P1D.28 **A Method for Adding Intensity Information to NWS Preliminary Tornado Damage Paths**

Christopher D. Karstens[1,2*],Bobby Saba[2], Bryan T. Smith[1], Richard L. Thompson[1],
Somer A. Erickson[3], Israel L. Jirak[1]

[1] NOAA/NWS/Storm Prediction Center, Norman, OK.
[2] School of Meteorology, University of Oklahoma, Norman, OK.
[3]Federal Emergency Management Agency, Washington D.C.

1. INTRODUCTION

In the immediate wake of a tornado event, there often exists a flurry of preliminary information about the tornado's intensity and the locations that were impacted. Such information is obtained from storm spotters or local emergency management personnel (EMs), and/or photography or videography ascertained via social media. However, the accuracy of such information, in terms of the exact time, location and particularly, intensity, is often of varying degrees of quality and is often incomplete.

Incident response personnel (i.e., first responders, local EMs) are tasked with quickly devising and implementing the appropriate tactical and strategic response to emergencies, including tornadoes. Tornado damage can be especially problematic due to 1) the rarity of such events, 2) the expansive spatial scales involved (which can overwhelm these services), and 3) the variation in impacts (which can vary with tornado intensity and population density). Therefore, access to timely and reliable, yet preliminary, reconnaissance information that is nationally consistent can help fill a critical information void and guide these efforts.

Tornado intensity diagnosis, let alone prediction, remains a challenge. The accepted method for diagnosing tornado intensity is through utilization of the EF-scale (WSEC 2006) in a detailed post-event damage survey. Surveyors identify individual damage indicators (DIs), if available, and assess the degree of damage (DOD) incurred to the DI to estimate the wind speed at the DI location. This process is repeated until a suitable number of DIs have been assessed to quantify the variation of tornado intensity along the entirety of the tornado track, culminating in a peak EF-scale rating. Damage surveys can take at least 24 hours to complete, and often require several days or even weeks for events that are more widespread or with DIs suggestive of EF4+ damage that may require detailed forensic analysis. Therefore, the post-event survey approach cannot (and should not) address the immediate needs of incident response personnel.

In the late 2000s, the National Weather Service (NWS) deployed an upgrade to the WSR-88D network to include super-resolution radar (Brown et al. 2005), and starting in the early 2010s, dual-polarization radar was introduced (e.g., Scharfenberg et al. 2005; Istok et al. 2009).

* *Corresponding author address:* Christopher D. Karstens, NOAA/NWS/Storm Prediction Center, National Weather Center, David L. Boren Blvd, Norman, OK 73072; email: chris.karstens@noaa.gov.

These enhancements allow for an improved ability to resolve mesocyclone and tornado vortex signatures, including tornado debris signatures. In the advent of these newly detectable signatures arose anecdotal evidence for relating such signatures to intuitive assertions of tornado intensity and related impacts, leading to the introduction of tornado impact tags in NWS tornado warnings (NWS 2011). With the use of WSR-88D radar data, NWS meteorologists are able to discriminate between weaker and more intense damage-based tornado wind speeds (Smith et al. 2020b) even without a consistent methodological approach.

To support the effort of real-time tornado intensity diagnosis, recent studies (e.g., Smith et al. 2020a; 2020b; 2022) have engaged in a detailed climatological analysis of 1) peak, scan-based and event-based WSR-88D rotational velocity (V_{rot}) signatures associated with tornadoes, with 2) a characterization of the near-storm environment of these events, as depicted by the effective Significant Tornado Parameter (STP), and 3) the associated population density on a 1 km grid, which are all linked to the 4) nearest-neighbor maximum DI-based wind speed. This effort has yielded 7513 V_{rot}/STP/population combinations from tornadoes occurring between 2009-2017, from which threshold dependent empirical damage-based wind speed distributions have been developed (Fig. 1; Smith et al. 2020a).

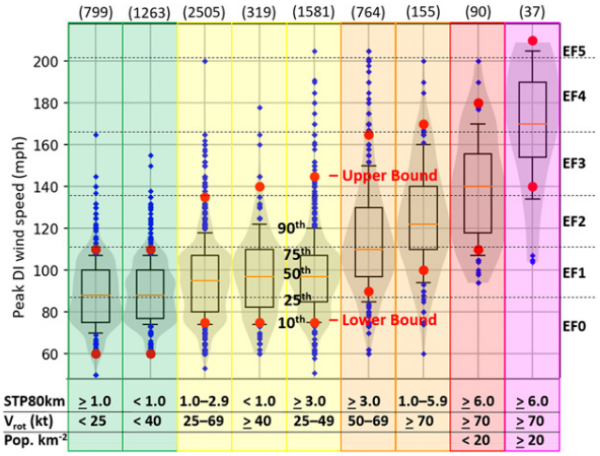

Figure 1. Empirical damage-based wind speed distributions within various combinations of peak Vrot signatures, effective STP, and population density. From Smith et al. (2020b), their Fig. 9.

In a general sense, these empirical distributions should help to guide and refine the aforementioned real-time intuition-based assertions of tornado impacts and implied intensity in a consistent manner nationwide. However, when applied immediately after a tornado event on a scan-by-scan basis, these distributions may also offer utility to address the needs of incident response personnel. This study describes a statistical model developed to create preliminary tornado damage paths that include spatial depictions and variations of damage-based wind speed. Deterministic output from this model was subjected to verification on an independent set of tornado events. Results from this effort are described in the context of ongoing and future plans.

2. METHODOLOGY

The approach developed at the time of this writing utilizes the following methodological approaches to develop a statistical model characterizing the estimated spatial extent and variation of damage-based tornado wind speeds. Three basic approaches have been taken to develop the statistical model, including 1) use of the Smith et al. (2020b) empirical wind speed distributions to obtain percentile-based estimates of the maximum wind speed near the ground, 2) use of an analytical vortex model to obtain an estimate of the variation of wind speed orthogonal to the tornado track, and 3) use of the operational NWS preliminary damage path polygon methodology (Smith and Speheger 2006) to obtain an estimate of the damage path spatial extent.

2.1 *Statistical Model*

Estimates of the peak damage-based wind speed at each WSR-88D V_{rot} centroid location were obtained from the Smith et al. (2020b) empirical wind speed distributions. The wind speed for each V_{rot}/STP/population combination used a percentile-based method, yielding 101 wind speed values. This allows for 1) consistency when composing deterministic depictions of wind speed along the track by fixing the relative statistical frequency of the wind speed across the V_{rot}/STP/population parameter space, and 2) the ability to derive probabilistic depictions of wind speed exceeding various thresholds. For this study, we only evaluate the deterministic approach, and the probabilistic approach, analogous to the work of Saba et al. (2022), will be the subject of future work.

Once estimates of the peak wind speed are obtained for point locations along the track, the next question to address is how to depict the variation of wind speed orthogonal to the tornado translation direction as implied by consecutive V_{rot} centroid locations. This was done by utilizing an analytical vortex model. This model uses a set of Rankine equations, which assume solid body rotation in the tornado core (interior to the radius of maximum winds) and logarithmic decay exterior to the tornado core (exterior to the radius of maximum winds). The radius of maximum winds was assumed to be roughly 20% of the radius of damaging winds (i.e., 65+ mph). Lastly, the ratio of the radial to tangential velocities was assumed to be 2:1, consistent with the findings of Karstens et al (2013).

An advantage of using the Rankine-based analytical approach is the ability to depict asymmetries in the resulting peak near-ground wind speeds as imposed from the implied tornado translation speed. This may better represent the physical characteristics of the damage gradients observed at ground level as opposed to using a more simplistic linear decay function emanating orthogonal from the V_{rot} centroid location. With that said, it is not clear whether the added complexity of the analytical model approach results in any additional skill in depicting near-ground wind speeds.

Lastly, an estimate of the variation in damage path width is needed to depict the spatial extent of damaging winds at ground level. The approach considered thus far is to utilize the same approach used operationally by the NWS to generate preliminary tornado damage path polygons for strong to violent tornadoes. These polygons are based on the method developed by Speheger and Smith (2006), and are intended to give an approximate 85% elliptical confidence interval for the depiction of the tornado track center relative to the V_{rot} centroid. Future efforts may consider an alternative or additional approach that uses historical tornado statistics relating maximum path width to peak tornado intensity.

Table 1. Distance criteria used for composing operational NWS preliminary tornado damage path polygons. Table is from Karstens et al. (2016).

Distance Criteria	Buffer Radius Equation
0 km $< d <$ 64.4 km (0 mi $< d <$ 40 mi)	buffer radius = 0.48 km + $d \times$ 0.005 (buffer radius = 0.3 mi + $d \times$ 0.005)
64.4 km $\leq d <$ 128.7 km (40 mi $\leq d <$ 80 mi)	buffer radius = 0.16 km + $d \times$ 0.01 (buffer radius = 0.1 mi + $d \times$ 0.01)
128.7 km $\leq d$ (80 mi $\leq d$)	buffer radius = $d \times$ 0.015 - 0.48 km (buffer radius = $d \times$ 0.015 - 0.3 mi)

To create a grid of wind speeds, the variational width relative to the track centerline is combined with the distance between consecutive V_{rot} centroids, and used as normalizing values for conducting distance weighting in the interior of the track. These weights are applied to bounding cross-sections of maximum wind speeds obtained from the analytical vortex model (with V_{rot} implied translation speed), thus revealing a smoothed depiction of wind speed variation at ground level. The model run time varies, depending on track length, but in general it can be executed in under 10-15 minutes (on a laptop) with a given set of V_{rot}/STP/Population data. Currently, the model grid is configured at 20 m spacing, however, this dimension is configurable. Future efforts may consider a slightly larger grid spacing to ease the computational demand associated with long-tracked tornado events.

2.2 *Verification*

To explore the feasibility of providing detailed preliminary estimates of tornado wind speeds at ground level using the method just described, peak observed damage-based wind speeds from 6425 DIs associated with 115 tornadoes occurring between 2020-2022 were obtained from the NWS Damage Assessment Toolkit (DAT; Camp et al. [2014]). These events were selected based on the availability of DIs that were present in the DAT for each event (Smith et al. 2022). As discussed in Smith et al. (2020a), a critical assumption in using DI wind speeds is the notion that such observations represent the maximum wind speed that occurred at the DI location. In reality, damage-based wind speeds represent 1) a lower-bound to the maximum wind speed, and 2) an expert-assessed estimate of the peak wind speed (i.e., not physically measured). With that said, ascertaining the true maximum wind speed at various points along a tornado track is an impossible endeavor. Even in situations where instrumentation has fortuitously sampled tornadoes, such instruments often do not survive the harsh, debris-ridden environment of a tornado. Thus, damage-based wind speeds remain the only feasible representation of the true wind speeds that are available in the current era. In the future, it would be advantageous to incorporate additional post-event methods of ascertaining tornado wind speeds as such methods become available (LaDue et al. 2022).

As previously mentioned, only the deterministic percentile-based wind speed tracks were evaluated in this study. Two approaches were taken to evaluate the quality, and to some extent the value, of the statistical model. The first is an evaluation of the variation in raw error between the model diagnostic maximum wind speed and the observed, damage-based DI wind speed. This was done for selected percentiles to evaluate the variation in error spanning the entirety of the percentile-based tracks. The second is an evaluation of the variation in diagnostic reliability to address the question of how good the statistical model is at depicting winds that were observed. This was evaluated in wind speed ranges corresponding to the EF-scale categories.

Both of these evaluations are conditional upon the statistical model depiction of wind speed > 0 mph. In addition, both of these evaluations were performed at each DI grid point, as well as evaluating the maximum diagnostic wind speed from the model within 500 m and 1000 m radii (i.e., neighborhoods) of the DI grid point. These latter two approaches were done to assess the value of the model's ability to depict wind speeds associated with DIs in which intense to violent damage occurred (i.e., EF3+), thus allowing some forgiveness for depicting strong winds in the vicinity of these DIs in a manner that is perhaps similar to its intended usage in a preliminary sense. In an ideal scenario, the verification technique would employ the same observational data (polygons as opposed to DI points) as used in Saba et al. (2022). However, such data were available for only a limited number of the 115 tornadoes used in this study, and thus, this type of analysis could not be conducted.

3. RESULTS

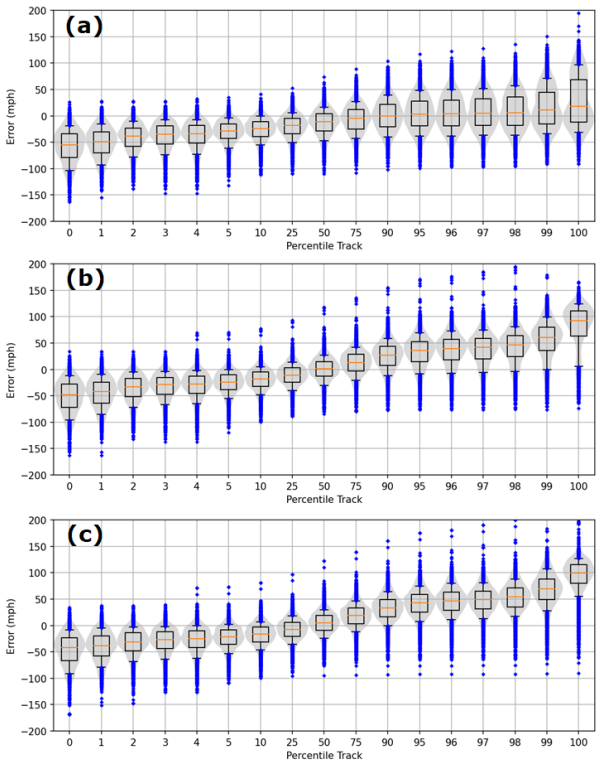

Figure 2. Violin plots of raw diagnostic wind speed error (mph; model estimated - damage estimated) for selected percentile-based tracks, evaluated (a) at each DI grid point, within (b) 500 m, and (c) 1000 m of each DI grid point.

The variation in raw diagnostic wind speed error (Fig. 2) demonstrates a general (and expected) trend of under estimation for low percentile tracks and over estimation for high percentile tracks. The error distributions center near zero around the 90th, 50th, and between the 25th and 50th percentile tracks when considering the maximum diagnostic wind speed depicted at each DI grid point (Fig. 2a), within 500 m of the DI (Fig. 2b), and within 1000 m of the grid point (Fig. 2c), respectively. In general, the error distributions exhibit an interquartile range of approximately 30 mph, which is about the range associated with each EF-scale category. The interquartile ranges increase with the lowest/highest percentile tracks, and decrease with larger neighborhoods.

From the results in Fig. 2, three percentiles were chosen for further evaluation in terms of diagnostic reliability of maximum wind speeds. These include the 50th, 75th, and 97th percentiles, which span the under/over diagnosis biases evident in Fig. 2 across the various verifying neighborhoods, with a slight subjective favoring of over-diagnosis due to the asymmetric penalty that is commonly associated with decision-making under uncertainty (Doswell 2004).

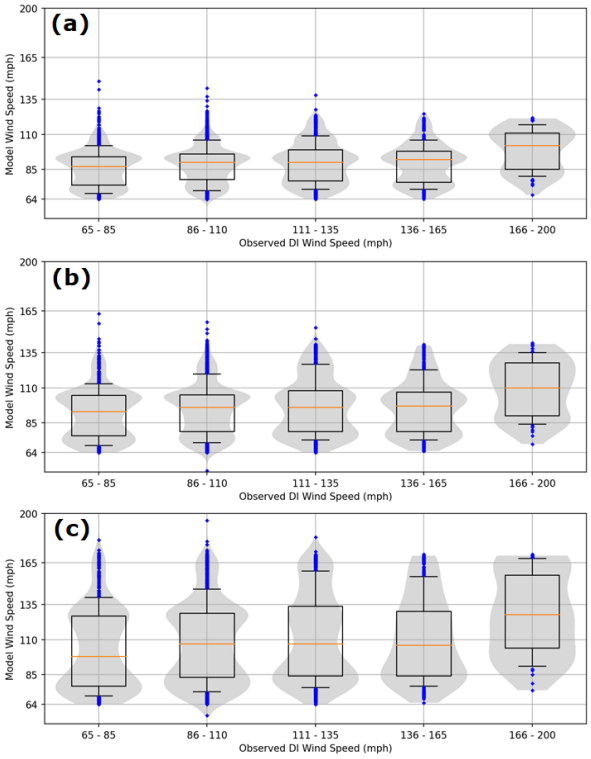

Figure 3. Violin plots of diagnostic reliability evaluated for the 50th percentile tracks (a) at each DI grid point, (b) within 500 m, and (c) within 1000 m of each DI grid point.

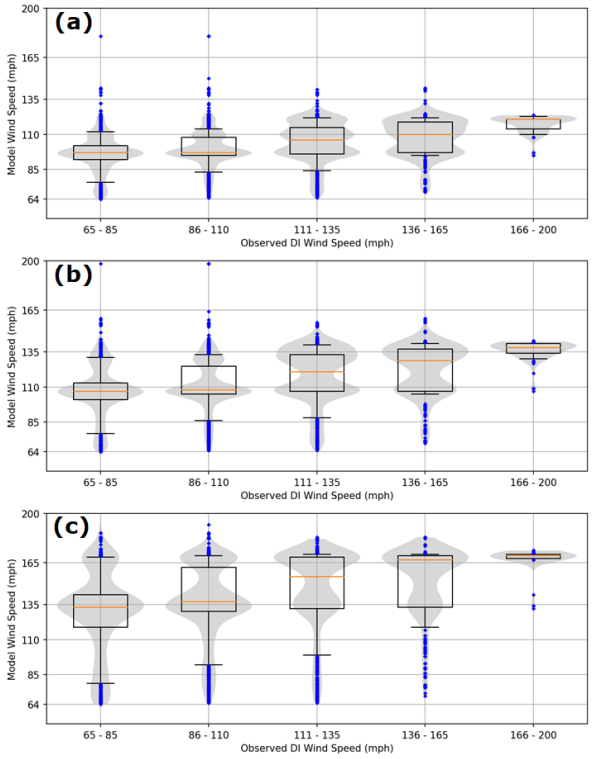

Figure 4. As in Fig. 3, except for the 75th percentile tracks.

The variation in diagnostic reliability for the 50th percentile tracks (Fig. 3) shows that for lower-range maximum wind speeds (i.e., EF0-EF1), the model is able to depict wind speeds at the DI grid points accurately (Fig. 3a). However, as DI wind speeds increase, the 50th percentile track does not produce a corresponding increase in maximum winds, but more or less similar lower-end wind speeds. An under-diagnosis bias for the 50th percentile tracks is also evident in Fig. 2a. This linear bias holds when considering the maximum diagnostic wind speed within 500 m (Fig. 3b) and 1000 m (Fig. 3c) but with an expanding upper tail to the distributions, resulting in a slight over/under diagnosis bias for lower/upper range wind speed distributions, respectively, and thus, total error distributions centered near zero (Figs. 2b and 2c). A notable upward shift of the distributions corresponding to DI wind speeds in the EF4 range can be noted as well in Figs. 3b and 3c, but the under-diagnosis bias is still present. This result implies that the 50th percentile tracks are able to depict lower-end maximum wind speeds quite accurately, particularly along the edges of the damage path where maximum wind speeds typically peak in the lower-range of the EF-scale, regardless of tornado intensity.

The variation in diagnostic reliability for the 75th percentile tracks (Fig. 4) shows trends similar to those in Fig. 3. but with accuracy for middle-range maximum wind speeds (i.e., EF2-EF3) at the DI grid points (Fig. 4a). Given the linear bias, there is an over-diagnosis of lower-range maximum wind speeds, and an under-diagnosis of higher-range maximum wind speeds at DI grid points, resulting in a very slight under-diagnostic bias (Fig. 2a). The most accurate depiction of maximum wind speeds occurs in the EF1 range at the DI grid points (Fig. 4a), EF2 range within 500 m of the DI grid point (Fig. 4b), and EF3/EF4 within 1000 m of the DI grid point (Fig. 4c). As the range of maximum wind speeds increases, the 75th percentile tracks tend to shift to an over-diagnostic bias (Figs. 2b and 2c) as the distributions in Figs. 4b and 4c shift upward, relative to those in Fig. 4a. These results support the notion that the 75th percentile tracks are most accurate at depicting middle-range maximum wind speeds at ground level, with perhaps some utility for upper-range maximum wind speeds when considering a large neighborhood.

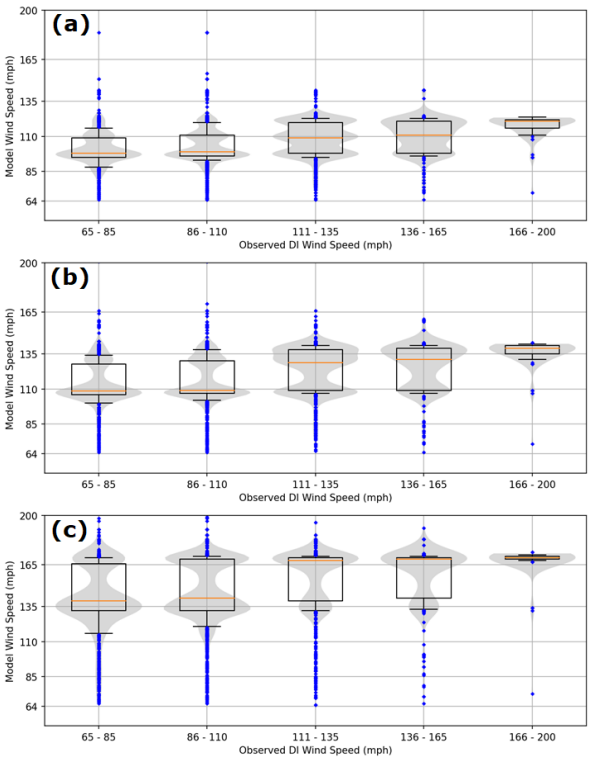

Figure 5. As in Fig. 3, except for the 97th percentile tracks.

Finally, the variation in diagnostic reliability for the 97th percentile tracks (Fig. 5) shows trends similar to those evident in Figs. 3 and 4. Similar to the 75th percentile tracks, the 97th percentile tracks exhibit accuracy in the EF1 range at the DI grid points (Fig. 5a), EF2 range within 500 m of the DI grid points (Fig. 5b), and EF3/EF4 range within 1000 m of the DI grid points (Fig. 5c). In general, an over-diagnostic bias is evident across all verifying neighborhoods in Fig. 2, although the 97th percentile distribution in Fig. 2a is centered very near zero. Comparing the distributions in the DI wind speeds in the EF3/EF4 range from the 75th and 97th percentile tracks, it is evident that the distributions from the 97th percentile tracks have a bit less spread, but at the expense of the aforementioned over-diagnostic bias in general. Thus, the 97th percentile tracks show accuracy for upper-range maximum wind speeds, but limited utility beyond that.

Example tracks for the 50th, 75th, and 97th percentiles are given in Figs. 6, 7, and 8 for three different tornado events in the evaluation dataset. These examples give a visual depiction of the representativeness of each track relative to the DIs present for each event and the neighborhoods used in this study.

4. DISCUSSION AND FUTURE WORK

The primary goal of this effort is to characterize the uncertainty of potential tornado intensity within a time frame that addresses the needs of incident response personnel. The results indicate that the present configuration of the statistical model has an ability to depict the variation of peak tornado intensity at ground level within approximately +/- 15 mph (i.e., approximately 1 EF-scale wind speed range) at varying neighborhoods from verifying DIs, when considering three deterministic outputs under a persistent linear diagnosis bias present for each track. To align the aforementioned goal and the results of this work with ongoing efforts occurring within the NWS to deliver similar preliminary characterizations of event magnitude/intensity (Waldstreicher et al. 2018), the three deterministic outputs can be categorized in the following manner.

1) Best-case scenario: 50th percentile track. Accuracy for lower-range maximum wind speeds but a low bias for depicting middle- and upper-range maximum wind speeds.

2) Most-likely scenario: 75th percentile track. Accuracy for middle- and upper-range maximum wind speeds, but with a high bias for depicting lower-range maximum wind speeds, and a slight low bias for depicting upper-range maximum wind speeds.

3) Worst-case scenario: 97th percentile track. Accuracy for upper-range maximum wind speeds, but with a high bais for depicting low- and middle-range maximum wind speeds.

These three percentiles were chosen to span the under/over diagnosis biases evident in Fig. 2 across the various neighborhoods from verifying DIs, with a slight subjective favoring of over-diagnosis due to the asymmetric penalty that is commonly associated with decision-making under uncertainty. The variation in diagnostic reliability results indicates that the 50th and 75th percentile tracks offer the best utility in estimating peak near ground wind speeds, on the whole, whereas the 97th percentile track is of limited additional value except in events in which upper-range maximum wind speeds occur.

Regardless of the subjectivity introduced to favor a slight over-diagnosis bias, it is interesting that the deterministic output that characterizes the uncertainty favors higher percentile values (50, 75, and 97) as opposed to percentiles that are statistically centered on the empirical distributions (e.g., 25, 50, 75). The most likely explanation for this effect is the process of adding spatial and temporal dimensionality to the empirically derived estimates of peak near-ground wind speed. This process necessarily introduces depictions of near-ground winds that are of lesser magnitude other than a very narrow corridor where the peak winds are represented at ground level. This result could serve as a basis for further investigating the representations of damage width, and perhaps the analytical model configurations. Future efforts will also include an investigation into probabilistic depictions of exceeding various wind speed thresholds, likely those

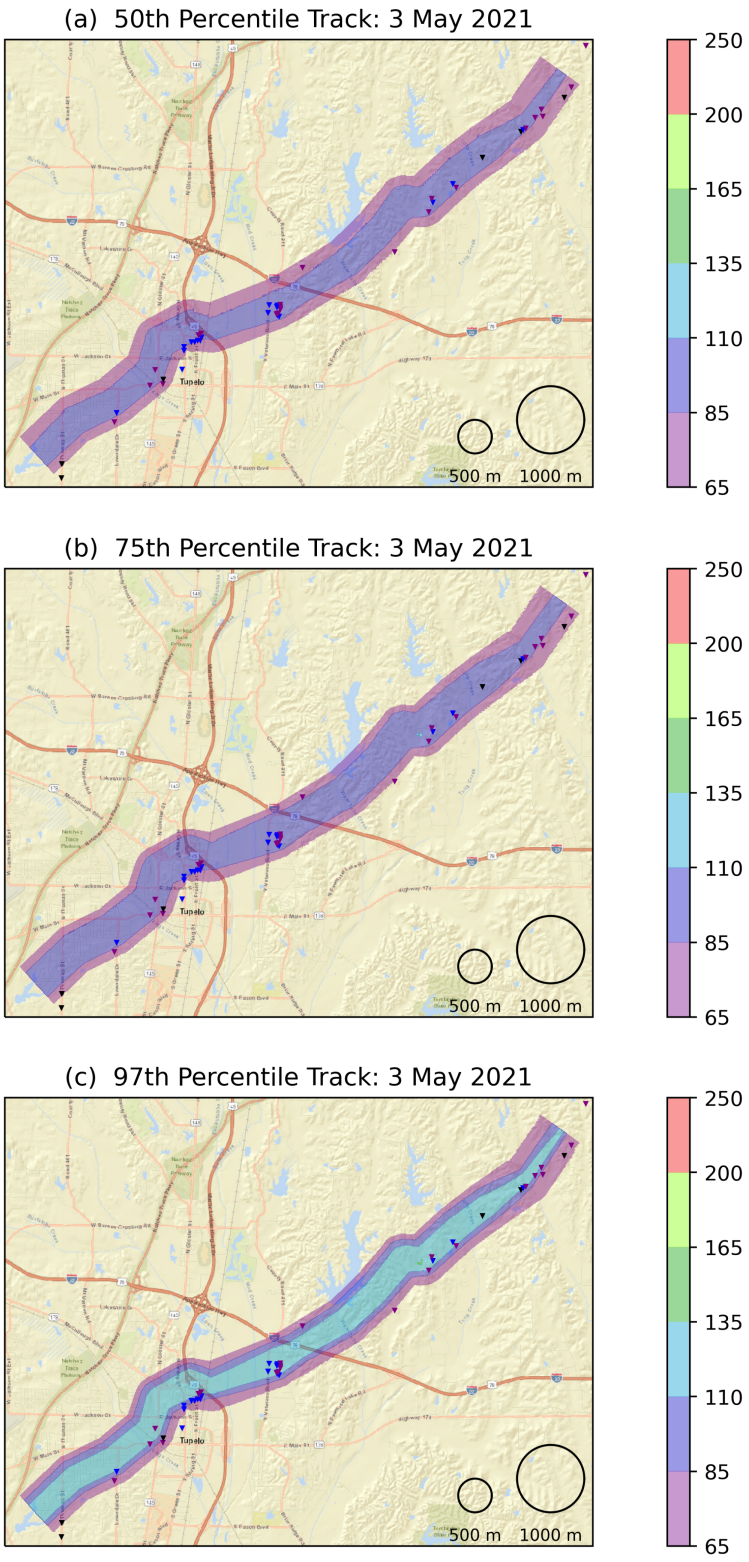

Figure 6. Example (a) 50th, (b) 75th, and (c) 97th percentile tracks for a tornado event on 3 May 2021. Colors correspond to wind speeds at ground level, using EF-scale categorical thresholds. Triangles represent DIs available for this event, colored using the same scheme as the model wind speeds. The 500 m and 1000 m neighborhoods used in this study are provided in the lower-right corner of each panel as a reference.

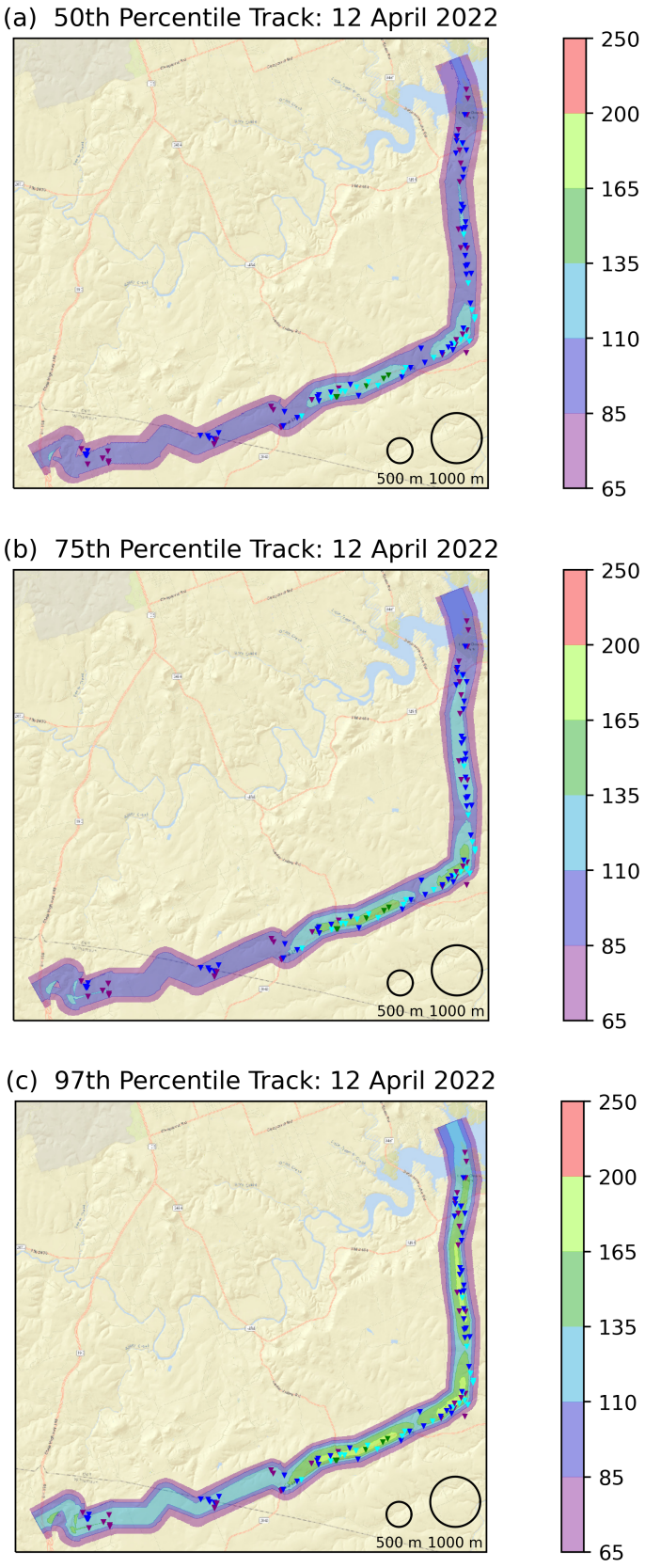

Figure 7. As in Fig. 6, except for a tornado event on 12 April 2022.

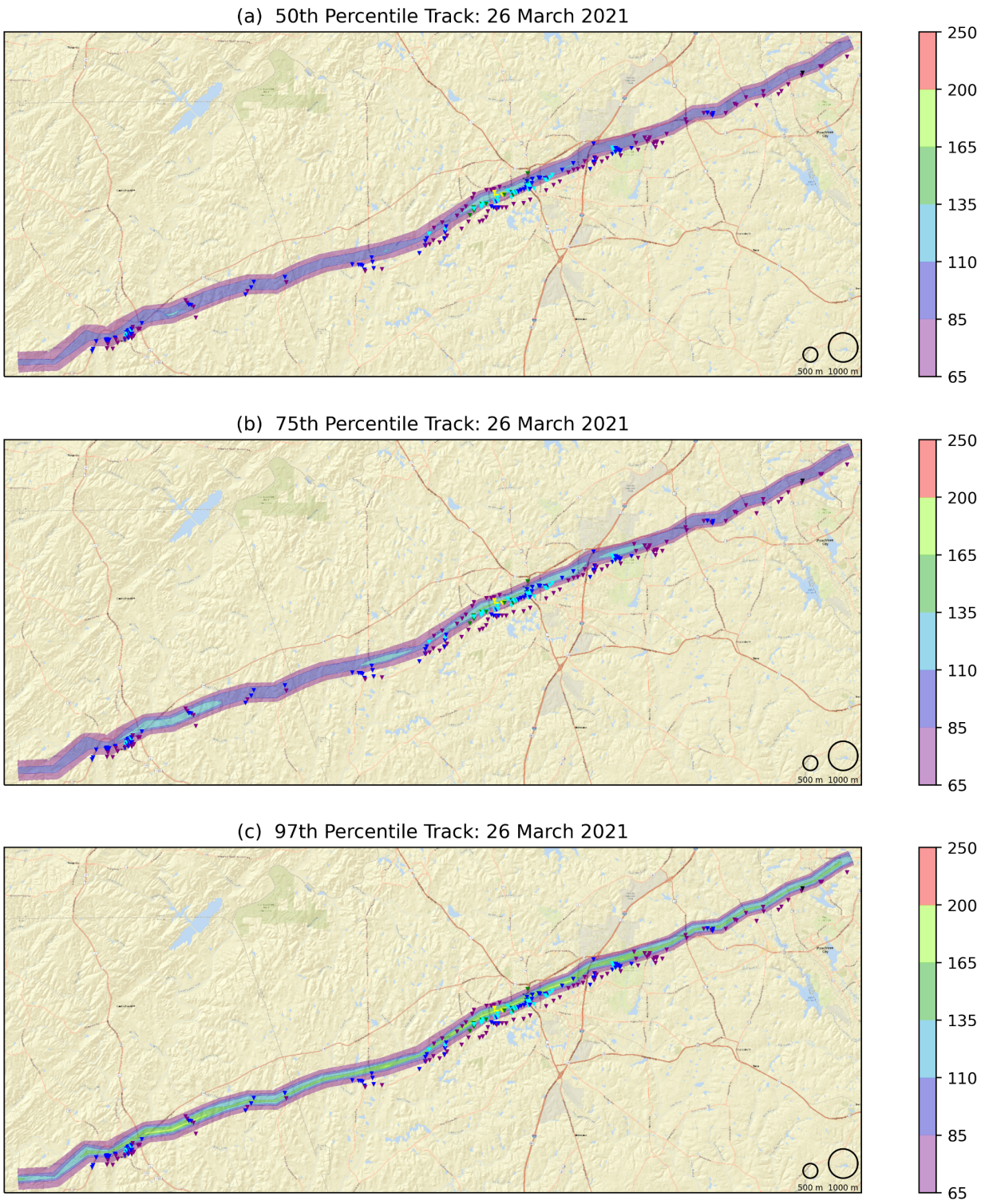

Figure 8. As in Fig. 6, except for a tornado event on 26 March 2021.

corresponding to the EF-scale categories. Saba et al. (2022) investigated such depictions for a limited sample of events, and thus, this approach will be expanded to include the 115 events used in the present work.

5. ACKNOWLEDGEMENTS

In addition to the list of coauthors, there have been several people who have provided valuable input to this project. These people include Patrick Marsh, Jim LaDue, Doug Speheger, and Harold Brooks.

6. REFERENCES

Brown, R. A., 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.

Burgess, D., and Coauthors, 2014: 20 May 2013 Moore, Oklahoma, tornado: Damage survey and analysis. *Wea. Forecasting,* **29,** 1229–1237.

Camp, J. P., L. P. Rothfusz, A. Anderson, D. Speheger, K. L. Ortega, and B. R. Smith, 2014: Assessing the Moore, Oklahoma (2013) tornado using the National Weather Service Damage Assessment Toolkit. Preprints, *Special Symposium on Severe Local Storms: The Current State of the Science and Understanding Impacts,* Atlanta, GA, Amer. Meteor. Soc., 830.

Doswell, C. A., 2004: Weather forecasting by humans—Heuristics and decision making. *Wea. Forecasting,* **19,** 1115–1126.

Istok, M. J., and Coauthors, 2009: WSR-88D dual polarization initial operational capabilities. *25th Conf. on Int. Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Phoenix, AZ, Amer. Meteor. Soc. 15.5.

Karstens, C. D., W. A. Gallus Jr., B. D. Lee, and C. A. Finley, 2013: Analysis of tornado-induced tree fall using aerial photography from the Joplin, Missouri, and Tuscaloosa–Birmingham, Alabama, tornadoes of 2011. *J. Appl. Meteor. Climatol.,* **52,** 1049–1068.

Karstens, C. D., K. Shourd, D. Speheger, A. Anderson, R. Smith, D. Andra, T. M. Smith, V. Lakshmanan, and S. A. Erickson, 2016: Evaluation of near real-time preliminary tornado damage paths, *J. of Oper. Meteor.,* **4,** 132-141.

LaDue, J. G., M. Levitan, T. M. Brown-Giammanco, D. W. Burgess, D. B. Roueche, J. Wurman, K. A. Kosiba, F. T. Lombardo, C. D. Karstens, and A. Merhi, 2022: Comparison of wind speed estimation methods applied to the Monroe, LA tornado of 12 April 2020. *30th Conf. on Severe Local Storms*, Santa Fe, NM, Amer. Meteor. Soc., 11.3A.

NWS, 2011: NWS Central Region service assessment: Joplin, Missouri, tornado. National Weather Service Service Assessment, 35 pp. [Available online at www.nws.noaa.gov/om/assessments/pdfs/Joplin_tornado.pdf.]

Saba, B., C. D. Karstens, B. T. Smith, R. L. Thompson, P. T. Marsh, D. A. Speheger, and J. G. LaDue, 2022: Verifying Tornado Intensity Added to NWS Preliminary Tornado Damage Paths. *10th Symposium on Building a Weather-Ready Nation:Enhancing Our Nation's Readiness, Responsiveness, and Resilience to High Impact Weather Events,* Houston, TX, Amer. Meteor. Soc., 363.

Scharfenberg, K. A., and Coauthors, 2005: The Joint Polarization Experiment: Polarimetric radar in forecasting and warning decision making. *Wea. Forecasting,* **20,** 775–788.

Smith, B.T., R.L. Thompson, 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.

Smith, B. T., R. L. Thompson, D. A. Speheger, A. R. Dean, C. D. Karstens, and A. K. Anderson-Frey, 2020b: WSR-88D tornado intensity estimates. Part II: Real-time applications to tornado warning time scales. *Wea. Forecasting,* **35,** 2493–2506.

Smith, B. T., R. L. Thompson, C. D. Karstens, J. Grams, A. Dean, and R. M. Mosier, 2022: Preliminary evaluation of a real-time diagnostic tornado damage intensity estimation tool used at the Storm Prediction Center. *30th Conf. on Severe Local Storms*, Santa Fe, NM, Amer. Meteor. Soc., 17.4B.

Speheger, D. A., and R. D. Smith, 2006: On the imprecision of radar signature locations and storm path forecasts. *Natl. Wea. Dig.,* **30,** 3–10.

Waldstreicher, J. S., D. R. Radell, and P. Schumacher, 2018: The NWS probabilistic snowfall forecast experiment: 2016−17 review and future plans. *13th Symposium on Societal Applications: Policy, Research and Practice,* Austin, TX, Amer. Meteor. Soc., 8.3.

WSEC, 2006: A recommendation for an enhanced Fujita scale (EF-Scale), revision 2. Wind Science and Engineering Center Rep., Texas Tech University, Lubbock, TX, 95 pp. [Available online at http://www.depts.ttu.edu/nwi/pubs/fscale/efscale.pdf.]