Examining Subdaily Tornado Warning Performance and Associated Environmental Characteristics

Makenzie J. Krocak,[a,b,c] Matthew D. Flournoy,[b] and Harold E. Brooks[d]

[a] *Center for Risk and Crisis Management, University of Oklahoma, Norman, Oklahoma*
[b] *Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma*
[c] *NOAA/NWS/Storm Prediction Center, Norman, Oklahoma*
[d] *NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*

ABSTRACT: Increasing tornado warning skill in terms of the probability of detection and false-alarm ratio remains an important operational goal. Although many studies have examined tornado warning performance in a broad sense, less focus has been placed on warning performance within subdaily convective events. In this study, we use the NWS tornado verification database to examine tornado warning performance by order-of-tornado within each convective day. We combine this database with tornado reports to relate warning performance to environmental characteristics. On convective days with multiple tornadoes, the first tornado is warned significantly less often than the middle and last tornadoes. More favorable kinematic environmental characteristics, like increasing 0–1-km shear and storm-relative helicity, are associated with better warning performance related to the first tornado of the convective day. Thermodynamic and composite parameters are less correlated with warning performance. During tornadic events, over one-half of false alarms occur after the last tornado of the day decays, and false alarms are 2 times as likely to be issued during this time as before the first tornado forms. These results indicate that forecasters may be better "primed" (or more prepared) to issue warnings on middle and last tornadoes of the day and must overcome a higher threshold to warn on the first tornado of the day. To overcome this challenge, using kinematic environmental characteristics and intermediate products on the watch-to-warning scale may help.

SIGNIFICANCE STATEMENT: This study examines the performance of tornado warnings during past severe weather events in an effort to better understand forecasting strengths and weaknesses. On days with multiple tornadoes, we find that the first tornado of the day is less likely to be warned and that, if it is warned, it has less lead time than the other tornadoes on the same day. Furthermore, there are some environmental factors (such as bulk wind shear) that influence the likelihood that the first tornado is warned. This study helps forecasters to understand which environmental traits may be more useful for better anticipating the first tornado of the day.

KEYWORDS: Forecast verification/skill; Forecasting; Operational forecasting

---

1. Introduction and background

Tornado warnings are one of the most well-known and well-studied products from the National Weather Service (NWS). Because of their time-sensitive nature and high cost associated with missed events, much work has gone into understanding tornado warning performance. In general, lead time has remained the same since the mid-1980s while probability of detection (POD) increased until around 2000, when it leveled out just under 70% until it began to decrease slightly in 2012 (Erickson and Brooks 2006; NOAA/NWS 2007; Brooks and Correia 2018).

Although there have been numerous studies that investigate overall tornado warning performance, many questions remain about how performance may vary by situational differences. Brotzge and Erickson (2009) stated that providing advanced warning on the first tornado of the day remains a difficult challenge. Using a dataset of tornado warnings from 2000 to 2004, Brotzge and Erickson (2010) found that the skill for warning the first tornado of the day is similar to the skill at warning isolated tornadoes. Similarly, another study found

that the first tornado of the day has lower lead time than other tornadoes on the same day (Bieringer and Ray 1996), although the dataset was more limited (<600 tornadoes). Andra et al. (2002) reached a similar conclusion in their overview of warning operations during the prolific 3 May 1999 tornado outbreak in central Oklahoma; the lead time for the first tornado of the first storm of the day was 4 mins (see their Fig. 5) whereas the median lead time for all tornadoes during the outbreak was around 23 mins. Multiple studies have found that the more tornadoes that occur on a given day, the higher the POD generally is (Brotzge and Erickson 2009; Anderson-Frey et al. 2018). These differences in performance are likely due to a combination of factors, from formal warning thresholds to the priming of warning forecasters to issue tornado warnings in more potent environments.

In addition to tornadoes being easier to warn on higher impact days, studies have also shown that certain environmental characteristics are related to warning performance. Guillot et al. (2008) found that forecast skill is related to storm type, with supercells and organized convective lines having higher POD and lower false-alarm ratio (FAR) than pulse or nonorganized storms. Using a dataset from 2003 to 2004, Brotzge et al. (2013) reached a similar conclusion in that supercell tornadoes were easier to warn for in terms of POD and

lead time than those from nondiscrete storms. Furthermore, Anderson-Frey et al. (2016) found that as MLCAPE and 0–6-km shear increased, so did warning performance (i.e., higher POD and lower FAR).

This work updates previous studies that have looked at warning performance by order within the day (e.g., Brotzge and Erickson 2009, 2010) by investigating 10 years of data during the polygon warning era. We also assess if and how warning performance varies by environmental characteristics with this same dataset. Our findings show that not all supercell tornadoes are created equal in terms of warning performance and that some environmental characteristics may be more useful than others in anticipating the first tornado of the day.

2. Methods

Tornadoes and associated warning data were downloaded from the NWS verification website (NOAA/NWS 2021). This dataset contains all tornadoes between 1 January 2008 and 31 December 2018. Then, storm mode and environmental data were obtained from the SPC report database (Smith et al. 2012; Thompson et al. 2012). This hand-compiled dataset includes 45 868 grid-hour severe events from all months across the contiguous United States from 2003 to 2019. These events are the maximum tornado rating or severe wind/hail magnitude per hour on a 40-km grid, and all tornado reports (enhanced Fujita scale EF0–5) are included during the entire time period.

The two datasets were merged by matching tornadoes based on location and timing information. To be considered a match, the reports must have occurred within 5 min of each other and within 0.03° latitude and longitude. Of the nearly 2700 tornadoes that were classified as coming from right-moving supercells in the SPC database, this method matched 2687 of them with reports from the NWS verification database. Of these 2687 reports, 1477 were rated EF0, 720 were rated EF1, 302 were rated EF2, 141 were rated EF3, 42 were rated EF4, and 5 were rated EF5.

Once the tornadoes, associated warning statuses, and environmental characteristics were compiled, report times were converted to convective outlook days (from 1200 to 1159 UTC). This conversion was done to reduce the chance that an event was split into two different days (based on local date), which can be problematic in the Southeast United States, where a higher proportion of events occur overnight (Krocak and Brooks 2018). As such, we define the first tornado of the day as the first report that occurred after 1200 UTC in each NWS county warning area (CWA). We chose to stratify by CWA (instead of just assessing the first tornado that occurred anywhere in the country on each day) because different forecasters issue warnings for each CWA, resulting in potentially more than one "first tornado warning decision of the day" for any event, depending on how widespread the event was. However, it is worth noting that NWS offices in adjacent CWAs often communicate with each other. Subsequently, "only" tornadoes are defined as single reports on a convective day and in a CWA, "last" events are defined as the latest report to occur on a convective day and in a CWA, and "middle" events are defined as reports that are not the first, last, or only reports on a convective day and in a CWA. Note that

each report was assigned to the CWA that it occurred within, such that storms that crossed CWA boundaries were assigned to multiple CWAs.

Additionally, tornadoes were considered "warned" only if the first segment was warned in advance. In other words, the initial lead time must be greater than zero minutes. In this work, we assess the proportion of warned events for all tornadoes based on the order in which they occurred (first tornado, only tornado, middle tornadoes, and last tornadoes of the day), median lead time based on this order, and the proportion of first tornadoes that were warned based on different environmental characteristics.

In addition to warned/unwarned tornadoes and lead time statistics, we also examine how FAR varies during convective events. A tornado warning is considered a false alarm if it did not contain any associated verifying tornado events. Based on this and our definition of "warned" tornadoes, this means that we do not analyze tornado warnings that contain negative lead time. We then separate these false alarms into three categories: those that occurred before the first tornado of the day, those that occurred after the last tornado of the day, and those that occurred between the first and last tornadoes of the day.

3. Results

a. Warning performance based on order within the day

Our results indicate that the quality of the warning is at least partially dependent on the order of the warning within the event timeline. Within convective days and NWS CWA boundaries, the first tornado and the only tornado are least likely to be warned (63% and 49% of the time, respectively). This is contrasted by the last and middle tornadoes, which are warned 80% and 87% of the time, respectively (Fig. 1).

We also examined the nature of false alarms during convective events containing at least one tornado. Out of these 4774 false alarms, 1227 (25.7%) occurred before the first tornado of the day, 856 (17.9%) occurred between the first and last tornado, and 2691 (56.4%) occurred after the last tornado. During tornadic events, false alarms are least likely to occur in between tornadoes and most likely to occur after tornado production has ceased. False alarms issued after the last tornado of the day account for more than one-half of all false alarms issued during tornadic events. We further discuss these findings and their implications in the last section.

The lead time for tornadoes also varies based on the order of the event. Note that we calculate lead time only from tornadoes that were warned in advance (hence, we eliminate all instances of zero lead time data points). First tornadoes have a median lead time of 14 min, which is similar to only tornadoes (13 min; Fig. 2). Middle tornadoes and last tornadoes have significantly higher lead times. Last tornadoes gain 5–6 min over first and only tornadoes, with a median lead time of 19 min. Middle tornadoes have an even longer median lead time of 22 min (Fig. 2). This results in an almost 70% increase in lead time between the only tornado of the day and a tornado that occurred in the middle of an event.
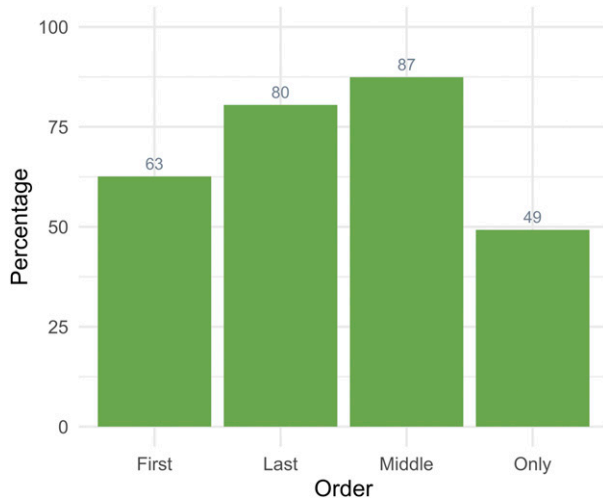
FIG. 1. The percentage of warned tornadoes as based on the order of the day in which they occurred.

FIG. 2. Median lead time of tornadoes that were warned in advance as based on the order of the day in which they occurred.

b. Environmental characteristics associated with first-tornado warning performance

In addition to understanding how the order of the tornado impacts the warning quality, we investigate how warning performance on the *first* tornado of the day differs in different environments. For example, are the first tornadoes of the day in "higher-end" environments more likely to be warned? To do this, we assessed the warned proportion of first tornadoes based on numerous environmental factors like different layers of shear and storm-relative helicity (SRH), mixed-layer convective available potential energy (MLCAPE) and lifted condensation level (MLLCL), and variables like the significant tornado and supercell composite parameters (STP and SCP, respectively). For each parameter, we examine how the percentage of the first tornadoes that are warned varies in different bins across the parameter space.

We find that kinematic variables, particularly in lower layers, perform "the best"; as the variable magnitude increases, so does the percentage of warned first tornadoes. Figure 3 shows this percentage across different bins of 0–1-km shear and SRH (calculated assuming Bunkers-right storm motion; Bunkers et al. 2000). In general, as either of these parameters increase, so does the percentage of warned first tornadoes. The trend is clearest for 0–1-km shear; however, this is magnified by the one event that was warned in an environment with immense 0–1-km shear (70–80 m s^{-1}). Most of the data fall within the 0–50 m s^{-1} range of 0–1-km shear values, which show a clear positive trend in the percentage of warned first tornadoes. The percentage of warned first tornadoes over a common range of 0–70 kt 0–6-km shear follows a similar trend (Fig. 3).

On the other hand, thermodynamic variables like MLCAPE and MLLCL (Fig. 4) do not show the same trend. Rather, as MLCAPE and MLLCL increase, the percentage of warned first tornadoes changes very little. This suggests that thermodynamic variables offer almost no guidance on warning the first tornado of the day than kinematic variables.
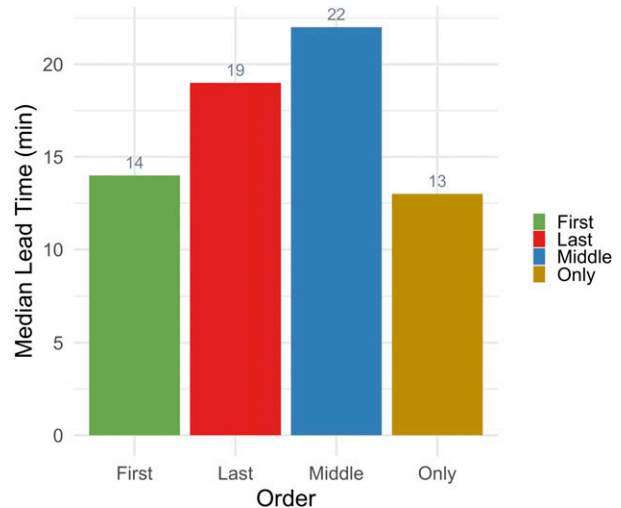
Composite parameters like the SCP and STP were formulated to combine the individual influences of kinematic and thermodynamic parameters on tornado potential (Thompson et al. 2003). As such, we hypothesized that these parameters would perform the best in terms of the percentage of warned first tornadoes. Figure 5 shows this percentage across the SCP and STP parameter spaces and in some ways is similar to the kinematic variables shown in Fig. 3. As SCP increases over a range of more commonly observed values (e.g., 0–10), the percentage of warned first tornadoes increases (Fig. 5a). As SCP increases from 10 to 30+, the percentage of warned first tornadoes plateaus and does not show a clear trend. The STP plot (Fig. 5b) shows a similar relationship, with the percentage of warned first tornadoes increasing as STP increases from 0 to around 2. A trend in the percentage of warned first tornadoes is not clear as STP increases from around 2 to 12.

In all, Figs. 3–5 suggest that the addition of thermodynamic information in the SCP and STP composite parameters does not provide any additional skill in helping forecasters anticipate the first tornado of the day. In other words, it would be better and simpler for forecasters to simply use kinematic variables like 0–1-km shear. However, this obviously does not discount the advantages of using composite parameters to predict the potential for an environment to support supercell and tornado formation (as they were designed to do).

4. Discussion and conclusions

Our results show that first tornadoes and only tornadoes of the day are successfully warned less often and with less lead time than tornadoes that occur in the middle or at the end of an event. There are a number of reasons that could explain this difference in warning performance. One reason is that forecasters may be unsure whether atmospheric conditions are currently conducive for tornadogenesis. This is consistent with our finding that only 25.7% of false alarms issued during tornadic events occurred prior to the first tornado of the day.
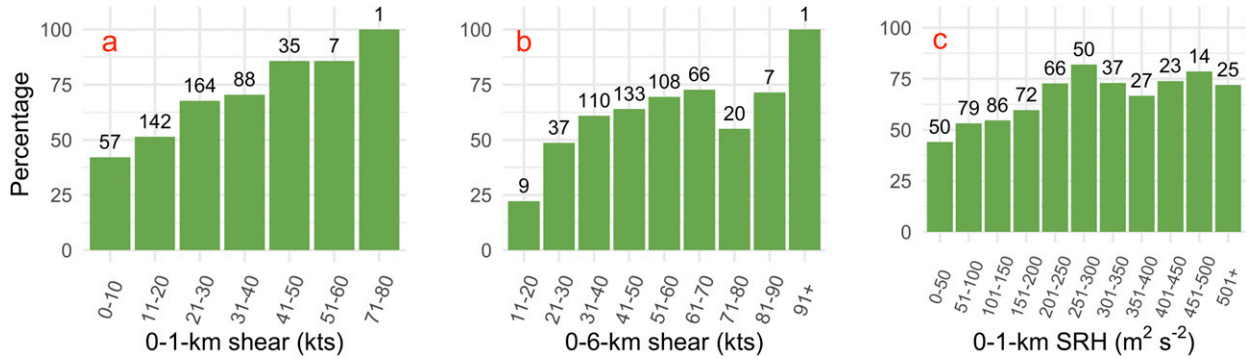
FIG. 3. The ratio of first tornadoes that were warned as based on corresponding kinematic environmental characteristics: (a) 0–1-km shear, (b) 0–6-km shear, and (c) 0–1-km SRH. The number of storms in each bin is reported above the bar.

After a tornado occurs, forecasters are likely more confident that the environment is supportive of tornadogenesis, and it may be easier to make the decision to continue warning that storm or to warn on an additional storm in the same area. In other words, the threshold to issue a warning likely decreases after the initial tornado occurs (e.g., Andra et al. 2002). Similarly, singular tornadoes may occur in more marginal environments, which can make the warning decision more difficult.

The fact that the minimum in FAR occurs "during" tornado production likely results from the time scale of typical tornado warnings relative to time scales of supercell and tornado evolution. Tornado warning durations have generally decreased from around 45 to 35 min in the last 30 years (Brooks and Correia 2018). Cycling mesocyclones can exhibit shorter lifetimes than this (e.g., Dowell and Bluestein 2002a,b; Beck et al. 2006; French et al. 2008). As a result, multiple mesocyclone and/or tornado cycles may occur within the same tornado warning, likely reducing FAR during that period of the supercell's life cycle.

A lowering warning threshold throughout the convective day is also consistent with our finding that over one-half of all false alarms issued during tornadic events occur after the demise of the last tornado. This is likely due to real-time forecasters not wanting to "miss" a tornado on a storm that has a history of producing tornadoes. This also may be influenced by our relative lack of understanding of processes influencing supercellular tornado decay, at least with respect to their genesis counterparts, as well as how changing environmental conditions may become less supportive of tornado production. Future study of tornado decay will yield more insight into these processes and how they may be better anticipated in real time.

It is interesting that the warning skill of the first tornadoes of the day increases across kinematic parameter spaces but not thermodynamic ones. It even appears that the addition of thermodynamic information in composite parameters like STP and SCP actually reduces the positive relationship between increasing kinematic parameters (like shear and SRH) and the percentage of warned first tornadoes (e.g., cf. Figs. 3 and 5). This is consistent with some recent modeling studies that suggest that tornado production may be more sensitive to the wind profile than the thermodynamic profile (e.g., Coffer and Parker 2018; Flournoy et al. 2020). Our findings complement these studies and show how further understanding of the
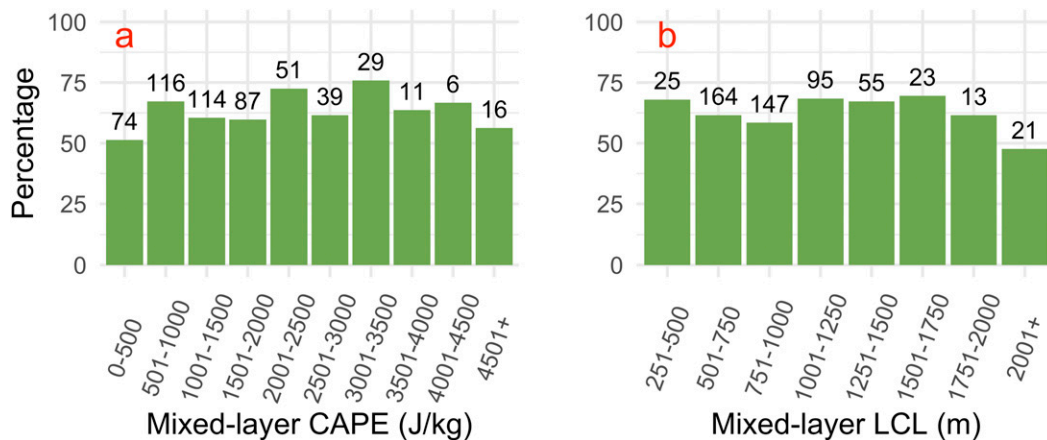

FIG. 4. The ratio of first tornadoes that were warned as based on corresponding thermodynamic environmental characteristics: (a) MLCAPE and (b) MLLCL. The number of storms in each bin is reported above the bar.
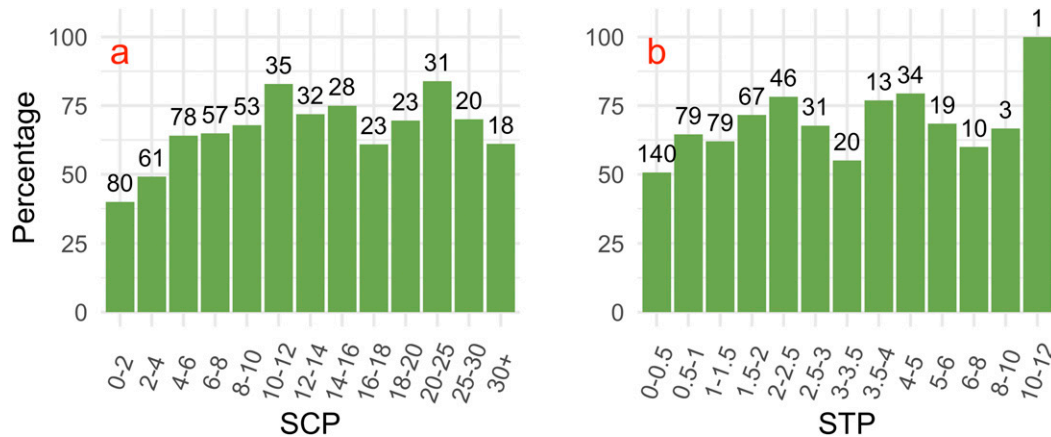
FIG. 5. The ratio of first tornadoes that were warned as based on corresponding composite parameters: (a) SCP and (b) STP. The number of storms in each bin is reported above the bar.

physical processes influencing tornado potential can continue to better inform operational forecasters.

Future forecasting systems should consider this warning challenge when developing intermediate products. In the current system, convective weather watches serve as one of these intermediate products that provide additional context for forecasters. Previous work has shown the value of these products in the form of forecaster readiness (Hales 1989), and ongoing work suggests additional value in terms of warning performance. Tornado warnings that occur in higher-end watches (e.g., tornado watches and ''Particularly Dangerous Situation'' tornado watches) tend to have higher POD and lower FAR (Krocak and Brooks 2021). These studies show the value of these intermediary products in the form of better-performing downstream products, although it is important to note that other factors (like differing environments) also play a role in the increase in warning performance.

Future work in this area will focus not just on the order of the warning by day and by CWA, but also the order of the tornado or warning by individual storm. This will reveal whether the relationships between tornadoes and warning performance found here exist on a storm-by-storm basis. Given our finding that warning performance on the first tornado of the day is related to the background environment, we expect that similar relationships will exist on the storm-scale. Furthermore, continued work should identify how intermediary products not only prime the forecaster to issue the warning, but also how these products can help members of the public prepare to make actionable decisions downstream. Although we know that the protective action decision making process is complicated (e.g., Lindell and Perry 2012), questions remain about what role intermediary products (like outlooks and watches) play in beginning the process prior to the warning. If people can prepare important items or inform friends and family prior to the warning stage, decisions made in that stage may be more informed and immediately actionable. Given the condensed decision making timeframe for tornadoes, this prewarning process is vitally important.

REFERENCES

Anderson-Frey, A. K., Y. P. Richardson, A. R. Dean, R. L. Thompson, and B. T. Smith, 2016: Investigation of near-storm environments for tornado events and warnings. *Wea. Forecasting*, **31**, 1771–1790, https://doi.org/10.1175/WAF-D-16-0046.1.

——, ——, ——, ——, and ——, 2018: Near-storm environments of outbreak and isolated tornadoes. *Wea. Forecasting*, **33**, 1397–1412, https://doi.org/10.1175/WAF-D-18-0057.1.

Andra, D. L., Jr., E. M. Quoetone, and W. F. Bunting, 2002: Warning decision making: The relative roles of conceptual models, technology, strategy, and forecaster expertise on 3 May 1999. *Wea. Forecasting*, **17**, 559–566, https://doi.org/10.1175/1520-0434(2002)017<0559:WDMTRR>2.0.CO;2.

Beck, J. R., J. L. Schroeder, and J. M. Wurman, 2006: High-resolution dual-Doppler analyses of the 29 May 2001 Kress, Texas, cyclic supercell. *Mon. Wea. Rev.*, **134**, 3125–3148, https://doi.org/10.1175/MWR3246.1.

Bieringer, P., and P. Ray, 1996: A comparison of tornado warning lead times with and without NEXRAD Doppler radar. *Wea. Forecasting*, **11**, 47–52, https://doi.org/10.1175/1520-0434(1996)011<0047:ACOTWL>2.0.CO;2.

Brooks, H. E., and J. Correia Jr., 2018: Long-term performance metrics for National Weather Service tornado warnings. *Wea. Forecasting*, **33**, 1501–1511, https://doi.org/10.1175/WAF-D-18-0120.1.

Brotzge, J., and S. Erickson, 2009: NWS tornado warnings with zero or negative lead times. *Wea. Forecasting*, **24**, 140–154, https://doi.org/10.1175/2008WAF2007076.1.

——, and ——, 2010: Tornadoes without NWS warning. *Wea. Forecasting*, **25**, 159–172, https://doi.org/10.1175/2009WAF2222270.1.

——, S. E. Nelson, R. L. Thompson, and B. T. Smith, 2013: Tornado probability of detection and lead time as a function of convective mode and environmental parameters. *Wea. Forecasting*, **28**, 1261–1276, https://doi.org/10.1175/WAF-D-12-00119.1.

Bunkers, M. J., B. A. Klimowski, J. W. Zeitler, R. L. Thompson, and M. L. Weisman, 2000: Predicting supercell motion using a new hodograph technique. *Wea. Forecasting*, **15**, 61–79, https://doi.org/10.1175/1520-0434(2000)015<0061:PSMUAN>2.0.CO;2.

Coffer, B. E., and M. D. Parker, 2018: Is there a ''tipping point'' between simulated nontornadic and tornadic supercells in VORTEX2 environments? *Mon. Wea. Rev.*, **146**, 2667–2693, https://doi.org/10.1175/MWR-D-18-0050.1.

Dowell, D. C., and H. B. Bluestein, 2002a: The 8 June 1995 McLean, Texas, storm. Part I: Observations of cyclic tornadogenesis. *Mon. Wea. Rev.*, **130**, 2626–2648, https://doi.org/10.1175/1520-0493(2002)130<2626:TJMTSP>2.0.CO;2.

——, and ——, 2002b: The 8 June 1995 McLean, Texas, storm. Part II: Cyclic tornado formation, maintenance, and dissipation. *Mon. Wea. Rev.*, **130**, 2649–2670, https://doi.org/10.1175/1520-0493(2002)130<2649:TJMTSP>2.0.CO;2.

Erickson, S. A., and H. E. Brooks, 2006: Lead time and time under tornado warnings: 1986–2004. *23rd Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., 11.5, https://ams.confex.com/ams/23SLS/techprogram/paper_115194.htm.

Flournoy, M. D., M. C. Coniglio, E. N. Rasmussen, J. C. Furtado, and B. E. Coffer, 2020: Modes of storm-scale variability and tornado potential in VORTEX2 near-and far-field tornadic environments. *Mon. Wea. Rev.*, **148**, 4185–4207, https://doi.org/10.1175/MWR-D-20-0147.1.

French, M. M., H. B. Bluestein, D. C. Dowell, L. J. Wicker, M. R. Kramar, and A. L. Pazmany, 2008: High-resolution, mobile Doppler radar observations of cyclic mesocyclogenesis in a supercell. *Mon. Wea. Rev.*, **136**, 4997–5016, https://doi.org/10.1175/2008MWR2407.1.

Guillot, E. M., T. M. Smith, V. Lakshmanan, K. L. Elmore, D. W. Burgess, and G. J. Stumpf, 2008: Tornado and severe thunderstorm warning forecast skill and its relationship to storm type. *24th Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, New Orleans, LA, Amer. Meteor. Soc., 4A.3, https://ams.confex.com/ams/88Annual/techprogram/paper_132244.htm .

Hales, J. E., Jr., 1989: The crucial role of tornado watches in the issuance of warnings for significant tornadoes. *Natl. Wea. Dig.*, **15**, 30–36.

Krocak, M. J., and H. E. Brooks, 2018: Climatological estimates of hourly tornado probability for the United States. *Wea. Forecasting*, **33**, 59–69, https://doi.org/10.1175/WAF-D-17-0123.1.

——, and ——, 2021: The influence of weather watch type on the quality of tornado warnings and its implications for future forecasting systems. *Wea. Forecasting*, **36**, 1675–1680, https://doi.org/10.1175/WAF-D-21-0052.1.

Lindell, M. K., and R. W. Perry, 2012: The protective action decision model: Theoretical modifications and additional evidence. *Risk Anal.*, **32**, 616–632, https://doi.org/10.1111/j.1539-6924.2011.01647.x.

NOAA/National Weather Service, 2007: NOAA's NWS national performance measures FY 2006–FY National Weather Service. Accessed 18 March 2021, http://www.nws.noaa.gov/com/files/All_GPRA2006.ppt.

——, 2021: Performance management web portal. Accessed 1 May 2021, https://verification.nws.noaa.gov/services/public/index.aspx.

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, https://doi.org/10.1175/WAF-D-11-00115.1.

Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261, https://doi.org/10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2.

——, B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, https://doi.org/10.1175/WAF-D-11-00116.1.