P2.19 **AN ANALYSIS OF CLUSTERED TORNADO EVENTS**

Andrew R. Dean*
NOAA/NWS/Storm Prediction Center, Norman, OK

1. INTRODUCTION

While any tornado event is dangerous, days with widespread tornado activity over a large area are particularly threatening to life and property. In this study, the subject of tornado “clusters” is investigated. For the purposes of this study, clusters are defined in the context of SPC probabilistic convective outlooks.

Once identified, these tornado clusters are analyzed, both in terms of a long-term climatology (1950-2009) using the Storm Data tornado database, and in terms of a shorter period (2003-2009). The latter corresponds to the time period of the SPC storm environment database, which includes an hourly archive of objectively analyzed convective parameters such as CAPE, bulk shear, storm-relative helicity, LCL height, and other fields. Spatial and temporal trends in the long-term record of these tornado clusters will be discussed, along with an analysis of the convective environments and storm modes associated with the more recent events.

Section 2 describes the clustering methodology. Section 3 looks at some aspects of the climatology of these tornado clusters over the period 1950-2009, and section 4 examines characteristics of the storm environment and convective mode of these clusters for the 2003-2009 period. Section 5 summarizes this preliminary work and presents future avenues of investigation into tornado clusters and their relation to SPC forecast products.

2. METHODOLOGY

This study focuses on the geographic clustering of tornado reports documented in *Storm Data* for the period 1950-2009. A “cluster” is defined in the context of SPC's convective outlooks, which forecast the probability of a severe weather event (all combined severe events for Day 3 and Day 2 outlooks, and stratified into severe hail, severe wind, and tornado events for Day 1) occurring within 40 km (25 mi) of a point. The focus here will strictly be on tornado events and forecasts.

SPC began issuing probabilistic outlooks in 1999, with probability thresholds for tornado forecasts defined at 2%, 5%, 15%, 25%, and 35%. In 2006, the thresholds were changed to 2%, 5%, 10%, 15%, 30%, 45%, and 60%. There is also a separate forecast category for significant tornadoes (EF2-EF5) that consists of a single threshold indicating a 10% or greater threat of these high-end events. In the remainder of this paper, probabilistic tornado outlooks will be abbreviated as “TOR” (i.e. 30% TOR area) and the strong tornado category will be abbreviated as “SIGTOR” (i.e. 10% SIGTOR area).

SPC has been issuing categorical outlook products since the 1950s. The category names have changed over the years, but over the last 30 years the categories have been defined as “Slight”, “Moderate”, and “High” risks (hereafter abbreviated as SR, MR, and HR, respectively). When the probabilistic outlooks were initiated, the categorical risks were statistically related to various probability thresholds (see http://www.spc.noaa.gov/misc/SPC_probotlk_info.html) . For example, a 30% TOR area corresponds to a categorical HR (prior to 2006, a 25% TOR area resulted in a HR). Strictly speaking, a 10% SIGTOR area is not required for a HR, but in practice, one is in place over all or most of a tornado-based HR area.

The motivation for this study is to examine historical tornado data and identify days where the tornado coverage over some sufficiently large area met the criteria for a HR outlook. It should be noted at the outset that SPC's forecasts should be as considered true probabilities, rather than as deterministic forecasts of areal coverage. Thus, it should not be implied that every event meeting the HR criteria in terms of areal coverage should have been covered by a HR forecast. However, all such events are of great importance to SPC's forecast mission and they should be examined in greater detail if the forecast did not adequately describe the observed threat.

The process for clustering the tornado reports is as follows:

1. Tornado reports are binned into “convective days”, defined as the period 12 UTC -12 UTC, which corresponds with time period covered by the initial Day 1 forecast.
2. A 40 km buffer is drawn around each tornado path (defined by a start and end point), which corresponds to the drawing of forecast probabilities “within 40 km of a point”. A 100-point polygon approximating a circle of radius 40 km is used. This approximates the area “affected” by tornadoes for the purposes of SPC's convective outlooks.
3. A second buffer of 80 km radius is drawn around each tornado path. This defines an area in which the tornado coverage within 40 km of a point will be at least 25%, since twice as large of a radius corresponds to four times as large of an area. Depending on the amount of

intersection between the 40 km areas, the actual coverage may be higher than 25%. Even though the high risk criterion is currently 30%, an area of 25% or greater coverage is used as a proxy for the purposes of this study.

4. The union of the 80 km buffer polygons is taken and all of the resulting discrete polygons are considered to be tornado clusters.
5. This process was used for all tornado reports, only F1 or greater reports, and only F2 or greater reports. For the F2+ reports, an outer buffer of 125 km (instead of 80 km) was used in order to examine clusters with 10% or greater coverage.

The process described above is illustrated in Fig. 1 for the 3 May 1999 case in Oklahoma and Kansas.

Over the last 10 year period, the median size of both 25%/30% TOR areas and 10% SIGTOR areas is near 100 000 km^2. This value was used a threshold and any tornado cluster smaller than this area was not considered for the purposes of this study. Thus, only the larger, more widespread tornado clusters that pose a greater societal risk are examined in the next section.

3. CLIMATOLOGY OF TORNADO CLUSTERS

3.1 Annual Trends

Any investigation into tornado climatology is complicated by the non-stationarity of the tornado record, which has been well documented in many previous studies (Brooks et al. 2003, Verbout et al. 2006, Doswell 2007). Fig. 2 shows the annual number of tornadoes for the period 1950-2009. While there has been a notable increase noted in the total number of tornadoes since the late 1980s, the number of F1-F5 (hereafter F1+) tornadoes has generally been stable, indicating that the increase is mainly due to the inflation in F0 tornado reports. When tornado days (convective days with at least one tornado report) are considered, as shown in Fig. 3, the trend is somewhat different, with overall tornado days relatively stable over the last 40 yrs of the period while a decrease is noted in the number of F1+ days. For F2-F5 (hereafter F2+) reports, a decrease is noted in both the number of reports and report days. The change in reporting trends in the middle 1970s is largely related to the different processes used to rate tornado damage, where prior to ~1975 they were based on historical accounts, rather than real time information.

Fig. 4 shows the annual number of days where a tornado cluster meeting the coverage (25% for F0+ and F1+, 10% for F2+) and area criteria was observed. An increase in the number of F0+ clusters and a decrease in the number of F2+ clusters is noted, while the number of F1+ clusters has been relatively more stable, corresponding to the trends in the overall number of reports. There is some indication of a shorter-term increase in F1+ and F2+ clusters since the mid-1990s, but establishing the significance of that trend is beyond the scope of this study.

Given the noted increase in the total number of tornado reports and corresponding number of F0+ clusters, for the purposes of examining the long-term climatology the primary focus here will be on clusters of F1-F5 reports. The frequency of F1+ clusters (generally 3-10 days per year) also corresponds well to the frequency of SPC high risk outlooks (not shown). The clustering of F2+ reports in areas of 10% coverage or greater will also be examined and the slight decrease in F2+ reports and clusters over the second half of the 1950-2009 period should be acknowledged when interpreting these results.

3.2 Regional and Seasonal Trends

For the period 1950-2009, a rather large area east of the Rockies and west of the Appalachians was prone to F1+ tornado clusters (for $\geq$ 25% coverage), as shown in Fig. 5. While this area includes the areas of the Plains commonly described as “Tornado Alley” (Brooks et al. 2003), it extends over a much broader area encompassing parts of the Southeast, Midwest, and Ohio Valley. Much of this area has seen 20-30 days in F1+ tornado clusters over the past 60 years, corresponding to an average event recurrence of once every 2-3 years. A similar pattern is noted for coverage of F2+ clusters (for $\geq$ 10 % coverage - Fig. 6), though with a more well-defined maximum running from OK eastward into parts of the Southeast. No tornado clusters were observed west of the Rockies and only a small number occurred east of the Appalachians.

Figs. 7-10 show the seasonal breakdown of clustered F1+ tornado events over the period 1950-2009. The proverbial “Tornado Alley” area is better defined when looking at the APR-JUN events (Fig. 8), though parts of the Tennessee and Ohio Valleys are nearly as active as the Plains during this period. The summer period JUL-SEP (Fig. 9) has a much lower frequency of occurrence of F1+ clusters, with events near the Gulf and Atlantic coasts at least partially related to tropical cyclone tornadoes, while clustered tornadoes decrease across the rest of the CONUS. A well-defined cool season/early spring maximum (Figs. 7 and 10) is noted over the Southeast. Figs. 11-14 show corresponding seasonal plots for F2+ clusters, which show a similar overall pattern with a well defined peak in occurrence during the APR-JUN period over central Oklahoma.

4. STORM ENVIRONMENTS AND CONVECTIVE MODE

The SPC has developed a database of convective environment data associated with each severe report for the period 2003-present, based on an objective analysis of surface observations using the RUC analysis as a first guess (Bothwell et al. 2002, Dean et al. 2006). In addition, a very detailed and comprehensive database of convective mode for a subset of reports from 2003-2009, including all tornadoes over that period, has recently been created (Smith et al. 2010, Thompson et al. 2010). This section will briefly examine tornado clusters over the period 2003-2009 in the context of this environment and mode data.

4.1. Storm Environments in Tornado Clusters

Figs. 15-17 show the distributions (in box plot form) of 100 mb mean mixed-layer CAPE (ML CAPE), 0-6 km bulk wind shear (SHR6), and 0-1 km storm-relative helicity (SRH1, uses the Bunkers et al. 2000 method for assumed storm motion) for clustered versus non-clustered reports across each category (F0+, F1+, F2+). While the median ML CAPE value in each category is slightly higher for clustered reports, there is substantial overlap in the distributions. Meanwhile, a more substantial increase in SHR6 and SRH1 is noted for clustered reports in each category, with a greater difference between median values and an offset of approximately one quartile between the distributions. While further analysis is required to generate more robust quantitative conclusions, including statistical significance testing, these results suggest that wind shear is a more important environmental parameter than buoyancy in terms of discriminating between environments of clustered versus non-clustered events.

4.2. Convective Modes in Tornado Clusters

Numerous convective mode types were specified in the comprehensive mode analysis by Smith et al. (2010) . For simplicity, the results presented here in Fig. 18 will focus on three basic modes: supercell (both discrete and in a line), linear non-supercell, and disorganized taken from the Smith et al. (2010) database. The vast majority of tornado events in all categories (clustered and non-clustered) are associated with supercells. For F0+ and F1+ clusters, there is a notable decrease in disorganized events, a slight decrease in linear non-supercell events, and an increase in supercellular events for clustered reports versus non-clustered. These results indicate a general increase in storm organization for clustered events, which is expected.

Interestingly, for F2+ events, there is actually a slight increase in the fraction of non-supercell linear events and a slight decrease in the fraction of supercellular events for clustered versus non-clustered tornado reports. Further investigation is required to determine why this might be the case. It is possible that it is a function of the lower coverage threshold (10% vs. 25%) for the F2+ clusters, which allows for more isolated events to be included in the sample.

Overall, around 10% of clustered events are non-supercellular. These events are also a likely target of future investigation, since non-supercell tornadoes typically pose a difficult challenge to forecast operations.

5. DISCUSSION AND SUMMARY

A clustering technique based on SPC's probabilistic outlook criteria was applied to tornado data over the period 1950-2009, with clusters comparable in size to a typical SPC “High Risk” area kept for analysis in this study. Tornado clusters consisting of any tornado report (F0+), F1+ reports, and F2+ reports were examined, with the latter having a lower coverage threshold (10% vs. 25% for F0+ and F1+) to correspond with SPC's 10% SIGTOR outlooks.

Annual trends in tornado cluster days during the 60 year data period generally correspond with the trends in tornado reports, with an increase in F0+ cluster days, a relatively steady occurrence of F1+ cluster days, and a slight decrease with time in F2+ cluster days. The frequency of tornado clusters was highest overall during the APR-JUN time frame, with an elevated frequency of occurrence over parts of the Southeast throughout the winter and early spring. A relatively high frequency of tornado clusters extended well east of the proverbial “Tornado Alley” region of the Plains.

Convective environment and storm mode for clustered events for the period 2003-2009 were briefly examined. While a slight increase in ML CAPE was noted for clustered versus non-clustered events, a more substantial increase was noted in 0-6 km bulk shear and 0-1 km SRH. Supercells were the dominant mode for all categories of tornado reports, with the fraction of supercell mode increasing further for clustered F0+ and F1+ reports. However, a slight increase in the fraction of linear non-supercell reports was found for clustered F2+ reports.

The results presented here a preliminary and there are many avenues of future investigation. Tornado clustering data have many potentially useful forecast verification applications. A few cluster-based verification results are presented in Davis et al. (2010). Report clusters could be used in an object-based verification scheme, which is potentially very useful but currently difficult to implement in the context of severe convection. Clustering could also be used to more closely examine of the tornado climatology, since it would allow investigation into the spatial relationship of historic reports, rather than simpler analyses using the number of reports or number of tornado days. Also, as SPC's storm environment and convective mode databases continue to evolve, forecast guidance focusing on the potential for clustered tornado events could be developed. Finally, a focused investigation into more isolated events needs to be undertaken as a companion study to this one, so that the full range of possible event coverage will be explored.

Acknowledgements: I would like to thank Steve Weiss for his thoughtful review of this paper. The statements, findings, conclusions, and recommendations are those of the author and do not necessarily reflect the views of NOAA or the Department of Commerce.

6. 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. Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc.

Brooks, H. E., C. A. Doswell III., and M. P. Kay, 2003: Climatological estimates of local daily tornado probability for the United States. *Wea. Forecasting*, **18**, 626–640.

Davis, J. M., A. R. Dean, and J. L. Guyer, 2010: A detailed analysis of SPC "high risk" outlooks, 2003-2009. Preprints, *25th Conf. Severe Local Storms*, Denver, CO, Amer. Meteor. Soc.

Dean, A.R., R.S. Schneider, and J.T. Schaefer, 2006: Development of a comprehensive severe weather forecast verification system at the Storm Prediction Center. Preprints, *23rd Conf. Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc.

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

Smith, B. T., R. L. Thompson, J. S. Grams, and C. Broyles, 2010: Convective modes associated with significant severe thunderstorms in the contiguous United States. Preprints, *25th Conf. Severe Local Storms*, Denver, CO, Amer. Meteor. Soc.

Thompson, R. L., B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2010: Climatology of near-storm environments with convective modes for significant severe thunderstorms in the contiguous United States. Preprints, *25th Conf. Severe Local Storms*, Denver, CO, Amer. Meteor. Soc.

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

7. FIGURES (see below)

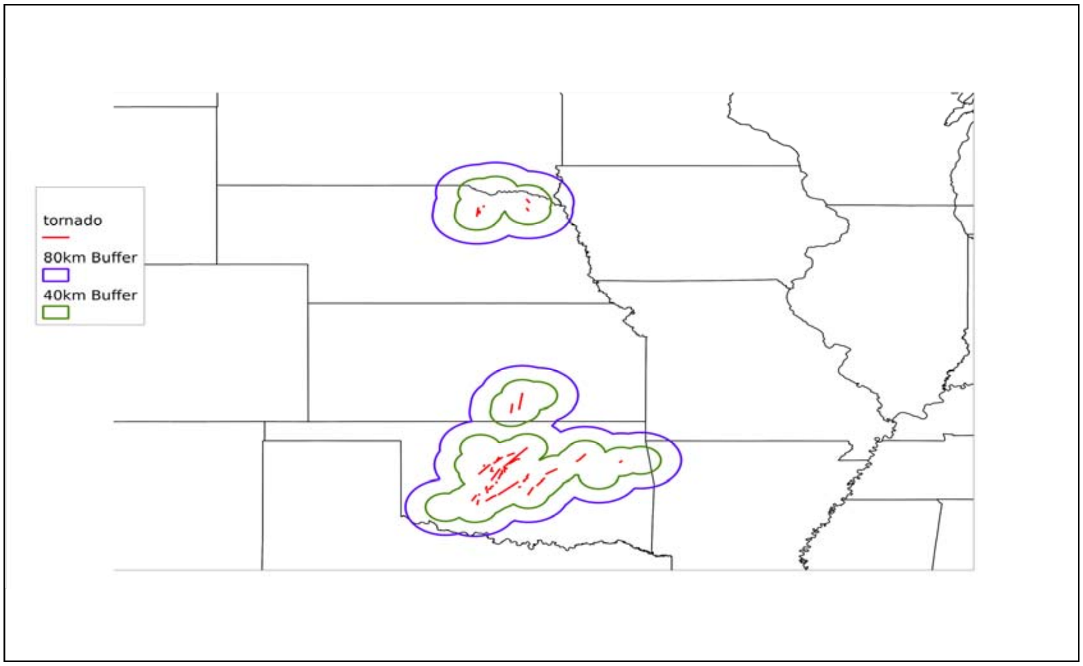

Fig. 1. Illustration of clustering technique for 3 May 1999 case. The red lines indicate tornado tracks, the area outlined in green represents the area within 40 km of the tornado tracks, and the area outlined in blue represents the area within 80 km of the tornado tracks. The blue area covering much of Oklahoma and part of Kansas is around 150 000 km^2 in area and is therefore large enough to be included in this study, while the area over Nebraska and extreme southeast South Dakota is too small to be considered.

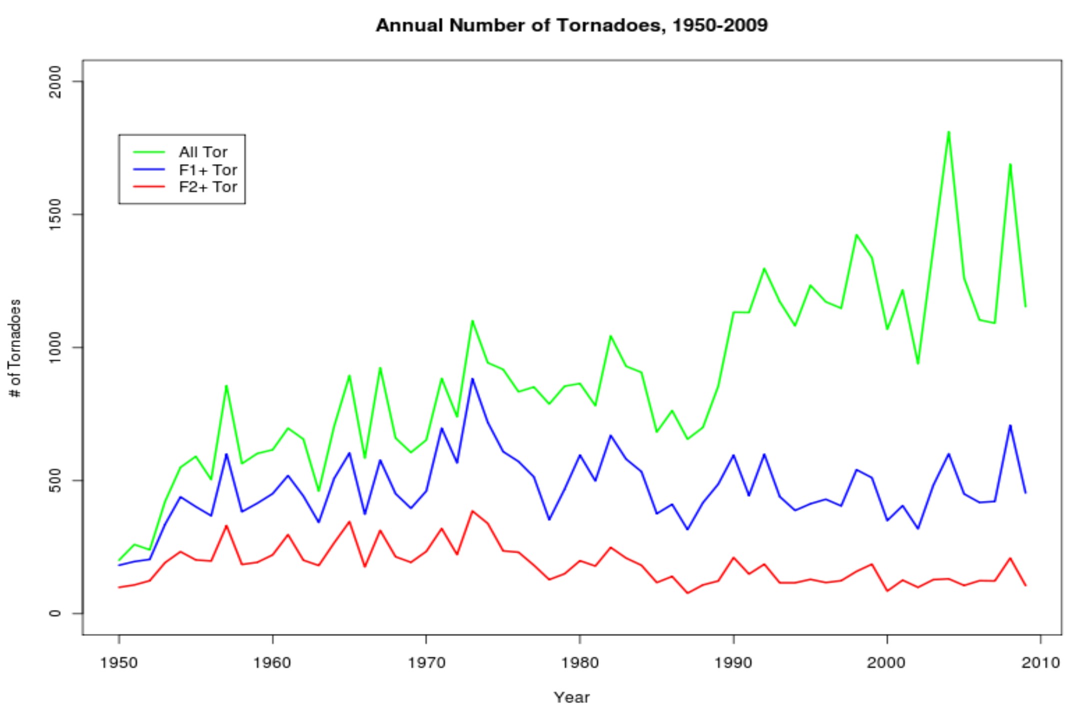

Fig. 2. Annual number of tornadoes, 1950-2009. All tornadoes are in green, F1+ tornadoes in blue, F2+ tornadoes in red.

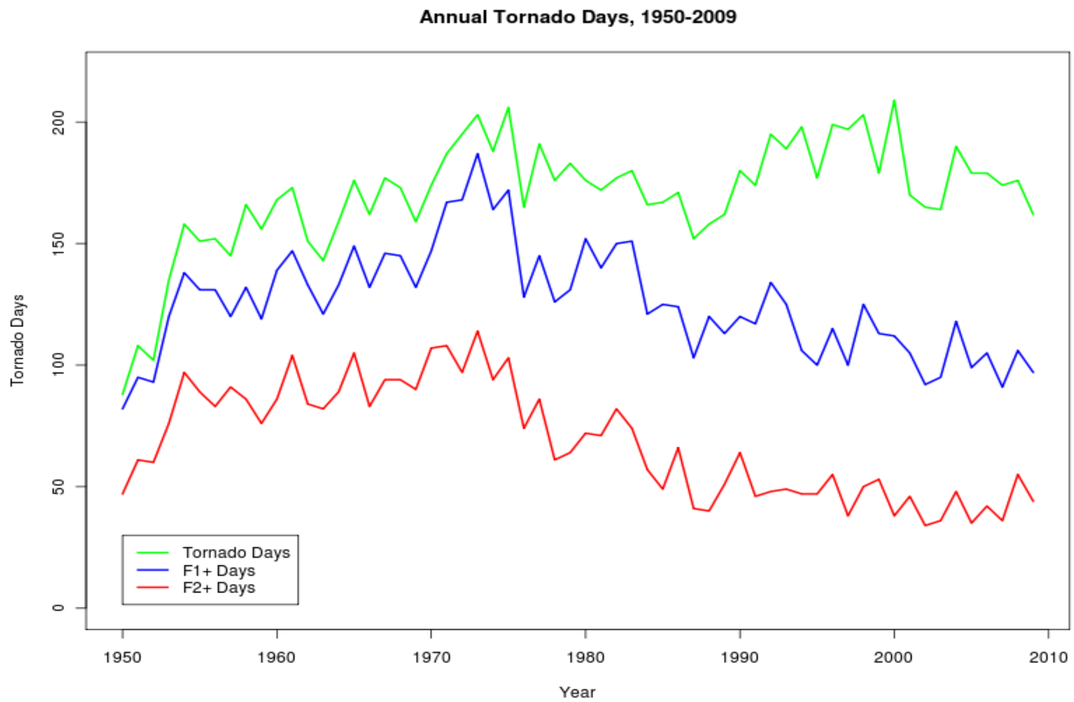

Fig. 3. Annual number of tornado days, 1950-2009. All tornado days in green, F1+ tornado days in blue, F2+ tornado days in red.

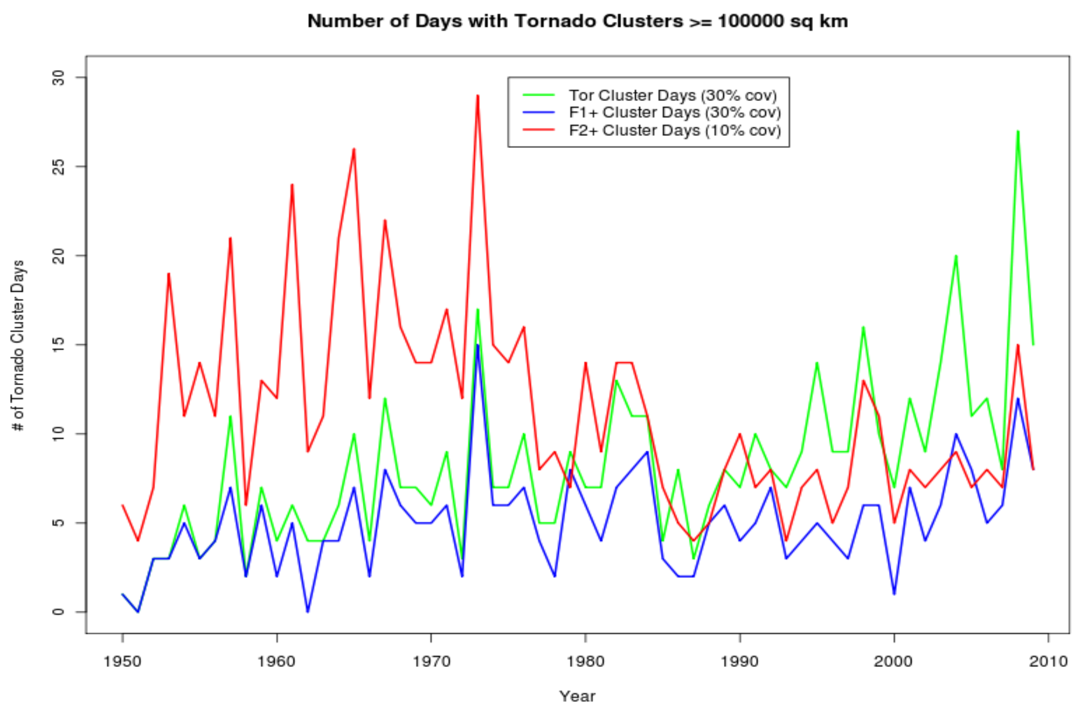

Fig. 4. Annual number of tornado cluster days, 1950-2009. Clusters of all tornado reports (F0+) in green, F1+ clusters in blue, F2+ clusters in red. F0+ and F1+ clusters have 25% or greater coverage of reports within 40 km of a point, while F2+ clusters have 10% or greater coverage of reports. Only clusters with area $\geq$ 100 000 km^2 were counted.

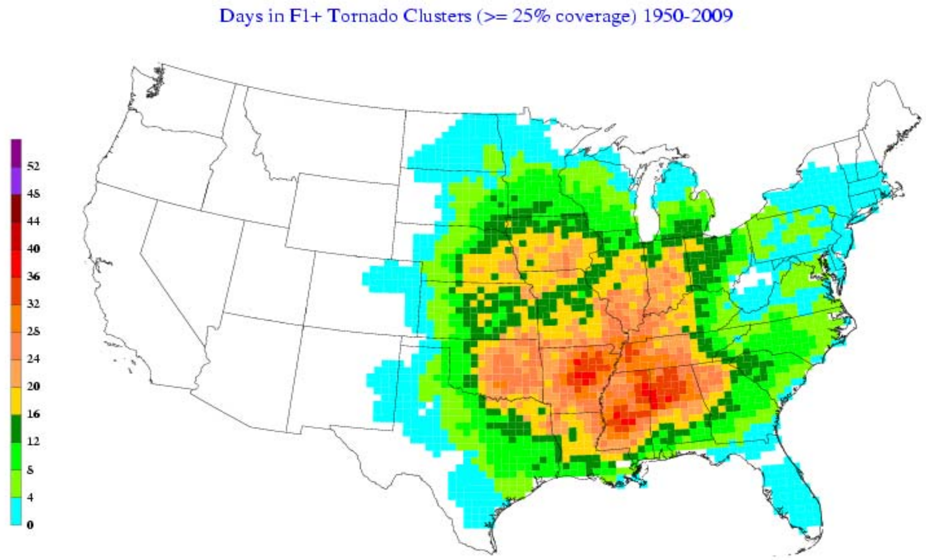

Fig. 5. Total number of convective days in F1+ tornado clusters, 1950-2009. Cluster polygons were mapped to a 40 km grid and the number of clusters that a given grid point was contained in was tabulated.

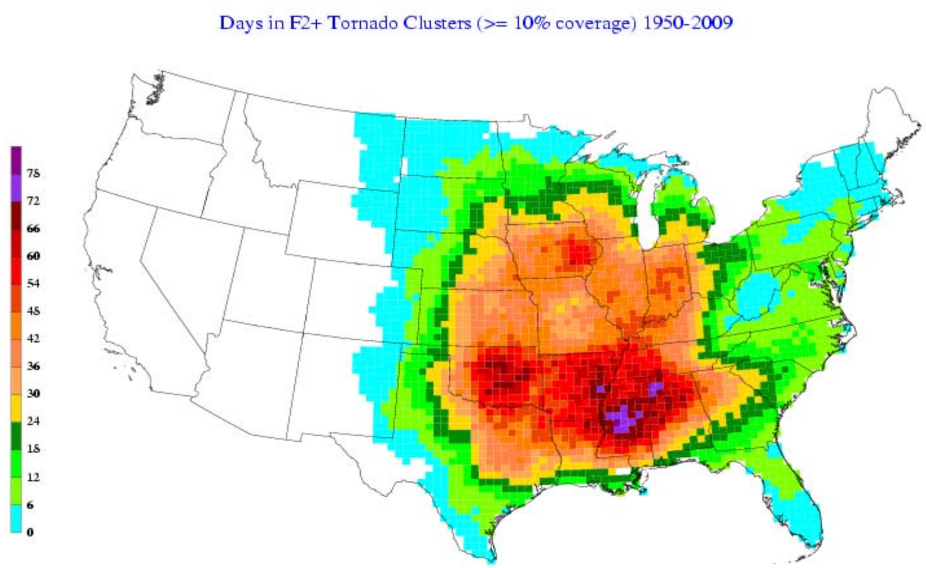

Fig. 6. As in Fig. 5, for F2+ tornado clusters. Note that the scale is different to account for the larger number of F2+ cluster events with ≥ 10% coverage.

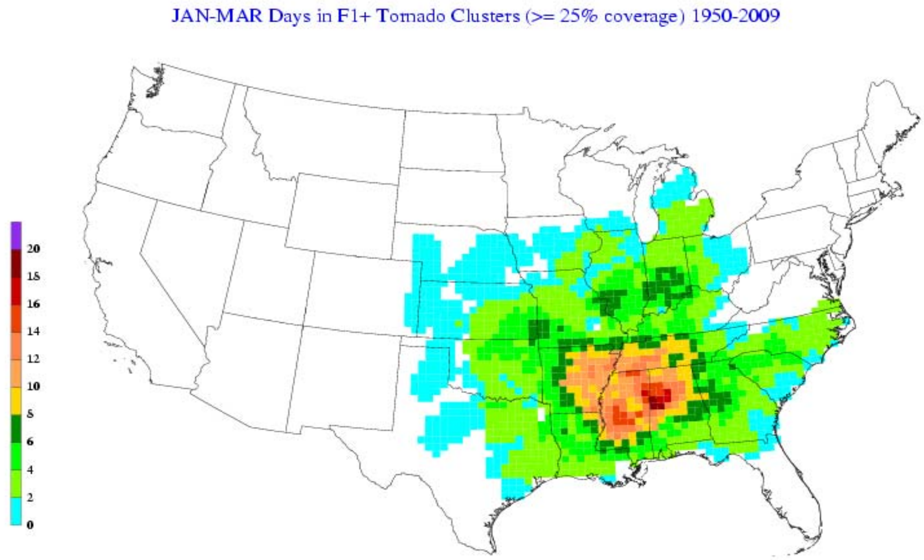

Fig. 7. F1+ tornado cluster days for the months of JAN-MAR, 1950-2009.

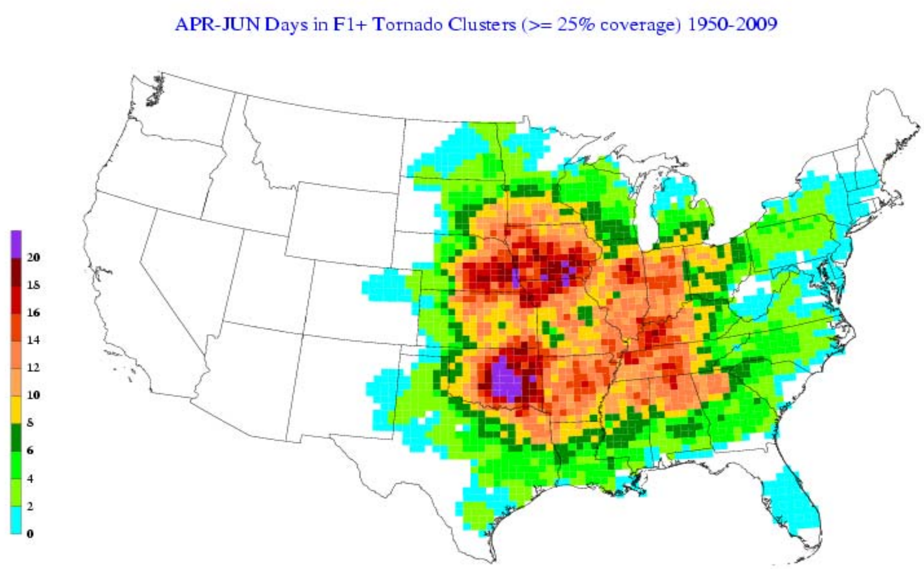

Fig. 8. F1+ tornado cluster days for the months of APR-JUN, 1950-2009.

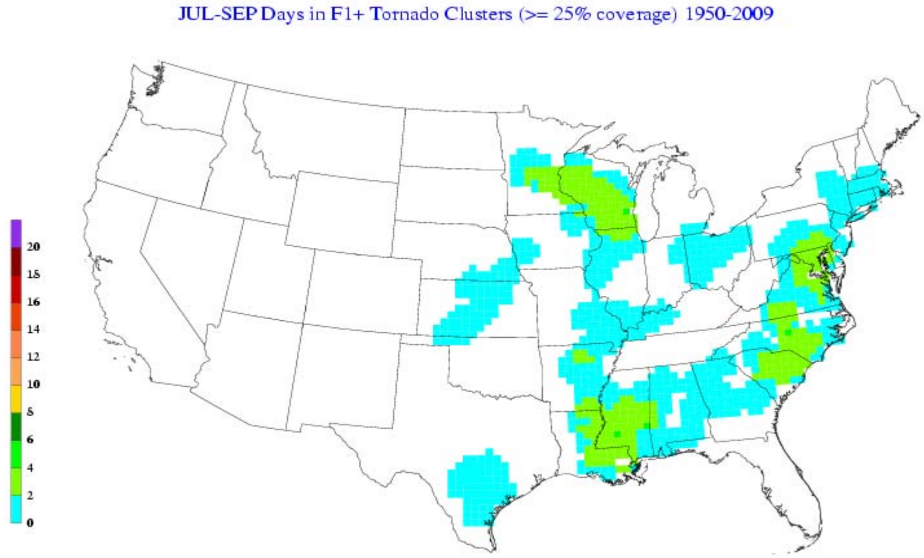

Fig. 9. F1+ tornado cluster days for the months of JUL-SEP, 1950-2009.

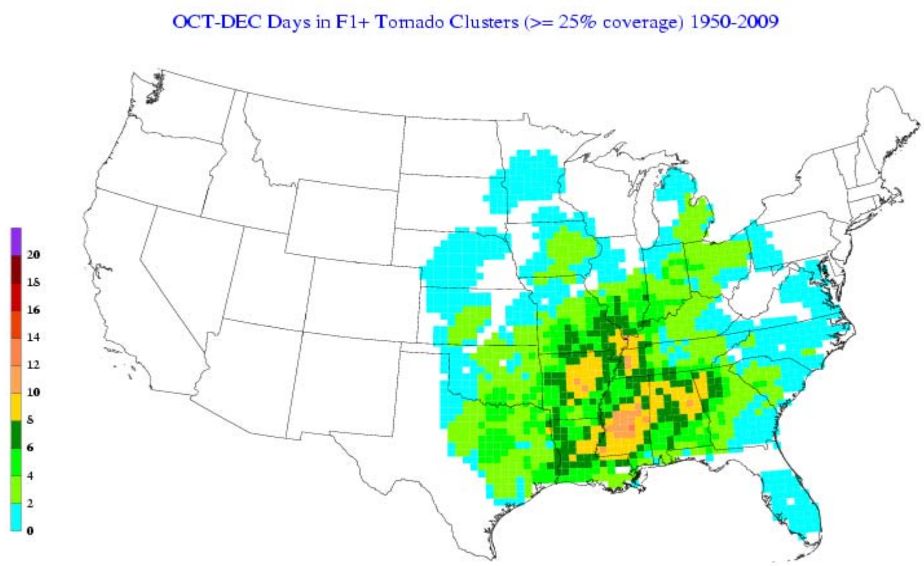

Fig. 10. F1+ tornado cluster days for the months of OCT-DEC, 1950-2009.

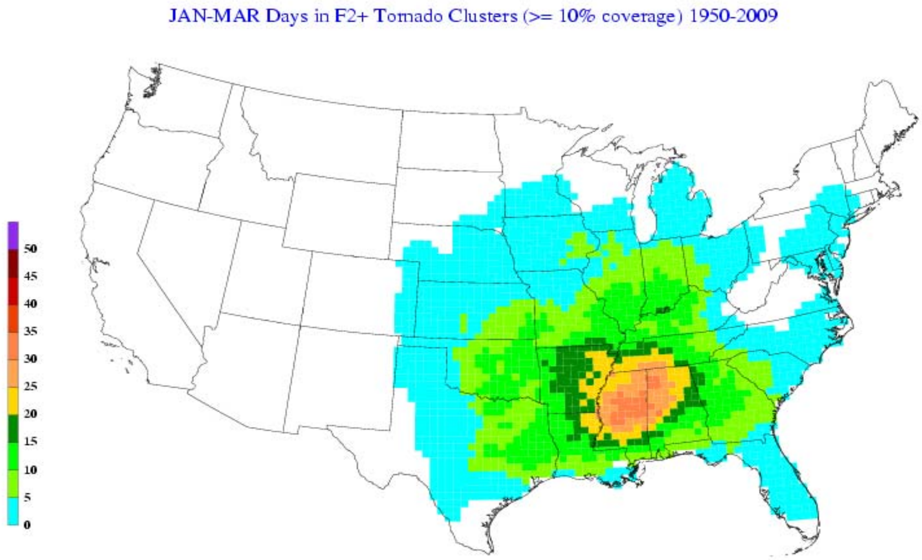

Fig. 11. F2+ tornado cluster days for the months of JAN-MAR, 1950-2009.

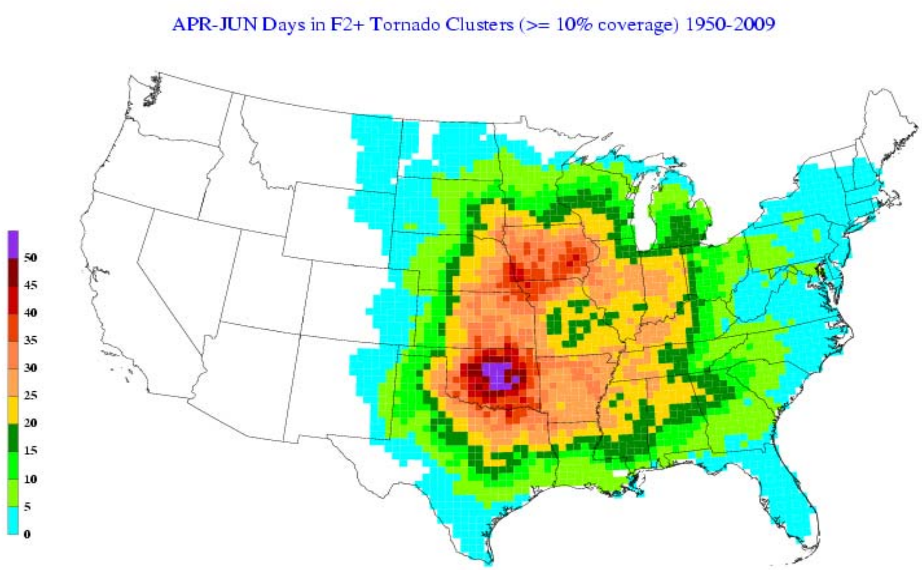

Fig. 12. F2+ tornado cluster days for the months of APR-JUN, 1950-2009.

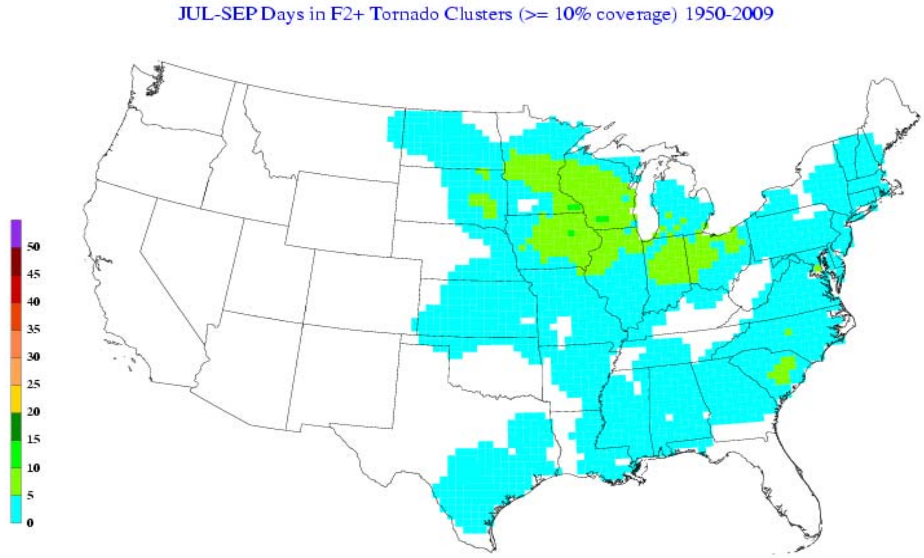

Fig. 13. F2+ tornado cluster days for the months of JUL-SEP, 1950-2009.

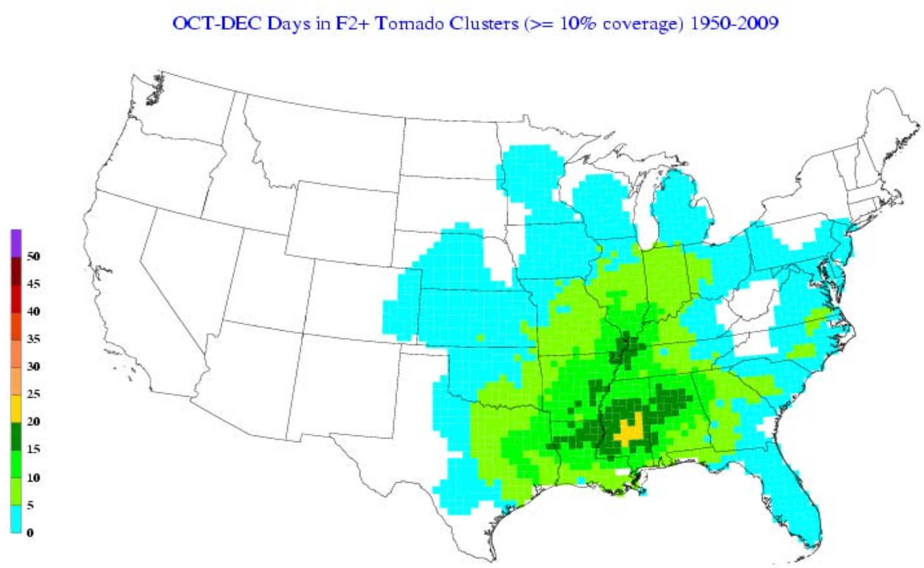

Fig. 14. F2+ tornado cluster days for the months of OCT-DEC, 1950-2009.

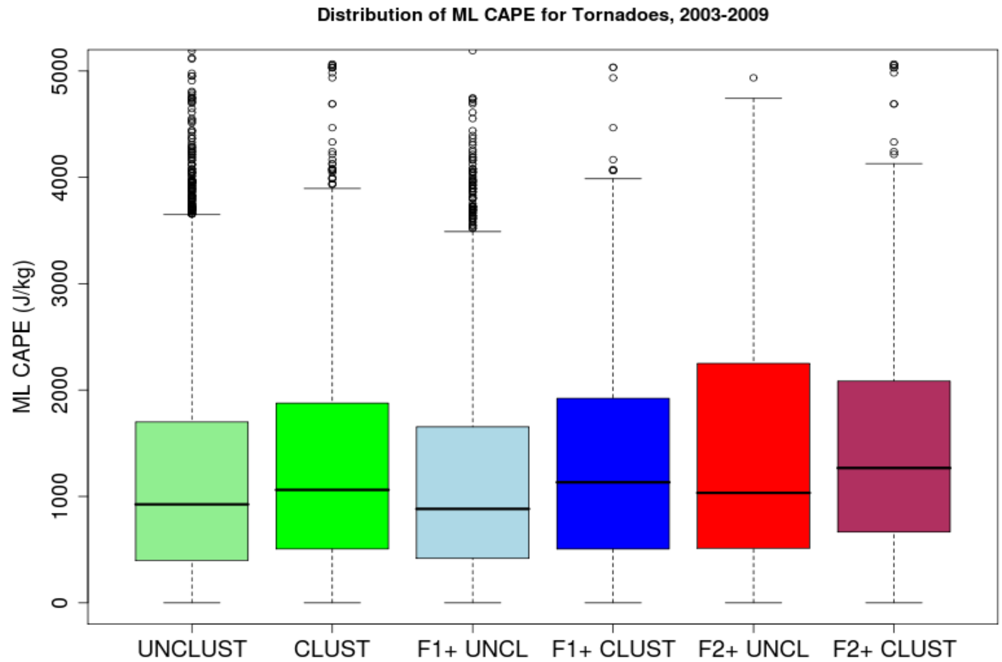

Fig. 15. Box plots showing distribution of ML CAPE for unclustered tornado reports (UNCLUST), clustered tornado reports (CLUST), unclustered F1+ tornado reports (F1+ UNCL), clustered F1+ tornado reports (F1+ CLUST), unclustered F2+ tornado reports (F2+ UNCL), and clustered F2+ tornado reports (F2+ CLUST). The box defines the area between the 25th and 75th percentiles, with the "whiskers" extending to 1.5 times the interquartile range (or to the most extreme data point, if it is within 1.5 times the IQR). Circles indicate outlier data points beyond the whisker extent.

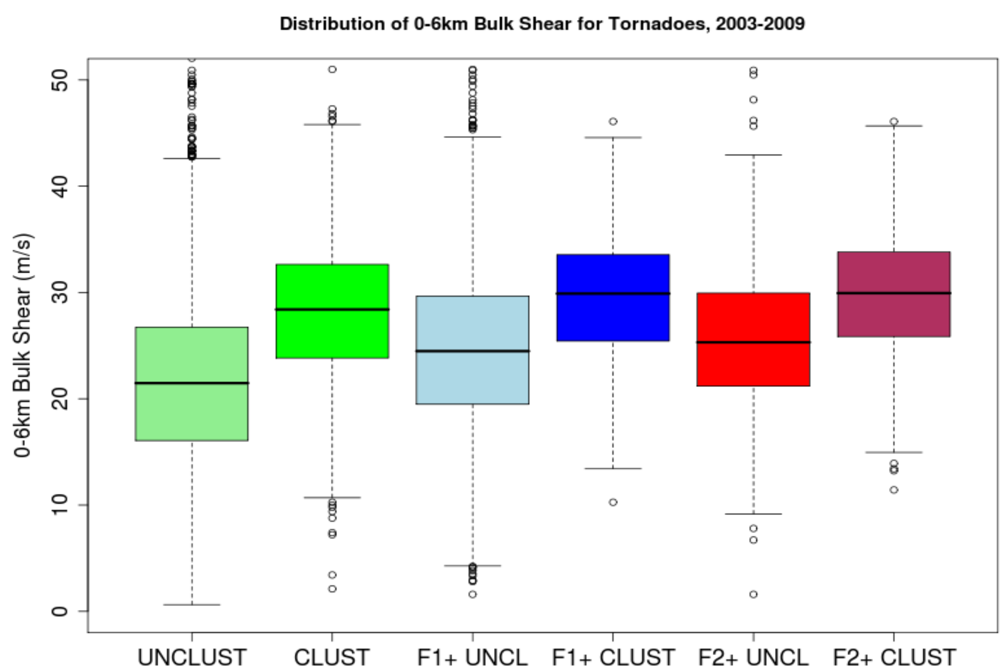

Fig. 16. As in Fig. 15, for 0-6 km bulk shear.

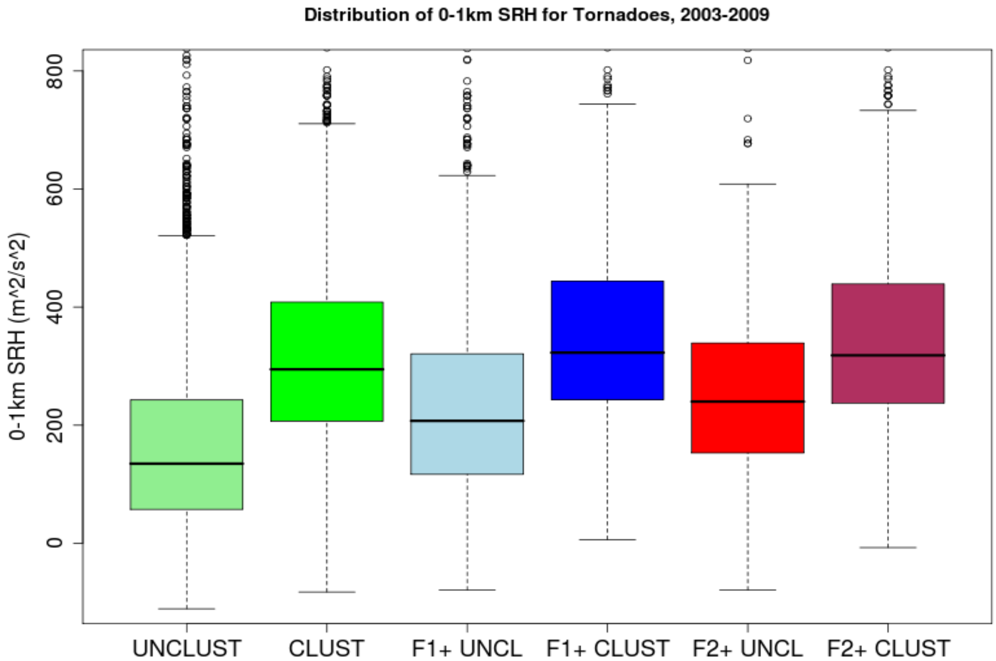

Fig. 17. As in Fig. 16, for 0-1 km storm-relative helicity.

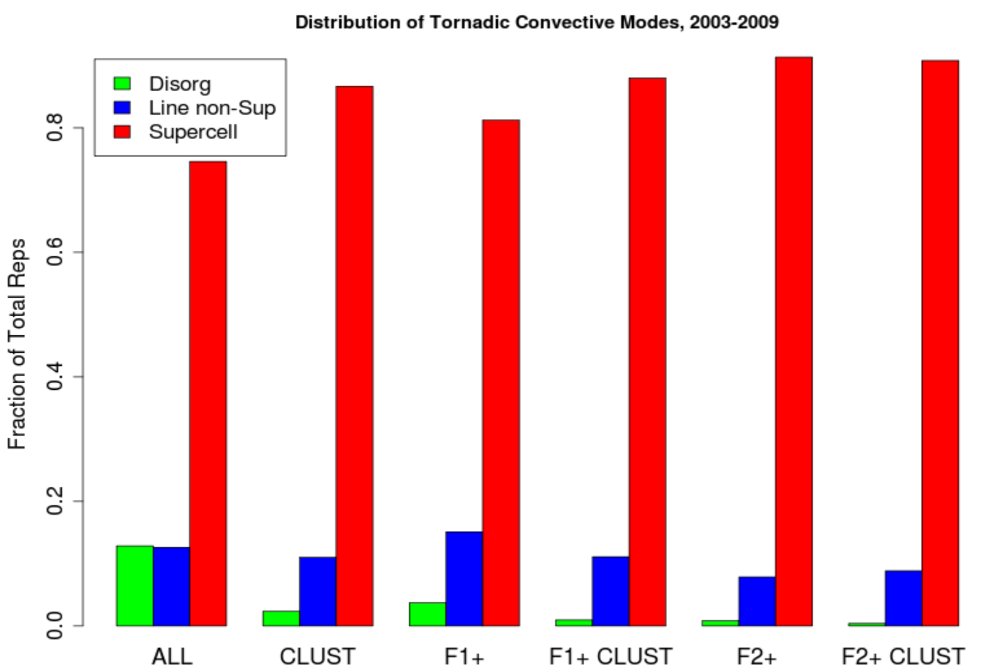

Fig. 18. Fraction of tornado reports by convective mode for each cluster category. ALL = all tornado reports, CLUST = clustered tornado reports, F1+ = all F1+ tornado reports, F1+ CLUST = clustered F1+ tornado reports, F2+ = all F2+ tornado reports, F2+ CLUST = clustered F2+ tornado reports. Green indicates disorganized mode, blue indicates linear non-supercell storm mode, and red indicates supercellular storm mode.