

# Multiscale EnKF Assimilation of Radar and Conventional Observations and Ensemble Forecasting for a Tornadoic Mesoscale Convective System

NATHAN SNOOK

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

MING XUE

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

YOUNGSUN JUNG

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

(Manuscript received 15 August 2013, in final form 26 November 2014)

## ABSTRACT

In recent studies, the authors have successfully demonstrated the ability of an ensemble Kalman filter (EnKF), assimilating real radar observations, to produce skillful analyses and subsequent ensemble-based probabilistic forecasts for a tornadoic mesoscale convective system (MCS) that occurred over Oklahoma and Texas on 9 May 2007. The current study expands upon this prior work, performing experiments for this case on a larger domain using a nested-grid EnKF, which accounts for mesoscale uncertainties through the initial ensemble and lateral boundary condition perturbations. In these new experiments, conventional observations (including surface, wind profiler, and upper-air observations) are assimilated in addition to the WSR-88D and the Center for Collaborative Adaptive Sensing of the Atmosphere (CASA) radar data used in the previous studies, better representing meso- and convective-scale features. The relative impacts of conventional and radar data on analyses and forecasts are examined, and biases within the ensemble are investigated.

The new experiments produce a substantially improved forecast, including better representation of the convective lines of the MCS. Assimilation of radar data substantially improves the ensemble precipitation forecast. Assimilation of conventional data together with radar observations substantially improves the forecast of near-surface mesovortices within the MCS, improves forecasts of surface temperature and dewpoint, and imparts a slight but noticeable improvement to short-term precipitation forecasts. Furthermore, ensemble analyses and forecasts are found to be sensitive to the localization radius applied to conventional data within the EnKF.

## 1. Introduction

The ensemble Kalman filter (EnKF), first developed by Evensen (1994, 2003), has been successfully applied to atmospheric data assimilation (DA) using both simulated and real data from a variety of observation platforms, for models ranging from global to convective storm scales (Houtekamer and Mitchell 1998; Hamill and Snyder 2000; Anderson 2001; Whitaker and Hamill 2002; Snyder and Zhang 2003; Dowell et al. 2004; Zhang

et al. 2004; Dirren et al. 2007; Tong and Xue 2008a; Xue et al. 2010; Dawson et al. 2012; Snook et al. 2011, hereafter SXJ11; Jung et al. 2012; Yussouf and Stensrud 2012; Yussouf et al. 2013). Though EnKF is rather expensive in terms of computation, requiring an ensemble of forecasts (typically using several dozen members), it provides flow-dependent multivariate background error covariances that less computationally intensive three-dimensional variational data assimilation (3DVAR) methods cannot. Cross covariances produced by the EnKF system are very valuable, especially for convective-scale DA, because state variables that are not directly observed can be retrieved (Tong and Xue 2005, 2008a). Further discussion of DA techniques commonly used for assimilation of weather observations,

---

*Corresponding author address:* Nathan Snook, Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David Boren Blvd., Rm. 5425, Norman, OK 73072.  
E-mail: nsnook@ou.edu

including 3DVAR, four-dimensional variational data assimilation (4DVAR) methods, and EnKF can be found in [Tong and Xue \(2005\)](#).

Analysis ensembles generated using EnKF are generally well suited as initial conditions for convective-scale ensemble forecasts. EnKF assimilation of Doppler radar data has proven to be effective in retrieving wind, temperature, and microphysical fields at the convective scale (e.g., [Dowell et al. 2004](#); [Tong 2006](#); [SXJ11](#); [Jung et al. 2012](#); [Putnam et al. 2014](#)). Furthermore, EnKF analyses, in principle, also characterize the analysis uncertainty; this is a particularly desirable quality in the ensemble forecast initial conditions. Forecast ensembles initialized from EnKF analyses have been shown to produce superior probabilistic predictions compared to ensembles initialized using traditional perturbation methods ([Houtekamer et al. 2005](#); [Hamill and Whitaker 2011](#)). EnKF analyses have been successfully applied to ensemble forecasts of convective systems, including supercell thunderstorms (e.g., [Aksoy et al. 2009, 2010](#); [Dawson et al. 2012](#)) and mesoscale convective systems (e.g., [Snook et al. 2012](#), hereafter [SXJ12](#); [Putnam et al. 2014](#)), as well as tropical cyclones (e.g., [Wu et al. 2010](#); [Aksoy et al. 2012, 2013](#)). As available computational power increases, it will become increasingly feasible to run a real-time convective-scale ensemble analysis system (e.g., [Xue et al. 2008](#)) incorporating EnKF DA (e.g., [SXJ12](#)), as envisioned in the “warn-on-forecast” paradigm being developed by the National Weather Service ([Stensrud et al. 2009](#)). For example, the Center for Analysis and Prediction of Storms (CAPS) is actively developing a quasi-operational cycled EnKF DA and ensemble forecast system.

In [SXJ11](#), an ensemble square root Kalman filter (EnSRF; [Whitaker and Hamill 2002](#)) is used together with the Advanced Regional Prediction System (ARPS; [Xue et al. 2000, 2001](#)) atmospheric model ([Tong and Xue 2005](#); [Xue et al. 2006](#); [Tong and Xue 2008b](#)) to assimilate radar reflectivity and radial velocity observations from multiple WSR-88D ([Crum et al. 1993](#)) S-band radars, and from the X-band radars deployed by the Center for Collaborative Adaptive Sensing of the Atmosphere (CASA; [McLaughlin et al. 2009](#)), for a tornadic mesoscale convective system (MCS) that occurred over Texas and Oklahoma on 9 May 2007. The 40-member ensemble mean analysis of [SXJ11](#) produces model storms whose geographic extent, convective mode, and intensity agree well with the radar observations. Furthermore, [SXJ11](#) finds that assimilation of CASA radar data improves the representation of near-surface circulations and cold pool structure. Ensemble forecasts initialized from the ensemble analyses of [SXJ11](#) are subsequently examined in [SXJ12](#). The

forecast ensembles of [SXJ12](#) produce skillful 0–3-h probabilistic forecasts for radar reflectivity and 2-h probabilistic forecasts of the presence and location of the tornadic mesovortex embedded within the MCS with probability maxima localized within several tens of kilometers of the observed tornadic mesovortex. The EnKF experiments of [SXJ11](#) and [SXJ12](#) did not assimilate any conventional observations, nor did they include any mesoscale perturbations in the initial ensemble or any perturbations to the lateral boundary conditions. The ensemble forecasts of [SXJ12](#) exhibited substantial high biases in heavy precipitation, as well as the development of spurious convection near the CASA radar network later in the forecast period. [SXJ12](#) also noted that the trailing convective line of the MCS dissipated too quickly near the southern model domain boundary in their forecast ensembles.

This study builds upon and extends the work of [SXJ11](#) and [SXJ12](#), addressing the shortcomings of those studies through an improved ensemble DA and forecast framework. The new experiments investigate the assimilation of both radar and conventional observations, including surface observations at 5-min intervals from the Oklahoma Mesonet. In their analysis and forecast study of the same case, [Schenkman et al. \(2011b\)](#) found that assimilating Oklahoma Mesonet observations via 3DVAR significantly improved the near-surface wind field within the model. Furthermore, the current study uses lateral boundary conditions that include mesoscale perturbations on the outer grid; the boundary conditions for the inner nest are interpolated from the outer-nest ensemble members. Recent studies (e.g., [Jung et al. 2012](#); [Yussouf et al. 2013](#)) have shown promising results for storm-scale data assimilation using similar ensemble designs. The geographic extent of inner-nest domain is also doubled in both horizontal directions compared to [SXJ12](#), reducing the potential negative impacts of boundary conditions.

This study will examine the relative and combined impacts of radar and conventional observations, assimilated using an EnKF, on the ensemble analyses and subsequent ensemble forecasts of the 9 May 2007 MCS. The remainder of this paper is organized as follows: [section 2](#) discusses the data assimilated, the ensemble DA, and the forecast experiments and methods. [Section 3](#) discusses the results of the experiments, focusing on the impacts of assimilating radar and/or conventional data, as well as ensemble forecast verification. Analyses and forecasts of radar reflectivity (as a proxy for precipitation), mesovortices (an indicator of tornado potential), surface temperature, and surface dewpoint are verified against radar and Oklahoma Mesonet observations, and forecast sensitivity to the assimilation

configuration of conventional observations is considered. Finally, section 4 contains a summary with conclusions.

## 2. Experiment setup and verification methodology

Similar to SXJ11 and SXJ12, EnKF analyses and 3-h storm-scale ensemble forecasts are generated for the tornadic MCS that occurred over Oklahoma and Texas on 8–9 May 2007. During this event, a line-end vortex (LEV) developed near the northern end of the MCS. This LEV moved through southwestern and central Oklahoma, producing two confirmed (enhanced Fujita scale) EF-1 tornadoes and one confirmed EF-0 tornado in central Oklahoma between 0354 and 0443 UTC. For additional details regarding the structure, evolution, and timing of the 8–9 May 2007 MCS we refer the reader to SXJ11.

SXJ11 and SXJ12 sought to assess the impact of two factors on their ensemble analyses and forecasts: 1) the assimilation of CASA X-band radar observations, and 2) the use of a mixed microphysics ensemble as a means to mitigate ensemble underdispersion. While the assimilation of WSR-88D data alone produced a reasonable analysis of the convective system, SXJ11 found that assimilating CASA X-band data in addition to WSR-88D data improved the resulting analysis, particularly with regard to the representation of near-surface circulations. Use of a mixed-microphysics ensemble was found to alleviate underdispersion by increasing the ensemble spread. SXJ12, who carried out ensemble forecasts initialized from the EnKF analyses of SXJ11, found that both assimilation of CASA data and the use of a mixed-microphysics ensemble improved 2-h forecasts of the tornadic mesovortex embedded within the MCS.

SXJ11 and SXJ12 used a single DA and forecast domain with a 2-km horizontal grid spacing; only radar data were assimilated. Lateral boundary conditions were provided by the NCEP NAM 6-hourly analyses and intervening 3-h forecasts. The initial ensemble was created by adding random perturbations with 6-km spatial decorrelation scales to a 1-h spinup forecast on the 2-km grid initialized from the 0000 UTC 9 May 2007 NCEP North American Mesoscale Model (NAM) analysis. While SXJ11 and SXJ12 produced encouraging analyses and forecasts, several deficiencies exist with their setup. Only storm-scale perturbations were used; no mesoscale perturbations were applied. Previous storm-scale data assimilation studies (e.g., Aksoy et al. 2009) suggest that proper structure in mesoscale uncertainty is highly important in obtaining good analyses and forecasts. Furthermore, the single DA domain used the same lateral boundary condition for all members, reducing ensemble spread near the upwind lateral domain boundaries and contributing to underdispersion in

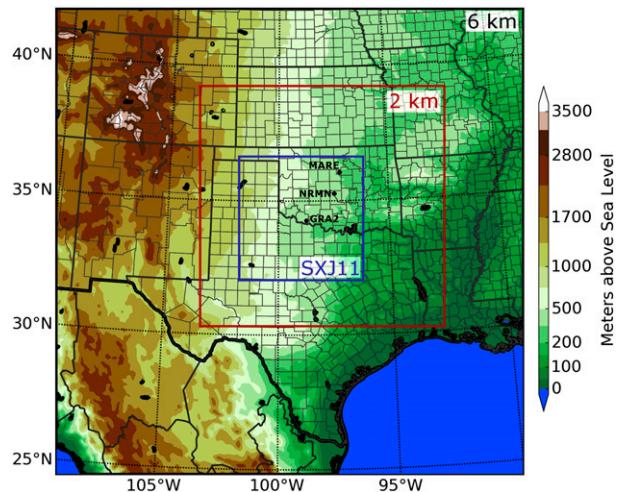


FIG. 1. Geographic extent of the outer model domain (6-km horizontal grid spacing) and the nested inner domain (2-km horizontal grid spacing). Surface elevation (in meters above mean sea level) is plotted for reference. Also shown for reference is the smaller 2-km domain used in SXJ11 and SXJ12. The three black dots in western and central Oklahoma indicate the locations of Oklahoma Mesonet stations MARE (Marena), NRMN (Norman), and GRA2 (Grandfield) used for time series verification in Figs. 15 and 16.

the ensemble analyses and forecasts. The geographic extent of the SXJ11 and SXJ12 domain was also rather limited, which caused detrimental interaction between the simulated MCS and the southern domain boundary. SXJ11 and SXJ12 also did not assimilate surface observations; SXJ12 found that convergence in the near-surface flow in the model contributed to the development of spurious convection in the forecast ensemble.

To improve upon the results of SXJ11 and SXJ12, several changes are implemented in this study. Most prominently, two grids are used: an outer  $300 \times 300 \times 40$  grid with 6-km horizontal spacing, and an inner  $512 \times 512 \times 40$  grid with 2-km horizontal spacing—the extent of the 2-km domain is substantially expanded from that of SXJ11 and SXJ12 (Fig. 1). Data assimilated on the inner grid include both radar and conventional observations, including surface observations at 5-min intervals from the Oklahoma Mesonet.

Uncertainties in the storm environment are taken into account by the ensemble nesting strategy used. Perturbations were applied to the initial outer-domain ensemble; the method is described in detail in the next paragraph. Member-by-member one-way nesting is applied from the outer to the inner domain—both ensembles contain 40 members. On both the outer and inner domains, a single-moment ice microphysics scheme based upon Lin et al. (1983) is used, with a rain intercept parameter of  $8.0 \times 10^5 \text{ m}^{-4}$ ; this value is reduced from the

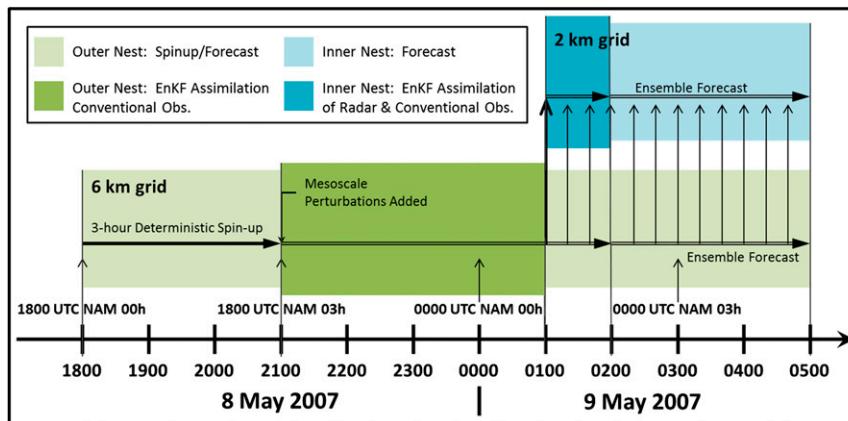


FIG. 2. Flow diagram for forecast experiments. The outer-nest forecast is initialized at 1800 UTC 8 May 2007 via interpolation from the 1800 UTC NAM analysis; 6-hourly NAM analyses and the intervening 3-h forecasts are used as boundary conditions for the outer nest. The inner nest is initialized at 0100 UTC 9 May 2007 using the outer-nest ensemble for initial and boundary conditions.

default value following the results of [Snook and Xue \(2008\)](#). Other model settings follow [SXJ12](#): radiation is parameterized using the NASA Goddard Space Flight Center long- and shortwave radiation parameterization, a two-layer soil model is used, surface fluxes are parameterized using predicted surface temperature and water content, a 1.5-order turbulent kinetic energy (TKE)-based subgrid-scale turbulence parameterization is applied, and model terrain is interpolated from global data with a resolution of 30 arcs.

As in [SXJ11](#), we use the ARPS EnSRF DA system ([Xue et al. 2006](#); [Tong and Xue 2008b](#)). The outer-grid ([Fig. 1](#)) forecast is first initialized at 1800 UTC 8 May 2007 from the 1800 UTC 8 May 2007 NCEP NAM analysis, and a single, 3-h preforecast is performed from this initial condition ([Fig. 2](#)). Lateral boundary conditions for the outer grid are provided from NAM analyses and intervening 3-h NAM forecasts from 1800 UTC 8 May to 0600 UTC 9 May. At 2100 UTC, an ensemble of 40 members is created by adding smoothed, random, Gaussian, mesoscale perturbations generated using a recursive filter to the deterministic forecast with decorrelation scales of 36 and 7.2 km in the horizontal and vertical, respectively, using the method of [Tong and Xue \(2008a\)](#). Perturbations are added to the horizontal wind ( $u$ ,  $v$ ) with a mean standard deviation of  $2 \text{ m s}^{-1}$ , to the

potential temperature ( $\theta$ ) with a mean standard deviation of 1 K, and to the mixing ratio of water vapor ( $q_v$ ) with a mean standard deviation of 10% of the  $q_v$  value at the given grid point.

On the outer domain, conventional observations are assimilated, including Automatic Surface Observing System (ASOS) and Automatic Weather Observing System (AWOS) observations, Oklahoma Mesonet observations, wind profiler data, and soundings at 0000 UTC 9 May 2007; these data are assimilated hourly from 2200 UTC 8 May to 0100 UTC 9 May 2007. Assumed observation errors used vary by observation type as indicated in [Table 1](#). To help maintain ensemble spread during the DA on the 6-km grid, we apply multiplicative covariance inflation to the prior ensemble over the entire domain with an inflation factor of 1.03 ([Anderson and Anderson 1999](#); [Tong and Xue 2005](#)). In addition, we also apply the relaxation technique of [Zhang et al. \(2004\)](#) with a coefficient of 0.5. Finally, 4-h ensemble forecasts are performed from the 0100 UTC ensemble analyses on the outer grid, producing forecasts until 0500 UTC; these forecasts are used to provide ensemble lateral boundary conditions for the inner-nest forecast ensembles.

The 2-km inner-grid EnKF DA experiments are initialized from the outer-grid ensemble analyses at 0100 UTC

TABLE 1. Assumed observation error magnitude for conventional observations.

Data type	$u$ ( $\text{m s}^{-1}$ )	$v$ ( $\text{m s}^{-1}$ )	Temperature (K)	Dewpoint (K)	Pressure (hPa)
Surface	1.5	1.5	1.5	2.0	2.0
Upper air	2.5	2.5	1.2	2.0	0.6
Profiler	2.5	2.5			

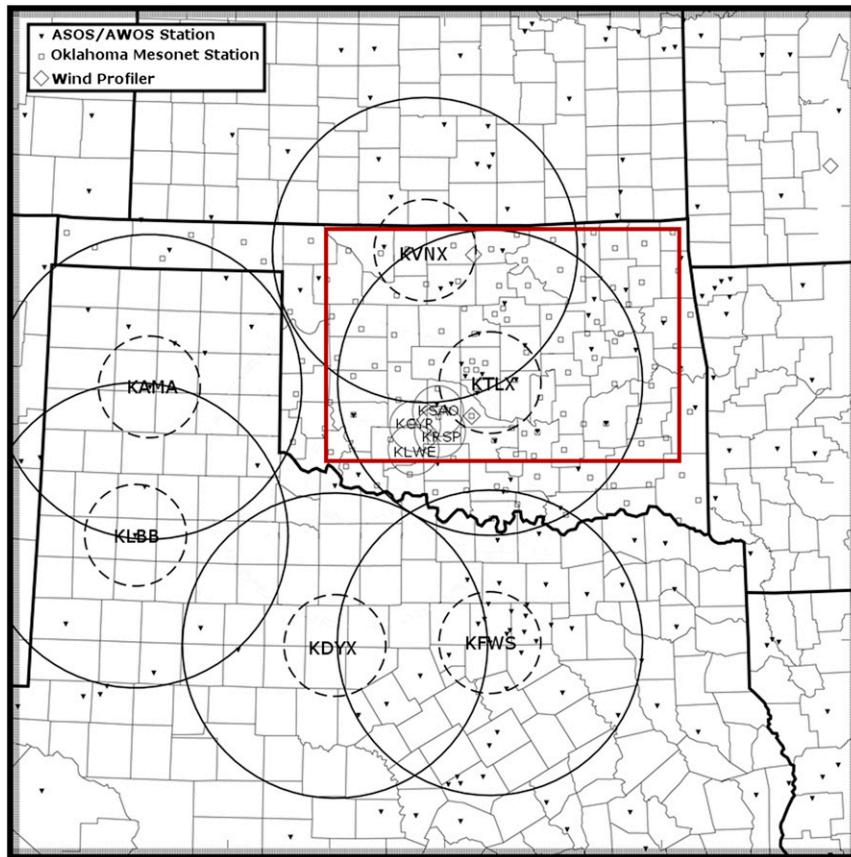


FIG. 3. Observations assimilated using EnKF on the inner-nested grid (2-km grid spacing). The dashed circles and large solid circles indicate 50- and 150-km radius range rings, respectively, for WSR-88D sites used. Small, thin circles indicate 30-km range rings for CASA X-band radar sites used, black triangles indicate ASOS and AWOS surface station sites, squares indicate Oklahoma Mesonet station sites, and diamonds indicate wind profiler sites. The red box indicates the Oklahoma verification subdomain.

via spatial interpolation. Lateral boundary conditions for inner-grid ensemble members are from the forecasts of corresponding outer-grid members at 15 min intervals. The inner-nest experiments assimilate data every 5 min from 0105 to 0200 UTC; the data assimilated include conventional data as described above, as well as radar reflectivity and radial velocity from WSR-88D and CASA radars. For radar data, observation error standard deviations are assumed to be  $2 \text{ m s}^{-1}$  for radial velocity and 3 dBZ for radar reflectivity; these values are increased from the  $1 \text{ m s}^{-1}$  and 2 dBZ used in SXJ11 following Jung et al. (2012). The observation operators used to map the model state to reflectivity and radial velocity observations follow Jung et al. (2008). As in Xue et al. (2006), a Gaussian power-gain function following Wood and Brown (1997) is used in the forward operator to sample radar data on the radar elevation angles. The horizontal and vertical covariance localization radii for radar data are set to 6 km. For conventional

data, the vertical localization radius is set to 6 km, and horizontal localization radii of 300 km for surface observations and 800 km for wind profiler and upper-air observations are used in the primary set of experiments; smaller localization radii are tested in sensitivity experiments on the inner nest. The locations of assimilated conventional observations and radar sites on the inner nest are shown in Fig. 3.

Four primary experiments are run on the inner grid to investigate the impacts of radar and conventional weather observations, assimilated individually or in combination, on the ensemble analyses and forecasts. These four experiments are summarized in Table 2. In experiment RADCONV, data from the WSR-88D and CASA network radars are assimilated, along with conventional observations. In experiment RAD, only radar data are assimilated; similarly, in experiment CONV radar data are omitted and only conventional data are assimilated. Finally, a control experiment (CNTL) is

TABLE 2. Summary of experiments.

Expt name	Radar data used?		Conventional data used?	
	Outer domain	Inner domain	Outer domain	Inner domain
	RADCONV	No	Yes	Yes
RAD	No	Yes	Yes	No
CONV	No	No	Yes	Yes
CNTL	No	No	Yes	No

performed in which no data of any kind are assimilated on the inner domain—the CNTL forecast ensemble is allowed to run freely from the initial ensemble states at 0100 UTC.

### 3. Results

In evaluating the results of the forecast experiments, we will focus primarily on the nature and extent of the individual and combined impacts of assimilated conventional and radar observations on the ensemble forecasts. Forecast verification is performed hourly between 0300 and 0500 UTC for radar reflectivity, surface dewpoint, and surface temperature. In addition, forecasts of low-level mesovortices are produced and verified at 0400 UTC, at which time a pronounced tornadic mesovortex was present in the observations (SXJ12). Radar reflectivity is chosen for verification because it serves as a proxy for precipitation and can be readily verified against WSR-88D observations spanning the full extent of the MCS. A neighborhood ensemble probability (NEP; Schwartz et al. 2010) method with a 5-km neighborhood radius is used for verification of radar reflectivity forecasts, and object-based probabilistic verification following the methodology of SXJ12 is used for the verification of mesovortex forecasts. The radar observations used for forecast verification are obtained by interpolating full radar reflectivity volumes to the model grid to produce a gridded radar reflectivity mosaic. Verification of surface temperature and dewpoint is performed by direct comparison of the forecast ensemble mean against Oklahoma Mesonet observations.

#### a. Impact of data sources during the analysis period

Experiments RAD, CONV, and RADCONV focus on the relative impact of conventional and radar data sources assimilated by the EnKF. CNTL, which is initialized from the same initial conditions as the other three experiments, but in which no additional observations are assimilated, provides a basis for comparison. We note, however, that all experiments (including CNTL) benefit from hourly EnKF DA of conventional observations on

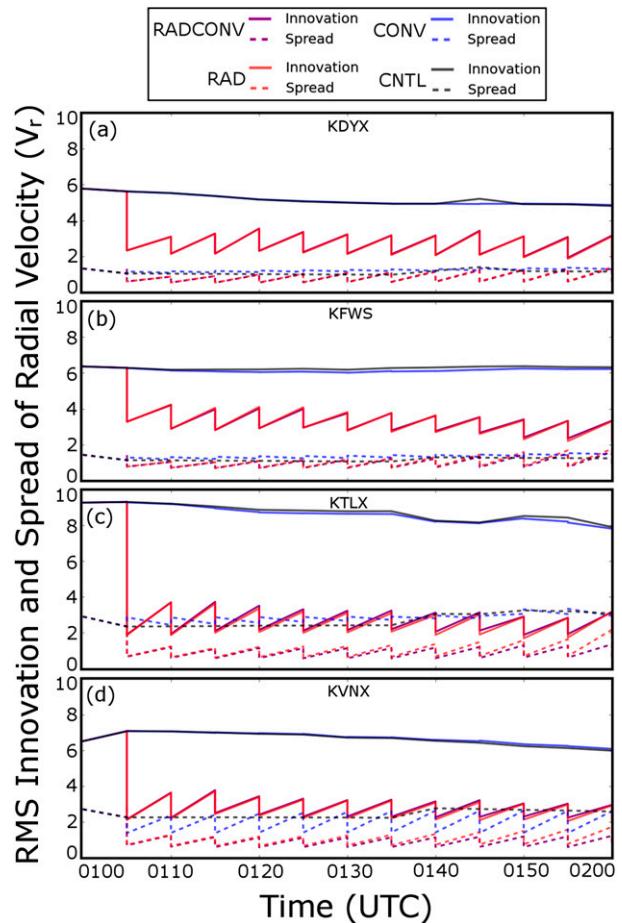


FIG. 4. Average root-mean-square (RMS) innovation (solid lines) of ensemble mean and the ensemble spread (dotted lines) of radial velocity ( $\text{m s}^{-1}$ ) over the observation region of four WSR-88Ds within the model domain from 0100 to 0200 UTC for all experiments. Calculations are limited to locations where observed and/or model (ensemble mean) reflectivity exceeds 15 dBZ.

the outer grid prior to 0100 UTC. All radar-assimilating experiments discussed in this study use both WSR-88D and CASA observations. The specific impact of CASA X-band radar data will not be considered in this paper—such experiments were performed, and they yielded results largely consistent with the findings of SXJ11 and SXJ12.

The assimilation of radar data and the assimilation of conventional observations each have positive impacts on the ensemble forecasts and analyses during the DA period. Compared to the CNTL ensemble, the radar-assimilating ensembles (RADCONV and RAD) exhibit substantially lower root-mean-square (RMS) innovation, as expected, in both radial velocity (Fig. 4) and radar reflectivity (Fig. 5) within the observational area of four WSR-88Ds close enough to observe the MCS. Furthermore, the probability-matched (Ebert 2001) ensemble

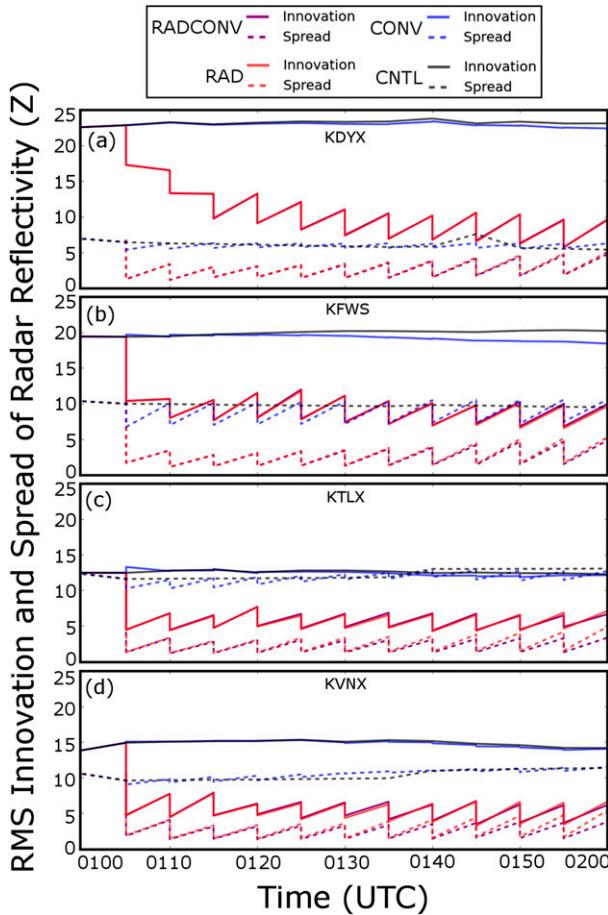


FIG. 5. As in Fig. 4, but for radar reflectivity (dBZ) instead of radial velocity.

mean radar reflectivity field of RADCONV and RAD (Figs. 6a,b) is structurally much closer to the observed radar reflectivity field (Fig. 6e) than that of CNTL (Fig. 6d). The ensemble spread in the radar-assimilating experiments is quickly reduced, both in terms of radial velocity (Fig. 4) and radar reflectivity (Fig. 5). Despite the use of multiplicative covariance inflation to maintain spread, the low spread in RADCONV and RAD indicates that the ensembles quickly become underdispersive; such underdispersion has often been noted in convective-scale ensembles that assimilate radar observations (e.g., Aksoy et al. 2009; Dowell and Wicker 2009; Jung et al. 2012; Yussouf et al. 2013).

Assimilation of conventional data alone in CONV results in a modest reduction in RMS innovation of  $Z$  late in the analysis period, compared to CNTL, against the observations of the KDYX and KFWS radars (Figs. 5a,b). CONV performs similarly to CNTL in terms of RMS innovation of  $Z$  when compared against KTLX and KVNX observations (Figs. 5c,d). KDYX and KFWS primarily observe the trailing stratiform precipitation and

trailing convective line between 0100 and 0200 UTC, while KTLX and KVNX mainly observe the leading portion of the MCS. Both the CNTL and CONV ensembles contain a large area of spurious precipitation in northern Oklahoma and southern Kansas, located within the observation areas of KTLX and KVNX (Figs. 6c–e). Assimilation of conventional observations alone could not suppress this region of spurious convection. RMS innovations of  $V_r$  in CONV are similar to or very slightly lower than those of CNTL (Fig. 4). The 0200 UTC probability-matched ensemble mean of radar reflectivity in CONV (Fig. 6c) shows greater coverage of precipitation in central and south-central Oklahoma compared to CNTL (Fig. 6d), but also contains more spurious precipitation, particularly to the east of the observed MCS (Figs. 6c,e). By contrast, in RAD and RADCONV, where radar data are assimilated alone or alongside conventional data, the probability-matched ensemble mean reflectivity at 0200 UTC (Fig. 6a) closely matches the structure of the observations (though the predicted intensity slightly lower than observed over portions of the MCS), and the spurious precipitation regions seen in CONV (Fig. 6c) are absent. Previous studies (e.g., Tong and Xue 2005) have shown the importance of assimilating radar data in clear-air regions in suppressing spurious precipitation during EnKF DA, consistent with the current results.

#### b. Impact of data sources on ensemble precipitation forecasts

NEP forecasts of  $P[Z > 25 \text{ dBZ}]$  at 0300, 0400, and 0500 UTC (Fig. 7), are generally skillful, particularly for the radar-assimilating experiments RADCONV (Figs. 7a,e,i) and RAD (Figs. 7b,f,j). RADCONV and RAD both predict regions of high  $P[Z > 25 \text{ dBZ}]$ , which closely match the region of precipitation exceeding 25 dBZ observed by the WSR-88D network, both in shape and in extent, particularly at 0300 and 0400 UTC. Decay of the southernmost portion of the trailing line is observed in many ensemble members at 0400 and 0500 UTC in RADCONV and RAD, though not to as great an extent as in SXJ12. The motion of the precipitation regions exceeding 25 dBZ in RADCONV and RAD matches well with the observed system (Figs. 7a–f).

In CONV and CNTL, where no radar data are assimilated, NEP forecasts of  $P[Z > 25 \text{ dBZ}]$  are less accurate than those in RADCONV and RAD. The region of highest probability in CONV and CNTL is located in a west-southwest to east-northeast-oriented streak in southern Kansas, near and just beyond the northern end of the observed region of precipitation exceeding 25 dBZ (Figs. 7g–l). Assimilation of conventional observations in CONV (Figs. 7c,g,k) results in an improved

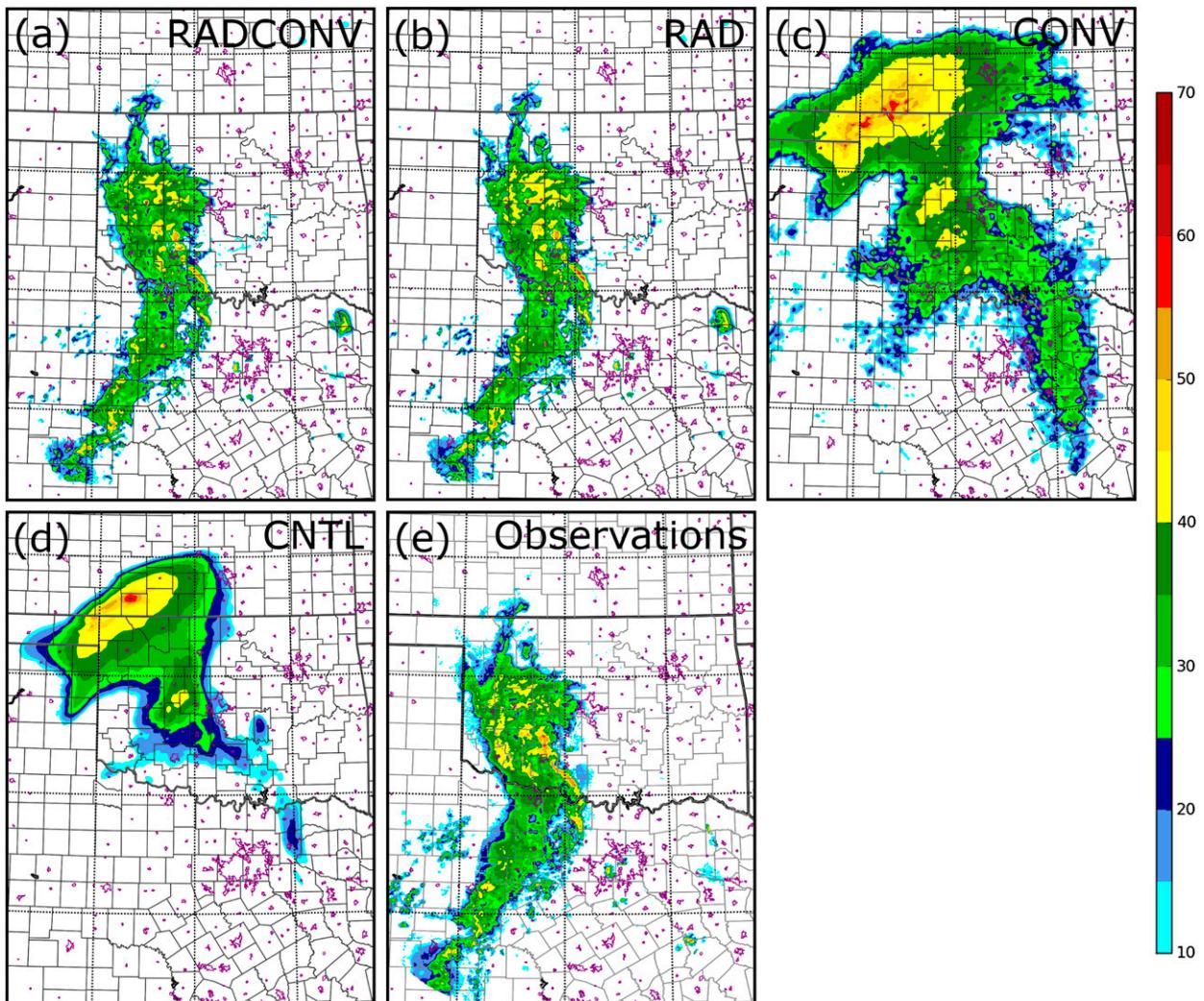


FIG. 6. Probability-matched ensemble mean reflectivity at model grid level 10 (approximately 2 km above the surface) for the 0200 UTC ensemble analyses of (a) RADCONV, (b) RAD, (c) CONV, and (d) CNTL. Also shown is (e) observed reflectivity at 0200 UTC from the WSR-88D network, interpolated to the model grid. Urban boundaries are shown in purple.

representation of the leading portion of the MCS over central and northern Oklahoma in the NEP forecast of  $P[Z > 25 \text{ dBZ}]$  compared to CNTL (Figs. 7d,h,l), as well as increased values of  $P[Z > 25 \text{ dBZ}]$  in central Oklahoma near the LEV, particularly at 0400 and 0500 UTC (Figs. 7g,k). Like CNTL, CONV contains some moderate to high values of  $P[Z > 25 \text{ dBZ}]$  away from any observed precipitation exceeding 25 dBZ—since no radar data were assimilated, this spurious convection could not be effectively suppressed. Overall, assimilation of conventional data alone improved the ensemble precipitation forecast modestly, but not nearly as much as assimilating radar observations.

The threshold of 25 dBZ is chosen to focus on all precipitation exceeding a light-to-moderate intensity. Depending upon the desired forecast focus, however,

a lower threshold may be used to include light precipitation in the NEP forecast, or a higher threshold may be chosen in order to focus exclusively on convective cores. To examine the impact of data sources on probabilistic forecasts of reflectivity with varying thresholds, the area under the relative operating characteristic curve (AUC) for NEP reflectivity forecasts with thresholds varying from 10 to 50 dBZ, calculated hourly over a subdomain encompassing the general region observed by the Oklahoma Mesonet (the red box in Fig. 3), is presented in Fig. 8 for all experiments. For each experiment the 5th–95th percentile range is also shown, obtained using a bootstrap method to produce 1000 randomly resampled 40 member ensembles, allowing for evaluation of the statistical significance of the differences between experiments.

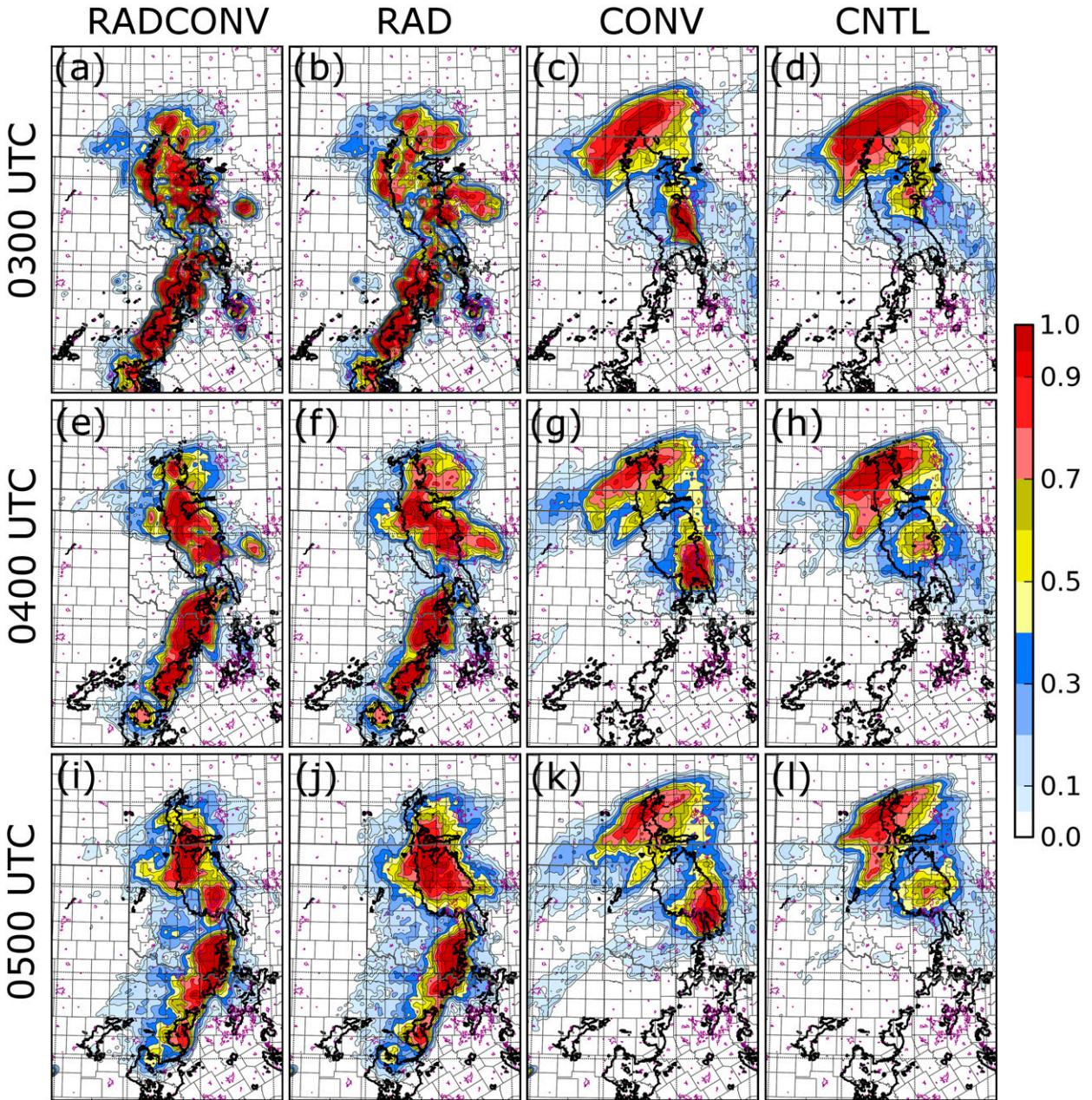


FIG. 7. Neighborhood ensemble probabilities (shaded) of radar reflectivity exceeding 25 dBZ,  $P[Z > 25 \text{ dBZ}]$ , at model grid level 10 (approximately 2 km above the surface) at (a)–(d) 0300, (e)–(h) 0400, and (i)–(l) 0500 UTC for RADCONV, RAD, CONV, and CNTL. The region of radar reflectivity exceeding 25 dBZ observed by the WSR-88D network at the corresponding time is outlined by a bold black contour. Urban boundaries are shown in purple.

For the radar-assimilating experiments (RAD and RADCONV), AUC in the 0200 UTC analysis (Fig. 8a) is very high (close to 1) for thresholds between 10 and 30 dBZ, declining to around 0.9 for higher thresholds, indicating a highly skillful ensemble analysis over the Oklahoma subdomain. Progressing through the forecast period, AUC generally decreases in RAD and RADCONV; the highest values (and thus most skillful

forecasts) are for thresholds between 15 and 35 dBZ. AUCs for thresholds above 40 dBZ decline quickly; this is to be expected, since at these thresholds only very intense reflectivity cores are being considered, and forecast skill is highly sensitive to displacement errors of these small, intense cores. Also, though RAD has a slightly higher AUC than RADCONV at 0200 UTC (Fig. 8a), particularly for higher thresholds, RADCONV outperforms RAD for all

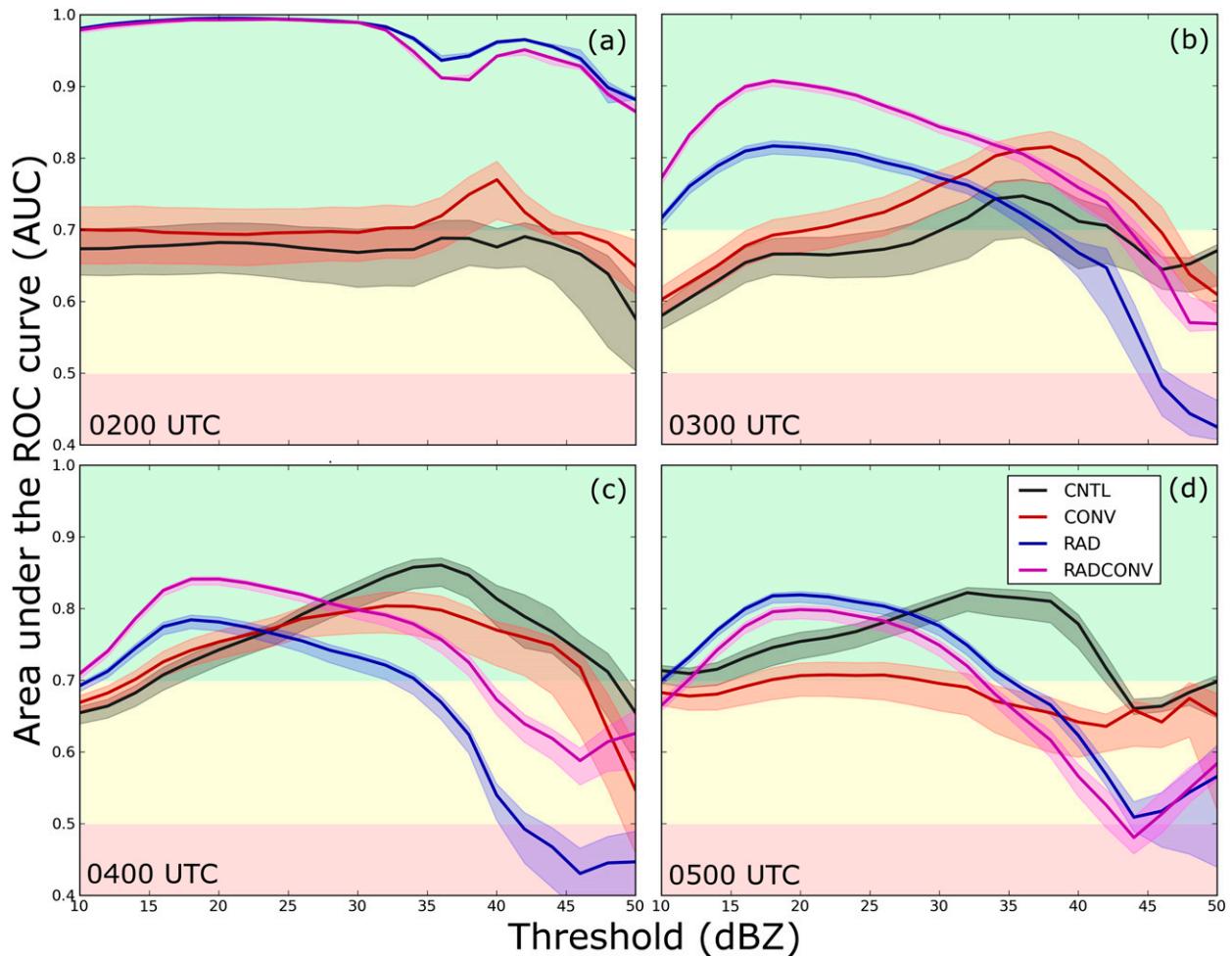


FIG. 8. Area under the relative operating characteristic (ROC) curve (AUC) (solid, bold lines) for all experiments for the analyses at (a) 0200 UTC, and forecasts at (b) 0300, (c) 0400, and (d) 0500 UTC for forecasts of radar reflectivity at vertical grid level  $k = 10$  (slightly more than 2 km above mean sea level) exceeding threshold values ranging from 10 to 50 dBZ at intervals of 2 dBZ. Also shown are the 5th–95th percentile ranges calculated using a 1000-member bootstrap to resample the ensemble (shaded regions). Calculations are performed over the Oklahoma verification subdomain (the red box in Fig. 2). In each panel, the green region indicates AUC values associated with an operationally useful forecast ( $AUC > 0.7$ ). The red region indicates forecasts with no skill ( $AUC < 0.5$ ).

thresholds at 0300 and 0400 UTC (Figs. 8b,c), and shows similar performance at most thresholds at 0500 UTC (Figs. 8c,d). Though RAD produces a better initial fit to the radar observations (note that a tighter fit of analysis to observations assimilated does not necessarily mean better analysis), the addition of conventional data in RADCONV results in more skillful 1- and 2-h forecasts. In both RAD and RADCONV, the 5th–95th percentile range is quite small at most thresholds, indicating relatively low spread within the ensemble. As noted earlier, underdispersion within the ensemble is a common issue when assimilating radar observations (Aksoy et al. 2009; Dowell and Wicker 2009; Jung et al. 2012; Yussouf et al. 2013).

When conventional data are assimilated alone in CONV, the 0200 UTC ensemble analysis of radar

reflectivity has a slightly higher AUC than control experiment CNTL, but a substantially lower AUC than the radar-assimilating experiments (Fig. 8a). AUC in CONV actually increases at high thresholds between 0200 and 0300 UTC (Figs. 8a,b), even outperforming RAD and RADCONV due to good placement of heavy convective cores in south-central Oklahoma and less overprediction of very intense rainfall cores (not shown), before declining at all thresholds between by 0500 UTC (Fig. 8d). Positive impact of conventional data in CONV on AUC (cf. CNTL) remains evident at 0300 UTC (Fig. 8b), but cannot be discerned at later times. AUC is strongly sensitive to the probability of false detection. Although the overall structure of MCS predicted by CNTL is poor (Fig. 7), relatively low probability of false

detection coupled with decent precipitation placement (Figs. 7h,i) leads to a deceptively high AUC score for high thresholds at later forecast hours.

Another method of assessing the skill of a probabilistic forecast is the reliability diagram, which compares the observed relative frequency of an event to the forecast probability. Because the reliability diagram is conditioned on the ensemble forecast, while AUC is conditioned on the observations, these two metrics complement one another and give a more complete assessment of forecast skill. Reliability diagrams, calculated over the Oklahoma verification subdomain (see Fig. 3) using forecast probability bins with a width of 0.05, are plotted for NEP forecasts of  $P[Z > 25 \text{ dBZ}]$  for all experiments in Fig. 9 to complement the analysis of AUC using the same subdomain presented in Fig. 8. In an ideal forecast, the observed frequency would be equal to the forecast probability, resulting in a straight reliability curve oriented along the  $45^\circ$  diagonal (indicated by the dotted lines in Fig. 9). The region below the diagonal indicates overforecasting of the event, while the area above the diagonal indicates underforecasting. Sharpness diagrams are also presented in Fig. 9, indicating the number of model grid points falling into each probability bin, and thus the overall distribution of probabilities in the forecast. Since the verification subdomain extends well outside of the MCS, these histograms contain many zero values.

In the ensemble analyses at 0200 UTC (Fig. 9a), the radar-assimilating experiments show good reliability in their forecasts of  $Z > 25 \text{ dBZ}$ , while CNTL and CONV exhibit substantial overprediction. In the radar-assimilating experiments (RADCONV and RAD), there is an overall monotonic increase in observed frequency as forecast probability increases, which is a desirable trait. This pattern is absent in the 0200 UTC analyses of CNTL and CONV for forecast probabilities of above approximately 0.4.

During the forecast period (Figs. 9b–d), there is a general trend toward overprediction of  $Z > 25 \text{ dBZ}$  in the data-assimilating experiments, particularly at 0400 and 0500 UTC. RADCONV shows good reliability throughout much of the forecast period, remaining near the diagonal except at the highest forecast probabilities, outperforming all other experiments. RAD and CONV substantially overpredict  $Z > 25 \text{ dBZ}$  for forecast probabilities above 0.4, but exhibit reliability curves that indicate a general increase of observed frequency as forecast probability increases, particularly at 0400 and 0500 UTC (Figs. 9c,d). Though CNTL shows good reliability for low-to-moderate forecast probabilities, its reliability is quite poor for forecast probabilities above about 0.6. The tendency toward greater overprediction of  $Z > 25 \text{ dBZ}$  during the forecast period in RAD and RADCONV is similar to that seen in Clark et al. (2009)

in their convection-allowing ensemble using 4-km horizontal grid spacing.

The radar-assimilating experiments (RAD and RADCONV) produce significantly more extreme probability values (near 0 or 1) than either CONV or CNTL (Fig. 9, right-hand side). The assimilation of radar data in these experiments results in strong agreement among the ensemble members in the structure of the MCS in the 0200 UTC analysis, while greater spread remains evident in the CONV and CNTL ensembles (see Fig. 5). The MCS evolves similarly in many RAD and RADCONV members during the forecast period (see Figs. 7a–f), causing this sharpness to persist to some extent throughout the forecast period.

### *c. Impact of data sources on precipitation forecast bias*

As in SXJ12, domain-wide histograms of radar reflectivity (Fig. 10) reveal persistent biases in the ensemble forecasts. The forecast histograms are obtained by counting occurrences of radar reflectivity values in each member separately. The resulting total in each bin is then divided by the number of members in the ensemble, so that the number of occurrences can be compared directly to reflectivity counts from radar observations interpolated to the model grid. In three of the experiments (RADCONV, RAD, and CNTL), there is a low bias for light precipitation (15–25 dBZ) which is most prevalent later in the forecast period at 0400 and 0500 UTC. In RAD and RADCONV there is an abundance of moderately intense precipitation (25–45 dBZ), resulting in a high bias. These biases are similar to those found in SXJ12, where assimilation of radar data resulted in overprediction of convective regions and underprediction of light precipitation in stratiform precipitation regions; this behavior is also present in radar-assimilating experiments RAD and RADCONV. The high bias in moderate precipitation is largely absent in CNTL, and manifests in RAD and RADCONV between the 0200 UTC analysis and the 1-h forecast at 0300 UTC, suggesting that the radar-assimilating forecasts may result from interaction of assimilated radar data with error in the single-moment microphysics scheme used in the ensembles. In a related study (Putnam et al. 2014) the high biases in moderately intense precipitation fields are reduced when a more sophisticated, two-moment scheme is used for this case.

At 0200 UTC, CONV exhibits a substantial high bias for both light and moderate precipitation (15–45 dBZ) mostly due to spurious precipitation causing overestimation of the extent of the precipitation area. By 0300 UTC, the high bias in light precipitation ( $Z < 20 \text{ dBZ}$ ) has disappeared, but the high bias in moderate

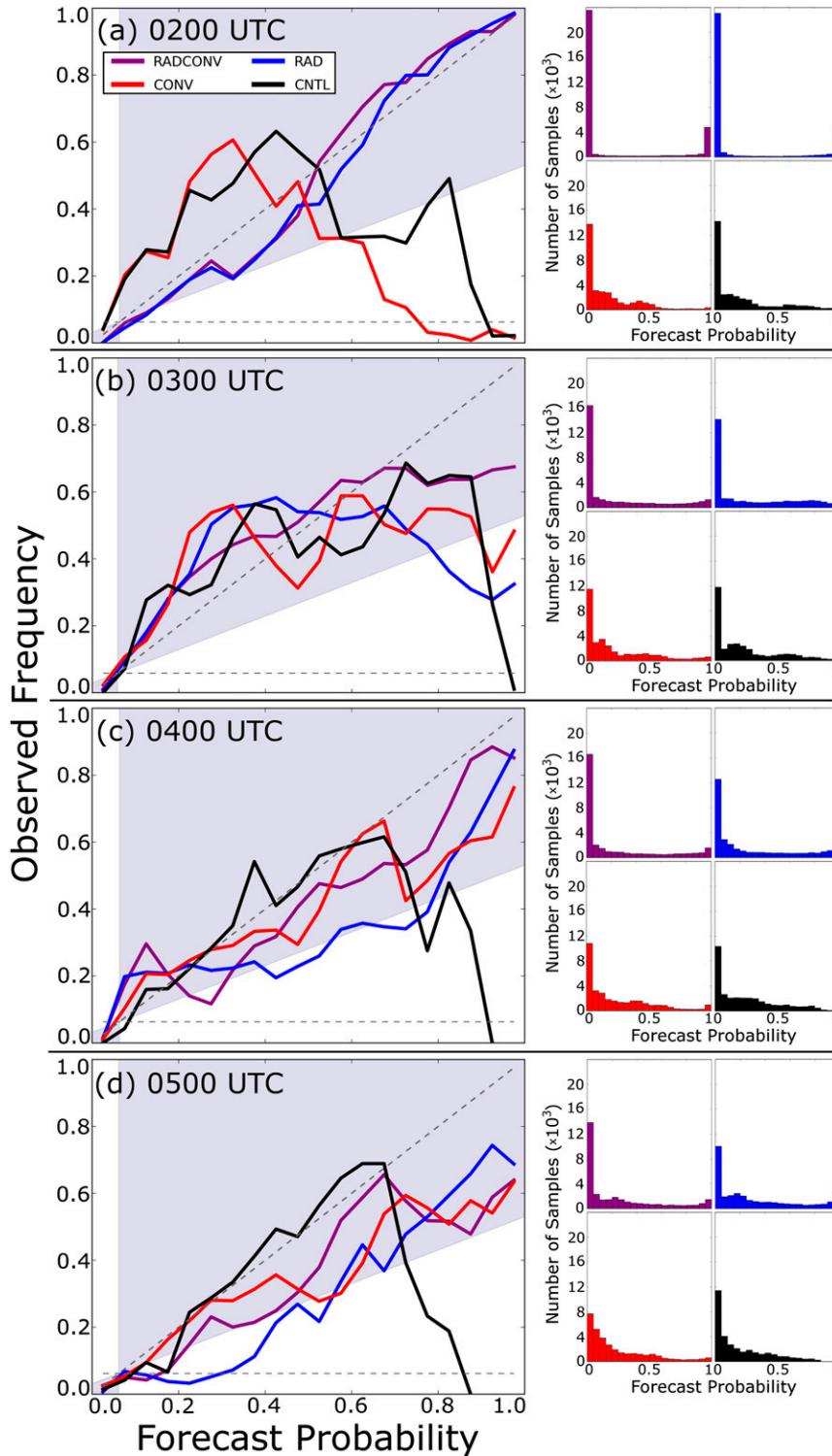


FIG. 9. Reliability and sharpness diagrams for NEP forecasts of  $P[Z > 25 \text{ dBZ}]$  for all experiments at (a) 0200, (b) 0300, (c) 0400, and (d) 0500 UTC calculated over the Oklahoma verification subdomain (the red box in Fig. 2). Forecast probability bins are spaced at intervals of 0.05.

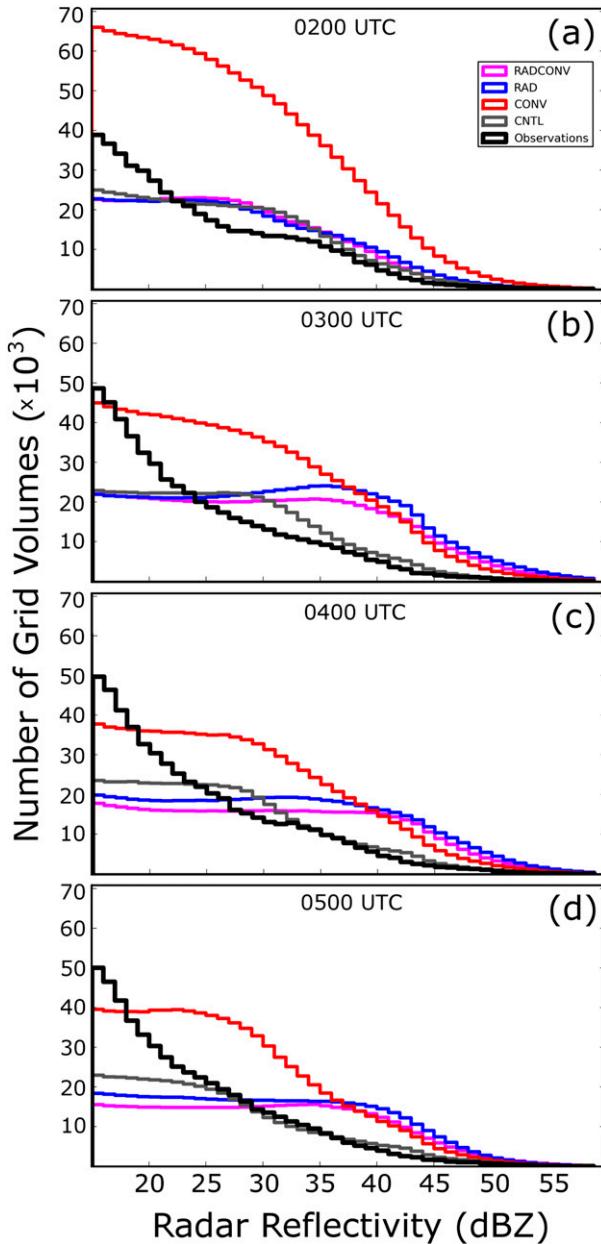


FIG. 10. Hourly, domain-wide histograms of forecast radar reflectivity for all ensemble forecast experiments, compared to WSR-88D observed reflectivity interpolated to the ensemble forecast grid. Bins are placed every 1 dBZ. The vertical axis indicates the number of model grid volumes within each bin; for the ensemble forecasts these values are normalized by the size of the ensemble.

to moderately intense precipitation (20–45 dBZ) remains through the rest of the forecast period. CONV shows very different bias behavior compared to RADCONV and RAD; the two radar-assimilating experiments have very similar bias behavior throughout the forecast period. From these results we can conclude that, at least for this case, conventional data have a different impact upon the

forecast bias of  $Z$  when they are assimilated alone; radar data, when they are assimilated, appear to be the dominant factor with regard to the bias of  $Z$  within the forecast ensemble.

*d. Impact of data sources on mesovortex prediction*

Object-based ensemble forecasts of the probability of low-level mesovortices within 25 km of a point are calculated using the 2-h ensemble forecasts valid at 0400 UTC. The methodology and criteria used to perform these forecasts follow that of SXJ12, and we refer the reader to SXJ12 for further details regarding the probability calculation. At 0400 UTC, a tornadic mesovortex was present west-southwest of the Oklahoma City metropolitan area, indicated by the green triangle in each panel of Fig. 11. All four ensemble forecast experiments (RADCONV, RAD, CONV, and CNTL) indicate a probability of at least 0.2 of a mesovortex being present in close proximity to the observed tornadic mesovortex; probability near the observed vortex location is highest in RADCONV and lowest in CNTL.

RADCONV produces the best probabilistic mesovortex forecast, with a region of moderate probability (maximum  $>0.6$ ) concentrated near the observed mesovortex location (Fig. 11a). RAD predicts a similarly shaped region of relatively low probability (maximum  $\approx 0.3$ ), centered 20–30 km northwest of the observed mesovortex location, as well as a region of probability between 0.05 and 0.20 in north-central Oklahoma (Fig. 11b). CNTL, which did not benefit from any assimilated observations on the inner grid, predicts a small region of low probability (maximum  $\approx 0.2$ ), also centered 30–40 km northwest of the observed mesovortex location (Fig. 11d), suggesting that at least some of the information needed to correctly predict the mesovortex in this MCS is captured in the outer-nest ensemble providing the initial and lateral boundary conditions for CNTL.

The better mesovortex prediction of RADCONV compared to RAD supports the findings of Schenkman et al. (2011a), who showed that the assimilation of Oklahoma Mesonet, CASA, and WSR-88D observations for this case using a 3DVAR and cloud analysis system yielded a better prediction of the low-level wind field and the tornadic mesovortex than when assimilating radar data alone. We note, however, that assimilation of conventional data alone does not substantially improve the probabilistic mesovortex forecast over CNTL. In short, for this case, assimilation of conventional data only results in an improved mesovortex forecast when radar data are also assimilated. Assimilation of radar data, which have relatively complete

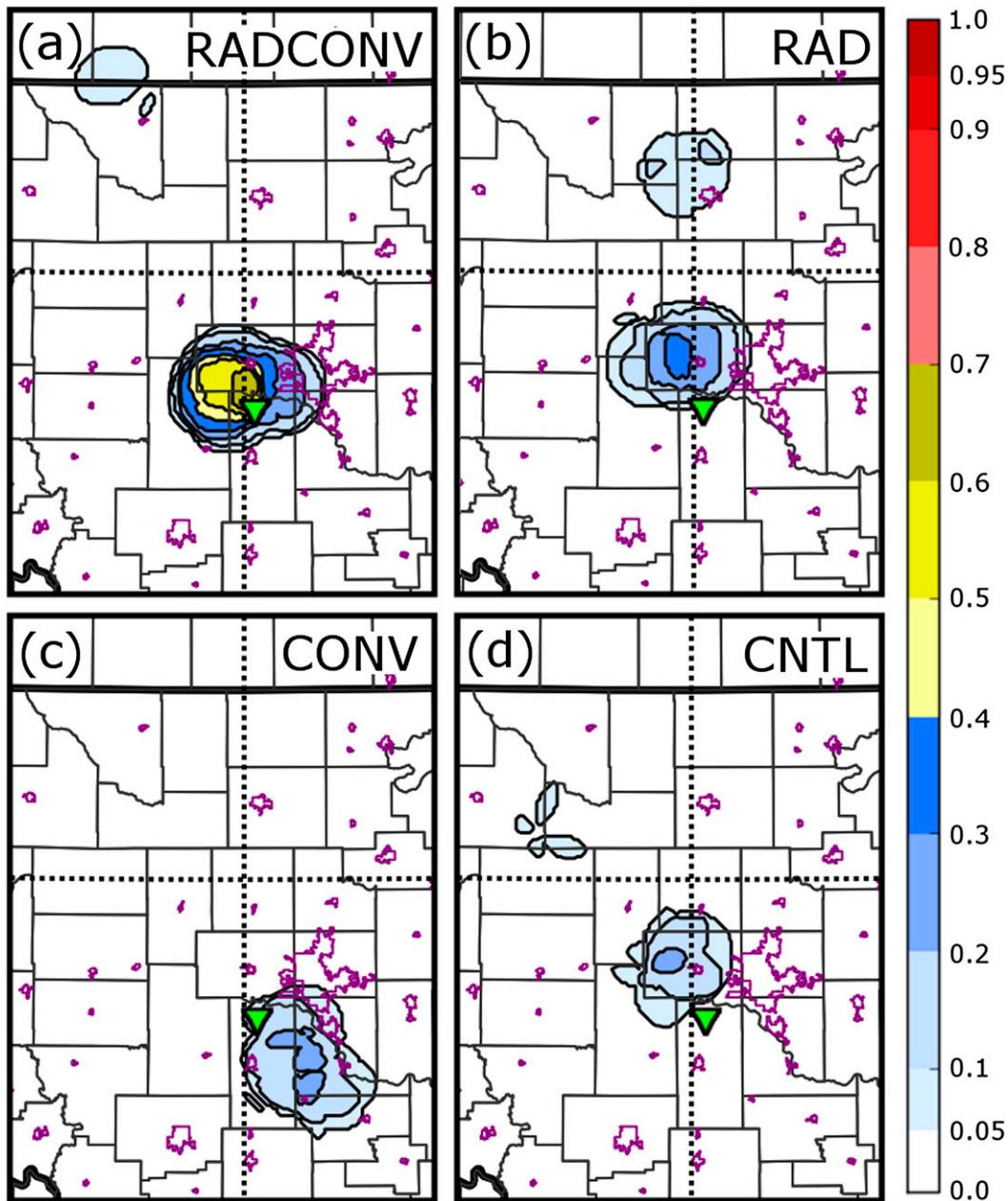


FIG. 11. Ensemble-based probability of a significant near-surface mesovortex occurring within 25 km of a point (shaded) at 0400 UTC for (a) RADCONV, (b) RAD, (c) CONV, and (d) CNTL. The location of the observed tornadic mesovortex (located within the line-end vortex of the MCS) at 0400 UTC is indicated by the green triangle in each panel. Urban boundaries are shown in purple.

volumetric coverage throughout the MCS, is necessary to produce a good initial representation of the storm within the model. Assimilating conventional observations alone (which are far coarser than the radar observations and most abundant at the surface) cannot substantially improve the storm-scale ensemble forecast or impart an accurate three-dimensional storm structure in this case.

#### *e. Verification of surface temperature and dewpoint*

The ability of radars to provide complete volumetric coverage is generally limited very close to the surface, since the curvature of the earth prevents radars from observing the near-surface region beyond a few tens of kilometers from the radar site. This limitation motivates the assimilation of conventional observations alongside

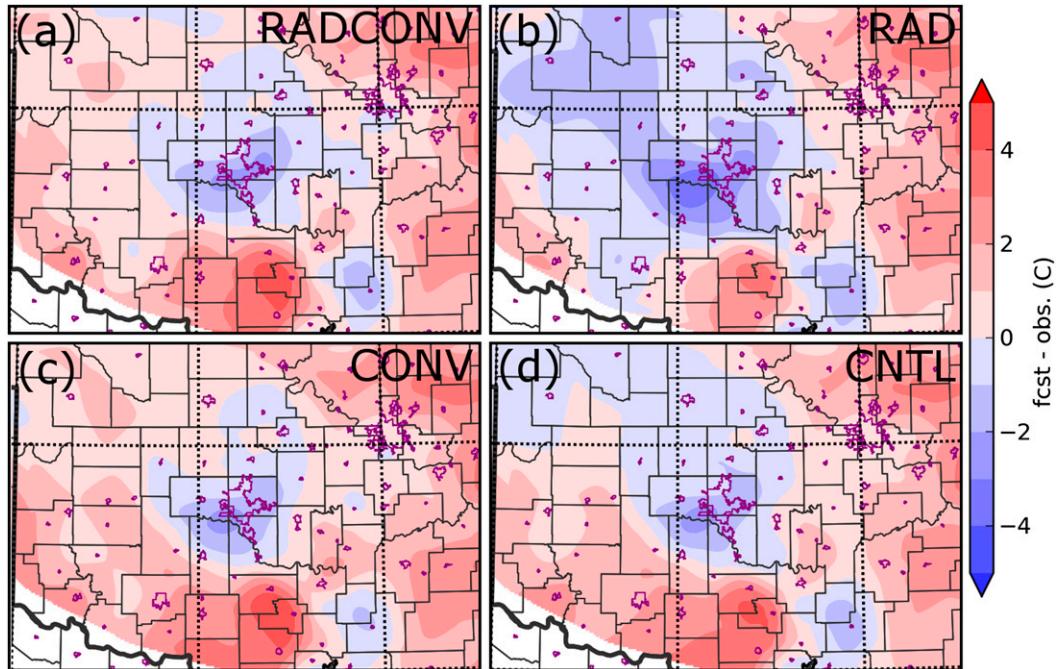


FIG. 12. Contours of the difference between ensemble mean surface (2 m) temperature forecast and Oklahoma Mesonet observations at 0400 UTC for (a) RADCONV, (b) RAD, (c) CONV, and (d) CNTL. Red areas indicate an ensemble mean temperature warmer than observed by the mesonet, while blue areas indicate an ensemble mean colder than observations.

radar, particularly when relatively dense surface observations such as those from the Oklahoma Mesonet are available. To assess the skill of the ensemble forecasts near the surface, surface temperature and dewpoint are compared against Oklahoma Mesonet observations at 0400 UTC (2 h of forecast time) in Fig. 12 and Fig. 13, respectively.

In all forecast experiments, the ensemble mean surface temperature is colder than observed near the LEV (located near the Oklahoma City metropolitan area at 0400 UTC; cf. Fig. 11), and warmer than observed to the east of the MCS and in the vicinity of the trailing convective line in southern Oklahoma (Fig. 12). The cold bias near the LEV is greatest in RAD (Fig. 12b), exceeding 3°C in places, and least in RADCONV and CONV (Figs. 12a,c). RADCONV (Fig. 12a) and CONV (Fig. 12c) overall show similar patterns of temperature forecast error.

When 2-h surface dewpoint forecasts are compared (Fig. 13), a prominent dry bias is present in the forecast ensembles over much of central and northwest Oklahoma. The dry bias is most intense in RAD (Fig. 13b), extending over nearly all of Oklahoma, and exceeding 6°C in northwest Oklahoma. In CONV and CNTL (Figs. 13c,d), which did not assimilate radar data, the dry bias is less intense, and more limited in extent. The overall bias is smallest in CONV. RADCONV (Fig. 13a) has a greatly

reduced dry bias compared to RAD (Fig. 13b). The reduced error in the surface temperature and surface dewpoint forecasts of RADCONV in comparison to RAD suggest that the conventional observations assimilated in RADCONV impart a substantial improvement to the surface thermodynamic and moisture fields of the ensemble that is maintained through the subsequent forecast period.

To more closely examine the ensemble behavior of surface fields, ensemble temperature and dewpoint fields are interpolated to the location of three Oklahoma Mesonet sites at 5-min intervals and compared against 5-min observations from the corresponding sites in Fig. 14 and Fig. 15. The Oklahoma Mesonet sites chosen are marked in Fig. 1 and include Marena (MARE), located in the northern stratiform region of the MCS; Norman (NRMN), which observes the passage of the convective line during the forecast period; and Grandfield (GRA2), which observed the poststorm environment during much of the forecast period. In addition to the ensemble mean, Fig. 14 and Fig. 15 also show the 5th–95th percentile range within the 40-member ensemble.

All experiments produce cold biases at NRMN (Figs. 14e–h) and warm biases at GRA2 (Figs. 14i–l) throughout the analysis and forecast period, while producing values near those observed at MARE (Figs. 14a–d). The ensembles of RADCONV and CONV (which

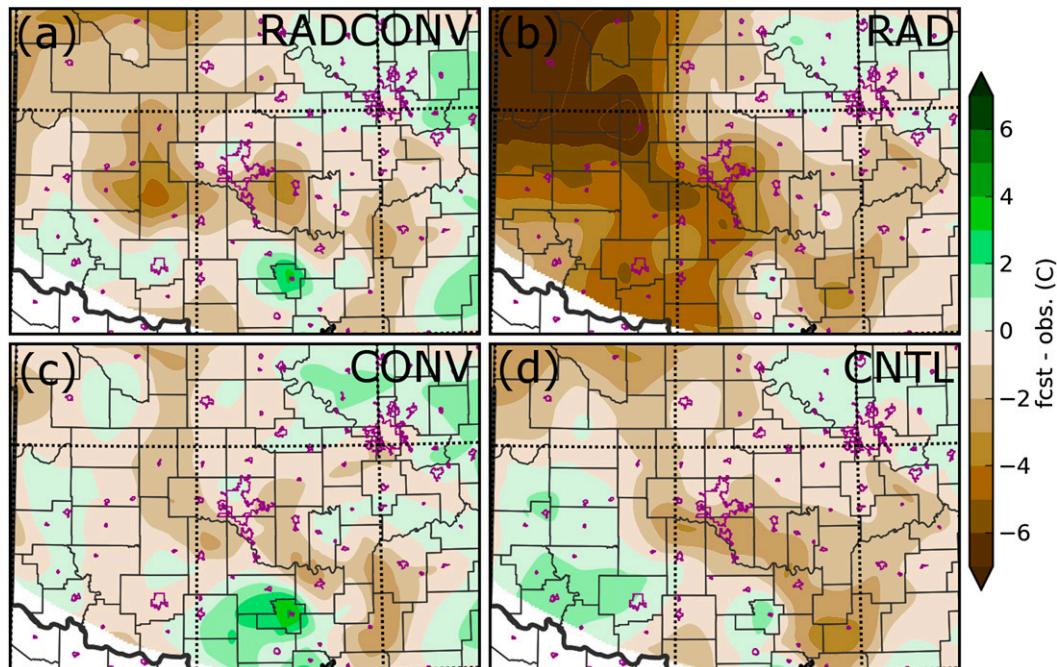


FIG. 13. Contours of the difference between ensemble mean surface (2 m) dewpoint forecast and Oklahoma Mesonet observations at 0400 UTC for (a) RADCONV, (b) RAD, (c) CONV, and (d) CNTL. Green areas indicate an ensemble mean dewpoint higher than observed by the mesonet, while brown areas indicate an ensemble mean dewpoint lower than observations.

assimilate conventional data) show more similar trends in temperature to the observed values than those of RAD and CNTL, particularly at GRA2 (Figs. 14i–l), where they show a drop in temperature between 0100 and 0200 UTC, consistent with observations. For dewpoint, the ensembles show dry biases at MARE and NRMN (Figs. 15a–h) throughout most of the assimilation and forecast period; this bias is greatest in RAD and least in CONV. RAD also shows a dry bias at GRA2 (Figs. 15i–l) during much of the forecast period; as noted earlier, RAD exhibits a pronounced dry bias over much of western Oklahoma (Fig. 13). For both temperature (Fig. 14) and dewpoint (Fig. 15), RAD has much greater spread within the ensemble, as evidenced by the very large 5th–95th percentile range in RAD compared to RADCONV and CNTL. The ensemble of RAD had a few ensemble members that produced very dry air near the surface over central and southwestern Oklahoma (not shown), contributing to the very large spread in the ensemble.

Because RAD does not assimilate surface observations, it relies on radar observations alone to adjust the surface thermodynamic fields. Though there is near-surface radar coverage over many portions of the domain, the fields produced in RAD based on the influence of radar covariance structure near the surface do not result in good forecasts in all ensemble members. This

result is consistent with Dowell et al. (2011), who also found that assimilation of radar reflectivity information using an EnKF did not result in positive impact upon verification against surface observations. When conventional data are assimilated alongside radar data in RADCONV, the bias of the ensemble is reduced and the ensemble spread in the surface fields is reduced to levels similar to CONV. We note, however that because the ensemble has a horizontal resolution of only 2 km and a minimum vertical spacing of 25 m, accurate prediction of surface fields may be beyond the capability of the ensemble, regardless of the data used. Yussouf et al. (2013) also noted relatively poor agreement between their storm-scale ensemble using 2-km horizontal grid spacing and individual surface observation time series.

#### f. Impact of localization radius for conventional observations

When using an EnKF that assimilates observations from multiple networks, the spatial covariance localization radii usually need to be tuned. The use of different, optimized radii for different observation types has been found to improve analyses and forecasts (Dong et al. 2011; Zhu et al. 2013). In addition to the four primary experiments presented above, several sensitivity experiments were performed using different covariance localization radii for conventional data assimilated in

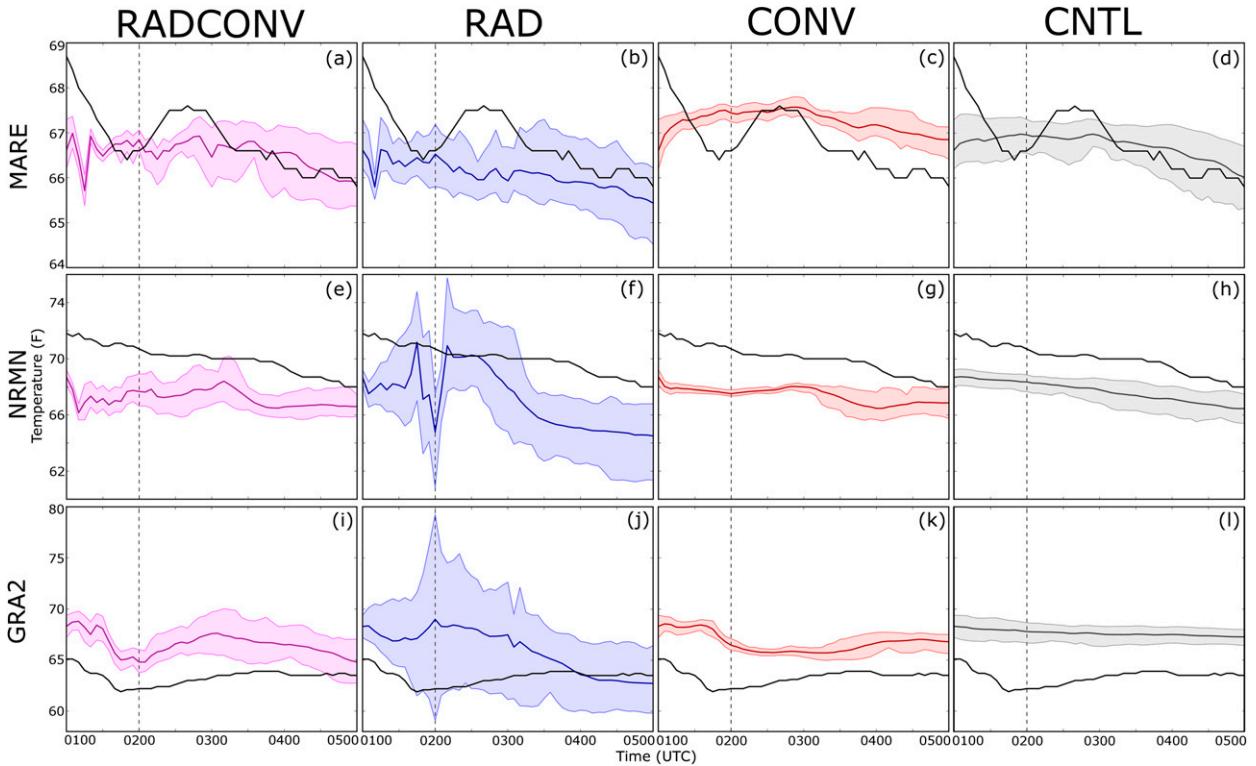


FIG. 14. Observed surface (2 m) temperature ( $^{\circ}\text{F}$ ) (solid black lines) at 5-min intervals between 0100 and 0500 UTC from Oklahoma Mesonet sites at (a)–(d) Marena (MARE), (e)–(h) Norman (NRMN), and (i)–(l) Grandfield (GRA2). Also shown are ensemble mean (thick colored lines) and the 5th–95th percentile range (shaded colored regions) for experiments RADCONV [purple; (a), (e), and (i)]; RAD [blue; (b), (f), and (j)]; CONV [red; (c), (g), and (k)]; and CNTL [gray; (d), (h), and (l)] interpolated to the location of the corresponding Oklahoma Mesonet sites. The vertical dotted line in each panel denotes the end of the DA period at 0200 UTC.

RADCONV and CONV. Horizontal localization radii from 100 to 800 km were tested for surface data, and radii from 150 to 1500 km were tested for upper-air observations; these ranges are based upon prior experimentation by the authors and typical values used in similar studies. Though all radii tested resulted in qualitatively similar 0200 UTC analyses (not shown), substantial differences were noted during subsequent ensemble forecasts. Because the forecast trend is monotonic as radii increase (for small to moderate radii), we only show results using the smallest radii tested and for the combination of radii that produced the most skillful forecasts (300 km for surface data and 800 km for upper-air data). Increasing influence radii beyond 300 km for surface data and 800 km for upper-air data resulted in forecasts whose skill decreased more quickly during the forecast period (not shown).

When only conventional data are assimilated, the quality of NEP forecasts of  $P[Z > 25 \text{ dBZ}]$  is better for larger localization radii. When a localization radius of 300 km is used for surface data and 800 km for upper-air data, as in CONV, regions of high probability are confined to a relatively tight north–south line near and

slightly to the east of the observed MCS location (Figs. 16d–f), and NEP near the LEV is high (in many places,  $>0.95$ ). By comparison, when reduced localization radii of 100 km for surface data and 150 km for upper-air data are used (Figs. 16a–c), the linear structure of the MCS is not as evident, particularly at 0500 UTC (Fig. 16c). The two distinct precipitation regions in the simulation using reduced localization radii are more similar to those of CNTL. When radar data are assimilated alongside conventional data, the positive impact of the assimilated observations is less when a smaller localization radius is used for conventional observations (not shown).

The tornadic mesovortex that was ongoing at 0400 UTC is also better predicted when larger localization radii are used to assimilate conventional data. Mesovortex probability predictions for RADCONV (Fig. 17a) and CONV (Fig. 17c), which use 300- and 800-km localization radii for surface and upper-air data, respectively, compare favorably to their counterparts using reduced localization radii for conventional DA (Figs. 17b,d). In particular, when radar data are assimilated alongside conventional data (Figs. 17a,b), the probability near the observed mesovortex location is increased. Spurious

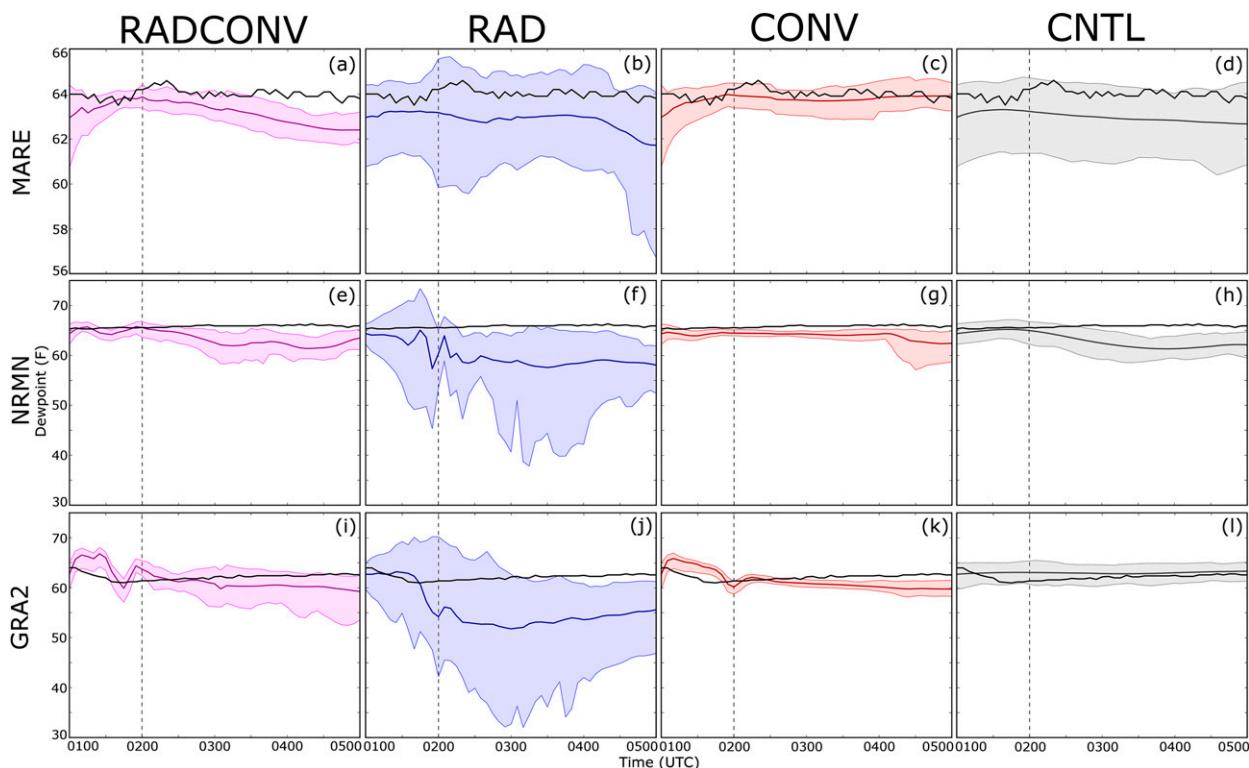


FIG. 15. As in Fig. 14, but for observed surface (2 m) dewpoint (solid black lines) and ensemble mean (thick colored lines) and the 5th–95th percentile range within the ensemble (shaded colored regions) of dewpoint interpolated to the corresponding Oklahoma Mesonet site.

detections are reduced when using the larger localization radii (Fig. 17a). Similar improvements are also present when conventional data are assimilated alone (Figs. 17c,d). The seemingly optimal 300-km localization radius used for the surface data appears rather large compared to the mean station spacing of the Oklahoma Mesonet, which is about 30 km. Dong et al. (2011) suggested optimal localization radii that are slightly larger than the mean station spacing of individual observation networks. We theorize that, since conventional observations were rather sparse above the surface and outside of the region covered by the much denser Oklahoma Mesonet, expanding the influence of conventional observations helps spread their positive impact upstream into regions with sparse data coverage during DA, thus increasing their ability to positively influence the ensemble forecast.

#### 4. Discussion and summary

The relative impacts of various data sources, and their effective assimilation, are important issues in convective-scale weather forecasting. At convective scales, Doppler radars provide temporally and spatially dense observations of radar reflectivity ( $Z$ ) and radial velocity ( $V_r$ ).

Conventional observations, including ASOS, AWOS, mesonet, wind profiler, and upper-air observations, offer sparser coverage than radar, but provide valuable information close to the surface and in clear-air regions that radar typically cannot.

In this study, we examined the individual and combined impacts of assimilating radar and/or conventional observations, using an ensemble square root Kalman filter (EnSRF), upon ensemble analyses and forecasts of a tornadic mesoscale convective system (MCS), focusing on prediction of radar reflectivity (a proxy for precipitation) and low-level mesovortices (a proxy for tornado potential). Conventional observations, radar observations, both, or neither were assimilated using a storm-scale EnKF nested within a mesoscale ensemble; 3-h ensemble forecasts were performed from the final ensemble analyses. Conventional and radar observations were both found to improve the analyses and forecasts, but in different ways.

All ensemble forecasts that assimilated conventional and/or radar data outperformed the control experiment (which assimilated neither) in terms of radar reflectivity probabilistic forecasts. The most skillful forecast, in terms of the area under the relative operating characteristic (ROC) curve (AUC), was produced by the

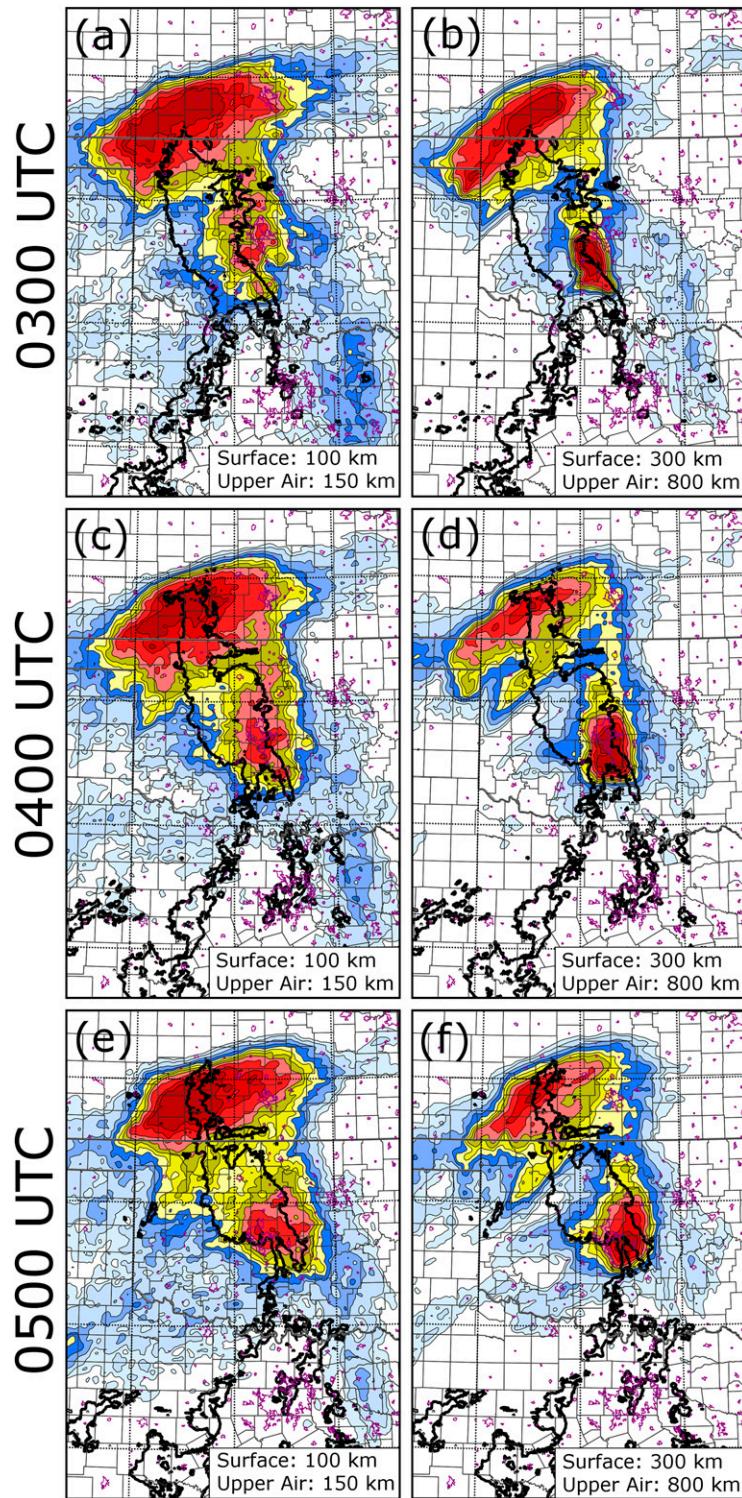


FIG. 16. Neighborhood ensemble probability (shaded) of radar reflectivity exceeding 25 dBZ,  $P[Z > 25 \text{ dBZ}]$ , at model grid level 10 (approximately 2 km above the surface) at (a),(b) 0300; (c),(d) 0400; and (e),(f) 0500 UTC for (right) CONV and (left) a variant of CONV using reduced localization radii for conventional data assimilation, at 0300, 0400, and 0500 UTC. The region of radar reflectivity exceeding 25 dBZ observed by the WSR-88D network at the corresponding time is outlined by a bold black contour. Urban boundaries are shown in purple. The localization radius used to assimilate surface and upper-air data is noted in each panel.

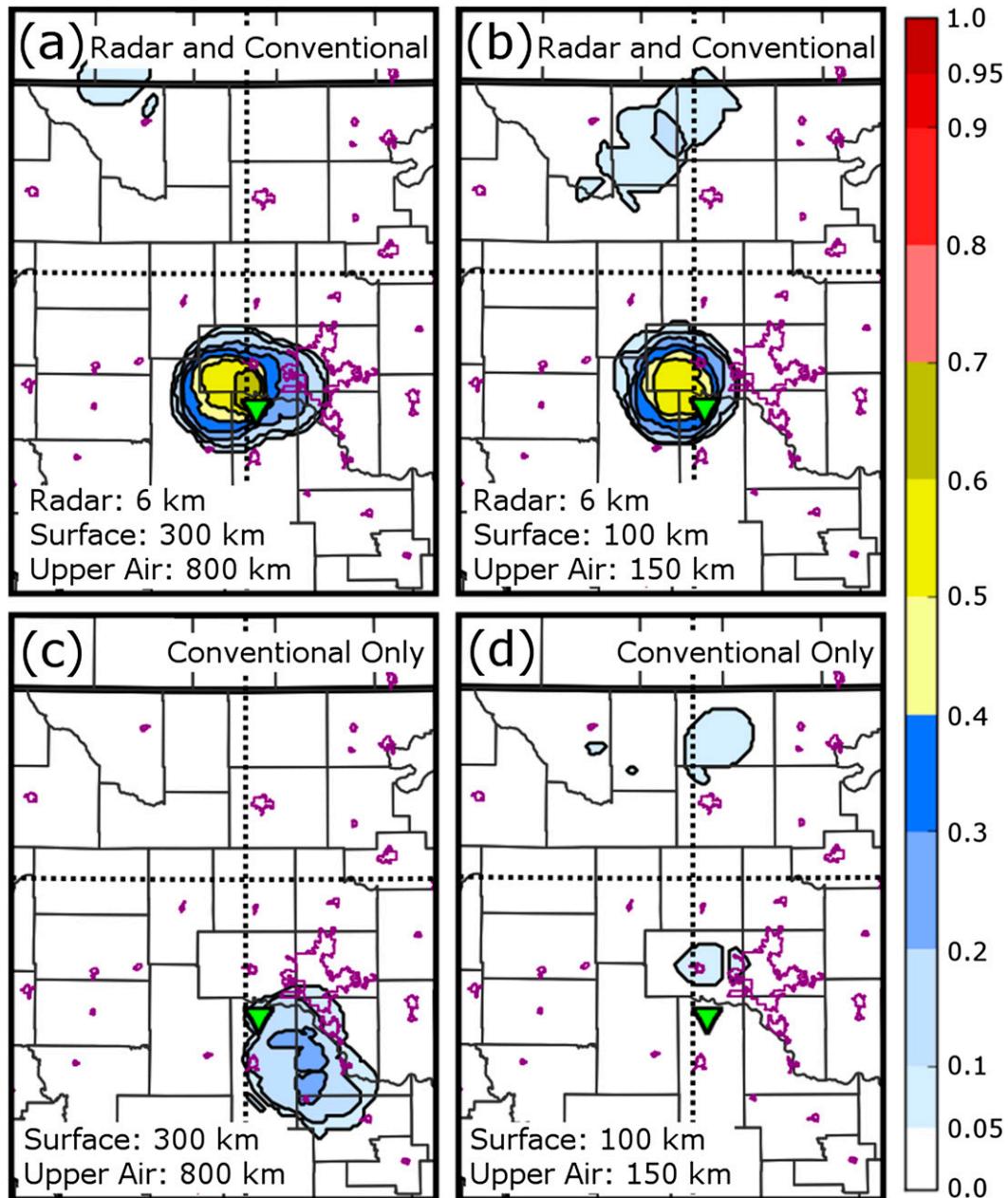


FIG. 17. Ensemble-based probability of a significant near-surface mesovortex occurring within 25 km of a point (shaded) at 0400 UTC for (a) RADCONV, (b) a variant of RADCONV using reduced localization radii for conventional data assimilation, (c) CONV, and (d) a variant of CONV using reduced localization radii for conventional data assimilation. The location of the observed tornadic mesovortex at 0400 UTC is indicated by the green triangle in each panel. Urban boundaries are shown in purple. The localization radius used to assimilate surface, upper-air, and (if used) radar observations, is noted in each panel.

ensemble assimilating both radar and conventional observations, though the result of assimilating radar data alone was almost as good. Assimilation of conventional observations alone resulted in modest improvement over the control experiment during the analysis period and first hour of the ensemble forecast.

In comparison, the positive impact of assimilating radar observations is retained throughout the 3-h forecast period. Assimilating radar data imparted the most skill to probabilistic reflectivity forecasts for thresholds that include light and/or moderate precipitation (i.e., 10–40 dBZ).

When radar and conventional data were assimilated together, the resulting probabilistic forecasts of  $Z$  closely resemble those obtained by assimilating radar data alone, in terms of structure, bias, and skill, although the positive impact of conventional data is evident in AUC skill assessments in the 1- and 2-h ensemble forecasts. Though conventional observations are relatively sparse, particularly above the surface, and primarily measure quantities only indirectly related to precipitation (such as temperature, humidity, and wind velocity), they provide valuable information on the near-surface temperature, moisture, and wind fields, all of which can help improve the accuracy of forecasts of convective storms.

Probabilistic forecasts in all experiments overpredicted the spatial extent of precipitation. A high-bias in heavy precipitation and a low-bias in light precipitation ( $>20$  dBZ) were noted in the radar-assimilating ensembles. The high bias in heavy precipitation was absent in the ensemble assimilating conventional data only, but a strong high bias in light-to-moderate precipitation was noted. When radar and conventional data were assimilated together, the bias behavior was similar to that of the ensemble assimilating radar data alone.

All four forecast ensembles predicted a discernible ( $>0.2$ ) probability of a near-surface mesovortex being present near the observed mesovortex location in a 2-h forecast. The experiment assimilating both radar and conventional observations produced the most accurate prediction, with a region of moderately high (maximum  $>0.6$ ) probability concentrated near and just to the northwest of the observed mesovortex. The forecast in the ensemble assimilating radar data alone yielded a more dispersed region of low-to-moderate (maximum  $<0.4$ ) probability near the observed mesovortex and to the north and west. Both radar-assimilating forecasts outperformed the other two ensembles (which assimilated conventional data only or no data at all). While conventional observations can help to improve the near-surface wind field, as found by Schenkman et al. (2011b), it appears that they are only beneficial if the ensemble already has a reasonably accurate representation of the MCS, which for this case requires the assimilation of radar observations. In short, for prediction of the mesovortex in this case, conventional data provided a strong benefit, but only when assimilated alongside radar observations. This is consistent with the findings of the observing system simulation experiments in Dong et al. (2011).

Adding conventional observations resulted in substantial improvements to 2-h ensemble-mean forecasts of surface dewpoint and temperature. Forecast error, compared against Oklahoma Mesonet observations, were reduced somewhat over the region near the MCS for 2-m temperature when both radar and conventional

data were assimilated compared to when just radar data were used. In the ensemble assimilating radar data alone, a substantial dry bias was present in the 2-h surface dewpoint forecast—this bias was greatly reduced in the ensemble assimilating both radar and conventional data. The lowest biases in those variables are found when conventional data were assimilated alone. These reductions in bias show the value of surface observations—particularly surface observations with high temporal and spatial resolution—for convective-scale forecasts.

The choice of horizontal localization radius is found to be important for assimilation of conventional observations. Localization radii of 100–300 km were tested for surface observations, and 150–800 km for upper-air observations. Larger radii were found to produce better ensemble forecasts for precipitation and mesovortices. We theorize that, for this case, since the bulk of conventional data was confined to the area covered by the Oklahoma Mesonet, using larger localization radii helped spread the information in the mesonet surface observations upstream, allowing its benefit to persist longer in the ensemble forecasts. Sobash and Stensrud (2013) found that convective-scale EnKF DA and forecasts of an MCS, obtained using an OSSE framework, were sensitive to the covariance localization used for radar observations. While we did not vary covariance localization radii for radar observations in this study, this topic has been actively investigated through numerical experiments at CAPS. At 1–2-km grid spacing, a localization radius of 6 km has been found to work well for radar observations (Xue et al. 2006).

Though the experiments presented in this paper yield skillful ensemble analyses and forecasts, we note that some shortcomings remain. The experiments presented in this study use a single-moment microphysical scheme; for this case, Putnam et al. (2014) found that using a dual-moment microphysical scheme resulted in improved representation of the trailing convective line, and better representation of dual-polarimetric radar signatures in emulated radar data obtained from model forecasts. Though the use of a dual-moment microphysical scheme increases the computational cost and complexity of DA and NWP, it offers a promising means of improving analysis and forecast quality.

Looking forward toward the implementation of real-time ensemble-forecast-based severe weather warnings, as envisioned in the warn-on-forecast paradigm (Stensrud et al. 2009), the relative benefit of various data sources should be considered together with their associated assimilation costs. As efficient parallel EnKF algorithms suitable for dense observations are developed (e.g., Wang et al. 2013), real-time implementation of such systems is increasingly within reach. Data assimilation experiments

run in a quasi-operational real-time environment could provide opportunities to examine the impact of multi-scale data sources on the forecasting of a large number of cases, leading to more robust conclusions. Such studies should be pursued in the future.

*Acknowledgments.* This work was primarily supported by NSF Grant EEC-0313747 as part of the ERC CASA program. The second author was also supported by NSF Grants AGS-0802888, AGS-1046171, AGS-0738370, AGS-0608168, AGS-0750790, AGS-0941491, and OCI-0905040. The authors would like to acknowledge Keith Brewster and Kevin Thomas for their assistance in data preparation, as well as Altug Aksoy and two anonymous reviewers for their constructive comments, which helped to improve this manuscript. Computing was performed primarily at the Kraken supercomputer of the National Institute for Computational Sciences (NICS).

#### REFERENCES

- Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, doi:10.1175/2008MWR2691.1.
- , —, and —, 2010: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Wea. Rev.*, **138**, 1273–1292, doi:10.1175/2009MWR3086.1.
- , S. Lorsolo, T. Vukicevic, K. J. Sellwood, S. D. Aberson, and F. Zhang, 2012: The HWRF Hurricane Ensemble Data Assimilation System (HEDAS) for high-resolution data: The impact of airborne Doppler radar observations in an OSSE. *Mon. Wea. Rev.*, **140**, 1843–1862, doi:10.1175/MWR-D-11-00212.1.
- , S. D. Aberson, T. Vukicevic, K. J. Sellwood, S. Lorsolo, and X. Zhang, 2013: Assimilation of high-resolution tropical cyclone observations with an ensemble Kalman filter using NOAA/AOML/HRD's HEDAS: Evaluation of the 2008–11 vortex-scale analyses. *Mon. Wea. Rev.*, **141**, 1842–1865, doi:10.1175/MWR-D-12-00194.1.
- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.
- , and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, doi:10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.
- Clark, A. J., W. A. Gallus Jr., M. Xue, and F. Kong, 2009: A comparison of precipitation forecast skill between small convection-permitting and large convection-parameterizing ensembles. *Wea. Forecasting*, **24**, 1121–1140, doi:10.1175/2009WAF2222222.1.
- Crum, T. D., R. L. Alberty, and D. W. Burgess, 1993: Recording, archiving, and using WSR-88D data. *Bull. Amer. Meteor. Soc.*, **74**, 645–653, doi:10.1175/1520-0477(1993)074<0645:RAAUWD>2.0.CO;2.
- Dawson, D. T., II, 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, doi:10.1175/MWR-D-11-00008.1.
- Dirren, S., R. D. Torn, and G. J. Hakim, 2007: A data assimilation case study using a limited-area ensemble Kalman filter. *Mon. Wea. Rev.*, **135**, 1455–1473, doi:10.1175/MWR3358.1.
- Dong, J., M. Xue, and K. K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phys.*, **112**, 41–61, doi:10.1007/s00703-011-0130-3.
- Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble forecasting and data assimilation. *J. Atmos. Oceanic Technol.*, **26**, 911–927, doi:10.1175/2008JTECHA1156.1.
- , —, and D. J. Stensrud, 2004: High-resolution analyses of the 8 May 2003 Oklahoma City storm. Part II: EnKF data assimilation and forecast experiments. *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., 12.5. [Available online at <https://ams.confex.com/ams/pdfpapers/81393.pdf>.]
- , —, and C. Snyder, 2011: Ensemble Kalman filter data assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, doi:10.1175/2010MWR3438.1.
- Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480, doi:10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2.
- 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.
- , 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, doi:10.1007/s10236-003-0036-9.
- Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919, doi:10.1175/1520-0493(2000)128<2905:AHEKFF>2.0.CO;2.
- , and J. S. Whitaker, 2011: What constrains spread growth in forecasts initialized from ensemble Kalman filters? *Mon. Wea. Rev.*, **139**, 117–131, doi:10.1175/2010MWR3246.1.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.
- , —, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620, doi:10.1175/MWR-2864.1.
- Jung, Y., G. Zhang, and M. Xue, 2008: Assimilation of simulated polarimetric radar data for a convective storm using ensemble Kalman filter. Part I: Observation operators for reflectivity and polarimetric variables. *Mon. Wea. Rev.*, **136**, 2228–2245, doi:10.1175/2007MWR2083.1.
- , M. Xue, and M. Tong, 2012: Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric radar data. *Mon. Wea. Rev.*, **140**, 1457–1475, doi:10.1175/MWR-D-11-00032.1.
- 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.

- 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.
- Putnam, B. J., M. Xue, Y. Jung, N. Snook, and G. Zhang, 2014: The analysis and prediction of microphysical states and polarimetric variables in a mesoscale convective system using double-moment microphysics, multinet radar data, and the ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 141–162, doi:10.1175/MWR-D-13-00042.1.
- Schenkman, A. D., M. Xue, A. Shapiro, K. Brewster, and J. Gao, 2011a: The analysis and prediction of the 8–9 May 2007 Oklahoma tornadic mesoscale convective system by assimilating WSR-88D and CASA radar data using 3DVAR. *Mon. Wea. Rev.*, **139**, 224–246, doi:10.1175/2010MWR3336.1.
- , —, —, —, and —, 2011b: Impact of CASA radar and Oklahoma mesonet data assimilation on the analysis and prediction of tornadic mesovortices in a MCS. *Mon. Wea. Rev.*, **139**, 3422–3445, doi:10.1175/MWR-D-10-05051.1.
- Schwartz, C. S., and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280, doi:10.1175/2009WAF2222267.1.
- Snook, N., and M. Xue, 2008: Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms. *Geophys. Res. Lett.*, **35**, L24803, doi:10.1029/2008GL035866.
- , —, and Y. Jung, 2011: Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Wea. Rev.*, **139**, 3446–3468, doi:10.1175/MWR-D-10-05053.1.
- , —, and —, 2012: Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data. *Mon. Wea. Rev.*, **140**, 2126–2146, doi:10.1175/MWR-D-11-00117.1.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, doi:10.1175/2555.1.
- Sobash, R. A., and D. J. Stensrud, 2013: The impact of covariance localization for radar data on EnKF analyses of a developing MCS: Observing system simulation experiments. *Mon. Wea. Rev.*, **141**, 3691–3709, doi:10.1175/MWR-D-12-00203.1.
- Stensrud, D. J., and Coauthors, 2009: Convective-scale Warn-on-Forecast System: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499, doi:10.1175/2009BAMS2795.1.
- Tong, M., 2006: Ensemble Kalman filter assimilation of Doppler radar data for the initialization and prediction of convective storms. School of Meteorology, University of Oklahoma, 243 pp.
- , and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, doi:10.1175/MWR2898.1.
- , and —, 2008a: Simultaneous estimation of microphysical parameters and atmospheric state with radar data and ensemble square-root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, **136**, 1630–1648, doi:10.1175/2007MWR2070.1.
- , and —, 2008b: Simultaneous estimation of microphysical parameters and atmospheric state with radar data and ensemble square-root Kalman filter. Part II: Parameter estimation experiments. *Mon. Wea. Rev.*, **136**, 1649–1668, doi:10.1175/2007MWR2071.1.
- Wang, Y., Y. Jung, T. A. Supinie, and M. Xue, 2013: A hybrid MPI/OpenMP parallel algorithm and performance analysis for an ensemble square root filter suitable for dense observations. *J. Atmos. Oceanic Technol.*, **30**, 1382–1397, doi:10.1175/JTECH-D-12-00165.1.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.
- Wood, V. T., and R. A. Brown, 1997: Effects of radar sampling on single-Doppler velocity signatures of mesocyclones and tornadoes. *Wea. Forecasting*, **12**, 928–938, doi:10.1175/1520-0434(1997)012<0928:EORSOS>2.0.CO;2.
- Wu, C.-C., G.-Y. Lien, J.-H. Chen, and F. Zhang, 2010: Assimilation of tropical cyclone track and structure based on the ensemble Kalman filter (EnKF). *J. Atmos. Sci.*, **67**, 3806–3822, doi:10.1175/2010JAS3444.1.
- Xue, M., 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.
- , and Coauthors, 2008: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment. *24th Conf. on Several Local Storms*, Savannah, GA, Amer. Meteor. Soc., 12.2. [Available online at [https://ams.confex.com/ams/24SL/techprogram/paper\\_142036.htm](https://ams.confex.com/ams/24SL/techprogram/paper_142036.htm).]
- , 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.
- , M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Oceanic Technol.*, **23**, 46–66, doi:10.1175/JTECH1835.1.
- , Y. Jung, and G. Zhang, 2010: State estimation of convective storms with a two-moment microphysics scheme and an ensemble Kalman filter: Experiments with simulated radar data. *Quart. J. Roy. Meteor. Soc.*, **136**, 685–700, doi:10.1002/qj.593.
- Yussouf, N., and D. J. Stensrud, 2012: Comparison of single-parameter and multiparameter ensembles for assimilation of radar observations using the ensemble Kalman filter. *Mon. Wea. Rev.*, **140**, 562–586, doi:10.1175/MWR-D-10-05074.1.
- , 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**, 3388–3411, doi:10.1175/MWR-D-12-00237.1.
- Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253, doi:10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.
- Zhu, K., Y. Pan, M. Xue, X. Wang, J. S. Whitaker, S. G. Benjamin, S. S. Weygandt, and M. Hu, 2013: A regional GSI-based ensemble Kalman filter data assimilation system for the Rapid Refresh configuration: Testing at reduced resolution. *Mon. Wea. Rev.*, **141**, 4118–4139, doi:10.1175/MWR-D-13-00039.1.