Digitizing and Geocoding Historical Records to Improve an Understanding of Tornado Climatology

Tyler Fricker
University of Louisiana Monroe, USA

Taylor Wells
Texas A&M University, USA

John Thomas Allemeier and Amelia Baxter
Florida State University, USA

Tornadoes have the potential to cause catastrophic destruction and mass casualties. What influence a changing climate has on tornado behavior remains a largely unanswered question. Much of the current knowledge of tornado climatology is linked to historical tornado reports that date back to 1950 and are maintained by the National Weather Service and the Storm Prediction Center. Here, we digitize and geocode significant and killer tornado reports over the period from 1880 to 1989 as a means of improving and expanding the data readily available to researchers. In total, 12,191 tornado reports were included to create a spatial data set that includes tornado characteristics consistent with the severe report database. Existing as both a straight-line track and an initial point model, the data set is broad enough to provide statistical analysis on tornado characteristics, geospatial analysis at the grid and county level, and point-based analysis at fine spatial scales. This flexibility in the data set allows for the application of methodologies across a range of disciplines that can improve studies related to tornado climatology and curriculum rooted in quantitative and geospatial analysis. **Key Words: climatology, physical geography, tornadoes.**

Tornadoes are violently rotating columns of wind in contact with the ground that have the potential to cause catastrophic destruction and mass casualties. Like many natural hazards, one of the enduring questions that interests researchers is the potential impact a changing climate has on tornado behavior. This problem is rooted in a deep understanding of the environmental variables attributed to the formation of rotating thunderstorms and the analysis of tornado characteristics (i.e., tornado frequency, intensity, and damage variables). Current knowledge of the study of tornado climatology often comes from one of two sources: (1) atmospheric observations and predictions—including global climate models and reanalysis data—and (2) historical tornado reports.

Environmental data on the atmosphere provides the physical basis for an understanding of tornado—or more appropriately convective storm—potential and occurrence. Of specific interest are measurements or estimates of convective available potential energy (CAPE), vertical wind shear or storm relative helicity (SRH), and convective inhibition, but other factors such as specific humidity and temperature lapse rate are also of some importance, as are proxy measures like significant tornado parameter (STP). Several studies have indicated that tornado-favorable environments have not clearly changed over the short term (i.e., late twentieth and early twenty-first centuries; Gensini and Ashley 2011; Allen, Tippett, and Sobel 2015; Tippett et al. 2015; Trapp and Hoogewind 2016). Other studies have suggested that future favorable tornado environments (i.e., late twenty-first century) could increase in frequency or duration (Trapp et al. 2007; Diffenbaugh, Scherer, and Trapp 2013; Seeley and Romps 2015; Hoogewind, Baldwin, and Trapp 2017). More recently, Moore, Clair, and McGuire (2022) evaluated a spatial climatology of atmospheric conditions favorable to tornadoes and found increases in CAPE and SRH in the Mid-South and Midwest United States over the period from 1980 to 2018. Similarly, Taszarek et al. (2021) and Koch et al. (2021) found increasing trends—especially across the northern Great Plains—in products of instability and wind shear (e.g., STP) over the period from 1979 to 2019.

Tornado reports provide the empirical basis for an understanding of tornado frequency and intensity. The United States is home to the most complete tornado data set in the world—dating back to 1950—which has led to a base-level knowledge of tornado behavior among scientists and the public alike. To this point, a broad understanding of locational tornado climatology is well defined. Most tornadoes in the United States occur across the Great

The Professional Geographer, 0(0) 2023, pages 1–13 © 2023 by American Association of Geographers.
Initial submission, July 2022; revised submission, September 2022; final acceptance, September 2022.
Published by Taylor & Francis Group, LLC.

Plains, Midwest, Mid-South, and Southeast. Seasonally, tornadoes occur in the greatest numbers across the Mid-South and Southeast during the winter and early spring before shifting northward—and westward—during the summer. This can be physically explained by the migrating jet stream (Brooks and Doswell 2001), changes in convective storm environments (Weaver, Baxter, and Kumar 2012; Bluestein 2013), and proximity to the Gulf of Mexico (Molina, Timmer, and Allen 2016). Together, this combination of occurrence and seasonality led to the quantification of the colloquial understood Tornado Alley in Brooks, Doswell, and Kay (2003), where daily probabilities of tornado climatology are evaluated. More recently, Elsner, Jagger, and Fricker (2016) provided a county-level tornado climatology through a statistical model accounting for changes in observational practices and population density over the period from 1970 to 2015, and Krocak and Brooks (2018) provided hourly tornado climatology estimates for fine-scale weather prediction using data from 1954 to 2015.

Less defined is an understanding of an intensity-based tornado climatology. This is due, in part, to differentiation in the definition of tornado intensity. When defining intensity by the Fujita (F) or Enhanced Fujita (EF) scale, most researchers look toward significant tornadoes—those rated F/EF2 or greater—as the underlying data (Concannon, Brooks, and Doswell 2000; Ashley 2007; Coleman and Dixon 2014). For example, Concannon, Brooks, and Doswell (2000) found that significant tornadoes over the period from 1921 to 1995 occurred most often in an L-shaped region from Iowa to Oklahoma to Mississippi, whereas Ashley (2007) found that significant tornadoes over the period from 1950 to 2004 occurred in a more linear pattern from Oklahoma through the Mid-South and Southeast. Coleman and Dixon (2014) incorporated path length into their definition of significant tornado climatology—or risk—and found a clear maximum from south-central Mississippi through northern Alabama over the period from 1973 to 2011. When defining intensity by some other metric, whether through a physical variable (e.g., energy [J] or power [W]) or not, similar spatial patterns exist. For instance, using the destruction potential index (Thompson and Vescio 1998), Coleman and Dixon (2014) found the most intense tornadoes occur in the Mid-South and Southeast. More recently, using total kinetic energy, Fricker and Elsner (2015) found the states of the southern Great Plains through the Mid-South and Southeast were affected by the most intense tornadoes over the period from 2007 to 2013.

Moving beyond the locational climatology of convective storm environments or tornado reports, it is possible to try and uncover any connection that might exist between changing tornado behavior and a changing climate. For example, recent work has found that tornadoes are occurring on relatively fewer days of the year (Elsner et al. 2014), whereas the number of days of the year with many tornadoes is increasing (Brooks, Carbin, and Marsh 2014; Elsner, Elsner, and Jagger 2015; Moore 2017; Moore and DeBoer 2019; Moore and Fricker 2020). In addition, the interannual variability—the volatility—of tornadoes has increased since the early twenty-first century (Tippett et al. 2015), as has the average annual power (Elsner, Fricker, and Schroder 2019). Spatially, the location of tornadoes in the United States appears to be shifting over time, with relatively more tornadoes occurring in the Mid-South and Southeast today than in years past (Agee et al. 2016; Gensini and Brooks 2018; Moore 2018). That said, although tornado behavior does show signs of changing, any definitive connection between these observed changes and a changing climate remains tenuous, as researchers still tease out the influence of technological advancements (i.e., better radar coverage and reporting, transition from F to EF scale, etc.; Edwards, Brooks, and Cohn 2021) and a population bias (Anderson et al. 2007; Elsner et al. 2013).

Perhaps the most effective way to continue to work toward a teasing out of the relationship between observed changes in tornado behavior and a changing climate is to produce more data—either through continued primary data collection or the chronicling of past tornadoes. With this in mind, we present the digitization of a historical tornado data set that includes all significant and killer—those producing at least one fatality—tornadoes over the period from 1880 to 1989. This data set exists in spatial terms as both a point (i.e., initiation point) and a line (i.e., straight-line track between the initiation and end points), allowing for the application of both geographic and statistical methodologies. In addition, we provide examples of the type of climatological analysis that can be performed on the data set as a means of corroborating, expanding, or strengthening our current understanding of tornado climatology. Our hope is that through the expansion of the current publicly available historical tornado data set, more creative ways to investigate tornado behavior and impacts in the United States can be developed in both educational- and research-based settings.

Data and Methods

Significant and Killer Tornado Data

The underlying attribute data digitized for this work come from a collection of historical tornado reports produced by Finley (1882) and Grazulis (1990a, 1990b). The documented tornado reports span several time periods, including an Early Period (1880–1915), a Middle Period (1916–1949), and a Modern

Period (1950–1989; Grazulis 1990a, 1990b). Through the detailed evaluation of annual tornado reports and local newspaper columns, Grazulis (1990a, 1990b) produced a significant tornado data set covering 1880 through 1989 that includes information on the date and local time, the number of fatalities and injuries, the length and width of the damage path, the estimated F-rating, and the state(s) and counties affected. Importantly, the data set includes only those tornadoes that can be described as significant—F2 or greater—or killer—producing at least one fatality.

Digitizing the Data

Using ABBYY FineReader PDF software, state-level tables found in Grazulis (1990a, 1990b) were digitized and converted into numeric and text-based documents—in a comma-separated values (CSV) file type. The resulting tables were then compared to the original data and cleansed of any numeric or text discrepancies or duplicates—for example, if a tornado affected two states and was included in each state-level table, only the record for the state of the initiation point was kept. In total, 12,191 tornado reports were digitized into state-level data sets that were combined to create a national data set that covers the contiguous United States.

Once the attribute tables were digitized into a usable format (i.e., CSV), each tornado track was geocoded according to the associated newspaper event narrative. For example, an F4 tornado affected Crawford County, Kansas, on 2 April 1880, at 6:30 pm local time. The event narrative for the tornado described a tornado moving northeast that began seven miles southwest of Girard and ended near the Missouri border. Thus, the tornado was geocoded from seven miles southwest of Girard, Kansas (37.43, –94.92 [latitude, longitude]) to near the Missouri border (37.65, –94.65). If no detailed spatial information was given in the associated event narrative, the tornado was geocoded as a point at the affected county seat for a single-county tornado or as a track from the initial county seat to the final county seat for a multiple-county tornado. This ensures that the affected counties are properly tabulated using any spatial methodologies.

The result of these geocoded reports is the creation of a significant and killer tornado data set over the period from 1880 to 1989 that is consistent, in structure, with the current Storm Prediction Center's (SPC) Severe Weather GIS data set (SVRGIS; see https://www.spc.noaa.gov/gis/svrgis/. To ensure that the structure of this Significant Tornado GIS data set (STORGIS) remains as consistent with the SPC SVRGIS as possible, local times were converted to Central Standard Time (CST) with spatial data given in latitude and longitude coordinates. In addition, like the SVRGIS, the STORGIS exists as both a straight-line track and an initial point model (Figure 1). We feel this allows for more flexibility in the methodologies (i.e., estimating tornado power; Fricker, Elsner, and Jagger 2017a; Elsner, Fricker, and Schroder 2019) that can be applied to the underlying data, opening the field of tornado climatology to a wider interdisciplinary audience.

Results

Preliminary Analysis of Tornado Characteristics

A basic analysis of tornado characteristics can provide insight into the history of significant tornado behavior in the United States. Here, 12,191 tornado reports are evaluated to uncover any preliminary patterns that might exist in the data. Over the period from 1880 to 1989, the annual number of recorded significant and killer tornadoes rose consistently through the mid-1970s before slightly decreasing thereafter (Figure 2). Interestingly, over the same period, the annual average strength—or power—of significant tornadoes was largely stable, showing little variation throughout time (Figure 2). From a human impact perspective, the number of annual tornado fatalities rose consistently through the 1920s before significantly decreasing thereafter (Figure 2), whereas the number of annual tornado injuries also rose consistently through the 1930s before stabilizing through the 1980s (Figure 2).

Seasonally, over the same period, the number of significant and killer tornadoes is highest in the late spring and early summer, with the most tornadoes in May, followed by April, June, and March (Figure 3). The number of significant and killer tornadoes is lowest in the late fall and early winter with the fewest tornadoes in December, followed by October, January, and September (Figure 3). Somewhat consistent with this finding is monthly average tornado power, which is highest in late spring and early summer, but lowest in mid-to-late summer (July and August; Figure 3). From a human impact standpoint, the number of tornado fatalities and injuries are by far the highest in early spring, with over three-fourths (79.6 percent) of all fatalities and injuries (76.8 percent) occurring in the months of March through June (Figure 3).

By time of day, significant and killer tornadoes over the period from 1880 to 1989 occurred in the greatest numbers in the mid- to late afternoon with nearly half of all tornadoes (42.6 percent) occurring between 3:00 pm and 6:59 pm CST. The fewest significant and killer tornadoes occurred during the overnight and early morning hours, with only 9.0 percent of all tornadoes occurring between 1:00 am and 7:59 am. When subset by F-rating, the vast majority (71.5 percent) of all significant and killer

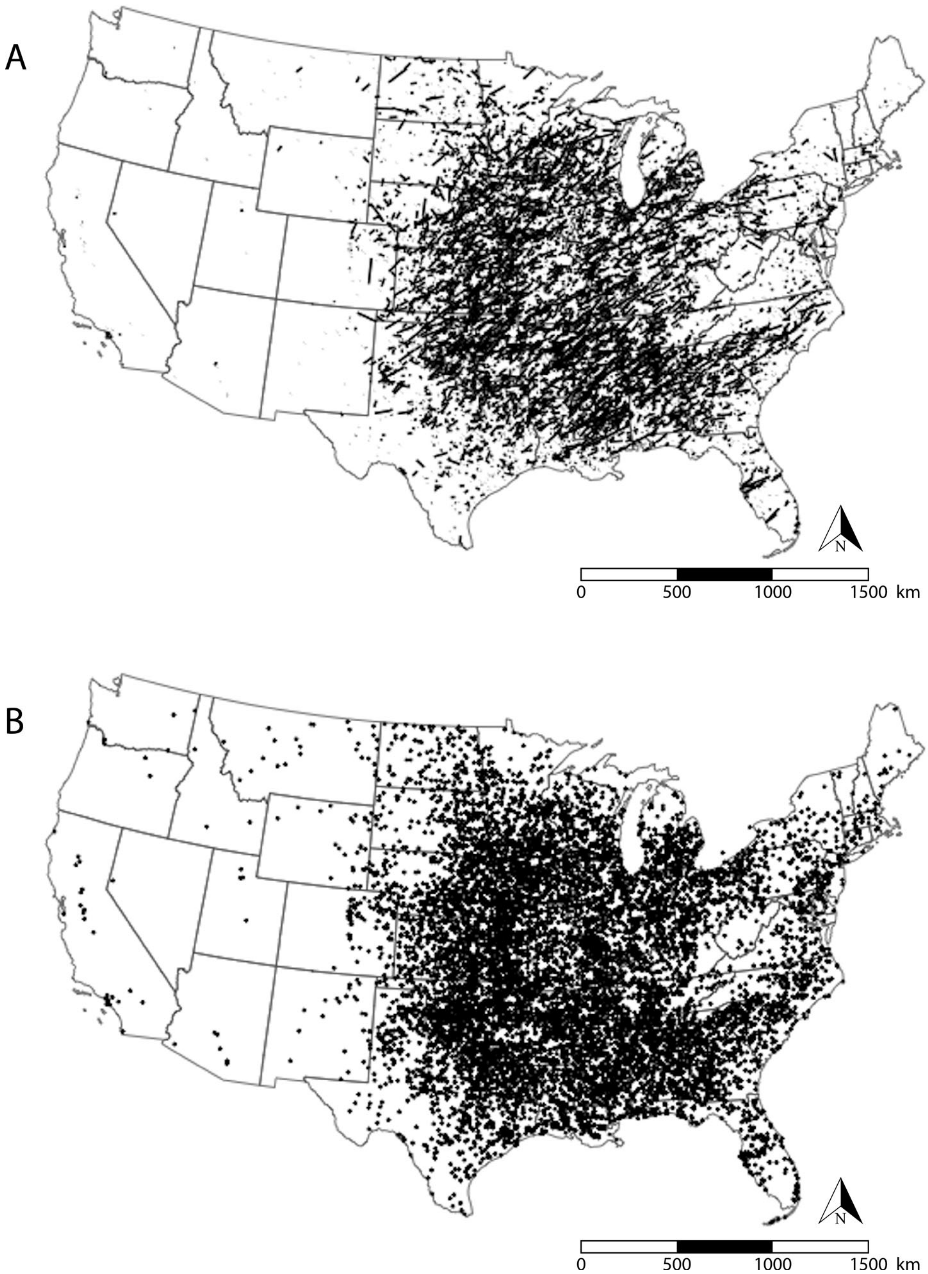

Figure 1 *(A) STORGIS straight-line tracks and (B) STORGIS initial points.*

tornadoes are rated F2, followed by those tornadoes rated F3 (18.7 percent) and those rated F4 (8.0 percent). Only 0.6 percent of all significant and killer tornadoes over the same period are rated F5, which is a smaller percentage (1.2 percent) than those F1 tornadoes that caused at least one fatality.

Grid- and County-Level Analysis

The utility of geocoded historical tornado records can be seen in the spatial evaluation of occurrence rates across the contiguous United States. Because the records have been digitized as a line, finding the number of tornadoes that have intersected grid cells or counties is rather straightforward. A simple intersect will provide a spatial representation of the number of tornadoes that have been recorded across a given area and when combined with spatial data at a similar scale, making sense of the relationship between significant and killer tornadoes and socioeconomic, demographic, and physical factors is possible.

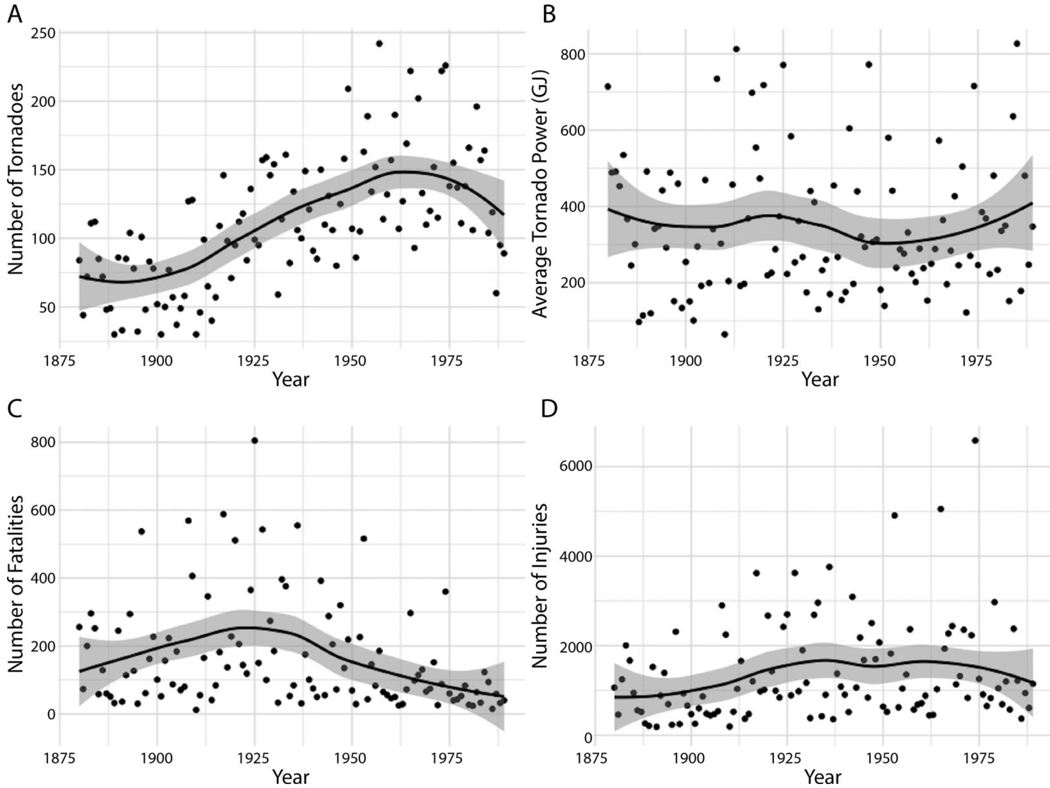

Figure 2 *(A) Number of significant tornadoes, (B) average tornado power, (C) number of tornado fatalities, and (D) number of tornado injuries by year. The black curve is a local regression and the gray band is the 95 percent uncertainty band on the curve.*

Using a 0.6° raster grid, a clear swath of high significant and killer tornado occurrence over the period exists in a tilted "L" shape extending from Iowa—south and west—into northern Texas—and east—through the Mid-South ending into northwest Georgia (Figure 4). The highest occurrence rates exist in the grid cells surrounding Oklahoma City, Oklahoma, Little Rock, Arkansas, and Birmingham, Alabama, with the grid cell over Norman, Oklahoma, affected by the highest number of significant and killer tornadoes (sixty-three). Nearly two-thirds (63.0 percent) of all grid cells were not affected by a single tornado over the period, although this concentration of cells exists in the areas west of the Rocky Mountains, surrounding the Appalachian Mountains, and in the far Northeast United States.

Similar spatial patterns in high significant and killer tornado occurrence exist at the county level, although a less defined swath is seen across the Mid-South (Figure 4). The highest occurrence rates again surround Oklahoma City, Little Rock, and Birmingham, with Pulaski County (Arkansas) and Oklahoma County (Oklahoma) tied as the most affected counties (forty-one). Interestingly, only 19.0 percent of total counties were not affected by a tornado over the period, which is a stark difference from the grid-level analysis, highlighting the large average area of western U.S. counties.

Beyond counts, it is possible to visualize other significant and killer tornado characteristics across space. For example, estimating tornado power allows for insights into the intensity of events rather than pure frequency. Because power is considered an extensive variable, it can be summed through space or calculated as an average. By total power, a spatial pattern of high intensity exists throughout much of the Southern Great Plains, Mid-South, and Midwest, with well-defined swaths from central Oklahoma through southern Kansas, from eastern Louisiana through northern Alabama, and from southeast Missouri through southern Indiana (Figure 5). By average power, the same general spatial distribution of high intensity exists with only a few areas standing out (e.g., northern Pennsylvania, south-central Texas, Oklahoma Panhandle, etc.; Figure 5), a result driven by the impacts of strong tornadoes occurring in small numbers.

Having spatial information attached to the characteristics of significant and killer tornadoes presents the opportunity to analyze statistical relationships on a map. When separated by median values of total power and average power, a clear swath of high total power and high average power exists in an "O"

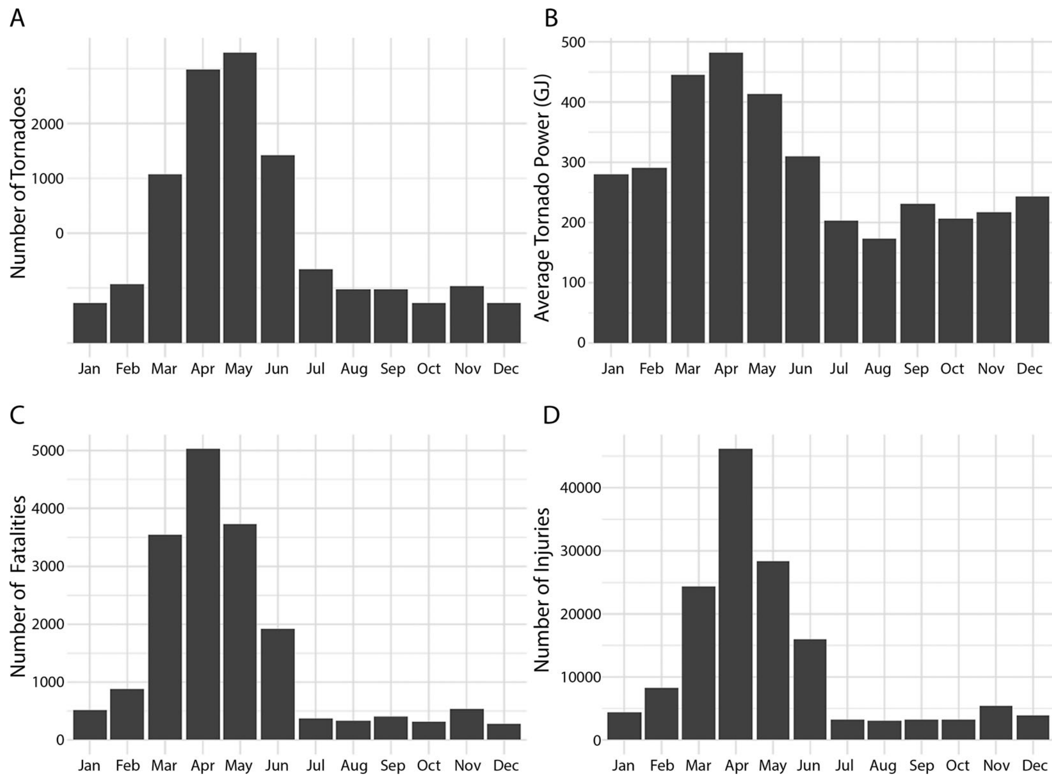

Figure 3 *(A) Number of significant tornadoes, (B) average tornado power, (C) number of tornado fatalities, and (D) number of tornado injuries by month.*

shape throughout the southern Great Plains, Mid-South, and Midwest, with only northern Missouri, southern Iowa, and northern Texas missing (Figure 5). When separated by mean values of total power and average power, a similar swath of high total power and high average power exists with additional gaps present in northern Texas, eastern Arkansas, and western Oklahoma (Figure 5).

Perhaps more interesting than the high–high relationship are the regions of mismatch. For example, the gaps seen in northern Texas, eastern Arkansas, and western Oklahoma suggest that high total intensity exists in these areas, but that they are often affected by many, relatively weaker tornadoes (high–low), whereas the gaps seen in the western edge of the Great Plains (western Kansas and western Nebraska) suggest that low total intensity exists in these areas, but that they are affected by relatively stronger tornadoes (low–high).

Point-Based Analysis

Investigation into the point patterns can also unveil information on the spatial distribution of significant and killer tornado frequency across space. Similar to the spatial analysis of tornadoes at the grid and county level, using the initiation point of significant and killer tornadoes can provide knowledge on high and low occurrence rates at a range of spatial scales. For example, instead of using grid cells or counties as a means of visualizing frequency, points can be overlayed onto city-level maps and smoothed to determine spatial occurrence rates at a finer resolution.

At the city level, high significant and killer tornado occurrence exists in different spatial patterns. For example, the Oklahoma City, Oklahoma, area has a main hot spot just south and east of downtown with a lower secondary hot spot over Shawnee, Oklahoma (Figure 6). The Dallas-Fort Worth, Texas, area has a main hot spot just north of downtown Dallas, with a wider secondary hot spot extending farther north and east of downtown Dallas and an additional secondary hot spot over Denton, Texas (Figure 6). The Birmingham, Alabama, area has two main hot spots over downtown Birmingham and over Cullman-Good Hope-Hanceville, Alabama, and the St. Louis, Missouri, area has a single hot spot over downtown (Figure 6).

Killer tornado occurrence rates show a slightly different spatial pattern at the city level than when all significant and killer tornadoes are considered. For example, Oklahoma City lacks a singular hot spot, although it does have a "U"-shaped pattern of high occurrence rates south of downtown (Figure 7). Dallas-Fort Worth shows three hot spots over

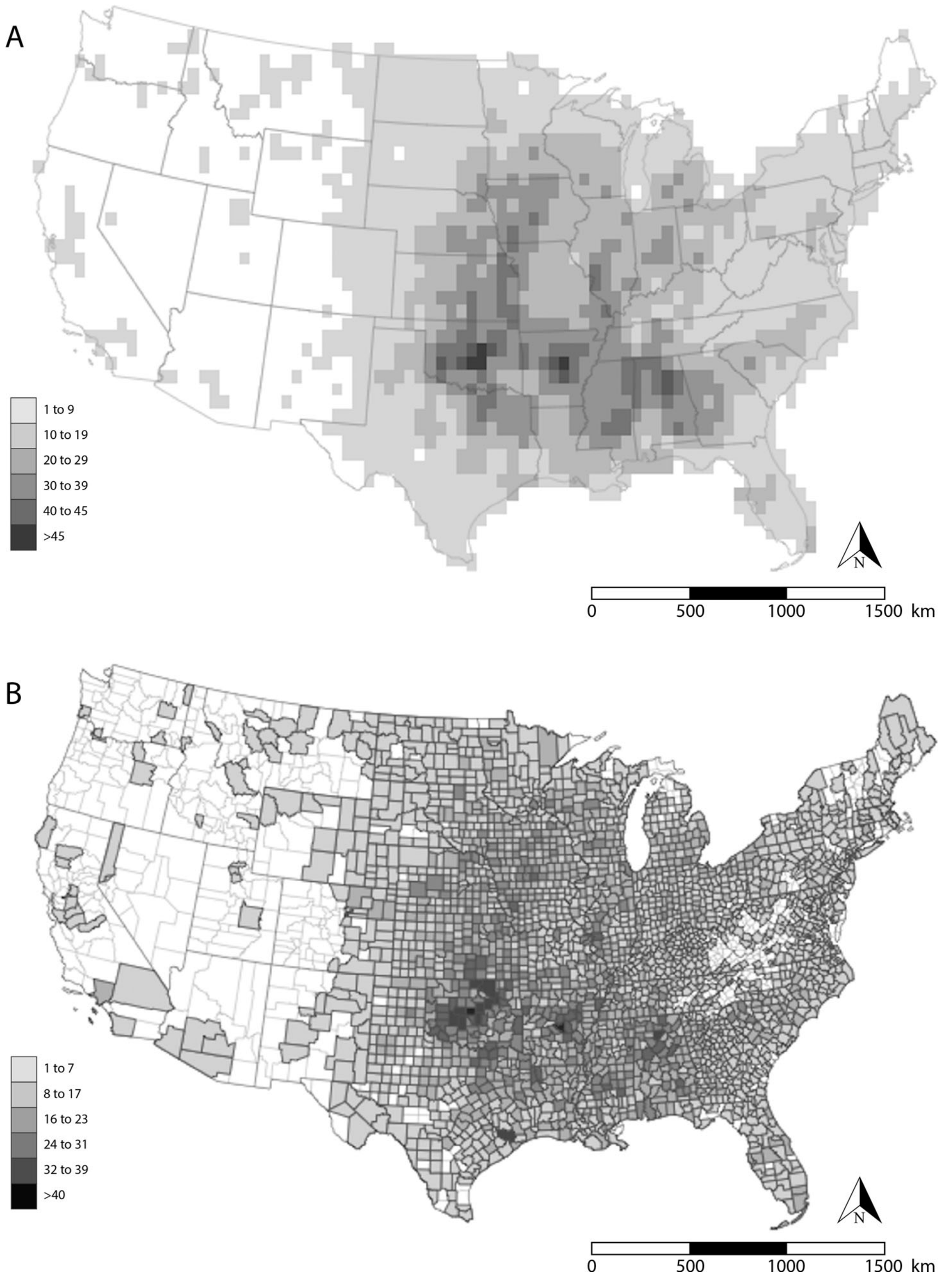

Figure 4 *The number of significant and killer tornadoes intersecting (A) spatial grid (0.6° resolution), and (B) contiguous U.S. counties.*

downtown Dallas, Denton, and Farmersville (Figure 7). In addition, Birmingham has a singular hotspot north and west of downtown, and St. Louis has a singular hot spot over downtown (Figure 7). Overall, these spatial patterns of killer tornado occurrence rates range from more concentrated than all tornado occurrence rates (e.g., Birmingham) to less concentrated than all tornado occurrence rates (e.g., Oklahoma City) to little change between both subsets (e.g., Dallas-Fort Worth and St. Louis).

Discussion and Conclusions

The United States is affected by more tornadoes than any other country on Earth. Because of this, the United States is also home to the most detailed accounts and records of tornadoes in the world. Much of what is known about tornado frequency and intensity and its changing—or unchanging—behavior or connection to climate change comes from the SPC's historical tornado database, dating

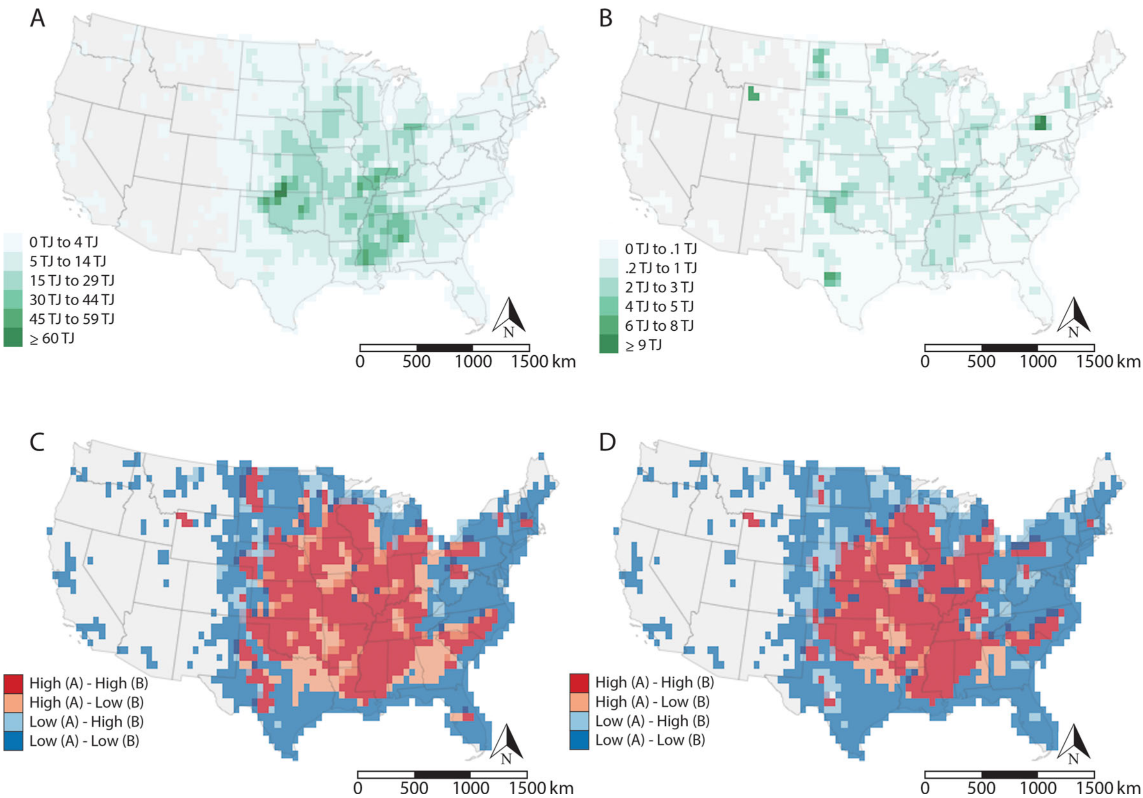

Figure 5 *(A) Total tornado power (TJ), (B) average tornado power (TJ), (C) swaths of high–low total tornado power and average tornado power by median values, (D) swaths of high–low total tornado power and average tornado power by mean values. The spatial grid exists at a 0.6° resolution.*

back to 1950. Missing from this knowledge, however, is a systematic understanding of tornado behavior prior to 1950. In response, here we present a digitized and geocoded significant and killer tornado record over the period from 1880 to 1989 as a means of providing additional information to the study of tornado climatology.

Tornado records were digitized from Grazulis (1990a, 1990b) using ABBYY FineReader PDF software and formatted into CSV files. Referring to newspaper event narratives, individual tornado paths were geocoded to provide a spatial component to the original attribute data set. Keeping consistent with the SPC's historical tornado database, local times were converted to CST with spatial data given in latitude and longitude coordinates. By doing this, both simple and detailed comparisons between the SPC database and this data set can be made for recorded tornado characteristics (i.e., path length and width, magnitude, timing, etc.) during an overlapping study period (1950–1989). For example, when subset by year, the Pearson correlation coefficient for counts between the SPC data set and the digitized data set is .87 (95 percent confidence interval [CI] [.77, .93], p value < 0.05). Similarly, when subset by F-rating, the Pearson correlation coefficient for counts between the SPC data set and the digitized data set is .99 [95 percent CI [.98, .99], p value < 0.05). Thus, on basic characteristic evaluation there is a tight relationship between the SPC data set and the digitized data set. More detailed analysis is needed, however, before a longitudinal study combining both data sets (i.e., tornado climatology from 1880–2020) is undertaken.

The STORGIS data set includes 12,191 significant and killer tornadoes that affected the contiguous United States. Over the study period, the annual number of recorded significant and killer tornadoes rose consistently through the mid-1970s before slightly decreasing thereafter. Explanations for this finding likely include some combination of an improvement in the detection of tornadoes, an increasing population, and the creation and implementation of the F scale (Brooks, Doswell, and Kay 2003; Verbout et al. 2006; Anderson et al. 2007; Elsner et al. 2013). Over the same period, the annual average power of significant tornadoes was largely stable with little variation throughout time, although a more detailed understanding of temporal changes in tornado strength remains plausible.

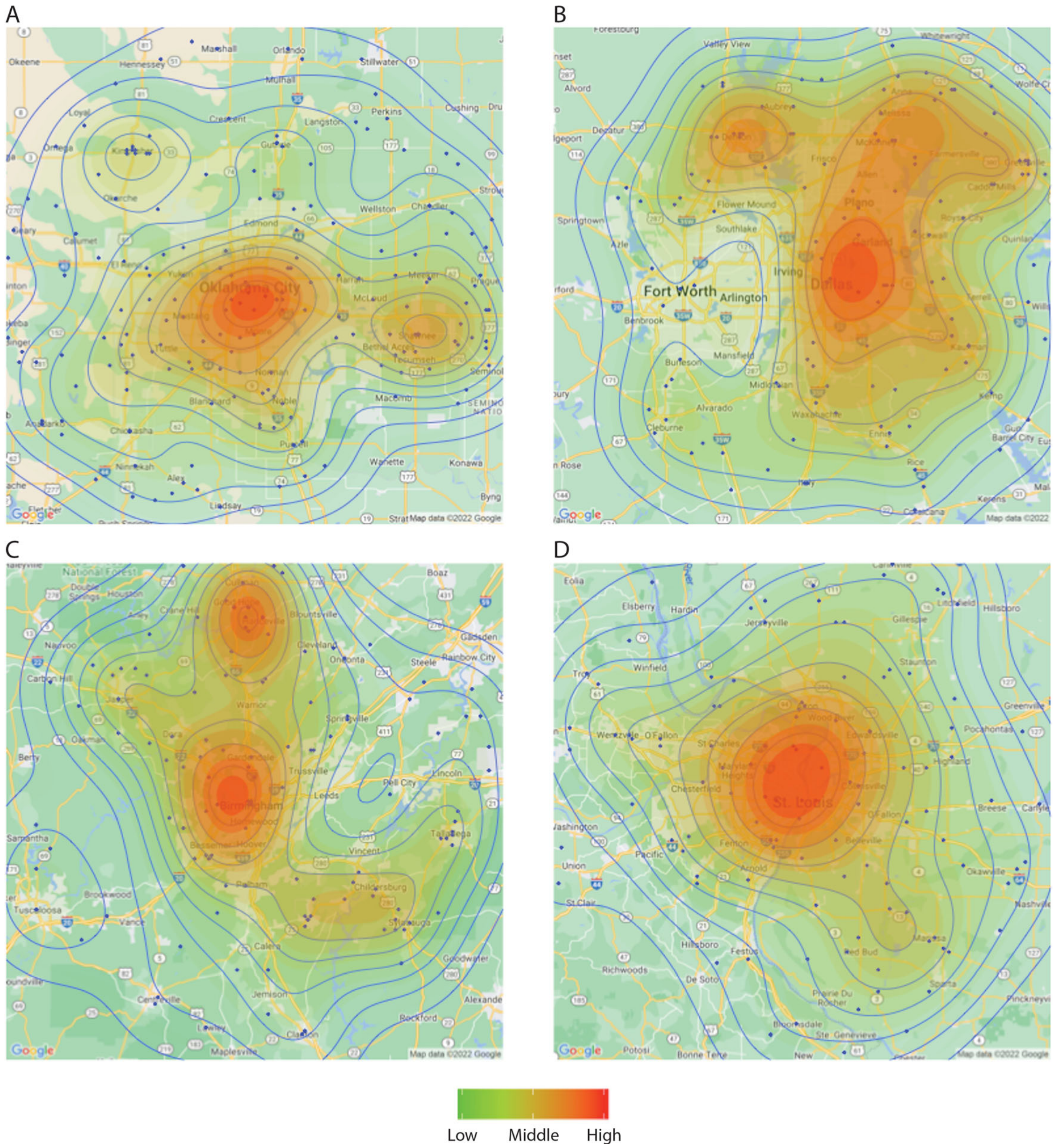

Figure 6 *Density heatmap of significant and killer tornado occurrence in (A) Oklahoma City, Oklahoma, (B) Dallas-Fort Worth, Texas, (C) Birmingham, Alabama, and (D) St. Louis, Missouri. Smoothing is performed using a two-dimensional kernel density estimation with fifteen contour bins. Low (25 percent), middle (50 percent), and high (75 percent) probabilities are shown with a color scale.*

Moving forward, analysis based on estimates derived from F scale categorization and path width measurements (i.e., tornado power) within this data set should contextualize the differences between record keeping before and after the adoption of the Fujita scale in the late 1970s (Edwards et al. 2013).

Seasonally, the number of significant and killer tornadoes is highest in the late spring and early summer and lowest in the late fall and early winter, whereas average tornado power is highest in late spring and early summer and lowest in mid- to late summer. Although these results are unsurprising, it should be noted that the physical drivers of tornadoes—a migrating jet stream, a clash of air masses, and atmospheric instability (Brooks and Doswell 2001; Schultz et al. 2014; Jagger, Elsner, and Widen 2015)—have remained stable over the past two centuries. That said, recent research has commented on

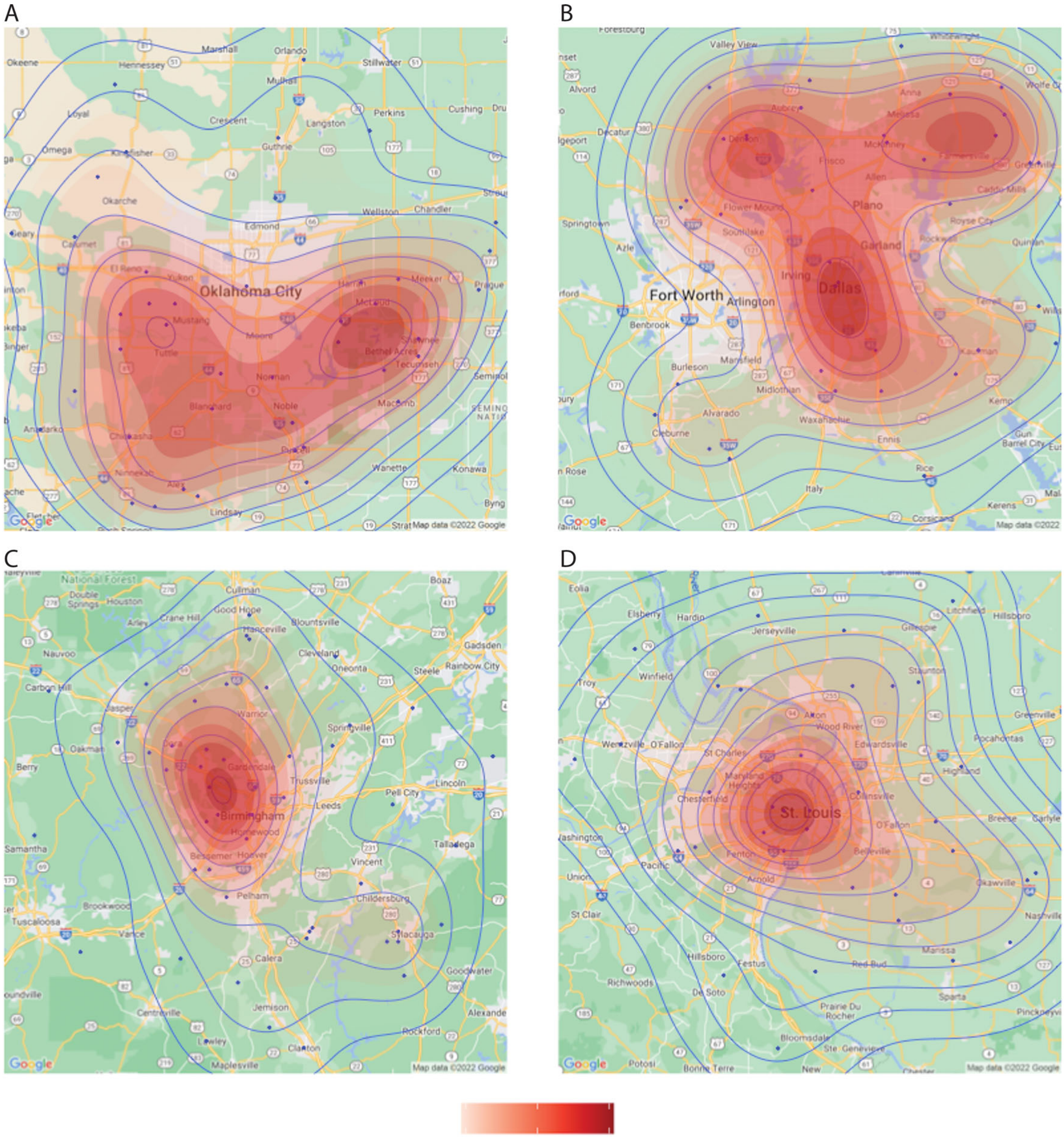

Figure 7 *Density heatmap of killer tornado occurrence in (A) Oklahoma City, Oklahoma, (B) Dallas-Fort Worth, Texas, (C) Birmingham, Alabama, and (D) St. Louis, Missouri. Smoothing is performed using a two-dimensional kernel density estimation with twelve contour bins. Low (25 percent), middle (50 percent), and high (75 percent) probabilities are shown with a color scale.*

a change in peak tornado activity with earlier maximums occurring in the later record (Long and Stoy 2014; Lu, Tippett, and Lall 2015), so it would be interesting to see if the same pattern exists dating back to the late nineteenth century.

Tornado casualties—injuries and fatalities—rose consistently in the data set through the 1930s with annual fatality rates dropping thereafter and injury rates stabilizing through the 1970s before dropping in the 1980s. These findings are in agreement with previous work (Brooks and Doswell 2002; Ashley 2007) and can largely be explained by a combination of the development and improvement in the tornado forecasting process (Doswell, Moller, and Brooks 1999), the creation of a national radar system, and improved communication of watches, warnings, and advisories. By season, tornado casualties occur in the highest numbers in early spring, with 79.6 percent of all fatalities and 76.8 percent of all injuries recorded in the months of March through June, a finding consistent with the work of Ashley (2007) and Fricker et al. (2017). Reasons for this strong seasonality in casualties likely include a combination of high tornado activity (Brooks, Doswell, and Kay 2003) and high tornado intensity (Fricker and Elsner 2015) interacting with a landscape (i.e., Mid-South and Southeast) at higher risk (Coleman and Dixon 2014) and vulnerability (Strader and Ashley 2018; Fricker and Elsner 2020) to negative tornado impacts.

Across the contiguous United States, a clear swath of high significant and killer tornado occurrence over the period exists in a tilted "L" shape extending from Iowa—south and west—into northern Texas—and east—through the Mid-South ending into northwest Georgia at the grid level. This titled "L" shape is similar in structure to the "C" described in Brooks, Doswell, and Kay (2003), but it does not include the upper Midwest. It is also similar to the "L" described in Concannon, Brooks, and Doswell (2000), although it extends farther east, and is close to the pattern seen in Ashley (2007), although the tilted "L" shape is not as linear as a swath from Oklahoma through the Mid-South. At the county level, a less defined swath is seen across the Mid-South, with Pulaski County (Arkansas) and Oklahoma County (Oklahoma) affected by the highest number of significant and killer tornadoes.

Given the STORGIS exists as both a straight-line track and an initial point model, it is possible to leverage point-based information to map occurrence rates at a fine-scale resolution. Here we selected four highly affected U.S. cities (Oklahoma City, Oklahoma; Dallas-Fort Worth, Texas; Birmingham, Alabama; and St. Louis, Missouri) and compared the spatial frequency of tornado initiation points across the urban landscape. Results highlight the potential influence population distribution has on local tornado records. For example, Oklahoma City is a centrally populated metro area with a high population density, whereas Dallas-Fort Worth is a far more sprawling metroplex with a relatively lower population density. This might explain why Oklahoma City has a main hot spot just south and east of downtown whereas Dallas-Fort Worth has a main hot spot just north of downtown Dallas, with a wider secondary hot spot extending farther north and east of downtown Dallas and an additional secondary hot spot over Denton, Texas.

Taking this a step further, questions remain surrounding low values of spatial frequency across Oklahoma City, Dallas-Fort Worth, Birmingham, and St. Louis. For example, it is curious to see no hot spot over Fort Worth given its current population. This result might be a matter of luck—the low probability of significant tornado impacts across space—or driven by the relative difference in population between Fort Worth (lower population and population density) and Dallas (higher population and population density) over the study period. Similarly, it is curious to see no hot spot over Tuscaloosa, Alabama, given its recent history with significant tornado impacts, but this might also be driven by a relatively low population with respect to neighboring Birmingham over the study period. Because tornado reports are often made due to human or property loss, it makes sense that more tornadoes are likely to be reported where more people live. When combined with knowledge on local topography, a deeper understanding of the relationship between tornado behavior, the physical environment, and human populations is possible.

Interestingly, the spatial distribution of killer tornado occurrence rates at the city level do not mirror the spatial distribution of all tornadoes in the four selected cities. For instance, Oklahoma City no longer shows a main downtown hot spot, although it does have a "U"-shaped pattern of high occurrence rates south of downtown. Dallas-Fort Worth shows three hot spots over downtown Dallas, Denton, and Farmersville, and Birmingham has a singular hot spot north and west of downtown—a difference from the two main hot spots over downtown Birmingham and over Cullman-Good Hope-Hanceville, Alabama, seen in all tornado reports. The only city that maintains a singular hot spot when subset by both killer tornadoes and all tornadoes is St. Louis, which underscores how centralized the history of urban development is in the city.

Future work leveraging the STORGIS should employ methods that account for a likely lower quality of data than the modern SPC historical data set (SVRGIS). Before a simple combination of the STORGIS and the SVRGIS is made, a comparison between the characteristics of both data sets should be undertaken. That said, when evaluating the STORGIS alone, we argue further advancement in knowledge of tornado climatology is possible. For example, the STORGIS could be evaluated alongside county-level population data to quantify the relationship between tornado occurrence rates and population over the period from 1880 to 1989. In some ways, this would be an extension of the quantified population bias seen in recent work (Anderson et al. 2007; Elsner et al. 2013; Potvin et al. 2019). In addition, the STORGIS data set could also be used in educational settings to provide baseline data for lessons in geospatial analysis. With a flexible data structure, the STORGIS is well equipped to provide multiple examples of raster and vector data analysis that can be applied in quantitative methods, GIS, and remote sensing courses. ■

Acknowledgments

We thank the two anonymous reviewers for their recommendations and feedback on an earlier draft.

Geocoding of the data was performed by the authors with help from undergraduate students in the Department of Geography at Texas A&M University and the Atmospheric Science Program at the University of Louisiana Monroe. The STORGIS data set is available (as both a straight-line track and an initial point model) from the corresponding author on reasonable request. The code used to produce the graphs and figures is available at https://github.com/tfricker/STORGIS-Climatology.

ORCID

Tyler Fricker http://orcid.org/0000-0002-8186-7399

Literature Cited

Agee, E., J. Larson, S. Childs, and A. Marmo. 2016. Spatial redistribution of US tornado activity between 1954 and 2013. *Journal of Applied Meteorology and Climatology* 55 (8):1681–97.

Allen, J. T., M. K. Tippett, and A. H. Sobel. 2015. Influence of the El Niño/Southern Oscillation on tornado and hail frequency in the United States. *Nature Geoscience* 8 (4):278–83.

Anderson, C. J., C. K. Wikle, Q. Zhou, and J. A. Royle. 2007. Population influences on tornado reports in the United States. *Weather and Forecasting* 22 (3):571–79.

Ashley, W. S. 2007. Spatial and temporal analysis of tornado fatalities in the United States: 1880–2005. *Weather and Forecasting* 22 (6):1214–28.

Bluestein, H. B. 2013. *Severe convective storms and tornadoes*. Chichester, UK: Springer.

Brooks, H. E., G. W. Carbin, and P. T. Marsh. 2014. Increased variability of tornado occurrence in the United States. *Science* 346 (6207):349–52. doi: 10.1126/science.1257460.

Brooks, H. E., and C. A. Doswell, III. 2001. Normalized damage from major tornadoes in the United States: 1890–1999. *Weather and Forecasting* 16 (1):168–76.

Brooks, H. E., and C. A. Doswell. III. 2002. Deaths in the 3 May 1999 Oklahoma City tornado from a historical perspective. *Weather and Forecasting* 17:354–61.

Brooks, H. E., C. A. Doswell, and M. P. Kay. 2003. Climatological estimates of local daily tornado probability for the United States. *Weather and Forecasting* 18 (4):626–40.

Coleman, T. A., and P. G. Dixon. 2014. An objective analysis of tornado risk in the United States. *Weather and Forecasting* 29 (2):366–76.

Concannon, P. R., H. E. Brooks, and C. A. Doswell, III. 2000. Climatological risk of strong and violent tornadoes in the United States. In *Preprints, 2nd symposium on environmental applications*, Vol. 9. Long Beach, CA: American Meteorological Society.

Diffenbaugh, N. S., M. Scherer, and R. J. Trapp. 2013. Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proceedings of the National Academy of Sciences* 110 (41):16361–66.

Doswell, C. A., III, A. R. Moller, and H. E. Brooks. 1999. Storm spotting and public awareness since the first tornado forecasts of 1948. *Weather and Forecasting* 14 (4):544–57.

Edwards, R., H. E. Brooks, and H. Cohn. 2021. Changes in tornado climatology accompanying the enhanced Fujita scale. *Journal of Applied Meteorology and Climatology* 60 (10):1465–82.

Edwards, R., J. E. LaDue, J. T. Ferree, K. Scharfenberg, C. Maier, and W. L. Coulbourne. 2013. Tornado intensity estimation: Past, present, and future. *Bulletin of the American Meteorological Society* 94 (5):641–53.

Elsner, J. B., S. C. Elsner, and T. H. Jagger. 2015. The increasing efficiency of tornado days in the United States. *Climate Dynamics* 45 (3):651–59.

Elsner, J. B., T. Fricker, and Z. Schroder. 2019. Increasingly powerful tornadoes in the United States. *Geophysical Research Letters* 46 (1):392–98.

Elsner, J. B., T. H. Jagger, and T. Fricker. 2016. Statistical models for tornado climatology: Long and short-term views. *PLoS ONE* 11 (11):e0166895.

Elsner, J. B., T. H. Jagger, H. M. Widen, and D. R. Chavas. 2014. Daily tornado frequency distributions in the United States. *Environmental Research Letters* 9 (2):024018.

Elsner, J. B., L. E. Michaels, K. N. Scheitlin, and I. J. Elsner. 2013. The decreasing population bias in tornado reports across the Central Plains. *Weather, Climate, and Society* 5 (3):221–32.

Finley, J. P. 1882. *Report on the character of six hundred tornadoes* (No. 7). Washington, DC: Office of Chief Signal Officer.

Fricker, T., and J. B. Elsner. 2015. Kinetic energy of tornadoes in the United States. *PLoS ONE* 10 (7):e0131090.

Fricker, T., and J. B. Elsner. 2020. Unusually devastating tornadoes in the United States: 1995–2016. *Annals of the American Association of Geographers* 110 (3):724–38.

Fricker, T., J. B. Elsner, and T. H. Jagger. 2017. Population and energy elasticity of tornado casualties. *Geophysical Research Letters* 44 (8):3941–49.

Fricker, T., J. B. Elsner, V. Mesev, and T. H. Jagger. 2017. A dasymetric method to spatially apportion tornado casualty counts. *Geomatics, Natural Hazards and Risk* 8 (2):1768–82.

Gensini, V., and W. S. Ashley. 2011. Climatology of potentially severe convective environments from the North American regional reanalysis. *E-Journal of Severe Storms Meteorology* 6 (8):1–40.

Gensini, V. A., and H. E. Brooks. 2018. Spatial trends in United States tornado frequency. *NPJ Climate and Atmospheric Science* 1 (1):1–5.

Grazulis, T. P. 1990a. *Significant tornadoes, 1880–1989: Discussion and analysis*, Vol. 1. St. Johnsbury, VT: Environmental Films.

Grazulis, T. P. 1990b. *Significant tornadoes, 1880–1989: A chronology of events*, Vol. 2. St. Johnsbury, VT: Environmental Films.

Hoogewind, K. A., M. E. Baldwin, and R. J. Trapp. 2017. The impact of climate change on hazardous convective weather in the United States: Insight from high-resolution dynamical downscaling. *Journal of Climate* 30 (24): 10081–10100.

Jagger, T. H., J. B. Elsner, and H. M. Widen. 2015. A statistical model for regional tornado climate studies. *PLoS ONE* 10 (8):e0131876.

Koch, E., J. Koh, A. C. Davison, C. Lepore, and M. K. Tippett. 2021. Trends in the extremes of environments associated with severe US thunderstorms. *Journal of Climate* 34 (4):1259–72.

Krocak, M. J., and H. E. Brooks. 2018. Climatological estimates of hourly tornado probability for the United States. *Weather and Forecasting* 33 (1):59–69.

Long, J. A., and P. C. Stoy. 2014. Peak tornado activity is occurring earlier in the heart of "Tornado Alley." *Geophysical Research Letters* 41 (17):6259–64.

Lu, M., M. Tippett, and U. Lall. 2015. Changes in the seasonality of tornado and favorable genesis conditions in the central United States. *Geophysical research Letters* 42 (10):4224–31.

Molina, M. J., R. P. Timmer, and J. T. Allen. 2016. Importance of the Gulf of Mexico as a climate driver for US severe thunderstorm activity. *Geophysical Research Letters* 43 (23):12–295.

Moore, T. W. 2017. On the temporal and spatial characteristics of tornado days in the United States. *Atmospheric Research* 184:56–65.

Moore, T. W. 2018. Annual and seasonal tornado trends in the contiguous United States and its regions. *International Journal of Climatology* 38 (3):1582–94.

Moore, T. W., J. M. S. Clair, and M. P. McGuire. 2022. Climatology and trends of tornado-favorable atmospheric ingredients in the United States. *Annals of the American Association of Geographers* 112 (2):331–49.

Moore, T. W., and T. A. DeBoer. 2019. A review and analysis of possible changes to the climatology of tornadoes in the United States. *Progress in Physical Geography: Earth and Environment* 43 (3):365–90.

Moore, T. W., and T. Fricker. 2020. Tornadoes in the USA are concentrating on fewer days, but their power dissipation is not. *Theoretical and Applied Climatology* 142 (3):1569–79.

Potvin, C. K., C. Broyles, P. S. Skinner, H. E. Brooks, and E. Rasmussen. 2019. A Bayesian hierarchical modeling framework for correcting reporting bias in the US tornado database. *Weather and Forecasting* 34 (1):15–30.

Schultz, D. M., Y. P. Richardson, P. M. Markowski, and C. A. Doswell. 2014. Tornadoes in the central United States and the "Clash of Air Masses." *Bulletin of the American Meteorological Society* 95 (11):1704–12.

Seeley, J. T., and D. M. Romps. 2015. The effect of global warming on severe thunderstorms in the United States. *Journal of Climate* 28 (6):2443–58.

Strader, S. M., and W. S. Ashley. 2018. Finescale assessment of mobile home tornado vulnerability in the central and southeast United States. *Weather, Climate, and Society* 10 (4):797–812.

Taszarek, M., J. T. Allen, H. E. Brooks, N. Pilguj, and B. Czernecki. 2021. Differing trends in United States and European severe thunderstorm environments in a warming climate. *Bulletin of the American Meteorological Society* 102 (2):E296–E322.

Thompson, R. L., and M. D. Vescio. 1998. The destruction potential index—A method for comparing tornado days. In *Preprints, 19th Conference on Severe Local Storms*, Vol. 280, 282. Minneapolis, MN: American Meteorological Society.

Tippett, M. K., J. T. Allen, V. A. Gensini, and H. E. Brooks. 2015. Climate and hazardous convective weather. *Current Climate Change Reports* 1 (2):60–73.

Trapp, R. J., N. S. Diffenbaugh, H. E. Brooks, M. E. Baldwin, E. D. Robinson, and J. S. Pal. 2007. Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. *Proceedings of the National Academy of Sciences* 104 (50):19719–23.

Trapp, R. J., and K. A. Hoogewind. 2016. The realization of extreme tornadic storm events under future anthropogenic climate change. *Journal of Climate* 29 (14):5251–65.

Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz. 2006. Evolution of the US tornado database: 1954–2003. *Weather and Forecasting* 21 (1):86–93.

Weaver, S. J., S. Baxter, and A. Kumar. 2012. Climatic role of North American low-level jets on US regional tornado activity. *Journal of Climate* 25 (19):6666–83.

TYLER FRICKER is an Assistant Professor of Geography in the Atmospheric Science Program at the University of Louisiana Monroe, Monroe, LA 71209. E-mail: tfricker@ulm.edu. His research interests include the environmental impacts of natural hazards on society and the connection between weather and climate.

TAYLOR WELLS is a Human Resources Specialist. E-mail: twells5@clemson.edu. She holds a Master of Public Administration degree from Clemson University and a BS in Environmental Studies from Texas A&M University.

JOHN THOMAS ALLEMEIER is a Graduate Student in Applied Statistics at Florida State University. E-mail: jta18d@fsu.edu. He holds a BS in Statistics from Florida State University. His research interests are in climate and measurement.

AMELIA BAXTER is a Development Associate for a Texas nonprofit. E-mail: ameliaabaxter@gmail.com. She holds a BS in International Affairs and a BS in Geography from Florida State University.