Sensitivity of Convective Initiation Prediction to Near-Surface Moisture when Assimilating Radar Refractivity: Impact Tests using OSSEs

Nicholas A. Gasperoni$^{1,2}$, Ming Xue$^{1,2}$, Robert D. Palmer$^{2,3}$, Jidong Gao$^4$

$^1$Center for Analysis and Prediction of Storms, Norman, OK 73072
$^2$School of Meteorology, University of Oklahoma, Norman, OK 73072
$^3$Atmospheric Radar Research Center, University of Oklahoma, Norman OK 73072
$^4$NOAA/OAR National Severe Storms Laboratory, Norman, OK 73072

February 2012
Submitted to
J. of Atmospheric and Oceanic Technology

Corresponding author address:

Dr. Ming Xue  
Center for Analysis and Prediction of Storms  
University of Oklahoma  
120 David L. Boren Blvd, Norman, OK 73072  
mxue@ou.edu
Abstract

The ARPS 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 results compared to the truth. A simulation for the May 19, 2010 Central Plain convection case is used for the OSSEs. It involves a large storm system, large convective available potential energy (CAPE), 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, to lead to improved CI forecast. 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. There is a larger sensitivity to the data coverage when a single radar is available.
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 is typical of boundary layer moisture (Weckwerth et al. 1996), can make the difference between storm initiation or not. Xue and Martin (2006a, 2006b) performed a high-resolution modeling study of the 24 May 2002 dryline CI case during the International H\(_2\)O Project (IHOP) 2002 (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 measurement is a major obstacle (e.g., Crook 1996; Dabberdt et al. 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 measurements to sufficiently sample boundary layer phenomena that lead to CI.

High-resolution observations of moisture are notoriously difficult to measure, however. 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. 2010; 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. 2010, 2009; Fabry 2004; Weckwerth et al. 2005). In particular, IHOP_2002 contained a convective initiation component (Weckwerth; Parsons 2006) that exploited the utilities of radar refractivity data also. Weckwerth et al. (2005) showed for IHOP_200 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.

At the University of Oklahoma, experiments have been ongoing collecting refractivity data from operational Oklahoma City (KTLX) and Frederick, OK (KFDR) WSR-88D radars. From 2007-2008, KTLX data were given to forecasters to evaluate, with feedback from surveys (Heinselman et al. 2009). Results were mixed; while forecasters saw an increased confidence in moisture trends by using refractivity data, they did not find that it added significant benefit to their forecasts over the Oklahoma Mesonet data. The study concluded that additional research is needed to better use the refractivity data. For example, refractivity still has not been directly assimilated into models nor evaluated for potential impact on convective-scale forecasts, at least not in the form of journal publications.

The current study focuses on the assimilation of refractivity measurements directly in a numerical weather prediction (NWP) framework, building upon the results of Shimose et al.
in a simpler 2D framework. 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; Xue et al. 2003; Xue et al. 2001) is used in this study. As an initial effort, we chose to use simulation refractivity data through Observing System Simulation Experiments (OSSEs), in order to test the sensitivity to a number of aspects related to refractivity measurements, such as instrument error and data coverage. With OSSE, the known and complete truth state also allows unambiguous evaluation of quality and sensitivity. The ARPS three-dimensional variational (3DVAR) (e.g., 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 (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, a forward model for radar refractivity measurements is developed and added to the ARPS 3DVAR. Data impact experiments are then performed using data simulated from the truth simulation, using the same refractivity observation operator.

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. Section 3 shows results from a simple zero-dimensional refractivity analysis problem to help understand the response to a refractivity measurement in an effectively optimal interpolation (OI) framework. The OSSE design is explained in section 4 and results are presented in section 5. Summary and conclusions are given in section 6.

2. Radar Refractivity Assimilation using ARPS 3DVAR

a. Refractivity Measurements
Typical Doppler weather radars work by sending out electromagnetic (EM) pulses through the atmosphere and receiving backscatter from targets through an antenna. When each EM pulse travels through the atmosphere, it travels roughly at 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,

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

(1)

where \( T \) is temperature in Kelvin, \( P \) is atmospheric pressure in hectopascals (hPa), and \( e \) is water vapor pressure also in hPA. The units of refractivity are typically referred to as “N-units”, though they are non-dimensional 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 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-integrated phase measurements \( \phi \) between the radar and ground clutter targets. To mitigate the issue of phase wrapping, reference phase measurements are taken at a reference observation time \( t_{\text{ref}} \) when the atmosphere is more spatially homogeneous and can therefore be well measured by surface networks. Taking a difference of phase measurements from some observation time, \( t \), and reference time, \( t_{\text{ref}} \), yields a phase change* measurement \( \Delta \phi \) which significantly reduces phase

* As in Shimose et al. (2012), we use ‘phase change’ to refer to difference in phase of a target between two different times, and ‘phase change difference’ (PCD) to refer to phase difference between two consecutive targets aligned
wrapping because phase change is typically much smaller than the absolute phase (see also discussion in Shimose et al. (2012)). Refractivity change measurements, $\delta N$, can be obtained by taking a range derivative of these $\delta \phi$,

$$
\delta N(r) \equiv N(r, t) - N(r, t_{\text{ref}}) = -10^6 \left( \frac{c}{4\pi f} \right) \frac{\partial}{\partial r} [\delta \phi(r)], \quad (2)
$$

where $f$ is the radar transmit frequency. These refractivity measurements are integrated quantities from radar to target site. By taking the phase change difference (PCD) at two range target values along the same radial,

$$
\Delta \phi(R_1, R_2) = \delta \phi(R_2) - \delta \phi(R_1) = -10^6 \left( \frac{4\pi f}{c} \right) \int_{R_i}^{R_f} \delta N(r') \, dr' \quad (3)
$$

where $\delta N(r') = N(r', t) - N(r', t_{\text{ref}})$, Fabry (2004) shows that localized gate-to-gate mean refractivity difference can then be calculated between $R_1$ and $R_2$,

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

assuming that the refractivity gradient is constant between targets. The two targets must be exactly collinear both spatially and vertically for (4) to work; however, real ground clutter targets do not have this property naturally, meaning the PCD measurements can be very noisy. Additional data processing techniques are required to reduce the noise and obtain reliable measurements.

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 along the same radial. Additionally, $\delta$ refers to time differencing and $\Delta$ refers to radial differencing of time-differenced measurements.
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 due to 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 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, \( J \), which 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

\[
J(x) = \frac{1}{2} (x - x_b)^T B^{-1} (x - x_b) + \frac{1}{2} (y_o - H(x))^T R^{-1} (y_o - H(x))
\]  

(5)

The first term is the background term, \( J_b \), containing the difference between analysis vector \( x \) and background vector \( x_b \), weighted by the inverse of the background error covariance matrix \( B \). The second term is the observation term, containing the differences between observation vector \( y_o \) and the state vector projected to the observation space, \( H(x) \), weighted by the inverse of observation error covariance matrix \( R \). \( R \) contains instrumental as well as representativeness errors. In this study, analysis vector \( x \) 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,

$$ J_N = \frac{1}{2} (N_o - H(x))^T R_N^{-1} (N_o - H(x)), $$

where $N_o$ is the vector of absolute refractivity observations from radar and $R_N$ 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,

$$ \nabla J_N(\delta x) = H^T R_N^{-1} (H \delta x - \delta N_o) $$

where $\delta x = x - x_b$ are the control increments; $\delta N_o = N_o - H(x_b)$ is the $N$ observation innovation vector; $H$ is the linearized version for observation operator, $H$, and $H^T$ is its transpose or adjoint.

The proper specifications of $R_N$ and $B$ are important to the 3DVAR analysis. $R_N$ is often comprised 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. Observation error correlation is difficult to determine in practice since the error includes both instrument and representativeness and sometimes post-processing errors. For remote sensing platforms such as radar, there is often error correlation in their measurements due to systematic biases. However, steps are normally taken to reduce such correlations via data thinning and bias removal (e.g., Dee 2005). $B$ matrix also plays an important role in the optimal analysis because its off-diagonal spatial covariance and cross-covariance terms control the amount of spread of increments spatially and across different variables. The cross-correlation terms of $B$ control balance properties of the fields, allowing for observations to be used more effectively via indirect corrections to other variables; however, flow-dependent spatial covariance structures are generally not available in a 3DVAR framework. In the ARPS 3DVAR system, the spatial correlation elements of $B$ 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. Liu and Xue (2006) and Liu et al. (2007) have discussions on the covariance modeling and the use of isotropic and anisotropic filters.

Because refractivity depends on $P$, $T$, and water vapor $P_e$, according to (1), the analysis of refractivity will directly influence those three variables. In ARPS 3DVAR, potential temperature $\theta$ and specific humidity $q_v$ are used as control variables which are linked to temperature and water vapor through the 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. Analysis increments from a single refractivity observation

Studies have shown that during the warm season, refractivity is most sensitive to moisture (Fabry 2004; Fabry et al. 1997; Gao et al. 2008; Weckwerth et al. 2005). Gao et al. (2008) show that the variability of refractivity with respect to dewpoint temperature is about 5 to 6 times greater in magnitude than with respect to $T$ when the base $T$ exceeds 30°C, and 2 to 4 times greater with a base $T$ between 10°C and 30°C. They further show that given a base state of 60% relative humidity (RH), 17°C $T$, and 1000 hPa $P$, a 1°C change in $T$ causes a magnitude 1.34 N-unit change in refractivity while a 1°C change in dewpoint causes a magnitude 4 N-unit change in $N$. Clearly, we should expect the 3DVAR analysis increments in moisture to be significantly higher than the analysis increments in either $T$ or $P$, given $N$ observations. Typically background
fields also contain useful information and their effects on the final analysis depend on their error magnitude relative to that of observations.

Insights can be gained by directly solving the 3DVAR equations for simple cases, with the simplest being the case with a single observation located at a grid point. The resulting analysis increment $\delta x$ should have the following functional dependence,

$$\delta x = f(\delta N_o, P_b, T_b, e, \sigma_N, \sigma_p, \sigma_T, \sigma_e)$$

where $\delta$ denotes the increment from the background, subscript $b$ denotes the background and $\sigma$ denotes the error standard deviations for the corresponding variables. For this point, all variables on the right are scalars and the analysis increment vector, $\delta x$, contains three increments.

To solve for $T$, $P$, and $e$, we minimize cost function $J$ by setting the gradients of $J$ with respect to each variable to zero and solve the resulting three equations with three unknowns. The moisture is specified in terms of relative humidity and can be readily converted to vapor pressure, $e$, or specific humidity, $q$, given the background temperature and pressure. Background errors $\sigma_p$, $\sigma_T$, and $\sigma_e$ need to be specified. The $T$ and $P$ errors are set at 3°C and 3 hPa, respectively. The moisture error is also specified in terms of relative humidity, $\sigma_{RH}$, which can be converted to vapor pressure error, $\sigma_e$, using

$$\sigma_e = \left(\frac{\sigma_{RH}}{100}\right) e^* (T_b).$$

where $e^*$ is the saturation vapor pressure, a function of the background temperature $T_b$. Relative humidity error is varied between 0 and 30% to examine the impact of background-observation relative error. For the tests, the radar observation innovation is +5 N-units above the background $N_b$ and $\sigma_N^2$ is set to 4 (N-units)$^2$. 

Fig. 1 shows that as background RH error increases, the moisture increment increases and the $P$ and $T$ increments decrease simultaneously. This result is expected based on optimal estimation (Kalnay 2002). Increasing the background error forces a better fit of the analysis to observation, and vice versa. Since the analysis updates all three state variables, increasing the background error in one variable increases the adjustment to that variable and reduces the need to adjust other variables to fit the observations, hence the reduced $P$ and $T$ increments observed.

Interestingly, under the typical warm-season conditions ($T$ at or above 20 °C), all analysis increments level off as the RH error increases beyond 10%; the $T$ and $P$ increments are on the order of -0.1 K and 0.1 hPa while specific humidity increments are between 0.6 and 0.85 g kg$^{-1}$. Thus, during warm season for a background with relative humidity error larger than 10% (typically true at the convective scale), we can expect the large majority of the adjustment by the optimal 3DVAR analyses will go to humidity, while the background $T$ and $P$ will remain largely unchanged by refractivity data assimilation (although large changes can occur through model adjustments).

It should be noted that in the pioneering paper of Fabry et al. (1997), a simple ‘rule of thumb’ states that a +1 N-unit change in refractivity corresponds with roughly either a decrease of 1 °C in temperature or an increase of 0.2 g kg$^{-1}$ in specific humidity at sea level during the warm season. The 3DVAR analysis increments in specific humidity seem to agree well with this estimate. If the background moisture field was perfect, the resulting temperature increments would be roughly 0.3 to 0.7 K per N-unit depending on background temperature and moisture, somewhat less than Fabry’s ‘rule of thumb’ but still within the same order of magnitude.

If RH error is below 5%, the analysis increments dramatically change in characteristics. The increment in moisture rapidly decreases to zero as RH error decreases while $P$ and $T$
increments rapidly increase in magnitude. This is the case because the 3DVAR is being told that
the background moisture value is very reliable, so it does not correct moisture as much. Fig. 1b
shows that the $P$ increment, although increased, remains under 1 hPa and relatively insignificant.
So most of the significant corrections are instead in the $T$ field, which generally are between a
decrease of 1 to 3°C, depending on background $T$ and RH. Not shown are two additional sets of
experiment with varying background $T$ and $P$ errors. Increasing $P$ error up to 10 hPa had very
little effect on the resulting increments. More effect was seen when increasing $T$ error up to 10
K; however, it was still insignificant compared to the effects of increasing moisture error.

4. 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).
May 19, 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 (MCS) left an outflow boundary which 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 the low 80s in Fahrenheit with dew point temperatures well into the 60s. 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 for CASA (Brewster et al. 2010; Brewster et al. 2008). It was initialized at 1930 from an ARPS 3DVAR analysis including available surface and Mesonet observations as well as available CASA and NEXRAD radar data, using a time-interpolated NAM 12-km analysis as the background. The domain is 450 km x 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 (LBCs) used for the truth
simulation were also taken from NAM 12-km analysis and forecasts every 3 hours. The truth simulation is run for 5 forecast hours, up until 0030 on 20 May 2010.

Fig. 2 shows the moisture and reflectivity fields between 2030 and 2330. The storm shown at 2030 is a cell which initiated within the first 15 minutes of the simulation, thus will be analyzed as a pre-existing 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. 2b 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 due to 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 Eq. (1) from the model truth fields at the surface. The resolution of simulated data is 4 km, taken every 4 grid points, which is similar to the effective resolution of the final processed radar refractivity measurements from KTLX and KFDR by Cheong et al. (2008). For a single radar, refractivity is assumed to be available within a radius of 50 km of radar, and the coverage is assumed to be continuous (Fig. 3). The addition of error and other uncertainty will be discussed in conjunction with sensitivity experiments.

c. List of OSSE experiments

A set of experiments with cycled analysis and forecasting are conducted. They all follow the general assimilation setup given in Fig. 4. The truth simulation is run for an hour to allow for
the model to “spin up,” developing more complex structures in all model fields (especially moisture) prior to introducing uncertainty and performing assimilation experiments.

To serve as a baseline, a “first-guess forecast” is initialized at 2030 UTC from the background and run for 4 hours (NoNd and NoNm in Table 1). The initial analysis background for first guess at 2030 is 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, which forces vertically correlated error in the boundary layer

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

(10)

Here, $q_{v}^{sfc}$ is the surface moisture after step 2, and subscript $truth$ denote the truth field and $D$ is set to 4 km. $z$ is height above ground. For $z > D$, $q_v = q_{v, truth}$. 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 truth. Experiments indicated an optimal horizontal de-correlation length scale of 8 km (not shown), in general agreement with Shimose et al. (2008). Additionally, the optimal vertical de-correlation length was found to be 6 vertical grid levels, or roughly 1 km.

From the assimilation experiments, cycled analysis and forecasts are performed for an hour between 2030 and 2130 at 10 minute intervals, yielding a total of 7 refractivity analyses. From the analysis at 2130, a 3-hour 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 (d for dry and m for moist). They are idealized experiments where the entire domain is completely covered by overlapping radars (Fig. 5a) and refractivity observations are available every 4 km and contain Gaussian errors with a 0.5 N-unit standard deviation. 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: (1) sensitivity to background field, (2) sensitivity to observation error, (3) sensitivity to realistic clutter coverage, and (4) sensitivity to radar domain coverage. For the second kind of sensitivity, 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 (3), realistic clutter coverage is applied to the full radar 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. 5b). For sensitivity to radar coverage, experiments include an isolated radar with 50-km radius refractivity coverage and the radar is placed near the supercell initiation point (see Fig. 2b).
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 the root-mean-square error (RMSE) and the equitable threat score (ETS). RMSE can be calculated against the truth according to,

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{i,\text{est}} - x_{i,\text{truth}})^2}$$ (11)

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

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

$$ETS = \frac{\text{hits} - \text{chance}}{\text{hits} + \text{misses} + \text{false alarms} - \text{chance}},$$ (12)

where hits are forecast and observed yesses, misses are forecast no’s and observed yesses, false alarms are forecasted yesses and observed no’s, and chance is the number of hits one would get by random chance,

$$\text{chance} = \frac{(\text{hits} + \text{misses})(\text{hits} + \text{false alarms})}{\text{hits} + \text{misses} + \text{false alarms} + \text{correct negatives}},$$ (13)

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.

5. 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. Fig. 6 shows time
series plots of RMSEs of specific humidity at the surface. The baseline forecasts NoNd (Fig. 6a) and NoNm (Fig. 6b) show similar characteristics, the RMSE starts at ~2 g kg$^{-1}$ due to the design of the first guess but through forecast decreases, to about 1 g kg$^{-1}$ after roughly 90 minutes. This reduction occurs due to convective mixing in the boundary layer, as can be seen in Fig. 7, because errors higher up are smaller. This surface error begins to increase again only within the last 30-60 minutes or in the fourth hour of forecast.

With 10-minute 3DVAR analysis cycles in the control experiments, the time series plots of RMSEs in Fig. 6 look similar to typical “sawtooth” plots seen in radar data assimilation studies using the ensemble Kalman filter (e.g., Tong; Xue 2005). Looking at the analysis cycling results in each of CNTLd and CNTLm, they show similar downward trends. Both CNTLd (Fig. 6a) and CNTLm (Fig. 6b) 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 ~0.1 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 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 corresponding baseline about 3 hours into the forecast while in CNTLm, the error remains lower all the time. 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. 7 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 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 k 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 3DVAR covariance now works less effectively. Despite that, the RMSEs of the control experiments are 0.5 – 1.0 g kg\(^{-1}\) below the corresponding baseline RMSE errors at 2130. Similar difference remains at the lower levels even after 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. 7a) remains about 0.3 g kg\(^{-1}\) less than the RMSE of NoNd even at 2230. However, the errors of NoNm (Fig. 7b) converge to those of CNTLm by 2130 at the higher levels, mainly due to much slower error growth in NoNm than in its dry counterpart. The effects of assimilating surface \(N\) data on the upper level moisture within the assimilation and forecast system are more complex than on the surface moisture itself.

The effect of refractivity assimilation on predicted storms is shown in Fig. 8, for 2200. The preexisting storm from Fig. 2a 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 8b and 8c 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 NW 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 \((x, y) = (150 \text{ km}, 270 \text{ km})\); however, these storms are 30-45 minutes late in initiation timing compared to the truth and quickly dissipate within an hour. Contrast this with NoNm (Fig. 8c), 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 minutes. The supercell area shows several points of CI that occur starting at 2055 in NoNm, which is 15 minutes early.

Figures 8d and 8e 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 data assimilation is the control forecasts has shown considerable positive impact on new points of initiation, to within less than 10 minutes of timing error.
These subjective results are confirmed with the ETS score values in Fig. 9 using a 30 dBZ threshold. Each ETS score starts at being perfect because hydrometeor fields in the initial background had no error with the preexisting storm. The score decreases with time due to error growth associated with new points of initiation and storm evolution. The baseline forecasts decrease to a relatively low score of 0.2 – 0.3 by 2200, while the control forecasts consistently score higher, at or above 0.5 throughout the 3 hours of forecast.

c. Results of sensitivity experiments

(1) Sensitivity to observation error

In the presence of observation error (Fig. 6a,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 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 errors in model variables to the error of refractivity, assuming model variables $P$, $T$, and $e$ are uncorrelated

$$
\sigma_N^2 = \left( \frac{\partial N}{\partial P} \right)^2 \sigma_P^2 + \left( \frac{\partial N}{\partial T} \right)^2 \sigma_T^2 + \left( \frac{\partial N}{\partial e} \right)^2 \sigma_e^2. \tag{14}
$$

Equation (14) is a first-order Taylor-series approximation of the propagation of error, considering each distribution is approximately normal. It is then rearranged to solve for vapor pressure error $\sigma_e$, 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 (approximately the same from RH to refractivity!) Compared to typical NWP background errors in RH that often are well above 5%, these error amounts are very small. Thus, observation errors up to 2 N-units should yield rather accurate 3DVAR 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.10a,b). Additionally, after just 30 minutes 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 effects on the moisture analysis and forecast. There are, however, some times where the 3DVAR analysis increases RMSE slightly in later cycles, such as FRADd2.0 at 2100 UTC (Fig 10a). 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 because with subsequent cycling the background error should be smaller as its accuracy increases. This is a general drawback of 3DVAR, the inability to dynamically update background error covariances with each new analysis.

Figure 10a shows reflectivity ETS scores of control experiment CNTLd, as well as experiments with greater error added to observations, FRADd1.0 and FRADd2.0, calculated over the entire domain. All experiments show higher scores than the baseline, similar to Fig. 9. There are only minor differences among the different assimilation experiments, no more than 0.1 in terms of ETS scores. The results are similar for the ‘moist’ experiments in Fig. 10b. When ETS is calculated over only the red box in Fig. 2b for the supercell (Fig. 11), much larger differences
exist among the experiments, all of which still yield higher ETS scores than the NoNd experiment, which is zero in the dry case and nearly zero in the moist case. Generally, CNTLd tends to have the best ETS scores (though not always) likely because of its relatively low amount of observation error, while FRADd2.0 has lower ETS scores (though still well above that of NoNd). The forecast of the supercell seems to be relatively insensitive to increased observation error of up to 2 N-units, and all assimilation experiment show improvement over the corresponding no data assimilation experiment in terms of ETS scores.

(2) Sensitivity to realistic clutter coverage

The sensitivity of analysis and forecast to realistic clutter coverage, through experiments FRADCLUTd and FRADCLUTm experiments is shown in Fig. 6c,d in terms of analysis and forecast RMSEs. The error evolutions look very similar to corresponding ones in Fig. 6a,b. 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.

ETS scores are lower in FRADCLUTd0.5, by roughly 0.1 point for the entire domain (Fig. 10). 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 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, then the impact of the data gaps
may be larger.

(3) Sensitivity to isolated radar domain coverage

The RMSEs of a single isolated radar experiments are shown in Fig. 6e,f. 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 indicate 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.

Fig. 12a depicts the ETS scores 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 moist-biased background field, the scores are
much lower and more comparable to NoNm than to CNTLm.

To look into this problem further, Fig. 13 plots reflectivity for the ‘dry’ experiments
CNTLd and 1RDd0.5 in 15 minute intervals, starting at 2115 UTC (approximate CI time of
truth). Experiment 1RDd0.5 recaptures the specific CI location and timing of the truth pretty
closely, within 10 km and less than 10 minutes 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 both CNTLd and 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, possibly due to the discontinuity at the edge of radar data coverage found in the analysis. 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.

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 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.

6. 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. A better understanding of the behaviors of the analysis increments of $P$, $T$, and $q_v$ within a variational framework was obtained first through the analysis of a single refractivity observation. It was found that under typical warm-season conditions ($T > 20 \, ^\circ C$), corrections to the specific humidity dominate the analysis while corrections to pressure and temperature are negligible. The resulting increments depended mostly on the background temperature and the background moisture error, due to the inherent sensitivity of refractivity to moisture at differing temperatures combined with relatively large errors typical of NWP background fields in moisture.
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 ETS scores calculated against the truth for both analysis and forecast period. The analysis period included cycling of 3DVAR and ARPS forecasts every 10 minutes for an hour, with a 3-hour ARPS forecast initialized from final analysis. The main results are as follows:

- Control experiments assimilated idealized refractivity available every 4 km in the entire domain with 0.5 N-unit error. Control experiment analyses reduced the moisture RMSE by about 1-2 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 ETS scores 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 later forecast times in the storm evolution showed ETS which did not remain above the no-data assimilation experiments. This should be due to the lack of observations outside the single radar coverage area.
• Single isolated radar analysis can lead to artificial moisture gradients on the boundaries of data coverage that can cause spurious cell behaviors. This together with the results of the previous bullet indicates the importance of spatial coverage, and importance of possible combination with measures from other surface networks, such as a mesonet.

• Generally, a larger benefit was seen in refractivity assimilation experiments where the background moisture is low-biased than the case when the background moisture is high-biased.

The OSSEs demonstrated a positive potential for refractivity data to improve CI forecasts, a precursor for studying the impact of real refractivity data. Using data from a dense radar network, such as CASA, should provide an advantage for real data studies due to the availability of a larger domain with continuous coverage. The assimilation could also benefit from using more advanced data assimilation methods, such as the ensemble Kalman filter or 4DVAR technique. They would allow multi-variate analysis that would directly improve other model fields not directly linked to refractivity. These can be topics for future research.

Acknowledgement: This work was primarily supported by NSF grant AGS-0750790. The second author was also supported by NSF grants OCI-0905040, AGS-0802888, AGS-0941491, AGS-1046171, and AGS-1046081. The authors thank Tian-You Yu, Richard Doviak, Dusan Zrnić, Jerry Brotzge, David Bodine, and Boon-leng Cheong for their input during project meetings. The author would also like to acknowledge Adam Clark for providing the ETS code used in this study.
References


Table 1. List of OSSE experiments. For abbreviations, FRAD refers to Full RADar domain (data available everywhere at 4km resolution, see Fig. 6a), 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 subtracted or added (respectively), and numbers 0.5, 1.0, and 2.0 refer to homogeneous standard error amounts added to all refractivity observations. FRADCLUT refers to full radar experiments with realistic CLUTter coverage (see Fig. 5b)

<table>
<thead>
<tr>
<th>Experiment Name</th>
<th>Refractivity assimilated?</th>
<th>Obs. Error Amt. (N-units)</th>
<th>Add or Subtract 2.0 g kg⁻¹ to Background $q_v$</th>
<th>Radar Domain Coverage</th>
<th>Realistic Clutter Coverage (data discontinuity)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNTLd or FRADd0.5</td>
<td>Yes</td>
<td>0.5</td>
<td>Subtract (dry)</td>
<td>Full</td>
<td>No</td>
</tr>
<tr>
<td>CNTLm or FRADm0.5</td>
<td>Yes</td>
<td>0.5</td>
<td>Add (moist)</td>
<td>Full</td>
<td>No</td>
</tr>
<tr>
<td>NoNd</td>
<td>No</td>
<td>N/A</td>
<td>Subtract (dry)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NoNm</td>
<td>No</td>
<td>N/A</td>
<td>Add (moist)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>FRADd1.0</td>
<td>Yes</td>
<td>1.0</td>
<td>Subtract (dry)</td>
<td>Full</td>
<td>No</td>
</tr>
<tr>
<td>FRADd2.0</td>
<td>Yes</td>
<td>2.0</td>
<td>Subtract (dry)</td>
<td>Full</td>
<td>No</td>
</tr>
<tr>
<td>FRADm1.0</td>
<td>Yes</td>
<td>1.0</td>
<td>Add (moist)</td>
<td>Full</td>
<td>No</td>
</tr>
<tr>
<td>FRADm2.0</td>
<td>Yes</td>
<td>2.0</td>
<td>Add (moist)</td>
<td>Full</td>
<td>No</td>
</tr>
<tr>
<td>FRADCLUTd0.5</td>
<td>Yes</td>
<td>0.5</td>
<td>Subtract (dry)</td>
<td>Full</td>
<td>Yes</td>
</tr>
<tr>
<td>FRADCLUTd1.0</td>
<td>Yes</td>
<td>1.0</td>
<td>Subtract (dry)</td>
<td>Full</td>
<td>Yes</td>
</tr>
<tr>
<td>FRADCLUTm0.5</td>
<td>Yes</td>
<td>0.5</td>
<td>Add (moist)</td>
<td>Full</td>
<td>Yes</td>
</tr>
<tr>
<td>FRADCLUTm1.0</td>
<td>Yes</td>
<td>1.0</td>
<td>Add (moist)</td>
<td>Full</td>
<td>Yes</td>
</tr>
<tr>
<td>1RDd0.5</td>
<td>Yes</td>
<td>0.5</td>
<td>Subtract (dry)</td>
<td>Isolated</td>
<td>No</td>
</tr>
<tr>
<td>1RDm0.5</td>
<td>Yes</td>
<td>0.5</td>
<td>Add (moist)</td>
<td>Isolated</td>
<td>No</td>
</tr>
</tbody>
</table>
List of figures

Fig. 1. Effect of background model relative humidity error on 3DVAR analysis increments in (a) temperature (K), (b) pressure (hPa), and (c) specific humidity (g kg$^{-1}$). These analysis increments correspond with a single +5 N-unit observation innovation. Shades of gray indicate experiments with a background temperature value of 0°C (light gray), 20°C (gray), and 40°C (black). Dashed, solid, and dash-dotted lines represent experiments with a background relative humidity of 0%, 50%, and 100%, respectively. For example, the black dotted curves represent experiments where the background temperature is 40°C and the background relative humidity is 100%.

Fig. 2. Composite reflectivity (colored, scale in dBZ) and surface specific humidity (lines & gray shaded contours, scale in g kg$^{-1}$) for the truth model simulation of 19 May 2010 at (a) 2030 UTC, (b) 2130 UTC, (c) 2230 UTC, and (d) 2330 UTC. The red box in (b) represents a verification subdomain used for ETS calculations, located around the supercell which initiates at roughly 2110 UTC in the truth simulation.

Fig. 3. Simulated refractivity data for a single radar with 50-km continuous data coverage centered near the CI location of storm 2 valid 19 May 2010 at (a) 1930 UTC, (b) 2000 UTC, (c) 2030 UTC, and (d) 2100 UTC. This hypothetical radar is centered at horizontal grid point pair $(i, j) = (100, 200)$, which is less than 50 km North of CASA radar KCYR (Cyril, OK).

Fig. 4. 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 minutes. At 2130 UTC, a forecast is run. Results of refractivity assimilation experiments are compared with forecast experiments without data assimilation (NoNd and NoNm is Table 1), which serve as baselines. All results are validated against the truth simulation using RMSE and ETS calculations.

Fig. 5. (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. 3).

Fig. 6. RMSE 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 (d-e) for sensitivity to isolated radar experiments. Note that in (a-d), RMSE was calculated over the whole domain, and in (e-f) RMSE was calculated only over the 50-km radius of the isolated radar.

Fig. 7. RMSE vertical profiles of specific humidity at 2030 UTC (first analysis, no marker), 2130 UTC (final analysis, dashed marker) and 2230 UTC (1-hour forecast, square marker) for the control (black) and no-data assimilation (gray) experiments: (a) CNTLd, NoNd, and (b) CNTLm, NoNm.

Fig. 8. Composite reflectivity (color fill, dBZ) and specific humidity (gray contours, g kg$^{-1}$) for (a) the truth field and experiments (b) NoNd, (c) NoNm, (d) CNTLd, and (e) CNTLm, valid 2200 UTC 19 May 2010.
Fig. 9. Timeseries of ETS, using threshold of composite reflectivity at 30 dBZ, for baseline experiments NoNd, NoNm (gray curves) and control experiments CNTLd, CNTLm (black curves), calculated over the entire domain.

Fig. 10. ETS time series calculated using threshold of composite reflectivity at 30 dBZ for sensitivity experiments listed in Table 1: (a) Dry-bias experiments, (b) Moist-bias experiments. ETS calculated over the entire domain, and includes pre-existing storm at initial time. Dashed, vertical line indicates final analysis time (initial forecast time, at 2130 UTC).

Fig. 11. ETS time series calculated over the verification subdomain (red box in Fig. 2b) for observation error and clutter field sensitivity experiments in Table 1: (a) Dry-bias experiments, (b) Moist-bias experiments. As in Fig. 10, dashed vertical line indicates final analysis time at 2130 UTC.

Fig. 12. ETS time series calculated over the verification subdomain (red box in Fig. 2b) for single radar coverage experiments in Table 1: (a) Dry-bias experiments, (b) Moist-bias experiments. Control and no-data-assimilation experiments plotted for reference.

Fig. 13. Composite reflectivity (color fill, dBZ) 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. 2b (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 15 min. intervals.
Fig. 1. Effect of background model relative humidity error on 3DVAR analysis increments in (a) temperature (K), (b) pressure (hPa), and (c) specific humidity (g kg$^{-1}$). These analysis increments correspond with a single $+5$ N-unit observation innovation. Shades of gray indicate experiments with a background temperature value of 0°C (light gray), 20°C (gray), and 40°C (black). Dashed, solid, and dash-dotted lines represent experiments with a background relative humidity of 0%, 50%, and 100%, respectively. For example, the black dotted curves represent experiments where the background temperature is 40°C and the background relative humidity is 100%.
Fig. 2. Composite reflectivity (colored, scale in dBZ) and surface specific humidity (lines & gray shaded contours, scale in g kg$^{-1}$) for the truth model simulation of 19 May 2010 at (a) 2030 UTC, (b) 2130 UTC, (c) 2230 UTC, and (d) 2330 UTC. The red box in (b) represents a verification subdomain used for ETS calculations, located around the supercell which initiates at roughly 2110 UTC in the truth simulation.
Fig. 3. Simulated refractivity data for a single radar with 50-km continuous data coverage centered near the CI location of storm 2 valid 19 May 2010 at (a) 1930 UTC, (b) 2000 UTC, (c) 2030 UTC, and (d) 2100 UTC. This hypothetical radar is centered at horizontal grid point pair \((i, j) = (100, 200)\), which is less than 50 km North of CASA radar KCYR (Cyril, OK)
Fig. 4. 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 minutes. At 2130 UTC, a forecast is run. Results of refractivity assimilation experiments are compared with forecast experiments without data assimilation (NoNd and NoNm is Table 1), which serve as baselines.
Fig. 5. (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. 3)
RMSE was calculated over the whole domain, and in (e) for sensitivity to observation error experiments, (d) for sensitivity to realistic clutter domain experiments, and (d-e) for sensitivity to isolated radar experiments. Note that in (a-d), RMSE was calculated over the whole domain, and in (e-f) RMSE was calculated only over the 50-km radius of the isolated radar.

Fig. 6. RMSE 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 (d-e) for sensitivity to isolated radar experiments.
Fig. 7. RMSE vertical profiles of specific humidity at 2030 UTC (first analysis, no marker), 2130 UTC (final analysis, dashed marker) and 2230 UTC (1-hour forecast, square marker) for the control (black) and no-data assimilation (gray) experiments: (a) CNTLd, NoNd, and (b) CNTLm, NoNm.
Fig. 8. Composite reflectivity (color fill, dBZ) and specific humidity (gray contours, g kg$^{-1}$) for (a) the truth field and experiments (b) NoNd, (c) NoNm, (d) CNTLd, and (e) CNTLm, valid 2200 UTC 19 May 2010.
Fig. 9. Timeseries of ETS, using threshold of composite reflectivity at 30 dBZ, for baseline experiments NoNd, NoNm (gray curves) and control experiments CNTLd, CNTLm (black curves), calculated over the entire domain.
Fig. 10. ETS time series calculated using threshold of composite reflectivity at 30 dBZ for sensitivity experiments listed in Table 1: (a) Dry-bias experiments, (b) Moist-bias experiments. ETS calculated over the entire domain, and includes pre-existing storm at initial time. Dashed, vertical line indicates final analysis time (initial forecast time, at 2130 UTC).
Fig. 11. ETS time series calculated over the verification subdomain (red box in Fig. 2b) for observation error and clutter field sensitivity experiments in Table 1: (a) Dry-bias experiments, (b) Moist-bias experiments. As in Fig. 10, dashed vertical line indicates final analysis time at 2130 UTC.
Fig. 12. ETS time series calculated over the verification subdomain (red box in Fig. 2b) for single radar coverage experiments in Table 1: (a) Dry-bias experiments, (b) Moist-bias experiments. Control and no-data-assimilation experiments plotted for reference. As in Fig. 10, dashed vertical line indicates final analysis time at 2130 UTC.
Fig. 13. Composite reflectivity (color fill, dBZ) 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. 2b (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 15 min. intervals.