Baerg, B. M., W. P. Gargan, A. E. Cohen, R. L. Thompson, B. T. Smith, A. E. Gerard, C. J. Schultz, L. A. Kelly, and H. V. Nepaul, 2020: Radar-based, storm-scale circulation and tornado-probability tendencies preceding tornadogenesis in Kansas and Nebraska. *Electronic J. Severe Storms Meteor.,* **15** (3), 1–23.

Electronic Journal of
SEVERE STORMS METEOROLOGY

Radar-Based, Storm-Scale Circulation and Tornado-Probability Tendencies Preceding Tornadogenesis in Kansas and Nebraska

BRYAN M. BAERG and WILLIAM P. GARGAN
NOAA/National Weather Service, Topeka, Kansas

ARIEL E. COHEN
NOAA/National Weather Service, Pueblo, Colorado

RICHARD L. THOMPSON and BRYAN T. SMITH
NOAA/NWS/NCEP/Storm Prediction Center, Norman, Oklahoma

ALAN E. GERARD
NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

CHAUNCY J. SCHULTZ
NOAA/National Weather Service, Bismarck, North Dakota

LAWRENCE A. KELLY and HEATHER V. NEPAUL
NOAA/National Weather Service, Miami, Florida

(Submitted 15 December 2019; in final form 4 August 2020)

ABSTRACT

This study analyzes the behavior of storm-scale circulations preceding initial tornadogenesis in 179 Kansas and Nebraska storms. Manually determined assessments of radar data for storm-scale circulations preceding the tornadoes are performed as far back in time prior to the tornado as a circulation is apparent, with average rotational velocity (V_{rot}), circulation diameter, and circulation clarity documented for the 0.5° elevation scan. These data are simultaneously combined with an indication of environmental conditions (as represented by the significant tornado parameter) to determine the tornado probability at each of these times based on a recently developed probabilistic model. By aggregating these parameters in time-range bins, subsequent statistical analyses portray the bulk variability of circulation characteristics and tornado probabilities preceding tornadogenesis. The blended approach for assessing tornado potential yields a stronger relative increase in tornado probabilities leading up to tornadogenesis than the sub-component of average V_{rot}; this is especially true within 15 min before tornadogenesis. Additionally, significant tornadoes are associated with more substantial increase in tornado probabilities preceding tornadogenesis compared to weak tornadoes, and smaller lead time to tornadogenesis for weak tornadoes. Also, a cycling pattern may appear in velocities prior to significant tornadoes, along with a relative decrease in pretornadic circulation diameter, especially for significant tornadoes. These findings are intended to highlight some of the behaviors of storm-scale circulations and their corresponding environments, which can be used to reinforce meteorologists' tornado threat assessment. Extending this work to encompass more convective-mode variability, null cases, and geographic expanse will be necessary for more overarching applicability.

---

1. Introduction

The past several years have featured a proliferation of research that addresses the attributes of storm-scale circulations associated with potentially high-impact convective weather. Within a large dataset of tornado and significant severe thunderstorm events from 2003–2011, Smith et al. (2012) and Thompson et al. (2012) thoroughly analyzed WSR-88D-based rotational velocity (V_{rot}) signatures associated with storm-scale circulations producing these events. In addition to quantifying the magnitude and size of these circulations, they incorporated other factors into their analysis to more fully resolve the context of these circulations: the effective-layer significant tornado parameter (STP; Thompson et al. 2007) to characterize the near-storm environment and manual classifications of the convective mode (Smith et al. 2012) typifying the parent convection. Since those works laid the foundation for contextualizing radar-based storm-scale attributes, a natural follow-up study by Smith et al. (2015) derived conditional probabilities of specific tornado ratings given a tornado, based on a large sample of tornadoes that occurred from 2009–2013. This was one of the first known attempts to link base radar data with both environmental conditions and tornado reports in the formulation of conditional tornado probabilities, effectively providing reproducible guidance for meteorologists to communicate tornado impacts diagnostically.

Building upon the aforementioned studies, Thompson et al. (2017) incorporated radar and environmental attributes associated with both tornadic and nontornadic convection to determine tornado probabilities. They investigated the following variables as separate constraints to derive these probabilities: peak low-level V_{rot}, height above radar level (ARL), circulation diameter, a subjectively determined "clear and/or tight" characterization of the circulation's appearance, the presence or absence of a dual-polarization (dual-pol) tornadic debris signature (TDS; Ryzhkov et al. 2005), along with STP. These variables are shown to explain differences in tornado-rating probabilities; Thompson et al. (2017) identified the relationship between each variable and tornado-rating probability, with each variable being investigated separately. For instance, they found that larger magnitudes of average V_{rot} and smaller diameters of the circulations were associated with higher tornado probabilities. A notable relationship appears between STP and tornado probability (i.e., rating of at least EF0).

The work leading up to and including Thompson et al. (2017) analyzed relationships between individual explanatory variables and tornado ratings. Subsequently, Cohen et al. (2018) identified a statistical method for mutually combining the variables that Thompson et al. (2017) documented, into deriving tornado-probability and tornadic windspeed approximations. Using a linear regression model incorporating multiple variables, Cohen et al. (2018) found predictors best related to tornado rating: height of the circulation observation ARL, average V_{rot}, STP, and the presence or absence of a TDS. Moreover, they used a binary logistic regression model to estimate tornado probabilities, with average V_{rot}, circulation diameter, subjectively determined circulation-appearance assessment, and STP best explaining tornado probability. The formulation of their statistical model follows in Eq. (1), whereby y corresponds to the predicted tornado probability, b corresponds to an intercept parameter, and each a_i represents a regression coefficient associated with predictor x_i offering the greatest amount of explanatory power for determining tornado potential:

$$y = \frac{1}{1+exp\{-[(\sum_{i=1}^{m} a_i x_i)+b]\}}. \qquad (1)$$

These coefficients are 0.0552 for average V_{rot} (measured in kt), –0.684 for circulation diameter (measured in nmi), 0.835 for the subjectively determined circulation-appearance assessment (unitless), and 0.0473 for STP (unitless), while the intercept parameter b is –2.51 (unitless).

By providing the predictor inputs corresponding to the variables on the right-hand-side of Eq. (1), which can be determined directly from 0.5° radar data and environmental conditions accessed from the hourly updating Storm Prediction Center mesoscale analysis system (Bothwell et al. 2002), a tornado probability can be calculated [i.e., the predictand on the left-hand side of Eq. (1)]. By employing the concept of radar teams in an operational setting, these predictor inputs can be assessed quickly by a mesoanalyst or radar assistant. This

---

Corresponding author address: Bryan Baerg, National Weather Service Topeka, KS, E-mail: Bryan.Baerg@noaa.gov

individual could enter these inputs into an interface performing mathematical calculations to compute the corresponding tornado probability, which could be shared with the primary radar operator. Such input may be valuable to performing tornado-threat assessment. While the assessment of the clearness and/or tightness of the circulation is subjective, it can be performed quickly by a radar interrogator. Thompson et al. (2017) distinguished between a “clear and/or tight” circulation and “nebulous and/or diffuse” based upon the meteorologist’s ability to identify maximum outbound and maximum inbound velocities comprising the storm-scale circulation. They suggested that if substantial effort is needed to identify these velocity components, then the “nebulous and/or diffuse" assessment applies (a binary value of “0”). This subjective assessment is inherent to the determination of other inputs for Eq. (1), and thus should require minimal time in incorporating into the tornado-probability assessment.

The Cohen et al. (2018) tornado probability model suggests that, among an independent sample of tornadic and nontornadic severe-thunderstorm cases, the upper extent of the interquartile range for nontornadic severe-thunderstorm cases nearly matches the lower extent of the interquartile range for tornado cases. Specifically, the 75th-percentile tornado probability among nontornadic events is 19.5%, whereas the 25th-percentile tornado probability among nontornadic events is 16.3%. This separation between tornado and non-tornado probability distributions suggests that the tornado-probability model can assist with identifying ongoing tornado potential accompanying rotating storms, especially given the robust sample size and multivariate foundation.

An investigation of the reliability of the calibration of the Cohen et al. (2018) tornado probability model has not been performed. However, this model benefits from a very large sample size, consisting of individual radar-based and human-analyzed storm-scale circulations that have been linked directly to severe-weather reports (analogous to considerations and operational practice integrated into the warning-decision-making process). The model also shows promise in distinguishing between tornadic and nontornadic events.

The work of Thompson et al. (2012), Smith et al. (2012), Smith et al. (2015), Thompson et al. (2017), and Cohen et al. (2018) provides a reproducible methodology for quantifying the present state of tornado threat and the potential strength of an associated tornado. The results from these studies can serve as diagnostic input for the meteorologist to communicate potential hazards accompanying ongoing circulations. The foundation for this input is a large and diverse sample size of past supercells and their severe-weather reports. More recently, Gibbs and Bowers (2019) present volumetric, radar-based methods for identifying forthcoming significant tornado (EF2+) events from supercells and quasilinear convective systems (QLCSs), quantifying the skill of particular signals.

The specific, aforementioned radar-based parameters are treated as proxies for tornado-threat assessment, as the WSR-88D rarely samples the processes related to the actual tornado circulation, except in cases when the mesocyclone is very close to the radome (e.g., Houser et al. 2015). Rather, the WSR-88D can detect processes aloft that can enhance or mitigate tornadogenesis via dynamic ascent (i.e., a supercell representing a perturbation pressure deficit).

Other tools exist to assist with the tornado-threat assessment process. For instance, the Cooperative Institute for Meteorological Satellite Studies (CIMSS) produces real-time severe-thunderstorm and tornado probabilistic guidance that incorporates radar, satellite, and environmental data, known as ProbSevere and ProbTor, respectively (Cintineo et al. 2018). Hart and Cohen (2016) address the Statistical Severe Convective Risk Assessment Model (SSCRAM), which produces individual severe-thunderstorm-hazard probabilities within a couple of hours into the future, based upon the storm environment. An interface also has been designed that can estimate tornado probabilities and wind speeds using the Cohen et al. (2018) models in real time (available online at http://arctic.som.ou.edu/tburg/products/R2O/torprob/).

Building upon these studies, the present work offers generalized analyses of the time-dependence of radar signatures leading to tornadogenesis. This is applied to a large number of cases, contextualized by the corresponding environment—effectively extending the work of Thompson et al. (2017) backward in time before tornado occurrence. This is intended to further help meteorologists identify forthcoming tornado potential based upon comparisons to past storm-scale circulation evolution. Compositing various components of storm-scale circulations, this work assesses the behavior of tornado probabilities, using the model developed by Cohen et al. (2018), to investigate the tornado-preceding evolution of a probabilistic parameter that simultaneously incorporates multiple aspects of the circulation.

The overall scope of previous research addressing WSR-88D data prior to tornadogenesis is an area of great potential growth. However, labor intensiveness of such work is a major challenge in its expansion—if using a method similar to Smith et al. (2012), Smith et al. (2015), and Thompson et al. (2017). In order to accomplish this work within a reasonable time, the dataset and statistical analysis for this preliminary study are confined to Kansas and Nebraska. However, the necessary conditions for tornadogenesis are geographically independent, thereby offering broader applicability. Nevertheless, this work motivates future studies that could investigate tornado-preceding parameter variability across a more geographically dispersive domain and including greater focus on convective-mode dependences. The approach followed herein substantiates communication of a more continuous delivery of pre-tornadic threat information, consistent with initiatives such as the Forecasting a Continuum of Environmental Threats (FACETs; Rothfusz et al. 2018) and probabilistic hazards information (PHI; Karstens et al. 2015).

Section 2 identifies the data and methods used to perform the present work. Results from the Kansas- and Nebraska-based pre-tornado analyses are presented in the form of statistical summaries in section 3, accompanied by an interpretation of the results integrated into a broader discussion. Conclusions regarding the entire scope of this work are provided in the final section 4.

2. Data and methodology

The process for determining storm-scale circulation attributes in the present study follows the procedure described by Thompson et al. (2017), which also provides the database used in the present work. Their dataset was filtered here for just Kansas and Nebraska. Maximum V_{rot} values were calculated within a maximum circulation diameter of 5 nmi (9.3 km), as in Thompson et al. (2017). Data were filtered further to ≤9900 ft (3 km) ARL, consistent with Thompson et al. (2017), and representative of sampling ranges of storm-scale circulations available to radar interrogators. Figure 1 depicts the relative frequency of observation heights within three layers below 9900 feet (3 km).

This yields 1788 assessments of storm-scale circulations from the 0.5° radar tilts preceding, and environmental conditions accompanying, 179 tornadoes (155 in Kansas and 24 in Nebraska), 11.7% of the Thompson et al. (2017) dataset size of 1530 tornadoes. No specific selection procedures were invoked in this study, and the distribution of tornado ratings is provided in Fig. 2. Official storm reports in final compiled form by the NCEI; NCEI (2018) were queried to determine the time of the first tornado associated with each convective element. Then, for each radar scan prior to the first tornado, WSR-88D scans at 0.5° were interrogated for as long as a circulation was evident as subjectively discernible as “clear and/or tight”, using the same radar regardless of its distance from the storm determined by Thompson et al. (2017). For each scan, the average V_{rot} was computed as follows:

$$V_{rot} = \frac{(V_{max} - V_{min})}{2}, \qquad (2)$$

where the maximum outbound and inbound storm-relative motion are represented by V_{max} and V_{min}, consistent with the work of Smith et al. (2015). Hereafter, average rotational velocity is referenced as V_{rot}. To ensure the maximum inbound and outbound velocities are associated with rotation as opposed to divergence or convergence signatures, the line connecting velocity extrema had to be within 45° of the line segment orthogonal to the beam centerline. In addition, circulation diameter and height ARL were determined for each scan. Only 11 cases (6.1%) of the 179 tornado events involved 0.5˚ radar data reaching over 9900 ft (3 km) ARL during the convective lifecycle. Any data above this level was omitted from subsequent analyses. Retained data then were combined at each time step prior to tornadogenesis to determine corresponding tornado probabilities based on the Cohen et al. (2018) probability model.

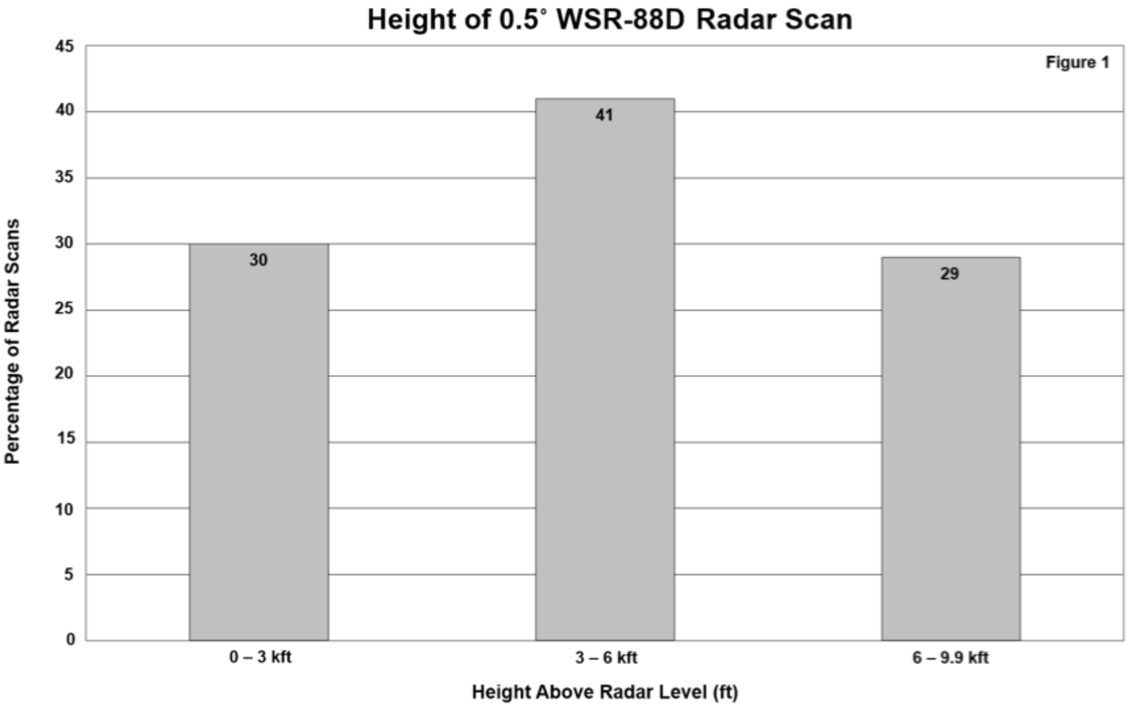


Figure 1: The height ARL for each scan binned into 3 ranges, with the corresponding percentage of the total data set (1788) labeled [e.g., 30% of the radar scans were <3000 ft (914 m) ARL].

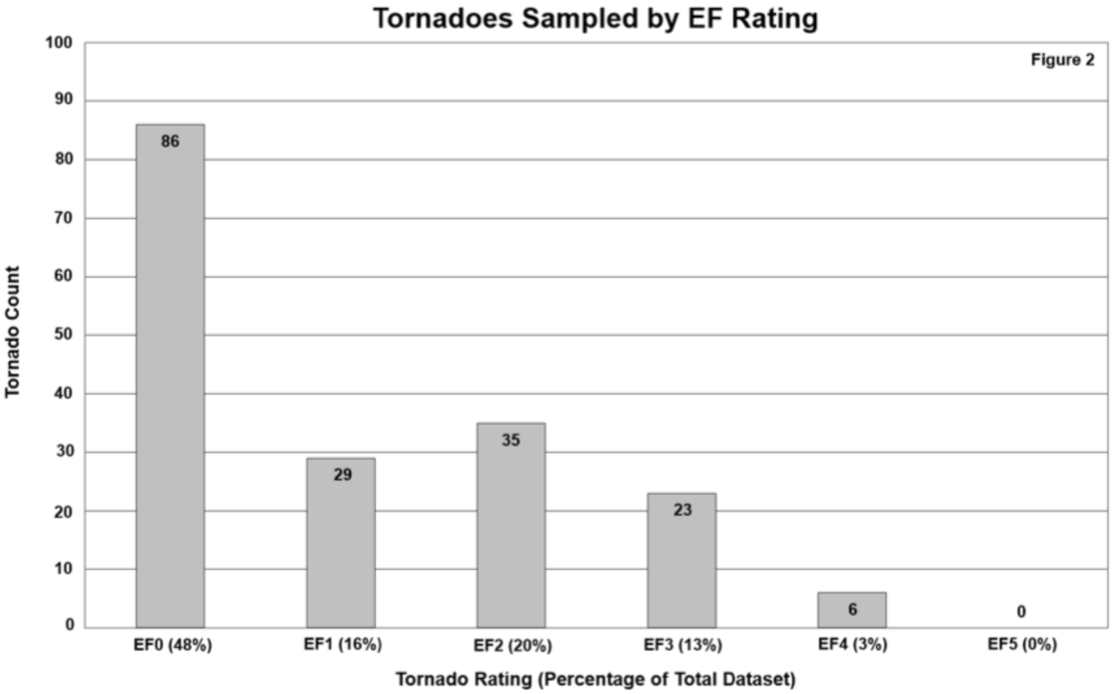


Figure 2: Distribution of the tornado dataset by EF rating. Sample size for each rating is listed above the bar in the chart, while the percent of each rating sample comprising the total Kansas/Nebraska tornado sample is listed below the *x*-axis.

The V_{rot} is dependent on sample averaging across a finite beam width, as well as the displacement between the circulation center and the center of the radar beam. While these are potential sources of inconsistencies among radar analyses, the large sample size of events studied is intended to represent the spectrum of such error inherent to operational radar assessments—rendering operational utility for performing tornado threat assessment. Also, these discrepancies would be inherent to real-time radar assessments, suggesting consistency between the data collection and operational practice.

The analysis of storm-scale circulation behavior prior to tornadogenesis is accomplished by investigating percentile ranks of V_{rot} and tornado probability. This is completed by first aggregating all V_{rot} and probability data within individual pre-tornado time-range bins. While such analysis does not account directly for individual-case pre-tornado behavior, it does permit a bulk evaluation of the relationships between pre-tornadic radar trends and mesoanalysis data. All distributions addressing time variance are plotted in the format of box-and-whisker diagram sequences, using 10th, 25th, 50th, 75th, and 90th percentiles. Moreover, within 30 min preceding tornadogenesis, the individual distributions by time were divided into 5-min increments. Meanwhile, for 30–60 min preceding tornadogenesis, the individual time-range bins were divided into 15-min increments, to account for the smaller sample sizes per 5-min bins long preceding tornadogenesis. To begin addressing the variability among individual cases, this study also analyzes the continuity of increase in tornado probabilities.

While STP is integrated into the tornado probability assessment, this study does not account for the variability of inflow-characterizing STP during the lifetime of the circulation preceding tornadogenesis. It merely treats the circulation-characterizing STP value from the Thompson et al. (2017) dataset as a constant throughout the duration of the circulation. In fact, Parker (2014) indicates STP that characterizes the distant inflow of tornadic storms varies by 3.1 as a composite value; the STP attributed to long-lasting pre-tornadic circulations may not necessarily characterize the actual near-storm environments around tornado occurrence (e.g., localized terrain effects on the low-level inflow, resulting in a more-favorable tornadic environment within the localized area). These factors mean that the present approach of applying a constant STP value to the lifecycle of the analyzed circulation imperfectly represent the near-storm environment. However, this methodology at least provides a rough approximation of the kinematic and thermodynamic properties of the inflow region, and is consistent with the documented environment around the time of tornado development. Errors associated with misrepresentations of near-storm environments for individual cases can become diluted by both large sample size and the blended approach to tornado-probability assessment.

Other sources of error or asymmetric sampling in this study stem from tornado rating and convective mode. While 64% of tornadoes addressed in the present study were rated as weak, the low population density of much of Kansas and Nebraska could result in rating underestimations from lack of damage indicators. Moreover, right-moving supercells characterize the convective mode of the vast majority of the 179 tornadoes, with the small remainder characterized by QLCS mode. Furthermore, the present study does not evaluate the convective-mode evolution preceding tornadogenesis, and does not distinguish between potentially variable convective modes prior to or including tornadogenesis. However, this work helps to assess if such an investigation would be worthwhile, by seeking to determine whether radar-based, storm-scale, pre-tornadic circulation trends are apparent.

3. Analysis and discussion

The first evaluation method incorporates environmental data (STP), trends in V_{rot}, along with other storm-scale circulation attributes, by using the tornado probability model from Cohen et al. (2018). Figure 3a suggests an increase in time-binned tornado probabilities preceding tornadogenesis—beginning as early as 16–20 min, and especially within the final 15 min, preceding tornadogenesis, for nearly all percentile rankings leading up to tornadogenesis.

From the 16–20-min interval, tornado probabilities increase with approach to the time of tornado development. For instance, among the 179 tornadoes analyzed, the median tornado probability 21–25 min prior to tornadogenesis is 8%, reaching 12% by 11–15 min prior to tornadogenesis, and surpassing 20% within 5 min before tornadogenesis.

As a practical example, Fig. 3a indicates that 50% of cases were preceded by tornado probabilities ≥12% within 11–15 min prior to tornado formation, with median probabilities steadily increasing leading up to and through this time window—as far back in time as 16–20 min preceding tornadogenesis. An operational meteorologist witnessing steadily increasing tornado probabilities ≥12% therefore may identify the majority of tornado-producing storms as early as 11–15 min before tornado formation. Even lower-probability magnitudes that trend upward could suggest an increase in tornado potential as early as 16–20 min beforehand. To increase operational usefulness, more work is needed to reveal tornado probability trends for nontornadic storms. Differentiating probability trends between nontornadic and tornadic storms could increase confidence on whether or not tornadogenesis is forthcoming. Furthermore, increased confidence for the warning meteorologist may enhance tornado warning lead time while reducing the false-alarm rate.

Most of the 179 tornadoes analyzed in the present study (Fig. 2) were rated as weak (i.e., EF0–EF1). The dataset was separated into weak versus significant tornadoes, to evaluate whether the behavior of tornado-preceding storm-scale attributes is related to subsequent ratings. This was motivated by previous studies linking these attributes to diagnostic assessments of tornado potential and rating (e.g., Smith et al. 2015; Thompson et al., 2017; Cohen et al. 2018). Weak tornadoes yield a less-amplified trend of tornado probabilities with time before tornadogenesis (Fig. 3b) compared to significant tornadoes (Fig. 3c). In fact, Fig. 3c indicates median tornado probabilities rising from 33%–45%, from the 11–15 min period to the 6–10 min bin before tornadogenesis. Tornado probabilities consistently rise for all percentile rankings beginning in the 16–20-min preceding-tornadogenesis time window. Time-binned tornado-probability behaviors associated with significant tornadoes appear to be vary more than for weak tornadoes.

In order to better understand the differences between the behavior of variables preceding weak versus significant tornadoes, one of the components of tornado probability often interrogated as part of tornado threat assessment, V_{rot}, is compared (Fig. 4a,b). For weak tornadoes, time-binned V_{rot} varies little prior to tornadogenesis (Fig. 4a) compared to significant tornadoes (Fig. 4b), consistent with the muted variability of corresponding tornado probability for weak tornadoes (Fig. 3b) compared to significant tornadoes (Fig. 4b). On the other hand, more amplitude appears in the tornado-probability cycle than in the V_{rot} cycle, preceding tornadogenesis in general (compare Fig. 3c to Fig. 4b and Fig. 3b to Fig. 4a). For instance, for significant tornadoes, from the 11–15-min to 6–10-min tornado-preceding time windows, the median value of V_{rot} increases from 43–44 kt (Fig. 4b), whereas the median tornado probability increases from 33%–45% (Fig. 3c). The practical signal reflected by the increase in tornado probability preceding tornado development is stronger than that in V_{rot}, where this change in V_{rot} may fall within sampling and interpretation error, based on the analysis of radial velocities.

These behaviors suggest a blended approach for investigating the behavior of tornado-preceding storm-scale circulations and their near-storm environment to better distinguish a tornado-preceding signal, such as the Cohen et al. (2018) model. This probability model simultaneously incorporate V_{rot} and STP, and includes the subjective classification of mesocyclone clarity and width—the feasibility of which is addressed in the introduction section, applying the notion of a radar team tasked with performing tornado threat assessment. As a result, the simultaneous incorporation of these storm-scale-circulation characteristics and near-storm environmental information may assist with forthcoming tornado threat assessment. Specifically, using the Cohen et al. (2018) model offers a stronger temporal trend preceding tornadogenesis, which potentially aids speculation regarding forthcoming tornado potential. These findings are consistent with the Gibbs and Bowers (2019) message regarding the importance of incorporating more holistic measures of storm-scale circulation to identify forthcoming tornado potential.

Figure 4b also illustrates that an increasing V_{rot} may be interrupted within about 0.5 h prior to significant tornadogenesis. Following a slight increase in these values up to around 0.5 h preceding tornadogenesis, a temporary downward shift appears in the interquartile range of V_{rot} around 30 min before tornadogenesis, followed by a more appreciable increase. This brief interruption is small, but more apparent for significant (Fig. 4b) than weak tornadoes (Fig. 4a) and could reflect the physical process of storm cycling. This process of supercell updraft regeneration can be associated with variable trends in the evolution of the convective element through its existence, which is especially apparent for longer-lived storms. Burgess et al. (1982) and Van Den Broeke (2017) address the concept of the multi-phase lifecycle of storms. This highlights a potentially misleading interpretation that a slight decrease in circulation strength spells decreasing tornado risk in a favorable environment for significant tornadoes.

Throughout this work, frequent references to “time binned” variable tendencies have been made, distinguishing the analyses herein from time-series analysis. This is because the analysis here combines different circulations and their properties into the same time-range bins, effectively masking some of the variability inherent to individual circulations. While considering this variability is important in assessing a storm’s tornado potential, the analyses herein offer bulk trends represented by time-range binning. Regardless, we take initial steps to analyze the variability characteristic of the individual circulations. For each circulation, the elapsed time during which tornado probabilities continuously increase immediately preceding tornadogenesis was determined. Weak and significant tornadoes then were compared, owing to the distinguishable tornado-probability behaviors between them. Figure 5(a,b) indicate that the continuous ramp-up in tornado probabilities was considerably more concentrated within 5 min before tornadogenesis for both weak and significant tornadoes. Over 70% of cases experienced continuously increasing probabilities in the 10 min preceding tornadogenesis.

Finally, the coefficient in Eq. (1) corresponding to pre-tornadic circulation diameter is negative. This means that increases in circulation diameter would correspond to increases in the denominator of the quotient in Eq. (1). As a result, all else held constant, the effect of a unit decrease in diameter would be a decrease in the denominator of this equation, and an increase in tornado probability. This is consistent with stronger azimuthal shear and related stronger dynamic lifting yielding surface vorticity intensification, in association with the circulation’s attendant perturbation pressure deficit.

Figure 6 illustrates the relationship between circulation diameter and tornado occurrence, which varies substantially between weak and significant tornadoes. For weak tornadoes, diameter modestly declines within the five min before tornadogenesis. A stronger decrease in circulation diameter is apparent for significant tornadoes extending back to around 20 min from tornadogenesis. This highlights the importance of monitoring the behavior of circulation size, as a contracting mesocyclone could offer a signal for forthcoming significant-tornado development with appreciable lead time. Furthermore, a slight increase in circulation diameter is observed in the 21–25 min time bin for significant tornadoes (Fig. 6b), corresponding to a slight decrease in Vrot (Fig. 4b). This result may represent a cycling process of the low-level mesocyclone.

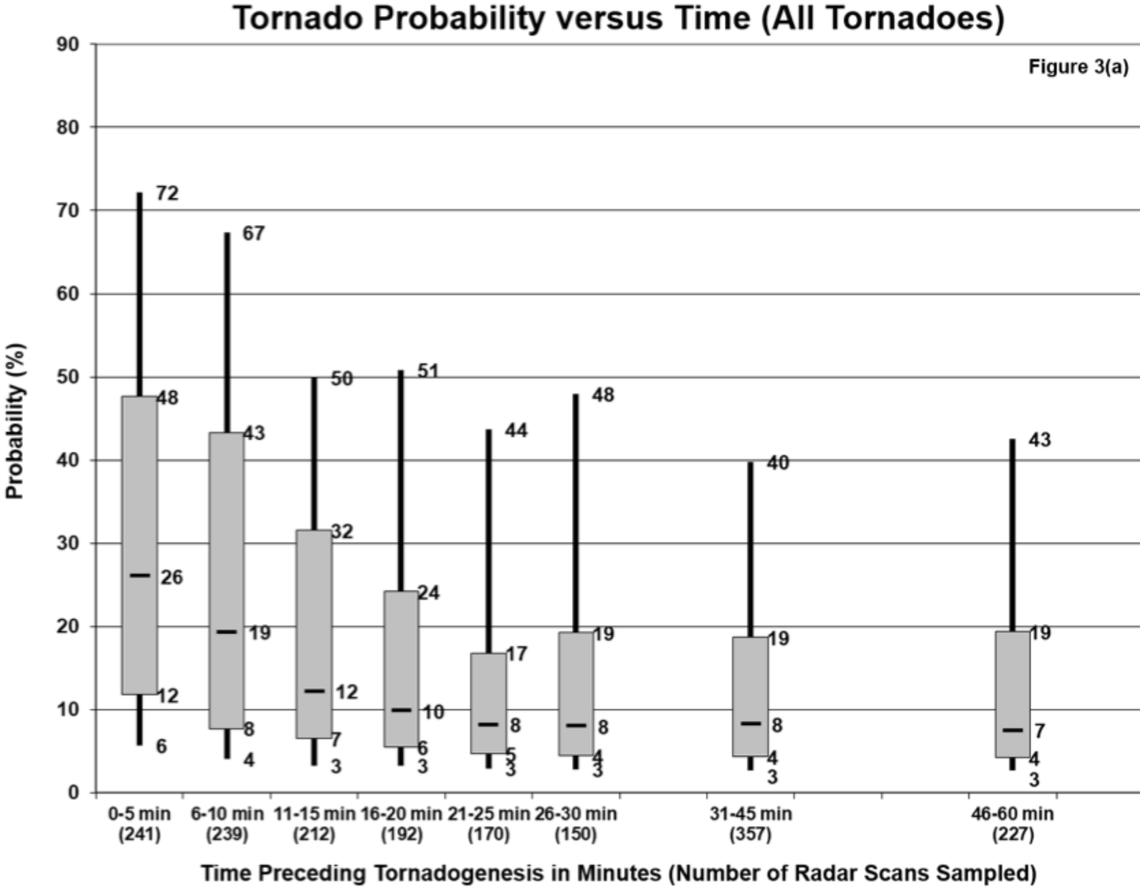
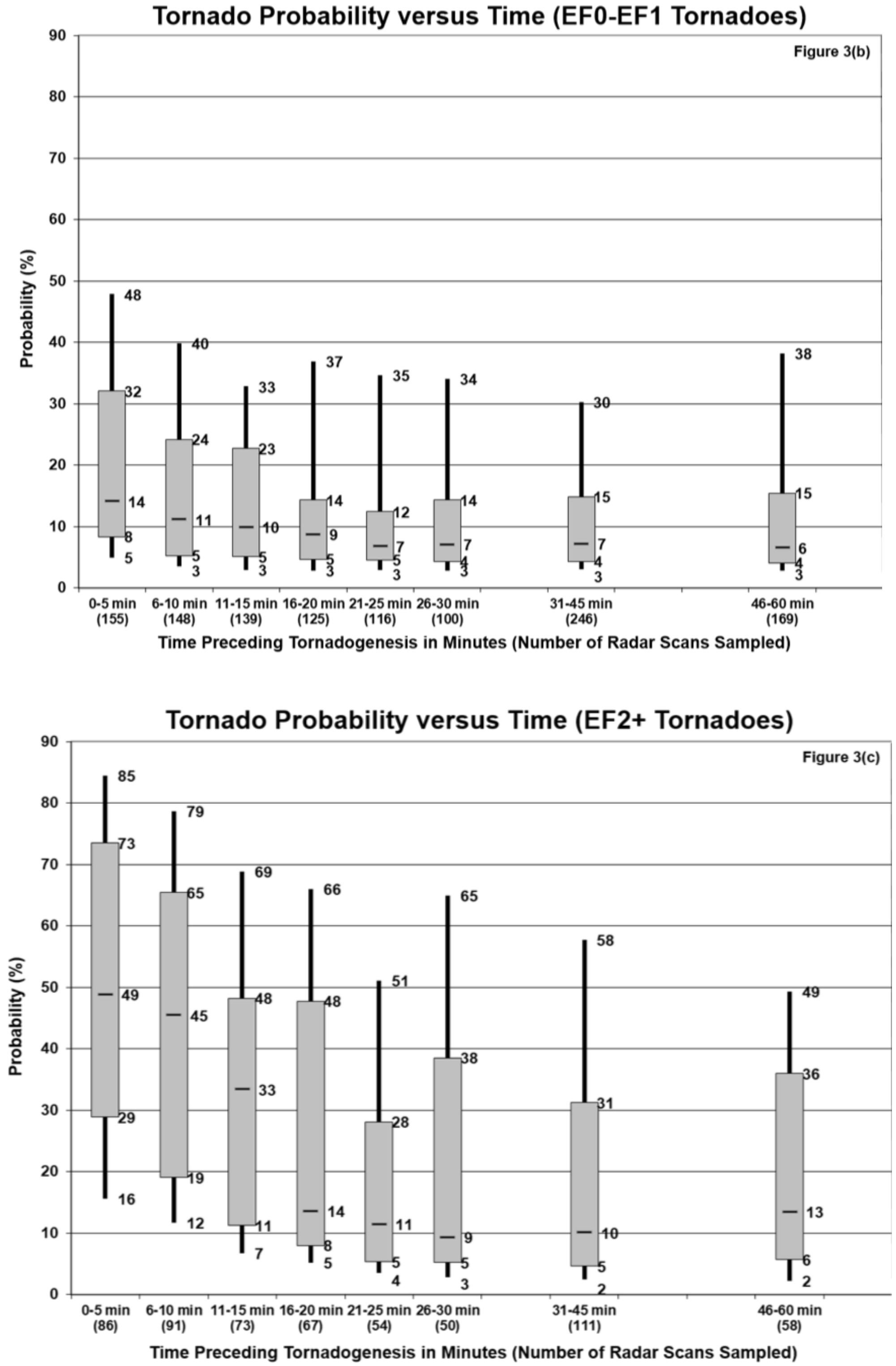
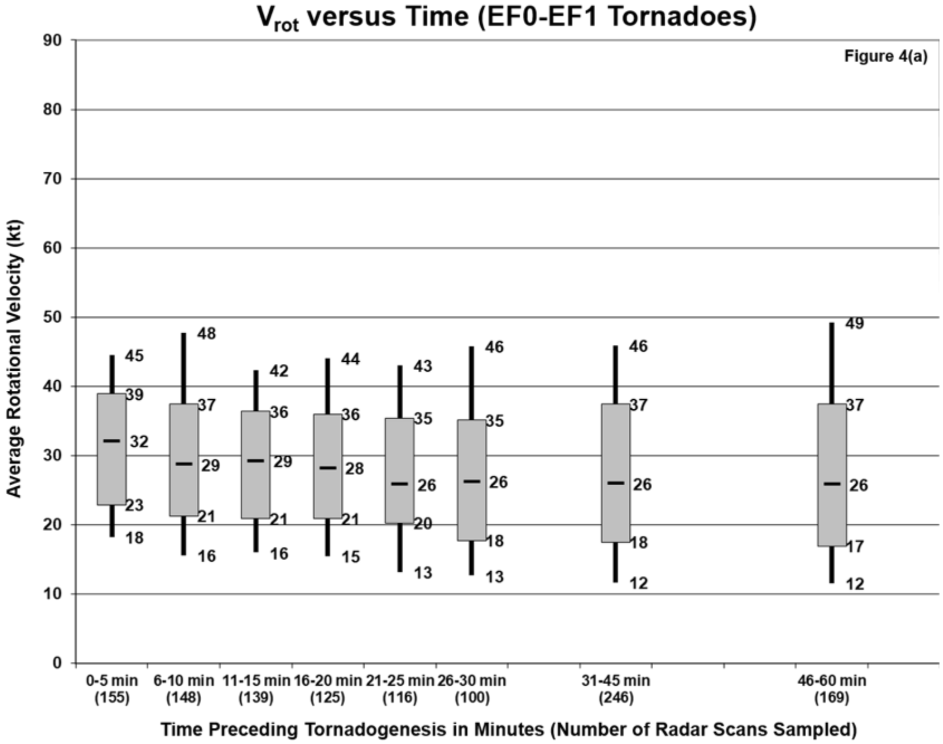


Figure 3: a) Distributions of pre-tornadic tornado probabilities accompanying storm-scale circulations sampled at 0.5° beam angle. Probabilities are computed based on Cohen et al. (2018). Data are binned by ranges of time (min) as follows, marked along the *x*-axis: 5-min bins prior to 30 min preceding tornadogenesis, and then 15-min bins from 30–60 min preceding tornadogenesis. Box-and-whisker plots for each time bin summarize the average tornado probability during that time bin, with the interquartile range depicted by gray shading extending from the 25th to 75th percentile. Median line is within the gray shading. Whiskers extending to the 10th and 90th percentiles. Sample sizes appear beneath each *x*-axis label.



<u>Figure 3 (continued)</u>: b) As in (a) except for weak tornadoes (EF0–EF1); and c) as in (a) except for significant tornadoes (EF2+).



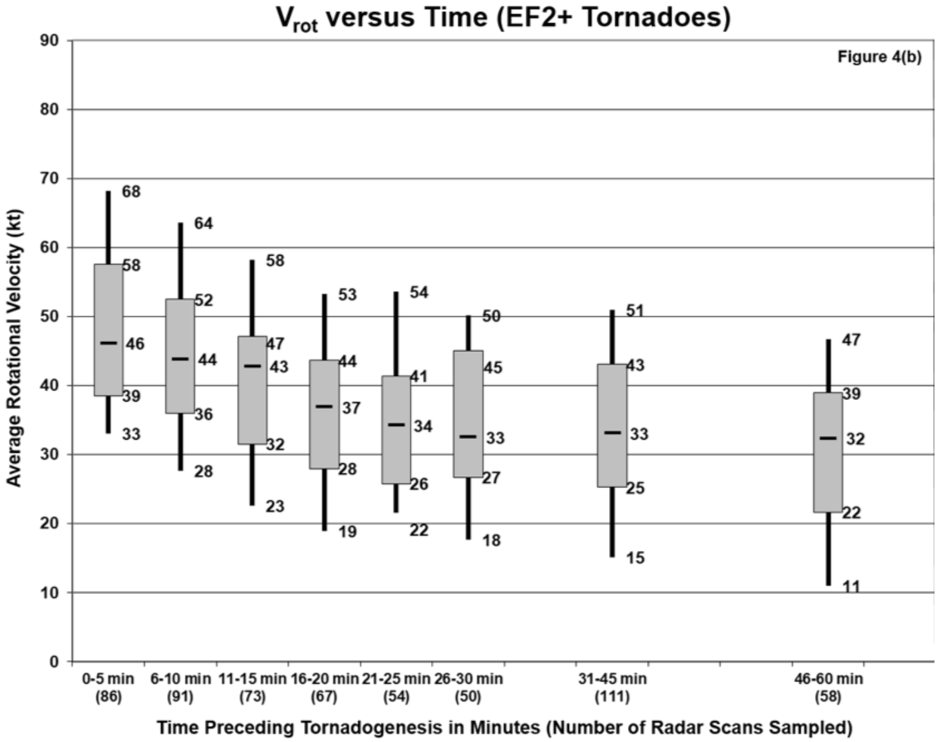


Figure 4: a) As in Fig. 3b, except for V_{rot} preceding weak tornadoes; b) As in Fig. 3c, except for V_{rot} preceding significant tornadoes.

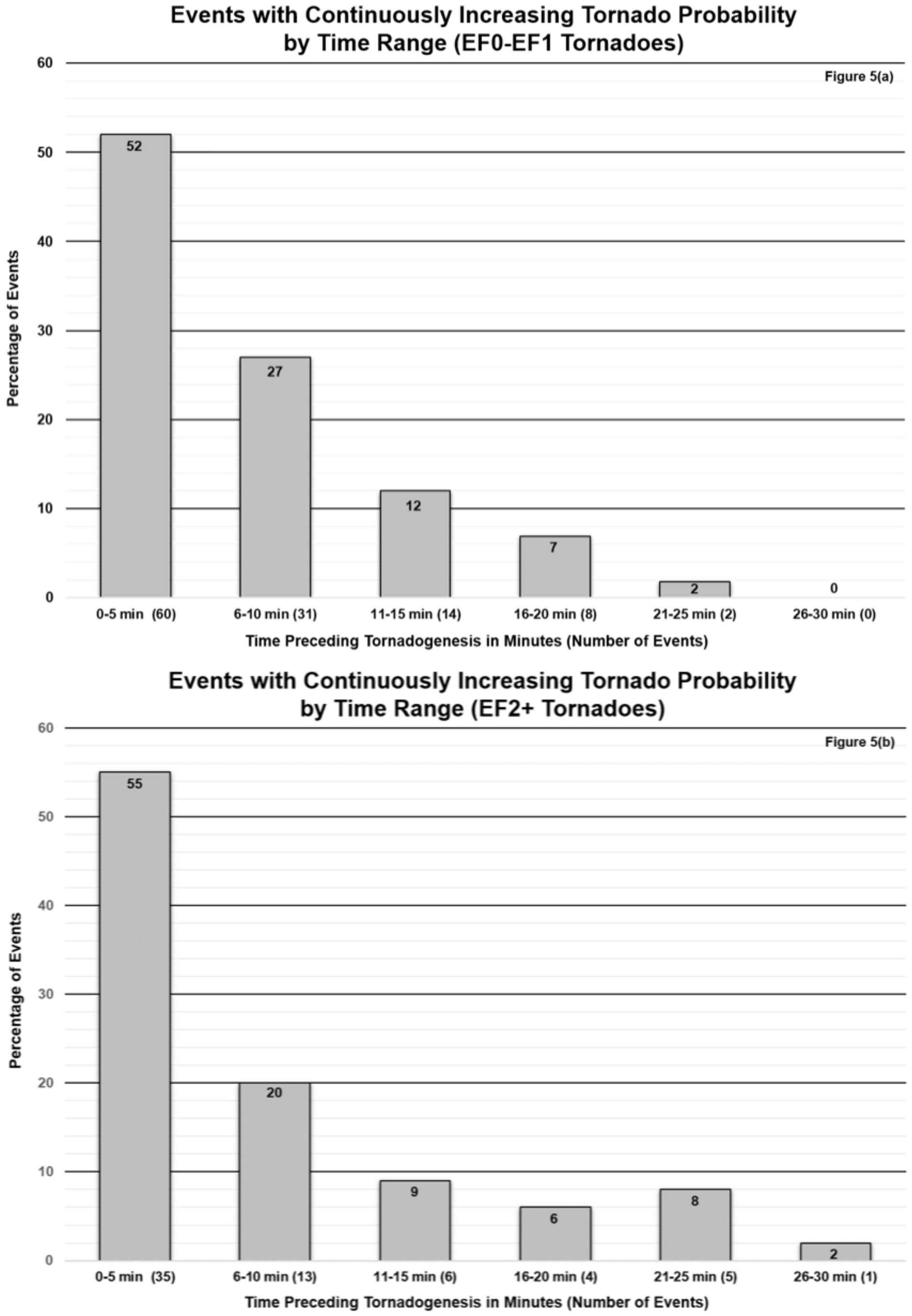


Figure 5: a) Percent of events, by time range preceding weak tornadoes, from which tornado probabilities continuously increase, with percent listed above individual bars and sample size listed in parentheses beneath *x*-axis; b) as in (a) except for significant tornadoes.

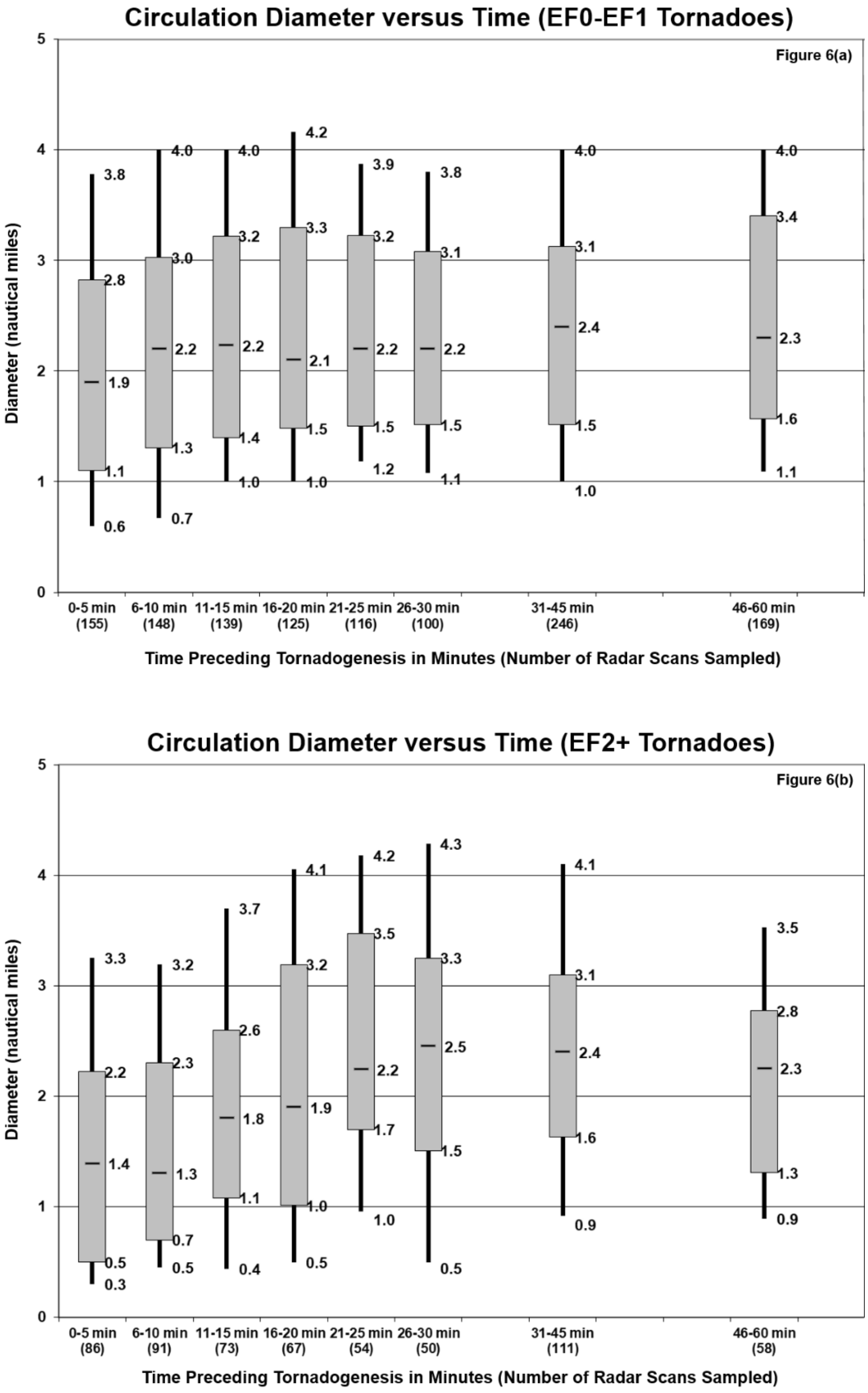


Figure 6: a) As in Fig. 3b, except for circulation diameter preceding weak tornadoes; b) As in Fig. 3c, except for circulation diameter preceding significant tornadoes.

4. Conclusions

This study assesses the behaviors of storm-scale circulations preceding tornadogenesis using WSR-88D data, environmental information from the Storm Prediction Center mesoanalysis fields, and the combination of these data via the tornado probability model provided by Cohen et al. (2018). The overall purpose of this work is to determine how tornado probabilities evolve leading up to tornadogenesis.

This preliminary analysis is comprised of an assessment of 1788 storm-scale circulations (based on the 0.5° radar tilt) and associated environments preceding 179 tornadoes. For weak tornadoes, percentile ranks representing the bulk distributions of V_{rot} only slightly increase leading up to tornadogenesis, principally within 5 min preceding tornadogenesis. This difference is practically very small, potentially existing within typical analysis error of Doppler velocities. For significant tornadoes, the trend in V_{rot} is modestly stronger in magnitude.

A more pronounced improvement in tornado predictability exists by considering a blended approach, simultaneously incorporating multiple properties of the storm-scale circulation and the environment Cohen et al. (2018). Notably increasing tornado probabilities characterize tornado events within 15 min of tornado formation, particularly for significant tornadoes. These results suggest that operationally meaningful trends become more apparent, and potentially at an earlier time, preceding tornado development for significant versus weak tornadoes. Using a blended approach of V_{rot} trends in conjunction with environmental data and other storm-scale attributes also can yield a stronger tornado-preceding signal than V_{rot} alone. Moreover, such a tool permits quantification of probabilities, which can be directly input to hazard-quantification initiatives such as FACETs (Rothfusz et al. 2018) and PHI (Karstens et al. 2015). Variations in circulation diameter also signal forthcoming tornadogenesis, especially for significant tornadoes. The majority of tornado events also are preceded by continuous increases in tornado probabilities within 10 min of tornadogenesis.

The preliminary findings here have unveiled additional context that operational meteorologists can consider in assessing forthcoming tornado potential. Subsequent work could consider a substantially more diverse storm-type cross section of tornadic circulations, to further generalize the results of predecessor variable behaviors. Moreover, while the necessary conditions for tornadogenesis are location-independent, a broader sample representing greater geographic diversity would help substantiate generalizations in future work. Stronger representation of significant tornadoes could help alleviate the asymmetry between weak and significant tornado sample sizes. Consideration of elevations above 0.5° could aid by incorporating more robust three-dimensional trends in convective behavior.

Future work could investigate prospects for automation of some of the storm-scale circulation characteristics manually classified in the present work. Very importantly, an investigation of null cases would be critical for evaluating false-alarm characteristics of parameter behaviors herein. The preliminary results in this study raise awareness for the types of operational signals that can potentially be assessed to anticipate and quantify forthcoming tornado potential, and the subsequent work will be building on this critical foundation.

ACKNOWLEDGMENTS

The authors are grateful for the oversight and review of this work by Jeff Manion of National Weather Service Central Region Headquarters. The authors are also thankful for the support of this work provided by Kris Craven of the National Weather Service in Topeka and Dr. Pablo Santos of the National Weather Service in Miami, and discussion about this work with Dr. Greg Mann of National Weather Service in Detroit/Pontiac, MI. The authors are appreciative of the comments and guidance provided by three reviewers—Corey Mead, Dr. Pam Heinselman, and Dr. Matthew Van Den Broeke—along with Editor Roger Edwards's contributions to improving this work. In addition, University of Oklahoma doctoral student Tomer Burg is graciously acknowledged for creating and running the website interface that provides estimates of tornado probabilities and wind speeds based on WSR-88D data. The second author, Dr. Ariel E. Cohen, is deeply appreciative of his father, Joel B. Cohen, who is also co-author of Cohen et al. (2018), for his assistance in interpreting, and identifying manners to operationally apply, the tornado-probability model presented therein. The scientific results and conclusions, as well as any

views or opinions expressed herein are those of the authors and do not necessarily reflect the views of NOAA or the Department of Commerce.

REFERENCES

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

Burgess, D. W., V. T. Wood, and R. A. Brown, 1982: Mesocyclone evolution statistics. Preprints, *12th Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., 422–424.

Cintineo, J. L., and Coauthors, 2018: The NOAA/CIMSS ProbSevere Model: Incorporation of total lightning and validation. *Wea. Forecasting*, **33**, 331–345.

Cohen, A. E., J. B. Cohen, R. L. Thompson, and B. T. Smith, 2018: Simulating tornado probability and tornado wind speed based on statistical models. *Wea. Forecasting*, **33**, 1099–1108.

Eilts, M. D., and S. D. Smith, 1990: Efficient dealiasing of Doppler velocities using local environment constraints. *J. Atmos. Oceanic Technol.*, **7**, 118–128.

Gibbs, J. G., and B. R. Bowers, 2019: Techniques and thresholds of significance for using WSR-88D velocity data to anticipate significant tornadoes. *J. Operational Meteor.*, **7 (9)**, 117–137.

Hart, J. A., and A. E. Cohen, 2016: The Statistical Severe Convective Risk Assessment Model. *Wea. Forecasting*, **31**, 1697–1714.

Houser, J. L., H. B. Bluestein, and J. C. Snyder, 2015: Rapid-scan, polarimetric, Doppler radar observations of tornadogenesis and tornado dissipation in a tornadic supercell: The “El Reno, Oklahoma” storm of 24 May 2011. *Mon. Wea. Rev.*, **143**, 2685–2710.

Jing, Z., and G. Wiener, 1993: Two-dimensional dealiasing of Doppler velocities. *J. Atmos. Oceanic Technol.*, **10**, 798–808.

Karstens, C. D., and Coauthors, 2015: Evaluation of a probabilistic forecasting methodology for severe convective weather in the 2014 Hazardous Weather Testbed. *Wea. Forecasting*, **30**, 1551–1570.

Kumjian, M. R., and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47**, 1940–1961.

——, and ——, 2009: Storm-relative helicity revealed from polarimetric radar measurements. *J. Atmos. Sci.*, **66**, 667–685.

Markowski, P. M., 2002: Hook echoes and rear-flank downdrafts: A review. *Mon. Wea. Rev.*, **130**, 852-876.

NCEI, 2018: Storm events database. [Available online at https://www.ncdc.noaa.gov/stormevents/.]

Parker, M.D., 2014: Composite VORTEX2 supercell environments from near-storm soundings. Mon. Wea. Rev., **142**, 508–529.

Romine, G. S., D. W. Burgess, and R. B. Wilhelmson, 2008: A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma City area tornadic supercell. *Mon. Wea. Rev.*, **136**, 2849–2870.

Rothfusz, L. P., R. Schneider, D. Novak, K. Klockow-McClain, A. E. Gerard, C. Karstens, G. J. Stumpf, and T. M. Smith, 2018: FACETs: A proposed next-generation paradigm for high-impact weather forecasting. *Bull. Amer. Meteor. Soc.*, **99**, 2025–2043.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, and D. S. Zrnic, 2005: Polarimetric tornado detection. *J. Appl. Meteor.*, **44**, 557–570.

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.

——, ——, 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.

Thompson, R. L., C. M. Mead, and R. Edwards, 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102–115.

——, 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.

——, ——, ——, ——, J. C. Picca, A. E. Cohen, E. M. Leitman, A. M. Gleason, and P. T. Marsh, 2017: Tornado damage rating probabilities derived from WSR-88D data. *Wea. Forecasting*, **32**, 1509–1528.

Van Den Broeke, M. S., 2017: Polarimetric radar metrics related to tornado life cycles and intensity in supercell storms. *Mon. Wea. Rev.*, **145**, 3671–3686.

REVIEWER COMMENTS

[Authors' responses in *blue italics.*]

REVIEWER A (Pamela L. Heinselman):

Initial Review:

Recommendation: Accept with minor revisions.

Substantive Comments: At first I was a bit concerned about the dependence on previous work referenced for understanding the methodology. However, after reading it carefully I believe that the content provided is likely sufficient and enables readers to dig deeper if interested. So not explicitly substantive but wanted you to be aware of this thought process.

Something that struck me in the analysis of trends in probability is what time interval/consistency in upward trend constitutes actionable lead time for a forecaster. For example, if an upward trend is seen for the first time between 11–15 min and 6–10 min, the actual lead time is at a maximum 6–10 min. It is unclear to me whether trends reported in the manuscript apply the former or subsequent time range. In my HWT experience (of course this group has much more experience!), most warning forecasters like to see consistency in a trend prior to taking action… would this bring the lead time to 0–5 min? Given the focus of this study on potential operational applications, I think it would be useful to discuss this challenge within the text.

This has been corrected. The probability increase begins as early as 21–25 min, but as you stated in an operational setting, this will only be realized in the 16–20 min time window.

One of the things that I learned working with rapid-update phased-array radar data is the usefulness of 1-min updates to detect trends in radar attributes. Given the implementation of new volume scans like SAILS within the NWS, as data samples sizes increase it might be of interest to conduct a similar study with 0.5° radial velocity data that updates every 2 min or so to see what those trends look like in comparison. Such a study would take longer if done manually, but could be discussed as potential future work if you think that could be useful to the field. Such a study could also help to inform the importance rapid-scan phased array or other radar technologies as an eventual replacement for the WSR-88D network.

Yes, small time bins could certainly be used in subsequent studies.

Another thing that struck me is the concept of radar teams in operations in which one person would input the V_{rot} information to assess probability using the equation developed by Cohen et al (2018). In my understanding of this scenario, a member of the radar team would produce a trend using a machine-human-mix approach. I am curious as to whether this concept has been tested in real time or pseudo-real-time operations; does such an interface to input V_{rot} into the equation exist? Can trends be displayed without the forecaster having to write them down? If not, do you envision conducting such a test in the future?

Yes, a radar assistant is encouraged to input the corresponding data and reveal the probabilities to the warning meteorologist. Furthermore, a web page has been developed recently with archiving capabilities—which allow for forecasters to see probability trends. The web page link has been added to the manuscript.

Because the estimate of V_{rot} is promoted be done by the forecaster (which I know is already part of operations), I sense concern that an automated algorithm estimating V_{rot} and computing the tornado probabilities would perform more poorly. Given the historic limitations of various radar-based algorithms designed to estimate mesocyclone and TVS intensity I can understand this reluctance. Yet I am curious as to the degree to which such an approach might reproduce the results found here.

This concept could be fodder for discussion near the end of the manuscript. This is a good idea for future work and has been added to the Conclusions section.

The probability of tornado trends shown for individual events—one significant and one significant wind—are both a bit noisy, which is what real trends tend to look like. There is a nice discussion of uncertainty of radar-based estimates within the manuscript. Do you have a sense for how this uncertainty might be accounted for in operations? Interestingly, the probability of tornado spikes just prior to the significant wind event. It might be worth mentioning this in the findings. Do you have a sense for why the timing of this trend occurs with this event? I have concern that, as written and with the small sample size of two storms, there is overconfidence expressed regarding the applicability of tornado probability trends being highly differentiable for non-tornadic storms. I suggest stating this as a hypothesis instead.

All text and associated images related to the null-case section has been removed.

The data set provided no statistical significance when compared to the tornado data set.
I noticed that no statistical significance tests or skill scores are computed within this study, and perhaps rightly so because of the study's purpose. At what point would you see efforts along this line being important to the implementation of methods like this in operations?

This has been addressed in the revised Conclusions section.

[Minor comments omitted...]

Second Review:

Recommendation: Accept with minor revisions.

General Comment: The authors have done a good job improving the quality and presentation of this study and I recommend one revision.

[Minor comments omitted...]

REVIEWER B (Matthew S. Van Den Broeke):

Initial Review:

Recommendation: Accept with major revisions.

General Comments: The manuscript is not publishable in EJSSM as currently written. It may be publishable after substantial revisions and additional analysis which establishes the significance of results.

The manuscript presents an examination of pre-tornado trends of tornado probability in a sample of Kansas and Nebraska supercell storms. The tornado probabilities are derived from both radial velocity signatures and a tornado probability model (Cohen et al. 2018) which uses environmental inputs. The authors show that tornado probabilities increase prior to tornadogenesis, especially prior to strong (EF2+) tornadoes.

The manuscript is a resubmission of one which was submitted to a different journal, and unfortunately most of the same deficiencies remain in this submission. The reviewer has several methodological concerns, does not believe the conclusions are supported by the evidence provided, and believes the operational significance of the results is overstated given the lack of comparison with nontornadic events. Specific major and minor comments are included below. A document with line numbers added has been attached to the review for ease of locating the items listed below. *[Editor's Note: EJSSM is flexible on reviewers' methods for presenting needed corrections to authors. Reviewers may add manuscript-embedded* *comments in Word or PDF form for minor comments, as well as line numbers to the paper, for a review draft.]*

Major Concerns:

1) The Introduction would benefit by a broader context and some reorganization. Specifically:

The discussion [dual-pol] is long and very detailed. While this work does provide the context for the methods in the present study, this level of detail does not appear warranted.

Mentions related to dual-pol have been shortened greatly.

In contrast, a broader context of the tornadogenesis predictability problem is generally lacking in the Introduction as currently written. This need not be long, but is important context.

Have added a portion related to tornadogenesis predictability from Hart and Cohen (2016).

[Several sentences] are unnecessarily long and repeat a few items. The final paragraph of this section touches on a conceptual model of tornadogenesis, which is not necessary in the context of the present study.

The mentioned paragraph has been removed.

If KR09 is mentioned in this context, Dawson et al. (2015; JAS) also must be cited and briefly discussed as an update of KR09.

This all collectively has been removed.

It seems like the motivation for the study and specific study goals should be introduced [later]. Perhaps the following information can be rearranged and trimmed accordingly.

This has been shortened and flows much better with the motivation/goals.

Why is Gibbs and Bowers (2019) introduced at the end of the Introduction, after a discussion of what will be done in the present study? This needs to be incorporated into the Introduction where appropriate rather than being added to the end.

This has been moved to an appropriate position in the complementary work on page 4.

2) Some aspects of the Methodology need to be better explained:

It is unclear why different EBWD thresholds were used in different years. For the study to be reproducible, this needs to be explained.

This section has been removed. The study simply used the data set from Thompson et al. (2017). The thresholds they used to create the dataset are irrelevant to this study. This study used the tornado reports and the associated STP values at that time.

It should be stated what percentage of KS/NE tornadoes are included in the sample. Given the regional radar network, the reviewer suspects a lot of cases were not included since they did not meet the 9900' ARL threshold, but a percentage should be stated. This would help readers decide if the study is relatively preliminary or relatively thorough for the sampled region.

Have added the percentage of the cases used in the study compared to the total data set (179/1025 = 17.5%). Furthermore, only 11 cases (6.1%) had radar scans reach 9,900 ft ARL, at which point the radar analysis for that tornado ended.

Tornadoes were also placed into weak and significant categories, but it should be addressed how land cover and population density may affect these categorizations, particularly on the Great Plains.

Have addressed this with including that the rural nature of KS/NE may result in tornadoes being rated lower than reality. This is due to ratings and associated wind speed estimates limited to damage indicators. More rural areas results in few damage indicators.

Choice of analysis times/radars should be better described.

a) Analysis times were basically any for which a circulation was visible and the storm met the required thresholds (e.g., 9900' ARL). Which radar was used for a given analysis time—the closest radar to the storm? In that case, presumably, you changed radars if a different radar became closer? Radar choice should be more carefully specified. If different radars were used for a single storm, it should be stated how often this was the case and what effect changing radars might have on the results (e.g., the discontinuity in ARL value between one scan and the next).

Have added that the choice of radar site was determined from the nearest radar site at the time of the tornado, which is documented with Thompson et al. (2017) data set. Furthermore, only one radar site was used for each tornado.

b) It should be indicated what percentage of scans were characterized by certain distances from the radar (e.g., ARL values). It's also worth exploring whether storms in different ARL bins have different predictability statistics. A similar study carried out in a different geographic location may have fairly different results because of the spacing of the radar network, and it's worth pointing this out and attempting some simple quantification of the effects.

Have added an image depicting the percentage of radar scans in 3 bins (0–3000 ft ARL, 3000–6900 ft ARL, and 6000–9900 ft ARL). Additionally, an analysis of tornado probabilities preceding tornadogenesis were conducted for the aforementioned bins and trends were nearly identical for all 3 bins, therefore this was not included within the manuscript.

c) Can you justify using a threshold of 9900' ARL when the low-level circulation is clearly no longer being sampled?

Using 9900 ft ARL as a maximum value adds diversity of sampling heights representative of what we may actually encounter in the real world (i.e., no obs below 9900 ft). This has been added.

3) <u>Addressing near-storm environmental variability</u>:

The authors note that the use of a constant STP value through time may not be representative because of mesoscale and storm-scale fluctuations in environment. The magnitude of the error could, however, be roughly approximated. The authors should at least bring in the results of prior work (e.g., Parker 2014; MWR) to help readers get a sense for typical variability of this variable in the near-storm environment.

This has been added, addressing the roughly 3.1 increase in STP that Parker 2014 found in the near-inflow region.

4) <u>Interpretation of a few figures appears to be misleading/not supported by the data or to have additional implications which need to be discussed</u>:

Figure 2a [and text] state that an increase in tornado probabilities begins 21–25 min prior to initial tornadogenesis, but this is misleading. This is the lowest point in terms of probabilities for all tornadoes, but there would be no way to know that the probabilities would subsequently increase. That is, operationally, one would not have any indication of increasing tornadogenesis potential in the 21–25 min prior. It is in the 16–20 min window that the probabilities begin to increase a little, which may be the first operationally-meaningful indication.

This has been updated to address that any operationally-meaningful trend begins at 16–20 min.

Figure 3b: The authors indicate a "marked uptick"; the reviewer is not sure they follow. The median is similar, and the 10th and 90th percentile values only increase a little. This doesn't seem to justify being called a marked uptick.

The verbage, "marked uptick" has been removed.

Figure 3b: This sinusoidal oscillation should be placed in the context of storm cycling, which is fairly well-known. The oscillation described makes good sense in terms of storm cycling, but this is not discussed in the manuscript. Doesn't storm cycling potentially represent a fundamental limitation to our ability to use such methods to anticipate tornadogenesis? If so, this is important to discuss. This result also appears to suggest that cycling behavior is more prominent in strongly-tornadic storms, which makes sense and is a nice result.

Have added discussion concerning the linkage between storm-cycling and V_{rot} *oscillations.*

5) [Null-event discussion]: As has been pointed out by reviewers of prior drafts, the discussion of null events is seriously flawed and should either be removed or substantially revised. This also includes the portion of the Conclusions. The most important problems with this analysis include:

- It's not clear how two events can be used to state that there is different behavior between tornadic and nontornadic storms. There is no way to know that these are representative events. They should at least have appeared similarly likely to produce a tornado at some point. A larger sample would alleviate this concern, ideally containing at least 70% the number of tornadic storms.

- It would also be more convincing if the comparison storms were from the same geographic region. One of the nontornadic storms in the current draft is from Oklahoma, which is irrelevant when all the other storms are from KS/NE.

- The comparison with wind and hail is not relevant in the first place. Since hail and wind occur in different points in the storm lifecycle than tornadoes, one would expect the radar signatures to be different leading up to hail/wind than leading up to tornadoes. Your result, then, is exactly what would be expected and does not indicate a difference between the two populations. It would be possible to alleviate this concern by selecting a sample of nontornadic storms and looking at the tornado probabilities leading up to tornadogenesis failure in them (the point with highest low-level radial velocity difference).

The null case section has been removed.

6) Figure 4: This may be the reviewer not correctly understanding the values in the figure, but something seems to be off with the values for n and percentage. As an easy example, in Fig. 4b, the second bar indicates that n = 13 and that 20% of events showed a continuous increase. This corresponds to 2.6 events. If 2 events showed this behavior, it would be 15%, and if 3 events showed it, the value would be 23%. These are different than the reported 20%.

The x-axis has been reworded to address the confusion.

[Minor comments omitted...]

Second Review:

Recommendation: Accept with minor revisions.

General Comment: The manuscript is a resubmission of a paper examining storm-scale circulations in pre-tornadic storms in Kansas and Nebraska. This reviewer feels that the prior reviewer comments were well-addressed, and I now recommend that the paper be published in EJSSM after (very) minor revisions. Larger issues with the methods and interpretation of findings have been largely resolved, and my small remaining items are mostly related to grammar and formatting.

[Minor comments omitted...]

REVIEWER C [Corey M. Mead]:

Initial Review:

Recommendation: Accept with minor revisions.

General comments: This study utilizes a subset of a comprehensive database to examine the behavioral characteristics of storm-scale circulations prior to tornadogenesis. A previously developed statistical model is leveraged to determine the probability of tornado occurrence, based on inputs of pre-tornadic radar data and proximity environmental considerations. As stated in the manuscript, "The overall purpose of this work is to determine how tornado probabilities evolve leading up to tornadogenesis." Admittedly, this study is preliminary in nature, with the present dataset lacking geographic and storm-type diversity. And, inclusion of an equally sized non-tornadic ("null") case set would aid in substantiating the presented results. Nonetheless, this work is an appreciated attempt at building a semi-automated guidance system which NWS warning forecast teams could use as a tool in the assessment of near-term tornado threat.

Substantive Comments: A critique I have of the manuscript is the comparison of only two significantly severe non-tornadic cases to the 179-case tornadic dataset in the analysis and discussion session. The authors assert that the existing work is preliminary, and that a more diverse dataset is necessary to draw more concrete conclusions. As such, my recommendation would be to either focus solely on the statistical analysis of the 179-case tornadic dataset, or create a similar-sized, non-tornadic case set which would allow for more rigorous testing of the devised system.

The null cases and associated text have been removed from the manuscript.

The authors have done a good job in documenting the lineage of studies that established conditional tornado probabilities from a large database linking severe weather reports to radar characteristics, storm type, and environment. And, they have importantly highlighted the ***diagnostic*** nature of that process. The compelling aspect of this study is the attempt apply similar tools in a ***prognostic*** sense. Following this line of thought, one could ask: are the same variables that have proven valuable in the identification of a phenomenon equally as valuable in its prediction? Specifically, the tornado-probability values derived by the Cohen et al. (2018) statistical model are derived from only 0.5-degree radar data. Diagnostically, that makes a whole lot of sense. But, what about from a prognostic standpoint, especially when the storm of interest is close to the RDA; say within 30–40 nm? In that case, would the evaluation of multiple-elevation data (e.g., 0.5 and 0.9, or 0.9 and 1.3, or 0.5, 0.9, and 1.3, etc.) improve the time-binned trends highlighted by Figs. 2 and 3? In other words, is it possible that the 0.5 degree scan "undershoots" critical tornadic precursor information when a storm is at close range? Clearly, doubling or tripling the manual input of data into the statistical model is not practical in an operational setting. I'm simply curious as to whether the additional data would improve the statistical results.

This is a good idea that we are addressing in the Conclusions. The incorporation of circulations that extend upwards of 9900 ft above radar level using 0.5° data also accounts for sampling at higher *elevations (e.g., RDA closer to storm). However, consideration of other elevations is being included in the updated Conclusions section.*

As discussed in the Introduction, Thompson et al. (2017) found that larger magnitudes of average rotational velocity **and** smaller circulation diameters tended to be associated with higher probabilities of tornado occurrence. Moreover, it was stated that Cohen et al. (2018) found that average rotational velocity and circulation diameter, among others, provided the greatest amount of explanatory power for tornado risk. However, the only component analysis performed and compared to the model-derived tornado probabilities was average rotational velocity. Given the Thompson et al. (2017) and Cohen et al. (2018) findings, was there any consideration of doing a similar component analysis on circulation diameter?

We have added a section and associated graphics related to circulation diameter.

I am a bit unclear on the discussion of the CIMSS ProbSevere and ProbTor models in the introduction. To me, it's odd that this guidance is solely highlighted as "additional tools in the tornado-threat-assessment process". Certainly, if the authors are only considering statistical modeling systems, the CIMSS ProbSevere model has shown promise in discriminating between severe and non-severe storms. If so, that needs to be explicitly stated. However, I'm not aware of any formal literature that has demonstrated the skill of the CIMSS ProbTor model in discriminating between tornadic and non-tornadic storms.

This discussion has been fully updated, as referenced within the in-text comments.

Technical Comments: *[Editor's Note: A few of the "technical" comments that this reviewer embedded in the paper are scientific in nature, and may qualify as "substantive" for the purpose of final review documentation, depending on the nature of their resolution by the authors.]*

With regard to "additional tools in the tornado-threat-assessment process," I'm surprised the authors chose to only mention ProbSevere and ProbTor. I was unable to find where in the cited reference either of the two statistical models were able to successfully discriminate between tornadic and nontornadic events.

This is a good point, and probably evolved from the earlier version of the paper trying to explain how this tornado probability is different from that produced by ProbTor. And of course, there are numerous sources of other information to support tornado threat assessment: observed-data, conceptual models, and other tools. We have adjusted the wording substantially in light of these considerations.

Median V_{rot} values increase from 32–33 kt with a slight increase in the 25th and 75th percentile values. Am I reading that correctly? If so, I wouldn't consider that a "marked uptick".

Have reworded to state, "near-steady to slight increase".

[Minor comments omitted...]

Second Review:

Recommendation: Accept with minor revisions.

General comments: The authors have satisfactorily addressed my substantive concerns, and only minor revisions are necessary prior to publication.

[Minor comments omitted...]