

ARTICLE

Received 22 Jul 2015 | Accepted 8 Jan 2016 | Published 29 Feb 2016

DOI: 10.1038/ncomms10668

OPEN

# Tornado outbreak variability follows Taylor's power law of fluctuation scaling and increases dramatically with severity

Michael K. Tippett<sup>1,2</sup> & Joel E. Cohen<sup>3,4</sup>

Tornadoes cause loss of life and damage to property each year in the United States and around the world. The largest impacts come from 'outbreaks' consisting of multiple tornadoes closely spaced in time. Here we find an upward trend in the annual mean number of tornadoes per US tornado outbreak for the period 1954–2014. Moreover, the variance of this quantity is increasing more than four times as fast as the mean. The mean and variance of the number of tornadoes per outbreak vary according to Taylor's power law of fluctuation scaling (TL), with parameters that are consistent with multiplicative growth. Tornado-related atmospheric proxies show similar power-law scaling and multiplicative growth. Path-length-integrated tornado outbreak intensity also follows TL, but with parameters consistent with sampling variability. The observed TL power-law scaling of outbreak severity means that extreme outbreaks are more frequent than would be expected if mean and variance were independent or linearly related.

<sup>1</sup>Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York 10027, USA. <sup>2</sup>Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah 21589, Saudi Arabia. <sup>3</sup>Laboratory of Populations, Rockefeller University, New York, New York 10065, USA. <sup>4</sup>The Earth Institute, Columbia University, New York, New York 10027, USA. Correspondence and requests for materials should be addressed to M.K.T. (email: mkt14@columbia.edu) or to J.E.C. (email: cohen@rockefeller.edu).

**H**azardous convective weather (tornadoes, hail and damaging wind) associated with severe thunderstorms affects large portions of the United States. Tornadoes cause particularly intense damage. Over a recent 10-year period (2005–2014), tornadoes in the United States resulted in an average of 110 deaths per year and annual losses ranging from \$500 million to \$9.6 billion<sup>1</sup>. The largest societal impacts from tornadoes are from ‘outbreaks’ in which multiple tornadoes occur in a single weather event. Tornado outbreaks across the eastern two-thirds of the United States were associated with 79% of all tornado fatalities over the period 1972–2010 (ref. 2) and are routinely responsible for billion-dollar loss events<sup>3</sup>.

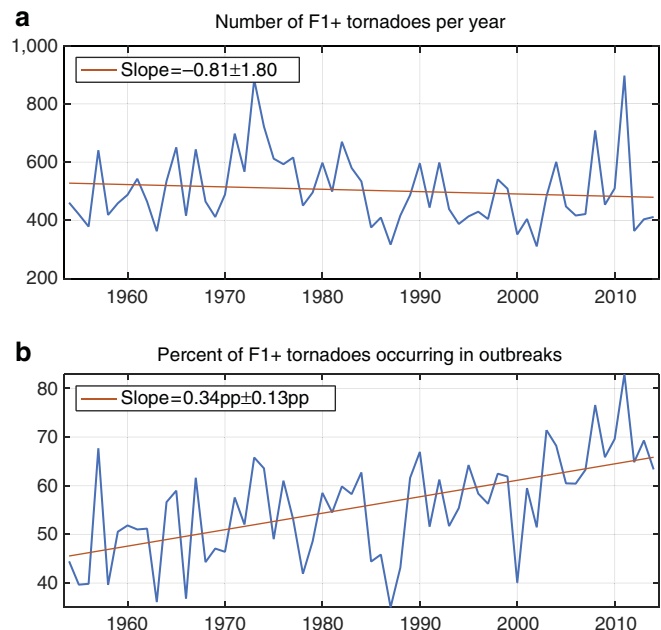
Whether US tornado activity will change in the future remains uncertain. Climate change projections indicate that environments favourable to severe thunderstorms will be more frequent in a warmer climate<sup>4</sup>, and high-resolution convection-permitting numerical modelling indicates increased activity and year-to-year variability of March–May US tornado occurrence as measured by severe weather proxies derived from explicitly depicted storms<sup>5</sup>. However, to date, there is no upward trend in the number of reliably reported US tornadoes per year<sup>6</sup>. Interpretation of the US tornado report data requires some caution. For instance, the total number of US tornadoes reported each year has increased dramatically over the last half century, but most of that increase is due to more reports of weak tornadoes and is believed to reflect changing reporting practices and other non-meteorological factors rather than increased tornado occurrence<sup>7</sup>. The variability of reported tornado occurrence has increased over the last few decades with more tornadoes being reported on days when tornadoes are observed<sup>8,9</sup>. In addition, greater year-to-year variability in the number of tornadoes reported per year has been associated with consistent changes in the monthly averaged atmospheric environments favourable to tornado occurrence<sup>10</sup>. Likewise, US large-event severe thunderstorm losses and the frequency of the most extreme environments have increased<sup>11</sup>. Regional changes in the seasonality of tornado occurrence have also been reported<sup>12,13</sup>.

Despite their importance, relatively little is known about the current or projected statistics of tornado outbreak severity beyond their climatology<sup>2</sup>. Most studies have considered only the statistics of tornadoes occurring during a single day<sup>14</sup> and have not considered outbreaks over multiple dates. Here we show that the annual mean number of tornadoes per outbreak increased during 1954–2014, and the annual variance increased more than four times faster than the mean. We show that the mean and variance of the number of tornadoes per outbreak are related by Taylor’s power law of fluctuation scaling (TL)<sup>15,16</sup> with parameters that are consistent with multiplicative growth<sup>17</sup>. TL scaling in tornado outbreak statistics was not previously known. Although power-law scaling is present in the probability distributions of various tornado characteristics<sup>18,19</sup>, the presence of power-law scaling in such probability distributions is neither necessary nor sufficient for TL scaling of the variance in relation to the mean (see also Supplementary Discussion for examples of TL scaling without power-law probability distributions). We also find that a tornado-related atmospheric proxy shows a similar power-law scaling and multiplicative growth. Path-length-integrated tornado outbreak intensity follows TL as well, but with parameters predicted by sampling variability<sup>20</sup>. The findings are similar when we restrict the analysis to the more recent period 1977–2014 and to more intense tornadoes. The observed TL scaling of outbreak severity means that extreme outbreaks are more frequent than would otherwise be expected if mean and variance were independent or linearly related.

## Results

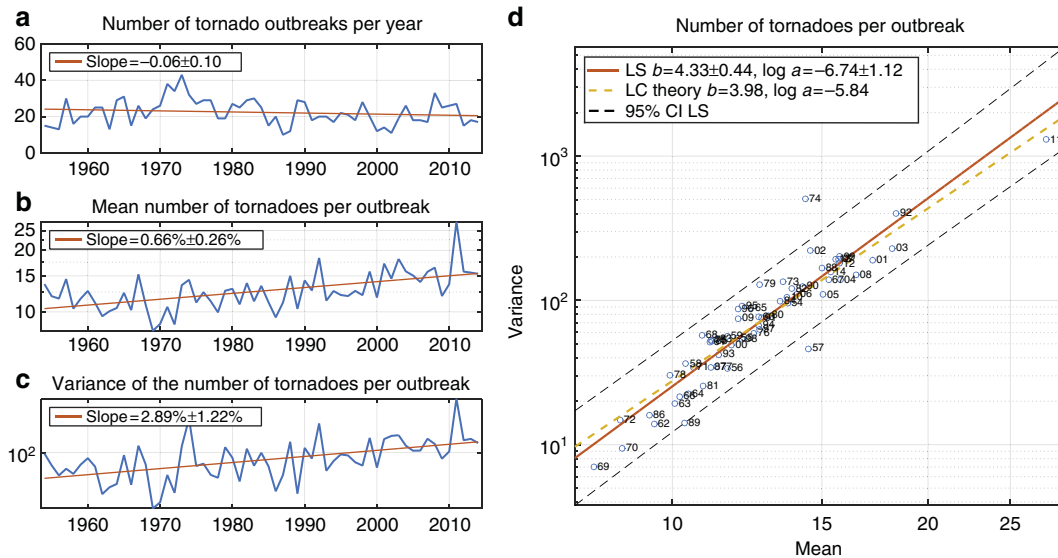
**Data and sensitivity analyses.** We use data from 1954 to 2014, which are generally considered reliable. Because of concerns regarding the data before 1977 (‘Methods’ section, Supplementary Fig. 1), we repeat some analysis using the more recent period 1977–2014 to test the robustness of the results (Supplementary Figs 2–5). We exclude the weakest tornadoes from our analysis and denote the remaining tornadoes as F1+ tornadoes (‘Methods’ section). We repeat some of the analysis restricted to more intense tornadoes (F2+; Supplementary Figs 7–10).

**Number of tornadoes per outbreak.** The annual number of F1+ tornadoes shows no significant trend over the period 1954–2014 (Fig. 1a). Generally, trends have not been found in the number of severe tornadoes when severity is defined using the Fujita scale, but upward trends have been found when severity is defined using path length<sup>18</sup>. The percentage of F1+ tornadoes that occur in outbreaks (‘Methods’ section) is increasing by 0.34 percentage points  $\pm$  0.13 percentage points per year (Fig. 1b), consistent with upward trends in the proportion of tornadoes occurring on days with many tornadoes<sup>8,9</sup>. Here and in all results,  $\pm$  intervals are 95% confidence intervals. The fraction of F1+ tornadoes that occur in outbreaks is less than one because not all F1+ tornadoes occur in outbreaks. During the period 1977–2014, the number of F1+ tornadoes also shows no significant trend, and the percentage of F1+ tornadoes occurring in outbreaks is also increasing, at a larger estimated rate (Supplementary Fig. 2). The US tornado reports show no statistically significant trend in the frequency of tornado outbreaks (Fig. 2a). Since the number of



**Figure 1 | Time series of counts and clustering of F1+ tornadoes**

**1954–2014 in the contiguous US. (a)** Number of F1+ tornadoes per year. The slope of the least-squares regression indicates that the number of F1+ tornadoes per year declined by 0.81 per year on average from 1954 to 2014 inclusive. This rate of decline is not statistically significantly different from 0 (no change). **(b)** Annual percentage of F1+ tornadoes occurring in outbreaks. The slope of the least-squares regression indicates that the percentage of F1+ tornadoes per year that occurred as part of outbreaks increased by 0.34 percentage points (pp) per year on average from 1954 to 2014 inclusive. This increase is statistically significantly greater than 0. In both **a** and **b**,  $\pm$  intervals are 95% confidence intervals.



**Figure 2 | Numbers of F1+ tornadoes per outbreak 1954-2014.** (a) Number of tornado outbreaks per year. The rate of decline is not statistically significantly different from 0 (no change). (b) Annual mean number of tornadoes per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (c) Annual variance of the number of tornadoes per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (d) Scatter plot of the annual mean number of tornadoes per outbreak versus the annual variance of the number of tornadoes per outbreak. Both axes are on logarithmic scale. The solid red line is the least-squares (LS) regression line (Taylor’s power law of fluctuation scaling) and the dashed yellow line has the slope and intercept predicted by LC theory<sup>17</sup>. The two-digit number following the plotting symbol ‘o’ gives the calendar year in the second half of the twentieth century or first half of the twenty-first century. In all the panels, ± intervals are 95% confidence intervals.

F1+ tornadoes and the number of outbreaks are not changing (on average, over time), the increasing percentage of F1+ tornadoes occurring in outbreaks means that the number of F1+ tornadoes per outbreak must be increasing, and indeed, the annual mean number of F1+ tornadoes per outbreak shows a significant upward trend (Fig. 2b). The annual mean number of tornadoes per outbreak is increasing by 0.66% ± 0.26% per year, and the variance is increasing more than four times as fast, 2.89% ± 1.22% per year (Fig. 2b,c) over the period 1954–2014. The growth rates are greater over the recent period 1977–2014, with similar ratio between the growth rates of mean and variance (Supplementary Fig. 3b,c).

The fact that the variance is increasing several times faster than the mean is especially noteworthy: it indicates a changing distribution in which the likelihood of extreme outbreaks is increasing faster than what the trend in mean alone would suggest. The coefficient of dispersion of a probability distribution with a positive mean is the ratio of its variance to its mean. Values greater than one (over-dispersion) indicate more clustering than a Poisson variable. For instance, European windstorms exhibit over-dispersion and serial clustering that increases with intensity<sup>21</sup> with implications for the return intervals of rare events<sup>22</sup>. Taylor’s law (TL) relates the mean and variance of a probability distribution by

$$\text{variance} = a(\text{mean})^b, a > 0, \tag{1}$$

where  $a$  and  $b$  are constants<sup>15,16</sup>. A value of  $b > 1$  indicates that the coefficient of dispersion increases with the mean. The annual mean and annual variance of the number of tornadoes per outbreak approximately satisfy TL with  $b = 4.3 \pm 0.44$  and  $\log a = -6.74 \pm 1.12$  (Fig. 2d); consistent values are seen over the period 1977–2014 (Supplementary Fig. 3d). (Throughout log is the natural logarithm.) The value of  $b$  here is remarkable since in most ecological applications, the TL exponent seldom exceeds 2. The TL exponent can be greater than 2 for lognormal

distributions with changing parameters (Supplementary Discussion and Supplementary Fig. 11). The TL scaling of tornado outbreak severity reveals a remarkably regular relation between annual mean and annual variance that extends over the full range of the data, even for years like 2011 which are extreme in mean and variance. The data from 1974 deviate most from TL scaling, with the excessive variance reflecting the 3–4 April ‘Super Outbreak.’

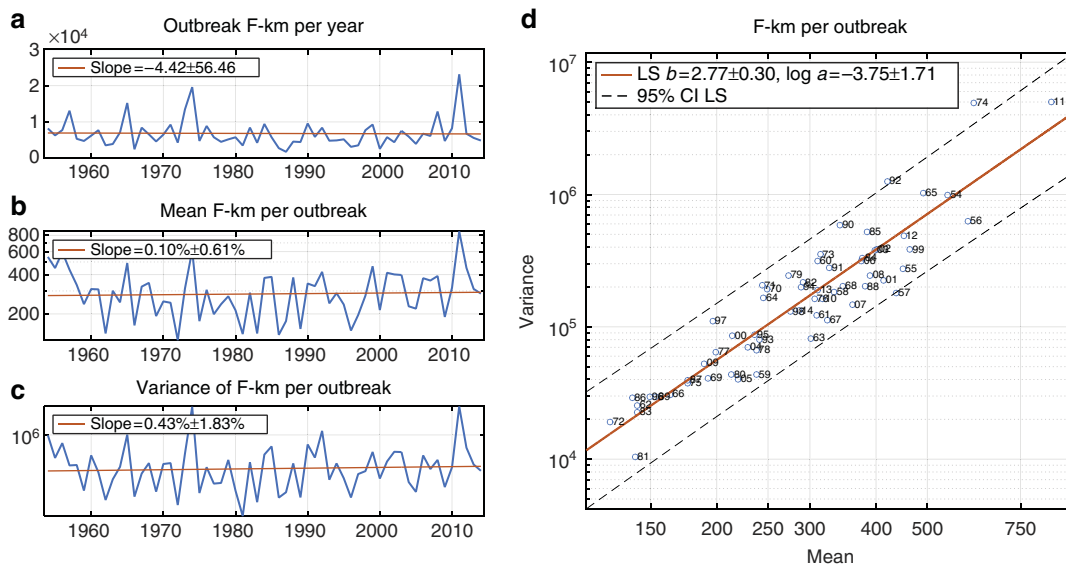
The upward trend in the number of tornadoes per outbreak provides an interpretation for the observed TL scaling since TL scaling arises in models of stochastic multiplicative growth<sup>17</sup>. In such models, the quantity  $N(t+1)$  at time  $t+1$  is related to its previous value  $N(t)$  by

$$N(t+1) = A(t)N(t), \tag{2}$$

where  $A(t)$  is the random multiplicative factor by which  $N(t)$  grows or declines from one time to the next. Here  $N(t)$  is the annual average number of tornadoes per outbreak, and each integer value of  $t$  represents one calendar year. The Lewontin–Cohen (LC) model for stochastic multiplicative growth assumes that the  $A(t)$  are independently and identically distributed for all  $t \geq 0$  with finite mean  $M > 0$  and finite variance  $V$ . If  $M \neq 1$ ,  $N(t)$  follows TL asymptotically with<sup>17</sup>

$$a = \frac{E[(N(0))^2]}{[E(N(0))]^2} \text{ and } b = \frac{\log[V + M^2]}{\log M}. \tag{3}$$

Here we estimate (‘Methods’ section)  $M = 1.03$  and  $V = 0.068$ , which leads to TL parameters  $b = 3.98$  and  $\log(a) = -5.84$ . Both values are consistent with the least-squares (LS) estimates of the corresponding parameters of TL (Fig. 2d). The LS estimates are also consistent with the values from LC theory during 1977–2014 (Supplementary Fig. 3d). 95% confidence intervals for  $M$  and  $V$  show that the hypothesis of no growth ( $M = 1$ ) under which equation (3) is not valid cannot be rejected (Supplementary Table 1). The Supplementary Discussion provides additional



**Figure 3 | F-km per outbreak 1954–2014.** (a) Total outbreak F-km per year. The rate of decline is not statistically significantly different from 0 (no change) among F1+ tornadoes. (b) Annual mean F-km per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is not statistically significantly greater than 0. (c) Annual variance of F-km per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is not statistically significantly greater than 0. (d) Scatter plot of the annual mean of F-km per outbreak versus the annual variance of F-km per outbreak. Both axes are on logarithmic scale. The solid red line is the least-squares (LS) regression line (Taylor's power law of fluctuation scaling). The two-digit number following the plotting symbol 'o' gives the calendar year in the second half of the twentieth century or first half of the twenty-first century. In all the panels,  $\pm$  intervals are 95% confidence intervals.

description of how the LC model leads to TL scaling with exponent approximately 4.

**Fujita-kilometers per outbreak.** Another measure of outbreak severity is Fujita-kilometers (ref. 2; F-km) which is the sum (over all tornadoes in an outbreak) of each tornado's path length in kilometers multiplied by its Fujita or Enhanced Fujita rating ('Methods' section). Annual totals of outbreak F-km, mean number of F-km per outbreak and the variance of F-km per outbreak do not show significant trends over the period 1954–2014 (Fig. 3a–c). The mean number of F-km per outbreak and the variance of F-km per outbreak show marginally significant trends over the recent period 1977–2014 (Supplementary Fig. 4a–c). The TL parameters relating the mean and variance of F-km per outbreak are  $b = 2.77 \pm 0.30$  and  $\log a = -3.75 \pm 1.71$  (Fig. 3d). The lack of robust trends means that LC theory is not appropriate to explain the TL scaling of F-km. However, TL scaling also arises from the sampling of stationary skewed distributions<sup>20</sup>. For a distribution with mean  $m$ , variance  $v$ , skewness  $\gamma_1$  and coefficient of variation  $CV$ , theory<sup>20</sup> predicts

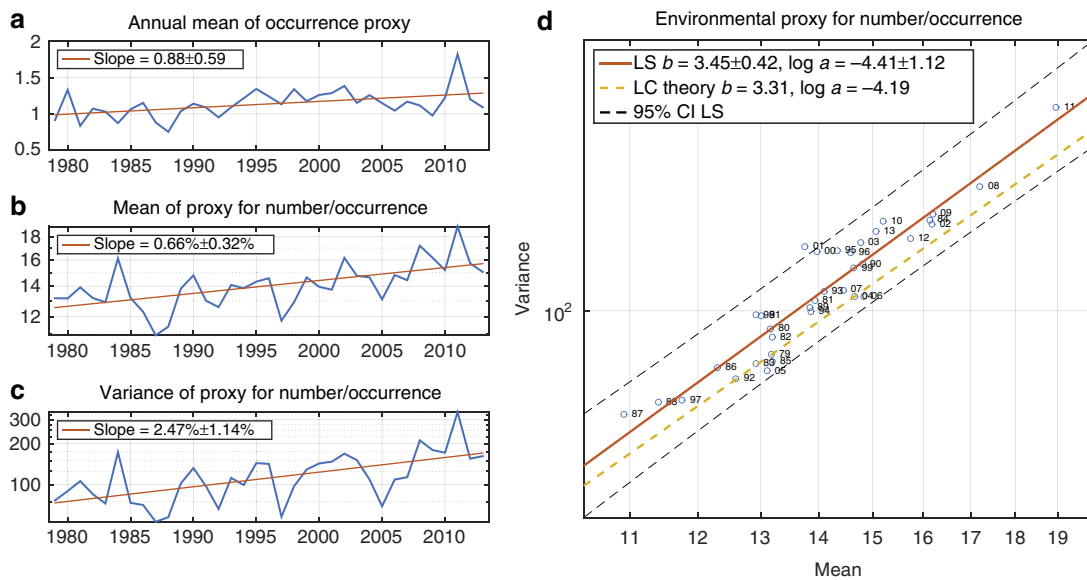
$$a = \log v - \frac{\gamma_1}{CV} \log m \text{ and } b = \frac{\gamma_1}{CV}. \quad (4)$$

Here, excluding two outlier outbreaks from the calculation of the distribution parameters ('Methods' section, Supplementary Fig. 5), equation (4) gives  $b = 2.71$  and  $\log a = -3.05$ , both of which are consistent with the LS estimates of the TL parameters for F-km per outbreak. (We use 'outlier' to indicate values far from other observations, not to suggest that the unusual values are the result of measurement error.) Therefore TL scaling of F-km per outbreak could be explained by sampling variability.

**A tornado environment proxy.** A reasonable concern is that the findings here represent properties of the tornado report database

that are not meteorological in origin, especially since other prominent features of the tornado report database are not meteorological in origin<sup>7</sup>. Environmental proxies for tornado occurrence and number of tornadoes per occurrence provide an independent, albeit imperfect, measure of tornado activity for the period 1979–2013 ('Methods' section). At a minimum, the environmental proxies provide information about the frequency and severity of environments favourable to tornado occurrence. The correlation between the annual average number of tornadoes per outbreak and the proxy for number of tornadoes per occurrence is 0.56 (Supplementary Fig. 6a). This correlation falls to 0.34, still significant at the 95% level, when the data from 2011 are excluded. Applying a 5-year moving average to the data highlights their common trends and increases the correlation to 0.88 (Supplementary Fig. 6b). The annual mean of the occurrence proxy, a surrogate for number of tornadoes per year, shows a marginally significant upward trend (Fig. 4a). The annual mean and annual variance of the proxy for number of tornadoes per occurrence show upward trends of  $0.63 \pm 0.30\%$  and  $2.43 \pm 1.12\%$ , respectively (Fig. 4b,c), values strikingly similar to those for number of tornadoes per outbreak (Fig. 2b,c). Moreover, the TL parameters of the proxy for number of tornadoes per occurrence are  $3.54 \pm 0.42$  and  $\log a = -5.87 \pm 1.43$ , which are consistent with the LC multiplicative growth theory estimates for the proxy (Fig. 4d) and quite similar to those for the number of tornadoes per outbreak. Extreme environments associated with tornado occurrence display the TL scaling and multiplicative growth similar to those of the number of tornadoes per outbreak. This similarity plausibly suggests that the changes in the number of tornadoes per outbreak reflect changes in the physical environment.

**Sensitivity to outbreak definition.** Another concern is that the results are sensitive to the details of the outbreak definition. We



**Figure 4 | Environmental proxies 1979–2013.** (a) Annual mean of occurrence proxy in per cent. The rate of increase is statistically significantly greater than 0. (b) Annual mean of the environmental proxy for number of tornadoes per occurrence. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (c) Annual variance of the environmental proxy for number of tornadoes per occurrence. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (d) Scatter plot of the annual mean of proxy for number of tornadoes per occurrence versus the annual variance of proxy for number of tornadoes per occurrence. Both axes are on logarithmic scale. The solid red line is the least-squares (LS) regression line and the dashed yellow line has slope and intercept predicted by LC theory<sup>17</sup>. The two-digit number following the plotting symbol 'o' gives the calendar year in the second half of the twentieth century or first half of the twenty-first century. In all the panels,  $\pm$  intervals are 95% confidence intervals.

assess the robustness of the results to the E/F1 threshold by repeating the analysis with tornadoes rated E/F2 and higher, denoted F2+ (Supplementary Figs 7–10). We use the period 1977–2014 because the annual number of F2+ tornadoes display a substantial decrease (not shown) around the 1970s that is likely related to the introduction of the F-scale. Overall the F2+ results are remarkably similar to the F1+ ones. The annual number of F2+ tornadoes has an insignificantly negative trend during 1977–2014 (Supplementary Fig. 7a), and the percentage of F2+ tornadoes occurring in F2+ outbreaks has a significant positive trend (Supplementary Fig. 7b). Although the number of F2+ outbreaks shows no significant trend, the mean number of tornadoes per F2+ outbreak and its variance both have significant upward trends (Supplementary Fig. 8a–c). The TL exponent for number of tornadoes per F2+ outbreak is 3.65 and is consistent with LC theory (Supplementary Fig. 8d). Annual totals of F2+ outbreak F-km have no significant trend (Supplementary Fig. 9a). Mean F-km per F2+ outbreak does have a significant upward trend, but variance does not (Supplementary Fig. 9b,c). The TL scaling of F2+ outbreak F-km (Supplementary Fig. 9d) is consistent with sampling variability when 2011 is excluded (Supplementary Fig. 10).

## Discussion

These findings have important implications for tornado risk in the United States and perhaps elsewhere, though we have examined only the US data. First, the number of tornadoes per outbreak is increasing. However, there is less evidence that F-km per outbreak are increasing. Both the number of tornadoes per outbreak and F-km per outbreak follow TL, which relates mean and variance. We find that TL scaling for the number of tornadoes per outbreak is compatible with multiplicative growth, and that TL scaling for F-km per outbreak could be due to

sampling variability. Finally, the key implication of TL scaling is that both number of tornadoes per outbreak and F-km per outbreak exhibit extreme over-dispersion which increases with mean. When the average tornado outbreak severity gets worse, the high extreme of severity rises even faster and the low extreme falls even faster, by either measure of severity.

## Methods

**Outbreak data.** Tornado reports come from the NOAA Storm Prediction Center (<http://www.spc.noaa.gov/wcm/#data>). There are no reports from either Alaska or Hawaii; one report from Puerto Rico is excluded. Tornadoes are rated by estimated or reported damage using the Fujita (F) scale, introduced in the mid-1970s, and since March 2007, the enhanced Fujita (EF) scale, with 0 being the weakest and 5 the strongest. Only reports of tornadoes rated F/EF1 or greater are used and are denoted as F1+. To compute outbreak statistics, tornado reports are first sorted in chronological order taking into account the time zone. Outbreaks are sequences of 6 or more F1+ tornadoes (regardless of location in the contiguous United States) whose successive start times are separated by no more than 6 h (ref. 2). A total 1,361 outbreaks are found over the period 1954–2014. Outbreaks spanning more than 1 day are possible with this definition. The median outbreak duration is about 8 h, and 95% of the outbreaks last less than 24 h. Additional climatological features can be found in ref. 2. The number of tornadoes and F-km are computed for each outbreak. Then the annual mean and variance of two outbreak severity measures, number of tornadoes per outbreak and F-km per outbreak, are calculated. The distribution of outbreak tornadoes by F/EF-scale shows considerable variation before 1977 (Supplementary Fig. 1) and may well reflect lower reliability in the earlier period<sup>23</sup> before damage surveys were a routine part of the rating procedure. Therefore, we repeat some of our analysis using the recent period 1977–2014 (Supplementary Figs 2,3).

The same outbreak calculation procedure is repeated but considering only reports of tornadoes rated F/EF2 or greater, denoted as F2+. These outbreak events are referred to as F2+ outbreaks (Supplementary Figs 7–10). Only F2+ tornadoes are used to calculate tornado numbers and F-km of F2+ outbreaks.

**Trends.** All trends and 95% confidence intervals are assessed using linear regression and ordinary least squares, assuming approximately normal distributions of residuals. The growth rates of the annual mean and variance of the outbreak severity measures in Fig. 2b,c, Fig. 3b,c and Supplementary Figs 3b,c; 4b,c; 8b,c and 9b,c are

computed by assuming exponential growth and fitting a linear trend to the logarithms of the data. All the other trends are fitted using untransformed data.

**TL parameters.** The TL parameters in Figs 2d, 3d and 4d and Supplementary Figs 3d, 4d, 8d and 9d and 95% confidence intervals are estimated using ordinary least-squares regression with the logarithms of the mean and variance.

**TL parameters implied by LC theory.** The growth factor in the LC model is computed from  $A(t) = N(t+1)/N(t)$ , where  $N(t)$  is the average number of tornadoes per outbreak in calendar year  $t$ . The 95% confidence intervals for the mean and variance of  $A(t)$  are computed from 10,000 bootstrap samples and reported in Supplementary Table 1. The mean and variance of  $A(t)$  are used to compute a prediction of the slope  $b$  using equation (3). The TL parameter  $a$  is estimated from equation (1) evaluated at the initial year, either 1954 or 1977.

**TL parameters implied by sampling variability.** There are 1,361 cases of outbreak F-km, and their distribution is highly right-skewed (Supplementary Figs 5a and 10a). The F-km values for the 1974 Super Outbreak and the 25–28 April 2011 tornado outbreak are more than 26 standard deviations above the mean of the data on an arithmetic scale and more than six standard deviations above the mean of the log-transformed data, when means and standard deviations are calculated after withholding the two extreme values. These outliers (values that are far from other observations) have a substantial impact on the estimates of the mean, variance and skewness of the F-km distribution. Despite their rarity, about 86% ( $1 - (1 - 2/1,361)^{1,361}$ ) of the bootstrap samples will contain one or both of these two events. The presence of the outliers results in bimodal distributions of the TL slope and intercept estimates (Supplementary Figs 5b,c and 10b,c) computed from equation (4), depending on whether or not the outlier values are in the particular bootstrap sample. Removal of the outliers results in unimodal distributions, whose ranges are consistent with the least-squares estimates of the TL slope and intercept for F-km (Supplementary Figs 5d,e and 10d,e).

**Environmental proxies.** We use two environmental proxies: one for tornado occurrence and one for the number of tornadoes. Tornadoes are often associated with elevated values of convective available potential energy (CAPE;  $\text{J kg}^{-1}$ ) and a measure of vertical wind shear, called storm relative helicity (SRH;  $\text{m}^2 \text{s}^{-2}$ ). Here 0–180 hPa CAPE and 0–3,000 m SRH data are taken from the North American Regional Reanalysis<sup>24</sup>. The data are interpolated to a  $1 \times 1$  degree latitude–longitude grid over the continental United States and daily averages are computed. The environmental proxy for tornado occurrence is defined using the energy-helicity index<sup>25</sup> (EHI), which is the product of CAPE and SRH divided by  $160,000 \text{ J kg}^{-1} \text{ m}^2 \text{ s}^{-2}$ . Values of EHI greater than one indicate the potential for supercell thunderstorms and tornadoes<sup>26</sup>, and accordingly we take our proxy for tornado occurrence to be the condition that EHI is greater than 1. Selection of an environmental proxy for the number of tornadoes is more challenging. The only example of a proxy calibrated to predict the number of tornadoes from the surrounding environment uses monthly averaged environment and monthly number of tornadoes<sup>27,28</sup>, and therefore is not suited for outbreaks, which last much less than a month. However, previous research does provide some indications for the functional form of such a proxy. For instance, on subdaily timescales the likelihood of significant severe weather is nearly twice as sensitive to vertical wind shear as to CAPE<sup>29,30</sup>. This sensitivity would argue for a proxy based on the product of CAPE and the square of vertical wind shear or equivalently the product of the square root of CAPE and vertical wind shear<sup>31</sup>. Likewise, proxies for the monthly number of tornadoes contain SRH with an exponent ranging from 1.89 to 4.36 depending on region<sup>27,28</sup>. The supercell composite parameter<sup>32</sup> is used in weather forecasting and is the product of CAPE, SRH and vertical shear, again indicating a scaling of severe weather with the square of vertical wind shear measures. On the basis of this evidence, we take as the proxy for number of tornadoes

$$\frac{\text{CAPE} \times \text{SRH}^2}{3,600,000 \text{ m}^4 \text{ s}^{-4}}, \quad (5)$$

which differs from the EHI in that the square of SRH appears. The normalizing factor is chosen to match overall annual outbreak numbers. The proxy for the number of tornadoes per occurrence is therefore the mean of equation (5) conditional on the occurrence proxy of EHI being greater than one.

## References

- Data from Summary of U.S. Natural Hazard Statistics 2005–2014 compiled by the National Weather Service Office of Climate, Water and Weather Services and the National Climatic Data Center. <http://www.nws.noaa.gov/om/hazstats.shtml>.
- Fuhrmann, C. M. *et al.* Ranking of tornado outbreaks across the United States and their climatological characteristics. *Wea. Forecasting* **29**, 684–701 (2014).
- Smith, A. & Matthews, J. Quantifying uncertainty and variable sensitivity within the US billion-dollar weather and climate disaster cost estimates. *Nat. Hazards* **77**, 1829–1851 (2015).
- Diffenbaugh, N. S., Scherer, M. & Trapp, R. J. Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proc. Natl Acad. Sci. USA* **110**, 16361–16366 (2013).
- Gensini, V. & Mote, T. Downscaled estimates of late 21st century severe weather from CCSM3. *Clim. Change* **129**, 307–321 (2015).
- Tippett, M. K., Allen, J. T., Gensini, V. A. & Brooks, H. E. Climate and hazardous convective weather. *Curr. Clim. Change Rep.* **1**, 60–73 (2015).
- Verbout, S. M., Brooks, H. E., Leslie, L. M. & Schultz, D. M. Evolution of the U.S. tornado database: 1954–2003. *Wea. Forecasting* **21**, 86–93 (2006).
- Brooks, H. E., Carbin, G. W. & Marsh, P. T. Increased variability of tornado occurrence in the United States. *Science* **346**, 349–352 (2014).
- Elsner, J. B., Elsner, S. C. & Jagger, T. H. The increasing efficiency of tornado days in the United States. *Clim. Dyn.* **45**, 651–659 (2015).
- Tippett, M. K. Changing volatility of U.S. annual tornado reports. *Geophys. Res. Lett.* **41**, 6956–6961 (2014).
- Sander, J., Eichner, J. F., Faust, E. & Steuer, M. Rising variability in thunderstorm-related U.S. losses as a reflection of changes in large-scale thunderstorm forcing. *Wea. Climate Soc.* **5**, 317–331 (2013).
- Long, J. A. & Stoy, P. C. Peak tornado activity is occurring earlier in the heart of ‘Tornado Alley’. *Geophys. Res. Lett.* **41**, 6259–6264 (2014).
- Lu, M., Tippett, M. & Lall, U. Changes in the seasonality of tornado and favorable genesis conditions in the central United States. *Geophys. Res. Lett.* **42**, 4224–4231 (2015).
- Doswell, III C. A., Edwards, R., Thompson, R. L., Hart, J. A. & Crosbie, K. C. A simple and flexible method for ranking severe weather events. *Wea. Forecasting* **21**, 939–951 (2006).
- Taylor, L. R. Aggregation, variance and the mean. *Nature* **189**, 732–735 (1961).
- Eisler, Z., Bartos, I. & Kertész, J. Fluctuation scaling in complex systems: Taylor’s law and beyond. *Adv. Phys.* **57**, 89–142 (2008).
- Cohen, J. E., Xu, M. & Schuster, W. S. F. Stochastic multiplicative population growth predicts and interprets Taylor’s power law of fluctuation scaling. *Proc. R. Soc. B* **280**, 20122955 (2013).
- Malamud, B. D. & Turcotte, D. L. Statistics of severe tornadoes and severe tornado outbreaks. *Atmos. Chem. Phys.* **12**, 8459–8473 (2012).
- Elsner, J. B., Jagger, T. H., Widen, H. M. & Chavas, D. R. Daily tornado frequency distributions in the United States. *Environ. Res. Lett.* **9**, 024018 (2014).
- Cohen, J. E. & Xu, M. Random sampling of skewed distributions implies Taylor’s power law of fluctuation scaling. *Proc. Natl Acad. Sci. USA* **112**, 7749–7754 (2015).
- Vitolo, R., Stephenson, D. B., Cook, I. M. & Mitchell-Wallace, K. Serial clustering of intense European storms. *Meteor. Z.* **18**, 411–424 (2009).
- Karremann, M. K., Pinto, J. G., Reyers, M. & Klawa, M. Return periods of losses associated with European windstorm series in a changing climate. *Environ. Res. Lett.* **9**, 124016 (2014).
- Agee, E. & Childs, S. Adjustments in tornado counts, F-scale intensity, and path width for assessing significant tornado destruction. *J. Appl. Meteor. Climatol.* **53**, 1494–1505 (2014).
- Mesinger, F. *et al.* North American regional reanalysis. *Bull. Am. Meteorol. Soc.* **87**, 343–360 (2006).
- Davies, J. M. in *17th Conference on Severe Local Storms*, 107–111 (Amer. Meteor. Soc., St. Louis, MO, 1993).
- Rasmussen, E. N. & Blanchard, D. O. A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting* **13**, 1148–1164 (1998).
- Tippett, M. K., Sobel, A. H. & Camargo, S. J. Association of U.S. tornado occurrence with monthly environmental parameters. *Geophys. Res. Lett.* **39**, L02801 (2012).
- Tippett, M. K., Sobel, A. S., Camargo, S. J. & Allen, J. T. An empirical relation between U.S. tornado activity and monthly environmental parameters. *J. Clim.* **27**, 2983–2999 (2014).
- Brooks, H. E., Lee, J. W. & Craven, J. P. The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.* **67–68**, 73–94 (2003).
- Brooks, H. E. Severe thunderstorms and climate change. *Atmos. Res.* **123**, 129–138 (2013).
- Gilleland, E., Brown, B. G. & Ammann, C. M. Spatial extreme value analysis to project extremes of large-scale indicators for severe weather. *Environmetrics* **24**, 418–432 (2013).
- Thompson, R. L., Mead, C. M. & Edwards, R. Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting* **22**, 102–115 (2007).

## Acknowledgements

M.K.T. was partially supported by a Columbia University Research Initiatives for Science and Engineering (RISE) award, the Office of Naval Research award N00014-12-1-091, NOAA’s Climate Program Office’s Modeling, Analysis, Predictions and Projections program award NA14OAR4310185, and the Willis Research Network. J.E.C. was partially supported by U.S. National Science Foundation grant DMS-1225529. The views

expressed herein are those of the authors and do not necessarily reflect the views of any of the sponsoring agencies.

### Author contributions

All the authors were involved in conducting the research and the scientific discussion equally throughout the study. The data were prepared by M.K.T. The manuscript was written, edited and reviewed by both the authors.

### Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

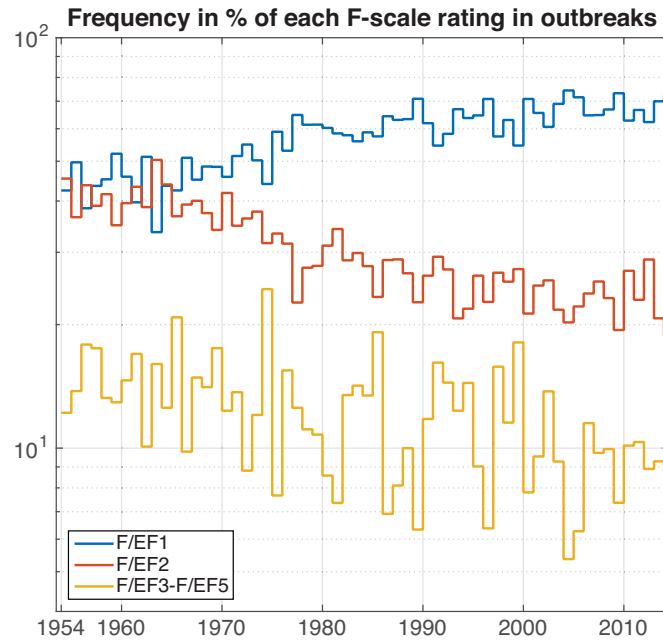
**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Tippett, M. K. & Cohen J. E. Tornado outbreak variability follows Taylor's power law of fluctuation scaling and increases dramatically with severity. *Nat. Commun.* 7:10668 doi: 10.1038/ncomms10668 (2016).

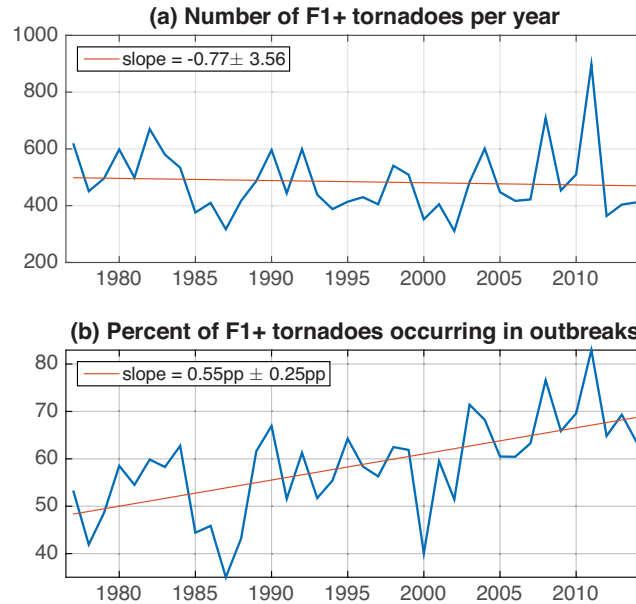


This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

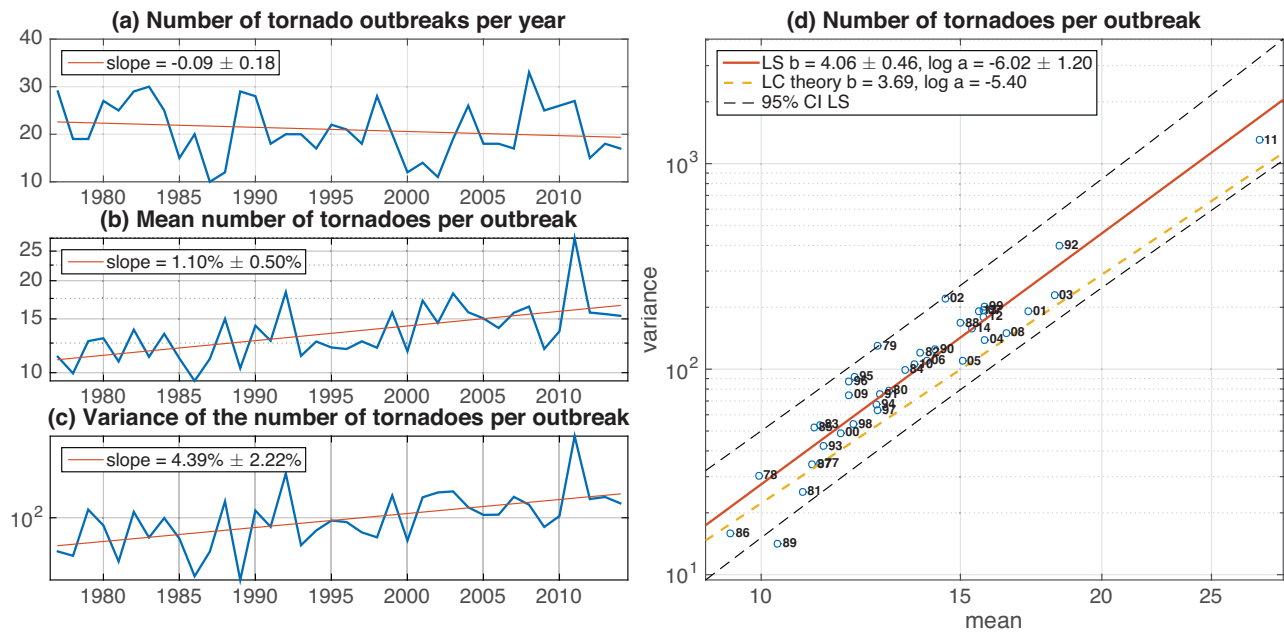


**Supplementary Figure 1 | Distribution of outbreak tornadoes by F/EF-scale rating.** Frequency (in percent per year) of F/EF1 (blue), F/EF2 (red) and F/EF3-F/EF5 (yellow) rated tornadoes in tornado outbreaks. Percentages sum to 100% in each year.

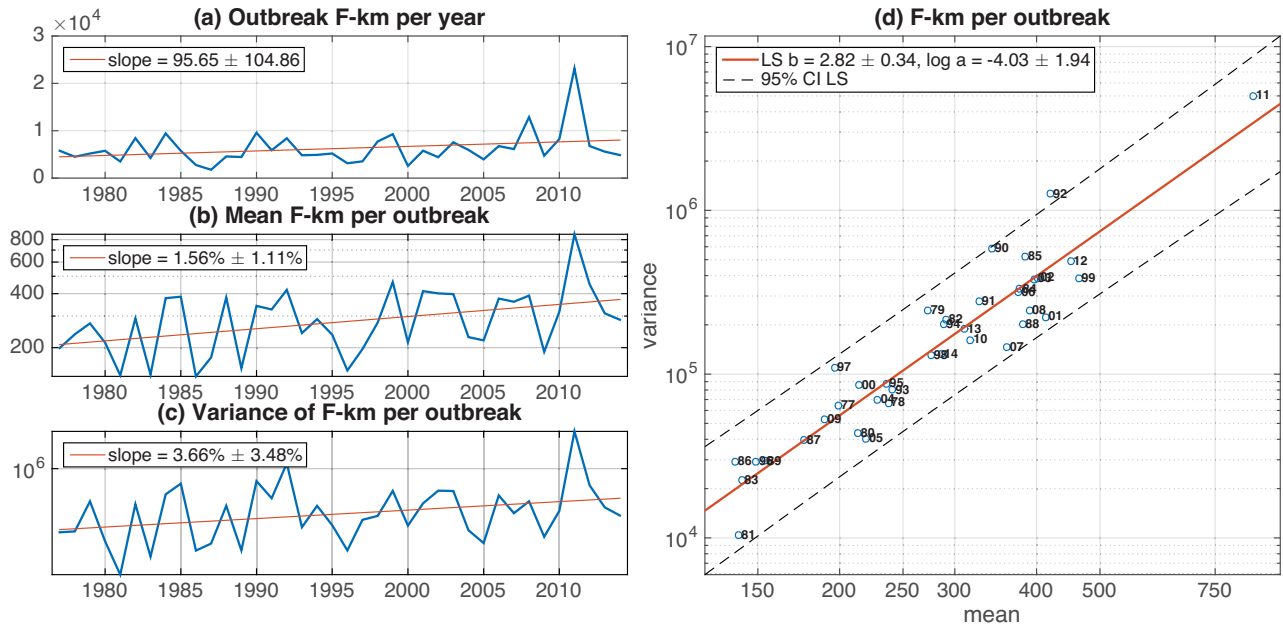




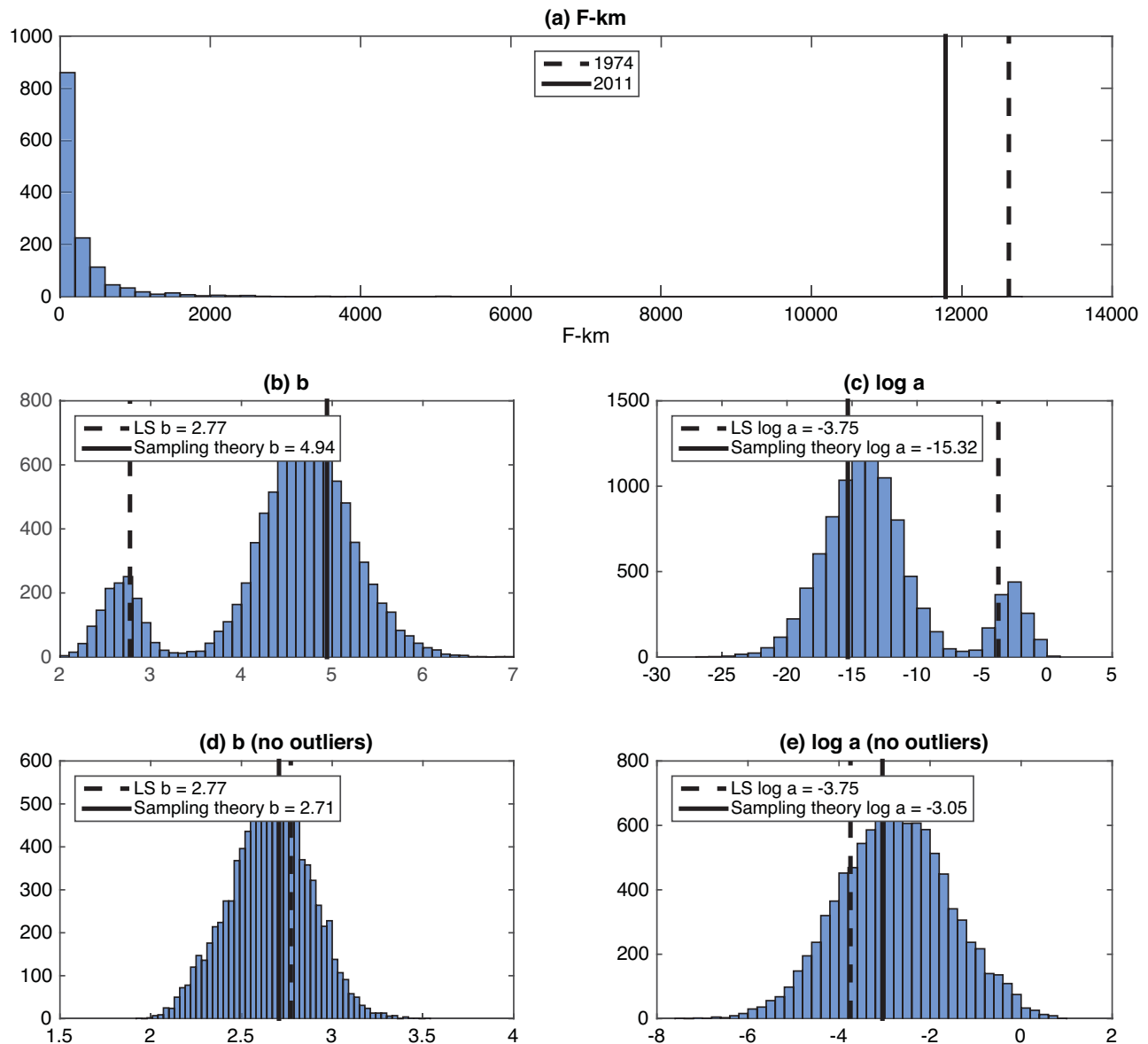
**Supplementary Figure 2 | Time series of counts and clustering of F1+ tornadoes 1977-2014 in the contiguous US.** (a) Number of F1+ tornadoes per year. The slope of the least-squares regression indicates that the number of F1+ tornadoes per year declined by 0.77 per year on average from 1977 to 2014 inclusive. This rate of decline is not statistically significantly different from 0 (no change). (b) Annual percentage of F1+ tornadoes occurring in outbreaks. The slope of the least-squares regression indicates that the percentage of F1+ tornadoes per year that occurred as part of outbreaks increased by 0.55 percentage points (pp) per year on average from 1977 to 2014 inclusive. This increase is statistically significantly greater than 0. In both (a) and (b),  $\pm$  intervals are 95% confidence intervals.



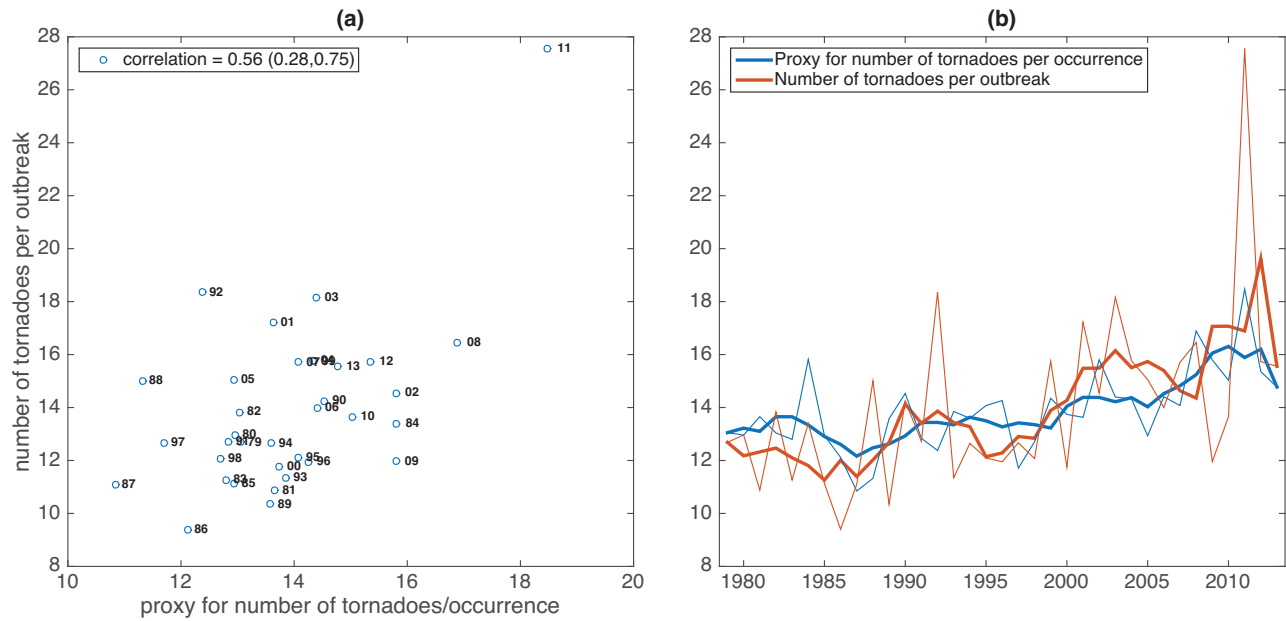
**Supplementary Figure 3 | Numbers of F1+ tornadoes per outbreak 1977-2014.** (a) Number of tornado outbreaks per year. The rate of decline is not statistically significantly different from 0 (no change). (b) Annual mean number of tornadoes per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (c) Annual variance of the number of tornadoes per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (d) Scatter plot of the annual mean number of tornadoes per outbreak versus the annual variance of the number of tornadoes per outbreak. Both axes are on logarithmic scale. The solid red line is the least-squares (LS) regression line (Taylor's power law of fluctuation scaling) and the dashed yellow line has the slope and intercept predicted by LC theory<sup>1</sup>. The two-digit number following the plotting symbol 'o' gives the calendar year in the second half of the 20th century or first half of the 21st century. In all panels,  $\pm$  intervals are 95% confidence intervals.



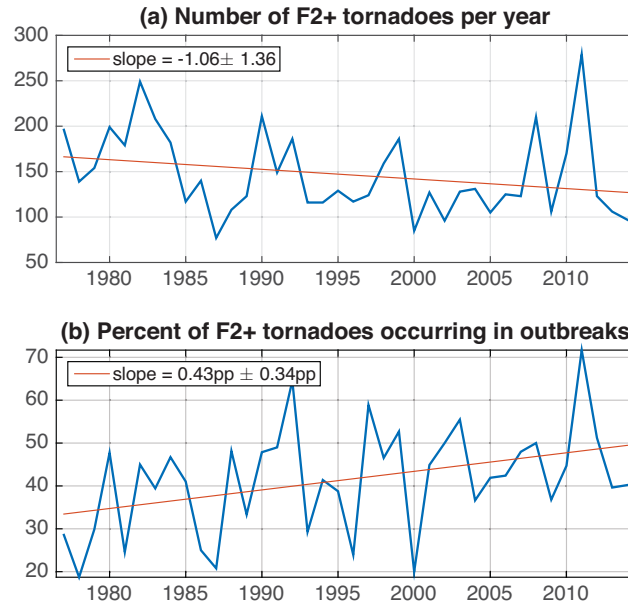
**Supplementary Figure 4 | F-km per outbreak 1977-2014.** (a) Total outbreak F-km per year. The rate of increase is not statistically significantly different from 0 (no change) among F1+ tornadoes. (b) Annual mean F-km per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (c) Annual variance of F-km per outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual variance is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (d) Scatter plot of the annual mean of F-km per outbreak versus the annual variance of F-km per outbreak. Both axes are on logarithmic scale. The solid red line is the least-squares (LS) regression line (Taylor's power law of fluctuation scaling). The two-digit number following the plotting symbol 'o' gives the calendar year in the second half of the 20th century or first half of the 21st century. In all panels,  $\pm$  intervals are 95% confidence intervals.



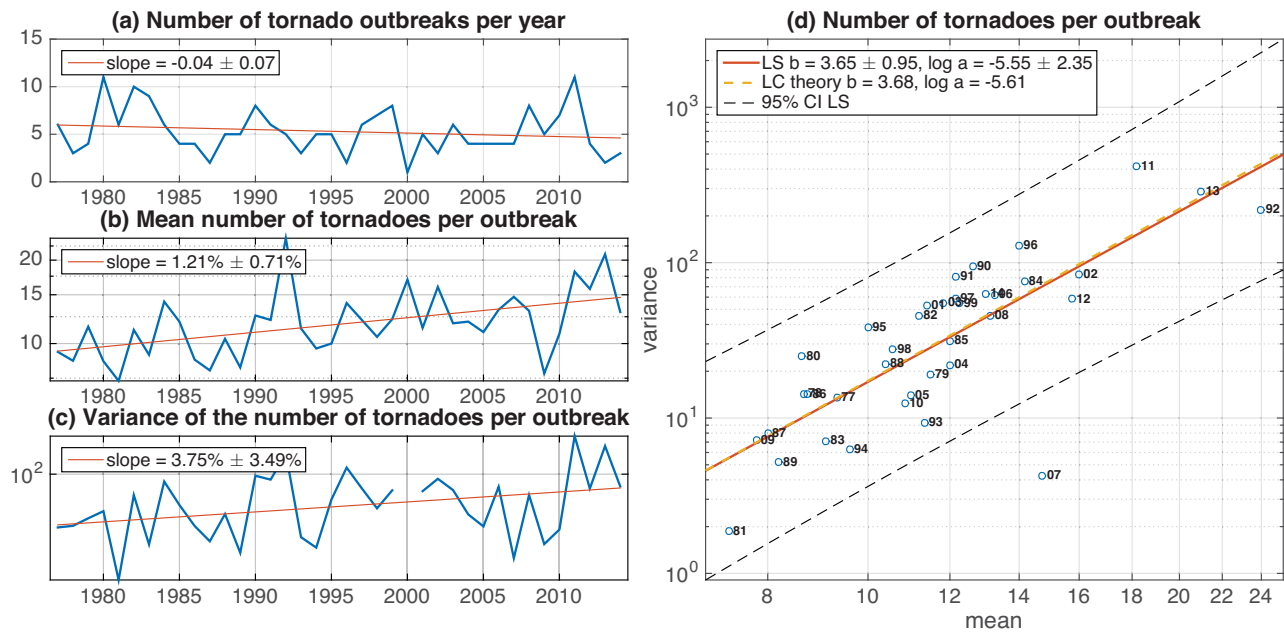
**Supplementary Figure 5 | Distribution of F-km and sampling theory estimates of TL parameters for outbreaks 1954-2014.** (a) Frequency distribution of F-km including the extreme values of 1974 and 2011. (b-c) Histogram of the TL parameters estimated from 10,000 bootstrap samples. (d-e) Histogram of the TL parameters estimated from 10,000 bootstrap samples excluding the extreme values of 1974 and 2011. The vertical axis is the frequency of occurrence among the 10,000 bootstrap samples.



**Supplementary Figure 6 | Environmental proxy.** (a) Scatter plot of the annual mean of number of F1+ tornadoes per outbreak and the annual average of the environmental proxy for number of tornadoes per occurrence for 1979-2013. Two-digit numbers indicate year. The linear correlation is 0.56. (b) Annual (thin lines) and 5-year moving average (thick lines) time series of the mean of number of tornadoes per outbreak and the average of the environmental proxy for number of tornadoes per occurrence for 1979-2013. The correlation of the 5-year moving averages is 0.88.



**Supplementary Figure 7 | Time series of counts and clustering of F2+ tornadoes 1977-2014 in the contiguous US.** (a) Number of F2+ tornadoes per year. The slope of the least-squares regression indicates that the number of F2+ tornadoes per year declined by 1.06 per year on average from 1977 to 2014 inclusive. The rate of decrease is not statistically significantly different from 0 (no change). (b) Annual percentage of F2+ tornadoes occurring in F2+ outbreaks. The slope of the least-squares regression indicates that the percentage of F1+ tornadoes per year occurring in outbreaks increased by 0.43 percentage points (pp) per year on average from 1977 to 2014 inclusive. This rate of increase is statistically significantly greater than 0. In both (a) and (b),  $\pm$  intervals are 95% confidence intervals.



**Supplementary Figure 8 | Numbers of F2+ tornadoes per F2+ outbreak 1977-2014. (a)**

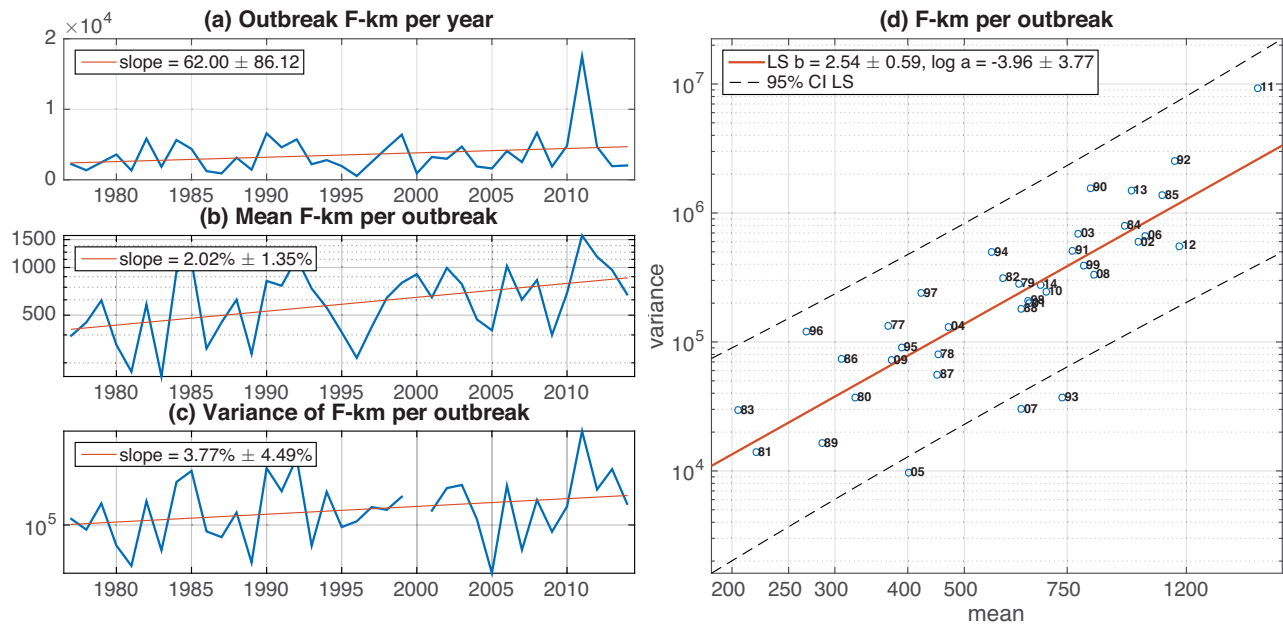
Number of F2+ outbreaks per year. The rate of decline is not statistically significantly different from 0 (no change).

(b) Annual mean number of tornadoes per F2+ outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0.

(c) Annual variance of the number of tornadoes per F2+ outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0.

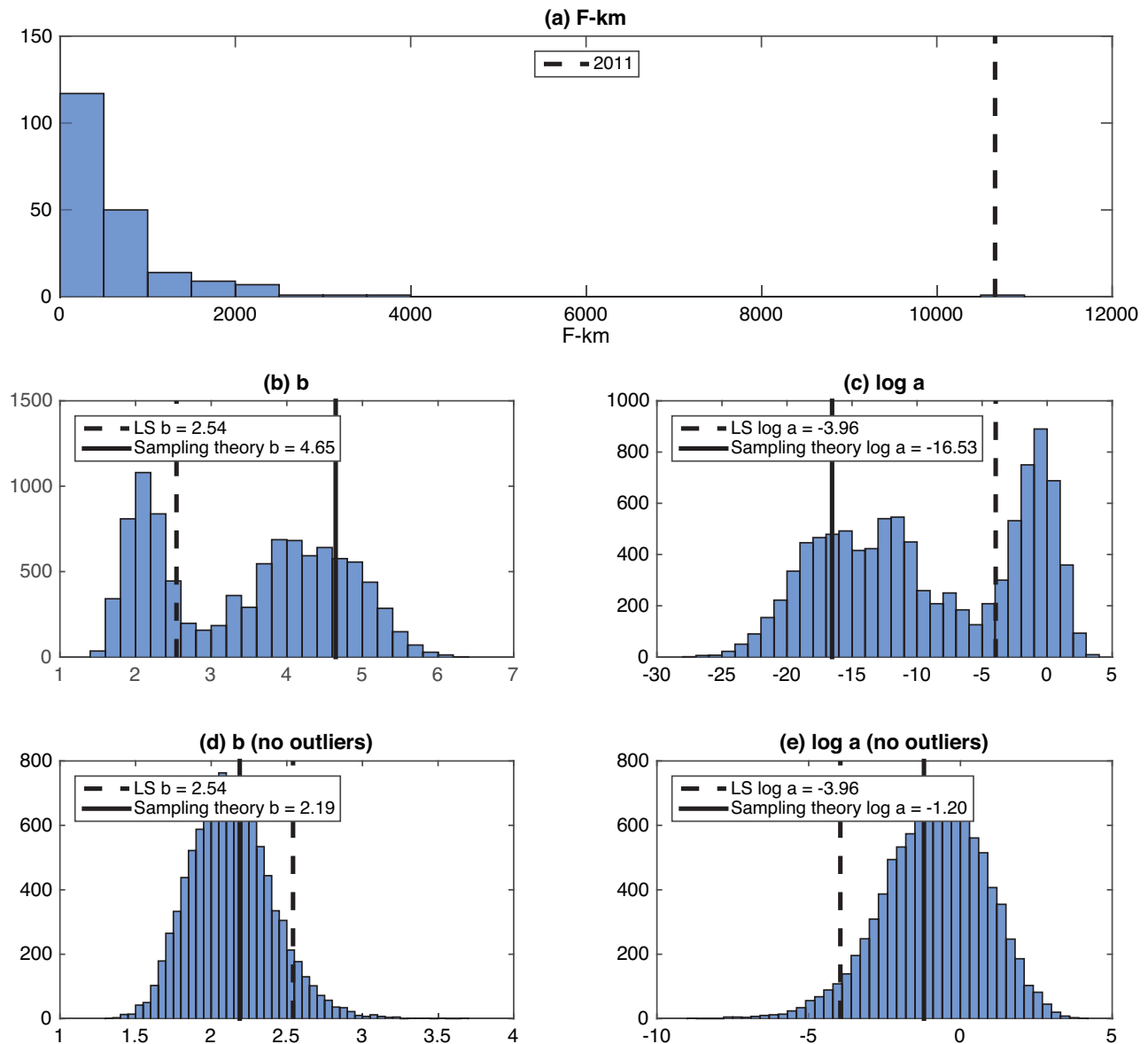
(d) Scatter plot of the annual mean number of tornadoes per F2+ outbreak versus the annual variance of the number of tornadoes per F2+ outbreak. Both axes are on logarithmic scale. The solid red line is the least-squares (LS) regression line (Taylor's power law of fluctuation scaling) and the dashed yellow line has the slope and intercept predicted by LC theory<sup>1</sup>. The two-digit number following the plotting symbol 'o' gives the calendar year in the second half of the 20th century or first half of the 21st century. In all panels, ± intervals are 95% confidence intervals.

There is one F2+ outbreak in 2000, and the variance of the number of tornadoes per F2+ outbreak in 2000 is not defined.

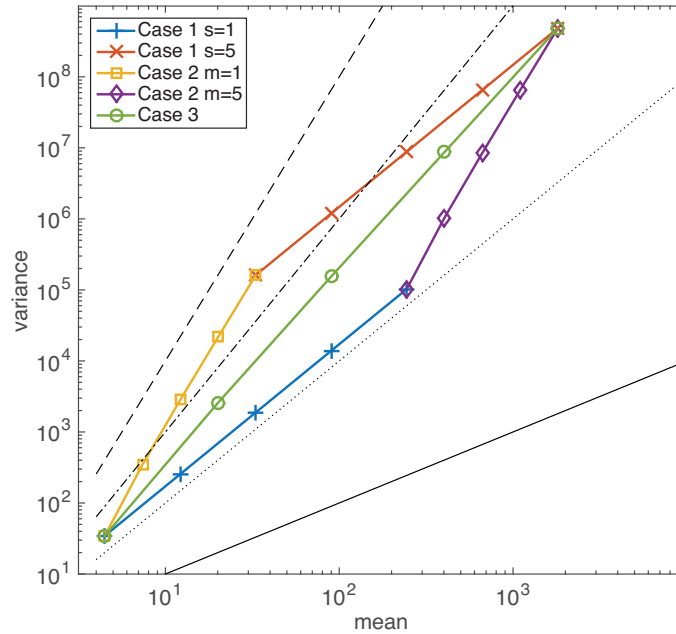


**Supplementary Figure 9 | F-km per F2+ outbreak 1977-2014.** (a) Total outbreak F-km per year. The rate of increase is not statistically significantly different from 0 (no change) among F2+ tornadoes. (b) Annual mean F-km per F2+ outbreak. Vertical axis is on logarithmic scale, so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is statistically significantly greater than 0. (c) Annual variance of F-km per outbreak. Vertical axis is on logarithmic scale so the rate of increase in the annual mean is expressed as a percentage per year. This rate of increase is not statistically significantly greater than 0. (d) Scatter plot of the annual mean of F-km per F2+ outbreak versus the annual variance of F-km per F2+ outbreak. Both axes are on logarithmic scale. The solid red line is the least-squares (LS) regression line (Taylor's power law of fluctuation scaling). The two-digit number following the plotting symbol 'o' gives the calendar year in the second half of the 20th century or first half of the 21st century. In all panels,  $\pm$  intervals are 95% confidence intervals. There is one F2+ outbreak in 2000, and the log variance of the number of tornadoes per F2+ outbreak in 2000 is not defined.





**Supplementary Figure 10 | Distribution of F-km and sampling theory estimates of TL parameters for F2+ outbreaks 1977-2014. (a)** Frequency distribution of F-km including the extreme value of 2011. **(b-c)** Histogram of the TL parameters estimated from 10,000 bootstrap samples. **(d-e)** Histogram of the TL parameters estimated from 10,000 bootstrap samples excluding the extreme value of 2011.



**Supplementary Figure 11 | Taylor's law scaling for lognormal distributions with changing parameters.** The thin black diagonal lines that pass through the origin have slopes 1 (solid line), 2 (dotted line), 3 (dash-dotted line), and 4 (dashed line), reading from the bottom to the top. These are guidelines for the eye. The thicker lines join markers which plot values of  $(x, y) = (\mu(m, s^2), \sigma^2(m, s^2))$ . Case 1 is illustrated by the two sides of the parallelogram with slope exactly 2, where  $m = 1, 2, 3, 4, 5$  and  $s^2 = 1$  (blue + markers, along the bottom of the parallelogram) and  $m = 1, 2, 3, 4, 5$  and  $s^2 = 5$  (red × markers, along the top of the parallelogram). Case 2 is illustrated by the two sides of the parallelogram with slope approximately 4, where  $m = 1$  and  $s^2 = 1, 2, 3, 4, 5$  (square yellow markers, along the left side of the parallelogram) and  $m = 5$  and  $s^2 = 1, 2, 3, 4, 5$  (diamond purple markers, along the right side of the parallelogram). Case 3 is illustrated by the diagonal of the parallelogram with slope approximately 8/3, where  $m = s^2 = 1, 2, 3, 4, 5$  (o green markers).

Mean (M) or variance (V) of multiplicative factors $A(t)$	Period	Point estimate	2.5%-tile	97.5%-tile
M	1954-2014	1.03	0.97	1.10
V	1954-2014	0.068	0.045	0.13
M	1977-2014	1.04	0.96	1.14
V	1977-2014	0.079	0.037	0.13

**Supplementary Table 1 | Estimates of the parameters from LC theory and their 95% confidence intervals over the periods 1954-2014 and 1977-2014.** The confidence intervals are based on 10,000 bootstrap samples.

## Supplementary Discussion

Here we give one possible interpretation of our finding that the mean and variance of the number of tornadoes per outbreak approximately satisfy TL with  $b = 4.3 \pm 0.44$ , i.e.,  $b$  does not differ significantly from 4. This behavior contrasts with that of the Poisson and negative binomial distributions where the variance is a linear and quadratic function of the mean, respectively. We shall show that  $b \approx 4$  is a consequence, under certain conditions which we shall specify, of the multiplicative population process proposed by Supplementary Ref. 2. The lognormal distribution is a limiting distribution of such processes.

By way of background, we recall that a positive-valued random variable  $X$  is lognormally distributed if  $\log X$  is normally distributed. Supplementary Reference 3 gives an informal introduction to the lognormal distribution. Since the normal distribution is fully specified by its mean  $m$  and its variance  $s^2$ , the lognormal distribution is fully specified by the same  $m$  and  $s^2$ . Explicitly,  $X(m, s^2)$  is lognormally distributed with parameters  $m$  and  $s^2$  if and only if  $\log(X(m, s^2))$  is normally distributed with mean  $m$  and variance  $s^2$ .

When the independent multiplicative factors  $A(t)$  have finite mean  $M$  and finite variance  $V$ , their cumulative product  $N(t)$  becomes lognormally distributed as  $t$  becomes large<sup>2</sup>. This convergence in distribution is a direct consequence of the central limit theorem applied to  $\log N(t+1) = \log N(0) + \log A(0) + \dots + \log A(t)$ , which follows from the definition of  $A(t)$  in equation (2) of our main text.

The following facts are standard:

$$E(X(m, s^2)) \equiv \mu(m, s^2) = \exp(m + s^2/2), \quad (1)$$

$$SD(X(m, s^2)) = (\text{Var}(X(m, s^2)))^{1/2} \equiv \sigma(m, s^2) = \exp(m + s^2/2) [\exp(s^2) - 1]^{1/2}, \quad (2)$$

$$CV(m, s^2) \equiv (\text{Var}(X(m, s^2)))^{1/2} / E(X(m, s^2)) = [\exp(s^2) - 1]^{1/2}, \quad (3)$$

which is independent of  $m$ .

A less-widely known but obvious relationship that is relevant to TL is that

$$\sigma^2(m, s^2) = \mu(m, s^2)^2 [\exp(s^2) - 1]. \quad (4)$$

Three cases of this relationship are illustrated in Supplementary Fig. 11.

Case 1. If  $m$  is changing and  $s^2$  is fixed,  $b = 2$ . Taylor's law holds exactly.

Case 2. If  $m$  is fixed and  $s^2$  is changing, then the variance is not exactly a power law function of the mean because the coefficient  $[\exp(s^2)-1]$  is changing. As  $s^2$  increases, both the mean  $\mu(m, s^2)$  and the variance  $\sigma^2(m, s^2)$  of  $X(m, s^2)$  increase. In this case, we can write the variance as a function of the mean since  $\mu(m, s^2)^2 = \exp(2m)\exp(s^2)$  and  $[\mu(m, s^2)\exp(-m)]^2 = \exp(s^2)$ , so

$$\exp(s^2) - 1 = [\mu(m, s^2)\exp(-m)]^2 - 1, \quad (5)$$

and

$$\sigma^2(m, s^2) = \mu(m, s^2)^2 [\exp(s^2) - 1] = \mu(m, s^2)^2 \{ [\mu(m, s^2)\exp(-m)]^2 - 1 \}. \quad (6)$$

Here the leading term is proportional to  $\mu(m, s^2)^4$ . For different values of  $s^2$ , Taylor's law holds approximately with exponent 4.

Case 3. Suppose that when  $m$  changes,  $s^2$  also changes according to  $s^2 = m$ . In this case,

$$\mu(m, m) = \exp(m+m/2) = \exp(3m/2), \quad (7)$$

$$\sigma(m, m) = \exp(3m/2)[\exp(m)-1]^{1/2} \quad (8)$$

and  $\sigma^2(m, m) = \mu(m, m)^2 [\mu(m, m)^{2/3} - 1]$ . In this case the leading term is equal to  $\mu(m, m)^{8/3}$ . Taylor's law holds not exactly but approximately and the exponent is  $2 + 2/3$ .

Why would  $s$  vary? In an ecological context, a spatial landscape could be composed of patches with varying degrees of disturbance. Highly disturbed patches would be expected to have higher  $s$  than less disturbed patches. In a meteorological context, different years might experience different levels of climate variability, e.g., La Nina events or fluctuations in jet streams. Highly disturbed years would be expected to have higher  $s$  than less disturbed years.

We suggest that our finding that  $b \approx 4$  could arise because, as time is passing,  $m$  is relatively steady but  $s^2$  is increasing.

## Supplementary References

- <sup>1</sup> Cohen, J. E., Xu, M. & Schuster, W. S. F. Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. *Proc. R. Soc. B*, **280**, 20122955 (2013).
- <sup>2</sup> Lewontin, R. C. & Cohen, D. On population growth in a randomly varying environment. *Proc. Natl. Acad. Sci. (USA)*, **62**, 1056-1060 (1969).
- <sup>3</sup> Limpert, E., Stahel, W., & Abbt, M. Log-normal distributions across the sciences: Keys and clues. *BioScience*, **51**, 341–352 (2001).