

# Direct Variational Assimilation of Radar Reflectivity and Radial Velocity Data: Issues with Nonlinear Reflectivity Operator and Solutions

CHENGSI LIU

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

MING XUE

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

RONG KONG

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

(Manuscript received 16 May 2019, in final form 12 December 2019)

## ABSTRACT

Radar reflectivity ( $Z$ ) data are either directly assimilated using 3DVar, 4DVar, or ensemble Kalman filter, or indirectly assimilated using, for example, cloud analysis that preretrieves hydrometeors from  $Z$ . When directly assimilating radar data variationally, issues related to the highly nonlinear  $Z$  operator arise that can cause nonconvergence and bad analyses. To alleviate the issues, treatments are proposed in this study and their performances are examined via observing system simulation experiments. They include the following: 1) When using hydrometeor mixing ratios as control variables (CVq), small background  $Z$  can cause extremely large cost function gradient. Lower limits are imposed on the mixing ratios (qLim treatment) or the equivalent reflectivity (ZeLim treatment) in  $Z$  observation operator. ZeLim is found to work better than qLim in terms of analysis accuracy and convergence speed. 2) With CVq, the assimilation of radial velocity ( $V_r$ ) is ineffective when assimilated together with  $Z$  data due to the much smaller cost function gradient associated with  $V_r$ . A procedure (VrPass) that assimilates  $V_r$  data in a separate pass is found very helpful. 3) Using logarithmic hydrometeor mixing ratios as control variables (CVlogq) can also avoid extremely large cost function gradient, and has much faster convergence. However, spurious analysis increments can be created when transforming the analysis increments back to mixing ratios. A background smoothing and a lower limit are applied to the background mixing ratios, and are shown to be effective. Using CVlogq with associated treatments produces better reflectivity analysis that is much closer to the observation without resorting to multiple analysis passes, and the cost function minimization also converges faster. CVlogq is therefore recommended for variational radar data assimilation.

## 1. Introduction

The numerical weather prediction (NWP) model is becoming increasingly important tool for forecasting convective-scale weather (Lilly 1990; Stensrud et al. 2013; Sun et al. 2014). To properly initialize convective storms and other precipitation systems within an NWP model, the assimilation of weather radar observations is critical because they provide the high temporal and spatial resolutions needed to observe such systems, and are the most important operational observation platform that is able to observe their interior structures

(Lilly 1990). The radial velocity ( $V_r$ ) and reflectivity ( $Z$ ) are the two main observed variables by Doppler weather radar. Among them, the assimilation of reflectivity is more challenging, especially in a variational framework, because of the high nonlinearity of its observation operator and its close involvement with the complex microphysics.

Over the past two to three decades, a number of methods have been proposed for assimilating radar reflectivity observations. These methods fall into the categories of semiempirical or indirect assimilation methods, variational, ensemble Kalman filter (EnKF), and hybrid ensemble–variational (EnVar) methods. In practice, to be able to analyze multiple hydrometeor

---

*Corresponding author:* Dr. Ming Xue, mxue@ou.edu

DOI: 10.1175/MWR-D-19-0149.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) ([www.ametsoc.org/PUBSReuseLicenses](https://www.ametsoc.org/PUBSReuseLicenses)).

species associated with ice microphysics schemes, semiempirical method such as the so-called complex cloud analysis schemes (Albers et al. 1996; Zhang et al. 1998) are often used together with another analysis scheme that assimilates other observations (Hu et al. 2006a; Xue et al. 2003). The cloud analysis schemes derive precipitating hydrometeors based on reflectivity and other cloud observations, and adjust temperature and moisture fields within precipitation regions, and have been shown to be effective in initializing convective storms, hurricanes and alleviating the precipitation spin up problem (e.g., Dawson and Xue 2006; Hu et al. 2006a; Jiang et al. 2015; Zhao and Xue 2009). Due to their relatively low computational cost, they have been routinely used in realtime forecasting systems (Kain et al. 2010; Xue et al. 2013) and operational forecasting (Hu et al. 2017; Weygandt et al. 2008). However, cloud analysis algorithms have their issues. They rely on empirical relations and assumptions to derive hydrometeor, temperature and moisture increments and do not adequately consider information and uncertainties contained in forecast background or observations, as optimization-theory-based modern data assimilation (DA) methods do. There is therefore a need to be able to assimilate reflectivity data directly using variational or ensemble-based methods.

Alternative to variational methods, EnKF has been shown to be quite effective in assimilating radar reflectivity observations into models that contain complex microphysics (Aksoy et al. 2009; Jung et al. 2008; Tong and Xue 2005; Xue et al. 2006; Xue et al. 2010). Because EnKF does not require the linear tangent or adjoint of the observation operators, it can handle nonlinear observation (such as radar reflectivity) operators without much technical difficulty. However, nonlinear operators do violate the Gaussian error distribution assumption inherent with Kalman filter, leading to suboptimal solutions, especially for dense highly nonlinear observations (Lorenc 2003). In addition, EnKF relies on ensemble-based background error covariance which is affected by sampling error and model error (Houtekamer and Mitchell 1998; Houtekamer et al. 2009); this problem can be alleviated to some extent by covariance localization techniques and/or by using a hybrid ensemble-variational (EnVar) framework that combines the ensemble-based and static background error covariances (Hamill and Snyder 2000). For example, the current operational North American Mesoscale Forecast System (NAM) and Rapid Refresh models of the U.S. National Weather Service use hybrid 3DVar DA method (Hu et al. 2017; Wu et al. 2017) except that their ensemble perturbations come from the EnKF system of

the Global Forecast System (Kleist and Ide 2015), and radar reflectivity data are assimilated using a cloud-analysis scheme. To directly assimilate reflectivity using either purely variational or hybrid EnVar method, adjoint of its observation operator is needed. Given that 3DVar with static background error covariance is a necessary building block for a hybrid EnVar system, it is important that reflectivity is assimilated as accurately as possible using 3DVar.

Attempts had been made to assimilate reflectivity data within variational frameworks. Although a nonlinear reflectivity observation operator can be used in variational DA, the linearization of the nonlinear observation operator can result in convergence problems and stability issues during variational minimization process. Within a four-dimensional variational Doppler radar analysis system (VDRAS) based on a cloud model that contained warm rain microphysics only, Sun and Crook (1997) compared direct assimilation of reflectivity with the assimilation of the rainwater mixing ratio retrieved from reflectivity (with the assumption of warm rain only, this can be easily done). They found that direct assimilation of reflectivity did not perform as well as assimilating derived rainwater, due to problems encountered during variational minimization that were shown to be related to the nonlinearity of the relation between reflectivity and rainwater mixing ratio. The assimilation of reflectivity within the WRF 3DVar and 4DVar systems follows a similar strategy, where rainwater mixing ratio converted from radar reflectivity is assimilated (Wang et al. 2013a,b). This leads to an indirect assimilation procedure that is also difficult to realize in the presence of ice microphysics unless empirical relations are introduced as in cloud analysis schemes. Earlier, a procedure for assimilating reflectivity data using 3DVar was implemented by Xiao et al. (2007) within the MM5 and later WRF 3DVar systems, assuming warm-rain microphysics only. Even though not discussed in their paper, problem associated with the nonlinear reflectivity operator should exist.

Gao and Stensrud (2012) directly assimilated reflectivity in the presence of multiphase hydrometeors within a 3DVar framework. To allow for physical partitioning of radar-observed precipitation among the hydrometeors, the reflectivity operator was modified so that its dependency on liquid or ice hydrometeor types was a function of temperature. Alternatively, Liu et al. (2019) chose to keep the original reflectivity operator unchanged, but employed a temperature-dependent hydrometeor background error profiles to achieve physical partitioning of radar-observed precipitation information among the liquid and ice hydrometeors. While both approaches produced reasonable analyses

of hydrometeors, there existed issues related to the nonlinearity of the observation operator of reflectivity in their variational minimization processes. Trying to reduce the issues, Carley (2012) used logarithmic mixing ratios of hydrometeors as the control variables (which we call CVlogq) instead of mixing ratios themselves (called CVq) within a hybrid 3D ensemble–variational (3DEnVar) system to assimilate radar reflectivity, but the approach also had its problems. Recognizing the problems, Wang and Wang (2017) proposed an alternative approach within an EnVar framework that does not use the linear tangent or adjoint of the reflectivity operator; in their system reflectivity instead of the hydrometeor mixing ratios is used as a state variable. However, with such an approach, the EnVar solution is effectively equivalent to that of EnKF. To be able directly assimilate reflectivity data within a hybrid EnVar DA system that includes static background covariance, linear tangent and adjoint of the observation operator will still be needed and the nonlinearity problem is therefore unresolved. Kong et al. (2018) compared direct assimilation of reflectivity between pure 3DEnVar that uses 100% ensemble covariance and EnKF through observing system simulation experiments (OSSEs). The analysis results from EnKF were better than those of pure 3DEnVar, which is affected by the convergence problem related to nonlinear reflectivity observation operator. Similar problem also existed when static background error is included in a hybrid 3DEnVar version (Kong et al. 2018).

In previous studies, some treatments have been proposed to deal with the problems related to the nonlinear reflectivity operator. In Sun and Crook (1997), a lower limit was added to rainwater mixing ratio (which we call the qLim approach) in their rainwater-only reflectivity operator. With the CVlogq approach, issues related to logarithmic transformation were noticed but not solved (Carley 2012; Wang and Wang 2017). Thus, the main goals of this study are to investigate different treatments specific to different control variables (i.e., the mixing ratios or logarithmic mixing ratios, and their relative performance when assimilating reflectivity directly within a 3DVar framework). OSSEs will be performed with simulated radar data for a model supercell storm closely following our earlier studies (Kong et al. 2018; Liu et al. 2019).

We note here that the choice of water vapor or hydrometeor variables in DA systems has received some attention in the literature. For example, Dee and Da Silva (2003) recommended the use of pseudo–relative humidity as the moisture analysis variable because of easier modeling of its background error and less coupling with temperature analysis. Fletcher and Zupanski (2006) explored variational assimilation of observations

whose error distribution is lognormal, which may be more suitable for certain moisture or precipitation observations. Ingleby et al. (2013) introduced a humidity control variable via nonlinear transform that is used in the Met Office variational DA system that seems to benefit the assimilation of humidity-sensitive satellite channels. These methods, however, do not directly address issues associated with the assimilation of radar reflectivity, whose observation operator is nonlinear and error distribution is non-Gaussian.

The rest of this paper is organized as follows. In section 2, reflectivity observation operators using CVq or CVlogq are introduced. Problems of using CVq or CVlogq within a 3DVar framework assimilating reflectivity data are discussed in section 3. The proposed treatments for directly assimilating reflectivity are described in section 4. The design of OSSEs is presented in section 5 while the results of the experiments with different treatments are discussed in section 6. Finally, conclusions are given in section 7.

## 2. Reflectivity observation operators

### a. Reflectivity observation operator using CVq

The reflectivity observation operator can be written in a general form:

$$Z = 10 \log_{10}(Z_e), \tag{1}$$

where  $Z_e$  is the equivalent reflectivity factor, which typically contains the contributions from three categories of hydrometeor mixing ratios [i.e., rainwater ( $q_r$ ), snow ( $q_s$ ), and hail ( $q_h$ )], according to the following equation:

$$Z_e = Z_{er}(q_r) + Z_{es}(q_s) + Z_{eh}(q_h). \tag{2}$$

The above assumes that there are three categories of precipitating hydrometeors within the microphysics scheme used in a numerical model and only the mixing ratios are predicted; this is true with the one-moment Lin ice microphysics scheme (Lin et al. 1983) used in this study. Based on the default values of the intercept parameters of hydrometeor size distributions and the default hail density of the Lin scheme, the equivalent reflectivity factors contributed from rainwater, snow, and hail are, respectively,

$$Z_{er} = 3.63 \times 10^9 \times (\rho q_r)^{1.75}, \tag{3}$$

$$Z_{es} = \begin{cases} 9.80 \times 10^8 \times (\rho q_s)^{1.75} & T \leq 0^\circ\text{C} \\ 4.26 \times 10^{11} \times (\rho q_s)^{1.75} & T > 0^\circ\text{C} \end{cases}, \tag{4}$$

$$Z_{eh} = 4.33 \times 10^{10} \times (\rho q_h)^{1.75}, \quad (5)$$

where  $T$  is air temperature in Celsius. These equations are the same as those given in [Tong and Xue \(2005\)](#) but with the default parameter values plugged in.

### b. Reflectivity observation operator using CVlogq

When using the logarithmic mixing ratios as the control variables (i.e., when using CVlogq), hydrometeor mixing ratios are transformed into the logarithmic space. Taking rainwater as an example, the transformation can be expressed as

$$\hat{q}_r = \log_{10}(q_r). \quad (6)$$

Thus, the equivalent reflectivity factor for rainwater can be written as

$$Z_{er} = 3.63 \times 10^9 \times (\rho 10^{\hat{q}_r})^{1.75}. \quad (7)$$

Similar transformation can be applied to the reflectivity components of snow and hail, effectively replacing  $q_s$  and  $q_h$  in Eqs. (4) and (5) with  $10^{\hat{q}_s}$  and  $10^{\hat{q}_h}$ , respectively. The total equivalent reflectivity for CVlogq is then

$$Z_e = Z_{er}(\hat{q}_r) + Z_{es}(\hat{q}_s) + Z_{eh}(\hat{q}_h). \quad (8)$$

We note here that given the observed reflectivity is the  $Z$  in Eq. (1), the log operation applied to equivalent reflectivity factor  $Z_e$  will make the relation between  $Z$  and  $\hat{q}$  linear, at least when there is only one hydrometeor species. This property likely can make the modeling of the error statistics of  $\hat{q}$  easier.

## 3. Problems with CVq and proposed treatments

### a. Extremely large gradient of 3DVar cost function for CVq

In 3DVar, the optimal analysis is obtained by minimizing the following cost function:

$$J = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}[H(\mathbf{x}) - \mathbf{y}]^T \mathbf{R}^{-1}[H(\mathbf{x}) - \mathbf{y}], \quad (9)$$

where  $\mathbf{x}$  is the analysis state vector and  $\mathbf{x}_b$  is the background state vector;  $\mathbf{y}$  is the observation vector;  $\mathbf{B}$  and  $\mathbf{R}$  are the background and observation error covariance matrices, respectively; and  $H$  is the possibly nonlinear observation operator, which typically includes spatial interpolation from model space to observation space, and transformation from model state variables to observation variables. The  $Z$  observation operator given in

the previous section is an example that transforms hydrometeor mixing ratios to reflectivity.

The minimization of the cost function requires the calculation of the gradient of the cost function  $J$  with respect to state vector  $\mathbf{x}$ , and the gradient is given by

$$\nabla_x J = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1}[H(\mathbf{x}) - \mathbf{y}], \quad (10)$$

where  $\mathbf{H}$  is the linear tangent version of operator  $H$  and  $\mathbf{H}^T$  is the adjoint of  $\mathbf{H}$ .

Based on Eqs. (1)–(3), the linear tangent of the  $Z$  operator or the gradient of  $Z$  with respect to the rainwater mixing ratio is

$$\frac{\partial Z}{\partial q_r} = \frac{\partial Z}{\partial Z_e} \frac{\partial Z_e}{\partial q_r} = \frac{6.35 \times 10^{10} \times \rho^{1.75} q_r^{0.75}}{\ln 10 \times Z_e}. \quad (11)$$

Very similar equations can be derived for the gradient with respect to snow and hail mixing ratios.

Considering the simpler situation when only rainwater contributes to reflectivity (e.g., at the low levels where only liquid rainwater is expected), the linear tangent of  $Z$  observation operator with respect to rainwater mixing ratio, after plugging in  $Z_{er}$  from Eq. (3), is

$$\frac{\partial Z}{\partial q_r} = \frac{17.5}{\ln 10 \times q_r}. \quad (12)$$

In Eq. (11) or Eq. (12), when  $Z_e$  or  $q_r$  is close to zero, the gradient becomes very large or approaches infinity and will cause problems to the convergence of cost function minimization ([Sun and Crook 1997](#); [Wang and Wang 2017](#)). The same is true in the case of having snow or hail mixing ratio only, or when the total  $Z_e$  including contributions from all species is close to zero. In addition, when both radial velocity and reflectivity observations are assimilated together, the much larger gradient of reflectivity relative to that of radial velocity will make the assimilation of radial velocity ineffective (i.e., little correction will be made to the wind components as the cost function sensitivity to these components is too small in relative terms). Without a better solution, one possibility of avoiding this problem is to assimilate radial velocity and reflectivity data separately, in two different passes (i.e., assimilating reflectivity and other observations first, then assimilate radial velocity in a second pass using the analysis of the first pass as the new background). While this can be theoretically problematic since the background error is typically not updated across the analysis passes, it does help realize the impact of  $V_r$  data, as will be seen later in [section 3b](#). Other solutions involve the choice of an alternative control variable and applying additional treatments, as will be discussed next.

b. Proposed treatments when using CVq

1) ADDING A LOWER LIMIT TO HYDROMETEOR MIXING RATIOS IN Z OPERATOR (QLIM)

As mentioned earlier, to avoid extremely large gradient of the cost function, Sun and Crook (1997) introduced a lower limit,  $\epsilon_q$ , to rainwater mixing ratio  $q_r$ , effectively setting the gradient of reflectivity with respect to  $q_r$  to zero in their rainwater-only reflectivity formula when  $q_r$  is less than lower limit. For multiphased microphysics, with a similar qLim treatment, the Z operator and its linear tangent term are written as

$$Z = H[\max(q, \epsilon_q)], \tag{13}$$

$$\frac{\partial Z}{\partial q} = \begin{cases} \frac{\partial Z(q)}{\partial q} & q > \epsilon_q \\ 0 & q \leq \epsilon_q \end{cases}, \tag{14}$$

where  $q$  can be any of the hydrometeor mixing ratios in the reflectivity formula.

2) MODIFYING EQUIVALENT REFLECTIVITY FOR SMALL MIXING RATIO (ZELIM)

The qLim treatment can greatly reduce the extremely large gradient of cost function when a large lower limit is used. However, with this treatment, reflectivity observations cannot be effectively assimilated in the area where the mixing ratio is no larger than the lower limit because the linear tangent is zero. To avoid this problem, an alternative treatment is proposed, which modifies the equivalent reflectivity for small mixing ratio in the Z operator, to the following:

$$Z_e = \begin{cases} Z_e(q) & Z_e \geq \epsilon_{Z_e} \\ Z_e(q) + \epsilon_{Z_e} & Z_e < \epsilon_{Z_e} \end{cases}, \tag{15}$$

where  $\epsilon_{Z_e}$  is a small equivalent reflectivity value. This treatment avoids the problem associated with  $Z_e$  in Eq. (11) going to zero.

3) USING A SEPARATE PASS TO ASSIMILATE  $V_r$  DATA (VRPASS)

As mentioned in section 3a, when using CVq, because the cost function gradient for the observation term corresponding to radial velocity  $V_r$  observation is much less than Z observation, especially when  $q$  is small, the impact of assimilating  $V_r$  observations is difficult to realize when both of  $V_r$  and Z observations are assimilated together (Sun and Crook 1997; Wang and Wang 2017). To alleviate this problem, we use two separate analysis passes, with  $V_r$  data analyzed in a second 3DVar pass.

Thus, the minimization procedures of the  $V_r$  cost function are not affected by the extremely large gradient of the cost function corresponding to the Z observation. This approach is theoretically problematic in a 3DVar framework, however, when the background error cannot be easily updated after Z and other observations are assimilated. When  $V_r$  data are assimilated in the second pass, if the wind fields have been updated by other observations, using the same background error of wind fields tend to overestimate the  $V_r$  data impact. In our case, the Z and  $V_r$  data can only directly affect some of the state variables (i.e., state variables that appear in their observation operators). The 3DVar background error covariance typically does not contain cross covariance among the hydrometeor and wind variables for convective-scale DA. In the experiments of this paper, only Z and  $V_r$  data are assimilated and the state variables they update do not overlap, therefore there is no theoretical problem employing the  $V_r$ Pass treatment. In practice, multiple analyses passes are often used within 3DVar in order to allow for the use of different background error correlation scales (e.g., Li et al. 2012; Schenkman et al. 2011). Performing separate analysis passes does increase computational cost. The increase will be even more in a hybrid EnVar system when the ensemble-based covariance processing carries significant cost. Furthermore, to be theoretically correct, the coupling EnKF system that provides the ensemble perturbations should also be performed using separate passes, assimilating the same set of observations in each pass, as was tested in Pan et al. (2014). This will create much more computational burden; therefore, a single pass that assimilates all observations is strongly preferred.

4) USING LOGARITHMIC MIXING RATIOS AS THE CONTROL VARIABLES (CVLOGQ)

An alternative way to avoid the extremely large gradient of the cost function for reflectivity DA is to use a logarithmic transformation of the hydrometeor mixing ratios for the control variables [i.e., using CVlogq, as given in Eq. (6)]. With CVlogq, the linear tangent of Z operator with respect to logarithmic rain mixing ratio can be written as

$$\frac{\partial Z}{\partial \hat{q}_r} = \frac{17.5 Z_{er}}{Z_e}. \tag{16}$$

If considering a single-phased hydrometeor only, the linear tangent term will be a constant, that is, 17.5 dBZ  $[\log(\text{kg kg}^{-1})]^{-1}$ . If taking into account multiphased hydrometeors, the linear tangent term will not be a constant but no larger than 17.5 dBZ  $[\log(\text{kg kg}^{-1})]^{-1}$ .

Thus, the extremely large gradient of the cost function can be avoided when using CVlogq.

$$\rho_{ij} = \frac{\delta q_i}{\delta q_j} \tag{20}$$

#### 4. Problems with CVlogq and treatments

##### a. Problem with spatial spreading of analysis increment

As mentioned in the previous section, using CVlogq can avoid the extremely large gradient of cost function for small values of mixing ratios or equivalent reflectivity, making the assimilation of larger reflectivity values (e.g., >35 dBZ) and radial velocity observations more effective. However, new problems can arise with the use of CVlogq. With CVlogq, typically, background errors with Gaussian spatial correlation (e.g., Huang 2000) for hydrometeor mixing ratios are simulated in log(*q*) space (Carley 2012). However, when transforming the Gaussian-shaped analysis increment in log(*q*) space back to the *q* space, the analysis increment is no longer Gaussian-like because of the high nonlinearity of the log transformation. With the nonlinear transform, strange looking/suspiciously shaped analysis increments can result in the increment of *q*. To illustrate, we examine the single-observation analysis problem. Assume that the background errors of hydrometeors are spatially homogenous, and the observation is located at grid point *j*. Since the observation operator is an identity matrix here, the solution of 3DVar for this single-observation test is the same as that of optimal interpolation (Kalnay 2002). Thus, the analysis increment of  $\hat{q} \equiv \log(q)$  at grid point *i* is

$$\delta \hat{q}_i = \hat{\sigma}_b^2 \hat{\rho}_{ij} (\hat{\sigma}_b^2 + \hat{\sigma}_o^2)^{-1} d, \tag{17}$$

where  $\hat{\sigma}_b^2$  and  $\hat{\sigma}_o^2$  are background and observation error variances, respectively; *d* is the observation innovation; and  $\hat{\rho}$  is background error correlation for log(*q*) (Kalnay 2002). At the observation point *j*, the analysis increment is

$$\delta \hat{q}_j = \hat{\sigma}_b^2 (\hat{\sigma}_b^2 + \hat{\sigma}_o^2)^{-1} d. \tag{18}$$

When dividing (17) by Eq. (18):

$$\hat{\rho}_{ij} = \frac{\delta \hat{q}_i}{\delta \hat{q}_j} \tag{19}$$

The above is also the effective spatial correlation of background error expressed in terms of the spatial spread of analysis increment of the control variable. Similarly, the background error correlation between *q* at points *i* and *j* is equal to

To convert the increment from log(*q*) space back to the *q* space, the following equation is used:

$$\delta q = 10^{\delta \hat{q} + \log q_b} - q_b. \tag{21}$$

Using Taylor series expansion, expanded around the value log $q_b$ ,  $10^{\hat{q}}$  ( $\hat{q} = \delta \hat{q} + \log q_b$ ) can be written as

$$10^{\delta \hat{q} + \log q_b} = 10^{\log q_b} + 10^{\log q_b} \ln 10 \times \delta \hat{q} + \sum_{k=2}^{\infty} 10^{\log q_b} (\ln 10)^k \delta \hat{q}^k (k!)^{-1}. \tag{22}$$

Substituting Eq. (22) into Eq. (21), the analysis increment of *q* given increment of  $\hat{q}$  is

$$\begin{aligned} \delta q &= 10^{\delta \hat{q} + \log q_b} - q_b \\ &= 10^{\log q_b} + 10^{\log q_b} \ln 10 \times \delta \hat{q} \\ &\quad + \sum_{k=2}^{\infty} 10^{\log q_b} (\ln 10)^k \delta \hat{q}^k (k!)^{-1} - q_b \\ &= q_b \ln 10 \times \delta \hat{q} + \sum_{k=2}^{\infty} q_b (\ln 10)^k \delta \hat{q}^k (k!)^{-1}. \end{aligned} \tag{23}$$

With Eq. (19), substituting Eq. (23) into Eq. (20), we obtain

$$\begin{aligned} \rho_{ij} = \frac{\delta q_i}{\delta q_j} &= \frac{q_{b_i} \ln 10 \times \delta \hat{q}_i + \sum_{k=2}^{\infty} q_{b_i} (\ln 10)^k \delta \hat{q}_i^k (k!)^{-1}}{q_{b_j} \ln 10 \times \delta \hat{q}_j + \sum_{k=2}^{\infty} q_{b_j} (\ln 10)^k \delta \hat{q}_j^k (k!)^{-1}} \\ &= \frac{\left[ 1 + \sum_{k=2}^{\infty} (\ln 10)^{k-1} \delta \hat{q}_i^{k-1} (k!)^{-1} \right]}{\left[ 1 + \sum_{k=2}^{\infty} (\ln 10)^{k-1} \delta \hat{q}_j^{k-1} (k!)^{-1} \right]} \times \frac{q_{b_i}}{q_{b_j}} \times \hat{\rho}_{ij}. \end{aligned} \tag{24}$$

Based on Eq. (24), we have the following discussions on the relationship between  $\rho$  and  $\hat{\rho}$ :

- 1) If the analysis increment of log(*q*) is very small, and the background is homogeneous (constant) (i.e., if  $\delta \hat{q} \rightarrow 0$  and  $q_{bi} = q_{bj}$ ), Eq. (24) can then be written as

$$\rho_{ij} \approx \hat{\rho}_{ij}, \tag{25}$$

which means that the effective spatial correlation in *q* space is almost the same as that in log(*q*) space.

- 2) If the analysis increment in log(*q*) space is not small compared to the background value, and the background is homogeneous, because the analysis increment near the observation point is always no greater

than that in the observation point (i.e.,  $\delta\hat{q}_i \leq \delta\hat{q}_j$ ), then Eq. (24) will give rise to

$$\rho_{ij} \leq \hat{\rho}_{ij}, \tag{26}$$

which means that if the background is homogeneous, the effective correlation of  $q$  is always no larger than that of  $\log(q)$ .

- 3) If the analysis increment in the  $\log(q)$  space is very small, but the background is inhomogeneous (i.e.,  $\delta\hat{q} \rightarrow 0$  and  $q_{b_i} \neq q_{b_j}$ ), Eq. (24) can be written as

$$\rho_{ij} \approx \frac{q_{b_i}}{q_{b_j}} \times \hat{\rho}_{ij}, \tag{27}$$

which means that when the background at the analysis grid point  $i$  is larger (or smaller) than that at the observation point  $j$ , the effective correlation in  $q$  space is larger (or smaller) than that in the  $\log(q)$  space. As a result, the analysis increment of  $q$  at grid point  $i$  will be larger (or smaller) than what the Gaussian correlation model would suggest.

In summary, when a Gaussian-like spatial correlation function is used to simulate the background error correlation in the space of control variable  $\log(q)$ , the effective spatial correlation in  $q$  will be almost Gaussian only when the analysis increments are very small and the background is homogenous within the radius of meaningful correlation. When either of the two conditions is not satisfied, the effective correlation function in  $q$  can deviate significantly from Gaussian, leading to strange looking/suspiciously shaped analysis increments in  $q$ .

To better understand the above discussions, single-pseudo-observation experiments in a one-dimensional space with homogenous and inhomogeneous background are conducted. The grid spacing is 2 km and the grid number is 30. A homogeneous background of rainwater mixing ratio  $q$  is set as  $0.3 \text{ g kg}^{-1}$ . To simulate an inhomogeneous background, a sectional function in terms of the gridpoint number  $n$  is assumed as follows:

$$q_b = \begin{cases} 0.1 & n < 14 \\ 0.3 & 14 \leq n \leq 16 \\ 0.7 & n > 17 \end{cases} \tag{28}$$

An observation of rainwater mixing ratio itself is placed at the 15th grid point and the observed value is 1 or  $0.6 \text{ g kg}^{-1}$ , respectively (Fig. 1a). The observation error is assumed to be  $0 \text{ g kg}^{-1}$ . Thus, the optimal analysis increments at the observation point should be 0.7 and  $0.3 \text{ g kg}^{-1}$ , in the two cases, respectively, according to Eq. (18) (as applied to  $q$ ), which represent large and small analysis increment relative to the background value (of  $0.3 \text{ g kg}^{-1}$ ) at the observation point, respectively. The

background error variance is assumed to be homogeneous, and a Gaussian function with 5 km de-correlation scale is used to construct the background error correlation model. A configuration summary for single-observation tests is given in Table 1.

The shape of the analysis increment using CVq for both observations is Gaussian-like (red line in Fig. 1b), and independent of the background. With CVlogq, the same Gaussian function is used to construct the background error correlation model. Thus, the shape of the analysis increment of  $\log(q)$  is also Gaussian-like as shown in Fig. 1c. As discussed earlier, the Gaussian-like spatial correlation function of  $\log(q)$  background error cannot be Gaussian-like in  $q$  space unless the analysis increment is very small (so that the problem is nearly linear) and the background is homogenous; this will reflect on the spatial spread of analysis increment in  $q$  space. As shown in Fig. 1b, when the analysis increment is small and the background is homogenous (CVlogq-IS-BH, black dash line in Fig. 1b), the shape of the analysis increment of  $q$  is essentially Gaussian-like and close to that of CVq (CVq-IS red dash line). Conversely, if the analysis increment is large, the spread of the analysis increment from CVlogq becomes narrower (CVlogq-IL-BH, black solid line). When the background is spatially inhomogeneous (blue line in Fig. 1a), the analysis increment in  $q$  is overestimated or underestimated (CVlogq-IS-BI and CVlogq-IL-BI, blue lines in Fig. 1b). This is because that even though the background error spatial correlation is assumed to be Gaussian for  $\log(q)$ , the spatial correlation for  $q$  is not Gaussian according to Eq. (27) when the background value of  $q$  is spatially varying. Such situation commonly occurs in convective regions, and the gradient in  $q$  can be especially large at the cloud boundaries. In short, the spatial spread of  $q$  analysis increment using CVlogq is consistent with earlier discussions on the spatial correlation of the background error.

*b. Treatment: Adding a smoothing function and a lower limit to background ( $X_bSL$ )*

To alleviate above problems caused by inhomogeneous background, a smoothing function (denoted  $f$ , a five-point smoothing function is used in our study) is used to smooth the background when converting  $\delta\hat{q}$  back to  $\delta q$ . A five-point smoother, also called the triangular smoother, is given by

$$\bar{q}_j = \frac{q_{j-2} + 2q_{j-1} + 3q_j + 2q_{j+1} + q_{j+2}}{9}, \tag{29}$$

where  $\bar{q}_j$  is the smoothed value at grid point  $j$ . In other words, when converting  $\delta\hat{q}$  to  $\delta q$ , the background  $q_b$  used in conversion is replaced by  $\bar{q}_b$ . It is important to

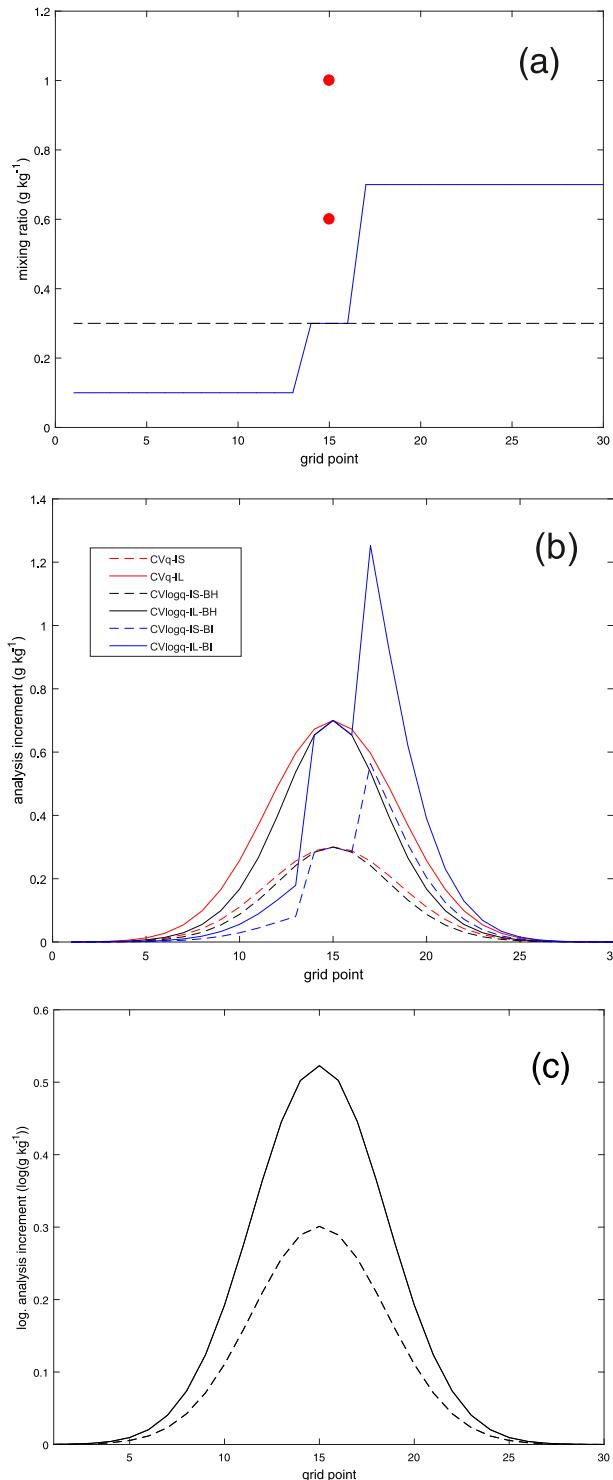


FIG. 1. Single-observation tests (see Table 1) with (a) two backgrounds (black dash line for homogeneous background, and blue line for inhomogeneous background) and two observations (red dots at 1 and 0.6 g kg<sup>-1</sup>). (b) Analysis increments of large (solid line) or small (dashed line) amplitude and using CVq (red line) or CVlogq for the case of spatially homogeneous background (black lines) and inhomogeneous background (blue

note that the background used for calculating innovation is still the original one (without applying the smoother). In addition, to avoid negative infinite value in the logarithmic transform, a lower limit is added to the background. Thus, Eq. (21) can be written as

$$\delta q = 10^{\delta q + \log\{f[\max(q_b, \varepsilon_b)]\}} - f[\max(q_b, \varepsilon_b)]. \quad (30)$$

This smoothing procedure avoids large discontinuity in the analysis increment of  $q$ .

To examine the impact of the above smoothing function, single-observation experiments similar to the previous section are conducted but using a smoothed inhomogeneous background (red line in Fig. 2a). Figure 2b shows that the problem of overestimating or underestimating analysis increments (blue line) are somewhat alleviated when the smoothed background is used when converting  $\delta \hat{q}$  back to  $\delta q$ .

## 5. Assimilation system and experimental design

The use of  $\log(q)$  control variables is first added to the ARPS 3DVar system (Xue et al. 2003; Gao et al. 2004), which is used to test our treatments proposed in section 4 through OSSEs. We note that no cross-variable background error covariance is included so the 3DVar system is more or less univariate, especially when the mass continuity constraint that couples three wind components together (Hu et al. 2006b) is not included, as is the case in this paper (so as to show more clearly the direct impact of assimilation  $V_r$  data). A supercell storm that occurred on 20 May 1977 in Del City, Oklahoma (Ray et al. 1981), is selected for the OSSEs. The model configurations are similar to those of Tong and Xue (2005) and Kong et al. (2018) except for a larger domain. The domain size is  $51 \times 51 \times 35$  grid points, with 2 km horizontal and 0.5 km vertical grid spacings, respectively. The recursive filter length scales for the background error spatial covariance in the horizontal and vertical directions are 4.1 and 1.4 km, respectively, which are optimal for ARPS 3DVar for the current OSSE problem (Kong et al. 2018). A limit of 200 iterations was imposed for the minimization and the convergence criteria for cost function were set to  $10^{-12}$ .

Temperature-dependent background error variance profiles for the hydrometeors (Liu et al. 2019), of which the error variance ranged from 0 to 1.2 k kg<sup>-1</sup>, are used for all experiments. Observation errors of radial velocity

←

lines). (c) Analysis increment in  $\log(q)$  space using CVlogq. In (b) and (c), solid lines are for the case of large analysis increments (for observation of 1 g kg<sup>-1</sup>) while dashed lines are for the case of small analysis increments (for observation of 0.6 g kg<sup>-1</sup>).

TABLE 1. Configuration summaries for single-observation tests.

Experiment name	Control variable	Increment	Background
CVq-IL	CVq	Large	Homogeneous or inhomogeneous
CVq-IS		Small	Homogeneous or inhomogeneous
CVlogq-IL-BH	CVlogq	Large	Homogeneous
CVlogq-IL-BI		Large	Inhomogeneous
CVlogq-IS-BH		Small	Homogeneous
CVlogq-IS-BI		Small	Inhomogeneous

and reflectivity are assumed to be  $1 \text{ m s}^{-1}$  and  $3 \text{ dBZ}$ , respectively. With these error standard deviations, unbiased normal distribution random errors are added to the simulated radial velocity and reflectivity to simulate observations errors. A single radar observes the simulated storm from the southwest corner of the model domain, and  $V_r$  and  $Z$  observations are simulated from model output at 5-min intervals. The  $V_r$  observations are assumed to be available at grid points where simulated reflectivity is no less than  $10 \text{ dBZ}$  while  $Z$  observations throughout the domain are assimilated. The assimilation of clear-air reflectivity helps suppress spurious reflectivity in the background (Tong and Xue 2005).

Using CVq, three sets of experiments are conducted with 1) qLim treatment (CVq-qLim), 2) ZeLim treatment (CVq-ZeLim), and 3) ZeLim and VrPass treatments together (CVq-ZeLim-VrP). When using qLim and VrPass treatments together (CVq-qLim-VrP), because there is no cross correlation between hydrometeor and wind variables, the wind analysis obtained in the second pass of CVq-qLim-VrP using  $V_r$  data will be the same as that in CVq-ZeLim-VrP since the results only depend on the background wind fields. Since the main purpose of the VrPass is to improve the wind analysis, experiment CVq-qLim-VrP is not performed. For CVlogq, two sets of experiments are conducted, which are 4) using CVlogq without treatment (CVlogq-NT) and 5) with  $X_b$ SL treatment (CVlogq- $X_b$ SL), respectively. A list of experiments using different treatments is given in Table 2.

With the above five combinations of control variables and treatments, we perform five cycled DA experiments, over a 1-h period, corresponding to simulated storm time of 25 min through 85 min. Our experiment setup closely follows that 3DVar experiment in Kong et al. (2018) except that a large domain is used to allow for the examination of longer forecasts. The 5-min spinup ensemble forecasts are first run from randomly perturbed initial conditions at 20 min, and the first analysis is performed at 25 min, using the ensemble mean forecast as the background. The DA cycles are run at 5-min intervals until 85 min when 60-min-long forecasts are launched. The final analyses and 30- and 60-min forecasts are compared.

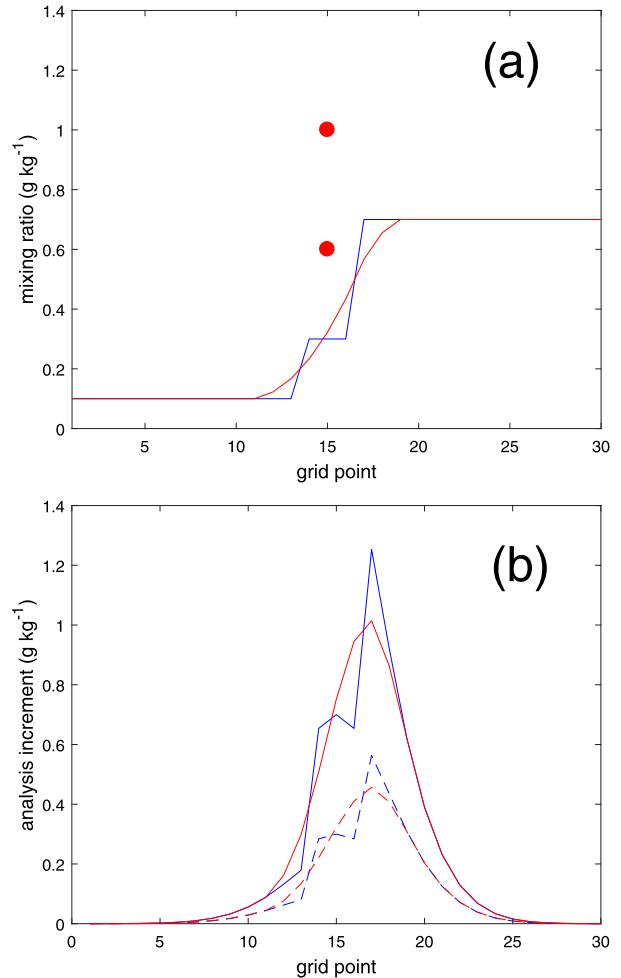


FIG. 2. As in the single-observation tests using inhomogeneous background in Fig. 1, but also showing (a) the smoothed background (red line) and (b) the corresponding analysis increments (blue lines) for observation of  $1$  (solid lines) and  $0.6 \text{ g kg}^{-1}$  (dashed lines).

The differences in the final analyses of the cycled experiments are the results of many DA cycles. To better identify the direct impact of the different control variables and treatments, we perform a set of single-time analysis experiments, where the forecast background from cycled experiment CVq-qLim at 85 min is used in all experiments. At this time the simulated supercell storm is well developed. The results of these single-time analyses are presented first in the next section.

## 6. Experimental results

### a. Results of single-time analyses

For experiments CVq-qLim, CVq-ZeLim, and CVlogq- $X_b$ SL, we first need to determine the optimal values of the lower thresholds, that is, of  $\epsilon_q$ ,  $\epsilon_{Z_e}$ , and  $\epsilon_b$ , respectively. We do this through the single-time analysis by trying different

TABLE 2. Configuration summaries for radar DA experiments using different treatments.

Experiment name	Control variable	Treatments
CVq-qLim	CVq	qLim
CVq-ZeLim		ZeLim
CVq-ZeLim-VrP		ZeLim+VrPass
CVlogq-NT	CVlogq	No treatment
CVlogq-X <sub>b</sub> SL		X <sub>b</sub> SL

values. For the qLim treatment using CVq, Fig. 3a shows that the magnitude of the cost function gradient decreases monotonically in the 3DVar minimization when  $\varepsilon_q$  is increased from  $10^{-8}$  to  $10^{-5} \text{ kg kg}^{-1}$ , indicating a clear impact of imposing qLim. However, because the gradient is effectively set to zero when the expected mixing ratio is no larger than the threshold in Eq. (14), a larger  $\varepsilon_q$  will lead to less chance of adjustment to mixing ratio in the analysis. The root-mean-square innovations (RMSI) of reflectivity with different  $\varepsilon_q$  values are calculated for all grid points where the observed reflectivity is greater than 15 dBZ. The number of observations used is on the order of 100. Figure 3b shows that when increasing the threshold, RMSI decreases until the threshold reaches  $10^{-5} \text{ kg kg}^{-1}$ . Even though a close fit to observations does not necessarily indicate a good analysis and may degrade the forecast, however, because convergence problem occurs with the 3DVar analysis when the background or expected analysis values of mixing ratio approaches zero, how well the analyzed mixing ratio field fits the observations gives a good idea on the success of the variational analysis. Based on the RMSIs in Fig. 3b,  $\varepsilon_q = 10^{-6} \text{ kg kg}^{-1}$  is considered the optimal choice and is there used in the cycled CVq-qLim experiment in the remainder of this paper. Using the same approach, the optimal thresholds of ZeLim and X<sub>b</sub>SL [i.e.,  $\varepsilon_{z_e}$  in Eq. (15) and  $\varepsilon_b$  in Eq. (30)], are found to be

$1 \text{ mm}^6 \text{ m}^{-3}$  and  $10^{-6} \text{ kg kg}^{-1}$ , respectively, and these values are used in the corresponding cycled experiments. For general applications, these threshold values should be further optimized.

The reflectivity RMSIs of the forecast background and the analyses of single-time analysis experiments CVq-qLim, CVq-ZeLim, CVlogq-NT, and CVlogq-X<sub>b</sub>SL are shown in Fig. 4. All analysis RMSIs are much smaller than that of background, indicating that the 3DVar is effective in assimilating the reflectivity observations in these experiments. When using CVq, in terms of RMSI, the ZeLim treatment is better than qLim, producing average RMSI that is about 1 dB lower. When using CVlogq without any special treatment in CVlogq-NT, the RMSI is similar to that of CVq-ZeLim while with additional X<sub>b</sub>SL treatment in CVlogq-X<sub>b</sub>SL, the analysis is further improved, although only by a small amount (Fig. 4).

Overall, the differences in reflectivity RMSIs of the final analyses using CVq and CVlogq are not that large (Fig. 4). A much bigger difference is in the 3DVar minimization convergence speed. This can be seen in the reduction rate of reflectivity RMSI as a function of minimization iterations, as shown in Fig. 5; it is clear that the convergence using CVlogq is much faster than using CVq. After about 15 iterations, the reflectivity RMSIs of CVlogq experiments become much lower than those of CVq experiments. For CVlogq experiments, the minimum of RMSI is essentially reached after 50 iterations while for CVq experiments, at least 200 iterations are needed, and the final RMSI of CVq-qLim stays higher than those of other experiments at 200 iterations (Fig. 5). The use of CVlogq generally reduces that range of values of the cost function gradient with respect to the control variable, leading to faster convergence. In general terms, the choice of control variable

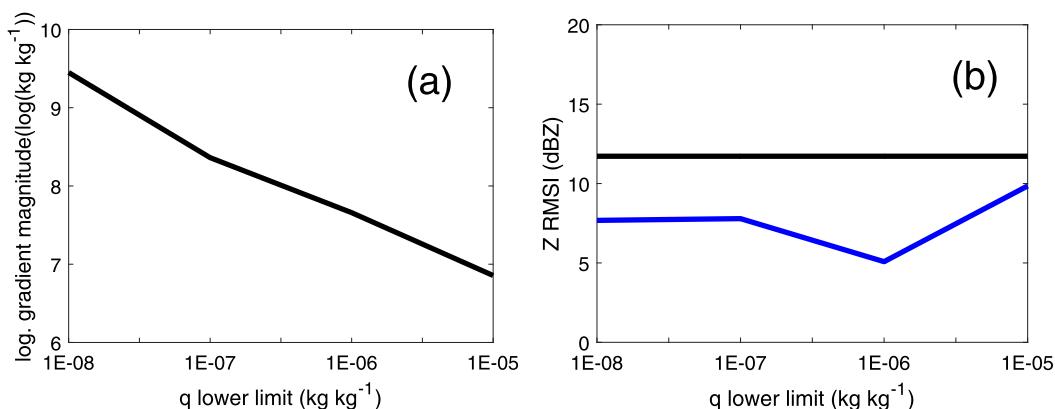


FIG. 3. (a) The total logarithmic magnitude of the cost function gradient with different lower limit of  $q$  or  $\varepsilon_q$  in Eq. (14) and (b) the reflectivity RMSIs for background (black line) and analysis (blue line) with different  $\varepsilon_q$ , for the single-time 3DVar analysis using CVq applying qLim.

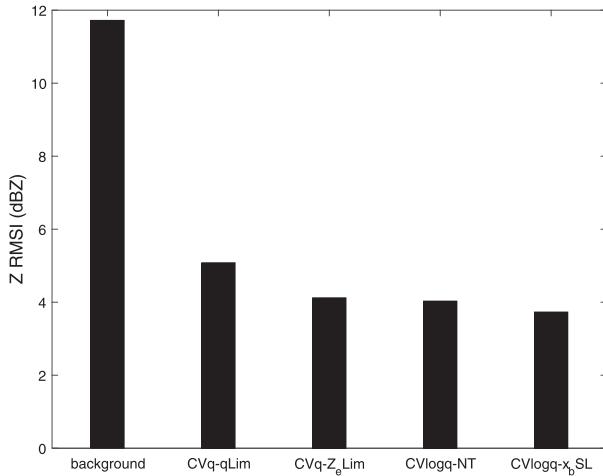


FIG. 4. The reflectivity RMSIs of the forecast background, and those of the final analyses of single-time analysis experiments CVq-qLim, CVq-ZeLim, CVlogq-NT, and CVlogq-X<sub>b</sub>SL.

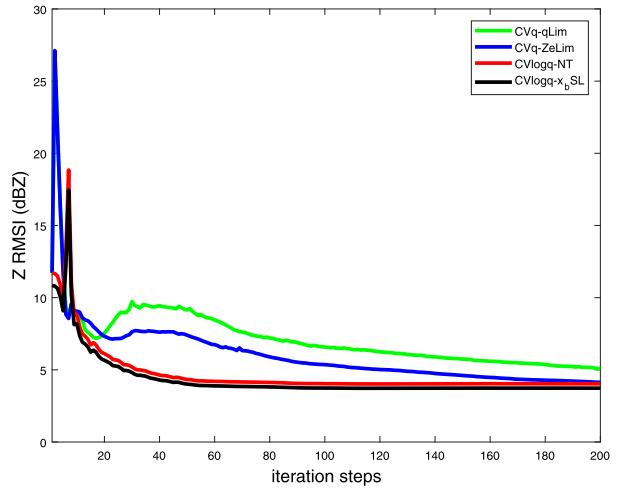


FIG. 5. Reflectivity RMSIs with different minimization iteration steps for CVq-qLim, CVq-ZeLim, CVlogq-NT, and CVlogq-X<sub>b</sub>SL.

and certain special treatment affects the condition number of the minimization problem (Bouttier and Courtier 1999).

The treatments related to hydrometeor mixing ratios or reflectivity or the choice of mixing ratio-related control variables have been shown to have direct effect on the analysis of hydrometeor fields which directly affect the analyzed reflectivity and therefore the fit of analyzed reflectivity to observations or reflectivity RMSI. The effects on the analyzed wind fields are not necessarily the same. The quality of the wind analysis is examined by looking at radial velocity RMSIs. Those for the forecast background, and the wind analyses of CVq-ZeLim, CVq-ZeLim-VrP, and CVlogq-NT are shown in Fig. 6. It can be seen that the assimilation of radial velocity observations together with the reflectivity data within the same pass produces a rather small impact on the background RMSI when using CVq (in CVq-ZeLim); this is due to, as suggested earlier, the gradient of the part of cost function associated with  $V_r$  observations is much smaller than that associated with reflectivity observations, making the assimilation of  $V_r$  data ineffective. When  $V_r$  data are assimilated in a separate pass after the  $Z$  observations are assimilated (i.e., with the VrPass treatment), the radial velocity RMSI is significantly reduced; this is because the cost function minimization is not affected by the gradient of  $Z$ -related cost function. Interestingly, when CVlogq is used, the problem does not occur even when both  $V_r$  and  $Z$  data are assimilated together. This suggests that CVlogq is a preferred choice, since assimilating different data in different passes creates new problems (e.g., with respect to the background error estimation), especially

when the analysis system is multivariate and/or when more observation types are included.

The composite reflectivity and horizontal wind vectors at 2.5 km height, and reflectivity and wind vectors in vertical cross sections through the maximum reflectivity of the right-moving and left-moving storms at 85 min from truth simulation, for forecast background, and for analyses from single-time analysis experiments CVq-qLim, CVq-ZeLim-VrP, CVlogq, and CVlogq-X<sub>b</sub>SL are shown in Fig. 7. In the truth simulation, the maximum reflectivity of left and right movers are over 65 dBZ (Fig. 7a). The main right moving cell is deep and the reflectivity shows bounded weak echo region (BWER) and an echo overhang in its vertical cross section (Fig. 7b) while the left mover does not exhibit such features. Compared to the truth, the reflectivity in the forecast background is weaker (less than 50 dBZ), especially for the left mover (Fig. 7d). In the background reflectivity of the main cell (Fig. 7e), there is no BWER or overhang echo. For the left-mover, there is not even clear updraft and the height of the echo top is at 9 km instead of the 12 km in truth simulation (Fig. 7f).

After assimilating radar data, the analyzed reflectivity fields of CVq-qLim (Fig. 7g) and CVq-ZeLim (Fig. 7j) are significantly improved over the background except for the central region of the main cell (Figs. 7h,k). As discussed earlier, it is a known problem with direct variational assimilation of reflectivity using CVq that very large gradients of cost function can be produced in areas of small  $Z_e$ , preventing efficient minimization convergence in areas of large  $Z_e$  (because gradients are relatively small there and variational adjustments are small and ineffective). Imposing lower limits on  $q$  or  $Z_e$  helps, but does not completely eliminates this problem.

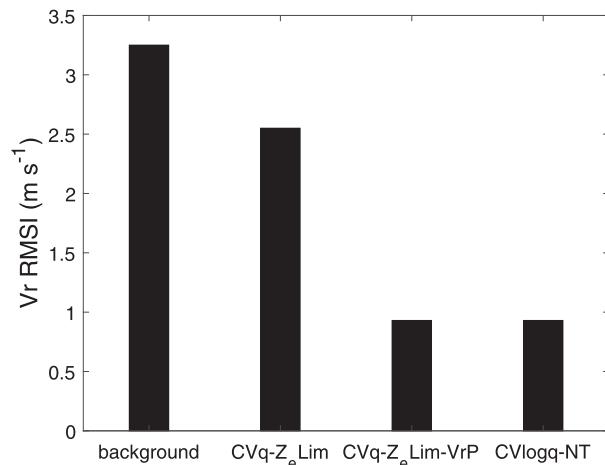


FIG. 6. The radial velocity RMSIs for forecast background, and the analyses of single-time analysis experiments CVq-ZeLim, CVq-ZeLim-VrP, and CVlogq-NT.

In CVq-qLim, some spurious reflectivity is present in expected clear air regions (Fig. 7g) because the gradient is set to zero when mixing ratio is no larger than the threshold in Eq. (15), rendering the assimilation of clear-air reflectivity data ineffective there. However, the ZeLim treatment in CVq-ZeLim does not have such problem (Fig. 7j).

For  $V_r$  data assimilated together with  $Z$  data, neither qLim nor ZeLim treatment helps improve wind field analysis. To more clearly see the impact of different control variables and treatments on wind analysis, the composite reflectivity and wind vectors are shown in Fig. 8 for a zoomed in region over the left moving cell. The wind fields from CVq-qLim and CVq-ZeLim (Figs. 8c,d) are hardly changed from the background (Fig. 8b). After applying VrPass treatment (i.e., when the  $V_r$  data are analyzed in a separate pass from the  $Z$  data), the horizontal winds in Fig. 8e are much closer to the true simulation (Fig. 8a), and the differences are most clear in the reflectivity core region of the left mover, where the background winds (Fig. 8b) are mostly from the south while those in truth (Fig. 8a) and CVq-ZeLim-VrP analysis (Fig. 8e) are mostly from the east. Since the vertical component in  $V_r$  is very small, the vertical velocity is not much changed by the  $V_r$  assimilation directly so the differences are less obvious in the vertical cross sections (Figs. 7n,o).

When using CVlogq, the intensity and location of the reflectivity cores are now more like the truth (Figs. 7p-r) than CVq-qLim or CVq-ZeLim; with CVlogq, the 3DVar is not affected by the extremely large gradient of the cost function as discussed in section 3a. For the same reason, even without the VrPass treatment, the wind analysis of CVlogq-NT (Fig. 8f) is still as good as

CVq-ZeLim-VrP (Fig. 8e). However, Fig. 7p does show that some spurious reflectivity analysis increments appear between the main cell and the left-mover. In addition, the reflectivity in the upper part of the main cell (Fig. 7q) and the left-mover (Fig. 7r) are overestimated and underestimated, respectively. Because there are fewer radar data covering the upper part of the storm, the analysis mainly depends on the upward spread of lower-level analysis increments. However, due to the logarithmic transformation from  $\log q$  to  $q$  increments in the presence of an inhomogeneous background, improper spreading of the analysis increment can occur as discussed in section 4b. After using the  $X_b$ SL treatment, the problems with overestimated or underestimated analysis increments are greatly alleviated, as shown in Figs. 7s-u. The spurious reflectivity between the right and left movers is now absent (Fig. 7s) and the reflectivity analyses at the upper levels are also improved (Figs. 7t,u).

#### b. Results of analyses and forecasts during the DA cycles

For easiness to see the direct effect of different treatments in the 3DVar, the results of single-time analyses are presented earlier. In practice, cycled DA is needed for better results, especially for 3DVar radar DA which does not include cross covariance in the background error. Multiple DA cycles are needed to propagate information in reflectivity and  $V_r$  observations to other state variables. To evaluate the quality of analyses and forecasts, we plot the RMSIs of the analyses and forecasts during the 60-long DA period with 5-min cycles in Fig. 9. The final analyses and 30- and 60-min forecasts in terms of reflectivity and wind vectors from different experiments are plotted in Fig. 10. The single-time analysis results in the previous section show that the ZeLim treatment is better than the qLim treatment, in terms of both RMSI and reflectivity analysis. Since both treatments aim to avoid the very large gradient of cost function, only the ZeLim treatment is kept for CVq in the following cycled DA and forecast experiments. Therefore, results of experiments CVq-ZeLim, CVq-ZeLim-VrP, CVlogq-NT, and CVlogq- $X_b$ SL are presented.

The RMSIs of analyses and forecasts during the DA cycles are calculated in the area where the observed reflectivity is greater than 15 dBZ (Fig. 9). Without the VrPass treatment, the  $V_r$  analysis RMSIs of CVq-ZeLim are clearly larger than those of other experiments (Fig. 9a). With the VrPass treatment, the  $V_r$  analysis RMSIs of CVq-ZeLim-VrP are much lower, and are very similar to those of CVlogq experiments. For reflectivity, the CVlogq-NT analyses are always the worst

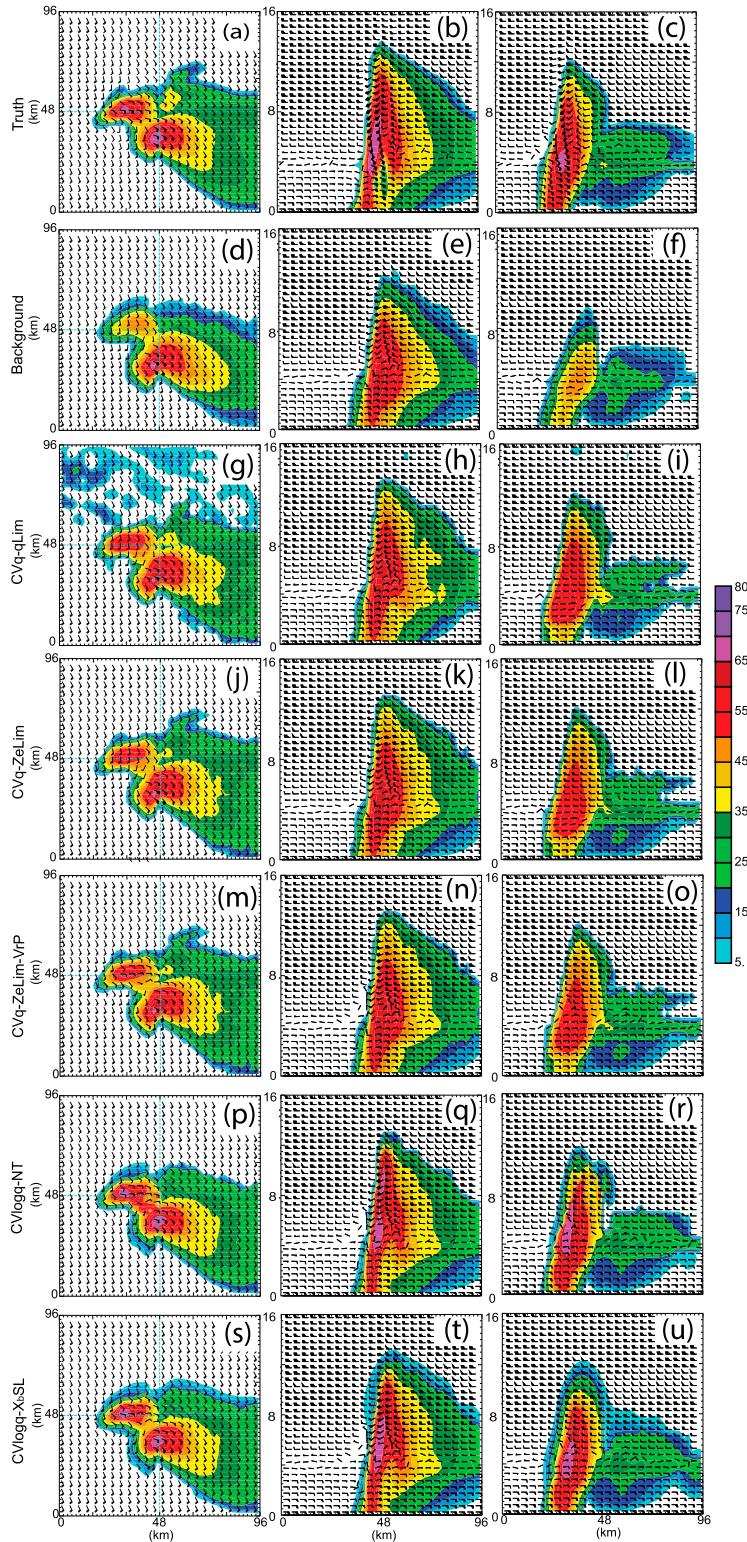


FIG. 7. (left) The composite reflectivity (shading) and wind vectors at 2.5 km AGL, and reflectivity and wind vectors in vertical cross sections through the maximum reflectivity of (middle) the main right-mover storm and (right) the left mover at 85 min from (a)–(c) truth simulation, (d)–(f) forecast background, and the analyses from single-time analysis experiments (g)–(i) CVq-qLim, (j)–(l) CVq-ZeLim, (m)–(o) CVq-ZeLim-VrP, (p)–(r) CVlogq, and (s)–(u) CVlogq-X<sub>b</sub>SL. The full ranges of model domain are plotted, in the x and y direction from 0 to 98 km, and in the vertical from surface to 16 km height.

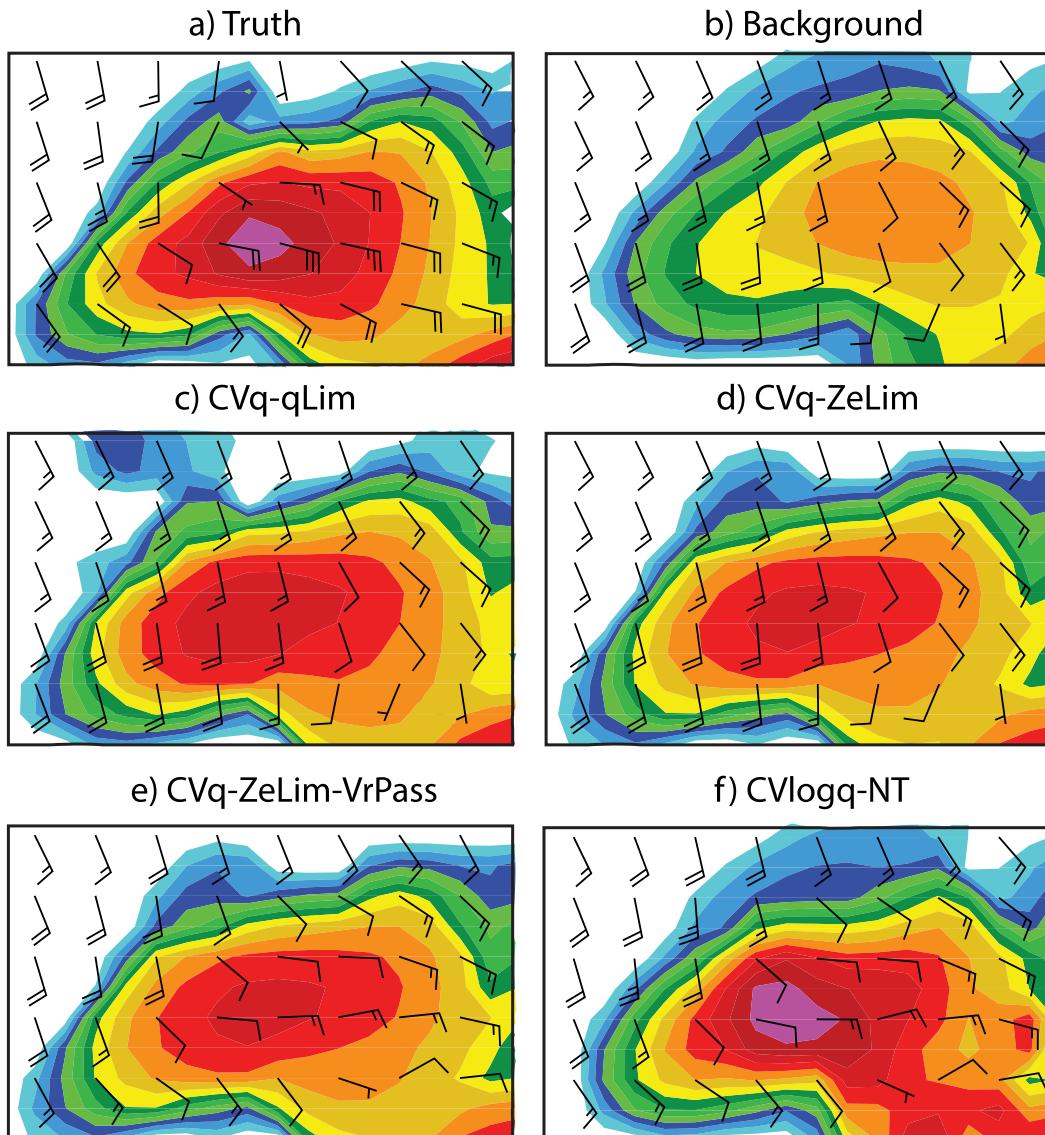


FIG. 8. As in the first column of Fig. 7, but only showing the left mover. Experiment CVlogq- $X_b$ SL is not included because its treatment does not affect wind analysis, which is the focus of this figure.

during the cycles, likely due to the improper increment spread issue, while the RMSIs of CVlogq- $X_b$ SL are clearly reduced and are the lowest among all experiments during the intermediate cycles. The forecast RMSIs of CVlogq-NT and CVlogq- $X_b$ SL, are much closer, however, within the cycles, suggesting that some of the discontinuities within the analyzed hydrometeor fields in CVlogq-NT may not be that harmful to forecasts as small-scale structures can be smoothed by the model time integration. The fact that the wind fields are analyzed the same way in the two experiments (Fig. 9a) also can help reduce the forecast differences between the two experiments. The RMSIs of CVq-ZeLim and CVq-ZeLim-VrP experiments are

generally between those of two CVlogq experiments for reflectivity (Fig. 9b).

The wind and composite reflectivity analyses at the end of DA cycles are shown in the first column of Fig. 10; they are very similar to the results of the corresponding single-time analyses shown in Fig. 7. Among them, there is some sign of spurious echo between the two main cells (Fig. 10j) in CVlogq-NT as in the single-time analysis case, and the maximum reflectivity at the core of the two cells are slightly overestimated. In general, the shape and intensity of the two cells are better analyzed in the CVlogq experiments (Figs. 10j,m) are better than those in the CVq experiments (Fig. 10d), compared to the truth (Fig. 10a).

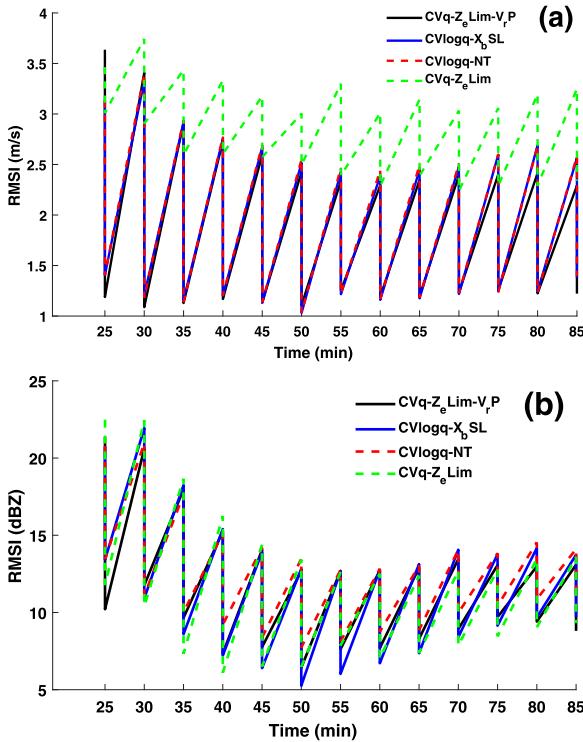


FIG. 9. RMSIs of (a) radial velocity and (b) reflectivity forecasts (higher values in the sawtooth shaped curves) and analyses (lower values) during 1-h-long DA cycles of CVq-ZeLim (green dashed line), CVq-ZeLim-VrP (black solid line), CVlogq-NT (red dashed line), and CVlogq-X<sub>b</sub>SL (blue solid line).

*c. Results of 1-h-long forecasts*

The 1-h-long free forecasts are produced from the final analyses at 85 min from various cycled DA experiments to see how the different treatments affect forecasts. The 30- and 60-min forecasts are shown in the second and third columns of Fig. 10. The intensity, structure and location of the right mover are reproduced quite well in all experiments throughout the hour; there are, however, large differences among the experiments for the prediction of the left mover. In the truth simulation, the left mover underwent another splitting between 115 and 145 min (Figs. 10b,c), which makes accurate prediction more difficult.

Without the VrPass treatment in CVq-ZeLim, because the  $V_r$  information is ineffectively assimilated and the circulation of the left-mover is not well analyzed, the left-mover basically dissipated by 30 min of forecast (Fig. 10e); by 60 min, a new left-mover developed at the original location (Fig. 10f). When using VrPass treatment, the prediction of left-mover in CVq-ZeLim-VrP is much improved (Figs. 10h,j), and appears comparable to those of CVlogq-NT and CVlogq-X<sub>b</sub>SL (Figs. 10k,j,n,o). While the X<sub>b</sub>SL treatment can greatly reduce the improper

analysis increment spreading in CVlogq-NT as shown in the previous section, the reflectivity forecast differences between CVlogq-NT and CVlogq-X<sub>b</sub>SL become very small after 30 min of forecast. The discontinuous analysis increments in CVlogq-NT may have been smoothed by the model integration so that the differences in the forecasts become small.

Figure 11 shows the RMSIs of  $V_r$  and Z forecasts (in areas of observed reflectivity exceeding 15 dBZ) at 5-min intervals starting from the final cycled analyses of CVq-ZeLim, CVq-ZeLim-VrP, CVlogq-NT, and CVlogq-X<sub>b</sub>SL. The CVq-ZeLim-VrP has the most accurate initial condition in terms of both  $V_r$  and Z, and it produces the best forecasts (Fig. 11). The reflectivity RMSI of CVq-ZeLim at the analysis time is slightly smaller than those of CVlogq-NT and CVlogq-X<sub>b</sub>SL but its analysis RMSI for  $V_r$  is much larger (Fig. 11a), leading to rapid increase in reflectivity forecast error (Fig. 11b) and its reflectivity forecast becomes the worst after 10 min while its  $V_r$  forecast is always the worse (Fig. 11a). The reflectivity RMSI of CVlogq-NT at analysis time is somewhat larger than that of CVlogq-X<sub>b</sub>SL and others but becomes very similar to or slightly smaller than that of CVlogq-X<sub>b</sub>SL after 20 min. Overall, the forecasts from the two CVlogq experiments are very similar, while that of CVq-ZeLim is clearly the worst. By analyzing  $V_r$  data in a second pass in CVq-ZeLim-VrP, the final analysis errors are slightly smaller than the two CVlogq experiments and such advantage is maintained throughout the 1-h-long forecast. About 30 min into the assimilation cycles, CVlogq-X<sub>b</sub>SL produces the lowest analysis errors though (Fig. 9b). Considering additional issues associated with the use of multiple passes, CVlogq-X<sub>b</sub>SL appears to be the preferred choice.

**7. Summary and conclusions**

In this study, 3DVar is used to directly assimilate radar radial velocity and reflectivity data for a supercell storm. A complex reflectivity observation operator that considers both liquid and ice phases is used. Several issues arise that are related to the high nonlinearity of the reflectivity observation operator. Special treatments are proposed to deal with the issues. Their performances are examined through OSSEs. For the OSSEs, radial velocity and radar reflectivity data are simulated from a single Doppler radar, sampling a simulated supercell storm. The radar data are assimilated using 3DVar with different treatments, at 5-min cycles over a 1-h period; the quality of analyses and forecasts during the assimilation cycles, and of 1-h-long forecasts are compared. To better examine the direct effects of different treatments, single-time analyses are also performed with the same forecast background.

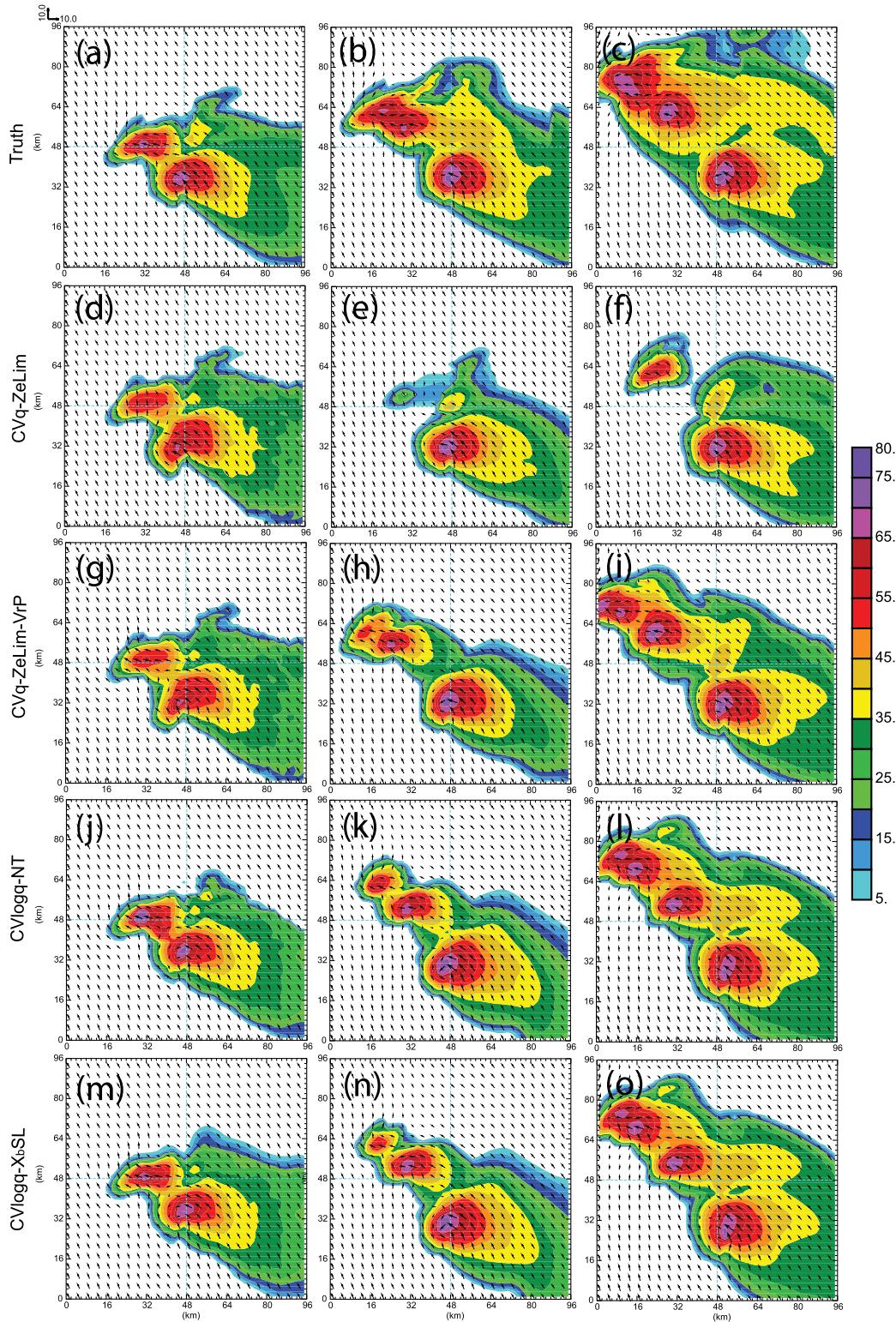


FIG. 10. The wind vectors at 2.5 km AGL and composite reflectivity (shading) at (left) 85, (middle) 115, and (right) 145 min of truth simulation time, for (a)–(c) the truth simulation, and (left) the analysis and (middle),(right) forecasts of (d)–(f) CVq-ZeLim, (g)–(i) CVq-ZeLim-VrP, (j)–(l) CVlogq-NT, and (m)–(o) CVlogq-X<sub>b</sub>SL.

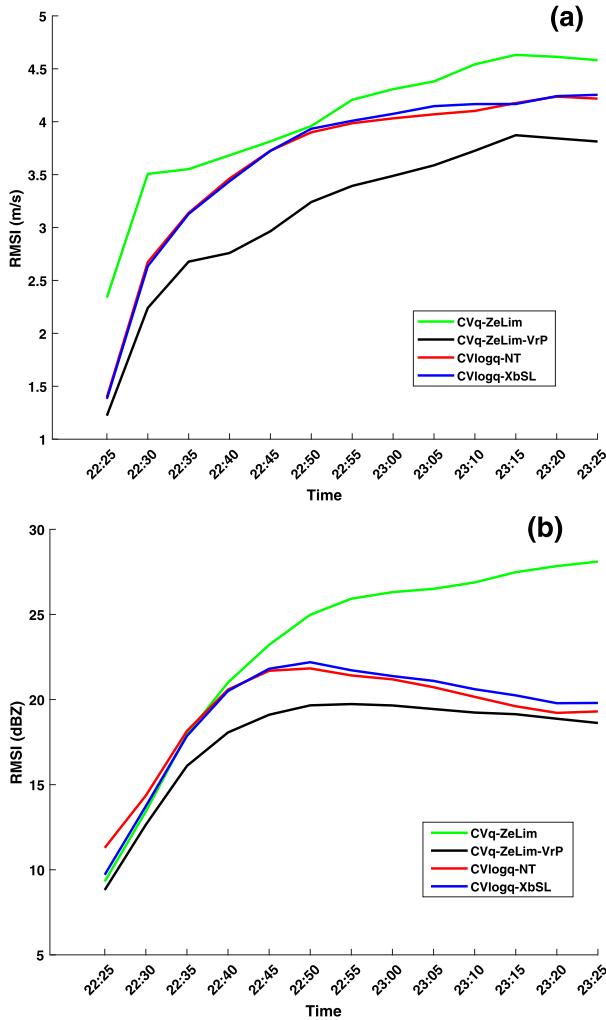


FIG. 11. RMSIs of 1-h forecasts for (a)  $V_r$  and (b)  $Z$  from the cycled analyses of CVq-ZeLim (green line), CVq-ZeLim-VrP (black line), CVlogq-NT (red line), and CVlogq-XbSL (blue line).

The issues, related treatments, and experimental results are summarized as follows:

1) When using the mixing ratios of hydrometeors as the control variables (denoted CVq), 3DVar suffers from extremely large gradient of the cost function due to the existence of very small background reflectivity, making the assimilation of reflectivity in many regions, especially within the core regions of convective cells, ineffective. To avoid such extremely large gradient of the cost function, a lower limit is imposed on the hydrometeor mixing ratios themselves (denoted treatment qLim). An alternative treatment is imposed on equivalent reflectivity factor in the reflectivity observation operator (denoted treatment ZeLim). Results show that both qLim and ZeLim treatments are effective in avoiding the

extremely large gradient of cost function and in improving the reflectivity analysis. However, with the qLim treatment, observations in regions where background hydrometeor mixing ratios are below the lower limit cannot be effectively assimilated because the gradient of the cost function corresponding to the observations becomes zero. Normally, the lower limits are chosen to be very small values, and their optimal values are found through experiments within this study. Experiments show that the ZeLim treatment performs better than qLim in producing analyses and forecasts that better fit observations.

- 2) When reflectivity and radial velocity are assimilated together using CVq, the assimilation of radial velocity is ineffective, because the gradient of the part of cost function corresponding to radial velocity data is much smaller than that corresponding to reflectivity data. Little adjustment is made to the wind fields by the radial velocity data. To be able to effectively assimilate radial velocity data, a separate analysis pass is introduced so that reflectivity and radial velocity data are analyzed in different 3DVar passes (denoted treatment VrPass). Results show that the radial velocity analysis is greatly improved with the VrPass treatment. The use of separate passes does include extra computational cost though, and can create theoretical problems if the background errors are not properly updated in subsequent passes.
- 3) An alternative way to alleviate the problem caused by extremely large gradient of the cost function is to use the logarithmic hydrometeor mixing ratios (denoted CVlogq) instead as the control variables. With CVlogq, the relationship between the reflectivity in dBZ and hydrometeor mixing ratios is much more linear. The magnitude of the gradient of cost function is greatly reduced compared to using CVq, so that radial velocity and reflectivity observations can be effectively assimilated together and no lower limit needs to be imposed on either  $q$  or equivalent reflectivity. In addition, the convergence of cost function minimization is much faster with CVlogq than with CVq.
- 4) However, other issues exist with CVlogq. The analysis increments of mixing ratios and hence reflectivity in certain regions of the storm can be over or underestimated. This problem is investigated through single-observation experiments. When the analysis background is inhomogeneous or has large spatial gradient, the analysis increment of hydrometeor fields can be improperly spread spatially due to the nonlinear transformation between increments of  $q$  and  $\log q$ , which leads to the observed

over or underestimation. To address this problem, a smoothing function and a lower limit are applied to the background hydrometeor mixing ratios when transforming the hydrometeor mixing ratio increments back to mixing ratio increments (denoted treatment  $X_b$ SL). This treatment is found to be able to suppress most of the spurious spreading of analysis increments.

- 5) At the end of the 1-h-long DA cycles, the analysis errors of both reflectivity and radial velocity are the lowest in the experiment using CVq with ZeLim and VrPass treatments, while the analysis errors in the middle of the DA cycles are smallest in the experiment using CVlogq with  $X_b$ SL treatment.
- 6) For the ensuing forecasts after 1-h-long DA cycles, the experiment using CVq with ZeLim treatment, with and without the VrPass treatment yield the best and worst forecasts, respectively, in terms of the errors in reflectivity and radial velocity. The forecasts of the experiments using CVlogq as control variables with and without  $X_b$ SL treatment are similar to that using CVq with ZeLim and VrPass treatments, especially in terms of reflectivity forecast error. Given the desire of using a single pass to assimilate all observations, and the much faster convergence speed using CVlogq, the use of logarithmic hydrometeor mixing ratios as control variables together with the proposed  $X_b$ SL treatments is recommended for directly assimilating radar reflectivity and radial velocity data using 3DVar.

In summary, several treatments are proposed and found to be helpful to dealing with problems arising from direct radar DA using 3DVar method. Although these treatments are tested here only for a 3DVar system, it is expected to benefit hybrid EnVar DA that share the same problems mentioned above when static background error covariance is included. In future studies, the proposed treatments will be applied to a hybrid EnVar DA system, such as that described in Kong et al. (2018), to further examine their impacts. We note here that the proposed treatments in this study are based on a single-moment microphysics scheme, for which the intercept parameters of particle size distributions are fixed. In reality, a small mixing ratio does not necessarily correspond to small reflectivity. In future studies, we will investigate the issues of nonlinear operator and necessary treatments when a multimoment microphysics scheme is used, where the reflectivity observation operator will contain both mixing ratio and number concentration of hydrometeors.

*Acknowledgments.* This research was primarily supported by the NOAA Warn-on-Forecast (WoF) Grant NA160AR4320115. The second author was also supported by NSF Grants AGS 0941491 and AGS-1046171.

Computational resources of the Oklahoma Supercomputing Center for Research and Education (OSCER), and the NSF Xsede Supercomputing Centers were used.

## 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, <https://doi.org/10.1175/2008MWR2691.1>.
- Albers, S. C., J. A. McGinley, D. A. Birkenheuer, and J. R. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analysis of clouds, precipitation and temperature. *Wea. Forecasting*, **11**, 273–287, [https://doi.org/10.1175/1520-0434\(1996\)011<0273:TLAAPS>2.0.CO;2](https://doi.org/10.1175/1520-0434(1996)011<0273:TLAAPS>2.0.CO;2).
- Bouttier, F., and P. Courtier, 1999: Data assimilation concepts and methods: Meteorological Training Course Lecture Series. European Centre for Medium-Range Weather Prediction, 58 pp., <https://www.ecmwf.int/node/16928>.
- Carley, J., 2012: Hybrid ensemble-3DVar radar data assimilation for the short-term prediction of convective storms. Ph.D. dissertation, Purdue University, 205 pp.
- Dawson, D. T., II, and M. Xue, 2006: Numerical forecasts of the 15–16 June 2002 southern plains mesoscale convective system: Impact of mesoscale data and cloud analysis. *Mon. Wea. Rev.*, **134**, 1607–1629, <https://doi.org/10.1175/MWR3141.1>.
- Dee, D. P., and A. M. Da Silva, 2003: The choice of variable for atmospheric moisture analysis. *Mon. Wea. Rev.*, **131**, 155–171, [https://doi.org/10.1175/1520-0493\(2003\)131<0155:TCOVFA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0155:TCOVFA>2.0.CO;2).
- Fletcher, S. J., and M. Zupanski, 2006: A data assimilation method for log-normally distributed observational errors. *Quart. J. Roy. Meteor. Soc.*, **132**, 2505–2519, <https://doi.org/10.1256/qj.05.222>.
- Gao, J., 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, <https://doi.org/10.1175/JAS-D-11-0162.1>.
- , M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. *J. Atmos. Oceanic Technol.*, **21**, 457–469, [https://doi.org/10.1175/1520-0426\(2004\)021<0457:ATVDAM>2.0.CO;2](https://doi.org/10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2).
- Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter—3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919, [https://doi.org/10.1175/1520-0493\(2000\)128<2905:AHEKFV>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2).
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, [https://doi.org/10.1175/1520-0493\(1998\)126<0796:DAUAEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2).
- , —, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126–2143, <https://doi.org/10.1175/2008MWR2737.1>.
- Hu, M., M. Xue, and K. Brewster, 2006a: 3DVAR and cloud analysis with WSR-88D Level-II data for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698, <https://doi.org/10.1175/MWR3092.1>.
- , —, J. Gao, and K. Brewster, 2006b: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: Impact

- of radial velocity analysis via 3DVAR. *Mon. Wea. Rev.*, **134**, 699–721, <https://doi.org/10.1175/MWR3093.1>.
- , S. G. Benjamin, T. T. Ladwig, D. C. Dowell, S. S. Weygandt, C. R. Alexander, and J. S. Whitaker, 2017: GSI three-dimensional ensemble–variational hybrid data assimilation using a global ensemble for the regional Rapid Refresh model. *Mon. Wea. Rev.*, **145**, 4205–4225, <https://doi.org/10.1175/MWR-D-16-0418.1>.
- Huang, X.-Y., 2000: Variational analysis using spatial filters. *Mon. Wea. Rev.*, **128**, 2588–2600, [https://doi.org/10.1175/1520-0493\(2000\)128<2588:VAUSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2588:VAUSF>2.0.CO;2).
- Ingleby, N. B., A. C. Lorenc, K. Ngan, F. Rawlins, and D. R. Jackson, 2013: Improved variational analyses using a nonlinear humidity control variable. *Quart. J. Roy. Meteor. Soc.*, **139**, 1875–1887, <https://doi.org/10.1002/qj.2073>.
- Jiang, H. L., and Coauthors, 2015: Real-time applications of the variational version of the Local Analysis and Prediction System (vLAPS). *Bull. Amer. Meteor. Soc.*, **96**, 2045–2057, <https://doi.org/10.1175/BAMS-D-13-00185.1>.
- Jung, Y., M. Xue, G. Zhang, and J. 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, <https://doi.org/10.1175/2007MWR2288.1>.
- Kain, J. S., and Coauthors, 2010: Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting–research environment. *Wea. Forecasting*, **25**, 1510–1521, <https://doi.org/10.1175/2010WAF2222405.1>.
- Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 341 pp.
- Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4D-EnVar and hybrid variants. *Mon. Wea. Rev.*, **143**, 452–470, <https://doi.org/10.1175/MWR-D-13-00350.1>.
- Kong, R., M. Xue, and C. S. Liu, 2018: Development of a hybrid En3DVar data assimilation system and comparisons with 3DVar and EnKF for radar data assimilation with observing system simulation experiments. *Mon. Wea. Rev.*, **146**, 175–198, <https://doi.org/10.1175/MWR-D-17-0164.1>.
- Li, Y., X. Wang, and M. Xue, 2012: Assimilation of radar radial velocity data with the WRF ensemble-3DVAR hybrid system for the prediction of Hurricane Ike (2008). *Mon. Wea. Rev.*, **140**, 3507–3524, <https://doi.org/10.1175/MWR-D-12-00043.1>.
- Lilly, D. K., 1990: Numerical prediction of thunderstorms—Has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779–798, <https://doi.org/10.1002/QJ.49711649402>.
- 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, [https://doi.org/10.1175/1520-0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2).
- Liu, C. S., M. Xue, and R. Kong, 2019: Direct assimilation of radar reflectivity data using 3DVAR: Treatment of hydrometeor background errors and OSSE tests. *Mon. Wea. Rev.*, **147**, 17–29, <https://doi.org/10.1175/MWR-D-18-0033.1>.
- Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203, <https://doi.org/10.1256/qj.02.132>.
- Pan, Y., K. Zhu, M. Xue, X. Wang, M. Hu, S. G. Benjamin, S. S. Weygandt, and J. S. Whitaker, 2014: A GSI-based coupled EnSRF-En3DVar hybrid data assimilation system for the operational Rapid Refresh model: Tests at a reduced resolution. *Mon. Wea. Rev.*, **142**, 3756–3780, <https://doi.org/10.1175/MWR-D-13-00242.1>.
- Ray, P., B. Johnson, K. Johnson, J. Bradberry, J. Stephens, K. Wagner, R. Wilhelmson, and J. Klemp, 1981: The morphology of several tornadic storms on 20 May 1977. *J. Atmos. Sci.*, **38**, 1643–1663, [https://doi.org/10.1175/1520-0469\(1981\)038<1643:TMOSTS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<1643:TMOSTS>2.0.CO;2).
- Schenkman, A., M. Xue, A. Shapiro, K. Brewster, and J. Gao, 2011: Impact of CASA radar and Oklahoma mesonet data assimilation on the analysis and prediction of tornadic mesovortices in a MCS. *Mon. Wea. Rev.*, **139**, 3422–3445, <https://doi.org/10.1175/MWR-D-10-05051.1>.
- Stensrud, D. J., and Coauthors, 2013: Progress and challenges with warn-on-forecast. *Atmos. Res.*, **123**, 2–16, <https://doi.org/10.1016/j.atmosres.2012.04.004>.
- Sun, J., 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, [https://doi.org/10.1175/1520-0469\(1997\)054<1642:DAMRFD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<1642:DAMRFD>2.0.CO;2).
- , and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, <https://doi.org/10.1175/BAMS-D-11-00263.1>.
- 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, <https://doi.org/10.1175/MWR2898.1>.
- Wang, H. L., J. Z. Sun, S. Y. 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, <https://doi.org/10.1175/JAMC-D-12-0120.1>.
- , —, X. Zhang, X. Y. 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, <https://doi.org/10.1175/MWR-D-12-00168.1>.
- Wang, Y. M., and X. G. Wang, 2017: 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 tornadic supercell. *Mon. Wea. Rev.*, **145**, 1447–1471, <https://doi.org/10.1175/MWR-D-16-0231.1>.
- Weygandt, S. S., S. G. Benjamin, 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, [https://ams.confex.com/ams/88Annual/techprogram/paper\\_134081.htm](https://ams.confex.com/ams/88Annual/techprogram/paper_134081.htm).
- Wu, W.-S., D. F. Parrish, E. Rogers, and Y. Lin, 2017: Regional ensemble–variational data assimilation using global ensemble forecasts. *Wea. Forecasting*, **32**, 83–96, <https://doi.org/10.1175/WAF-D-16-0045.1>.
- Xiao, Q., Y.-H. Kuo, J. Sun, W.-C. Lee, D. M. Barker, and E. Lim, 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, <https://doi.org/10.1175/JAM2439.1>.
- 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, <https://doi.org/10.1007/s00703-001-0595-6>.

- , M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Oceanic Technol.*, **23**, 46–66, <https://doi.org/10.1175/JTECH1835.1>.
- , Y. Jung, and G. Zhang, 2010: State estimation of convective storms with a two-moment microphysics scheme and an ensemble Kalman filter: Experiments with simulated radar data. *Quart. J. Roy. Meteor. Soc.*, **136**, 685–700, <https://doi.org/10.1002/QJ.593>.
- , F. Kong, K. A. Brewster, K. W. Thomas, J. Gao, Y. Wang, and K. K. Droegemeier, 2013: Prediction of convective storms at convection-resolving 1 km resolution over continental United States with radar data assimilation: An example case of 26 May 2008 and precipitation forecasts from spring 2009. *Adv. Meteor.*, **2013**, 259052, <https://doi.org/10.1155/2013/259052>.
- Zhang, J., F. Carr, and K. Brewster, 1998: ADAS cloud analysis. Preprints, *12th Conf. on Numerical Weather Prediction*, Phoenix, AZ, Amer. Meteor. Soc., 185–188.
- Zhao, K., and M. Xue, 2009: Assimilation of coastal Doppler radar data with the ARPS 3DVAR and cloud analysis for the prediction of Hurricane Ike (2008). *Geophys. Res. Lett.*, **36**, L12803, <https://doi.org/10.1029/2009GL038658>.