# Sensitivity of Convective Initiation Prediction to Near-Surface Moisture When Assimilating Radar Refractivity: Impact Tests Using OSSEs

NICHOLAS A. GASPERONI AND MING XUE

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

ROBERT D. PALMER

*School of Meteorology, and Advanced Radar Research Center, University of Oklahoma, Norman, Oklahoma*

JIDONG GAO

*NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*

ABSTRACT

The Advanced Regional Prediction System (ARPS) three-dimensional variational (3DVAR) system is enhanced to include the analysis of radar-derived refractivity measurements. These refractivity data are most sensitive to atmospheric moisture content and provide high-resolution information on near-surface moisture that is important to convective initiation (CI) and precipitation forecasting. Observing system simulation experiments (OSSEs) are performed using simulated refractivity data. The impacts of refractivity on CI and subsequent forecasts are investigated in the presence of varying observation error, radar location, data coverage, and different uncertainties in the background field. Cycled refractivity assimilation and forecasts are performed and the results compared to the truth. In addition to the *perfect model* experiments, *imperfect model* experiments are performed where the forecasts use the Weather Research and Forecasting (WRF) model instead of the ARPS. A simulation for the 19 May 2010 central plain convection case is used for the OSSEs. It involves a large storm system, large convective available potential energy, and little convective inhibition, allowing for CI along a warm front in northern Oklahoma and ahead of a dryline later to the southwest. Emphasis is placed on the quality of moisture analyses and the subsequent forecasts of CI. Results show the ability of refractivity assimilation to correct low-level moisture errors, leading to improved CI forecasts. Equitable threat scores for reflectivity are generally higher when refractivity data are assimilated. Tests show small sensitivity to increased observational error or ground clutter coverage, and greater sensitivity to the limited data coverage of a single radar.

## 1. Introduction

One of the most important variables related to convective-scale forecasting is the near-surface moisture field. The timing and location of convective initiation (CI) is often highly sensitive to moisture within the boundary layer (BL). Variations as small as $1\,g\,kg^{-1}$ in specific humidity, which are typical of boundary layer moisture (Weckwerth et al. 1996), can make the difference in whether or not storm initiation occurs. Xue and Martin (2006a,b) performed a high-resolution modeling study of the 24 May 2002 dryline CI case during the International H$_2$O Project (IHOP_2002; Weckwerth et al. 2004). Results of their assimilation and forecast experiments show a strong link between the low-level moisture pattern and the CI location and timing. Several studies have discussed and shown that high spatial resolution and accuracy in low-level moisture is key to CI forecasting, and the absence of such measurements is a major obstacle (e.g., Crook 1996; Dabberdt and Schlatter 1996; Emanuel et al. 1995; Koch et al. 1997; Weckwerth 2000). Weckwerth (2000) concluded that 100-m spatial and 10-min temporal resolutions are required for moisture

*Corresponding author address:* Dr. Ming Xue, Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: mxue@ou.edu

measurements to sufficiently sample boundary layer phenomena that lead to CI.

High-resolution observations of moisture are, however, notoriously difficult to measure. Surface observation networks, such as the Oklahoma Mesonet (hereafter Mesonet; Brock et al. 1995; McPherson et al. 2007), provide mesoscale observations on the order of 20–30-km spatial resolution at best. Measurements at even higher resolutions will typically require remote sensing approaches, because of the costs and practical limitations deploying very dense in situ observing instruments.

Fabry et al. (1997) developed a technique for obtaining atmospheric refractivity measurements from radar using phase measurements of stationary ground clutter targets. Since refractivity is most sensitive to moisture during the warm season, it is often used as a proxy for moisture (Bodine et al. 2011; Fabry 2004; Fabry et al. 1997; Gao et al. 2008). Processed refractivity data from radar generally have a resolution of about 2–4 km spatially and 4–10 min temporally, depending on radar type, scan strategy, and clutter target density. Given the high demand for high-resolution moisture measurements, studies have been conducted concerning the meteorological applications of refractivity measurements (e.g., Bodine et al. 2009, 2011; Fabry 2004; Weckwerth et al. 2005). In particular, IHOP_2002 contained a convective initiation component (Weckwerth and Parsons 2006) that also exploited the utilities of radar refractivity data. Weckwerth et al. (2005) showed for IHOP_2002 cases good correlations between the refractivity-based moisture measurements and surface station observations. In particular, they showed that refractivity measurements were representative of the lower boundary layer (approximately the lowest 250 m), especially under well-mixed conditions.

Studies examining the impact of refractivity data on numerical prediction of convective storms have been limited. Refractivity-derived mixed layer moisture was used to initialize a cloud model in Montmerle et al. (2002) and was believed to lead to more vigorous storm development that was in a better agreement with observations. In the review article of Sun (2005), the impact of including refractivity data on the 30-min prediction of a convective storm was briefly discussed. It was suggested that both the increased magnitude and variation of humidity due to refractivity data contributed to the initiation of part of a convective storm in the case presented. These represent preliminary studies examining the impact of refractivity data on convective storm initiation and prediction.

The current study focuses on the assimilation of refractivity measurements directly in a numerical weather prediction (NWP) framework, building upon the results of a simpler 2D framework by Shimose et al. (2013). The focus is to determine the potential impact of assimilating refractivity observations on CI and subsequent storm forecasts in an NWP model. The Advanced Regional Prediction System (ARPS; Xue et al. 2000, 2001, 2003) is used to create the truth simulation. During the data assimilation cycle and the ensuring forecast, either the ARPS or the Weather Research and Forecasting (WRF) model (Skamarock et al. 2005) is used, depending on whether the observing system simulation experiment (OSSE) uses a *perfect* (the same as that of truth simulation) model or an *imperfect* (different from the truth simulation) model. Both models are run using the Lin ice microphysics scheme (Lin et al. 1983). The planetary boundary layer (PBL) parameterization in the ARPS is based on a 1.5-order turbulent kinetic energy (TKE) scheme (Deardorff 1980) with special vertical mixing length treatment within the convective boundary layer after Sun and Chang (1986). WRF forecasts use the Mellow–Yamada–Janjić (MYJ) PBL scheme (Janjić 1990). As an initial effort, we chose to use simulation refractivity data through OSSEs, in order to test the sensitivity to a number of aspects related to refractivity measurements, such as instrument error and data coverage. With OSSEs, the known and complete truth state also allows unambiguous evaluation of quality and sensitivity. The ARPS three-dimensional variational (3DVAR; Gao et al. 2004) system is used for the data assimilation because of its simplicity and low cost. This system has been applied successfully to the assimilation of radar radial velocity and reflectivity data in many studies (e.g., Hu et al. 2006a,b; Schenkman et al. 2011a,b), but it is the first time that it is applied the refractivity data assimilation problem. To do that, an observation operator for radar refractivity measurements is developed and added to the ARPS 3DVAR. OSSEs are then performed using data created from the truth simulation using the same observation operator. Impacts of refractivity assimilation are also evaluated in terms of the forecasts produced by the ARPS and WRF models, in so-called perfect model and imperfect model forecast experiments, respectively.

The rest of this paper is organized as follows. In section 2, refractivity measurements are defined and the 3DVAR system and its configurations are described. The OSSE design is explained in section 3, and results are presented in section 4. The summary and conclusions are given in section 5.

## 2. Radar refractivity assimilation using ARPS 3DVAR

### a. Refractivity measurements

Doppler weather radars send out electromagnetic pulses that travel through the atmosphere at roughly the

speed of light $c$, modified only by the atmosphere's refractive index $n$. Although the refractive index is nearly constant at 1.003, deviations on the order of $10^{-6}$ near the surface are related to temperature, pressure, and water vapor pressure variations. For convenience, Bean and Dutton (1968) define refractivity $N$ in terms of refractive index $n$, which is related to meteorological variables as follows:

$$N = (n - 1) \times 10^6 = 77.6 \frac{P}{T} + 3.73 \times 10^5 \frac{e}{T^2}, \quad (1)$$

where $T$ is temperature (K), $P$ is atmospheric pressure (hPa), and $e$ is water vapor pressure (hPa). The units of refractivity are typically referred to as "$N$ units," though they are nondimensional in nature. At higher temperatures, refractivity is most sensitive to moisture (see Fig. 2 of Fabry et al. 1997) and is often used as a proxy for near-surface moisture.

Fabry et al. (1997) and Fabry (2004) describe in detail the basic concept and application of using Doppler weather radars to retrieve measurements of refractivity from the atmosphere, discussed in brief here. Radar refractivity retrievals are obtained from path-averaged phase measurements $\phi$ between the radar and ground clutter targets. Absolute phase measurements wrap every 5 cm for a 10-cm wavelength S-band radar, which presents a problem for obtaining meaningful measurements at typical range gate distances. To help mitigate this issue, reference phase measurements are taken at a reference observation time $t_{ref}$ when the atmosphere is relatively spatially homogeneous and can, in principle, be measured quite accurately by surface stations so that absolute refractivity can later be obtained. Given the refractivity at the reference time, subsequent measurements are the phase changes from the reference rather than the absolute phase. Since phase changes are usually much smaller than the absolute phase, phase wrapping will occur much less frequently. Taking a difference of phase measurements from some observation time $t$ and reference time $t_{ref}$ yields a phase change[1] measurement $\delta\phi$,

$$\delta\phi(r) = \phi(r,t) - \phi(r, t_{ref})$$
$$= -10^{-6} \frac{4\pi f}{c} \int_0^r [N(r',t) - N(r', t_{ref})] \, dr', \quad (2)$$

in terms of the phase change that is typically much smaller than the absolute phase [see also discussion in Shimose et al. (2013)]. Refractivity change measurements $\delta N$ can be obtained by taking a range derivative of these $\delta\phi$,

$$\delta N(r) = N(r,t) - N(r, t_{ref}) = -10^6 \left( \frac{c}{4\pi f} \right) \frac{\partial}{\partial r} [\delta\phi(r)], \quad (3)$$

where $f$ is the radar transmit frequency. These refractivity measurements are path-averaged quantities from radar to target site. One problem in applying (3) with a radar is that range $r$ in the measurement space is not continuous; phase change measurements are sampled at range gates. To get around this problem, Fabry (2004) showed that this derivative can be estimated using finite differences of phase change measurements between consecutive targets at ranges $R_1$ and $R_2$ according to

$$\overline{\delta N(R_1, R_2)} = \delta N(R_2) - \delta N(R_1)$$
$$\approx -10^6 \left( \frac{c}{4\pi f} \right) \frac{\delta\phi(R_2) - \delta\phi(R_1)}{R_2 - R_1}. \quad (4)$$

The two targets and the radar should be aligned along the same straight line for (4) to work; however, real ground clutter targets do not have this property naturally, meaning the phase change difference measurements can be very noisy. Additional sources of noise include differing clutter target heights within one range gate (e.g., building, trees); during times with high vertical refractivity gradients, target sway from vegetation, propagation delay of the radar beam, etc. (see, e.g., Bodine et al. 2011). Fabry (2004) found that vegetation sway is one of the largest error sources, possibly affecting refractivity from a single target by as much as $\pm 10$ $N$-units. Additional data processing techniques are required to reduce the noise and obtain reliable measurements, such as smoothing and censoring of bad-quality targets.

Absolute refractivity can be determined by summing a reference field of refractivity with the refractivity difference measurements. The refractivity algorithm used at the University of Oklahoma, developed by Cheong et al. (2008), uses an objective analysis of Mesonet data for a reference refractivity field, chosen from a time when the low-level refractivity is nearly homogeneous ($<5$-$N$-unit range over radar domain), the wind speed is low ($<5 \, m \, s^{-1}$), and there is no rainfall, for at least 10 consecutive radar scans to roughly ensure a steady state (Bodine et al. 2011). Even with these conditions, reference measurements may still yield poor-quality reference maps because of variations in clutter coverage. Prior to processing the PCD measurements into refractivity, a phase unwrapping procedure is applied. The algorithm includes censoring of poor-quality clutter targets

---

[1] As in Shimose et al. (2013), we use the term *phase change* to refer to the difference in phase of a target between two different times and *phase change difference* (PCD) to refer to the phase difference between two consecutive targets aligned along the same radial.

and smoothing to reduce noise, such that the final refractivity measurements have a spatial resolution of approximately 4 km, similar to that presented in Weckwerth et al. (2005).

### b. Three-dimensional variational analysis

In this study, ARPS is used as the simulation and forecast model. ARPS is a compressible nonhydrostatic storm-scale NWP model, and the system includes a 3DVAR data assimilation package developed by Gao et al. (2004). The ARPS 3DVAR minimizes a cost function $J$ that includes background, observation, and mass continuity constraint terms. Following standard convention defined in Ide et al. (1997), the cost function used in this work can be written as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)$$
$$+ \frac{1}{2}[\mathbf{y}_o - H(\mathbf{x})]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{y}_o - H(\mathbf{x})]. \qquad (5)$$

The first term is the background term $J_b$ containing the difference between state vector $\mathbf{x}$ and background vector $\mathbf{x}_b$, weighted by the inverse of the background error covariance matrix $\mathbf{B}$. The second term is the observation term, containing the differences between observation vector $\mathbf{y}_o$ and the state vector projected to the observation space, $H(\mathbf{x})$, weighted by the inverse of observation error covariance matrix $\mathbf{R}$ which contains instrumental as well as representativeness errors. In this study, state vector $\mathbf{x}$ includes the three wind components ($u$, $v$, and $w$), potential temperature $\theta$, pressure $P$, and specific humidity $q_v$. Here, the observation term includes only refractivity,

$$J_N = \frac{1}{2}[\mathbf{N}_o - H(\mathbf{x})]^{\mathrm{T}}\mathbf{R}_N^{-1}[\mathbf{N}_o - H(\mathbf{x})], \qquad (6)$$

where $\mathbf{N}_o$ is the vector of absolute refractivity observations from radar and $\mathbf{R}_N$ is the error covariance matrix of $N$ measurements. Additionally, the gradient of the $N$ cost function component needed by the ARPS 3DVAR is given in an incremental form,

$$\nabla J_N(\delta\mathbf{x}) = \mathbf{H}^{\mathrm{T}}\mathbf{R}_N^{-1}(\mathbf{H}\delta\mathbf{x} - \delta\mathbf{N}_o), \qquad (7)$$

where $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$ is the incremental state vector; $\delta\mathbf{N}_o = \mathbf{N}_o - H(\mathbf{x}_b)$ is the $N$ observation innovation vector; $\mathbf{H}$ is the linearized version for observation operator, $H$; and $\mathbf{H}^{\mathrm{T}}$ is its transpose or adjoint. The proper specifications of $\mathbf{R}_N$ and $\mathbf{B}$ are important to the 3DVAR analysis; $\mathbf{R}_N$ is often composed of only observation error variances as diagonal elements; off-diagonal covariance elements are commonly assumed to be zero, based on the assumption that observation errors are uncorrelated. In the ARPS

3DVAR system, the spatial correlation elements of $\mathbf{B}$ are modeled using recursive filters applied in each of the three directions. The spatial covariance is assumed to be Gaussian, isotropic, and spatially homogeneous. The reader is referred to Gao et al. (2004) for details of the recursive filter used in ARPS 3DVAR, as well as cost function minimization and preconditioning of the control variables.

Because refractivity depends on $P$, $T$, and $e$, according to (1), the analysis of refractivity will directly influence those three variables. In ARPS 3DVAR, pressure $P$, potential temperature $\theta$, and specific humidity $q_v$ are used as control variables that are linked to $P$, $T$, and $e$ through observation operator. For the current study, because no link exists between those three state variables and the wind field in the 3DVAR system, wind is not changed by the 3DVAR analysis directly, but changed through mutual adjustments in the assimilation cycles. Gradient checks were formed to ensure the new added codes for $N$ are correct.

## 3. OSSE experimental design

Before being entangled with uncertainties and unknown sources of error of refractivity observations, which has proven to be a significant issue (Bodine et al. 2011), we perform OSSEs (e.g., Atlas 1997; Lord et al. 1997) first, by creating simulated refractivity data from model fields of some "truth" or "nature run" simulation. Knowledge of the reference truth allows for unambiguous assessment of the quality of 3DVAR analyses and the subsequent model forecasts. The case of 19 May 2010 is chosen for our OSSEs. It was a day of widespread severe weather for much of Oklahoma and included several storms initiating ahead of a warm front and dryline.

One advantage of OSSEs is the ability to study individual components in the data assimilation and prediction system that may affect the analysis and subsequent forecast. In our case, sensitivities to observation error, domain coverage, and data discontinuity via realistic clutter fields are examined. Since this is a sensitivity study of CI to near-surface moisture, differing background moisture fields are used as the first guess in the 3DVAR analysis. Knowledge of the reference truth allows for direct calculation of verification statistics, such as root-mean-square error (RMSE) and the equitable threat score (ETS).

### a. Case overview, numerical simulation setup, and truth simulation

19 May 2010 was a day of severe weather outbreak in western and central Oklahoma, as there were more than 70 Storm Prediction Center (SPC) severe weather reports, including 16 tornadoes in Oklahoma alone. A

large deep-layer cyclone moved through the central and southern Great Plains, with an upper-level low pressure tracking through central Kansas. A well-defined warm front, dryline, and cold front are all evident by this time, with a clear triple-point intersection in west-central Oklahoma by 2100 UTC (all times will be in UTC, which will be omitted hereafter). An overnight mesoscale convective system left an outflow boundary that aided CI from the dryline–cold front triple point and along the warm front. With afternoon diabatic heating, surface temperature in the warm sector rose to 25°–30°C with dewpoint temperatures of 15°–20°C. Little convective inhibition (less than 25 J kg$^{-1}$) and large surface-based CAPE values (as high as 3000 J kg$^{-1}$) provided a favorable environment for convective initiation. As such, storms were initiated near the triple point between 1930 and 2030, with additional storm development east along the warm front and ahead of the dryline later on between 2130 and 2230.

The truth simulation uses the same configurations as those used by the Center for Analysis and Prediction of Storms (CAPS) for rapidly updated real-time forecasts (Brewster et al. 2010, 2008) for the Engineering Research Center (ERC) for Collaborative Adaptive Sensing of the Atmosphere (CASA; McLaughlin et al. 2009). It was initialized at 1930 from an ARPS 3DVAR analysis including available surface and Mesonet observations as well as available CASA and Next Generation Weather Radar (NEXRAD) data, using a temporally and spatially interpolated National Centers for Environmental Prediction (NCEP) operational North American Mesoscale Model (NAM) 12-km analysis as the background. The domain is 450 km × 420 km and has a horizontal resolution of 1 km. In the vertical, a stretched grid is used with a resolution of 20 m near the surface and 800 m at the model top, and an average vertical grid spacing of 400 m. The lateral boundary conditions used for the truth simulation were also taken from the NAM 12-km analysis and forecasts every 3 h. The truth simulation is run for five forecast hours, up until 0030 on 20 May 2010.

Figure 1 shows the moisture and reflectivity fields of the truth simulation between 2030 and 2330. The storm shown at 2030 is a cell that initiated within the first 15 min of the simulation; thus, it will be analyzed as a preexisting storm in subsequent assimilation experiments. Additional points of initiation occur in the model to the west-southwest of the preexisting storm between 2130 and 2200. The storm located within the red box in Fig. 1b is a single cell whose evolution is supercellular, and it initiates at about 2105, making it the first point of CI in the truth field (aside from the preexisting storm). This storm initiates close to, but ultimately ahead of, the dryline, in an area of enhanced moisture convergence

because of abundant moisture (above 15 g kg$^{-1}$) combined with localized wind convergence (not shown).

### b. Generation of simulated refractivity data

Simulated refractivity measurements are generated by simply calculating radar refractivity according to (1) from the model truth fields at the surface. The resolution of simulated data is 4 km, taken every four grid points, which is similar to the effective resolution of the final processed radar refractivity measurements from radar sites KTLX and KFDR by Cheong et al. (2008). For single radar, refractivity is assumed to be available within a 50-km radius of radar, and the coverage is assumed to be continuous (Fig. 2). The addition of error and other uncertainty will be discussed in conjunction with sensitivity experiments.

### c. OSSEs

A set of experiments with cycled analysis and forecasting are conducted, each following the general assimilation setup given in Fig. 3. The truth simulation is run for an hour to allow for the model fields to "spin up," that is, to develop more complex structures in all model fields (especially moisture) as permitted by the high-resolution model, prior to introducing uncertainty and performing assimilation experiments.

For comparison purposes, a pair of "baseline forecasts" without data assimilation, NoNd and NoNm (d for dry and m for moist; see Table 1), are initialized at 2030 from the error-containing analysis backgrounds and run for 4 h. The pair of analysis backgrounds used as the first guess for 3DVAR analysis at 2030 are created by introducing near-surface moisture error to the truth. The procedure includes three steps:

1) Smooth surface $q_v$ 50 times using a 25-point smoother;
2) Add or subtract 2 g kg$^{-1}$ to surface $q_v$ everywhere;
3) Introduce error into the boundary layer by spreading the specified surface error vertically according to the following equation:

$$q_v = q_{v,\text{truth}} + (1 - z/D)^2 (q_v^{\text{sfc}} - q_{v,\text{truth}}^{\text{sfc}}). \quad (8)$$

Here, $q_v^{\text{sfc}}$ is the surface moisture after step 2, subscript "truth" denotes the truth field, $z$ is height above ground, and $D$ is set to 4 km. For $z > D$, $q_v = q_{v,\text{truth}}$. Therefore, the moisture error decreases quadratically with height and becomes zero at 4 km AGL. Most of the error is confined below 1 km AGL. All other variables in the initial background are set to the values of the truth.

The background errors created above contain smaller-scale errors of random nature due to the smoothing
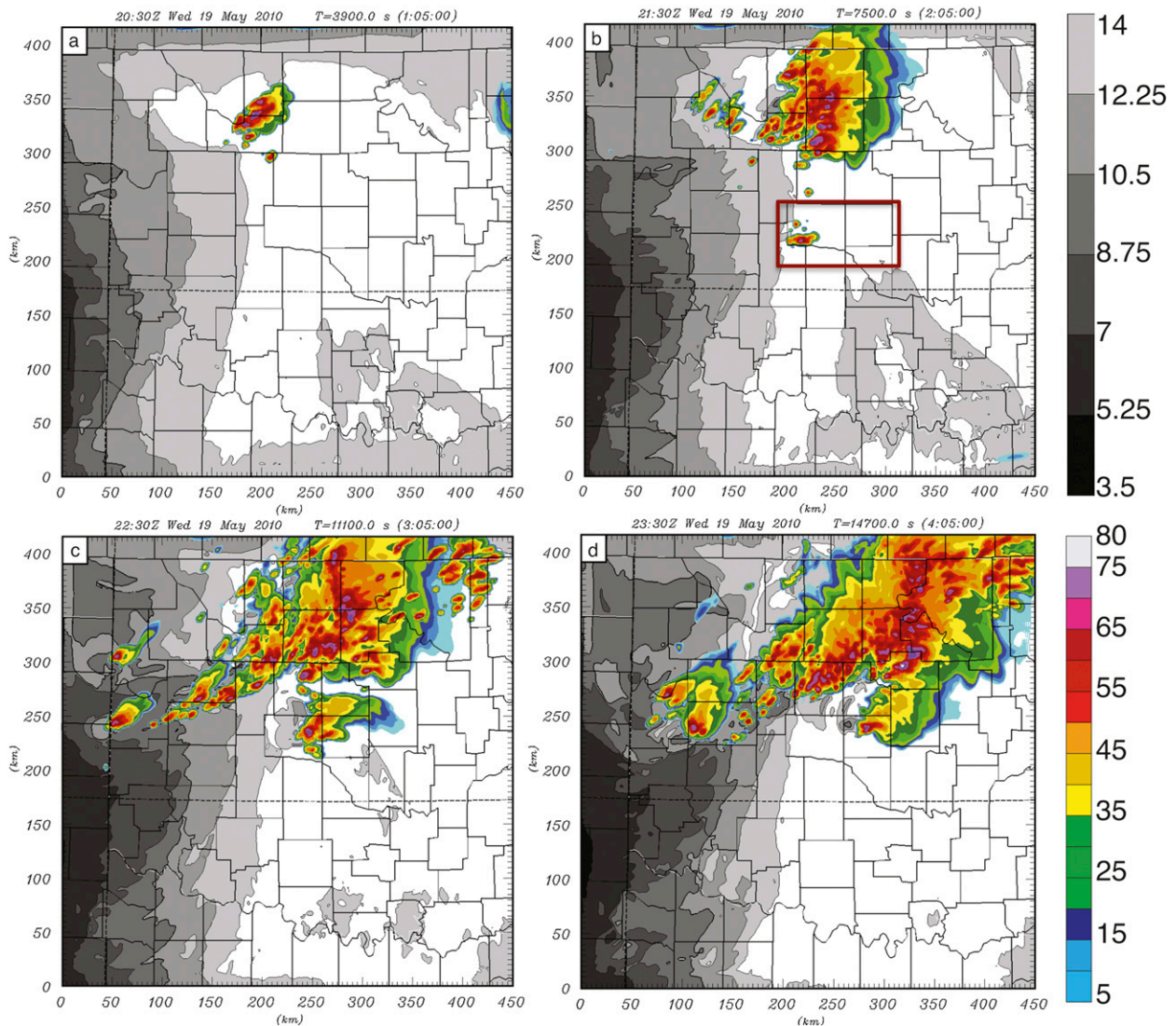
FIG. 1. Composite reflectivity (colored, dBZ) and surface specific humidity (lines and gray shaded contours, g kg$^{-1}$) for the truth model simulation of 19 May 2010 at (a) 2030, (b) 2130, (c) 2230, and (d) 2330 UTC. Red box in (b) represents a verification subdomain used for ETS calculations, located around the supercell that initiates at roughly 2105 in the truth simulation.

applied in the first step. Step 2 creates horizontally cor-related error of significant magnitude at the surface, while step 3 creates vertically correlated error in the lower atmosphere. These errors are representative of ran-dom errors, synoptic-scale horizontally correlated error, and boundary layer vertically correlated error, respec-tively. Because our experiments focus on the impact of refractivity data, no other types of observations are in-cluded; therefore, we have no way of correcting error in the upper atmosphere. For this reason, we limit the background moisture error to the lower atmosphere. Also, because we do not have other observations while the refractivity measurement is most sensitive to mois-ture, we limit the simulated background error to the

moisture field only. These are caveats of the current study that a reader should keep in mind, and more so-phisticated experiments with additional observation types using more advanced data assimilation methods should be explored in the future. As the first OSSE study, we choose to start with a relatively simple setup here.

The ARPS 3DVAR system assumes Gaussian error background covariance (e.g., Huang 2000; Liu and Xue 2006), and the spatial decorrelation scale $L$ in the Guas-sian correlation model is specified by the user (Gao et al. 2004). In practice, $L$ is often chosen to be comparable to the mean spacing of the observation network. Experi-ments indicated an optimal horizontal decorrelation length scale of 8 km in the horizontal (not shown) for
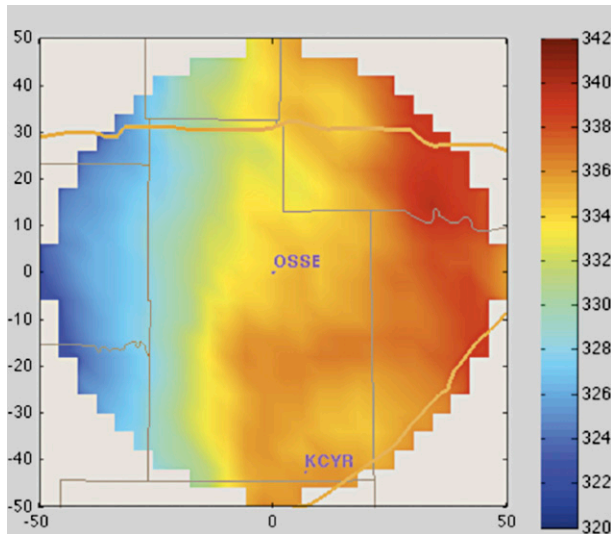
FIG. 2. Simulated refractivity data for a single radar with 50-km continuous data coverage centered near the CI location of the isolated supercell, valid 2030 UTC 19 May 2010. This hypothetical radar is centered at horizontal grid point pair $(i, j) = (200, 200)$, which is less than 50 km north of CASA radar KCYR (Cyril, OK).

our refractivity data, in general agreement with Shimose et al. (2013). Additionally, the optimal vertical decorrelation length was found to be six vertical grid levels, or roughly 1 km, for the model configuration used.

From the assimilation experiments, cycled analysis and forecasts are performed for an hour between 2030 and 2130 at 10-min intervals,[2] yielding a total of seven analyses. From the analysis at 2130, a 3-h forecast is run, and the results are compared to the baseline forecasts and the truth. Table 1 summarizes each of these OSSE experiments. With two background fields with plus and minus surface error, respectively, there are two control experiments, CNTLd and CNTLm. They are idealized experiments where the entire domain is completely covered by overlapping radars (Fig. 4a) and refractivity observations are available every 4 km and contain Gaussian errors with a 0.5-$N$-unit standard deviation. The value of 0.5 $N$-units was chosen in part due to assuming preprocessing using the algorithm of Cheong et al. (2008) censors out bad-quality target data. Some data are left with higher error, and for this reason sensitivity experiments with increasing standard deviation of error are included (explained below). The control experiments

(CNTLd and CNTLm) are compared to their respective baseline experiments with no data assimilation (NoNd and NoNm) and verified against the truth fields.

Other experiments listed in Table 1 are sensitivity experiments. We examine four kinds of sensitivities: (i) sensitivity to background field, (ii) sensitivity to observation error, (iii) sensitivity to realistic clutter coverage, and (iv) sensitivity to radar domain coverage. For (ii), there are three cases in which observation errors were randomly drawn from a zero-mean Gaussian distribution having standard deviations 0.5, 1.0, and 2.0 $N$-units, respectively. For (iii), realistic clutter coverage is applied to the full radar (FRAD) tests to study the effect of data discontinuity. First, a typical KTLX clutter domain is remapped to a Cartesian 4-km grid. This clutter domain is used for each radar in the full domain, with the addition of random rotation to avoid regular patterns in the clutter field. The resulting data domain coverage is roughly 50%–60% of the continuous full domain coverage (Fig. 4b). For sensitivity to radar coverage, experiments include an isolated radar with 50-km radius refractivity coverage as in Fig. 2, and the radar is placed near the supercell initiation point (see Fig. 1b).

Because the truth simulation is produced by the same model used in the assimilation and forecasting experiments, the OSSEs described so far are considered perfect model experiments. To avoid the caveats of such "identical twin" experiments that often lead to too optimistic conclusions, a second set of imperfect model experiments are conducted— one for each experiment in Table 1. For the imperfect experiments, the WRF serves as the forecast model within the assimilation cycles as well as for the final free forecast. Since the 3DVAR analysis is native to the ARPS system, this requires extra steps of interpolating the model fields from the ARPS grid to the WRF grid and back within each analysis cycle; because the setups of the horizontal grids of WRF and ARPS are identical, interpolation and variable conversion occur only in the vertical direction (WRF uses mass-based vertical coordinates, while ARPS uses height-based vertical coordinates). Because a different model is used, the forecast model WRF is considered to contain "error" compared to the truth simulation model, which is the ARPS.

### d. Verification methods—RMSE and ETS

A distinct advantage of OSSEs is the knowledge of the truth, which allows for direct calculation of verification statistics such as RMSE and ETS. RMSE can be calculated against the truth according to

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_{i,\text{est}} - x_{i,\text{truth}})^2}{n}}, \qquad (9)$$

---

[2] Note that the short 10-min assimilation cycles have the potential of introducing imbalance noise that may take longer than 10 min to settle down in the forecast model, although convective-scale data assimilation studies (e.g., Hu and Xue 2007) have successfully used such short intervals. A digital filter (Lynch and Huang 1992) can be applied in the forecasting step to reduce potential noise.
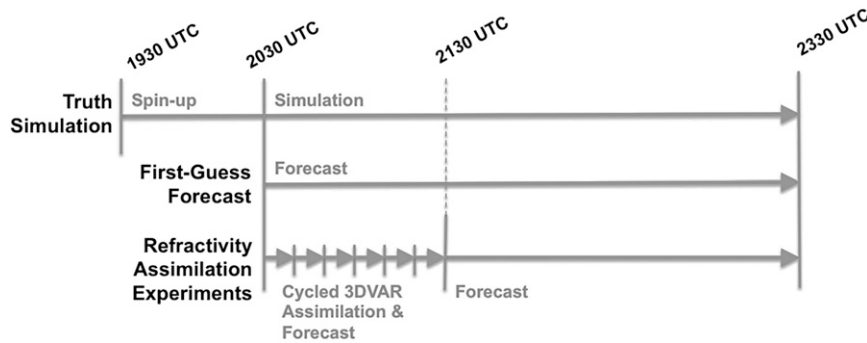
FIG. 3. Cycled assimilation and forecast experiment design. Truth model is initialized from a 3DVAR analysis without refractivity data at 1930 UTC and allowed to spin up for the first hour to give more complex structure to the moisture field. At 2030 UTC, the background is produced for refractivity assimilation experiments. Cycling window is an hour long, with analysis frequency of 10 min. At 2130 UTC, a forecast is run. Results of refractivity assimilation experiments are compared with forecast experiments without data assimilation (NoNd and NoNm in Table 1), which serve as baselines.

where $x$ is some model variable (specific humidity $q_v$ in this case), subscript $i$ refers to the grid point index, and $n$ is the number of grid points for which RMSE is calculated over.

To compare results of multiple sensitivity experiments, ETS (Schaefer 1990) is used as verification on the composite reflectivity fields. ETS is given by

$$ETS = \frac{hits - chance}{hits + misses + false\ alarms - chance}, \quad (10)$$

where hits are simultaneous forecast and observed events, misses are observed events that were not forecasted, false alarms are forecasted events that are not observed, and

TABLE 1. List of OSSE experiments. For abbreviations, FRAD refers to Full RADar domain (data available everywhere at 4-km resolution; see Fig, 4a); CNTL refers to CoNTroL experiments; NoN refers to no data assimilation experiments; 1RD refers to experiments with a single, isolated radar; "d" and "m" refer to backgrounds with $2\,g\,kg^{-1}$ *subtracted* or *added* (respectively); and numbers 0.5, 1.0, and 2.0 refer to homogeneous standard deviation error amounts added to all refractivity observations. FRADCLUT refers to full radar experiments with realistic CLUTter coverage (see Fig. 4b).

| Expt name | Refractivity assimilated? | Obs error ($N$ units) | *Add* or *subtract* $2.0\,g\,kg^{-1}$ for background $q_v$ at 2030 UTC | Radar domain coverage | Realistic clutter coverage (data discontinuity)? |
|---|---|---|---|---|---|
| CNTLd or FRADd0.5 | Yes | 0.5 | Subtract (dry) | Full | No |
| CNTLm or FRADm0.5 | Yes | 0.5 | Add (moist) | Full | No |
| NoNd | No | N/A | Subtract (dry) | N/A | N/A |
| NoNm | No | N/A | Add (moist) | N/A | N/A |
| FRADd1.0 | Yes | 1.0 | Subtract (dry) | Full | No |
| FRADd2.0 | Yes | 2.0 | Subtract (dry) | Full | No |
| FRADm1.0 | Yes | 1.0 | Add (moist) | Full | No |
| FRADm2.0 | Yes | 2.0 | Add (moist) | Full | No |
| FRADCLUTd0.5 | Yes | 0.5 | Subtract (dry) | Full | Yes |
| FRADCLUTd1.0 | Yes | 1.0 | Subtract (dry) | Full | Yes |
| FRADCLUTm0.5 | Yes | 0.5 | Add (moist) | Full | Yes |
| FRADCLUTm1.0 | Yes | 1.0 | Add (moist) | Full | Yes |
| 1RDd0.5 | Yes | 0.5 | Subtract (dry) | Isolated | No |
| 1RDm0.5 | Yes | 0.5 | Add (moist) | Isolated | No |
| FRADd-20Z | Yes | 0.5 | Subtract $q_v$ error at 2000 UTC, then forecast for 30 min to 2030 UTC using ARPS | Full | No |
| FRADm-20Z | Yes | 0.5 | Add $q_v$ error at 2000 UTC, then forecast for 30 min to 2030 UTC using ARPS | Full | No |
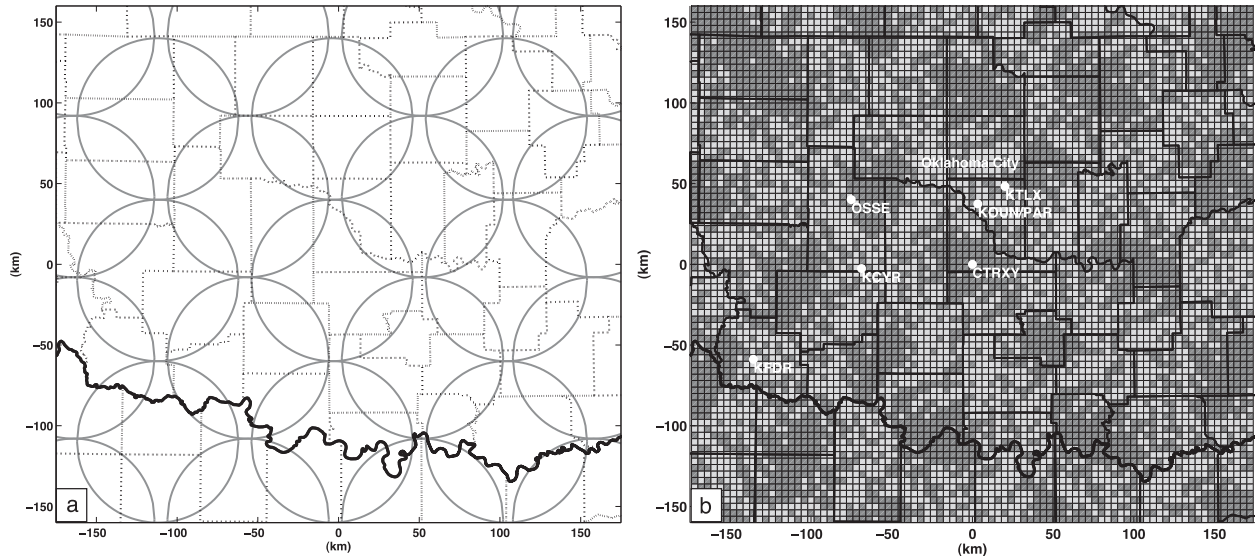
FIG. 4. (a) Depiction of the domain completely filled with radar coverage, for both control experiments as well as the FRAD experiments. (b) Clutter field corresponding to each "radar" in (a), where dark gray refers to points where refractivity is available from each radar, used for FRADCLUT sensitivity experiments. OSSE denoted in (b) refers to the location of the single radar for 1RD experiments (example data in Fig. 2).

chance is the number of hits one would get by random chance, defined as

$$\text{chance} = \frac{(\text{hits} + \text{misses})(\text{hits} + \text{false alarms})}{\text{hits} + \text{misses} + \text{false alarms} + \text{correct negatives}}.$$
(11)

The ETS is a common verification score used in meteorology, where a score of 0 shows no skill and a score of 1 is a perfect forecast.

## 4. Results

### a. Results of baseline and control experiments

The impacts of refractivity assimilation in CNTLd and CNTLm are best evaluated relative to their respective no-N-assimilation baseline experiments, NoNd and NoNm. Figure 5 shows time series plots of RMSEs of specific humidity at the surface. The baseline forecasts NoNd (Fig. 5a) and NoNm (Fig. 5b) show similar characteristics—the RMSE starts at $\sim 2\,\text{g kg}^{-1}$ because of the design of the first guess but through forecast decreases—to about $1\,\text{g kg}^{-1}$ after roughly 90 min. This reduction occurs because of convective mixing in the boundary layer, as can be seen in Fig. 6, because errors higher up are smaller. This surface error begins to increase again only within the last 30–60 min or in the fourth hour of forecast.

With 10-min 3DVAR analysis cycles in the control experiments, the time series plots of RMSEs in Fig. 5 look similar to typical "sawtooth" plots seen in radar data assimilation studies using the ensemble Kalman filter (e.g., Tong and Xue 2005). Looking at the analysis cycling results in each of CNTLd and CNTLm, they show similar downward trends. Both CNTLd (Fig. 5a) and CNTLm (Fig. 5b) reduce the moisture errors well below their respective background values in the first analysis at 2030, although in CNTLd it takes a second cycle to reduce most of the error and prevent much error growth in the subsequent forecast. Both control experiments reduce the RMSEs to $\sim 0.1\,\text{g kg}^{-1}$, a full order of magnitude smaller than the baseline error. With the free forecast beginning from final analysis at 2130, the RMSE begins to grow; however, assimilating refractivity causes the forecast error to remain below the corresponding baseline forecast. Two hours into the forecast, CNTLd and CNTLm have errors about $0.4\,\text{g kg}^{-1}$ less than in NoNd and NoNm, at which point the error growth becomes faster. In CNTLd, the forecast error exceeds that of the corresponding baseline about 3 h into the forecast, while in CNTLm the error remains lower throughout. Refractivity assimilation has thus a positive impact on moisture for at least 2–3 hours past the final analysis time.

To further assess the impact of $N$ on the moisture analysis, Fig. 6 shows vertical profiles of RMSE in moisture for the control and baseline experiments. Because of the error correlation in the vertical, the 3DVAR is able to correct moisture error throughout the boundary layer with a proper vertical background error correlation scale. In CNTLd and CNTLm, the moisture error is reduced to
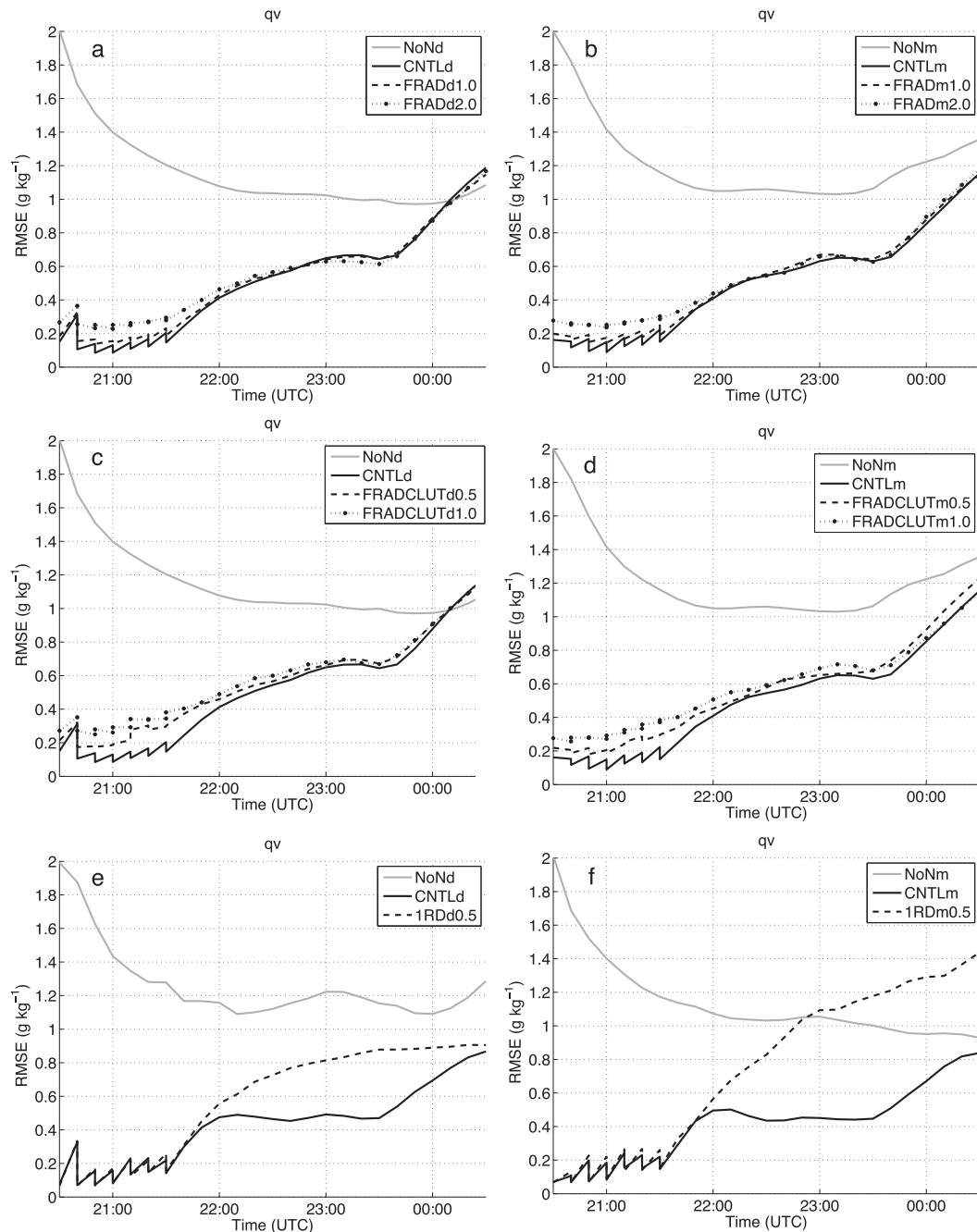
FIG. 5. RMSEs of the specific humidity field (g kg$^{-1}$) for sensitivity experiments listed in Table 1: (a),(b) for sensitivity to observation error experiments; (c),(d) for sensitivity to realistic clutter domain experiments; and (e),(f) for sensitivity to isolated radar experiments. Note that in (a)–(d), RMSEs were calculated over the whole domain, and in (e),(f) RMSEs were calculated only over the 50-km radius of the isolated radar.

less than 0.25 g kg$^{-1}$ at all levels at 2030, after the first analysis. This quite uniform reduction of the error in the vertical from surface observations benefited from our knowledge about the initial background error structure—the error was strongly correlated with the surface error. At 2130, the surface error remains low, but the errors away

from the surface are larger, having a maximum of about 0.7 g kg$^{-1}$ at 1.5 km AGL. The less effective removal of error in the later cycles is a result of less accurate modeling of the background error. During the assimilation cycles, the vertical error correlation structure is no longer simple—the static Gaussian covariance now works less effectively.
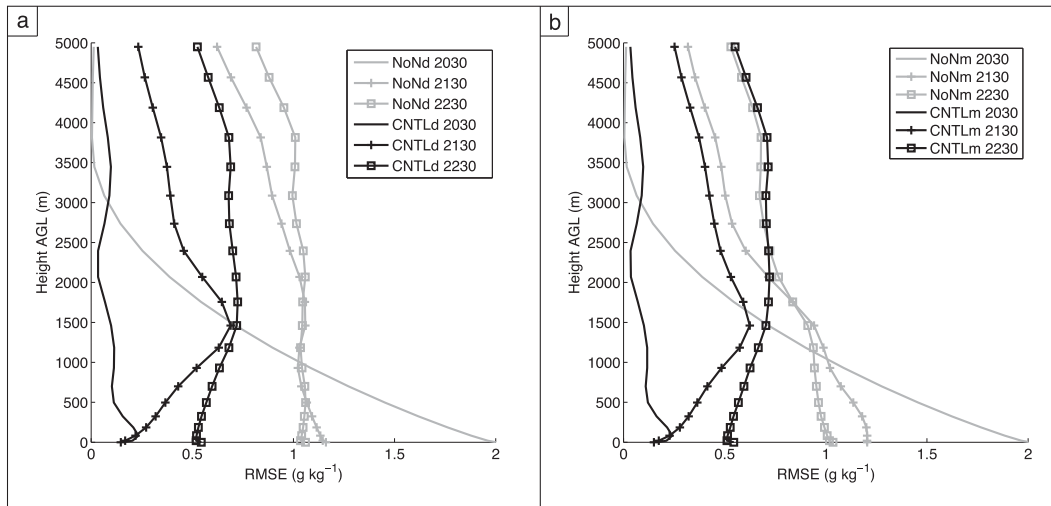
FIG. 6. RMSE vertical profiles of specific humidity at 2030 UTC (first analysis, no marker), 2130 UTC (final analysis, dashed marker) and 2230 UTC (1-h forecast, square marker) for the control (black) and no-data-assimilation (gray) experiments: (a) CNTLd, NoNd and (b) CNTLm, NoNm.

Despite that, the RMSEs of the control experiments are $0.5–1.0\,\mathrm{g\,kg^{-1}}$ below the corresponding baseline RMSE errors at 2130. A similar difference remains at the lower levels even after an hour of free forecast, at 2230. Above 1.5 km AGL, however, we begin to see some differences in the dry experiments versus the moist experiments. The RMSE in CNTLd (Fig. 6a) remains about $0.3\,\mathrm{g\,kg^{-1}}$ less than the RMSE of NoNd even at 2230. However, the errors of NoNm (Fig. 6b) converge with those of CNTLm by 2130 at the higher levels, mainly because of much slower error growth in NoNm than in its dry counterpart. Because of the limited ability of the refractivity data in directly correcting errors at the upper levels beyond the vertical error decorrelation scale (which is approximately 1 km in the boundary layer), the errors found at the upper levels contain accumulated effects of forecast error growth, and propagation of errors initially found at the lower levels. The vertical error propagation is also directly linked to convection within the model domain.

The effect of refractivity assimilation on predicted storms is shown in Fig. 7 for 2200. The preexisting storm from Fig. 1a has grown into a large multicell feature, with several points of new storm initiation to the west and south, including the single supercell that initiates just ahead of the dryline near the center of the domain. Figures 7b and 7c show the baseline forecasts for NoNd and NoNm, respectively. There is a clear contrast in these two experiments. As we expect, the dry baseline shows reduced convection, while the moist baseline shows enhanced convection. In particular, the differences seem to be mainly in new points of initiation, not in the preexisting storm. Whereas there are several new

points of initiation in the truth within the northwest portion of the domain, NoNd does not capture these storms. Reducing the moisture near the surface effectively suppresses these new storms, most notably the supercell in the center of the domain. Only a couple of new initiations occur in NoNd between 2215 and 2230 near $(x, y) = (150\,\mathrm{km}, 270\,\mathrm{km})$; however, these storms are 30–45 min late in initiation timing compared to the truth and quickly dissipate within an hour. Contrast this with NoNm (Fig. 7c), where not only are all points of initiation captured, but the storms have evolved more aggressively and there are several new points of spurious convection, especially near the supercell. Additionally, many points of CI occur earlier than in their respective locations in the truth, by as much as 30 min. The supercell area shows several points of CI that occur starting at 2055 in NoNm, which is 15 min early.

Figures 7d and 7e show the corresponding control forecasts from CNTLd and CNTLm. Both CNTLd and CNTLm show results that are qualitatively closer to the truth than either baseline experiment NoNd or NoNm. There are still a few points of spurious convection in CNTLm, though not nearly as many as in NoNm. Experiment CNTLd has recaptured the points of initiation to the west of the preexisting storm. The supercells in CNTLd and CNTLm are a bit different in general shape and size, with the supercell in CNTLm covering a larger area than in CNTLd; however, neither is clearly subjectively closer to the truth. All in all, refractivity assimilation has shown considerable positive impact on correctly capturing new points of initiation, to within less than 10 min of timing error.
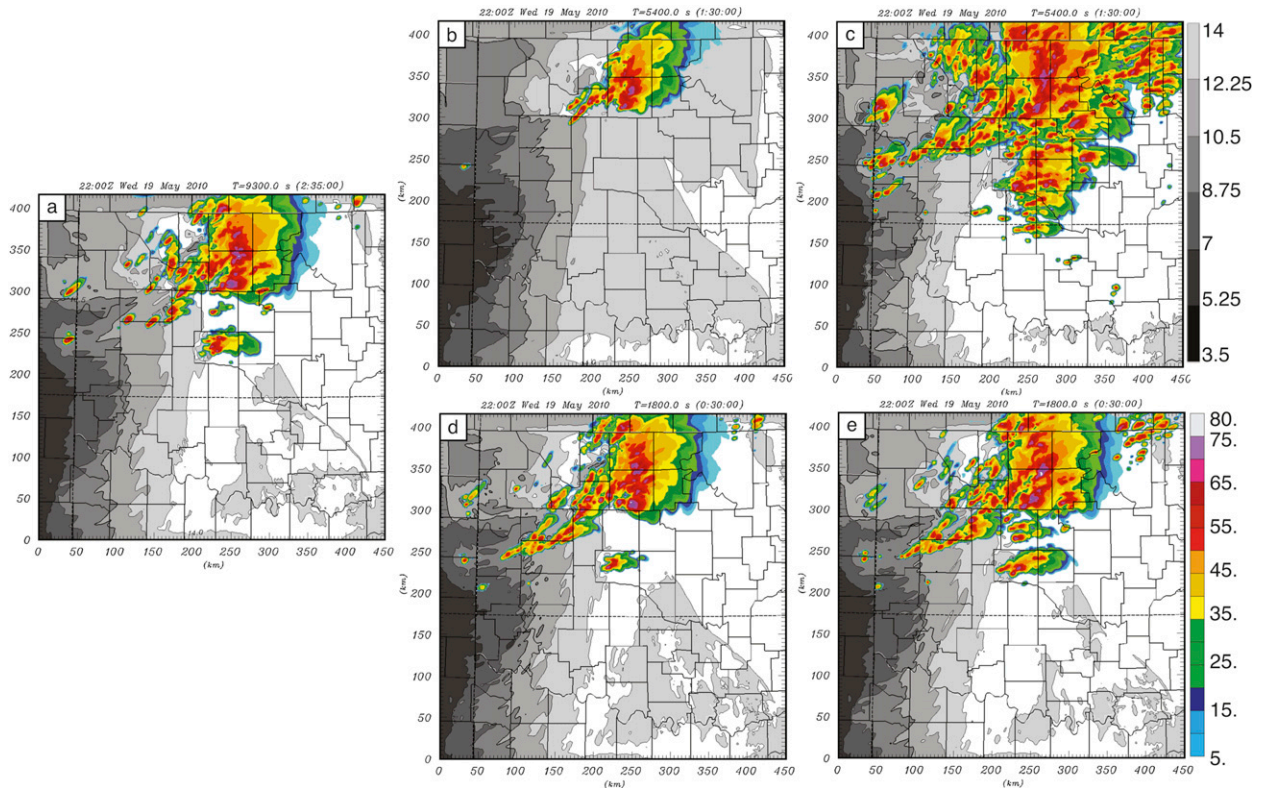
FIG. 7. Composite reflectivity (color fill, dBZ) and specific humidity (gray contours, $g\,kg^{-1}$) for (a) the truth field and perfect model experiments, (b) NoNd, (c) NoNm, (d) CNTLd, and (e) CNTLm, valid 2200 UTC 19 May 2010.

## b. Results of perfect model sensitivity experiments

### 1) SENSITIVITY TO OBSERVATION ERROR

In the presence of observation error (Figs. 5a,b), the results of moisture analysis cycling and forecast show significantly reduced RMSE compared to the baseline fields. For example, at 0.5 $N$-units of standard deviation error for the $N$ observations, the RMSE of CNTLd at 2100 is $0.1\,g\,kg^{-1}$. For 1.0 $N$-units it is $0.15\,g\,kg^{-1}$ and for 2.0 $N$-units it is $0.24\,g\,kg^{-1}$. Each of these RMSE values is an order of magnitude less than the baseline.

What do the error amounts of 0.5, 1.0, and 2.0 $N$-units mean in terms of moisture? To answer this question, we start with a formula converting standard deviation errors in model variables to the error of refractivity, assuming model variables $P$, $T$, and $e$ are uncorrelated:

$$\sigma_N^2 = \left(\frac{\partial N}{\partial P}\right)^2 \sigma_P^2 + \left(\frac{\partial N}{\partial T}\right)^2 \sigma_T^2 + \left(\frac{\partial N}{\partial e}\right)^2 \sigma_e^2. \quad (12)$$

Equation (14) is a first-order Taylor series approximation of the propagation of error, considering each distribution is approximately normal. It is rearranged to solve for vapor pressure deviation $\sigma_e$, assuming $T$ and $P$

errors are zero. Using saturation vapor pressure, this value can be converted to RH error $\sigma_{RH}$. For a warm-season $T$ of 20°C, RH errors of 0.5%, 1%, and 2% are approximately 0.5, 1.0, and 2.0 $N$-units, respectively, in terms of refractivity. Compared to typical NWP background errors in RH often are well above 5%, these error amounts are very small. Thus, observation errors up to 2 $N$-units should yield accurate analyses of moisture.

As expected, with increased observation error the analysis is unable to reduce RMSE as much as the control experiments; however, they still are significantly reduced over the baseline experiments NoNd and NoNm (Figs. 5a,b). Additionally, after just 30 min of forecast starting from the 2130 analyses, the RMSE curves converge to nearly the same values. So, increasing observation error up to 2 $N$-units has little detrimental effect on the moisture analysis and forecast. There are, however, some times when the 3DVAR analysis *increases* RMSE slightly in later cycles, such as FRADd2.0 at 2100 (Fig. 5a). In general, if 3DVAR is working properly this should not happen, assuming all errors are specified correctly. However, in these experiments the background error in moisture was kept at a static value of 20%, which is not valid after a few cycles of refractivity assimilation
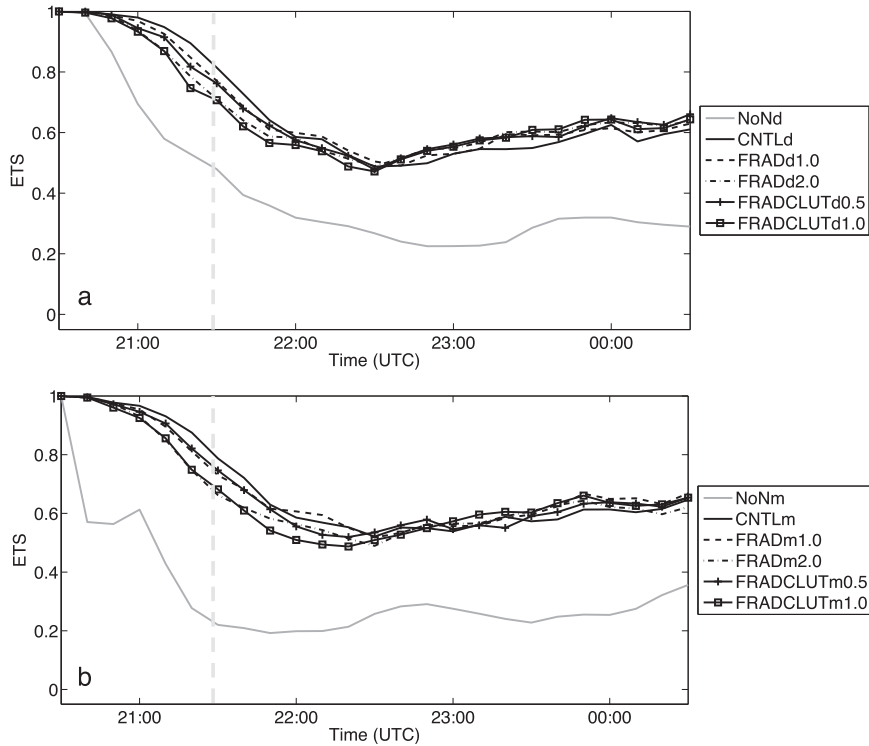
FIG. 8. ETS time series calculated using threshold of composite reflectivity at 30 dB$Z$ for sensitivity experiments listed in Table 1: (a) dry-bias experiments and (b) moist-bias experiments. ETS is calculated over the entire domain, and includes preexisting storm at initial time. Dashed vertical line indicates final analysis time at 2130 UTC used to initialize a 3-h free forecast.

because with subsequent cycling the background error should decrease as its accuracy increases. This is a general drawback of 3DVAR—the inability to dynamically update background error covariances with each new analysis.

Figure 8a shows reflectivity ETSs for the 30-dB$Z$ reflectivity threshold of control experiment CNTLd, as well as experiments with greater error added to observations, FRADd1.0 and FRADd2.0, calculated over the entire domain. Each ETS starts at a perfect 1.0 because hydrometeor fields in the initial background had no error with the preexisting storm. All experiments show higher scores than the baseline. There are only minor differences among the different assimilation experiments, no more than 0.1 in terms of ETSs. The results are similar for the moist experiments in Fig. 8b. Generally, CNTLd tends to have the best ETSs (though not always), likely because of its relatively low amount of observation error, while FRADd2.0 has lower ETSs (though still well above that of NoNd). The forecast of the supercell seems to be insensitive to increased observation error of up to 2 $N$-units, and all assimilation experiments show improvement over the corresponding no-data-assimilation experiment in terms of ETSs.

## 2) SENSITIVITY TO REALISTIC CLUTTER COVERAGE

The sensitivity of analysis and forecast to realistic clutter coverage, through experiments FRADCLUTd and FRADCLUTm, is shown in Figs. 5c and 5d in terms of analysis and forecast RMSEs. The error evolutions look very similar to corresponding ones in Figs. 5a and 5b. With realistic ground clutter having incomplete coverage, the RMSE is increased by ~0.1 g kg$^{-1}$ but the error converges after a short time in the forecast to the control experiment values. The effect of incomplete data coverage on the final RMSE by 2130 is slightly higher than the effect of increased observation error but, in general, realistic clutter coverage has little to no detrimental effect on the moisture analysis and forecast.

ETSs in Fig. 8 are lower for FRADCLUTd0.5, by roughly 0.1 point for the entire domain. However, the scores are still consistently above the baseline forecast, meaning that the use of realistic clutter coverage does not significantly impact the subsequent forecast. The ETS results are similar for FRADCLUTm0.5. The worst scores occur with FRADCLUTd1.0 and FRADCLUTm1.0, though the differences are still relatively minor. The
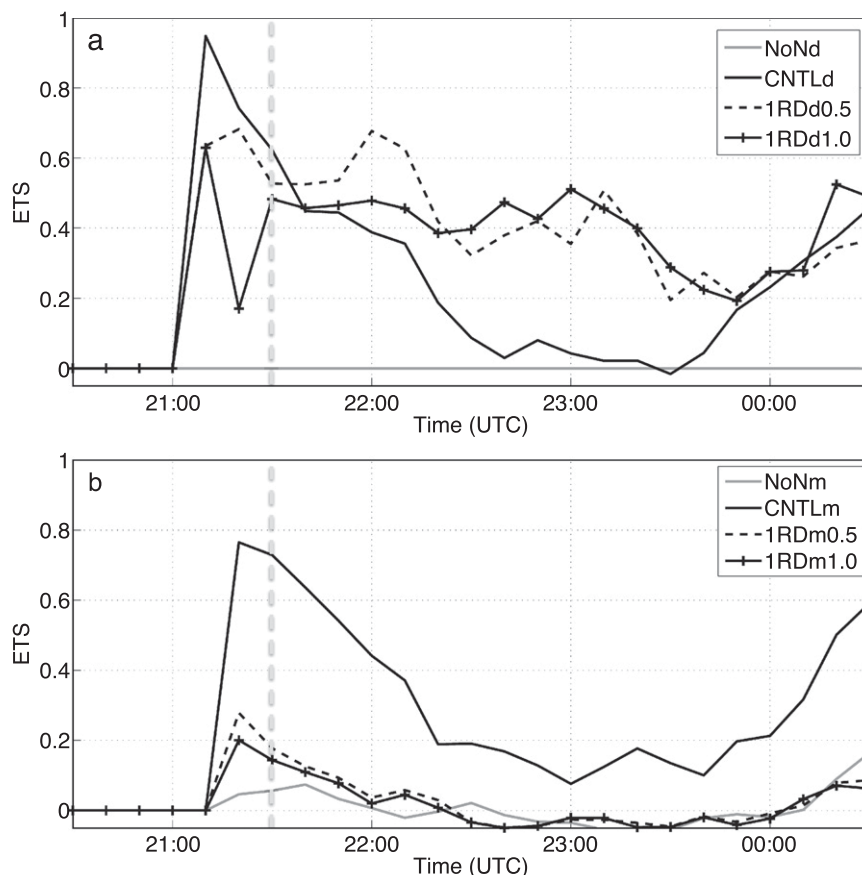
FIG. 9. ETS time series calculated over the verification subdomain (red box in Fig. 1b) for single-radar coverage experiments in Table 1: (a) dry-bias experiments and (b) moist-bias experiments. Control and no-data-assimilation experiments plotted for reference. Dashed vertical line indicates final analysis time at 2130 UTC as in Fig. 8.

results of realistic clutter coverage experiments tend to follow the general results of observation error experiments. Although a forecaster may be thrown off by the discontinuous appearance of clutter domain refractivity coverage at KTLX and KFDR, the data assimilation system is rather robust from a data assimilation and numerical forecasting standpoint. The caveat is that inevitably a real storm may initiate within a region where there is no clutter coverage, and then the impact of the data gaps may be larger.

### 3) SENSITIVITY TO ISOLATED RADAR DOMAIN COVERAGE

The RMSEs of the single isolated radar (1RD) experiments are shown in Figs. 5e and 5f. The RMSE calculations were done only over the actual coverage of the isolated radar. Experiments 1RDm0.5 and 1RDd0.5 have nearly identical RMSEs as their corresponding control experiment, having the same observation error. However, during the subsequent forecast the error curves

of 1RD experiments increase sharply, much more so than those of the CNTL and NoN experiments (either dry or moist). From 2300 on, 1RDm0.5 shows *greater* errors than baseline experiment NoNm. This indicates some negative effects of having only partial radar domain coverage; large gradients at the edge of the data coverage can cause structure differences in the simulated storms.

Figure 9a depicts the ETSs for the dry isolated radar experiments compared to the control and baseline experiments for the verification subdomain near the supercell. The 1RDd experiments show scores comparable to the CNTLd experiments, which are improved over the no-skill baseline. For the same experiments using a moist-biased background field (Fig. 9b), the scores are much lower and more comparable to NoNm than to CNTLm.

To look into this problem further, Fig. 10 plots reflectivity for the dry experiments CNTLd and 1RDd0.5 in 15-min intervals, starting at 2115 (approximate CI time of truth). Experiment 1RDd0.5 recaptures the specific CI location and timing of the truth pretty closely,
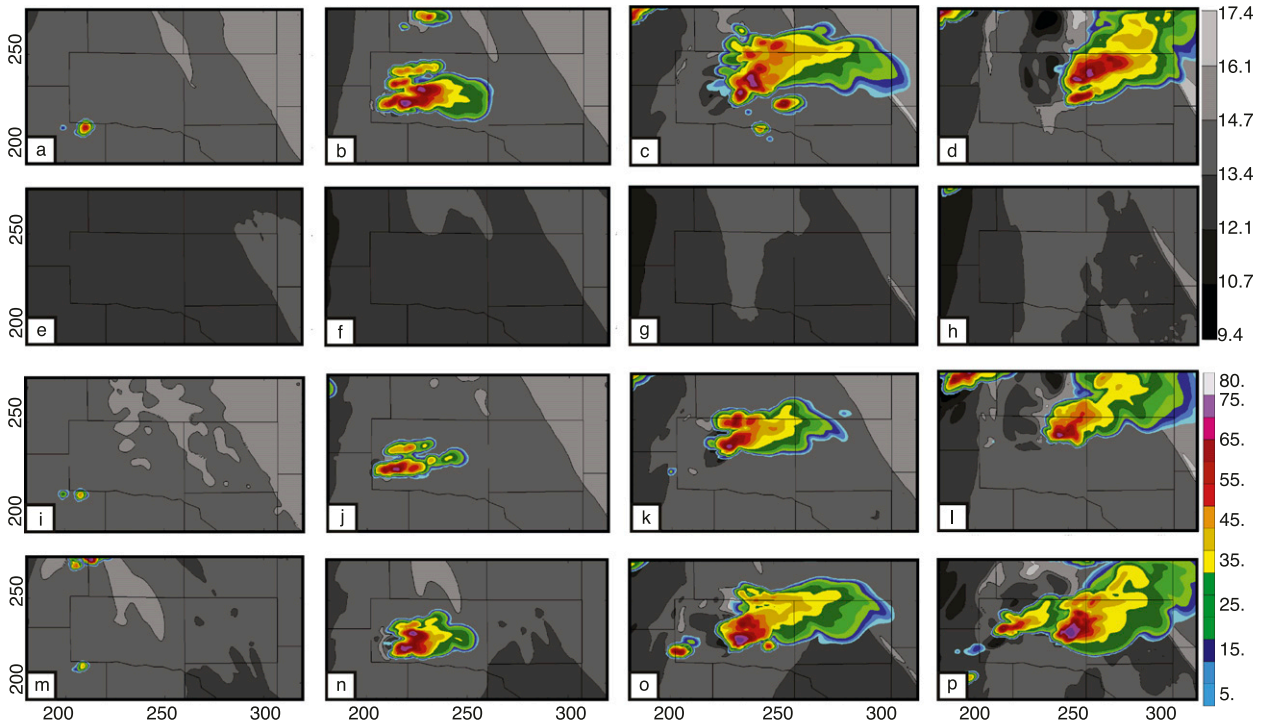
FIG. 10. Composite reflectivity (color fill, dB$Z$) and specific humidity (gray contours, g kg$^{-1}$) for (a)–(d) the truth field, (e)–(h) baseline experiment NoNd, (i)–(l) CNTLd experiment, and (m)–(p) 1RDd0.5. Plotted over verification subdomain in Fig. 1b (red box). Each column is valid at the same time, starting from 2115 UTC 19 May 2010 in the first column to 2245 UTC in the last column, plotted at 30-min intervals.

within 10 km and less than 10 min late. Only minor differences in CI timing and location can be seen between experiments 1RDd0.5 and CNTLd, both close to the truth. However, the evolutions of the storms are quite different. In 1RDd0.5, the storm does not grow as large as in CNTLd or the truth, owing to less moisture available near the

surface outside the radar range. As such, it develops more slowly and takes a northern path. Additionally, in experiment 1RDd0.5, there are two additional spurious points of initiation to the west and southwest of the supercell. This spurious activity appears to be related to the discontinuity found in the analysis at the edge of
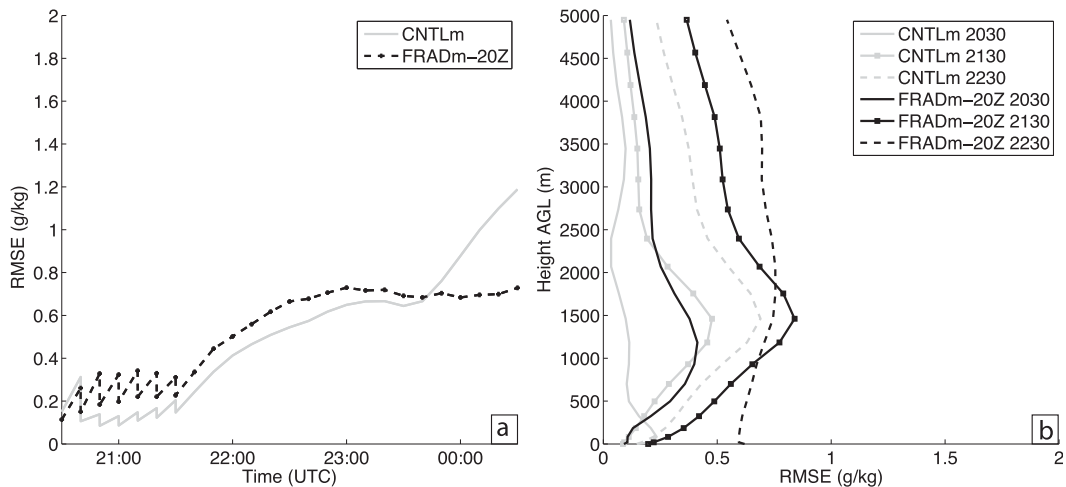


FIG. 11. (a) Time series RMSE of moisture for experiments FRADm-20Z and CNTLm. (b) Vertical profile of moisture RMSE for experiments FRADm-20Z and CNTLm, plotted at initial analysis time (2030 UTC); final analysis time (2130 UTC); and 1 h into the free forecast off of the final analysis (2230 UTC).
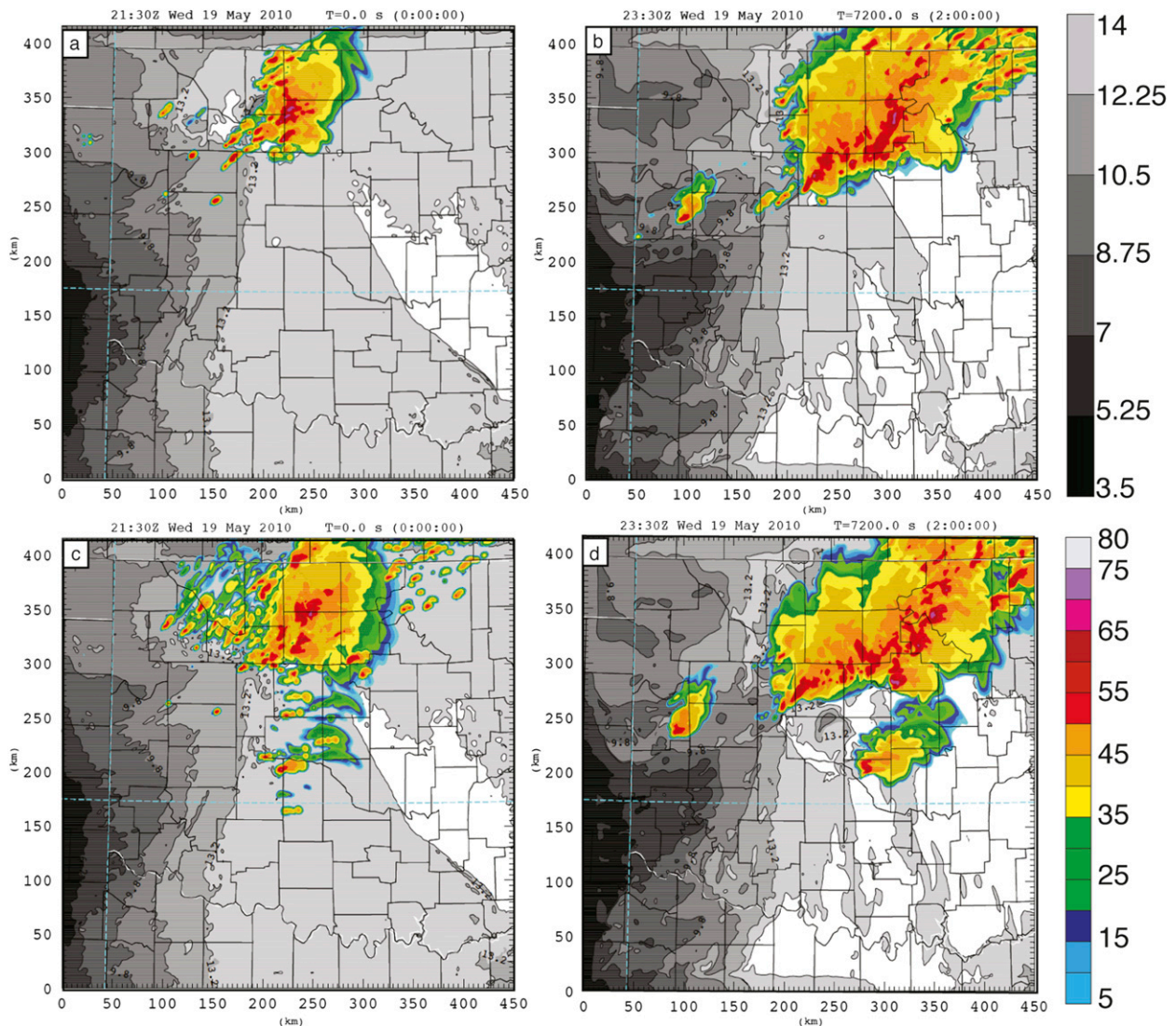
FIG. 12. Composite reflectivity (colored, dBZ) and surface specific humidity (lines and gray shaded contours, g kg$^{-1}$) for experiments (a),(b) FRADd-20Z and (c),(d) FRADm-20Z are valid at (a),(c) 2130 and (b),(d) 2330 UTC 19 May 2010.

radar data coverage. These results suggest that complete data coverage is much preferred and when complete coverage is unavailable, efforts should be made to reduce the effects of data edges in creating artificial moisture gradients that can cause forecast deterioration. Furthermore, a complete coverage has an equally important benefit of allowing for the correction of moisture error in the entire domain.

Overall, there are some improvements by isolated radar data coverage over the baseline, especially for the dry-biased experiments, but the impact is complicated by the gradient issue and a much smaller impact is obtained with the moist-biased case. The conclusion is that the storm forecast with assimilation of refractivity is most sensitive to isolated radar domain coverage.

### c. Results of experiments with time-evolved forecast background

The analysis background used in the experiments up to 1RDm0.5 in Table 1 was artificially constructed to mimic typical errors in a well-mixed convective boundary layer, and the error is limited to the moisture field only. The error is mostly removed by refractivity assimilation after just one cycle. Two additional experiments, FRADd-20Z and FRADm-20Z, are introduced to include more realism to the background error (see Table 1). For these experiments, initial error in the moisture field is introduced the same way as in the earlier experiments, as described in section 3, though 30 min earlier at 2000. The ARPS model is integrated from this initial
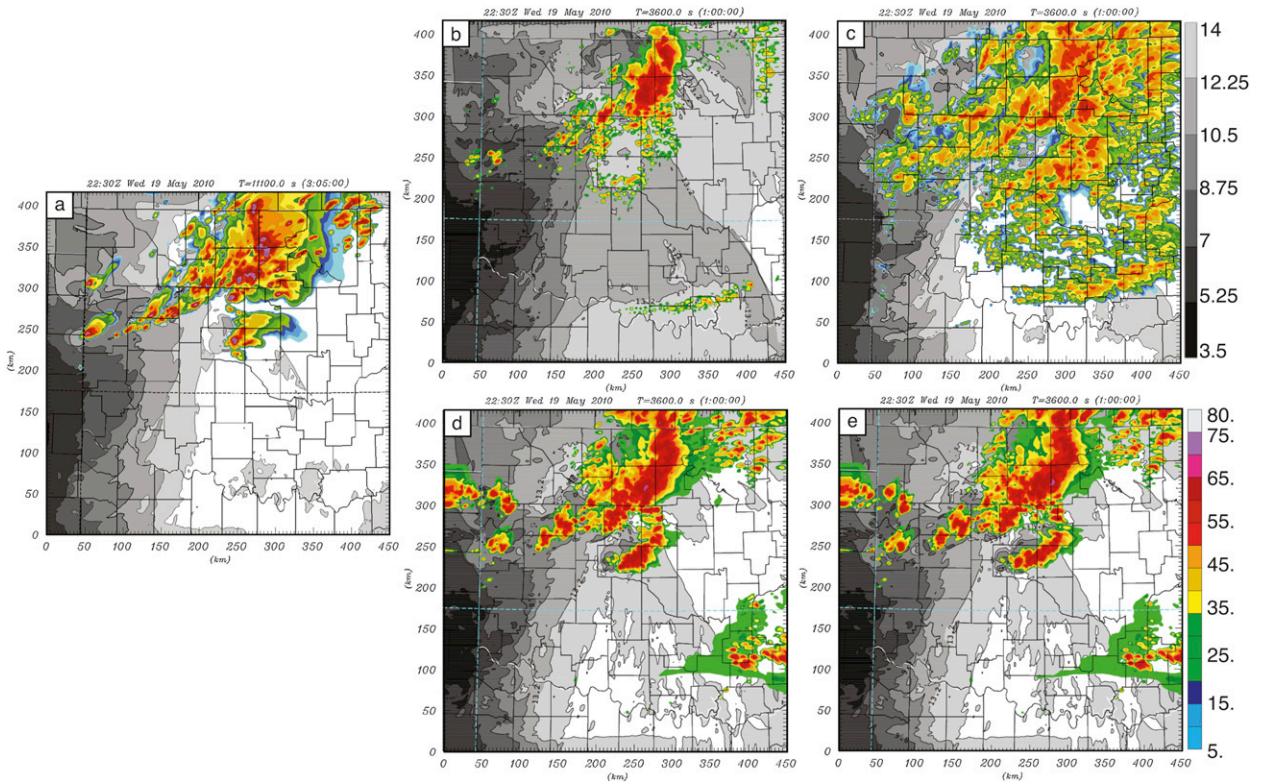
FIG. 13. Composite reflectivity (color fill, dB$Z$) and specific humidity (gray contours, g kg$^{-1}$) for (a) the truth field and imperfect model (using WRF) experiments (b) NoNd, (c) NoNm, (d) CNTLd, and (e) CNTLm, valid at 2230 UTC 19 May 2010 (1 h into WRF forecast run launched from final refractivity 3DVAR analysis).

background for 30 min until 2030, and this forecast serves as the new initial background for data assimilation starting at 2030 for the two new experiments. Such a time-evolved background would contain errors in other fields in addition to moisture due to error propagation through model integration. Experiments FRADd-20Z and FRADm-20Z are otherwise the same as experiments CNTLd and CNTLm, respectively.

For experiment FRADm-20Z, the surface RMSEs of moisture both during cycling and the free forecast are comparable in order of magnitude to the corresponding control experiment CNTLm (Fig. 11a). However with the added uncertainty in the forecast background, experiment FRADm-20Z contains about 0.1 g kg$^{-1}$ more error throughout the cycling window and free forecast, though interestingly the error becomes *lower* than in CNTLm after roughly 2345. While the analysis at the surface is consistent with other assimilation experiments, greater differences exist in the vertical profiles (Fig. 11b). By the end of the cycling window at 2130, experiment FRADm-20Z shows as much as 0.4 g kg$^{-1}$ greater error than in CNTLm in the vertical profile. Compared to experiment NoNm, however, there is still a significant improvement in the vertical profile of BL moisture up to

about 1.5 km even an hour into the forecast (cf. to NoNm in Fig. 6b). Experiment FRADd-20Z exhibits similar characteristics as FRADm-20Z in moisture error (not shown).

Figure 12 shows the reflectivity at the final analysis time (2130) as well as 2 h into the forecast (2330) for FRADd-20Z and FRADm-20Z. Unlike CNTLd and CNTLm, even after assimilation some effects of the background error remain in the predicted storm. In other words, FRADd-20Z (Fig. 12a) and FRADm-20Z (Fig. 12c) show less and more convection, respectively, than in the truth at 2130 (Fig. 1b). Though the preexisting storm complex develops similarly as the truth, the isolated supercell (Fig. 1b) is not captured at all in FRADd-20Z, which is different from the results of CNTLd (Fig. 7d). Clearly, errors in more fields can affect the impact of refractivity assimilation. On the other hand, in FRADm-20Z, refractivity assimilation does show a positive impact for the isolated supercell. Even before the refractivity assimilation window, the positive bias in moisture caused many spurious cell initiation points south of the preexisting storm. However, these cells quickly weakened during the cycling window to the point where by 2330 (Fig. 12d) only a single developed
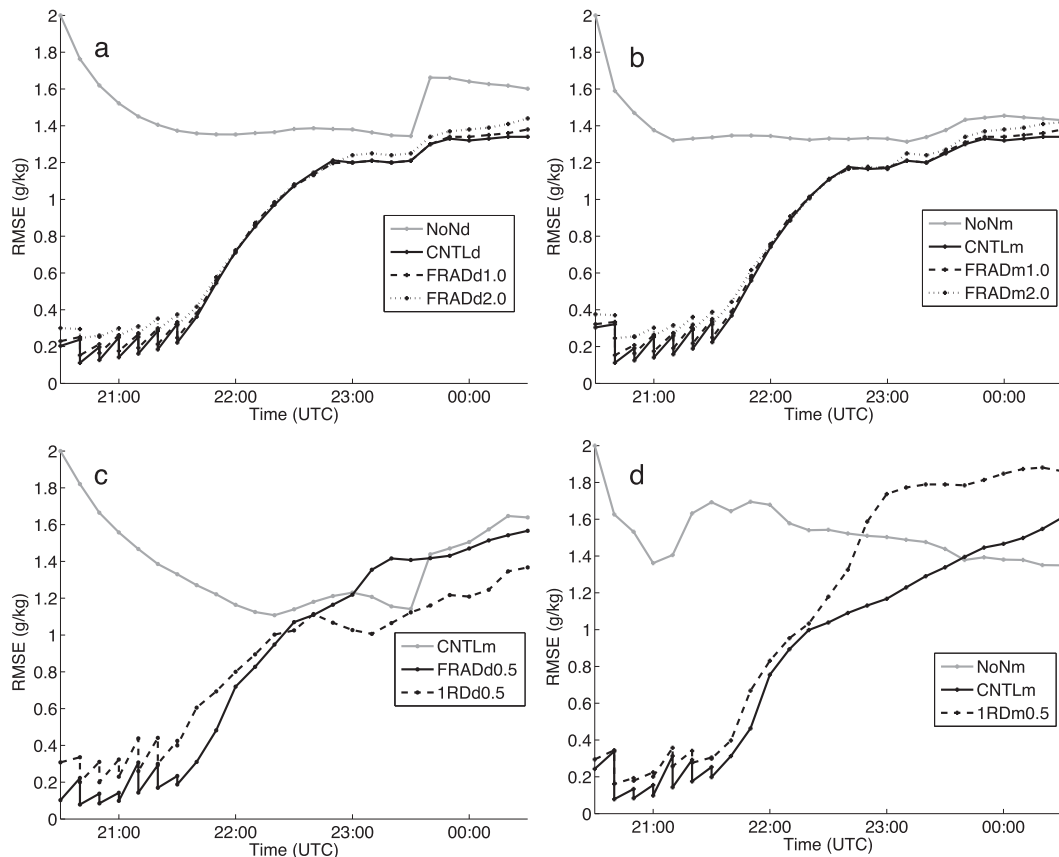
FIG. 14. RMSEs of the specific humidity field $(g\,kg^{-1})$ for the imperfect model cycled analysis and forecast sensitivity experiments listed in Table 1. Note that the WRF forecast was initialized at 2130 UTC from the final 3DVAR analysis (a),(b) for sensitivity to observation error experiments, and (c),(d) for sensitivity to isolated radar experiments. Note that in (a),(b) RMSEs were calculated over the whole domain and in (c),(d) RMSEs were calculated only over the 50-km radius of the isolated radar.

supercell remained. This supercell is roughly 30–40 km south of the supercell in the truth (Fig. 1d) but otherwise appears to show similar shape and strength.

### d. Results of imperfect model experiments

Results from the imperfect model assimilation cycles and forecasts show similar tendencies of improvement because of refractivity assimilation as in the perfect model forecasts, but with some important differences. In the reflectivity plots of Fig. 13, an hour into the forecast, the general shape of the convective complex in control experiments is similar to that of the truth. The baseline experiments NoNd (Fig. 13b) and NoNm (Fig. 13c) show significantly reduced and increased convection, respectively, as compared to the truth (Fig. 13a). There are many more locations of convection east and south of this complex than were present in the truth simulation for all imperfect model experiments, and in general the storms do not develop as strongly as in the truth. Comparing the RMSEs of moisture for the WRF forecasts in

Fig. 14 to RMSEs of perfect model forecasts in Fig. 5, it is clear that forecast error growth is up to 2–3 times faster in the WRF forecasts than in the ARPS forecasts, especially within the first 30 min. The errors of the refractivity data assimilation experiments in Fig. 14 do, however, tend to stay below their respective baseline forecast throughout the 3-h forecast period, despite this large error growth.

ETSs of reflectivity calculated over the full domain (Fig. 15a) show similar levels for full-domain refractivity assimilation experiments, decreasing from 0.7 to 0.8 at the initial forecast time (2130) to about 0.4 an hour into the forecast. By 0000, the experiments score at or below 0.3, an indication that the effect of refractivity assimilation on convection is lost after roughly 90–120 forecast minutes. Baseline experiment NoNd shows generally lower scores from around 0.7 at the initial time down to 0.2 an hour into the forecast. NoNm shows initially the lowest scores at 0.1–0.2 until 2330, because of the widespread spurious convection caused by the positive
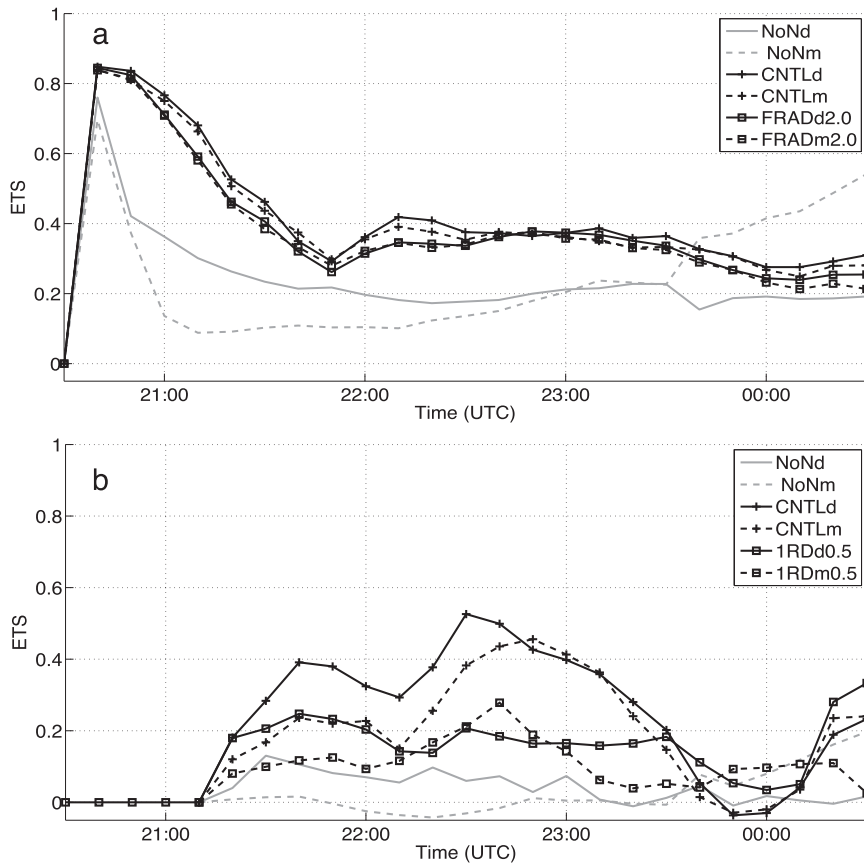
FIG. 15. ETS time series of imperfect (WRF) model analysis and forecast experiments. (a) For NoN, CNTL, and FRAD experiments, calculated over the whole domain; and (b) for NoN, CNTL, and single radar (1RD) experiments, calculated over the verification subdomain shown in Fig. 1b.

moisture bias in the background. The ETSs for imperfect model CNTL experiments are slightly lower than those of perfect model forecast in Fig. 8. The results for forecast experiments of increased observational error are consistent with the perfect model forecast results.

ETSs calculated over the subverification domain of Fig. 1b near the isolated supercell in the truth (Fig. 15b) are lower in general. Both baseline experiments remain near 0 throughout the forecast, as neither captures the supercell initiation and development and have many spurious cells. Single-radar experiment 1RD0.5m shows the lowest impact on the forecast above the baselines, as ETS remains around 0.1 throughout. Since the assimilation only covers a 50-km radius of a single radar, the bias of high moisture in the rest of the field outside data coverage still dominates the forecast, causing widespread aggressive storm development. Thus, the impact of assimilating refractivity is significantly reduced in experiment 1RD0.5m. Dry-bias single-radar experiment 1RD0.5d shows higher scores closer to 0.2. Both single-radar experiments score generally lower than their corresponding

control experiments, an indication of the impact of the loss in refractivity data coverage.

To further look into what is occurring within these imperfect model experiments near the isolated supercell, Fig. 16 plots composite reflectivity over the box in Fig. 1b for dry-bias experiments NoNd, CNTLd, and 1RDd0.5. As in the perfect model experiments, the imperfect model forecasts of CNTLd and 1RDd0.5 show similar development of the supercell as in the truth at 2145 (recall the supercell initiates at 2105 in the truth), whereas the baseline NoNd does not show corresponding initiation. However, 30 min later significant differences begin to appear, as the imperfect model shows increased spurious convection, even in baseline NoNd. CNTLd shows a strong supercell that is similar to that in the truth, though it has a broader area of high reflectivity. Experiment 1RDd0.5 also shows a storm that appears to be similar to the supercell in the truth, though its supercell characteristics are unclear and it is not as large. The spurious storm development may be affecting the growth of the single supercell in 1RDd0.5. Positive
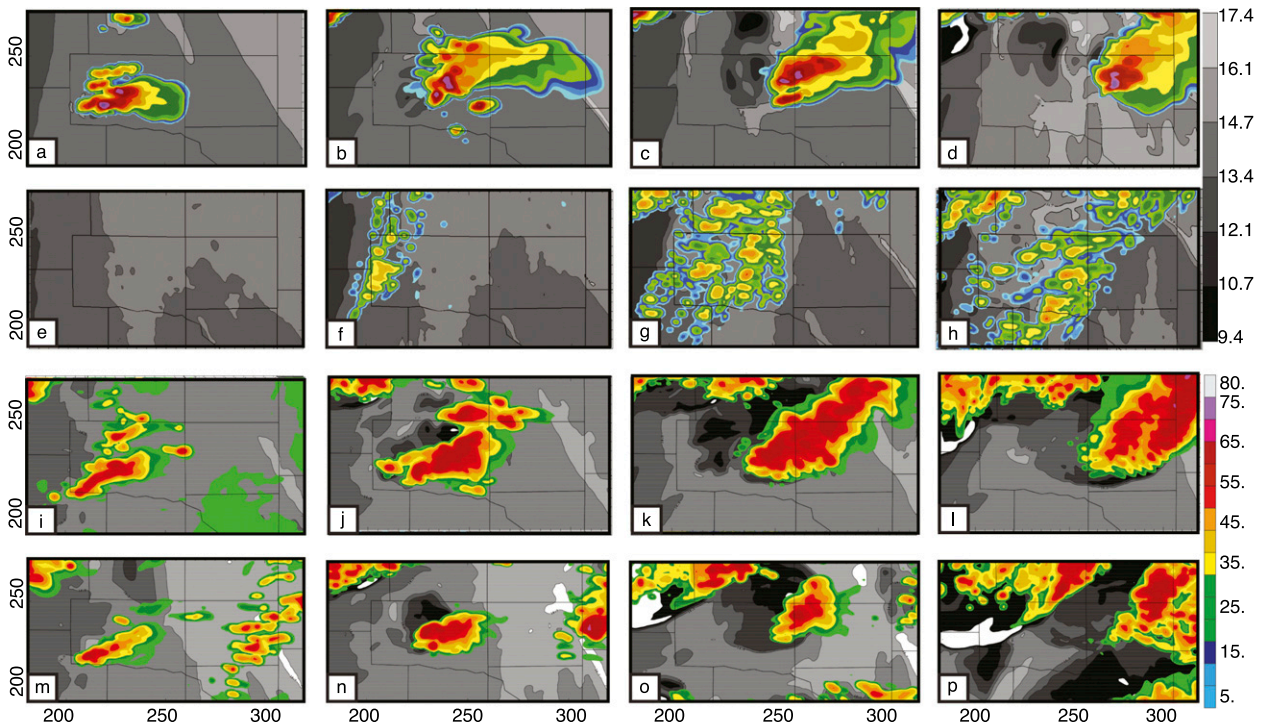
FIG. 16. Composite reflectivity (color fill, dB$Z$) and specific humidity (gray contours, $g\,kg^{-1}$) for (a)–(d) the truth field, (e)–(h) imperfect model baseline experiment NoNd, (i)–(l) imperfect model control CNTLd experiment, and (m)–(p) imperfect model experiment 1RDd0.5. Plotted over verification subdomain in Fig. 1b (red box). Each column is valid at the same time, starting from 2145 UTC 19 May 2010 in the first column to 2315 UTC in the last column, plotted at 30-min intervals.

impacts can be seen initially, but model error grows larger than the differences because of initial conditions, masking the impact later on. In imperfect forecasts, the impact of refractivity assimilation on convection lasts no longer than 30–60 min, whereas in perfect model forecasts the impact lasts for at least an hour, and in some experiments up to 3 h.

## 5. Summary and conclusions

The main goal of this study was to develop and test a method to assimilate radar-derived refractivity measurements. The ARPS 3DVAR system was enhanced to include the assimilation of refractivity measurements directly. Observing system simulation experiments (OSSEs) were conducted using simulated refractivity data from radar networks to assess their impact on the analysis and forecast of convective storms, including their initiation and subsequent evolution. The assimilation experiments were performed in pairs, using a background containing positive or negative $2\,g\,kg^{-1}$ near-surface moisture error. Analysis experiments were conducted with simulated observations of differing error statistics and data coverage, with moisture RMSE and reflectivity ETSs calculated against the truth for

both analysis and forecast periods. The analysis period included cycling of 3DVAR and ARPS forecasts every 10 min for an hour, with a 3-h ARPS forecast initialized from the final analysis. The main results are as follows:

- Control experiments assimilated idealized refractivity available every 4 km in the entire domain with random errors of 0.5-$N$-unit standard deviation. Analyses of the control experiment reduced the moisture RMSE by about one to two orders of magnitude after just one analysis. The RMSEs of forecast moisture remained below those of the baseline experiments for the duration of the forecast, and the forecast reflectivity ETSs are higher overall.
- Sensitivity experiments with increased observational error and more realistic data discontinuities due to ground clutter coverage yielded only minor degradations in analysis and forecast quality compared to the control experiments.
- Experiments assimilating refractivity from a single isolated radar show similar RMSE downward trends in the analysis period to control experiments over the radar domain; however, the forecast period error growth is much faster. Additionally, the timing and location of CI are captured well, but the storm evolution

showed ETSs that did not remain above the no-data-assimilation experiments. This should be due to the lack of observations outside the single-radar coverage area.

- Generally, a larger benefit was seen in refractivity assimilation experiments where the background moisture is low biased than when the background moisture is high biased.
- When the initial analysis background contains errors in other fields in addition to the moisture field, the impact of assimilating refractivity is somewhat reduced, but still evident.
- Imperfect model OSSEs—where the assimilation and forecasting model used is completely different from the truth simulation model—show more limited benefit of refractivity assimilation, particularly with the high-moisture-biased experiments. This is because in the imperfect model experiments, forecast errors due to model error tend to become dominant after 30–60 min of integration, overwhelming the improvements from refractivity data assimilation. Still, positive impacts of refractivity data are present.

In brief, this proof-of-concept study using OSSEs demonstrated a positive *potential* for refractivity data to improve CI forecasts, a precursor for studying the impact of real refractivity data. Dense radar networks, such as that of Collaborative Adaptive Sensing of the Atmosphere (CASA), should be most advantageous, since they can provide continuous refractivity data coverage over large areas. Given that the refractivity data impact can be case dependent, more case studies should be examined to fully determine the general impact of the data. The assimilation can also benefit from using more advanced data assimilation methods, such as the ensemble Kalman filter or the 4DVAR technique, that allow for better multivariate analysis and adjustments to model fields not directly linked to refractivity. These can be topics for future research.

## REFERENCES

Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan,* **75,** 111–130.

Bean, B. R., and E. J. Dutton, 1968: *Radio Meteorology.* Dover Publications, 435 pp.

Bodine, D., P. L. Heinselman, R. D. Palmer, B. L. Cheong, and D. Michaud, 2009: Survey of applications of radar refractivity retrievals. Preprints, *34th Conf. on Radar Meteorology,* Williamsburg, VA, Amer. Meteor. Soc., P6.4. [Available online at https://ams.confex.com/ams/34Radar/techprogram/paper_155495.htm.]

——, and Coauthors, 2011: Understanding radar refractivity: Sources of uncertainty. *J. Appl. Meteor. Climatol.,* **50,** 2543–2560.

Brewster, K. A., J. Brotzge, K. W. Thomas, Y. Wang, M. Xue, J. Gao, D. Weber, and K. Howe, 2008: High resolution assimilation of CASA and NEXRAD radar data in near-real time: Results from spring 2007 and plans for spring of 2008. Preprints, *12th Conf. IOAS-AOLS,* New Orleans, LA, Amer. Meteor. Soc., 15B.7. [Available online at http://twister.ou.edu/papers/BrewsterIOAS2008.pdf.]

——, 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. Preprints, *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.]

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.

Cheong, B. L., R. D. Palmer, C. Curtis, T.-Y. Yu, D. S. Zrnic, and D. Forsyth, 2008: Refractivity retrieval using the phased array radar: First results and potential for multi-mission operation. *IEEE Trans. Geosci. Remote Sens.,* **46,** 2527–2537.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.,* **124,** 1767–1785.

Dabberdt, W. F., and T. W. Schlatter, 1996: Research opportunities from emerging atmospheric observing and modeling capabilities. *Bull. Amer. Meteor. Soc.,* **77,** 305–324.

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

Emanuel, K., and Coauthors, 1995: Report of the first Prospectus Development Team of the U.S. Weather Research Program to NOAA and the NSF. *Bull. Amer. Meteor. Soc.,* **76,** 1194–1208.

Fabry, F., 2004: Meteorological value of ground target measurements by radar. *J. Atmos. Oceanic Technol.,* **21,** 560–573.

——, C. Frush, I. Zawadzki, and A. Kilambi, 1997: On the extraction of near-surface index of refraction using radar phase measurements from ground targets. *J. Atmos. Oceanic Technol.,* **14,** 978–987.

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.

——, K. Brewster, and M. Xue, 2008: Variation of radio reflectivity with respect to moisture and temperature and influence on radar ray path. *Adv. Atmos. Sci.,* **25,** 1098–1106.

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

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

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

Huang, X.-Y., 2000: Variational analysis using spatial filters. *Mon. Wea. Rev.,* **128,** 2588–2600.

Ide, K., P. Courtier, M. Ghil, and A. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan,* **75,** 181–189.

Janjić, Z. I., 1990: The step–mountain coordinates: Physical package. *Mon. Wea. Rev.,* **118,** 1429–1443.

Koch, S. E., A. Aksakal, and J. T. McQueen, 1997: The influence of mesoscale humidity and evapotranspiration fields on a model forecast of a cold-frontal squall line. *Mon. Wea. Rev.,* **125,** 384–409.

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,** 1065–1089.

Liu, H., and M. Xue, 2006: Retrieval of moisture from slant-path water vapor observations of a hypothetical GPS network using a three-dimensional variational scheme with anisotropic background error. *Mon. Wea. Rev.,* **134,** 933–949.

Lord, S. J., E. Kalnay, R. Daley, G. D. Emmitt, and R. Atlas, 1997: Using OSSEs in the design of the future generation of integrated observing systems. Preprints, *First Symp. on Integrated Observation Systems,* Long Beach, CA, Amer. Meteor. Soc., 45–47.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.,* **120,** 1019–1034.

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.

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.

Montmerle, T., A. Caya, and I. Zawadzki, 2002: Short-term numerical forecasting of a shallow storms complex using bistatic and single-Doppler radar data. *Wea. Forecasting,* **17,** 1211–1225.

Schaefer, J. T., 1990: The critical success index as an indicator of forecasting skill. *Wea. Forecasting,* **5,** 570–575.

Schenkman, A., M. Xue, A. Shapiro, K. Brewster, and J. Gao, 2011a: 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.

——, ——, ——, ——, and ——, 2011b: 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.

Shimose, K., M. Xue, R. D. Palmer, J. Gao, B. L. Cheong, and D. J. Bodine, 2013: Two-dimensional variational analysis of near-surface moisture from simulated radar refractivity-related phase change observations. *Adv. Atmos. Sci.,* **30,** 291–305.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp.

Sun, J., 2005: Convective-scale assimilation of radar data: Progress and challenges. *Quart. J. Roy. Meteor. Soc.,* **131,** 3439–3463.

Sun, W.-Y., and C.-Z. Chang, 1986: Diffusion model for a convective layer: Part I: Numerical simulation of a convective boundary layer. *J. Climate Appl. Meteor.,* **25,** 1445–1453.

Tong, M., 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.

Weckwerth, T. M., 2000: The effect of small-scale moisture variability on thunderstorm initiation. *Mon. Wea. Rev.,* **128,** 4017.

——, and D. B. Parsons, 2006: A review of convection initiation and motivation for IHOP_2002. *Mon. Wea. Rev.,* **134,** 5–22.

——, J. W. Wilson, and R. M. Wakimoto, 1996: Thermodynamic variability within the convective boundary layer due to horizontal convective rolls. *Mon. Wea. Rev.,* **124,** 769–784.

——, and Coauthors, 2004: An overview of the International $H_2O$ Project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.,* **85,** 253–277.

——, C. R. Pettet, F. Fabry, S. Park, M. A. LeMone, and J. W. Wilson, 2005: Radar refractivity retrieval: Validation and application to short-term forecasting. *J. Appl. Meteor.,* **44,** 285–300.

Xue, M., and W. J. Martin, 2006a: A high-resolution modeling study of the 24 May 2002 case during IHOP. Part I: Numerical simulation and general evolution of the dryline and convection. *Mon. Wea. Rev.,* **134,** 149–171.

——, and ——, 2006b: A high-resolution modeling study of the 24 May 2002 case during IHOP. Part II: Horizontal convective rolls and convective initiation. *Mon. Wea. Rev.,* **134,** 172–191.

——, K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.,* **75,** 161–193.

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

——, D. Wang, J. 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.