz, D. M. (2006). Evolution of the U.S. Tornado Database: 1954–2003. *Weather and Forecasting*, 21, 86–93.

Wood, N. J. & Good, J. W. (2004). Vulnerability of Port and Harbor Communities to Earthquake and Tsunami Hazards: The Use of GIS in Community Hazard Planning. *Coastal Management*, 32(3), 243–269.

Wu, S., Yarnal, B., & Fisher, A. (2002). Vulnerability of coastal communities to sea-level rise: a case study of Cape May County, New Jersey, USA. *Climate Research*, 22(3), 255–270.

Wurman, J., Robinson, P., Alexander, C., & Richardson, Y. (2007). Low-Level Winds in Tornadoes and Potential Catastrophic Tornado Impacts in Urban Areas. *Bulletin of the American Meteorological Society*, 88(1), 31–46.

Zahran, S., Tavani, D., & Weiler, S. (2013). Daily variation in natural disaster casualties: Information flows, safety, and opportunity costs in tornado versus hurricane strikes. *Risk Analysis*, 33(7), 1265–1280.