8.3B ERRORS, ODDITIES and ARTIFACTS in U.S. TORNADO DATA, 1995–2021

Roger Edwards[1], Matthew S. Elliott and Patrick T. Marsh
Storm Prediction Center, Norman, Oklahoma

Douglas A. Speheger
NWS Weather Forecast Office, Norman, Oklahoma

1. BACKGROUND and DATA CHARACTERISTICS

Tornado data in the U.S. have been used for decades for research into climatological occurrence characteristics (e.g., Doswell and Brooks 2001; Brooks et al. 2014; Elsner et al. 2014a) and to assess risk to society and industry (e.g., Thom 1963; Grazulis 1993; Coleman and Dixon 2014). Official tornado records follow a stepwise documentation process, from preliminary reporting to "final" codification in a unified national dataset. The Storm Prediction Center (SPC) and its predecessor, the National Severe Storms Forecast Center (NSSFC), have maintained a national tornado database since late 1954 (Pautz 1969), extended back to 1950 by manual polling of the former Environmental Data Service publication *Climatological Data* (a predecessor to *Storm Data*; Kelly et al. 1978). Details discussed by Schaefer and Edwards (1999), who explained the history and nature of the data before then, are summarized and updated in this section.

Local National Weather Service (NWS) weather forecast offices (WFOs) are responsible for initially documenting each tornado within their county warning areas (CWAs). WFOs usually provide preliminary local storm reports (LSRs) of possible tornadoes over a wide time range—from real time, as tornadoes are still underway with eyewitnesses reporting them and/or radar tornadic debris signatures (Rhyzhkov et al. 2005) detect them, through ensuing days. LSRs are most common the days of and after a tornadic event, and form the basis for the daily SPC reports website. These reports are highly preliminary and subject to changes, including confirmation, disconfirmation, magnitude (damage-rating updates), path character (length, width), or event-type reclassification (e.g., from tornado to thunderstorm wind, or vice versa), as damage surveys and other quality control occur locally.

Not all tornadoes appear in LSRs. Some may not be confirmed (or disconfirmed) until days to weeks later, as previously unknown tornadic damage, media and social-media reports, and/or public photo/video documentation reach WFOs. Tornadoes not already logged in LSRs also may be included in WFO public information statements, web pages and direct contributions to NCEI *Storm Data*, bypassing the LSR stage altogether. This contributes positively to tornado numbers, compared to LSRs. Local surveys and mapping finalize tornado paths and consolidate duplicate reports of the same tornado. This contributes negatively to total tornado counts, though surveys also may discover previously unknown tornadoes. Full preliminary data are available starting in 2005. From 2006 onward, tornado LSRs as a whole have overstated "final" whole-tornado numbers (Fig. 1) by an average of 308 reports per year.

NCEI collects monthly WFO reports of tornadoes and other severe weather. Those have been filed into both a legacy *Storm Data* monthly publication—discontinued after November 2018—and a Storm Events Database website that can be polled for specific events or groups of events by geographic area or keyword. Since the same data inform both NCEI portals, *Storm Data* hereafter can refer to either. WFOs provide tornado-path records by county, and they remain segmented that way in *Storm Data*. Therefore, as of this writing, *Storm Data* represents whole tornado paths only when entirely within a county. *Storm Data* inflates tornado counts because it segments paths crossing county lines. The Storm Prediction Center (SPC) gathers *Storm Data* tornado segments yearly and stitches them together into single-tornado tracks, forming a dataset known as "ONETOR".

Storm Data and ONETOR explicitly offer these attributes for county segments and whole-tornado paths (not all-inclusive): date, tornadogenesis time in local standard time (Storm Data) or CST (ONETOR), state, damage rating, human injury and fatality counts, coded economic crop and non-crop losses, starting and ending latitude and longitude, path width in yd, and path length in mi. Damage ratings came from the Fujita (F; Fujita 1971, 1981) scale prior to February 2007, and the enhanced Fujita (EF; Doswell et al. 2009; Edwards et al. 2013; Marshall et al. 2022) scale since. Coded cost estimates can be adjusted for inflation and wealth in postprocessing (Brooks and Doswell 2001) to normalize economic-impact estimates of tornadoes over time, but ONETOR and Storm Data use available estimates in ranges of event-year dollar values. A full documentation of ONETOR formatting, as of this writing, can be found on the SPC WCM website, along with finalized

---

[1] *Corresponding author address:* Roger Edwards, Storm Prediction Center, National Weather Center, 120 Boren Blvd #2300, Norman, OK 73072; E-mail: roger.edwards@noaa.gov

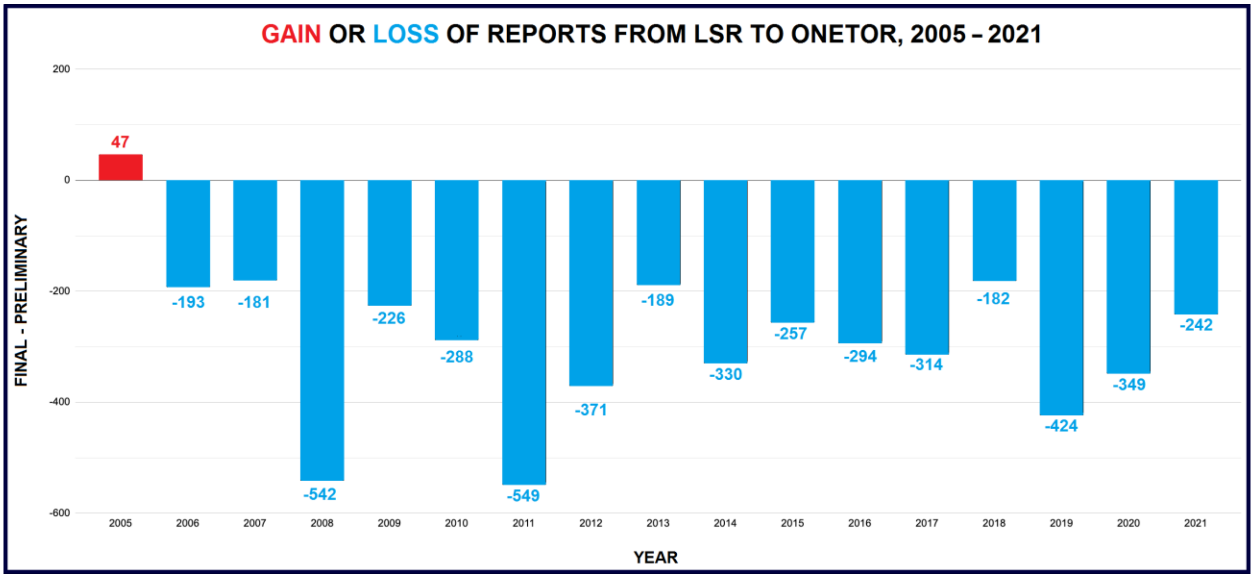


Figure 1: Bar chart showing difference in yearly tornado counts between final ONETOR and the preliminary "rough log" of LSRs that form the daily SPC website plots. Blue (red) is net loss (gain) from LSRs to ONETOR.

Table 1: 2016 algorithm applied to 1953–1982 tornadoes originally rated F-unknown (–9) in ONETOR. Dollar values derive from *Storm Data* categories, not inflation- nor wealth-adjusted as with Brooks and Doswell (2001).

IF property loss is equal to:	**THEN set F-scale equal to:**	**IF path length <=5 miles add:**	**IF path length >5 miles add:**
0,1 (<$50)	0	0	+1
2,3 (up to $5K)	1	-1	+1
4,5 (up to $500K)	2	-1	+1
6,7 (up to $50M)	3	-1	+1
8,9 (up to $5B)*	4	-1	+1

** No F=-9 tornado records met the 8,9 property loss criteria.*

ONETOR data since 1950, in both comma-separated value (CSV) and geographic information systems (GIS)-compatible formats.

The nationwide deployment of the WSR-88D network in the early–mid 1990s, as part of NWS modernization and restructuring, increased emphasis on warning verification and storm-spotter ground truth. Most of the radars were operational by 1995, which also marks the start of the era of using max path width instead of mean width for tornado records (Agee et al. 2014). This also was conterminous with the dawn of both the World Wide Web and the growth of cellular communications for more expedient reporting of all tornadoes—but especially weak ones that might have gone unreported in the pre-WSR-88D era. Documentation of weak (especially F/EF0) tornadoes increased markedly (e.g., Verbout et al. 2006), though upper-end strong to violent (F/EF3–5) tornadoes did not. The weak-tornado report inflation represented a climatological "shock" to the data (Thorne and Vose 2010; Edwards et al. 2021, hereafter EBC21). Since then, a period of relatively stable *Storm Data*/ONETOR characteristics mostly has persisted. As such, and to follow a related study (EBC21) on bulk tornado-data changes associated directly with the F/EF transition, 1995–2021 reports for the conterminous U.S. are emphasized herein.

Finally, in 2016, SPC retroactively rerated 1864 F-unknown ONETOR entries from 1953–1982, long before the current, operational EF-unknown category. The technique involved recorded economic-loss bins and path length as described in Table 1. Of the F-unknown entries, 1038 (55.5%) became F0, 742 (40.1%) became F1, 26 (1.3%) became F2, 52 (2.7%) became F3, and 6 (0.3%) became F4. No F5 tornadoes were added. Pre-2016 versions of ONETOR lack this adjustment, so an "fc" column was inserted, with a value of 1 identifying each of the rerated tornadoes. These changes also are documented in an untitled online file with no listed authorship, but written by G. Carbin.

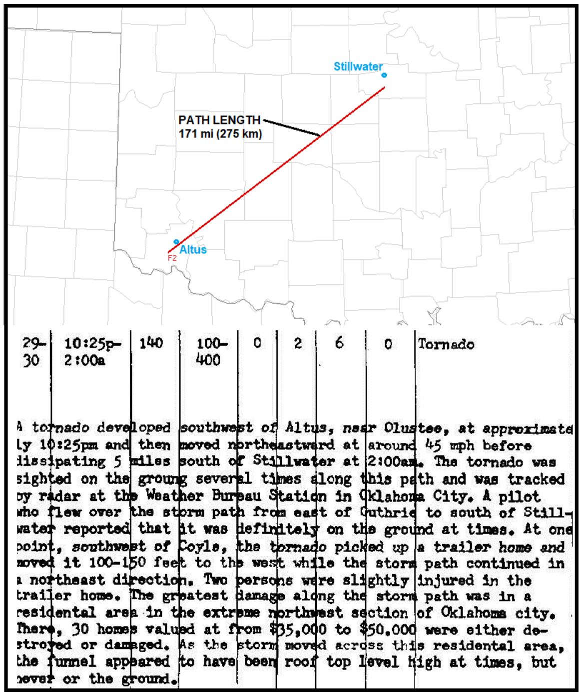
In section 2, we present examples of various kinds of apparently secular (nonmeteorological) oddities, errors and artifacts in ONETOR, especially 1995–2021. Section 3 discusses implications of such anomalies for data users and researchers.

2. DATA ANOMALIES

a. Pre-1995

Although the emphasis of our analyses is tornado data since 1995, crucial and potentially erroneous aspects of the data already documented in the literature will be summarized in this subsection, for the benefit of researchers either comparing eras or analyzing across them. Kelly et al. (1978), Grazulis (1993), Schaefer and Edwards (1999), Speheger (2001), Verbout et al. (2006), Agee et al. (2014), and Grazulis (2022, personal communication) elucidated anomalies and shifts in the pre–1994 "official" (ONETOR) data that may affect bulk analysis, depending on sampling, as summarized here:

- Kelly et al. (1978) "meticulously screened" and permanently eliminated as "fallacious" 20% of what now would be ONETOR reports discussed in Pautz (1969), but were ambiguous on the criteria for removal. "Very few tornadoes were added or deleted" based on further cross-referencing with local newspaper accounts, found by paid research assistants in each state. The accordingly filtered data formed the foundation (and remain part of) modern ONETOR. Characteristics of removed tornado reports now are largely unknown.
- Tornado-rating practices using the Fujita (F) scale shifted through the pre-WSR-88D era, enough to influence the bulk characteristics of rating categories.
- NSSFC contracted students to check records from before about 1978 (the timeline varying by state) against newspaper accounts and photos. This resulted in an unknown total number of changed ratings (often becoming overratings), and adjustments to times, dates and counties of tornadoes based on "judgment calls".
- Grazulis was contracted to reconcile differences between his data, Fujita's University of Chicago tornado dataset and often student-rated paper forms input to ONETOR. Rating differences <1 F-scale level were disused for ONETOR purposes, at the insistence of the era's NSSFC Techniques Development Unit chief. Most of the unused rating changes were downgrades from entries originally classified as "significant" (≥F2). One accepted one-level downgrade was the Lubbock, TX tornado of 11 May 1970 from its original assignment of "F6" (a nonexistent rating) to F5.
- Archaic *Storm Data* descriptions of tornadoes as "treetop level" or "rooftop level" (e.g., Fig. 2, and in Barnes 1978) may indicate either 1) true path discontinuity or 2) simple misunderstanding of what a tornado is (by ground-contact definition).
- Paper-form ratings of tornadoes commonly had stated justifications inconsistent with damage indicators of then or today. An example is the Valley Mills, TX tornado of 6 May 1973, rated F5 explicitly because it threw a pickup truck ½ mi (0.8 km). Vehicles were not an official damage indicator in the F or EF eras (Edwards et al. 2013), though this is expected to change with the next EF version (Marshall et al. 2022). Numerous other tornadoes with similar vehicular effects (e.g., El Reno, OK, 31 May 2013; Wurman et al. 2014) have not garnered F/EF5 ratings.
- Counties in multiple states have been incorrect in FIPS coding instances in ONETOR and Storm Data, with obviously affected counties (via geography or text descriptions) missing from either. This sort of error occurred alongside excessive counties listed without evidence of being affected, and/or inconsistencies between database latitude/longitude and listed counties (Speheger 2001).
- Location discrepancies remain between 16 Oklahoma *Storm Data* and ONETOR entries for this era, but with insufficient evidence to judge which is correct. The national extent of such irregularities has not been investigated, but likely are not confined just to one state.



29- 30	10:25p- 2:00a	140	100- 400	0	2	6	0	Tornado

A tornado developed southwest of Altus, near Olustee, at approximate ly 10:25pm and then moved northeastward at around 45 mph before issipating 5 miles south of Stillwater at 2:00am. The tornado was sighted on the groung several times along this path and was tracked y radar at the Weather Bureau Station in Oklahoma City. A pilot who flew over the storm path from east of Guthrie to south of Still-water reported that it was definitely on the ground at times. At one oint, southwest of Coyle, the tornado picked up a trailer home and moved it 100-150 feet to the west while the storm path continued in a northeast direction. Two persons were slightly injured in the trailer home. The greatest damage along the storm path was in a residental area in the extreme northwest section of Oklahoma city. There, 30 homes valued at from $35,000 to $50.000 were either de-stroyed or damaged. As the storm moved across this residental area, the funnel appeared to have been roof top level high at times, but ever or the ground.

Figure 2: Plotted ONETOR path (top) and screenshot of *Storm Data* entry (bottom) for a purportedly 171-mi (275-km) Oklahoma tornado path from 0430–0800 UTC 30 April 1973. Altus and Stillwater—near the beginning and end of the path—are mapped.

Additionally, In analyzing *Climatological Data*/*Storm Data* tornado entries from 1953–1972, Howe (1974) discovered improbable path dimensions attributable to clerical error, including: "…the 100-mi-long Missouri tornado listed as 8 yards wide and the 7000-yard-wide New York tornado 90 mi long." Howe also excluded an unspecified number of tornadoes for which only one dimension (length or width) was missing. He also suggested, similarly to the "tornado family" categorizations of Grazulis (1993), that numerous tornadoes described in the comments as "skipping" or "hopping" were multiple different tornadoes, instead of the longer paths entered in official data of that era.

In an NWS WFO Norman County Warning Area-based examination, Speheger (2001) noted seven tornadoes in Oklahoma with similarly exaggerated lengths, based on Storm Data verbiage implying gaps in damage during "skipping" or as "funnels". For example, a tornado on 30 April 1973 (UTC time) supposedly traveled 171 mi (275 km) from southwest of Altus to south of Stillwater (Fig. 2), and is plotted in ONETOR as continuous, despite the text description of the tornado as being "on the ground at times". Only two damage locations are mentioned explicitly along the entire path, one of which (in northwestern

Oklahoma City) *Storm Data* described this way: "The funnel appeared to have been roof top level high at times, but never or [sic] the ground," despite damaging about 30 houses. Barnes (1978) and Grazulis (1993), by contrast, discussed two separate significant tornadoes along or near this path after midnight CDT (0500 UTC). Grazulis rated both F2—the first as a 47-mi (76-km) "family" in Grady, Canadian and Oklahoma Counties, the second as a different 3-mi (4.8-km) path very near the first in Oklahoma County (the damaged houses). Another tornado on 25 May 1973 was recorded as continuous across Pontotoc and Seminole Counties despite the following *Storm Data* description: "A tornado touched down briefly 4 miles southwest of Konawa...The funnel went aloft before reaching Konawa and then touched down again 1 mile northeast of Konawa." This entry also missed an intervening county (Pottawatomie) that would be necessary to include, if the path were truly continuous.

In short, *tornadoes cannot skip*. A new path is, by strict physical definition involving ground contact, a new tornado. Truly discontinuous paths (as opposed to inaccessible and unknown, or continuous but non-damaging due to lack of damage indicators) must be multiple tornadoes, even if from the same supercell or even the same mesocyclone. A correctly documented example of the latter is the spatiotemporal gap from the F3 "Chickasha" tornado to the F5 "Bridge Creek/Moore" tornado on 3 May 1999, shown by survey (Speheger et al. 2002), and seen by numerous eyewitnesses, including this preprint's lead author. Still, Speheger et al. (2002) acknowledged that uncertainty exists in tornado-path discontinuity, such as when the same apparent tornadic vortex (with or without visible funnel) stops producing damage, but persists, then causes damage again. Literal adherence to the ground-contact definition of a tornado means that one *apparent* tornadic vortex could yield two or more tornadoes via survey. This, and other interpretive path aspects, influence an unknown number of path dimensions, particularly in the pre-1995 period when in-person NWS surveys were much less common and consistent (Schaefer and Edwards 1999).

A thorough filtering for all text descriptions of single-tornado path discontinuities in *Storm Data* comments is beyond the scope of this work, but we have found them randomly throughout the pre-1995 era, and even on a few entries since. Forensic analyses (e.g., Ostuno 2008, for the Lower Michigan event of 3 April 1956) have uncovered others. For bulk-analysis purposes, the extent is unknown that the overstatement of lengths resulting from "skipping", and the underrepresentation of curved paths due to straight-line ONETOR plotting (next section), can counterbalance each other. Even in the pre-*Storm Data* era, forensic work has uncovered path discontinuities that render uncertainty to notable tracks (e.g., the Tri-State tornado of 18 March 1925; Johns et al. 2013). Other discontinuities and close separations, such as satellite tornadoes (Edwards 2014), were documented rarely and poorly prior to 1995, also likely due to a relative lack of both surveys and direct eyewitness reports.

In addition to width and length, path locations and/or orientations sometimes are incorrect. Errors in tornado start or end points likely exist in the pre-1995 data, as well as the period since (next section, part 3), given that the city-location data used for *Storm Data* predate the "modernized" NWS era. A relative dearth of comparative surveys, however, makes finding such paths challenging. While imperfect in terms of likelihood of detection, mapping of multiple tornado paths from the same day's event can help researchers to find anomalous path positions, lengths and/or alignments for deeper scrutiny.

b. WSR-88D era: 1995–2021

1) Pre-analysis filtering

The errors and anomalies described below also appeared in the 24-y 1995–2018 subset of EBC21 tornado data—some of which also were noted by them. For analysis purposes herein, the raw 1995–2021 ONETOR data were filtered to remove 35 tornadoes outside the conterminous U.S. (CONUS) and all state-segment duplicates.

In ONETOR, each tornado-state segment appears as a separate entry (row), usually bracketing a combined row representing the true starting and ending point, with the starting state labeled. As long done for the official dataset of tropical cyclone (TC) tornadoes (Edwards 2010; Edwards and Mosier 2022), we eliminated all single-state segments (rows) and kept the whole-path entry, then added the postal abbreviation for the second state (and if applicable, third) with slashes in the "state" column. For example, three rows covering a tornado crossing from Texas into Oklahoma becomes a single row labeled TX/OK in our modified analysis dataset, truly representing one tornado. A few tornadoes crossed three state lines. The most segmented entry was for a *two-state* tornado on 30 July 2009. That tornado crossed the frequently bending state line—which follows current and former Mississippi River channels—four times, starting and ending in Arkansas. That event yielded four segments (removed) plus the entry for the whole tornado, and was labeled AR/MS/AR/MS/AR in the state(s) column.

A total of 346 tornadoes crossed a state line at least once, seven of which affected parts of three states each. Only the full entry for each multistate tornado was retained. This process removed 715 excess "tornadoes" (state segments) from the data—an average of about 26/y. Maximum path width was greater in five state segments than their combined listing; in each case, the latter entry was reassigned the greater segment width. Three multistate entries had endpoints erroneously at the state line, instead of within the next state; their true ending latitude/longitude pairs were pasted into the full entry from the second state's segment, before deleting the latter. Five raw multistate entries were missing one of two possible state segments already. One entry had no state segments, but its genesis and dissipation points crossed state lines (path already was combined in the data, contrary to convention). Ten segments erroneously had rows of zeroes for all these categories: start latitude, start longitude, end latitude, end longitude, length, and width. Of those, two segments contained F/EF4 ratings: one in Minnesota on 29 March 1998, and the other in Mississippi on 24 April 2010. Seven more state segments had zeroes for width.

Combining state-crossing tornadoes is logically consistent with combining county-crossing segments, if the aim implied by the ONETOR name is a true "one tornado" dataset agnostic to geopolitical artifice. Users performing state-segment filtering should note that chronological (and thus row) separation sometimes exists between state-segment listings (e.g., when other tornadoes begin before the timeline of the second or third state segment). As such, some state-segment rows of a single tornado are not adjacent to each other in the raw ONETOR data.

For latitude and longitude, 31 (unsegmented) tornadoes had values of zero, which literally would place them in the Gulf of Guinea, off Africa's Atlantic Coast. Of those, two were significant (EF2 and EF3, as defined by Hales 1988). One of the filtered zero-latitude/zero-longitude events had a zero for number of states, even though it did occur in one: Mississippi (24 February 2001 in CST, 25 February in UTC). Of the remaining whole tornadoes, 555 still had unphysical zero width and/or length (Fig. 3), and also were filtered for analysis purposes. Of those, 361 (65%) were in one year, 1999; no known explanation exists for that oddity.

Figure 3: Locations of 555 whole ONETOR paths with zero length and/or width, 1995–2021.

After the filtering described above, 32 775 whole-tornado, CONUS-only entries remained, averaging 1214 per year. Several forms of oddities and apparent errors remained, which will be discussed in the next subsection.

2) Quantifiable changes, errors and anomalies

Perhaps the most prominent data change in the post-1995 era involves apparent impacts of the EF scale on tornadic path characteristics already documented by EBC21, which are summarized here as well. Though the EF scale was intended to calibrate with the F scale and have minimal cross-transitional data impact (Edwards et al. 2013), this was not necessarily the case. The greatest of the changes involved significant, systematic increases in path width, across multiple time and rating bins, following the onset of the EF scale in 2007 (Fig. 4). Lesser, but still noticeable increases occurred in path length, especially for weak tornadoes, such that (for example) *no* 3-y bin in the EF era had a shorter average weak-tornado path than *any* 3-y bin in the F era. No physical explanation exists in the real atmosphere for tornadic size growth so abruptly across the F/EF transition. That indicates a profound secular influence whose true cause(s) remain in the realm of speculation. As noted in EBC21, further changes to the EF scale (e.g., Marshall et al. 2022) also may affect the path data in unknown or unplanned ways. Researchers analyzing tornado data across any such EF upgrade should be ready to document and account for any associated data discontinuities.

The rating category EF-unknown (EFU) is intended for tornadoes previously rated EF0 that are 1) without known damage, 2) so remote as to be inaccessible for surveying, and/or 3) damaging only nonstandard indicators to which a rating cannot be applied. This category was valid before 2013 (Edwards et al. 2013), but unused in practice until 2016 due to software constraints and lack of an explicit policy directive (NWS 2021). Yearly EFUs nationwide increased by an order of magnitude starting in 2019, for unknown reasons (Fig. 5a). This category has been concentrated strongly in the central and southern Plains States, as well as Iowa and Illinois (Fig. 5b). Those are areas of relatively high climatological tornado frequency, open land and low density of EF damage indicators. Since state segments already are filtered, the total EFU count of 606 (8.1% of all tornadoes in the same years) does not include an EFU state segment on 15 December 2021, for which the whole tornado path earned a numeric rating (thanks to another state segment rated EF0). Perhaps the most anomalous EFU path is the longest one: a 23.2-mi (37.4-km) long, 40-yd (36.6-m) wide entry in Wyoming on 12 June 2017. An EFU in northeastern California, on 15 August 2020, had an unusually precise max width of 156 yd (143 m). The mean path length and width for EFUs were 1.2 mi (2 km) and 51 yd (47 m), respectively.

Physical improbabilities define some other apparent errors in the data. For example, 48 tornadoes have a listed maximum path width of 1 yd (0.9 m). Of those, 33 are rated EF0, with ten EFU, four at EF1, and one EF2. The latter, occurring on 12 August 2004 in northern Florida in tropical cyclone (TC) Bonnie, is the highest-rated of any tornado <10 yd (9.1 m) wide. All three of Bonnie's Florida tornadoes were entered as 1 yd wide, unlike 13 tornadoes from the same system in other states. Another 1-yd-wide TC tornado, from Jeanne on 26 September 2004, also in northern Florida, has a listed path length of 21 mi (34 km). These indicate local data-entry errors and not true widths. Nine more ONETOR tornadoes are listed at 2 yd (1.8 m) wide. The longest-path TC tornado (Edwards and Mosier 2022, this conference) had an improbable combination of 52-mi (84-km) path length and 10-yd (9.1-m) max width.

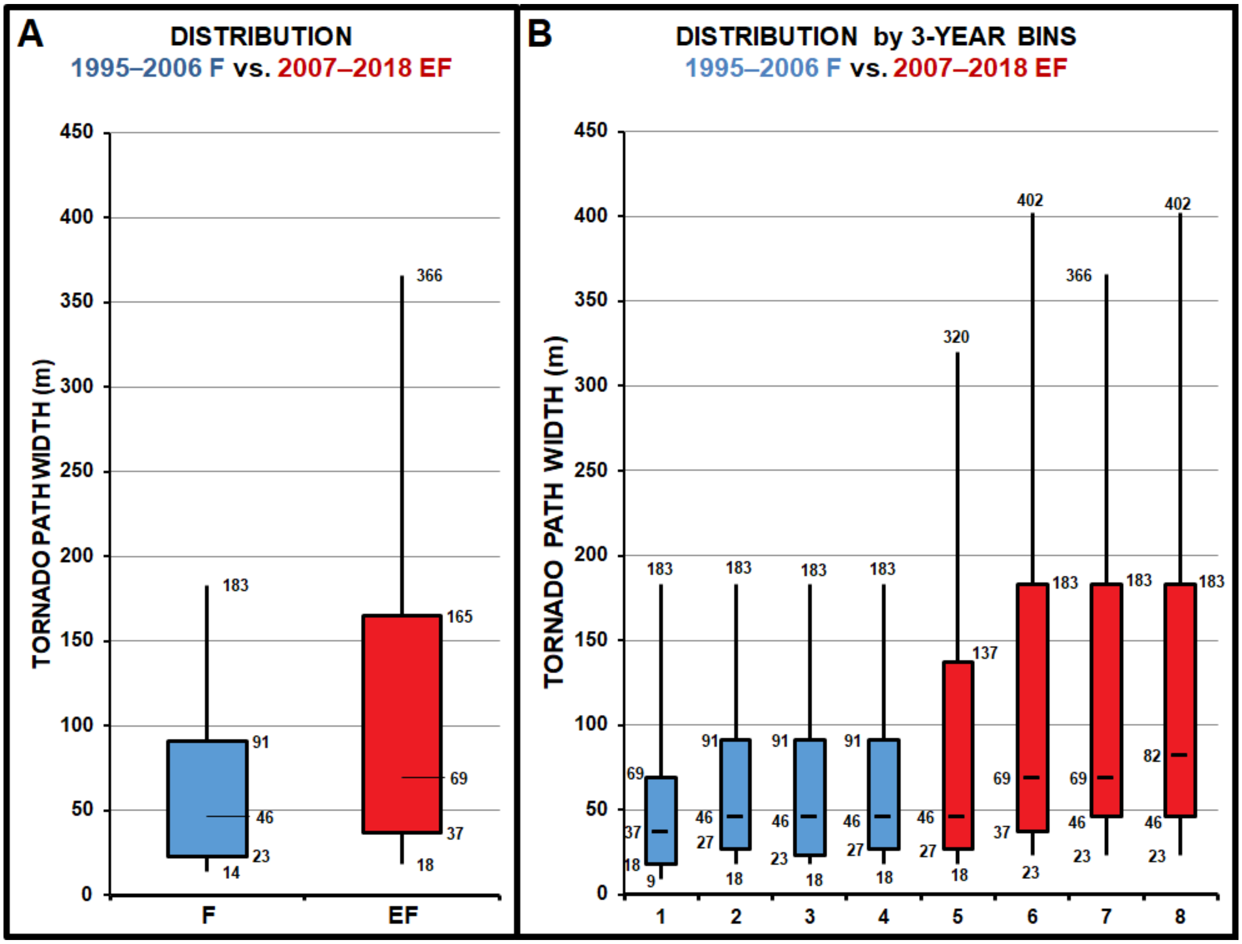


Figure 4: Figure 6 from EBC21, showing distributions of path width (m) in the F (blue) and EF (red) portions of the period 1995–2018, in increments of a) 12-y and b) 3-y. Boxes represent 25th and 75th percentile, with whiskers to 10th and 90th percentiles, all with values labeled. Labeled median bars reside in the boxes.

3) Artifacts of uncertain quantity

Clerical data-entry errors have continued into the WSR-88D era. For example, radar-echo comparisons by Smith et al. (2012) and Edwards et al. (2012) revealed one-hour time errors for a minority of nontropical and TC tornadoes, respectively. Such errors most commonly arise either from conversion failures between daylight and standard time, or in WFO jurisdictions straddling two time zones. During their research on long-path tornadoes, Straka and Kanak (2022) discovered an 80-mi (127-km) track on 22 April 2004, in one Oklahoma county whose maximum dimension cannot fit that length. Its ultimately correct latitude/longitude coordinates indicated a 0.8-mi (1.3-km) length—a difference of exactly two orders of magnitude. One of that paper's reviewers, also the lead author of this one, documented the discrepancy to be a clerical error upon consultation with the WFO warning coordination meteorologist (E. Calianese, personal communication); the width in yd had been entered inadvertently as the length in mi. The result was submitted to *Storm Data* for a much-belated correction. Less-obvious data-entry errors of this sort—whether transpositions or mere typographical mistakes—are difficult to detect using automated means, and are of unknown frequency. As such, only painstaking manual quality control or happenstance will reveal most clerical errors.

As noted in section 2a, latitude and/or longitude errors have been discovered in some listings of the city database used for *Storm Data*. Most (but not all) of these found so far have been in Oklahoma and New Mexico, given the NWS WFO Norman data domain that is quality-controlled by one of the authors (DAS), and efforts at NWS Albuquerque in 2013 to correct sometimes extreme errors in that state (J. Shoemake, personal communication). Misplaced reference cities have persisted into the WSR-88D era, and affect tornado paths when one or both endpoints are converted from city-referenced azimuths and ranges in *Storm Data* to latitude and longitude. An outstanding example is in Fig. 6, which compares the incorrect ONETOR path (resulting from a bad latitude for the tornadogenesis reference town) to the actual, surveyed track (the *Storm Data* text description is correct). The *Storm Data* town location was incorrect for San Jon, NM, affecting tornado points referencing that town, such as a case on 14 August 2004. A correction was submitted to Storm Data for that event, but in the final ONETOR data, only the ending point was corrected. The beginning was not.

The largest town-location error found anywhere so far was for El Huerfano, in northwestern New Mexico

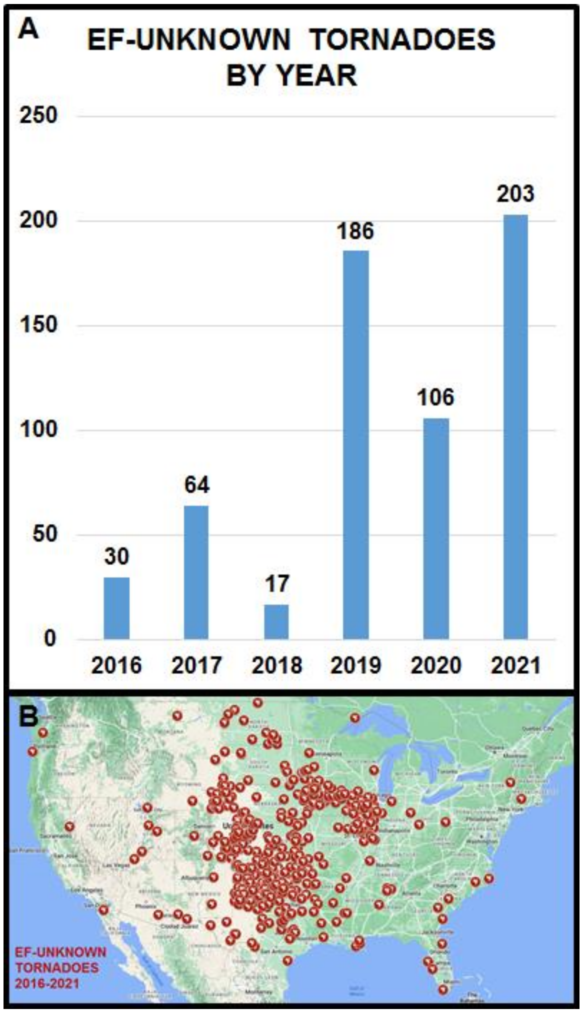


Figure 5: EF-unknown (EFU) tornadoes as: a) bar chart by year, since initial use in 2016, and b) mapped over the CONUS. Background map courtesy Google.

near Farmington, misplaced 275 mi (443 km) due south of its real location by another apparent latitudinal-entry error in the cities data (not shown). Fortunately, no known tornado records referenced that town. A tornado near Butterfield, MN, on 8 May 2014, has the correct path latitude and longitude in ONETOR, but an incorrect town location and reference in the *Storm Data* publication (not shown). In some instances, *Storm Data* and ONETOR positions are inconsistent, but it is unclear which one is correct, if either. The Geary, OK, location error in Fig. 6, and other known New Mexico and Oklahoma city-placement errors, have been fixed in *Storm Data*. Figure 6 also illustrates the shape and length difference between rigidly linear ONETOR tracks and surveyed paths—which can be complex and curving. The number of erroneous ONETOR paths that are based on city-location errors also is unknown nationally. So is the true difference between actual and ONETOR path length for most tornadoes.

3. SUMMARY and DISCUSSION

The objectives of this effort do not include documenting every individual glitch, nor uncovering all possible anomalies or kinds of errors in the data. In fact, the nature of some errors (e.g., undetected typos such as mis-entry of one digit in a fractional path mile, latitude, or longitude, erroneous time conversions, or other times and location errors not revealed via radar checks) renders complete error discovery and documentation practically impossible. Instead, we illustrate types and examples of data problems conducive to ready filtering, along with oddities that are interesting and potentially may contaminate analyses. Small subsampling increases the potential impact of such errors on tornado-data analysis and interpretation (Doswell 2007). The goal is to motivate the variety of users of the data—from climatic and severe-weather researchers to casual users, insurance, reinsurance and risk-reduction interests, and the scientific news media—to examine the data with fullest possible understanding of its vagaries, to exercise due diligence and caution in its use, and to filter and/or detrend the data reproducibly.

Future format modernization, event reexaminations, error corrections, and richer metadata texturing should *reduce* problems in the data. Preliminary plans have been devised to convert ONETOR to an open-source data format, more flexible to revision. Under this concept, evidence-based submissions of corrections, additions and subtractions would be reviewed by a small, expert committee, which then would make approved changes. Nonetheless, a presence of some historic errors, and occasional future ones, are inevitable in a dataset gathered by hundreds of fallible people across several decades, containing ~10^4 reports as of this writing, and likely reaching ~10^5 events in the 2030s at a rate of ≈1200 per year. Uncertainties are inherent to storm datasets derived through subjective influences, and not just with tornadoes (e.g., Torn et al. 2013, for hurricane data).

We strongly caution tornado-data analysts to quality-control the data for themselves, line by line via either automated or manual means if feasible, for anomalies of any type that may contaminate their results. Researchers should document all such procedures thoroughly in their reports or journal papers, for reproducibility's sake. For example, records with path lengths or widths of zero, or with length less than width, should be located and expunged from the analytic dataset, where informing bulk indices or other examinations of path dimensions. Examples include use of the Destruction Potential Index (DPI; Thompson and Vescio 1998), or statistical applications to estimate tornadic "power" or intensity from path characteristics (e.g., Elsner et al. 2014b). Zero path-length or -width values also are incompatible with logarithmic analyses (e.g., of highly nonlinear variables like DPI, as in the logarithmic-ordinate boxplot of EBC21's Fig. 17).

Users of ONETOR also should understand that, despite or within guidelines in both F and EF scales, *tornado ratings are subjective* (Doswell and Burgess 1988; Edwards 2003; McCarthy 2003; Edwards et al. 2013). Because tornadoes are rated by their peak damage level—even if just one such damage indicator is found somewhere in the path—a tornado's nominal rating alone conveys little information about:

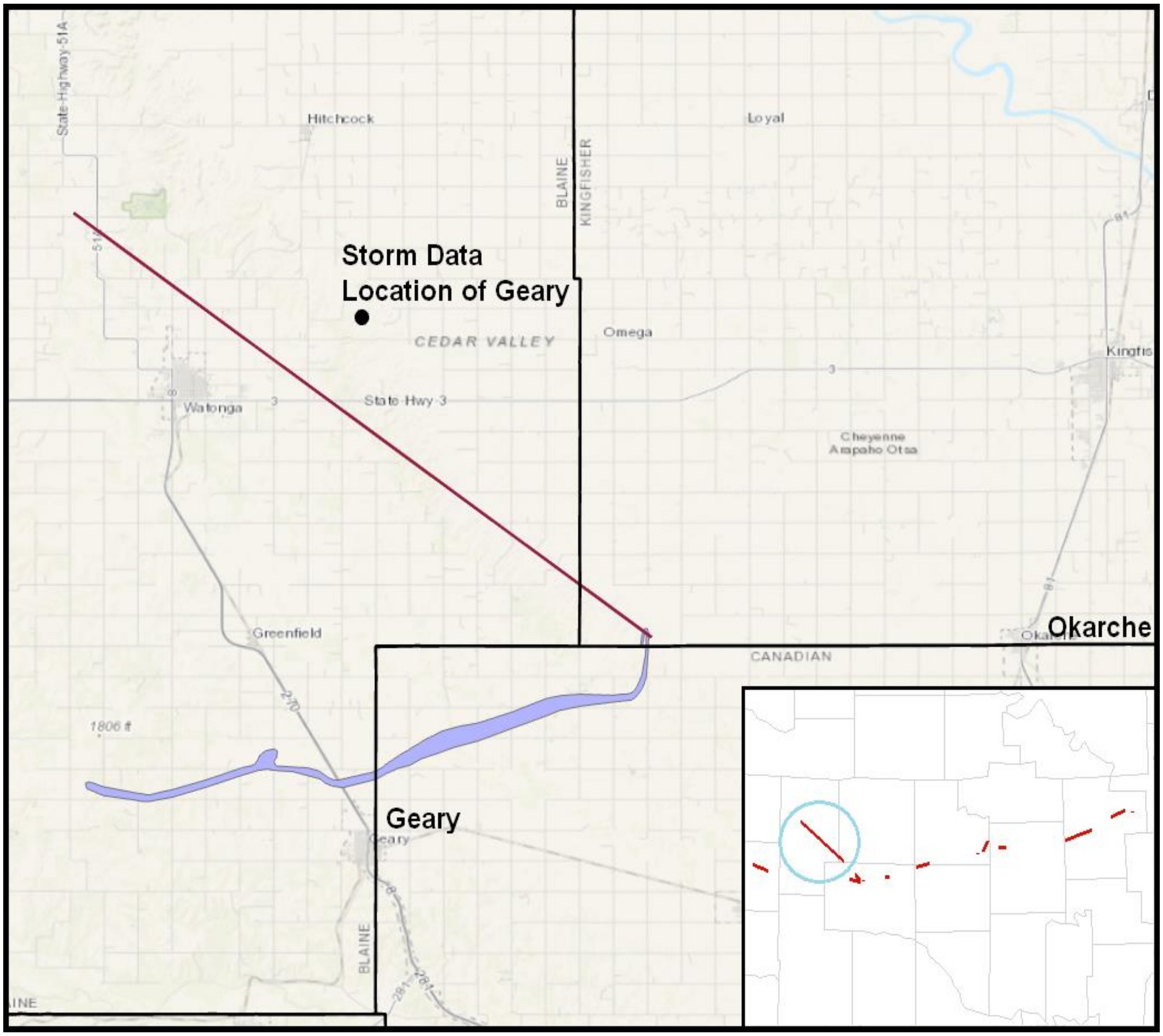


Figure 6: Tornado path (purple) of 29 May 2004 "Geary tornado", as surveyed from 9 mi (14.5 km) west-northwest of Geary, OK, to 11 mi (17.7 km) west of Okarche, OK. The ONETOR path (red) is shown with the beginning point displaced 18 mi (29 km) too far north, due to incorrect latitude of Geary in the *Storm Data* cities database at the time the report was filed. Reference towns and false *Storm Data* location of Geary labeled in bold black. Counties labeled in gray along black county lines. Small squares in the main map are 1 mi (1.6 km) across. White inset map shows ONETOR Geary path (circled) in context of others from the same supercell.

- Variations in damage and implied windspeed ranges elsewhere in the path,
- Total amount of damage, or
- The texture of damage ratings or wind estimates over an entire path.

Single-indicator ratings, especially on long paths, also may cause overrepresentation of tornadoes when using implied indices of “power” or “energy dissipated” for a whole tornado.

In a similar vein, tornado *damage rating* also should not be conflated in scientific writing with tornado *intensity*. An F or EF rating only represents a remotely assessed or directly surveyed *estimate of windspeed ranges* at individual point(s) in the path (Doswell and Burgess 1998; Edwards et al. 2013 respectively). Finescale mapping and geotagged documentation of damage and path curvature has become more common over the past decade with the Damage Assessment Toolkit (Camp et al. 2014), but was rare before the 2010s, and has not been done for all tornadoes since.

Meanwhile, ONETOR has maintained an often-misleading, strictly linear path-mapping process by offering only a starting and ending point for each tornado. Paths plotted therefrom, that have accurate start and end points (which is not a given, as we have shown) are the shortest possible. As such, *tracks plotted from ONETOR underrepresent true length for paths containing any curvature*. By extension, *ONETOR should not be used for precise calculations of tornadic translation speed*, which can vary greatly along a path (e.g., Wurman et al. 2006; Wurman et al. 2013), and which the straight line will exaggerate for a curved path. While an automated algorithm has yet to be applied to ONETOR to detect suspicious path characteristics—especially anomalous orientations relative to others on a given day (e.g., Fig. 6)—use of machine-learning training datasets to find and flag suspicious historic paths for further investigation may be a substantial time- and effort-saver in the future.

Furthermore, documented systemic changes in data-gathering and recording practices (e.g., Schaefer and Edwards 1999; Verbout et al. 2006; Agee and Chlds 2014) and damage-rating procedures (e.g., Doswell et al. 2009; Edwards et al. 2013; EBC21) can impart major shifts, discontinuities across time, or "shocks" (Thorne and Vose 2010) to the data from their inception onward. Researchers using bulk tornado data should be aware of, and explicitly address the effects of, data shocks in their own work, and as applicable to statistical objectives, numerically normalize for them.

Based on data problems documented here, and in the various citations throughout, we recommend an overarching best practice for scientific use of ONETOR. ***In all tornado-data analyses, the presence of errors and artifacts introduces uncertainty that should be divulged and, to the greatest extent possible, quantified.***

ACKNOWLEDGMENTS

We gratefully acknowledge previous stewards of ONETOR, including those before 1995 at SPC's predecessor in Kansas City, for making its existence possible. They laid a progressively deeper foundation for ongoing improvements and future format modernization (in alphabetic order): Greg Carbin, Joe Galway, Leo Grenier, John Halmstad, Dave Higginbotham, Don Kelly, Larry Lee, Dan McCarthy, Allen Pearson, Joe Schaefer, and Dave Swenson. The impact of their stewardship is a substantial net positive, enabling the presence of a large, nationwide, whole-tornado dataset. The glitches discussed herein are not intended to detract from their monumental efforts and legacy in building and maintaining the data, but instead, to motivate research diligence and further data improvement. Since the late 1980s, Tom Grazulis has given the lead author numerous insights into tornado history and data, and specific to this effort, contributed a sampling of original, 1970s-era NSSFC rating/rerating worksheets for tornadoes. Harold Brooks, Ed Calianese, Andy Dean, Chuck Doswell, Chris Nowotarski, Thea Sandmael, Russ Schneider, Jennifer Shoemake, Bryan Smith, Justin Spotts, and Steve Weiss, have engaged us with stimulating and insightful discussions, attention to errors and oddities, and suggestions for improvement to the data. Israel Jirak (SPC) provided a beneficial review of this paper.

REFERENCES

Agee, E., and S. Childs, 2014: Adjustments in tornado counts, F-scale intensity, and path width for assessing significant tornado destruction. *J. Appl. Meteor. Climatol.,* **53**, 1494–1505.

Barnes, S. L., 1978: Oklahoma thunderstorms on 29–30 April 1970. Part III: Tornado characteristics inferred from damage tracks. *Mon. Wea. Rev.*, **106**, 697–703.

Brooks, H. E., and C. A. Doswell III, 2001: Normalized damage from major tornadoes in the United States: 1890–1999. *Wea. Forecasting*, **16**, 168–176.

——, G. W. Carbin, and P. T. Marsh, 2014: Increased variability of tornado occurrence in the United States. *Nature*, **346**, 349–352.

Camp, J. P., L. Rothfusz, A. Anderson, D. A. Speheger, K. L. Ortega, and B. R. Smith, 2014: Assessing the Moore, Oklahoma (2013) tornado using the National Weather Service Damage Assessment Toolkit. Proc., *Special Symp. on Severe Local Storms: The Current State of the Science and Understanding Impacts*, Atlanta, GA, Amer. Meteor. Soc., 830.

Coleman, T. A., and P. G. Dixon, 2014: An objective analysis of tornado risk in the United States. *Wea. Forecasting*, **29**, 366–376.

Doswell, C. A. III, 2007: Small sample size and data quality issues illustrated using tornado occurrence data. *Electronic J. Severe Storms Meteor.*, **2** (5), 1–16.

——, and D. W. Burgess, 1988: On some issues of United States tornado climatology. *Mon. Wea. Rev.*, **116**, 495–501.

——, H. E. Brooks, and N. Dotzek, 2009: On the implementation of the enhanced Fujita scale in the USA. *Atmos. Res.*, **93**, 554–563.

Edwards, R., 2003: Rating tornado damage: An exercise in subjectivity. Preprints, *1st Symp. on F-Scale and Severe-Weather Damage Assessment*, Long Beach CA, Amer. Meteor. Soc., P1.2.

——, 2010: Tropical cyclone tornado records for the modernized NWS era. Proc., 25th Conf. on Severe Local Storms, Denver, CO, Amer. Meteor. Soc., P3.1.

——, 2014: Characteristics of supercellular satellite tornadoes. Proc., *27th Conf. on Severe Local Storms*, Madison, WI, Amer. Meteor. Soc., 17.5.

——, and R. M. Mosier, 2022: Over a quarter century of TCTOR: Tropical cyclone tornadoes in the WSR-88D era. Proc., 30th Conf. on Severe Local Storms, Santa Fe, NM, Amer. Meteor. Soc., P171.

——, A. R. Dean, R. L. Thompson, and B. T. Smith, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part III: Tropical cyclone tornadoes. *Wea. Forecasting*, **27**, 1507–1519.

——, J. G. LaDue, J. T. Ferree, K. Scharfenberg, C. Maier, and W. L. Coulbourne, 2013: Tornado intensity estimation: Past, present, and future. *Bull. Amer. Meteor. Soc.*, **94**, 641–653.

——, H. E. Brooks, and H. Cohn, 2021: Changes in tornado climatology accompanying the enhanced Fujita scale. *J. Appl. Meteor. Climatol.*, **60**, 1465–1482.

Elsner, J. B., S. C. Elsner, and T. H. Jagger 2014a: The increasing efficiency of tornado days in the United States. *Climate Dyn.* **45**, 651–659.

——, T. H. Jagger, and I. J. Elsner, 2014b: Tornado intensity estimated from damage path dimensions. *PLOS One*, **9**, e107571.

Fujita, T. T., 1971: Proposed characterization of tornadoes and hurricanes by area and intensity. University of Chicago SMRP Research Paper 91, 42 pp.

——, 1981: Tornadoes and downbursts in the context of generalized planetary scales. *J. Atmos. Sci.*, **38**, 1511–1534.

Grazulis, T. P., 1993: *Significant Tornadoes, 1680-1991*. Environmental Films, St. Johnsbury, VT, 1326 pp.

Howe, G. M., 1974: Tornado path sizes. *J. Appl. Meteor.*, **13**, 343–347.

Johns, R. H., D. W. Burgess, C. A. Doswell III, M. S. Gilmore, J. A. Hart, and S. F. Piltz, 2013: The 1925 Tri-State tornado damage path and associated storm system. *Electronic J. Severe Storms Meteor.*, **8** (2), 1–33.

Kelly, D. L., J. T. Schaefer, R. P. McNulty, C. A. Doswell III, and R. F. Abbey Jr., 1978: An augmented tornado climatology. *Mon. Wea. Rev.*, **106**, 1172–1183.

Marshall, T. P., T. M. Brown-Giammanco, S. Krautwurst, and N. L. de Toledo, 2022: On the current revision of the Enhanced Fujita (EF) scale. Proc., *30th Conf. on Severe Local Storms*, Santa Fe, NM, Amer. Meteor. Soc., 11.1A.

McCarthy, D. W., 2003: NWS tornado surveys and the impact on the national tornado database. Preprints, *1st Symp. on F-scale and Severe Weather Damage Assessment*, Long Beach, CA, Amer. Meteor. Soc., 3.2.

NWS, 2021: *Storm Data* preparation. Performance and evaluation, NWSPD 10-16. [Available online at https://www.nws.noaa.gov/directives/sym/pd01016005curr.pdf.]

Ostuno, E. J., 2008: A case study in forensic meteorology: Investigating the 3 April 1956 tornadoes in southwest Lower Michigan. *Electronic J. Severe Storms Meteor.*, **3** (1), 1-33.

Pautz, M. E., 1969: Severe local storm occurrences 1955–1967. ESSA Tech. Memo. WBTM FCST 12, 77 pp.

Schaefer, J. T., and R. Edwards, 1999: The SPC tornado/severe thunderstorm database. Preprints, *11th Conf. on Applied Climatology*, Dallas, TX, Amer. Meteor. Soc., 6.11.

Speheger, D. A., 2001: Corrections to the historic tornado database. NWS Southern Region Tech. Memo. NWS SR-209, 24 pp. [Available online at https://repository.library.noaa.gov/view/noaa/6487/Share.]

——, C. A. Doswell III, and G. J. Stumpf, 2002: The tornadoes of 3 May 1999: Event verification in central Oklahoma and related issues. *Wea. Forecasting*, **17**, 362–381.

Straka, J. M., and K. M. Kanak, 2022: A climatology of long-track tornadoes. *Electronic J. Severe Storms Meteor.*, **17** (1), 1–49.

Thom, H. C. S., 1963: Tornado probabilities. *Mon. Wea. Rev.*, **91**, 730–736.

Thompson, R. L., and M. D. Vescio, 1998: The Destruction Potential Index: A method for comparing tornado days. Preprints, *19th Conf. on Severe Local Storms*, Minneapolis, MN, Amer. Meteor. Soc., 280–282.

Thorne, P. W., and R. S. Vose, 2010: Reanalysis suitable for characterizing long-term trends: Are they really achievable? *Bull. Amer. Meteor. Soc.*, **91**, 353–361.

Torn, R. D., and C. Snyder, 2013: Uncertainty in tropical cyclone best-track information. *Wea. Forecasting*, **27**, 715–729.

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

Wurman, J., J. M. Straka, and E. N. Rasmussen, 1996: Fine-scale Doppler radar observations of tornadoes. *Science*, **272**, 1774–1777.

——, K. Kosiba, P. Robinson, and T. P. Marshall, 2014: The role of multiple-vortex tornado structure in causing storm researcher fatalities. *Bull. Amer. Meteor. Soc.*, **95**, 31–45.