Sensitivities of 1-km Forecasts of 24 May 2011 Tornadic Supercells to Microphysics Parameterizations

DEREK R. STRATMAN[a]

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

KEITH A. BREWSTER

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

(Manuscript received 22 July 2016, in final form 23 March 2017)

ABSTRACT

On 24 May 2011, Oklahoma experienced an outbreak of tornadoes, including one rated EF5 on the enhanced Fujita (EF) scale and two rated EF4. The extensive observation network in this area makes this an ideal case to examine the impact of using five different microphysics parameterization schemes, including single-, double-, and triple-moment microphysics, in an efficient high-resolution data assimilation system suitable for nowcasting and short-term forecasting with low latencies. Additionally, the real-time configuration of the 1-km ARPS, which assimilated increments produced by 3DVAR with cloud analysis using incremental analysis updating (IAU), had success providing a good baseline forecast. ARPS forecasts of 0–2 h are verified using observation-point, neighborhood, and object-based verification techniques. The object-based verification technique uses updraft helicity fields to represent mesocyclone centers, which are verified against tornado locations from three supercells of interest. Varying levels of success in the forecasts are found and appear to be dependent on the complexity of the storm interaction, with early forecasts of isolated storms exhibiting the most success. Verification scores indicate that the multimoment microphysics schemes tend to produce better forecasts of tornadic supercells. However, some of the forecasts from the single-moment microphysics schemes perform as well as or better than the forecasts from the multimoment microphysics schemes.

1. Introduction

On the afternoon of 24 May 2011, an outbreak of 12 tornadoes, including two EF4 [on the enhanced Fujita (EF) scale] tornadoes and one EF5 tornado, afflicted northern and central Oklahoma within the National Weather Service (NWS) Norman, Oklahoma, Weather Forecast Office's (WFO) County Warning Area (CWA; NWS 2016). Unfortunately, this outbreak caused numerous deaths and injuries along with considerable damage. An extensive observation network was in place in this area during the spring of 2011, allowing for ample research opportunities. Several observational (Brotzge and Luttrell 2015; French et al. 2015; Heymsfield et al. 2013; Houser et al. 2015; Xu et al. 2015) and numerical weather prediction (NWP; Clark et al. 2013; Fierro et al. 2012; Jones et al. 2015; Shapiro et al. 2015; Tanamachi et al. 2015) studies have taken advantage of the available observational data from the 24 May 2011 severe weather outbreak. The observational studies used different sources of radar data to analyze various facets of this outbreak. For example, Heymsfield et al. (2013) analyzed measurements collected when flying a dual-frequency (Ku and Ka band) nadir-pointing Doppler radar through storms containing hail and graupel, while Houser et al. (2015) used data collected by a mobile, rapid-scan, X-band, polarimetric, Doppler radar (RaXPol) to understand the dynamical processes leading to the genesis of the EF5 tornado a few minutes after the dissipation of an EF3 tornado.

Other studies used the observational data with NWP to show the potential benefits of assimilating lightning data into initial analyses (Fierro et al. 2012), reveal the

---

[a] Current affiliation: National Severe Storms Laboratory, Norman, Oklahoma.

Corresponding author: Dr. Derek R. Stratman, derek.stratman@noaa.gov

DOI: 10.1175/MWR-D-16-0282.1

benefits of assimilating both satellite and radar data on analyses and subsequent forecasts (Jones et al. 2015), assess the impact of advection correction in trajectory calculations (Shapiro et al. 2015), and explore the role of a storm merger on a tornado handoff (Tanamachi et al. 2015). Clark et al. (2013) assessed the tornado pathlength forecasts via the updraft helicity (UH; Kain et al. 2008) field from the Storm Scale Ensemble Forecast (SSEF) system for several cases in 2011, including 24 May. That study showed the 4-km SSEF forecasted tornado pathlengths quite well for 24 May 2011, but placed the UH tracks too far to the east as a result off the mishandling of the dryline. Regardless of the dryline placement, the tight clustering of the tornadic and nontornadic supercells and storms on 24 May 2011 made the forecasting of supercell tracks challenging for storm-scale models. The Center for Analysis and Prediction of Storms (CAPS) real-time 1-km forecasting system (Brewster et al. 2010) had good success in predicting the potential severity and location of these storms (not shown). However, forecasts of supercell evolution and motions (i.e., location and timing) might be further improved by using more sophisticated parameterizations of microphysical processes.

Cloud and precipitation microphysics parameterizations play a large role in all scales of NWP. For instance, microphysics parameterizations dictate the extent and location of latent heating and cooling associated with water-phase changes (i.e., freezing, melting, evaporation, condensation, deposition, and sublimation) and precipitation type and amount. These microphysical adjustments can then impact near-surface conditions. For example, varying properties of downdrafts can alter near-surface temperatures and/or wind speeds and directions, and cloud cover can indirectly alter near-surface temperatures through blocking or trapping of longwave and shortwave radiation. From a short-term, small-scale convective forecasting perspective, these microphysics parameterization impacts can affect simulated storm intensity, motion (i.e., speed and direction), and mode (e.g., supercell versus MCS), which are key attributes of severe weather forecasting.

Numerous studies have investigated the sensitivities of cloud and precipitation microphysics parameterization schemes on storm-scale forecasts (e.g., Dawson et al. 2010, 2015; Putnam et al. 2014; Wainwright et al. 2014; Wheatley et al. 2014; Yussouf et al. 2013). These microphysics-related studies comparing microphysics schemes come to the general conclusion that multimoment microphysics schemes produce better forecasts of idealized and real supercells (Dawson et al. 2010, 2015; Wainwright et al. 2014; Yussouf et al. 2013) and mesoscale convective systems (MCSs; Putnam et al. 2014; Wheatley et al. 2014) than do single-moment microphysics schemes. However, the previous studies using simulations of supercells to compare microphysics schemes only focused on a single, isolated supercell. This study will evaluate microphysics schemes using a spectrum of storm-interaction complexity ranging from an isolated supercell to several supercells coexisting as close as 30 km to each other.

Microphysics parameterization schemes are generally characterized as bin and bulk schemes. Bin schemes use discrete bins to form particle size distributions for various particle sizes and types, while bulk schemes use continuous functional forms to describe particle size distributions. Bulk schemes are largely preferred for NWP and storm simulation studies over bin schemes as a result of the bin schemes being computationally more expensive and having less success at predicting changes in ice particle concentrations (Stensrud 2007). Therefore, this study will focus on the impact of various bulk schemes on storm-scale forecasts. Specifically, research experiments are done using five different microphysics parameterization schemes: Lin 3-ice microphysics scheme (LIN3; Lin et al. 1983), Weather Research and Forecasting (WRF) Model single-moment 6-class microphysics scheme (WSM6; Hong and Lim 2006), Milbrandt and Yau single-moment bulk microphysics scheme (MYSM), Milbrandt and Yau double-moment bulk microphysics scheme (MYDM), and Milbrandt and Yau triple-moment bulk microphysics scheme (MYTM; Milbrandt and Yau 2005a,b).

All five of these schemes are based on particle size distributions most generally described by a gamma distribution:

$$N_x(D) = N_{0x} D^{\alpha_x} e^{-\lambda_x D}, \qquad (1)$$

where N_x is the number of particles per unit volume (m^{-4}) for species x, D is the particle diameter (m), N_{0x} is the intercept parameter, α_x is the spectral shape parameter, and λ_x is the slope parameter. LIN3, WSM6, and MYSM only predict the mixing ratios for each species, namely, water vapor (q_v), cloud water (q_c), rain (q_r), snow (q_s), ice (q_i), hail (q_h; LIN3 and MYSM only), and graupel (q_g; WSM6 and MYSM only). For these single-moment schemes, $\alpha_x = 0$, N_{0x} is constant (Table 1), and λ_x varies according to the mixing ratio (i.e., inverse-exponential distribution). MYDM predicts both mixing ratios (q_x) and number concentrations (N_x) for the same species as MYSM, so while $\alpha_x = 0$ for this study, both N_{0x} and λ_x can vary with q_x and N_x. Finally, MYTM predicts q_x, N_x, and reflectivities (Z_x) for the same species as MYSM. The addition of the third

TABLE 1. Intercept parameter values for the simulations using the single-moment microphysics parameterization schemes. MD means monodisperse distribution, and $f(T)$ signifies that N_0 is based on a function of temperature [see Hong and Lim (2006) for WSM6's N_{0s} and Milbrandt and Yau (2005b) for MYSM's N_{0i}].

	LIN3	WSM6	MYSM
N_{0r} ($\times 10^6$ m^{-4})	8.0	8.0	8.0
N_{0i} ($\times 10^6$ m^{-4})	MD	MD	$f(T)$
N_{0s} ($\times 10^6$ m^{-4})	3.0	$f(T)$	3.0
N_{0g} ($\times 10^6$ m^{-4})	—	4.0	0.4
N_{0h} ($\times 10^6$ m^{-4})	0.04	—	0.04

moment allows for α_x to vary independently of N_{0x} and λ_x (Milbrandt and Yau 2005a).

For this study, an object-based verification technique is utilized in a new way to determine location and timing errors for simulated circulation centers, which are typically associated with severe weather. Additionally, all of the projects employ for the first time the recently updated Advanced Regional Prediction System's (ARPS; Xue et al. 2000; Xue et al. 2001; Xue et al. 2003) data assimilation system (ADAS) complex cloud analysis package (Hu et al. 2006a,b; Brewster and Stratman 2015). In section 2, an overview of the 24 May 2011 Oklahoma tornado outbreak is presented. Verification metrics and methods are described in section 3. The experiment design and model configurations are detailed in section 4. The verification results of the experiment are presented in section 5. Finally, section 6 contains a summary and discussion of the results from this study.

2. Overview of event

On 24 May 2011, a well-forecasted severe weather outbreak plagued parts of Kansas, Missouri, Oklahoma, Arkansas, and Texas with all facets of severe weather (SPC 2016). Tornadic and nontornadic supercells developed along the dryline and trekked northeast across Oklahoma, which led to the issuance of 34 tornado warnings by the Norman, Oklahoma, WFO (Fig. 1). Within the NWS Norman WFO's CWA, the outbreak was responsible for a dozen tornadoes comprising one EF5 (S1b), two EF4 (S2a and S3a), two EF3 (S0a and S1a), two EF2 (S1c and S3c), two EF1 (S3b and S4a), and three EF0 (S0b; other two not annotated) tornadoes (Fig. 1) and, unfortunately, 11 deaths and 342 injuries (NCDC 2011). Damaging wind and large hail up to 3 in. (7.62 cm) in diameter also wreaked havoc on many areas (SPC 2016).

Severe weather episodes tend to transpire when key atmospheric ingredients (i.e., shear, lift, instability, and moisture) come together in space and time. As Brotzge and Luttrell (2015) described, this severe weather outbreak was a classic high-shear, high-instability event with >25 m s^{-1} of effective bulk shear (Fig. 2a) and >3000 J kg^{-1} of CAPE (Fig. 2b), which collectively support supercells, and significant tornadoes were possible given >200 m^2 s^{-2} of 0–1-km storm-relative helicity (Fig. 2c; Thompson et al. 2003; Markowski et al. 2003). Upper-level divergence and the dryline provided large-scale ascent and focused lift, respectively, while a stout ~800-hPa warm nose mostly prevented widespread convection across the warm sector, especially during the first couple of hours after convective initiation. Ample low-level moisture aided in low lifting-condensation levels (i.e., <1000 m), which benefit the generation of tornadoes (Fig. 2d; Rasmussen and Blanchard 1998). Isolated storm modes were favored because of the deep-layer mean wind vector being mostly perpendicular to the dryline (Bluestein and Weisman 2000; Dial and Racy 2004). With these severe weather ingredients forecasted and observed, NWS and SPC forecasters were confident that the potential for an outbreak of tornadic supercells existed across the southern plains region, especially Oklahoma, as reflected in the convective outlooks leading up to the event (e.g., the 0600 UTC day 1 convective outlook contained a high risk for a majority of the eastern two-thirds of Oklahoma; SPC 2016). With such potential for severe weather over a wide area, it is important to try to more precisely identify areas at risk of damage and/or life-threatening conditions.

3. Verification methodology

In this study, a few different verification techniques are employed to properly assess the forecasts of tornadic supercells. First, root-mean-square errors (RMSEs) for four near-surface variables are computed using the following equation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(f_k - o_k)^2}, \qquad (2)$$

where n is the total number of observation locations and f_k and o_k are the forecasted and observed variables, respectively, at the kth observation location. For consistency of observing instruments, the RMSEs are calculated using only the Oklahoma Mesonet dataset, which features 5-min averages for 1.5-m temperature T, 1.5-m dewpoint temperature T_d, and 10-m u- and v-component wind [u and v, respectively; Brock et al. (1995); McPherson et al. (2007)]. The forecasted values f_k for T, T_d, u, and v are first linearly interpolated in the

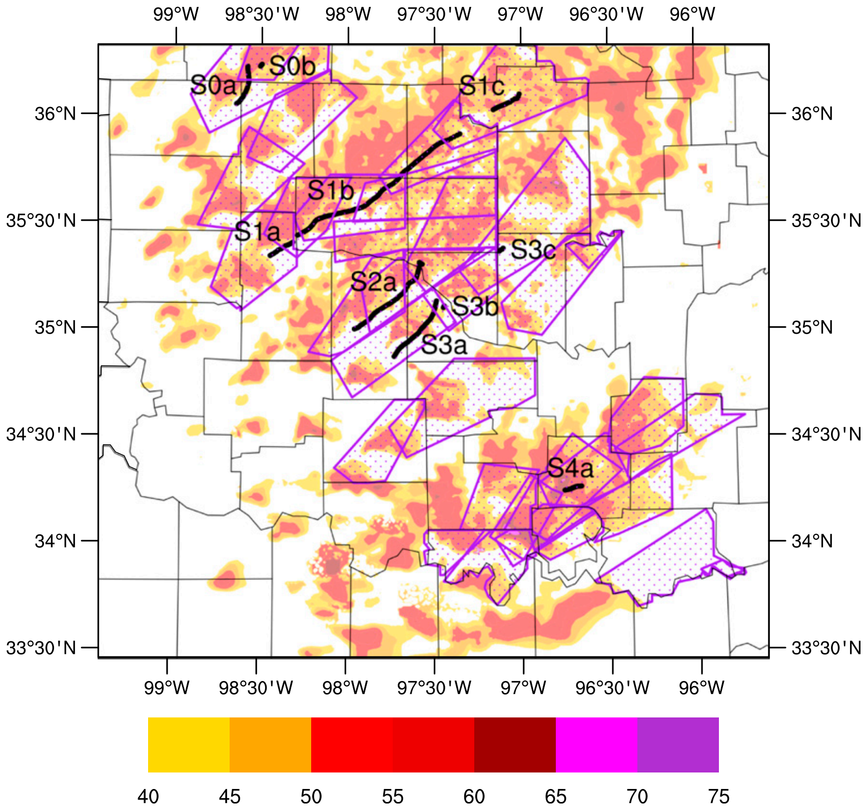


FIG. 1. Maximum 0.5°-tilt reflectivity (dBZ) above 40 dBZ from the Oklahoma City radar (KTLX) for output times 1902, 1932, 2002, 2032, 2102, 2132, 2201, 2231, 2301, and 2331 UTC on 24 May 2011 and 0000, 0030, and 0101 UTC on 25 May 2011 are plotted with color shading at 50% opacity. Tornado tracks (black) are labeled by storm number and lettered sequentially. Tornado warnings issued by the Norman WFO during the 24 May 2011 tornado outbreak are represented by magenta polygon outlines and stipples.

vertical to 2, 2, 10, and 10 m AGL, respectively, and are then bilinearly interpolated to the n mesonet locations. For interpolation below the lowest model grid point at ~10 m AGL, the potential temperature and specific humidity are assumed constant and pressure is taken to be hydrostatic.

In addition to RMSE, linear regression slopes, coefficients of determination, and probability distribution functions (PDFs) are calculated for further comparisons. The linear regression slopes are computed using the following equation:

$$m = \frac{\sum_{k=1}^{n}[(o_k-\overline{o})(f_k-\overline{f})]}{\sum_{k=1}^{n}(o_k-\overline{o})^2}, \qquad (3)$$

where $\overline{f}$ and $\overline{o}$ are the means of the forecasted and observed values, respectively. The coefficients of determination, which are the squares of the Pearson product-moment correlation coefficient, are computed as

$$R^2 = \frac{m^2\sum_{k=1}^{n}(o_k-\overline{o})^2}{\sum_{k=1}^{n}(f_k-\overline{f})^2}, \qquad (4)$$

where the numerator is the regression sum of squares and the denominator is the total sum of squares.

To determine how much a forecast of composite reflectivity needs to be smoothed to exhibit "useful" skill, the fractions skill score (FSS) is computed for 30-, 40-, and 50-dBZ thresholds using the following equation:

$$\mathrm{FSS} = 1 - \frac{\frac{1}{N}\sum_{N}(P_f-P_o)^2}{\frac{1}{N}\left(\sum_{N}P_f^2+\sum_{N}P_o^2\right)}, \qquad (5)$$

where P_f and P_o are the fractional coverages of forecasted and observed values, respectively, exceeding a threshold within a neighborhood window and N is the

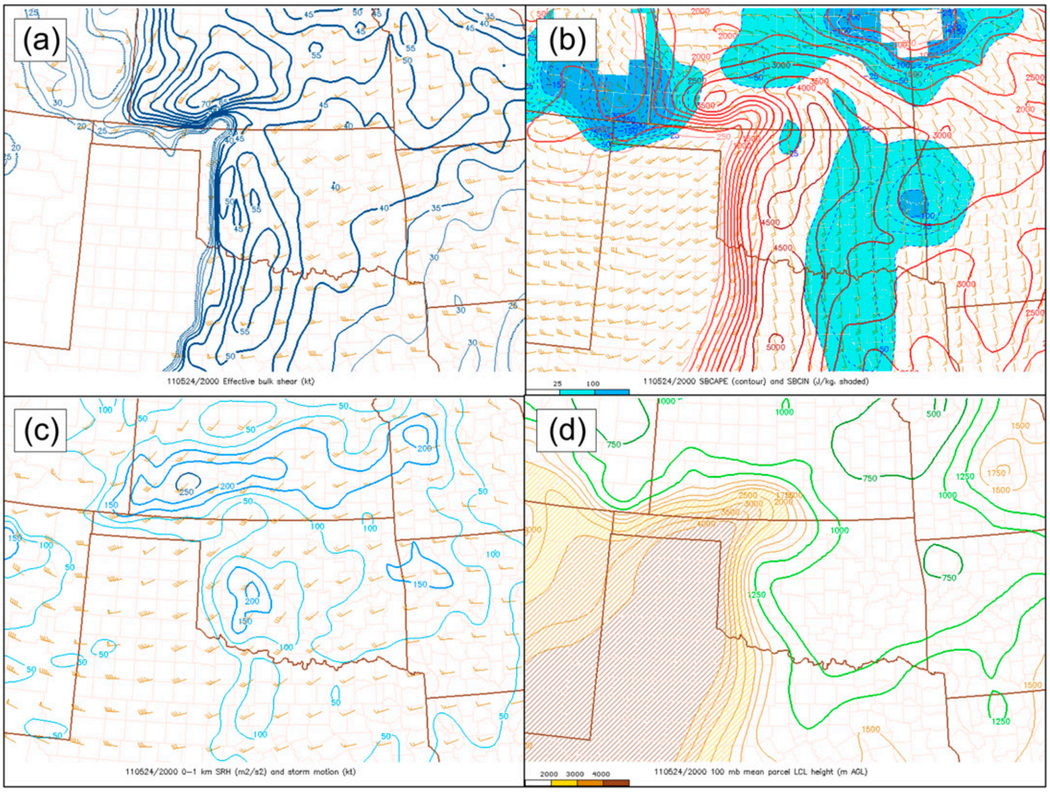


FIG. 2. (a) Effective bulk shear (kt, where 1 kt = 0.5144 m s^{-1}; wind barbs and contours); (b) surface-based CAPE (J kg^{-1}; red contours), surface-based CIN (J kg^{-1}; blue shading), and surface wind (kt; wind barbs); (c) 0–1-km storm-relative helicity (m^2 s^{-2}; blue contours) and storm motion (kt; wind barbs); and (d) 100-mb mean parcel lifting condensation level heights (m AGL; contours and shading) from the SPC Mesoanalysis Archive Data website (http://www.spc.noaa.gov/exper/mesoanalysis/archive/) at 2000 UTC 24 May 2011.

number of neighborhood windows used for each neighborhood size (Ebert 2008, 2009). The FSS is a measure of the differences between the probabilities of a forecast field exceeding a threshold and the probabilities of an associated observed field exceeding a threshold within subsets of the model domain (i.e., neighborhood windows). For this study, the neighborhood windows are $n \times n$ gridpoint squares on the 1-km model grid, where $n = 1, 3, 5, 9, 17, 33, 65$, and 129 grid points. In addition, the FSS for the full domain ($N = 1$) is calculated. The FSS ranges from 0 (no skill) to 1 (perfect skill). Roberts and Lean (2008) defined the "useful" FSS as $\mathrm{FSS}_{\mathrm{useful}} = 0.5 + f_o/2$, which is the halfway point between a random forecast ($\mathrm{FSS}_{\mathrm{random}} = f_o$) and a perfect forecast (FSS = 1). The observed base rate (i.e., the ratio between the number of grid points with values exceeding a threshold and the total number of grid points) is f_o. This study will use $\mathrm{FSS}_{\mathrm{useful}}$ to determine the smallest scales with potentially "useful" skill. Square neighborhood windows are employed, and when grid points within neighborhood windows extend beyond the edge of the domain, those grid points are assigned a value of 0. The observed composite reflectivity from the National Severe Storms Laboratory's national mosaic and quantitative precipitation estimation (NMQ; Zhang et al. 2011) system is used as the observation field.

An object-based verification technique was developed to assess model performance by verifying simulated mesocyclone centers, via the UH field, with estimated tornado points. The locations of the six tornadoes (i.e., S1a, S1b, S1c, S2a, S3a, and S3b in Fig. 1) associated with the three storms of interest are estimated every minute based on NWS damage surveys, radar data, and high-resolution aerial photos from Google Maps. Two adjacent layers of UH (viz., 0–1 and 1–6 km) are used in the verification process to confirm the simulations and are calculated as

$$\mathrm{UH} = \int_{z_1}^{z_2} w\zeta \, dz\,, \qquad (6)$$

which is the integral through the depth, z_1–z_2, of the product of vertical velocity w and the vertical component of relative vorticity ζ. These two layers are

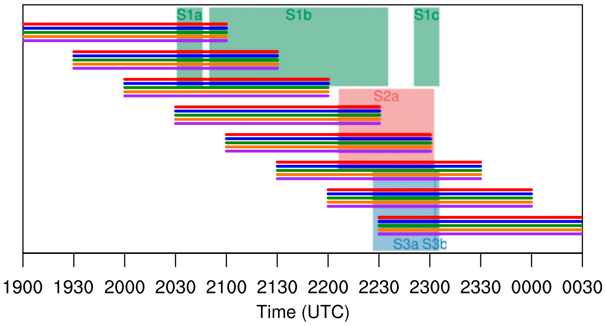


FIG. 3. Schematic of the timeline for all 2-h simulations (LIN3, red; WSM6, blue; MYSM, green; MYDM, orange; and MYTM, purple) and the six tornadoes from the three storms of interest.

intended to represent simulated low- and midlevel mesocyclones, respectively. Kain et al. (2008) used hourly UH from 2 to 5 km AGL to signify midlevel mesocyclones, but for this study, a deeper layer of instantaneous UH is utilized to give more robust UH values by capturing more of the simulated midlevel mesocyclones. The 0–1-km UH value is used to pinpoint rotation near the ground and is expected to be more closely associated with damage reports.

Since UH is a 2D field and not point data, a 2D object-based technique is utilized to find UH-weighted centers (analogous to mass-weighted centers), which will be compared to the estimated tornado points. The selection of objects is similar to other object-based verification studies (e.g., Davis et al. 2006a,b), but convolution is not applied to the UH fields (only thresholding) because of the need for UH objects to be less contiguous than precipitation fields. For this study, a search radius of 4 km (i.e., four grid points) is used to isolate 1–6-km (0–1 km) UH maxima that are greater than or equal to 300 $\mathrm{m^2\,s^{-2}}$ (15 $\mathrm{m^2\,s^{-2}}$) and their surrounding gridpoint values. A maximum UH value is considered a UH-center candidate if four out of eight (one out of eight) of the adjacent gridpoint values equal or exceed 150 $\mathrm{m^2\,s^{-2}}$ (10 $\mathrm{m^2\,s^{-2}}$). Once a UH-center candidate is determined, the UH-weighted center is computed using a radius of 3 km (2 km) extending from the grid point with the maximum UH value. The 0–1-km UH-weighted centers are filtered by requiring a 1–6-km UH-weighted center to concurrently exist within 5 km. This is meant to separate centers of strong low-level updraft rotation that are also associated with a substantial midlevel rotating updraft from more benign low-level shear maxima or rotation centers that are commonly found along gust fronts. The thresholds were determined through a subjective assessment of what UH entities should be considered objects (not shown). However, the qualitative results remain similar and independent of the thresholds.

With the filtered 0–1-km UH-weighted center locations, an objective verification technique is used to quantify location and timing errors. This technique is similar to methods used to verify simulated tropical cyclone tracks and intensities (e.g., Xue et al. 2013) and precipitation objects (e.g., Davis et al. 2006b). First, distance errors are computed between 0–1-km UH (0–1UH) center locations and the nearest estimated tornado locations at coincident times (referred to as ‘‘same time,’’ or ST, for the rest of this paper). Second, distance and timing errors are calculated between 0–1UH center locations and the nearest estimated tornado locations at any time during the life of the tornadoes of interest (referred to as ‘‘anytime,’’ or AT, for the rest of this paper). The 0–1UH centers with the smallest distance errors at each forecast time for each storm of interest are used in the statistics. Positive (negative) timing errors indicate that the simulated 0–1UH centers are too fast (too slow). For example, if the closest tornado point to a 0–1UH center occurred at 2140 UTC while the 0–1UH center occurred at 2130 UTC, the 0–1UH center is 10 min too fast. The average maximum 0–1UH value for each center and the total number of 0–1UH centers are also computed for further evaluation.

4. Experiment design

For this experiment, eight microphysically diverse sets of simulations are run using a potential future real-time, storm-scale forecasting framework. In short, each set of simulations is initialized every 30 min from 1855 to 2225 UTC, spun up during a 5-min data assimilation window, and then integrated out to 120 min. Within this framework, four sets of simulations provide forecasts encompassing each observed storm of interest’s first tornadogenesis (Fig. 3). The first storm (S1; storms depicted in Fig. 4) developed and stayed outside the Engineering Research Center for Collaborative Adaptive Sensing of the Environment (CASA) radar network and produced two tornadoes, including the outbreak’s only EF5 tornado. The second and third storms (S2 and S3, respectively) developed in the CASA radar network and both produced EF4 tornadoes, which dissipated before impacting the Oklahoma City, Oklahoma, metropolitan area.

Recently, the ARPS (Xue et al. 2000; Xue et al. 2001; Xue et al. 2003) ADAS complex cloud analysis package (Hu et al. 2006a,b) was updated for several microphysics schemes, including the five in this study (Brewster and Stratman 2015). The goal of this update was to improve analyses of hydrometeors using scheme-specific reflectivity inversion equations. The 1800 UTC 12-km North American Mesoscale Forecast System (NAM) forecasts (output every 3 h) are interpolated to the

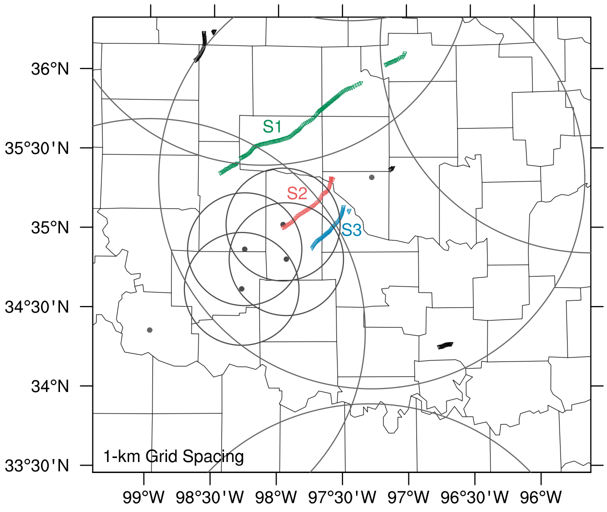


FIG. 4. Domain of numerical simulations with WSR-88D and CASA radar locations (gray dots), 40-km range rings for CASA radars (small gray circles), 150-km range rings for WSR-88D radars (large gray circles), and estimated tornado points for S1 (green), S2 (red), and S3 (blue).

initialization times (i.e., 1855, 1925, 1955, 2025, 2055, 2125, 2155, and 2225 UTC) and used as background fields in the 3DVAR data assimilation (Gao et al. 2004) and complex cloud analysis (Hu et al. 2006a, Brewster et al. 2015) to produce initial analyses on a 323 × 353 gridpoint domain with 1-km horizontal grid spacing (Fig. 4) and 53 vertically stretched levels with a minimum vertical grid spacing of 20 m at the bottom.

Three analysis passes with 20, 50, and 50 iterations for minimizations and horizontal influence radii of 45, 2, and 1 km, respectively, are used to produce the 3DVAR analyses through the minimization of the cost function. Surface observations from NWS and FAA METARs and Oklahoma Mesonet stations are utilized in the first and third passes. The radial wind and reflectivity data from the WSR-88D [Dallas/Fort Worth, Texas (KFWS); Dodge City, Kansas (KDDC); Frederick, Oklahoma (KFDR); Tulsa, Oklahoma (KINX); Twin Lakes, Oklahoma (KTLX); Vance Air Force Base, Oklahoma (KVNX); and Wichita, Kansas (KICT)] and CASA IP-1 X-band [Chickasha, Oklahoma (KSAO); Cyril, Oklahoma (KCYR); Lawton, Oklahoma (KLWE); and Rush Springs, Oklahoma (KRSP); see Fig. 4] radar networks (McLaughlin et al. 2009) are applied in the second and third passes. In addition, a 3D mass divergence constraint is utilized to couple the wind components together (Hu et al. 2006b).

After the 3DVAR analyses are produced, the ARPS model simulations are integrated forward from the interpolated NAM forecasts to produce forecasts out to 125 min. During the first 5 min, incremental analysis updating (IAU; Bloom et al. 1996) is performed by introducing fractional analysis increments every 20 s (i.e., specific fractions of the total increments are added directly to the various variable fields). The fractional increments have a triangular time-weighting pattern to slowly ramp up and then down with the maximum around the midpoint of the assimilation window (Fig. 5). The increments are applied to all fields except for vertical velocity and pressure since those two fields are not directly observed in 3D and will quickly respond to the other fields to create a balanced state.

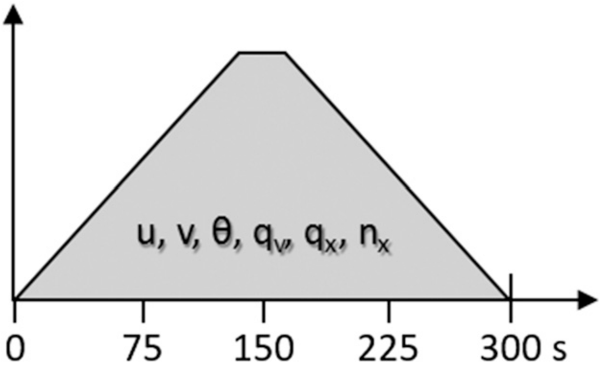


FIG. 5. Schematic of the temporally weighted distributions of the analysis increments.

The simulations proceed without further data insertion or adjustment for the remaining 120 min. During the integration of ARPS, a big and small time step of 2.0 and 0.5 s, respectively, are employed in the leapfrog time formulation. The 1800 UTC 12-km NAM forecasts are used for the lateral boundary conditions. Some other model details include the 4th-order momentum advection in both the horizontal and vertical directions, scalar advection using Zalesak's multidimensional version of flux-corrected transport (Zalesak 1979), 1.5-order TKE closure based on Sun and Chang (1986), 4th-order computational mixing, Rayleigh damping beginning at 12 km AGL, National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) atmospheric parameterization of longwave and shortwave radiation processes [Chou (1990, 1992) and Chou and Suarez (1994), respectively], surface fluxes calculated from stability-dependent surface drag coefficients using predicted surface temperature and volumetric water content, and a two-layer force–store soil model based on Noilhan and Planton (1989).

5. Results

a. Observation-point results

Since microphysics parameterization schemes can have an impact on near-surface variables, the RMSE is

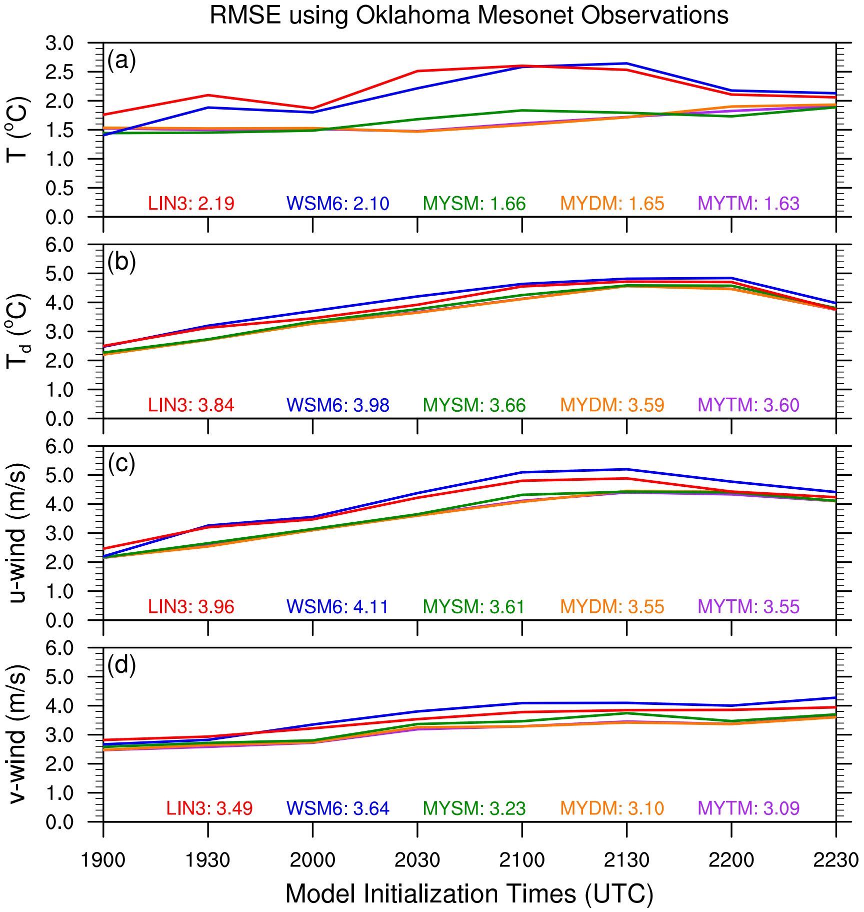


FIG. 6. Average RMSEs plotted for each simulation for near-surface (a) temperature (°C), (b) dewpoint temperature (°C), (c) u-component wind (m s^{-1}), and (d) v-component wind (m s^{-1}). Average RMSEs for all simulations for each microphysics scheme are annotated along the bottom of the plots.

computed for each forecast time in 5-min intervals for near-surface temperature, dewpoint temperature, pressure, and u- and v-wind components. MYTM has the smallest average RMSE for temperature for all simulations and forecast times with 1.63°C, but MYDM and MYSM's average RMSEs for temperature are within 0.03°C (Fig. 6a). LIN3 and WSM6 both have average RMSEs for temperature around 0.5°C larger than the MY schemes. The RMSEs for dewpoint temperature are approximately twice the RMSEs for temperature, but once again, the MY schemes have the smallest average RMSEs, though the differences among the microphysics schemes are smaller than for temperature (Fig. 6b).

The RMSEs for u wind gradually increase with later model initialization times while the RMSEs for the v wind largely remain within the 2–4 m s^{-1} range (Figs. 6c,d). Interestingly, the average RMSEs for the u wind are about 0.5 m s^{-1} larger than the RMSEs for the v wind. This difference could be due to incorrect surface roughness lengths and boundary layer and turbulence parameterization schemes undermixing westerlies aloft down to the surface in areas of downdrafts or deep, well-mixed boundary layers (note the wind vectors in western Oklahoma in Fig. 9). To explain the difference in the RMSEs, assume an observed wind is from the west (W)-southwest (SW) (240°) at 10 m s^{-1}, and the corresponding forecasted wind is from the south (S)-SW (210°) at 5 m s^{-1} (i.e., both slower and more easterly). This example yields observed u and v winds of ~8.7 and 5.0 m s^{-1}, respectively, and forecasted u and v winds of 2.5 and ~4.3 m s^{-1}, respectively. As this simple example indicates, the errors are larger for u winds (~6.2 m s^{-1}) than v winds (~0.7 m s^{-1}) when the forecasted winds are slower and more easterly. As with the other variables, the MY schemes generally have the smallest RMSEs for both wind components (Figs. 6c,d). For all five near-surface variables, MYDM and MYTM exhibit very

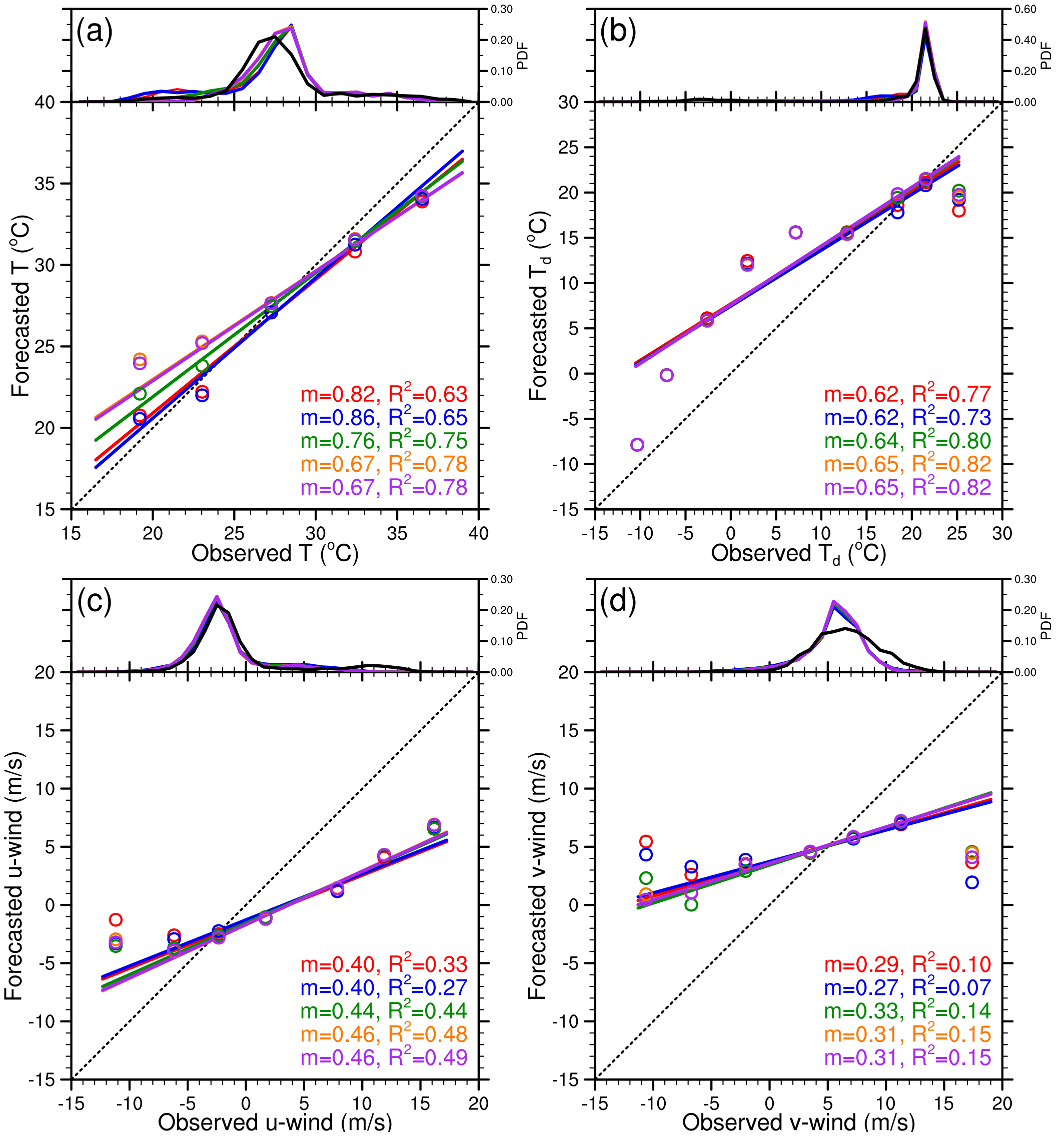


FIG. 7. Linear regression slopes (m) and R^2 for all paired observations and forecasts for (a) temperature (°C), (b) dewpoint temperature (°C), (c) u-component wind (m s^{-1}), and (d) v-component wind (m s^{-1}) are plotted and annotated for LIN3 (red), WSM6 (blue), MYSM (green), MYDM (orange), and MYTM (purple). Colored circles depict averages for observations in distinct five-unit bins and their associated forecast values. Dashed black line represents a perfect, unbiased forecast. Probability distribution functions for the number of forecasted (nonblack lines) and observed (black line) values within one-unit bins are plotted in the top portion of each variable's plot window.

similar RMSEs, with MYSM having only slightly larger average RMSEs. Compared to WSM6, LIN3 has smaller average RMSEs for all variables except for temperature, for which the difference is less than 0.1°C.

For further comparison, linear regression slopes, R^2 values, and PDFs are computed to assess biases, correlations, and distributions, respectively. For temperature, MYDM and MYTM have the largest R^2 values while also having the smallest slopes (Fig. 7a). This is due to MYDM and MYTM having a warm bias for observed temperatures below 25°C (i.e., within cold pools). WSM6's slope for temperature is closest to 1, with LIN3 having a slightly smaller slope, but their R^2 values are substantially smaller than those of the MY schemes (Fig. 7a). MYSM's slope falls in between the two pairs of slopes, but the R^2 value is closer to MYDM and MYTM. All schemes shift the distribution peak to the bin that is 1°C warmer than the bin containing the observed

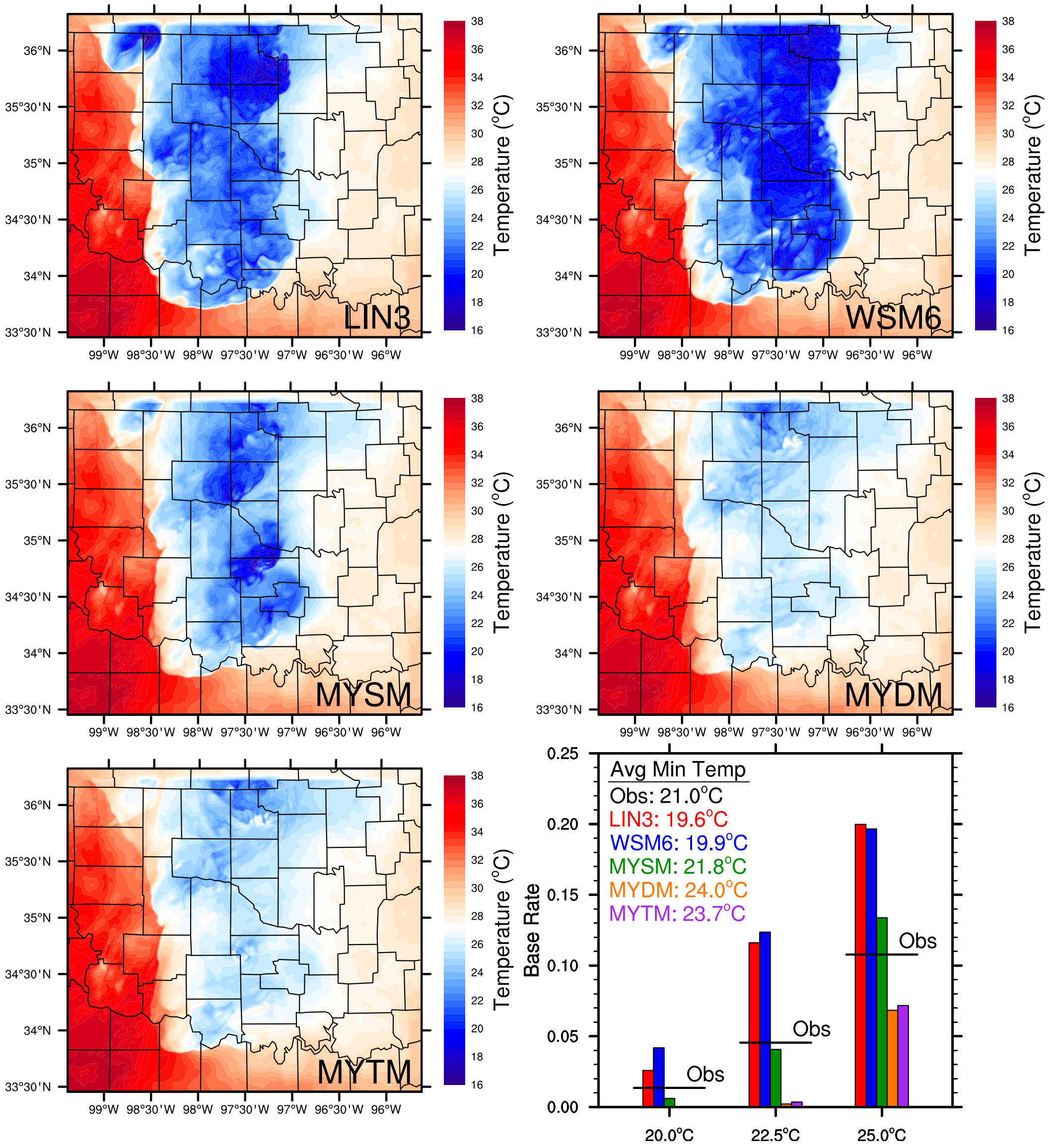


FIG. 8. The 90-min forecasts of near-surface temperatures (°C) from the 2130 UTC simulations are plotted for each of the microphysics schemes. The bar graph depicts the average base rates of temperatures below 20.0°, 22.5°, and 25.0°C at the Oklahoma Mesonet locations for the different microphysics schemes for all simulations. Horizontal black lines represent the observed base rates. Average minimum temperatures for all simulations are annotated in the plot window.

distribution peak. Also, LIN3 and WSM6 exhibit higher probabilities in the bins less than 23°C, indicating larger and colder cold pools. An example forecast of near-surface temperatures and average base rates at Oklahoma Mesonet stations for temperatures below 20°–25°C reveals that is indeed the case for LIN3 and WSM6 (Fig. 8). Conversely, MYDM and MYTM yield cold pools that are too warm, while MYSM is the closest to the observations. On average, all forecasts lead to forecasted surface temperatures cooler than the observed temperatures for temperatures greater than about 30°C, but since the microphysics is not active in these warm regions, the bias is not due to the microphysics. For dewpoint temperature, all forecasts tend to be too moist for observed dewpoint temperatures less than 15°C (i.e., the dry side of the dryline where microphysics is not active; Fig. 7b). The MY schemes have the largest R^2 values, and their slopes are slightly closer to 1 than the other microphysics schemes.

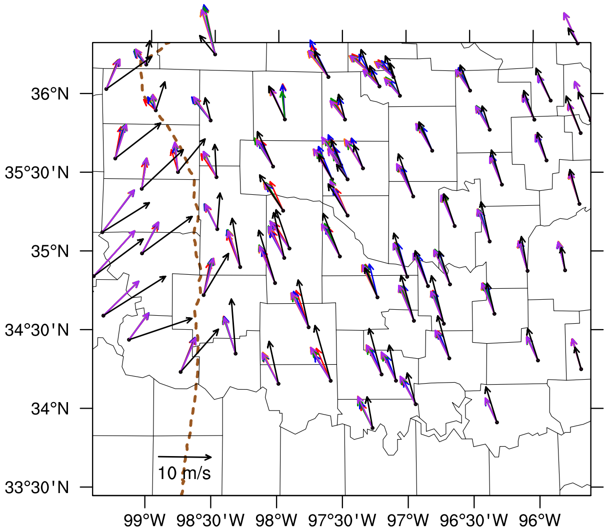


FIG. 9. Average forecasted (using LIN3, red; WSM6, blue; MYSM, green; MYDM, orange; and MYTM, purple) and observed (black) wind vectors at all Oklahoma Mesonet stations within the model domain. An example wind vector at the bottom left of the plot represents a 10 m s^{-1} wind vector. The average analyzed 18°C isodrosotherm (brown dashed line) for 1900–2230 UTC represents the dryline.

The forecasted u-wind speeds tend to be less westerly for winds with a positive u component and more westerly for observed u components less than about -5 m s^{-1} (Fig. 7c). The PDFs are similar among the observations and schemes' forecasts with the majority of the u-wind values existing within the -5 to 0 m s^{-1} range, which is due to the relatively expansive warm, moist sector. The MY schemes' simulations produce the steepest slopes and largest R^2 values for the u wind, but both values remain near or below 0.5 (Fig. 7c). Similar to the u wind, the forecasted v wind speeds are generally less southerly (more southerly) for observed v winds greater than (less than) about 5 m s^{-1} (Fig. 7d). All schemes' simulations are similar and substantially worse at predicting the v wind than the u wind based on shallow slopes (i.e., <0.35), small R^2 values (<0.20), and forecasted distributions of the v wind being too narrow.

Overall, the comparison between the Oklahoma Mesonet observations and forecasted values using various statistics has highlighted some microphysics successes and failures. The MY schemes, especially MYDM and MYTM, exhibit the smallest RMSEs for all five near-surface variables, but the largest RMSE differences between the MY schemes and LIN3 and WSM6 are for the near-surface temperatures as a result of LIN3 and WSM6 producing cold pools that are too cold and too large. The large differences between the observed and forecasted wind fields (i.e., direction and magnitude) largely stems from areas of observed and forecasted convection and areas west of the dryline (Fig. 9); the latter of which is due to background/model errors and not microphysical processes. These various air masses within the model domain likely contribute to the nonlinear nature of the biases (i.e., overforecasting and underforecasting). The differences among the microphysics schemes' forecasts of all four near-surface variables are almost entirely confined to areas of convection (see Figs. 8 and 9). This indicates the MY microphysics schemes contribute to better forecasts of storm-related processes, despite the warmer cold pools in MYDM and MYTM. While the multimoment microphysics schemes generally outperformed the single-moment microphysics schemes, MYSM performed more similarly to MYDM and MYTM than LIN3 and WSM6.

b. *Neighborhood results*

Observed composite reflectivity is a spatially dense, remotely sensed field and, thus, provides a unique opportunity to spatially verify simulated convective systems on the entire grid domain. At the 30-dBZ threshold, LIN3 transitions to useful skill at the smallest scale (~16 km), with MYDM exhibiting useful skill beginning around 19 km (Fig. 10a). MYSM performs the worst, with useful skill only at scales larger than ~41 km. Interestingly, MYTM is more similar to MYSM than MYDM at this threshold. LIN3 depicts the smallest scales for the transition from no useful skill to useful skill for five out of eight sets of simulations, with average useful skill beginning at scales less than 9 km twice (Fig. 10b). MYSM performs the worst, with average useful skill never occurring at scales less than 17 km. In general, the earlier initialized simulations exhibit useful skill at smaller scales than the later initialized simulations likely because of the increasing number of storms and the complex nature of the storm interactions.

At the 40-dBZ threshold, MYDM begins to show useful skill at scales around 12 km (Fig. 10c). Once again, MYSM performs the worst with useful skill on average not existing for scales below ~41 km. MYTM transitions to useful skill around 18 km, which is closer to MYDM than MYSM. While below our useful skill threshold, MYDM and MYTM both depict more skill than the other microphysics schemes at the smaller scales, but at larger scales, LIN3 and WSM6 match and/or beat MYDM and MYTM by this metric. For all initialization times, MYDM on average has the smallest scale at which useful skill begins, with over half of the runs falling at or below 9 km (Fig. 10d). Conversely, MYSM starts exhibiting useful skill at scales larger than the other microphysics schemes for all initialization times except for the 1900 UTC simulations. Once again, there is a slight upward trend in the scales where useful skill begins with later

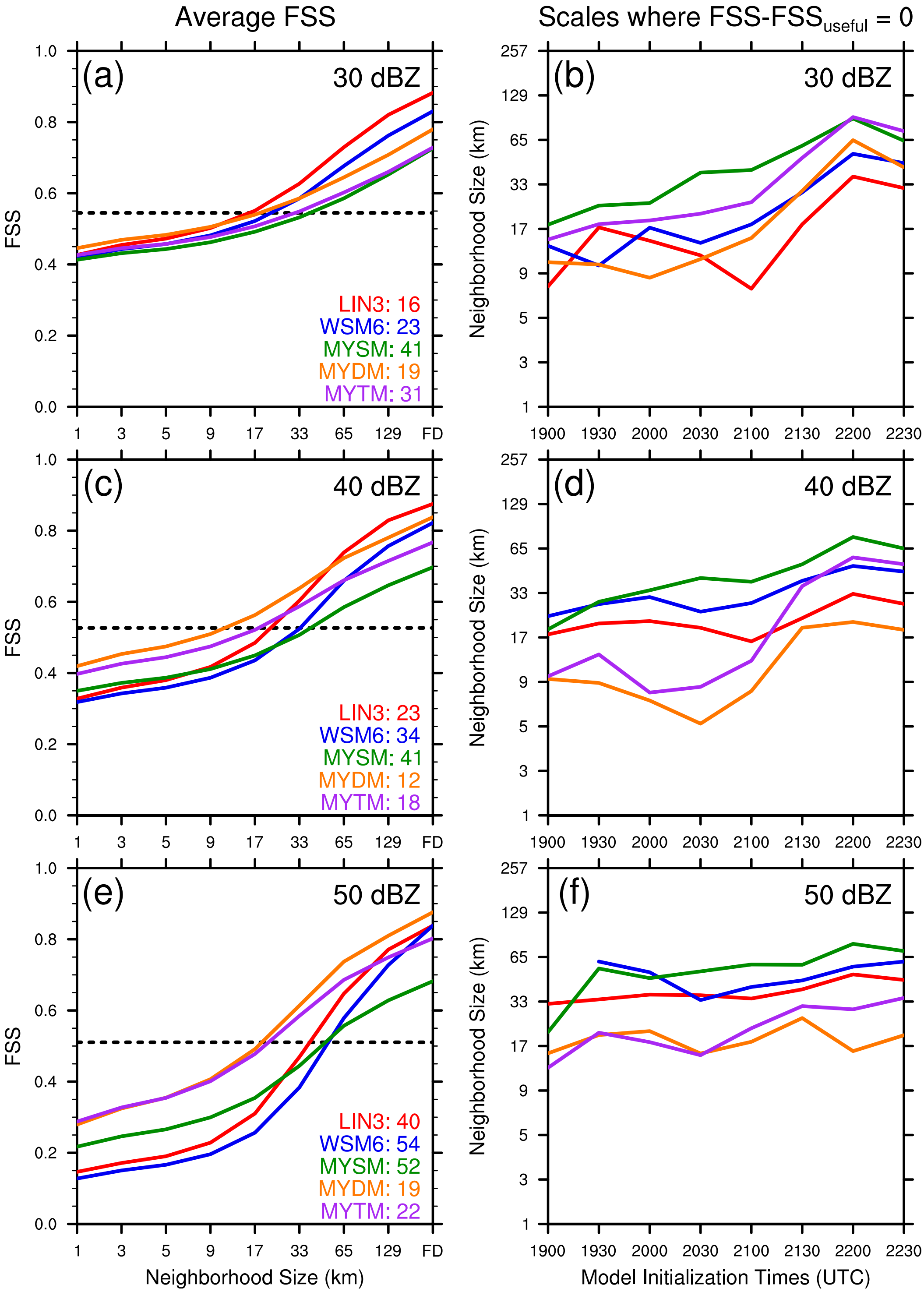


FIG. 10. Average FSS across all simulations and scales (note that FD = full domain) for LIN3 (red), WSM6 (blue), MYSM (green), MYDM (orange), and MYTM (purple) are plotted for thresholds of (a) 30, (c) 40, and (e) 50 dBZ. Average neighborhood sizes (km) at which FSS − FSS$_{useful}$ = 0 for each microphysics scheme are annotated in (a), (c), and (e). The FSS$_{useful}$ results are represented by the horizontal dashed black lines in (a), (c), and (e). Average neighborhood sizes (km) at which FSS − FSS$_{useful}$ = 0 for each model initiation time (UTC) and microphysics scheme are plotted for thresholds of (b) 30, (d) 40, and (f) 50 dBZ.

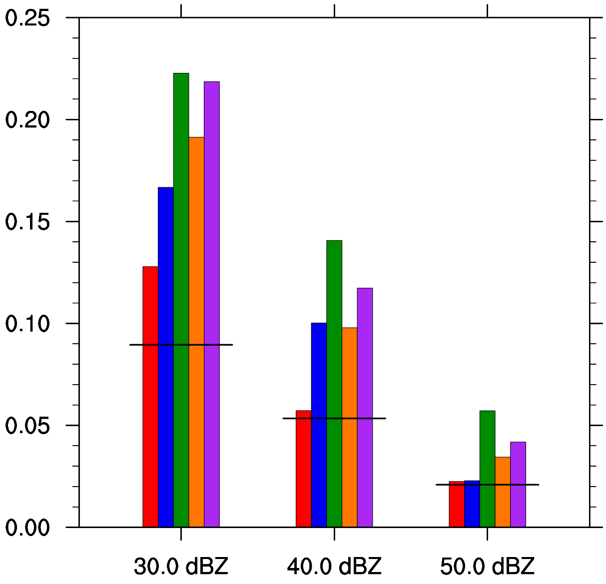


FIG. 11. Bar graph of the forecasted (LIN3, red; WSM6, blue; MYSM, green; MYDM, orange; and MYTM, purple) and observed (black horizontal lines) base rates of composite reflectivity for the 30-, 40-, and 50-dBZ thresholds.

initialization times likely because of an increase in the areal coverage of reflectivity.

For the 50-dBZ threshold, MYDM and MYTM both depict useful skill starting around 19 and 22 km, respectively, while the other three microphysical schemes lead to useful skill beginning around 40 km and larger (Fig. 10e). For the first six model initialization times, MYDM and MYTM depict similar average scales where useful skill begins while WSM6 and MYSM usually perform the worst (Fig. 10f). As an overview, all of the schemes generally become less skillful with increasing reflectivity threshold at the smaller scales, but at the larger scales, LIN3, WSM6, and MYSM remain similar for all thresholds while MYDM and MYTM increase in skillfulness with increasing thresholds (Figs. 10a,c,e). Also, the differences among the microphysics schemes' simulations are greater at larger thresholds, but this is likely due to the variability in the number of events exceeding the thresholds among the different schemes, which is investigated next.

Overall, MYDM and MYTM performed the best at the 40- and 50-dBZ thresholds, which are the most impactful levels, and MYSM consistently performed the worst at this verification metric. As compared to the observed base rates, all of the microphysics schemes, especially the MY schemes, yield larger forecast base rates for each threshold (Fig. 11). However, LIN3's base rates have the smallest biases, so this success likely results in the best FSSs for the larger scales at 30 and 40 dBZ. Conversely, MYSM's base rates are the largest for each threshold, which likely contributes to the poor skill relative to the other microphysics schemes. Cross sections of a simulated storm's reflectivity (Fig. 12) and mixing ratios reveal that the overforecasting of reflectivity by MYSM is mostly due to large values of snow collocated with graupel in the forward flank of the storm (Fig. 13). MYDM and MYTM do not share this result, but those two schemes do advect hail and graupel farther downwind than the other schemes, as previous studies have noted (e.g., Dawson et al. 2010, 2015; Wainwright et al. 2014). Beyond that, the advected hail and graupel fields appear to be contributing to larger areas of light to moderate rain as they fall and melt. Even with larger forecast base rates than LIN3 and WSM6, MYDM and MYTM still exhibit the most forecast skill at the smaller scales at the 40- and 50-dBZ thresholds, which are generally associated with the most impactful weather when deep, moist convection is present.

c. *Object-based results*

Supercell mesocyclones are discontinuous in space, so simulated mesocyclones can be treated as objects using the UH field for verification purposes. The thresholds detailed in section 3 define what constitutes a simulated mesocyclone here; the number of simulated mesocyclones can be greatly altered by adjusting the thresholds. For this study, fairly low thresholds are utilized to potentially produce more robust results and to agree with the visual assessment of what constitutes a mesocyclone. However, the thresholds have minimal impact on the overall scores and comparisons (not shown). As stated before, 0–1UH centers are used to represent low-level mesocyclones and are filtered by requiring pairing with a 1–6-km UH center to ensure that each valid low-level feature has support from a midlevel mesocyclone.

Using all simulations' 5-min interval output, several hundred 0–1UH centers were identified as objects for each microphysics scheme (Fig. 14a). Both S1 and S2 have numerous 0–1UH centers near their tornado points, but S3 has substantially fewer 0–1UH centers nearby, which indicates relatively poor detection of S3's low-level circulation by the forecast system. The 0–1UH centers generally surround S1's tornado points, but are largely clustered to the north of S2's tornado locations. All of the simulations in this experiment produce a similar number of 0–1UH centers, with LIN3 yielding the most and MYDM yielding the fewest (Fig. 14b). Based on the estimated tornado points, the total number of 0–1UH centers should potentially be 219 (119 for S1, 45 for S2, and 32 for S3). However, the total number of 0–1UH centers for each microphysics scheme is 4–5 times that value. This difference is likely a result of the combination of the low thresholds used in defining UH objects, not accounting for observed mesocyclones that did not produce a tornado (e.g., southern Oklahoma; see Fig. 1), and some microphysics schemes

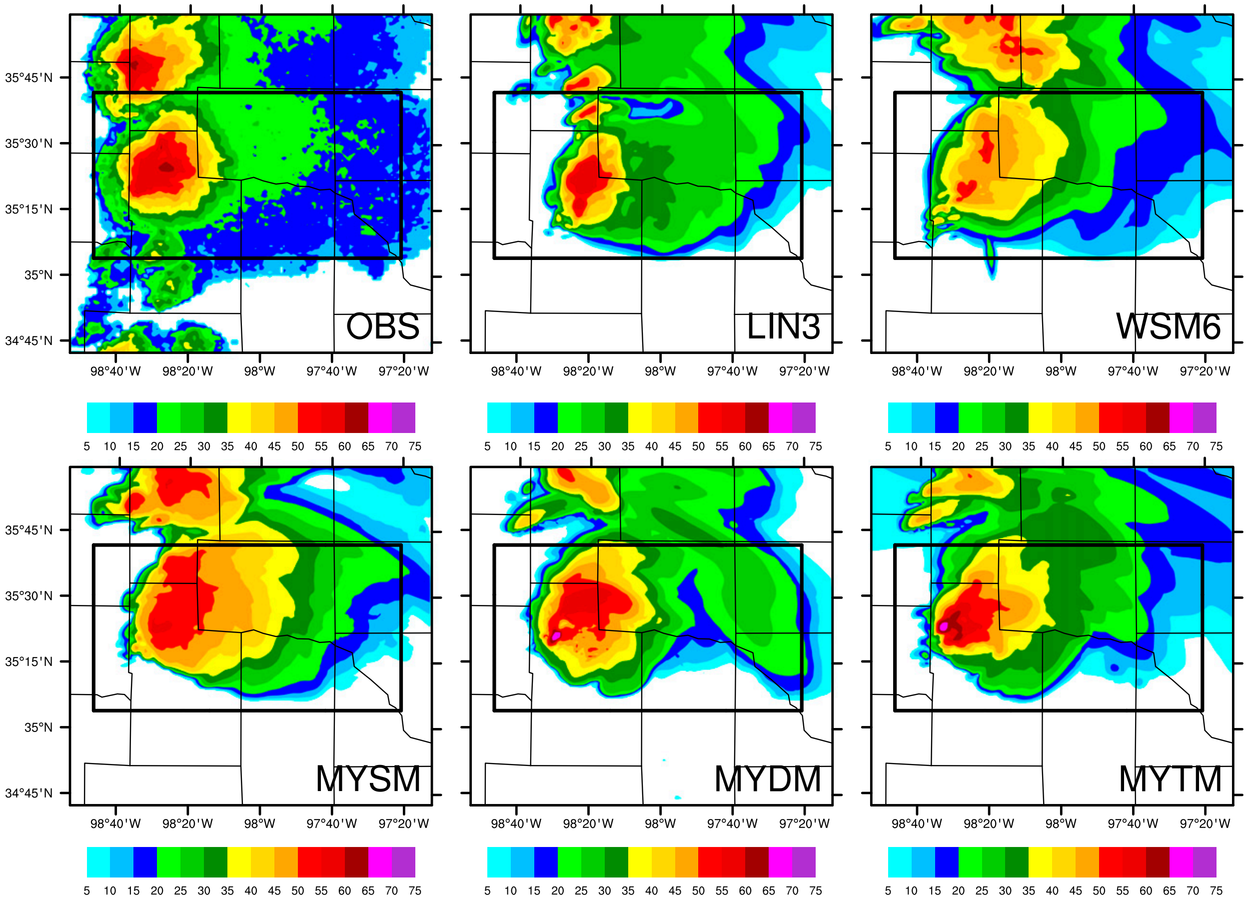


FIG. 12. Observed (OBS) mosaic composite reflectivity (dBZ) for 2030 UTC and 90-min forecasts (LIN3, WSM6, MYSM, MYDM, and MYTM) of composite reflectivity (dBZ) from the 1900 UTC simulations. Black box indicates the portion of the domain used in Fig. 13.

leading to overforecasting of 0–1UH centers (more likely for LIN3 and WSM6). The simulations initialized at 2100 UTC produce the largest number of 0–1UH centers with ~150–200 centers (Fig. 14b). MYDM has the largest average maximum 0–1UH for each center, and WSM6 has the smallest average maximum 0–1UH for each center (Fig. 14c). For most initialization times, the MY schemes, especially for MYDM and MYTM, have distinctly larger average maximum 0–1UH values for each center than LIN3 and WSM6, which are very similar in this respect. These 0–1UH intensity differences are probably related to the cold pool strengths, where colder cold pools inhibit more intense low-level circulations. This finding is supported by numerous previous studies (e.g., Markowski and Richardson 2014; Dawson et al. 2015). To briefly summarize, the MY schemes produce stronger and fewer 0–1UH centers than LIN3 and WSM6.

1) STORM 1

As seen previously, numerous 0–1UH centers are generally close to S1's tornado locations with many centers within 10 km. This demonstrates that the use of a computationally efficient 3DVAR and IAU system, as CAPS is running in real time today, can produce remarkably accurate predictions of an isolated supercell. Digging into the details, LIN3 and WSM6 lead to same-time 0–1UH centers that tend to be slightly fast and somewhat south of the tornado locations (Fig. 15a). The MY schemes have same-time 0–1UH centers that are generally much closer to the tornado locations and each other (Fig. 15a). For anytime 0–1UH centers, all schemes contribute to a similar number of centers, which are evenly spread north and south of S1's tornado locations (Fig. 15b).

The MY schemes have average ST distance errors less than 11 km for 0–1UH, but LIN3 and WSM6 have average ST distance errors 5–10 km larger (Fig. 15c). The ST distance errors for each microphysics scheme are fairly consistent across all simulations. It is worth noting that for the 1900 UTC simulations all of the microphysical schemes, except for LIN3, have average ST distance errors around or less than 5 km, which is

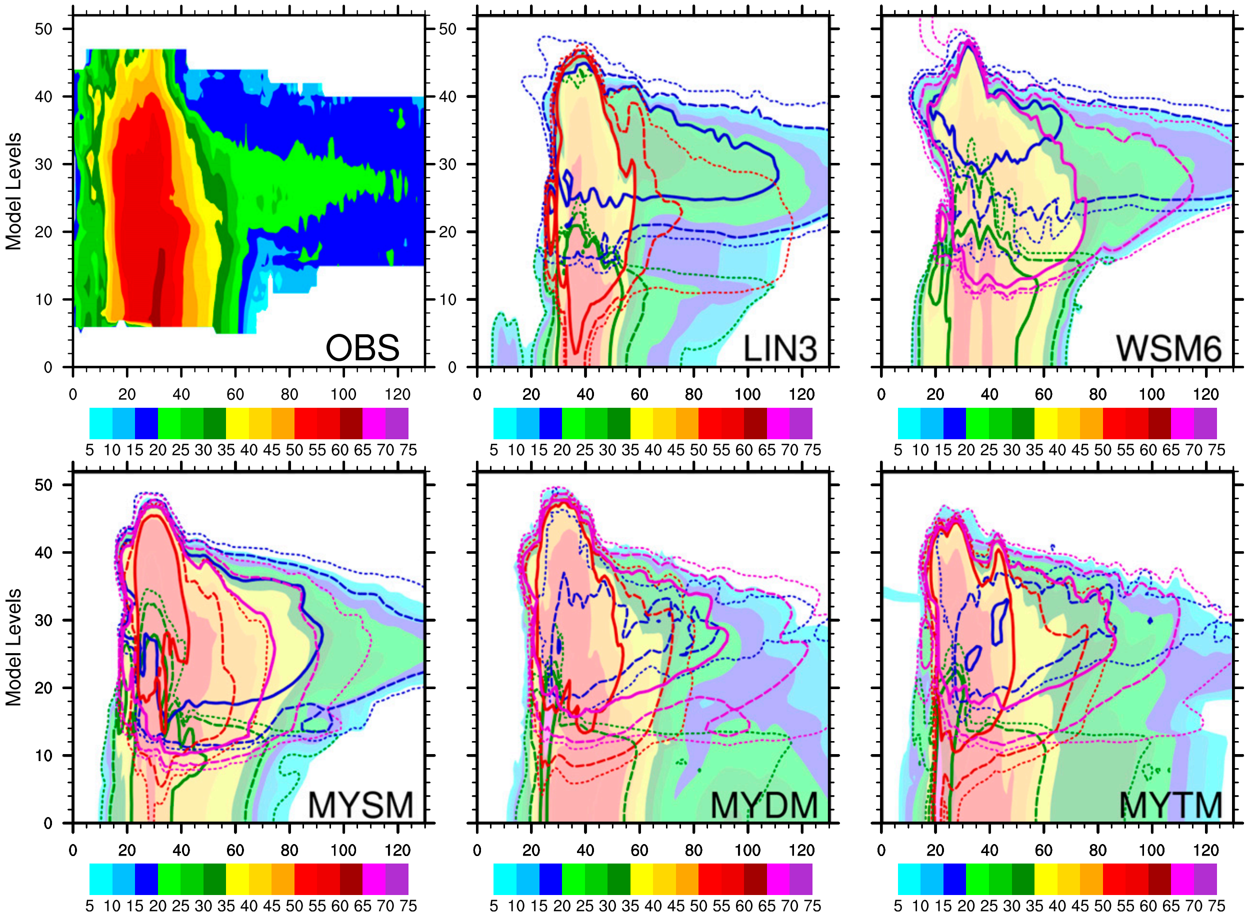


FIG. 13. East–west cross sections of the maximum reflectivity (dB*Z*; color shading) existing between the top and bottom of the black boxes in Fig. 12. For the simulations, rain (green), snow (blue), hail (red), and graupel (magenta) mixing ratios are contoured at 0.01 (short dashed lines), 0.1 (long dashed lines), and 1.0 (solid lines) g kg^{-1}.

excellent given the nearly 90-min lead time. As an example, the simulated reflectivity and 0–1UH fields from the MYDM run are plotted in Fig. 16 and depict very successful forecasts of S1 at 1- and 2-h lead times. The differences in the AT distance errors among the five schemes are smaller those for the ST distance errors, and the average AT distance errors are less than 9 km for all of the schemes (Fig. 15d). Even though LIN3 and WSM6 exhibit AT distance errors similar to the MY schemes, those schemes' AT timing errors are on average mostly 15–20 min too fast (Fig. 15e). The MY schemes have substantially smaller AT timing errors with most 0–1UH centers occurring within ±10 min (i.e., too fast and too slow) of the estimated tornado times.

Overall for S1, all of the schemes contributed to very successful forecasts of low-level circulations near S1's tornado points. The MY schemes tend to produce the smallest distance and timing errors, and WSM6 generally has the largest errors with the LIN3 results being not too different. The largest issue with LIN3 and WSM6 is their tendency to produce 0–1UH centers that move too fast. As previous studies (e.g., Dawson et al. 2010, 2015) and analysis of the near-surface variables have suggested, this is likely due to those schemes producing cold pools that are too cold and thus inclined to advancing too quickly. Dawson et al. (2015) attributed the stronger cold pools to the excessive evaporation of rain in the downdrafts resulting in stronger, deeper downdrafts.

2) STORM 2

While the maximum number of possible 0–1UH centers is less for S2 than S1 (i.e., shorter-lived tornado), a plethora of 0–1UH centers still exist near S2's tornado points for all schemes. Given that a 1-km model is trying to accurately forecast the locations of tornadic circulations, these forecasts are considered successful. The MY schemes' ST 0–1UH centers are generally 5–10 km north

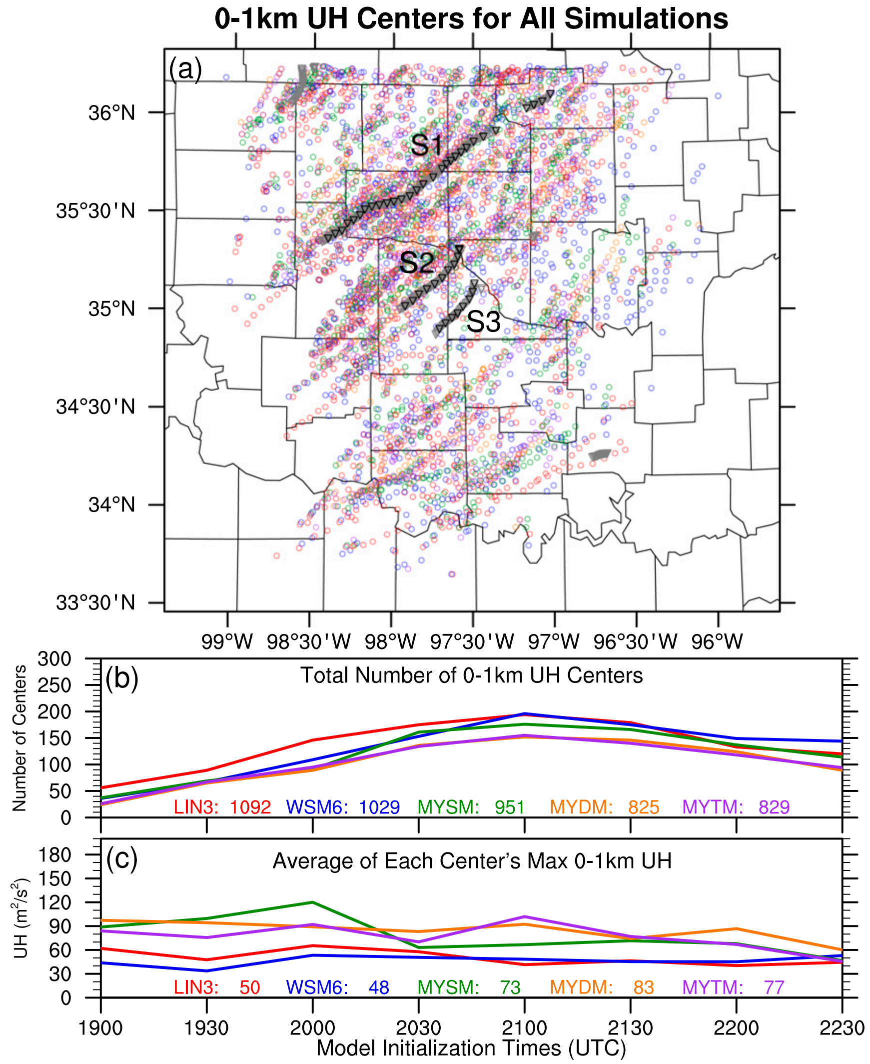


FIG. 14. (a) Filtered 0–1UH centers for LIN3 (red), WSM6 (blue), MYSM (green), MYDM (orange), and MYTM (purple) for all simulations. Gray upside-down triangles represent tornado locations every minute, while black upside-down triangles represent tornado locations occurring at the same times as the forecast output. (b) Line graph of the total number of 0–1UH centers for all simulations with the total number of centers across all simulations annotated along the bottom of the plot. (c) Line graph of the average of each center's maximum 0–1UH ($m^2 s^{-2}$) for each simulation with the overall average of each center's maximum 0–1UH annotated along the bottom of the plot.

of the tornado locations (Fig. 17a). Conversely, WSM6's 0–1UH centers are too fast and too far south, while LIN3's 0–1UH centers are mostly just too fast. For anytime centers, the majority of the 0–1UH centers occur near the tornado locations with a bias to the north, except for WSM6, which also has the smallest number of centers (Fig. 17b). The earlier initiated sets of simulations exhibit more variability in ST distance errors among the different microphysics schemes than the later simulation runs (Fig. 17c). Average ST distance errors for 0–1UH centers range from about 15–20 km for the MY schemes to ~29 km for LIN3 and WSM6. The 0–1UH AT distance errors are substantially smaller than the ST distance errors with average errors less than 15 km for all schemes, and the differences among the schemes are smaller, as well (Fig. 17d). However, timing errors range from ~14 min with MYDM and MYTM, which tend to be too slow, to nearly 30 min with LIN3, which is generally too fast (Fig. 17e).

Overall for S2, the MY schemes mostly outperform LIN3 and WSM6, with mostly smaller distance and timing errors. As mentioned before, there are fewer

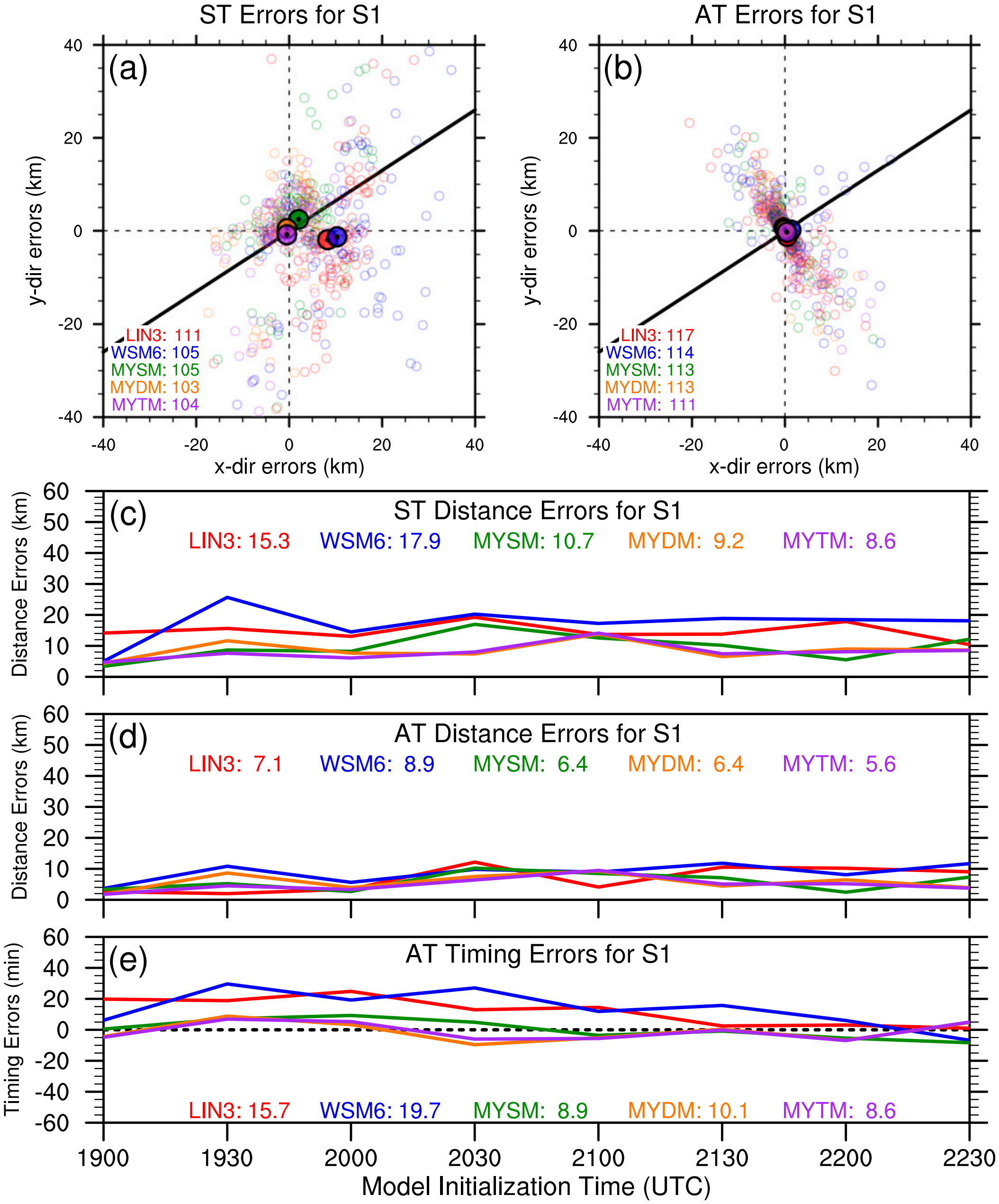


FIG. 15. (a) ST and (b) AT 0–1UH centers for S1 from all simulations with LIN3 (red), WSM6 (blue), MYSM (green), MYDM (orange), and MYTM (purple). Larger colored-filled circles represent the average locations of the microphysical schemes' 0–1UH centers. The number of 0–1UH centers within the plot window are annotated in (a) and (b). Line graphs show the average (c) ST distance errors (km), (d) AT distance errors (km), and (e) AT timing errors (min) for all simulations' 0–1UH centers. The respective averages across all simulations for each microphysics scheme are annotated within the plots.

0–1UH centers near S2 than near S1 even though S2 developed within the CASA radar network. However, S1 was, at least initially, a more isolated storm than S2, which in reality was influenced by storm interactions and mergers. This complication likely led to larger distance and timing errors, but the forecast system still managed to produce successful forecasts of low-level circulations.

3) STORM 3

While the microphysically diverse set of simulations had good success with forecasting 0–1UH centers for

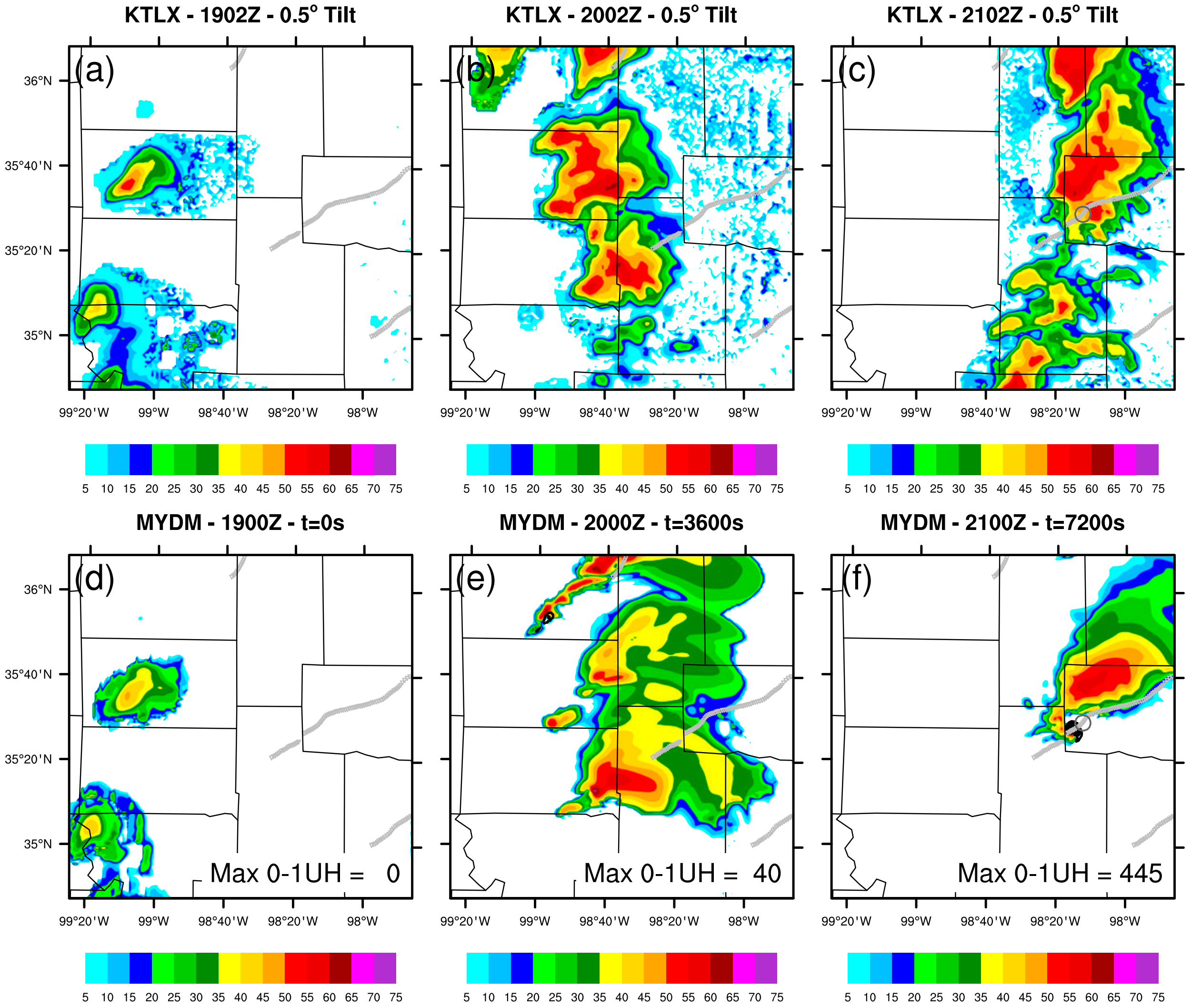


FIG. 16. Observed 0.5°-tilt reflectivity (dBZ) from KTLX for (a) 1902, (b) 2002, and (c) 2102 UTC are interpolated and plotted onto the 1-km model grid. Forecasts of 1 km AGL simulated reflectivity (dBZ; colored) from the 1900 UTC MYDM simulation are plotted for (d) 1900 UTC ($t = 0$ s), (e) 2000 UTC ($t = 3600$ s), and (f) 2100 UTC ($t = 7200$ s). The 0–1UH results are contoured in black in (d)–(f) from 10 to 210 $m^2 s^{-2}$ with an interval of 25 $m^2 s^{-2}$. The maximum UH value ($m^2 s^{-2}$) in each plot window is annotated near the bottom of each plot. Light gray upside-down triangles depict the estimated tornado locations every minute, and darker gray circles in (c) and (f) indicate the estimated location of the tornado occurring at 2100 UTC.

S1 and S2, most of the simulations struggle with forecasting 0–1UH centers near S3's tornado locations, as depicted in Fig. 14. Few ST 0–1UH centers exist within 10 km of S3's tornado points, and average 0–1UH centers indicate a southeast bias due to simulated mesocyclones existing ~30 km to the southeast of S3 (Fig. 18a). This bias is even more evident when timing differences are ignored (Fig. 18b). However, these simulated mesocyclones are not spurious since another supercell was observed propagating through the same area (see Fig. 1). Not surprisingly, ST and AT distance errors for 0–1UH are generally greater than 20 km for all microphysics schemes (Figs. 18c,d). All simulations produce 0–1UH centers that are mostly within 20 min of the tornado occurrence times, and WSM6 is the only scheme that is mostly too fast (Fig. 18e).

Since very few 0–1UH centers were forecasted near S3, the results above are not too meaningful except that the forecast system as a whole struggled with S3 even though the storm developed and propagated through the CASA radar network. Other model configuration improvements, besides using more advanced microphysics schemes, need to be explored to improve the forecasts of S3, which was directly affected by storm interactions and mergers.

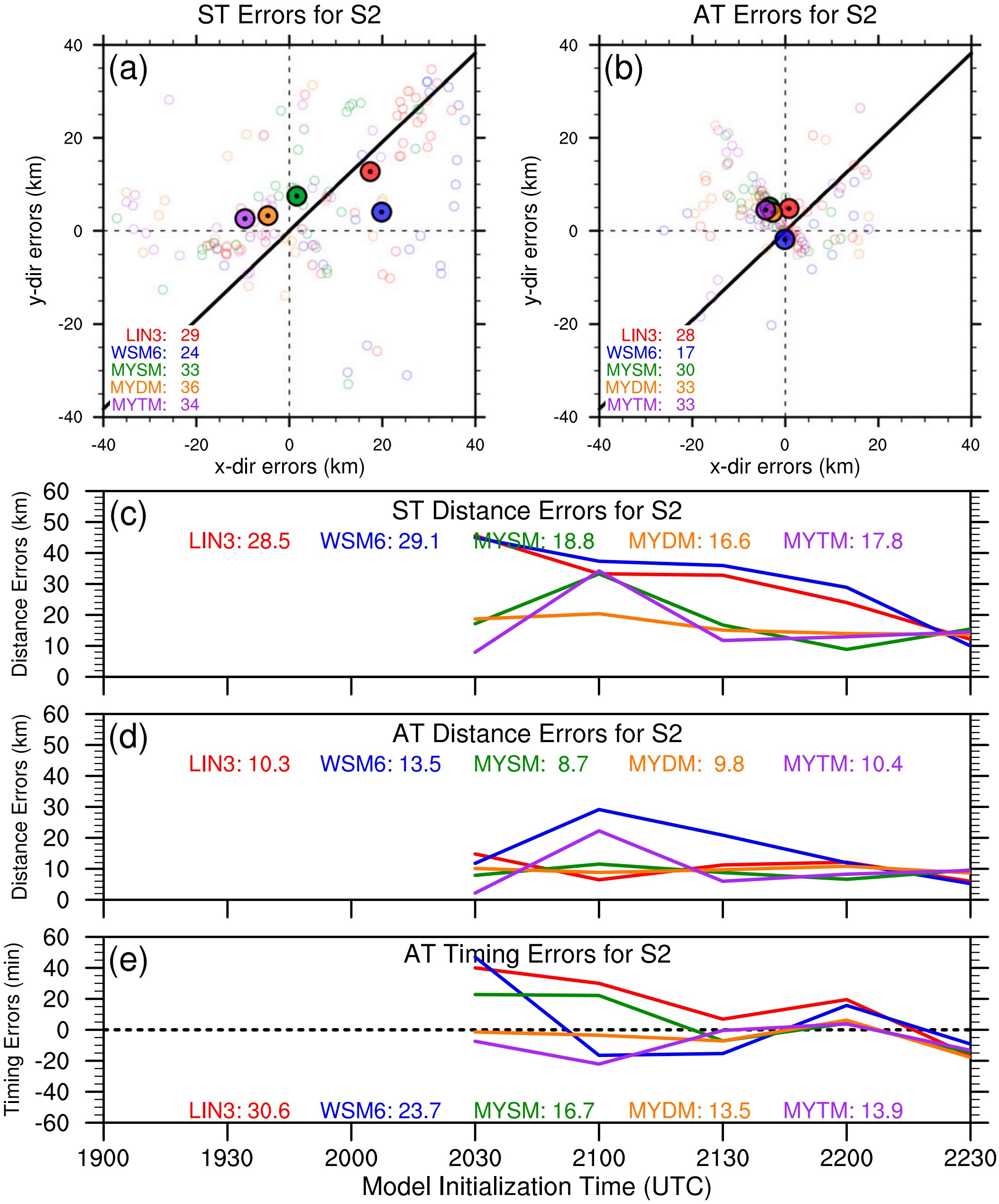


FIG. 17. As in Fig. 15, but for S2.

6. Summary and discussion

On 24 May 2011, an outbreak of tornadic supercells propagated across central Oklahoma producing several strong to violent tornadoes. Unfortunately, the tornadoes and supercells caused numerous deaths and injuries along with widespread damage. However, the extensive observation network across central Oklahoma during the spring of 2011 provides an opportunity for forecast system sensitivity studies, such as this one, to explore ways to potentially improve short-term, storm-scale forecasts of severe convection. Data from numerous observation sources in the area were used in the data assimilation process, which consisted of using 3DVAR with ADAS complex cloud analysis, to produce model state variable increments, which were introduced during an IAU window, for an ARPS 1-km model. In conjunction with this study, the ADAS complex cloud analysis (Brewster and Stratman 2015) and IAU (Brewster et al. 2015; Brewster and Stratman

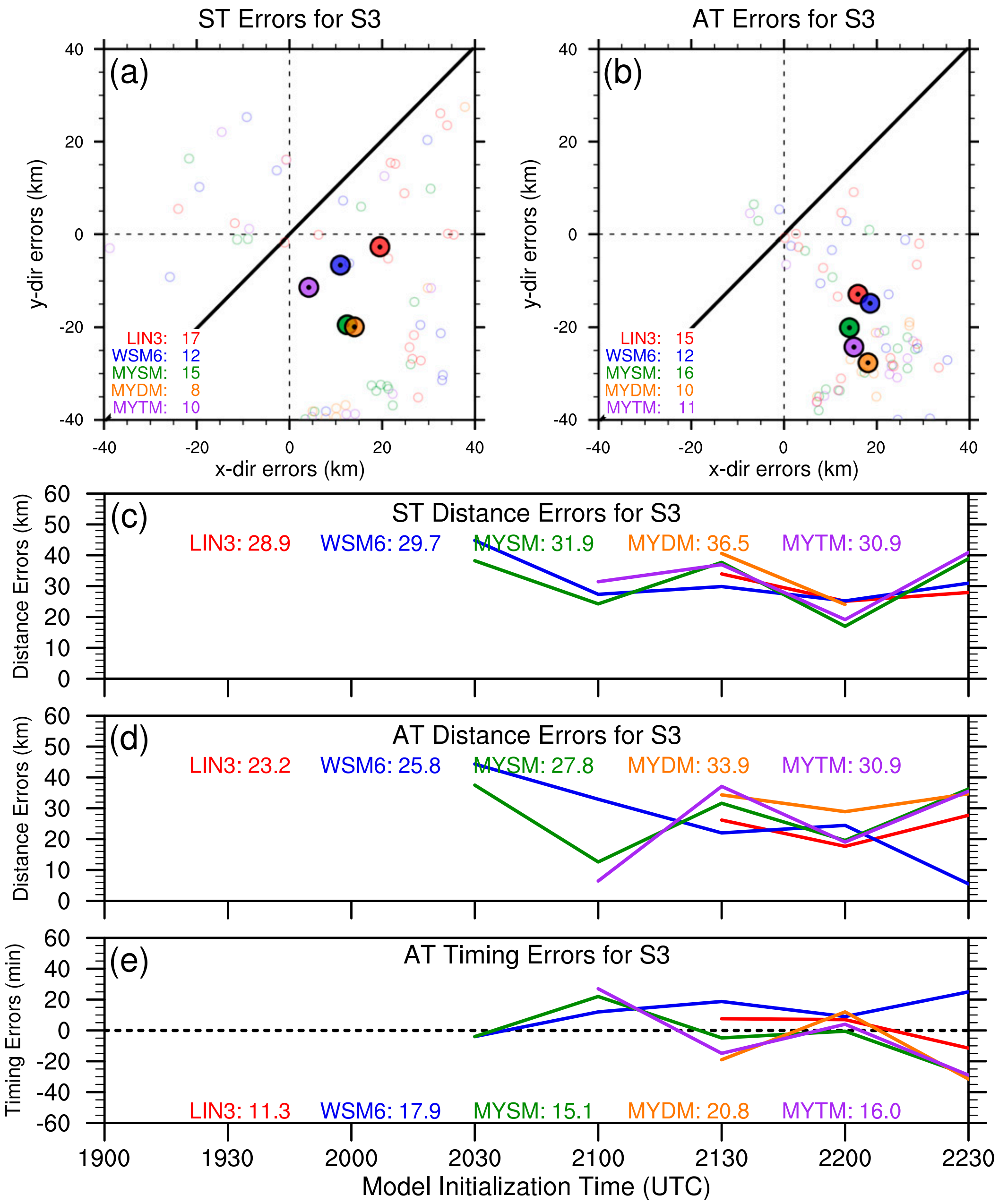
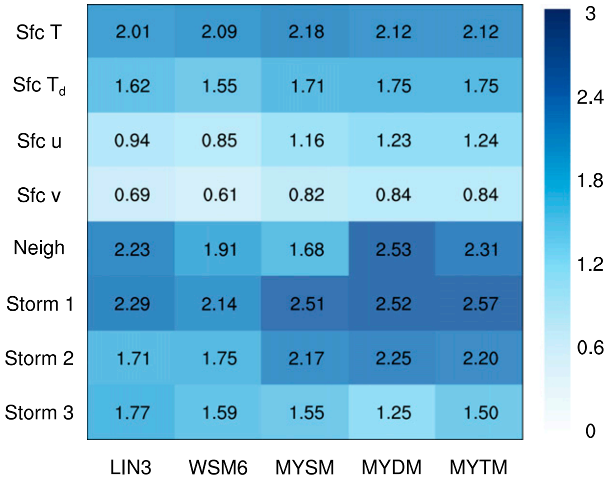


FIG. 18. As in Fig. 15, but for S3.

2016) packages within the ARPS system were both updated.

To emulate a real-time, continuous storm-scale forecasting system, eight microphysically diverse sets of simulations were initialized every 30 min starting at 1900 UTC and integrated out to 2 h. This forecast framework allowed for in-depth examinations of sensitivities related to microphysics parameterization schemes. Simulations with different microphysics schemes (i.e., LIN3, WSM6, MYSM, MYDM, and MYTM) were compared to assess the characteristics of each scheme and to determine which scheme leads to better forecast results. Observation-point, neighborhood, and object-based verifications techniques were utilized to quantitatively evaluate the forecasts.

Prior studies investigating the differences between single-moment and multimoment microphysics schemes' impacts on forecasts of supercells (Dawson et al. 2010, 2015; Yussouf et al. 2013) and MCSs (Wheatley et al. 2014) have generally found that the multimoment microphysics



	LIN3	WSM6	MYSM	MYDM	MYTM
Sfc T	2.01	2.09	2.18	2.12	2.12
Sfc T_d	1.62	1.55	1.71	1.75	1.75
Sfc u	0.94	0.85	1.16	1.23	1.24
Sfc v	0.69	0.61	0.82	0.84	0.84
Neigh	2.23	1.91	1.68	2.53	2.31
Storm 1	2.29	2.14	2.51	2.52	2.57
Storm 2	1.71	1.75	2.17	2.25	2.20
Storm 3	1.77	1.59	1.55	1.25	1.50

FIG. 19. Summary of the forecast skill of each microphysics scheme using summed values (annotated and shaded in blue based on the value) of the various normalized verification metrics.

schemes yield better forecasts of storm intensity, structure, and evolution. Hence, the expectation was that MYDM and MYTM would outperform LIN3, WSM6, and MYSM. To aggregate and simplify all of the verification results, unity-based normalization is first used to normalize all of the verification metrics' average values for each microphysics schemes' simulations onto a scale from 0 to 1, with a value of 1 indicating a perfect forecast (Fig. 19). The perfect skill (no skill) values used for this normalization include 5 (0) for RMSE, 1 (0) for slopes and R^2 values, 1 (100) for FSS − $\text{FSS}_{\text{useful}}$ = 0 scales, 0 (56.6) for the distance errors, and 0 (60) for the timing errors. Next, each verification group's three normalized values for each microphysics scheme are summed together to give a value between 0 and 3, with 3 indicating perfect forecast skill. For the near-surface variables, the normalized values for RMSE, slope, and R^2 are summed together. For the neighborhood metric, the normalized values for the FSS − $\text{FSS}_{\text{useful}}$ = 0 scales at the three reflectivity thresholds are summed together. For the storm location and timing errors, the normalized values of the ST and AT distances and timing errors are summed. Comparisons can only be made among the different microphysics schemes using variables with the same units (e.g., the normalized values of the neighborhood and object-based metrics cannot be compared with each other).

For near-surface temperatures, it could be argued that MYSM outperformed MYDM and MYTM as a result of having slopes closer to 1 (Fig. 19). However, MYDM and MYTM have smaller RMSEs and larger R^2 values, so MYSM's success over MYDM and MYTM is limited. For the other near-surface variables, MYDM and MYTM do show improvement over the single-moment schemes, but it is worth noting that MYSM's forecasts of near-surface variables end up closer to the MYDM and MYTM forecasts than the LIN3 and WSM6 forecasts, which tend to have colder and larger cold pools (see Fig. 8). By these metrics, it appears the MY schemes, as a group, outperformed LIN3 and WSM6.

As a result of having the largest forecasted base rates for composite reflectivity (see Fig. 11), MYSM exhibited the least ''useful'' skill among the five microphysics schemes (Fig. 19). LIN3 has only a slightly smaller overall score than MYTM for the neighborhood-based verification technique, likely because of MYTM's larger forecasted base rates, but MYTM generally depicted more skill at the 40- and 50-dB*Z* thresholds at the smaller scales. Even with commonly higher forecasted base rates than LIN3 and WSM6 (but less than MYSM and MYTM), MYDM usually exhibited the most skill at all three reflectivity thresholds.

In regard to the 0–1UH centers, the MY schemes resulted in substantially fewer and stronger 0–1UH centers than LIN3 and WSM6, so in this respect, no inference can be made that separates the single-moment microphysics schemes from the multimoment microphysics schemes. For S1, the MY schemes predominately performed similarly and largely outperformed LIN3 and WSM6, but overall, the entire set of simulations had remarkably good success at forecasting S1's circulation locations (Fig. 19). Even though the simulations had somewhat less success at forecasting S2's circulation locations, the MY schemes still substantially outperformed LIN3 and WSM6 (Fig. 19). Conversely, all of the microphysics schemes using the base data assimilation scheme struggled with forecasting the more complex S3 event and mostly yielded somewhat similar errors (Fig. 19).

Overall, the simulations using the MY schemes outperformed the simulations using LIN3 and WSM6, so based on this finding, the differences between the various types of microphysics schemes (i.e., LIN3 and WSM6 versus the MY schemes) are larger than the differences between single-moment and multimoment microphysics schemes. Even though the MY schemes exhibited similar abilities to forecast the tornado locations from the three storms of interest, the multimoment MY schemes do tend to better forecast the near-surface variables and simulated reflectivity coverage. Additionally, the computational cost of the MY schemes is almost double that of LIN3 and WSM6, and while not as large, the differences in computational cost among the MY schemes are also noticeable (Fig. 20). All of these factors must be taken into account when designing a model configuration best suited for storm-scale,

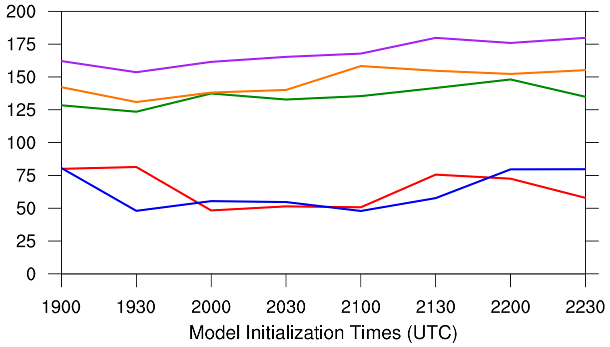


FIG. 20. Line graph of the total CPU time (core hour) used for each LIN3 (red), WSM6 (blue), MYSM (green), MYDM (orange), and MYTM (purple) simulation. Note that for these experiments, 50 cores of Intel Xeon "Sandy Bridge" processors were used, so 100 core hours represent approximately 2 h of wall-clock time. The forecasts are generally scalable by core. With the current ARPS real-time system using 200 cores, LIN3 and WSM6 would have wall-clock times of ~15–25 min for the 2-h runs, and the MY schemes would need ~35–55 min of wall-clock time to complete the same 2-h forecast.

short-term forecasts of severe convection within the confines of computational resources and the desired low latency.

Even though meaningful results were revealed in this study using just one case, the lack of knowledge of the statistical significance is a major caveat. To potentially address this concern, additional ARPS 1-km simulations of tornadic supercell events, such as the 26 December 2015 tornadoes near Dallas, Texas, using the same data assimilation techniques and model configurations could be executed to fully understand the microphysics sensitivities tested in this experiment. Also, ensemble forecasts, which would use initial analyses produced by either a 3DVAR or ensemble Kalman filters (EnKF) based system, could be utilized to determine statistical significance. Although the five microphysics schemes used in this study provided a decent spread of solutions, other single-moment and multimoment schemes (e.g., Thompson microphysics scheme; Thompson et al. 2008) currently not in the ARPS package should be added and tested alongside the current set of microphysics schemes.

The object-based verification technique utilized in this study to determine distance and timing errors proved to be highly effective in discerning the model's skill at forecasting circulation locations. However, perhaps other observed or analyzed objects, such as radar-detected rotation tracks (Lakshmanan et al. 2013) or analysis locations of low-level vorticity or UH, can be used instead of the estimated tornado locations to streamline the verification process. Furthermore, the object-based verification technique can be employed to verify, for example, forecasted hail and graupel mixing ratios with radar-detected Z_{DR} columns (Kumjian and Ryzhkov 2008). With respect to the neighborhood-based verification technique, most of the simulations tended to substantially overforecast the composite reflectivity field, which is likely due to both the microphysics scheme and the relatively large 1-km grid spacing (i.e., coarser than reality), so in addition to looking at cross sections of reflectivity at various heights (Figs. 12 and 13), using percentiles or other precipitation fields, such as precipitation rates, could be explored in determining skill despite the overforecasting bias. However, an overmoistening in downdrafts in the cloud analysis scheme used in this work has recently been identified, and modifications to the cloud analysis scheme to address this issue have been made. These may reduce the reflectivity overforecast bias when the model is initialized with ongoing convection.

This study highlights the ability of a 1-km forecast system to successfully predict potentially tornadic supercells on short time scales using increments derived from analyses produced by the 3DVAR and ADAS complex cloud analysis package, so when considering the computational cost of real-time data assimilation systems, less computationally intensive systems, such as the one used in this study, should be considered alongside more computationally expensive systems, such as EnKFs (Evensen 1994). A study comparing both data assimilation systems is needed to better assess any potential benefits each data assimilation system may provide to short-term, storm-scale forecasts of severe convection.

Acknowledgments. The authors thank Ming Xue, Fred Carr, and Michael Richman for their discussions and suggestions. The authors also thank the three anonymous reviewers for providing comments and suggestions that greatly improved this paper. Finally, the authors thank everyone else who provided feedback on this research during seminars and scientific meetings. This work is supported by the National Science Foundation via CASA ERC Grant EEC 03-13747 and by the National Oceanic and Atmospheric Administration via Grant NA11OAR4320072, and related support provided by the University of Oklahoma (OU). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Most of the computing for this project was performed at the OU Supercomputing Center for Education and Research (OSCER). This work utilized data from the Oklahoma Mesonet provided by the Oklahoma Climatological Survey.

REFERENCES

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271, doi:10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2.

Bluestein, H. B., and M. L. Weisman, 2000: The interaction of numerically simulated supercells initiated along lines. *Mon. Wea. Rev.*, **128**, 3128–3149, doi:10.1175/1520-0493(2000)128<3128:TIONSS>2.0.CO;2.

Brewster, K. A., and D. R. Stratman, 2015: An updated high resolution hydrometeor analysis system using radar and other data. *27th Conf. on Weather Analysis and Forecasting/23rd Conf. on Numerical Weather Prediction*, Chicago, IL, Amer. Meteor. Soc., 31. [Available online at https://ams.confex.com/ams/27WAF23NWP/webprogram/Paper273811.html.]

——, and ——, 2016: Tuning an analysis and incremental analysis updating assimilation for an efficient high resolution forecast system. *20th Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, New Orleans, LA, Amer. Meteor. Soc., 10.6. [Available online at https://ams.confex.com/ams/96Annual/webprogram/Paper289235.html.]

——, K. W. Thomas, J. Gao, J. Brotzge, M. Xue, and Y. Wang, 2010: A nowcasting system using full physics numerical weather prediction initialized with CASA and NEXRAD radar data. *25th Conf. on Severe Local Storms*, Denver, CO, Amer. Meteor. Soc., 9.4. [Available online at https://ams.confex.com/ams/25SLS/techprogram/paper_176053.htm.]

——, F. H. Carr, K. W. Thomas, and D. R. Stratman, 2015: Utilizing heterogeneous radar systems in a real-time high resolution analysis and short-term forecast system in the Dallas/Fort Worth Testbed. *37th Conf. on Radar Meteorology*, Norman, OK, Amer. Meteor. Soc., 14A.3. [Available online at https://ams.confex.com/ams/37RADAR/webprogram/Paper276230.html.]

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.*, **12**, 5–19, doi:10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2.

Brotzge, J. A., and C. M. Luttrell, 2015: Genesis of the Chickasha, Oklahoma, tornado on 24 May 2011 as observed by CASA radar and Oklahoma Mesonet. *J. Oper. Meteor.*, **3**, 59–69, doi:10.15191/nwajom.2015.0306.

Chou, M.-D., 1990: Parameterizations for the absorption of solar radiation by O_2 and CO_2 with application to climate studies. *J. Climate*, **3**, 209–217, doi:10.1175/1520-0442(1990)003<0209:PFTAOS>2.0.CO;2.

——, 1992: A solar radiation model for use in climate studies. *J. Atmos. Sci.*, **49**, 762–772, doi:10.1175/1520-0469(1992)049<0762:ASRMFU>2.0.CO;2.

——, and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, Vol. 3, 85 pp. [Available online at https://permanent.access.gpo.gov/gpo60401/19950009331.pdf.]

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

Davis, C., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methods and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, doi:10.1175/MWR3145.1.

——, ——, and ——, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.*, **134**, 1785–1795, doi:10.1175/MWR3146.1.

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, doi:10.1175/2009MWR2956.1.

——, ——, ——, and A. Shapiro, 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, doi:10.1175/MWR-D-14-00279.1.

Dial, G. L., and J. P. Racy, 2004: Forecasting short term convective mode and evolution for severe storms initiated along synoptic boundaries. *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., 11A.2. [Available online at https://ams.confex.com/ams/11aram22sls/techprogram/paper_81495.htm.]

Ebert, E. E., 2008: Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework. *Meteor. Appl.*, **15**, 51–64, doi:10.1002/met.25.

——, 2009: Neighborhood verification: A strategy for rewarding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, doi:10.1175/2009WAF2222251.1.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, doi:10.1029/94JC00572.

Fierro, A. O., E. R. Mansell, C. L. Ziegler, and D. R. MacGorman, 2012: Application of a lightning data assimilation technique in the WRF-ARW model at cloud-resolving scales for the tornado outbreak of 24 May 2011. *Mon. Wea. Rev.*, **140**, 2609–2627, doi:10.1175/MWR-D-11-00299.1.

French, M. M., P. S. Skinner, L. J. Wicker, and H. B. Bluestein, 2015: Documenting a rare tornado merger observed in the 24 May 2011 El Reno–Piedmont, Oklahoma, supercell. *Mon. Wea. Rev.*, **143**, 3025–3043, doi:10.1175/MWR-D-14-00349.1.

Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. *J. Atmos. Oceanic Technol.*, **21**, 457–469, doi:10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2.

Heymsfield, G. M., L. Tian, L. Li, M. McLinden, and J. I. Cervantes, 2013: Airborne radar observations of severe hailstorms: Implications for future spaceborne radar. *J. Appl. Meteor. Climatol.*, **52**, 1851–1867, doi:10.1175/JAMC-D-12-0144.1.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Houser, J. H., H. B. Bluestein, and J. C. Snyder, 2015: Rapid-scan, polarimetric, Doppler radar observations of tornadogenesis and tornado dissipation in a tornadic supercell: The "El Reno, Oklahoma" storm of 24 May 2011. *Mon. Wea. Rev.*, **143**, 2685–2710, doi:10.1175/MWR-D-14-00253.1.

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

——, ——, J. Gao, and K. Brewster, 2006b: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: Impact of radial velocity analysis via 3DVAR. *Mon. Wea. Rev.*, **134**, 699–721, doi:10.1175/MWR3093.1.

Jones, T. A., D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165–194, doi:10.1175/MWR-D-14-00180.1.

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, doi:10.1175/WAF2007106.1.

Kumjian, M. R., and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47**, 1940–1961, doi:10.1175/2007JAMC1874.1.

Lakshmanan, V., M. Miller, and T. Smith, 2013: Quality control of accumulated fields by applying spatial and temporal constraints. *J. Atmos. Oceanic Technol.*, **30**, 745–758, doi:10.1175/JTECH-D-12-00128.1.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

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

——, C. Hannon, J. Frame, E. Lancaster, A. Pietrycha, R. Edwards, and R. L. Thompson, 2003: Characteristics of vertical wind profiles near supercells obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1262–1272, doi:10.1175/1520-0434(2003)018<1262:COVWPN>2.0.CO;2.

McLaughlin, D., and Coauthors, 2009: Short-wavelength technology and the potential for distributed networks of small radar systems. *Bull. Amer. Meteor. Soc.*, **90**, 1797–1817, doi:10.1175/2009BAMS2507.1.

McPherson, R. A., and Coauthors, 2007: Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.*, **24**, 301–321, doi:10.1175/JTECH1976.1.

Milbrandt, J. A., and M. K. Yau, 2005a: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62**, 3051–3064, doi:10.1175/JAS3534.1.

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

NCDC, 2011: *Storm Data*. Vol. 53, No. 5, 944 pp.

Noilhan, J., and S. Planton, 1989: A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.*, **117**, 536–549, doi:10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2.

NWS, 2016: The tornado outbreak of May 24, 2011. [Available online at www.weather.gov/oun/events-20110524-fastfacts.]

Putnam, B. J., M. Xue, Y. Jung, N. Snook, and G. Zhang, 2014: The analysis and prediction of microphysical states and polarimetric radar variables in a mesoscale convective system using double-moment microphysics, multinetwork radar data, and the ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 141–162, doi:10.1175/MWR-D-13-00042.1.

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

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, doi:10.1175/2007MWR2123.1.

Shapiro, A., S. Rahimi, C. K. Potvin, and L. Orf, 2015: On the use of advection correction in trajectory calculations. *J. Atmos. Sci.*, **72**, 4261–4280, doi:10.1175/JAS-D-15-0095.1.

SPC, 2016: Severe weather event: May 24, 2011. [Available online at http://www.spc.noaa.gov/exper/archive/event.php?date=20110524.]

Stensrud, D. J., 2007: *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*. Cambridge University Press, 459 pp.

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

Tanamachi, R. L., P. L. Heinselman, and L. J. Wicker, 2015: Impacts of a storm merger on the 24 May 2011 El Reno, Oklahoma, tornadic supercell. *Wea. Forecasting*, **30**, 501–524, doi:10.1175/WAF-D-14-00164.1.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, doi:10.1175/2008MWR2387.1.

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, doi:10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2.

Wainwright, C. E., D. T. Dawson II, M. Xue, and G. Zhang, 2014: Diagnosing the intercept parameters of the exponential drop size distributions in a single-moment microphysics scheme and impact on supercell storm simulations. *J. Appl. Meteor. Climatol.*, **53**, 2072–2090, doi:10.1175/JAMC-D-13-0251.1.

Wheatley, D. M., N. Yussouf, and D. J. Stensrud, 2014: Ensemble Kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. *Mon. Wea. Rev.*, **142**, 3243–3263, doi:10.1175/MWR-D-13-00260.1.

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

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

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

——, D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170, doi:10.1007/s00703-001-0595-6.

——, J. Schleif, F. Kong, K. W. Thomas, Y. Wang, and K. Zhu, 2013: Track and intensity forecasting of hurricanes: Impact of convection-permitting resolution and global ensemble Kalman filter analysis on 2010 Atlantic season forecasts. *Wea. Forecasting*, **28**, 1366–1384, doi:10.1175/WAF-D-12-00063.1.

Yussouf, N., E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics schemes. *Mon. Wea. Rev.*, **141**, 3388–3412, doi:10.1175/MWR-D-12-00237.1.

Zalesak, S. T., 1979: Fully multidimensional flux-corrected transport algorithms for fluids. *J. Comput. Phys.*, **31**, 335–362, doi:10.1016/0021-9991(79)90051-2.

Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, doi:10.1175/2011BAMS-D-11-00047.1.