

# Direct Assimilation of Radar Reflectivity without Tangent Linear and Adjoint of the Nonlinear Observation Operator in the GSI-Based EnVar System: Methodology and Experiment with the 8 May 2003 Oklahoma City Tornado Supercell

YONGMING WANG AND XUGUANG WANG

*School of Meteorology, University of Oklahoma, Norman, Oklahoma*

(Manuscript received 17 June 2016, in final form 6 December 2016)

## ABSTRACT

A GSI-based EnVar data assimilation system is extended to directly assimilate radar reflectivity to initialize convective-scale forecasts. When hydrometeor mixing ratios are used as state variables (method mixing ratio), large differences of the cost function gradients with respect to the small hydrometeor mixing ratios and wind prevent efficient convergence. Using logarithmic mixing ratios as state variables (method logarithm) fixes this problem, but generates spuriously large hydrometeor increments partly due to the transform to and from the logarithmic space. The tangent linear of the reflectivity operators further contributes to spuriously small and large hydrometeor increments in method mixing ratio and method logarithm, respectively. A new method is proposed by directly adding the reflectivity as a state variable (method dBZ). Without the tangent linear and adjoint of the nonlinear operator, the new method therefore avoids the aforementioned problems.

The newly proposed method is examined on the analysis and prediction of the 8 May 2003 Oklahoma City tornadic supercell storm. Both the probabilistic forecast of strong low-level vorticity and maintenance of strong updraft and vorticity in method dBZ are more consistent with reality than in method logarithm and method mixing ratio. Detailed diagnostics suggest that a more realistic cold pool due to the better analyzed hydrometeors in method dBZ than in other methods leads to constructive interaction between the surface gust front and the updraft aloft associated with the midlevel mesocyclone. Similar low-level vorticity forecast and maintenance of the storm are produced by the WSM6 and Thompson microphysics schemes in method dBZ. The Thompson scheme matches the reflectivity distribution with the observations better for all lead times, but shows more southeastward track bias compared to the WSM6 scheme.

## 1. Introduction

Owing to high spatial and temporal resolution capable of sampling the structure of the convective storms, Doppler radar radial velocity and reflectivity are the most commonly used observations for storm-scale numerical weather prediction (NWP) (Sun 2005; Dowell et al. 2011). Numerous studies (Daley 1991; Sun and Crook 1997, 1998; Snyder and Zhang 2003; Dowell et al. 2004; Tong and Xue 2005; Caya et al. 2005; Xiao et al. 2005; Gao and Xue 2008; Xu et al. 2008; Jung et al. 2008; Dowell and Wicker 2009; Aksoy et al. 2009; Yussouf and Stensrud 2010; Lu and Xu 2009; Xue et al. 2009; Zhang et al. 2009; Dowell et al. 2011; Gao and Stensrud 2012; Sun and Wang 2013; Wang et al. 2013; Yussouf et al. 2013; Johnson et al. 2015) investigated the impact of radar data assimilation and indicated the benefits from

assimilating these data for convective-scale analysis and prediction. However, challenges for the assimilation of radar data still remain, especially for the reflectivity observations (Dowell et al. 2011; Gao and Stensrud 2012).

Early studies employed the cloud analysis method to adjust some of the first-guess variables, such as cloud liquid water, cloud ice, and precipitation species (rain, snow, and hail), based on the observed reflectivity to initialize the convective-scale NWP (Albers et al. 1996; Brewster 1996; Zhang et al. 1998; Zhang 1999; Souto et al. 2003; Xue et al. 2003; Hu et al. 2006). To mitigate the spinup problem in the short-range forecast associated with the cloud analysis, several methods were implemented to adjust the associated thermodynamical variables. These methods include a moisture and diabatic initialization scheme developed by Zhang et al. (1998) and Zhang (1999), an in-cloud temperature adjustment method modified by Hu et al. (2006), and a digital filter by Weygandt et al. (2008). Current operational or

---

Corresponding author e-mail: Dr. Xuguang Wang, xuguang.wang@ou.edu

semi-operational meso- and convective-scale NWP systems such as the Rapid Refresh (RAP), the High-Resolution Rapid Refresh (HRRR), and the North American Mesoscale Rapid Refresh (NAMRR) adopt the cloud analysis method and a digital filter technique to derive the hydrometeor mixing ratios from the reflectivity observations and to adjust the latent heating to match with these derived fields (Weygandt and Benjamin 2007; Weygandt et al. 2008; Carley et al. 2015; Benjamin et al. 2016). Although a positive impact of using an improved cloud analysis is found based on the description of the important characteristics of storms in early studies (e.g., Hu et al. 2006), the cloud analysis method relies on empirical algorithms to relate the hydrometeor variables and the reflectivity, and requires tuning of many uncertain parameters (e.g., Gao and Stensrud 2012).

Recently, an increasing number of studies have started to use the ensemble Kalman filter (EnKF) to directly assimilate reflectivity (e.g., Dowell et al. 2004; Tong and Xue 2005; Jung et al. 2008; Lei et al. 2009; Dowell et al. 2011; Yussouf et al. 2013; Johnson et al. 2015). EnKF estimates the covariance between the observed variables and the model state variables, including unobserved variables, through an ensemble of forecasts. Through the ensemble covariances, the reflectivity is used to directly update hydrometeor, thermal, and dynamical fields in a dynamically and thermodynamically coherent fashion (e.g., Dowell et al. 2004). These previous studies have demonstrated encouraging results using EnKF to assimilate radar data including direct assimilation of reflectivity, to improve the storm-scale analysis and forecast. As discussed in Dowell et al. (2011), efforts are still needed to further improve the EnKF assimilation of reflectivity. Such efforts include developing covariance localization methods that can be adaptive to the clear-air and precipitating regions, improving model simulated prior reflectivity by further reducing errors in microphysics parameterization scheme and observation operator, and developing methods to take into account the non-Gaussianity of both the background and observation errors, which violate the assumption in the EnKF.

Unique difficulties also remain in the variational data assimilation framework to directly assimilate reflectivity observations. One difficulty is associated with the tangent linear and adjoint of the nonlinear observation operator, which are typically used explicitly in the variational minimization. First, Sun and Crook (1997) found that in their four-dimensional variational data assimilation (4DVar) system, a large gradient of the observation term of the cost function with respect to the small rainwater mixing ratio prevents efficient convergence. Therefore, worse results were obtained from

direct assimilation of reflectivity than from assimilation of retrieved rainwater mixing ratios in their studies. Second, Wang et al. (2013a,b) found in their 3DVar and 4DVar systems that there is a large difference between the nonlinear reflectivity operator and its tangent linear approximation, especially for the small rainwater mixing ratios. This deficiency can lead to overestimation of the nonlinear reflectivity perturbations and therefore an underestimation of the rainwater content in the analysis. Therefore, instead of direct assimilation of the reflectivity observations, rainwater and water vapor were first retrieved from radar reflectivity before being assimilated by their 3DVar and 4DVar systems. However, due to the large errors associated with the rainwater retrieval for the strong convective region, only reflectivity less than 55 dBZ was used in their study. Carley (2012) also mentioned a similar problem of overestimated reflectivity perturbations if the hydrometeor mixing ratios are used as the state variables. In addition, Carley (2012) also discussed that using the logarithm of the mixing ratio can reduce this error but can also lead to underestimation of the nonlinear reflectivity perturbations. Third, the reflectivity operator depends on the microphysics schemes. Therefore, efforts to redevelop or update the tangent linear and adjoint of the reflectivity operator would be needed whenever the microphysics scheme is changed.

In the present study we propose a new method to overcome the above problems to achieve direct assimilation of radar reflectivity observations in the variational data assimilation framework. This new method extends the state variables to include the observed variable, reflectivity, and therefore does not need the tangent linear and adjoint of the reflectivity operator. This method thus avoids the aforementioned problem for direct reflectivity assimilation in the traditional variational method. The new method is derived and applied in the EnVar [see Lorenc (2013) and Wang and Lei (2014) for nomenclature]. Unlike the traditional variational system, EnVar implicitly uses ensemble covariances to estimate the background error covariance. This feature means that the new method does not require the use of the adjoint of the observation operator to spread the observed information to the model variables as used in the traditional variational system.

The new method is implemented in the GSI-based EnVar system. GSI-based EnVar and its variants have been successfully implemented for the NCEP GFS global forecast (Wang et al. 2013; Wang and Lei 2014; Kleist and Ide 2015; Kutty and Wang 2015). Efforts have been made since to further extend the system for convective-scale data assimilation. Lu et al. (2017) extended the GSI-based EnKF and EnVar system to

assimilate airborne radar radial wind observations for the Hurricane Weather Research and Forecasting Model (HWRF) for hurricane initialization and prediction. Johnson et al. (2015) extended the GSI-based EnKF to assimilate ground-based radar wind and reflectivity observations for multiple complex, convectively active cases with multiple storm modes and interactions in a heterogeneous environment. Carley (2012) assimilated radar observations in the GSI-based hybrid for the Non-hydrostatic Multiscale Model on the B grid (NMMB) through passive use of the coarse-resolution ensemble available from an independent source. The current study further extends the EnVar capability in GSI for convective-scale data assimilation by solving the aforementioned problems of direct assimilation of radar reflectivity observations in the variational framework.

The new method is applied to the analysis and prediction of the 8 May 2003 Oklahoma City (OKC) tornadic supercell storm. This study is therefore among the first published studies to investigate the aforementioned problems associated with the tangent linear and adjoint of the nonlinear operator within the coupled ensemble-variational (EnVar) method for direct convective-scale radar reflectivity assimilation using a real-data case.

The GSI-based EnVar algorithm, the development of the radar radial velocity assimilation, and the development of the reflectivity assimilation with various hydrometeor state variables including the newly proposed method in GSI-based EnVar are provided in section 2. Problems with using the hydrometeor mixing ratios and the logarithmic hydrometeor mixing ratios state variables for the direct reflectivity assimilation in GSI-based EnVar and how the new method overcomes these problems are investigated through analytical approaches in section 3. An overview of the 8 May 2003 Oklahoma City tornadic supercell storm, and the design of the experiments are then presented in the section 4. Section 5 describes the results of using various hydrometeor state variables including the new approach for the reflectivity assimilation on the analysis and prediction of the 8 May tornadic supercell storm case. A summary and discussion are given in the section 6.

## 2. Methodology

### a. GSI-based EnVar formulation

Lorenc (2013) and Wang and Lei (2014) defined the acronyms of various flavors of the coupled ensemble-variational data assimilation. This paper focuses on the EnVar flavor. Different from a hybrid system, an EnVar system utilizes a pure ensemble covariance rather than blending the ensemble covariance

with the static covariance. Wang (2010) described the mathematical details of the GSI-based EnVar formulations. For consistency, the description of the formula herein parallels that of Wang (2010).

Briefly, the analysis increment  $\mathbf{x}'$  is defined as

$$\mathbf{x}' = \sum_{k=1}^K (\mathbf{a}_k \circ \mathbf{x}_k^e). \quad (1)$$

In Eq. (1),  $\mathbf{x}_k^e$  denotes the  $k$ th ensemble perturbation divided by  $\sqrt{K-1}$ . The  $K$  denotes the ensemble size. The vectors  $\mathbf{a}_k$ ,  $k = 1, \dots, K$ , are the unitless control vectors (CV) for each ensemble member, and form the vector  $\mathbf{a}$ . The symbol  $\circ$  is the Schur product. The following cost function is minimized to obtain the analysis increment:

$$J(\mathbf{a}) = 0.5(\mathbf{a})^T \mathbf{A}^{-1}(\mathbf{a}) + 0.5(\mathbf{y}' - \mathbf{H}\mathbf{x}')^T \mathbf{R}^{-1}(\mathbf{y}' - \mathbf{H}\mathbf{x}'). \quad (2)$$

The first term of Eq. (2) on the right-hand side is associated with the flow-dependent covariance. Here  $\mathbf{A}$  is a block-diagonal matrix and defines the localization applied to the ensemble covariance. The last term is the observation term ( $J_o$ ), where  $\mathbf{R}$  is the observation error covariance,  $\mathbf{H}$  is the linearized observation operator, and  $\mathbf{y}'$  is the innovation vector. The gradient of the cost function [Eq. (2)] with respect to control variables  $\mathbf{a}$  is given as

$$\nabla_{\mathbf{a}} J = \mathbf{A}^{-1} \mathbf{a} + \mathbf{D}^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}' - \mathbf{y}'), \quad (3)$$

where  $\mathbf{D}$  is denoted as  $[\text{diag}(\mathbf{x}_1^e) \dots \text{diag}(\mathbf{x}_K^e)]$ . Note that the GSI-based EnVar houses both a three-dimensional algorithm called 3DEnVar (Wang et al. 2013) and its four-dimensional extension (4DEnVar; e.g., Wang and Lei 2014). As detailed in the appendix of Wang and Lei (2014), 4DEnVar extends both the increment [Eq. (1)] and the observation term in the cost function [Eq. (2)] from 3D to 4D. The new algorithm method dBZ in this study applies for both 3DEnVar and 4DEnVar. Therefore, in the algorithm description we do not specify 3DEnVar or 4DEnVar. It is noted, however, that the experiments in this study are all conducted using 3DEnVar.

### b. Developing radar radial velocity assimilation in GSI-based EnVar

Johnson et al. (2015) described extending the GSI-based EnKF to assimilate radar radial velocity and reflectivity. This paper focuses on introducing the extension of the GSI-based EnVar to assimilate these data directly. Both the forward and adjoint of the radial

velocity ( $V_r$ ) observation operator are implemented in EnVar. The forward operator for  $V_r$  is calculated as

$$V_r = u \sin \varphi \cos \mu + v \cos \varphi \cos \mu + w \sin \mu, \quad (4)$$

where  $\mu$  is the elevation angle and  $\varphi$  is the azimuth angle of radar beams (Sun 2005; Gao and Stensrud 2012); and  $u$ ,  $v$ , and  $w$  represent zonal, meridional, and vertical velocities, respectively. In Eq. (1),  $u$ ,  $v$ , and  $w$  are included in  $\mathbf{x}'_k$ .

### c. Developing radar reflectivity assimilation in GSI-based EnVar

Reflectivity is a measure of the signals returned from the remote hydrometeors to the radar receivers (Doviak and Zrnić 1993). To investigate the impacts of the different hydrometeor state variables for direct assimilation of reflectivity, hydrometeor mixing ratios (method mixing ratio), logarithmic hydrometeor mixing ratios (method logarithm), and the newly proposed method of augmenting the state variables to directly use the reflectivity itself as the state variable (method dBZ) are implemented in the GSI-based EnVar system with the Advanced Research Weather Research and Forecasting Model (WRF-ARW). To illustrate these methods, the radar reflectivity observation operator consistent with the WSM6 microphysics scheme (Hong and Lim 2006) is used following Lin et al. (1983), Gilmore et al. (2004), Tong and Xue (2005), Dowell et al. (2011), and Johnson et al. (2015).

#### 1) USE OF HYDROMETEOR MIXING RATIOS AS STATE VARIABLES (METHOD MIXING RATIO)

In this method, rainwater  $q_r$ , snow  $q_s$ , and graupel  $q_g$  mixing ratios are assigned in  $\mathbf{x}'$ . Both the tangent linear and adjoint of the reflectivity observation operator with respect to these hydrometeor mixing ratios are implemented in the GSI-based EnVar.

The observation operator for reflectivity follows Johnson et al. (2015). The operator was developed for the WSM6 microphysics scheme based on Lin et al. (1983), Tong and Xue (2005), Dowell et al. (2011), and Gao and Stensrud (2012). Mathematically,

$$H(q_r, q_s, q_g) = Z_{\text{dB}} = 10 \log Z_e, \quad (5)$$

where  $Z_e$  is the equivalent radar reflectivity factor and is contributed by the equivalent radar reflectivity factor from all three hydrometeors—rainwater  $Z_r$ , snow  $Z_s$ , and graupel  $Z_g$  as follows:

$$Z_e = Z_r + Z_s + Z_g. \quad (6)$$

Hereafter, to be concise, only the contribution from graupel is used as a representative example. The

equivalent reflectivity factor contributed from the graupel mixing ratio is defined as

$$Z_g = 4.33 \times 10^{10} \times (\rho q_g)^{1.75}, \quad (7)$$

where  $\rho$  is the air density and  $q_g$  is the graupel mixing ratio. The tangent linear of  $H(q_r, q_s, q_g)$  in Eq. (5) is derived by adding a small perturbation to the hydrometeor mixing ratio and by only maintaining the linear term after the Taylor expansion. For example, the tangent linear of  $H(q_r, q_s, q_g)$  with respect to  $q_g$  is

$$\mathbf{H}_{q_g} = \frac{10 \times 4.33 \times 10^{10} \times \rho^{1.75} \times 1.75 \times (q_g)^{0.75}}{\log(10) \times Z_e}. \quad (8)$$

Note that Eq. (8) is valid only when  $Z_e$  is greater than zero. In this study, the rainwater, snow, and graupel mixing ratios are reset to be  $10^{-6}$ ,  $10^{-9}$ , and  $10^{-8} \text{ kg kg}^{-1}$  if the values are smaller than these thresholds. Similar values are used in the WRF WSM6 microphysics scheme.

#### 2) USE OF LOGARITHMIC TRANSFORM OF HYDROMETEOR MIXING RATIOS AS STATE VARIABLES (METHOD LOGARITHM)

In method logarithm, the reflectivity observation operator is the same as method mixing ratio. The primary difference lies in the hydrometeors used in the  $\mathbf{x}'$  term in Eq. (2). Instead of assigning the mixing ratios to  $\mathbf{x}'$ , the logarithmic transform is applied to the hydrometeor mixing ratios and these transformed variables are introduced into  $\mathbf{x}'$  in the variational minimization. In this study, this method is implemented in the GSI-based EnVar for WRF-ARW. A similar approach was implemented for the NMMB model (Carley 2012).

We denote  $\hat{q}_r$ ,  $\hat{q}_s$ , and  $\hat{q}_g$  for the rainwater, snow, and graupel mixing ratios in logarithmic form, respectively, which are defined as  $\hat{q}_r = \log(q_r)$ ,  $\hat{q}_s = \log(q_s)$ , and  $\hat{q}_g = \log(q_g)$ . The equivalent reflectivity factor  $Z_g$  contributed from the graupel mixing ratio is then defined as

$$Z_g = 4.33 \times 10^{10} \times (\rho \times 10^{\hat{q}_g})^{1.75}. \quad (9)$$

The tangent linear of  $H(\hat{q}_r, \hat{q}_s, \hat{q}_g)$  with respect to  $\hat{q}_g$  is

$$\mathbf{H}_{\hat{q}_g} = \frac{10 \times 4.33 \times 10^{10} \times \rho^{1.75} \times 1.75 \times (10^{\hat{q}_g})^{0.75}}{Z_e}. \quad (10)$$

Both the forward and tangent linear of the reflectivity operator in method logarithm are not valid when  $Z_e$  is zero. Similar to rainwater, snow and graupel mixing ratios thresholds as in method mixing ratio are used to avoid this problem.

### 3) AUGMENTING STATE VARIABLES WITH REFLECTIVITY (METHOD dBZ)

As mentioned by Sun and Crook (1997) and discussed in detail in the next section, the variational minimization in method mixing ratio has difficulty to converge. Together with the linear approximation of the nonlinear operator, method mixing ratio produces spuriously small hydrometeor increments. As shown in this study, method logarithm is able to alleviate this large gradient problem. However, method logarithm produces spuriously large hydrometeor increments. A new method (method dBZ) is introduced and described in this subsection to solve these problems. Method dBZ augments the state variables in  $\mathbf{x}'$  in Eq. (2) by directly including the reflectivity variable as a state variable. Therefore, the tangent linear of  $H$  with respect to reflectivity is an identity:

$$\mathbf{H}_{\text{dBZ}} = \mathbf{I}. \quad (11)$$

Since the tangent linear and adjoint (TLA) of an identity matrix is also equal to an identity, the TLA of the reflectivity operator no longer exists in the EnVar minimization in method dBZ. In other words, method dBZ allows an EnVar to avoid the TLA of the nonlinear operator by extending the model state variables to include the observed variables. Similar extension of the state variable was adopted for a pure EnKF with a different purpose, which was to enable a parallel implementation of a serial ensemble filter (Anderson and Collins 2007). Independent studies by Zupanski (2005) and Liu et al. (2008) suggested ways of using an ensemble in the variational framework to avoid the use of the TLA of the observation operator. Unlike this study, neither Zupanski (2005) nor Liu et al. (2008, 2009) includes the ensemble covariance localization inside the variational minimization. It can be proven (not shown) that their solutions are ultimately equivalent to Eq. (11) (i.e., the algorithm used in method dBZ) if covariance localization is not implemented inside the variational minimization. In addition, unlike Zupanski (2005) and Liu et al. (2008), this study motivates the new method dBZ not only for the convenience of avoiding the development of the TLA of a nonlinear observation operator, but also more importantly through revealing the fundamental, methodological problems of using the TLA of the nonlinear operator in EnVar.

The reflectivity operator is often dependent on the microphysics schemes. Method dBZ by design does not use the TLA of the reflectivity operator. Therefore, once the forward reflectivity is obtained, the variational minimization in method dBZ does not depend on the

microphysics scheme. Method dBZ also saves the efforts of developing the TLA of the nonlinear operator, which is nontrivial especially for complicated operators.

### 3. Problems with hydrometeor mixing ratio (method mixing ratio) and logarithmic hydrometeor mixing ratio (method logarithm) in GSI-based EnVar

#### a. Cost function gradient

Sun and Crook (1997) mentioned the difficulty of direct assimilation of reflectivity observations in their 4DVar system because the low values of rainwater mixing ratios can lead to large gradients of the observation term  $J_o$  of the cost function, which can prevent efficient convergence during the variational minimization. Therefore, in their 4DVar system, the direct assimilation of reflectivity performed worse than assimilating the rainwater mixing ratios retrieved from reflectivity. In this subsection, the impact of using the three methods of hydrometeor variables on the gradient of the cost function for the GSI-EnVar is revealed using both an analytical approach and an approach with the sample of data collected over a region of active convection (35.14°–36.07°N, 97.82°–96.81°W, hereafter “active region”) from the 8 May tornadic supercell experiment.

Figure 1 shows the gradient of  $J_o$  with respect to  $\mathbf{a}$  for the three different reflectivity DA methods. The following procedures are followed to produce Fig. 1. In the  $\nabla_{\mathbf{a}} J$  term of Eq. (3), the gradient of  $J_o$  is given by

$$\nabla_{\mathbf{a}} J_o = \mathbf{D}^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}' - \mathbf{y}^{o'}), \quad (12)$$

and the term  $\mathbf{H}\mathbf{x}' - \mathbf{y}^{o'}$  is assumed to be 25 dBZ for the reflectivity and  $4 \text{ m s}^{-1}$  for the radial velocity. The observation error standard deviation is assumed to be 5 dBZ for the reflectivity and  $2 \text{ m s}^{-1}$  for the radial velocity, respectively. Similar observation error values were used in early studies for storm-scale radar data assimilation experiments (Dowell et al. 2004; Yussouf et al. 2013; Johnson et al. 2015). To illustrate the problem of the cost function gradient for method mixing ratio, a 45-member first-guess ensemble valid at 2150 UTC 8 May 2003 over the active region is used. Each grid point is assumed to have one radial velocity and one reflectivity observation located at an elevation angle of  $10^\circ$  and an azimuth angle of  $45^\circ$ . The gradient of  $J_o$  with respect to the extended control variables  $\mathbf{a}$  is calculated for each grid point and each ensemble member in this region. The contributions of the radial velocity and the radar reflectivity to this gradient are separately plotted in Fig. 1. In other words,

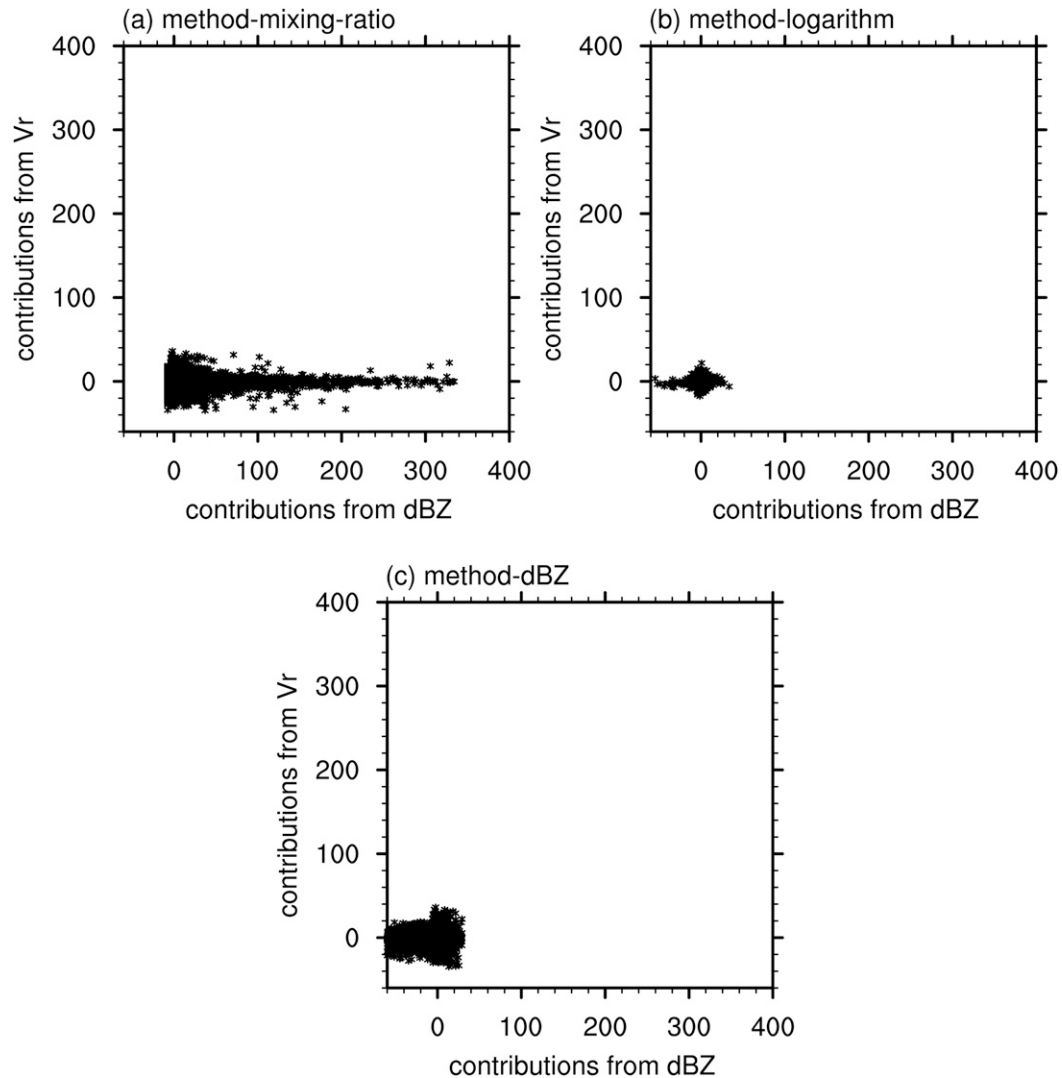


FIG. 1. Contributions (unit: 1) to the gradient of  $J_o$  from reflectivity (dBZ,  $x$  axis) and radial velocity ( $V_r$ ,  $y$  axis) for (a) method mixing ratio, (b) method logarithm, and (c) method dBZ.

each point in Fig. 1 is a pair of  $J_o$  gradients contributed from the radial velocity and reflectivity.

The gradients of  $J_o$  contributed from radial velocity for all three reflectivity DA methods are the same. The  $\mathbf{D}^T$  in Eq. (12) comprises the ensemble perturbations of zonal  $u$ , meridional  $v$ , and vertical  $w$  velocities. The  $\mathbf{H}^T$  in Eq. (12) is the transpose of the tangent linear of the radial velocity forward operator defined in Eq. (4). Similar procedures are applied to calculate the gradient of  $J_o$  contributed from dBZ for all three reflectivity DA methods. In method mixing ratio,  $\mathbf{D}^T$  is formed from the ensemble perturbations of mixing ratios for rain  $q_r$ , snow  $q_s$ , and graupel  $q_g$ . The  $\mathbf{H}^T$  is the transpose, or adjoint, of the tangent linear of  $H(q_r, q_s, q_g)$  with respect to  $q_r$ ,  $q_s$ , and  $q_g$ . An example of the tangent linear of  $H(q_r, q_s, q_g)$  with respect to  $q_g$ ,  $\mathbf{H}_{q_g}$  is given by Eq. (8). A

similar derivation is applied for  $\mathbf{H}_{q_r}$  and  $\mathbf{H}_{q_s}$ . In the calculation of the tangent linear of the reflectivity operator, the ensemble mean field is used as the reference.

In method logarithm, the ensemble perturbations of the logarithmic hydrometeor mixing ratios are used to define  $\mathbf{D}^T$ . The  $\mathbf{H}^T$  is the transpose of the tangent linear of  $H(\hat{q}_r, \hat{q}_s, \hat{q}_g)$  with respect to logarithmic rainwater  $\mathbf{H}_{\hat{q}_r}$ , snow  $\mathbf{H}_{\hat{q}_s}$ , and graupel  $\mathbf{H}_{\hat{q}_g}$  mixing ratios. The derivations for  $\mathbf{H}_{\hat{q}_s}$  and  $\mathbf{H}_{\hat{q}_r}$  are similar to Eq. (10) for  $\mathbf{H}_{\hat{q}_g}$ .

In method dBZ,  $\mathbf{D}^T$  includes the ensemble perturbations of the reflectivity. As shown in section 2c(3), since the tangent linear of the reflectivity operator are identities in method dBZ,  $\mathbf{H}^T$  is an identity.

As shown in Fig. 1, in method mixing ratio, the gradient of  $J_o$  contributed from reflectivity is dominantly large, where the largest value is greater than 300. In

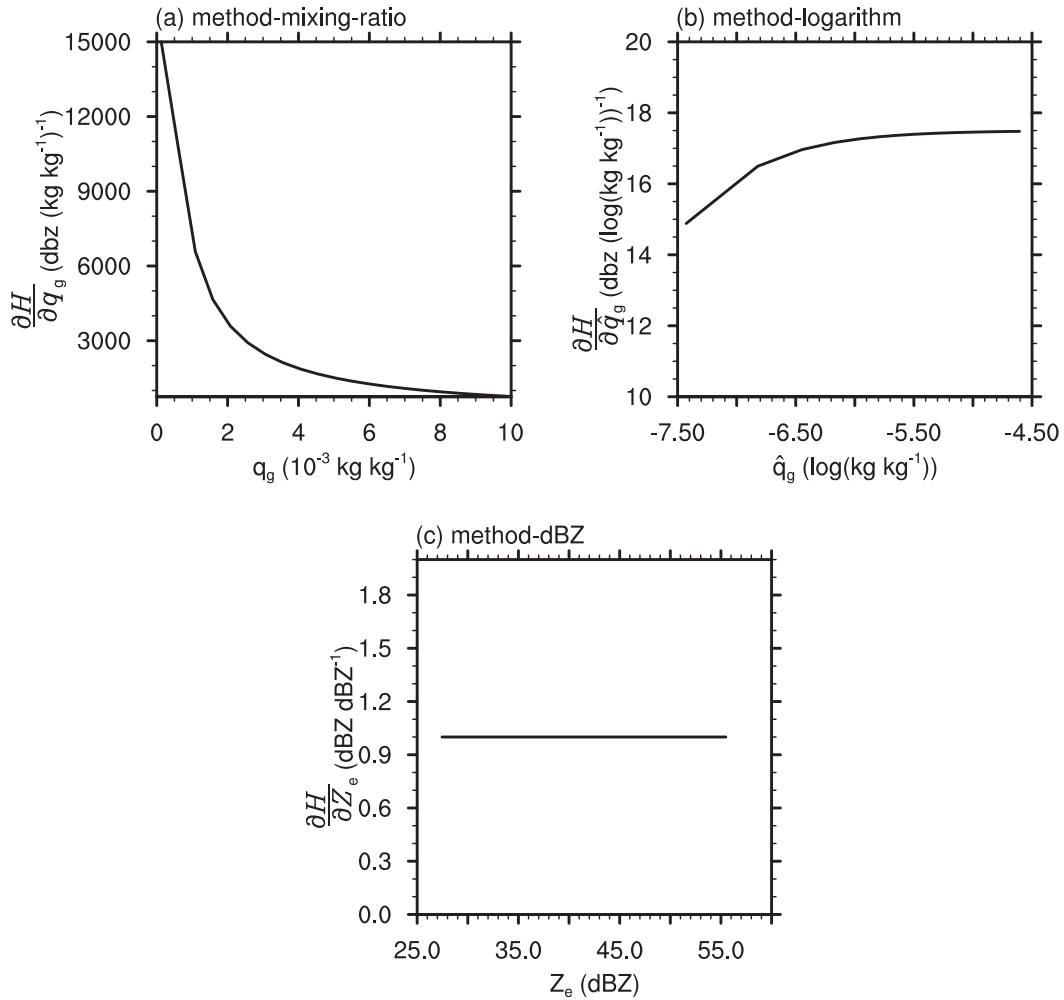


FIG. 2. Derivative of reflectivity with respect to (a) graupel mixing ratio, (b) logarithmic graupel mixing ratio, and (c) reflectivity itself. The figure uses the same mixing ratio for rainwater and snow, which is  $1 \times 10^{-4} \text{ kg kg}^{-1}$  and graupel mixing ratio varies from  $1 \times 10^{-4}$  to  $1 \times 10^{-2} \text{ kg kg}^{-1}$  for the calculation.

comparison, the maximum gradient of  $J_o$  contributed from radial velocity is less than 60. Such large differences in the cost function gradients can cause imbalance during the variational minimization and prevent efficient convergence. This result is consistent with that of 4DVar in Sun and Crook (1997). The gradients of  $J_o$  contributed from reflectivity are greatly reduced in both method logarithm and method dBZ. Both maxima are less than 60, which is similar to the gradient of  $J_o$  contributed from radial velocity.

The following is an analytical explanation of the differences of the  $J_o$  gradients among the three hydrometeor variable methods. Figure 2a shows the tangent linear of the reflectivity operator with respect to the graupel mixing ratio in method mixing ratio [Eq. (8)]. The tangent linear of the reflectivity operator in method mixing ratio has large values, especially with small values of

graupel mixing ratio. For example, when the graupel mixing ratio is  $1 \times 10^{-4} \text{ kg kg}^{-1}$ , the  $\mathbf{H}_{q_g}$  term is larger than  $15000 \text{ dBZ (kg kg}^{-1})^{-1}$ . This is further confirmed by examining Eq. (8). For example, if only the graupel mixing ratio contributes to the reflectivity, the simplified Eq. (8) will be

$$\mathbf{H}_{q_g} = \frac{17.5}{\log(10) \times (q_g)}. \tag{13}$$

Because the variable  $q_g$  is in the denominator, a small  $q_g$  when the reflectivity is close to zero can lead to a very large value returned from the linearized operator. A similar derivation of Eq. (13) was employed in Sun and Crook (1997) and Wang et al. (2013a) in their 4DVar and 3DVar systems. Sun and Crook (1997) have attempted to alleviate the large gradient problem by setting the gradient of

$J_o$  to be zero if the mixing ratio is less than a specified value. However, they found that this modification of the gradient problem led to large forecast errors in the regions where the mixing ratio is lower than the specified value.

After applying the logarithmic transform to the hydrometeor mixing ratios in method logarithm, the value returned from the linearized observation operator is greatly reduced to be no larger than about  $17.5 \text{ dBZ} [\log(\text{kg kg}^{-1})]^{-1}$  (Fig. 2b). According to Eq. (10), when the graupel mixing ratio is relatively large, the reflectivity contributed from rainwater and snow mixing ratios can be neglected in the denominator. Because the logarithmic mixing ratios related terms in the numerator and denominator almost cancel each other out, the values returned from the linearized operator approximate a constant  $17.5 \text{ dBZ} [\log(\text{kg kg}^{-1})]^{-1}$ . When the graupel mixing ratio is small and equivalent to the other types of hydrometeor mixing ratios, the value returned from the linearized operator can be less than  $17.5 \text{ dBZ} [\log(\text{kg kg}^{-1})]^{-1}$  (Fig. 2b). As discussed in section 2c(3), the observation operator in method dBZ is an identity. Therefore, its tangent linear value is reduced to be equal to  $1 \text{ dBZ dBZ}^{-1}$  (Fig. 2c).

In Eq. (12), the innovation term  $\mathbf{H}\mathbf{x}' - \mathbf{y}'$  and the observation error variances  $\mathbf{R}$  are fixed. Therefore, the large gradient of  $J_o$  in method mixing ratio is attributed to  $\mathbf{D}^T$  and  $\mathbf{H}^T$ . Further analysis of the results in Figs. 1 and 2 shows that the anomalously large  $J_o$  gradient contributed from the reflectivity in Fig. 1a in method mixing ratio is primarily attributed to the large value of the gradient of the reflectivity operator (i.e.,  $\mathbf{H}^T$ ) with respect to small hydrometeor mixing ratio values.

### b. Problem with the logarithmic transform

In method logarithm, the logarithmic hydrometeor mixing ratios are updated and then transformed back into the mixing ratio space. Since the logarithmic transform is nonlinear, the analysis produced in logarithmic space is not equivalent to the analysis produced without the logarithmic transform. To illustrate the impact of the logarithmic transform on the increments, single-pseudo-observation experiments are used following this equation,

$$\delta\mathbf{x} = \frac{\mathbf{P}^b \mathbf{H}^T}{\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}} (\mathbf{y} - \mathbf{H}\mathbf{x}_b), \quad (14)$$

where  $\mathbf{P}^b$  is the background error covariance,  $\mathbf{H}$  is the observation operator,  $\mathbf{y}$  is the observation vector,  $\mathbf{R}$  is the observation error covariance, and  $\mathbf{x}_b$  is the background (Lorenz 1986). This formula allows the impact of the logarithmic transform to be isolated without being affected by the linear approximation of the nonlinear reflectivity operator, which will be discussed section 3c.

In the pseudo-observation experiments, a radar reflectivity observation with an innovation of  $10 \text{ dBZ}$  and an observation error of  $5 \text{ dBZ}$  is assimilated. In other words,  $\mathbf{y} - \mathbf{H}\mathbf{x}_b$ ,  $\mathbf{H}\mathbf{P}^b\mathbf{H}^T$ , and  $\mathbf{R}$  are the same for method mixing ratio, method logarithm, and method dBZ. Through Eq. (14), when a reflectivity observation is assimilated method mixing ratio and method dBZ update rainwater, snow, and graupel mixing ratios directly through the cross correlation between reflectivity and hydrometeor mixing ratios estimated by the flow-dependent ensemble error covariance. The increments of hydrometeor mixing ratios in method mixing ratio and method dBZ are the same. In method logarithm, rainwater, snow, and graupel mixing ratios in logarithmic form are updated through the cross correlation between reflectivity and logarithmic hydrometeor mixing ratio estimated by the flow-dependent ensemble error covariance. To collect a large sample of ensemble covariances from the 8 May case experiments, a number of pseudo-single-observations located at all the model grid points over the active region are created. Note that the ensemble from which the ensemble covariance is derived for method mixing ratio and method logarithm is the same. Here the graupel mixing ratio valid at 2150 UTC is chosen as an example in Fig. 3.

Figure 3 shows that the increments in method logarithm is far greater than that in method mixing ratio and method dBZ, and the maximum increments in method logarithm are almost 2 times larger than in the other two methods.

To further explain the cause of such differences, we choose a point valid at 2150 UTC near the hook echo and calculate the increments using Eq. (14) as a function of a range of innovation values from  $1$  to  $35 \text{ dBZ}$  for method mixing ratio and method logarithm, respectively. Note again,  $\mathbf{y} - \mathbf{H}\mathbf{x}_b$ ,  $\mathbf{H}\mathbf{P}^b\mathbf{H}^T$ , and  $\mathbf{R}$  are the same for both method mixing ratio and method logarithm. Therefore, the differences seen are solely attributed to the difference between using the mixing ratios or their logarithmic forms as state variables and the forward and backward transform to and from the logarithm. As shown in Fig. 4, the increment is amplified during the transform. For example, when the graupel mixing ratio increment is  $0.60 \times 10^{-3} \text{ kg kg}^{-1}$  in method mixing ratio, the corresponding graupel mixing ratio increment in logarithmic space is  $0.57 \log(\text{kg kg}^{-1})$  in method logarithm, as shown on the left axis in Fig. 4. After transforming back to the mixing ratio space, the mixing ratio increment corresponding to  $0.57 \log(\text{kg kg}^{-1})$  is  $1.23 \times 10^{-3} \text{ kg kg}^{-1}$  as shown on the right axis in Fig. 4. This is twice as large as the increment by method mixing ratio of  $0.60 \times 10^{-3} \text{ kg kg}^{-1}$ . Method mixing ratio and method dBZ do not share the same problem as no logarithmic transform is involved.



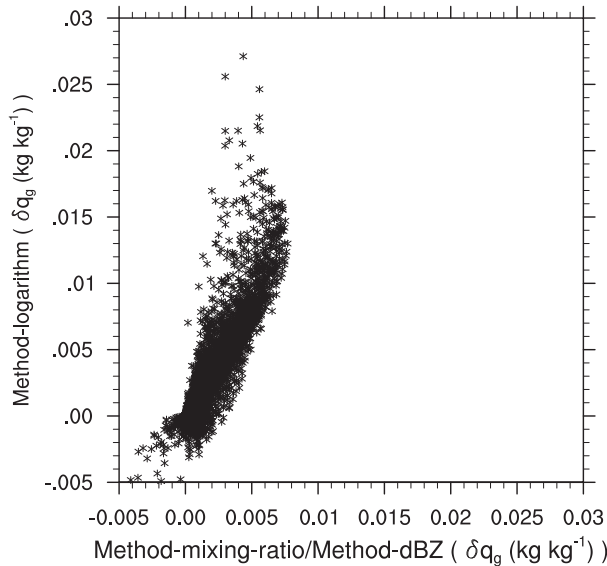


FIG. 3. Graupel mixing ratio increment in method logarithm (y axis) against that in method mixing ratio and method dBZ (x axis). See text for more details.

*c. Problem associated with the linear approximation of the nonlinear reflectivity operator*

Another cause of errors in the hydrometeor increments is specific to the use of the tangent linear of the reflectivity operator (i.e., the linear approximation of a nonlinear operator in the variational minimization). Within the variational method, the tangent linear approximation  $\mathbf{H}$  is applied in the observation term of the cost function [Eq. (2)]. During the minimization,  $\mathbf{H}\mathbf{x}'$  is iteratively adjusted toward the innovation  $\mathbf{y}^o$ . However, since  $\mathbf{H}$  is a linear approximation of the full nonlinear operator, such an approximation may lead to errors in the increment  $\mathbf{x}'$ . Such error applies to both method mixing ratio and method logarithm because both involve the tangent linear approximation. Such an approximation is not applied in method dBZ since its operator is an identity. Figure 5 shows the tangent linear reflectivity perturbations  $\mathbf{H}_{q_g} \times \delta q_g$  and  $\mathbf{H}_{\hat{q}_g} \times \delta \hat{q}_g$ , respectively, for method mixing ratio and method logarithm as a function of the graupel mixing ratio increments. The averaged mixing ratios of rainwater, snow, and graupel from the active region valid at 2150 UTC are used to calculate  $\mathbf{H}_{q_g}$  and  $\mathbf{H}_{\hat{q}_g}$  in Fig. 5. The corresponding true reflectivity perturbation using the nonlinear (NL) operator is also shown. In Fig. 5, when the reflectivity perturbation is small, such as less than 4 dBZ, the increments of graupel mixing ratio in all three methods are almost equal. This result is within expectation as the error associated with the linear approximation increases with an increasing increment. When the innovation increases from 4 to

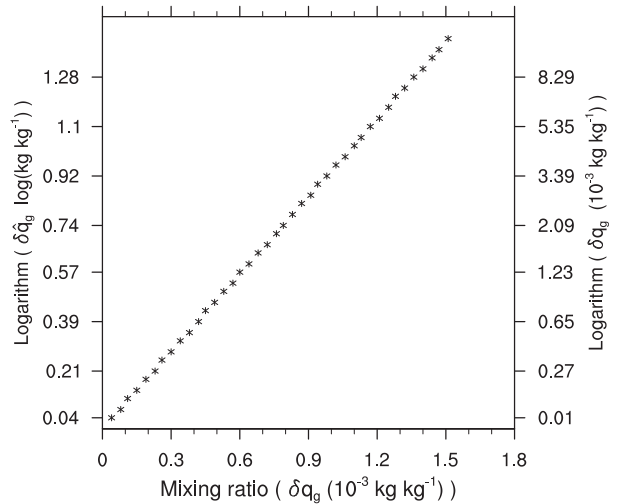


FIG. 4. Graupel mixing ratio increment in method logarithm (y axis) against that in method mixing ratio/method dBZ (x axis) corresponding to the innovation increasing from 1 to 35 dBZ. The left y axis represents the increments in logarithmic form and the right y axis is the mixing ratio increments backward transformed from the logarithm space. The graupel mixing ratio first guess is  $4.53 \times 10^{-4} \text{ kg kg}^{-1}$  for the calculation in this figure.

8 dBZ, the increments of the graupel mixing ratios in method mixing ratio are still a close approximation to NL, but the increments in method logarithm are larger than NL. In other words, the graupel mixing ratios are overly adjusted. For example, when the innovation of the reflectivity is 7 dBZ, the increments of the graupel mixing ratios in method mixing ratio and NL are  $\sim 0.9 \times 10^{-3} \text{ kg kg}^{-1}$ , while the increment in method logarithm is  $\sim 2.0 \times 10^{-3} \text{ kg kg}^{-1}$ . Further increasing the innovation over 10 dBZ, the linear approximation of the observation operator in method mixing ratio underestimates the increments of the graupel mixing ratios while method logarithm still overestimates the increments. The underestimated hydrometeor mixing ratio increments in method mixing ratio revealed in Fig. 5 are also explained by Wang et al. (2013a) in their 3DVar system. The overestimated hydrometeor mixing ratio increments in method logarithm demonstrated in Fig. 5 are consistent with that briefly discussed in Carley (2012). In method dBZ, the reflectivity itself is applied as the state variable. As a result, the errors associated with the linearization of the nonlinear reflectivity operator are avoided. In addition, because this new method method dBZ does not use the tangent linear of the operator, it avoids the outer loops, which are used in traditional variational systems (Rosmond and Xu 2006) to reduce errors associated with the linear approximation of the nonlinear operator. More detailed discussion is offered in section 6. Note also that the problems

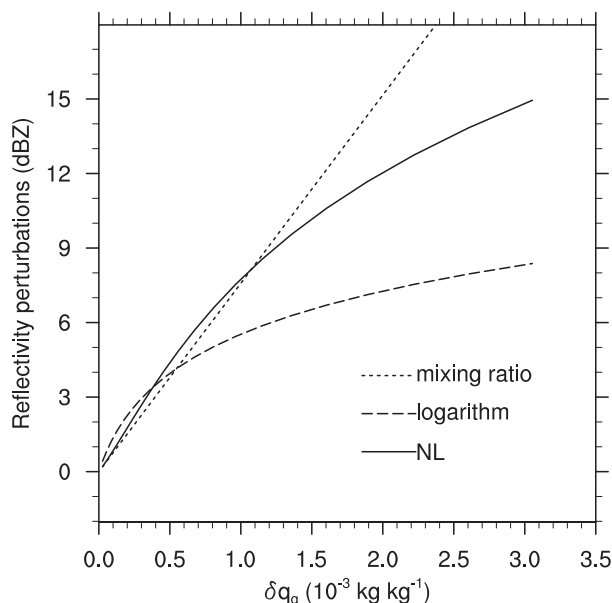


FIG. 5. Relationship between radar reflectivity perturbations and graupel mixing ratio increments for the tangent linear reflectivity operator w.r.t hydrometeor mixing ratios (dotted line) and logarithm of hydrometeor mixing ratio (dashed line). The nonlinear (NL) reflectivity perturbation (solid line) used in method dBZ is shown as a reference.

associated with method logarithm and method mixing ratio revealed in this section do not depend on microphysics schemes.

#### 4. Experiment design

In this section, GSI-based EnVar with the three different reflectivity DA methods is applied to analysis and prediction of the 8 May 2003 Oklahoma City supercell case. The impacts of using the three methods are studied from both the rationality of the analyses and the longevity of the predicted storms. The experiment design is described in this section followed by the results in section 5.

##### a. Overview of the 8 May 2003 Oklahoma City tornadic supercell case

The 8 May 2003 Oklahoma City (OKC) isolated supercell produced a violent tornado rated F-4 on the Fujita scale in Moore, Oklahoma. On 8 May 2003, the overall synoptic-scale environment over Oklahoma was governed by a typical weather pattern conducive for tornadic storms. A dryline was located in west-central Oklahoma with several storm cells along it by 2050 UTC, and a new cell that emerged by  $\sim$ 2100 UTC. This new cell developed into an isolated supercell in the next hour. The supercell moved northeastward and spawned a

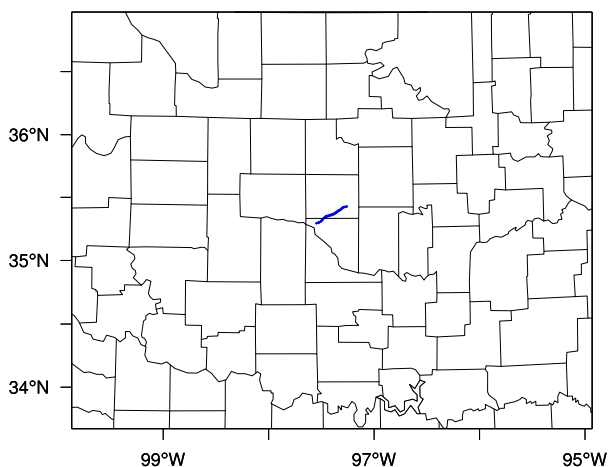


FIG. 6. The domain for the storm-scale ensemble analyses and forecasts. The overlaid blue line in the figure is the NWS-observed tornado damage track during 2210–2238 UTC.

violent tornado from 2210 to 2238 UTC, which tracked east-northeastward for about 30 km (Yussouf et al. 2013). The supercell storm developed and moved east-northeastward and lasted for more than 3 h starting from  $\sim$ 2100 UTC, weakened by 2300 UTC 8 May, and finally dissipated after 1 h at  $\sim$ 0020 UTC 9 May (Hu and Xue 2007).

##### b. Design of assimilation and forecasts experiments

In this study, a single domain (Fig. 6), which is configured the same as the inner domain of Yussouf et al. (2013), is adopted. This domain uses a grid spacing of 2 km with  $226 \times 181$  horizontal grid points and 50 vertical levels. A 45-member ensemble is run. A similar number of ensemble members were used in earlier convective scale radar data assimilation (e.g., Yussouf et al. 2013; Johnson et al. 2015). Only radar data are assimilated in this domain. The initial and lateral boundary conditions (ICs and LBCs) for this domain are interpolated from the mesoscale ensemble from the outer domain of Yussouf et al. (2013). Briefly, this mesoscale domain covers the continental United States (CONUS) with a horizontal grid spacing of 18 km and a  $267 \times 178 \times 50$  grid. An ensemble adjustment Kalman filter (EAKF) is used to assimilate the routinely available observations from Meteorological Assimilation Data Ingest System (MADIS) into the ensemble.

Version 3.5.1 of WRF-ARW is used in the present study. The model physics configuration includes the Kain–Fritsch cumulus parameterization (Kain and Fritsch 1993), the Mellor–Yamada–Janjić planetary boundary layer scheme (Mellor and Yamada 1982; Janjić 1990, 1994, 2002), the WRF single-moment 6-class microphysics



FIG. 7. Schematics of data assimilation experiment configuration. The radar data assimilation adopts the 5-min cycling, driven by the mesoscale analyses and forecasts. The 1-h forecast is then initialized by the analysis at 2200 UTC after 1-h assimilation.

scheme (WSM6; Hong and Lim 2006), the Noah land surface model (Chen and Dudhia 2001), the Dudhia shortwave radiation scheme (Dudhia 1989), and the Rapid Radiative Transfer Model (RRTM) longwave radiation scheme (Mlawer et al. 1997).

A two-way coupled GSI-based EnVar data assimilation method is applied for the storm-scale assimilation. Detailed procedures of the two-way coupled EnVar can be found in Wang et al. (2013). The procedure of radar data assimilation and the subsequent forecast is shown in Fig. 7. Starting at 2100 UTC 8 May 2003, the radar data are assimilated every 5 min for a 1-h period out to 2200 UTC. The 1-h forecasts are then launched from each of the 45-member analyses and the control analysis initialized at 2200 UTC. The control forecast is updated using the EnVar where ensemble perturbations are applied to estimate the background error covariance through the use of the extended control variable method. Cutoff distances of 12 km and 1.1 scale height (e.g., a difference in natural log of pressure of 1.1)<sup>1</sup> are used for the horizontal and vertical localizations, respectively. Similar cutoff distances are used in earlier studies assimilating radar observations using an ensemble-based method (e.g., Yussouf et al. 2013; Johnson et al. 2015). The ensemble perturbations are updated by the ensemble smoother version of the square root filter algorithm (EnSRF; Whitaker and Hamill 2002). The analysis ensemble is further recentered around the control analysis to obtain the final ensemble analysis. Covariance localization and inflation are used to account for system errors when updating the perturbations with EnSRF. Cutoff distances in the horizontal and vertical are the same as those used in the EnVar. The posterior ensemble spread is relaxed back to 90% of the prior ensemble spread using the relaxation to prior spread (RTPS; Whitaker and Hamill 2012) inflation method. To further account for the deficiency of the spread of the first-guess ensemble in the EnSRF, both the constant inflation and additive noise (Whitaker et al. 2008; Dowell and Wicker 2009; Dowell et al. 2011; Dawson et al. 2012; Jung et al.

2012; Wang et al. 2013; Yussouf et al. 2013) are applied wherever the observed radar reflectivity is greater than 25 dBZ. The constant inflation is applied with a coefficient of 1.04 to each ensemble perturbation in every assimilation cycle. The additive noise is applied to the horizontal winds, temperature, and dewpoint analyses at the first six cycles, and their standard deviations are  $0.5 \text{ m s}^{-1}$ , 0.5 K, and 0.5 K, respectively. The horizontal and vertical length scales of 3 km are used for the perturbation smoothing function. These parameters are chosen based on our sensitivity tests. The additive perturbations are added to the analyses in order to establish the flow-dependent structure of these perturbations during the 5-min model integration (Wang et al. 2013).

The NEXRAD level-2 data archived at the National Climatic Data Center (NCDC) are used to provide the radar reflectivity and radial velocity observations for this study. The Warning Decision Support System-Integrated Information (WDSSII; Lakshmanan et al. 2007) software is used to quality control these data, as in Johnson et al. (2015). Reflectivity values below or equal to 5 dBZ and missing values are considered to be “no precipitation” observations and are set to 0 dBZ (Yussouf et al. 2013). Spurious convection generated in the model can be suppressed through assimilating these no precipitation observations. The same observation errors in section 3,  $2 \text{ m s}^{-1}$  and 5 dBZ, for radial velocity and reflectivity are applied in the real data experiments.

Three EnVar experiments are conducted using the three different reflectivity DA methods described in section 2. All three experiments use the WSM6 (Hong and Lim 2006) microphysics scheme. Another experiment using the Thompson (Thompson et al. 2008) microphysics scheme is included in section 5f to further evaluate the impact of this microphysics scheme and the flexibility of method dBZ.

## 5. Results

### a. Observation-space diagnostics

We use observation-space diagnostics (Dowell et al. 2004; Dowell and Wicker 2009; Dowell et al. 2011; Yussouf et al. 2013), including the root-mean-square of

<sup>1</sup> For example, if the observation is at 900 hPa, the corresponding pressure above 900 hPa associated with a 1.1 scale height cutoff distance is 300 hPa [absolute value of  $\ln(300/900)$  is 1.1].

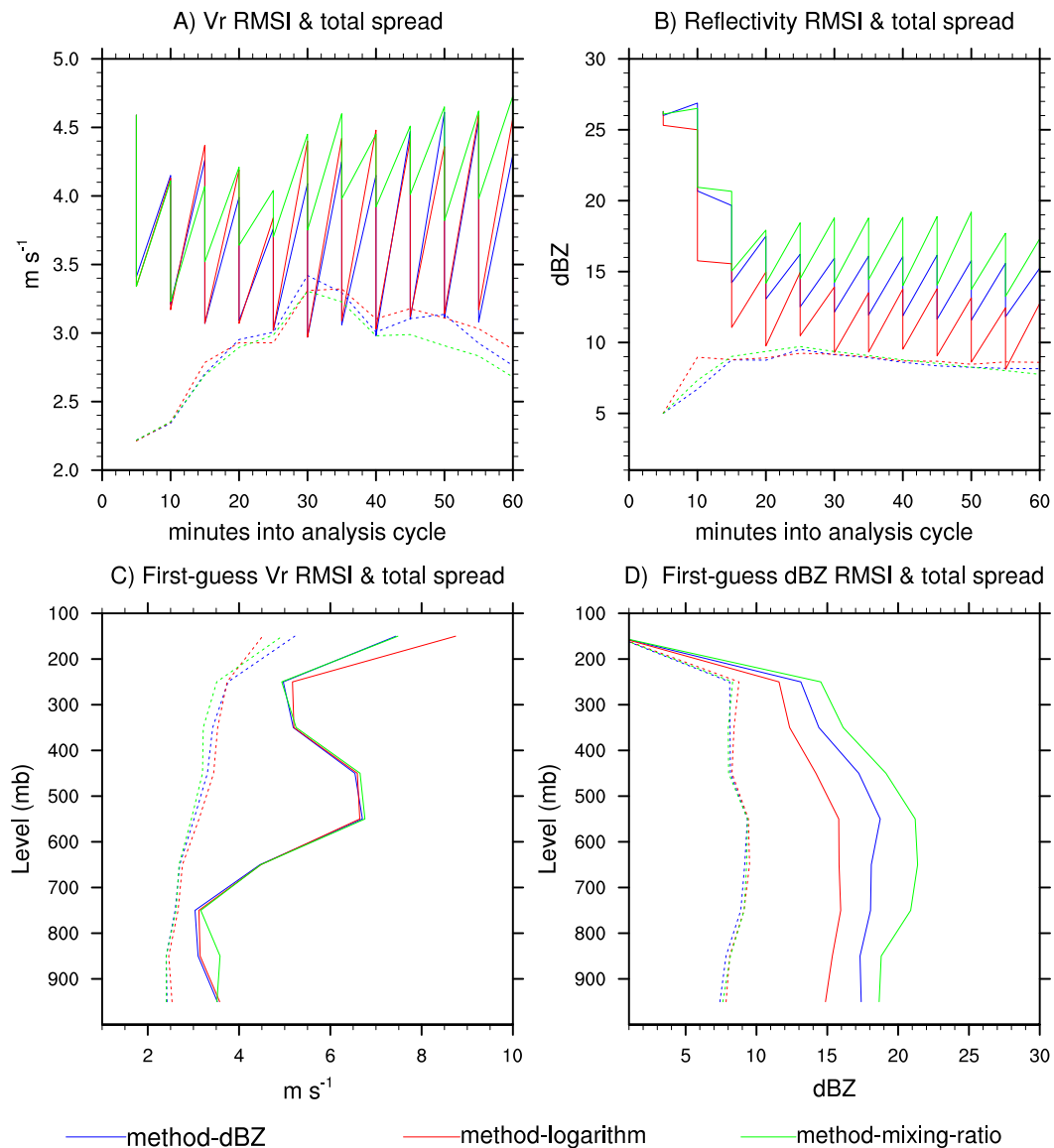


FIG. 8. Observation-space diagnostic statistics for the assimilated KTLX (a) radial velocity ( $V_r$ ) and (b) reflectivity (dBZ) for method dBZ (blue), method logarithm (red), and method mixing ratio (green) during the 1-h data assimilation period. The upper points and lower points of the sawtooth patterns in (a) and (b) correspond to the background and analysis statistics, respectively. The statistics include RMSI (solid) and total spread (dotted). (c),(d) As in (a),(b), but for first-guess vertical profiles.

innovation (RMSI) and total ensemble spread (total spread), to compare the influences of the three different reflectivity DA methods. The innovation here denotes the observation minus control forecast mapped to the observation location. The RMSI indicates the agreement between the forecasts and the observations, and the total spread provides information about the estimated forecast uncertainty (Yussouf et al. 2013). RMSI and total spread statistics are shown in Fig. 8 for the assimilated KTLX (Oklahoma City, Oklahoma) radial velocity and reflectivity observations during the 1-h

assimilation period for the 8 May 2003 supercell storm. The root-mean-square fit of the analysis to the observations is also included. Similar to Yussouf et al. (2013) and many other publications, only the locations with observed reflectivity greater than 10 dBZ are used to calculate the reflectivity statistics and the radial velocity statistics are calculated at the available observed locations without any threshold.

The RMSI of the posterior in method mixing ratio is larger than that from method logarithm and method dBZ, especially for the reflectivity observations. This

result is consistent with the results in sections 3a and 3c. The cost function gradient contributed by the reflectivity is much larger than that contributed by the radial velocity, which prevents efficient convergence of variational minimization. The deficiency of ill-conditioned convergence in method mixing ratio leads to less fit of the analysis to both the radial velocity and reflectivity observations. The lower fit of the posterior to the reflectivity is further exacerbated by the underestimate of the hydrometeor increments due to the errors associated with the linear approximation of the nonlinear operator in method mixing ratio. As a result, method mixing ratio produces less fit of the subsequent 5-min forecast to observations, especially for reflectivity.

The RMSI of the posterior for the radial velocity observations is similar in method logarithm and method dBZ, with a slight reduction of forecast fit to observations by method dBZ. However, for the reflectivity, the analysis and forecast in method logarithm fit the observations more than those in method dBZ. This result is consistent with the findings in sections 3b and 3c, where method logarithm is shown to overestimate the hydrometeor mixing ratio increments. As shown later in section 5e, the anomalously large hydrometeor increments by method logarithm, and therefore excessive fit to observations, can cause an overly strong and widespread cold pool in the forward flank. The overly strong cold pool cuts off the constructive interaction between the surface gust front and the updraft aloft associated with the midlevel mesocyclone, producing a simulated supercell much shorter lived than in reality. The underdispersive ensemble is a common problem in radar data assimilation at convective scales as shown in early studies (Dowell and Wicker 2009; Aksoy et al. 2009; Dowell et al. 2011; Snook et al. 2011, 2012; Yussouf et al. 2013). Consistent with these earlier works, the total ensemble spreads for both reflectivity and radial velocity are consistently smaller than the RMSI in all three experiments. The total spreads from all three reflectivity DA methods are similar in magnitude for reflectivity (Figs. 8b,d) after 15 min when the RMSI and spread statistics become stable. The spreads for radial velocity are also similar, except that method mixing ratio has a slightly smaller spread after 40 min (Fig. 8a), which mainly results from the smaller spread from 450 to 250 hPa (Fig. 8c).

The reflectivity analyses are further compared with the observations for the first 30 min for all three methods in Fig. 9 to reveal the spinup process during the data assimilation cycling. Here the spinup time is defined as the time needed for the system to reach a stable or saturated value. Consistent with Fig. 8b, all three methods reach their stable reflectivity value at about 2115 UTC, the third DA cycle. Note that at the second cycle

(2110 UTC), method logarithm shows a larger reflectivity value. However, this value is still smaller than its own saturated value. Therefore, no method stands out showing a faster spinup than the others, although their final saturated values are different. Specifically, method logarithm has the largest saturation value, method dBZ the second largest, and method mixing ratio has the smallest saturation value. As documented in section 3 and discussed above for Fig. 8 and section 5e, due to the methodological problems of method logarithm and method mixing ratio, the overly large and small reflectivity values in their analyses lead to the shorter-lived supercell than in reality.

#### *b. Ensemble probabilistic forecasts of low-level vorticity*

To evaluate the impact of the choice of the different reflectivity DA methods on the forecasts, we first verify the forecast against the observed tornado track. Since the 2-km grid spacing is not able to explicitly depict a tornado circulation, the presence of low-level rotation (vorticity) instead is used to represent a forecast tornado probability (Stensrud and Gao 2010; Dawson et al. 2012; Stensrud et al. 2013; Yussouf et al. 2013; Thompson et al. 2015). Forecast probabilities (Fig. 10) of the vorticity exceeding  $0.003 \text{ s}^{-1}$  at 150 m AGL are calculated at a neighborhood with a 6-km radius around each grid point. The neighborhood method is used to account for the small displacement errors across the ensemble members (Theis et al. 2005; Mittermaier 2007; Ebert 2009; Schwartz et al. 2010; Bouallegue et al. 2011; Johnson and Wang 2012; Yussouf et al. 2013). Results indicate that the low-level mesocyclone persisted during the entire 1-h forecast for method dBZ. The high probabilities (above 95%) of significant mesocyclones correlate well with the observed tornado track and extend beyond the track (Fig. 10a). In comparison, the forecast probabilities of vorticity in method logarithm are much lower after the 25-min lead time. In method mixing ratio, the swaths of high probability of vorticity also match well with the observed tornado track at early lead time, but become lower than 95% after the 45-min forecast time. As discussed in section 4a, the tornado-producing supercell lasts beyond the track of the tornado. Therefore, method dBZ maintains the supercell more consistently with the reality than method mixing ratio. The threshold of vorticity should only show up around the observed tornado track. However, spurious vorticity spreads over almost the entire western domain in method logarithm (Fig. 10b), which is more extensive than both method dBZ (Fig. 10a) and method mixing ratio (Fig. 10c). Method mixing ratio shows the least spurious vorticity. It is speculated that

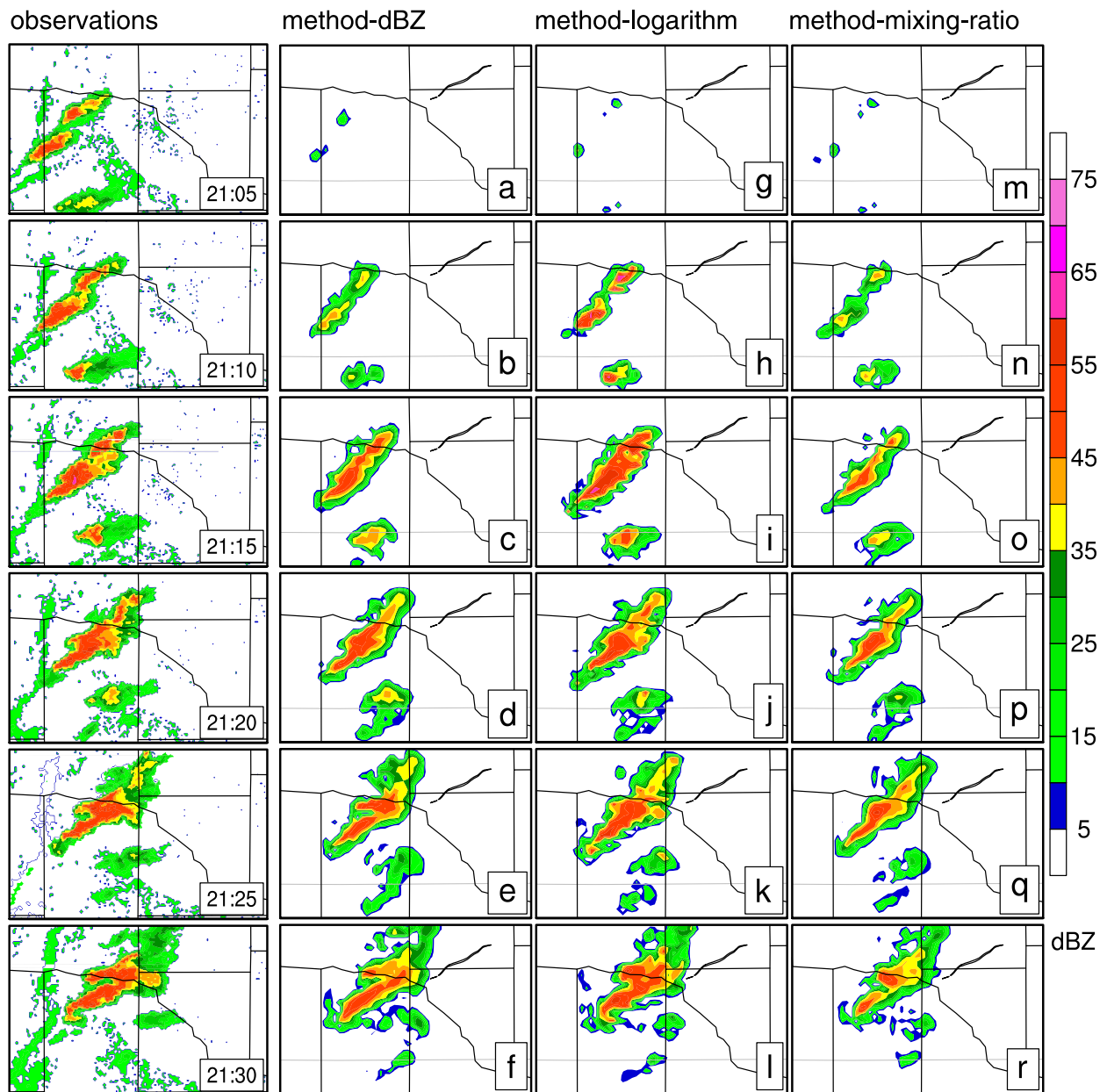


FIG. 9. The reflectivity (colors; dBZ) analysis at 1 km AGL for (a)–(f) method dBZ, (g)–(l) method logarithm, and (m)–(r) method mixing ratio from each data assimilation cycle during the first 30-min data assimilation period. The first column is the corresponding reflectivity observations from KTLX.

the extensively spurious vorticity in method logarithm is due to the strong and widespread cold pool, resulting from the overestimated hydrometeor increments, leading to the development of new spurious convective cells.

### c. Forecasts of midlevel updraft and vorticity

We further explore the influences of the three different reflectivity DA methods on the maintenance of the

supercell storm by examining the vertical velocity and vorticity in midlevels (4 km) during a 1-h forecast period (2200–2300 UTC) in Fig. 11.

All three experiments generate strong mesocyclones at the analysis time (2200 UTC), with the maximum updraft larger than  $30 \text{ m s}^{-1}$  and the maximum vorticity greater than  $1.2 \times 10^{-2} \text{ s}^{-1}$  at the 4-km height. The primary differences lie in their capability of maintaining the tornado-producing supercell during the entire 1-h

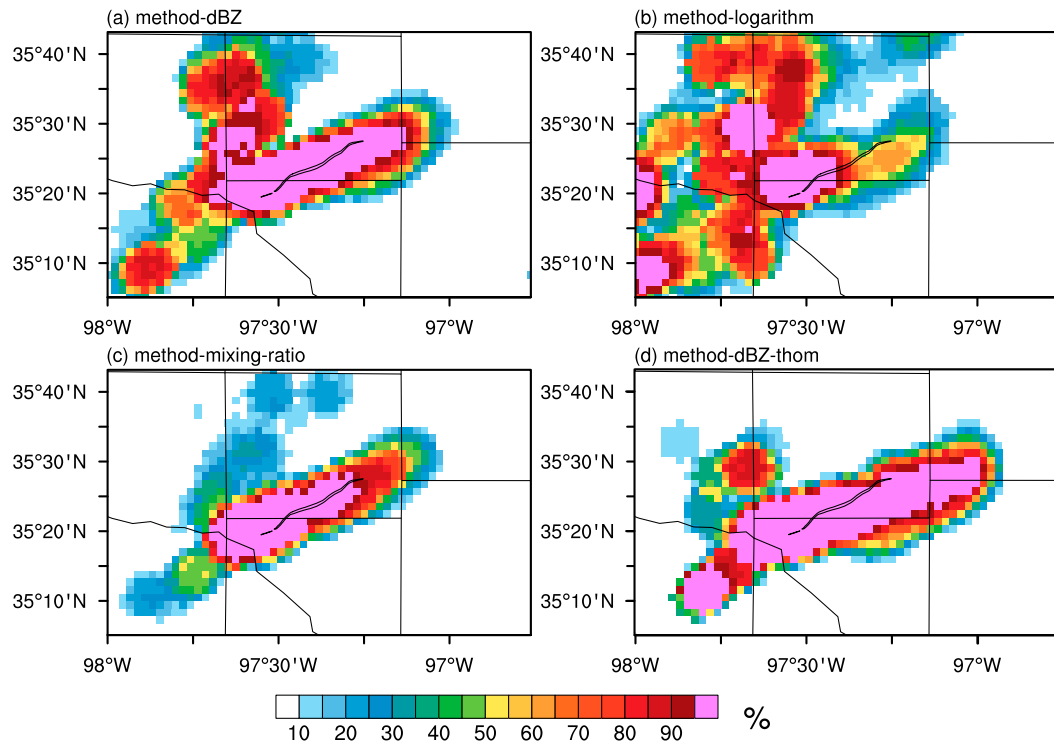


FIG. 10. Neighborhood ensemble probability (%) of vorticity exceeding thresholds of  $0.003 \text{ s}^{-1}$  at 150 m AGL during 1-h forecast period initialized from the analyses at 2200 UTC and ending at 2300 UTC for the (a) method dBZ, (b) method logarithm, (c) method mixing ratio, and (d) method dBZ thom. The overlaid squiggly line in each panel is the NWS-observed tornado damage track during 2210–2238 UTC.

forecast period. The storm in method dBZ persists during the entire 1-h forecast period. The maximum updraft in method dBZ is above  $30 \text{ m s}^{-1}$  for the entire forecast period. The maximum vorticity in method dBZ remains above  $1.0 \times 10^{-2} \text{ s}^{-1}$  until the 50-min forecast lead time. For method logarithm, the supercell storm starts to weaken at the 5–10-min forecast lead time. For example, the maximum updraft is reduced to below  $25 \text{ m s}^{-1}$  at about 5 min, and the maximum vorticity is reduced to below  $1.0 \times 10^{-2} \text{ s}^{-1}$  at about 10 min. The storm further dissipates at the 20-min forecast lead time. A spurious cell develops and follows the main supercell, with its updraft much weaker than method dBZ at the 1-h lead time. In method mixing ratio, while the storm is maintained much longer than method logarithm, it weakens more quickly than method dBZ. Specifically, the maximum updraft is reduced to below  $25 \text{ m s}^{-1}$  after the 10-min lead time and the maximum vorticity weakened to below  $1.0 \times 10^{-2} \text{ s}^{-1}$  at the 20-min lead time.

#### d. Reflectivity analyses and forecasts

Distributions of reflectivity analyses and forecasts are widely used to examine the overall location and

characteristics of the simulated storms (Xiao et al. 2007; Hu and Xue 2007; Lei et al. 2009; Gao and Stensrud 2012; Yussouf et al. 2013). For comparison purpose, we first present the KTLX-observed reflectivity at a 15-min interval during the forecast period from 2200 to 2300 UTC in Fig. 12. The reflectivity observations are interpolated to 1 km AGL using the NOAA's Weather and Climate Toolkit (WCT), distributed from NOAA/NCDC. Evolutions of reflectivity at 1 km AGL during the 1-h forecast period for the three experiments are plotted in Figs. 13a–o.

After a 1-h data assimilation, reflectivity analyses from all three experiments are able to reproduce the hook-echo structure at approximately the correct location (Figs. 13a,f,k), as well as the strong low-level rotation at 1 km AGL, which compare well with the KTLX radar observations (Fig. 12a). The reflectivity distribution in the forward-flank region in method dBZ (Fig. 13a) is closer to the size and shape of the observed forward-flank regions than in method logarithm (Fig. 13f) and method mixing ratio (Fig. 13k). Method logarithm produces reflectivity analyses with the largest area of reflectivity higher than 50 dBZ; while method mixing ratio has the smallest region. This is mainly due

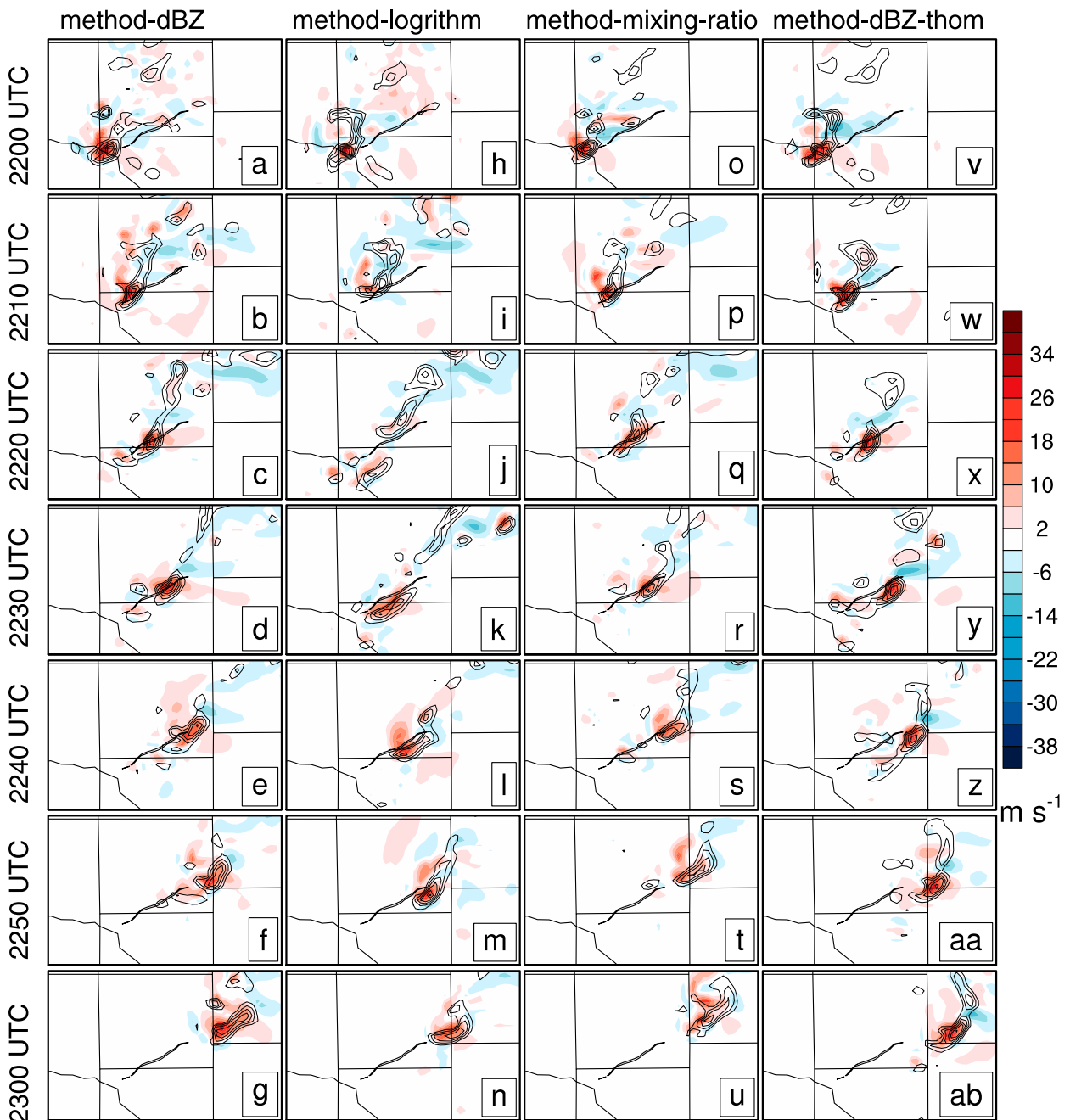


FIG. 11. The evolution of vertical velocity (colors;  $\text{m s}^{-1}$ ) and vorticity (contours starting from  $0.002 \text{ s}^{-1}$  in  $0.002 \text{ s}^{-1}$  interval) at 4 km AGL during a 1-h forecast period starting from the analyses at 2200 UTC and ending at 2300 UTC from the (a)–(g) method dBZ, (h)–(n) method logarithm, (o)–(u) method mixing ratio, and (v)–(z),(aa),(ab) method dBZ thom. The top row represents the last analysis time valid at 2200 UTC and the second to seventh rows are for the forecast time valid at 2210, 2220, 2230, 2240, 2250, and 2300 UTC, respectively. The overlaid line in each panel is as in Fig. 7.

to the different distributions of the analyzed hydrometeors in these experiments. The distributions of rainwater and snow mixing ratios (not shown) are similar with the graupel mixing ratio. As an example, the distributions of the graupel mixing ratio, from the analyses

at around the melting layer (5 km) are shown in Fig. 14. Consistent with the discussion in sections 3b and 3c that method logarithm produces overly large hydrometeor mixing ratio increments, method logarithm produces the greatest graupel mixing ratio in the analysis (Fig. 14b).



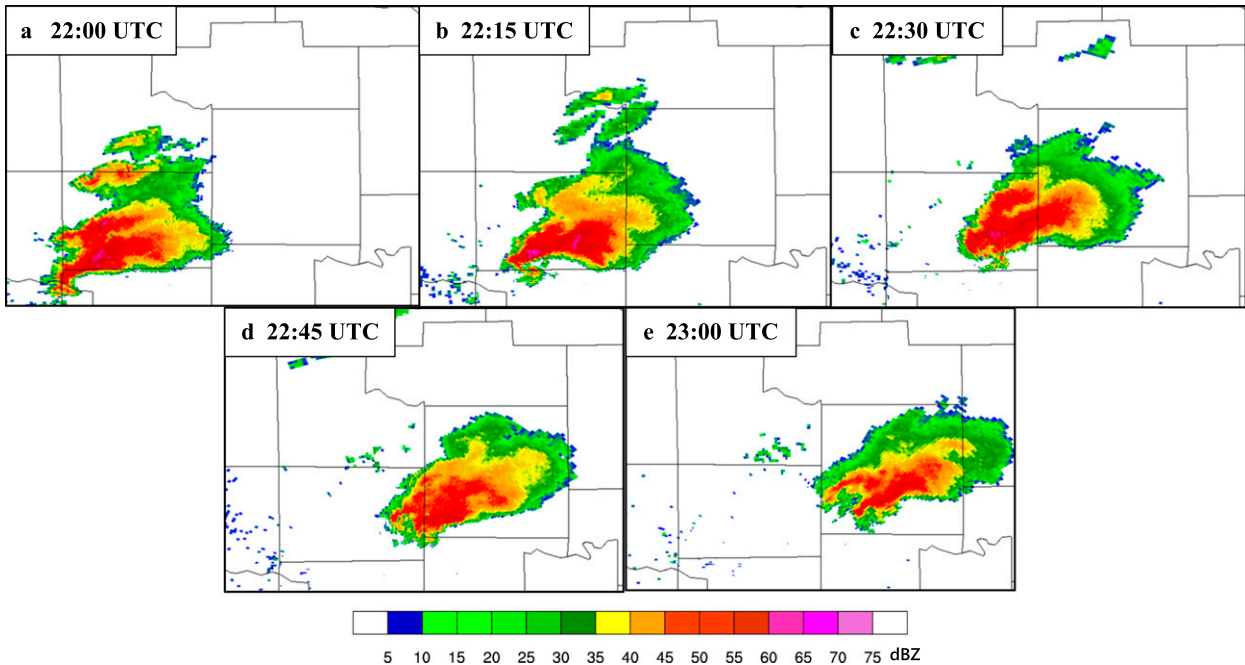


FIG. 12. The KTLX reflectivity observations at 1 km AGL at 15-min interval during the forecast period from 2200 to 2300 UTC.

Method mixing ratio generates the smallest amount of graupel mixing ratios (Fig. 14c), consistent with the discussion in sections 3a and 3c.

The predicted supercell reflectivity in method dBZ maintains its hook-echo structure until 45 min (Fig. 13d). Method logarithm and method mixing ratio lose their hook-echo structures by the 15-min (Fig. 13g) and 30-min (Fig. 13m) forecast lead times, respectively. Compared to method dBZ, the flanks gradually separate into several patches in method mixing ratio and method logarithm. In all three methods, the reflectivity distributions in the forecast are further to the northeast and narrower than the reflectivity observations after a 30-min lead time (Figs. 13c–e, h–j, m–o). As shown and discussed in section 5f, this result is largely due to the use of the WSM6 microphysics scheme. Although the predicted reflectivity does not compare well with the observed reflectivity, the maximum vorticity at 1 km AGL near the hook echo from method dBZ remains above  $6 \times 10^{-3} \text{ s}^{-1}$  during the entire 1-h forecast period (Figs. 13a–e). In method logarithm, the rotation at 1 km AGL starts to weaken at the 30-min lead time, with the maximum vorticity near the southwestern end of the storm less than  $5 \times 10^{-3} \text{ s}^{-1}$  (Fig. 13h). The rotation dissipates and is replaced by the rear spurious rotation. For method mixing ratio, the rotation at 1 km AGL keeps weakening over the 1-h forecast period with the maximum vorticity near the hook echo reduced to below  $5 \times 10^{-3} \text{ s}^{-1}$  by the 15-min lead time (Fig. 13l).

#### e. Impacts of the cold pool

This subsection explores further how the use of the different hydrometeor state variables affects the maintenance of the tornadic supercell during the forecast from a storm dynamics point of view. Surface temperatures from all three experiments are shown in Fig. 15. Method mixing ratio generates a smaller and weaker cold pool compared to method dBZ, while the cold pool from method logarithm is more extensive and stronger than from method dBZ. Generally, the cold pool is developed through the sublimation, evaporation, and melting of hydrometeors (Dowell et al. 2011). The cold pool in all three experiments corresponds well with the magnitude of the graupel mixing ratios aloft in their own simulations (Fig. 14). For method logarithm, the greatest graupel mixing ratios aloft lead to the greatest cooling and the largest precipitation loading compared to the other two experiments. The weakest cold pool from method mixing ratio results from the least graupel mixing ratio aloft with the least loading and cooling.

To investigate the impact of the simulated cold pool on the intensity of the supercell, vertical cross sections along the strong updraft cores are plotted in Fig. 16. It is hypothesized that the realistic cold pool in method dBZ leads to the constructive interaction between surface gust front and the updraft aloft associated with the midlevel mesocyclone, which maintains the supercell

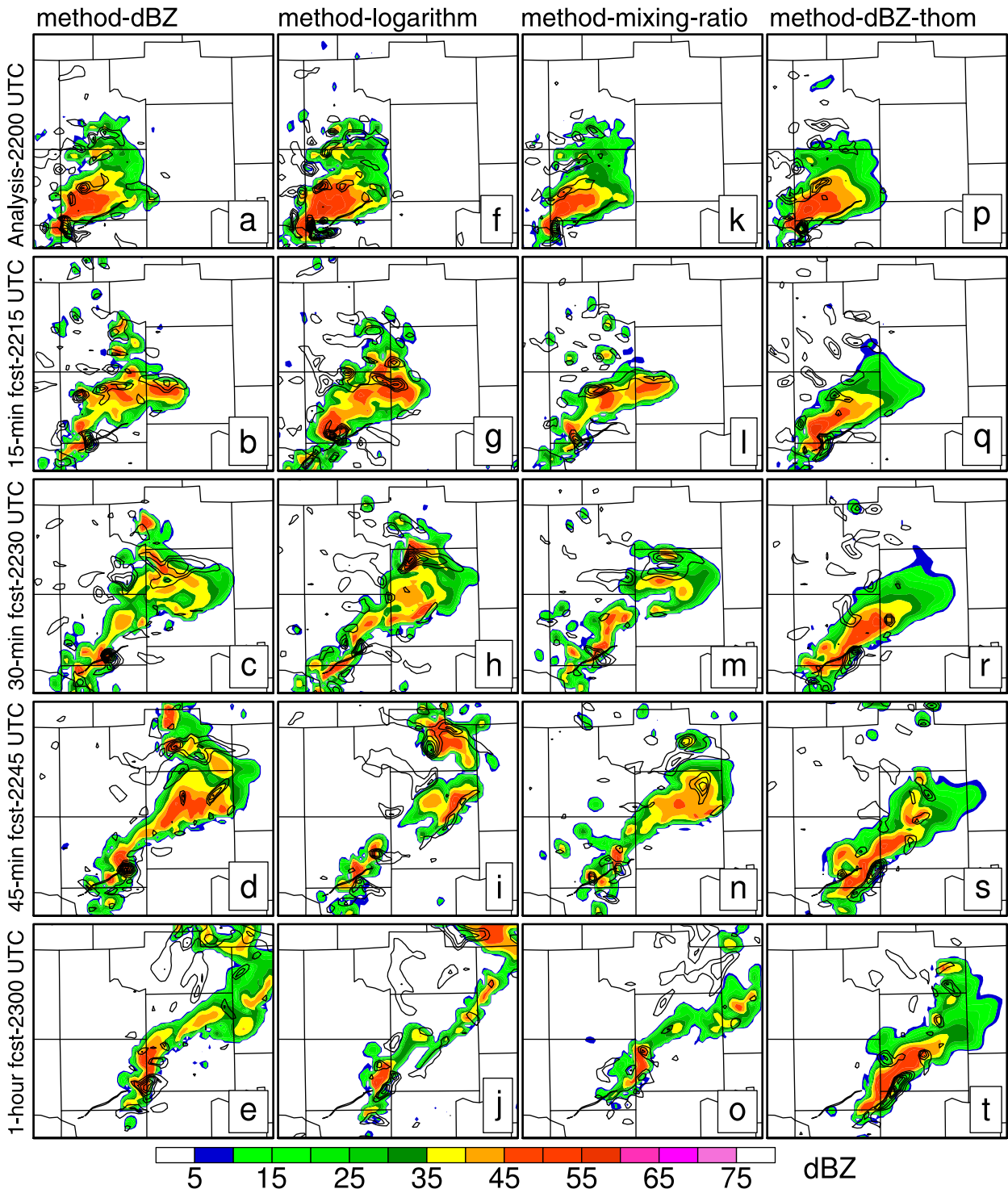


FIG. 13. The reflectivity (colors; dBZ) and vertical vorticity (contours from  $0.001$  to  $0.01 \text{ s}^{-1}$  at  $0.001 \text{ s}^{-1}$  interval) at  $1 \text{ km}$  AGL for (a)–(e) method dBZ, (f)–(j) method logarithm, (k)–(o) method mixing ratio, and (p)–(t) method dBZ thom. The analyses at  $2200 \text{ UTC}$  are shown in the first row;  $15\text{-min}$ ,  $30\text{-min}$ ,  $45\text{-min}$ , and  $1\text{-h}$  forecasts are shown in the second to fifth rows, respectively. The overlaid line in each panel is as in Fig. 7.

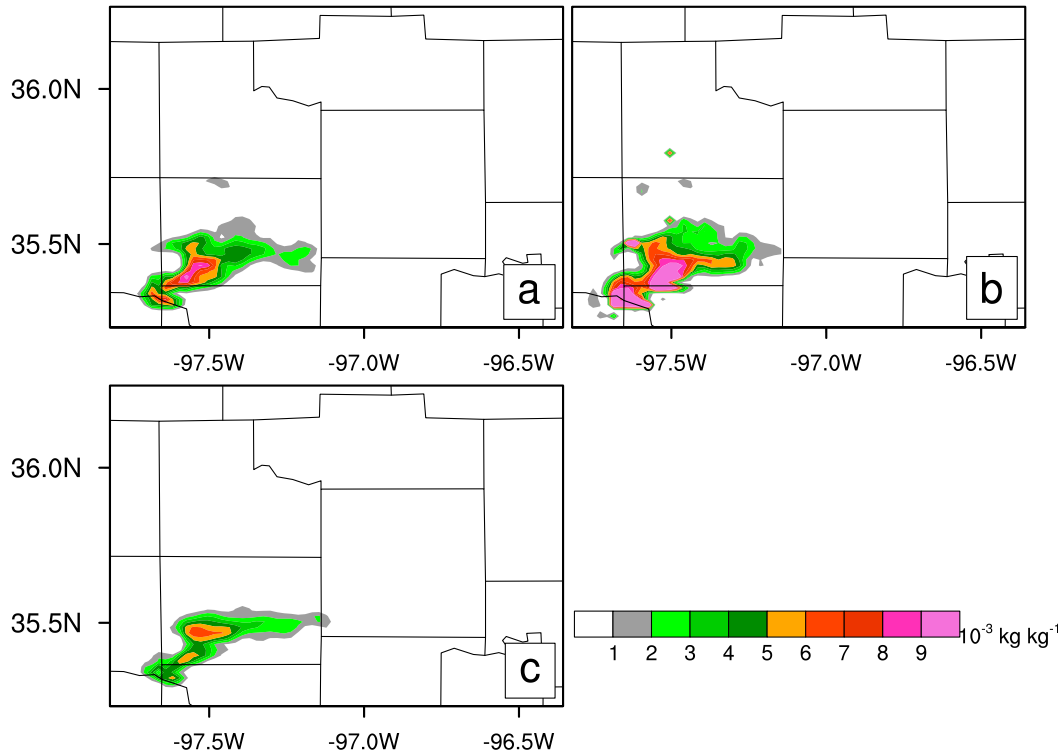


FIG. 14. The graupel mixing ratios (colors,  $10^{-3} \text{ kg kg}^{-1}$  interval) at 5 km AGL from the analyses of (a) method dBZ, (b) method logarithm, and (c) method mixing ratio.

during the entire 1-h period. Compared to method dBZ, method logarithm has a stronger cold pool, and its updraft is tilted by the gust front produced by the cold pool (Fig. 16b), which prevents the low-level circulation from constructively interacting with the midlevel mesocyclones (Engerer et al. 2008; Zheng and Chen 2014); while method mixing ratio has the weakest cold pool, which is not able to enhance the midlevel mesocyclone due to the small temperature gradients across the outflow boundaries (Markowski et al. 1998).

#### f. Application of method dBZ with the Thompson microphysics scheme

As discussed in Yussouf et al. (2013), the quality of the supercell forecast can be dependent on the adopted microphysics scheme. To reveal to what extent the less ideal reflectivity forecast shown in Fig. 13 is due to the use of the single-moment microphysics scheme, and also to demonstrate the easiness of using method dBZ without involving the TLA of a more complicated reflectivity operator associated with a more complex microphysics scheme, in this section the Thompson microphysics scheme is applied in place of WSM6 for method dBZ, denoted as method dBZ thom. The simulated reflectivity is computed within WRF-ARW and is

consistent with the Thompson scheme. Method dBZ is chosen for such experiments because method dBZ experiment with the WSM6 microphysics scheme produced the best analysis and forecast.

Forecast probabilities of vorticity exceeding  $0.003 \text{ s}^{-1}$  at 150 m AGL from method dBZ thom are shown in Fig. 10d. Method dBZ thom produces a similar low-level vorticity probabilistic forecast compared to method dBZ, but the vorticity swaths are displaced southeast beyond the 30-min lead time compared to method dBZ, which uses the WSM6 scheme. Evolution of the vertical velocity and vorticity at 4-km AGL from the control analysis and forecast of method dBZ thom are plotted in Figs. 11v–z,aa,ab. Similar to method dBZ, method dBZ thom maintains the storm during the entire 1-h forecast. Method dBZ thom is also able to reproduce the hook-echo structure (Fig. 13p) as well as the strong rotation at 1 km AGL, and compares well with the KLTX radar observation (Fig. 12a). Compared to method dBZ, the simulated reflectivity in the forward-flank region is much closer to the distributions and intensities of the observed forward-flank regions in method dBZ thom during the entire forecast period (Figs. 13q–t). This results from the greater flexibility in the hydrometeor size distributions in the Thompson microphysics scheme (Yussouf et al. 2013).

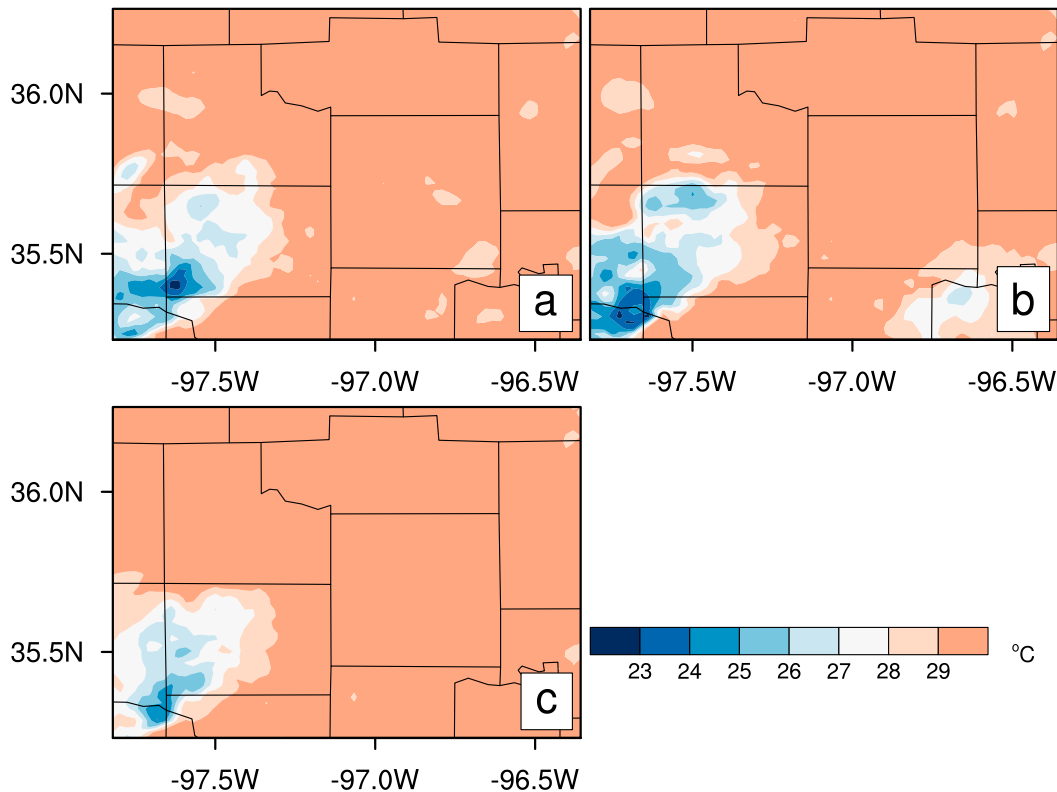


FIG. 15. Surface temperature (colors, 1°C interval) at the analyses time of 2200 UTC from (a) method dBZ, (b) method logarithm, and (c) method mixing ratio.

## 6. Conclusions and discussion

A series of studies have proven the usefulness of radar data assimilation for convective-scale NWP. However, when the reflectivity observation is directly assimilated, variational methods suffer from several problems associated with the use of the tangent linear approximation of the nonlinear reflectivity observation operator. This study reveals and demonstrates these problems in the GSI-based EnVar. When a hydrometeor mixing ratio is used as the state variable (method mixing ratio), due to the small hydrometeor mixing ratios, the gradient of the cost function contribution from reflectivity can be much larger than the gradient contribution from radial velocity. Such significant differences of the gradients among different variables can prevent efficient convergence. The logarithmic hydrometeor mixing ratio method (method logarithm) can help to alleviate this issue. However, the tangent linear of the reflectivity observation operator associated with the logarithmic hydrometeor mixing ratio produces spuriously large hydrometeor increments. In addition, the linear approximation of the operators contributes to the overestimate and underestimate of hydrometeor increments in method logarithm and method mixing ratio,

respectively. In this study, a new method to directly assimilate the reflectivity is introduced in the EnVar framework to solve the above problems. The new method augments the state vector to include reflectivity as a state variable (method dBZ). Therefore, neither the tangent linear nor the adjoint of the reflectivity operator are needed in the variational framework. This new method is derived and implemented in the GSI-based EnVar data assimilation system. It is revealed that method dBZ solves the aforementioned problems in method mixing ratio and method logarithm.

The new method (method dBZ), together with method mixing ratio and method logarithm, is further examined using the analysis and prediction of the 8 May 2003 Oklahoma City tornadic supercell storm. The experiments are conducted with WRF-ARW at a 2-km convection allowing resolution with the WSM6 microphysics scheme. A 45-member ensemble is used in the GSI-based EnVar system. The assimilation starts at 2100 UTC 8 May 2003. The initial ensemble and control background are downscaled from ensemble analyses and their mean of the mesoscale ensembles, respectively. Radar observations including both reflectivity and radial velocity are assimilated every 5 min for a total period of 1 h. The best results are obtained with method

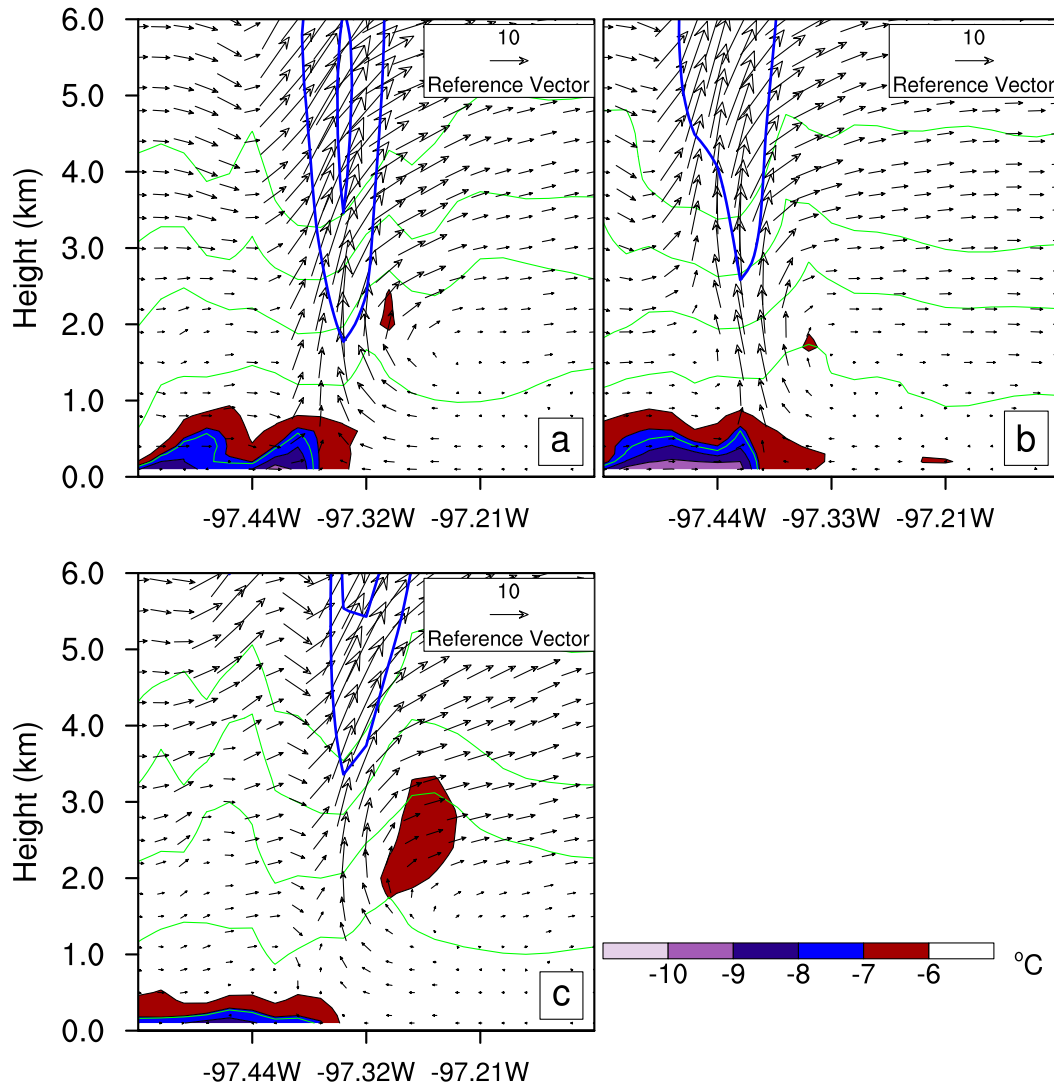


FIG. 16. Vertical cross sections along the strong updraft cores of winds (vector;  $\text{m s}^{-1}$ ) and near-surface cold pool (shaded;  $^{\circ}\text{C}$ ) valid at 30-min forecast from the analyses of (a) method dBZ, (b) method logarithm, and (c) method mixing ratio. Midlevel updraft (contours from 15 to  $25 \text{ m s}^{-1}$  at  $10 \text{ m s}^{-1}$  interval) and potential temperature (contours from 300 to 320 K at 4-K interval) are indicated by blue lines and green lines, respectively.

dBZ. The probabilistic forecast of strong low-level vorticity derived from the ensemble from method dBZ follows with the observed tornado track for both the location and longevity of the storm much better than method mixing ratio and method logarithm. The predicted tornadic supercell initialized from the analysis using method dBZ maintains the strong midlevel updraft and vorticity during the entire 1-h forecast period. In comparison, the supercell storm dissipates after around 5–10 and 10–20 min, respectively, for method logarithm and method mixing ratio. Detailed diagnostics reveal that method dBZ more correctly analyzes hydrometeor fields such as graupel mixing ratio. Such an analysis of the hydrometeor fields leads to constructive interaction of the

cold pool, the surface gust front, and the updraft associated with the midlevel mesocyclone. In contrast, such constructive interaction is missing due to overly extensive and overly constrained hydrometeor analyses in method logarithm and method mixing ratio, respectively.

The method dBZ experiment was repeated with Thompson microphysics scheme. Similar to the WSM6 scheme, the Thompson scheme maintains the tornado-producing supercell for the entire 1-h forecast period. The reflectivity distribution in the forward flank matches the reflectivity observations better with the Thompson scheme than with the WSM6 scheme. However, the Thompson scheme has more track forecast errors with surface vorticity swaths displaced more southeastward

than both the reality and the WSM6 scheme after the 30-min forecast.

As pointed out in section 2a, in this study the GSI-based EnVar system is employed where the background error covariance is fully contributed by the ensemble. As discussed in Michel et al. (2011) and Johnson et al. (2015), for convective-scale DA, the static covariance, if constructed properly, may help produce reflectivity more efficiently at locations where the ensemble background is “clear air” but where the observation is precipitating. Michel et al. (2011) and Descombes et al. (2015) provided a promising approach for including hydrometeor mixing ratios in the static background error covariance where covariance between hydrometeor mixing ratio and other variables were considered. A similar approach can be applied when the reflectivity is included in the state variables. The impact of including the static covariance in method dBZ is left for future study.

In a traditional incremental variational framework such as GSI, outer loops are used to reduce the errors associated with the tangent linear of the nonlinear observation operator (Rosmond and Xu 2006). In the new method, the reflectivity is included as part of the state variable by precalculating the reflectivity from the first guesses outside the minimization. In other words, inside the minimization there are no such steps as applying nonlinear reflectivity operator and its associated TLA. Therefore, the new method avoids the need of using outer loops in both 3DEnVar and 4DEnVar.

The newly proposed method is still based on the cost function derived from the Gaussian assumption. For convective-scale data assimilation where nonlinearity and non-Gaussianity are present in the error distributions for both the prior and the observation, methods need to be developed to rigorously take into account the deviation from the Gaussian distribution assumption. Methods that consider higher-order moments (e.g., Hodyss 2012) or relax the parametric distribution assumption (e.g., Poterjoy 2016) become increasingly computationally feasible. Initial studies of examining such methods for radar DA have only just begun (Poterjoy et al. 2016). Comparison of the proposed method with such non-Gaussian filters will be left for future studies.

While encouraging results are obtained by examining the newly proposed method with one supercell case, more robust conclusions should be drawn with systematic experiments using more cases. The comparison of the proposed EnVar without TLA of the nonlinear operator with other ensemble-based data assimilation methods such as the pure EnKF is also left for future studies.

*Acknowledgments.* This study is supported by the NOAA/Office of Oceanic and Atmospheric Research

under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce and NSF Awards AGS-1359703, AGS-1046081, and NOAA Award NA15OAR4590193. The computing for this project was performed at the OU Supercomputing Center for Education and Research (OSCER) at the University of Oklahoma (OU). This work also used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation Grant ACI-1053575, and the Yellowstone machine (ark:/85065/d7wd3xhc) at NCAR’s Computing Computational and Information Systems Laboratory, sponsored by the National Science Foundation. Ting Lei is acknowledged for initial efforts of the study. We thank Nusrat Yussouf for providing the mesoscale ensembles, Jacob Carley for discussion during early stages of the work, and Aaron Johnson for helping to proofread the manuscript. Part of the conclusion of the work was included in the abstract of the authors’ 2016 AMS annual conference presentation, and shared internally through progress reports to the funding agency.

## REFERENCES

- Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicaser comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, doi:10.1175/2008MWR2691.1.
- Albers, S. C., J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analyses of clouds, precipitation, and temperature. *Wea. Forecasting*, **11**, 273–287, doi:10.1175/1520-0434(1996)011<0273:TLAAPS>2.0.CO;2.
- Anderson, J. L., and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.*, **24**, 1452–1463, doi:10.1175/JTECH2049.1.
- Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, doi:10.1175/MWR-D-15-0242.1.
- Bouallegue, Z. B., S. Theis, and C. Gebhardt, 2011: From verification results to probabilistic products: Spatial techniques applied to ensemble forecasting. *Proc. Fifth Int. Verification Methods Workshop*, Melbourne, Australia, CAWCR, CAWCR Tech. Rep. 046, 4–5.
- Brewster, K., 1996: Application of a Bratseth analysis scheme including Doppler radar data. Preprints, *15th Conf. on Weather Analysis and Forecasting*, Norfolk, VA, Amer. Meteor. Soc., 92–95.
- Carley, J. R., 2012: Hybrid ensemble-3DVar radar data assimilation for the short-term prediction of convective storms. Ph.D. dissertation, Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, 205 pp.
- , E. Rogers, S. Liu, B. Ferrier, E. Aligo, M. Pyle, X. Zhang, and G. DiMego, 2015: A status update for the NAMRR, an hourly-updated version of NAM forecast system. *19th Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, Phoenix, AZ, Amer. Meteor. Soc., 4.4. [Available online at <https://ams.confex.com/ams/95Annual/webprogram/Paper263759.html>.]

- Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.*, **133**, 3081–3094, doi:10.1175/MWR3021.1.
- Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 471 pp.
- Dawson, D. T., II, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140**, 696–716, doi:10.1175/MWR-D-11-00008.1.
- Descombes, G., T. Auligné, F. Vandenberghe, D. M. Barker, and J. Barré, 2015: Generalized background error covariance matrix model (GEN\_BE v2.0). *Geosci. Model Dev.*, **8**, 669–696, doi:10.5194/gmd-8-669-2015.
- Doviak, R. J., and D. S. Zrnić, 1993: *Doppler Radar and Weather Observations*. Dover, 562 pp.
- Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.*, **26**, 911–927, doi:10.1175/2008JTECHA1156.1.
- , F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005, doi:10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2.
- , L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, doi:10.1175/2010MWR3438.1.
- Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.
- Ebert, E., 2009: Neighborhood verification: A strategy for re-warding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, doi:10.1175/2009WAF2222251.1.
- Engerer, N. A., D. J. Stensrud, and M. C. Coniglio, 2008: Surface characteristics of observed cold pools. *Mon. Wea. Rev.*, **136**, 4839–4849, doi:10.1175/2008MWR2528.1.
- Gao, J., and M. Xue, 2008: An efficient dual-resolution approach for ensemble data assimilation and tests with simulated Doppler radar data. *Mon. Wea. Rev.*, **136**, 945–963, doi:10.1175/2007MWR2120.1.
- , and D. J. Stensrud, 2012: Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification. *J. Atmos. Sci.*, **69**, 1054–1065, doi:10.1175/JAS-D-11-0162.1.
- Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation and evolution sensitivity in simulated deep convective storms: Comparisons between liquid-only and simple ice and liquid phase microphysics. *Mon. Wea. Rev.*, **132**, 1897–1916, doi:10.1175/1520-0493(2004)132<1897:PAESIS>2.0.CO;2.
- Hodys, D., 2012: Accounting for skewness in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 2346–2358, doi:10.1175/MWR-D-11-00198.1.
- Hong, S. Y., and J. O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42** (2), 129–151.
- 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, doi:10.1175/MWR3313.1.
- , —, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698, doi:10.1175/MWR3092.1.
- Janjić, Z. I., 1990: The step-mountain coordinate: Physical package. *Mon. Wea. Rev.*, **118**, 1429–1443, doi:10.1175/1520-0493(1990)118<1429:TSMCPP>2.0.CO;2.
- , 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.
- , 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp.
- Johnson, A., and X. Wang, 2012: Verification and calibration of neighborhood and object-based probabilistic precipitation forecasts from a multimodel convection-allowing ensemble. *Mon. Wea. Rev.*, **140**, 3054–3077, doi:10.1175/MWR-D-11-00356.1.
- , —, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143**, 3087–3108, doi:10.1175/MWR-D-14-00345.1.
- Jung, Y., M. Xue, G. Zhang, and J. M. Straka, 2008: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Mon. Wea. Rev.*, **136**, 2246–2260, doi:10.1175/2007MWR2288.1.
- , —, and M. Tong, 2012: Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric data. *Mon. Wea. Rev.*, **140**, 1457–1475, doi:10.1175/MWR-D-11-00032.1.
- Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models*, Meteor. Monogr., No. 46, Amer. Meteor. Soc., 165–170.
- Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DVar and hybrid variants. *Mon. Wea. Rev.*, **143**, 452–470, doi:10.1175/MWR-D-13-00350.1.
- Kutty, G., and X. Wang, 2015: A comparison of the impacts of radiosonde and AMSU radiance observations in GSI-based 3DVar and 3DVar data assimilation systems for NCEP GFS. *Adv. Meteor.*, **2015**, 280546, doi:10.1155/2015/280546.
- Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The warning decision support system–integrated information. *Wea. Forecasting*, **22**, 596–612, doi:10.1175/WAF1009.1.
- Lei, T., M. Xue, and T. Yu, 2009: Multi-scale analysis and prediction of the 8 May 2003 Oklahoma City tornadic supercell storm assimilating radar and surface network data using EnKF. *13th Conf. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, Phoenix, AZ, Amer. Meteor. Soc., 6.4. [Available online at [https://ams.confex.com/ams/89annual/techprogram/paper\\_150404.htm](https://ams.confex.com/ams/89annual/techprogram/paper_150404.htm).]
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl.*

- Meteor.*, **22**, 1065–1092, doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.
- Liu, C., Q. Xiao, and B. Wang, 2008: An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. *Mon. Wea. Rev.*, **136**, 3363–3373, doi:10.1175/2008MWR2312.1.
- , —, and —, 2009: An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing System Simulation Experiments with the Advanced Research WRF (ARW). *Mon. Wea. Rev.*, **137**, 1687–1704, doi:10.1175/2008MWR2699.1.
- Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194, doi:10.1002/qj.49711247414.
- , 2013: Recommended nomenclature for EnVar data assimilation methods. Research Activities in Atmospheric and Oceanic Modelling, WGNE, 2 pp. [Available online at [http://www.wcrp-climate.org/WGNE/BlueBook/2013/individual-articles/01\\_Lorenc\\_Andrew\\_EnVar\\_nomenclature.pdf](http://www.wcrp-climate.org/WGNE/BlueBook/2013/individual-articles/01_Lorenc_Andrew_EnVar_nomenclature.pdf).]
- Lu, H., and Q. Xu, 2009: Trade-offs between measurements accuracy and resolutions in configuring phased-array radar velocity scans for ensemble-based storm-scale data assimilation. *J. Appl. Meteor. Climatol.*, **48**, 1230–1244, doi:10.1175/2008JAMC2009.1.
- Lu, X., X. Wang, Y. Li, M. Tong, and X. Ma, 2017: GSI-based ensemble-variational hybrid data assimilation for HWRF for hurricane initialization and prediction: Impact of various error covariances for airborne radar observation assimilation. *Quart. J. Roy. Meteor. Soc.*, **143**, 223–239, doi:10.1002/qj.2914.
- Markowski, P. M., E. N. Rasmussen, and J. M. Straka, 1998: The occurrence of tornadoes in supercells interacting with boundaries during VORTEX-95. *Wea. Forecasting*, **13**, 852–859, doi:10.1175/1520-0434(1998)013<0852:TOOTIS>2.0.CO;2.
- Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875, doi:10.1029/RG020i004p00851.
- Michel, Y., T. Auligné, and T. Montmerle, 2011: Heterogeneous convective-scale background error covariances with the inclusion of hydrometeor variables. *Mon. Wea. Rev.*, **139**, 2994–3015, doi:10.1175/2011MWR3632.1.
- Mittermaier, M., 2007: Improving short-range high-resolution model precipitation forecast skill using time-lagged ensembles. *Quart. J. Roy. Meteor. Soc.*, **133**, 1487–1500, doi:10.1002/qj.135.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, doi:10.1029/97JD00237.
- Poterjoy, J., 2016: A localized particle filter for high-dimensional nonlinear systems. *Mon. Wea. Rev.*, **144**, 59–76, doi:10.1175/MWR-D-15-0163.1.
- , J. Anderson, and R. Sobash, 2016: Convective-scale data assimilation in the Weather Research and Forecasting model using a nonlinear ensemble filter. *EGU General Assembly Conference Abstracts*, Vol. 18, Vienna, Austria, EGU, 13429.
- Rosmond, T., and L. Xu, 2006: Development of NAVDAS-AR: Non-linear formulation and outer loop tests. *Tellus*, **58A**, 45–58, doi:10.1111/j.1600-0870.2006.00148.x.
- Schwartz, C. S., and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280, doi:10.1175/2009WAF2222267.1.
- Snook, N., M. Xue, and Y. Jung, 2011: Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Wea. Rev.*, **139**, 3446–3468, doi:10.1175/MWR-D-10-05053.1.
- , —, and J. Jung, 2012: Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data. *Mon. Wea. Rev.*, **140**, 2126–2146, doi:10.1175/MWR-D-11-00117.1.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, doi:10.1175//2555.1.
- Souto, M. J., C. F. Balseiro, V. Pérez-Muñuzuri, M. Xue, and K. Brewster, 2003: Impact of cloud analysis on numerical weather prediction in the Galician region of Spain. *J. Appl. Meteor.*, **42**, 129–140, doi:10.1175/1520-0450(2003)042<0129:IOCAON>2.0.CO;2.
- Stensrud, D. J., and J. Gao, 2010: Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short range forecasts. *Mon. Wea. Rev.*, **138**, 1250–1272, doi:10.1175/2009MWR3027.1.
- , and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16, doi:10.1016/j.atmosres.2012.04.004.
- Sun, J., 2005: Convective-scale assimilation of radar data: Progress and challenges. *Quart. J. Roy. Meteor. Soc.*, **131**, 3439–3463, doi:10.1256/qj.05.149.
- , and N. A. Crook, 1997: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: Model development and simulated data experiments. *J. Atmos. Sci.*, **54**, 1642–1661, doi:10.1175/1520-0469(1997)054<1642:DAMRFD>2.0.CO;2.
- , and —, 1998: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part II: Retrieval experiments of an observed Florida convective storm. *J. Atmos. Sci.*, **55**, 835–852, doi:10.1175/1520-0469(1998)055<0835:DAMRFD>2.0.CO;2.
- , and H. Wang, 2013: Radar data assimilation with WRF 4D-Var. Part II: Comparison with 3D-Var for a squall line over the U.S. Great Plains. *Mon. Wea. Rev.*, **141**, 2245–2264, doi:10.1175/MWR-D-12-00169.1.
- Theis, S. E., A. Hense, and U. Damrath, 2005: Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteor. Appl.*, **12**, 257–268, doi:10.1017/S1350482705001763.
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. R. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, doi:10.1175/2008MWR2387.1.
- Thompson, T. E., L. J. Wicker, X. Wang, and C. Potvin, 2015: A comparison between the Local Ensemble Transform Kalman Filter and the Ensemble Square Root Filter for the assimilation of radar data in convective-scale models. *Quart. J. Roy. Meteor. Soc.*, **141**, 1163–1176, doi:10.1002/qj.2423.
- 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, doi:10.1175/MWR2898.1.
- Wang, H., J. Sun, S. Fan, and X.-Y. Huang, 2013a: Indirect assimilation of radar reflectivity with WRF 3D-Var and its impact on prediction of four summertime convective events. *J. Appl. Meteor. Climatol.*, **52**, 889–902, doi:10.1175/JAMC-D-12-0120.1.
- , —, X. Zhang, X. Huang, and T. Auligne, 2013b: Radar data assimilation with WRF 4D-Var. Part I: System development



- and preliminary testing. *Mon. Wea. Rev.*, **141**, 2224–2244, doi:10.1175/MWR-D-12-00168.1.
- Wang, X., 2010: Incorporating ensemble covariance in the Grid-point Statistical Interpolation (GSI) variational minimization: A mathematical framework. *Mon. Wea. Rev.*, **138**, 2990–2995, doi:10.1175/2010MWR3245.1.
- , and T. Lei, 2014: GSI-based four dimensional ensemble-variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System. *Mon. Wea. Rev.*, **142**, 3303–3325, doi:10.1175/MWR-D-13-00303.1.
- , D. Parrish, D. Kleist, and J. S. Whitaker, 2013: GSI 3DVar-based ensemble-variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, doi:10.1175/MWR-D-12-00141.1.
- Weygandt, S. S., and S. G. Benjamin, 2007: Radar reflectivity-based initialization of precipitation systems using a diabatic digital filter within the Rapid Update Cycle. Preprints, *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., 1B.7. [Available online at [https://ams.confex.com/ams/22WAF18NWP/techprogram/paper\\_124540.htm](https://ams.confex.com/ams/22WAF18NWP/techprogram/paper_124540.htm).]
- , —, T. G. Smirnova, and J. M. Brown, 2008: Assimilation of radar reflectivity data using a diabatic digital filter within the Rapid Update Cycle. *12th Conf. on IOAS - AOLS*, New Orleans, LA, Amer. Meteor. Soc., 8.4. [Available online at <https://ams.confex.com/ams/pdfpapers/134081.pdf>.]
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.
- , and —, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, doi:10.1175/MWR-D-11-00276.1.
- , —, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482, doi:10.1175/2007MWR2018.1.
- Xiao, Q., Y. Kuo, J. Sun, W. Lee, E. Lim, Y.-R. Guo, and D. M. Barker, 2005: Assimilation of Doppler radar observations with a regional 3DVAR system: Impact of Doppler velocities on forecasts of a heavy rainfall case. *J. Appl. Meteor.*, **44**, 768–788, doi:10.1175/JAM2248.1.
- , —, —, —, D. M. Barker, and L. Eunha, 2007: An approach of radar reflectivity data assimilation and its assessment with the inland QPF of Typhoon Rusa (2002) at landfall. *J. Appl. Meteor. Climatol.*, **46**, 14–22, doi:10.1175/JAM2439.1.
- Xu, Q., H. Lu, S. Gao, M. Xue, and M. Tong, 2008: Time-expanded sampling for ensemble Kalman filter: Assimilation experiments with simulated radar observations. *Mon. Wea. Rev.*, **136**, 2651–2667, doi:10.1175/2007MWR2185.1.
- Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170, doi:10.1007/s00703-001-0595-6.
- , M. Tong, and G. Zhang, 2009: Simultaneous state estimation and attenuation correction for thunderstorms with radar data using an ensemble Kalman filter: Tests with simulated data. *Quart. J. Roy. Meteor. Soc.*, **135**, 1409–1423, doi:10.1002/qj.453.
- Yussouf, N., and D. J. Stensrud, 2010: Impact of phased-array radar observations over a short assimilation period: Observing system simulation experiments using an ensemble Kalman filter. *Mon. Wea. Rev.*, **138**, 517–538, doi:10.1175/2009MWR2925.1.
- , E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics schemes. *Mon. Wea. Rev.*, **141**, 3388–3412, doi:10.1175/MWR-D-12-00237.1.
- Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125, doi:10.1175/2009MWR2645.1.
- Zhang, J., 1999: Moisture and diabatic initialization based on radar and satellite observation. Ph.D. thesis, University of Oklahoma, 194 pp.
- , F. Carr, and K. Brewster, 1998: ADAS cloud analysis. Preprints, *12th Conf. on Numerical Weather Prediction*, Phoenix, AZ, Amer. Meteor. Soc., 185–188.
- Zheng, K., and B. Chen, 2014: Sensitivities of tornadogenesis to drop size distribution in a simulated subtropical supercell over eastern China. *Adv. Atmos. Sci.*, **31**, 657–668, doi:10.1007/s00376-013-3143-7.
- Zupanski, M., 2005: Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Wea. Rev.*, **133**, 1710–1726, doi:10.1175/MWR2946.1.