

A Spatial and Intensity-Based Tornado Outbreak Index

Adonte Knight

ak1817@msstate.edu

Mississippi State University

Andrew Mercer

Mississippi State University

Article

Keywords:

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11159032/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

1 **A Spatial and Intensity-Based Tornado Outbreak Index**

2
3
4 Adonte N. Knight

5 *Applied Research Collaboratory, Mississippi State University, Starkville, MS, USA*

6 *Department of Geosciences, Mississippi State University, Starkville, MS, USA*

7 Andrew E. Mercer

8 *Department of Geosciences, Mississippi State University, Starkville, MS, USA*

9
10 ,

11 Submitted to npj Natural Hazards

12 20 September 2026

13
14
15
16
17
18
19
20

21 Corresponding author address: Adonte Knight, 2 Research Boulevard, Starkville, MS 39759

22 Email: ak1817@msstate.edu

23 Contributing author email: mercera@ngi.msstate.edu

ABSTRACT

Historic tornado outbreak (TO) intensity metrics traditionally emphasize counts on discrete tornado days, providing limited spatial context. This study introduces a spatially aware tornado outbreak intensity index for TOs occurring east of the Rocky Mountains between 1960 and 2024. Kernel density estimation (KDE) is used to identify spatially coherent TO regions, while detrended tornado counts and Adjusted Fujita miles are combined within each KDE region to quantify outbreak intensity and spatial extent. The resulting TOs are grouped into four intensity categories and further subclassified using k-means clustering. The clustering identifies distinct outbreak modes, including events dominated by weak tornadoes, mixed-intensity outbreaks, and extreme outbreaks characterized by numerous strong and violent tornadoes. The most intense TOs, including the 1974 and 2011 Super Outbreaks, form a distinct cluster, highlighting historical rarity. The proposed index provides an objective framework for tornado outbreak classification and applications in outbreak climatology and hazard assessment.

1. Introduction

Tornado outbreaks (hereafter TOs), defined loosely as “multiple tornado occurrences within a single synoptic-scale system” (Glickman 2000), are one of the most impactful atmospheric hazards, despite remaining understudied owing to the lack of a consistent, specific definition (Cwik et al. 2021). The American Meteorological Society definition (Glickman 2000), as well as other TO definitions offered in Cwik et al. (2021), offers insufficient information regarding tornado counts, spatial extent, or intensity, making it difficult to differentiate TOs from other tornado-producing events. A consistent framework for quantifying TO intensity would facilitate forecast evaluation and improve understanding of meteorological factors governing TOs of varying intensity.

Early TO studies defined outbreaks primarily using tornado counts with measures of spatial extent. Beebe (1956) defined a TO as three or more tornadoes, with at least two separated by 100 miles or more. Galway (1975), by contrast, characterized a TO (their “family outbreak”) as five or more tornadoes associated with a weather system on a given day. More recent work utilized tornado metrics derived from tornado report databases to quantify TO intensity. Schneider et al. (2004) focused on tornado days (local midnight to local midnight) using counts of significant to violent tornadoes. Edwards et al. (2004) quantified outbreak intensity using outbreak days (defined as 1200 UTC – 1159 UTC the next day). They incorporated

tornado counts (including significant and violent tornado counts), destructive potential index (DPI; Thompson and Vescio 1998), tornado path length, deaths, killer tornadoes, and the number of tornado tracks exceeding 80 km. Their work introduced important questions regarding tornado spatial scatter and grouping and conceptualized the idea of outbreak regions. Forbes (2006) first attempted to address tornado spatial scatter within a TO by considering reports in bordering states within 6 hours of each other. This approach was limited by artificial state boundaries and varying state sizes, but it was a critical first effort to identify the spatial extent of tornado reports within a TO.

Later TO intensity work built heavily on the foundation established in Edwards et al. (2004). Doswell et al. (2006) defined two severe weather outbreak intensity indices, one for TOs and one for non-tornadic outbreaks, based on 14 outbreak characteristics, including those proposed by Edwards et al. (2004) as well as a density parameter and a newly defined middle 50% parameter aimed at quantifying TO spatial spread. They also observed that the Fujita scale rating along each tornado path likely overestimated the tornado's intensity and that many of their tornado diagnostics exhibited large secular trends due to changes in population density, storm surveys, and reporting practices (Weiss et al. 2002; Verbout et al. 2006; McCarthy 2003; Schaefer et al. 2002). They addressed the latter by scaling each tornado diagnostic by a log-linear transformed annual mean that removed the secular trend, yielding standardized indices similar to those of Edwards et al. (2004).

Shafer and Doswell expanded the work of Doswell et al. (2006) through a series of studies addressing report clustering, tornado density, and outbreak duration. Shafer and Doswell (2011 – hereafter SD11) used kernel density estimation (KDE- Wilks 2019) regions to delineate objective spatial boundaries on outbreak regions from severe weather report density. Shafer and Doswell (2012) expanded this approach, using spatio-temporally overlapping KDE regions to identify multi-day outbreaks. In contrast, Fuhrmann et al. (2014; F14) focused primarily on outbreak intensity rather than spatial and temporal delineation. They combined tornado path length with F/EF-scale rating to develop the adjusted Fujita miles metric, a measure of accumulated outbreak intensity. These two approaches (SD11 and F14) offer a complementary

framework for establishing TO intensity by establishing outbreak regions and quantifying outbreak intensity.

Despite these advances, critical issues remain:

- 1) Existing detrending approaches in both studies were developed assuming monotonic temporal growth in tornado reports. However, recent decades exhibit substantial deviations from the assumptions presented in SD11 and F14 (Fig. 1), including a flattening of weak (EF0-EF1) tornado report counts and persistent discontinuities associated with the implementation of the Fujita scale in 1974 and the Enhanced Fujita Scale in 2007 (Fujita 1974, 1975; Agee & Childs 2014).
- 2) Although SD11 and F14 provide complementary perspectives on TO characterization, each possesses methodological limitations that may be mitigated through a unified framework.

The primary objective of this study is to leverage the strengths of the methodologies presented in SD11 and F14 to obtain an updated measure of TO intensity that characterizes both the intensity and spatial extent of the outbreak. This index also needs to be straightforward to update as new data become available. To accomplish this, we update existing detrending methods using a piecewise detrending methodology associated with known historic breakpoints in the tornado record that better fits the report databases used to quantify TO intensity. This method is combined with the existing KDE methods outlined in SD11 to present a spatially aware and impact-based index characterizing outbreak intensity. This intensity index will be used to categorize TOs into discrete intensity groups, enabling identification of characteristic outbreak features across outbreaks of varying strength. Ultimately, this index will be able to serve as a cohesive basis for quantifying TO impacts and intensity in future research.

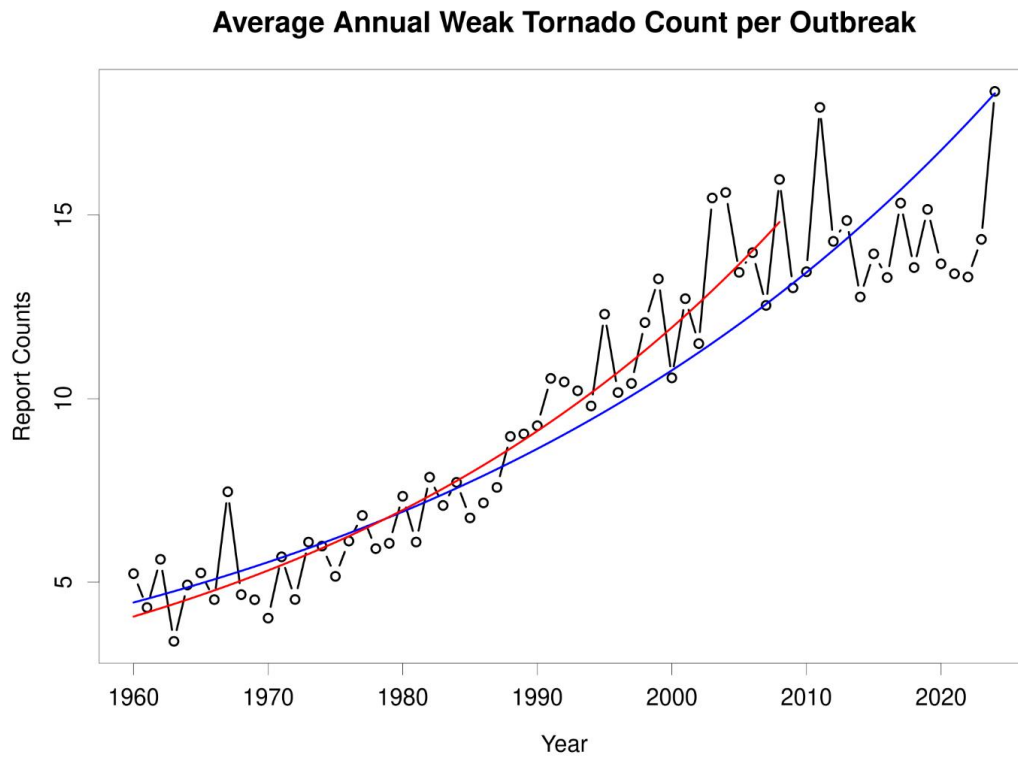


Figure 1. Average annual weak (F/EF0-1) tornado count per outbreak for SD11's period of study (1960-2008 red regression line) and the current period of study (1960-2024 blue regression line) following the detrending methodology described in SD11. Note the divergence in trend line slope from the observed tornado report record in the last 15 years of the time series, demonstrating the need for a new detrending approach.

2. Results

a. Detrending Results

To quantify the suitability of our piecewise detrending approach for weak tornado counts in the TO regions, we compared it directly against a linear and exponential fit like those used in SD11. The piecewise method ($R^2 \sim 0.913$, Fig. 2A) characterized the fit (Fig. 2) marginally better than the linear trend ($R^2 \sim 0.902$, Fig. 2C) and the exponential trend ($R^2 \sim 0.866$, Fig. 2E). The linear detrending approach yielded artificially increased variability in the detrending in the Pre-Fujita period, while the exponential fit showed a clear divergence between the time series trend and the observed flattening seen in the last decade of dataset, an issue alluded to in SD11. Importantly, a notable limitation of the piecewise approach is a slight underestimation of weak tornado counts from 1974 to 1978.

The strong tornado count (EF2-EF3) detrending results (Fig. 3) benefited greatly from the piecewise detrending approach ($R^2 \sim 0.361$, Fig. 3A) owing to the impact of retroactive tornado ratings before 1974's introduction of the Fujita scale (Pre-Fujita). The Pre-Fujita strong tornado report mean was $\sim 50\%$ greater than that of after 1974 (Post-Fujita), with a notable discontinuity at 1974. This discontinuity supported the use of a piecewise detrending method here as well and yielded a substantially better fit than either the linear ($R^2 \sim 0.172$, Fig. 3C) or exponential ($R^2 \sim 0.236$, Fig. 3E) method. Importantly, the piecewise detrending approach removed a statistical U-shaped artifact that was present in the monotonically increasing methods in the middle decades of the detrended time series (Fig. 3D and 3F). The piecewise approach successfully accounted for the Fujita-scale inflection point in the strong tornado count dataset.

Violent tornado counts exhibited the weakest secular trend among all tornado counts (Fig. 4), suggesting secular influences on the violent tornado record may be smaller than for weaker tornado cases. However, a notable drop in violent tornado counts is observed with the implementation of the Fujita Scale (Agee and Childs 2014, Fig. 4A). After this post-Fujita implementation drop, there was a weakly negative slope in the annual violent tornado count mean, with relatively low values in the final decade (consistent with the observed relative minimum in EF4+ tornado activity in the recent decade). This weakly negative slope is likely a result of stricter enforcement of the Fujita scale and more frequent and detailed surveying (Doswell & Burgess 1988; Doswell et al. 2009). Relatively low violent tornado frequencies during the final decade reduced the statistical fit of the piecewise model ($R^2 \sim 0.042$) when compared with the linear ($R^2 \sim 0.100$) and exponential ($R^2 \sim 0.118$) methods. However, those control methods did not characterize the 1974 breakpoint well and overestimated violent tornado counts per outbreak after the Fujita scale implementation (Fig. 4C and 4E). Note that some subtle differences in the time series plots (Fig. 4A, 4C, and 4E) exist owing to the sensitivity of the TO definition to the detrending methodology with violent tornadoes (that have inherently limited sample sizes). Additionally, due to the extreme outlier nature of 1974, it was not included in either the Pre-Fujita or Post-Fujita time series blocks in the piecewise regression approach.

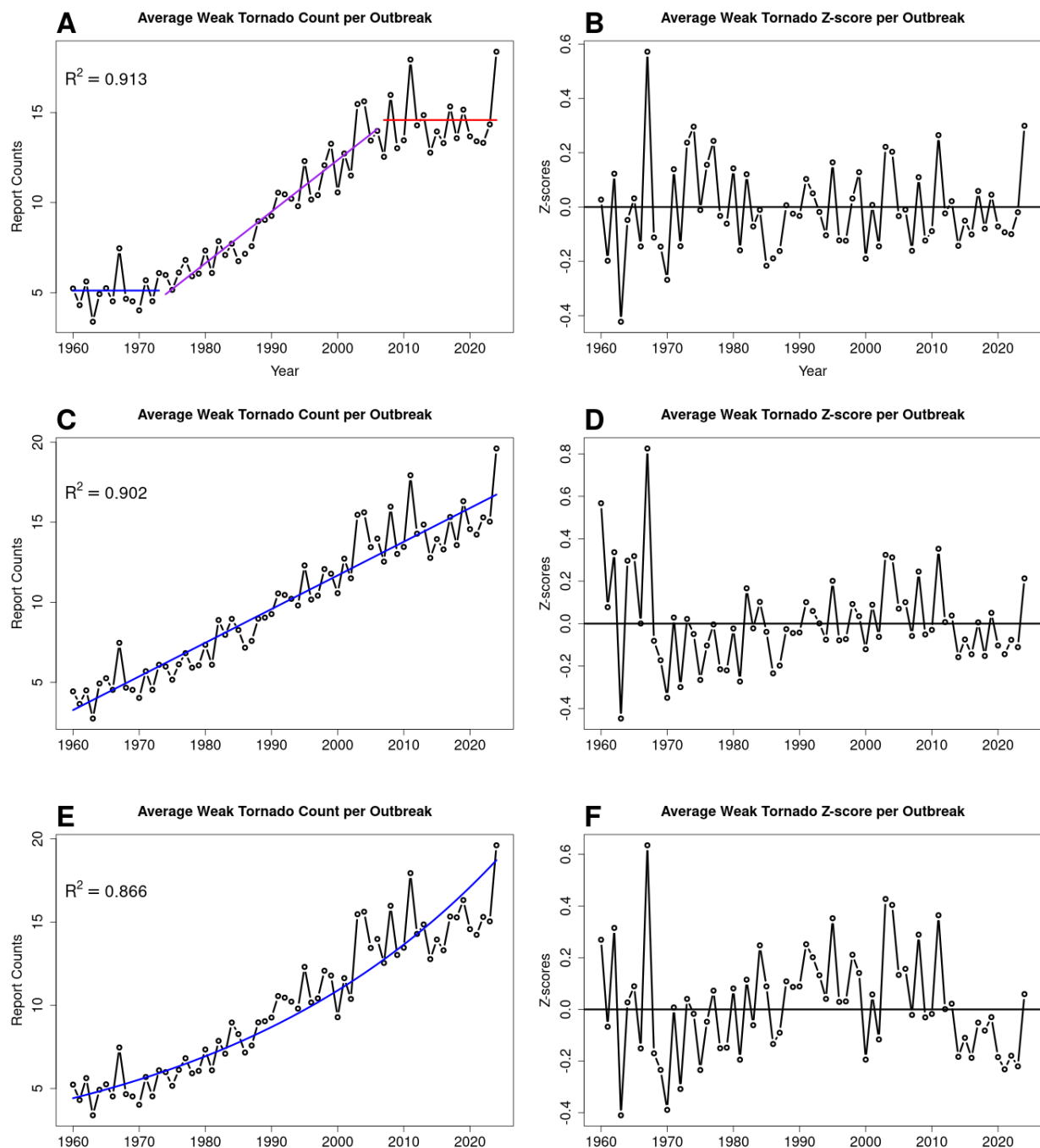


Figure 2. Weak tornado means and modeled means plotted against the annual data (left), and resulting z-scores are plotted about zero (right) for each of the 3 methods: piecewise (top), linear (middle), and exponential (bottom).

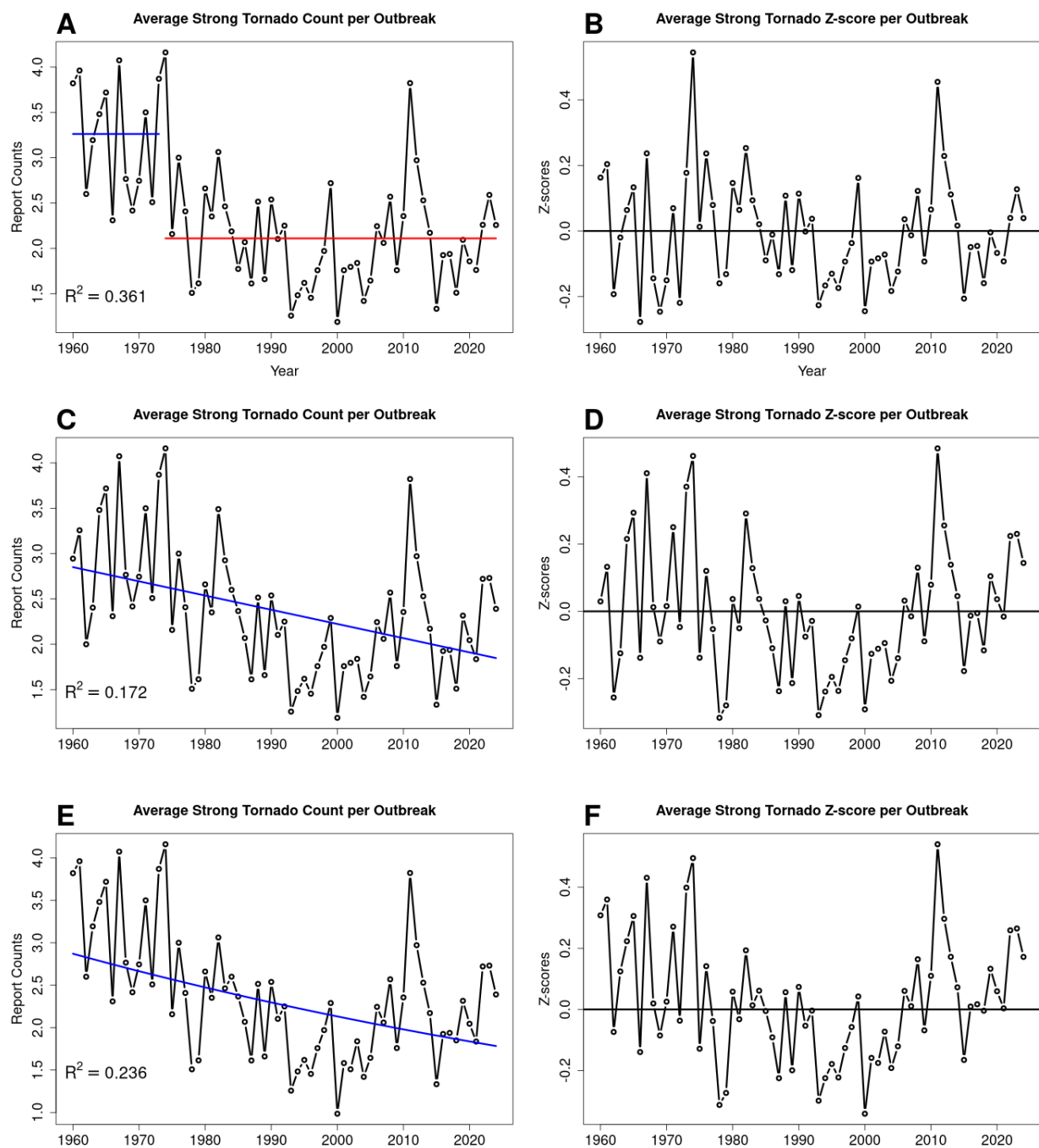


Figure 3. Same as Figure 2, but for strong tornadoes.

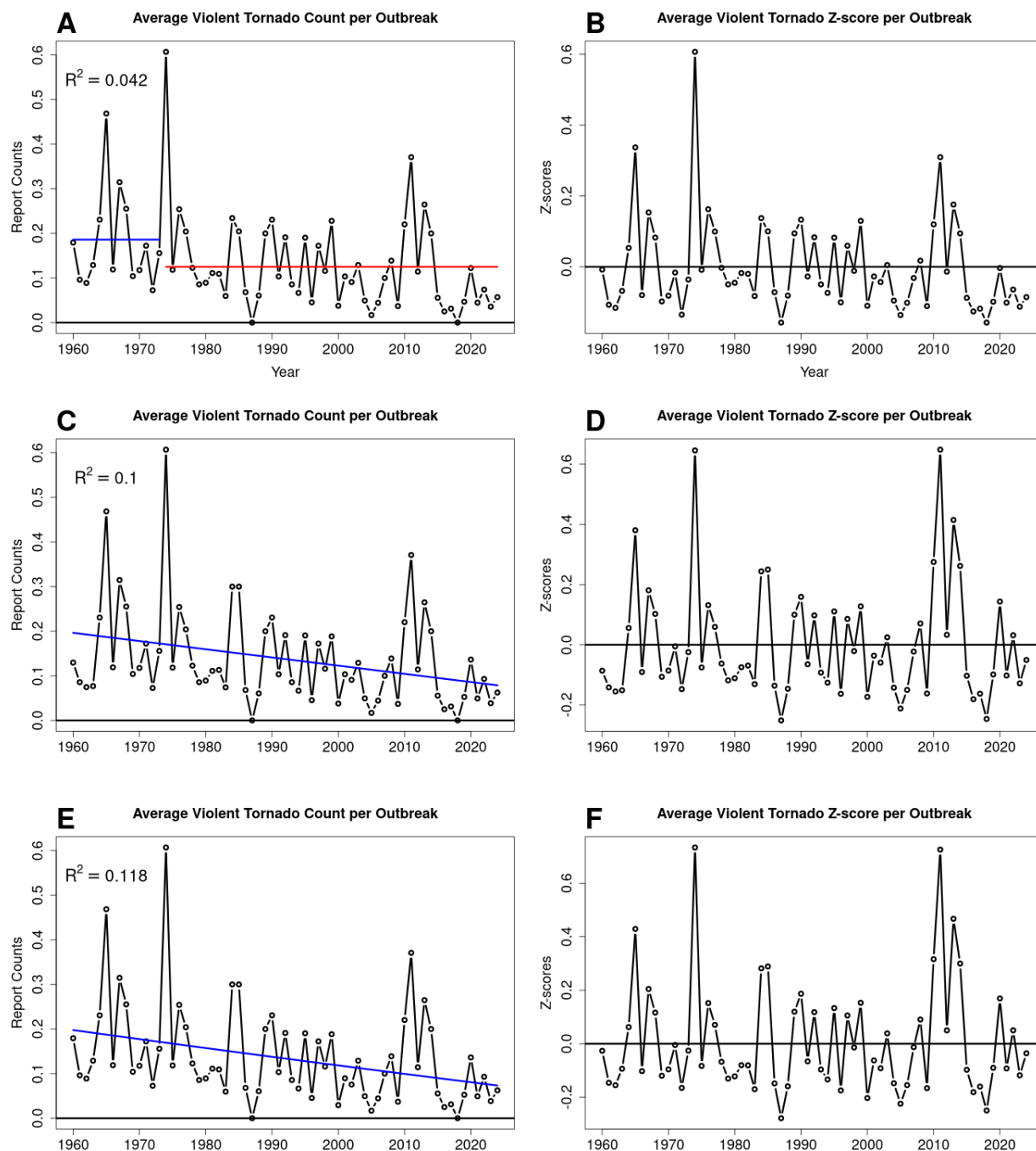


Figure 4. Same as Figure 2, but for violent tornadoes. Note that 1974 was not included in the piecewise approach owing to bias introduced in the detrending from its anomalously high number of violent tornadoes.

b. Index Results

While all previously presented detrending methods had some limitations, the piecewise approach yielded the overall strongest statistical fit for weak tornadoes (which had the strongest secular trend – Verbout et al. 2006) while maintaining data consistency among the breakpoints for each tornado report type. As a result, the intensity index I was constructed from the piecewise detrended outbreak parameters. The sorted I distribution showed heavy right skew (Fig. 5) with greater than 73% of TOs assigned a negative index value. This skew is evidenced by a large range (0.02-26.35) of 4th quartile I values in Fig. 5, while quartiles 1 and 2 (the weakest TOs) only have a total I range of 0.183.

The upper tail of the I distribution provides insight into the sensitivity of the most extreme events to the selected detrending approach. The top 10 outbreaks with the largest I magnitudes from each detrending method are provided in Table 1. The I magnitudes are generally similar, though the piecewise approach yielded slightly smaller values owing to the observed decrease in strong and violent tornadoes post-Fujita. For example, z_v for the “Super Outbreak” of 3 April 1974 was notably consistent among the three detrending methods (26.35 for piecewise, 27.35 for linear and 28.49 for exponential). In contrast, the historic 27 April 2011 TO exhibited considerable variability in z_v among the three methods (18.73 for piecewise, 35.10 for linear, and 39.54 for exponential). Regardless of the detrending method selected, the top 4 outbreaks sorted by I remain fixed in their relative ranking, with only some subtle shifts in the remaining top 10 TOs observed (Table 1). The robustness of the rankings among the most intense TOs regardless of detrending method, coupled with the statistical superiority of the piecewise detrending methodology support its use in the construction of the final TO intensity index I .

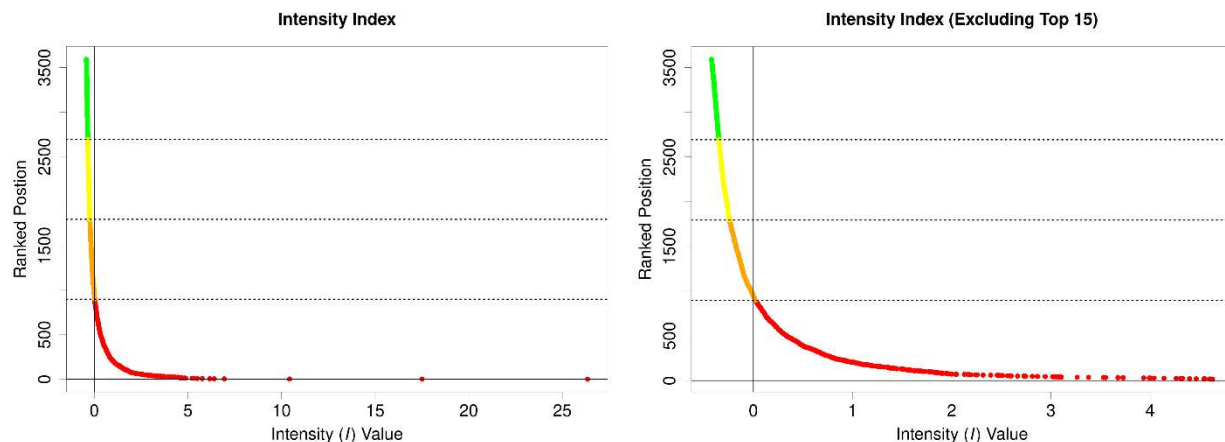


Figure 5. The outbreak intensity I values (piecewise detrending method) are plotted for all outbreaks meeting the mean tornado report threshold. Colors represent individual quartiles of I . Both the full set (left panel) and the set excluding the top 15 outbreaks (right panel) are provided to show detail in the remaining TO I distribution. Most of the outbreaks have a negative index value between 0 and -2. Nearly 95% of the remaining outbreaks with positive I are in Category 4.

Table 1. The top 10 most intense outbreaks for each method. Each outbreak is colored uniquely to demonstrate deviations in position among the different detrending methods tested.

	Piecewise		Linear		Exponential	
	Date	Index	Date	Index	Date	Index
1	April 3 rd , 1974	26.35	April 3 rd , 1974	27.35	April 3 rd , 1974	28.49
2	April 27 th , 2011	17.51	April 27 th , 2011	20.62	April 27 th , 2011	21.31
3	April 11 th , 1965	10.43	April 11 th , 1965	11.12	April 11 th , 1965	11.50
4	March 31 st , 2023	6.944	March 31 st , 2023	7.159	March 31 st , 2023	7.173
5	March 13 th , 1990	6.409	April 12 th , 2020	6.934	March 13 th , 1990	7.079
6	April 12 th , 2020	6.158	March 13 th , 1990	6.791	April 12 th , 2020	7.032
7	December 10 th , 2021	5.768	June 2 nd , 1990	6.400	June 2 nd , 1990	6.821
8	June 2 nd , 1990	5.766	December 10 th , 2021	6.289	February 5 th , 2008	6.530
9	May 4 th , 2003	5.496	February 5 th , 2008	6.243	May 4 th , 2003	6.432
10	May 31 st , 1985	5.302	May 4 th , 2003	6.143	December 10 th , 2021	6.348

When exploring the intensity of individual outbreaks using I , the “Super Outbreak” of 3 April 1974 stands out as the most intense TO in the historical record, with the 27 April 2011 Southeastern United States TO ranking a clear second. Other notable highly-ranked TOs included the 11 April 1965 Palm Sunday TO and the major 31 March 2023 TO. Two other notable recent events also appeared in the top-10 I index values: the 20 December 2021 TO that included the historically long-track Mayfield, Kentucky, tornado and the 12 April 2020 outbreak that impacted locations near the Gulf Coast. Note that none of these post

2020 events were included in the TO intensity climatologies developed by SD11 or F14, providing important context for the impacts of these recent events in the broader TO record.

The TOs presented in Table 2 exhibit substantial variability in the factors contributing to their I values. Some outbreaks, such as 20 December 2021, possess a large I magnitude due to a small number of strong and violent long-track tornadoes, while other events (e.g., 31 March 2023) include extreme numbers of weak tornadoes with very few violent tornadoes. This variability suggests TOs with similar overall intensity may arise through different combinations of outbreak characteristics. We formulated cluster analysis to identify unique TO types within each quartile of the sorted I distribution (Figure 5) as a first attempt at identifying types of TOs by their intensity. Such an analysis can then be used in future work to explore meteorological differences in these outbreak types to better understand physical processes underlying each one.

Table 2. The top 10 most intense outbreaks from the piecewise detrending method for computing I . Both z-scores and raw values for tornado counts and AFM are shown.

	Date	Index Value	Weak Tornadoes		Strong Tornadoes		Violent Tornadoes		AFM	
			Raw	Z-score	Raw	Z-score	Raw	Z-score	Raw	Z-score
1	April 3 rd , 1974	26.35	46	11.390	66	16.970	30	37.610	7052	30.710
2	April 27 th , 2011	17.51	120	8.324	36	9.002	15	18.730	5301	22.990
3	April 11 th , 1965	10.43	14	2.167	15	3.421	17	20.050	2875	12.310
4	March 31 st , 2023	6.944	96	6.429	41	10.330	1	1.102	1882	7.934
5	March 13 th , 1990	6.409	32	3.122	23	5.549	4	4.879	1965	8.302
6	April 12 th , 2020	6.158	80	5.165	30	7.409	3	3.620	1651	6.918
7	December 10 th , 2021	5.768	45	2.402	21	5.018	2	2.361	1959	8.275
8	June 2 nd , 1990	5.766	36	3.677	22	5.283	7	8.656	1366	5.660
9	May 4 th , 2003	5.496	59	4.518	21	5.018	4	4.879	1485	6.187
10	May 31 st , 1985	5.302	10	0.318	13	2.893	7	8.656	1590	6.649

c. Cluster Analysis

To identify these TO types, we first discretized the sorted TO distribution into quartiles, labeled as category 1 (minor – first quartile), category 2 (significant – second quartile), category 3 (major – third quartile), and category 4 (extreme – fourth quartile), shown in Fig. 5. We then clustered the TO intensity diagnostics for each outbreak in the given category to obtain sub-groups of unique structure for each quartile. As k -means requires a guess of the number of clusters to formulate, the elbow method (Wilks 2019)

was used to identify the optimal number of clusters to retain for each I category. These plots (Fig. 6) were sufficient for selecting a number of clusters for categories 1 and 2 (Fig. 6A and B); however, categories 3 and 4 (Fig. 6C and 6D) did not present a clear elbow. For category 3, the within sum-of-squares (WSS) suggested minimal variability in the character of the clusters when $k = 3$. Additionally, when testing $k \geq 4$ for category 3, one cluster was defined exclusively by TOs with a single violent tornado while other outbreak features in that cluster varied widely. Category 4 TOs were challenging to cluster owing to the extreme magnitudes of a handful of the top outbreaks (e.g., Table 2). While $k = 4$ for Category 4 TOs from the elbow plot in Fig. 6D is a reasonable estimate, this k value effectively sorted outbreaks by I , not providing distinct TO types. Selecting $k = 5$ could not effectively remove this bias, placing approximately 64% (572/898) of the outbreaks in the lowest intensity cluster, limiting the uniqueness of the resulting clusters. However, with $k = 6$, intensity bias is minimized by extending the weakest category 4 TOs across 2 clusters while only slightly reducing the I range of cases in the weakest cluster (1.25 to 1.17). As a result, we elected to use $k = 6$ clusters. In total, the number of retained clusters and their individual components are summarized in the results in Table 3. Note that while the detrending substantially reduced temporal inhomogeneities, residual biases with historical tornado practices remain and should be considered when interpreting cluster-level tornado counts. We briefly discuss the cluster results for each category below.

i. Category 1(Minor) TOs

Category 1 primarily comprises outbreaks that possess tornado counts just surpassing the annual mean cutoff threshold value used to define a TO, with most of these outbreaks also containing only weak tornadoes. In fact, two resulting clusters (3 and 4) are entirely composed of weak tornadoes. The prevalence of weak tornadoes means the minor TO clusters are highly dependent on their associated time period. For example, cluster 3 does not contain a TO before 1981, while cluster 4 does not contain an outbreak after 1974. Cluster 4 has slightly higher mean and median z_w than cluster 3, attributed to the slight underestimation of weak tornadoes during the Pre-Fujita period observed in Agee and Childs (2014). Cluster 2 behaves similarly to 3 and 4, clustering TOs with tornado counts near the mean cutoff threshold value. Cluster 1 is the most statistically consistent with the overall tornado report behavior

documented earlier, showing Pre-Fujita outbreaks with 1-2 strong tornadoes and Post-Fujita outbreaks with fewer strong tornadoes, on average. In total, these cluster results suggest two general outbreak types in Category 1: those that contain roughly 1 strong tornado with multiple weak tornadoes and those that only contain weak tornadoes.

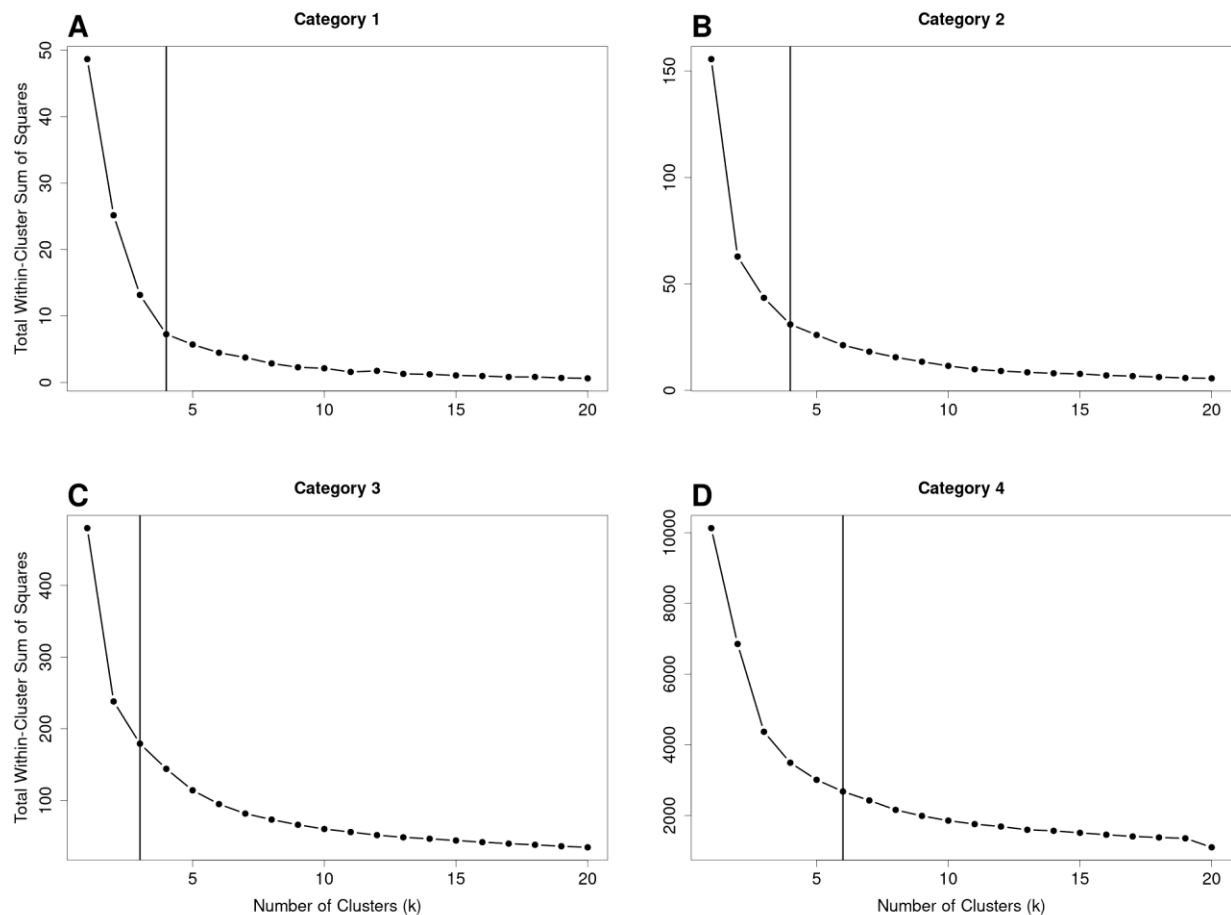


Figure 6. WSS elbow plots for Category 1 (panel A), 2 (panel B), 3 (panel C), and 4 (panel D) TOs. (least to most intense)

Table 3. Average weak, strong, violent, and AFM, as well as intensity index I for each cluster in each category, as well as the number of TOs in each cluster. Clusters are listed below in order of decreasing intensity in their given category. The two most common clusters from each intensity category are highlighted in red. Note that these tornado counts are not detrended but taken directly from the SPC record for interpretation purposes only. The term Q1 refers to category 1 TOs, Q2 for category 2, etc.

		Weak		Strong		Violent		AFM		Index	
		M	Range	M	Range	M	Range	M	Range	M	Range
Q4	G1 (3)	46	14-120	36	15-66	17	15-30	5301.6	2874.9 - 7052.1	17.505	10.427 - 26.349
	G2 (38)	32.5	3-96	21	6-41	3	0-7	1223.0	569.2 - 1965.4	4.362	2.867 - 6.944
	G3 (64)	45.5	18-88	9	0-20	0	0-2	298.9	7.6 - 913.7	1.656	0.567 - 3.074
	G4 (136)	14	1-44	10	1-23	0	0-4	380.3	97.3 - 1219.2	1.354	0.461 - 3.079
	G5 (259)	22	8-55	3	0-11	0	0-2	82.2	1.1 - 491.7	0.297	0.022 - 1.584
	G6 (398)	8	0-25	4	0-12	0	0-4	144.7	11.6 - 510.9	0.238	0.020 - 1.191
	Total	4	0-120	14	0-66	5	0-30	156.4	1.1-7052.1	0.414	0.020-26.349
Q3	G1 (222)	15	6-37	1	0-4	0	-	16.4	0 - 86.2	-0.111	-0.239 - 0.020
	G2 (300)	4	0-12	3	1-8	0	0-1	56.8	1.2 - 175.3	-0.139	-0.240 - 0.020
	G3 (376)	9	3-19	2	0-6	0	0-1	40.1	0.8 - 127.1	-0.150	-0.240 - 0.019
	Total	8	0-37	2	0-8	0	0-1	40.0	0-175.3	-0.136	-0.240-0.020
Q2	G1 (174)	8	5-18	0	0-1	0	-	4.3	0 - 56.2	-0.276	-0.344 - (-0.241)
	G2 (128)	3	0-7	2	2-5	0	-	19.2	0.5 - 60.3	-0.287	-0.347 - (-0.242)
	G3 (272)	5	2-10	1	0-3	0	-	18.8	0.3 - 73.8	-0.307	-0.347 - (-0.240)
	G4 (324)	9	3-13	0	0-2	0	-	8.8	0 - 69.6	-0.316	-0.347 - (-0.240)
	Total	6	0-18	1	0-5	0	-	12.1	0-73.8	-0.304	-0.347 - (-0.240)
Q1	G1 (115)	4	0-7	1	1-4	0	-	5.9	0.2 - 23.3	-0.363	-0.402 - (-0.347)
	G2 (313)	7	3-10	0	0-2	0	-	1.9	0 - 26.8	-0.369	-0.410 - (-0.347)
	G3 (397)	6	4-8	0	-	0	-	1.7	0 - 25.0	-0.397	-0.423 - (-0.347)
	G4 (74)	4	4-5	0	-	0	-	1.079	0 - 17.4	-0.414	-0.418 - (-0.354)
	Total	6	0-10	0	0-4	0	-	2.1	0-26.8	-0.383	-0.423 - (-0.347)

ii. *Category 2 (Significant) TOs*

Category 2 TOs revealed a different clustering structure than category 1 owing to a smaller influence of temporal reporting bias with these stronger outbreaks. Many outbreaks in category 2 were well above the minimum tornado count threshold used in our TO definition. The remaining TOs that were near this threshold typically had 1 or more strong tornadoes or occurred in the Pre-Fujita period. There are 2 clusters (clusters 1 and 4) with high weak tornado counts and lower AFMs. While both have similar median tornado counts, cluster 1's Pre-Fujita median weak tornado count is lower than cluster 4 owing to the smaller number of Post-Enhanced-Fujita TOs in cluster 1 (14.5 versus 11 for cluster 4). While this temporal distribution is partly responsible for the clustering, there appears to be a meaningful difference in cluster intensity as well. Cluster 2 generally comprises outbreaks with a high ratio of strong to weak tornadoes while cluster 3 shows a moderate number of weak tornadoes and

a median of 1 strong tornado. Overall, category 2 TOs fall into two broad archetypes. The first (based on clusters 1 and 4) have TOs dominated by large numbers of weak tornadoes and relatively low AFM values. The second (clusters 2 and 3) contains outbreaks where a small number of strong tornadoes contribute more substantially to outbreak intensity, either through elevated strong-to-weak tornado ratios (cluster 2) or a more balanced measure of strong and weak tornadoes (cluster 3).

iii. Category 3 (Major) TOs

Category 3 comprises outbreaks ranging from numerous weak tornadoes to outbreaks of a few strong to violent tornadoes, with Pre-Fujita tornado totals ranging from 6-37 before detrending. The Category 3 TO group yielded three clusters (Table 3) with distinct characteristics. Cluster 1 possesses a high weak tornado count (Pre-Enhanced-Fujita range of 18-37), as well as small AFMs and strong tornado counts. In contrast, cluster 2 contains higher strong tornado counts (Post-Enhanced-Fujita range: 1-6) and AFMs with minimal numbers of weak tornadoes. Cluster 3 lies between these two extremes, once again suggesting two primary types of major TOs: those that produce large numbers of weak tornadoes and those that produce a smaller number of long-track strong tornadoes. Note that while outbreaks in category 3 did include violent tornadoes, the median violent tornado count remained 0 for all three clusters, suggesting violent tornadoes are a rare occurrence with this category of outbreak intensity.

iv. Category 4 (Extreme) TOs

Category 4 TOs are typically characterized by exceptionally large tornado counts, elevated numbers of strong and violent tornadoes (or both), and a high overall median tornado count (20) with an $I = 0.414$. However, the highly right-skewed nature of the I (Fig. 5) distribution results in a large range of intensity values for Category 4 TOs (Table 3). Cluster 1 only contains the top 3 TOs in the dataset (the Super Outbreak, 27 April 2011, and the Palm Sunday outbreak), all of which are historically significant events (Fujita et al. 1970; Fujita 1974, 1975; Knupp 2014). Regardless of the clustering approach we

utilized, these three outbreaks consistently formed a distinct cluster, highlighting their exceptional nature. Outside of Cluster 1, Category 4 TOs separated into two similar cluster pairs (clusters 3 and 4 and clusters 5 and 6, respectively). Within each pair, one cluster was characterized by larger numbers of weak tornadoes, whereas the other contained fewer weak tornadoes with approximately one more strong tornado. The shift in weak tornado counts when comparing these cluster pairs is again partially attributed to the breakpoints in the TO time series. Cluster 2 was second only to cluster 1 in overall intensity and included the remainder of the top 10 outbreaks in Table 2. Additionally, greater than 35% (177 of 493) of all violent tornadoes in the record were produced by the 41 outbreaks in clusters 1 and 2. This result is consistent with the strong contribution of violent tornadoes to TO severity and highlights their association with the most extreme TOs in the historical record.

3. Discussion and Conclusions

In this study, we quantify TO intensity by combining two seminal outbreak intensity frameworks (S11 and F14) using an improved detrending methodology that accounts for non-meteorological breakpoints in the tornado record. Our results demonstrate the value of a piecewise detrending approach in removing these secular trends while providing a more temporally consistent representation of outbreak tornado counts across the full tornado record. Specifically, we find that no statistically significant trends exist in mean weak tornado counts (EF0-EF1) per TO prior to the Fujita Scale implementation in 1974 nor following the Enhanced Fujita Scale implementation in 2007, with a stark linear trend observed between these two periods (Fig. 2). Non-meteorological impacts such as surveying efforts and warning verification, population growth, and, most importantly, the introduction of the WSR-88D, contributed heavily to the dramatic upward trend between 1974 and 2007 and motivated our decision to utilize a piecewise temporal distribution. We also observe that trends in strong and violent tornado counts were closely tied to the introduction of the Fujita scales in 1974 and 2007. As a result, strong and violent tornado counts exhibit discontinuities in the transition between the Pre-Fujita and Fujita periods, requiring additional detrending.

Our intensity index I successfully reproduces several expected features of a tornado outbreak climatology. The most historically notable TOs (Table 3) in the report record are consistently associated with the largest I values (category 4 clusters 1 and 2). Our accompanying cluster analysis reveals that TOs in the lower three quartiles of I possess two primary characteristics: numerous weak tornadoes or few weak tornadoes with an increase in strong and/or violent tornadoes. These characteristics remain remarkably consistent across categories 1 through 3, suggesting that outbreaks of similar overall intensity can be achieved in multiple ways. Some outbreaks derive their intensity by producing relatively large numbers of weak tornadoes, while others are characterized by a greater proportion of stronger and long-track tornadoes. Such a result introduces the idea of outbreak types within TOs, which is a potential area of future research. Our index also reflects the rare-event nature of violent tornadoes, which are barely present in category 3 TOs (and non-existent in Category 1 and 2 TOs). Instead, a large percentage ($>35\%$) of violent tornadoes are associated with only the strongest I values in our category 4 (extreme) outbreaks. Since these violent tornadoes are associated with the greatest societal impacts, substantial damage, and typically longer and wider tracks that yield large AFM values (Doswell et al. 2006, Straka and Kanak 2022, F14), their appearance in an outbreak often elevated them to the highest intensity category.

By including diagnostics from both SD11 and F14, our blended I index provides a more comprehensive representation of TO intensity than either approach alone. The framework formalizes a categorization of outbreaks based on tornado composition while preserving key outbreak characteristics such as tornado intensity and path length, as well as the outbreak's geographic extent. Moreover, the index appropriately captures the disproportionate contribution of violent tornadoes to the most extreme outbreaks.

While the results presented herein demonstrate the effectiveness of I in characterizing TO properties, it is important to note the limitations of our approach that may have introduced bias into the results. The results are based on an incomplete tornado record that relies heavily on tornado observation before an event can be included. This likely contributes to the non-meteorological trends in the tornado count record, though we expect the effect is minimized in the Post-Enhanced Fujita period owing to the abundance of observations available today. Additionally, while our piecewise regression approach dramatically reduces

the temporal biases that exist within the tornado count records, some artifacts remain in the data. For example, violent tornado counts in the Post-Fujita era exhibit a weakly negative slope. Two major outlier events (the Super Outbreak and the 27 April 2011 outbreak) contain unusually large numbers of violent tornadoes that substantially increase the variance within the record and likely have influenced this trend. Raw tornado counts also serve as an initial criterion to select which events are included in our detrending analysis, despite the Pre-Fujita period having a larger ratio of strong to weak tornadoes than later years. This likely contributes to an overrepresentation of strong tornadoes during the Pre-Fujita period. Finally, our selection of a piecewise detrending methodology, while consistent with known non-meteorological shifts in the tornado record, could introduce some bias in the resulting I metric. These issues introduce uncertainty into our analysis, though they largely reflect inherent limitations in the historic tornado record and its associated non-meteorological factors.

Despite these limitations, this updated TO intensity method offers a framework for future analysis investigating meteorological conditions underlying outbreaks of varying impact. This in turn could improve forecasting of TO intensity, not just occurrence, potentially yielding substantial benefits for impact mitigation and risk reduction. Further, this framework offers a methodological approach that could be employed in other hazard fields where clear secular trends exist within their observational record, especially those with spatial aspects to the hazard impact distribution. Finally, this work can naturally be expanded to other important severe weather hazards, including hail and straight-line winds, ultimately supporting the development of a unified severe weather outbreak intensity framework. Ultimately, our approach offers an updated, temporally consistent and physically meaningful characterization of TO intensity, establishing a foundation for future climatological, forecasting, and hazard-impact studies.

4. Methods

a. Data

Tornado report data (1960-2024) were obtained from the Storm Prediction Center (SPC) Severe Weather Report database (Schaefer and Edwards 1999, Edwards 2022). Despite its prevalence, the SPC

Severe Weather Report database still contains known biases that will have important impacts on work such as this study. These include:

- 1) Tornado path width reporting: In 1994, the National Weather Service began requiring the reporting of max path width instead of mean path width, creating a temporal discontinuity.
- 2) Introduction of the EF-scale: In 2007, the updated EF scale created a second temporal discontinuity in the tornado intensity measures and introduced the EF-Unknown category (EFU) for tornadoes that were not readily quantified in terms of damage and impact (Edwards et al. 2013, Edwards et al. 2021, Edwards 2022).
- 3) Non-meteorological trends in the data: Non-meteorological trends (Brooks 2004, Agee and Childs 2014, McCarthy 2003, Edwards 2022) from reporting biases owing to the introduction of updated technology (i.e., cell phones, radar) created an artificial increase in tornado reporting that was highly statistically significant yet non-meteorological in nature.

These issues, while known, are limitations in the database that are not easily overcome. However, it is critical these issues are addressed in the construction of an updated TO climatology. We address these issues utilizing the methods presented below.

b. Defining a TO

As seen above, a consistent definition of a TO is not present in the literature. To align our work closely with more recent TO intensity studies (Edwards et al. 2004; Doswell et al. 2006), we utilize outbreak days to quantify individual TOs. Our outbreak days are constrained to 24-hour periods from 1200 UTC to 1159 UTC the following day and include all reported tornadoes occurring between 25 and 50° latitude and between -110 and -70° longitude (Fig. 7). Once the tornado reports are sorted into outbreak days, each day could be comprised of multiple areas of enhanced tornado activity (i.e. SD11, their Fig. 2). To establish spatial boundaries for each TO's region, we employed the same kernel density estimation (KDE) as in SD11. Modifying KDE parameters such as bandwidth and contour level can adjust the outbreak region size and shape, so we utilized the same values as SD11 (grid spacing of 0.25° with a bandwidth of 1 and KDE

contour level of 0.001) to ensure consistency with their method This approach yielded 16,136 total KDE outbreak regions across our study period.

After identifying the tornado activity KDE regions, we computed the mean total tornado counts for each year's KDE regions. All KDE regions whose total tornado counts were smaller than the detrended annual mean tornado count for the given year (Fig. 8) were culled as they were deemed too small be considered an outbreak. Given the observed strong right skew to the tornado report distributions per year (not shown), this reduced the number of retained KDE regions (hereafter TOs) down roughly 78% to only 3,593. This ensured only the most tornado-active KDE regions were considered TOs for the purposes of our intensity index.

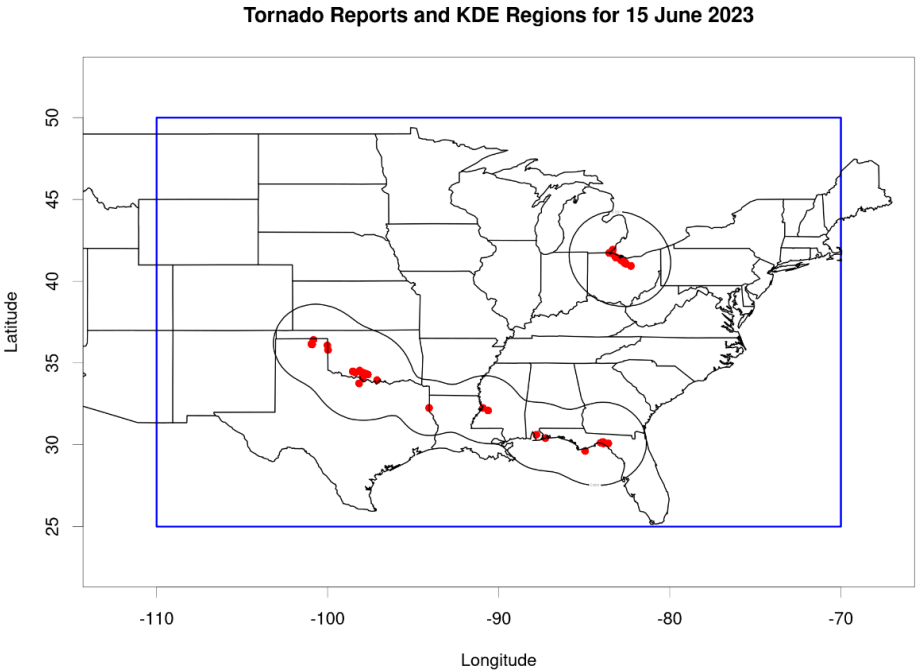


Figure 7. The study region (blue box) and an example of KDE regions on a multi-outbreak day (15 June 2023).

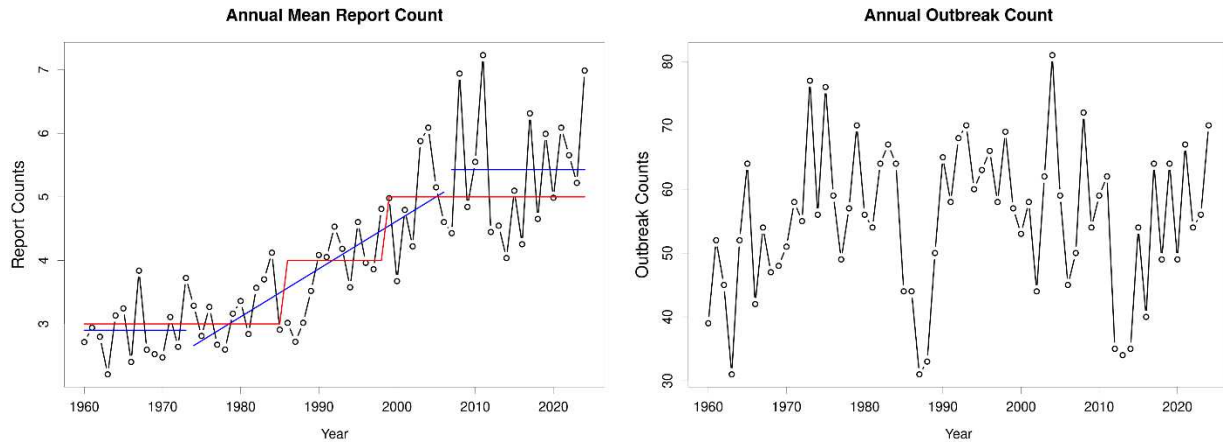


Figure 8. Annual mean reports with regression (blue) and rounded threshold values (red) overlay (Left). Report regions per year following removal of mean report values (Right)

c. Variables

Many previous studies utilized either single variables comprised of multiple tornado characteristics (F14) or multiple individual parameters (SD11) to quantify TO intensity. These studies have a commonality; the tornado diagnostics (e.g., path length, Fujita scale, DPI) shared a high correlation ($r > 0.8$) which introduces multicollinearity and complicates interpretation of the resulting intensity indices. Our study blends these two approaches by incorporating the adjusted Fujita mile (AFM) diagnostic from F14 alongside tornado counts within three distinct EF-scale intensity categories. AFMs have advantages over other previous indices (such as those in SD11) that address overall outbreak impact as they do not rely on path width (a limitation in the Severe Weather Report database) and they possess a strong relationship with societal impacts without directly requiring societal impact data (F14). However, AFMs also possess two key limitations: (1) AFMs likely overestimate the damage rating on a full tornado track where the EF scale is not well known along the entire track (F14), and (2) the AFM calculation in F14 excluded EF0 tornadoes. To address these limitations, we include counts of weak (EF0 – EF1), strong (EF2 – EF3), and violent (EF4+) tornadoes as additional intensity diagnostics. This approach mitigates multicollinearity concerns in the SD11 and F14 methods (Table 4) while preserving both tornado intensity by tornado counts and path length information using AFMs. We did observe a high correlation

(0.81) between violent tornadoes and AFM, but we deemed this limitation acceptable given the rare nature of violent tornadoes (roughly 1.1% of all tornadoes) in the overall outbreak database.

Table 4. Correlation matrix of tornado index variables before detrending. Only outbreaks kept for the detrending process are included in this data. All are statistically significant ($p < 0.05$).

	Weak	Strong	Violent	AFM	Path Length
Weak					
Strong	0.463				
Violent	0.224	0.522			
AFM	0.456	0.779	0.810		
Path Length	0.630	0.810	0.669	0.950	

d. Trends and detrending

As our study period spanned 65 years, severe weather reporting capabilities and practices, surveying, and detection capabilities have notably shifted over that period. Notably, from 1970 to 1999, the number of severe local storm events (of which tornadoes are one component) recorded in the Severe Weather Report database increased by nearly a full order of magnitude. This is attributed in prior studies to a multitude of non-meteorological factors, including the development of trained spotter programs, the national warning verification program, the NEXRAD radar network, population increase, and an increase in weather awareness (Weiss et al. 2002).

Despite these non-meteorological factors, tornadoes and tornado track trends are a more nuanced topic owing to their unique measures of path length and width. Additionally, tornado intensity is not scaled based on a physically quantifiable characteristic of the tornado, like size or duration. Instead, it is measured using the Fujita/Enhanced Fujita scales that have undergone multiple adjustments over the study period. Two key breakpoints exist in the Fujita/Enhanced Fujita scale definitions: (1) the implementation of the Fujita scale in National Weather Service offices in 1974 and (2) the introduction of the Enhanced Fujita Scale (EF-scale) in 2007. Consistent with Agee and Childs (2014), the annual TO tornado count statistics presented herein reveal clear discontinuities near both the Fujita scale implementation in 1974 and the Enhanced Fujita scale implementation in 2007 (Fig. 9). Specifically, while F/EF2 tornadoes initially decrease post-implementation (Fig. 9B), F/EF1 tornadoes increase, suggesting an overestimation of F/EF2 tornado

intensity in the pre-Fujita period. Additionally, the increase in F/EF0 tornadoes appears to be associated with population growth prior to 1990, which represented the nationwide implementation of the Weather Surveillance Radar-1988 Doppler (WSR-88D) network. This implementation coincided with a massive annual increase of greater than 200 F0 tornadoes per year over the span of 2 years and preceded an increase in annual F0 variability (Agee and Childs 2014). They go on to state that F0 and F1 damage reports greatly increased when Weather Forecasting Offices incorporated the services of storm spotters. While the increase in total annual tornadoes mostly flattens in the 1990s, weak tornadoes specifically associated with outbreak regions continued to increase until roughly 2007, corresponding to the implementation of the EF-scale described above (Fig. 9A). Violent tornadoes exhibit substantially weaker non-meteorological trends (Fig. 9C), and AFMs are minimally impacted by these issues (Fig. 9D and F14). These breakpoints suggest that a single monotonic correction (such as a linear or exponential trend removal method) is unlikely to account for the temporal structure within the tornado record, motivating the use of a piecewise detrending methodology. AFMs exhibited negligible secular trends when compared with tornado count diagnostics (Fig. 4D), so this piecewise detrending was applied only to the tornado-count variables used in the construction of the intensity index.

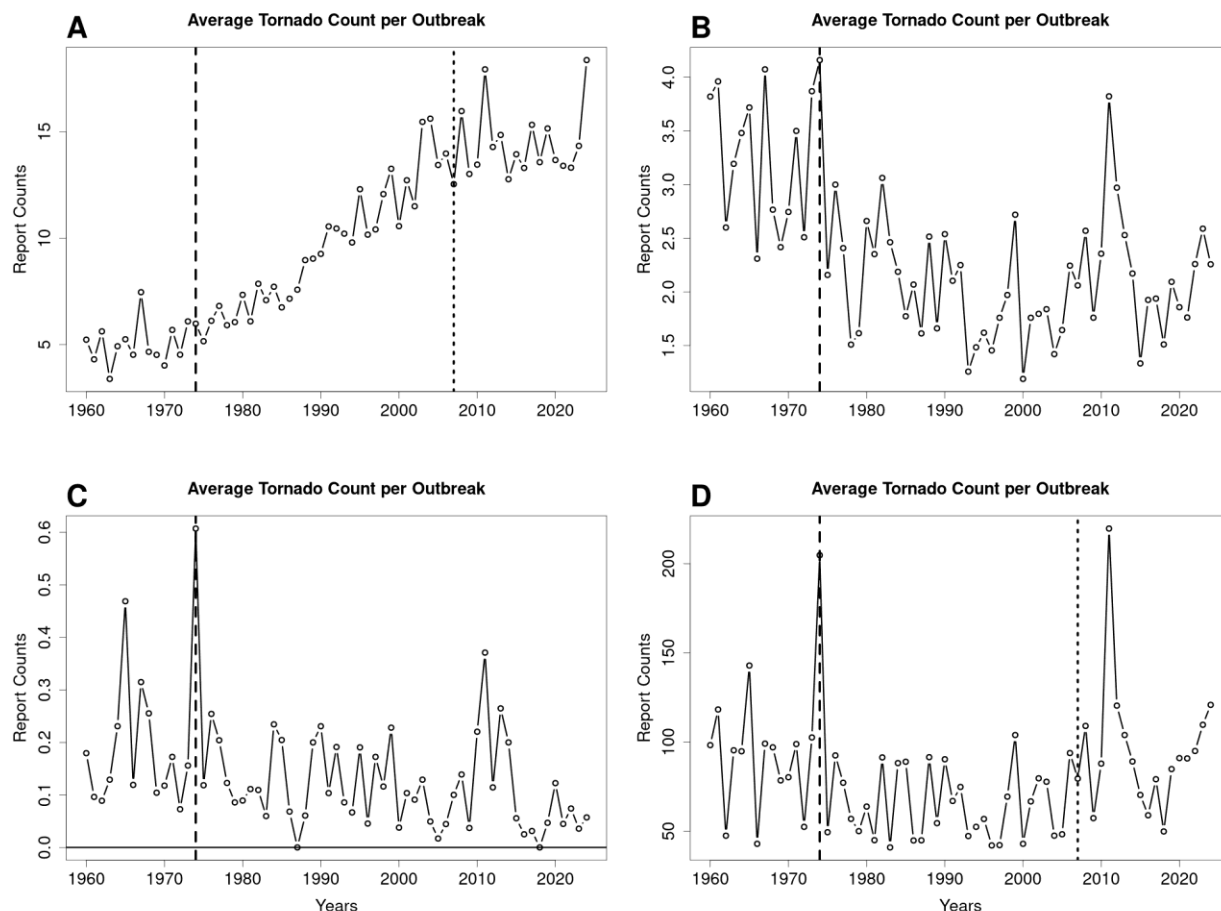


Figure 9. Average annual weak tornadoes (A), strong tornadoes (B), violent tornadoes (C), and AFMs (D). Dotted lines are plotted for 1974 (NWS started implementing the Fujita Scale) and 2007 (NWS adopted the new Enhanced Fujita Scale).

The staggered structure of the weak tornado database (Fig. 9A) motivated the utilization a piecewise detrending methodology to remove secular trends. For the weak tornado counts (Fig. 9A), we separated the time series into three distinct periods: Pre-Fujita (1960-1973), Post-Fujita (1974 – 2006), and Post-Enhanced-Fujita (2007 onward). The weak tornado Pre-Fujita and Post-Enhanced Fujita trends were non-significant ($p > 0.05$) and thus were not corrected, while the Post-Fujita period (1974-2006) exhibited a highly statistically significant positive trend ($p < 0.05$). Additionally, strong tornado trends within the Pre-Fujita and Post-Fujita periods were not statistically significant, though a breakpoint exists in strong tornado counts with the implementation of the Fujita scale in 1974. While Pre-Fujita violent tornado trends are also not statistically significant, Post-Fujita violent tornado counts are only non-significant when

removing the Super Outbreak from 3 April 1974, which was a massive outlier for violent tornadoes. For this reason, the Post-Fujita violent tornado trend is assumed to be non-significant. To address this pre-Fujita/post-Fujita breakpoint, we only use 1974 to separate the strong and violent tornado record for both strong and violent tornadoes (i.e., there is no second breakpoint introduced in the Post Enhanced-Fujita period). This distinction helped reduce biases that lingered in tornado reporting prior to the implementation of the Fujita scale.

Following identification of the breakpoint structure, univariate linear regression is applied to the Post-Fujita significant trend in weak tornadoes (1974-2007 weak tornadoes), while all remaining non-significant segments in all three tornado intensity record time series remain represented by temporally stationary means. An analogous procedure is applied to annual standard deviations, which exhibit highly similar temporal structures to the mean counts (not shown). These annual detrended means and standard deviations are then used to scale each diagnostic into a detrended z -score (equation 1):

$$z = \frac{x - \hat{x}}{\hat{s}} \quad (1)$$

where $\hat{x}$ is the detrended mean from the piecewise regression and $\hat{s}$ is the detrended standard deviation. These z -scores are then able to be summed to construct an intensity index. It is important to note that the inclusion of a massive outlier TO for violent tornadoes (the Super Outbreak of 3 April 1974) inflates the raw report standard deviation of the Post-Fujita period for violent tornadoes. Consequently, the resulting violent tornado z -scores exhibit reduced variability during the Post-Fujita period. These detrended z -scores, alongside z -scores for AFMs that were not temporally detrended, were used in the construction of the intensity index.

e. Variable Weighting

To compute the intensity index, we utilized a simple summed weighting scheme, specifically splitting the index into 50% weight based on tornado counts and 50% weight based on AFMs. For the 50%

weight for tornado counts, we utilized detrended z scores of weak, strong, and violent tornadoes equally, each weighted 1/3 the relative weight of the AFMs (equation 2):

$$I = \frac{3z_A + z_W + z_S + z_V}{6} \quad (2)$$

where z_A represents the z -score for the AFM, z_W is the detrended z -score for weak tornadoes, z_S the detrended z -score for strong tornadoes, and z_V the detrended z -score for violent tornadoes. This equal weighting scheme was selected to balance the tornado count information offered by the report diagnostics as well as path length information yielded by the AFMs. A single value of I was constructed for all 3,593 TOs discussed above, and these values served as the final intensity index for each TO in the study. This simple scheme is utilized to maximize flexibility such that a different weighting approach, if desirable and beneficial, could be implemented in a future study.

e. Cluster Analysis

Once I was obtained, we quantify the relative impact of the tornado diagnostics on the outbreak intensity to identify unique outbreak types. SD11 originally applied a k -means cluster analysis to a four-dimensional decomposition of the severe weather index values (tornado component, hail component, wind component, and an “other” component) to address this in their work, further decomposing their results into minor and major outbreaks. Here we are exclusively exploring TOs, so we utilized a similar approach directly on the diagnostics comprising I . All four diagnostics (weak, strong, and violent tornado z -scores as well as AFM z -scores) are used in the k -means cluster analysis (Wilks 2019). To formulate clusters, the I distribution was sorted and discretized into quartiles, where each quartile (deemed category 1 through 4 TOs in the results) was clustered independently to identify unique structures within these intensity categories. Elbow plots were used to determine the number of retained clusters for each TO category. The resulting

527 clusters were used to identify recurring outbreak structures within each intensity category and
528 examine how TO composition varied with increasing values of I .

529

530 **Funding Declaration**

531 The authors declare that this research received no specific funding.

References

1. 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).
2. Beebe, R. G. Tornado Composite Charts. *Mon. Wea. Rev.* **84**, 127–142 (1956).
3. Brooks, H. E. On the Relationship of Tornado Path Length and Width to Intensity. *Weather Forecast.* **19**, 310–319 (2004).
4. Ćwik, P., McPherson, R. A. & Brooks, H. E. What is a tornado outbreak?: Perspectives through time. *Bull. Amer. Meteorol. Soc.* **102**, E817–E835 (2021).
5. Doswell, C. A., III, Edwards, R., Thompson, R. L., Hart, J. A. & Crosbie, K. C. A simple and flexible method for ranking severe weather events. *Weather Forecast.* **21**, 939–951 (2006).
6. Doswell, C. A., III & Burgess, D. W. On some issues of United States tornado climatology. *Mon. Wea. Rev.* **116**, 495–501 (1988).
7. Doswell, C. A., III, Brooks, H. E. & Dotzek, N. On the implementation of the Enhanced Fujita Scale in the USA. *Atmos. Res.* **93**, 554–563 (2009).
8. Edwards, R., Elliott, M. S., Marsh, P. T. & Speheger, D. A. Errors, oddities and artifacts in U.S. tornado data, 1995–2021. In *30th Conference on Severe Local Storms*, 8.3B (American Meteorological Society, 2022).
9. Edwards, R. et al. Tornado Intensity Estimation: Past, Present, and Future. *Bull. Am. Meteorol. Soc.* **94**, 641–653 (2013).
10. Edwards, R., Thompson, R. L., Crosbie, K. C., Hart, J. A. & Doswell, C. A., III. Proposals for modernizing the definitions of tornado and severe thunderstorm outbreaks. In *22nd Conference on Severe Local Storms* (Amer. Meteor. Soc., Hyannis, MA, 2004).
11. Forbes, G. S. Meteorological aspects of high-impact tornado outbreaks. In *Symposium on the Challenges of Severe Convective Storms* (Amer. Meteor. Soc., Atlanta, GA, 2006).
12. Fujita, T. T. Jumbo tornado outbreak of 3 April 1974. *Weatherwise* **27**, 116–126 (1974).
13. Fujita, T. T., Bradbury, D. L. & Van Thullenar, C. F. Palm Sunday tornadoes of April 11, 1965. *Mon. Wea. Rev.* **98**, 29–69 (1970).
14. Fujita, T. T. New evidence from April 3–4, 1974 tornadoes. In *Preprints, 9th Conference on Severe Local Storms* 248–255 (Amer. Meteor. Soc., 1975).
15. Fuhrmann, C. M. et al. Ranking of tornado outbreaks across the United States and their climatological characteristics. *Weather Forecast.* **29**, 684–701 (2014).
16. Galway, J. G. Relationship of Tornado Deaths to Severe Weather Watch Areas. *Mon. Wea. Rev.* **103**, 737–741 (1975).
17. Glickman, T. S., ed. *Glossary of Meteorology*, 2nd edn. (American Meteorological Society, 2000).
18. Knupp, K. R. et al. Meteorological overview of the devastating 27 April 2011 tornado outbreak. *Bull. Am. Meteorol. Soc.* **95**, 1041–1062 (2014).
19. McCarthy, D. W. NWS tornado surveys and the impact on the National Tornado Database. *Preprints, First Symp. on F-Scale and Severe-Weather Damage Assessment*, Long Beach, CA, Amer. Meteorol. Soc., CD-ROM, 3.2 (2003).
20. Mercer, A. E., Shafer, C. M., Doswell, C. A., III, Leslie, L. M. & Richman, M. B. Objective classification of tornadic and nontornadic severe weather outbreaks. *Mon. Weather Rev.* **137**, 4355–4368 (2009).

21. Schneider, R. S., Schaefer, J. T. & Brooks, H. E. Tornado outbreak days: An updated and expanded climatology (1875–2003). In *22nd Conference on Severe Local Storms* (Amer. Meteor. Soc., Hyannis, MA, 2004).
22. Schaefer, J. T. & Edwards, R. The SPC tornado/severe thunderstorm database. In *Preprints, 11th Conference on Applied Climatology* 603–606 (Amer. Meteor. Soc., Dallas, TX, 1999).
23. Schaefer, J. T., Schneider, R. S. & Kay, M. P. The robustness of tornado hazard estimates. In *Preprints, 3rd Symposium on Environmental Applications* 35–41 (Amer. Meteor. Soc., Orlando, FL, 2002).
24. Shafer, C. M. & Doswell, C. A., III. Identifying and ranking multi-day severe weather outbreaks. In *26th Conference on Severe Local Storms* (Amer. Meteor. Soc., Nashville, TN, 2012).
25. Shafer, C. M. & Doswell, C. A., III. Methods to rank and classify severe weather outbreaks. In *9th Conference on Artificial Intelligence and Its Applications to the Environmental Sciences* (Amer. Meteor. Soc., Seattle, WA, 2011).
26. Straka, J. M. & Kanak, K. M. A climatology of long-track tornadoes. *Electron. J. Severe Storms Meteor.* **17**, 1–49 (2022).
27. Thompson, R. L. & Vescio, M. D. The destruction potential index—A method for comparing tornado days. In *Preprints, 19th Conference on Severe Local Storms* 280–282 (Amer. Meteor. Soc., Minneapolis, MN, 1998).
28. Verbout, S. M., Brooks, H. E., Leslie, L. M. & Schultz, D. M. Evolution of the U.S. tornado database: 1954–2003. *Weather Forecast.* **21**, 86–93 (2006).
29. Weiss, S. J., Hart, J. A. & Janish, P. R. An examination of severe thunderstorm wind report climatology: 1970–1999. In *21st Conference on Severe Local Storms*, 446–449 (American Meteorological Society, 2002).
30. Wilks, D. S. *Statistical Methods in the Atmospheric Sciences*, 4th edn. (Elsevier, 2019).