

An Analysis of 2016–18 Tornadoes and National Weather Service Tornado Warnings across the Contiguous United States

EVAN S. BENTLEY,^a RICHARD L. THOMPSON,^a BARRY R. BOWERS,^b JUSTIN G. GIBBS,^c AND STEVEN E. NELSON^d

^aNOAA/NWS/NCEP/Storm Prediction Center, Norman, Oklahoma

^bNational Weather Service Forecast Office, Norman, Oklahoma

^cNOAA/NWS Warning Decision Training Division, Norman, Oklahoma

^dNational Weather Service, Peachtree City, Georgia

(Manuscript received 30 December 2020, in final form 26 July 2021)

ABSTRACT: Previous work has considered tornado occurrence with respect to radar data, both WSR-88D and mobile research radars, and a few studies have examined techniques to potentially improve tornado warning performance. To date, though, there has been little work focusing on systematic, large-sample evaluation of National Weather Service (NWS) tornado warnings with respect to radar-observable quantities and the near-storm environment. In this work, three full years (2016–18) of NWS tornado warnings across the contiguous United States were examined, in conjunction with supporting data in the few minutes preceding warning issuance, or tornado formation in the case of missed events. The investigation herein examines WSR-88D and Storm Prediction Center (SPC) mesoanalysis data associated with these tornado warnings with comparisons made to the current Warning Decision Training Division (WDTD) guidance. Combining low-level rotational velocity and the significant tornado parameter (STP), as used in prior work, shows promise as a means to estimate tornado warning performance, as well as relative changes in performance as criteria thresholds vary. For example, low-level rotational velocity peaking in excess of 30 kt (15 m s^{-1}), in a near-storm environment, which is not prohibitive for tornadoes ($\text{STP} > 0$), results in an increased probability of detection and reduced false alarms compared to observed NWS tornado warning metrics. Tornado warning false alarms can also be reduced through limiting warnings with weak ($< 30 \text{ kt}$), broad ($> 1 \text{ n mi}$; $1 \text{ n mi} = 1.852 \text{ km}$) circulations in a poor ($\text{STP} = 0$) environment, careful elimination of velocity data artifacts like sidelobe contamination, and through greater scrutiny of human-based tornado reports in otherwise questionable scenarios.

SIGNIFICANCE STATEMENT: Recent studies have explored the radar signatures associated with severe storms and tornadoes and recently these radar signatures have been correlated to surveyed damage from tornadoes. However, to date, there is no known research relating the radar signatures that prompt National Weather Service tornado warnings to the verification of those warnings. This research accomplished this goal and showed that the most skillful warning thresholds match current guidance for National Weather Service forecasters. Typically, an increase in POD will result in an increase in FAR and vice versa. However, this research showed there may be opportunities to improve POD and FAR with minimal negative consequences by focusing on the tails of the distribution (poor environment/weak signature and favorable environment and strong signature).

KEYWORDS: Radars/Radar observations; Mesoscale forecasting; Operational forecasting

1. Introduction

The National Weather Service (NWS) issues tornado warnings across the United States and its territories for the protection of life and property. Tornado warnings are meant to precede tornado formation, or the arrival of an ongoing tornado, within the warned area. The format of the warnings transitioned from county-based to storm-based polygons in 2007 (Ferree et al. 2006). The justification for tornado warnings includes observations of tornadoes from storm spotters, but the majority of tornado warnings with lead time in the United States are issued primarily on the basis of WSR-88D data. The tornado warning process is complex and nuanced, and Brotzge and Donner (2013) provided an overview of the end-to-end tornado warning

process, with discussion of the challenges facing operational forecasters in both the prediction and detection phases of the warning process. Brooks and Correia (2018) documented long-term increases in probability of detection, gradual reductions in false alarms, and relatively steady lead time in NWS tornado warnings from 1986 to 2011. Thereafter, lead time and probability of detection has been reduced by shorter-duration warnings, with an apparent emphasis on reducing false alarms.

Within NWS warning procedures, forecasters are expected to estimate the potential impacts of tornadoes in warning polygons, as part of the impact-based warning program (IBW; Wagenmaker et al. 2014). As of March 2016, IBW was adopted nationwide for all tornado warnings, which includes warning tags for a “radar indicated” or “observed” tornado [based on reliable spotter reports or a tornadic debris signature (TDS)]. NWS tornado warnings do not include explicit estimates of

Corresponding author: Evan S. Bentley, evan.bentley@noaa.gov

DOI: 10.1175/WAF-D-20-0241.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

tornado intensity. However, in certain circumstances, the tornado damage threat can be specified as “considerable” when a tornado is thought to be capable of producing EF2+ damage with a preference for a confirmed tornado, or “catastrophic” when there is a confirmed tornado with strong evidence of EF2+ intensity and a significant enough population footprint that catastrophic impacts to life and/or property are imminent (Warning Decision Training Division 2021).

The ability of the WSR-88D to detect potentially tornadic circulations has improved markedly since the late 2000s with the advent of super-resolution velocity data in 2007 (Brown et al. 2002, 2005; Torres and Curtis 2007), dual polarization data in 2012 (Saxion and Ice 2012), and the supplemental adaptive intravolume low-level scans (SAILS)/multiple elevation scan option for SAILS (MESO-SAILS) scanning strategies by 2014 (Chrisman 2011, 2014) for more frequent lowest-scan updates. The super-resolution velocity data, with effective 0.5° beamwidth, allows for more precise velocity observations and resultant estimates of storm-scale rotation magnitude. Dual polarization data, specifically cross-polar correlation coefficient data, allow identification of a TDS (Ryzhkov et al. 2005; Schultz et al. 2012a,b; Bodine et al. 2013; Van Den Broeke and Jauernic 2014; Snyder and Ryzhkov 2015). The MESO-SAILS scan strategies result in lowest-scan updates on the order of 1–2 min, compared to 4–6-min update frequencies with legacy scan strategies, though at the expense of longer gaps in full volume scan updates. The net result of these upgrades to the WSR-88D has been to improve detection of velocity and reflectivity signatures associated with tornadoes in both supercells and quasi-linear convective systems (QLCSs), though the WSR-88D does not explicitly resolve the majority of tornadoes due to beamwidth, beam height, and range limitations. Tornadoes are not fully resolved in WSR-88D data the vast majority of the time due to the aforementioned limitations, but recent work (e.g., Toth et al. 2013; Smith et al. 2015; Kingfield and LaDue 2015; Thompson et al. 2017; Smith et al. 2020a,b) has revealed clear relationships between stronger WSR-88D measurements of storm-scale rotation and both stronger tornadoes and greater probabilities of tornado occurrence.

Taking advantage of these improvements to the NEXRAD network, multiple studies have focused on tornado discrimination and tornado intensity estimates, in support of NWS tornado warnings. Kingfield and LaDue (2015) evaluated discrimination between weak (EF0–1) and strong (EF2+) tornadoes via automated calculations of low-level rotational velocity, with skill maximized at a threshold of 23 m s^{-1} ($\sim 45 \text{ kt}$). Smith et al. (2015) considered a large sample of observed tornadoes from 2009 to 2013, encompassing the first 5 years of nationwide super-resolution velocity data. Smith et al. (2015) found that manual calculations of maximum low-level rotational velocity during tornado lifetime tended to increase as maximum tornado damage intensity increased. Gibbs (2016) evaluated the skill of several techniques in discriminating between observed weak (EF0–1) and significant (EF2+) tornadoes in impact-based warnings issued by the NWS. Real-time discrimination between weak and significant tornadoes was most skillful at a low-level rotational velocity threshold of 40 kt ($1 \text{ kt} \approx 0.51 \text{ m s}^{-1}$) combined with the presence of a TDS. Thompson et al. (2017, hereafter T17), building on the prior work by Smith et al. (2012, 2015) for tornadic

storms, considered a large sample of nontornadic, cyclonic velocity couplets associated with severe thunderstorms [right-moving supercells and QLCSs producing hail $\geq 1 \text{ in.}$ (2.54 cm) diameter, measured or estimated gusts $\geq 50 \text{ kt}$ ($\sim 25 \text{ m s}^{-1}$), or wind damage] in an effort to estimate tornado probabilities based on low-level rotational velocity. Gibbs and Bowers (2019), using the Smith et al. (2015) and T17 case samples, focused on lead time to the onset of EF2+ tornado damage, based on peak rotational speed, circulation diameter, and mesocyclone depth. Smith et al. (2020a,b) refined the work of Smith et al. (2015) by showing that low-level rotational velocity and near-storm environment can be combined to estimate tornado intensity on a scan-by-scan basis with WSR-88D data.

The Warning Decision Training Division (WDTD) is tasked with training NWS forecasters in tornado warning issuance, and has compiled and organized the findings of the aforementioned studies (Fig. 1). The recommendations shown in Fig. 1 were largely based on work done by Gibbs (2016) and T17, where the probability of a tornado and estimates of tornado intensity were calculated using near-storm environmental data and WSR-88D storm-scale signatures. The data provided by Gibbs (2016) and T17 allowed for calibrated, quantitative estimates of tornado probabilities, whereas most prior estimates were more subjective and/or qualitative. Over 1100 NWS employees have taken the in-person (Radar Applications Course) training or on-line IBW training since 2018 when this quantitative guidance was added.

While the recommended criteria from WDTD and the WSR-88D-based tornado probabilities calculated by T17 are optimized for peak forecast skill, tornadoes can occur in a wide range of scenarios and warnings also require consideration of lead time. In addition, there are differing warning philosophies and thresholds applied by each human forecaster. In fact, Karstens et al. (2018) showed that when NWS forecasters were asked to identify the probability of a tornado associated with the outline of an experimental warning plume in the Hazardous Weather Testbed PHI experiment, the answers ranged from 0% to 80% with a majority of the answers between 20% and 60% (Fig. 2).

The variability shown in Fig. 2 provides motivation to establish consistent and reproducible tornado warning criteria within the NWS. The primary goal of this work is to quantify the radar and environmental characteristics related to all tornado warnings, both with and without tornadoes, as well as unwarned tornadoes in the contiguous United States from 2016 to 2018. The metrics calculated for NWS tornado warnings during this 3-yr period will be compared to estimated warning performance based on the frequency of occurrence of severe, nontornadic storms, in an effort to calibrate current NWS warning performance and examine opportunities for potential tornado warning performance improvements.

2. Data and methods

Radar-identified convective mode, low-level rotational velocity, and near-storm environmental data were assigned to 6881 NWS tornado warnings issued between 2016 and 2018

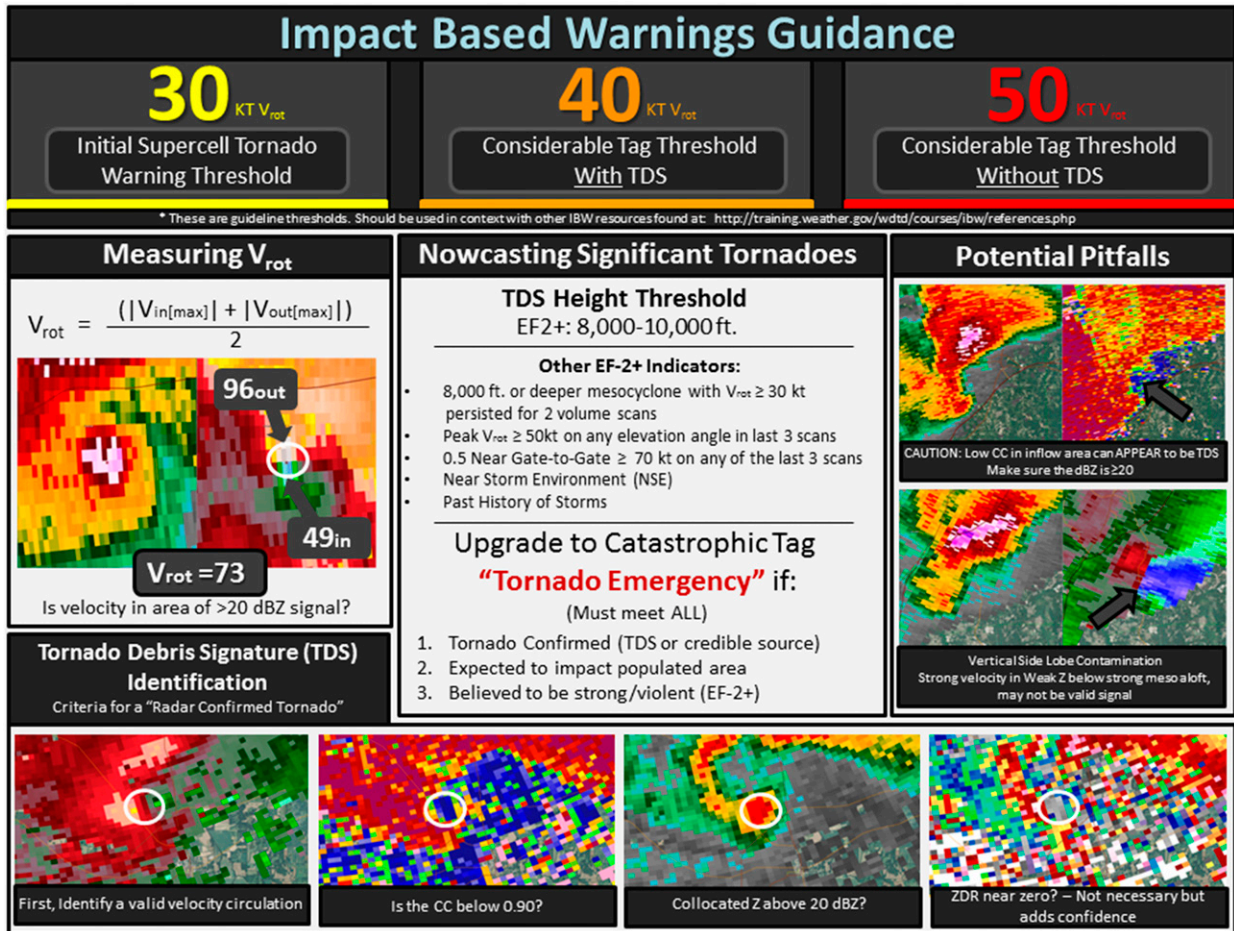


FIG. 1. One page “quick guide” created by the WDTD for NWS severe weather warning forecasters to use as a reference for recommended impact based tornado warning guidance.

using a similar methodology to T17. The warnings were collected from the NWS Performance Management web page, and the Iowa State Environmental Mesonet archive was used to collect 93 tornado warnings that were missing from the NWS Performance Management database. Convective mode was assigned to each tornado event via manual examination of full volumetric WSR-88D data in the 10 min before tornado warning issuance, and 0.5° elevation scan peak rotational velocity (hereafter referred to as V_{rot}) was calculated manually using super-resolution radar data in the same 10-min time period before tornado warning issuance, where

$$V_{rot} = (\text{maximum outbound} - \text{maximum inbound})/2.$$

The maximum inbounds and outbound velocities must have been within 5 n mi (1 n mi = 1.852 km) of one another and a line connecting the velocity maxima must have been within 45° of the radial through the velocity couplet centroid. The 10-min time window was chosen to represent the period immediately preceding a warning decision. Supercell-related convective parameters from the hourly Storm Prediction Center (SPC)

objective analyses for the preceding hour,¹ on a 40-km horizontal grid (Bothwell et al. 2002), accompanied each tornado warning or unwarned tornado.

a. Tornado warning verification

Warning verification data were collected from the Performance Management web page (NWS 2011). A total of 93 tornado warnings were missing from the Performance Management website, and these warnings were verified manually using time matching and geolocation between warning and tornado data. In addition, there were 38 apparent tornado report errors in Storm Data (i.e., time was entered in daylight time rather than standard time)—in these 38 cases, the times/locations were modified to match the centroid of V_{rot} signatures from WSR-88D data. This resulted in 26 additional verified tornado warnings. Finally, 35 warnings that had a clear TDS (similar to those observed in Edwards and Picca 2016) within them were counted as verified

¹ Note: warnings issued in the first ~10 min of a new hour would not have the most recent hourly analysis available to the radar operator.

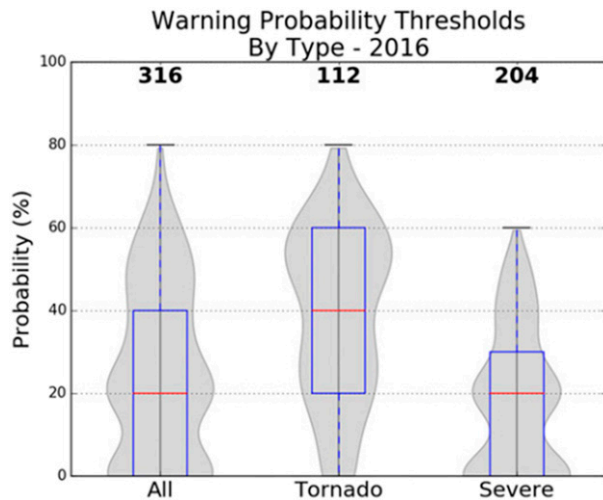


FIG. 2. Violin plots [reproduced from Karstens et al. (2018)] quantifying probability values selected by NWS forecaster during HWT PHI experimental warning plumes.

regardless of whether or not there was a tornado report in *Storm Data* (e.g., Schultz et al. 2012b). After these adjustments, the overall POD and FAR for our database for 2016–18 was POD = 0.62 and FAR = 0.69 despite NWS *Storm Data* (NWS 2011) showing POD = 0.58 and FAR = 0.70. Therefore, future comparisons to current NWS warning performance will be made to these adjusted metrics of POD = 0.62 and FAR = 0.69.

b. Radar characteristics

The Gibson Ridge Level II Analyst radar-viewing software (http://grlevelx.com/gr2analyst_2/) was used to analyze archived

WSR-88D level-II single-site radar data from the NEXRAD archive hosted by Amazon AWS. The closest radar to the storm was used to calculate the peak V_{rot} in the 10 min prior to each tornado warning issuance, and during the life of the tornado for the missed events. Once the peak V_{rot} radar scan was identified, all other radar characteristics were derived from this radar scan including convective mode, maximum V_{rot} circulation diameter, the presence or absence of a TDS, and the number of SAILS/MESO SAILS scans between full volume updates. Convective modes were classified as discrete, cell in cluster, or cell in line with storm types of supercell, QLCS, or marginal supercell using the same methodology outlined in Smith et al. (2012). Potentially contaminated velocity bins as a result of sidelobe contamination (Doviak and Zrnić 1993; Piltz and Burgess 2009; see Fig. 3), three-body scatter spikes (Zrnić 1987), or apparent velocity dealiasing problems, were documented, if present.

Potentially contaminated velocity data were not excluded because there were relatively frequent occurrences (26%) of tornado warnings with accompanying velocity signatures that were impacted by velocity contamination (i.e., sidelobe shown in Fig. 3). Potential data quality concerns were noted with each case, but the strongest V_{rot} was used in conjunction with each tornado warning to reflect the potential for the contaminated velocity data to influence warning decisions. Likewise, we focused on the peak V_{rot} in the 10 min prior to tornado warning issuance, to reflect data that were available to the NWS forecaster leading up to the warning decision. The diameter of the circulation (restricted to the same ≤ 5 n mi diameter as Smith et al. 2015 and T17) was recorded by measuring the distance between the center of the strongest inbound velocity pixel and the strongest outbound velocity pixel used for the V_{rot} calculation. In the 10 min prior to tornado warning issuance, reflectivity, velocity, and

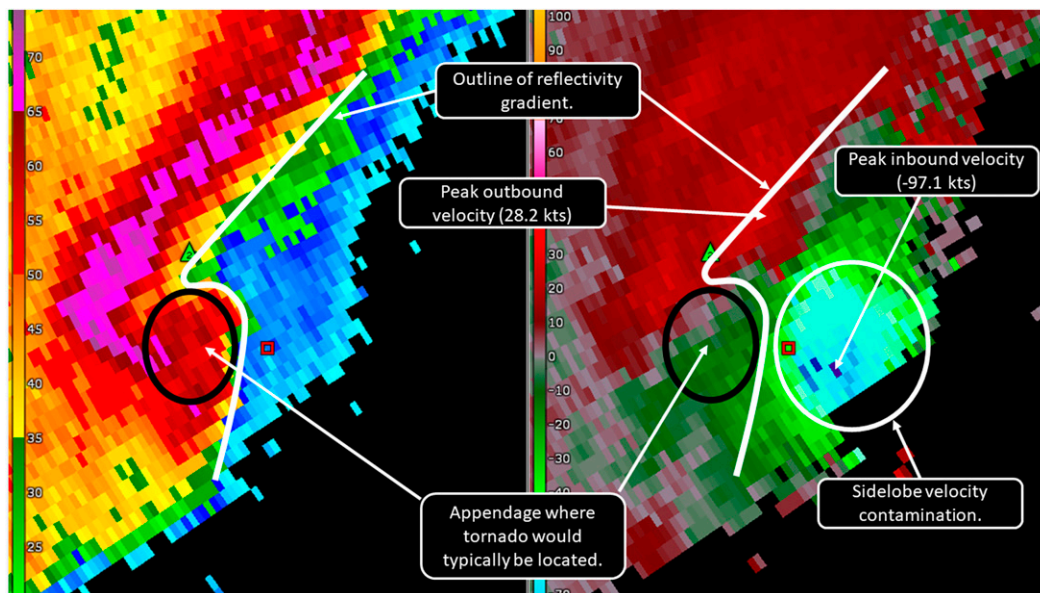


FIG. 3. Example of a two-panel (left) reflectivity (dBZ) and (right) velocity (kt) display showing sidelobe velocity contamination and the effect it can have on properly identifying the V_{rot} for a given storm.

TABLE 1. Table showing the confirmed source for all NWS tornado warnings with the confirmed tag and the number of warnings, verified warnings, and FAR for each.

Source	Warnings			Verified			FAR
	TOR	SVS	Total	TOR	SVS	Total	
Radar	158	236	393	109	219	328	0.17
Trained spotter/storm chaser/NWS employee	286	228	514	173	184	357	0.30
Observed/reports/public/broadcast media	50	34	84	25	25	50	0.40
Emergency management/law enforcement	130	78	208	51	58	109	0.48
Total	624	576	1194	358	486	844	0.29

correlation coefficient radar data were examined to determine whether there was a TDS present prior to tornado warning issuance. Unlike the other radar characteristics, which were identified at the scan with the highest V_{rot} , the presence of a TDS was examined in all scans present in the 10 min prior to tornado warning issuance, and for a few scans after tornado dissipation. This was done to account for delays in TDS appearance and to determine the verification rate of the “TORNADO...CONFIRMED” IBW tags. Last, the height above radar level (ARL) and the number of SAILS/MESO-SAILS scans active² were recorded at the time of the maximum V_{rot} in the 10 min preceding a tornado warning, or during the lifespan of an unwarned tornado.

c. Velocity contamination

As noted above, maximum inbound and outbound velocities were used for V_{rot} calculation regardless of data quality. Velocity contamination was identified using all tilts of reflectivity, velocity, and correlation coefficient radar data. The most common type of velocity contamination, which corresponded to more than 90% of the questionable velocity signatures, was sidelobe contamination (Fig. 3). Sidelobe contamination was typically identified as an area of (usually) greater velocity in the weak reflectivity inflow region in the low levels of a supercell, with much greater reflectivity overhang aloft. The other types of velocity contamination present were dealiasing issues, unknown radar issues that caused single high velocity pixels that were not supported by any nearby pixels, three body scatter spikes, and nonuniform beam filling.

3. Results

a. IBW tags and report sources

There were 1205 warnings (17.5%) that had the confirmed tag in them at some point during the time span of the warning [initial warning or follow-up severe weather statements (SVS)]. Of these, 814 (68%) of the tornado warnings verified with a tornado during the valid time of the warning. When the confirmed tag was introduced in the initial warning, that

warning verified 54% of the time. However, when the confirmed tag was introduced within an SVS, the warning verified 82% of the time. This discrepancy can largely be explained by tornadoes that dissipate in the period between the report and warning dissemination, or tornadoes that dissipate before entering the warning (in the case of downstream tornado warning issuance), both of which would not be an issue for SVS verification. This is a somewhat common occurrence as documented by Blair and Leighton (2014).

The confirmed tags were also examined by the sources of the tornado reports. After combining similar sources (i.e., spotter and trained spotter, and the use of plural), there were 10 confirmed sources that were used in warnings. These sources were radar, trained spotter, storm chaser, NWS employee, observed, reports, public, broadcast media, emergency management, and law enforcement. Due to limited data from some of these categories, they were further combined into four categories: radar, trained spotter (storm chaser, NWS employee), public (broadcast media, observed, reports), and law enforcement (emergency management). The confirmed source that was most likely to verify with a tornado in a tornado warning was radar (83%), and the least likely to verify was law enforcement (52%). Of the 393 warnings that used the radar confirmed tag, 328 verified with a tornado report or TDS, and 66 did not (Table 1).

Many of these unverified tornado warnings with a confirmed tag were issued due to a valid tornado report, but the tornado ended before it entered the warning area, or just prior to warning issuance. Therefore, additional analysis was done to determine how often there was no tornado confirmed within 10 mi and 30 min of the time the confirmed tag was added to a warning in an attempt to quantify how often the source of the confirmed tag was incorrect. Of the tornado warnings issued on the basis of a reported tornado, 136 (17%) were apparently based on false reporting (Table 2). The highest percentage of false reports used for tornado confirmed tags came from emergency management and law enforcement (25%). We speculate that the perceived authority that accompanies tornado reports from emergency management and law enforcement, both of which are core partners of the NWS, also drives the higher frequency of false alarm tornado warnings from the NWS. Since many of the tornado reports relayed by law enforcement and emergency management are second hand, there appears to be an opportunity for

² SAILS adds one additional 0.5° scan per full volume update, and MESO-SAILS adds two or three additional 0.5° scans.

TABLE 2. Total number of unverified tornado warnings with confirmed tag by confirmation source.

Source	Unvalidated report			Percent of time it is unvalidated	Percent of all unvalidated reports
	TOR	SVS	Total		
Trained spotter	34	31	65	13.05%	47.79%
Public	12	6	18	22.50%	13.24%
Emergency management	11	4	15	22.39%	11.03%
Law enforcement	29	9	38	26.95%	27.94%
Total	86	50	136	17.00%	

the NWS to improve the integrity of “confirmed” tornado warnings by further scrutinizing tornado reports without clear corroborating evidence from radar and/or the near-storm environment.

There were 393 radar confirmed tornado warnings issued during the study period and 317 of these radar confirmed warnings were found to have a TDS. This leaves 76 (19.3%) of the radar confirmed tornado warnings that were issued without evidence of a TDS (Table 3). The most common characteristic of these misidentified TDS signatures was reflectivity < 20 dBZ, within the supercell inflow region in the low levels. The sample size is somewhat small, but there does not appear to be a discernible trend in the number of misidentified TDS signatures through time.

The average peak V_{rot} in the 10 min prior to the addition of a “considerable” tag was 56.1 kt ($\sim 29 \text{ m s}^{-1}$), and above the WDTD recommended threshold of 50 kt (26 m s^{-1} ; without a TDS) the majority of the time (67%). Similarly, for the “catastrophic” tag, the average V_{rot} was 63.5 kt ($\sim 33 \text{ m s}^{-1}$) and the V_{rot} was >50 kt a clear majority of the time (79%).

b. Radar characteristics

1) STORM MODE AND V_{ROT}

Tornado warning verification varied substantially based on convective storm mode. POD and FAR were higher for right-moving supercells (POD = 0.76 and FAR = 0.70), and lower for QLCS events (POD = 0.52 and FAR = 0.62), though the critical success index (CSI; Schaefer 1990) was nearly equal for both (0.277 and 0.278, respectively). Unsurprisingly, the tornadoes and tornado warnings with marginal supercells (defined by a $V_{rot} < 20$ kt) were characterized by low POD (0.31) and a high FAR (0.87).

The average maximum V_{rot} in the 10 min prior to warning issuance was larger for tornado warnings verified with tornado reports (40 kt; 21 m s^{-1}) than for false alarm tornado warnings (34 kt; 17 m s^{-1}). The average maximum V_{rot} for tornadoes without tornado warnings (28 kt; 14 m s^{-1}) was lower than for the average of storms with false alarm tornado warnings. Figure 4 shows the full sample by binned V_{rot} values. The FAR remains quite high for $V_{rot} < 30$ kt (15 m s^{-1}) before dropping significantly in the range of 30–35 kt ($15\text{--}18 \text{ m s}^{-1}$), and then continuing to drop as V_{rot} increases. In addition, the true skill statistic (TSS) was calculated for each bin using the method first described by Wilks (1935). The correct null forecasts were estimated based on the conditional probability of a tornado per binned value of V_{rot} (using the raw data displayed in Fig. 8 of

T17). The TSS shows peak skill in the 30–35- and 35–40-kt bins, where POD is reasonably large, FAR is reasonably low, and a majority of correct null cases are maintained for the V_{rot} values < 30 kt.

The individual V_{rot} bins are combined in Fig. 5 to estimate potential tornado warning performance metrics for V_{rot} thresholds. Per Fig. 5, POD = 0.57 and FAR = 0.78 at a V_{rot} threshold ≥ 30 kt, which is where TSS is maximized, both of which are worse performance than the 2016–18 NWS warning performance. Therefore, it can be reasonably assumed the human forecaster does add value to the current warning process over strict adherence to a V_{rot} threshold (i.e., $V_{rot} \geq 30$ kt) (Fig. 6). It is important to note that there is no unique database of rotating storms, regardless of tornado production. T17 examined a 2-yr sample of nontornadic, severe storms across the contiguous United States in an effort to estimate the relative frequency of tornadoes in association with particular ranges of cyclonic V_{rot} . For example, their Fig. 7 suggests that $\sim 85\%$ of all severe storms with V_{rot} values that peak in the 20–29-kt ($10\text{--}14 \text{ m s}^{-1}$) range were nontornadic. Likewise, their results suggest that tornado warnings issued for all storms where V_{rot} peaks in the 20–29-kt range would result in $\sim 85\%$ false alarms. Thus, we used the T17 results to estimate the number of potentially nontornadic storms that would be false alarms (or correct null cases), if tornado warnings were issued strictly on the basis of V_{rot} thresholds.

Compact circulations on radar verified with tornadoes more often than broader diameter velocity couplets (Fig. 7). However, the FAR does not increase linearly with circulation diameter. The FAR is lowest with circulation diameters less than 0.5 n mi,

TABLE 3. Table of total “radar confirmed” warnings with the percentages that were correctly or incorrectly identified TDS signatures. Data are displayed by year in the lower half of the table.

	Cases	Percent of all radar	
True TDS	317	80.66%	
Misdiagnosed TDS	76	19.34%	
Total	393		
Misdiagnosed TDS by year			
	Misdiagnosed	Correct	Percent
2016	28	93	23.14%
2017	20	121	14.18%
2018	28	103	21.37%

2016-2018 Tornado Warning Performance by V_{rot} Bin

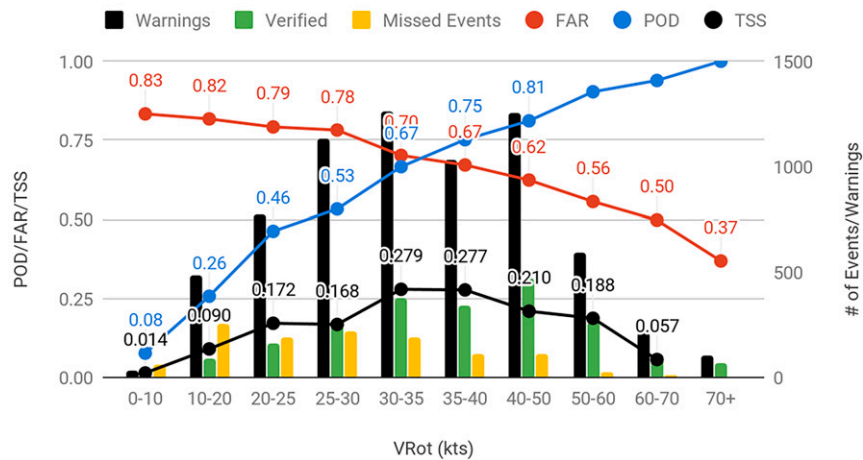


FIG. 4. Tornado warning FAR (red), POD (blue), and TSS (black) per V_{rot} bin. Bar charts show all tornado warnings (black), verified tornado warnings (green), and missed events (yellow). FAR (assuming warning issuance for all cases meeting the threshold values) is estimated using the T17 null case frequency for each V_{rot} bin. Actual NWS FAR is lower because tornado warnings are not issued for every single rotating storm, especially for the low V_{rot} values (<20 kt).

increases some between 0.5 and 1.0 n mi (but is still low), and the FAR changes very little for circulations broader than 1.0 n mi in diameter. There was no discernible trend in missed events based on circulation diameter with a relatively narrow POD range of 0.63–0.7.

The FAR increases as the radar range increases, with the lowest FAR for tornado warnings on storms close to the radar site (i.e., <2000 ft ARL). This is likely related to radar resolution

in addition to the height of the beam in sampling the low-level mesocyclone. At 2000 ft ARL, the radial distance between 2 pixels is roughly 0.25 n mi. However, 4 pixels reside within the same radial distance around 500 ft ARL. Surprisingly, missed events show little if any correspondence to radar beam height with nearly uniform POD across all radar height bins below 8000 ft ARL (Fig. 8). FAR is lower for storms being sampled below 2000 ft ARL (close to the radar), but POD shows little

2016-2018 Tornado Warning Performance by V_{rot} Threshold

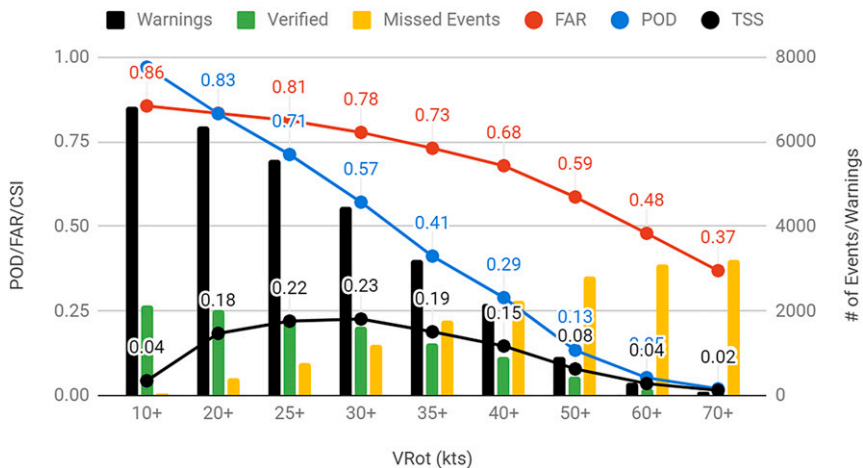


FIG. 5. Estimated tornado warning FAR (red), POD (blue), and TSS (black) per V_{rot} threshold, assuming warnings were issued in every case of $V_{rot} \geq 10$ kt. Bar charts show all estimated tornado warnings (black), verified tornado warnings (green), and missed events (yellow). FAR (assuming warning issuance for all cases meeting the threshold values) is estimated using the T17 null case frequency for each V_{rot} bin. Actual NWS FAR is lower because tornado warnings are not issued for every single rotating storm, especially for the low V_{rot} values (<20 kt).

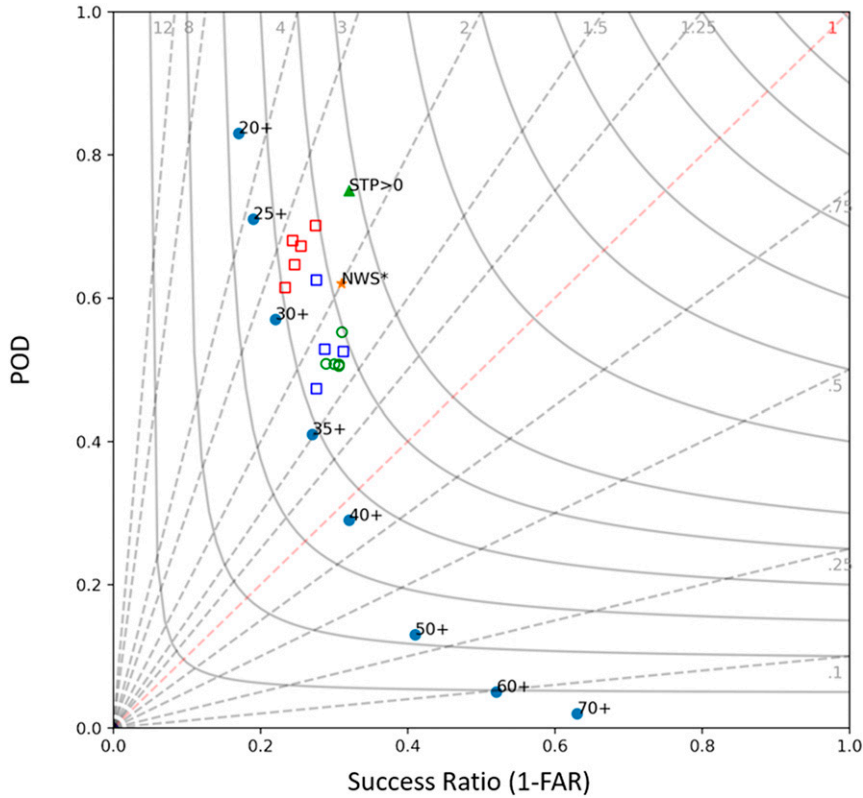


FIG. 6. Performance diagram depicting POD and success ratio for V_{rot} thresholds from Fig. 6 (blue filled circles), Brooks and Correia (2018) yearly NWS stats (unfilled symbols, red: 2007–11, blue: 2012–16, green: 2017–20), and STP > 0 (green triangle) from Fig. 4. NWS overall 2016–18 performance (after adjusted for storm data and missing warnings errors), shown by the orange star.

difference. The reason for this is unclear; however, it could be related to damage survey practices for tornado warnings with little or no damage, especially for cases of relatively brief/weak tornadoes with little corroborating evidence.

The number of SAILS/MESO-SAILS scans active (based on the radar SAILS metadata setting) at the time of the warning or tornado seems to correspond with tornado warning performance. Both FAR and percentage of tornadoes that were

2016-2018 Tornado Warning Performance by Circulation Diameter

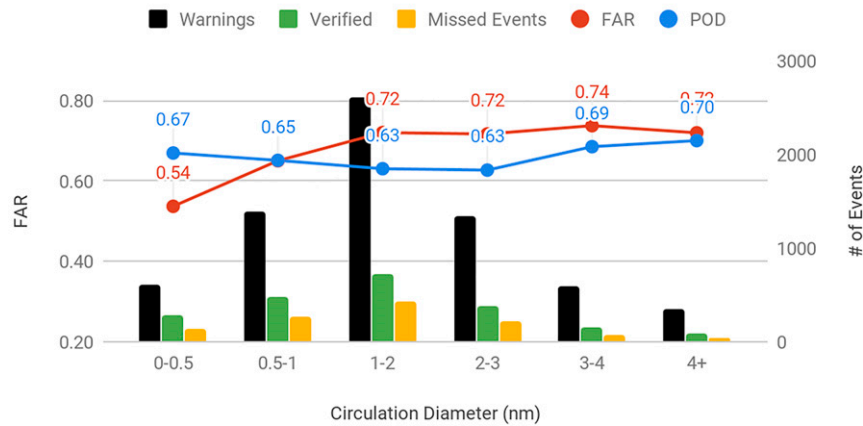


FIG. 7. POD (blue) and FAR (red) per circulation diameter bins (n mi). Bar charts show warnings (black), verified warnings (green), and missed events (yellow).

2016-2018 Tornado Warning Performance by Radar Range



FIG. 8. FAR (red) and POD (blue) per height ARL range bin. Bar chart also shows warnings (black), verified warnings (green), and missed events (yellow).

missed decreased as the number of SAILS scans increased (Fig. 9). However, there are likely other factors that impact this finding. For example, the average STP (~2.2) for tornado warnings issued with SAILS_3 active (three extra 0.5° scans per full volume update) is more than twice as large as STP (~1.1) when the SAILS scan strategies are not in use. Therefore, at least a portion of tornado warning performance related to the number of SAILS scans can be attributed to the near-storm environment and the expectation for tornadoes—a near-storm environment more favorable for tornadic supercells typically corresponds to a larger number of SAILS scans. A larger number of SAILS scans would seem to be more favorable most of the time; however, there is an opportunity cost of fewer full volumetric scans that have been shown to have some correlation to significant tornado development (Gibbs and Bowers

2019). Mid-volume rescans of low-level elevations (MRLE) has been added after the study period and helps add more midlevel and low-level scans, but some opportunity cost with full volumetric data still exists.

2) VELOCITY CONTAMINATION

Velocity contamination was found to be a frequent occurrence, especially in supercells. Contaminated velocity data were identified in the 10 min prior to warning issuance for about a quarter (26%) of all tornado warnings. When the storm mode is supercell, this frequency increases to 29%. Also, with supercell storm mode, the height ARL does seem to have an impact on the frequency of velocity contamination. Between 4000 and 8000 ft ARL, the frequency of storms with velocity contamination increases to 31%–32%, while contaminated

2016-2018 Tornado Warning Performance by SAILS Scans

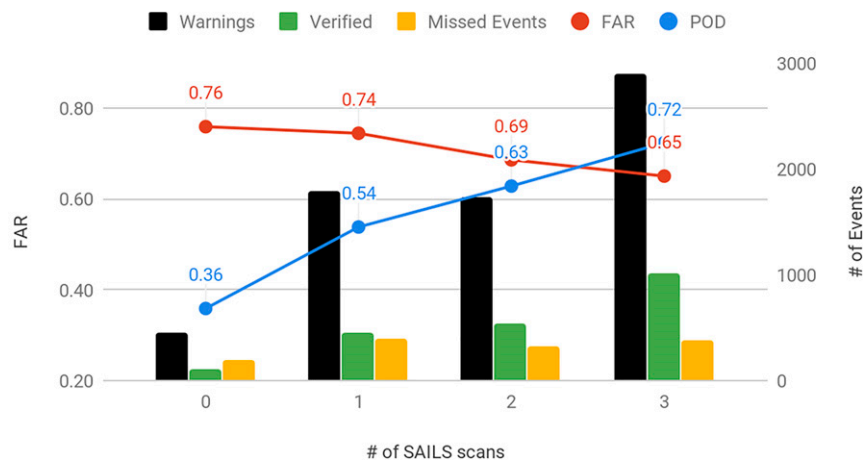


FIG. 9. FAR (red), POD (blue) per number of SAILS scans. Bar chart also shows warnings (black), verified warnings (green), and missed events (yellow).

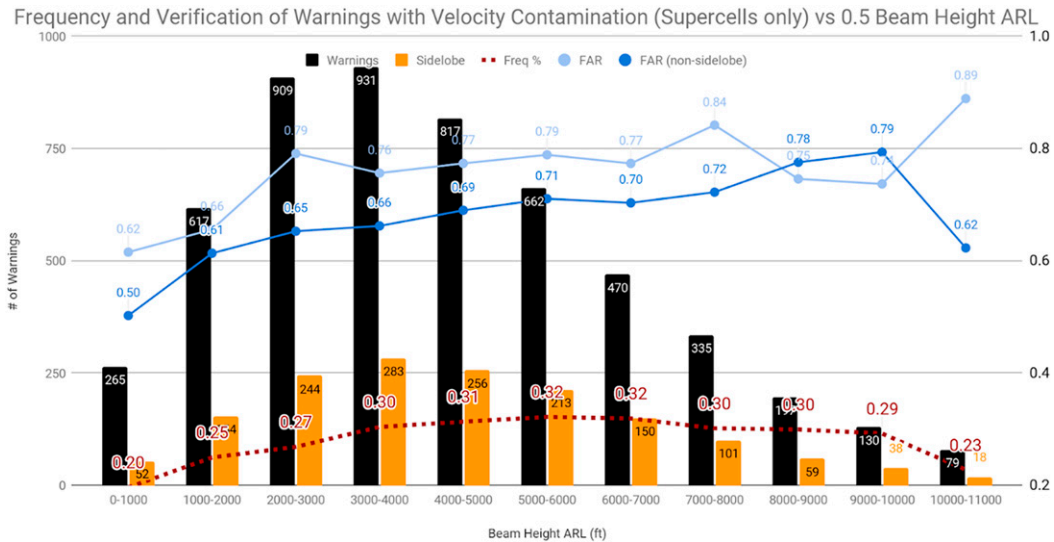


FIG. 10. Frequency (red dotted line) of warnings that are issued when sidelobe contamination is present within 10 min of the warning issuance per height ARL range bin. Also shown are bar charts of warnings (black), and warnings with sidelobe (yellow). The blue line shows the FAR for warnings with no sidelobe present while the light blue lines show the FAR for warnings with sidelobe present.

velocity data are less common (<20%) when the lowest scans observe the storm below 1000 ft ARL (Fig. 10). This increased frequency of poor velocity data observed between 4 and 8 kft (in the lowest scan) is likely attributable to the typical supercell where the radar sidelobe samples velocity data with the high-reflectivity core and echo overhang aloft, above the low-level storm inflow region with little reflectivity. Velocity contamination appears to drive an increase in tornado warning FAR (0.79 compared to only 0.66 for tornado warnings with high-quality velocity data). The majority of contaminated velocity data are caused by the radar sidelobe, which almost always results in seemingly stronger velocities within the inflow region of the storm, stronger (albeit inaccurate) V_{rot} , and an overestimate of the tornado threat within the 10 min prior to tornado warning issuance.

c. Mesoscale environment

Each tornado warning was paired to the nearest gridpoint parameter values from the SPC hourly mesoanalysis system (Bothwell et al. 2002; Schneider and Dean 2008), for the hour immediately prior to each warning or unwarned tornado. The FAR for three percentile rank values from a standard box-and-whisker plot (i.e., the bottom whisker, the median, and the top whisker) is shown in Table 4 for a set of convective parameters. In general, tornado warning FAR is largest for the lowest (10th percentile) values of the shear-related parameters, and for the highest values (90th percentile) of the moisture variables [either high lifting condensation level (LCL) heights/large dewpoint depressions, or very moist environments with the largest precipitable water (PWAT) values]. The greatest reduction in FAR across the spectrum of values occurs as the composite parameters increase from marginal (10th percentile) to extreme (90th percentile) values, led by the supercell composite

parameter (SCP; Thompson et al. 2003) and the significant tornado parameter (STP; Thompson et al. 2012).

Our analysis, in the context of tornado warnings, provides additional support for the use of STP as a relatively simple representation of the near-storm environment and associated tornado potential, as discussed by Thompson et al. (2012). While STP provides discrimination throughout the entire spectrum of values, especially to discriminate tornado intensity, the greatest FAR reduction occurs for any STP values greater than zero (Fig. 11). Therefore, a threshold value of $>0^3$ will be used with STP to discriminate between environments that are prohibitive for tornado development and those that are more favorable.

4. Discussion and summary

This analysis of radar signatures associated with tornado warnings and missed tornado events provides a robust dataset, which can be used to estimate tornado warning verification characteristics based on thresholds of V_{rot} . If the desired goal is to increase POD, increases in FAR can be predicted based on how much the V_{rot} threshold is lowered. Conversely, if reducing FAR is the goal, the lowering of POD can likewise be predicted based on the thresholds of V_{rot} . In addition to V_{rot} , measures of the near-storm environment can be combined with the radar characteristics to improve warning performance.

Actual NWS warning performance differs from our estimates in that warning issuance depends on more than just

³The more familiar STP threshold of 1, dating to Thompson et al. (2003), was meant to identify EF2+ tornadoes, while the majority of tornado warnings and associated tornadoes are weaker (EF0–1).

TABLE 4. Table showing multiple SPC mesoanalysis parameters associated with tornadoes and tornado warnings from 2016 to 2018 listed in order from those that show the greatest to the least reduction in FAR between the 10th and 90th percentiles of each parameter. The “ML” prefix denotes the use of the lowest 100-hPa mean lifted parcel. SRH calculations used the supercell motion estimate from Bunkers et al. (2014). Of the three percentile rank values displayed, the lowest FAR for each parameter is set in bold font.

Parameter	FAR			
	10th percentile	50th percentile	90th percentile	10th–90th difference
STP	0.78	0.71	0.59	−0.19
SCP	0.75	0.71	0.60	−0.15
Effective bulk shear	0.76	0.70	0.64	−0.12
Effective SRH	0.75	0.64	0.65	−0.10
700–500-hPa lapse rate	0.74	0.68	0.65	−0.09
MLCAPE	0.71	0.69	0.62	−0.09
0–3-km bulk shear	0.75	0.67	0.67	−0.08
0–3-km MLCAPE	0.75	0.67	0.67	−0.08
0–3-km lapse rate	0.73	0.69	0.67	−0.06
MLCIN	0.73	0.68	0.68	−0.05
DCAPE	0.68	0.72	0.66	−0.02
Dewpoint depression	0.73	0.66	0.72	−0.01
MLLCL	0.72	0.66	0.72	0.00
PWAT	0.67	0.65	0.75	+0.08

V_{rot} , especially when considering warning lead time. For realistic tornado warning performance estimates, the number of potential false alarm cases (or correct null cases) must be known. Unfortunately, there is no known, comprehensive documentation of all rotating storms that do not produce tornadoes. The work of T17 did address this concern by calculating tornado frequency of occurrence as a function of V_{rot} , at least for reported severe storms in the contiguous United States. We used the independent sample from T17 to estimate the frequency of nontornadic, severe storms with cyclonic rotation in the low levels, instead of the overwhelming task of examining radar imagery for all thunderstorms during 2016–18. The combined approach of examining all 2016–18 tornadoes and warnings, with the T17 estimates of potential false alarms, allows for robust, reproducible estimates of POD and FAR for tornado warnings based on uniform application of V_{rot} thresholds.

During the period 2016–18, NWS tornado warnings were associated with a POD = 0.62 and a FAR = 0.69, resulting in a CSI = 0.26. The most skillful V_{rot} threshold was 30 kt, which matches well with current WDTD guidance (Fig. 1). Therefore, no overall recommendations for change seem necessary, as any change in current warning practices to either increase POD or reduce FAR would likely have a direct positive (negative) impact on the opposite variable. However, greater attention to the tails of the distribution could both increase POD and decrease FAR. These tail cases, which should receive greater attention, are as follows.

- 1) Storms with a $V_{\text{rot}} \geq 30$ kt and STP ≥ 0 should strongly be considered for a warning. If tornado warnings were issued anytime these criteria were met, approximately 292 additional tornado warnings and 109 fewer missed events would be expected each year, with associated POD = 0.68 and FAR = 0.63—both of which are improvements on the overall 2016–18 database numbers (Fig. 12).
- 2) Tornado warnings should be avoided for storms with a $V_{\text{rot}} < 30$ kt, STP = 0, couplet diameter > 1 n mi, and no

credible report or TDS. If tornado warnings were not issued for these lower-criteria events, there would be approximately 124 fewer tornado warnings and 17 additional missed events expected each year. Tornado warnings issued in this part of the parameter space result in FAR = 0.86, which is notably higher than the database average (Fig. 13).

Making these slight modifications could increase the POD to 0.72 and decrease the FAR to 0.67 for an overall CSI increase to ~ 0.29 . In addition, this would lead to 277 fewer missed events with 116 fewer EF0 misses, 147 fewer EF1 misses, 13 fewer EF2 misses, and 1 fewer EF3 miss. The 17 additional missed events are likely an overestimate as some of these circulations likely strengthened beyond the original <30-kt V_{rot} during the life of the warning. It is important to note the fewer missed tornado numbers are likely unrealistic with the assumption that a tornado warning could be issued at the exact minute a V_{rot} reaches 30+ kt, which is not always feasible. However, even if more than half of these previously missed events could be warned, an improvement in overall POD can be accomplished. The V_{rot} threshold proposed here matches recent WDTD guidance (Fig. 1), and additional specific environment/diameter thresholds included here could be added to this guidance to further refine the initial tornado warning criteria. Additional factors that may also improve tornado warning performance:

- Proper diagnosis of velocity data
- Proper TDS diagnosis
- More rigorous vetting of tornado/funnel cloud reports
- Longer tornado warning lead times for long-track/strong tornadoes where downstream persistence is anticipated.

The quality of radar data plays a primary role in warning performance, in the absence of corroborating tornado reports. From T17 (their Fig. 7), the probability of a tornado increases dramatically as peak V_{rot} increases from ~ 30 to ~ 50 kt. The T17 sample excluded obvious cases of velocity

2016-2018 Tornado Warning Performance by STP threshold

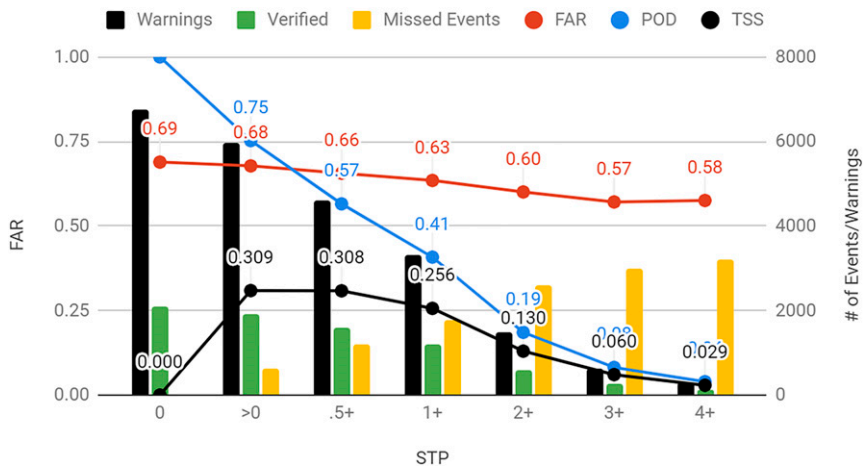
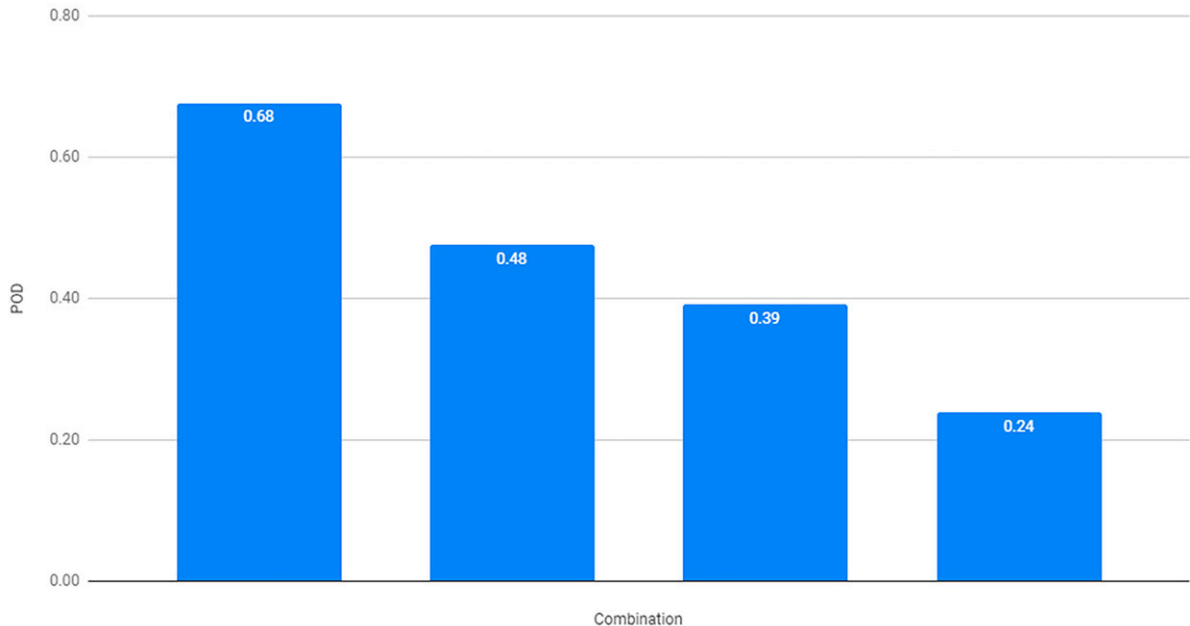


FIG. 11. Tornado warning FAR (red) by binned values of STP. Bar charts show all tornado warnings (black), verified tornado warnings (green), and missed events (yellow).

contamination, which is important since common problems like sidelobe contamination can drive erroneously high V_{rot} calculations, and resultant overestimates of tornado probabilities. FAR can be reduced substantially by identifying and

eliminating warnings associated with velocity contamination, which also lack any other corroborating evidence of a tornado (per Fig. 10). Identification of sidelobe velocity contamination is a first step in tempering warning decisions

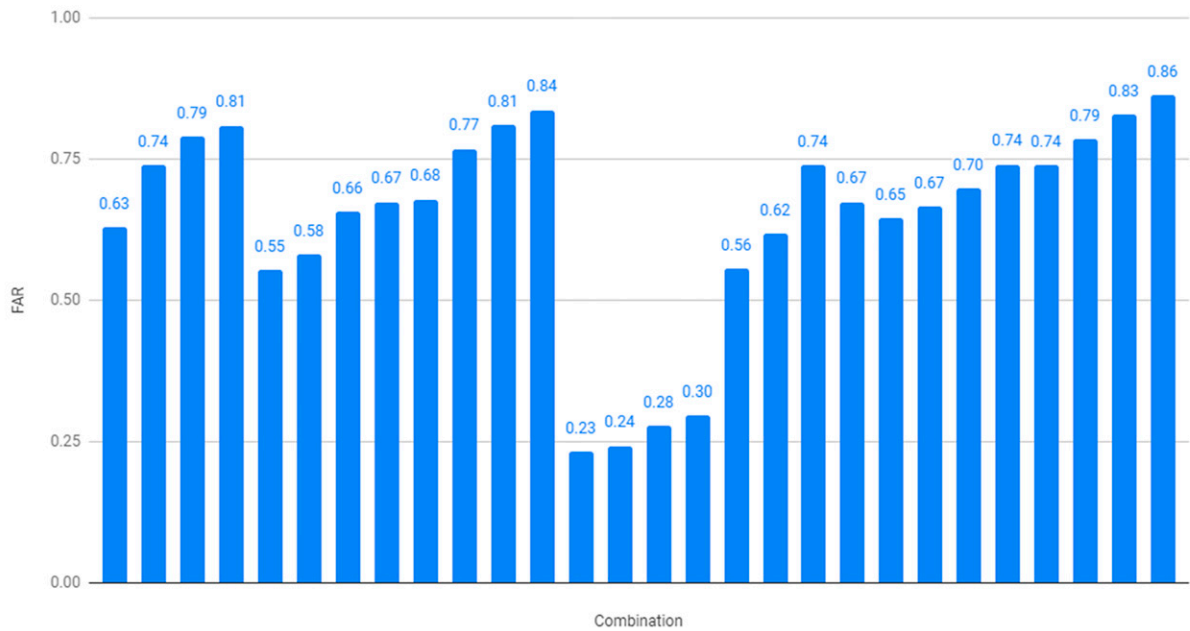
POD when combining thresholds of V_{rot} and STP



POD	0.68	0.48	0.39	0.24
V_{rot} (kts)	>30	<30	>30	<30
STP	>0	>0	0	0

FIG. 12. POD (blue bars) for different combined thresholds of V_{rot} and STP (chart below colored with green as thresholds that are typically more favorable for tornadoes and red for variables that are typically less favorable), showing POD is maximized for $V_{rot} > 30$ kt and STP > 0 .

FAR when combining thresholds of Vrot, STP, Circulation Diameter, and Report



FAR	0.63	0.74	0.79	0.81	0.55	0.58	0.66	0.67	0.68	0.77	0.81	0.84	0.23	0.24	0.28	0.30	0.56	0.62	0.74	0.67	0.65	0.67	0.70	0.74	0.74	0.79	0.83	0.86
Vrot (kts)	>30	>30	<30	<30	>30	<30	>30	>30	<30	>30	<30	<30	>30	<30	>30	>30	<30	<30	<30	>30	<30	>30	<30	>30	<30	>30	<30	<30
STP	>0	0	>0	0	>0	>0	0	0	0	>0	0	>0	>0	0	>0	0	>0	0	0	>0	>0	>0	0	0	0	0	>0	0
Diameter (nm)			<1	<1	>=1	<1	<1	>=1	>=1	>=1	<1	<1	<1	>=1	>=1	>=1	<1	>=1	<1	<1	>=1	<1	<1	>=1	<1	>=1	>=1	>=1
Report													yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no

FIG. 13. FAR (blue bars) for different combined thresholds of V_{rot} , STP, circulation diameter, and report yes/no (chart below with green as thresholds that are typically more favorable for tornadoes and red for variables that are typically less favorable), showing FAR is the greatest for $V_{rot} < 30$ kt, STP = 0, diameter ≥ 1 n mi, and no report.

that are based primarily on radar data from the low levels of a supercell. To maintain lead time, warning decisions (in the face of contaminated velocity data) may need to rely more heavily on the near-storm environment (such as STP) and longer-term expectations of tornado potential (i.e., high tornado probabilities in convective outlooks and tornado watches, versus low tornado probabilities and severe thunderstorm (or no) watches (Krocak and Brooks 2021).

Likewise, the source of a “confirmed” tornado is quite important when considering how the sources influence warning outcomes. Most tornadoes are short lived, the appearance of a TDS in radar data can be delayed (~2–3 min after tornadogenesis, on average, if sufficient debris is present to be lofted), and the time required to issue a tornado warning can result in both a missed event and a false alarm. Figures 14 and 15 show the expected FAR if a tornado warning were issued for all missed events with a TDS from 2016 to 2018. These were grouped in 1-min bins up to 5-min post-TDS detection to identify how much time a warning forecaster has after initial TDS detection to issue a warning without substantially increasing the FAR. Figure 15 shows that a warning should be issued for all TDSs (even 4–5 min after first appearance) if the environment is favorable for significant tornadoes (STP ≥ 1). However, in weaker

environments (STP < 1), the warning should be issued within the first 2–3 min before FAR starts to increase significantly (Fig. 14). The FAR was high for tornado warnings following TDS detection with weak (< 20 kt) accompanying V_{rot} , since these tornadoes tended to dissipate prior to or soon after TDS detection and prior to warning issuance.

By focusing on opportunities to issue/not issue tornado warnings on the tails of the distribution based on above guidance, improvements in tornado warning performance are possible with an overall improvement in POD, reduction in FAR, and ~90 fewer missed events per year. The perception of warning verification is necessarily a multi-dimensional problem with more nuance than just simple “hits,” “misses,” or “false alarms” (Barnes et al. 2007). Also, tornado warning lead time, especially with the initial warnings for a particular storm, requires the addition of human expertise in the areas of full volumetric radar interpretation, and comparison of the radar structures with expectations based on storm mode and the near-storm environment, as discussed in Gibbs and Bowers (2019).

Beyond the presence of a TDS or credible tornado reports, radar and environmental characteristics can be combined to provide a recipe for the most effective tornado warning verification. The highest POD tornado warnings are associated

Tornado warning FAR if warning issued "x" minutes after TDS (STP<1)

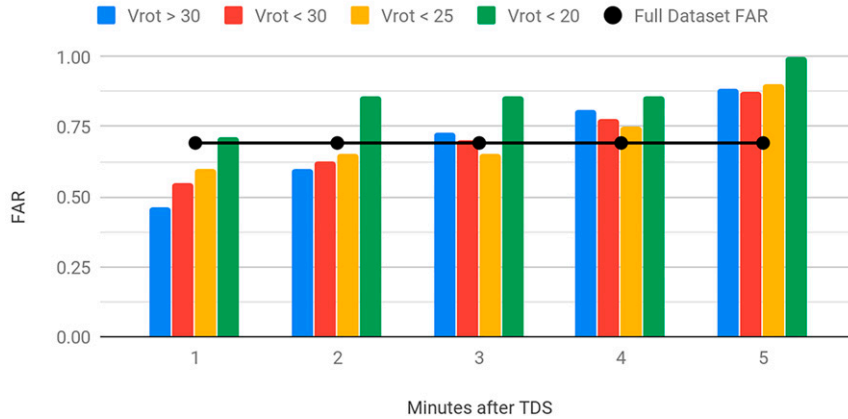


FIG. 14. FAR per time after TDS appearance that warning was issued binned into <20 kt (green), <25 kt (yellow), <30 kt (red), and greater than 30 kt (blue) for STP less than 1. Full dataset FAR in black.

with larger (>30 kt) V_{rot} , a tight circulation (<1 n mi diameter) observed in the low levels, in an environment with the $STP > 0$. This recipe for above average POD tornado warnings is consistent with the work of Brotzge et al. (2013) and Anderson-Frey et al. (2016, 2019), where POD for tornadoes is greatest in environments with larger buoyancy and vertical wind shear. The highest FAR tornado warnings are characterized by weak V_{rot} (<30 kt), a broad circulation (>1 n mi diameter) observed far from the radar, in an environment with the $STP = 0$. The majority of missed tornadoes (that are generally weak) and false alarms occur in more marginal near-storm environments (e.g., Anderson-Frey et al. 2016, 2019). Tornado warning false alarms are most common on

days when tornadoes do not occur (Brotzge et al. 2011), and there will continue to be challenges posed to NWS meteorologists in warning for lone tornado events (Brotzge and Erickson 2009, 2010)

Overall, this study found that current NWS warning practices and WDTD criteria closely match the most skillful statistical results. However, additional improvements in FAR and POD are possible by focusing on more frequent warnings on strong circulations in favorable environments, and less frequent warnings on weak, broad circulations in poor environments. While fewer missed events has an obvious positive impact, it can also be beneficial for NWS forecasters to keep FAR from becoming too large, as large

Tornado warning FAR if warning issued "x" minutes after TDS (STP>1)

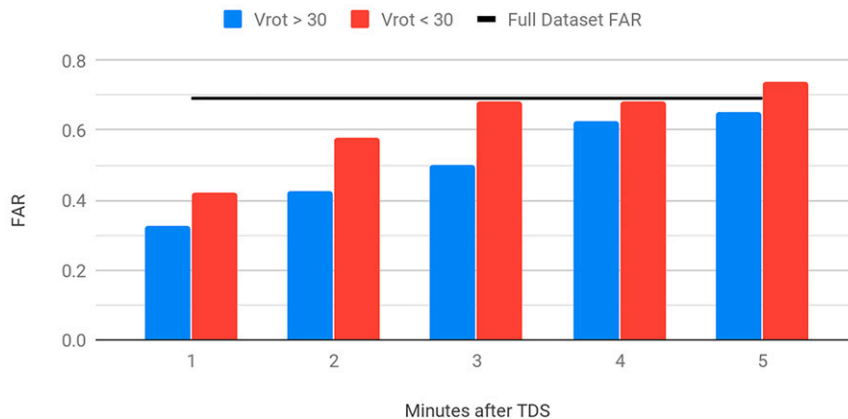


FIG. 15. FAR per time after TDS appearance that warning was issued binned into <20 and <25 kt (not shown, sample size too small), <30 kt (red), and >30 kt (blue) with $STP > 1$. Full dataset FAR in black.

FAR can reduce the credibility of NWS tornado warnings, and may have undesirable effects on warning responses and resulting casualties (Simmons and Sutter 2009; NWS 2011; Ripberger et al. 2015).

Acknowledgments. The authors thank Israel Jirak (SPC Science and Operations Officer) for his helpful suggestions to clarify and strengthen our presentation, and Bill Bunting (SPC Operations Branch Chief) for encouraging our pursuit of this project. In addition, special thanks are extended to Andy Dean (SPC Techniques Development Meteorologist) who collected SPC mesoanalysis data for these cases. Also, thank you to the anonymous reviewer who provided the performance diagram to better display the results in Fig. 6.

REFERENCES

- Anderson-Frey, A., Y. Richardson, A. Dean, R. Thompson, and B. 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 —, 2019: Characteristics of tornado events and warnings in the southeastern United States. *Wea. Forecasting*, **34**, 1017–1034, <https://doi.org/10.1175/WAF-D-18-0211.1>.
- Barnes, L., E. Grunfest, M. Hayden, D. Schultz, and C. Benight, 2007: False alarms and close calls: A conceptual model of warning accuracy. *Wea. Forecasting*, **22**, 1140–1147, <https://doi.org/10.1175/WAF1031.1>; Corrigendum, **24**, 1452–1454, <https://doi.org/10.1175/2009WAF2222300.1>.
- Blair, S. F., and J. W. Leighton, 2014: Assessing real-time tornado information disseminated through NWS products. *Wea. Forecasting*, **29**, 591–600, <https://doi.org/10.1175/WAF-D-13-00126.1>.
- Bodine, D. J., M. R. Kumjian, R. D. Palmer, P. L. Heinselman, and A. V. Ryzhkov, 2013: Tornado damage estimation using polarimetric radar. *Wea. Forecasting*, **28**, 139–158, <https://doi.org/10.1175/WAF-D-11-00158.1>.
- 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. *21st Conf. on Severe Local Storms/19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., JP3.1, https://ams.confex.com/ams/SLS_WAF_NWP/techprogram/paper_47482.htm.
- 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>.
- , and W. Donner, 2013: The tornado warning process. A review of current research, challenges, and opportunities. *Bull. Amer. Meteor. Soc.*, **94**, 1715–1733, <https://doi.org/10.1175/BAMS-D-12-00147.1>.
- , S. Erickson, and H. Brooks, 2011: A 5-yr climatology of tornado false alarms. *Wea. Forecasting*, **26**, 534–544, <https://doi.org/10.1175/WAF-D-10-05004.1>.
- , S. Nelson, R. Thompson, and B. 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>.
- Brown, R. A., V. T. Wood, and D. Sirmans, 2002: Improved tornado detection using simulated and actual WSR-88D data with enhanced resolution. *J. Atmos. Oceanic Technol.*, **19**, 1759–1771, [https://doi.org/10.1175/1520-0426\(2002\)019<1759:ITDUSA>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<1759:ITDUSA>2.0.CO;2).
- , B. A. Flickinger, E. Forren, D. M. Schultz, D. Sirmans, P. L. Spencer, V. T. Wood, and C. L. Ziegler, 2005: Improved detection of severe storms using experimental fine-resolution WSR-88D measurements. *Wea. Forecasting*, **20**, 3–14, <https://doi.org/10.1175/WAF-832.1>.
- Bunkers, M. J., D. A. Barber, R. L. Thompson, R. Edwards, and J. Garner, 2014: Choosing a universal mean wind for supercell motion prediction. *J. Oper. Meteor.*, **2**, 115–129, <https://doi.org/10.15191/nwajom.2014.0211>.
- Chrisman, J. N., 2011: Supplemental Adaptive Intra-Volume Low-Level Scan (SAILS). NOAA/NWS, 13 pp., http://www.roc.noaa.gov/wsr88d/PublicDocs/NewTechnology/SAILS_Initial_Presentation_Sep_2011.pdf.
- , 2014: Multiple elevation scan option for SAILS (MESO-SAILS). NOAA/NWS, 27 pp., http://www.roc.noaa.gov/wsr88d/PublicDocs/NewTechnology/MESO-SAILS_Description_Briefing_Jan_2014.pdf.
- Doviak, R. J., and D. S. Zrnić, 1993: *Doppler Radar and Weather Observations*. Academic Press, 562 pp.
- Edwards, R., and J. C. Picca, 2016: Tornadoic debris signatures in tropical cyclones. Preprints, *28th Conf. Severe Local Storms*, Portland, OR, Amer. Meteor. Soc., 162, <https://ams.confex.com/ams/28SLS/webprogram/Paper300633.html>.
- Ferree, J. T., J. M. Looney, and K. R. Waters, 2006: NOAA/National Weather Services' storm-based warnings. *23rd Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., 11.6, https://ams.confex.com/ams/23SLS/techprogram/paper_115513.htm.
- Gibbs, J. G., 2016: A skill assessment of techniques for real-time diagnosis and short-term prediction of tornado intensity using the WSR-88D. *J. Oper. Meteor.*, **4**, 170–181, <https://doi.org/10.15191/nwajom.2016.0413>.
- , and B. R. Bowers, 2019: Techniques and thresholds of significance for using WSR-88D velocity data to anticipate significant tornadoes. *J. Oper. Meteor.*, **7**, 117–137, <https://doi.org/10.15191/nwajom.2019.0709>.
- Karstens, C. D., and Coauthors, 2018: Development of a human-machine mix for forecasting severe convective events. *Wea. Forecasting*, **33**, 715–737, <https://doi.org/10.1175/WAF-D-17-0188.1>.
- Kingfield, D. M., and J. G. LaDue, 2015: The relationship between automated low-level velocity calculations from the WSR-88D and maximum tornado intensity determined from damage surveys. *Wea. Forecasting*, **30**, 1125–1139, <https://doi.org/10.1175/WAF-D-14-00096.1>.
- Krocak, M. J., and H. E. Brooks, 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>.
- NWS, 2011: NWS Central Region service assessment. NOAA/NWS, Joplin, MO, accessed 17 June 2019, <https://www.weather.gov/publications/assessments>.
- Piltz, S. F., and D. W. Burgess, 2009: The impacts of thunderstorm geometry and WSR-88D beam characteristics on diagnosing supercell tornadoes. *34th Conf. on Radar Meteorology*, Williamsburg,

- VA, Amer. Meteor. Soc., P6.18, https://ams.confex.com/ams/34Radar/techprogram/paper_155944.htm.
- Ripberger, J. T., C. L. Silva, H. C. Jenkins-Smith, D. E. Carlson, M. James, and K. G. Herron, 2015: False alarms and missed events: The impact and origins of perceived inaccuracy in tornado warning systems. *Risk Anal.*, **35**, 44–56, <https://doi.org/10.1111/risa.12262>.
- Ryzhkov, A., T. J. Schuur, D. W. Burgess, and D. S. Zrnić, 2005: Polarimetric tornado detection. *J. Appl. Meteor.*, **44**, 557–570, <https://doi.org/10.1175/JAM2235.1>.
- Saxion, D. S., and R. L. Ice, 2012: New science for the WSR-88D: Status of the dual-polarization upgrade. *28th Conf. on Interactive Information Processing Systems*, New Orleans, LA, Amer. Meteor. Soc., 5, https://ams.confex.com/ams/92Annual/webprogram/Manuscript/Paper197645/NEXRAD_DP_Status_28th_IIPS_Jan2012.pdf.
- Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575, [https://doi.org/10.1175/1520-0434\(1990\)005<0570:TCSIAA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2).
- Schneider, R. S., and A. R. Dean, 2008: A comprehensive 5-year severe storm environment climatology for the continental United States. *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., 16A.4, <http://ams.confex.com/ams/pdfpapers/141748.pdf>.
- Schultz, C. J., and Coauthors, 2012a: Dual-polarization tornadic debris signatures Part I: Examples and utility in an operational setting. *Electron. J. Oper. Meteor.*, **13**, 120–137, <http://nwafiles.nwas.org/ej/pdf/2012-EJ9.pdf>.
- , and Coauthors, 2012b: Dual-polarization tornadic debris signatures Part II: Comparisons and caveats. *Electron. J. Oper. Meteor.*, **13**, 138–150, <http://nwafiles.nwas.org/ej/pdf/2012-EJ10.pdf>.
- Simmons, K. M., and D. Sutter, 2009: False alarms, tornado warnings, and tornado casualties. *Wea. Climate Soc.*, **1**, 38–53, <https://doi.org/10.1175/2009WCAS1005.1>.
- Smith, B. T., R. L. Thompson, J. S. Grams, and J. C. Broyles, 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>.
- , —, A. R. Dean, and P. T. Marsh, 2015: Diagnosing the conditional probability of tornado damage rating using environmental and radar attributes. *Wea. Forecasting*, **30**, 914–932, <https://doi.org/10.1175/WAF-D-14-00122.1>.
- , —, D. A. Speheger, A. R. Dean, C. D. Karstens, and A. K. Anderson-Frey, 2020a: WSR-88D tornado intensity estimates. Part I: Real-time probabilities of peak tornado wind speeds. *Wea. Forecasting*, **35**, 2479–2492, <https://doi.org/10.1175/WAF-D-20-0010.1>.
- , —, —, —, —, and —, 2020b: WSR-88D tornado intensity estimates. Part II: Real-time applications to tornado warning timescales. *Wea. Forecasting*, **35**, 2493–2506, <https://doi.org/10.1175/WAF-D-20-0011.1>.
- Snyder, J., and A. V. Ryzhkov, 2015: Automated detection of polarimetric tornadic debris signatures using a hydrometeor classification algorithm. *J. Appl. Meteor. Climatol.*, **54**, 1861–1870, <https://doi.org/10.1175/JAMC-D-15-0138.1>.
- Thompson, R., R. Edwards, J. Hart, K. 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](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: Supercells and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, <https://doi.org/10.1175/WAF-D-11-00116.1>.
- , and Coauthors, 2017: Tornado damage rating probabilities derived from WSR-88D data. *Wea. Forecasting*, **32**, 1509–1528, <https://doi.org/10.1175/WAF-D-17-0004.1>.
- Torres, S., and C. Curtis, 2007: Initial implementation of super-resolution data on the NEXRAD network. *23rd Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, San Antonio, TX, Amer. Meteor. Soc., 5B.10, <https://ams.confex.com/ams/87ANNUAL/webprogram/Paper116240.html>.
- Toth, M., R. J. Trapp, J. Wurman, and K. A. Kosiba, 2013: Comparison of mobile-radar measurements of tornado intensity with corresponding WSR-88D measurements. *Wea. Forecasting*, **28**, 418–426, <https://doi.org/10.1175/WAF-D-12-00019.1>.
- Van Den Broeke, M. S., and S. T. Jauernic, 2014: Spatial and temporal characteristics of polarimetric tornadic debris signatures. *J. Appl. Meteor. Climatol.*, **53**, 2217–2231, <https://doi.org/10.1175/JAMC-D-14-0094.1>.
- Wagenmaker, R., G. Mann, and M. Hudson, 2014: NWS Central Region Impact-Based Warning project introduction and overview. NOAA, 17 pp.
- Warning Decision Training Division, 2021: Warning content: Impact-based warnings. Accessed 19 May 2021, https://training.weather.gov/wdtd/courses/rac/warnings/warn-content/presentation_html5.html.
- Wilks, S. S., 1935: The likelihood test of independence in contingency tables. *Ann. Math. Stat.*, **6**, 190–196, <https://doi.org/10.1214/aoms/117732564>.
- Zrnić, D. S., 1987: Three-body scattering produces precipitation signature of special diagnostic value. *Radio Sci.*, **22**, 76–86, <https://doi.org/10.1029/RS022i001p00076>.