CONVECTIVE MODE CLASSIFICATION AND CLIMATOLOGY OF US TORNADO EVENTS 2000-2021

8.1B

ANDREW D. LYONS*, BRYAN T. SMITH, RICHARD L. THOMPSON ANDREW R. DEAN
NOAA/NWS/NCEP/Storm Prediction Center *Norman*, *Oklahoma.*

1. INTRODUCTION

Tornadoes and severe thunderstorms represent a significant threat to life and property in the contiguous United States (CONUS) annually. This study builds on prior work from Smith et al. (2012) which featured manual, single-site radar storm-mode classification of 10724 tornado events in the CONUS from 2003-2011. An additional 9961 tornado events were added (N=19941) from the years of 2012-2021. Events from 2000-2002 were examined, but ultimately excluded due to poor data quality - see results section for additional information. Various statistics along with spatial and temporal analyses were completed and compared to the original sample to build a robust multidecadal climatology of tornadoes and convective mode across the CONUS.

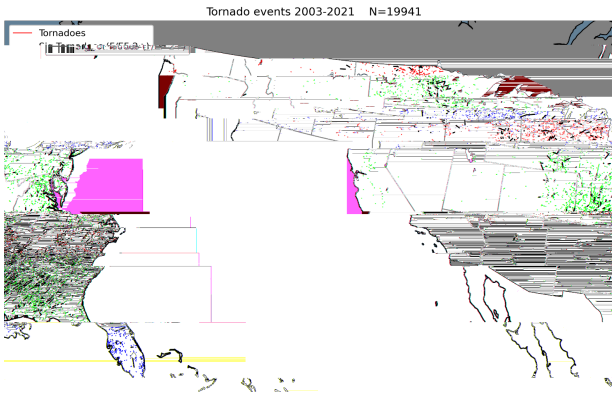
1.1 BACKGROUND

Thunderstorms across the CONUS vary to a wide degree in convective mode, organization and intensity. Tornadoes occur with all storm types, but occur more frequently with several organized modes, predominantly supercells. Smith et al. (2012) developed estimated climatologies for severe storms and tornadoes using nine years of manually-analyzed, single-site radar imagery and storm report data filtered by hour on a 40-km horizontal grid. Though their sample of events was large, the sample size was still limited in terms of multi-decadal climatology information related to convective mode. Additionally, interdecadal variability could not be assessed due to the limited sample. This study's goal was to produce a comprehensive, multi-decadal storm-mode climatology for the majority of the Weather Surveillance Radar 1988 Doppler (WSR-88D) era, building on the original work by Smith et al. (2012). Of particular interest are spatial estimates of events per decade, and changes in the distribution of tornado reports by convective mode over the nineteen-year period.

2. METHODOLOGY

Tornado reports over the research period were gathered from NOAA NCEI's publication Storm Data (2021) and filtered for duplicates/errors. Tornado events were matched with grid hours of SPC RUC/RAP mesoanalysis environmental information (Bothwell et al. 2002) and combined into a large database. Single-site radar imagery corresponding to each event was gathered approximately 1 hour before and 1 hour after the grid hour reported time of an event to ensure adequate temporal coverage. Radar data was obtained from NOAA's NCEI Level-2 radar archive and analyzed by Storm Prediction Center (SPC) forecasters using Gibson Ridge Software LLC GRLevel-2 Radar viewer. SPC Forecasters made a subjective classification of the storm mode into 5 main categories for a given event. The five main mode categories are; right- or left-moving supercell (RM or LM), quasi-linear convective system (QLCS), supercell/QLCS hybrid, marginal supercell, and disorganized. Several sub modes were also included for the main categories; discrete cells, cell in cluster and cell in line. Discrete cell and cell in cluster applied to RM/LM and disorganized modes, while cell in line was limited to RM/LM in line or QLCS (including the hybrid designation). The QLCS mode also allowed for the bow echo sub-designation.

Storms were sorted into the listed mode and sub-mode categories using the WSR-88D data and report data. Supercell mesocyclones were also categorized as weak, moderate and strong, based on rotational velocities and circulation diameter of associated low and mid-level mesocyclones. Approximately 23,453 events were initially classified from 2000-2021. Filtering of the events occurred due to errors with reports, duplications, or missing radar data. Where Level-2 radar data was missing, level-3 single-site or composite radar imagery was used. This left 19941 events over the nineteen years from 2003-2021 to be included in the final dataset (**Fig. 1**).



(**Fig. 1** Tracks of all tornado (significant F/EF-2+ in black) events in the CONUS 2003-2021)

2.1 Spatial Analysis methods

After the classification and filtering of the dataset was completed, tornado events were separated into various groups by convective mode, seasonal occurrence, magnitude, etc.. The Python package SciPy was used to sort tornado events from the database onto a 40x40 km rectilinear grid for spatial analysis. Grid cell counts of tornadoes were used to calculate relative frequencies as well as averages of events per square kilometer. Where data were sparse with low counts per grid cell, super-sampling (grid search radii of 80 km) was used to boost grid-cell counts with the information being projected back onto the 40 km grid.

Gridded data were also used to construct Gaussian kernel density estimates (KDE) of events per decade of various convective modes, seasons, and event magnitudes (e.g., Brooks 2003). These plots were compared across the two samples, and as a whole to assess changes in spatial distributions and occurrence of events and changes between the new and original datasets.

3. Results

After sorting and classification of all tornado events into the respective convective modes for the years, 2000-2021 convective mode data was analyzed and compared to the original nine-year sample, and in aggregate to identify significant spatial, temporal, and frequency trends of different convective modes and magnitudes of tornado events. Substantial effort was undertaken to extend the original dataset to overlap with the majority of the WSR-88D radar era Smith et al., (2012). The period from 2000-2002 was examined with approximately 2768 tornado events. Significant data quality issues were found with the single-site radar archive from 2000-2002. Approximately 19% of all single-site radar data were missing for tornado events. In addition to the missing archive information, radar-data errors including; insufficient data quality for the determination of mode, or various errors in radial velocity that made mesocyclone strength, tornado location, or storm mode challenging to assess were found. In the event that no data could be found for a corresponding event, Level-3 composite radar data was used to make the best guess determination of storm mode. However, this resulted in no information being available for mesocyclone strength determination given the lack of radial velocity information. Data quality was found to degrade further prior to the year 2000, seriously limiting useful analysis. In addition to poor radar imagery data quality, SPCRUC/RAP mesoanalysis archives only extended back to 2003, further limiting environmental analysis. Despite data quality issues, estimated storm mode data from 2000-2002 was completed with missing events removed. This data was compared to the three most recent years of the original sample (2003-2005) from Smith et al. (2012) as a check for consistency. Analysis revealed 2000-2003 data to be similarly distributed to the original sample. Similar relative frequencies of supercells, QLCS, and hybrid tornado events were found when comparing the original sample and the additional events. However, strong interannual variability was present and the 2000-2002 period saw much lower occurrence of tornado events per year, contributing to some differences in event counts (Storm Data 2021). Despite similarities with the longer sample, data from 2000-2003 were not included in the final analysis due to the aforementioned data quality issues. Final filtering of the entire database from 2000-2021 resulted in only events from 2003 onward being used for consistency, with 443 event errors removed for N=19941 events.

Mode 2003-11	Number of Events	Fraction
Supercell	8055	0.75
QLCS	1485	0.14
Disorganized	1184	0.11
Total	10723	1

Mode 2012-21	Number of Events	Fraction
Supercell	6099	0.63
QLCS	2645	0.27
Disorganized	917	0.09
Total	9661	1

Mode	Number of Events	Fraction
[illegible]	[illegible]	[illegible]
0.21	QLCS	4097
0.1	Disorganized	2078
1	Total	19941

Tables 1, 2 & 3. Counts and relative fractions of tornado events by convective mode; 2003-2011, 2012-2021, and 2003-2021.

Supercells accounted for the majority of tornado occurrences across the CONUS with an average relative frequency from 2003-2021 of 69% (**Table 3**). QLCS tornadoes also account for approximately 21% of all tornadoes in the US, with hybrid and disorganized modes making up the remaining 10% of all events (**Table 3**). Events between the two-decade samples showed similar frequency distributions of convective modes. However, key differences between supercells and QLCS modes were noted (**cf. Tables 1 & 2**). Disorganized modes saw the least amount of change over the two samples. A decrease in the relative frequency of supercell events was noted from 2012-2021 when compared to the original 2003-2011 sample. Interdecadal variability in the occurrence of tornado events likely played some role in reducing the occurrence of supercell events. For example, many years in the original nine-year sample were above average in tornado activity (Storm Data 2021) while the opposite was true from 2012 to 2021 with below-average tornado reports for several years. Thus, lower tornado counts and changes in observing practices for other modes potentially resulted in lower fractions of supercell events. QLCS events saw a large increase in the relative frequency of occurrence from 2012 to 2021. See section **3.2** for further analysis of changes in QLCS distributions.

3.1 Kernel Density Estimates

KDEs of events per decade provided a useful metric to assess the spatial and temporal trends in tornado events by various convective modes, seasons and other characteristics . An initial goal was to recreate original KDE figures from Smith et al. (2012). This was completed successfully and similar spatial results were found using the original 2003-2011 sample (**Fig. 2a in Appendix)**. Key similarities in the recreated figure corroborated Smith et al. (2012) findings that the maximum tornado occurrence was across portions of the lower Mississippi Valley and southeastern CONUS, with a large area of six to seven events per decade. Frequent tornado occurrences (four to five events per decade) also extended away from the maximum over portions of the Mid-Mississippi Valley, southern Great Plains, and Midwest.

2012-2021 tornado events were analyzed using the same KDE method as used in the recreation of the original figures. These data were found to be similarly distributed, with a maximum in tornado frequency centered over portions of the southeastern CONUS. Variability was noted in the estimate of events per decade between the two samples, primarily over portions of the Midwest and Great Plains. Coverage across these regions was spatially similar to 2003-2011, but featured slightly lower magnitudes at three to four events per decade. This change was likely due to lower overall tornado numbers on average over the period due to interdecadal variability. However, this was a small difference suggesting that the original Smith et al. (2012) estimates were quite accurate. (See **Figs. 2b2c in Appendix**)

In addition to the individual decade samples, 2003-2021 KDE data was viewed in aggregate to provide a robust multi-decadal estimate of events per decade (**Fig. 2d**)

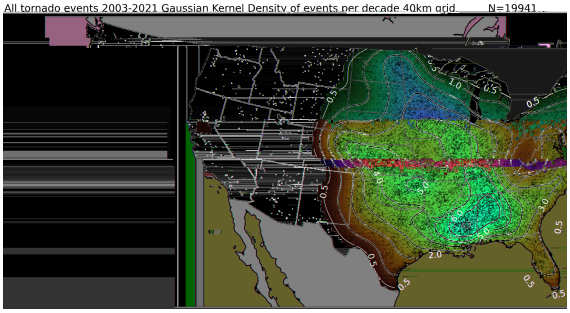


Fig 2d. Gaussian Kernel Density Estimate of tornado

events per decade from all convective modes 2003-2021 Color fill begins at 0.4 events per decade.

The aggregate data retains many features of the individual decade samples, including the maximum tornado occurrence of seven events per decade over portions of the lower Mississippi Valley and Southeast. Estimates of four and five events per decade are common over the Great Plains and Midwest, but decrease with the western extent toward the Rocky Mountains. Prominent minima over the Appalachians and west of the Continental divide are also present and well-represented in previous publications (e.g., Brooks 2003).

Using the combined sample, KDEs were used to estimate tornado occurrence by different convective modes (**Fig. 3a**). Supercell events had maxima over the Southeast and portions of the central and southern Plains. The maximum in QLCS events extended from the lower Mississippi Valley, across the Southeast, into portions of the lower Ohio Valley, and upper Midwest. Hybrid supercell and QLCS events, unsurprisingly, occurred where the greatest overlap of supercells and QLCSs were observed, primarily in the southeastern CONUS. Disorganized storm events also displayed a robust spatial signal with the majority of events occurring within the High Plains, upper Midwest, and Florida.

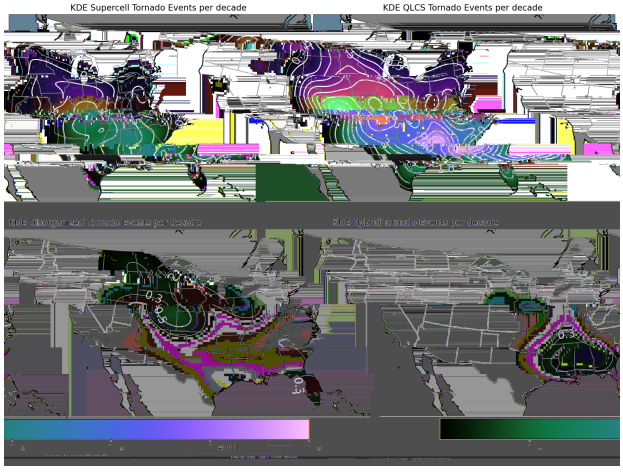


Fig. 3a. Gaussian KDE of events per decade of tornadoes by different convective modes. Color fill begins at 0.1. Top left to bottom right: Supercells, QLCS, Hybrid Supercell/QLCS, and Disorganized

The convective mode also varied spatially with a strong seasonal component. Supercells displayed the most spatial variability by season with disorganized modes also displaying strong seasonal trends. Supercells were most common near the Gulf of Mexico through the cool season before transitioning to the Great Plains and Midwest while increasing in occurrence through boreal spring (MAM; **Fig. 3b in Appendix**). In the summer (JJA) supercells are most common across the upper Midwest and High Plains before returning to the Southeast and Gulf Coast during the fall (SON). Disorganized modes followed a similar but weaker overall trend, migrating from the Gulf Coast and Florida to the upper Midwest and High plains during the warm season (not shown). QLCS events displayed little seasonal spatial change but did become a larger fraction of tornado events through the fall, winter, and early spring. (**Figs, 4a-c in Appendix**) which agrees with Trapp et al., (2005) showing QLCS events in the cool season..

3.2 Differences in Convective Modes QLCS

Large changes in the relative frequency of tornadoes associated with QLCS modes were observed between the two decades. (**Tables 1-2**) QLCS tornado events increased by 78%, or roughly 100 more events per year on average through the period from 2012-2021. Much of this change is likely the result of an increased emphasis on NWS warning practices in detecting and classifying QLCS tornado events with the advent of dual polarimetric radar Schaumann and Przybylinski (2012). Interannual and interdecadal variability may also explain some changes in the frequency, as supercell tornado counts also decreased from 2012-2021. Increased detection in QLCS tornadoes with a simultaneous decrease in overall tornado counts supports higher relative fractions of QLCS events. While some annual/decadal variability is to be expected, spatial distributions of the changes in QLCS tornadoes were not consistent across the CONUS. Large changes in the spatial occurrence were noted around areas of dense population, particularly across portions of the Great Lakes, upper Midwest, Mid Mississippi Valley,

Southeast, and Northeastern US (**Fig.5**).

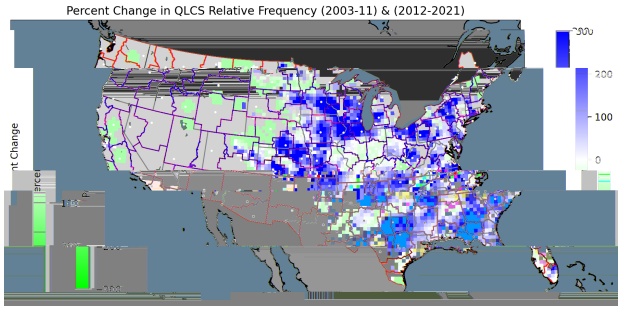


Fig. 5. Percent Change in QLCS relative frequency per grid cell between 2003-2011 and 2012-2021. Events were super-sampled at 80km and plotted on a 40-km grid. Green polygons are NWS County Warning Area boundaries and white dots indicate the location of NWS WSR-88Ds.

The non-uniform changes in relative frequency across the CONUS are heavily correlated with population centers, radar locations, and political boundaries, including NWS County Warning Areas. Supercell events (not shown) display no such similar trend over the same time period. This strongly suggests that recent changes in QLCS tornado occurrence are related not only to changes in observations with dual-pol radar, but possibly to individual NWS office reporting approaches. Thus, while significant changes have been noted in this dataset, conclusions about the natural spatial and temporal occurrence and variability of QLCS tornadoes are difficult to quantify. A more in-depth study is likely needed.

3.3 Gridded Observations

In addition to smoothed KDEs, tornado events were gridded and their spatial distributions were analyzed for significant trends. For example, grid cells where tornado occurrence was high provided a robust estimate of the relative frequency of tornadoes by supercell convective mode across the Great Plains, where approximately 80% of all tornadoes occurred with supercells **(Fig. 6**). Much lower relative frequencies of 40-50% were found across portions of the mid-Mississippi and lower Ohio River Valleys, colocated with the maximum in QLCS events (**cf. Figs. 3a & 6**).**.** These results were found to be extremely similar to KDE frequency estimates from Smith et al. (2012). However, the use of actual grid cell counts provides a more robust estimate due to the true occurrence of events per grid cell rather than KDEs. Several other relative maxima in supercell events were noted across portions of the Gulf Coast and in the lee of the central Appalachians corresponding to maxima in the supercell event KDE from Section **3.1**.

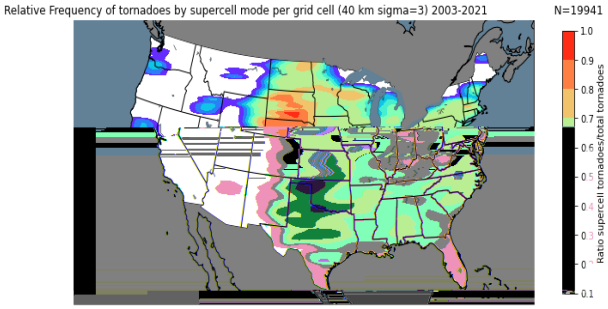


Fig. 6. Smoothed gridded relative frequency analysis of tornadoes by supercell mode per 40 km grid cell with a 120-km Gaussian smoother from 2003-2021.

Gridded counts of significant (F/EF-2+) tornadoes (N=2633) were also compiled from 2003-2021. Many significant tornadoes were found to occur near or east of the Mississippi River Valley, across portions of the Mid-South and southeastern US. Relative frequencies for significant tornadoes peaked near 20% in this region, significantly higher than portions of the Great Plains and Midwest (**Fig. 7)**. This coincides with the maximum tornado occurrence and increased coverage of damage indicators and population. Tornadoes across the Plains displayed a similar low rating bias to Wurman et al. (2021) despite the occurrences of tornadoes displaying significant VROT signatures Smith et al. (2020). Our data support that significant tornado ratings are related to the coverage of damage indicators and population and are therefore subject to being biased to locations near and east of the Mississippi River. However, care should be taken in the interpretation of results with significant tornadoes due to the limited sample size of events.

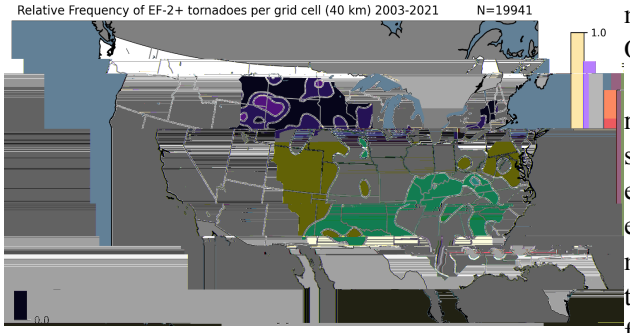


Fig. 7. Gridded relative frequencies of significant (F/EF-2+) tornadoes per 40 km grid cell smoothed using a 120-km Gaussian filter from 2003-2021. Cells with less than 5 tornado events are filtered to increase statistical robustness.

4. Summary and Discussion

In summary, the addition of data from 2012-2021 positively benefited the original Smith et al. (2012) sample by adding 10 additional years and 9961 tornado events. While 2000-2002 data was not used in the final database due to poor radar quality, tornado events by convective mode were similarly distributed seasonally, by fraction of mode, and spatially to the original sample.

Recreation of original KDE figures from Smith et al. (2012) was successful and similar spatial trends were found with the dataset post-2012 (**Figs. 2a-d**). KDE analysis showed tornado events within the period from 2003-2021 were most common over portions of the lower Mississippi Valley and southeastern CONUS, outside of traditional tornado alley. However, areas of the Great Plains and Midwest did see relatively frequent tornado events (four to five per decade) over the nineteen-year period of study. Supercells accounted for the majority of tornadoes across the Great Plains and Southeast with relative fractions often greater than 70%. QLCS events were most common over the Southeast and Ohio River Valleys. Strong indications of interannual and interdecadal spatial variability were present within the individual samples. However, the aggregate dataset was relatively consistent in showing tornado frequency across the Plains, Midwest, and Southeast. Within individual mode data, QLCS tornado fractions likely increased with improved radar observations, complex geopolitical factors, and lower overall tornado counts. However, a more rigorous study is needed to quantify changes in QLCS event distributions and trends.

Significant tornado (F/EF-2+) events regardless of convective mode displayed a significant spatial trend with higher relative frequencies and estimated recurrence per decade located along and east of the Mississippi River (**Figs. 2d and 7**). This matches other work suggesting a systemic bias in tornado rating across the Southeast where more frequent occurrences and denser coverage of damage indicators/population exist. However, significant events were a small subset of the dataset and further study is likely needed.

The authors thank Israel Jirak (SPC) for helping with data interpretation and for providing a thorough review of this manuscript. The authors also thank Kelton Halbert and the Science Support Branch for their assistance with plotting and data preparation.

5. References

Bothwell, P. D., J. A. Hart, and R. L. Thompson, 2002: An integrated three-dimensional objective analysis scheme in use at the Storm Prediction Center. Preprints, 21st Conf. on Severe Local Storms, San Antonio, TX, Amer. Meteor. Soc., JP3.1. 1134 WEATHER AND FORECASTING VOLUME 27

Brooks, H. E., Doswell , C. A., III, & Kay, M. P. (2003). Climatological Estimates of Local Daily Tornado Probability for the United States, Weather and Forecasting, 18(4), 626-640.

DOC/NOAA/NESDIS/NCEI > National Centers for Environmental Information, NESDIS, NOAA, U.S. Department of Commerce. 'Storm Data and Unusual Weather Phenomena'. (2021)

Smith, B. T., Thompson, R. L., Grams, J. S., Broyles, C., & Brooks, H. E. (2012). Convective Modes for Significant Severe Thunderstorms in the Contiguous United States. Part I: Storm Classification and Climatology, *Weather and Forecasting*, *27*(5), 1114-1135.

Smith, B.T., Thompson, R.L., Speheger, D.A., Dean, A.R., Karstens, C.D., & Anderson-Frey, A.K. (2020). WSR-88D Tornado Intensity Estimates. Part I: Real-Time Probabilities of Peak Tornado Wind Speeds. *Weather and Forecasting, 35*, 2479-2492.

Smith, B. T., Thompson, R. L., Speheger, D. A., Dean, A. R., Karstens, C. D., & Anderson-Frey, A. K. (2020). WSR-88D Tornado Intensity Estimates. Part II: Real-Time Applications to Tornado Warning Time Scales, *Weather and Forecasting*, *35*(6), 2493-2506.

Schaumann, J. S., and R. W. Przybylinski, 2012: Operational application of 0–3 km bulk shear vectors in assessing quasi linear convective system mesovortex and tornado potential. 26th Conf. on Severe Local Storms, Nashville, TN, Amer. Meteor. Soc., 142, https://ams.confex.com/ams/26SLS/webprogram/Paper212008.html.

Thompson, R. L., Smith, B. T., Grams, J. S., Dean, A. R., & Broyles, C. (2012). Convective Modes for Significant Severe Thunderstorms in the Contiguous United States. Part II: Supercell and QLCS Tornado Environments, *Weather and Forecasting*, *27*(5), 1136-1154.

Trapp, R. J., and M. L. Weisman, 2003: Low-level mesovortices within squall lines and bow echoes. Part II: Their genesis and implications. Mon. Wea. Rev., 131, 2804–2823.

Trapp, R. J., Tessendorf, S. A., Godfrey, E. S., & Brooks, H. E. (2005). Tornadoes from Squall Lines and Bow Echoes. Part I: Climatological Distribution, Weather and Forecasting, 20(1), 23-34.

Wurman J, Kosiba K, White T, Robinson P. Supercell tornadoes are much stronger and wider than damage-based ratings indicate. Proc Natl Acad Sci USA. 2021 Apr 6;118(14):e2021535118. doi: 10.1073/pnas.2021535118. PMID: 33753558; PMCID: PMC8040662.

6. Image Appendix

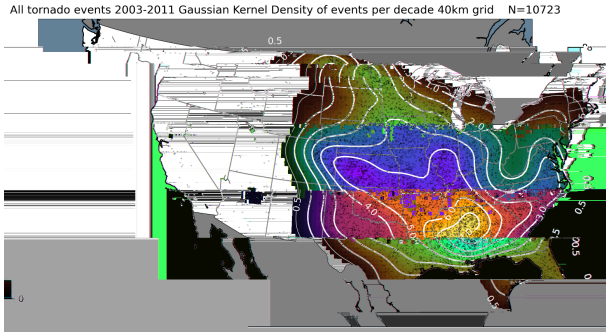


Fig 2a. Gaussian Kernel Density Estimate of tornado events per decade from all convective modes 2003-2011 Color fill begins at 0.4 events per decade.

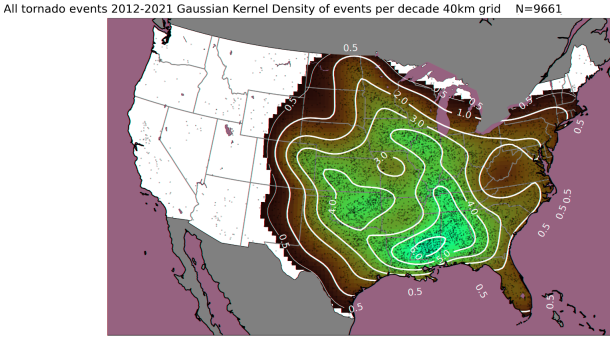


Fig 2. Same as Fig. 2a, except for 2012-2021.

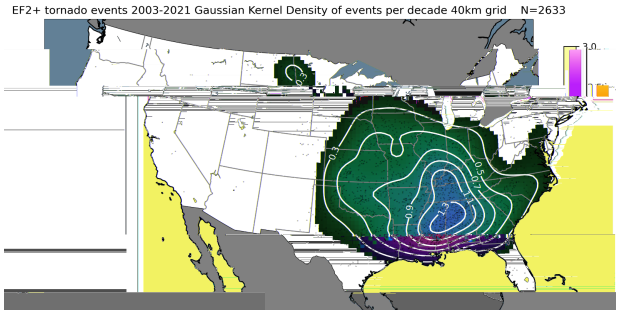


Fig 2c. Same as Fig. 2a, except for significant (F/EF-2+) tornado events per decade from all convective modes 2003-2021.

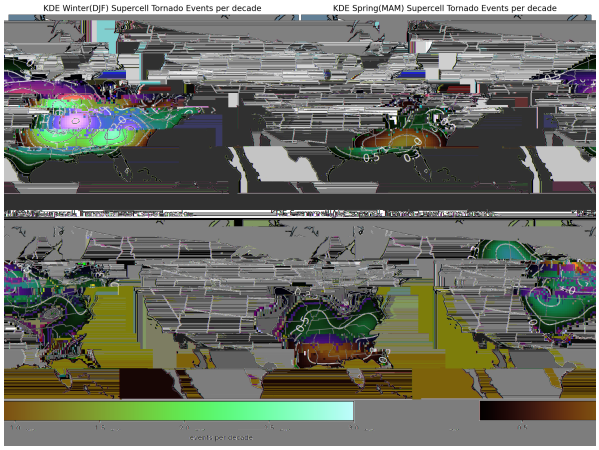


Fig 3b. Gaussian KDE of events per decade of supercell events by season. Top left to bottom right: Winter (DJF), Spring (MAM), Summer (JJA), and Fall (SON).

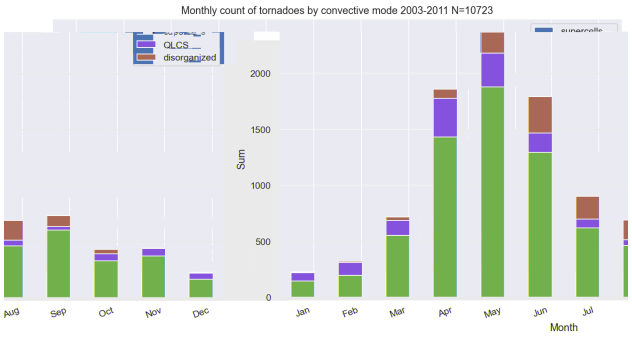


Fig 4a. Bar chart of monthly counts of tornado occurrence by convective mode 2003-2011.

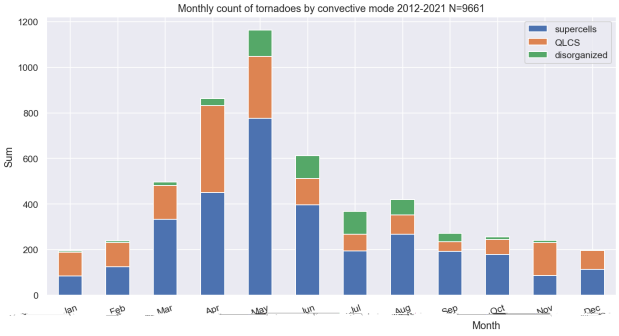


Fig 4b. Same as Fig. 4a, except for 2012-2021.

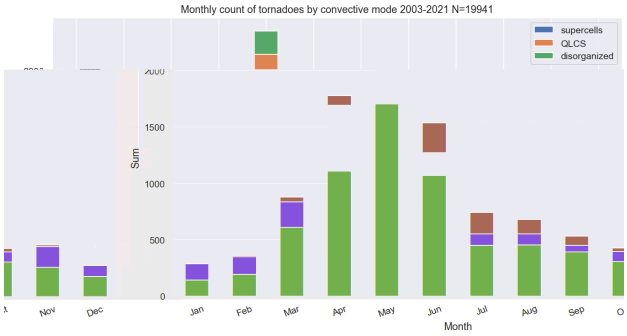


Fig 4c. Same as Fig. 4a, except for 2003-2021