Doppler Lidar and Mobile Radiosonde Observation-Based Evaluation of Warn-on-Forecast System Predicted Near-Supercell Environments during TORUS 2019

JORDAN J. LASER,[a] MICHAEL C. CONIGLIO,[b,a] PATRICK S. SKINNER,[a,c,b] AND ELIZABETH N. SMITH[b,a]

[a] *School of Meteorology, University of Oklahoma, Norman, Oklahoma*
[b] *National Severe Storms Laboratory, Norman, Oklahoma*
[c] *Cooperative Institute for Severe and High Impact Weather Research and Operations, Norman, Oklahoma*

(Manuscript received 30 November 2021, in final form 18 July 2022)

ABSTRACT: Observational data collection is extremely hazardous in supercell storm environments, which makes for a scarcity of data used for evaluating the storm-scale guidance from convection allowing models (CAMs) like the National Oceanic and Atmospheric Administration (NOAA) Warn-on-Forecast System (WoFS). The Targeted Observations with UAS and Radar of Supercells (TORUS) 2019 field mission provided a rare opportunity to not only collect these observations, but to do so with advanced technology: vertically pointing Doppler lidar. One standing question for WoFS is how the system forecasts the feedback between supercells and their near-storm environment. The lidar can observe vertical profiles of wind over time, creating unique datasets to compare to WoFS kinematic predictions in rapidly evolving severe weather environments. Mobile radiosonde data are also presented to provide a thermodynamic comparison. The five lidar deployments (three of which observed tornadic supercells) analyzed show WoFS accurately predicted general kinematic trends in the inflow environment; however, the predicted feedback between the supercell and its environment, which resulted in enhanced inflow and larger storm-relative helicity (SRH), were muted relative to observations. The radiosonde observations reveal an overprediction of CAPE in WoFS forecasts, both in the near and far field, with an inverse relationship between the CAPE errors and distance from the storm.

SIGNIFICANCE STATEMENT: It is difficult to evaluate the accuracy of weather prediction model forecasts of severe thunderstorms because observations are rarely available near the storms. However, the TORUS 2019 field experiment collected multiple specialized observations in the near-storm environment of supercells, which are compared to the same near-storm environments predicted by the National Oceanic and Atmospheric Administration (NOAA) Warn-on-Forecast System (WoFS) to gauge its performance. Unique to this study is the use of mobile Doppler lidar observations in the evaluation; lidar can retrieve the horizontal winds in the few kilometers above ground on time scales of a few minutes. Using lidar and radiosonde observations in the near-storm environment of three tornadic supercells, we find that WoFS generally predicts the expected trends in the evolution of the near-storm wind profile, but the response is muted compared to observations. We also find an inverse relationship of errors in instability to distance from the storm. These results can aid model developers in refining model physics to better predict severe storms.

KEYWORDS: Storm environments; Supercells; Lidars/Lidar observations; Radiosonde/rawinsonde observations; Ensembles; Model evaluation/performance

1. Introduction

For several decades, numerical weather prediction systems have shown the potential to reproduce the structure of supercell thunderstorms and their associated severe weather hazards (Klemp and Wilhelmson 1978; Weisman and Klemp 1982; Wicker and Wilhelmson 1995). With these early, idealized studies came the desire to explicitly predict supercells and their hazards before severe weather occurs (Lilly 1990). Continued advancement in computer power has allowed for high-resolution (3–4-km horizontal grid spacing, >50 vertical levels), deterministic convection allowing models (CAMs) to produce forecasts rapidly enough to provide useful information to forecasters in real time (Pielke and Carbone 2002; Weisman et al. 2008).

As advancements in modeling and data assimilation of atmospheric observations continued, attention turned to an ensemble approach to account for rapid, flow-dependent convection-scale error growth in CAMs (Hohenegger and Schar 2007; Zhang et al. 2016; Judt 2020). One such ensemble numerical weather prediction system being developed within the National Oceanic and Atmospheric Administration (NOAA), the Warn-on-Forecast System (WoFS), aims to provide short-term (0–6 h) probabilistic guidance for severe weather events like hail, flash floods and tornadoes (Stensrud et al. 2009; Lawson et al. 2018). WoFS prototypes have produced skillful predictions of severe hail (Snook et al. 2016), tornadic mesocyclones (Dawson et al. 2012; Yussouf et al. 2013), and flash flooding events (Yussouf et al. 2016). Recent work shows that real-time WoFS runs, from 2016 onward (Wheatley et al. 2015; Jones et al. 2016, 2020), can produce accurate forecasts of thunderstorm and mesocyclone position on spatiotemporal scales typical of National Weather Service (NWS) warnings (Skinner 2016; Skinner et al. 2018; Flora et al. 2019).

Corresponding author: Michael Coniglio, Michael.Coniglio@noaa.gov

DOI: 10.1175/WAF-D-21-0190.1

An important component of WoFS development is the validation of forecasts to inform system developers of model deficiencies and to inform researchers on how to best communicate findings that have relevant implications for the performance of the model. These verification efforts have focused on comparing simulated and observed storm characteristics, like storm placement, intensity, and rotation (Skinner et al. 2018; Flora et al. 2019). However, an important component of model performance is the depiction of the storm environment (Stensrud and Gao 2010; Coniglio et al. 2013, 2019), which has a large control on storm characteristics and evolution (e.g., Weisman and Klemp 1982; Rasmussen and Blanchard 1998; Markowski and Richardson 2014). In particular, a significant source of environment error in CAMs is thought to be from the planetary boundary layer (PBL) scheme, which is necessary to parameterize the effects of subgrid-scale turbulence on gridscale variables. The two categories of PBL schemes are referred to as local and nonlocal and, with their contrasting biases, provide spread around the ensemble mean. Local schemes generally undermix the PBL, leading to negative errors in temperature and PBL height[1] and positive errors in moisture. The nonlocal schemes typically predict more mixed, deeper, warmer, and drier PBLs (see Cohen et al. 2015 for a summary of these biases in CAM forecasts). As described in more detail in section 2, WoFS uses three PBL schemes (two local, one nonlocal) that produce known biases in temperature and humidity in the lower atmosphere. Some of these biases may have noticeable impacts on storm-scale characteristics in WoFS (Potvin et al. 2020), which provides motivation to continue exploring PBL scheme errors in WoFS.

Another motivation for this study is that storms can modify their ambient environment through diabatic temperature changes and dynamic effects associated with deep convection, such as rotation in updrafts (Rotunno and Klemp 1982; Klemp and Rotunno 1983). For instance, low pressure that forms beneath a strong rotating updraft causes an acceleration of air toward the storm and enhances the wind field near the supercell, increasing local storm-relative helicity (SRH) (Parker 2014; Wade et al. 2018; Flournoy et al. 2020). These feedbacks can be depicted in CAM forecasts (Kerr et al. 2019; Nowotarski and Markowski 2016, Potvin et al. 2019, 2020). Other than Kerr et al. (2019), studies that evaluate model depictions of storm environments focus on "undisturbed" conditions away from deep convection in an effort to sample the ambient environment in which the storm is thought to develop and respond. However, Kerr et al. (2019) presents evidence that feedbacks to the environment from the storm can alter the storm itself. Therefore, if feedbacks can be influential in storm evolution, it is important to understand how well CAMs depict storm–environment feedbacks to more fully explore storm–environment errors.

Prior studies have also tended to explore conditions thought to be in undisturbed environments because direct observations of conditions close to storms are relatively scarce, especially above the ground. This scarcity is due in part to practical difficulties in obtaining those observations; Potvin et al. (2010) show that storms are much more likely to influence the environment within 40 km of the updraft, usually termed the "near-field"[2] of the storm. Observations that could be used to validate environments within this distance of storms are mostly relegated to those collected from specialized field projects [e.g., VORTEX2 (Wurman et al. 2012), MPEX (Weisman et al. 2015), PECAN (Geerts et al. 2017)]. The overlap of routinely available CAM ensembles designed to predict the short-term evolution of storms (like WoFS) and field projects that can provide the necessary observations to validate storm environments is small and limited to the last 5–10 years.

The Targeted Observations with UAS and Radar of Supercells (TORUS, https://www.eol.ucar.edu/field_projects/torus) field campaign took place in 2019, during a period when WoFS ran routinely over the same areas targeted by TORUS, and provides a set of unique observations to validate storm–environment feedbacks in WoFS. Therefore, a goal of this study is to use these observations to explore how WoFS depicts these feedbacks.

Past studies of ambient supercell-environments (Rasmussen and Blanchard 1998; Thompson et al. 2003; Potvin et al. 2010; Parker 2014; Wade et al. 2018), supercell characteristics (Klemp and Wilhelmson 1978; Coniglio and Parker 2020; Flournoy et al. 2020), and the feedbacks between the two (Markowski and Richardson 2014) motivate investigating how accurately WoFS is predicting this complex system. The two-way feedbacks between a supercell and its inflow environment necessitate verification of numerical weather prediction systems' ability to resolve these processes, which, in turn require specialized observations of the supercell and its inflow environment. Knowledge of the model error characteristics and evolution of supercell inflow environments informs development of WoFS and other convection-allowing systems. This paper aims to add to the studies of convection-allowing numerical weather prediction systems, specifically regarding PBL scheme performance. Many past verification studies typically use large numbers of operational radiosonde observations for the verification dataset; however, as described in more detail below, this research is the first to compare WoFS environmental forecasts to dense spatiotemporal observations of the inflow of supercells using both Doppler wind lidar and mobile radiosondes. Additionally, to the authors' knowledge, this study is the first to use high temporal resolution Doppler wind lidar to verify predictions of mesoscale modification of the near-storm wind field of supercells.

[1] For this study, the top of the PBL is understood as the height at which lapse rates drop below 2 K km^{-1}; typically this height coincides with pressure between 700 and 900 hPa.

[2] Herein, the "near-field" is used to describe distances within 40-km of the storm updraft and "far-field" refers to distances greater than 40 km.

2. Data and methods

a. *TORUS observations*

TORUS is a collaborative project including the University of Nebraska, University of Colorado, Texas Tech University, University of Oklahoma, the National Severe Storms Laboratory (NSSL), and the Cooperative Institute for Mesoscale Meteorological Studies (CIMMS)[3] and is motivated by improving the current conceptual model of supercell thunderstorms. The primary aim of TORUS is to reveal the four-dimensional character of storm-generated boundaries and coherent structures that are crucial to tornadogenesis and the evolution of supercells.

A plethora of observational platforms were used for TORUS: Doppler wind lidar, mobile mesonets (Straka et al. 1996; Waugh 2021), mobile radiosondes, uncrewed aerial systems (UAS) (Argrow and Houston 2007), Ka-band (Weiss et al. 2009; Duncan et al. 2019) and X-band (Burgess et al. 2010) mobile Doppler radars, and the NOAA P-3 Orion Hurricane Hunter aircraft (Collins and Flaherty 2014). These platforms were deployed close to and in the near-field inflow region of supercells. The first TORUS field phase occurred from 15 May to 15 June 2019, during which 19 supercells were sampled. Previous studies have used large (>100) radiosonde datasets to statistically evaluate model representation of the PBL in CAMs (Coniglio 2012; Coniglio et al. 2013; Evans et al. 2018). However, similar to Kerr et al. (2019), this study uses a case study approach to inspect WoFS forecasts of PBL characteristics given the unique opportunity to evaluate how a CAM predicts the near-storm environment. Three days of tornadic supercells from TORUS are the focus in this study: 17, 20, and 23 May.

This study focuses on profiles of the lower atmosphere collected with Doppler wind lidar and mobile radiosondes because these data provide the best, highest-resolution (in time and height above the ground) datasets available from TORUS of the PBL.[4] There were 118 radiosondes launched and 37 deployments of the Doppler wind lidar, of which a subset of 28 radiosonde launches and 8 lidar deployments occurred across the three case studies examined here.

Temperature, relative humidity, and winds (through GPS) were collected every 1 s with Vaisala RS41-SGP radiosondes (Vaisala 2018), which yields a vertical resolution of 4–5 m given the nominal balloon ascent rate of 4–5 m s^{-1}. Vaisala applies proprietary filters to the raw radiosonde data to produce the profiles used in this study as truth. Each sounding was inspected manually by TORUS principal investigators (PIs) to ensure data consistency and reliability. Observations collected outside of the inflow to the target supercell or any observations prior to WoFS initialization at 1900 UTC were not considered.

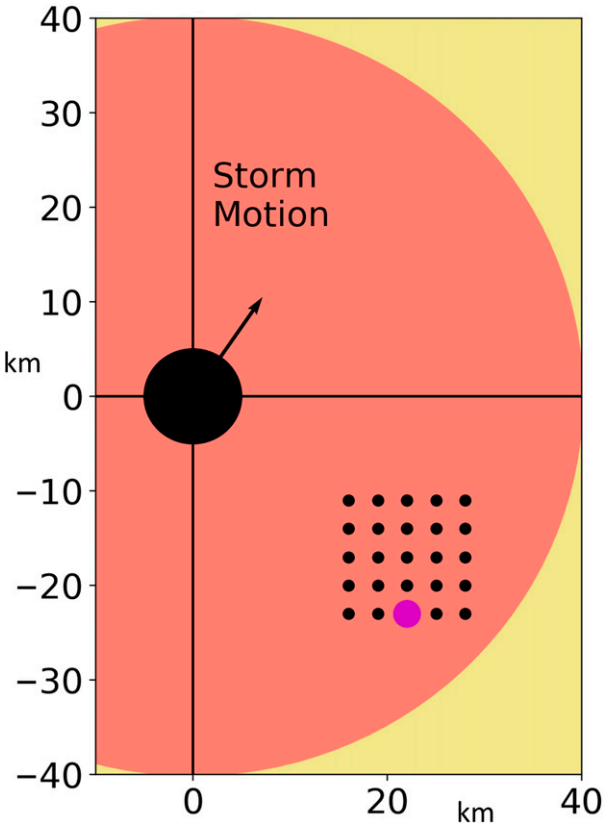


FIG. 1. Small black dots represent all the WoFS grid points included in the convective parameter comparisons. The magenta dot represents the nearest WoFS grid point to the near-ground location of the sounding observation. The large black dot and arrow represents the storm location and direction of motion. The salmon filled circle shows distances from the storm within 40 km (the near field); yellow is outside 40 km.

A sounding was considered to be in the inflow if the mean storm-relative PBL wind has a component that points toward the supercell updraft. Radiosondes are carried by the mean flow and can cover >50 km in horizontal distance during their ascent to the tropopause. Considering WoFS outputs convective parameters on a grid fixed to the ground, potential position errors caused by balloon drift and environmental variability are accounted for by comparing the observed values to a distribution of modeled convective parameter values from a 15 × 15 km^2 grid north of the observation and centered zonally (Fig. 1). In general, the winds change direction with height, but there is a strong southerly component to the wind throughout most of the atmosphere across all three cases. This motivates the meridionally symmetric and northern selection of the WoFS grid points used in the parameter analysis. This approach additionally accounts for small-scale environmental heterogeneity and allows radiosonde observations to be compared to a distribution of WoFS values.

The Doppler wind lidar is a Halo Photonics Streamline XR that was mounted in the bed of a pickup truck and was stationary and level with the ground while scanning. The

[3] CIMMS is now known as the Cooperative Institute for Severe and High-Impact Weather Research and Operations (CIWRO).

[4] Note that the data collected in the PBL by TORUS from the other observational platforms (e.g., UASs, P-3 compact Raman lidar, and dual-Doppler analyses) continued to undergo quality control (QC) and were not available for use at the time of this study.

standard lidar scanning strategy during TORUS was a 70° elevation plan position indicator (PPI) scan (including 12 azimuths) every 2 min with vertical stares in between. The vertical profile of horizontal wind components were derived by applying a velocity–azimuthal display (VAD) technique to the PPI scans (Liang 2007). The lidar parameters allow for high vertical and temporal resolution wind profiles; approximately 18 m in the vertical (starting at ~60 m AGL) and 30–45 s in time.

The data collection strategy varied by vehicle: one platform (termed the far-field sounding vehicle, or FFS) was tasked with releasing soundings every 45 min within the inflow at 40–80 km from the supercell. Two mobile mesonet vehicles also occasionally deployed radiosondes while performing transects of the near-field inflow. The lidar vehicle would position itself to capture the transition from the far-field inflow environment sampled by FFS to the potentially storm-modified near-field inflow environment and deployed radiosondes in concert with lidar deployments. The horizontal wind profiles obtained by the lidar have been shown to agree well with the canonical radiosonde vertical profiles in the lower levels (Smith et al. 2020).

The lidar data were quality controlled with a signal-to-noise ratio filter, typically leaving the first available data at around 60 m AGL and terminating somewhere below 2000 m AGL (due to loss of scattering particles or the presence of cloud water, which quickly attenuates the "laser" pulse). To create observational datasets comparable to the WoFS grid, both lidar and radiosonde datasets are resampled at the height (pressure) corresponding to vertical levels in the WoFS for kinematic (thermodynamic) quantities using a 100-m averaging window. For example, observed data between 150 and 250 m AGL are averaged to compare to a WoFS vertical level at 200 m AGL. This averaging also reduces noise that could contaminate point-to-point comparison.

b. *WoFS forecasts*

WoFS is a Weather Research and Forecasting (WRF) Model and Advanced Research version of WRF (WRF-ARW)-based (Skamarock and Klemp 2008) CAM ensemble designed to provide probabilistic guidance for severe weather threats (Wheatley et al. 2015; Jones et al. 2016). The version used for this study ran in 2019 over a 900 × 900 km^2 domain with 3-km horizontal grid spacing. WoFS uses a 51-level stretched vertical grid with grid spacing of approximately 100 m in the boundary layer stretching to 1 km at model top (10 hPa). WSR-88D reflectivity and radial velocity along with satellite observations of cloud water path and clear sky radiances are assimilated every 15 min while surface observations are assimilated every hour using a customized version of the ensemble Kalman filter in the Community Gridpoint Statistical Interpolation (GSI) software package (Kleist et al. 2009; Hu et al. 2016). Complete details of the WoFS configuration used in 2019 are available in Jones et al. (2020). WoFS forecasts use 36 unique members for analyses and issues 18-member forecasts differentiated by their PBL, longwave, and shortwave radiation schemes. Past studies have shown negligible differences in forecast quality across radiation schemes (Potvin et al. 2020), motivating the focus here on the different PBL schemes.

Three PBL schemes are distributed among the 18 unique forecast members in WoFS: the Yonsei University (YSU) (Hong et al. 2006), the Mellor–Yamada–Nakanishi–Niino (MYNN; Nakanishi and Niino 2004, 2006), and the Mellor–Yamada–Janjić (MYJ; Janjić 2002) schemes. These schemes parameterize the larger-scale impacts of turbulence differently. The local (MYJ and MYNN) schemes constrain the turbulent-mixing equations by indexing only adjacent vertical grid points in calculations of state variables, while the nonlocal (YSU) scheme indexes beyond adjacent grid points to incorporate the impacts of deep eddies in convective boundary layers (Stull 1991; Cohen et al. 2015). Numerous studies in the past decades have explored general characteristics of these schemes in CAMs (e.g., Coniglio et al. 2013; Cohen et al. 2015; Evans et al. 2018), with a general finding that nonlocal (local) schemes tend to over(under)mix the PBL in CAM forecasts, leading to warmer (cooler) and drier (more moist) PBLs in the nonlocal (local) schemes. The impact of enhanced mixing in the nonlocal schemes, as shown by Potvin et al. (2020), is that they produce environments that are less favorable for storm intensification and yield weaker storms than the local schemes. However, the impact of these errors on parameters related to convection, like CAPE and SRH, and on the convection itself do not seem to show clear biases across these studies, with perhaps the most consistent finding being an underprediction of CIN. For the WoFS, Potvin et al. (2020) confirm these findings and suggests the underprediction of CIN leads to an initial overprediction of storm coverage in all three PBL schemes in WoFS. These latter two results provide motivation to continue exploring the relationship between errors in the PBL and WoFS storm-scale forecasts.

In the spring of 2019, WoFS forecasts were produced every 30 min in the afternoon to early evening hours. For this study, the forecast initialization and valid times are chosen by selecting the nearest forecast valid time to the observation, then selecting the nearest associated initialization time, provided that the valid forecast is between 15 and 45 min after the initialization time (e.g., a 2236 UTC observation is compared to a WoFS forecast initialized at 2200 UTC and valid at 2235 UTC). The 15–45-min range of lead time is chosen because these forecasts will be the most up-to-date forecasts in a real forecasting situation, given that WoFS forecasts become available roughly 15 min after initialization time and new forecasts are available every half-hour. WoFS output is available every 5 min, resulting in a maximum time difference between the nominal observation time and valid forecast time of 2.5 min. Hereafter, references to the observation–WoFS forecast pairs will appear as follows: observation time, valid WoFS forecast time used in comparison (e.g., 2302, 2300 UTC).

Because initial and boundary conditions [provided by the High-Resolution Rapid Refresh Ensemble (HRRRE)] vary and errors in CAMs grow rapidly, the 18 members in WoFS predict storms that vary in location and strength compared to the observed storms. Owing to these differences in storm location, simply pulling the WoFS grid point nearest to the

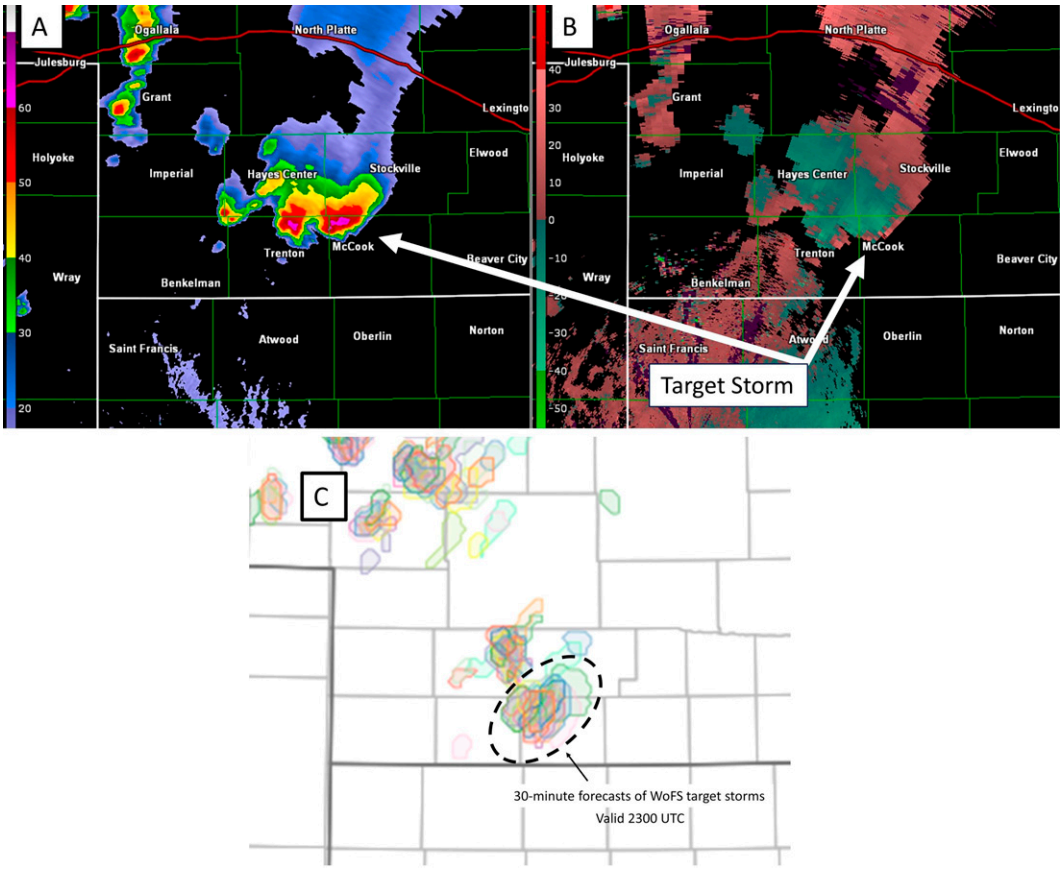
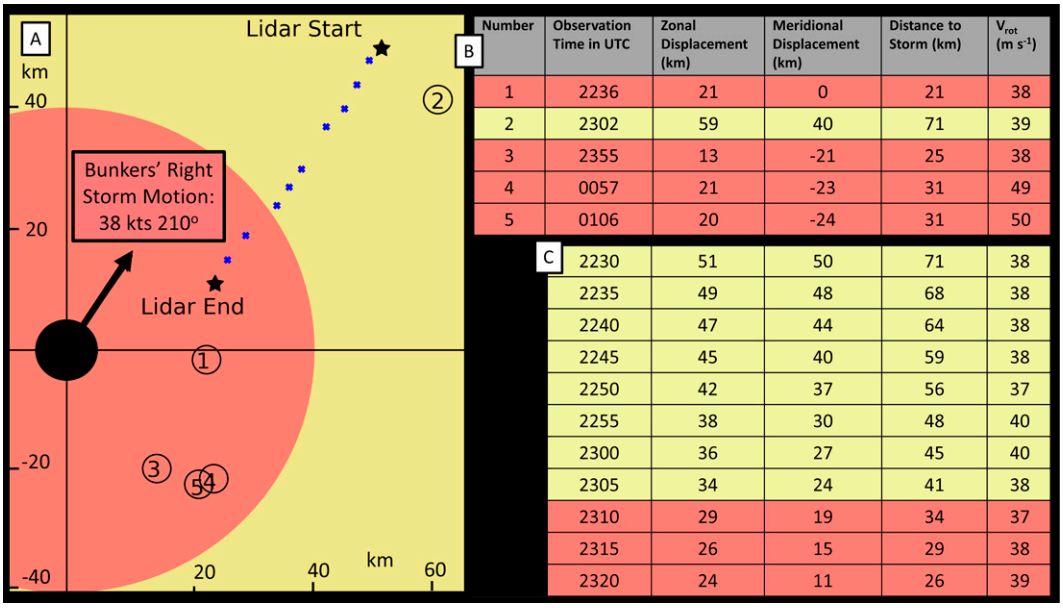


FIG. 2. Radar images from KGLD of the 17 May McCook, NE, supercell at 2300 UTC; 0.5° elevation angle. (a) Reflectivity (dBZ) and (b) storm-relative velocity (kt; 1 kt ≈ 0.51 m s^{-1}). (c) The dashed oval encompasses contours of 45-dBZ simulated composite reflectivity from 30-min forecasts from all 18 WoFS members valid at 2300 UTC and initialized at 2230 UTC 17 May. Different color shading depicts different ensemble members.

location of the observation may yield errors caused by incorrect storm locations in the model that are larger than the errors in the environment. Therefore, a storm-relative analysis is employed in this study to maintain a consistent distance between the storms and the observation location. This approach leaves out a spatially aligned verification, wherein WoFS would be assessed in its ability to forecast the location of storms. Instead, this method provides insight into the accuracy of WoFS to forecast the ambient environment in which the storms exist.

The storm-relative position within the model is determined by calculating the displacement vector from the observed storm to the observation, then adding this vector to the location of the modeled storms. To do so, a WoFS storm matching the observed storm must be found at every forecast time of interest. This was done manually with guidance provided by a storm-detection algorithm described in Britt et al. (2020). The algorithm creates storm objects using Python's scikit-image library (van der Walt et al. 2014) to identify coherent regions of 0–2-km updraft helicity (UH) at each 5-min output time in WoFS output and return the centroids. The initial objects are then sorted by proximity to the observed storm location (<30 km), and the first author manually selected the object that most resembled the observed storm. The location of the matching observed storm was determined using data from the nearest WSR-88D within Gibson Ridge radar analysis software to pull the latitude and longitude from the manually identified location of the mesocyclone. While it is usually difficult to match modeled and observed storms in longer-term (3+ h) forecasts of convection because of storm-scale error growth, WoFS is designed to assimilate and predict ongoing individual storms (Skinner et al. 2018) and does so accurately enough (position errors are generally <30 km) in the three case studies used here to present little uncertainty in matching the correct modeled storm to the observed storm.

Once the storm-relative location is determined in all ensemble members, a model sounding is extracted from the WoFS output for each time needed for the comparison to observations. Environmental variables like CAPE and SRH are extracted from the grid shown in Fig. 1, while wind, temperature, and dewpoint values are extracted from the single storm-relative grid point. The data are then compared to the



B

Number	Observation Time in UTC	Zonal Displacement (km)	Meridional Displacement (km)	Distance to Storm (km)	V_{rot} (m s^{-1})
1	2236	21	0	21	38
2	2302	59	40	71	39
3	2355	13	-21	25	38
4	0057	21	-23	31	49
5	0106	20	-24	31	50

C

2230	51	50	71	38
2235	49	48	68	38
2240	47	44	64	38
2245	45	40	59	38
2250	42	37	56	37
2255	38	30	48	40
2300	36	27	45	40
2305	34	24	41	38
2310	29	19	34	37
2315	26	15	29	38
2320	24	11	26	39

FIG. 3. Storm-relative locations of observations collected on 17 May used in this study. (a) A similar configuration as in Fig. 1 and displays storm-relative radiosonde (numbers) and lidar (stars and X marks) observation locations. (b),(c) The five radiosonde and lidar times, storm-relative coordinates and distance, and V_{rot} of the targeted mesocyclone. In all panels, observations made farther than 40 km from the storm (far field) are shaded in yellow; observations made closer to the storm are shaded in red (near field). Units of zonal, meridional displacement, and distance to storm are in kilometers and the V_{rot} of the mesocyclone, a proxy for supercell strength, is measured in m s^{-1}.

resampled and averaged observations and differences are calculated (model − observation) for every member at every vertical grid point for which observational data are available.

To assess the strength of both the observed and modeled supercells, rotational velocity (V_{rot}) and 0–2-km UH are investigated, respectively. WoFS is able to provide a reliable proxy for mesocyclone strength (UH); however, this variable is not observable with traditional observing systems. Radial velocity derived from WSR-88D data (0.5° elevation angle) then serves as the observable proxy for mesocyclone strength. Updraft helicity contains direct information about updraft strength that V_{rot} does not; however, both metrics are proportional to mesocyclone strength. The strength of the mesocyclone is proportional to the low-level pressure perturbations induced by the mesocyclone and the inflow low that contributes to modifying the near-storm wind field.[5] Note that V_{rot} and updraft helicity are physically different quantities, so their magnitudes should not be directly compared and should only be used to compare the relative rotational evolution between the modeled and observed storms over time.

Rotational velocity V_{rot} is defined as

$$V_{rot} = \frac{V_{out} - V_{in}}{2}, \tag{1}$$

where V_{out} and V_{in} were manually chosen by selecting two gates with the largest velocity difference centered on the rotational couplet and not farther than 3 km apart. While UH and V_{rot} are proportional to the low pressure perturbation induced by the mesocyclone strength and thus the magnitude of storm modification, they both have limitations. For example, V_{rot} is an approximation of vertical vorticity assuming solid body rotation and UH is limited by the dependence on vertical velocity and by the grid spacing of WoFS.

The updraft helicty (UH) is defined as

$$\mathrm{UH} = \int_{0\,\mathrm{km}}^{2\,\mathrm{km}} w\zeta\, dz, \tag{2}$$

where w is the vertical velocity of the updraft and ζ is the vertical component of vorticity. The storm object location chosen to calculate UH was selected with the algorithms used by Britt et al. (2020) as discussed above. A single UH value is assigned to a storm by taking the maximum UH value within the storm object.

3. Results

a. 17 May 2019 McCook to Farnam, Nebraska, supercell

The first supercell examined was observed by TORUS and two WSR-88D radars at Goodland, Kansas (KGLD, Fig. 2), and North Platte, Nebraska (KTLX). Convection initiated along a dryline in western Kansas and evolved into a supercell near Goodland, Kansas, by 2100 UTC. The storm moved at an average heading and speed of 210° and ~20 m s^{-1} over the

[5] The size of the mesocyclone influences pressure perturbations (Markowski and Richardson 2010) and this is not directly accounted for in UH or V_{rot}; however, it is impractical to estimate the effects of mesocyclone size across all WoFS forecasts.

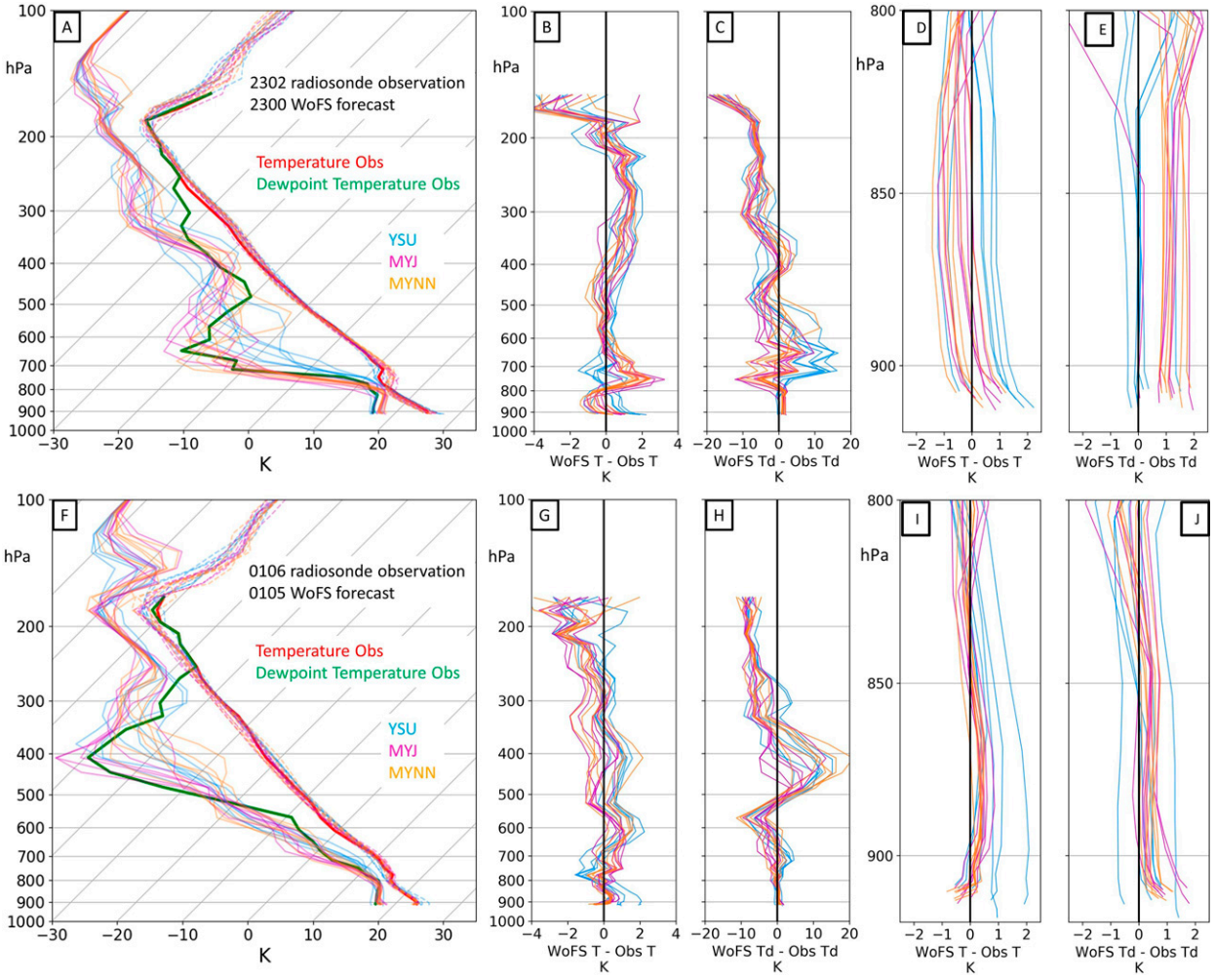


FIG. 4. (a) A sounding of the 2302 UTC radiosonde launch compared to the 2300 UTC WoFS forecast, initialized at 2230 UTC 17 May 2019. The red and green bold lines are the observed temperature and dewpoint temperature, respectively. (b),(c) Vertical profiles of the error in both temperature and dewpoint temperature (K). (d),(e) Plots of (b) and (c) zoomed in on the PBL for 2230 UTC. (f)–(j) As in (a)–(e), but for the 0106 UTC radiosonde launch and the 0105 UTC forecast initialized at 0030 UTC. Members are color-coded based on PBL scheme: YSU is blue, MYJ is magenta, and MYNN is orange.

following five hours and produced multiple tornadoes from 2240 to 0120 UTC ranging from EF0 to EF3 (NOAA/NCEI 2020). This case study focuses on observations collected by two radiosondes that made it to the equilibrium level (EL), launched at 2302 and 0106 UTC by FFS and on the one available lidar observation, which observed the low-level wind field nearly continuously between 2230 and 2320 UTC (Fig. 3). The lidar remained stationary during this period and entered the near-field of the approaching storm at approximately 2310 UTC.

The sounding comparisons at times 2302 UTC (taken in the far-field; Figs. 4a–e) and 0106 UTC (taken in the near-field; Figs. 4f–j) illustrate the typical WoFS mean errors in temperature among all five radiosondes for this case; they are mostly <2 K in magnitude throughout most of the troposphere (Figs. 4b–e,g–j). In both comparisons the ensemble members that use nonlocal YSU and local MYJ and MYNN schemes are behaving as anticipated, with the nonlocal scheme forecasting a drier, warmer, and more mixed boundary layer (Cohen et al. 2015, 2017). The YSU scheme forecasted the smallest errors in temperature (<1 K) near the location of the isothermal layer between 750 and 700 hPa, with MYJ and MYNN forecasts being too warm in this layer. However, as this isothermal layer becomes shallower and lowers later in the evening (Figs. 4f), the distribution of member temperature values become noticeably tighter. These comparisons illustrate that incorporating members that use the YSU scheme acts to provide spread to the temperature forecasts such that the observation falls within the WoFS members distribution much more frequently than they would in a MYJ/MYNN-only ensemble.

The dewpoint temperature errors are sensitive to the location of a dry layer from 700 to 500 hPa at 2302 UTC and from 550 to 350 hPa at 0106 UTC and result in mean errors as large

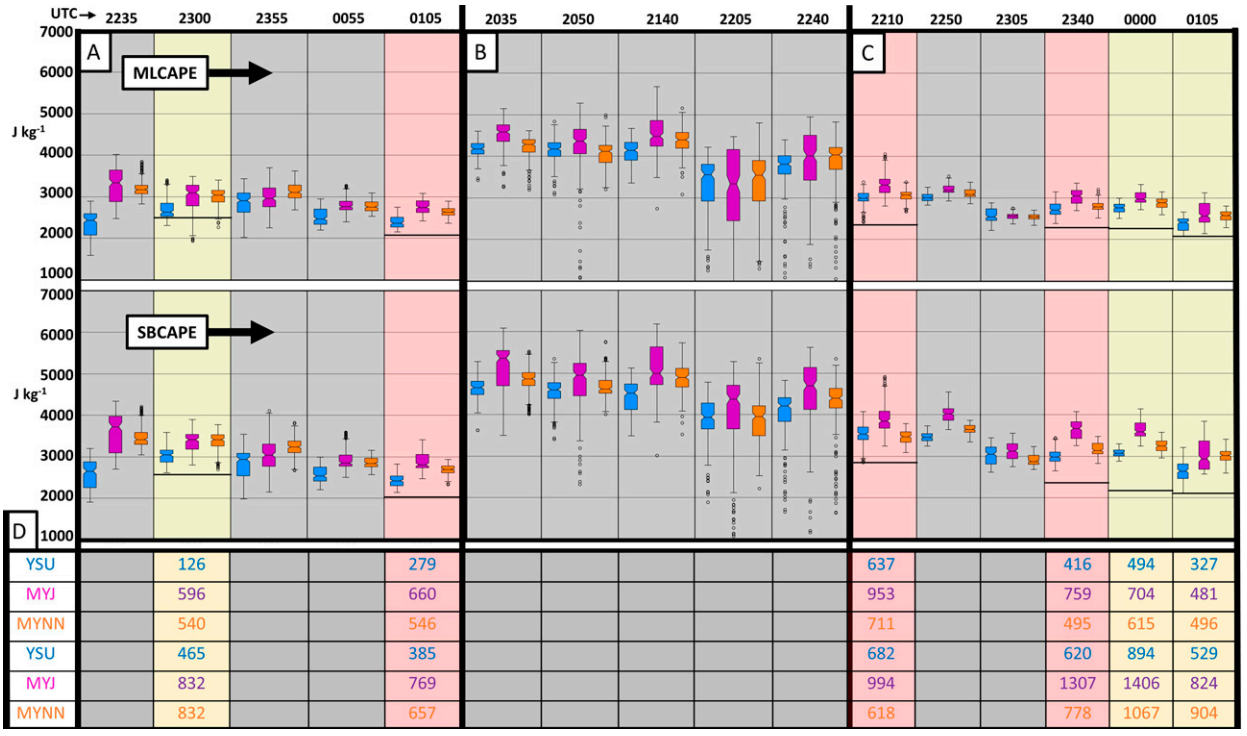


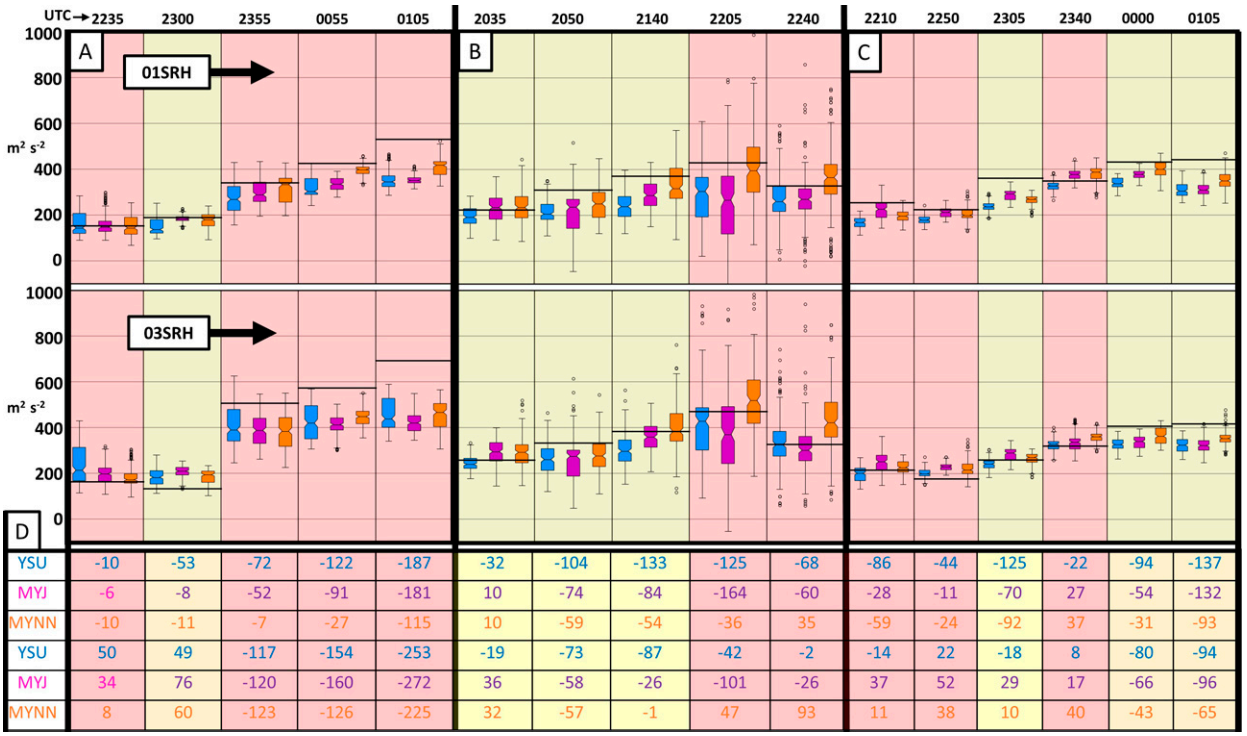
FIG. 5. Box-and-whisker plots of (top) MLCAPE (J kg^{-1}) and (middle) SBCAPE (J kg^{-1}) on days (a) 17 May, (b) 20 May, and (c) 23 May collected over all grid points within a 15 km × 15 km area near the radiosonde observation (Fig. 1). The notch in the box designates the median value; circles represent outliers. (d) The PBL scheme mean parameter value minus the observation value. Boxplots and table data are color coded by PBL scheme: YSU in blue, MYJ in magenta, and MYNN in orange. All columns refer to the respective time shown above at the top of the figure. The nearest forecast times are used for comparisons [e.g., WoFS forecasts initialized at 2200 UTC valid at 2235 UTC for an observation at 2236 UTC; observation–WoFS pairs are shown in the format of radiosonde (WoFS)]. Radiosondes 2236 (2235), 2355 (2355), and 0106 (0105) did not reach the EL so the respective CAPE plots are shaded gray and do not include an observed value, which are otherwise displayed as a horizontal black bar. In all panels, observations made farther than 40 km from the storm (far-field) are shaded in yellow; observations made closer to the storm (near field) are shaded in red.

as 10–20 K (Figs. 4c,h). At both times, the observed dewpoint temperature falls within WoFS member values at most altitudes below 300 hPa, but do so in the midtroposphere only through large spread among the members of 25 K. Overmixing in the YSU members results in overprediction of the dewpoint above the observed PBL top (in the 800–650-hPa layer), while the local schemes have smaller mean errors that are slightly below zero in this layer. This indicates, again, that having ensemble members that use the YSU scheme contribute to the ensemble spread such that the observation becomes more centered in the WoFS distribution of forecasts.

The observed profile saturates with respect to ice at 250 hPa as the radiosondes enter the anvil of the storm, but this is not forecasted in WoFS by any members (as identified using simulated satellite fields). The under forecast of anvil coverage should result in a positive PBL temperature error as a result of more surface heating than in reality (Frame and Markowski 2010, Frame and Markowski 2013); however, the observation still falls within the predicted temperatures of WoFS ensemble members (Figs. 4b,g). This may suggest there is a source of a negative temperature error in the PBL balancing the over prediction of surface heating.

Of the five radiosonde observations (1 far-field, 4 near-field; Fig. 3), again, only two ascended to the EL and can be used to compare the observed CAPE to WoFS predicted CAPE. The ensemble mean WoFS over forecasted mixed-layer (ML) CAPE (computed from a parcel averaged over the lowest 100 hPa) and surface-based (SB) CAPE in the near and far field (Fig. 5a). The CAPE decrease from 2300 to 0100 UTC of ~300 J kg^{-1} is well forecasted in WoFS; however, an approximately constant positive error of 300 (500) J kg^{-1} was present in YSU (MYJ/MYNN), implying that WoFS over predicted MLCAPE and SBCAPE at both 2300 and 0105 UTC.

WoFS 0–1-km SRH (01SRH) and 0–3-km SRH (03SRH) errors (Fig. 6) appear to depend more on time than distance from the storm. Observations made during the earlier two times (2235 and 2300 UTC) fall within the ensemble spread with relatively smaller mean errors compared to the observations made at the later three times (2355, 0055 and 0105 UTC). All ensemble members under predicted SRH at these times, with 01SRH (03SRH) errors from −100 to −200 m^2 s^{-2} (from −200 to −300 m^2 s^{-2}) at 0105 UTC. Although it is difficult to separate the influence of the storm from the usual increase in



YSU	-10	-53	-72	-122	-187	-32	-104	-133	-125	-68	-86	-44	-125	-22	-94	-137
MYJ	-6	-8	-52	-91	-181	10	-74	-84	-164	-60	-28	-11	-70	27	-54	-132
MYNN	-10	-11	-7	-27	-115	10	-59	-54	-36	35	-59	-24	-92	37	-31	-93
YSU	50	49	-117	-154	-253	-19	-73	-87	-42	-2	-14	22	-18	8	-80	-94
MYJ	34	76	-120	-160	-272	36	-58	-26	-101	-26	37	52	29	17	-66	-96
MYNN	8	60	-123	-126	-225	32	-57	-1	47	93	11	38	10	40	-43	-65

FIG. 6. As in Fig. 5, but showing (top) 01SRH ($m^2\ s^{-2}$) and (middle) 03SRH ($m^2\ s^{-2}$).

SRH toward/after sunset (Coffer and Parker 2015), SRH values in the far-field increase by 300 $m^2\ s^{-2}$ between 2200 and 0100 UTC, which show that WoFS is forecasting SRH increases as a result of the early evening transition (EET). The forecasted increase in SRH from the EET is still less than the observed increase in SRH, suggesting that WoFS could be under predicting the magnitude of storm-modification to the environment in the near field. Note that although the early observation at 2236 UTC was taken in the near field, the WoFS errors of SRH were much smaller, perhaps because the storm was in an earlier stage of maturity in which storm modifications did not extend far from the storm (Peters et al. 2022).

Evidence for weaker than observed storm modification is also seen when comparing the lidar observed wind profiles to the WoFS wind profiles (Fig. 7). The winds at all levels of lidar data were observed to back (turn counterclockwise with time) as the storm approached (Figs. 7a–c). Since the lidar was positioned to the northeast of the storm, this is likely a response to the storm-induced low pressure perturbation as the storm approached the stationary lidar. Observed V_{rot} estimated from the nearest WSR-88D (KGLD) over this time period is nearly constant at 40 m s^{-1} (Fig. 7f; beam heights were in the 2–3-km range), likely indicating that significant storm rotation, and associated pressure falls, were mainly responsible for the backing winds (Davies-Jones 2002). The winds also back in WoFS forecasts over this period; however, not nearly as much as in the observations—the predicted u component changes not more than ~3 m s^{-1} during the duration of the lidar deployment and the observed u component changes by ~8 m s^{-1} during the same time. At 2230 UTC, winds are more backed in WoFS in the lowest 500 m AGL than in the observations [negative u errors in green in Fig. 7e(iii)], but WoFS winds are veered relative to lidar observations by the end of the deployment owing to greater backing in the lidar data winds with the approach of the storm. WoFS winds are generally too strong above 500 m, with v component errors of +1–3 m s^{-1} [Fig. 7e(ii)].

The trend in UH over the period of the lidar deployment is examined to explore if the weaker storm modification of the low-level wind field in WoFS is the result of a predicted storm too weak to induce sufficient pressure perturbations (Fig. 7f). Although the UH of the modeled storm cannot be compared directly to observed V_{rot}, the lack of a significant backing signal in WoFS does not seem to be caused by a storm that is too weak; 0–2-km UH increases to a mean of 90 $m^2\ s^{-2}$ by the end of the period (Fig. 7f). These 0–2-km UH values are well above 30 $m^2\ s^{-2}$, which is roughly the minimum of the 0–3-km UH thresholds considered useful as proxies for severe weather for horizontal grid spacing of 3 km (Sobash et al. 2016), despite integrating UH over a shallower layer. WoFS may be simply under forecasting the storm-scale feedback between the low-level mesocyclone and the ambient wind field.

b. 20 May 2019 Mangum, Oklahoma, supercell

The supercell that would eventually produce a tornado near Mangum, Oklahoma, at 2210 UTC initiated ~30 km east of Lubbock, Texas, at 1700 UTC and was observed by the Frederick, Oklahoma (KFDR), WSR-88D radar during the

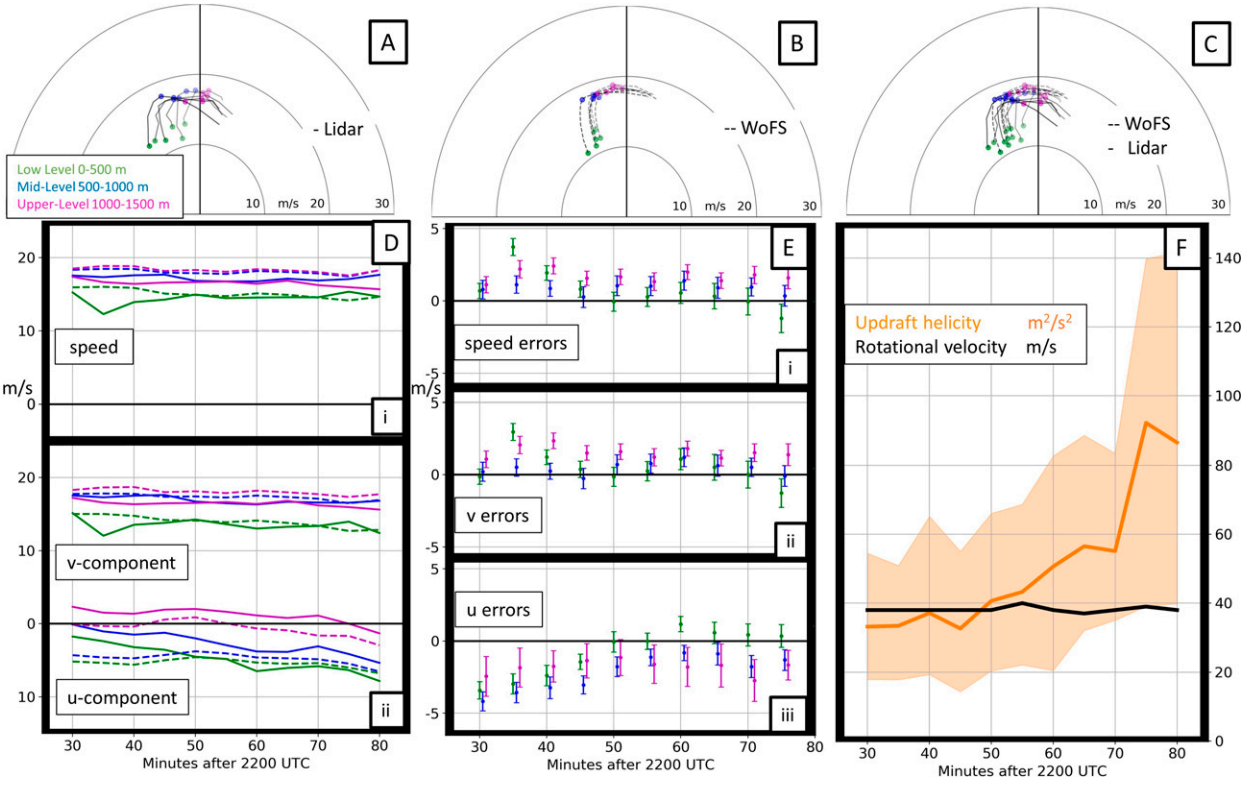


FIG. 7. Hodograph and wind-component errors plotted against time for the lidar observation from 2230 to 2320 UTC. (a)–(c) Lidar observed winds, WoFS predicted winds, and a stacked plot of the two at 1-min intervals, respectively. Hodographs are shown for 2230, 2240, 2250, 2300, 2310, and 2320 UTC, with times increasing from lighter to darker shades. In all panels, wind at 0–500 m AGL are represented with green, winds at 500–1000 m AGL are represented with blue, and winds at 1000–1500 m AGL are represented with magenta. The green, blue, and magenta studs in (a)–(c) highlight wind speed at 0, 500, and 1000 m, respectively. (d) Observed u and v wind components and speed. (e) The error in the u and v wind components and speed as a function of time (m s^{-1}). (f) 0–2-km updraft helicity (m^2 s^{-2}) and observed V_{rot} of the mesocyclone (m s^{-1}). Shading indicates one standard deviation among the WoFS members.

time period TORUS observations were made (Fig. 8). Out of 11 radiosondes available in this period, 5 were made during WoFS forecasts and are used for analysis, 3 in the far field and 2 in the near field (Fig. 9). Signal was lost from all five radiosondes as they were ingested into the target supercell at ~450 hPa, which leaves no CAPE values to compare to WoFS. The lidar was deployed twice, observing both the near and far field from 2030 to 2135 UTC, then for a shorter deployment in the near field while the tornado near Mangum was ongoing.

As for the 17 May case, the WoFS temperature errors are less than 2 K in magnitude between the surface and 700 hPa (Fig. 10). However, the WoFS forecasts near the ground are too warm at 2035 UTC in most members, with errors up to 2 K in some of the YSU and MYJ members (Figs. 10b,d,g,i). A possible explanation is a lack of anvil shading in the WoFS forecasts (evident in WoFS simulated infrared satellite) compared to observations. The 2052, 2050 UTC pair also shows a positive temperature error near the ground (Figs. 10g,i), but it extends up to 700 hPa at this location, which is peculiar considering this radiosonde is 68 km from the supercell (Fig. 9) and thus likely experiences minimal impacts from the storm. Given the substantial depth to the positive temperature error, it is possible that the temperature errors were inherited from the background conditions provided by HRRRE. Given the presence of clouds from 900 to 700 hPa, reflecting the moist environment (Figs. 10a,f), a lack of clouds in some HRRRE members (seen indirectly by many unsaturated WoFS member profiles in this layer) may have resulted in larger insolation than reality.

All radiosondes ascended above 3 km so observed SRH up to this level can be compared to WoFS forecasts of 0–1-km SRH and 0–3-km SRH (Fig. 6b). For the observation farthest from the storm (2034 UTC), the mean of the WoFS forecasts for both 01SRH and 03SRH are close to the observed values with mean errors generally less than 40 m^2 s^{-2}, and each of the five soundings lies within the distribution of WoFS predicted SRH values. The YSU members forecast SRH values that are consistently lower and with smaller spread than the values forecasted by the other two PBL schemes. MYNN mean errors are the smallest with a relatively large spread. A consistent, negative SRH bias emerges for the YSU and MYJ schemes for the four other soundings. These errors appear to have no dependence on distance from or orientation with

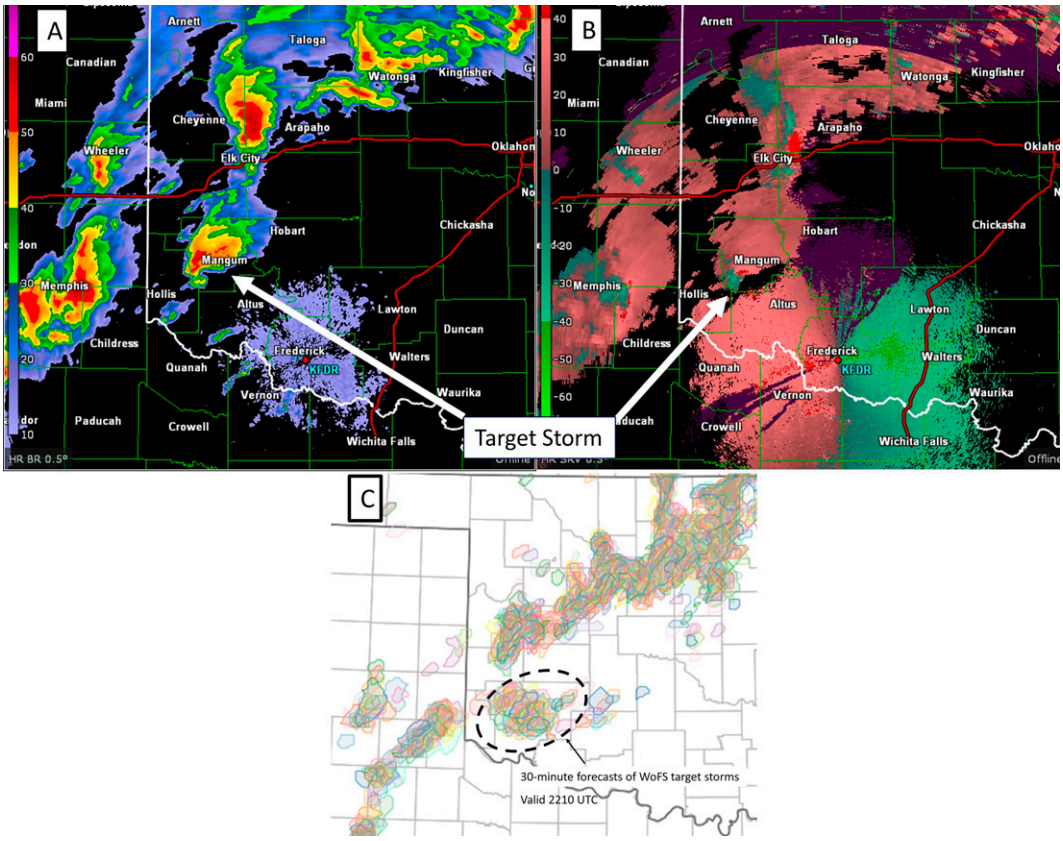
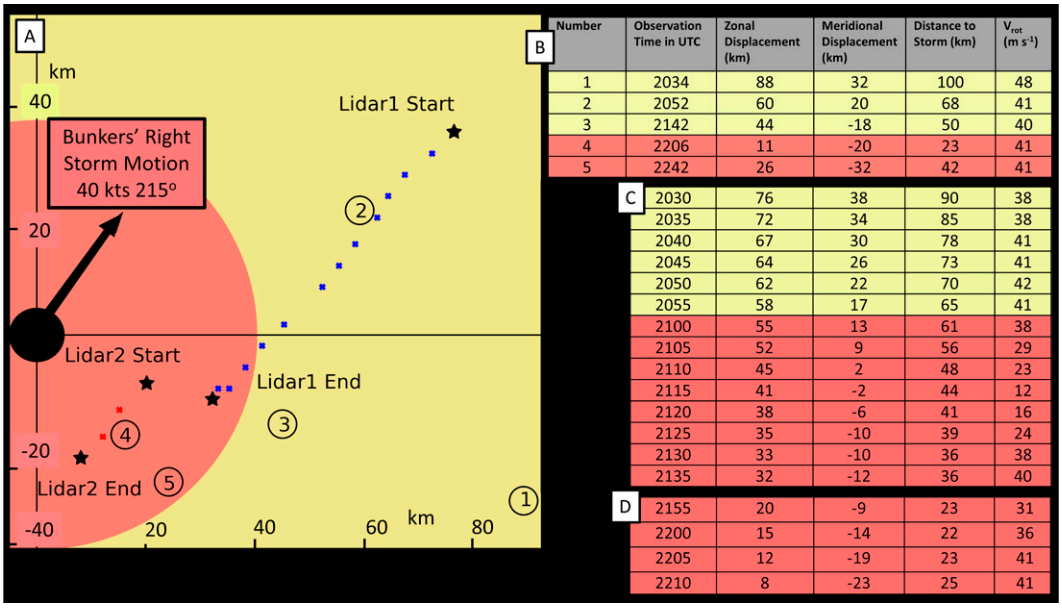


FIG. 8. Radar images from KFDR of the 20 May 2019 Mangum, OK, supercell at 2210 UTC, the time of the Mangum tornado report. (a) Reflectivity (dBZ) and (b) storm-relative velocity (kt). (c) The dashed oval encompasses contours of 45-dBZ simulated composite reflectivity from 30-min forecasts from all 18 WoFS members valid at 2210 UTC and initialized at 2130 UTC. Different color shading depicts different ensemble members.

respect to the target supercell. As in Cohen et al. (2017), YSU members predict smaller SRH values than the local schemes, and MYJ consistently forecasts smaller SRH values than MYNN. MYNN displays a small underforecast bias for 01SRH, but is actually a little biased toward overforecasting 03SRH values. Again, the variability in SRH is largest for the MYNN forecasts for these other four soundings.

During the first of two lidar deployments, the winds were observed to remain nearly constant with time (Fig. 11) with only a small backing and strengthening in the v component of 3–5 m s^{-1} noted above 500 m AGL after ~2110 UTC. The lidar was positioned mostly in the far field with the lidar still 35 km away from the storm at the end of the deployment, perhaps explaining the lack of a stronger modification to the winds from the mesocyclone. Also, the deployments occur during the late afternoon, so any nocturnal cooling related modifications of the wind field are not expected. An additional potential contributor to the relatively stable wind profile in this period is a temporary decrease in observed V_{rot} (Fig. 11f) between 2100 and 2130 UTC as a storm merger occurred, which suggests smaller pressure perturbations within the mesocyclone and limited modification of the near-storm environment. As an example of how pressure perturbations modify the near-storm environment, a uniform pressure gradient of 1 hPa over 10 km acting on a friction-less solid-body requires about 10 min to result in a wind speed change of 5 m s^{-1}. This speed increase would be lower in magnitude and take longer if accounting for friction effects (turbulence), but the calculation gives an order-of-magnitude estimate to the rapid response time of supercell environments to changes to the pressure field within the storm. Although the winds in the WoFS forecasts were weaker than observed toward the end of the deployment (see u and v component errors after 75 min in Fig. 11e), the wind profiles evolved similarly to the observations, with winds backing with time (a little more so than in observations near the ground).

For the near-field deployment (about 20–25 km from the updraft; Fig. 12), the WoFS forecast also evolved similarly to the observations; the observed inflow wind speeds appeared to peak from 2155 to 2205 UTC before weakening as the storm moved farther away from the lidar (Fig. 12). The evolution of the WoFS hodographs are similar (Fig. 12b) with weakening of the winds (and a small veering of the near-surface winds) with time. However, the predicted wind speeds are too weak overall in the WoFS forecasts, by up to 6–8 m s^{-1}. The wind speed errors in WoFS were mostly



B

Number	Observation Time in UTC	Zonal Displacement (km)	Meridional Displacement (km)	Distance to Storm (km)	V_{rot} (m s^{-1})
1	2034	88	32	100	48
2	2052	60	20	68	41
3	2142	44	-18	50	40
4	2206	11	-20	23	41
5	2242	26	-32	42	41

C

2030	76	38	90	38
2035	72	34	85	38
2040	67	30	78	41
2045	64	26	73	41
2050	62	22	70	42
2055	58	17	65	41
2100	55	13	61	38
2105	52	9	56	29
2110	45	2	48	23
2115	41	-2	44	12
2120	38	-6	41	16
2125	35	-10	39	24
2130	33	-10	36	38
2135	32	-12	36	40

D

2155	20	-9	23	31
2200	15	-14	22	36
2205	12	-19	23	41
2210	8	-23	25	41

FIG. 9. As in Fig. 3, but for 20 May 2019.

driven by errors in the v component (Fig. 12e). This indicates that the WoFS forecasts again evolved similarly to the observations near the storm, but with a lower amplitude relative to the observed winds. In this case, the forecasts were unable to depict the observed strong low-level inflow of 25–30 m s^{-1} at 500–1000 m AGL into the ongoing tornadic supercell from within 40 km from the storm. These lower wind speed values may indicate weaker predicted mesocyclones and associated storm-modification than in the observations.

c. 23 May 2019 Pampa, Texas, supercells

TORUS targeted multiple supercells on 23 May, all of which were tornadic at some point in their life cycles. The first of the storms initiated close to Amarillo, Texas, over a southwest–northeast outflow boundary from morning convection. This supercell was nontornadic during the lidar observation period of 2226–2308 UTC and tracked northeast through the Texas Panhandle. A second supercell initiated off the outflow of the first a few miles south of Amarillo and moved to the northeast away from the boundary. This supercell produced a tornado from 0000 to 0030 UTC during the second lidar deployment from 2330 to 0030 UTC (Fig. 13). A third supercell tracked along a path just south and east of the first two supercells and approached the lidar location at 0120 UTC. This later deployment was not used in this analysis owing to an outflow boundary moving over the lidar position at 0010 UTC. All three supercells were observed by the Amarillo, Texas, WSR-88D (KAMA; Fig. 14).

The environment of the first two storms was sampled by three radiosondes for each storm. Of the four radiosondes that made it to the EL, WoFS substantially over forecasts both MLCAPE and SBCAPE at all observation times (Fig. 5c). The errors are consistently smallest for the YSU members and consistently largest for the MYJ members. These CAPE errors appear to be related mostly to low-level temperature errors that are as large as 4 K in magnitude at 2210 UTC (Fig. 5c). The MLCAPE and SBCAPE (and low-level temperature) errors are smallest for the last radiosonde observation taken at 0104 UTC, which was by far the farthest from the target storms (Fig. 13). The lack of saturation near the top of the PBL in the WoFS forecasts compared to the near-saturated conditions in the observations (with some contribution from an under forecast of anvil shading in the model forecasts, particularly at 2210 UTC) again suggests that low-level cloudiness was under predicted and led to the CAPE errors (Fig. 15).

Above the PBL in the 2210 UTC observation (800–600 hPa) and in the 2340 UTC observation (800–700 hPa), however, the later WoFS forecasts are notably warmer and drier than observed. The observation shows near saturation in these layers and is suggestive of mesoscale lifting in the storm inflow, similar to that seen in the near field radiosonde observations of Wade et al. (2018) (although a contribution from precipitation evaporating into the inflow from above cannot be ruled out). This possible lifting also is seen in the 2252 UTC observation (another near field sounding; not shown); however, the WoFS forecasts appear to represent this cooled/moistened environment more accurately, perhaps owing to additional assimilation cycles resulting in more accurate analyses of the storms and their influence on the near-storm environment.

WoFS more accurately forecasts both 01SRH and 03SRH on 23 May than on 17 and 20 May. All six WoFS ensemble mean forecasts have mean errors less than 100 m^2 s^{-2} and commonly less than[6] 50 m^2 s^{-2} (Fig. 6c). However, the tendency of an under forecast of SRH seen in the other two cases

[6] These values are typically 10%–15% of observed values and are considered relatively small given the strong (>200 m^2 s^{-2}) 0–1- and 0–3-km SRH in the environment.

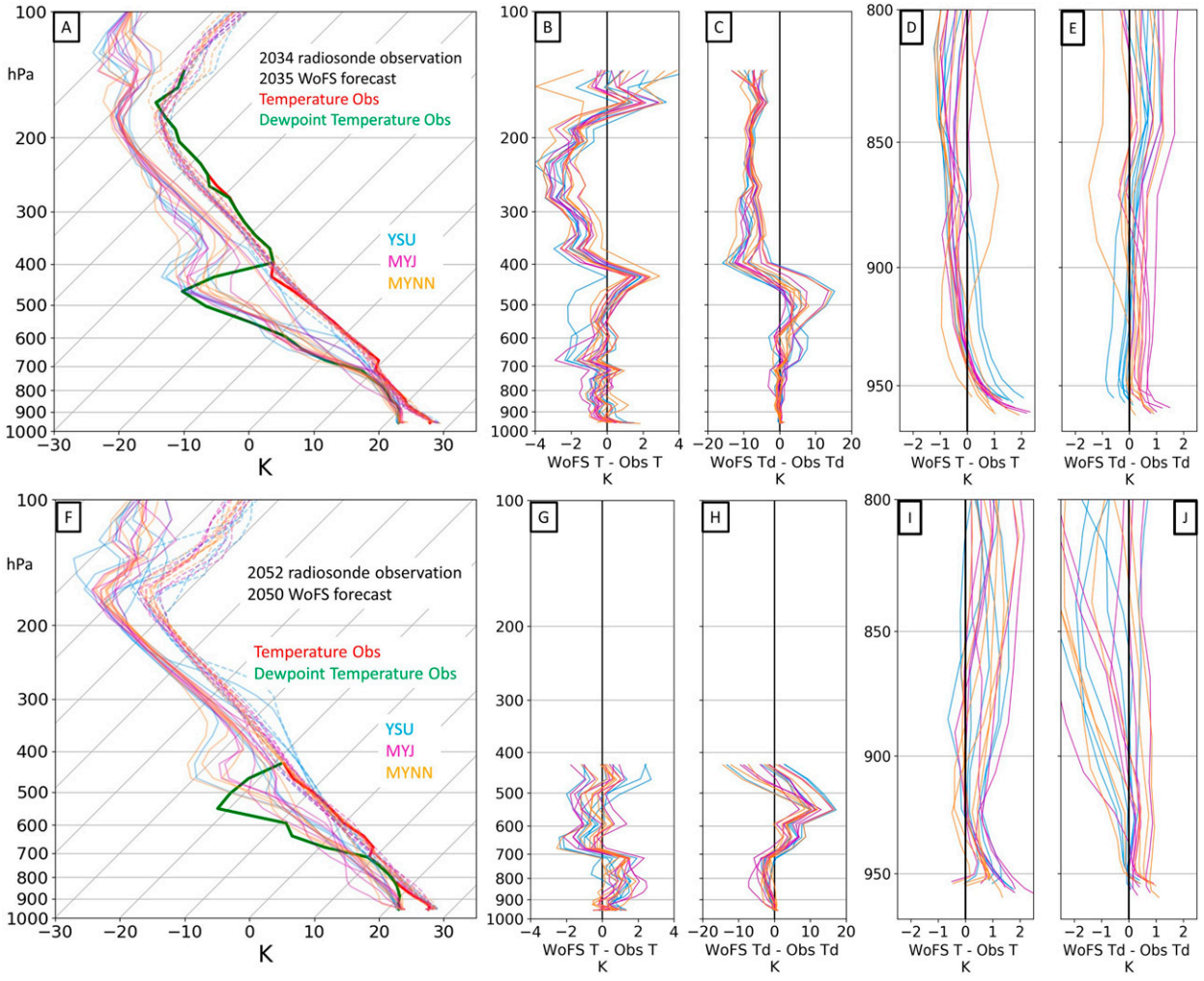


FIG. 10. As in Fig. 4, but for the (a)–(e) 2034 UTC radiosonde launch compared to the 2035 UTC WoFS forecast and (f)–(j) 2052 UTC radiosonde launch compared to the 2050 UTC WoFS forecast on 20 May 2019.

also is seen for this case. Peculiarly, WoFS tends to under predict SRH in the far field more so than in the near field (see the 0000 and 0105 UTC columns on Fig. 6c). The reason why WoFS is under predicting SRH in the far field is not clearly understood, but a possible explanation for this finding, noting that the far field observations were made later in the evening and the observed increase in SRH with time, is that the forecasted EET timing is off. Another possibility could be that an accurate WoFS prediction of storm modification is distorted by initial errors in the mesoscale background forecast provided by the HRRRE. Among all the soundings, the YSU scheme commonly under forecasts mean SRH (particularly in the 0–3-km layer) more so than the other schemes, which was also found in the first two cases (Fig. 6).

The first lidar deployment lasted 35 min and observed the near-field environment to the east-northeast of the second of the three supercells (Fig. 13). The lidar was only able to sample winds up to 500 m AGL because of low clouds, but observed a backing and strengthening in the winds below that level (Fig. 16a), as would be expected with the decreasing zonal distance to the strengthening storm and increasing V_{rot} (Fig. 16f). WoFS accurately predicts a small amount of backing; however, as in the 20 May case, the increase in low-level wind speeds at the time the lidar is closest to the storm [a 21 m s^{-1} south-southeast wind at 500 m AGL (Fig. 16a)] is somewhat under predicted by WoFS. This can be seen by the errors from -2 to -3 m s^{-1} in u and v at 2305 UTC (Fig. 16e). The V_{rot} of the storm increases from approximately 2230 to 2245 UTC, while the UH of WoFS storms is decreasing, implying that the observed storm is strengthening relative to the WoFS storms. This could potentially explain some of the under prediction in the 500 m AGL inflow wind speed at this time.

The second lidar deployment lasted from 2330 to 0030 UTC and sampled the near field environment of the second tornadic supercell (third overall supercell) in the line (Fig. 17). This deployment begins to the northeast and ends almost due south of the supercell (Fig. 13). This is a particularly interesting

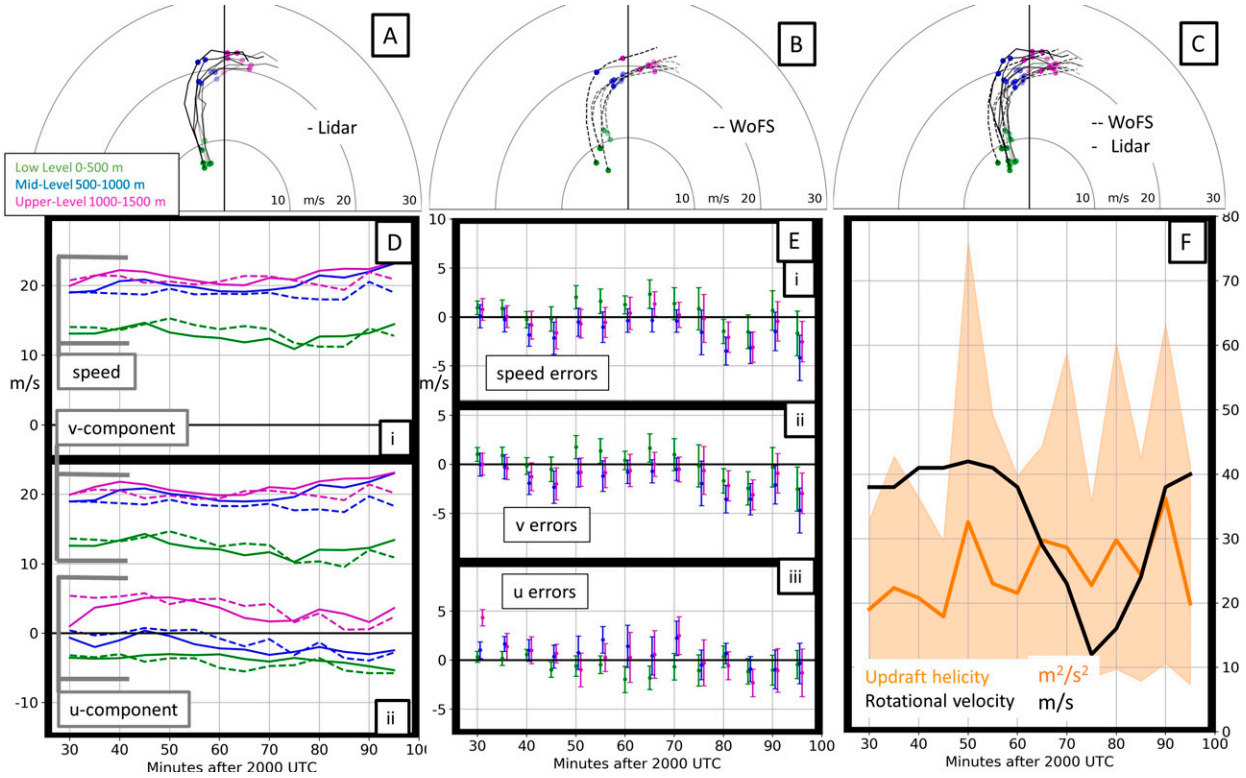


FIG. 11. As in Fig. 7, but on 20 May 2019. Times shown in (a)–(c) are 2030, 2040, 2050, 2100, 2110, 2120, and 2130 UTC.

deployment as the winds backed and strengthened early on, particularly above 250 m AGL, then veered as the storm moved north of the lidar after 0000 UTC, causing a strong meridional increase in the winds at 0005 UTC (Figs. 17a,d). WoFS initially had small errors in the wind, but forecasts the opposite trend to what is seen in the lidar observations, with a veering and weakening of the winds during the first 30 min of the deployment (Fig. 17a). This leads to some of the largest wind errors among the three case studies with 5–7 m s^{-1} errors common in both the u and v components toward 0000 UTC (Fig. 17e). After this time, WoFS then backs and strengthens the low-level winds (resulting in a sudden decrease in the v component errors after 0005 UTC), but the winds are still much weaker in WoFS than in observations. The time-dependent changes before 0005 UTC suggest that the storms in WoFS may have been weaker, leading to a lack of a backing signal in the modeled winds and under forecasting the increase in wind speed after 0005 UTC. WoFS under predicting the strength of the wind field relative to observations could be a result of a weaker storm (as suggested by the decreasing UH values over this time period; Fig. 17f), but some of the under prediction is likely attributable to an inability of WoFS to fully resolve the perturbation pressure field in mesocyclones with 3-km horizontal grid spacing. Finer scale pressure perturbations may be concentrated regions of dramatically lower pressure (e.g., tornadoes) that contribute substantially to storm environment modification (Wade et al. 2018).

4. Summary and discussion

With all predictive models, an evaluation of the predictions relative to the observed events must be done to assess the quality of the system. To this end, the overlap between WoFS forecasts and TORUS 2019 operations allowed for a rare opportunity to evaluate the quality of WoFS predictions in hazardous severe weather settings, with specific emphasis on the feedback between supercells and their environments. This paper is the first to use Doppler wind lidar observations to capture this dynamic process with high temporal resolution time series of vertical profiles of wind that are used for a time-dependent evaluation of WoFS environmental predictions. In addition, mobile radiosonde observations provide thermodynamic information that aids in assessing the impact on WoFS forecasts from each of the PBL schemes.

Due to the limited sample size available, the observed trends and potential error sources discussed in this paper cannot be generalized, yet still provide unique information about WoFS forecast quality. The radiosonde observations collectively show that WoFS tended to over forecast CAPE in both the near and far field inflow. It seems there may be an inverse relationship between WoFS over prediction of CAPE and distance from the storm (i.e., WoFS errors in CAPE becomes larger closer to the storm). The CAPE errors were most often attributable to WoFS being too warm near the ground close to the storms (e.g., Fig. 4), although small moisture biases also contributed. The over prediction of WoFS surface temperature and dewpoint likely relates to WoFS under predictions of

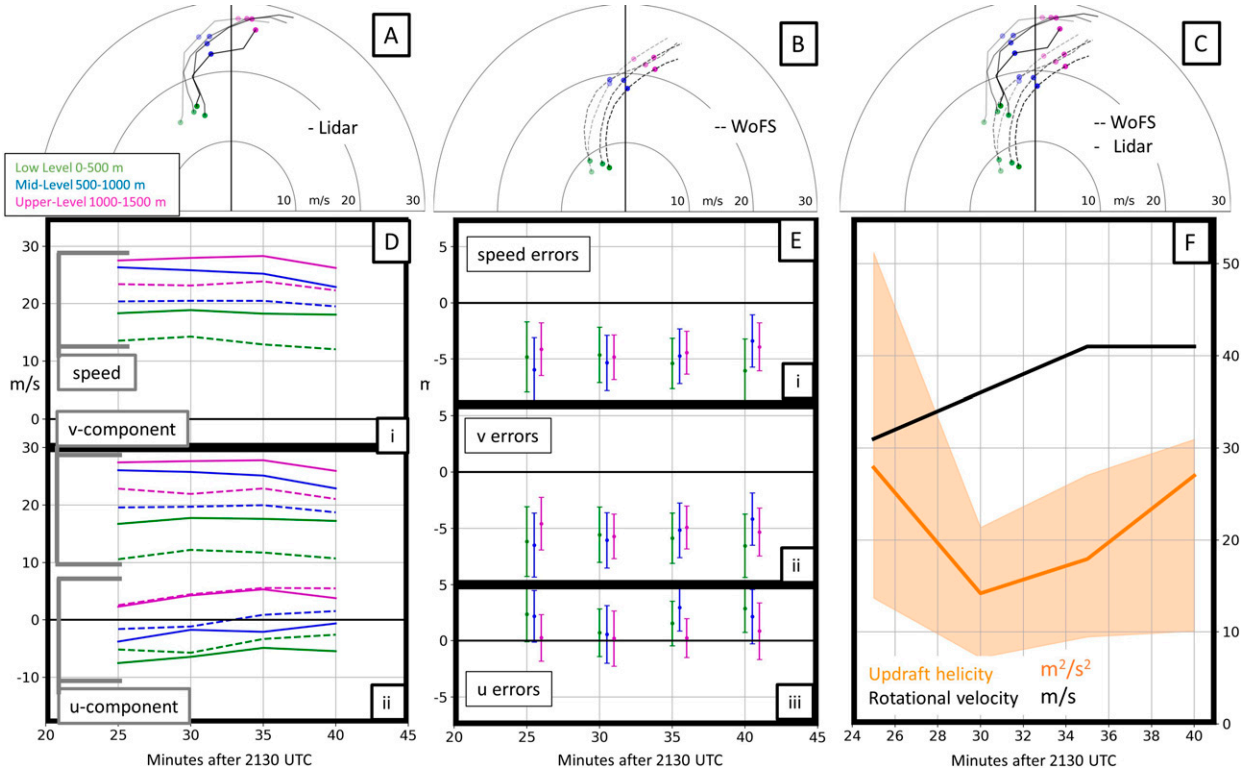
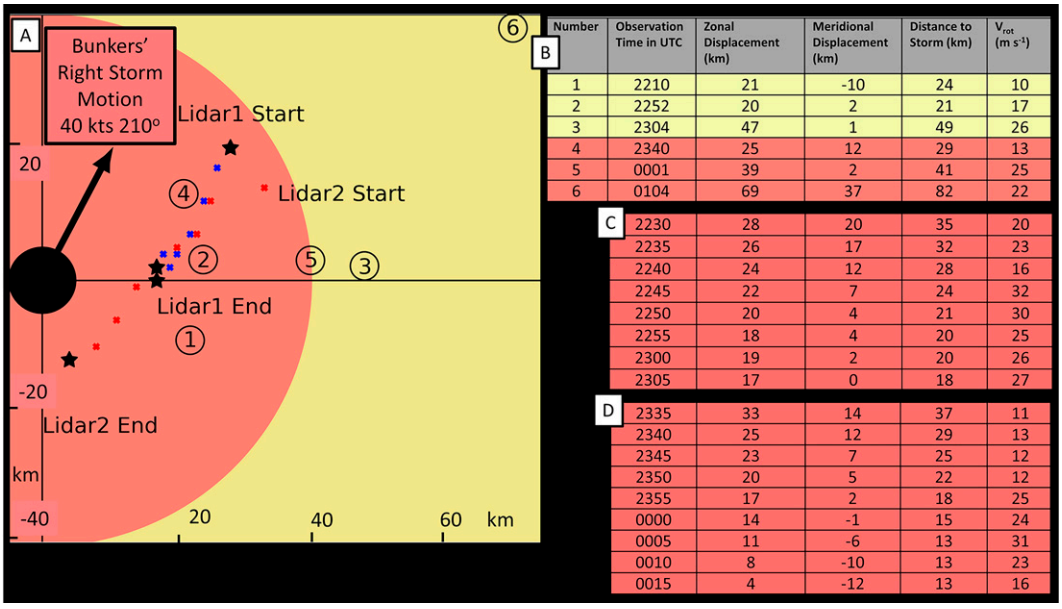


FIG. 12. As in Fig. 7, but on 20 May 2019. Times shown in (a)–(c) are 2155, 2200, 2205, and 2210 UTC.

anvil-level and low-level cloud cover, leading to larger insolation and evapotranspiration than reality. The YSU scheme predicted lower CAPE and SRH than MYJ and MYNN, supporting the work of Cohen et al. (2015, 2017) and Potvin et al. (2020) where they found the nonlocal schemes often under predicted 03SRH and MLCAPE. The under prediction of SRH and CAPE in the YSU scheme is thought to be a consequence of more aggressive mixing in the nonlocal versus local



Number	Observation Time in UTC	Zonal Displacement (km)	Meridional Displacement (km)	Distance to Storm (km)	V_{rot} ($m\,s^{-1}$)
1	2210	21	-10	24	10
2	2252	20	2	21	17
3	2304	47	1	49	26
4	2340	25	12	29	13
5	0001	39	2	41	25
6	0104	69	37	82	22
C	2230	28	20	35	20
	2235	26	17	32	23
	2240	24	12	28	16
	2245	22	7	24	32
	2250	20	4	21	30
	2255	18	4	20	25
	2300	19	2	20	26
	2305	17	0	18	27
D	2335	33	14	37	11
	2340	25	12	29	13
	2345	23	7	25	12
	2350	20	5	22	12
	2355	17	2	18	25
	0000	14	-1	15	24
	0005	11	-6	13	31
	0010	8	-10	13	23
	0015	4	-12	13	16

FIG. 13. As in Fig. 3, but for 23 May 2019. The blue dots correspond to the first lidar deployment, which observed the first supercell, and the red dots correspond to the second lidar deployment, which observed the second supercell.

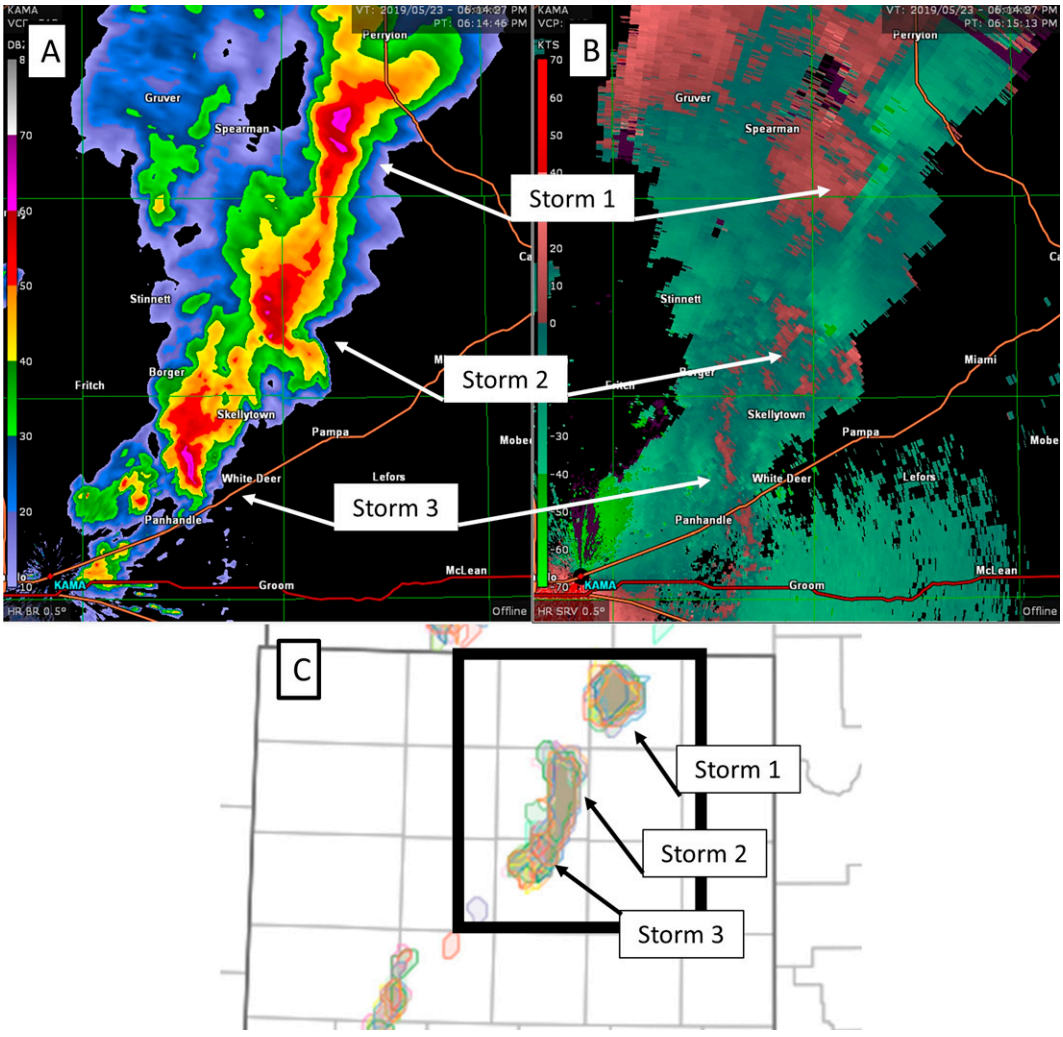


FIG. 14. Radar images from KAMA of the 23 May Pampa, TX, supercell at 2315 UTC, the time of the Pampa tornado report. (a) Reflectivity (dBZ) and (b) storm-relative velocity (kt). (c) 45-dBZ simulated composite reflectivity from 30-min forecasts from all 18 WoFS members initialized at 2300 UTC, valid at 2315 UTC 23 May. Different color shading depicts different ensemble members. Black contour outlines the same area as in (a) and (b).

schemes. Although YSU consistently predicts the warmest and driest conditions, the error of YSU predictions relative to the other members was case dependent.

Several radiosonde observations provide examples of how the PBL physics diversity in WoFS provides increased ensemble spread (Figs. 4 and 10a–e). The local schemes predict cooler, more humid PBLs with negative and positive errors, respectively, while the YSU scheme forecasts a warmer, drier PBL with positive and negative errors, respectively. The radiosonde observations fall within the ensemble spread more often than with any single scheme, leading to ensemble mean errors smaller than those from the individual members. This increased tendency to predict the possibility of observations falling in the spread of YSU, MYJ, and MYNN forecasts often extends above the PBL; YSU tends to entrain relatively warm, dry air down into the PBL more than MYJ and MYNN does, which results in cooler temperatures above the PBL; the WoFS ensemble mean errors also tend to be smaller than the individual members in this layer.

Although WoFS often accurately forecasts SRH with mean error magnitudes less than 50 $m^2 s^{-2}$ (generally less than 15% of ensemble mean SRH), there were several instances, both from the radiosonde and lidar observations, when the predicted storm-relative winds were weaker and less backed than observations in the lowest 1 km. These SRH errors tended to be largest later in the deployments, suggesting a relationship between SRH errors and supercell duration. Given that two of the deployments lasted until after 0100 UTC, it is also possible that some under prediction of the effects of the frictional decoupling of the surface layer with the PBL during the EET (e.g., Wingo and Knupp 2015) contribute to the lack of backing of low-level winds in WoFS forecasts (e.g., Fig. 7). However, all of the near field conditions were observed before local sunset before the majority of horizontal wind speed increases with the

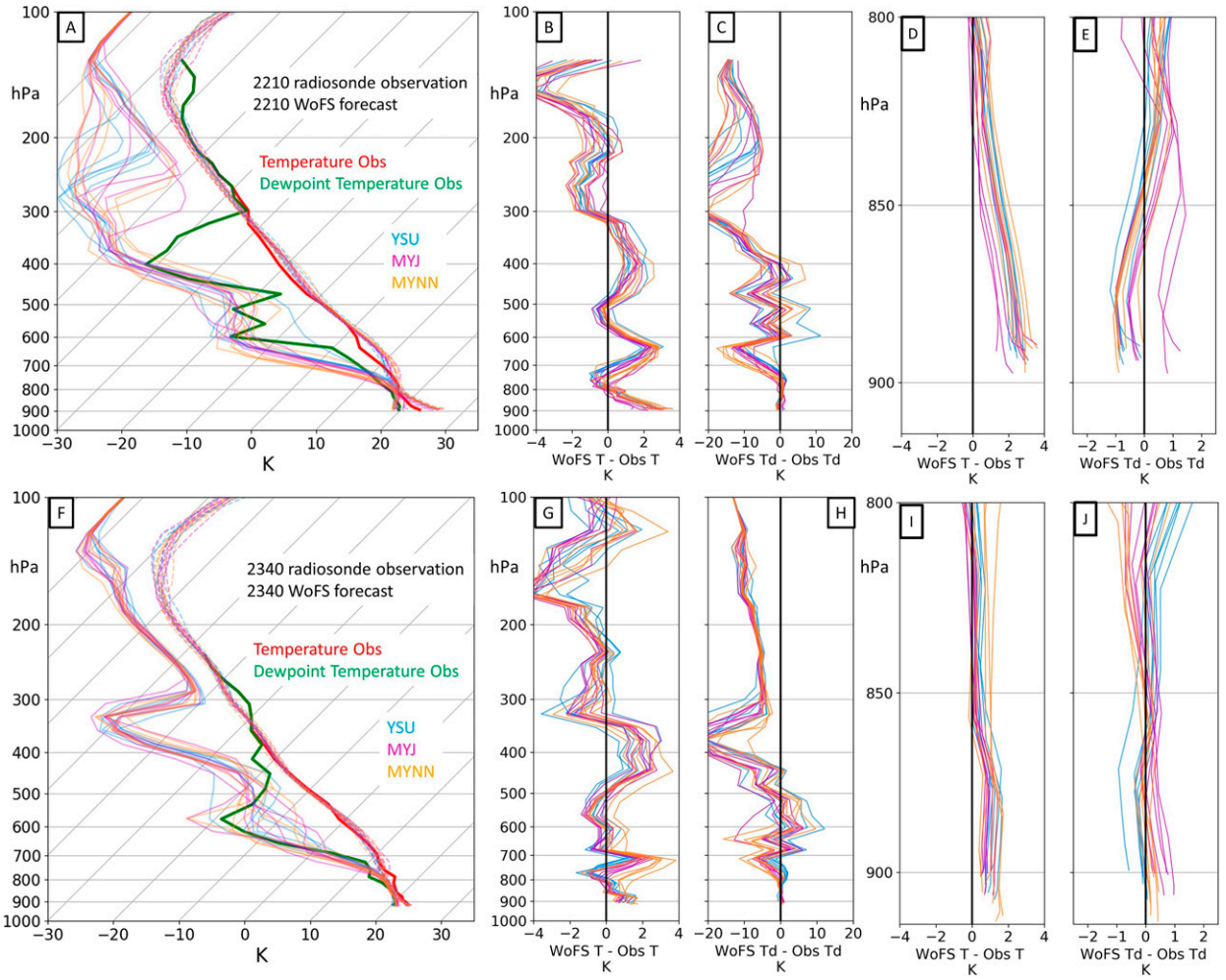


FIG. 15. As in Fig. 4, but for the (a)–(e) 2210 UTC radiosonde launch compared to the 2210 UTC WoFS forecast and (f)–(j) 2340 UTC radiosonde launch compared to the 2340 UTC WoFS forecast on 23 May 2019.

EET tend to occur (Mahrt 1981; Smith et al. 2019). Also, the 20 May case displayed clear signs of a low-level wind response to a low pressure perturbation with the storm as early as 2100 UTC, so it is likely that storm-induced flow modifications are present to some extent in all of the near-field observations. The presented analysis of the five lidar deployments across the three cases studies collectively illustrate, and add to the presented evidence from the radiosonde observations, that the wind field in WoFS appears to be modified by the modeled storms, mostly in the sense that the wind field is strengthened in the direction of the supercell due to the low pressure perturbation induced by the rotating updraft. However, the extent of the response, seen in both the changes in wind direction (e.g., backing) and speed, is muted compared to what was observed. Possible explanations for this muted response could be a result of errors in the prediction of the size and/or intensity of the mesocyclones or perhaps insufficient grid spacing to reproduce the observed storm modification.

It was anticipated that SRH errors related to storm modification would be more strongly dependent on distance from the storm, but the first radiosonde observations on 17 and 23 May were near field (both approximately 25 km from the storm), yet showed the smallest errors for each day. This suggests WoFS may be under forecasting storm modification of the environment, not solely in a spatial sense, but in a temporal sense as storm modifications of the environment need time to develop fully. This would suggest the differences seen in Wade et al. (2018) could be related to the time needed for pressure gradient accelerations to increase wind velocities to make an appreciable difference in SRH, given a supercell that is changing strength at the time of the observation. This trend is seen clearly on 17 May in Fig. 6 where WoFS predicts increased SRH later in the evening, but predictions remain lower than the observations.

5. Future research

Extending this analysis to include more case studies from TORUS 2019 would be informative. Exploring those data for trends like the ones found in these case studies could help determine if the behavior identified in this work is representative of WoFS more broadly and in a quantitative sense. Furthermore,

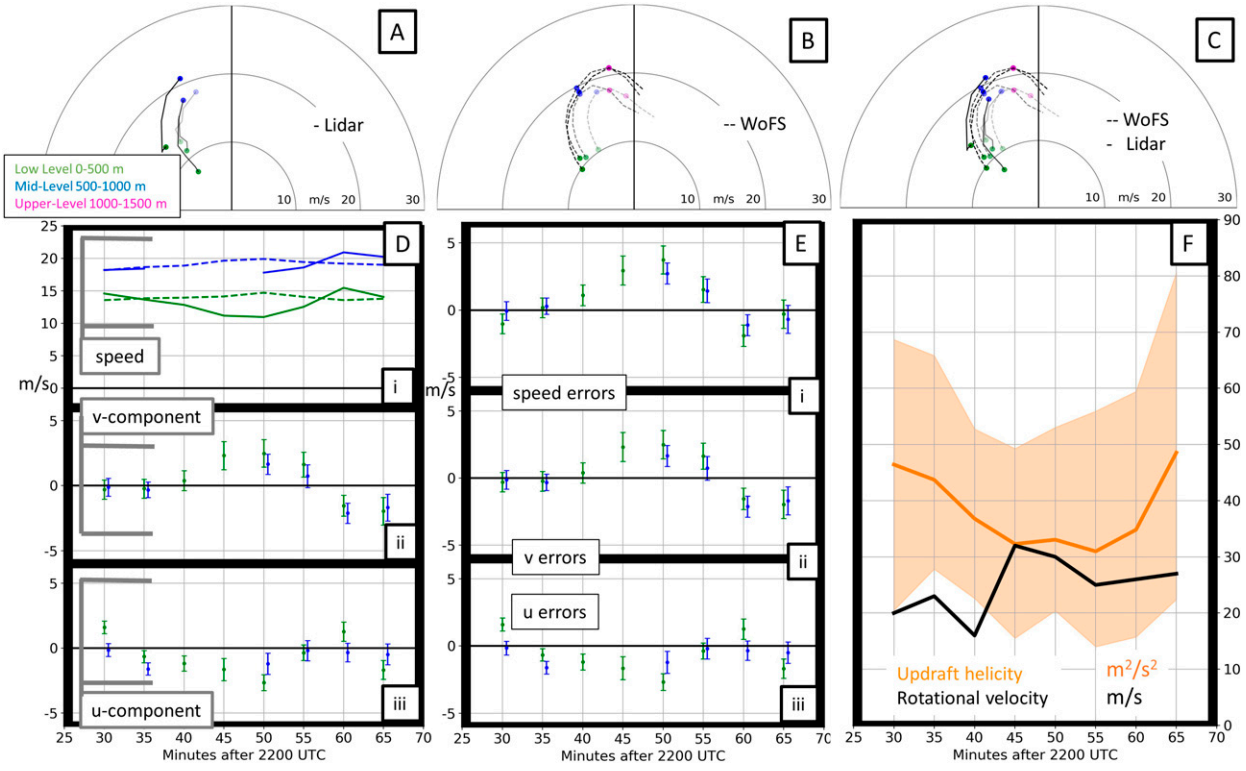


FIG. 16. As in Fig. 7, but on 23 May 2019. Times shown in (a)–(c) are 2230, 2240, 2250, and 2300 UTC.

Cohen et al. (2017) highlighted the need for regionally dependent studies as models may behave differently based on typical environmental setups for severe convective weather (e.g., higher SRH and lower CAPE in the Southeast United States). This study analyzed WoFS environmental forecasts in late afternoon-to-early evening supercells during the springtime in the U.S. Great Plains. Exploring WoFS environmental forecasts near either tornadic or nontornadic supercells in other parts of the country at different times of the day or year would provide WoFS developers with a better understanding of the biases inherent in WoFS forecasts among a variety of severe weather regimes (e.g., squall lines). With enough observations, the analysis could be split based on if the supercells are tornadic or not to investigate if WoFS predicted-environments—and their modifications—may have important implications on the WoFS ability to discern between tornadic and nontornadic storms similar to the near-field composites presented in Coniglio and Parker (2020).

The largest errors in the predicted near-storm environment in this study occurred in regions where WoFS incorrectly predicted storm-scale features such as cloud cover or intensity trends. It is likely that improved storm-scale forecasts in WoFS will in turn improve representation of storm-modifications to the surrounding environment. In particular, more accurate prediction of cloud cover near predicted storms offers the potential to improve representation of the near-storm thermodynamics (Jones et al. 2018).

The importance of near-storm modification to the ambient environment on subsequent supercell evolution (Markowski and Richardson 2010; Flournoy et al. 2020) suggests that WoFS storm-scale forecasts could be improved with a better representation of the near-storm environment in the ensemble. As WoFS underpredicts the intensity or evolution of storm modification to the inflow wind field in each of the three cases examined, numerical experiments on WoFS configuration could identify the source of these errors. Furthermore, it was found that WRF-ARW simulations produced stronger near-storm modifications to the wind field compared to the FV3 core (Potvin et al. 2019), suggesting that considerations of storm modification of the surrounding environment will be important for successful transition to the Unified Forecasting System.

Acknowledgments. The first author was provided support for this work by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma (OU) Cooperative Agreement NA16OAR4320115, U.S. Department of Commerce. Radiosonde data were collected in the field with support from NOAA discretionary funding and National Science Foundation Award AGS-1824811. WoFS data were generated and provided by the WoFS team at CIWRO and NSSL. All analyses and figures were created using the open-source Anaconda Python distribution and SciPy, Matplotlib, basemap, netcdf4, and sharppy libraries. Radar images were generated with GRLevel3. A great number of scientists and engineers contributed value to this project. We appreciate the efforts of TORUS 2019 field scientists that operated in

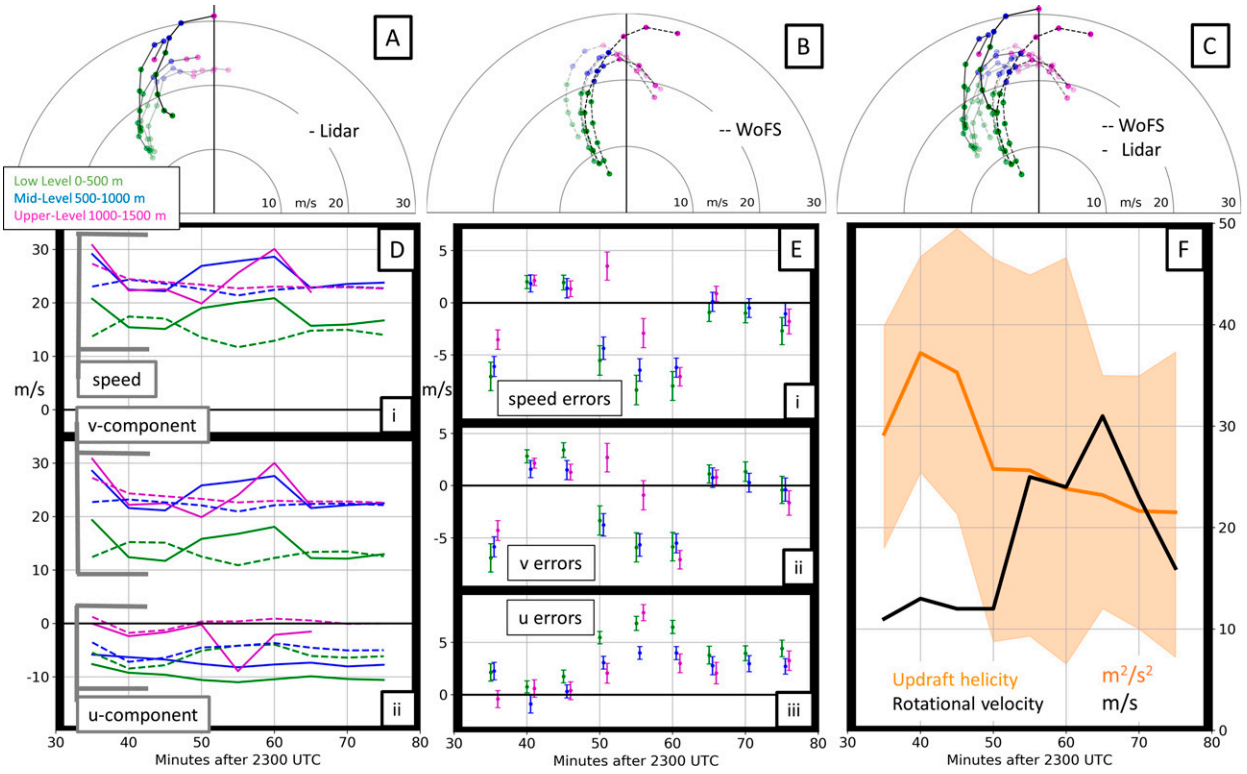


FIG. 17. As in Fig. 7, but on 23 May 2019. Times shown in (a)–(c) are 2335, 2345, 2355, 0005, and 0015 UTC.

dangerous environments to collect observational data, with special thanks to Sean Waugh for his efforts in maintaining much of the observational equipment. We thank the University of Oklahoma, the University of Colorado, the University of Nebraska, and Texas Tech University for their part in the TORUS 2019 field mission. We thank those at NSSL for their review of this work. Kelsey Britt provided the storm-object selection code, and Noah Brauer provided the sounding plotting code. We also thank the reviewers for their time and effort ensuring the quality and consistency of this work.

Data availability statement. The lidar and radiosonde data used during this study are openly available from the NCAR Earth Observing Laboratory Data Archive at https://data.eol.ucar.edu/. These data were reanalyzed by the second author. All of the data analyzed for this study and further information about data processing is available upon request from the National Severe Storms Laboratory through the second author (Michael.Coniglio@noaa.gov).

REFERENCES

Argrow, B., and A. Houston, 2007: UAS for in situ sensing of an atmospheric airmass boundary. AIAA 2007 Conf., Rohnert Park, CA, AIAA, 8 pp., https://citeseerx.ist.psu.edu/viewdoc/citations?doi=10.1.1.577.94.

Britt, K. C., P. S. Skinner, P. L. Heinselman, and K. H. Knopfmeier, 2020: Effects of horizontal grid spacing and inflow environment on forecasts of cyclic mesocylogenesis in NSSL's Warn-on-Forecast System (WoFS). *Wea. Forecasting*, **35**, 2423–2444, https://doi.org/10.1175/WAF-D-20-0094.1.

Burgess, D. W., E. R. Mansell, C. M. Schwartz, and B. J. Allen, 2010: Tornado and tornadogenesis events seen by NOXP X-Band, dual-polarization radar during VORTEX2 2010. *25th Conf. on Severe Local Storms*, Denver, CO, Amer. Meteor. Soc., 5.2, https://ams.confex.com/ams/25SLS/webprogram/Paper176164.html.

Coffer, B. E., and M. D. Parker, 2015: Impacts of increasing low-level shear on supercells during the early evening transition. *Mon. Wea. Rev.*, **143**, 1945–1969, https://doi.org/10.1175/MWR-D-14-00328.1.

Cohen, A. E., S. M. Cavallo, M. C. Coniglio, and H. E. Brooks, 2015: A review of planetary boundary layer parameterization schemes and their sensitivity in simulating southeastern U.S. cold season severe weather environments. *Wea. Forecasting*, **30**, 591–612, https://doi.org/10.1175/WAF-D-14-00105.1.

——, ——, ——, ——, and I. L. Jirak, 2017: Evaluation of multiple planetary boundary layer parameterization schemes in southeast U.S. cold season severe thunderstorm environments. *Wea. Forecasting*, **32**, 1857–1884, https://doi.org/10.1175/WAF-D-16-0193.1.

Collins, J., and P. Flaherty, 2014: The NOAA hurricane hunters: A historical and mission perspective. *Fla. Geogr.*, **45**, 14–27.

Coniglio, M. C., 2012: Verification of RUC 0–1-h forecasts and SPC mesoscale analyses using VORTEX2 soundings. *Wea. Forecasting*, **27**, 667–683, https://doi.org/10.1175/WAF-D-11-00096.1.

——, and M. D. Parker, 2020: Insights into supercells and their environments from three decades of targeted radiosonde

observations. *Mon. Wea. Rev.*, **148**, 4893–4915, https://doi.org/10.1175/MWR-D-20-0105.1.

——, J. Correia, P. T. Marsh, and F. Kong, 2013: Verification of convection-allowing WRF Model forecasts of the planetary boundary layer using sounding observations. *Wea. Forecasting*, **28**, 842–862, https://doi.org/doi:10.1175/waf-d-12-00103.1.

——, G. S. Romine, D. D. Turner, and R. D. Torn, 2019: Impacts of targeted AERI and Doppler lidar wind retrievals on short-term forecasts of the initiation and early evolution of thunderstorms. *Mon. Wea. Rev.*, **147**, 1149–1170, https://doi.org/10.1175/MWR-D-18-0351.1.

Davies-Jones, R., 2002: Linear and nonlinear propagation of supercell storms. *J. Atmos.Sci.*, **59**, 3178–3205, https://doi.org/10.1175/1520-0469(2003)0592.0.CO;2.

Dawson, D. T., 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.

Duncan, J. B., B. D. Hirth, and J. L. Schroeder, 2019: Doppler radar measurements of spatial turbulence intensity in the atmospheric boundary layer. *J. Appl. Meteor. Climatol.*, **58**, 1535–1555, https://doi.org/10.1175/JAMC-D-18-0151.1.

Evans, C., S. J. Weiss, I. L. Jirak, A. R. Dean, and D. S. Nevius, 2018: An evaluation of paired regional/convection-allowing forecast vertical thermodynamic profiles in warm-season, thunderstorm-supporting environments. *Wea. Forecasting*, **33**, 1547–1566, https://doi.org/10.1175/waf-d-18-0124.1.

Flora, M. L., P. S. Skinner, C. K. Skinner, A. E. Reinhart, T. A. Jones, N. Yussouf, and K. H. Knopfmeier, 2019: Object-based verification of short-term, storm-scale probabilistic mesocyclone guidance from an experimental Warn-on-Forecast System. *Wea. Forecasting*, **34**, 1721–1739, https://doi.org/10.1175/WAF-D-19-0094.1.

Flournoy, M. D., M. C. Coniglio, E. N. Rasmussen, J. C. Furtado, and B. E. Coffer, 2020: Modes of storm-scale variability and tornado potential in 2 VORTEX2 near- and far-field tornadic environments. *Mon. Wea. Rev.*, **148**, 4185–4207, https://doi.org/10.1175/MWR-D-20-0147.1.

Frame, J., and P. Markowski, 2010: Numerical simulations of radiative cooling beneath the anvils of supercell thunderstorms. *Mon. Wea. Rev.*, **138**, 3024–3047, https://doi.org/10.1175/2010MWR3177.1.

——, and ——, 2013: Dynamical influences of anvil shading on simulated supercell thunderstorms. *Mon. Wea. Rev.*, **141**, 2802–2820, https://doi.org/10.1175/MWR-D-12-00146.1.

Geerts, B., and Coauthors, 2017: The 2015 Plains Elevated Convection At Night Field Project. *Bull. Amer. Meteor. Soc.*, **98**, 767–786, https://doi.org/10.1175/BAMS-D-15-00257.1.

Hohenegger, C., and C. Schar, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.*, **88**, 1783–1794, https://doi.org/10.1175/BAMS-88-11-1783.

Hong, S. Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Hu, M., H. Shao, D. Stark, K. Newman, C. Zhou, and X. Zhang, 2016: Gridpoint statistical interpolation user's guide version 3.5. Developmental Testbed Centre, 148 pp., https://dtcenter.org/sites/default/files/community-code/gsi/docs/users-guide/GSIUserGuide_v3.5.pdf.

Janjić, Z. I., 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp., http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf.

Jones, T. A., K. Knopmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast System. Part II: Combined radar and satellite data experiments. *Wea. Forecasting*, **31**, 297–327, https://doi.org/10.1175/WAF-D-15-0107.1.

——, P. Skinner, K. Knopmeier, E. Mansell, P. Minnis, R. Palikonda, and W. Smith Jr., 2018: Comparison of cloud microphysics schemes in a Warn-on-Forecast System using synthetic satellite objects. *Wea. Forecasting*, **33**, 1681–1708, https://doi.org/10.1175/WAF-D-18-0112.1.

——, and Coauthors, 2020: Assimilation of *GOES-16* radiances and retrievals into the Warn-on-Forecast System. *Mon. Wea. Rev.*, **148**, 1829–1859, https://doi.org/10.1175/MWR-D-19-0379.1.

Judt, F., 2020: Atmospheric predictability of the tropics, middle latitudes, and polar regions explored through global storm-resolving simulations. *J. Atmos. Sci.*, **77**, 257–276, https://doi.org/http://dx.doi.org/10.1175/JAS-D-19-0116.1.

Kerr, C. A., D. J. Stensrud, and X. Wang, 2019: Diagnosing convective dependencies on near-storm environments using ensemble sensitivity analyses. *Mon. Wea. Rev.*, **147**, 495–517, https://doi.org/10.1175/MWR-D-18-0140.1.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.

Klemp, J. B., and R. B. Wilhelmson, 1978: The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.*, **35**, 1070–1096, https://doi.org/10.1175/1520-0469(1978)035<1070:TSOTDC>2.0.CO;2.

——, and R. Rotunno, 1983: A study of the tornadic region within a supercell thunderstorm. *J. Atmos. Sci.*, **40**, 359–377, https://doi.org/10.1175/1520-0469(1983)040<0359:ASOTTR>2.0.CO;2.

Lawson, J. R., J. S. Kain, N. Yussouf, D. C. Dowell, D. M. Wheatley, and K. H. Knopfmeier, 2018: Advancing from convection-allowing NWP to Warn-on-Forecast: Evidence of progress. *Wea. Forecasting*, **33**, 599–607, http://dx.doi.org/10.1175/waf-d-17-0145.1.

Liang, X., 2007: An integrating velocity–azimuth process single-Doppler radar wind retrieval method. *J. Atmos. Oceanic Technol.*, **24**, 658–665, https://doi.org/10.1175/JTECH2047.1.

Lilly, D. K., 1990: Numerical prediction of thunderstorms—Has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779–798, https://doi.org/10.1002/qj.49711649402.

Mahrt, L., 1981: The early evening boundary layer transition. *Quart. J. Roy. Meteor. Soc.*, **107**, 329–343, https://doi.org/10.1002/qj.49710745205.

Markowski, P. M., and Y. P. Richardson, 2010: *Mesoscale Meteorology in Midlatitudes*. Wiley, 430 pp.

——, and ——, 2014: The influence of environmental low-level shear and cold pools on tornadogenesis: Insights from idealized simulations. *J. Atmos. Sci.*, **71**, 243–275, https://doi.org/10.1175/JAS-D-13-0159.1.

Nakanishi, M., and H. Niino, 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31, https://doi.org/10.1023/B:BOUN.0000020164.04146.98.

——, and ——, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional

prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407, https://doi.org/10.1007/s10546-005-9030-8.

NOAA/NCEI, 2020: Storm events database, version 3.0. NOAA/National Climatic Data Center, accessed 16 December 2020, https://www.ncdc.noaa.gov/stormevents/versions.jsp.

Nowotarski, C. J., and P. M. Markowski, 2016: Modifications to the near-storm environment induced by simulated supercell thunderstorms. *Mon. Wea. Rev.*, **144**, 273–293, https://doi.org/10.1175/MWR-D-15-0247.1.

Parker, M. D., 2014: Composite VORTEX2 supercell environments from near-storm soundings. *Mon. Wea. Rev.*, **142**, 508–529, https://doi.org/10.1175/MWR-D-13-00167.1.

Peters, J. M., H. Morrison, T. C. Nelson, J. N. Marquis, J. Mulholland, and C. J. Noqotarski, 2022: The influence of shear on deep convection initiation. Part II: Simulation. *J. Atmos. Sci.*, **79**, 1691–1711, https://doi.org/doi:10.1175/JAS-D-21-0144.1.

Pielke, R., and R. E. Carbone, 2002: Weather impacts, forecasts, and policy: An integrated perspective. *Bull. Amer. Meteor. Soc.*, **83**, 393–403, https://doi.org/10.1175/1520-0477(2002)083<0393:WIFAP>2.3.CO;2.

Potvin, C. K., K. L. Elmore, and S. J. Weiss, 2010: Assessing the impacts of proximity sounding criteria on the climatology of significant tornado environments. *Wea. Forecasting*, **25**, 921–930, https://doi.org/10.1175/2010WAF2222368.1.

——, and Coauthors, 2019: Systematic comparison of convection-allowing models during the 2017 NOAA HWT Spring Forecasting Experiment. *Wea. Forecasting*, **34**, 1395–1416, https://doi.org/10.1175/WAF-D-19-0056.1.

——, and Coauthors, 2020: Assessing systematic impacts of PBL schemes on storm evolution in the NOAA Warn-on-Forecast System. *Mon. Wea. Rev.*, **148**, 2567–2590, https://doi.org/10.1175/MWR-D-19-0389.1.

Rasmussen, E. N., and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148–1164, https://doi.org/10.1175/1520-0434(1998)013<1148:ABCOSD>2.0.CO;2.

Rotunno, R., and J. B. Klemp, 1982: The influence of the shear-induced pressure gradient on thunderstorm motion. *Mon. Wea. Rev.*, **110**, 136–151, https://doi.org/10.1175/1520-0493(1982)110<0136:TIOTSI>2.0.CO;2.

Skamarock, W. C., and J. B. Klemp, 2008: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J. Comput. Phys.*, **227**, 3465–3485, https://doi.org/10.1016/j.jcp.2007.01.037.

Skinner, P. S., 2016: Application of two spatial verification methods to ensemble forecasts of low-level rotation. *Wea. Forecasting*, **31**, 713–735, https://doi.org/10.1175/WAF-D-15-0129.1.

——, D. M. Wheatley, K. H. Knopfmeier, K. H. Reinhard, A. E. Choate, J. J. Jones, and R. Palikonda, 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.

Smith, E. N., J. G. Gebauer, P. M. Klein, E. Fedorovich, and J. A. Gibbs, 2019: The Great Plains low-level jet during PECAN: Observed and simulated characteristics. *Mon. Wea. Rev.*, **147**, 1845–1869, https://doi.org/10.1175/MWR-D-18-0293.1.

——, M. C. Coniglio, and S. Waugh, 2020: TORUS Doppler lidar and radiosonde wind observation intercomparison. *10th Symp. on Lidar Atmospheric Applications*. Boston, MA, Amer. Meteor. Soc., 366169, https://ams.confex.com/ams/2020Annual/webprogram/Paper366169.html.

Snook, N., 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., C. S. Schwartz, G. S. Romine, K. R. Fossell, and M. L. Weisman, 2016: Severe weather prediction using storm surrogates from an ensemble forecasting system. *Wea. Forecasting*, **31**, 255–271, https://doi.org/http://dx.doi.org/10.1175/WAF-D-15-0138.1.

Stensrud, D. J., 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**, 1487–1500, https://doi.org/10.1175/2009BAMS2795.1.

Straka, J. M., E. N. Rasmussen, and S. E. Fredrickson, 1996: A mobile mesonet for finescale meteorological observations. *J. Atmos. Oceanic Technol.*, **13**, 921–936, https://doi.org/10.1175/1520-0426(1996)013<0921:AMMFFM>2.0.CO;2.

Stull, R. B., 1991: A comparison of parameterized vs. measured transilient mixing coefficients for a convective mixed layer. *Bound.-Layer Meteor.*, **55**, 67–90, https://doi.org/10.1007/BF00119327.

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

van der Walt, S., J. L. Schönberger, J. N. Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, 2014: scikit-image: Image processing in Python. *PeerJ*, **2**, e453, https://doi.org/10.7717/peerj.453.

Wade, A. R., M. C. Coniglio, and C. L. Ziegler, 2018: Comparison of near- and far-field supercell inflow environments using radiosonde observations. *Mon. Wea. Rev.*, **146**, 2403–2415, https://doi.org/10.1175/MWR-D-17-0276.1.

Waugh, S. M., 2021: The "U-tube": An improved aspirated temperature system for mobile meteorological observations, especially in severe weather. *J. Atmos. Oceanic Technol.*, **38**, 1477–1489, https://doi.org/10.1175/JTECH-D-21-0008.1.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520, https://doi.org/10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2.

——, C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW Model. *Wea. Forecasting*, **23**, 407–437, https://doi.org/10.1175/2007WAF2007005.1.

——, and Coauthors, 2015: The Mesoscale Predictability Experiment (MPEX). *Bull. Amer. Meteor. Soc.*, **96**, 2127–2149, https://doi.org/10.1175/BAMS-D-13-00281.1.

Weiss, C. C., J. L. Schroeder, J. Guybes, P. S. Skinner, and J. Beck, 2009: The TTUKa mobile Doppler radar: Coordinated radar and in situ measurements of supercell thunderstorms during Project VORTEX2. *34th Conf. on Radar Meteorology,* Williamsburg, VA, Amer. Meteor. Soc., 11B.2, https://ams.confex.com/ams/34Radar/webprogram/Paper155425.html.

Wheatley, D. M., K. H. Knopmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0043.1.

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.

Wingo, S. M., and K. R. Knupp, 2015: Multi-platform observations characterizing the afternoon-to-evening transition of the planetary boundary layer in northern Alabama. *Bound.-Layer Meteor.*, **155**, 29–53, https://doi.org/10.1007/s10546-014-9988-1.

Wurman, J., D. Dowell, Y. Richardson, P. Markowski, E. Rasmussen, D. Burgess, L. Wicker, and H. Bluestein, 2012: The Second Verification of the Origins of Rotation in Tornadoes Experiment: VORTEX2. *Bull. Amer. Meteor. Soc.*, **93**, 1147–1170, https://doi.org/10.1175/BAMS-D-11-00010.1.

Yussouf, N., J. Gao, D. J. Stensrud, and G. Ge, 2013: The impact of mesoscale environmental uncertainty on the prediction of a tornadic supercell storm using ensemble data assimilation approach. *Adv. Meteor.*, **2013**, 731647, https://doi.org/10.1155/2013/731647.

——, J. S. Kain, and A. J. Clark, 2016: Short-term probabilistic forecasts of the 31 May 2013 Oklahoma tornado and flash flood event using a continuous-update-cycle storm-scale ensemble system. *Wea. Forecasting*, **31**, 957–983, https://doi.org/10.1175/WAF-D-15-0160.1.

Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2016: Intrinsic predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma at storm scales. *Mon. Wea. Rev.*, **144**, 1273–1298, https://doi.org/doi.org/10.1175/MWR-D-15-0105.1.