Tornado-Resolving Ensemble and Probabilistic Predictions of the 20 May 2013 Newcastle–Moore EF5 Tornado

NATHAN SNOOK

Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma

MING XUE

Center for Analysis and Prediction of Storms, and School of Meteorology, University of Oklahoma, Norman, Oklahoma

YOUNGSUN JUNG

Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma

ABSTRACT

An ensemble of 10 forecasts is produced for the 20 May 2013 Newcastle–Moore EF5 tornado and its parent supercell using a horizontal grid spacing of 50 m, nested within ensemble forecasts with 500-m horizontal grid spacing initialized via ensemble Kalman filter data assimilation of surface and radar observations. Tornadic circulations are predicted in all members, though the intensity, track, and longevity of the predicted tornado vary substantially among members. Overall, tornadoes in the ensemble forecasts persisted longer and moved to the northeast faster than the observed tornado. In total, 8 of the 10 ensemble members produce tornadoes with winds corresponding to EF2 intensity or greater, with maximum instantaneous near-surface horizontal wind speeds of up to $130 \, \mathrm{m \, s^{-1}}$ and pressure drops of up to 120 hPa; values similar to those reported in observational studies of intense tornadoes. The predicted intense tornadoes all acquire well-defined two-cell vortex structure, and exhibit features common in observed tornadic storms, including a weak-echo notch and low reflectivity within the mesocyclone. Ensemble-based probabilistic tornado forecasts based upon near-surface wind and/or vorticity fields at 10 m above the surface produce skillful forecasts of the tornado in terms of area under the relative operating characteristic curve, with probability swaths extending along and to the northeast of the observed tornado path. When probabilistic swaths of 0–3- and 2–5-km updraft helicity are compared to the swath of wind at 10 m above the surface exceeding $29 \, \mathrm{m \, s^{-1}}$, a slight northwestward bias is present, although the pathlength, orientation, and the placement of minima and maxima show very strong agreement.

1. Introduction

To increase the lead time and accuracy of tornado warnings, it is essential that we shift the paradigm of issuing tornado warnings from one based on extrapolation from current observations (warn-on-detection) to one based on short-term, high-resolution numerical weather prediction (NWP) model forecasts (Warn-on-Forecast; Stensrud et al. 2009, 2013). The current warn-on-detection paradigm has enabled warning lead times of about 14 min on average for tornadoes, but such lead times are likely near the upper limit (Stensrud et al.

2013; Xue et al. 2014). Because forecast errors grow very quickly, and large uncertainty exists in NWP forecasts at the scales necessary to resolve tornadoes, ensemble forecasts are necessary for capturing the possible range of intensity, location and timing of predicted tornadoes, as well as their uncertainties.

Because of the very high computational costs of explicitly resolving tornadoes (or tornado-like vortices) within NWP models, much recent work has focused on the prediction of the parent storms of tornadoes, especially tornadic supercells. Such studies have mostly focused on data assimilation techniques, forecast model configurations, and proxy products (e.g., Dowell et al. 2004; Hu et al. 2006; Hu and Xue 2007; Jung et al. 2012;

Dawson et al. 2012; Clark et al. 2012, 2013; Tanamachi et al. 2013; Snook et al. 2016; Supinie et al. 2016; Skinner et al. 2018; Flora et al. 2018). These studies have demonstrated that, with proper assimilation of radar observations, skillful prediction of tornadic supercells as well as mesocyclone circulations within tornadic supercells is possible on time scales of up to a few hours. Similar successes have been achieved with tornadoes developing within a mesoscale convective system, where mesovortices of similar spatial scales develop in place of mesocyclones (Schenkman et al. 2011; Snook et al. 2012; X. Xu et al. 2015). The aforementioned studies typically use horizontal grid spacing ranging from 500 m to 3 km. Such grids are capable of simulating the tornadic parent storms but insufficient to capture the circulations of a tornado or tornado-like vortex; for these, a horizontal grid spacing of 100 m or less is generally necessary.

Previous tornado simulations have often used idealized model configurations in which a single sounding is used to define a horizontally homogeneous storm environment while the storm is initialized by a thermal bubble or other artificial means (e.g., Wicker and Wilhelmson 1995). Snook and Xue (2008) simulated tornado-like vortices in idealized supercells using 100-m horizontal grid spacing and studied the impact of microphysical drop size distribution parameters on tornadogenesis. Lerach et al. (2008), another idealized simulation study, used 111-m grid spacing and focused on aerosol influences on tornadogenesis. More recently, the single-sounding-based simulation studies of Roberts et al. (2016) and Roberts and Xue (2017) using a 50-m grid spacing examined the effects of surface drag and identified vorticity generated from surface drag as the primary source of vorticity for mesocyclone and tornadogenesis in their simulations.

Hu and Xue (2007) and Xue et al. (2014) were among the first studies to produce realistic forecasts of observed tornadoes—specifically, an F4 intensity tornado that occurred on 8 May 2003 in a supercell thunderstorm near Oklahoma City, Oklahoma. Using multiple nested grids, and a 50-m horizontal grid spacing on the innermost grid, their simulations produced two tornadoes which tracked within 10 km of the observed tornado path. By analyzing the simulation data of Xue et al. (2014), Schenkman et al. (2014) found that surface drag was an important source of low-level vorticity for the developing tornado. Building upon this work, Schenkman et al. (2016) further examined the origins of internal outflow surges within the 8 May 2003 simulation dataset, and found that these surges played a key role in instigating tornadogenesis. Using a similar simulation methodology, Schenkman et al. (2012) successfully produced simulations of tornadoes forming within

a mesoscale convective system using 100-m grid spacing that allowed for investigation of tornadogenesis dynamics. The simulations of these studies assimilated Doppler weather radar data for storm initialization.

Starting from initial conditions from an operational mesoscale forecasting system, Mashiko et al. (2009) obtained successful simulations of a tornado associated with a supercellular storm within a landfalling tropical cyclone; quadruply nested grids were used to reach a 50-m grid spacing over the tornadic storm. More recently, Mashiko (2016a, b) obtained reasonable simulations for a supercell tornado of F3 intensity that occurred in Tsukuba, Japan, using 50-m grid spacing, again starting from operational mesoscale analysis. Most recently, Yokota et al. (2018) produced ensemble forecasts at 50-m grid spacing for tornadoes occurring within a supercell storm over Japan, focusing on tornadogenesis dynamics.

The tornado-resolving simulation studies discussed above have, with the exception of Yokota et al. (2018), all been based on deterministic prediction of tornadoes or tornado-like vortices, and have primarily focused on investigation of processes responsible for tornadogenesis within individual tornado forecasts. While deterministic forecasts can provide useful insights into dynamical and thermodynamic processes occurring within the model storm, Warn-on-Forecast predictions for tornadoes will require an ensemble-based probabilistic approach to forecasting in order to properly characterize the uncertainty that tends to grow very fast in forecasts at the subkilometer scale (Lorenz 1969). With this in mind, in this study we produce and verify ensemble forecasts using a 50-m tornado-resolving grid spacing for the Newcastle–Moore EF5 tornado of 20 May 2013 and its parent supercell. This ensemble is nested within the storm-scale ensemble forecasts of Snook et al. (2016), which used a 500-m horizontal grid spacing with initial conditions from ensemble Kalman filter (EnKF) data assimilation of radar and conventional weather observations. To our knowledge, the current study represents the first attempt to produce and evaluate ensemble predictions of real tornadoes, at a resolution sufficient to explicitly resolve tornadoes. Prediction of tornadoes using ensembles capable of directly resolving tornadoes avoids the need for relying on proxy parameters (such as UH) to infer the potential of tornadoes when the models are incapable of resolving tornadoes directly. Furthermore, when tornadoes are directly resolved within the forecasts, we will be able to investigate the actual tornado formation, maintenance and decay processes, and the predictability of actual tornadoes, including their timing, intensity, track length, etc. (Sun et al. 2018,
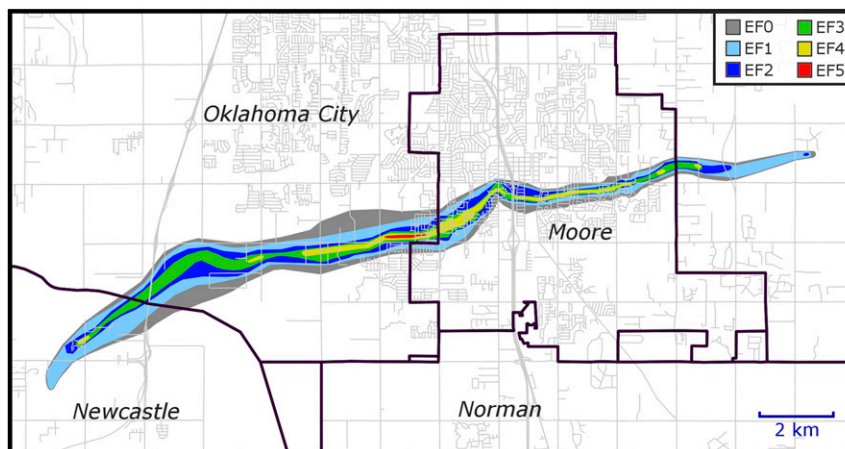
FIG. 1. Approximate path and intensity of the 20 May 2013 Newcastle–Moore EF5 tornado based upon the results of a National Weather Service damage survey. Tornado intensity, in terms of EF scale rating, is indicated by color shading. City borders are plotted as thick black lines—relevant cities have been labeled—and roads are shown in light gray.

manuscript submitted to *Geophys. Res. Lett.*). Establishing the feasibility of explicit ensemble probabilistic prediction of tornadoes is an important goal of this paper, which is also the vision and long-term goal of the NOAA Warn-on-Forecast project (Stensrud et al. 2009).

The remainder of this paper is organized as follows: section 2 contains a brief overview of the 20 May 2013 Newcastle–Moore EF5 tornado case. Section 3 details the experiment design and verification techniques used in the study. Results of the ensemble forecast experiment are presented and ensemble-based probabilistic forecast products are verified in section 4, and a general summary and discussion are presented in section 5.

2. A brief overview of the 20 May 2013 Newcastle–Moore EF5 tornado

Early in the afternoon of 20 May 2013, a 500-hPa trough was positioned from eastern South Dakota extending to northeastern New Mexico. Conditions over central Oklahoma were favorable for the development of supercell storms; a sounding from Norman, OK (KOUN), taken at 1800 UTC indicated mixed-layer CAPE exceeding $3000\,\mathrm{J\,kg^{-1}}$, along with 0–1-km storm-relative helicity of approximately $130\,\mathrm{m^2\,s^{-2}}$. A distinct dryline was present near and west of Interstate 35 in central and southwest Oklahoma. Severe weather, including tornadoes, was expected by operational forecasters; the 1630 UTC Storm Prediction Center convective outlook included a moderate risk region with a 10% or greater probability of EF2–EF5 tornadoes along and east of the dryline in Oklahoma.

Between 1830 and 1930 UTC, several supercell storms developed and rapidly intensified ahead of the dryline. The supercell storms moved to the northeast, producing reported severe hail and several tornadoes as they passed over central Oklahoma. The most damaging of the tornadoes produced on 20 May 2013 was an EF5 tornado that passed through the cities of Newcastle, Moore, and Oklahoma City in central Oklahoma.

The Newcastle–Moore EF5 tornado began near Newcastle at approximately 1956 UTC, and moved initially northeastward and then east-northeastward through Oklahoma City and Moore, eventually dissipating in a rural area east of Moore at 2035 UTC (Burgess et al. 2014). Both the tornado and its parent storm were well-observed by NWS WSR-88D KTLX at Twin Lakes, Oklahoma, which is located approximately 20 km east of Moore. The damage path of the tornado was also observed in great detail by a National Weather Service storm survey of the tornado damage; based on the results of this survey, the tornado traversed a path 23 km in length with a maximum width of over 1700 m, and resulted in 24 fatalities, 212 injuries, and thousands of damaged or destroyed structures (Burgess et al. 2014). The path and intensity of the tornado are shown in Fig. 1. An in-depth discussion of the tornado and the NWS damage survey can be found in Burgess et al. (2014); additional information on the storms of 20 May 2013 can also be found in Zhang et al. (2015).

3. Ensemble design and methodology

In this study, an ensemble of tornado-resolving forecasts is run using 50-m grid spacing; the 50-m forecasts

are one-way nested within an ensemble of forecasts at a coarser 500-m horizontal grid spacing. The 500-m ensemble consists of 40 members, while the 50-m ensemble, due to the high computational cost, consists of ten members nested within selected members of the 500-m ensemble. Both ensembles are run using the Advanced Regional Prediction System (ARPS) (Xue et al. 2000, 2001). The ARPS ensemble Kalman filter (EnKF) (Xue et al. 2006; Tong and Xue 2008; Wang et al. 2013) is used to perform data assimilation (DA) on the 500-m grid. The specifics of the 500- and 50-m grids are discussed in detail below.

As described in Snook et al. (2016), EnKF DA cycles are performed on the 500-m grid, and are initialized at 1800 UTC 20 May 2013 from the 0000 UTC 20 May 2013 experimental 40-member storm-scale ensemble forecasts of CAPS (SSEFs) (Kong et al. 2014) that were run in real time as part of the 2013 NOAA Hazardous Weather Testbed (HWT) spring experiment. The 2013 CAPS SSEFs used version 3.4.1 of the WRF-ARW (Skamarock et al. 2008). The CAPS SSEF used multiple planetary boundary layer schemes to help increase ensemble spread, and the initial and boundary condition perturbations of the CAPS SSEF came from the operational Short-Range Ensemble Forecasting (SREF) system; details regarding specific model settings and member configurations for the 2013 CAPS SSEFs can be found in CAPS (2013). Lateral boundary conditions for the 500-m grid during the DA cycles and subsequent forecasts are also provided by the CAPS SSEFs.

The 500-m domain has $603 \times 653 \times 63$ grid points and covers portions of Oklahoma and far northern Texas (Fig. 2). The model grid is stretched in the vertical, with a minimum vertical grid spacing of 50 m near the surface and an average vertical spacing of 425 m. Other settings are chosen following Snook et al. (2015). Physics options within the ARPS model are identical for all members, and include the NASA Goddard Space Flight Center long- and shortwave radiation parameterization, a combination of a 1.5-order Deardorff (1980) type subgrid-scale turbulence closure scheme and the 1D convective boundary layer turbulence closure scheme of Sun and Chang (1986), a two-layer soil model, and stability-dependent surface flux parameterization using predicted surface temperature and soil moisture content. The 1D boundary layer turbulence closure scheme is only active when the planetary boundary layer is convectively unstable. While this scheme should not be necessary on the 50-m grid, including it makes the transition of boundary layer solutions across the 50- and 500-m boundaries more smooth. Surface terrain is from the USGS terrain elevation model with a resolution of 30 arc s. All ensemble members use the two-moment
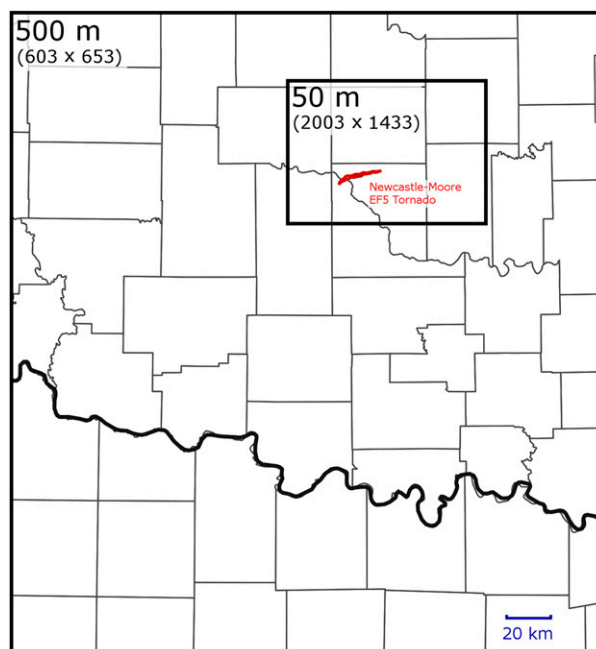

FIG. 2. Spatial extent of the 500- and 50-m grids. The number of horizontal grid points in each domain is noted (603×653 grid points for the 500-m grid, and 2003×1433 grid points for the 50-m grid). The track of the 20 May 2013 Newcastle–Moore EF5 tornado is shown in red.

version of Milbrandt and Yau (2005) microphysics scheme, on both grids during the EnKF DA cycles and subsequent forecasts. Details on the ARPS physics schemes can be found in Xue et al. (2001) and the performance of the Milbrandt and Yau (2005) for supercell storm simulations are documented in Dawson et al. (2010, 2015).

On the 500-m grid, the ARPS EnKF DA system is used to assimilate radar observations of radar reflectivity factor Z and radial velocity V_r, as well as available surface observations, including those from the Oklahoma Mesonet, at 5-min intervals from 1830 to 1930 UTC. Data from five WSR-88Ds are assimilated; the radar sites are: Oklahoma City, OK (KTLX); Vance Air Force Base, OK (KVNX); Frederick, OK (KFDR); Fort Worth, TX (KFWS); and Dyess Air Force Base, TX (KDYX). The ensemble analyses and forecasts from the 500-m grid generally capture the three primary supercell thunderstorms, including the Newcastle–Moore storm, reasonably well in terms of location, size, and evolution (Snook et al. 2016). For a detailed description of the EnKF settings used, as well as an in-depth discussion of analysis and forecast results from the 500-m ensemble, we refer the reader to Snook et al. (2016).

The domain extent and membership size of the 50-m ensemble are limited due to the extreme computational

expense of running NWP forecasts at such high resolution. For the configuration used in this study, running a 90-min forecast for a single member of the 50-m ensemble requires approximately 100 h of wallclock time on 2860 processors on the University of Texas Advanced Computing Center (TACC) Stampede2 supercomputer. For this reason, we chose 10 out of the 40 (in total) 500-m ensemble members to nest the 50-m ensemble forecasts within. The 500-m ensemble forecasts valid at 1950, 1955, 2000, and 2005 UTC (just before and during the initial stages of the observed tornado) were subjectively evaluated in terms of the intensity of near-surface circulation in the parent storm of the Newcastle–Moore tornado and its placement relative to the observed tornado track. To help maintain the ensemble spread, and to facilitate future investigations of practical tornado predictability, we selected five members which, based on the aforementioned evaluations, represented the structure and placement of the near-surface circulation of the Newcastle–Moore storm relatively well on the 500-m grid and five which did not represent those features as well to serve as initial and boundary conditions for the 50-m forecasts. In discussions of the 50-m ensemble, we will label these members as 1–10, with the five members initialized from "good" 500-m members being labeled 1–5, and the five members initialized from "poor" 500-m members being labeled 6–10. Though it would be better to have a larger ensemble so that more possibilities could be sampled and the probabilistic forecasting skill might increase (e.g., Clark et al. 2011), the resource constraints prevent this in the current study. We note that operational convection-allowing forecast ensembles, such as version 2 of the High Resolution Ensemble Forecast (HREFv2), have provided operationally useful forecasts of severe weather using approximately 10 members.

The 50-m ensemble is initialized at 1930 UTC, by interpolating the final 500-m EnKF analyses to the 50-m grid (the observed Newcastle–Moore EF5 tornado begins at 1956 UTC, 26 min after this initialization time). This timing for initialization of the 50-m grid is chosen to provide sufficient lead time for prediction of the Newcastle–Moore tornado. The 50-m grid has 2003 × 1433 × 63 grid points covering the Oklahoma City metropolitan area and portions of several nearby counties (Fig. 2)—the size and position of this grid were contain the Newcastle–Moore tornado and its parent storm to the best extent possible while remaining within the constraints of computational expense. Stretched vertical grid spacing is used, with a minimum vertical grid spacing of 20 m near the surface and an average vertical grid spacing of 340 m. Within the lowest kilometer above the surface, the average vertical grid spacing is 114 m. The smaller minimum vertical grid spacing at the surface than that of 500-m grid allows for better treatment of near surface processes, including frictionally generated vorticity that has been found to be important for tornadogenesis in our earlier studies (Schenkman et al. 2014; Roberts et al. 2016; Roberts and Xue 2017). Physics parameterization settings are the same as on the 500-m grid. On the 50-m grid, fourth-order computational mixing (operating in the ARPS model computational space) is used with a coefficient of $2.0 \times 50^4 \times 10^{-3} \, \text{m}^4 \, \text{s}^{-1}$ in the horizontal and $1.0 \times 340^4 \times 10^{-3} \, \text{m}^4 \, \text{s}^{-1}$ in the vertical to help suppress near-grid-scale noise. The 50-m forecasts are run for 90 min, ending at 2100 UTC (approximately 25 min after the dissipation of the observed Newcastle–Moore tornado). Model outputs are saved at one-minute intervals.

4. Forecast results

a. Overview of tornadoes predicted by the 50-m ensemble

Tornado-like vortices were predicted in all ten members of the 50-m ensemble, with eight members predicting tornadoes with winds exceeding the enhanced Fujita scale (American Meteorological Society 2013) EF2 threshold ($50 \, \text{m} \, \text{s}^{-1}$) and four members predicting tornadoes with winds exceeding the EF5 threshold ($90 \, \text{m} \, \text{s}^{-1}$). A paintball plot of wind exceeding the EF0 threshold ($29 \, \text{m} \, \text{s}^{-1}$) at the first model level (about 10 m) above the surface during the 90-min forecast period is presented in Fig. 3. As Fig. 3 plots any winds exceeding the EF0 threshold, regions of strong nontornadic winds appear in rare instances (such as the small north–south region located to the east of Norman, OK, in member 9), but the vast majority of the winds plotted in Fig. 3 are associated with tornadoes produced by the Newcastle–Moore storm. The discrete nature of the wind swaths in Fig. 3, particularly notable near the observed tornado track in members 3 and 4, and near the northeast end of the wind swaths of members 8, is an artifact of the one-minute data output interval. In all but one of the ensemble members, tornadic activity continues after the end of the observed tornado at 2035 UTC, with the bulk of the tornadic activity in the ensemble occurring between 2020 and 2100 UTC (Fig. 3). Coupled with faster than observed storm motion to the northeast within the ensemble, the tracks of many of the tornadoes predicted within the ensemble extend well to the northeast of the observed tornado track (Fig. 3). Also, the motion of the predicted tornadoes exhibits a somewhat more northward component than that of the observed tornadic
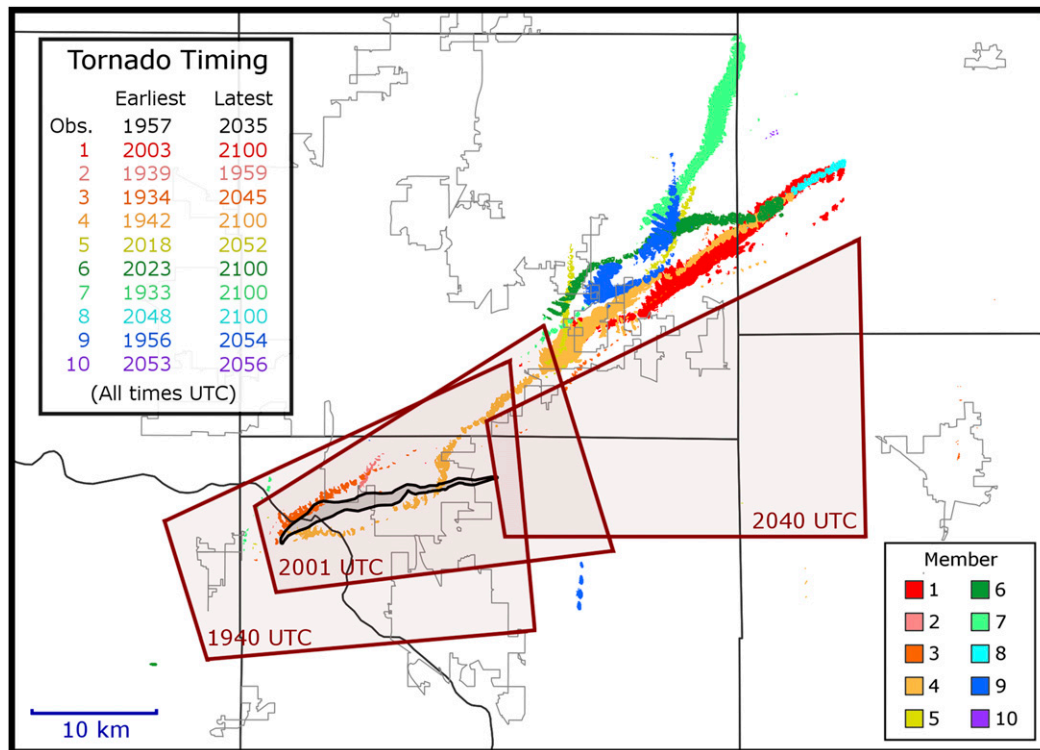
FIG. 3. Swaths of wind speed exceeding the EF0 threshold ($29 \, \mathrm{m \, s^{-1}}$) at the first model level above the surface for each of the ten members of the 50-m ensemble, color-coded by ensemble member. Urban areas are outlined in gray, county boundaries are denoted by thin black lines, and the observed extent of EF0 or greater damage from the Newcastle–Moore tornado is indicated by a thick black contour. Tornado warnings issued between 1930 and 2100 UTC by the NWS Norman WFO, labeled by time of issuance, are plotted (dark red boxes) for comparison. The earliest and latest times at which tornadoes were present for the observed tornado ("Obs.") and for each member ("1"–"10") are listed in the upper-left inset as a general reference for timing.

mesocyclone, as suggested by the locations of NWS warnings for the storm (Fig. 3).

The duration of tornadoes in each of the ensemble members, as well as that of the observed Newcastle–Moore tornado, is plotted in Fig. 4. The observed tornado duration is based on the times reported in the NWS damage survey (Burgess et al. 2014), while the duration of tornadoes in the ensemble members is defined to be all times at which wind speed exceeds $29 \, \mathrm{m \, s^{-1}}$ (EF0) and vertical vorticity exceeds $0.4 \, \mathrm{s^{-1}}$ in the vicinity of the supercell updraft at the first level (~10 m) above the model surface—this vorticity threshold was chosen empirically via preliminary evaluations of model forecast fields and structures. Duration of tornadoes within the ensemble ranges from 3 min (member 10) to more than an hour (member 4), with four ensemble members (1, 4, 5, and 6) predicting tornadoes that last at least as long as the observed tornado (about 40 min) (Fig. 4).

Multiple tornadic events associated with the Newcastle–Moore storm, in terms of the surface wind speed, are

predicted in seven of the 10 ensemble members (Fig. 4). Members 1, 2, 4, and 8 exhibit two separate tornadoes, members 7 and 9 each exhibit four, and member 3 exhibits five separate tornadic events. In each of the members exhibiting multiple tornadoes, a relatively strong near-surface circulation remains continuously present in the vicinity of the simulated hook echo between the tornadic events, but the wind speed associated with the circulation drops below the threshold of $29 \, \mathrm{m \, s^{-1}}$ used as a criterion for the presence of a tornado during the gaps indicated in Fig. 4. In other words, some of the separate occurrences of "tornadic events" are actually part of the same continuous tornado vortex.

The initial time of tornado formation varied widely among the ensemble members, with four members (2, 3, 4, and 7) predicting tornadoes that began at least 20 min prior to the onset of the observed tornado, and four members (5, 6, 8, and 10) predicting tornadoes that began at least 20 min later than observed. In member 4, where the strongest tornadic circulation forms near the
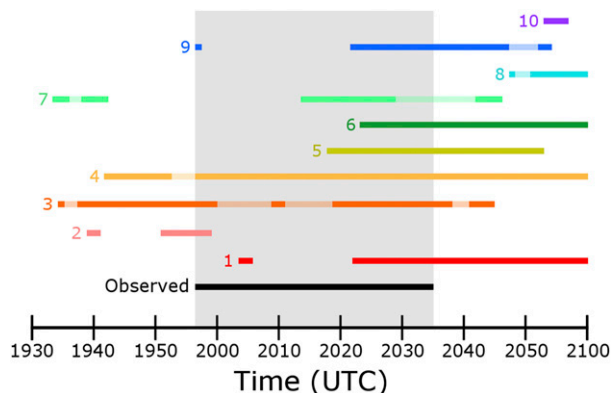
FIG. 4. Duration of the observed Newcastle–Moore tornado and predicted tornadoes from all ensemble members (1–10). Tornadoes are indicated as being present in ensemble members at times when, associated with the tornadic circulation, winds at the first model level above the surface exceeded $29 \, \text{m s}^{-1}$ and maximum vorticity exceeded $0.4 \, \text{s}^{-1}$. The faded line segments for members 3, 4, 7, 8, and 9 indicate the presence of a continuous, subtornado-intensity, near-surface circulation between tornadoes (i.e., a single, continuous vortex rather than fully discrete tornadic events). The duration of the observed tornado is also shaded in light gray to facilitate comparison with the ensemble members.

start of the observed tornado, the predicted tornado does not dissipate as quickly as the observed tornado; the tornado in member 4 is present through the end of the forecast period at 2100 UTC, while the observed tornado dissipates by 2035 UTC (Fig. 4). In contrast, member 3 predicts a tornado which forms very near the start of the observed tornado track and moves only slightly to the north of the observed track (Fig. 3), persisting for 23 min. Because of the relatively fast storm motion of the ensemble members, though, this tornado occurs almost entirely prior to the start of the observed tornado (Fig. 4). The faster movement speed of predicted supercells than observed is consistent with a number of earlier studies (e.g., Stensrud and Gao 2010; Dawson et al. 2012; Xue et al. 2014; Yussouf et al. 2015; Supinie et al. 2017), and this issue will be discussed in more detail in the concluding section.

Particularly in members 1, 4, and 6, the predicted tornadoes become quite intense. Time series of instantaneous domain-wide maximum horizontal wind speed and most extreme negative pressure perturbation at a data interval of 0.2 s, as well as 3-s running averages of these fields, are plotted for all 10 ensemble members in Fig. 5. The 3-s averages are plotted to correspond to the standard of "maximum 3-s wind gust at 10 m" used by the EF scale (as was done in, e.g., Dahl et al. 2017). While the values shown are obtained from domain-wide data, these extrema are almost all closely collocated with the near-surface circulation of the model supercell

and occur near the first level above the model surface (~10 m), particularly at times when a tornado is present. Peak 10-m wind speeds exceed $100 \, \text{m s}^{-1}$ in members 1, 4, and 6 in both the instantaneous data (Fig. 5a) and the 3-s running average (Fig. 5b), with the most intense tornadoes occurring later in the forecast period (2020–2100 UTC). The most intense tornado, predicted by member 1, has a peak instantaneous 10-m wind speed of nearly $140 \, \text{m s}^{-1}$ (Fig. 5a). Significant variation in maximum wind speed is evident during times when an intense tornado is present; the most intense winds occur in peaks lasting only a few seconds, and peak winds within the intense tornadoes of members 1, 4, and 6 often vary by more than $20 \, \text{m s}^{-1}$ over time spans of <10 s (Fig. 5a). These extreme variations are somewhat lessened, and the magnitude of the maximum wind slightly decreased, in the 3-s average (Fig. 5b), though fluctuations exceeding $10 \, \text{m s}^{-1}$ on time scales of several seconds remain even in the running average.

During the times that intense tornadoes are present in members 1, 4, and 6, large negative perturbations, exceeding 60 hPa in magnitude, are present in air pressure at 10 m above the surface (Fig. 5c). The most extreme negative pressure perturbation, more than 120 hPa, occurs in member 1 late in the forecast period, while members 4 and 6 exhibit maximum negative pressure perturbations of approximately 95 and 90 hPa respectively (Fig. 5c). Pressure extrema for these members in the 3-s running average (Fig. 5d) are only slightly reduced in magnitude. These pressure deficits are similar to those reported in previous studies for measurements taken by in situ instruments in intense tornadoes (e.g., Karstens et al. 2010) and Dawson et al. (2016) showed this pressure deficit can contribute significantly to the positive buoyancy of air parcels entering the tornado vortex.

b. Structure of intense tornadoes in the model

Simulated near-surface radar reflectivity fields in the vicinity of the model tornadoes exhibit very distinct hook echoes; examples from the predicted tornadoes in members 1 and 4 at 2040 UTC are plotted in Fig. 6a and Fig. 7a. Members 1 and 4 are among the four members that predict tornadoes of EF5 intensity (Fig. 5b); at 2040 UTC, the tornado in member 1 is near its maximum intensity, while the tornado in member 4 is less intense, typical of the type of circulation seen in many ensemble members (Fig. 5). In both members, the predicted tornado is located within the hook, just to the south of its tip (indicated by the black contours of vorticity $>0.4 \, \text{s}^{-1}$ in Fig. 6a and Fig. 7a. These simulated near-surface reflectivity fields closely resemble near-surface radar observations of tornadoes
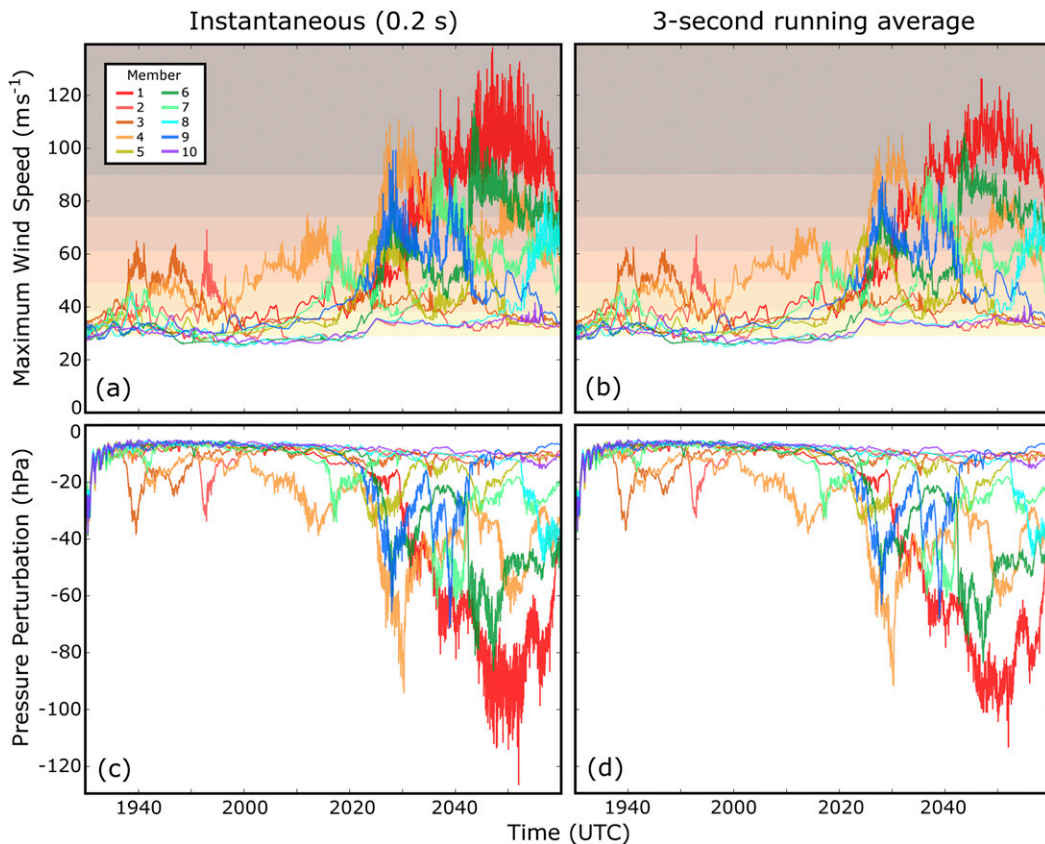
FIG. 5. Time series of (a),(b) domain-wide maximum horizontal wind speed (m s^{-1}) and (c),(d) domain-wide maximum negative pressure perturbation (hPa). Data are plotted (a),(c) at 0.2 s intervals, and (b),(d) as 3-s running averages. Shaded regions in (a) and (b) indicate wind thresholds associated with enhanced Fujita scale ratings, from EF0 (lightest) to EF5 (darkest).

reported in previous studies (e.g., Bluestein et al. 2007; Tanamachi et al. 2012; Wurman and Kosiba 2013), and contain several of the prominent features reported in those studies, including the weak-echo notch, and a region of relatively low reflectivity within the tornadic circulation itself.

Observed radar reflectivity (Fig. 8a) and radial velocity (Fig. 8b) from WSR-88D radar site KTLX for the Newcastle–Moore supercell at 2013 UTC, when the tornado was near its peak intensity, are shown in Fig. 8. The overall supercell structure, and the size and location of the tornado circulations in members 1 (Fig. 6) and 4 (Fig. 7) are similar to those of the observed tornado and its parent storm (Fig. 8). In terms of reflectivity structure, it should be noted that the simulated reflectivity from the ensemble members (Figs. 6a and 7a) lacks the ability to capture nonmeteorological radar returns from lofted dust or debris, as such processes are not parameterized or simulated by the model; the simulated hook echoes are similar in structure to those seen in previous model studies (e.g., Xue et al. 2014). In observed

tornadoes, lofted dust and debris often contribute to an enhanced reflectivity ring; the absence of this process in the simulated tornadoes likely contributes to the lack of a distinct "eye" in the center of the tornado (Bluestein et al. 2007). The effect of debris is evident in KTLX reflectivity observations of the Newcastle–Moore storm (Fig. 8a), in the form of an enhanced region of reflectivity (debris ball signature) collocated with the tornado vortex (Fig. 8b).

Updraft velocity, vertical vorticity, and horizontal wind speed at 10 m above the surface are plotted in the immediate vicinity of the tornadic vortices in Figs. 6b–d and Figs. 7b–d. The tornadic vortex is well-defined in each of these model fields at 2040 UTC in both member 1 (Figs. 6b–d) and member 4 (Figs. 7b–d). In both members, the radius of maximum wind of the tornado at 10 m is approximately 400 m (Figs. 6d and 7d), or approximately eight times the model grid spacing, suggesting that the tornado is relatively well resolved on the 50-m grid. The highest wind speeds are located on the southeast side of the tornadic circulation, where the
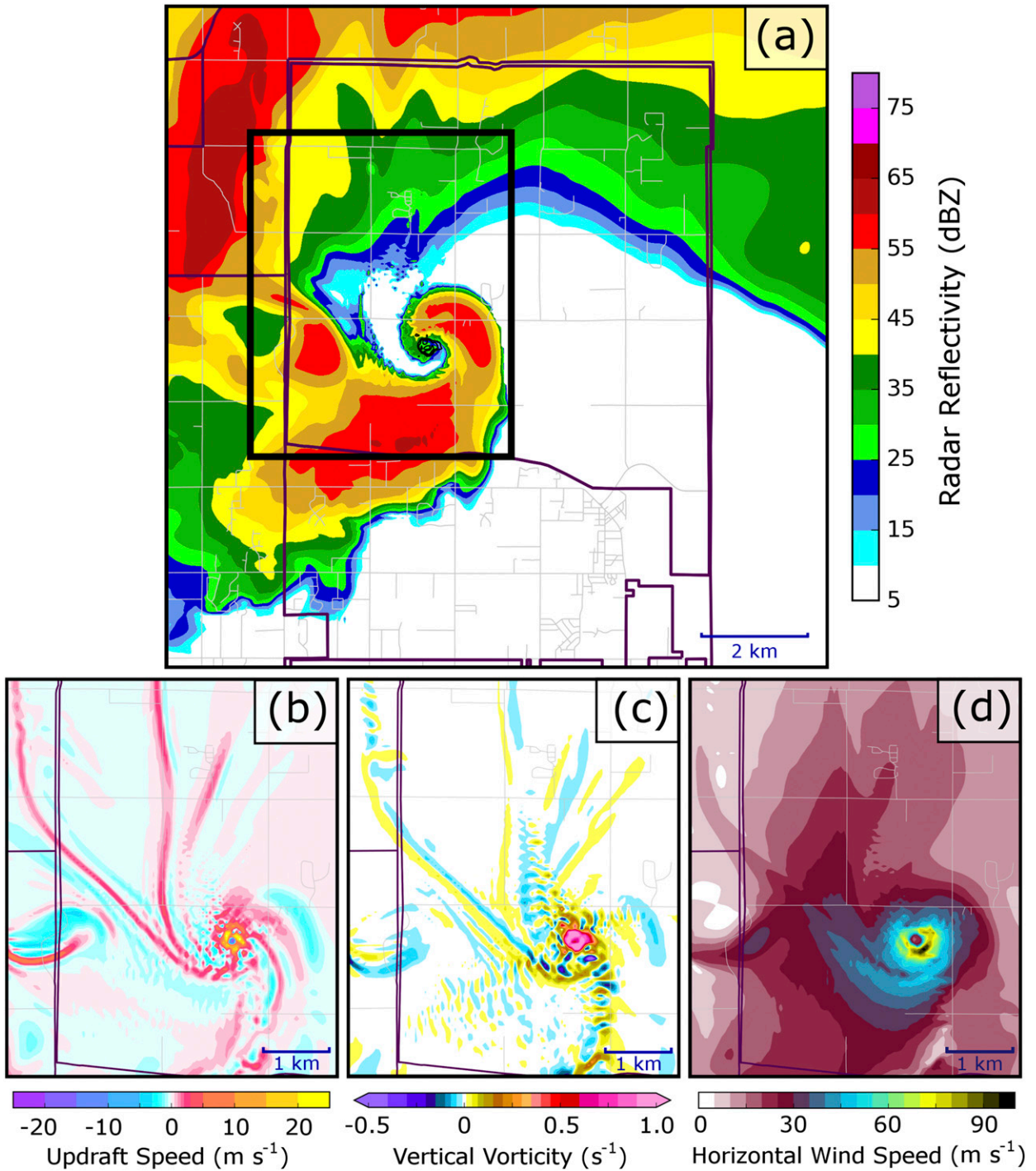
FIG. 6. Plots of (a) radar reflectivity, (b) vertical wind speed, (c) vertical vorticity, and (d) horizontal wind speed at 10 m above the surface in the region of the tornadic circulation of ensemble member 1 at 2040 UTC. The thick black box shown in (a) indicates the region shown in (b)–(d). City borders are indicated in dark purple, and roads are shown in light gray. Vertical vorticity contours at intervals of $0.3\,s^{-1}$ starting from $0.4\,s^{-1}$ are overlaid in (a).

wind speed from the rotation of the vortex is aligned with the direction of motion of the tornado, and exceed $90\,m\,s^{-1}$ at this time in member 1 (Fig. 6d) and $70\,m\,s^{-1}$ in member 4 (Fig. 7d).

The structure of the tornadic vortex and the flow in its vicinity exhibit features consistent with those reported in previous observational studies of tornadoes. In both member 1 (Fig. 6b) and member 4 (Fig. 7b), updraft
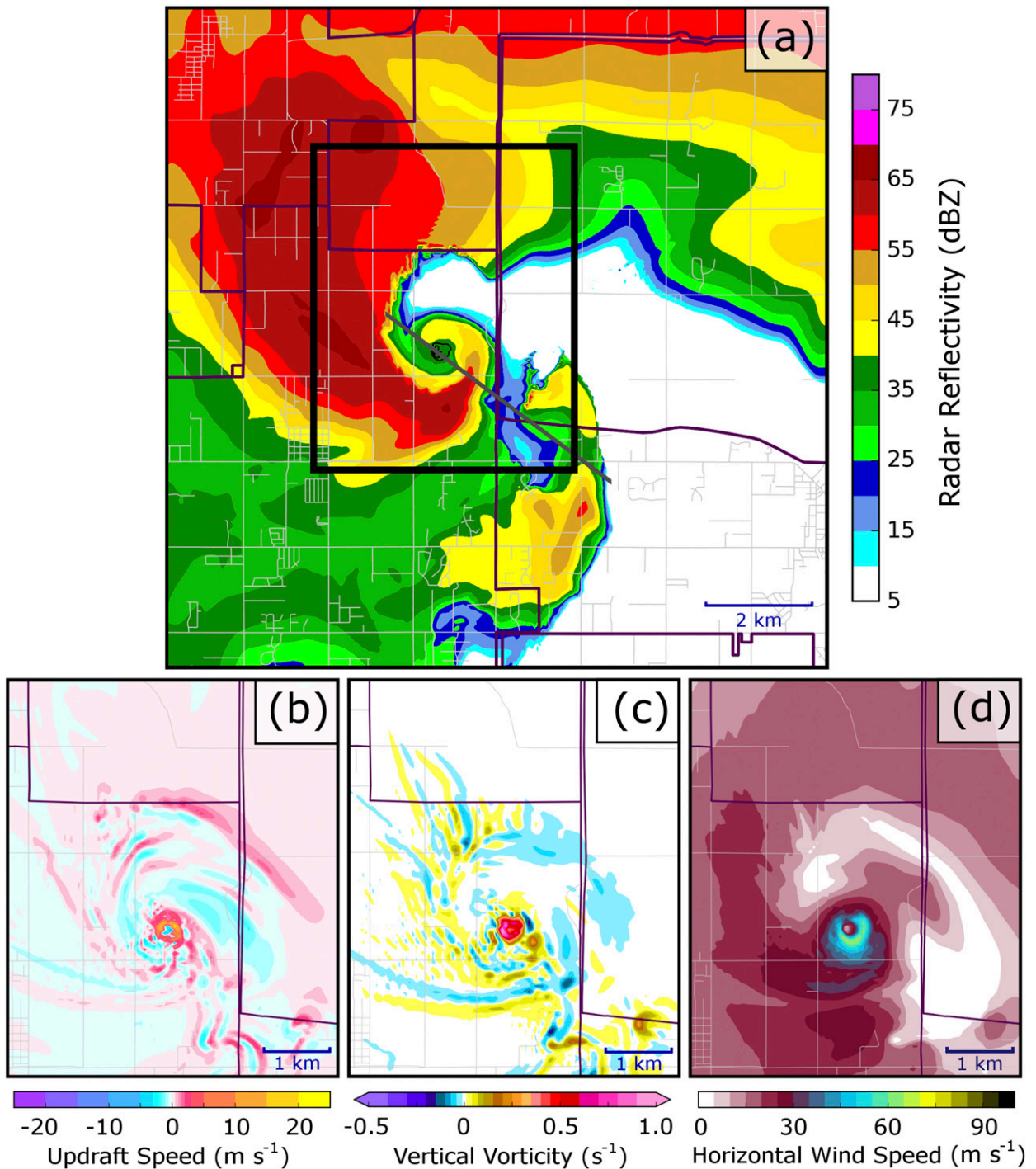
FIG. 7. As in Fig. 6, but for member 4 at 2040 UTC. The dark gray line in (a) indicates the path of the vertical cross section plotted in Fig. 9.

speeds at 10 m above the surface in the tornado are at a maximum in an annular ring approximately located immediately inside of the radius of maximum horizontal wind. The updraft within this annulus is relatively intense, exceeding 20 m s^{-1} in member 1 (Fig. 6b), and 15 m s^{-1} in member 4 (Fig. 7b), even though the level is

only 10 m above ground, indicating very large vertical acceleration and therefore stretching near the surface. Within the ring of updrafts, there is a strong downdraft in the core of the tornadic circulation, in excess of 20 m s^{-1} in member 1 (Fig. 6b), and 10 m s^{-1} in member 4 (Fig. 7b). This vertical wind structure, with a central
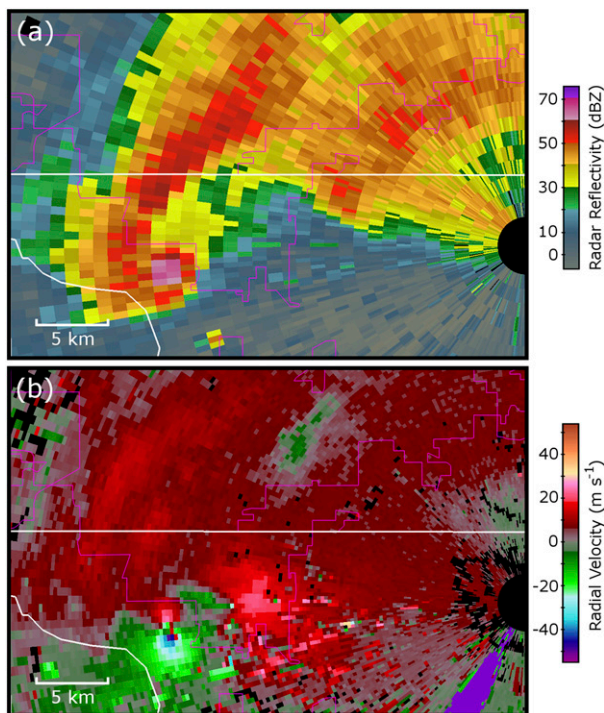
FIG. 8. Observed (a) radar reflectivity and (b) radial velocity from the 0.5° tilt of WSR-88D KTLX at 2013 UTC 20 May 2013 for the region of the Newcastle–Moore tornado and its parent supercell.

downdraft surrounded by a ring of updrafts, is consistent with a two-cell vortex (Davies-Jones 1986; Davies-Jones et al. 2001), a structure which has been reported in prior tornadic simulations (e.g., Wood and Brown 2011; Xue et al. 2014); the downward pressure gradient force associated with larger negative pressure perturbations within the vortex near ground is the primary cause for this downdraft. The two-cell structure of the tornado vortex has many similarities with the inner core structures of an intense typhoon/hurricane having an axisymmetric eyewall (Houze 2010). In member 4 (Fig. 7), a satellite vortex (Agee et al. 1976) is present at 2040 UTC about 1.5 km to the east-southeast of the primary vortex, as evidenced by a slight notch in the reflectivity field (Fig. 7a), and local maxima in the updraft, vertical vorticity, and horizontal wind fields (Figs. 7b–d)—similar satellite vortices (in terms of relative positioning and scale) have been observed via radar in prior observational studies (e.g., Wurman and Kosiba 2013).

To the west and northwest of the tornado in member 1, streamer-like regions of enhanced vorticity are evident, extending from northwest of the tornado around its west edge and merging into the tornado circulation on its south side (Fig. 6c). These vorticity streamers are collocated with similar streamer-like regions of enhanced updraft (Fig. 6b) and slightly enhanced horizontal wind speed (Fig. 6d). These features appear to be quite similar to the ''vorticity river'' features reported by Dahl et al. (2014). In future work, the cause of these streamer-like features will be analyzed, and their possible role in tornadogenesis investigated.

Vertical cross sections through the main tornado and the satellite vortex in member 4 (from WNW to ESE) at 2040 UTC (see Fig. 7) are presented in Fig. 9 for updraft velocity (Fig. 9a) and horizontal wind speed (Fig. 9b), with vertical vorticity as a contoured overlay. Thresholds for horizontal wind speed shading correspond to wind speed values associated with EF scale ratings from EF0-EF5. The two-celled nature of the primary tornado vortex is apparent in the vertical cross section, with downdrafts of $15–20 \, \mathrm{m \, s^{-1}}$ extending from very near the surface (approximately 350 m above mean sea level) to over 1000 m above mean sea level (Fig. 9a). Updrafts surround the downdraft region, with updraft speeds near the surface exceeding $20 \, \mathrm{m \, s^{-1}}$ on the western side of the vortex (Fig. 9a). The horizontal winds are strongest on the eastern side of the vortex (Fig. 9b), due in part to the translational motion of the storm. The satellite vortex appears as an extension toward the surface of positive updraft speed and higher horizontal wind speed, approximately 2.5 km east of the primary tornado vortex (Fig. 9).

c. Ensemble and probabilistic tornado forecasts

Ensemble probabilistic prediction of tornadoes, and meaningful verification of the resultant forecasts, is very challenging due to the highly localized nature of tornadoes and often substantial differences in the timing (see Fig. 4) and location/path (see Fig. 3) of the predicted tornadoes among the ensemble members. To address these challenges, we will consider probabilistic forecasts of the presence of a tornado within a given distance of a point at any time during the forecast—an approach that considers the full swath of the tornado in this manner allows for variation in timing among members, and the use of a ''neighborhood'' allows for error in the tornado track. A variety of neighborhood radii are considered, ranging from 100 m to 20 km. A Gaussian smoother with a standard deviation equal to the neighborhood radius is applied to the raw probabilistic forecasts to further account for positional uncertainty and variation among members. This approach is similar to that used quasi-operationally in the NOAA Hazardous Weather Testbed (HWT) spring experiment for probabilistic predictions of severe weather hazards (Jirak et al. 2012; CAPS 2016). Schwartz and Sobash (2017) present a review on the neighborhood approaches.
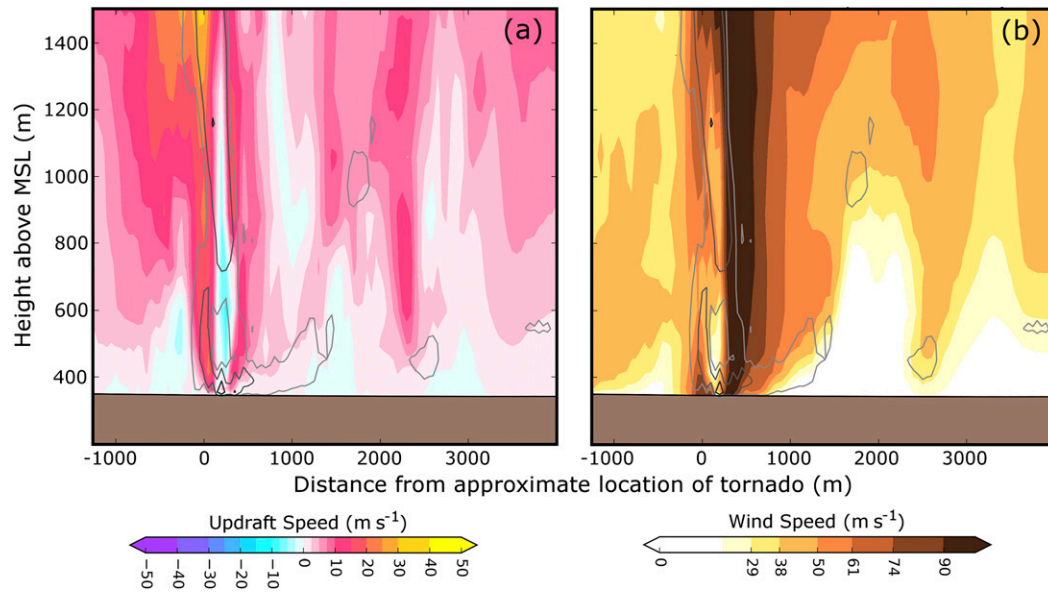
FIG. 9. Vertical cross sections through the tornado and satellite vortex in member 4 at 2040 UTC (from northwest to southeast along the dark gray line in Fig. 7a). Shown are (a) updraft velocity (shaded) and vertical vorticity (contoured), and (b) horizontal wind speed (shaded) and vertical vorticity (contoured). Vertical vorticity is contoured, from lightest to darkest contours, at thresholds of 0.1, 0.2, and $0.4\,s^{-1}$. The area near the bottom of each panel, shaded in brown, is below the ground.

While tornadoes can be identified in many model fields, we consider probabilistic forecasts using two of the most directly relevant model fields: horizontal wind speed and vertical vorticity. For wind speed, we consider $29\,m\,s^{-1}$, the minimum wind value associated with the EF0 range of the enhanced Fujita scale (American Meteorological Society 2013), as the minimum threshold for the presence of a tornado. For vertical vorticity, we use a threshold of $0.4\,s^{-1}$; as mentioned in section 4a, this threshold was chosen empirically via preliminary evaluations of model forecast fields and structures. All forecasts will consider values of the wind and/or vorticity fields at the first model level, or about $10\,m$ above ground.

Ensemble-based forecasts of the probability of a tornado occurring within $2\,km$ of a point during the forecast period are presented in Fig. 10 using wind alone as a criterion (Fig. 10a), vorticity alone as a criterion (Fig. 10b), and requiring both wind and vorticity criteria to be met (Fig. 10c). In each case, swaths of nonzero tornado probability extend from near the start of the observed tornado track, following the observed track eastward for most of its length, then extending northeast beyond the end of the observed track through the eastern portions of the Oklahoma City metropolitan area and into regions just to the east (Fig. 10). The highest probabilities are located in the northeastern portion of the predicted swath, with maxima ranging from

approximately 0.3 in the forecast requiring both wind and vorticity criteria to be met (Fig. 10c) to nearly 0.4 in the forecast requiring only the wind criterion to be met (Fig. 10a). The predicted swath is slightly wider and probabilities are generally slightly higher when only one criterion is applied (Figs. 10a,b) than when both criteria are used together (Fig. 10c), but all three forecasts largely agree in terms of swath length and position. Given this general agreement, and given that the wind field is most directly related with the societal impacts of tornadoes, using explicit forecasts of 10-m wind speed as a basis for meaningful probabilistic tornado forecasts may be the most useful method in a Warn-on-Forecast context, and is feasible for ensembles similar to the one used in this study.

To evaluate the impact of the size of the neighborhood used in probabilistic forecast product generation on the skill of the tornado probabilistic forecast, we verify forecasts of tornado probability using the same set of criteria as in Fig. 10 for distances ranging from $100\,m$ to $20\,km$ against the observed path of EF0 or greater damage from the damage survey of the Newcastle–Moore tornado. The results are presented in terms of area under the relative operating characteristic curve (AUC; Mason 1982; Mason and Graham 1999) in Fig. 11. The Gaussian filter to smooth the probabilistic forecasts verified in Fig. 11 uses a standard deviation equal to the neighborhood radius. A tornado is
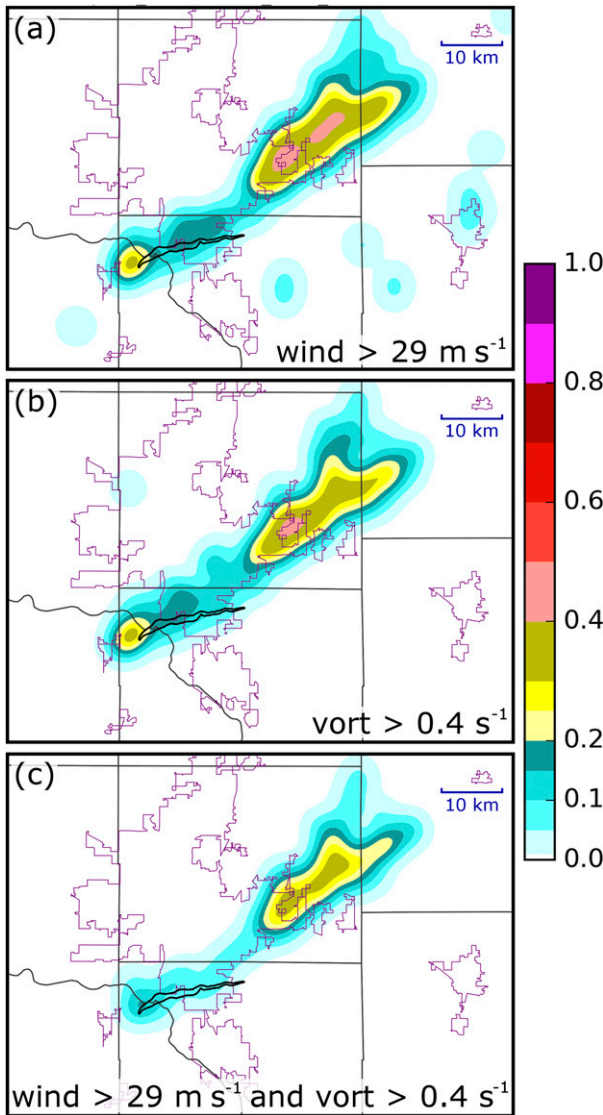
FIG. 10. Probability of (a) horizontal wind speed at 10 m above the surface exceeding $29 \, \text{m s}^{-1}$, (b) vertical vorticity at 10 m above the surface exceeding $0.4 \, \text{s}^{-1}$, and (c) both the wind and vorticity criteria being met within 2 km of a point, with a Gaussian filter with a standard deviation of 2 km applied. Also shown are county boundaries (in black) and urban boundaries (in purple).
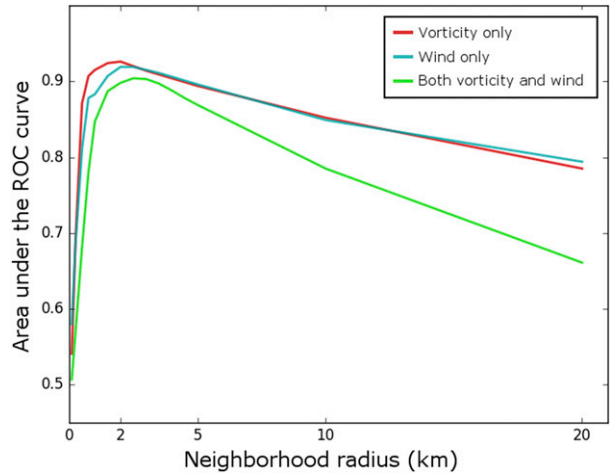

FIG. 11. Area under the relative operating characteristic curve (AUC) of forecasts of the probability of a tornado within a given radius of a point, smoothed using a Gaussian filter with a standard deviation of the same radius, as a function of radius using criteria of vorticity $> 0.4 \, \text{s}^{-1}$ (red), horizontal wind speed $> 29.0 \, \text{m s}^{-1}$ (blue), and both the vorticity and wind criteria together (green), at the first vertical model level above the surface, to indicate the presence of a tornado within the ensemble members.

considered to have occurred in the model at a given grid point if a tornado is present, according to the criteria being evaluated, anywhere in that grid point's neighborhood.

Regardless of whether the wind criterion, the vorticity criterion, or both are applied, the forecasts exhibit the greatest skill in terms of AUC for neighborhood radii of 2.0 to 2.5 km, peaking at AUC of approximately 0.9. Forecast skill is slightly higher when only one of the wind or vorticity criteria are applied than when both are, and skill is slightly greater for the

vorticity criterion than for the wind criterion at radii of 2.0 km or less (Fig. 11). The difference in skill between the wind- and vorticity-based forecasts at small neighborhood radii results, at least in part, from slightly reduced incidence of false alarms to the northeast of the observed track when using the vorticity criterion compared to the wind criterion (see Figs. 10a,b), and small differences in the positioning of the wind and vorticity swaths within individual member forecasts. For neighborhood radii exceeding 2.0 km, the wind and vorticity criteria, applied individually, produce very similar results in terms of AUC. Regardless of the criteria used, all tornado forecasts evaluated exhibited useful operational skill (AUC > 0.7; Buizza et al. 1999) for all neighborhood radii tested from 0.75 to 10.0 km. We note, however, that several forecast members (i.e., members 1, 4, 6, and 8) had ongoing tornadoes at the end of the forecast period; had the forecast period been extended beyond 2100 UTC (which would also have required a larger 50-m domain), this could have led to additional false alarms which would have negatively impacted the forecast skill.

To better illustrate the impact of the choice of neighborhood size on the probabilistic tornado forecasts, forecasts of the probability of a tornado within 0.5, 2.0, and 5.0 km of a point, calculated using the wind criterion, are compared in Fig. 12. When a very small neighborhood of 0.5 km is used (Fig. 12a), the paths of tornadoes within individual members are evident in the
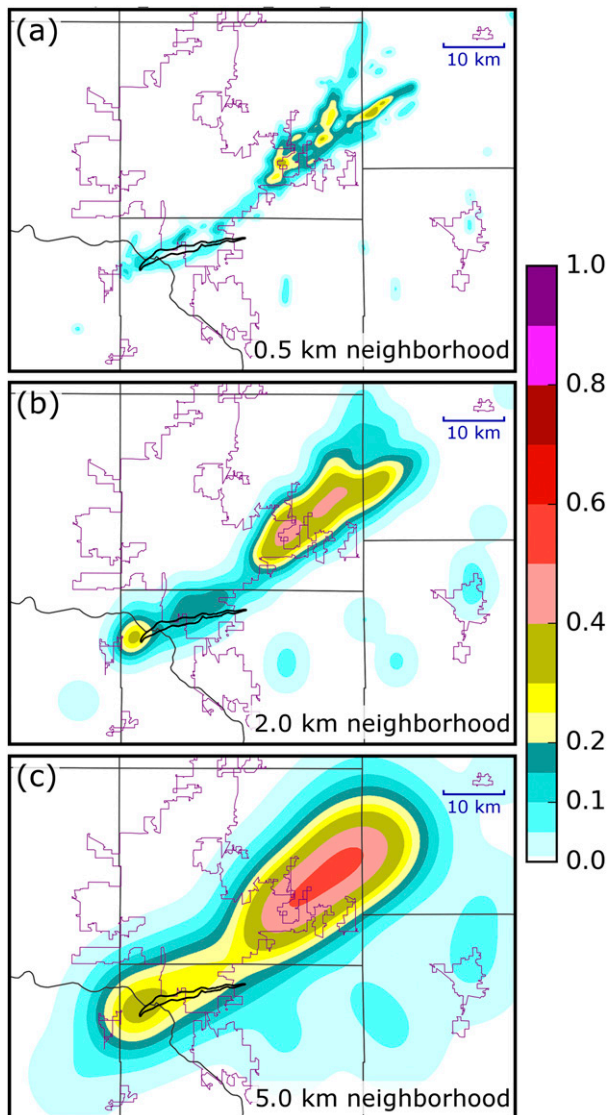
FIG. 12. Probability of horizontal wind speed at 10 m above the surface exceeding 29 m s^{-1} within a radius of (a) 0.5, (b) 2.0, and (c) 5.0 km of a point, with a Gaussian filter with a standard deviation of the same radius applied. Also shown are county boundaries (in black) and urban boundaries (in purple).

probability field, particularly near the observed tornado track. Probabilities also remain relatively low when a very small neighborhood radius is used, as tornadoes in separate members must cross paths or pass within just a few hundred meters of a given point to generate probabilities exceeding 0.1 (Fig. 12a). When the neighborhood is increased to 2.0 km (Fig. 12b), individual member tornado paths are no longer evident in the probability field, and probabilities are generally higher within the region where tornadoes occurred in the ensemble, particularly near and to the northeast of the observed tornado track (Fig. 12b). This pattern

continues as the neighborhood is expanded to 5.0 km (Fig. 12c), with still higher probabilities; at this larger radius, however, the probability field extends well away from the track of tornadoes within the ensemble (cf. Fig. 12c and Fig. 3), and small areas with wind exceeding 29 m s^{-1} away from the primary tornado track are exaggerated into large regions of low probability, particularly in the southeast portion of the domain (cf. Fig. 12a and Fig. 12c).

The comparison in Fig. 12 helps to explain the pattern of skill shown in Fig. 11. When a very small neighborhood is used (Fig. 12a), areas of nonzero probability are closely confined to the tornado tracks within each member, leading to gaps in the probability field near the observed tornado track that greatly decreased the probability of detection (and as a result, the AUC) for small neighborhood radii. As the neighborhood was increased to 2.0 (Fig. 12b) and 5.0 km (Fig. 12c), these gaps in the probability field disappeared, but the region of false alarms associated with the tornadoes that occurred to the northeast of the observed track in many members had an increasingly negative impact on the AUC as the neighborhood radius increased, as did the increasing impact of small areas of wind exceeding 29 m s^{-1} away from the primary observed tornado track (Fig. 12c). The balance of these two factors resulted in the AUC being maximized for neighborhood radii of approximately 2.0 km. Considering this, and taking into account operational concerns (e.g., how far from the tornado would be desirable for a probabilistic operational product to extend), as well as forecast uncertainties, a neighborhood radius of anywhere from 2 to 10 km would likely be appropriate for future operational (or quasi operational) tornado-resolving systems. The societal implications of the radius chosen, and skill over a range of cases would need to be considered to arrive at an operationally optimal neighborhood radius, though such considerations are beyond the scope of this study.

d. Comparison of explicit tornado forecasts to forecasts of updraft helicity

In convection-permitting ensembles with resolution sufficient to capture supercell thunderstorms but insufficient to capture tornadoes or tornado-like vortices (typically using horizontal grid spacing ranging from 500 m to 4 km), updraft helicity (UH; Kain et al. 2008) is often used as a proxy for tornado potential associated with supercell thunderstorms (e.g., Kain et al. 2010; Sobash et al. 2011; Jirak et al. 2014; Gallo et al. 2016). UH is computed as the product of the vertical component of vorticity and updraft speed integrated vertically over a specified depth of the atmosphere,

often between 2 and 5 km above the surface. As a product of a quantity measuring rotation (vertical vorticity) and updraft speed, UH directly indicates the presence of rotating updrafts within a model forecast, and has been shown to have considerable skill as a proxy for forecasts of tornadic storms (e.g., Gallo et al. 2016). Because of the exceptionally high computational cost of tornado-resolving forecasts, however, direct comparisons of predicted UH swaths to the paths of tornadoes within the same model forecast have, to the knowledge of the authors, not previously been made.

In Fig. 13, we perform such a comparison for UH over two separate layers: 0–3 km (Fig. 13a) and the more commonly used 2–5 km (Fig. 13b), with the goal of quantifying the utility and possible biases of the probabilistic UH forecasts compared to explicit forecasts of 10-m wind exceeding the tornadic threshold (calculated as in Fig. 10a). The probabilistic UH forecasts shown in Fig. 13 are calculated using the same procedure as the probabilistic forecasts of wind and vorticity presented in Fig. 10. UH is typically used as an indicator of features on the scale of mesocyclones; to ensure we are focusing on structures on this scale, and not smaller, tornado-scale structures, we calculate UH from data horizontally averaged to a grid spacing of 1 km. UH is calculated using a neighborhood of 2 km, a threshold of $800 \, \mathrm{m^2 \, s^{-2}}$ for UH in the 0–3-km layer (Fig. 13a) and a threshold of $1800 \, \mathrm{m^2 \, s^{-2}}$ for UH in the 2–5-km layer (Fig. 13b). The differing thresholds used for the 0–3- and 2–5-km layers are necessitated by the fact that UH is substantially higher over the 2–5-km layer due to substantially stronger updrafts in the midlevels of the troposphere compared to those near the surface. We also note that the UH field used in Fig. 13 is calculated at the full vertical resolution allowed by the model rather than by averaging over coarser vertical layers, such as the three 1-km-deep layers used in Kain et al. (2008). UH is calculated in this manner to allow us to focus on biases present in the model fields rather than those introduced by vertical interpolation.

Overall, the probabilistic forecast swaths for UH, both in the 0–3-km layer (Fig. 13a) and the 2–5-km layer (Fig. 13b) are qualitatively quite similar to the probabilistic forecast swath for 10-m wind exceeding $29 \, \mathrm{m \, s^{-1}}$. The extent and path of both swaths is similar, and maxima and minima are similarly located for both probabilistic fields. For both UH layers, there is a slight northwestward bias of the UH field relative to the probabilistic forecast of tornadic winds at the surface, ranging from a few hundred meters to around 1.5 km (Fig. 13); this bias is largest for the western portion of the predicted swath, and slightly larger for UH over the 2–5-km layer (Fig. 13b).
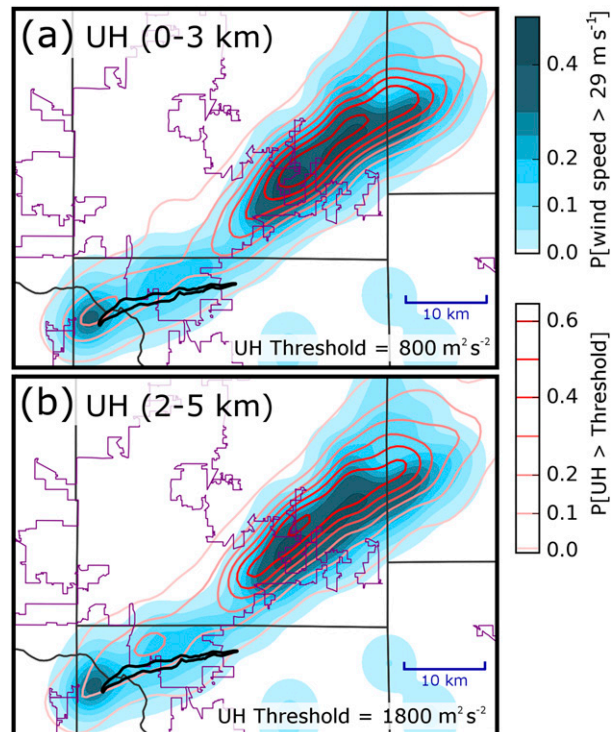

FIG. 13. Contours of probability of (a) 0–3-km updraft helicity exceeding $800 \, \mathrm{m^2 \, s^{-2}}$ and (b) 2–5-km updraft helicity exceeding $1800 \, \mathrm{m^2 \, s^{-2}}$ within 2 km of a point, calculated using data smoothed to a horizontal grid spacing of 1 km, with a Gaussian filter with a standard deviation of 2 km applied. The probability of the 10-m horizontal wind speed exceeding $29 \, \mathrm{m \, s^{-1}}$, calculated as in Fig. 12b using unsmoothed 50-m data, is included in each panel (shaded) to facilitate comparison.

than for UH over the 0–3-km layer (Fig. 13a). This small but persistent bias is a result of both tilt of the storm updraft with height and the position of the tornado at the surface relative to the center of the parent mesocyclone (UH, particularly over the 2–5-km layer, is maximized over the midlevel mesocyclone circulation, and previous studies (e.g., Kain et al. 2008) attribute the skill of UH as a predictor of tornadic potential to its ability to detect mesocyclones). The northwestward tilt of the tornado and mesocyclone circulations with height implied by these results agrees with previously reported radar observations of tornadic supercells (e.g., Mahre et al. 2018); such tilt may be linked to the balance between low-level inflow and cold pool, as suggested by Snook and Xue (2008). Despite the aforementioned bias, predicted swaths of tornadic wind at the surface and of UH within the ensemble show remarkable similarity in terms of these probabilistic forecasts. Such results are expected, as a strong low-level rotating updraft (as measured by UH) would promote strong vertical velocity acceleration and vertical vorticity stretching near the surface,

and rapid intensification of surface vortex, leading to the genesis and intensification of tornado.

5. Summary and discussion

In this study, we perform a ten-member ensemble forecast for the 20 May 2013 Newcastle–Moore EF5 tornado and its parent storm using a model domain with 50-m horizontal grid spacing nested within a 500-m grid spacing storm-scale ensemble (Snook et al. 2016) which assimilated radar and surface observations using an ensemble Kalman filter. Forecasts were initialized at 1930 UTC, approximately 26 min prior to the formation of the Newcastle–Moore tornado, and were run for a period of 90 min, ending at 2100 UTC. From this ensemble of tornado-resolving forecasts, ensemble-based probabilistic forecasts of tornadic circulations at the surface were produced and verified, using both wind and vorticity as criteria, and these forecasts were compared to ensemble-based probabilistic forecast swaths of updraft helicity (UH) over the 0–3- and 2–5-km layers. The structure of tornadic vortices in individual members was also examined.

Tornadic circulations were present in all forecast members, with four of the 10 ensemble members predicting tornadoes with wind speeds associated with EF5 tornadoes (90 m s^{-1}). Among the 10 ensemble members, there was substantial variation in the tornado intensity, duration, and path. Peak tornado intensity, as estimated from wind speed, varied from EF0 in member 10 to EF5 in members 1, 4, 6, and 7. In all members, the predicted supercell thunderstorm moved to the northeast more quickly than the observed Newcastle–Moore storm. In members predicting intense tornadoes, this overly fast northeastward motion, along with the persistence of the tornado within the forecast members well after the end of the observed tornado at 2035 UTC, resulted in forecast swaths of tornadic winds that extend well to the north and east of the observed tornado damage path.

The problem of storms moving too quickly within storm-scale NWP forecasts, particularly for supercell thunderstorm cases, has been consistently noted in quite a few prior studies (e.g., Stensrud and Gao 2010; Dawson et al. 2012; Xue et al. 2014; Yussouf et al. 2015; Supinie et al. 2017). Several theories have been put forward to explain the bias within NWP models toward overly fast storm motion, including inability of the model to adequately replicate the pressure perturbations and vertical accelerations responsible for the deviation of supercell storm motion from the mean flow (Stensrud and Gao 2010), issues relating to turbulence parameterization (Fiori et al. 2010), and errors in the analyzed near-storm environment (Xue et al. 2014).

While identifying and correcting the source(s) of error responsible for model storms moving too quickly remains an important research topic, it is beyond the scope of this study. Nonetheless, the relatively small total area covered by model predicted tornadoes (both in terms of wind swaths and probabilistic predictions) compared to the area covered by NWS tornado warnings issued for this storm suggests that, if such displacement and timing errors can be remedied in future ensemble tornado prediction studies, a Warn-on-Forecast approach using a tornado-resolving ensemble forecast has the potential to lead to more precise warnings and reduced occurrence of false alarms. Also of note is that the model predictions presented in this study were initialized at 1930 UTC, 10 minutes before the first tornado warning was issued at 1940. If tornado-resolving ensemble forecasts can be produced quickly enough, they will be able to provide much longer lead time than today's detection-based warnings; this is the key concept behind Warn-on-Forecast.

When time series of maximum wind speed within the tornado were examined, the highest wind speeds predicted by the ensemble members containing EF5 tornadoes (100–130 m s^{-1}), are similar to or slightly higher than those reported in radar observations from high-end intense tornadoes (e.g., Burgess et al. 2002). Pressure perturbations within the most intense tornadoes in the ensemble ranged between approximately 80 and 120 hPa; such deviations are similar to those observed by in situ pressure measurements within intense tornadoes (e.g., Karstens et al. 2010). The structure of intense tornadoes within the ensemble members was similar between members, consisting of a central downdraft region surrounded by a ring of updrafts near the radius of maximum wind (approximately 400 m), typical of a two-cell vortex. Key features present in observed tornadoes, including a weak echo notch and a region of relatively low reflectivity within the tornado itself, were present in the model-derived reflectivity forecasts.

Ensemble-based probabilistic forecasts of tornadoes being present within a given distance of a point (using neighborhood radii varying from 0.1 to 20.0 km) were produced using both wind and vorticity criteria; the probabilistic forecast swaths were qualitatively similar for predictions made using either of these fields or both fields together. While skillful forecasts, in terms of area under the relative operating characteristic curve, were produced for all neighborhood radii between 0.75 and 20.0 km, forecast skill was maximized for neighborhood radii of 2.0 to 2.5 km. Though a neighborhood of approximately 2.0 km produced the most skillful forecasts for this case, we note that the optimal neighborhood radius would likely vary depending on the level of

spread within the ensemble and other parameters of the case chosen (including the width of tornadoes produced within the ensemble and the speed at which the storms are moving). Additional cases should be examined to evaluate the impact of these factors and to enable a more general recommendation for the neighborhood parameters for similar forecasts. Furthermore, if a method similar to this is used to produce operational probabilistic tornado forecasts at some point in the future, additional factors will need to be considered in the choice of neighborhood radius, including the tolerance for false alarms and misses in an operational context. The ensemble size will also likely be a factor to consider. Performing similar ensemble forecast studies using different cases is also important to evaluate the robustness of the findings of this study, as a single case study alone is insufficient to establish that tornado-resolving ensembles will generally be able to capture the intensity, timing, and location of observed tornadoes.

Ensemble-based probabilistic forecasts of UH exceeding a specified threshold over the 0–3- and 2–5-km layers produced swaths that were qualitatively very similar to those of tornadic wind and vorticity at the surface, with similar placement of maxima and minima. A slight northwestward bias, ranging from a few hundred meters to around 1.5 km, was present in the UH swath relative to those predicted using the surface wind and vorticity fields, attributable to the tilt of the storm updraft with height and the placement of the tornado at the surface relative to the center of its parent mesocyclone. The strong agreement between the UH forecasts and the forecasts of surface variables explicitly indicative of a tornado (i.e., wind and vorticity) suggests that UH is a rather good proxy for the path of tornadic winds at the surface, at least for UH based on the tornado-resolving forecasts. Comparisons of UH to surface-level tornadic features in more tornado-resolving forecasts will be needed to evaluate the robustness of this result. Furthermore, because real time tornado-resolving ensemble forecasts are still far from reality due to extreme computational cost, practical model-based predictions will remain at nontornado-resolving resolutions in the near future. For this reason, the tornado-prediction skill of convective storm-resolving forecasts with resolutions on the order of 1 km based on proxy parameters such as UH is worth investigation, in comparison to the direct predictions of tornado-resolving forecasts. In a future study, we will examine this topic based on ensemble forecasts produced at 500-m grid spacing, and grid spacings between 500 and 50 m.

The substantial differences in the tornadoes predicted between the members of our forecast ensemble, as well as errors in speed, placement, and timing of the predicted tornadoes (which occur well later than and to the northeast of the observed tornado in several members), highlights the need for further study into the practical and intrinsic predictability of tornadoes, with the help of very-high-resolution convective forecast ensembles. The sensitivity of the tornado forecast to initial condition errors and differences, in both the storm regions and in the storm environments, also require investigation. In addition, we note that there may be substantial sensitivity of the tornado forecast to timing of initialization of the 50-m grid, or to the timing of data assimilation on the outer (500 m) grid; these factors also require further investigation. We also note that, because the 500-m run cannot explicitly resolve tornadoes, there appears not to be a direct relation between the simulated tornado on the 50-m grid and the simulated supercell characteristics on the 500-m grid. Investigating which aspects of the parent storm determine the characteristics of embedded tornado is a separate research question related to both tornado dynamics and predictability; for example, Coffer et al. (2017) and Yokota et al. (2018) found that the vertical perturbed pressure gradient force resulting from the mesocyclone to be vital for tornadogenesis, and Roberts and Xue (2017) found that the ingestion of frictionally generated vorticity into the low-level mesocyclone led to lowering of the mesocyclone circulation and the creation of a strong upward pressure gradient force, ultimately contributing to tornadogenesis. Investigating these topics will require a separate study. Studies on these topics could be particularly relevant for the planned development of operational Warn-on-Forecast systems capable of predicting tornadic hazards (Stensrud et al. 2009). In addition to predictability considerations, examination of the sources of model error and model error growth, including sensitivity to the choice of microphysical scheme, could be valuable in producing more skillful tornado forecasts.

We note that limitations in available computational resources remain a strong limiting factor in our ability to perform tornado-scale ensemble forecast studies. Each ensemble member forecast produced during this study required more than 100 000 core-hours on a high-end supercomputing system to complete, limiting the size of the ensemble that could be produced at a reasonable computational cost. While this study suggests that tornado-resolving ensembles can provide operationally useful probabilistic forecast information regarding tornadic hazards, the immense computational expense required to produce them renders such a system, now and for the foreseeable future, impractical for real-time operational use. Future work investigating the skill of probabilistic predictions of tornado-like vortices in

ensembles using somewhat coarser grids (with horizontal spacing on the order of 100–500 m) is under way and could provide valuable information into the grid spacing needed to provide skillful tornado probabilistic forecasts. If forecasts on coarser grids prove capable of producing comparable results to those of the 50-m grids used in this study, the transition of such a system into an operational or quasi-operational computing environment would be far more feasible.

Finally, we note that the 50-m ensemble dataset produced in this study could be used for diagnostic analyses aimed at better understanding tornadogenesis dynamics, as was done in the most recent study of Yokota et al. (2018) for their ensemble forecasts at 50-m grid spacing of tornadoes occurring within a supercell storm over Japan. In future work, we plan to perform diagnostic analyses on the ensemble simulations for the understanding of tornadogenesis dynamics, following the approaches of Yokota et al. (2018) and other recent dynamical studies of tornadoes and tornadogenesis processes (e.g., Schenkman et al. 2014; Dawson et al. 2016; Roberts et al. 2016; Roberts and Xue 2017). Other possible future work using this dataset will include using the vortex wind analysis method of Q. Xu et al. (2015, 2017) to generate simulated observations of the tornado from individual 50-m forecasts, which will be used in data assimilation studies with the goal of improving the prediction (in terms of track and timing) of the tornado and its parent supercell, and performing detailed dynamical analyses of substorm-scale features (such as satellite vortices and near-surface vorticity streamers).

REFERENCES

Agee, E. M., J. T. Snow, and P. R. Clare, 1976: Multiple vortex features in the tornado cyclone and the occurrence of tornado families. *Mon. Wea. Rev.*, **104**, 552–563, https://doi.org/10.1175/1520-0493(1976)104<0552:MVFITT>2.0.CO;2.

American Meteorological Society, 2013: Enhanced Fujita scale. *Glossary of Meteorology*, Accessed 16 August 2016, http://glossary.ametsoc.org/wiki/Enhanced_Fujita_Scale.

Bluestein, H. B., C. C. Weiss, M. M. French, E. M. Holthaus, R. L. Tanamachi, S. Frasier, and A. L. Pazmany, 2007: The structure of tornadoes near Attica, Kansas, on 12 May 2004: High-resolution, mobile, Doppler radar observations. *Mon. Wea. Rev.*, **135**, 475–506, https://doi.org/10.1175/MWR3295.1.

Buizza, R., A. Hollingsworth, F. Lalaurette, and A. Ghelli, 1999: Probabilistic predictions of precipitation using the ECMWF Ensemble Prediction System. *Wea. Forecasting*, **14**, 168–189, https://doi.org/10.1175/1520-0434(1999)014<0168:PPOPUT>2.0.CO;2.

Burgess, D. W., M. A. Magsig, J. Wurman, D. C. Dowell, and Y. Richardson, 2002: Radar observations of the 3 May 1999 Oklahoma City tornado. *Wea. Forecasting*, **17**, 456–471, https://doi.org/10.1175/1520-0434(2002)017<0456:ROOTMO>2.0.CO;2.

——, and Coauthors, 2014: 20 May 2013 Moore, Oklahoma, tornado: Damage survey and analysis. *Wea. Forecasting*, **29**, 1229–1237, https://doi.org/10.1175/WAF-D-14-00039.1.

CAPS, 2013: 2013 CAPS Spring Forecast Experiment Program Plan. Accessed 16 August 2016, 24 pp., http://forecast.caps.ou.edu/SpringProgram2013_Plan-CAPS.pdf.

——, 2016: CAPS 2016 HWT Spring Experiment Ensemble Forecast Products. Accessed 19 September 2016, http://www.caps.ou.edu/~fkong/sub_atm/spring16.html.

Clark, A. J., and Coauthors, 2011: Probabilistic precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble. *Mon. Wea. Rev.*, **139**, 1410–1418, https://doi.org/10.1175/2010MWR3624.1.

——, J. S. Kain, P. T. Marsh, J. Correia Jr., M. Xue, and F. Kong, 2012: Forecasting tornado pathlengths using a three-dimensional object identification algorithm applied to convection-allowing forecasts. *Wea. Forecasting*, **27**, 1090–1113, https://doi.org/10.1175/WAF-D-11-00147.1.

——, J. Gao, P. T. Marsh, T. Smith, J. S. Kain, J. Correia Jr., M. Xue, and F. Kong, 2013: Tornado path length forecasts from 2010–2011 using ensemble updraft helicity. *Wea. Forecasting*, **28**, 387–407, https://doi.org/10.1175/WAF-D-12-00038.1.

Coffer, B. E., M. D. Parker, J. M. L. Dahl, L. J. Wicker, and A. J. Clark, 2017: Volatility of tornadogenesis: An ensemble of simulated nontornadic and tornadic supercells in VORTEX2 environments. *Mon. Wea. Rev.*, **145**, 4605–4625, https://doi.org/10.1175/MWR-D-17-0152.1.

Dahl, J. M., M. D. Parker, and L. J. Wicker, 2014: Imported and storm-generated near-ground vertical vorticity in a simulated supercell. *J. Atmos. Sci.*, **71**, 3027–3051, https://doi.org/10.1175/JAS-D-13-0123.1.

Dahl, N. A., D. S. Nolan, G. H. Bryan, and R. Rotunno, 2017: Using high-resolution simulations to quantify underestimates of tornado intensity from in situ observations. *Mon. Wea. Rev.*, **145**, 1963–1982, https://doi.org/10.1175/MWR-D-16-0346.1.

Davies-Jones, R. P., 1986: Tornado dynamics. *Thunderstorm Morphology and Dynamics*, E. Kessler, Ed., University of Oklahoma Press, 197–236.

——, R. J. Trapp, and H. B. Bluestein, 2001: Tornadoes and tornadic storms. *Severe Convective Storms, Meteor. Monogr.*, No. 50, Amer. Meteor. Soc., 167–222.

Dawson, D. T., II, M. Xue, J. A. Milbrandt, and M. K. Yau, 2010: Comparison of evaporation and cold pool development between single-moment and multimoment bulk microphysics schemes in idealized simulations of tornadic

thunderstorms. *Mon. Wea. Rev.*, **138**, 1152–1171, https://doi.org/10.1175/2009MWR2956.1.

——, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140**, 696–716, https://doi.org/10.1175/MWR-D-11-00008.1.

——, M. Xue, A. Shapiro, and J. A. Milbrandt, 2015: Sensitivity of real-data simulations of the 3 May 1999 Oklahoma City tornadic supercell and associated tornadoes to multimoment microphysics. Part I: Storm- and tornado-scale numerical forecasts. *Mon. Wea. Rev.*, **143**, 2241–2265, https://doi.org/10.1175/MWR-D-14-00279.1.

——, ——, ——, and ——, 2016: Sensitivity of real-data simulations of the 3 May 1999 Oklahoma City tornadic supercell and associated tornadoes to multimoment microphysics. Part II: Analysis of buoyancy and dynamic pressure forces in simulated tornado-like vortices. *J. Atmos. Sci.*, **73**, 1039–1061, https://doi.org/10.1175/JAS-D-15-0114.1.

Deardorff, J. W., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteor.*, **18**, 495–527, https://doi.org/10.1007/BF00119502.

Dowell, D., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005, https://doi.org/10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2.

Fiori, E., A. Parodi, and F. Siccardi, 2010: Turbulence closure parameterization and grid spacing effects in simulated supercell storms. *J. Atmos. Sci.*, **67**, 3870–3890, https://doi.org/10.1175/2010JAS3359.1.

Flora, M. L., C. K. Potvin, and L. J. Wicker, 2018: Practical predictability of supercells: Exploring ensemble forecast sensitivity to initial condition spread. *Mon. Wea. Rev.*, **146**, 2361–2379, https://doi.org/10.1175/MWR-D-17-0374.1.

Gallo, B. T., A. J. Clark, and S. R. Dembek, 2016: Forecasting tornadoes using convection-permitting ensembles. *Wea. Forecasting*, **31**, 273–295, https://doi.org/10.1175/WAF-D-15-0134.1.

Houze, R. A., 2010: Clouds in tropical cyclones. *Mon. Wea. Rev.*, **138**, 293–344, https://doi.org/10.1175/2009MWR2989.1.

Hu, M., and M. Xue, 2007: Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case. *Mon. Wea. Rev.*, **135**, 507–525, https://doi.org/10.1175/MWR3313.1.

——, ——, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698, https://doi.org/10.1175/MWR3092.1.

Jirak, I. L., S. J. Weiss, and C. J. Melick, 2012: The SPC storm-scale ensemble of opportunity: Overview and results from the 2012 Hazardous Weather Testbed Spring Forecasting Experiment. *26th Conf. on Severe Local Storms,* Nashville, TN, Amer. Meteor. Soc., 137, https://ams.confex.com/ams/26SLS/webprogram/Paper211729.html.

——, C. J. Melick, and S. J. Weiss, 2014: Combining probabilistic ensemble information from the environment with simulated storm attributes to generate calibrated probabilities of severe weather hazards. *27th Conf. on Severe Local Storms*, Madison, WI, Amer. Meteor. Soc., 2.5, https://ams.confex.com/ams/27SLS/webprogram/Paper254649.html.

Jung, Y., M. Xue, and M. Tong, 2012: Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric radar data. *Mon. Wea. Rev.*, **140**, 1457–1475, https://doi.org/10.1175/MWR-D-11-00032.1.

Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, https://doi.org/10.1175/WAF2007106.1.

——, S. R. Dembek, S. J. Weiss, J. L. Case, J. J. Levit, and R. A. Sobash, 2010: Extracting unique information from high-resolution forecast models: Monitoring selected fields and phenomena every time step. *Wea. Forecasting*, **25**, 1536–1542, https://doi.org/10.1175/2010WAF2222430.1.

Karstens, C. D., T. M. Samaras, B. D. Lee, W. A. Gallus, and C. A. Finley, 2010: Near-ground pressure and wind measurements in tornadoes. *Mon. Wea. Rev.*, **138**, 2570–2588, https://doi.org/10.1175/2010MWR3201.1.

Kong, F., and Coauthors, 2014: CAPS storm-scale ensemble forecasting system: Impact of IC and LBC perturbations. *26th Conf. on Weather Analysis and Forecasting/22nd Conf. on Numerical Weather Prediction*, Amer. Meteor. Soc., 119, https://ams.confex.com/ams/94Annual/webprogram/Paper234762.html.

Lerach, D. G., B. J. Gaudet, and W. R. Cotton, 2008: Idealized simulations of aerosol influences on tornadogenesis. *Geophys. Res. Lett.*, **35**, L23806, https://doi.org/10.1029/2008GL035617.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, https://doi.org/10.3402/tellusa.v21i3.10086.

Mahre, A., J. M. Kurdzo, D. J. Bodine, C. B. Griffin, R. D. Palmer, and T. Yu, 2018: Analysis of the 16 May 2015 Tipton, Oklahoma, EF-3 tornado at high spatiotemporal resolution using the Atmospheric Imaging Radar. *Mon. Wea. Rev.*, **146**, 2103–2124, https://doi.org/10.1175/MWR-D-17-0256.1.

Mashiko, W., 2016a: A numerical study of the 6 May 2012 Tsukuba City supercell tornado. Part I: Vorticity sources of low-level and midlevel mesocyclones. *Mon. Wea. Rev.*, **144**, 1069–1092, https://doi.org/10.1175/MWR-D-15-0123.1.

——, 2016b: A numerical study of the 6 May 2012 Tsukuba City supercell tornado. Part II: Mechanisms of tornadogenesis. *Mon. Wea. Rev.*, **144**, 3077–3098, https://doi.org/10.1175/MWR-D-15-0122.1.

——, H. Niino, and K. Teruyuki, 2009: Numerical simulations of tornadogenesis in an outer-rainband minisupercell of Typhoon Shanshan on 17 September 2006. *Mon. Wea. Rev.*, **137**, 4238–4260, https://doi.org/10.1175/2009MWR2959.1.

Mason, I. B., 1982: A model for the assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291–303.

Mason, S. J., and N. E. Graham, 1999: Conditional probabilities, relative operating characteristics, and relative operating levels. *Wea. Forecasting*, **14**, 713–725, https://doi.org/10.1175/1520-0434(1999)014<0713:CPROCA>2.0.CO;2.

Milbrandt, J. A., and M. K. Yau, 2005: A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. *J. Atmos. Sci.*, **62**, 3065–3081, https://doi.org/10.1175/JAS3535.1.

Roberts, B., and M. Xue, 2017: The role of surface drag in mesocyclone intensification leading to tornadogenesis within an idealized supercell simulation. *J. Atmos. Sci.*, **74**, 3055–3077, https://doi.org/10.1175/JAS-D-16-0364.1.

——, ——, A. D. Schenkman, and D. T. Dawson II, 2016: The role of surface friction in tornadogenesis within an idealized

supercell simulation. *J. Atmos. Sci.*, **73**, 3371–3395, https://doi.org/10.1175/JAS-D-15-0332.1.

Schenkman, A., M. Xue, A. Shapiro, K. Brewster, and J. Gao, 2011: The analysis and prediction of the 8–9 May 2007 Oklahoma tornadic mesoscale convective system by assimilating WSR-88D and CASA radar data using 3DVAR. *Mon. Wea. Rev.*, **139**, 224–246, https://doi.org/10.1175/2010MWR3336.1.

Schenkman, A. D., M. Xue, and A. Shapiro, 2012: Tornadogenesis in a simulated mesovortex within a mesoscale convective system. *J. Atmos. Sci.*, **69**, 3372–3390, https://doi.org/10.1175/JAS-D-12-038.1.

Schenkman, A., M. Xue, and M. Hu, 2014: Tornadogenesis within numerically simulated 8 May 2003 Oklahoma City tornadic supercell storm. *J. Atmos. Sci.*, **71**, 130–154, https://doi.org/10.1175/JAS-D-13-073.1.

Schenkman, A. D., M. Xue, and D. T. Dawson, 2016: The cause of internal outflow surges in a high-resolution simulation of the 8 May 2003 Oklahoma City tornadic supercell. *J. Atmos. Sci.*, **73**, 353–370, https://doi.org/10.1175/JAS-D-15-0112.1.

Schwartz, C. S., and R. A. Sobash, 2017: Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: A review and recommendations. *Mon. Wea. Rev.*, **145**, 3397–3418, https://doi.org/10.1175/MWR-D-16-0400.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Skinner, P. S., and Coauthors, 2018: Object-based verification of a prototype Warn-on-Forecast system. *Wea. Forecasting*, **33**, 1225–1250, https://doi.org/10.1175/WAF-D-18-0020.1.

Snook, N., and M. Xue, 2008: Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms. *Geophys. Res. Lett.*, **35**, L24803, https://doi.org/10.1029/2008GL035866.

——, ——, and Y. Jung, 2012: Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data. *Mon. Wea. Rev.*, **140**, 2126–2146, https://doi.org/10.1175/MWR-D-11-00117.1.

——, ——, and ——, 2015: Multiscale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Mon. Wea. Rev.*, **143**, 1035–1057, https://doi.org/10.1175/MWR-D-13-00262.1.

——, Y. Jung, J. Brotzge, B. Putnam, and M. Xue, 2016: Prediction and ensemble forecast verification of hail in the supercell storms of 20 May 2013. *Wea. Forecasting*, **31**, 811–825, https://doi.org/10.1175/WAF-D-15-0152.1.

Sobash, R. A., J. S. Kain, D. R. Bright, A. R. Dean, M. C. Coniglio, and S. J. Weiss, 2011: Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. *Wea. Forecasting*, **26**, 714–728, https://doi.org/10.1175/WAF-D-10-05046.1.

Stensrud, D. J., and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16, https://doi.org/10.1016/j.atmosres.2012.04.004.

——, and J. Gao, 2010: Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short-range forecasts. *Mon. Wea. Rev.*, **138**, 1250–1272, https://doi.org/10.1175/2009MWR3027.1.

——, and Coauthors, 2009: Convective-scale Warn-on-Forecast system: A Vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499, https://doi.org/10.1175/2009BAMS2795.1.

Sun, W., and C. Chang, 1986: Diffusion model for a convective layer. Part I: Numerical simulation of convective boundary layer. *J. Climate Appl. Meteor.*, **25**, 1445–1453, https://doi.org/10.1175/1520-0450(1986)025<1445:DMFACL>2.0.CO;2.

Supinie, T., Y. Jung, M. Xue, D. Stensrud, M. French, and H. Bluestein, 2016: Impact of VORTEX2 observations on analyses and forecasts of the 5 June 2009 Goshen County, Wyoming, supercell. *Mon. Wea. Rev.*, **144**, 429–449, https://doi.org/10.1175/MWR-D-15-0171.1.

Supinie, T. A., N. Yussouf, Y. Jung, M. Xue, J. Cheng, and S. Wang, 2017: Comparison of the analyses and forecasts of a tornadic supercell storm from assimilating phased-array radar and WSR-88D observations. *Wea. Forecasting*, **32**, 1379–1401, https://doi.org/10.1175/WAF-D-16-0159.1.

Tanamachi, R. L., H. B. Bluestein, J. B. Houser, S. J. Frasier, and K. M. Hardwick, 2012: Mobile, X-band, polarimetric Doppler radar observations of the 4 May 2007 Greensburg, Kansas, tornadic supercell. *Mon. Wea. Rev.*, **140**, 2103–2125, https://doi.org/10.1175/MWR-D-11-00142.1.

——, L. J. Wicker, D. C. Dowell, H. B. Bluestein, D. T. Dawson II, and M. Xue, 2013: EnKF assimilation of high-resolution, mobile Doppler radar data of the 4 May 2007 Greensburg, Kansas, supercell into a numerical cloud model. *Mon. Wea. Rev.*, **141**, 625–648, https://doi.org/10.1175/MWR-D-12-00099.1.

Tong, M., and M. Xue, 2008: Simultaneous estimation of microphysical parameters and atmospheric state with radar data and ensemble square-root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, **136**, 1630–1648, https://doi.org/10.1175/2007MWR2070.1.

Wang, Y., Y. Jung, T. A. Supinie, and M. Xue, 2013: A hybrid MPI/OpenMP parallel algorithm and performance analysis for an ensemble square root filter suitable for dense observations. *J. Atmos. Oceanic Technol.*, **30**, 1382–1397, https://doi.org/10.1175/JTECH-D-12-00165.1.

Wicker, L. J., and R. B. Wilhelmson, 1995: Simulation and analysis of tornado development and decay within a three-dimensional supercell thunderstorm. *J. Atmos. Sci.*, **52**, 2675–2703, https://doi.org/10.1175/1520-0469(1995)052<2675:SAAOTD>2.0.CO;2.

Wood, V. T., and R. A. Brown, 2011: Simulated tornadic vortex signatures of tornado-like vortices having one- and two-celled structures. *J. Appl. Meteor. Climatol.*, **50**, 2338–2342, https://doi.org/10.1175/JAMC-D-11-0118.1.

Wurman, J., and K. Kosiba, 2013: Finescale radar observations of tornado and mesocyclone structures. *Wea. Forecasting*, **28**, 1157–1174, https://doi.org/10.1175/WAF-D-12-00127.1.

Xu, Q., L. Wei, and K. Nai, 2015: Analyzing vortex winds in radar-observed tornadic mesocyclones for nowcast applications. *Wea. Forecasting*, **30**, 1140–1157, https://doi.org/10.1175/WAF-D-15-0046.1.

——, ——, and ——, 2017: A three-step method for estimating vortex center locations in four-dimensional space from radar-observed tornadic mesocyclones. *J. Atmos. Oceanic Technol.*, **34**, 2275–2281, https://doi.org/10.1175/JTECH-D-17-0123.1.

Xu, X., M. Xue, and Y. Wang, 2015: Mesovortices within the 8 May 2009 bow echo over the central United States: Analyses of the characteristics and evolution based on Doppler radar observations and a high-resolution model simulation. *Mon. Wea. Rev.*, **143**, 2266–2290, https://doi.org/10.1175/MWR-D-14-00234.1.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.*, **75**, 161–193, https://doi.org/10.1007/s007030070003.

——, and Coauthors, 2001: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76**, 143–165, https://doi.org/10.1007/s007030170027.

——, M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Oceanic Technol.*, **23**, 46–66, https://doi.org/10.1175/JTECH1835.1.

——, M. Hu, and A. D. Schenkman, 2014: Numerical prediction of the 8 May 2003 Oklahoma City tornadic supercell and embedded tornado using ARPS with the assimilation of WSR-88D data. *Wea. Forecasting*, **29**, 39–62, https://doi.org/10.1175/WAF-D-13-00029.1.

Yokota, S., H. Niino, H. Seko, M. Kunii, and H. Yamauchi, 2018: Important factors for tornadogenesis as revealed by high-resolution ensemble forecasts of the Tsukuba supercell tornado of 6 May 2012 in Japan. *Mon. Wea. Rev.*, **146**, 1109–1132, https://doi.org/10.1175/MWR-D-17-0254.1.

Yussouf, N., D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**, 3044–3066, https://doi.org/10.1175/MWR-D-14-00268.1.

Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2015: Practical predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma: Sensitivity to synoptic timing and topographical influence. *Mon. Wea. Rev.*, **143**, 2973–2997, https://doi.org/10.1175/MWR-D-14-00394.1.