A Model for U.S. Tornado Casualties Involving Interaction between Damage Path Estimates of Population Density and Energy Dissipation

JAMES B. ELSNER, TYLER FRICKER, AND WILLIAM D. BERRY

Florida State University, Tallahassee, Florida

(Manuscript received 13 April 2018, in final form 2 July 2018)

ABSTRACT

A recent study showed the importance of tornado energy as a factor in a model for tornado deaths and injuries (casualties). The model was additive under the assumption of uniform threat. Here, we test two explicit hypotheses designed to examine this additive assumption. The first hypothesis concerns energy dissipation's effect conditional on population density and the second concerns population's effect conditional on energy. Both hypotheses are tested using a regression model that contains the product of population density and energy dissipation. Results show that the elasticity of casualties with respect to energy dissipation increases with population density. That is, the percentage increase in casualties with increasing energy dissipation *increases* with population density. Similarly, the elasticity of casualties with respect to population density increases with energy dissipation. That is, the percentage increase in casualties with increasing population density *increases* with energy dissipation. Allowing energy and population elasticities to be conditional rather than constant provides a more complete description of how tornado casualties are influenced by these two important factors.

1. Introduction

Nearly one-fifth of all natural-hazard fatalities in the United States are the direct result of tornadoes (National Oceanic and Atmospheric Administration 2015). A tornado that hits a city is capable of inflicting hundreds or thousands of casualties (death or injury). Data from the Storm Prediction Center of the National Oceanic and Atmospheric Administration show that the 27 April 2011 Sawyerville–Eoline, Alabama, tornado produced 1564 casualties, with 64 of them resulting in death. Less than a month later, the 22 May 2011 Joplin, Missouri, tornado produced 1308 casualties, with 158 of them resulting in death. More recently, the 26 December 2015 Garland–Rowlett, Texas, tornado resulted in 10 deaths and 468 injuries.

Casualties occur when a tornado strikes people at home, work, school, and so on. Extreme wind speeds, relatively short warning lead times, and low-quality built environments increase the potential for casualties (Greenough et al. 2001). Risk factors for a casualty include taking shelter in a manufactured home, being an older adult, and inaccessibility to a safe room in a basement or in a reinforced structure (Centers for Disease Control and Prevention 2012). Early warnings and strong shelters are the most effective way to reduce deaths and injuries (Bohonos and Hogan 1999). The potential for casualties is not uniform across the country, however (Ashley and Strader 2016). Southern states are particularly at risk because of the abundance of manufactured homes and the frequency of reduced-visibility conditions (trees, hills, and heavy rain).

Our goal is to better understand how factors related to tornado exposure and ferocity combine to determine the number of casualties. The objective is a statistical model that describes the rate of casualties given known tornado path characteristics and the estimated number of people within the path. Our exclusive focus is on casualty-producing tornadoes. Given that a tornado produces a casualty, how do population density and energy dissipation combine to explain the number of casualties?

2. Hypotheses

Fundamentally, it is clear that a tornado's *casualty rate* (how many people get injured or killed) depends on the *density of population* in the tornado's path and the tornado's *strength* (i.e., the power of the winds that it generates, or, using more technical language, the amount of

Corresponding author: James B. Elsner, jelsner@fsu.edu

DOI: 10.1175/JAMC-D-18-0106.1

atmospheric energy that it dissipates).[1] Fricker et al. (2017) estimate the elasticity of tornado casualties with respect to each of population density and energy dissipation. They use these estimates to calculate that, when the population density within its path is held constant, a doubling of a tornado's energy increases casualties, on average, by 33% (±3%) and, when the energy dissipation of a tornado is held constant, a doubling of population density along its path increases casualties by an average of 21% (±3%).

However, the statistical analysis yielding these findings is based on a model that assumes the effect of a tornado's strength on its casualty rate is the same regardless of the population density along its path. Yet, it seems more reasonable to expect that the effect of tornado strength on the casualty rate is conditional on population density. When the population density along a tornado's path is zero, by definition, the casualty rate must be zero—since regardless of the strength of the tornado, there are no people present to be injured or killed. This guarantees that tornado energy has no effect on the casualty rate when population density is zero. In contrast, when population density along a tornado's path is very high, tornado energy should have a strong effect on the casualty rate—with low-energy storms uniformly producing a small number of casualties and high-energy storms often resulting in a large number of casualties. Indeed, as the population density in the path of a tornado rises from zero, the effect of tornado energy on casualties should gradually grow in magnitude. We can formalize this reasoning about the way in which the effect of tornado strength on the casualty rate is conditional on population density in the following two-part testable hypothesis—the ''energy's effect conditional on population'' hypothesis: *1) As long as the population density along a tornado's path is greater than zero, the effect of the tornado's energy on its casualty rate is positive regardless of the density of population in its path. However, 2) the effect of tornado energy on the casualty rate increases in magnitude as population density rises.*

If the effect of tornado strength on the casualty rate is conditional on population density, the symmetry of conditionality guarantees that the effect of population density on the casualty rate is conditional on tornado strength (Berry et al. 2012). However, the guaranteed conditionality does not establish the sign of the effect of population density at any level of tornado energy. Additional reasoning is required. Since even the weakest of tornadoes is strong enough to kill and injure, at any tornado energy level, an increase in the population density along the tornado's path puts more people ''in harm's way'' and should result in an increase in the casualty rate. As a tornado's energy rises, the potential of the storm to inflict casualties on the population in its path should increase, and so the magnitude of the positive effect of population density on the casualty rate should increase.

The expected positive effect of population density on the casualty rate is not due solely to the increase in the number of people in harm's way. Low population density is most likely to be found in rural areas, and high density is most likely to be present in urban areas; for several reasons, an urban area tends to be more prone to tornado casualties than a rural area that is impacted by a tornado of identical strength. Donner (2007) argues that the types of relationships that rural communities create, as well as the awareness that people in such regions hold of their habitats, might help to protect otherwise vulnerable populations. He maintains that strong social bonds expand the scope and availability of potential sources of weather information, access to shelters, and other resources that help with protective action. He also speculates that urban areas are more vulnerable to casualties because the rare occurrence of tornadoes in urban communities leads to lack of preparedness among residents, along with skepticism about the risks. Moreover, the presence of tall buildings and other objects may make it difficult to confirm visually the presence of tornadoes in urban communities. Without an environmental cue, warning responses tend to be lower (Mileti and O'Brien 1992).

We can formalize our reasoning about the way in which the effect of population density on the casualty rate is conditional on a tornado's strength in the following two-part testable hypothesis—the ''population's effect conditional on energy'' hypothesis: *1) The effect of population density along the path of a tornado on the tornado's casualty rate is positive regardless of the tornado's energy. However, 2) the effect of population density increases in magnitude as tornado energy rises.*

3. Model

We test our two hypotheses by modifying a model introduced by Fricker et al. (2017). They estimate a negative binomial regression with the tornado as the unit of analysis and the event count, that is, the number of casualties C resulting from the tornado, as dependent variable:

$$C \sim \mathrm{NegBin}(\hat{\mu}, n) \quad \text{and} \tag{1a}$$

[1] We treat as synonymous the terms tornado *strength*, tornado *energy*, and energy *dissipation*.

$$\ln(\mu) = \ln(\beta_0) + \beta_P \ln(P) + \beta_E \ln(E). \qquad (1b)$$

The casualty rate C is assumed to be adequately described by a negative binomial (NegBin) distribution with a rate parameter μ and a size parameter n. The variable P denotes population density, and E denotes energy dissipation. The coefficient β_P for $\ln(P)$ represents the elasticity of the tornado casualty rate with respect to population density (hereinafter *population elasticity*) and represents the percentage change in the casualty rate that is associated with a 1% increase in population density. Similarly, β_E represents the elasticity of the tornado casualty rate with respect to energy dissipation (hereinafter *energy elasticity*) and reflects the percentage change in the casualty rate that is associated with a 1% increase in tornado energy. The functional form of Eq. (1b) imposes the assumption that energy elasticity is the same regardless of the value of population density and that population elasticity is the same regardless of the value of energy dissipation. These assumptions are directly contrary to our two conditional hypotheses.

However, we can adapt the Fricker et al. (2017) model to be compatible with our conditional hypotheses by adding a multiplicative term, $\ln(P) \times \ln(E)$, to Eq. (1b) to yield the model

$$C \sim \text{NegBin}(\hat{\mu}, n) \qquad (2a)$$

$$\ln(\mu) = \ln(\beta_0) + \beta_P \ln(P) + \beta_E \ln(E) + \beta_{P-E}[\ln(P) \times \ln(E)]. \qquad (2b)$$

To reveal how this revised model makes energy elasticity conditional on population density, we fix P in Eq. (2b) at some constant P^*:

$$\ln(\mu) = \ln(\beta_0) + \beta_P \ln(P^*) + \beta_E \ln(E) + \beta_{P-E}[\ln(P^*) \times \ln(E)]. \qquad (3)$$

Rearranging terms yields

$$\ln(\mu) = [\ln(\beta_0) + \beta_P \ln(P^*)] + [\beta_E + \beta_{P-E} \ln(P^*)] \times \ln(E). \qquad (4)$$

In this equation, each of $[\ln(\beta_0) + \beta_P \ln(P^*)]$ and $[\beta_E + \beta_{P-E} \ln(P^*)]$ is a constant, and therefore $\ln(\mu)$ is expressed as a linear function of $\ln(E)$. Accordingly, the coefficient for $\ln(E)$, $[\beta_E + \beta_{P-E} \ln(P^*)]$, is the elasticity of the tornado casualty rate with respect to energy dissipation when population density P is equal to P^*. Thus, it is clear that Eq. (2b) specifies energy elasticity as conditional on the value of population density; more specifically, energy elasticity is a linear function of the logarithm of population density, that is, $\ln(P)$. Similarly, we can fix E in Eq. (2b) at the constant E^* and manipulate terms to show that $[\beta_P + \beta_{P-E} \ln(P^*)]$ is the elasticity of the tornado casualty rate with respect to population density when energy dissipation E is equal to E^*—implying that population elasticity is conditional on the value of energy dissipation.

Because Eq. (2b) models the elasticity of the tornado casualty rate with respect to each of energy dissipation and population density as conditional on the value of the other, we can describe the equation as assuming that energy dissipation and population density *interact* in influencing tornado casualties. This is in contrast to Eq. (1b), which assumes that energy dissipation and population density are *additive* in their influence on tornado casualties. The coefficients β_P, β_E, and β_{P-E} from Eq. (2b) can be used to calculate the elasticity of the tornado casualty rate with respect to each of energy dissipation and population density at any specified value of the other variable, thereby permitting an empirical test of each of our two hypotheses. Moreover, the coefficient β_{P-E} in Eq. (2b) measures the magnitude of the interaction between population density and energy dissipation, that is, the magnitude of the relationship between 1) population elasticity and the log of energy dissipation and 2) energy elasticity and the logarithm of population density.

Our overall modeling approach is similar to that of recent work that examines factors related to tornado casualties (Lim et al. 2017; Zahran et al. 2013; Simmons and Sutter 2008; Donner 2007). However, it differs in several key ways. First, we use a continuous measure of a tornado's energy dissipation—rather than path area, enhanced Fujita-scale (EF) rating, or total damage—as the indicator of tornado strength. Correlation among these variables is high, but energy dissipation has a better physical justification as it is wind power that is proximal to the cause of casualties. Second, we exclude tornadoes without casualties, focusing on factors influencing the casualty rate among those tornadoes producing at least one casualty. This makes our models simpler, and the interpretation of the results more straightforward, since for a variety of reasons most tornadoes do not result in casualties.[2] Third, our study focuses on the interaction between the environment (tornado energy) and a demographic factor (population density) in explaining tornado casualties. A focus on

[2] The question of *whether* a tornado produces casualties is interesting but has been addressed elsewhere (e.g., Ashley et al. 2008).

the statistical *interaction* between environmental and demographic factors is unique in research about tornado casualties and provides a new lens through which to view factors that influence the rate of injuries and death from tornadoes.

In brief, the results from this study offer support for both our conditional hypotheses. We find that the elasticity of the tornado casualty rate with respect to energy dissipation is positive and statistically significant across most of the range of population density values, and that this elasticity increases with population density. We also find that the elasticity of the tornado casualty rate with respect to population density is positive and statistically significant across most of the range of energy dissipation values, and that this elasticity increases with tornado energy. Supplementary analysis shows that if our results are biased as a result of error in measuring population density, this bias may result in an *underestimation* of the extent of interaction between energy dissipation and population density. The paper continues in section 4 with a description of the indicators that we use to measure tornado casualties, tornado energy, and population density. Results are presented in section 5. A summary of what was done and prospects for future work are given in section 6.

4. Data and variables

a. Tornado casualty rates

Interest centers on the count of injuries and fatalities listed in the Storm Prediction Center's tornado database. The database is compiled from the National Weather Service's (NWS) *Storm Data* and includes all known tornadoes dating back to 1950. A tornado casualty is defined as a fatality or injury directly attributable to a tornado or to impact by airborne, falling, or moving debris. An example of a direct fatality would be a driver killed when a motor vehicle is tossed over. To be considered a direct injury, the injury must require treatment by a first responder or at a medical facility. An example would be a person treated for a laceration caused by flying debris. Indirect casualties are not considered. An example of an indirect casualty would be electrocution during debris removal.

NWS *Storm Data* also includes for each tornado the straight-line track, start location, damage path dimensions of length and maximum width, and the maximum EF rating (wind speed damage rating on an ordered scale of categories from 0 to 5). Reports in the database are compiled by the NWS offices and reviewed by the National Climatic Data Center (Verbout et al. 2006). We consider only tornadoes occurring within the conterminous United States.

During the 22-yr period 1995–2016, there are 26 863 tornadoes recorded in the conterminous United States. Of these, 2208 are linked to 25 968 casualties. Only 6.7% of all casualties lead to death. Most casualty-producing tornadoes result in just a few casualties. Indeed, even among casualty-producing tornadoes, the median number of casualties is 3.

A relatively small number of casualty-producing tornadoes result in many casualties. The Tuscaloosa–Birmingham, Alabama, tornado of 27 April 2011 and the Joplin tornado of 22 May 2011 top the list of the most casualties in the period. The Tuscaloosa–Birmingham tornado was one of 362 tornadoes occurring between 25 and 28 April in the largest outbreak in recorded U.S. history. The Joplin tornado is the deadliest in the modern record-keeping era (since 1950).

During the period 1995–2016, tornado casualties were most common across the southeastern quarter of the country (Fig. 1). Alabama had the most (3937), followed by Oklahoma (2500), Missouri (2025), Arkansas (1759), Texas (1704), and Tennessee (1678). Rounding out the top 10 are Georgia (1610), Mississippi (1400), Kentucky (1051), and North Carolina (985). Outside the South, Massachusetts (ranked 22nd) stands out in New England with 232 casualties from only three tornadoes.

The number of casualties is positively related to a tornado's EF (wind speed) rating (Table 1). Among casualty-producing tornadoes, on average an EF0 tornado results in two casualties, an EF1 tornado results in 3.1 casualties, and an EF2 tornado results in 6.5 casualties. Beyond EF2, expected casualties increase dramatically. On average, an EF3 tornado results in 20 casualties, an EF4 tornado results in 62 casualties, and an EF5 tornado results in 236 casualties.

b. Population density

We create a damage path from the tornado track using the width variable and then estimate the total number of people within the path boundary. Population data are obtained from the Gridded Population of the World, version 4 (GPWv4), from the Socioeconomic Data and Applications Center at Columbia University. The database contains decennial census density estimates for 1990, 2000, and 2010 represented as residential population per square kilometer. The native cell resolution is 0.0083° latitude/longitude, which at 36°N latitude means a cell having the dimension of 0.9 km in the north–south direction and 0.7 km in the east–west

FIG. 1. Location of casualty-producing tornadoes (1995–2016) by EF rating. The size of the circle is proportional to the number of casualties.

direction. For each tornado we assign the population density corresponding to the closest decennial estimate so that, for example, a tornado that occurred in 1996 uses the 2000 decennial estimate whereas a tornado that occurred in 2012 uses the 2010 estimate.[3]

For the set of 2192 tornadoes with at least one casualty[4] the median population density per tornado is 31.9 people per km^2, with an interquartile range between 9.88 and 137 people per km^2. There are seven orders of magnitude separating the lowest and highest

[3] Linear interpolation of the population to the tornado year had no significant influence on the model results.

[4] Sixteen of the 2208 casualty-producing tornadoes occurred over areas without population and are removed from further analysis.

TABLE 1. Casualties by EF rating. Data period: 1995–2016. Here, avg indicates the average casualties per tornado.

EF	Casualties	Tornadoes	Fatalities	Injuries	Avg
0	362	183	8	354	2.0
1	2300	750	79	2221	3.1
2	4888	753	218	4670	6.5
3	7933	392	536	7397	20.2
4	7187	116	457	6730	62.0
5	3298	14	432	2866	235.6

TABLE 2. Descriptive statistics for the variables used to estimate the model of Eqs. (2a) and (2b) observed over the 2192 tornadoes in our sample. Data period: 1995–2016.

Variable	Value
No. of tornadoes	2192
Avg no. of casualties per tornado	12
Median no. of casualties per tornado	3
Highest no. of casualties	1564
Lowest no. of casualties	1
Std dev of casualties	53
Avg population density per tornado (people per km^2)	223
Greatest population density (people per km^2)	13 949
Least population density (people per km^2)	0.003
Std dev of population density (people per km^2)	708
Avg energy dissipation per tornado (GW)	855
Max energy dissipation (GW)	66 200
Min energy dissipation (GW)	0.0057
Std dev of energy dissipation (GW)	2900

per-tornado population densities (Table 2; Fig. 2). We estimate that the EF2 Brooklyn, New York, tornado of 8 August 2007 that injured nine had the highest population density in its path at 13 949 people per km^2. Of the 10 casualty-producing tornadoes with the highest population densities, 3 occurred in New York, 2 occurred in Florida, and 2 occurred in California. Pennsylvania, Massachusetts, and Texas each had 1.

The distribution of population and population density in the path of tornadoes by EF rating is shown in Table 3. We estimate that about 2.5 million people have been within the path of a tornado and 1.5 million people have been within the path of a casualty-producing tornado. Because of their frequency and path area, EF1 tornadoes account for the largest accumulated number of people at risk. Given a casualty-producing tornado, however, it is the EF3 tornadoes that, as a group, affected the most people. The number of people affected is a strong function of EF rating since more-damaging tornadoes tend to have larger damage paths (Brooks 2004; Elsner et al. 2014; Ashley and Strader 2016). There is no strong relationship between median population density and EF rating. There are 30 noncasualty-producing tornadoes without an EF rating with an estimated total population of 13 people within the paths.

c. Energy dissipation as a measure of tornado strength

Tornadoes dissipate a large amount of atmospheric energy (Fricker et al. 2014, 2017). Following what was done with hurricanes (Emanuel 2005), the energy dissipation of a tornado is computed as

$$E = A_p \rho \sum_{j=0}^{5} w_j v_j^3, \qquad (5)$$

where the summation is over the six possible EF ratings (0, 1, 2, 3, 4, and 5), A_p is the area of the tornado's path (m^2), ρ is air density (1 kg m^{-3}), v_j is the midpoint wind speed (m s^{-1}) for each damage rating (EF scale) j, w_j is the corresponding fraction of path area by damage rating, and J is the maximum damage rating. Multiplying the units from the individual terms results in E being measured in a unit of power [kg m^2 s^{-3} = joules per second = watts (W)].

The fraction of path area is that recommended by the U.S. Nuclear Regulatory Commission (see Fricker and Elsner 2015), which combines a Rankine vortex with empirical estimates (Ramsdell and Rishel 2007). Threshold wind speeds for the EF ratings are a 3-s gust. With no upper bound on the EF5 wind speeds, the midpoint wind speed is set at 97 m s^{-1} (7.5 m s^{-1} above the threshold wind speed consistent with the EF4 midpoint speed relative to its threshold). Additional details and justification for energy dissipation as a valid measure of tornado strength are given in Fricker et al. (2017).

For the set of 2192 tornadoes with at least one casualty, the median energy dissipation is 97.2 GW, with an interquartile range between 14 and 514 GW. The Tallulah–Yazoo City–Durant tornado (Louisiana and Mississippi) of 24 April 2010 that killed 10 and injured 146 has the highest energy dissipation at 66 200 GW. There are seven orders of magnitude separating the lowest from the highest per-tornado energy dissipation (Fig. 2).[5]

5. Results

a. Visual inspection of the data

Figure 3 plots the tornado casualty rate for each tornado path in the sample against both the storm's energy

[5] The distributions of population density and energy dissipation are approximately normal on a logarithmic scale.

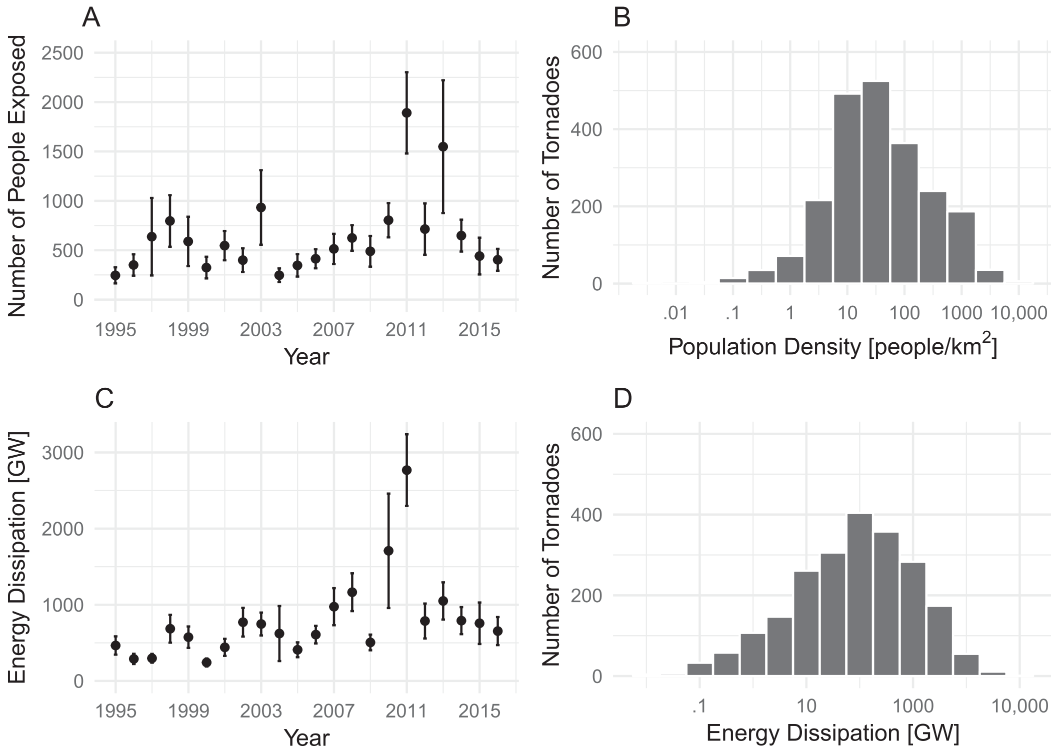

FIG. 2. People exposed and energy dissipation for all casualty-producing tornadoes (1995–2016). (a) Annual average number of people exposed per tornado, (b) distribution of per-tornado population density, (c) annual average energy dissipation per tornado, and (d) distribution of per-tornado energy dissipation. The vertical bars in (a) and (c) are ±2 standard deviations from the mean.

dissipation and the population density in its path. The correlation (Pearson) between energy dissipation and population density is −0.07. Given a tornado with at least one casualty, the number of casualties increases with the estimated number of people in the path and with energy dissipation. It is rare for a tornado with less than 0.1 GW of energy dissipation to produce more than a dozen casualties when the population density is less than 100 people per km^2. Casualty-producing tornadoes with an order of magnitude more energy dissipation tend to produce many more casualties, especially when affecting areas with population densities exceeding 10 people per km^2.

Energy dissipation depends strongly on the EF damage rating. Tornadoes that hit more targets in developed

TABLE 3. Population and population density in the path of tornadoes (casualty producing and all) by EF damage rating. Data period: 1995–2016.

EF	No. of tornadoes	Median path length (km)	Total people	Avg No. people per tornado	Median No. people per tornado	Avg population density (per km^2)	Median population density (per km^2)
				Casualty-producing tornadoes			
0	183	2	9386	51	4	440	55
1	750	5	217 053	289	17	293	37
2	753	11	397 781	528	72	186	29
3	392	24	492 471	1256	261	88	24
4	116	34	289 722	2498	824	103	27
5	14	54	89 981	6427	3318	82	25
				All tornadoes			
0	16 453	1	196 183	12	0	128	5
1	7463	4	698 008	94	5	139	14
2	2143	9	633 813	296	29	119	15
3	620	21	555 967	897	125	77	17
4	140	28	294 953	2107	550	91	18
5	14	54	89 981	6427	3318	82	25

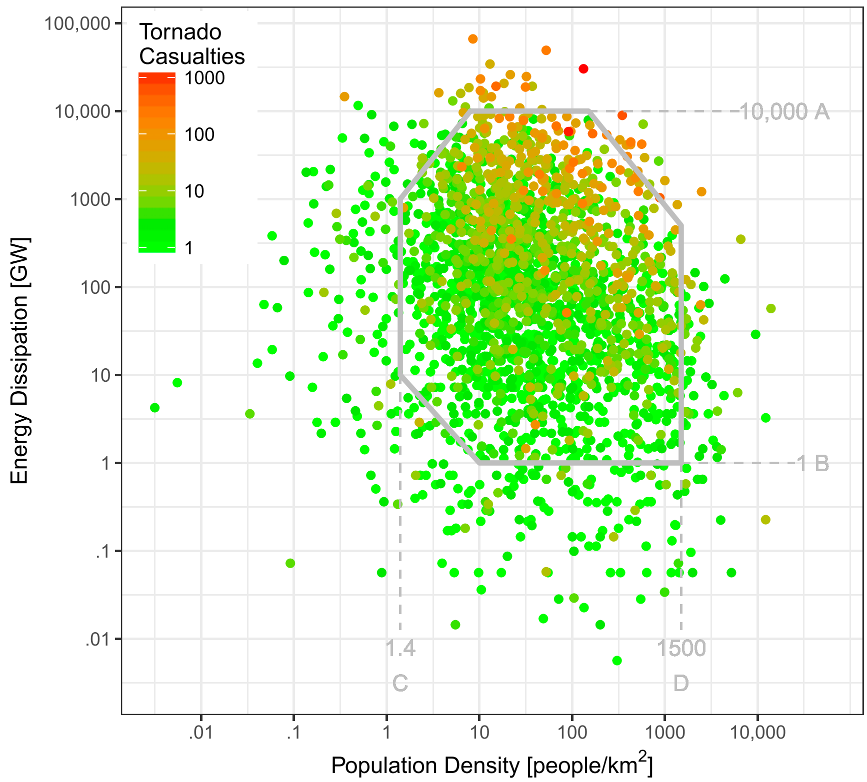

FIG. 3. Scatterplot showing casualty-producing tornadoes by population density and energy dissipation on logarithmic scales. The number of casualties per casualty-producing tornado is given in color, with the color ramp being on a logarithmic scale.

areas (built environments) have the potential to be rated higher, all else being equal, because the rating comes from the worst damage. In built environments with more targets, there are more people and therefore energy dissipation might contain some population effect. However, the correlation between energy dissipation and population density is small, indicating that this is not a problem.

b. *Model results*

We estimate the interactive negative binomial regression model of Eqs. (2a) and (2b) using our dataset of 2192 tornadoes.[6] The resulting coefficient estimates are presented in the top portion of Table 4. To understand the empirical results, we use these coefficients to estimate the 1) energy elasticity at each possible value of population density and 2) population elasticity at each possible value of energy dissipation. However, the model's estimated elasticities are more trustworthy for values of energy dissipation and population density near which there are abundant data, since estimates "far from the data" are less influenced by data and more influenced by the assumed functional form of the model (King and Zeng 2006). Thus, in all figures displaying model estimates below, we restrict attention to values for population density and energy dissipation over which the density of observations in the sample of tornadoes is relatively large.

In particular, we focus our attention on estimates over the region internal to the gray polygon superimposed over the joint distribution for population density and energy dissipation in Fig. 3. This region was identified by visual inspection of the plot as excluding pairs of values for population density and energy dissipation near which there are very few observations. The polygon—which contains 83.4% of the tornadoes in our sample—excludes all values of population density below 1.4 and above 1500 people per km^2 (a range containing 92.3% of the tornadoes in the sample) and all values of energy dissipation below 1 and above 10 000 GW (a range

[6] All empirical analysis is performed using the open-source R language for statistical computing (https://www.R-project.org) with freely available data. All figures were made using functions from the ggplot2 package (Wickham 2009). The R code and links to the data for this study are available online (https://rpubs.com/jelsner/interaction).

TABLE 4. Coefficient estimates for the interactive model of Eqs. (2a) and (2b) and the additive model of Eqs. (1a) and (1b) using data for $n = 2192$ observed tornadoes. Std error is the standard error of the estimate, and $\Pr(>t)$ is the probability of the t value more extreme under the null hypothesis that the coefficient is zero. The AIC and BIC for the interactive model are 13 223 and 13 246, respectively; those for the additive model are 13 323 and 13 340, respectively. The coefficients were determined via an iterative maximum likelihood approach with the "glm.nb" function from the Modern Applied Statistics for S (MASS) package for R (Venables and Ripley 2002; http://www.stats.ox.ac.uk/pub/MASS4).

	Estimate	Std error	t value	$\Pr(>t)$
		Interactive model		
$\hat{\beta}_0$	−3.0268	0.9606	−3.15	0.0016
$\hat{\beta}_P$	−0.8121	0.2294	−3.54	0.0004
$\hat{\beta}_E$	0.1661	0.0377	4.41	<0.0001
$\hat{\beta}_{P-E}$	0.0415	0.0092	4.51	<0.0001
		Additive model		
$\hat{\beta}_0$	−6.8825	0.5302	−13.0	<0.0001
$\hat{\beta}_P$	0.2133	0.0248	8.62	<0.0001
$\hat{\beta}_E$	0.3212	0.0196	16.4	<0.0001

containing 92.7% of tornadoes). Effectively, our choice is the minimum area encompassing at least 83% of the data values. Estimates are the same if we do not omit any data, but uncertainty is higher for estimates made outside this area.

Estimated elasticities are used to construct Fig. 4.[7] The black solid lines in Figs. 4a and 4b show, respectively, estimated energy elasticity (along with a 95% confidence band) conditional on population density and estimated population elasticity conditional on energy dissipation. Figure 4a shows, for example, that when population density is equal to 100 people per km^2 then energy elasticity is estimated to equal 0.357 (with a confidence interval extending from 0.315 to 0.400)—a value that indicates that a 1% increase in energy dissipation is associated, on average, with a 0.357% increase in the number of casualties.

The *energy's effect conditional on population* hypothesis predicts that 1) the effect of energy dissipation on the casualty rate is positive at any population density exceeding zero but that 2) the magnitude of this effect grows as population density rises. With regard to prediction 1, Fig. 4a shows that energy elasticity is positive and statistically significant at each population density value over which elasticities are displayed. If we extend the focus to the entire range of values for population density in the sample (i.e., from 0.003 to 13 949), estimated energy elasticity is positive and statistically significant at all population densities greater than 0.21 people per km^2; this is a range of values in which 99% of the tornadoes in the sample lie. In contrast, there are no values for population density at which estimated energy elasticity is negative and statistically significant.

Given the functional form of Eq. (2b), prediction 2 of the *energy's effect conditional on population* hypothesis requires that the coefficient for the multiplicative term β_{P-E} be positive, and indeed the estimate of this coefficient is positive and statistically significant (with a significance value p of less than 0.0001), implying that energy elasticity increases as population density rises. More specifically, the point estimate for β_{P-E} of 0.0415 indicates that the elasticity of the casualty rate with respect to energy dissipation increases by 0.0288 with each

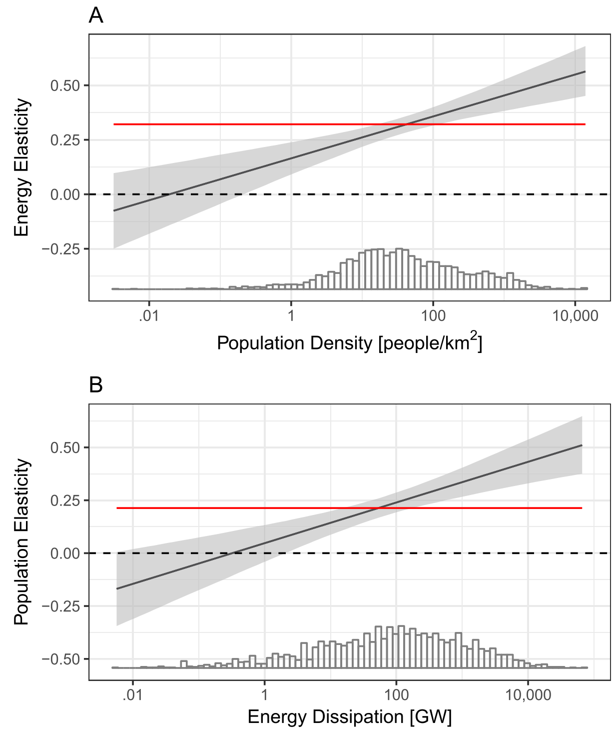

FIG. 4. Estimated relationship (a) between population density and the elasticity of tornado casualties with respect to energy dissipation and (b) between energy dissipation and the elasticity of tornado casualties with respect to population density [on the basis of coefficient estimates for the interactive model of Eqs. (2a) and (2b) found in the top portion of Table 4]. The plotted solid black lines denote point estimates of elasticities, and the gray ribbons indicate the 95% uncertainty bound around the point estimates. The red lines denote the point estimates of elasticities from the additive model of Eqs. (1a) and (1b) as presented in the bottom half of Table 4. A histogram showing the distribution of the horizontal-axis variable is displayed at the bottom of (a) and (b).

[7] Although Eq. (2b) is framed using natural logarithms, to facilitate clarity the plots in Fig. 4 and elsewhere have axes that are scaled by using (more widely familiar) common logarithms.

doubling of population density.[8] Figure 4a helps to clarify the extent to which energy elasticity varies with the value of population density. The plot shows that a shift in population density from 1.4 people per km^2 (roughly the density in Plainview, Arizona) to 1500 people per km^2 (roughly the density in Eastlake, Ohio) results in an increase in estimated energy elasticity from 0.179 (0.110, 0.247) to a value more than 2.5 times in magnitude: 0.470 (0.393, 0.549).

Figure 4b allows an evaluation of the *population's effect conditional on energy* hypothesis—which predicts that 1) the impact of population density on the casualty rate is positive at any value of tornado energy but that 2) the strength of the impact rises as energy dissipation increases. For prediction 1, Fig. 4b reveals that population elasticity is positive and statistically significant at each energy dissipation value of greater than 2.1 GW;[9] this is a range of tornado energy values in which 90.6% of the tornadoes in the sample lie. In contrast, there are no energy dissipation values at which estimated population elasticity is negative and statistically significant.

The statistically significant positive coefficient estimate for the multiplicative term $\beta_{P\text{–}E}$ is evidence that population elasticity increases as energy dissipation rises—consistent with prediction 2 of the *population's effect conditional on energy* hypothesis. Furthermore, Fig. 4b clarifies the implications of the coefficient for the multiplicative term about the degree to which population elasticity varies with the level of tornado energy. Increasing energy dissipation in the sample from 1 to 10 000 GW prompts an increase in estimated energy elasticity from 0.047 (−0.041, 0.133) to a value that is nearly 10 times its size: 0.432 (0.330, 0.537).

To better understand the extent of interaction between energy dissipation and population density in influencing the tornado casualty rate, we can use the coefficients for the model of Eqs. (2a) and (2b) to compute the estimated tornado casualty rate conditional on a tornado's values for energy dissipation and population density. Figure 5 displays the estimated relationship between energy dissipation and the casualty rate at three different values for population density: the minimum value contained within the polygon in Fig. 3 (1.4 people per km^2), the maximum value within the polygon (1500 people per km^2), and the median value across tornadoes in the sample (31.9 people per km^2).

[8] To get the value 0.0288, one multiplies the coefficient $\beta_{P\text{–}E}$ by ln(2): 0.0415 × 0.6931 = 0.0288.

[9] For comparison a 40 m s^{-1} hurricane dissipates about 12.5 GW of kinetic energy over the average lifetime of a tornado (12 min; see Emanuel 1999).

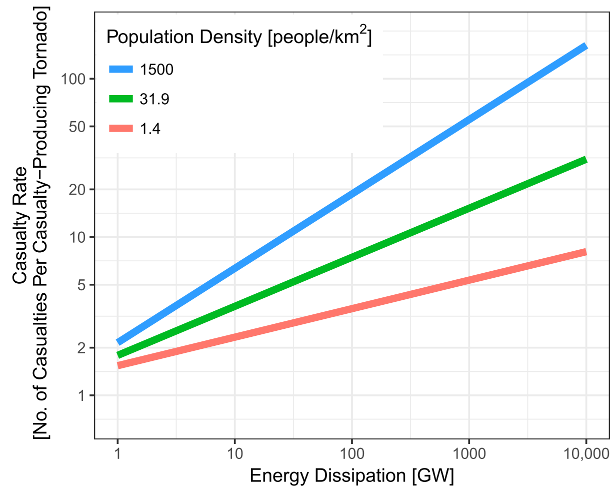

FIG. 5. Estimated effect of tornado energy dissipation on the casualty rate at various levels of population density (colors) on the basis of coefficient estimates for the interactive model of Eqs. (2a) and (2b).

As energy dissipation increases from 1 to 10 000 GW, the casualty rate is expected to increase by a difference of

1) 6.55 (from 1.54 to 8.08) when population density is at its *minimum*,
2) 29.2 (from 1.79 to 31.0) when population density is at its *median*, and
3) 160.5 (from 2.15 to 162.6) when population density is at its *maximum.*

Thus, Figs. 4 and 5 provide strong empirical evidence for both of our hypotheses: the *energy's effect conditional on population* hypothesis and the *population's effect conditional on energy* hypothesis. Specifically, the effect on a tornado's casualty rate of each of energy dissipation and population density is positive at nearly any value of the other variable, but with a magnitude that grows with the value of the other variable.

c. The consequences of shifting from an additive model to an interactive model

The conditionality in the effects of energy dissipation and population density would be undetectable if we were to analyze the data using the additive model, since the additive model assumes that the elasticity of tornado casualties with respect to each of energy dissipation and population density is the same regardless of the value of the other variable. Yet, it is interesting to assess more systematically the implications for our understanding of the consequences of tornadoes of abandoning the additive model [Eqs. (1a) and (1b)] in favor of an interactive model [Eqs. (2a) and (2b)].

Using the same data we employ to estimate the interactive model, we estimate the additive model [i.e., the model created by deleting the multiplicative term in Eq. (2b)]; the resulting coefficient estimates are presented in the bottom half of Table 4. The statistical significance of the coefficient for the multiplicative term in the interactive model constitutes initial evidence that the interactive model provides a better fit to the data than the additive model. However, it may be that the improvement in fit by specifying that population density and energy dissipation interact is too small to justify abandoning the simpler additive model in favor of the more complex interactive model. Thus, we conduct two tests suitable for comparing additive and interactive models in a way that balances model fit and model complexity by imposing a penalty for greater complexity: Akaike's information criterion (AIC) and the Bayesian information criterion (BIC). Both criteria indicate the superiority of the interactive model—with AIC decreasing from 13 323 to 13 223 and BIC decreasing from 13 340 to 13 246—when the multiplicative term is added to the model.

We can analyze the impact of shifting from an additive model to an interactive model on the magnitude of estimated elasticities. The horizontal red lines in Fig. 4 plot the estimated value of energy elasticity (Fig. 4a) and population elasticity (Fig. 4b) obtained from the additive model. The fact that each red line is horizontal is a consequence of the additive model's assumption that the elasticity of the casualty rate with respect to each of energy dissipation and population density is constant. We saw above (from Fig. 4a) that the interactive model yields the estimate that, as population density increases from 1.4 to 1500 people per km^2, energy elasticity increases from 0.179 to 0.470. This contrasts with the additive model's estimate of a constant energy elasticity of 0.321. The interactive model also estimates that as energy dissipation increases from 1 to 10 000 GW, population elasticity increases from 0.047 to 0.432, in contrast with the additive model's estimate of a constant population elasticity of 0.213.

For a final analysis of the implications of shifting from an additive model to an interactive model, we use the coefficients for each model to compute the predicted casualty rate for each tornado path in the estimation dataset given the tornado's values for energy dissipation and population density. Then, for each tornado path, we compute the absolute difference between the two models' predicted casualty rates. The median absolute difference is 0.480, and the mean is 1.37. These values may, at first glance, seem small. However, the typical tornado produces relatively few casualties; even if we exclude the majority of tornadoes that cause no casualties, the median casualty rate across all (casualty producing) tornadoes is 3. Given the median of 3, we view an absolute difference of 0.480 as consequential. For 11% of tornadoes, the absolute error exceeds 3, and for 5.5%, the absolute difference is greater than 5. Moreover, a nontrivial share of tornadoes are characterized by an absolute difference that is very high; for 1.6% of tornadoes (36), the absolute difference is greater than 10. Clearly, modifying the statistical analysis of tornadoes to allow estimated energy and population elasticities to be conditional rather than constant yields a much more nuanced, and richer, characterization of the effects of a tornado's strength and of the population in its path on the resulting number of casualties.

d. A potential bias from the population data

Ideally, our measure of population density should reflect the number of people in the path of the tornado *when the tornado strikes*. However, because we do not have access to this information, we rely on the number of people that *reside* within a tornado's path—ignoring the fact that, when the tornado strikes, some people residing in the path will be away from home, and others who live outside the path will be within the path. It seems reasonable to assume that residential population more closely approximates the number of persons in a tornado's path when the tornado strikes at night (and the vast majority of people are at home) than when it strikes during the day. If so, estimating the interactive model on the subset of nighttime tornadoes could be conceived as estimating the model using those tornadoes in our sample for which population density in the path is measured with the least error. The top half of Table 4 shows that when the model is estimated using all tornadoes the coefficient for the multiplicative term—which reflects the extent of interaction between energy dissipation and population density in influencing the tornado casualty rate—is 0.042. In contrast, when the model is estimated using *just* nighttime tornadoes (i.e., those between sunset and sunrise), the coefficient for the multiplicative term increases in magnitude to 0.049. This supplementary analysis suggests that if our main results are biased as a result of measurement error then this bias results in an *underestimation* of the extent of interaction between population density and energy dissipation on casualties.

6. Summary

A recent study showed the importance of energy dissipation on tornado casualties using an additive regression model (Fricker et al. 2017). In this paper we test two explicit hypotheses designed to examine the additive assumption implicit in this earlier study. The first hypothesis concerns energy's effect conditional on population, and the second concerns population's effect conditional on energy. The hypotheses are tested with a regression model that contains a term as the product of population density and energy dissipation.

New results here show that energy elasticity increases with population density. That is, the percentage increase in casualties with increasing energy dissipation *increases* with population density. Similarly, population elasticity increases with energy dissipation. That is, the percentage increase in casualties with increasing population density *increases* with energy dissipation. The conclusion is unambiguous in that allowing energy and population elasticities to be conditional rather than constant provides a more complete description of how casualties are influenced by these two factors as indicated by more-accurate predictions of casualty rates.

This study focused on the two dominant factors that determine the number of casualties resulting from a tornado: tornado energy and population density. It is undoubted that casualties are influenced by other variables as well. Future work will use the modeling framework described here to address the role specific socioeconomic and demographic variables have in statistically explaining tornado casualties. Toward that end it will be necessary to obtain consistent estimates of these variables at the tornado level, and the model will need to include random effects related to when the tornado occurs (such as month or time of day).

Acknowledgments. The work was partially supported by crowdfunding to the second author (TF) through https://experiment.com.

REFERENCES

Ashley, W. S., and S. M. Strader, 2016: Recipe for disaster: How the dynamic ingredients of risk and exposure are changing the tornado disaster landscape. *Bull. Amer. Meteor. Soc.*, **97**, 767–786, https://doi.org/10.1175/BAMS-D-15-00150.1.

——, A. J. Krmenec, and R. Schwantes, 2008: Vulnerability due to nocturnal tornadoes. *Wea. Forecasting*, **23**, 795–807, https://doi.org/10.1175/2008WAF2222132.1.

Berry, W. D., M. Golder, and D. Milton, 2012: Improving tests of theories positing interaction. *J. Polit.*, **74**, 653–671, https://doi.org/10.1017/S0022381612000199.

Bohonos, J. J., and D. E. Hogan, 1999: The medical impact of tornadoes in North America. *J. Emerg. Med.*, **17**, 67–73, https://doi.org/10.1016/S0736-4679(98)00125-5.

Brooks, H. E., 2004: On the relationship of tornado path length and width to intensity. *Wea. Forecasting*, **19**, 310–319, https://doi.org/10.1175/1520-0434(2004)019<0310:OTROTP>2.0.CO;2.

Centers for Disease Control and Prevention, 2012: Tornado-related fatalities—Five states, southeastern United States, April 25–28, 2011. *Morbidity and Mortality Weekly Report*, Vol. 61, No. 28, Centers for Disease Control and Prevention, Atlanta, GA, 529–533, https://www.cdc.gov/mmwr/PDF/wk/mm6128.pdf.

Donner, W. R., 2007: The political ecology of disaster: An analysis of factors influencing U.S. tornado fatalities and injuries, 1998–2000. *Demography*, **44**, 669–685, https://doi.org/10.1353/dem.2007.0024.

Elsner, J. B., T. H. Jagger, H. M. Widen, and D. R. Chavas, 2014: Daily tornado frequency distributions in the United States. *Environ. Res. Lett.*, **9**, 024018, https://doi.org/10.1088/1748-9326/9/2/024018.

Emanuel, K. A., 1999: The power of a hurricane: An example of reckless driving on the information superhighway. *Weather*, **54**, 107–108, https://doi.org/10.1002/j.1477-8696.1999.tb06435.x.

——, 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686–688, https://doi.org/10.1038/nature03906.

Fricker, T., and J. B. Elsner, 2015: Kinetic energy of tornadoes in the United States. *PLOS ONE*, **10**, e0131090, https://doi.org/10.1371/journal.pone.0131090.

——, ——, P. Camp, and T. H. Jagger, 2014: Empirical estimates of kinetic energy from some recent U.S. tornadoes. *Geophys. Res. Lett.*, **41**, 4340–4346, https://doi.org/10.1002/2014GL060441.

——, ——, and T. H. Jagger, 2017: Population and energy elasticity of tornado casualties. *Geophys. Res. Lett.*, **44**, 3941–3949, https://doi.org/10.1002/2017GL073093.

Greenough, G., M. McGeehin, S. M. Bernard, J. Trtanj, J. Riad, and D. Engelbert, 2001: The potential impacts of climate variability and change on health impacts of extreme weather events in the United States. *Environ. Health Perspect.*, **109** (Suppl.), 191–198, https://doi.org/10.2307/3435009.

King, G., and L. Zeng, 2006: The dangers of extreme counterfactuals. *Polit. Anal.*, **14**, 131–159, https://doi.org/10.1093/pan/mpj004.

Lim, J., S. Loveridge, R. Shupp, and M. Skidmore, 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.

Mileti, D. S., and P. W. O'Brien, 1992: Warnings during disasters: Normalizing communicated risk. *Soc. Probl.*, **39**, 40–57, https://doi.org/10.2307/3096912.

National Oceanic and Atmospheric Administration, 2015: Weather fatality, injury, and damage statistics. National Weather Service, http://www.nws.noaa.gov/om/hazstats/resources/weather_fatalities.pdf.

Ramsdell, J. V., Jr., and J. P. Rishel, 2007: Tornado climatology of the contiguous United States. Pacific Northwest National Laboratory Tech. Rep. NUREG/CR-4461, PNNL-15112, 260 pp., https://www.nrc.gov/docs/ML0708/ML070810400.pdf.

Simmons, K. M., and D. Sutter, 2008: Tornado warnings, lead times, and tornado casualties: An empirical investigation. *Wea. Forecasting*, **23**, 246–258, https://doi.org/10.1175/2007WAF2006027.1.

Venables, W. N., and B. D. Ripley, 2002: *Modern Applied Statistics with S.* 4th ed., Springer, 495 pp.

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

Wickham, H., 2009: *ggplot2: Elegant Graphics for Data Analysis*. Springer, 213 pp.

Zahran, S., D. Tavani, and S. Weiler, 2013: Daily variation in natural disaster casualties: Information flows, safety, and opportunity costs in tornado versus hurricane strikes. *Risk Anal.*, **33**, 1265–1280, https://doi.org/10.1111/j.1539-6924.2012.01920.x.