



Spatial–temporal clustering of tornadoes

Bruce D. Malamud¹, Donald L. Turcotte², and Harold E. Brooks³

¹Department of Geography, King's College London, London, WC2R 2LS, UK

²Department of Geology, University of California Davis, CA 95616, USA

³National Severe Storm Laboratory, National Oceanic and Atmospheric Administration, Norman, OK 73072, USA

Correspondence to: Bruce D. Malamud (bruce.malamud@kcl.ac.uk)

Received: 3 March 2016 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: 23 March 2016

Revised: 30 October 2016 – Accepted: 7 November 2016 – Published: 21 December 2016

Abstract. The standard measure of the intensity of a tornado is the Enhanced Fujita scale, which is based qualitatively on the damage caused by a tornado. An alternative measure of tornado intensity is the tornado path length, L . Here we examine the spatial–temporal clustering of severe tornadoes, which we define as having path lengths $L \geq 10$ km. Of particular concern are tornado outbreaks, when a large number of severe tornadoes occur in a day in a restricted region. We apply a spatial–temporal clustering analysis developed for earthquakes. We take all pairs of severe tornadoes in observed and modelled outbreaks, and for each pair plot the spatial lag (distance between touchdown points) against the temporal lag (time between touchdown points). We apply our spatial–temporal lag methodology to the intense tornado outbreaks in the central United States on 26 and 27 April 2011, which resulted in over 300 fatalities and produced 109 severe ($L \geq 10$ km) tornadoes. The patterns of spatial–temporal lag correlations that we obtain for the 2 days are strikingly different. On 26 April 2011, there were 45 severe tornadoes and our clustering analysis is dominated by a complex sequence of linear features. We associate the linear patterns with the tornadoes generated in either a single cell thunderstorm or a closely spaced cluster of single cell thunderstorms moving at a near-constant velocity. Our study of a derecho tornado outbreak of six severe tornadoes on 4 April 2011 along with modelled outbreak scenarios confirms this association. On 27 April 2011, there were 64 severe tornadoes and our clustering analysis is predominantly random with virtually no embedded linear patterns. We associate this pattern with a large number of interacting supercell thunderstorms generating tornadoes randomly in space and time. In order to better understand these associations, we also applied our approach to the Great Plains tornado outbreak of 3 May 1999. Careful

studies by others have associated individual tornadoes with specified supercell thunderstorms. Our analysis of the 3 May 1999 tornado outbreak directly associated linear features in the largely random spatial–temporal analysis with several supercell thunderstorms, which we then confirmed using model scenarios of synthetic tornado outbreaks. We suggest that it may be possible to develop a semi-automated modelling of tornado touchdowns to match the type of observations made on the 3 May 1999 outbreak.

1 Introduction

The touchdown of a tornado is a point event in space and time in analogy to the initial point of rupture of an earthquake. The path length of tornado touchdowns is a measure of the strength of the tornado, in analogy to the Richter magnitude of an earthquake. In this paper, we consider the spatial and temporal statistics of tornado touchdowns for three USA tornado outbreak events from 1999 and 2011. We restrict our attention to severe tornadoes, those tornadoes with path lengths $L \geq 10$ km. The available data in the USA are quite complete for these severe tornadoes.

This paper takes a methodology for spatial–temporal clustering analysis developed by Zaliapin et al. (2008) for seismicity and applies it to tornadoes. Their methodology considers the times of occurrence and locations of point events. All pairs of events are considered and the spatial lag d (distance between a pair of events) is plotted against the temporal lag τ (the time difference between the pair of events). The methodology was developed to decluster earthquake aftershocks from background seismicity. All earthquakes have aftershock sequences; the aftershocks are clusters close to the

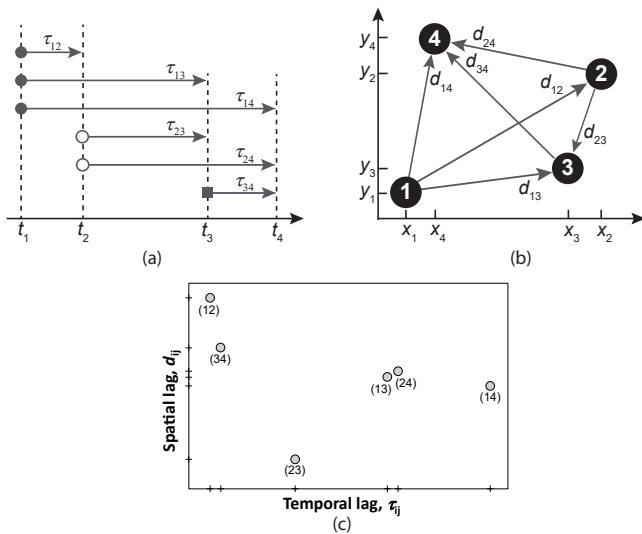


Figure 1. Illustration of our clustering analysis methodology. **(a)** A sequence of four point events occur at times t_1, t_2, t_3, t_4 . The six temporal lags $\tau_{12}, \tau_{13}, \tau_{14}, \tau_{23}, \tau_{24}, \tau_{34}$ are shown. **(b)** The two-dimensional locations of the four point events are shown. The six spatial lags, $d_{12}, d_{13}, d_{14}, d_{23}, d_{24}, d_{34}$ are also shown. **(c)** The six spatial lags d_{ij} are shown as a function of the corresponding temporal lags τ_{ij} , where i is the first event and j is the second event in time.

main shocks in both time and space. However, background seismicity (other main shocks) will also occur in the region and time interval in which the aftershocks occur. The background seismicity will occur randomly in space and time, whereas the aftershocks of each background earthquake will be tightly clustered in space and time. It is important to separate the aftershocks from the background seismicity in order to study the statistics of the aftershocks. Zaliapin et al. (2008) demonstrated that plots of spatial lag vs. temporal lag clearly separated the two groups of earthquakes.

Here we consider the time and place of the touchdown of a tornado as a point event. Our studies will be concentrated on tornado outbreaks. An outbreak is a sequence of several to hundreds (Fuhrmann et al., 2014) of spatially correlated tornadoes that occur in a relatively short period of time, typically a day, with generally fewer tornadoes at night as severe convection is inhibited. In contrast, earthquake aftershock sequences are unrestricted in time by convective activity, and a severe earthquake of $M = 6$ would typically generate thousands of aftershocks down to $M = 2$ (e.g. Utsu, 1970) compared to a severe tornado outbreak involving just hundreds of tornadoes. Despite differences in process and scales between tornadoes and earthquakes, as discussed above, both tornado touchdowns and earthquake aftershocks can be considered as point events, and the spatial–temporal methodology developed by Zaliapin et al. (2008) for seismicity is a very useful analysis for tornado outbreaks.

In this paper we will give several examples of tornado outbreaks, including maps of the tornado touchdown points as well as a clustering analysis of the dependence of spatial lag d_{ij} between the touchdowns of two tornadoes on the temporal lag τ_{ij} between the touchdown times of the same two tornadoes. We consider each tornado i and measure distance and times to each subsequent tornado j in the sequence. If the tornadoes occur randomly in space and time, the dependence of d_{ij} on τ_{ij} will also be random. Alternatively, a tornado outbreak could be a near-linear sequence of tornado touchdowns produced by a single supercell thunderstorm moving at a near-constant velocity. In this case, the dependence of d_{ij} on τ_{ij} is approximately linear, and the slope is the velocity of the convective cell.

To illustrate this clustering analysis methodology applied to tornadoes, we consider a sequence of four point events that occur at successive times t_1, t_2, t_3, t_4 and two-dimensional locations $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$, as illustrated in Fig. 1. The temporal lags (time differences) are $\tau_{12} = t_2 - t_1$, $\tau_{13} = t_3 - t_1$, $\tau_{14} = t_4 - t_1$, $\tau_{23} = t_3 - t_2$, $\tau_{24} = t_4 - t_2$ and $\tau_{34} = t_4 - t_3$. The corresponding spatial lags (spatial separations) are $d_{12} = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{0.5}$ and $d_{13}, d_{14}, d_{23}, d_{24}$ and d_{34} determined in a similar way. The temporal lags τ for our four point events are illustrated in Fig. 1a and the spatial lags d in Fig. 1b. The dependence of the spatial lags d on the temporal lags τ are given in Fig. 1c. In this paper, we will show the dependence of spatial lags on temporal lags for pairs of tornado touchdowns.

Studies of the statistics of tornadoes are limited by the problems associated with the quantification of tornado intensity. Ideally, tornado intensities would be based on wind speed measurements. However, as noted by Doswell et al. (2009), high-resolution Doppler measurements of wind velocities in tornadoes are not possible at this time. Currently, the Enhanced Fujita scale is the standard measure of tornado intensities (Edwards et al., 2013). Tornadoes are classified on a scale of EF0 to EF5 based on a qualitative measure of damage. An alternative measure of tornado intensity is the tornado path length, L . In the United States, the NOAA (2015) Storm Prediction Center Severe Weather Database (SPC–SWD) provides Enhanced-Fujita-scale values and path lengths for tornadoes. Brooks (2004) has provided a detailed study of the statistical correlations between the Fujita-scale intensities and the path length. Malamud and Turcotte (2012) extended these studies and defined a severe tornado to be a tornado with path length $L \geq 10$ km. In the studies reported in this paper, we will retain this definition and consider only tornadoes with $L \geq 10$ km. A path length of $L = 10$ km corresponds roughly to an EF2 tornado (Malamud and Turcotte, 2012). Most severe tornadoes are generated by supercell thunderstorms (Doswell et al., 1993). A supercell thunderstorm can be defined as a long-lived (> 1 h) thunderstorm with a high degree of spatial correlation between its mesocyclone (the vortex of air within the storm) and updraught (Davies-Jones et al., 2001).

The objective of this paper is to study the clustering statistics of tornado outbreaks. However, it must be recognized that the definition of a tornado outbreak is somewhat arbitrary (Mercer et al., 2009). Ideally, the definition of a tornado outbreak would be the occurrence of multiple tornadoes within a particular synoptic-scale weather system, but the spatial and temporal limits on the weather system are subject to arbitrary distinction (Glickman, 2000). Galway (1977) classified tornado outbreaks into three types: (i) a local outbreak with a radius less than 1000 miles (1609 km), (ii) a progressive outbreak moving from west to east in time and (iii) a line outbreak associated with a single moving supercell thunderstorm. Unfortunately, the NOAA (2015) NWS-SPC database does not associate individual tornadoes with a specific tornado outbreak using any of these three (or other) classifications.

There is a strong diurnal variability in tornado occurrence associated with solar heating. For these reasons, Doswell et al. (2006) defined a tornado outbreak to include all tornadoes in the continental USA in a convective day, i.e. the 24 h period from 12:00 UTC (Coordinated Universal Time) of a given day to 12:00 UTC of the following day, with 12:00 UTC (04:00 to 08:00 local time in the continental USA depending on month and location) corresponding to the approximate time of the daily minimum in tornado occurrence. The Severe Weather Database that we use in our analyses list most tornadoes in Central Standard Time (CST), so we will consider tornadoes in a convective day as 06:00–06:00 CST. However, consistent with the studies of severe tornado outbreaks given by Malamud and Turcotte (2012), we will consider a severe tornado outbreak to include only those tornadoes with path lengths $L \geq 10$ km. Elsner et al. (2015) developed a method for separating distinct spatial clusters of tornado touchdowns during a convective day. Our methods differ in that we consider both space and time and are searching for near-linear features.

2 Clustering analysis of tornadoes

To illustrate our clustering analysis methodology for tornadoes, we will first consider the intense tornado outbreaks in the central United States on 26 and 27 April 2011. The tornado outbreaks in the spring of 2011 have been discussed in detail by Doswell et al. (2012). They concluded that ideal conditions for severe tornado outbreaks occurred during the last 2 weeks of April 2011, and that the supercell thunderstorms responsible for the tornadoes were generated by a sequence of extratropical cyclones. In this paper, we focus our attention on the outbreaks that occurred on 26 and 27 April 2011. Although these outbreaks were certainly related to the same synoptic-scale weather pattern, we will treat the two outbreaks separately for our statistical studies. We will consider severe ($L \geq 10$ km) tornadoes on convective days: (i) 06:00 on 26 April to 06:00 CST on 27 April 2011 (i.e. a

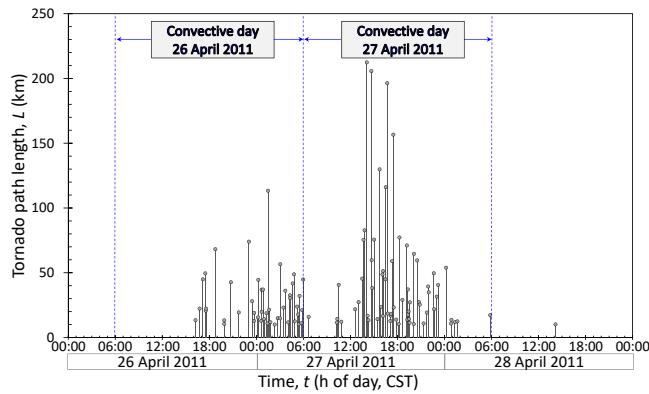


Figure 2. Tornado outbreak on 26–28 April 2011. Times of touchdown and path lengths of severe tornadoes ($L \geq 10$ km) that occurred on 26, 27 and 28 April 2011. There were 45 severe tornadoes on 26 April (convective day, 06:00 CST to 06:00 CST the following day) and 64 severe tornadoes on 27 April (convective day). Data were obtained from NOAA (2015).

convective day equivalent to 12:00 on 26 April to 12:00 UTC on 27 April 2011) and (ii) 06:00 on 27 April to 06:00 CST on 28 April 2011.

In Fig. 2 we give touchdown times t and path lengths L for the 45 severe ($L \geq 10$ km) tornadoes that occurred on 26 April 2011 (convective day, 06:00 to 06:00 CST of the following day) and for the 64 severe tornadoes that occurred on 27 April 2011 (convective day). In Malamud and Turcotte (2012), we suggested that a quantitative measure of the strength of a severe tornado outbreak is the total path length L_D of all severe ($L \geq 10$ km) tornadoes in a convective day in the continental USA. By this measure the strongest tornado outbreak during the 60-year period 1954–2013 was on 3 April 1974 (convective day) with 105 severe ($L \geq 10$ km) tornadoes and a total tornado path length $L_D = 3852$ km. For the two outbreaks illustrated in Fig. 2, the outbreak on 26 April 2011 with 45 severe tornadoes had a total tornado path length $L_D = 1239$ km, the fifth strongest outbreak during this same 60-year period, 1954–2013. The outbreak on 27 April 2011 with 64 severe tornadoes had a total path length $L_D = 2815$ km, the second strongest outbreak during this period.

We next consider the spatial distributions of the tornado touchdown points for both the 26 and 27 April 2011 outbreak events. In Fig. 3a we give a map of the tornado paths of the 45 severe ($L \geq 10$ km) tornadoes that occurred on 26 April 2011 (convective day) and in Fig. 3b, the 64 severe tornadoes that occurred on 27 April 2011 (convective day). In Fig. 3, the tornado touchdowns are given by symbols and the paths by lines. The symbols for tornado touchdowns are given by shapes and colours, with combinations defining eight 3 h periods for the initial touchdown times. The lines for each tornado path length illustrate the overall tornado movements. Tornado path lengths vary from $L = 10$ km (our lower cut-

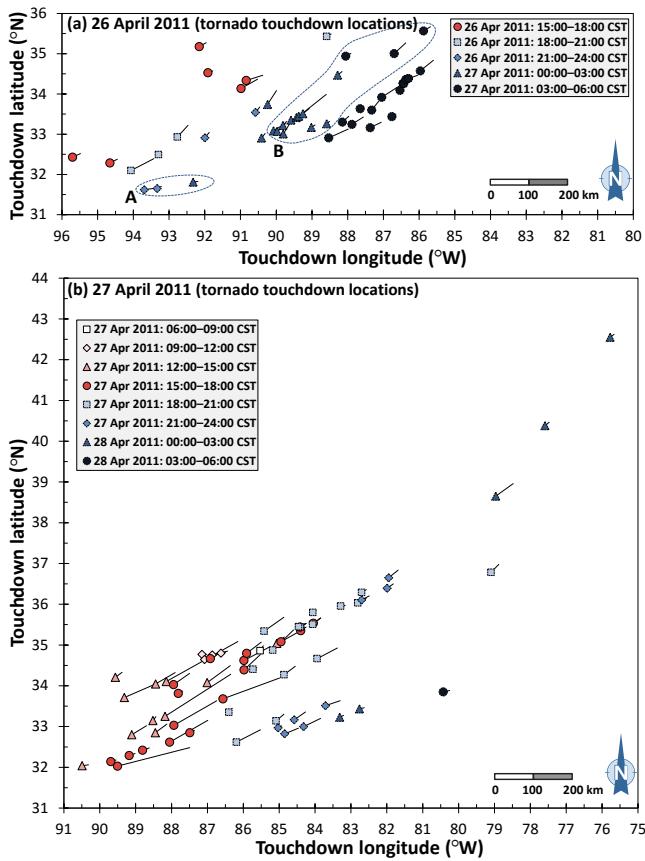


Figure 3. Tornado outbreak on 26–27 April 2011. Touchdown locations of (a) 45 severe ($L \geq 10$ km) tornadoes that occurred on 26 April 2011 (convective day, 06:00–06:00 CST) and (b) 64 severe ($L \geq 10$ km) tornadoes that occurred on 27 April 2011 (convective day). The touchdowns points for each tornado are given by colours and shapes (as given in the legend), representing successive 3 h intervals. The tornado path lengths for each tornado are given by thin black lines. In (a) the tornadoes outlined in the regions A and B will be discussed in a later section. Data were obtained from NOAA (2015).

off for a severe tornado) to 113.3 km (26 April 2011) and 212.4 km (27 April 2011). We will postpone a discussion of the regions A and B that are indicated on Fig. 3a until a later section. In Fig. 3a, although there tends to be a south-west to north-east trend to the 26 April 2011 touchdowns, the spatial distribution appears visually to be diffuse. In Fig. 3b, the south-west to north-east trend of the 27 April 2011 touchdowns is visually less diffuse than in Fig. 3a.

We now turn to our clustering analyses of the two tornado outbreaks on 26 and 27 April 2011. From the times of occurrence given in Fig. 2 and the spatial locations of tornado touchdowns given in Fig. 3a and b, we obtain the temporal and spatial lags using the method illustrated in Fig. 1. In Fig. 4a we give the spatial–temporal lag correlations of all pairs of the 45 severe ($L \geq 10$ km) tornado touchdowns that occurred on 26 April 2011 (convective day). The number of

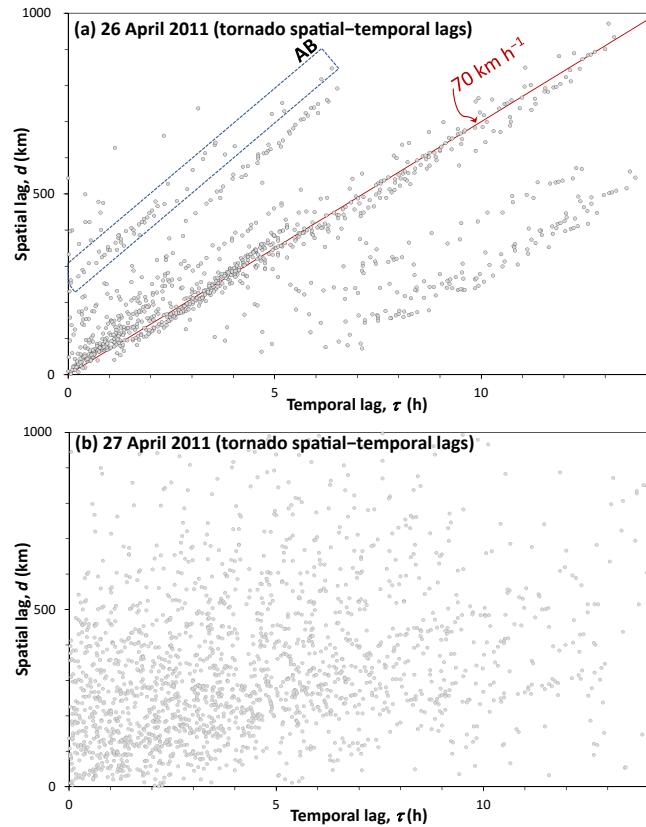


Figure 4. Tornado outbreak on 26–27 April 2011. Spatial–temporal lag correlations between the touchdowns for (a) 45 severe ($L \geq 10$ km) tornadoes that occurred on 26 April 2011 (convective day, 06:00–06:00 CST) and (b) 64 severe ($L \geq 10$ km) tornadoes that occurred on 27 April 2011 (convective day, 06:00–06:00 CST). The spatial lag d is plotted against the temporal lag τ for each of the (a) $N_p = 990$ pairs of tornado touchdowns and (b) $N_p = 2016$ pairs of tornado touchdowns. So that (a) and (b) have the same spatial–temporal limits, 147 (7 %) of the 2016 data points for (b) that have large spatial or temporal values are not included. The data points in Region AB in (a) are correlations between the spatial–temporal lags for the tornadoes in Region A and Region B in Fig. 3a.

pairs are $N_p = 1 + 2 + \dots + (N_T - 1)$, with N_T the number of tornadoes considered. With $N_T = 45$ tornadoes, we have $N_p = 990$ data points on the plot. There are quite clear near-linear trends to the d (spatial lags) vs. τ (temporal lags) data given in Fig. 4a, with the spatial lags increasing with the temporal lags.

Consider the spatial–temporal lag correlation associated with a series of tornadoes generated by a point source at a constant velocity v . The correlation will be a linear trend passing through the origin with the slope giving the velocity v . This behaviour will be demonstrated in some detail in the next section. This behaviour may explain the strong linear trend passing through the origin in Fig. 4a. We fit a straight line to this trend in Fig. 4a and obtain a velocity

$v = 70 \text{ km h}^{-1}$. A possible association is with the south-west to north-east movement of a single cell thunderstorm.

We next turn our attention to one of the near-linear trends in Fig. 4a that does not pass through the origin, indicated by the rectangular region AB. We return to Fig. 3a, where in Region A we outlined a spatial cluster of the touchdowns for three severe tornadoes that occurred on 26 April 2011 and, in Region B, a spatial cluster of the touchdowns for 14 severe tornadoes. In the rectangular region AB, given in Fig. 4a there are 51 data points of which 42 (82 %) represent all of the pairs of tornado touchdowns between the two regions A and B in Fig. 3a, with none of the data points in box AB associated with pairs of tornadoes within Region A or pairs of tornadoes in Region B. We find that this explanation of correlations between tornadoes generated by two separate single cell thunderstorms (the spatial regions A and B in Fig. 3a) provides a similar explanation for the near-linear trends of spatial and temporal lags observed in Fig. 4a.

In Fig. 4b, we give the spatial lag vs. temporal lag for each of the pairs of the 64 severe ($L \geq 10 \text{ km}$) tornado touchdowns that occurred on 27 April 2011 (convective day). In this case, there are $N_p = 2016$ pairs. Comparing Fig. 4b with Fig. 4a, there are striking differences. Specifically, in Fig. 4b, there is no clear near-linear trend of the spatial–temporal lag data, whereas in Fig. 4a, this linear trend both through the origin and in other spatial–temporal lag regions of the plot is dominant. The near-random distribution of data points in Fig. 4b can be associated with the simultaneous generation of tornadoes by several separately defined supercell thunderstorms. The resulting random generation of tornadoes both in space (the several supercells) and time (for each supercell) would lead to a near-random distribution of data points. The thunderstorms on 26 April 2011 are much less likely to be supercellular than those on 27 April 2011 (Knupp et al., 2013).

To further illustrate the relationship between tornadoes and storms, we apply our clustering analysis to severe tornadoes that developed during a tornado outbreak that occurred in the south-east of the USA on 4 April 2011. During this outbreak, an extensive squall line developed along and ahead of a cold front extending from Ohio in a south-westerly direction to Mississippi and Louisiana (Corfidi et al., 2015). The environment proved suitable for the development of thunderstorms within the largely linear convective band (Aon Benfield, 2011). The 4 April 2011 tornado outbreak is recognized as a derecho event (Aon Benfield, 2011; NOAA, 2011; Corfidi et al., 2015), that is, a near-linear squall line dominated by straight-line high winds rather than cyclonic winds dominant in supercell thunderstorms. We consider six severe ($L \geq 10 \text{ km}$) tornadoes that occurred between 13:42 and 18:43 CST. Three severe tornadoes on that day that were spatially distant ($> 600 \text{ km}$ from any of the six tornadoes) were not considered. The touchdown locations and tracks are given in Fig. 5a. In Fig. 5b, the spatial lag d is plotted against the temporal lag τ for each of these 15 pairs of tornado touchdown points, with a good linear correlation found.

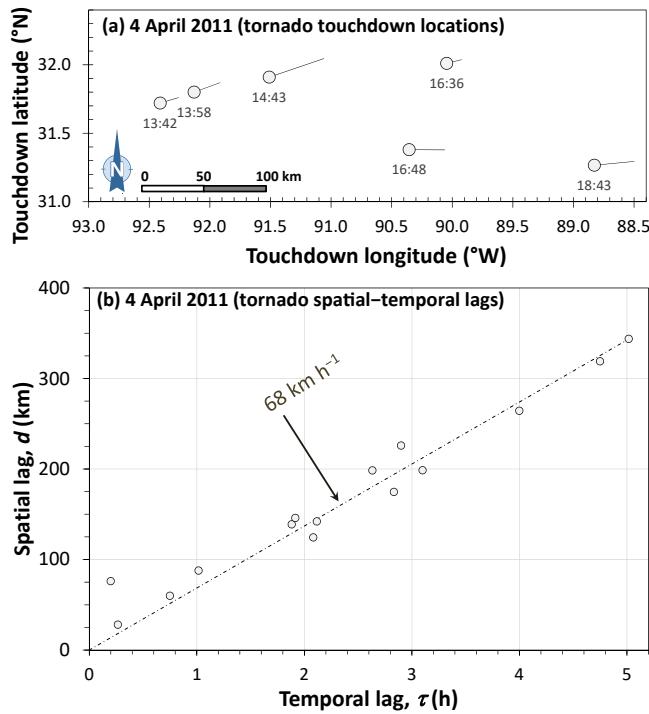


Figure 5. Tornado outbreak on 4 April 2011. (a) Touchdown locations of six severe ($L \geq 10 \text{ km}$) tornado that occurred on 4 April 2011 (convective day, 06:00 CST–06:00 CST) with touchdown times from 13:42 to 18:43 CST. (b) The spatial lag d is plotted against the temporal lag τ for each of the 15 pairs of tornado touchdown points. The straight-line fit to the data passing through the origin gives a velocity $v = 68.5 \text{ km h}^{-1}$.

We compare the data values in Fig. 5b with a least-squares fit to a linear correlation passing through the origin, resulting in a supercell velocity of $v = 68.5 \text{ km h}^{-1}$ (Spearman rank correlation coefficient, $r^2 = 0.92$).

One hypothesis for the 4 April 2011 tornado outbreak is that the tornadoes touched down randomly along the squall line. However, this hypothesis is not consistent with the data in Fig. 5b. Random spatial and random temporal touchdowns produce the random distribution of data points seen in Fig. 4b. The data in Fig. 5 require that the tornadoes are produced at a near-stationary point on the squall line as the squall line migrates at a near-uniform velocity.

As a final application of our cluster analysis, we will consider the Great Plains tornado outbreak of 3 May 1999. On this day, multiple supercell thunderstorms produced many large and damaging tornadoes in central Oklahoma. With additional tornadoes in south-central Kansas and northern Texas, over 70 tornadoes were observed during this event. This outbreak has been discussed in detail by Thompson and Edwards (2000) and is of particular interest to us because a detailed association of each tornado with specific supercells has been given (NOAA, 1999). We consider only the 18 severe ($L \geq 10 \text{ km}$) tornadoes that occurred during this

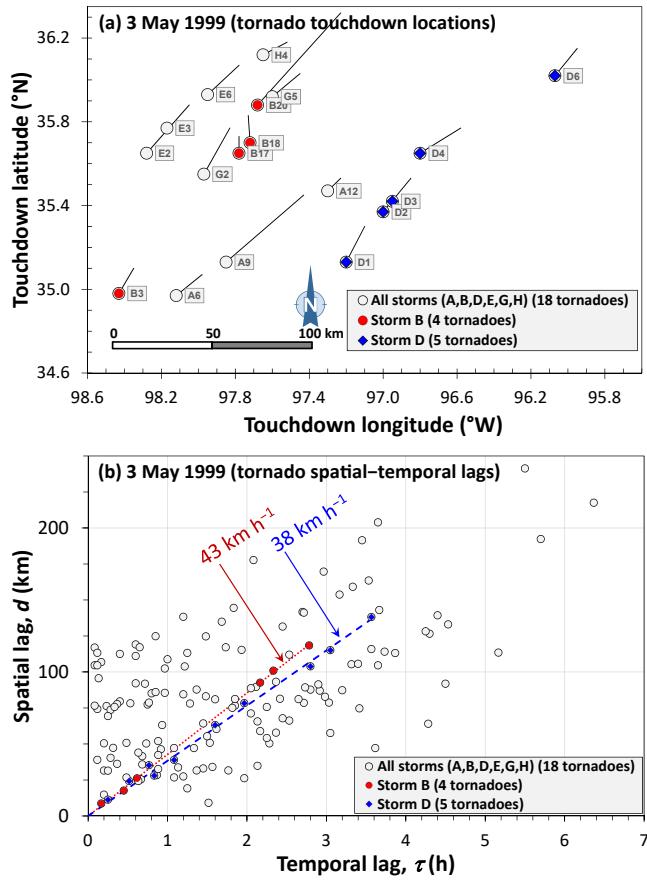


Figure 6. Tornado outbreak on 3 May 1999 (convective day, 06:00–06:00 CST). (a) Touchdown locations of 18 severe ($L \geq 10 \text{ km}$) tornadoes, with storms A, B, D, E, G, H responsible for the tornadoes identified. (b) Spatial–temporal lag correlations between the touchdowns for (i) all 18 tornadoes given in (a), (ii) four tornadoes from Storm B, (iii) five tornadoes from Storm D. Also given are the best-fit lines for spatial–temporal lags for both Storms B and D. Data were obtained from NOAA (1999, 2015).

outbreak. The touchdown locations and tracks are given in Fig. 6a. Each of the tornadoes is associated with one of the six supercell thunderstorms designated A, B, D, E, G, H. The spatial and temporal lags have been obtained for all pairs of the 18 severe tornadoes using the method illustrated in Fig. 1. The results for the 153 spatial–temporal pairs are given in Fig. 6b. There are no clear linear patterns.

We will now focus our attention on the four severe tornadoes associated with supercell B (shown in red in Fig. 6a) and the five tornadoes associated with supercell D (shown in blue in Fig. 6a). In Fig. 6b we designate spatial–temporal lags associated with supercell B (shown in red) and supercell D (shown in blue). Clear linear patterns for the spatial–temporal lags associated with each of the two supercells are obtained. Also included are the best-fit lines for the spatial–temporal lags; for supercell B and D the velocities (slopes) are 43 and 38 km h^{-1} respectively.

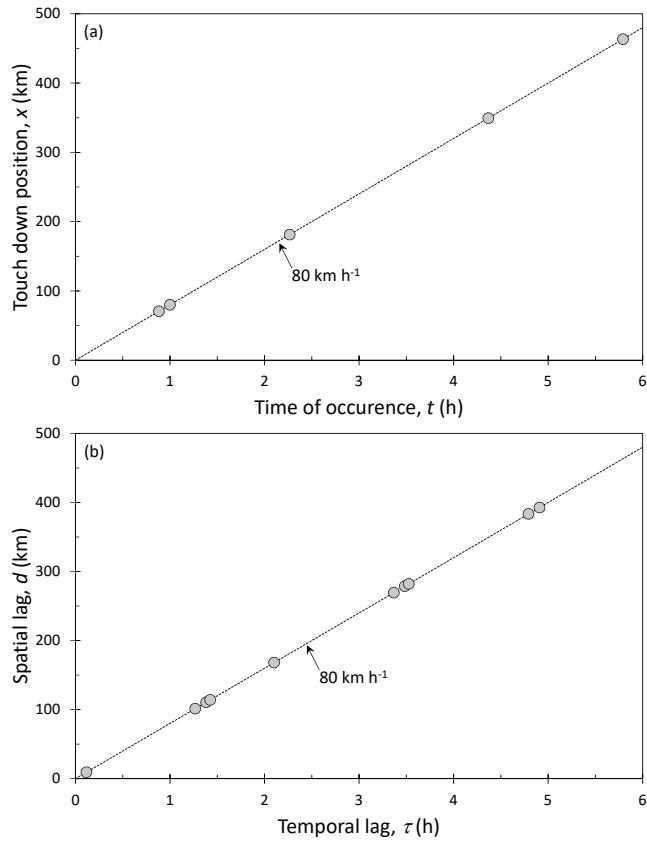


Figure 7. Five model tornado touchdown points located randomly in time during a 6 h time window along a linear track. (a) The touch down positions x along the track are shown as a function of the random times t of occurrence. The model supercell thunderstorm responsible for the tornadoes moves along the track at a velocity $v = 80 \text{ km h}^{-1}$. (b) Spatial–temporal lag correlations between the 5 model tornadoes shown in Fig. 7a. The spatial lag d is plotted against the temporal lag τ for each of the 10 pairs of model tornado touchdown points. The data points again lie on a straight line with a slope of 80 km h^{-1} .

In order to better understand the implications of our spatial–temporal lag correlations, we consider two idealized models. The first is a model for a sequence of tornadoes generated by a single supercell thunderstorm moving at a constant velocity v along a linear track. In this first model, we take a 6 h time window during which the point source moves at a uniform velocity $v = 80 \text{ km h}^{-1}$ along a linear track with a length of 480 km. We assume that five model tornadoes touch down at random times during the 6 h time period. The times t and locations of touchdowns are illustrated in Fig. 7a. The spatial–temporal lag correlations between the model tornado touchdowns in Fig. 7a are shown in Fig. 7b. The spatial lag d is plotted against the temporal lag τ for each of the 10 pairs of tornadoes. The data fall on a straight line that defines the velocity $v = 80 \text{ km h}^{-1}$. We suggest that spatial–temporal lag correlations that fall on or close to a straight line going

through the origin are indicative of a progressive tornado outbreak, possibly from a single supercell thunderstorm.

Confirmation of this behaviour has been obtained in our treatment of the 3 May 1999 outbreak. Severe tornadoes previously associated with single supercell thunderstorms generate spatial–temporal lag correlations that are well approximated by straight lines as illustrated in Fig. 6b. A similar linear correlation was shown for the six severe tornadoes we studied from the 4 April 2011 outbreak illustrated in Fig. 5b. We also suggest that the strong linear trends seen in the spatial–temporal correlation data for the 26 April 2011 outbreak (Fig. 4a) may be associated with tornadoes generated by one or more single cell storms.

In order to further address the large difference in spatial–temporal correlations in the data illustrated in Fig. 4, we consider a second model, more complex than the one just given. In our second idealized model, we consider a quasi-linear vertical (north–south, y) “squall” line moving to the east at constant velocity $v = 80 \text{ km h}^{-1}$ over an $800 \text{ km} \times 800 \text{ km}$ region and a 10 h period (for tornado touchdowns). Tornadic cells are distributed along the near-linear squall line with an approximate spacing Δy . Tornadoes are assumed to touch down at equally spaced time intervals (plus some noise ε that we introduce) $\Delta t + \varepsilon$. The ratio $\Delta y/\Delta t$ defines a characteristic velocity. Our hypothesis is that the non-dimensional velocity ratio $B = (\Delta y)/(v\Delta t)$ defines the behaviour of the system. If B is large ($B > 1$), quasi-linear behaviour is observed in the spatial–temporal lag domain. If B is small ($B < 1$), quasi-random behaviour is observed in the spatial–temporal lag domain. In Figs. 8 and 9, we give two model scenario examples. In Fig. 8a we consider four tornadic cells (each cell represented by a horizontal set of 10 circles, top to bottom) for which the vertical spacing between tornadic cells is $\Delta y \approx (800 \text{ km})/4 = 200 \text{ km}$, and there are 10 tornado touchdowns from each cell so that $\Delta t \approx 1 \text{ h}$ with touchdown times indicated below each circle. The timing of the first tornado touchdown for each tornadic cell is chosen randomly within the first 1 h $[(10 \text{ h})/(10 \text{ tornadoes})]$ plus some noise ε and spatially along the tornadic cell such that the first tornado touchdown occurs horizontally within the first 80 km ($800 \text{ km} 10 \text{ tornadoes}^{-1}$). The non-dimensional parameter $B = (\Delta y)/(v\Delta t) = (200 \text{ km})/[(800 \text{ km})/(10 \text{ tornadoes})] = 2.5$. In Fig. 8b we consider 10 tornadic cells so that $\Delta y \approx (800 \text{ km})/10 = 80 \text{ km}$ and four tornado touchdowns from each cell so that $\Delta t \approx 2.5 \text{ h}$ with the first tornado touchdown for each tornadic cell occurring randomly within the first 2.5 h $[(10 \text{ h})/(4 \text{ tornadoes})]$ and horizontally within the first 200 km $[(800 \text{ km})/(4 \text{ tornadoes})]$. The non-dimensional parameter $B = (\Delta y)/(v\Delta t) = (80 \text{ km})/[(80 \text{ km h}^{-1})(2.5 \text{ h})] = 0.4$. In Fig. 8 we give the tornado touchdown locations and times for both model scenarios 1 and 2, and in Fig. 9 the spatial–temporal lag results. For scenario 1 in Fig. 9a, with $B = 2.5$, we obtain quasi-linear behaviour in the spatial–temporal lag

domain, which we consider further. The four tornadic lines from top to bottom in Fig. 8a are referred to as A ($10 \text{ tornadoes along } y = 725 \text{ km}$), B ($y = 475 \text{ km}$), C ($y = 300 \text{ km}$) and D ($y = 90 \text{ km}$). In Fig. 9a, the spatial–temporal lag domain, the linear correlation passing through the origin results from lags within each of the four tornadic lines A, B, C and D. The next higher set of correlations in the spatial–temporal lag domain in Fig. 9a (starting at about $d = 200 \text{ km}$ and $\tau = 0 \text{ h}$) is a set of three lines adjacent to each other, which are the result of correlations between tornadic lines A and B, B and C and C and D. Similarly, the two sets of adjacent lines in Fig. 9a (starting at about $d = 400 \text{ km}$ and $\tau = 0 \text{ h}$) are the results of correlations between tornadic lines A and C and B and D. Finally, the single line in Fig. 9a (starting at about $d = 600 \text{ km}$ and $\tau = 0 \text{ h}$) is the result of correlation between the tornadic lines A and D in Fig. 8a. The model we consider is idealized, but we believe it illustrates conditions favourable for linear features (i.e. 26 April 2011) vs. more random features (i.e. 27 April 2011).

We now return to a discussion of the well-defined linear trends in the spatial–temporal correlation given in Fig. 4a. The first linear trend, extending from the origin with a slope corresponding to $v = 70 \text{ km h}^{-1}$, can be explained as we explained the similar linear trends in Figs. 5–7. For the second linear trend within the box AB of Fig. 4a, we determined that these points were the result of spatial–temporal correlations between the tornadoes in boxes A and B in Fig. 3a. Most of the data points (82 %) in box AB in Fig. 4a were the result of spatial–temporal lag correlations between boxes A and B in Fig. 3a. The approximately 300 km vertical offset distance at zero time lag in Fig. 4a between the origin and box AB is approximately the distance between the nearest touchdown locations between Region A and Region B in Fig. 3a. We attribute the curvature of secondary correlations in Fig. 9a to the initiation of the linear tracks in Fig. 8a at short time intervals. If the initiation of the tracks were offset for relatively large times, then straighter correlations would be expected.

We next introduce a measure of the combined spatial–temporal separation of pairs of tornado touchdowns, for which the spatial–temporal separation ψ is given by the following:

$$\psi = \tau + \frac{d}{v_c}, \quad (1)$$

where, as previously, τ and d respectively represent the temporal and spatial lags between the tornado touchdowns, and v the average supercell velocity, which we take here to be $v = 80 \text{ km h}^{-1}$. Small values of both temporal and spatial lag result in small values of the spatial–temporal separation. For example, if the lags between two tornado touchdowns are $\tau = 2 \text{ h}$ and $d = 160 \text{ km}$, then the spatial–temporal separation $\psi = (2 \text{ h}) + (160 \text{ km})/(80 \text{ km h}^{-1}) = 4 \text{ h}$. We consider the statistical distribution of the values of ψ by introducing

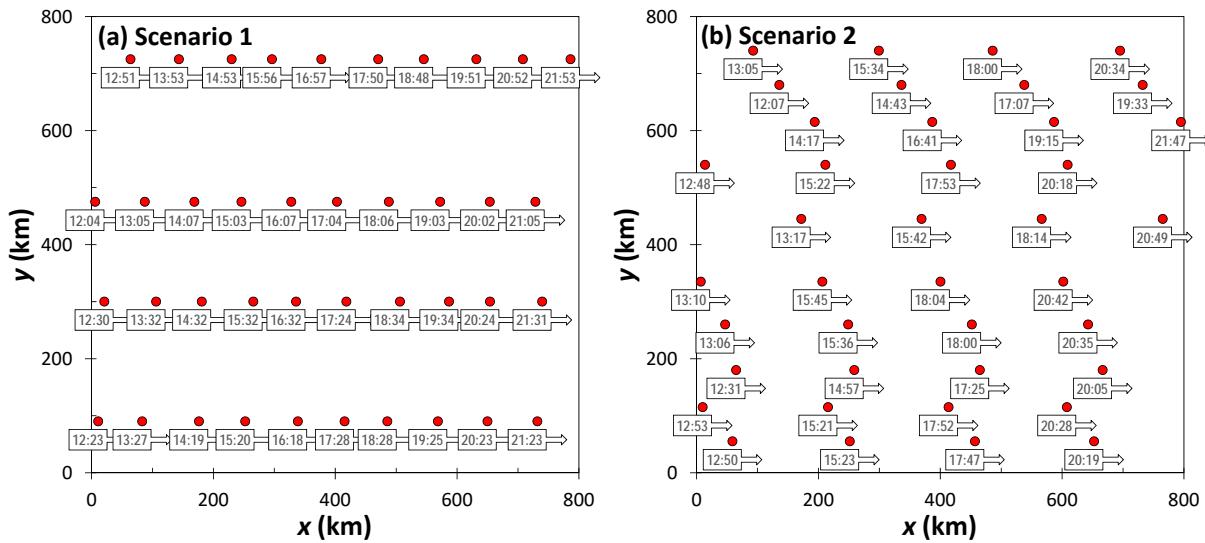


Figure 8. Two model scenarios for 40 tornadoes in an $800 \text{ km} \times 800 \text{ km}$ region over a time period of 10 h. (a) Four parallel supercells moving at about 80 km h^{-1} with 10 tornadoes each. (b) Ten parallel supercells moving at about 80 km h^{-1} with 4 tornadoes each.

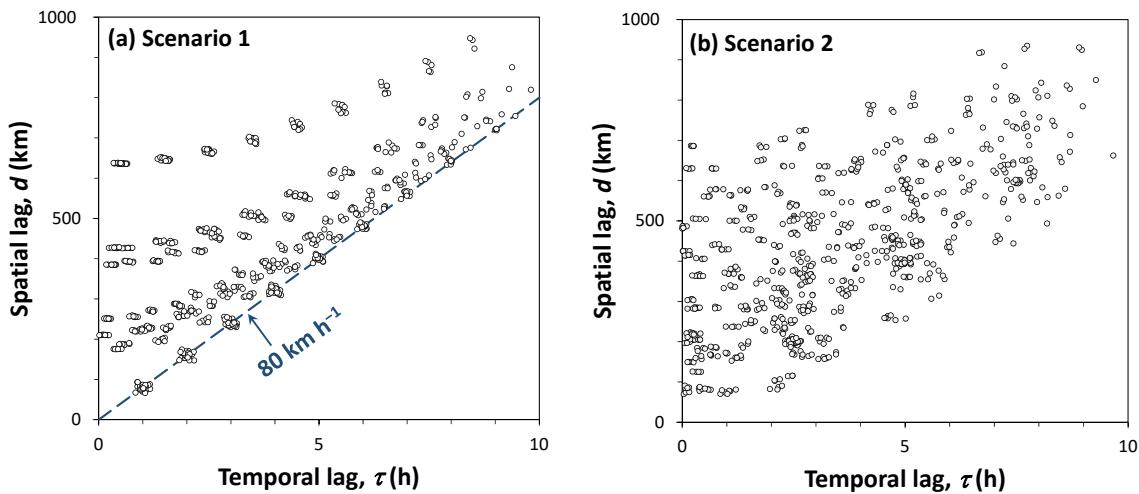


Figure 9. Spatial–temporal lag diagrams for the two model scenarios given in Fig. 8.

the normalized cumulative probability, defined as follows:

$$P(<\psi) = \frac{N_C(<\psi)}{N_T}, \quad (2)$$

with $N_C(<\psi)$ the number of tornado touchdown pairs with spatial–temporal separation values less than ψ and N_T the total number of pairs considered.

In Fig. 10 we plot the normalized cumulative probability $P(<\psi)$ as a function of the spatial–temporal separations ψ . We consider the data for the two tornado outbreaks in the USA on 26 and 27 April 2011 (convective days) and utilize the values given in Fig. 3 for spatial–temporal separations $\psi < 4 \text{ h}$. We have not considered data for the 4 April 2011 outbreak given in Fig. 5, because of the very small number of data points.

We see that the sets of normalized cumulative probability values for the two outbreaks given in Fig. 10 have a very different pattern, one linear and the other a power law. For the 26 April 2011, our data set consisted of 45 severe tornadoes resulting in $N_P = 990$ pairs of tornado touchdowns of which 245 spatial–temporal separations are illustrated in Fig. 10. The least-squares best-fit linear correlation for the spatial–temporal separations for 26 April 2011, over the range $0.0 < \psi < 4.0 \text{ h}$, gives the following:

$$P(<\psi) = 0.0671\psi - 0.0241, \quad (3)$$

which is in excellent agreement with the data in the range $0.6 < \psi < 4.0 \text{ h}$. For the 27 April 2011 outbreak, our data set consisted of 64 severe tornadoes resulting in $N_P = 2016$ pairs of tornado touchdowns of which 330 spatial–temporal sep-

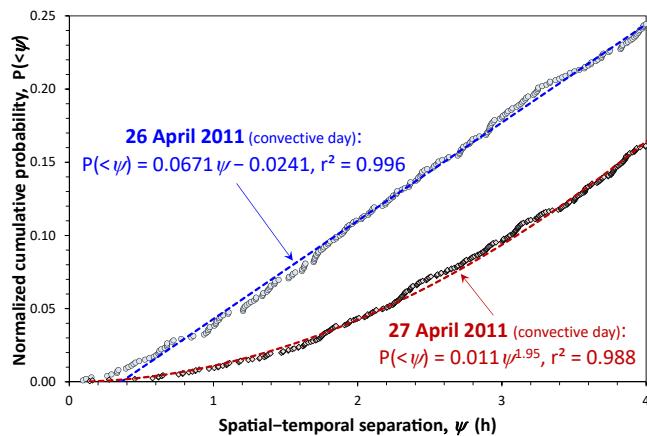


Figure 10. Tornado outbreak on 26–27 April 2011. Normalized cumulative probability $P(<\psi)$ of spatial–temporal separations ψ between pairs of severe tornado (path length $L \geq 10$ km) touchdowns during tornado outbreaks in the USA on 26 and 27 April 2011 (convective days; see Figs. 2 and 3). Cumulative probabilities are given for spatial–temporal separations $\psi < 4$ h. The least-square best-fit line (blue dashed line) and power law (red dashed line) to the data are shown in this figure for 26 and 27 April 2011 respectively. Data were obtained from NOAA (2015).

arations are illustrated in Fig. 7. The least-squares best-fit power-law correlation for the spatial–temporal separations for 27 April 2011, over the range $0.0 < \psi < 4.0$ h, gives the following:

$$P(<\psi) = 0.011\psi^{1.95}, \quad (4)$$

which is in excellent agreement with the data and has an exponent close to 2.

We now give an explanation for the linear and power-law correlations that we have found. If the tornado touchdowns occur randomly along a path for relatively small values of spatial–temporal separations ψ , then a linear correlation of normalized cumulative probability $P(<\psi)$ with spatial–temporal separation ψ is expected to be a good approximation. In contrast, if the tornado touchdowns occur randomly in both space and time, then it is expected that $P(<\psi)$ is proportional to ψ^2 , i.e. the area of the segment of a circle of possible touchdown locations. The transition from linear to random behaviour indicated by the data in Fig. 10 is consistent with our previous qualitative discussion of the data given in Fig. 4.

3 Discussion

Unlike many other natural hazards, it is difficult to quantify strong tornadoes precisely. For hurricanes, there are extensive data on wind speeds and barometric pressures along the path of the storm. For floods, flood gauges provide a quantitative measure of the flow rate in a river. For earthquakes,

seismographs give measures of shaking intensity. Quantifying volcanic eruptions and landslides is more difficult but volumes of material involved can be estimated. It is not possible to reliably measure the wind speeds or pressure changes in tornadoes. The standard measure of tornado intensity used today is the Enhanced Fujita scale. This scale is based qualitatively on the damage caused by a tornado. An alternative measure of tornado intensity is the tornado path length L .

Malamud and Turcotte (2012) showed that records of tornado path lengths from the 1990s to the present appear to be relatively complete for severe tornadoes (defined to be $L \geq 10$ km) in the United States. They also showed that the number-length scaling of severe tornado touchdowns is well approximated by a power-law distribution. Elsner et al. (2014) showed that the distribution of daily tornado counts in the United States is also well approximated by a power-law relationship. Malamud and Turcotte (2012) also studied the statistics of recent severe tornado outbreaks. They quantified the strength of a severe tornado outbreak to be the total tornado path length L_D of the severe tornadoes occurring during a convective day. They showed that the number-length scaling of severe tornado outbreaks is also well approximated by a power-law distribution.

Another important aspect of tornado outbreaks is the distribution of touchdown points in space and time. In terms of expectations for these data, there are two limiting cases.

- i. Tornadoes occur randomly in space and time during a specified spatial region and time interval. In this case the touchdown points will be randomly distributed in space by interacting supercell thunderstorms.
- ii. Tornadoes are generated by a single cell thunderstorm moving on a near-linear path at a constant velocity. In this case the touchdown points will approximately be on a linear path.

Actual tornado outbreaks will generally be a complex combination of these two limiting cases.

The statistics of the touchdown points of a tornado outbreak can certainly be studied using a spatial map of the touchdown points. However, this does not directly incorporate the time of the touchdowns. In this paper, we have considered an alternative statistical measure for tornado touchdowns by applying a spatial–temporal clustering analysis originally developed by Zaliapin et al. (2008) for earthquakes. The sequence of severe tornado touchdowns occurring during a convective day is considered to be a sequence of point events in space and time. All pairs of these point events are considered and a plot produced of the spatial lag d (i.e. spatial distance between the touchdown points for a pair of events) vs. the temporal lag τ (difference in touchdown times between the same pair of events).

A principal focus of this paper is the application of a clustering analysis to several observed tornado outbreaks. It is expected that a small outbreak of severe tornadoes in a given

convective day could be associated with tornadoes generated randomly along a linear squall line progressing at a near-constant velocity. In this case the tornado touchdowns occur randomly both for space and time, and the cluster plot of d vs. τ would also be random. Alternatively, the tornadoes could be generated by a single cell thunderstorm moving in a near-linear path at a near-constant velocity. We have shown in Fig. 7 that in this case the points in a d vs. τ plot lie approximately on a straight line through the origin, with the slope equal to the velocity of the thunderstorm. As a specific example, we considered six severe tornado touchdowns associated with the 4 April 2011 derecho event in the south-eastern USA. The severe tornadoes could have been generated randomly along the squall line. However, in Fig. 5b, we see that the d vs. τ data points lie approximately on a straight line through the origin with a slope of $v = 68 \text{ km h}^{-1}$. This suggests that these six tornadoes were generated by a single large thunderstorm or several closely spaced thunderstorms moving at a velocity of about 68 km h^{-1} .

To further illustrate the applicability of our clustering analysis to severe tornado touchdowns, we considered the Great Plains tornado outbreak of 3 May 1999. Careful studies have associated individual tornadoes in the outbreak with specific supercell thunderstorms as shown in Fig. 6b. When all 18 severe tornadoes are considered the data are quite randomly distributed. However, when two sets of tornadoes are considered that are associated with two supercell thunderstorms, clear linear patterns are obtained with slopes of 43 and 38 km h^{-1} .

We also applied our clustering analysis to the intense tornado outbreaks in the central United States on 26 and 27 April 2011, with 45 and 64 severe tornadoes occurring respectively (convective days) and more than 300 fatalities. For each pair of tornadoes on the two separate days, the severe tornado touchdown spatial lags are given as a function of their temporal lags in Fig. 4. The observed patterns are very different. The results for 26 April 2011 (convective day) in Fig. 4a are dominated by a complex sequence of linear tracks that we have previously discussed. Knupp et al. (2013) suggest that this 26 April outbreak of tornadoes was associated with a quasi-linear convective squall line. The pattern seen in Fig. 4a has similarities to that seen in Fig. 5b but is clearly more complex. We suggest that on 26 April 2011 groups of these tornadoes were associated with one large thunderstorm or several closely spaced thunderstorms but there was a small number of groups that generated the complexity. This pattern is consistent with the movement of a discrete set of thunderstorms moving from the south-west to the north-east at velocities near 70 km h^{-1} . The observed pattern for 27 April 2011 (convective day) given in Fig. 4b is quite different. It is predominantly random with virtually no embedded linear patterns. We suggest that this is due to a relatively large number of supercell thunderstorms generating tornadoes randomly in space and time.

In order to better understand the roles of supercell thunderstorms in generating random and linear patterns in our spatial-temporal lag diagrams, we studied two model scenarios for tornado generation. In the first model scenario (Fig. 8a), four supercell thunderstorms originating in a squall line each generated 10 tornadoes randomly. In the second model scenario (Fig. 8b), 10 supercell thunderstorms originating in a squall line each generated four tornadoes randomly. The corresponding spatial-temporal lag diagrams for these two model scenarios are given in Fig. 9. The first scenario generated linear-type features; the second appeared random.

Although there are no physical processes directly in these two model scenarios, the statistical processes represent a variety of scales of processes that are important in tornado outbreaks. In general, the synoptic scale provides the background that leads to convection over a broad area (e.g. Knupp et al., 2014). The spacing between storms and the timing of initiation depends upon relationships between the synoptic-scale and smaller-scale features. Lilly (1979), Bluestein and Weisman (2000) and Lee et al. (2006) modelled the complexity of behaviour of storms that were initiated along lines; interaction included both constructive and destructive ones that can lead to the characteristic spacing associated with a particular event. Within a single supercell itself, the distance in time and space for repeated tornado genesis is a function of the storm motion (related to the large-scale environment in which the storm forms) and within-storm processes that lead to the distribution of precipitation and temperature leading to the birth and death of rotation features in the storm (Burgess et al., 1982; Alderman et al., 1999). For a particular tornado outbreak, the exact details depend upon the full range of atmospheric processes. Confidence is greatest in the understanding that certain large-scale environments are more likely to lead to outbreaks occurring, with details of individual storm occurrence and within-storm features becoming increasingly less certain.

We suggest that it may be possible to generate a large number of different model scenarios of this type, with corresponding spatial-temporal diagrams and compare them to spatial-temporal diagrams of observations in a semi-automated way. The objective would be to take the type of observed data illustrated in Fig. 6a and determine association of tornadoes with postulated supercell thunderstorm tracks semi-automatically. This application has the potential to provide constraints on simulated tornado outbreaks that are made by insurance and reinsurance modellers to create scenarios estimating risk of property loss. More realistic portrayals of tornado outbreaks could be important for the setting of optimal rates.

4 Data availability

NOAA (National Oceanic and Atmospheric Administration) Storm Prediction Centre (SPC), Tornado, Hail, and Wind Database, available at: www.spc.noaa.gov/wcm/.

Acknowledgements. The authors thank J. Elsner and one anonymous referee for their helpful and constructive suggestions.

Edited by: R. Trigo

Reviewed by: J. Elsner, H. Brooks¹, and one anonymous referee

References

- Adlerman, E. J., Droege, K. K., and Davies-Jones, R.: A numerical simulation of cyclic mesocyclogenesis, *J. Atmos. Sci.*, 56, 2045–2069, doi:10.1175/1520-0469(1999)056<2045:ANSOCM>2.0.CO;2, 1999.
- Aon Benfield: United States April & May 2011 Severe Weather Outbreaks, Impact Forecasting, Aon Benfield (Chicago, USA) report, available at: http://www.aon.com/attachments/reinsurance/201106_us_april_may_severe_weather_outbreaks_recap.pdf (last access: 22 June 2016), 2011.
- Bluestein, H. B. and Weisman, M. L.: The interaction of numerically simulated supercells initiated along lines, *Mon. Weather Rev.*, 128, 3128–3149, doi:10.1175/1520-0493(2000)128<3128:TIONSS>2.0.CO;2, 2000.
- Brooks, H. E.: On the relationship of tornado path length and width to intensity, *Weather Forecast.*, 19, 310–319, doi:10.1175/1520-0434(2004)019<0310:OTROTP>2.0.CO;2, 2004.
- Burgess, D. W., Wood, V. T., and Brown, R. A.: Mesocyclone evolution statistics. 12th Conf. on Severe Local Storms, San Antonio, TX, Amer. Meteor. Soc., 422–424, 1982.
- Corfidi, S. F., Coniglio, M. C., Cohen, A. E., and Mead, C. M.: A proposed revision to the definition of “derecho”, *B. Am. Meteorol. Soc.*, 97, 935–949, doi:10.1175/BAMS-D-14-00254.1, 2015.
- Davies-Jones, R., Trapp, R. J., and Bluestein, H. B.: Tornadoes and tornadic storms, in: *Severe Convective Storms*, edited by: Doswell III, C. A., Meteorological Monographs, 167–221, Am. Meteorol. Soc., 2001.
- Doswell III, C. A. and Burgess, D. W. Tornadoes and tornadic storms: A review of conceptual models, in: *The Tornado: Its Structure, Dynamics, Prediction, and Hazards*, edited by: Church, C., Burgess, D., Doswell III, C., and Davies-Jone, R., *Geophys. Monogr.*, 79, 161–172, American Geophysical Union, Washington, D.C., 1993.
- Doswell III, C. A., Edwards, R., Thompson, R. L., Hart, J. A., and Crosbie, K. C.: A simple and flexible method for ranking severe weather events. *Weather Forecast.*, 21, 939–951, doi:10.1175/WAF959.1, 2006.
- Doswell III, C. A., Brooks, H. E., and Dotzek, N.: On the implementation of the enhanced Fujita scale in the USA, *Atmos. Res.*, 93, 554–563, doi:10.1016/j.atmosres.2008.11.003, 2009.
- Doswell III, C. A., Carbin, G. W., and Brooks, H. E.: The tornadoes of spring 2011 in the USA: An historical perspective, *Weather*, 67, 88–94, doi:10.1002/wea.1902, 2012.
- Edwards, R., LaDue, J. G., Ferree, J. T., Scharfenberg, K., Maier, C., and Coulbourne, W. L.: Tornado intensity estimation: Past, present, and future, *B. Am. Meteorol. Soc.*, 94, 641–653, doi:10.1175/BAMS-D-11-00006.1, 2013.
- Elsner, J. B., Jagger, T. H., Widen, H. M., and Chavas, D. R.: Daily tornado frequency distributions in the United States, *Environ. Res. Lett.*, 9, 024018, doi:10.1088/1748-9326/9/2/024018, 2014.
- Elsner, J. B., Elsner, S. C., and Jagger, T. H.: The increasing efficiency of tornado days in the United States, *Clim. Dynam.*, 45, 651–659, doi:10.1007/s00382-014-2277-3, 2015.
- Fuhrmann, C. M., Konrad, C. E., Kovach, M. M., McLeod, J. T., Schmitz, W. G., and Dixon, P. G.: Ranking of tornado outbreaks across the United States and their climatological characteristics, *Weather Forecast.*, 29, 684–701, doi:10.1175/WAF-D-13-00128.1, 2014.
- Galway, J. G.: Some climatological aspects of tornado outbreaks, *Mon. Weather Rev.*, 105, 477–484, doi:10.1175/1520-0493(1977)105<0477:SCAOTO>2.0.CO;2, 1977.
- Glickman, T. S. (Ed.): *Glossary of Meteorology*, 2nd Edn., American Meteorological Society, 782 pp., 2000.
- Knupp, K. R., Murphy, T. A., Coleman, T. A., Wade, R. A., Mullins, S. A., Schultz, C. J., Schultz, E. V., Carey, L., Sherrer, A., McCaul Jr., E. W., Carcione, B., Latimer, S., Kula, A., Laws, K., Marsh, P. T., and Klockow, K.: Meteorological overview of the devastating 27 April 2011 tornado outbreak, *B. Am. Meteorol. Soc.*, 95, 1041–1062, doi:10.1175/BAMS-D-11-00229.1, 2013.
- Lee, B. D., Jewett, B. F., and Wilhelmson, R. B.: The 19 April 1996 Illinois tornado outbreak. Part I: Cell evolution and supercell isolation, *Weather Forecast.*, 21, 433–448, doi:10.1175/WAF944.1, 2006.
- Lilly, D. K.: The dynamical structure and evolution of thunderstorms and squall lines, *Annu. Rev. Earth Planet. Sc.*, 7, 117–161, 1979.
- Malamud, B. D. and Turcotte, D. L.: Statistics of severe tornadoes and severe tornado outbreaks, *Atmos. Chem. Phys.*, 12, 8459–8473, doi:10.5194/acp-12-8459-2012, 2012.
- Mercer, A. E., Shafer, C. M., Doswell III, C. A., Leslie, L. M., and Richman, M. B.: Objective classification of tornadic and nontornadic severe weather outbreaks, *Mon. Weather Rev.*, 137, 4355–4368, doi:10.1175/2009MWR2897.1, 2009.
- NOAA (National Oceanic and Atmospheric Administration): Normal Oklahoma National Weather Service Weather Forecast Office, The Great Plains tornado outbreak of May 3–4, 1999, available at: <https://www.weather.gov/oun/events-19990503> (last access: 1 July 2016), 1999.
- NOAA (National Oceanic and Atmospheric Administration): The Southeast US “Derecho” of April 4–5, available at: www.spc.noaa.gov/misc/AbtDerechos/casepages/apr042011page.htm (last access: 1 April 2016), 2011.
- NOAA (National Oceanic and Atmospheric Administration): Storm Prediction Centre (SPC), Tornado, Hail, and Wind Database, available at: www.spc.noaa.gov/wcm/, last access: 24 May 2015.
- Thompson, R. L. and Edwards, R.: An overview of environmental conditions and forecast implications of the 3 May 1999 tornado outbreak, *Weather Forecast.*, 15, 682–699, doi:10.1175/1520-0434(2000)015<0682:AOOECA>2.0.CO;2, 2000.

¹H. Brooks reviewed the discussion paper and was added as a co-author during the revision.

Utsu, T.: Aftershocks and earthquake statistics (1): Some parameters which characterize an aftershock sequence and their interrelations. *J. of the Faculty of Science, Hokkaido University, Series 7 (Geophysics)*, 3, 129–195, 1970.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H: Clustering analysis of seismicity and aftershock identification, *Phys. Rev. Lett.*, 101, 018501, doi:10.1103/PhysRevLett.101.018501, 2008.