

# Development and Analysis of a Database of Tornado Impacts on US Critical Facilities

Nico de Toledo <sup>a1</sup>, Marc L. Levitan <sup>a2</sup>, Jamil Malik <sup>a3</sup>, Warren Stewart <sup>a4</sup>, Katherine J. Johnson <sup>a5</sup>, P. Shane Crawford <sup>ab6</sup>

<sup>a</sup> National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA
<sup>b</sup> Federal Emergency Management Agency (FEMA), Washington, DC, USA
<sup>1</sup> <u>nicholas.detoledo@nist.gov</u>, <sup>2</sup> <u>marc.levitan@nist.gov</u>, <sup>3</sup> <u>jamil.m4371@gmail.com</u>,
<sup>4</sup> <u>wsstewart268@gmail.com</u>, <sup>5</sup> <u>katherine.johnson@nist.gov</u>, <sup>6</sup> <u>shane.crawford@fema.dhs.gov</u>

ABSTRACT: Despite the significant hazard that tornadoes pose, much remains to be learned about the impacts of tornadoes on critical facilities in the United States. While post-storm investigation reports have documented the damage from individual tornadoes or tornado outbreaks, there is a major gap in the availability of data describing the cumulative national impacts of tornadoes on critical facilities. In response to this gap, we are creating a database of tornado impacts on critical facilities (e.g., schools, hospitals, etc.) by mining National Weather Service narratives from the National Centers for Environmental Information Storm Events Database. The first stage in the creation of the critical facility impacts database includes documentation of preschool-12th grade schools hit by tornadoes recorded in the Storm Events Database between 1993 and 2020. To date, we have identified 669 schools that were struck by tornadoes, for an average of approximately 24 per year. The NWS narratives do not include an exhaustive list of facilities affected by each tornado, therefore it is likely that other schools were hit by tornadoes during this period but not recorded, and the true number of schools struck by tornadoes is higher. To overcome underestimation inherent in the narrative analysis, we are developing a complementary GIS method to identify critical facilities struck by tornadoes by combining tornado damage polygons from the National Oceanic and Atmospheric Administration's Damage Assessment Toolkit with various critical facility location datasets. Preliminary results from applying this method to schools indicate a high likelihood of capturing additional impacted facilities not found through mining the NWS narratives.

KEYWORDS: tornado, tornado damage, school damage, NCEI Storm Events Database, critical facilities, tornado database.

# 1 INTRODUCTION<sup>1</sup>

The geospatial footprints of single tornadoes are much smaller than those of earthquakes and hurricanes, and individual tornadoes often receive less national media coverage, so it can be difficult to grasp their overall impacts. While not well understood, the cumulative impacts of the more than

<sup>&</sup>lt;sup>1</sup> Disclaimer: Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

1,250 reported tornadoes per year in the US (NOAA 2022b) are quite significant. For instance, from 1950 through 2014, tornadoes caused more fatalities in the US than earthquakes and hurricanes combined (USGS 2016, NOAA 2022d)<sup>2</sup>, and from 1997 to 2016, tornadoes and tornadic storms caused more catastrophe insured losses than hurricanes and tropical storms combined, according to the Insurance Information Institute (2021). A critical step in mitigating these devastating impacts is to design the most important community facilities to resist tornadoes. Tornado-resistant engineering design of these critical facilities (e.g., schools, fire stations, and hospitals) is emerging in the US, with the introduction of requirements for tornado loads in the new American Society of Civil Engineers ASCE 7-22 Standard on minimum design loads for buildings (ASCE 2021). This is in addition to the expanding construction of tornado shelters<sup>3</sup> and safe rooms<sup>4</sup> at schools and emergency response facilities across the central and southeastern US. A more thorough understanding of the history and impacts of tornado strikes on critical facilities in the US will be valuable as state and local governments consider adopting building codes and standards that include tornado-related requirements and as communities and school boards consider constructing tornado shelters at new or existing schools. However, to our knowledge, there is no systematic national database documenting tornado impacts on critical facilities. This paper details our work to build such a database.

While some publications document tornado damage to critical facilities in great detail, these are typically focused on a single tornado or outbreak (e.g., NIST 2013 and 2014, FEMA 2012, Ramseyer et al. 2019). The most comprehensive source of tornado records in the US, the Storm Events Database (SED) hosted by The National Centers for Environmental Information (NCEI) (NOAA 2021), includes National Weather Service (NWS) narratives describing the impacts of all recorded US tornadoes. However, the SED is not specifically designed to document tornado strikes on critical facilities, and accounting for every impact to critical facilities mentioned in the NWS narratives requires an extensive review of the SED. Mining these narratives for critical facility impacts is the first part of our effort to construct a database of tornado impacts on these important community facilities.

The NWS tornado narratives often summarize the most severe tornado damage but do not exhaustively account for impacts on critical facilities, and there is likely an underreporting of critical facilities that experienced less damage than other buildings in the paths of the tornadoes. To account for tornado impacts on critical facilities more thoroughly, we plan to complement the NWS narrative mining with a GIS method for identifying critical facilities hit by tornadoes. This method consists of overlaying geospatial tornado damage polygons from the NWS Damage Assessment Toolkit (DAT) (NWS 2022) with critical facility location datasets.

This paper focuses on the initial results of a dual-method approach to documenting tornado impacts on critical facilities by 1) mining NWS narratives in the SED for tornado strikes on schools, and 2) performing a GIS analysis of the intersection of schools and tornado damage polygons. These initial efforts support a broader effort to understand the cumulative impacts of tornadoes on US infrastructure that will enable better decision-making to mitigate tornado impacts.

<sup>&</sup>lt;sup>2</sup> There were approximately 5,757 fatalities from tornadoes (NOAA 2022d), 459 from earthquakes (USGS 2016), and 3,107 from hurricanes (NOAA 2022d).

<sup>&</sup>lt;sup>3</sup> Tornado shelters are constructed in accordance with the ICC 500 Standard for the Design and Construction of Storm Shelters (ICC 2020, or prior editions).

<sup>&</sup>lt;sup>4</sup> Tornado safe rooms are constructed following guidance provided by the Federal Emergency Management Agency (FEMA 2021). Safe rooms must meet all ICC 500 requirements, plus several additional FEMA funding criteria.

#### 2 METHOD 1- STORM EVENTS DATABASE MINING

#### 2.1 Data Acquisition

To develop the database of tornado impacts on critical facilities, we identified the SED as the best source of tornado records because it is the most comprehensive and systematic. Many publications include granular detail on the impacts of specific tornadoes or tornado outbreaks, whereas the SED stores information for all tornadoes recorded by the NWS going back to 1950 (NOAA 2021)<sup>5</sup>.

Each tornado record in the SED specifies the date, time, and location of a tornado or "tornado segment" (NOAA 2022a)<sup>6</sup>. Beginning with the year 1993, the SED also includes NWS narratives describing the impacts of the tornadoes (referred to as event narratives) and/or the storms associated with the tornadoes (episode narratives). Our work to create a refined database of critical facility impacts was focused on these narratives, spanning the period of 1993 through 2020. Digital versions of narratives for tornadoes occurring before 1993 were not available in the SED.

The event narratives range widely from one sentence to multiple paragraphs in length and highlight significant information about each tornado's path and impacts, and may include notes from field damage surveys. While most tornado impact information is typically included in the event narratives, some episode narratives also mention critical facility impacts. Not every record has both an event and episode narrative, and 1,154 out of the 37,354 total records from 1993 through 2020 do not have any narratives.

Example narratives from the EF-3 tornado on 3/1/2007 in Webster, Sumter, and Macon counties in Georgia are provided below, with emphasis added to highlight impacts on critical facilities.

#### *Example Episode Narrative:*

"A major, negatively tilted and closed upper trough rotated through the mid-south and southeast U.S. on March 1st....The strong upper dynamics present over this region combined with the instability just south of the wedge provided a very favorable environment for long lived, strong tornadoes. A total of 14 tornadoes affecting 17 counties tracked across central and east central Georgia and within the Peachtree City, Georgia county warning area during the late afternoon and evening hours of March 1st. This was the second greatest number of tornadoes recorded to have occurred in the Peachtree City, Georgia forecast area within a 24-hour period, second only to the 16 tornadoes, affecting 18 counties, associated with Hurricane Katrina on August 29, 2005. The March 1st tornadoes spanned an area from Stewart county in the far southwest part of the county warning area to Warren county in the far east central portion of the county warning area. The first tornado touched down in Stewart county at 4:11 pm EST and the last tornado lifted in Marion county at 10:55 pm EST. By far the hardest hit county was Sumter county, and especially the city of Americus, where hundreds of homes and business, including the regional hospital, were heavily damaged or destroyed..."

<sup>&</sup>lt;sup>5</sup> The SED website has a disclaimer stating "some information appearing in Storm Data may be provided by or gathered from sources outside the National Weather Service (NWS), such as the media, law enforcement and/or other government agencies, private companies, individuals, etc. An effort is made to use the best available information but because of time and resource constraints, information from these sources may be unverified by the NWS. Therefore, when using information from Storm Data, customers should be cautious as the NWS does not guarantee the accuracy or validity of the information" (NOAA 2022a).

<sup>&</sup>lt;sup>6</sup> According to the SED website, individual tornadoes may be recorded in multiple segments in some situations, for example when a tornado crosses county or state lines, or when it "lifts of the ground for less than 4 minutes or 2 miles" (2 miles = 3.2 km) (NOAA 2022a).

#### *Example Event Narrative:*

"A damage survey conducted by the National Weather Service in Peachtree City, Georgia concluded that the EF3 tornado that first touched down in southeast Webster county continued to track northeast from the southwest to the northeast corner of Sumter county and then continued into extreme southern Macon county before finally lifting. The tornado tracked a total distance on the ground of approximately 40 miles. The tornado entered Sumter county about 4.75 miles southwest of Plains in southwest Sumter county and exited the northeast part of the county about 9 miles north of Methvins in northeast Sumter county. This was by far the most violent and devastating tornado of the March 1st outbreak. The tornado tracked roughly 32 miles across Sumter county with a maximum path width of 1.0 mile wide, which occurred in the Americus area. Damage within the city of Americus was extensive, although structures, trees, and power lines were down along the entire path of the tornado. The most significant damage in Americus was to the Sumter Regional Hospital, of which a significant portion was destroyed. Hundreds of homes and businesses in the Americus area were either totally destroyed or sustained significant damage. [Hundreds] of vehicles were also damaged or destroyed by debris and/or tossed about like matchsticks. The majority of the damage was within the vicinity of the Sumter Regional Hospital. Two deaths were reported at a home in Americus when a wall collapsed on a 43 year-old male and a 53-year old female. At least eight injuries were documented, but there may have been more. Within the city of Americus, proper, there were 1235 total structures damaged or destroyed, including 217 businesses, 993 residences, 3 cemeteries, 10 churches, 1 fire station, 1 hospital, 8 recreational facilities/parks, and 2 schools. There were 75 structures (42 businesses, 31 residences, 1 hospital, 1 church) were destroyed. There were 148 structures (27 businesses, 116 residences, 3 recreational facilities/parks, 2 churches) with major damage. There were 331 structures (60 businesses, 260 residences, 3 recreational facilities/parks, 5 churches, 1 school, 2 cemeteries) with moderate damage and 681 structures (88 businesses, 586 residences, 2 recreational facilities/parks, 2 churches, 1 school, 1 cemetery, 1 fire station) with minor damage."

Some narratives are briefer, for example the event narrative from the 1997 F1 tornado in Mc Adams, Mississippi (emphasis added):

"This short track tornado moved east-southeast through the town of Mc Adams. Many trees were blown down or uprooted. Some of the trees fell on buildings. Minor damage was down to a few buildings and *many windows were blown out of a school.*"

## 2.2 Narrative Mining Process

To begin to create the database, we reviewed the NWS narratives from 1993 through 2020 to identify records where critical facilities were mentioned. To attempt to find all instances where critical facilities were mentioned in the NWS narratives, we created a list of keywords for each critical facility type that accounted for possible synonyms that could occur within the narratives. For example, to identify schools, we included terms such as "academy" and "learning center."

We filtered both the event and episode narratives using the list of keywords and then validated the filtered list of records through manual inspection. This was a key step because the narratives often contained false positives where keywords were used in contexts unrelated to impacted critical facilities. For example, "Most of the damage occurred on Mayberry *School* Road and Dodson Branch Road," returned a false positive when inspecting for school facility impacts and "The *fire department* reported a narrow debris path crossing state highway 70 south of Sweetwater," returned a false positive for fire station impacts.

We considered a tornado to have hit a facility if the tornado crossed or occurred anywhere on the campus, including outdoor areas like fields or parking lots. This is because the goal of this project is to identify the total number of critical facilities hit by tornadoes, not just the number damaged by tornadoes, to establish a more accurate understanding of the risk from tornadoes to critical facilities. Additionally, a facility could have sustained damage that was not observed during the NWS damage survey due to line-of-sight issues (e.g., minor damage to a flat roof).

Once we confirmed that a narrative described a tornado hitting a critical facility, we recorded whether the narrative reported damage to the facility, and if so, the type and extent of damage (if any such information was provided). We used other sources including NOAA's DAT, Google Maps, the National Center for Education Statistics (NCES) school search tools (NCES 2021a, 2021b)<sup>7</sup>, and traditional and social media reports to fill in additional information about the facilities, including addresses and websites. During this process, we occasionally discovered records in the literature or media of other critical facilities hit by tornadoes. We added these to the database as well, noting the source and flagging them. In situations where a narrative mentioned a tornado hitting multiple facilities, we created a record for each impacted facility. Some narratives mentioned a tornado passing near a critical facility but did not clarify whether it actually struck the facility; these reports were flagged as unknowns for future investigation. If the narrative did not mention details such as the facility name, and we could not find this information online, we recorded it as "unknown."

# 2.3 Application to Schools

For the first stage of database construction, we applied this method to schools (public and private) serving students from preschool through 12th grade, including vocational institutes providing classes for high school students. To create a list of search keywords, we reviewed lists of many real school names. We selected the keywords "academy," "learning center," and "vocational" in addition to "school." We tested other keywords (shown in Table 1) but these either did not capture any results or did not return any results that were not already found with the words "school" or "academy."

Table 1. Keywords we used for the SED narrative mining (left), and keywords that we tested but were unproductive or superfluous (right).

Selected keywords	Other keywords considered
School, academy, learning center, vocational	Institute, prep, department, hall, charter, boarding, magnet, primary, secondary, Montessori, pre-k, kinder- garten, nursery, vo tech, votech, vo-tech, k-12

The school database has a record for each school impacted by a tornado. When a single tornado hit multiple schools, we created a record for each school. Each record retains the original tornado

<sup>&</sup>lt;sup>7</sup> The NCES public schools search tool was based on data for the 2018-2019, 2019-2020, and/or 2020-2021 school year (depending on the date accessed) from the Department of Education's Common Core of Data (NCES 2021a). The NCES private schools search tool was based on "data for the 2017-2018 school year" from the NCES Private School Universe Survey (PSS) (NCES 2021b).

information from NCEI (such as date and location details) and has additional fields for the school impact details. Table 2 shows some of the key fields in the database.

Field name	Description
Campus Damage Reported?	In addition to stating that the tornado hit the campus,
	did the narrative report any damage?
Extent of Damage	Qualitative damage categorization based on the narra-
	tive: unknown, minor, moderate, major, or demolished
What Was Damaged?	List of the components of the school campus and/or
	building that were damaged (e.g. roof, scoreboard)
Name	Name of the school
School Type	Preschool, elementary, elementary-middle, elementary-
	middle-high, middle, middle-high, high, or unknown
Address	Address of the school (usually found on the school's
	website, Google Maps, or the NCES search tools
Website	The school's website, usually found on Google Maps
	or by a web search for the school's name
Notes	Any important notes about uncertainty, assumptions, or
	methods that were applied
Potential Additional Schools?	Names of any schools that might have been hit by the
	tornado, but the narrative does not provide enough in-
	formation to confirm a hit
Additional Sources	Sources (other than Google Maps or the NCES search
	tools) consulted for missing information (if applicable)

Table 2. Sample of key fields in the database.

After making one pass through the narratives to confirm that each one mentioned a school impact, and if affirmative, fill out the school impact details, many records were missing information, such as the school type or name, because these were not in the corresponding narratives. There were also records with ambiguous situations that we needed to address before finalizing the database, such as determining the school type when there was an unusual combination of grades, and determining the number of schools impacted when there were multiple schools-within-a-school, or two schools with separate addresses on a single campus. Often these two challenges were interconnected, for example, if there were a middle school and an elementary school in the same building, but they were listed as two different schools on the school district website, it was not clear whether these should be recorded as a single elementary-middle school, or as separate elementary and middle schools. To complicate matters more, sometimes the narrative mentioned a single school that we discovered was part of a multiple-school complex on Google Maps when looking up the school's name or address.

Before going back through these records, we refined our methods for classifying school types and the number of schools involved in an incident to ensure consistency. We applied these new procedures to all the records that had missing/unclear information, and also to all the records involving multiple schools. We did not apply these new procedures to records that weren't missing information and only involved single schools.

We designed the method for dealing with multiple schools carefully to meet our goal of most accurately documenting the number of schools hit. If the multiple schools in question had unique physical addresses, we kept them as separate records in the database. If they had the same physical address, we considered them to be a single school, documented with a single record in the database, with two exceptions:

If the narrative referred to them as separate schools, e.g. "the tornado hit the town middle school and then it hit the high school," we deferred to the narrative and kept them as separate records.

If the school(s) appeared to be part of a large complex of separate buildings on Google Maps satellite imagery, we considered them to be multiple schools and documented them with separate records, even if they had the same address. However, in this case, all schools on the campus that were not mentioned as being struck by the tornado in the narrative were marked as unknowns, and were moved to a separate spreadsheet for potential future analysis and not included in the total count of impacted schools.

If the narrative stated that multiple schools were hit by the tornado, but there was not enough information to determine the number, we left the incident as a single record in the database and noted that the number of schools hit by the tornado was uncertain. During the final totaling of schools in the database, we made the most conservative estimate of the number of schools impacted in each of these instances. For example, if the number was two to four schools, we counted them as two, and if the number was some unknown non-zero quantity, we counted it as one.

Below is an example narrative (with emphasis added) from the April 3rd, 2012 tornado in Sowers, TX that was found through mining the NWS narratives for school keywords:

"An EF-0 tornado with maximum estimated around 85 mph occurred in west Irving. The twister damaged a metal awning at a house and uprooted trees along Avenida Loop by the Irving Mall. Continuing to the northeast, the twister damaged part of the roof of the Goodwill store on N Belt Line Rd and blew HVAC units off the rooftop of the store. Wood power poles were leaning near W Rochelle Rd and N Belt Line Rd. Near the end of the track, trees were uprooted *and minor damage was noted at Townsell Elementary School. Two air conditioning units at the school were damaged and so was a gas line.* The path length was non-continuous and was approximately 1.3 miles long with a narrow width of 30 yards."

The narrative clearly describes a tornado hitting part of the school's campus, one of the buildings, so we created one school record and noted that there was damage reported. We recorded the name (Townsell) and type of the school (elementary), noted what was damaged (AC units and a gas line) and that the extent of the damage was "minor." Finally, we looked up the address and website online and added those to the entry. There is nothing unusual to note for this entry, and no potential additional schools, so we left those fields blank.

The following is an example of a less straightforward narrative from the EF-2 tornado on May 30th, 2013 near Broken Arrow, OK (emphasis added):

"...The tornado continued to move eastward, crossing the Muskogee Turnpike where it snapped numerous large trees. *Two school buildings lost a number of windows as the tornado approached S 257th E Avenue*. A home was damaged and a large outbuilding was destroyed in this area. Trees were snapped along and near Joy Avenue before the tornado dissipated. Maximum estimated wind in this segment of this tornado based on this damage was 125 to 135 mph."

It is clear from the narrative that at least one school was hit by the tornado, however, it is not clear whether two buildings at one school were damaged, or two separate schools were damaged. Because the school name and type were not given, we used the date and location information in combination with the DAT and Google Maps to identify that a school called Oneta Ridge Middle School was hit by the tornado. We filled out the details for Oneta Ridge Middle School, put a yes in the "Campus Damage Reported?" field, noted the extent of damage qualitatively, and that the damage was to windows. We noted that the number of schools hit by the tornado was uncertain because of the ambiguity in the narrative, and counted one school for this event in the final tally.

#### 3 METHOD 2- GIS ANALYSIS

The method discussed so far has a significant weakness that will lead to underreporting of critical facilities struck by tornadoes; inclusion of critical facility impacts in the narratives was voluntary and up to the discretion of each author, not a uniform requirement.

To overcome the limitations of the SED mining analysis, we are in the process of developing a complementary GIS method for finding critical facilities hit by tornadoes. This method may help clarify uncertain records flagged in the SED mining analysis, such as the example with the ambiguous number of schools in the last section. The method consists of overlaying geospatial records of tornado damage paths documented by NWS field surveyors that are stored in the DAT, with various geospatial critical facility datasets, such as those stored in the Department of Homeland Security's Homeland Infrastructure Foundation-Level Data (HIFLD) geoplatform<sup>8</sup>. We plan to use this method to collect another list of critical facilities hit by tornadoes that we will compare to the list from the Storm Events Database mining from the same period to 1) identify additional schools that were potentially struck by tornadoes, and 2) better understand the proportion of school strikes that are not captured in the narratives.

The DAT includes three types of geospatial damage records for tornadoes: damage points, tracks, and polygons. Damage points record observed damage at specific locations (such as a fallen tree or damage to a specific building) and include metadata fields such as the applicable damage indicator (WISE 2006), date and time the point was logged, a wind speed estimate, and text information about the type and severity of the damage. Damage tracks record linear centerline path estimates of tornadoes along with metadata such as length and width estimates and EF-Scale rating. Damage polygons record estimates of the areas damaged by tornadoes in the form of geospatial polygons (Fig. 1), and include metadata such as EF-ratings and event comments which describe impacts. Tornadoes in the DAT can have multiple damage polygons corresponding to different levels of damage on the EF-Scale (Fig. 1).

In the initial application of the GIS method, we plan to only use damage polygons to account for the spatial extent of the tornado, and will overlay these polygons (NWS 2020) with the critical facility datasets. We will associate the damage polygon metadata (e.g. EF-ratings, event comments) with critical facilities that spatially intersect them. Critical facility location datasets are typically archived as point locations although the facilities that they represent are often large areas. Therefore, we are currently testing various approaches to better represent the critical facilities, including using either buffered points or building footprints (Fig. 2).

We are focusing the GIS analysis on the years 2010 through 2020 because the creation of damage polygons does not seem to have been widespread until around 2010. Even so, there are only 5,196 damage polygons for this period, representing approximately 40% of the 13,003 tornadoes recorded in the SED for the same time (NOAA 2022b)<sup>9</sup>. It's also important to note that according to the metadata for the DAT features, "while the data has been quality controlled, it is still considered preliminary" (NOAA 2022c).

<sup>&</sup>lt;sup>8</sup> https://hifld-geoplatform.opendata.arcgis.com/

<sup>&</sup>lt;sup>9</sup> The count of tornadoes for 2020 is still listed as preliminary.



Figure 1. Examples of tornado damage polygons in the Damage Assessment Toolkit (DAT): a single polygon from one tornado (top), and multiple polygons representing different levels of damage for another individual tornado (bottom).



Figure 2. Examples of the GIS intersection methods we are testing: intersecting damage polygons with buffered critical facility points (left), and with building footprints (right). A nonzero intersection area would mean that the facility was hit by the tornado based on our assumptions.

The GIS method relies on the assumption that if some or all of a critical facility's campus is inside a damage polygon, then it was hit by the associated tornado. One challenge with this is that critical facility inventories change over time as new facilities are constructed and old ones are torn down or reused for other functions, so it is important to use critical facility datasets that are relevant to the varying dates of the tornadoes being studied. Another challenge is that damage polygons are best estimates of tornado damage swaths, but they might include areas that were not impacted by the tornado or not include all areas impacted by the tornado.

# 3.1 Application to Schools

In the next section, we present two examples of the GIS method applied to schools. We used tornadoes that were recorded in the SED for these examples to demonstrate the insights that this method may add to the results from the SED mining. The first example is the 2013 EF-2 tornado near Broken Arrow, OK. The event narrative for this tornado was included in the Method 1 section of this paper. There was uncertainty about how many schools were hit by the tornado. We show how the GIS method can clarify the number of schools that were struck, by applying it to this incident using the corresponding damage polygon from the DAT.

The second example is a 2020 tornado in Tennessee with narratives that only mention two school impacts. For this example, we overlaid the DAT damage polygons for that tornado with a school dataset in ArcMap 10.8 (ESRI 2019), to show how this method can be used to identify additional impacted schools that were not mentioned in the narratives. We used the Private Schools dataset from the HIFLD database (DHS 2020).

## 4 RESULTS

## 4.1 Method 1: Storm Events Database Mining Results- Schools

This method identified 669 schools that were hit by tornadoes from 1993 through 2020, for an average of nearly 24 schools per year. The database results also show that 546 tornadoes (or tornado segments) from the SED hit schools, and 56 of those hit multiple schools (not including incidents with an uncertain number of schools hit). The count of 669 schools comprises 648 schools that were found through mining the narratives, plus an additional 21 we encountered in reports/media while gathering more data to complete records where the narratives were missing school names or other information. High schools were most frequently struck, followed by elementary schools. A map of the frequency of school strikes by state is shown in Figure 3. Texas, Tennessee, and Alabama had the most schools hit by tornadoes. The next ten highest frequency states are also located in the Midwest and Southeast. As the primary data source was not specifically designed to capture all schools struck by tornadoes, it should be noted that this map represents a lower bound and there are certainly more.

Some of the most commonly reported damage was to roofs, windows, awnings, trees, fences around athletic fields, and bleachers. Rating the level of damage to the schools proved to be difficult because the narratives often only included a brief (sometimes limited to a single adjective) and highly subjective description of the degree of damage.



Figure 3. Number of schools hit by tornadoes in each state during the period from 1993-2020, based on mining of NWS narratives from the Storm Events Database. Note that Texas, Tennessee, and Alabama had the most schools hit by tornadoes.

## 4.2 Method 2: GIS Analysis Results- Schools

## 4.2.1 Example 1: Broken Arrow, OK EF-2 Tornado

The narrative for the 2013 EF-2 tornado near Broken Arrow, OK includes "Two school buildings lost a number of windows," but is not clear whether these were two buildings on one campus or two different schools. We used the location description in the narrative to find the approximate location of the school(s) and found a damage point for the tornado in the DAT for a building in that general location (Fig. 4), which was identified as Oneta Ridge Middle School according to Google Maps (Fig. 5). While we were able to identify this one school, we were not able to determine if there were other schools impacted using Method 1. We recorded the total number of schools hit by this tornado as uncertain.

The spatial comparison of tornado damage polygons to school locations used in Method 2 shows that the nearby Highland Park Elementary school which was marked on Google Maps (Fig. 5), is also inside a damage polygon for the tornado (Fig. 6), even though it was not marked with a damage point on the DAT, or explicitly mentioned in the narrative.

While we noticed Highland Park Elementary on recent aerial imagery, we were able to verify using Google Earth historical imagery (Google Earth 2012) that the school was indeed in the same location at the time of the tornado. This example demonstrates the utility of the GIS method for clarifying and validating results from the SED mining.



Figure 4. DAT damage point and details box used to identify the school recorded for the 2013 EF-2 tornado near Broken Arrow, OK.

# 4.2.2 Example Two: TN Multi-County EF-3 Tornado

The NWS event narrative for the long-track EF-3 tornado that hit Davidson, Wilson, and Smith counties, TN on 3/3/2020 mentions two schools hit by the tornado: Stoner Creek Elementary and West Wilson Middle (shown in Google Maps aerial imagery in Figure 7). However, by overlaying the DAT damage polygons for the tornado and the HIFLD 2020 Private Schools dataset, it appears

that Mt. Juliet Christian Academy was also hit by the tornado (Fig. 8), despite not being mentioned in either the event or episode narratives. According to an article in The Wilson Post (2020), "Mt. Juliet Christian Academy was BATTERED by Tuesday's tornado." Figure 9 shows NOAA (2020) post-event aerial imagery that confirms significant damage to the school.



Figure 5. Google Maps imagery and labels used to identify the school recorded for the 2013 EF-2 tornado near Broken Arrow, OK during Method 1 implementation (Oneta Ridge Middle School, left), and used to identify a potential additional school (Highland Park Elementary, right) hit by the tornado using Method 2. Imagery credits: Google, ©2022 Maxar Technologies, USDA Farm Service Agency; Map data ©2022 Google.

![](_page_12_Figure_3.jpeg)

Figure 6. A portion of the damage polygon on the DAT for the 2013 Broken Arrow tornado that clearly includes both schools shown in Figure 5.

![](_page_13_Picture_0.jpeg)

Figure 7. Screenshot from Google Maps showing West Wilson Middle School, Stoner Creek Elementary School, and Mt. Juliet Christian Academy. Imagery credits: Google, ©2022 Maxar Technologies, USDA Farm Service Agency; Map data ©2022 Google.

![](_page_13_Figure_2.jpeg)

Figure 8. The NWS event narrative for the EF-3 tornado that hit Davidson, Wilson, and Smith counties, TN on 3/3/2020 only mentions two schools hit by the tornado: Stoner Creek Elementary and West Wilson Middle (labeled). However, by overlaying the HIFLD Private Schools dataset with the DAT polygons for this tornado (there are several polygons representing different levels of damage) we can see that Mt. Juliet Christian Academy was also hit. Imagery credits: ESRI, Maxar, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community. *Note:* 1 *ft* = 0.3048 *m*.

![](_page_14_Picture_0.jpeg)

Figure 9. Post-event aerial imagery of the damage to Mt. Juliet Christian Academy (NOAA 2020). Notice that the roof is completely removed from part of the building in the center of the image.

# 5 DISCUSSION

# 5.1 Storm Events Database Mining for School Impacts- Results

The SED mining cataloged 669 schools that were hit by tornadoes from 1993 through 2020, for an average of nearly 24 schools per year. These results provide evidence that tornadoes hit US schools with a higher frequency than is perhaps commonly assumed. The spatial distribution of schools hit by tornadoes is somewhat as expected since the Midwest and the South experience the majority of US tornadoes. However, Figure 3 shows that the distribution skews more towards the South and is not centered in the area of the highest tornado wind speeds (generally Kansas, Missouri, Arkansas, and Oklahoma) according to the tornado wind speed maps in Chp. 32 of the ASCE 7-22 Standard (ASCE 2021). This may be due to various factors including differences in population density (and resulting school density), school campus size, and damage vulnerability of the school buildings.

## 5.2 Storm Events Database Mining for School Impacts- Uncertainty

There are several sources of uncertainty in the results from the SED mining, the majority of which bias the results towards undercounting the true number of schools hit by tornadoes in the US during this period. One of the most significant sources of error is that the SED was not designed to be used for recording school tornado impacts. Whether a school ended up being described in an NWS narrative was up to the author. Because the narratives are brief statements meant to capture the most significant impacts of each tornado or overall weather event, there could have been numerous situations where schools that were hit were not mentioned because the damage was not notable in comparison to that of other structures, especially if the school was not occupied during the event. Some narratives were unclear about whether a school was actually hit by the tornado, and we set

these records aside for future investigation. However, it is likely that in some of these situations the school was indeed hit. Additionally, if field surveyors could not see damage due to line-of-sight issues, they might not have recorded the school as being hit by the tornado.

Furthermore, the SED is not an exhaustive list of every tornado that has occurred in the US since 1950, as it only includes tornadoes that were reported. Through Bayesian modeling of "expected tornado counts" for 1975-2016, Potvin et al. (2019) estimated that the NCEI SED significantly underreports the tornadoes that occurred during this period. Their "results suggest only ~45% of tornadoes that occurred within the central US analysis domain during 1975–2016 were reported." Hence, there may be schools that were hit by tornadoes that were never recorded in the SED

Even for tornadoes that were recorded in the SED, for some entries, it is impossible to tell if schools were impacted because they are missing narratives. For example, out of the 37,354 tornado events that were recorded in the SED from 1993 through 2020, 1,154 had neither episode nor event narratives. Furthermore, 4,519 were missing event narratives, and the event narratives are where school impacts are most frequently mentioned. All these factors combine to make our total for the number of schools hit by tornadoes from 1993 to 2020 a potentially significant underestimate.

A limitation of this type of post-facto database construction is that school details such as address and type (e.g., elementary, high) are often based on current online information, which may be different than at the time of the tornado. This limitation may affect the total count of schools hit by tornadoes because we used recent online sources when applying the multiple schools procedure (which required looking up the school addresses and viewing maps or imagery of the school locations). It also adds uncertainty to the numbers of each school type that were hit. It is also important to note that some of the names of the schools in the database may not be correct, because school names change over time, and these were generally based on current information unless the name was specified in the narrative or media reports.

A methodological factor that could potentially skew our count of schools from the SED mining towards a higher number is our procedure for dealing with multiple schools hit by the same tornado. We only applied this procedure to records that had missing/unclear information, and also to all the records involving multiple schools, together comprising roughly half of all the records. A limitation of the procedure is that we determined the number of schools hit by the tornado based on current information, usually on Google Maps and the school websites. This information may not have been accurate at the time of the tornado. For example, if we found a middle school next to a high school on Google Maps with separate addresses, we marked them as two separate schools hit by the tornado, but one of them could have been built after the tornado. The converse situation is also possible, where there were multiple schools at the time of the tornado but one was torn down or the addresses were consolidated, which would lead to an undercount of the number of schools hit. It is also possible that a tornado could pass over just one of a few adjacent schools. Another methodological source of uncertainty is our choice of keywords and the potential for spelling errors to cause our filter to miss relevant records. Additionally, it's possible that in a few instances we misread and then misclassified a few uncertain or false positive records as confirmed hits. It is likely that the methodological sources of uncertainty have a smaller contribution to the total uncertainty than the effects of underreporting of schools in the narratives.

# 5.3 GIS Analysis of School Impacts- Preliminary Findings and Methodological Challenges

Our two examples for the GIS method demonstrate positive results for clarification, validation, and expansion of results from the SED mining approach, although new challenges are presented. A considerable limitation to the GIS analysis approach for identifying critical facilities affected by

tornadoes is that there are no damage polygons available for approximately 60% of all tornadoes in the SED from 2010 through 2020.

The GIS analysis requires the assumption that the damage polygons accurately portray the bounds of tornado damage, but this may not be the case because the spatial determination of damage can be very difficult in situations without thorough ground surveys or remote sensing data, and where damage indicators are sparse. This means that it would not be completely certain that a school inside a damage polygon was hit by the tornado.

A significant challenge with the GIS method is that many schools were likely built, demolished, moved, and renamed during the study period. It is difficult to assemble a combination of school datasets to intersect with the damage polygons from widely varying dates that will not result in errors like concluding that a school was hit by a tornado before the school existed. One approach we are exploring is using the union of school datasets from the beginning and end of the study period to create a single dataset of schools that did not move or close during the study period. However, this will result in undercounting because schools that were hit by tornadoes but did not exist throughout the whole study period will not be discovered. This is particularly an issue for finding severely damaged schools may then be rebuilt in different locations, and even if these schools are rebuilt in the same locations, the new schools may have distinct records in the school datasets. However, severely damaged schools are also more likely to have been reported in the NWS narratives and accounted for in Method 1.

## 5.4 Ongoing and Future Analyses

We plan to apply the SED mining method and the GIS method to several other critical facility types. We are continuing to assess and improve the accuracy of the database through: 1) tracking media reports of tornado impacts as they occur and comparing these to the impacts recorded in the SED during the same period; and 2) comparing the lists of critical facilities found using the SED mining to those found using the GIS method. This second tactic will allow us to assess the proportion of the true population of schools hit by tornadoes captured by the mining of narratives. To continue to develop the GIS method, we are testing the use of critical facility building footprints vs. buffered critical facility points for intersection with the damage polygons.

To overcome the limitation caused by using recent information (e.g. Google Maps imagery) to determine which and how many schools were impacted by tornadoes in uncertain situations, future analyses could involve checking historical imagery for all the school records to confirm the number and locations of the schools at the time of each tornado. We did not do this for all the schools in the initial stage of database construction because it would have been too time-intensive to perform in addition to the other online research required for the hundreds of schools. Another potential improvement on the SED mining method would be to use natural language processing to increase the efficiency of the process.

# 6 CONCLUSION

Tornadoes affect much smaller areas than earthquakes and hurricanes but the cumulative impacts of the over 1,250 US tornadoes per year equal or exceed the average annual impacts of earthquakes and hurricanes. With the emerging adoption of engineering design for tornadoes, including the new tornado load provisions in ASCE 7-22 that specifically apply to most critical facilities, thor-

ough documentation and quantification of past tornado impacts on critical facilities will be valuable for local and regional risk assessment in support of decision-making regarding the adoption of tornado-resistant codes and standards and installation of public tornado shelters.

We are creating a database of tornado strikes on critical facilities through mining NWS narratives from the NCEI SED. For the initial stage of database construction, we applied these methods to preschool-12th grade schools. We considered a tornado to have hit a school if the tornado crossed or occurred anywhere on the campus. We found that at least 669 schools in the US were hit by tornadoes from 1993 through 2020, for an average of nearly 24 schools per year. These results demonstrate that tornadoes have a very significant cumulative impact on schools. Texas, Tennessee, and Alabama had the most schools hit by tornadoes. Overall, school-tornado impacts were concentrated in the South and Midwest.

A limitation of this type of post-facto database construction is that school details such as address and type, and the number of schools in a given location (sometimes necessary to determine the number of schools hit) often had to be based on current online information. Underreporting of tornadoes, and the fact that mentioning school impacts in the narratives is not required both bias the results towards undercounting the true number of schools hit by tornadoes in the US during this period.

To verify the results from the SED mining, and to find more of the critical facilities that have been hit by tornadoes, we are developing a GIS analysis method involving the intersection of tornado damage polygons from the DAT with various school location datasets. There are two main challenges with this method: dealing with the varying accuracy of the damage polygons, and assembling a set of school location datasets from different years that will most accurately match the numerous dates of the damage polygons.

We plan to apply both the SED mining method and the finalized GIS analysis to other critical facility types in addition to schools. We also plan to publish the database so that it may be used to improve risk assessment and cultivate more accurate public and decision-maker understandings of tornado impacts on these important community facilities.

## 7 ACKNOWLEDGEMENTS

We would like to thank Stuart Hinson from NCEI for providing us with the Storm Events Database tornado data that were essential to this project. We gratefully acknowledge our peer reviewers, Camila Young from NIST, James LaDue from the National Weather Service, and Gordon Strassberg (Storm Data Program Manager for the National Weather Service), for their contributions.

## 8 REFERENCES

- ASCE (American Society of Civil Engineers), ASCE/SEI 7-22 Standard: Minimum Design Loads and Associated Criteria for Buildings and Other Structures, 2021. <u>https://doi.org/10.1061/9780784415788.</u>
- DHS (US Department of Homeland Security), Homeland Infrastructure Foundation-Level Data (HIFLD), Private Schools, updated November 12, 2020, Washington, DC, US Department of Homeland Security, October 30, 2020. Online. <u>https://hifld-geoplatform.opendata.arcgis.com/datasets/private-schools/explore</u>.

ESRI (Environmental Systems Research Institute), ArcGIS Release 10.8, Redlands, CA, 2019.

FEMA (Federal Emergency Management Agency), P-908: Mitigation Assessment Team Report – Spring 2011 Tornadoes: April 25-28 and May 22, (2012). <u>https://www.fema.gov/sites/default/files/documents/fema\_p\_908\_tornado\_mat\_combined\_0.pdf</u>.

FEMA, Safe Rooms for Tornadoes and Hurricanes: Guidance for Community and Residential Safe Rooms, P-361, 4th ed., Washington, DC, 2021. <u>https://www.fema.gov/emergency-managers/risk-management/safe-rooms</u>.

- Google, Google Earth Pro Version 7.3.4.8248. (August 9, 2012). Broken Arrow, OK. 36°02'13.82" N, 95°42'08.11" W, Eye altitude 3165 ft. Landsat/Copernicus. Accessed 2/3/2022.
- Institute of Education Sciences, National Center for Education Statistics (NCES), Search for Public Schools. <u>https://nces.ed.gov/ccd/schoolsearch/</u>. Accessed summer 2021(a).
- Institute of Education Sciences, NCES, Search for Private Schools. <u>https://nces.ed.gov/surveys/pss/privateschool-search/</u>. Accessed summer 2021(b).
- Insurance Information Institute, Spotlight on: Catastrophes: Insurance Issues, 13 December 2021. https://www.iii.org/article/spotlight-on-catastrophes-insurance-issues.
- WISE (Wind Science and Engineering Center), Texas Tech University (TTU), A Recommendation for an Enhanced Fujita Scale (EF-Scale), Submitted to the United States National Weather Service and Other Interested Users, Revision 2, October 10, 2006, WISE, TTU, Lubbock, TX, 2006. <u>http://www.depts.ttu.edu/nwi/Pubs/FScale/EF-Scale.pdf</u>.
- ICC (International Code Council), ICC/NSSA Standard for the Design and Construction of Storm Shelters, ICC 500-2020, ICC and National Storm Shelter Association, 2020. <u>https://codes.iccsafe.org/content/ICC5002020P1</u>.
- NOAA (National Oceanic and Atmospheric Administration), National Geodetic Survey (NGS), 2020 NOAA NGS Emergency Response Imagery: Nashville, TN Tornado, Silver Spring, MD, NOAA, 2020, Online. https://storms.ngs.noaa.gov/storms/nashville/index.html#9/36.1800/-86.2000. Accessed 8 Feb. 2022.
- NOAA, National Centers for Environmental Information (NCEI), Storm Events Database, (Years 1993 through 2020), Washington, DC, NOAA, 2021. <u>https://www.ncdc.noaa.gov/stormevents/details.jsp</u>.
- NOAA, NCEI, Storm Events Database FAQ. <u>https://www.ncdc.noaa.gov/stormevents/faq.jsp</u>. Accessed 3 Feb. 2022(a).
- NOAA, NCEI, US Tornadoes. https://www.ncdc.noaa.gov/societal-impacts/tornadoes/. Accessed 1 Feb. 2022(b).
- NOAA, National Weather Service (NWS), ESRI ArcGIS Online, Damage Viewer Overview. <u>https://noaa.maps.arcgis.com/home/item.html?id=b6b9dd55450a4f64bcafb3cd6e005a5c</u>. Accessed 3 Feb. 2022(c).
- NOAA, NWS, 80-Year List of Severe Weather Fatalities. <u>https://www.weather.gov/media/hazstat/80years\_2020.pdf</u>. Accessed 17 Feb. 2022(d).
- NWS, Damage Assessment Toolkit (DAT) Damage Viewer, Damage Polygons SDE feature layer, updated 5/5/2021, Silver Spring, MD, NWS, & Redlands, CA, ESRI ArcGIS Online, 29 June 2020, Online. <u>https://services.dat.noaa.gov/arcgis/rest/services/nws\_damageassessmenttoolkit/DamageViewer/FeatureServer</u>. Accessed Feb. 2022.
- NWS, Damage Assessment Toolkit (DAT). <u>https://apps.dat.noaa.gov/stormdamage/damageviewer/</u>. Accessed 2021 and 2022.
- NIST (National Institute of Standards and Technology), Preliminary Reconnaissance of the May 20, 2013, Newcastle-Moore Tornado in Oklahoma, NIST Special Publication 1164, Gaithersburg, MD, 2013. <u>https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1164.pdf</u>.
- NIST, Final Report: NIST Technical Investigation of the May 22, 2011, Tornado in Joplin Missouri, NCSTAR 3, Gaithersburg, MD, 2014. <u>https://doi.org/10.6028/NIST.NCSTAR.3</u>.
- C.K. Potvin, C. Broyles, P.S. Skinner, H.E. Brooks, and E. Rasmussen, A Bayesian Hierarchical Modeling Framework for Correcting Reporting Bias in the US Tornado Database, Weather and Forecasting, vol. 34, no. 1, Feb. (2019) pp. 15–30. <u>https://doi.org/10.1175/WAF-D-18-0137.1</u>.
- C. Ramseyer, L. Holliday, and S.T. Sherry, Lessons Learned from Two Elementary School Collapses during the May 20, 2013 Moore Tornado, Journal of Performance of Constructed Facilities 33.1 (2019). https://doi.org/10.1061/(ASCE)CF.1943-5509.0001228.
- The Wilson Post Staff Reports, MJCA Suffers Serious Tornado Damage, The Wilson Post, 4 Mar. 2020. https://www.wilsonpost.com/news/mjca-suffers-serious-tornado-damage/article\_b2b7d934-5e4c-11ea-be55cff197b9b004.html. Accessed 7 Feb. 2022.
- USGS (US Geological Survey), Deaths from US Earthquakes: 1811-2014, 23 Mar. 2016. <u>https://web.ar-chive.org/web/20160323192352/https://earthquake.usgs.gov/earthquakes/states/us\_deaths.php</u>. Accessed 17 Feb. 2022.