Research Article

The Development of a Hybrid EnKF-3DVAR Algorithm for Storm-Scale Data Assimilation

Jidong Gao,[1,2] **Ming Xue,**[1,3] **and David J. Stensrud**[2]

[1] *Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK 73072, USA*
[2] *NOAA/National Severe Storm Laboratory, National Weather Center 120 David L. Boren Boulevard, Norman, OK 73072, USA*
[3] *School of Meteorology, University of Oklahoma, Norman, OK 73072, USA*

Correspondence should be addressed to Jidong Gao; jidong.gao@noaa.gov

A hybrid 3DVAR-EnKF data assimilation algorithm is developed based on 3DVAR and ensemble Kalman filter (EnKF) programs within the Advanced Regional Prediction System (ARPS). The hybrid algorithm uses the extended alpha control variable approach to combine the static and ensemble-derived flow-dependent forecast error covariances. The hybrid variational analysis is performed using an equal weighting of static and flow-dependent error covariance as derived from ensemble forecasts. The method is first applied to the assimilation of simulated radar data for a supercell storm. Results obtained using 3DVAR (with static covariance entirely), hybrid 3DVAR-EnKF, and the EnKF are compared. When data from a single radar are used, the EnKF method provides the best results for the model dynamic variables, while the hybrid method provides the best results for hydrometeor related variables in term of rms errors. Although storm structures can be established reasonably well using 3DVAR, the rms errors are generally worse than seen from the other two methods. With two radars, the results from 3DVAR are closer to those from EnKF. Our tests indicate that the hybrid scheme can reduce the storm spin-up time because it fits the observations, especially the reflectivity observations, better than the EnKF and the 3DVAR at the beginning of the assimilation cycles.

1. Introduction

The effective assimilation of radar data into a numerical weather prediction (NWP) model requires advanced data assimilation (DA) techniques, such as variational and ensemble Kalman filter methods. A three-dimensional variational (3DVAR) system, which includes a mass continuity equation and other appropriate model equations as weak constraints, has been developed in recent years [1–5]. This system was designed with special considerations for assimilating radar data into a convective-scale nonhydrostatic model—the Advanced Regional Prediction System (ARPS)—and has been used to provide initial conditions for numerous real-time convective-scale data forecasts. These forecasts have been produced since 2008 using grid spacing that varied from 4 to 1 km for domains covering the entire continental United States as part of the NOAA Hazardous Weather Testbed (HWT) Spring Experiments [6, 7]. For the HWT Spring Experiments, Level-II radial velocity and reflectivity data

from over 120 operational Weather Surveillance Radar-1988 Doppler (WSR-88D) radars were analyzed using the 3DVAR system, and ensemble forecasts were produced by adding additional initial condition perturbations to this 3DVAR analysis. The ARPS 3DVAR system has also been used in a large number of real case studies with encouraging results [2, 3, 8, 9]. Barker et al. [10] and Xiao et al. [11] also applied the 3DVAR method to assimilate Doppler radar observations into the Weather Research and Forecasting (WRF) model [12]. The major advantage of the 3DVAR method is its computational efficiency and the ease by which weak constraints can be included. However, the truly flow-dependent background error covariances were not included in either ARPS 3DVAR or WRF 3DVAR systems at that time.

Compared to 3DVAR, the more advanced 4DVAR technique incorporates the full prediction model into the assimilation system and implicitly includes the effects of flow-dependent error covariances through the use of both the

forward and backward models. In recent years, the 4DVAR technique has helped improve global forecasts at several operational NWP centers, including the European Centre for Medium-Range Weather Forecasts, Meteo-France, Meteorological Service of Canada, and Japan Meteorological Agency (JMA) [13]. Research has also focused on storm-scale radar data assimilation using the 4DVAR method by Sun and Crook [14–16]. In these studies, both radial velocity and reflectivity data were assimilated into a convective cloud-resolving model. Despite some encouraging results, 4DVAR for convective-scale applications has been limited to the use of simple microphysics in almost all cases because the strong nonlinearity within sophisticated microphysics schemes makes the minimization process difficult. Honda and Koizumi [17] report difficulties, including slow convergence, when including complex ice microphysics within the inner loop of the 4DVAR system when using a nonhydrostatic model at JMA.

The ensemble Kalman filter (EnKF) is an advanced data assimilation method that shares many of the advantages of 4DVAR. It has gained considerable popularity in recent years in meteorology and oceanography since first proposed by Evensen [18]. For convective storms, very encouraging results have been obtained in recent studies using the ensemble Kalman filter method in analyzing wind, temperature, moisture fields, and even microphysics variables from radar observations of convective storms [19–26]. One of the advantages of the EnKF method over variational methods is that it can explicitly evolve and carry the background error covariances through the assimilation cycles. However, one of the major sources of error with ensemble-based DA is covariance matrix rank deficiency or sampling error as a result of a relatively small ensemble size [27, 28]. This problem can be more severe with storm-scale data assimilation because the degrees of freedom of the system are typically even larger relative to the practical ensemble size. The commonly utilized remedy to the rank deficiency problem is to apply covariance localization by a Schur product as introduced by Houtekamer and Mitchell [27]. This solution, however, prevents the use of distant correlations that are physically meaningful. Further, the modification to the spatial covariances within a cut-off radius by a Schur product also introduces imbalances, and the effect is more substantial when the localization is more restrictive [28]. This problem may be remedied or reduced when using a hybrid 3DVAR and EnKF method.

As discussed above, the 3DVAR method is attractive for convective scale assimilation because of its computational efficiency and the ease by which weak constraints can be added. However, the major shortcoming is that the background error covariances are stationary and isotropic and error covariances related to the model equations cannot be simply defined. In addition, for convective-scale radar data assimilation, only observations of radial velocity and reflectivity are typically measured, while all other state variables have to be "retrieved"; in this case, the flow-dependent background error covariances, such as that derived from a forecast ensemble, are especially important. One way to blend the advanced features of both variational and EnKF methods and to overcome their respective shortcomings is to employ a hybrid ensemble 3DVAR framework. In such a framework, a combination of the static background error covariance and the flow-dependent error covariance derived from an ensemble is used within the variational analysis. For large-scale data assimilation, such an approach was initially demonstrated for a quasigeostrophic system by Hamill and Snyder [29] and further developed by Lorenc [30], Buehner [31], and Zupanski [32] with different formulations. Another relatively new approach estimates the four-dimensional background-error covariances from the ensemble members to produce a 4D analysis with the variational data assimilation approach. In this method, the tangent-linear or adjoint versions of the forecast model are no longer needed. This approach was called the En-4DVar approach [33–35] but was recently renamed as 4DEnVar [36].

Wang et al. [37] showed that the formulations proposed by Hamill and Snyder [29], Lorenc [30], and Buehner [31], though different in implementation and computational cost, are mathematically equivalent. Barker et al. [38], Li et al. [39], and Zhang et al. [40] recently reported the capability of the WRF hybrid system for mesoscale applications. Further studies have demonstrated the potential advantages of the hybrid method over both the pure variational and pure ensemble methods for mesoscale and global applications, especially for small ensemble size [41–44]. However, the application of hybrid methods to convective scale data assimilation has so far been limited. The purpose of this paper is to demonstrate the potential usefulness of the hybrid EnKF-3DVAR method for convective scale data assimilation, especially when assimilating radar data.

The rest of this paper is organized as follows. In Section 2, we introduce the hybrid EnKF-3DVAR system developed in this study. Section 3 describes the DA experiment design. Experiment results and quantitative performance are assessed in Section 4. We conclude in Section 5 with a summary and outlook for future work.

2. The Hybrid EnKF-3DVAR Scheme

In the implementation of the hybrid method for convective scale, the ensemble covariance is incorporated in the variational framework using the extended control variable method [30, 31, 37]. A convenient approach, initially suggested by Buehner [31], is to combine the ensemble-derived and static covariance matrices through the augmentation of state vector, from $\mathbf{v}$ to $(\mathbf{v}, \mathbf{w})$ within the 3DVAR cost function, which can be written as

$$
\begin{aligned}
J = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\left[H\left(\overline{\mathbf{x}}^b + \Delta\mathbf{x}\right) - \mathbf{y}^o \right]^T \\
\times \mathbf{R}^{-1}\left[H\left(\overline{\mathbf{x}}^b + \Delta\mathbf{x}\right) - \mathbf{y}^o \right] + J_c,
\end{aligned}
\tag{1}
$$

where

$$
\Delta\mathbf{x} = \Delta\mathbf{x}_1 + \Delta\mathbf{x}_2 = \beta_1 \mathbf{B}^{1/2}\mathbf{v} + \beta_2 \mathbf{P}^{1/2}\mathbf{w}
\tag{2}
$$

is the analysis increment of state vector $\mathbf{x}$, $\mathbf{B}$ is the static 3DVAR background error covariance matrix, and $\mathbf{P}$ is the covariance matrix derived from an ensemble of forecasts. The

control variable $\mathbf{v}$ is defined in association with $\mathbf{B}$, and $\mathbf{w}$ is the augmented control vector associated with $\mathbf{P}$. The size of $\mathbf{v}$ is the number of analysis variables multiplied by their dimension, and the size of $\mathbf{w}$ is the ensemble size multiplied by the dimension of variables. By using control variables $\mathbf{v}$ and $\mathbf{w}$, instead of $\Delta\mathbf{x}_1$ and $\Delta\mathbf{x}_2$ in (2), the minimization procedure is preconditioned by $\mathbf{B}^{1/2}$ and $\mathbf{P}^{1/2}$, respectively. This technique was first proposed in the context of data assimilation by Derber and Rosati [45]. The definition of $(\mathbf{B})^{1/2}$ is the same as Gao et al. [1]. If no localization is applied to the ensemble covariance, $\mathbf{P}^{1/2}$ is simply a rectangular matrix whose columns are the ensemble perturbation vectors divided by $\sqrt{N-1}$, where N is the ensemble size. The localization of the ensemble covariance in a variational system with preconditioning is discussed in Lorenc [30], Buehner [31], and Wang et al. [37]. The procedure and cost of doing so were also discussed in these papers. For computational efficiency, we also use the recursive filter for covariance localization, as suggested in Wang et al. [41].

In (2), there are two factors β_1 and β_2 that define the weights placed on the static background error covariance and the ensemble covariance. To conserve the total background-error variance, β_1 and β_2 are constrained by

$$\beta_1^2 + \beta_2^2 = 1. \tag{3}$$

A similar constraint was applied in Hamill and Snyder [29]. This approach for combining two covariance matrices to form a hybrid covariance provides flexibility since it allows for different relative contributions from two covariance matrices. When $\beta_1 = 1$, the analysis is back to a 3DVAR analysis scheme, when $\beta_2 = 1$, the analysis is mathematically equivalent to a EnKF scheme, and in between, we have a hybrid scheme that incorporates a mixture of both static and flow-dependent error covariances. When $\beta_2 = 1$, the scheme is essentially a variational formulation of an ensemble-based analysis scheme, and it can be called 3DEnVAR. Though the dimension of the control variables is increased, the form of the background term of the cost function remains unchanged from that of 3DVAR, so that codes from an existing 3DVAR system can readily be utilized [30].

In the current study, the hybrid system will assimilate both radar reflectivity and radial velocity data. Within this system, flow-dependent background-error covariances, in particular cross covariances between microphysical and dynamic variables, will be derived and utilized. The single-resolution version of the EnKF system of Gao and Xue [46] is used for updating the ensemble perturbations in the data assimilation cycles. In Gao and Xue [46], an efficient dual-resolution (DR) data assimilation algorithm was developed based on the ensemble square root Kalman filter method and tested using simulated radar radial velocity data for a supercell storm. Within the algorithm, radar observations were assimilated on both high-resolution and lower-resolution grids using ensemble Kalman filter algorithms and the flow-dependent background error covariance estimated from the lower resolution ensemble. In that paper, the DR method was compared to a standard full-resolution ensemble square root Kalman filter method which is used in this study.

Different from other hybrid systems [40, 41], for this hybrid method, an extra model integration for the length of the analysis cycle is needed to produce a control forecast and analysis cycle. The EnKF analyses are performed to update analysis perturbations for each ensemble member. Then, the cost function (1) is minimized to obtain optimal analyses of control vectors $\mathbf{v}$ and $\mathbf{w}$, and the optimal analysis increment, $\Delta\mathbf{x}$, is derived from (2). The ensemble mean analysis is replaced with the hybrid EnKF-3DVAR analysis. Finally, the initial conditions for the ensemble and one control forecast are obtained. The above steps are repeated for each data assimilation cycle (Figure 1).

3. Model and Experimental Design

3.1. Prediction Model and Truth Simulation for OSSEs. We test our hybrid EnKF-3DVAR algorithm and compare its results with those of 3DVAR and EnKF schemes, using simulated data from a classic supercell storm of May 20, 1977, near Del City, Oklahoma [47]. The ARPS prediction model is used in a 3D cloud model mode, and the prognostic variables include three velocity components u, v, and w, perturbation potential temperature θ', pressure p, and six categories of water substances, that is, water vapor specific humidity q_v, and mixing ratios for cloud water q_c, rainwater q_r, cloud ice q_i, snow q_s, and hail q_h. The microphysical processes are parameterized using the single-moment, three-category ice scheme of Ying Lin et al. [48]. More details on the model can be found in Xue et al. [49, 50].

For our experiments, the model domain is $57 \times 57 \times 16\,\mathrm{km}^3$. The horizontal grid spacing is $1\,\mathrm{km}$, and the mean vertical grid spacing is $500\,\mathrm{m}$. The truth simulation run is initialized from a modified real sounding plus a $4\,\mathrm{K}$ ellipsoidal thermal bubble centered at $x = 48$, $y = 16$, and $z = 1.5\,\mathrm{km}$, with radii of $10\,\mathrm{km}$ in x and y and $1.5\,\mathrm{km}$ in the z direction. Open conditions are used at the lateral boundaries. The length of simulation is 2 hours. A constant wind of $u = 3\,\mathrm{ms}^{-1}$ and $v = 14\,\mathrm{ms}^{-1}$ is subtracted from the observed sounding to keep the primary storm cell near the center of model grid. The evolution of the simulated storms is similar to those documented in Xue et al. [50]. During the truth simulation, the initial convective cell strengthens over the first 30 min. The strength of the cell then decreases over the next 30 min or so, which is associated with the splitting of the cell at around 55 min. The right moving (relative to the storm motion vector which is towards north-northeast) cell tends to dominate the system, and its updraft reaches a peak value of over $40\,\mathrm{ms}^{-1}$ at 90 min. The initial cloud starts to form at about 10 min, and rainwater forms at about 15 min. Ice phase fields appear at about 20 min. A similar truth simulation was also used in Gao et al. [51], Tong and Xue [21], and Gao and Xue [46].

3.2. Simulation of Radar Observations. The simulated radial velocity observations are assumed to be available on the grid points. The simulated radial velocity, v_r, is calculated from

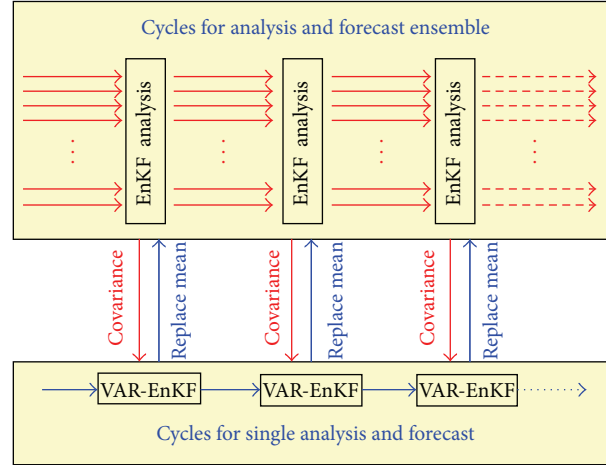$$v_r = u\sin\phi\cos\mu + v\cos\phi\cos\mu + w\sin\mu, \tag{4}$$

FIGURE 1: Illustration of cycle used in a hybrid EnKF-3DVAR analysis scheme.

where μ is the elevation angle ϕ is the azimuth angle of radar beams, and u, v, and w are the model-simulated velocities interpolated to the scalar points of the staggered model grid. Random errors drawn from a normal distribution with zero mean and a standard deviation of $1\,\text{ms}^{-1}$ are added to the simulated data. Since v_r is sampled directly from the model velocity fields, hydrometeor sedimentation is not involved. The ground-based radar is located at the southwest corner of the computational domain, that is, at the origin of the x-y coordinates. The simulated reflectivity observations are calculated based on Smith et al. [52] and Ferrier [53]. For reflectivity, random errors drawn from a normal distribution with zero mean and a standard deviation of 3 dBZ are added to the simulated data. The radial velocity data are assimilated and are only available where the truth reflectivity is greater than zero in the analysis domain. We also use only the data at every other grid point from the 1 km truth simulation grid in horizontal, so that the total data used are one-fourth of total model grid points.

3.3. Design of Assimilation Experiments. We start the initial ensemble forecast at 20 min of the model integration time when the storm cell is well developed. To initialize the ensemble members, random noise is first added to the initially horizontally homogeneous first guess defined using the environmental sounding. A 2D five-point smoother is applied to the resultant fields, similar to a method used by Zupanski et al. [54]. The random noise is sampled from Gaussian distributions with zero mean and standard deviations of $5\,\text{ms}^{-1}$ for u, v, and w and 3 K for potential temperature. These perturbation variances are somewhat larger than those used in Tong and Xue [21], but the standard deviation of the final perturbations is not necessarily larger because of the smoothing. Other variables, including the microphysical variables, are not perturbed at the initial time. The radial and reflectivity observations are calculated and assimilated using a 5 min cycle in all three data assimilation schemes. The first analysis is performed at 20 min, and 20 ensemble members are used. A cut-off radius of 8 km is used in most of our experiments.

We perform two set of experiments. The first group of experiments is performed to compare the performance of three different schemes when observations from a single Doppler radar are used. The second group of experiments will be performed when observations from two Doppler radars are used. For comparison purposes, all three methods (3DVAR, EnKF, and Hybrid EnKF-3DVAR) are performed with 16 data assimilation cycles where each cycle has a 5 min analysis-prediction interval. The total assimilation period is 75 min.

4. Results

4.1. Single Observation Experiment. Figure 2 provides analysis results of a single observation with three model variables, showing that ensemble information can provide flow-dependent estimates of the background-error covariance and that both the EnKF and hybrid 3DVAR-EnKF methods can utilize such information to provide flow-dependent analysis increments. Because mass continuity equation is used as a weak constraint in 3DVAR [1], the 3DVAR method can also provide a kind of flow-dependent anisotropic non-Gaussian type covariance structure for both u component and w component (Figures 2(a) and 2(b)). However, the 3DVAR cannot provide increments for potential temperature (Figure 2(d)), though updated potential temperature can be obtained through a cycled 3DVAR analysis (built up by integration of a convective NWP model, ARPS in this study). The EnKF provides a flow-dependent covariance structure (Figures 2(b), 2(e), and 2(h)), and the hybrid 3DVAR-EnKF provides a covariance structure in between the other two structures. In addition, both EnKF and hybrid 3DVAR-EnKF can provide increments for unobserved variables, such as potential temperature which is not directly related to radial velocity (Figures 2(h) and 2(i)). Because the mass continuity equation is used as a weak constraint in 3DVAR, this actually provides a physical constraint for three components of wind field. Similar to Buehner [31] and to take advantage of both 3DVAR and EnKF methods, 50/50 weightings are chosen in the cost function.
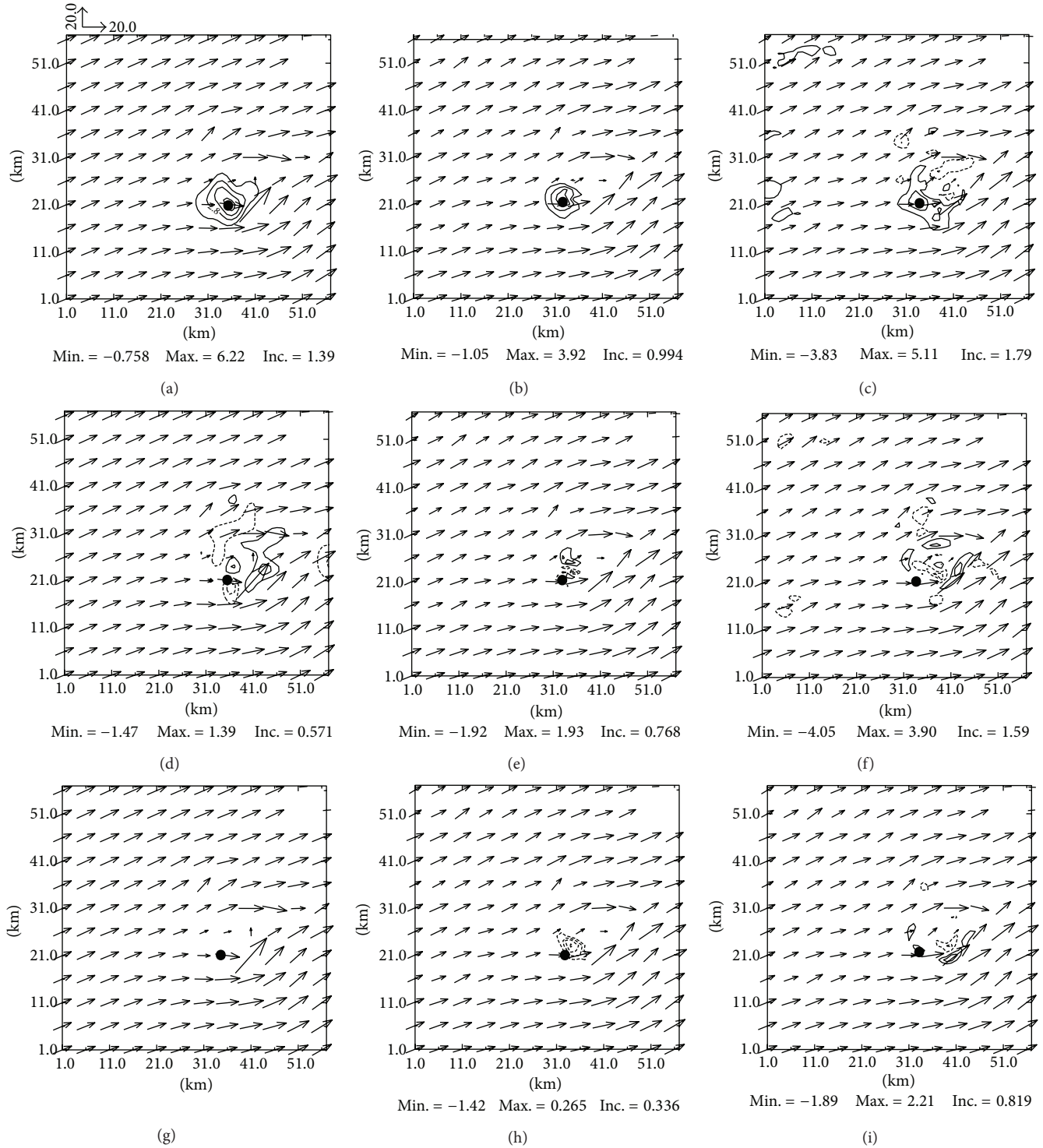
FIGURE 2: Wind vectors: u-component increment by using (a) 3DVAR, (b) EnKF, and (c) hybrid 3DVAR-EnKF, w-component increment by using (d) 3DVAR, (e) EnKF, and (f) hybrid 3DVAR-EnKF, and potential temperature increment by using (g) 3DVAR, (h) EnKF, and (i) hybrid 3DVAR-EnKF by assimilating a single radial velocity at the black dot.

4.2. Experiments with Single Radar. As stated above, the first group of experiments is performed with radial velocity and reflectivity data from a single radar. Figure 3 shows the final assimilation results after 16 assimilation cycles with 5 min prediction-analysis intervals. The low-level flow, reflectivity patterns, and the strength of the cold pool from both EnKF and hybrid EnKF-3DVAR agree very well with the simulated truth (Figure 3(a)) and are better than the result using 3DVAR (Figure 3(b)), although this 3DVAR can also establish the storm structures reasonably well. The most obvious difference
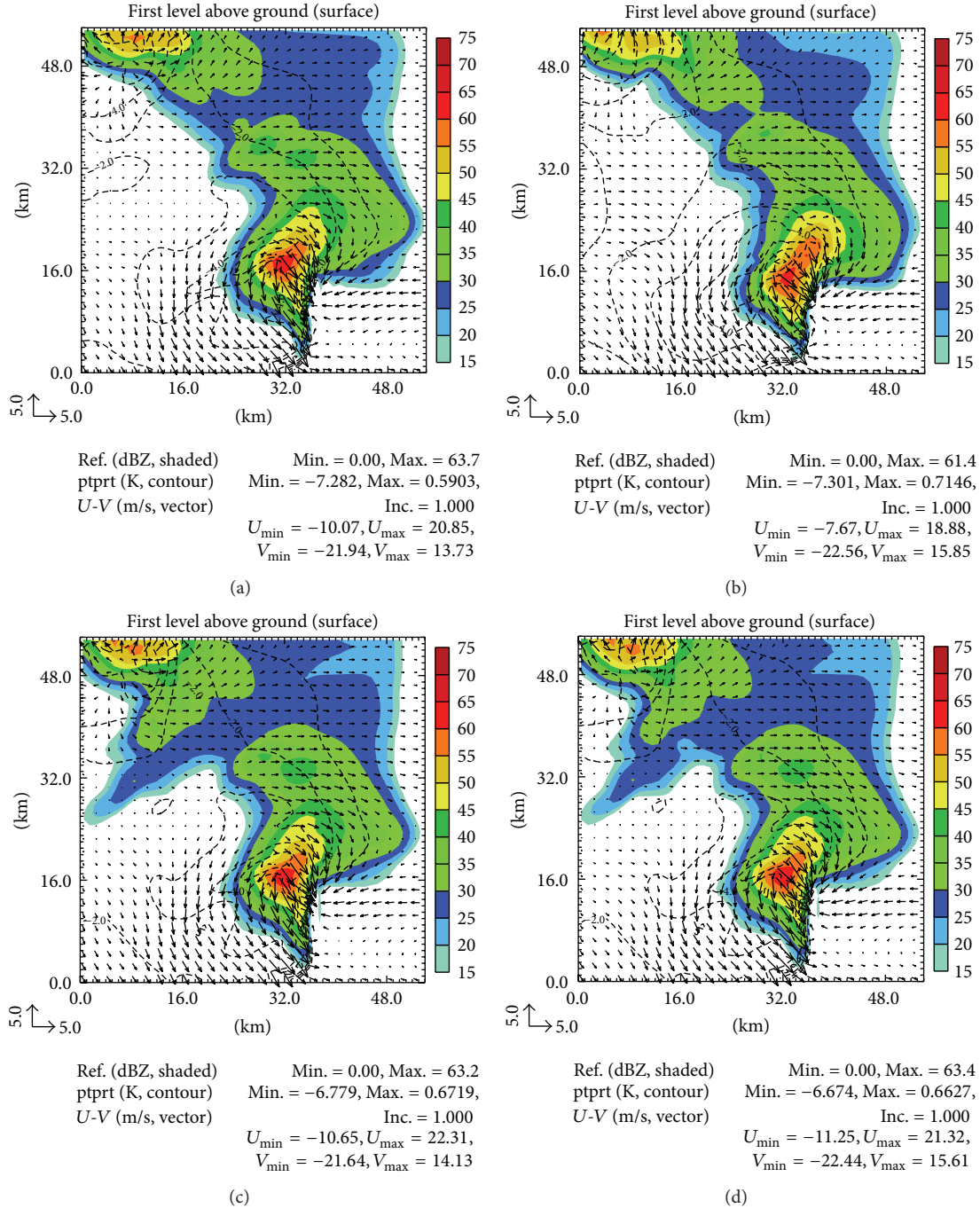
Ref. (dBZ, shaded)   Min. = 0.00, Max. = 63.7
ptprt (K, contour)   Min. = −7.282, Max. = 0.5903,
U-V (m/s, vector)   Inc. = 1.000
$U_{min} = -10.07, U_{max} = 20.85,$
$V_{min} = -21.94, V_{max} = 13.73$

(a)

Ref. (dBZ, shaded)   Min. = 0.00, Max. = 61.4
ptprt (K, contour)   Min. = −7.301, Max. = 0.7146,
U-V (m/s, vector)   Inc. = 1.000
$U_{min} = -7.67, U_{max} = 18.88,$
$V_{min} = -22.56, V_{max} = 15.85$

(b)

Ref. (dBZ, shaded)   Min. = 0.00, Max. = 63.2
ptprt (K, contour)   Min. = −6.779, Max. = 0.6719,
U-V (m/s, vector)   Inc. = 1.000
$U_{min} = -10.65, U_{max} = 22.31,$
$V_{min} = -21.64, V_{max} = 14.13$

(c)

Ref. (dBZ, shaded)   Min. = 0.00, Max. = 63.4
ptprt (K, contour)   Min. = −6.674, Max. = 0.6627,
U-V (m/s, vector)   Inc. = 1.000
$U_{min} = -11.25, U_{max} = 21.32,$
$V_{min} = -22.44, V_{max} = 15.61$

(d)

FIGURE 3: Horizontal winds (vectors; ms^{-1}), perturbation potential temperature (contours at 1-K intervals), and simulated reflectivity (shaded contours; dBZ) at 250 m AGL for (a) the truth simulation; (b) the 3DVAR analysis; (c) the EnKF analysis; and (d) the hybrid EnKF-3DVAR analysis for the single radar experiment. The time shown is at 100 min (the end of data assimilation cycles). Wind vectors are shown every 2 km.

is the reflectivity field in the center of model domain. The area of reflectivity values greater than 55 dBZ is over extended in a peanut-shaped region for 3DVAR. The spread of potential temperature is little bit far to the south-southwest direction in the southwest corner (Figure 3(b)). But the strength of the cold pool in 3DVAR, as indicated by minimum perturbation potential of −7.30°, is closer to the truth simulation (−7.28°) than seen in either EnKF or the hybrid EnKF-3DVAR.

The rms errors of the analyzed fields with data from a single radar are shown in Figure 4. The rms error calculation is limited to the regions where the truth reflectivity exceeds 10 dBZ. Figure 4 shows that the rms errors for model variables u, v, w, θ, and q_v and reflectivity Z (derived from the hydrometeor mixing ratios) generally decrease with the cycles in all three experiments. The errors for 3DVAR decrease more slowly and remain at a higher level at the end of assimilation
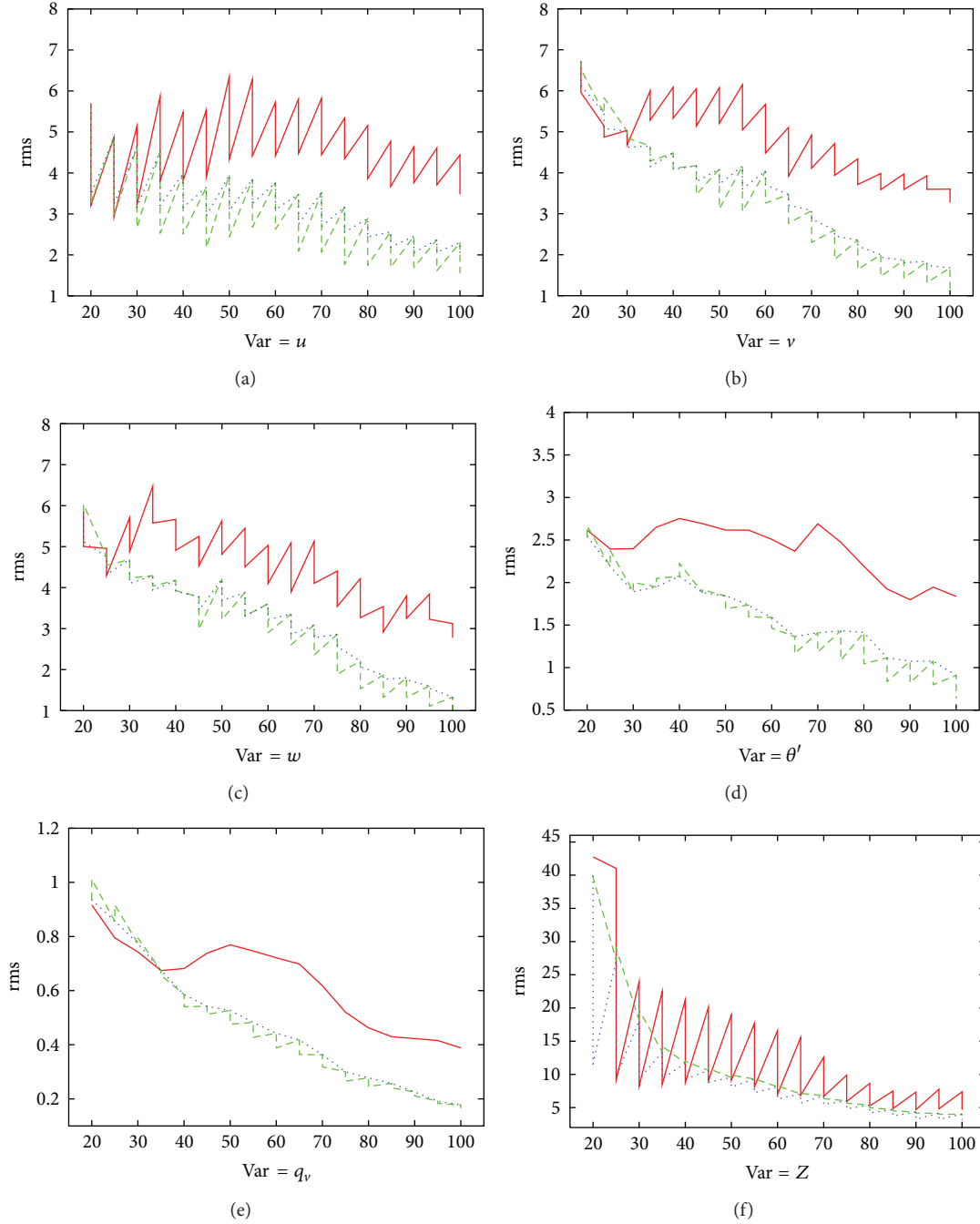
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 4: The rms errors of the analysis and forecast for the 3DVAR, (red) EnKF, (green) hybrid EnKF-3DVAR, (blue) methods averaged over points at which the reflectivity is greater than 10 dBZ for (a) u-wind component, (b) v-wind component, (c) vertical wind speed, (d) potential temperature, (e) water vapor mixing ratio, and (f) reflectivity.

cycles than those for the ensemble based methods for most of model variables. For example, the rms error of w is close to 3 ms^{-1} at 100 min for 3DVAR method, while that in EnKF and hybrid EnKF-3DVAR is close to 1.3 ms^{-1}. The rms errors of q_v for 3DVAR is 0.4 g/kg, and that in °EnKF and hybrid EnKF-3DVAR is below 0.2 k/kg. While these differences are significant, the error levels late in the assimilation period for EnKF and hybrid EnKF-3DVAR are unrealistically low due to the perfect model assumption. For real data cases where model error exists, the analysis errors are likely to be much

larger (see, for example, Dowell et al. [22, 23]). For systems containing discrete intense updrafts, the rms error tends to exaggerate errors because of small spatial displacement and/or structure discrepancies, such as those seen in **Figure 4**. So the results for 3DVAR may still be reasonable. It should be noted that for most of model variables, the performance of EnKF and hybrid methods is very close to each other, with EnKF a little bit better. Interestingly, the differences among the rms errors for Z in different experiments are smallest (**Figure 4(f)**). The rms error of Z is decreased to about

First level above ground (surface)

(a)

Ref. (dBZ, shaded)	Min. = 0.00, Max. = 63.7
ptprt (K, contour)	Min. = −7.282, Max. = 0.5903,
U-V (m/s, vector)	Inc. = 1.000

$U_{min} = -10.07, U_{max} = 20.85,$
$V_{min} = -21.94, V_{max} = 13.73$

First level above ground (surface)

(b)

Ref. (dBZ, shaded)	Min. = 0.00, Max. = 63.7
ptprt (K, contour)	Min. = −6.549, Max. = 0.6395,
U-V (m/s, vector)	Inc. = 1.000

$U_{min} = -11.15, U_{max} = 21.52,$
$V_{min} = -21.55, V_{max} = 14.42$

First level above ground (surface)

(c)

Ref. (dBZ, shaded)	Min. = 0.00, Max. = 63.4
ptprt (K, contour)	Min. = −6.536, **Max.** = 0.7219,
U-V (m/s, vector)	Inc. = 1.000

$U_{min} = -10.06, U_{max} = 21.50,$
$V_{min} = -21.14, V_{max} = 14.39$

First level above ground (surface)

(d)

Ref. (dBZ, shaded)	Min. = 0.00, Max. = 63.8
ptprt (K, contour)	Min. = −6.567, Max. = 0.6563,
U-V (m/s, vector)	Inc. = 1.000

$U_{min} = -10.77, U_{max} = 21.79,$
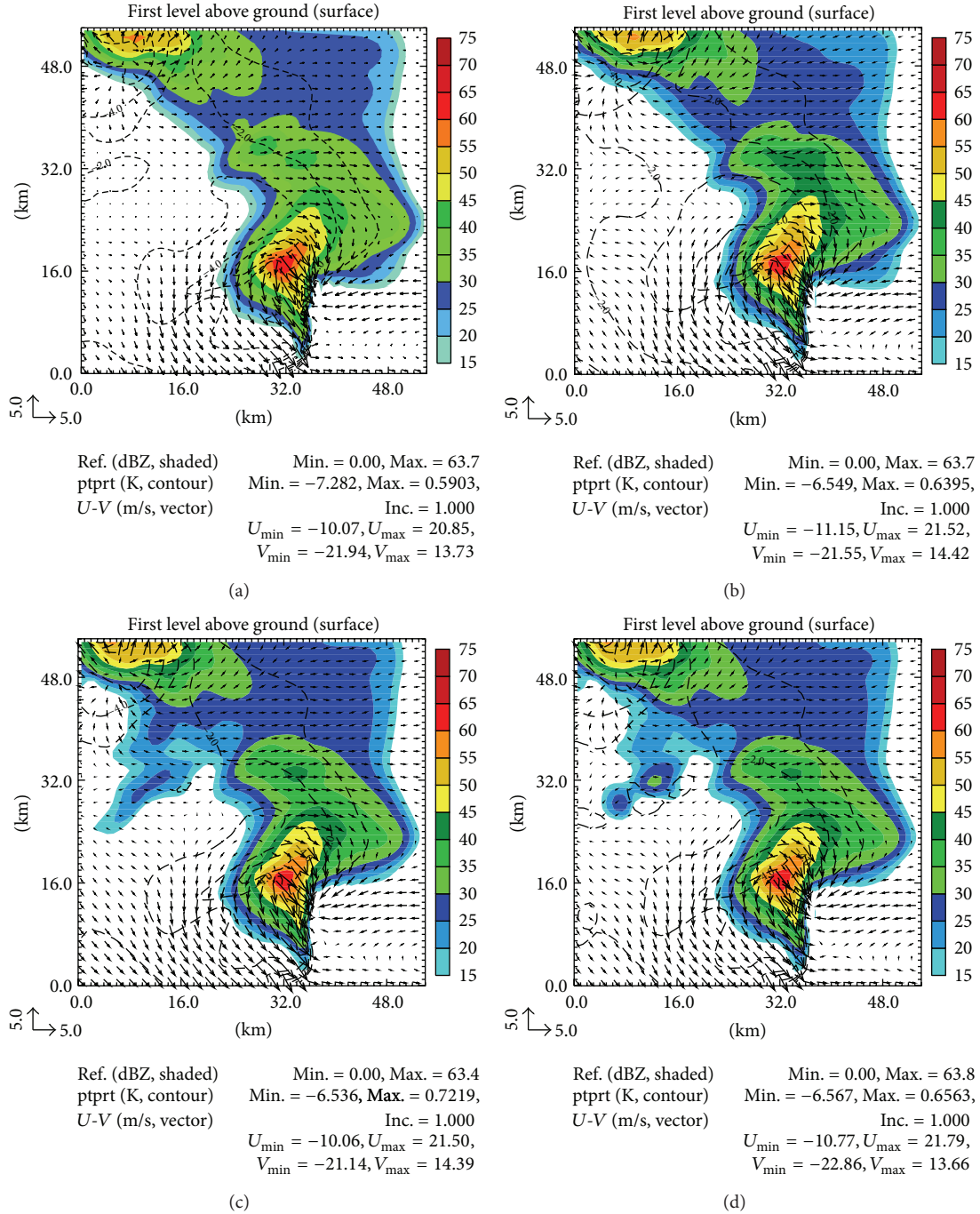$V_{min} = -22.86, V_{max} = 13.66$

FIGURE 5: The same as **Figure 3** but for the experiment with two radars.

5 dBZ in all three experiments. The variation of rms errors is volatile for 3DVAR, especially near the very beginning of the assimilation. The method can decrease the errors from about 40 dBZ to 10 dBZ in two data assimilation cycles, but the errors quickly increase to above 20 dBZ after the 5 min model integration step. The rms errors for the EnKF method decrease more smoothly throughout the data assimilation cycles because of its statistical nature. Perhaps the advantage of hybrid method is most obvious for reflectivity, as it fits the observed reflectivity field more closely than the other two

methods. Though the evolution of rms errors is also volatile for the first 10 minutes, it quickly settles down, and its rms errors are the lowest among all three methods.

4.3. Experiments with Two Radars. The second group of experiments is performed with radar data from two simulated Doppler radars. **Figure 5** shows the final assimilation results after 16 assimilation cycles. As expected, the low-level flow, reflectivity patterns, and the strength of the cold pool look
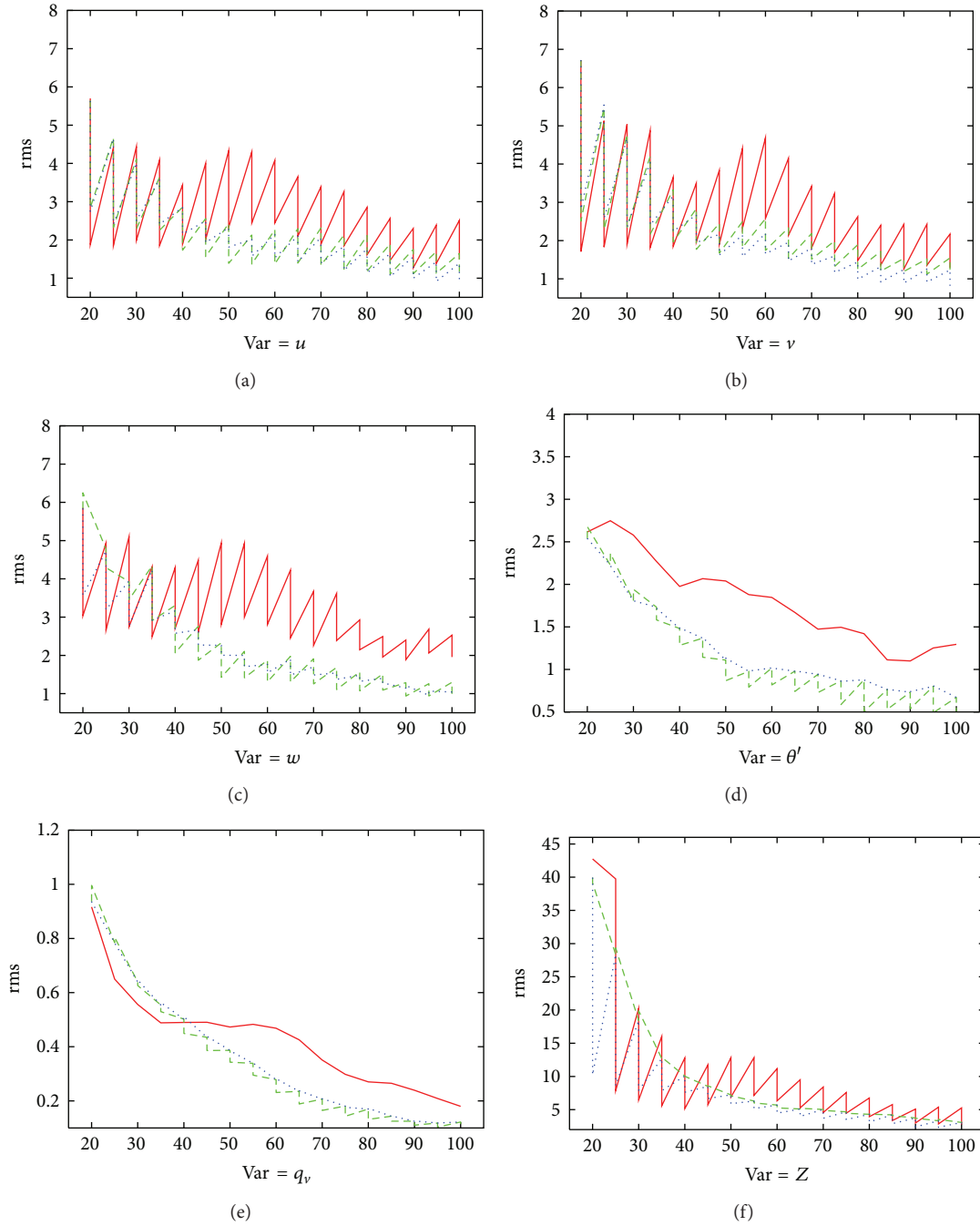
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 6: The same as Figure 4 but for the experiment with two radars.

much better (Figure 5(b)) for 3DVAR (Figure 5(b)), especially for the reflectivity field. The pattern for potential temperature is improved when compared with the single radar experiment (Figure 3(b)) but is still not as good as the truth simulation (Figure 5(a)) and that for EnKF (Figure 5(c)) and the hybrid EnKF-3DVAR (Figure 5(d)). So with more data used, the results for 3DVAR are improved. Again, the most obvious improvement is for the reflectivity field in the center of model domain. The area with reflectivity values larger than 55 dBZ is more similar to the shape of truth simulation. The storm structure for all three methods is well established by

the end of data assimilation at 100 min of reference model assimilation time. The variation of rms errors for the analyzed fields using data from two radars is shown in Figure 6. It is not surprising that the rms errors for model u and v are much improved for 3DVAR. For the first several data assimilation cycles, the errors for 3DVAR are the lowest. With more cycles, the errors for the hybrid method become the lowest among three methods. For most of variables (except potential temperature), the errors for 3DVAR decrease more quickly than seen in the other two methods for the first several data assimilation cycles but then remain at higher levels for later

DA cycles. The variation of rms errors is less volatile when data from two radars are used compared to when data from a single radar is used for 3DVAR. The other features are quite similar to the cases when data from a single radar are used.

5. Summary and Future Work

A hybrid EnKF-3DVAR data assimilation system has been developed based on existing 3DVAR and ensemble Kalman filter (EnKF) programs within the ARPS model. The algorithm uses the extended control variable approach to combine the static and ensemble-derived flow-dependent forecast error covariances [30, 31, 37].

The method is applied to the assimilation of radar data from a simulated supercell storm. Two groups of experiments are performed using different amounts of radar data. Results obtained using 3DVAR (with static covariances entirely), hybrid EnKF-3DVAR, and EnKF are compared. When data from a single radar are used, results show that after 16 cycles of data assimilation, the EnKF and hybrid schemes provide similar results. When evaluated in term of rms errors, the EnKF provides slightly better results for the model dynamic variables, while the hybrid provides slightly better results for the hydrometeor related variables. Though the storm structures can be established reasonably well using 3DVAR, its rms errors are generally worse than those from the other two methods. When data from two radars are used, the rms errors for the hybrid method are smallest for most of the model variables. With two radars, the results from 3DVAR are close to those from EnKF. These tests also indicate that the hybrid scheme can reduce the storm spin-up time because it fits the observations, especially the reflectivity observations, better than the EnKF and the 3DVAR at the beginning of the assimilation cycles. Thus, precipitation exists from the beginning of the model integration.

Our future studies will try to answer a number of key questions within the hybrid EnKF-3DVAR framework just described. They include the following. (1) What is the optimal choice for the relative weight of the static and flow-dependent covariances for storm scale radar data assimilation? (2) What is the optimal combination of ensemble size and grid spacing for a specific computational cost? (3) How does the overall performance of the proposed method compare with 3DVAR and EnKF methods when model error is present? More sensitivity experiments will be performed to answer these questions in the near future, and results will likely help us to solve the challenges of applying this method to real-world scenarios. Even if these questions are successfully answered, the high computational cost of this method is still likely to be a big hurdle. For this, we will apply the dual-resolution strategy as developed for the EnKF scheme in Gao and Xue [46]. A new strategy for hybrid data assimilation proposed by Penny [55] also will be tested within a storm scale data assimilation framework in the near future.

Acknowledgments

References

[1] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.

[2] M. Hu, M. Xue, J. Gao, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part IIimpact of radial velocity analysis via 3DVAR," *Monthly Weather Review*, vol. 134, no. 2, pp. 699–721, 2006.

[3] M. Hu and M. Xue, "Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma city tornadic thunderstorm case," *Monthly Weather Review*, vol. 135, no. 2, pp. 507–525, 2007.

[4] G. Ge and J. Gao, "Latest development of 3DVAR system for ARPS and its application to a tornadic supercell storm," in *Proceedings of the 22nd Conference on Weather Analysis and Forecasting/18th Conference on Numerical Weather Prediction*, 2007.

[5] G. Ge, J. Gao, and M. Xue, "Diagnostic pressure equation as a weak constraint in a storm-scale three dimensional variational radar data assimilation system," *Journal of Atmospheric and Oceanic Technology*, vol. 29, pp. 1075–1092, 2012.

[6] M. Xue, F. Kong, K. Thomas et al., "CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA hazardous weather testbed spring experiment," in *Proceedings of the 24th Conference on Severe Local Storms*, AMS, Savannah, Ga, USA, 2008.

[7] F. Kong, M. Xue, K. W. Thomas et al., "A realtime storm-scale ensemble forecast system: 2009 spring experiment," in *Proceedings of the 23rd Conference on Weather Analysis and Forecasting/19th Conference on Numerical Weather Prediction*, American Meteorological Society, Omaha, Neb, USA, 2009.

[8] A. D. Schenkman, M. Xue, A. Shapiro, K. Brewster, and J. Gao, "Impact of CASA radar and Oklahoma Mesonet data assimilation on the analysis and prediction of tornadic mesovortices in an MCS," *Monthly Weather Review*, vol. 139, no. 11, pp. 3422–3445, 2011.

[9] A. D. Schenkman, M. Xue, A. Shapiro, K. Brewster, and J. Gao, "The analysis and prediction of the 8-9 May 2007 Oklahoma tornadic mesoscale convective system by assimilating WSR-88D and CASA radar data using 3DVAR," *Monthly Weather Review*, vol. 139, no. 1, pp. 224–246, 2011.

[10] D. M. Barker, W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. Xiao, "A three-dimensional variational data assimilation system for MM5: implementation and initial results," *Monthly Weather Review*, vol. 132, pp. 897–914, 2004.

[11] Q. Xiao, Y.-H. Kuo, J. Sun et al., "Assimilation of Doppler radar observations with a regional 3DVAR system: impact of Doppler velocities on forecasts of a heavy rainfall case," *Journal of Applied Meteorology*, vol. 44, no. 6, pp. 768–788, 2005.

[12] W. C. Skamarock, J. B. Klemp, J. Dudhia et al., "A description of the advanced research WRF," Version 2, 2005.

[13] F. Rabier, "Overview of data assimilation developments in numerical weather prediction centers," in *Proceedings of the 4th WMO International Symposium: Assimilation of Observations in Meteorology and Oceanography*, Prague, Czech Republic, 2005.

[14] J. Sun and N. A. Crook, "Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: model development and simulated data experiments," *Journal of the Atmospheric Sciences*, vol. 54, no. 12, pp. 1642–1661, 1997.

[15] J. Sun and N. A. Crook, "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," *Journal of the Atmospheric Sciences*, vol. 55, no. 5, pp. 835–852, 1998.

[16] J. Sun, "Initialization and numerical forecasting of a supercell storm observed during STEPS," *Monthly Weather Review*, vol. 133, no. 4, pp. 793–813, 2005.

[17] Y. Honda and K. Koizumi, "The impact of the assimilation of precipitation data and Radar reflectivity with a pre-operational 4DVAR for the JMA nonhydrostatic model," in *Proceedings of the 10th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, Land Surface (IOAS-AOLS '06)*, 2006.

[18] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, no. 5, pp. 10143–10162, 1994.

[19] C. Snyder and F. Zhang, "Assimilation of simulated Doppler radar observations with an ensemble Kalman filter," *Monthly Weather Review*, vol. 131, no. 8, pp. 1663–1677, 2003.

[20] F. Zhang, C. Snyder, and J. Sun, "Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filte," *Monthly Weather Review*, vol. 132, pp. 1238–1253, 2004.

[21] M. Tong and M. Xue, "Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments," *Monthly Weather Review*, vol. 133, no. 7, pp. 1789–1807, 2005.

[22] D. Dowell, F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, "Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: ensemble Kalman filter experiments," *Monthly Weather Review*, vol. 132, pp. 1982–2005, 2004.

[23] D. C. Dowell, L. J. Wicker, and C. Snyder, "Ensemble kalman filter assimilation of radar observations of the 8 may 2003 oklahoma city supercell: influences of reflectivity observations on storm-scale analyses," *Monthly Weather Review*, vol. 139, no. 1, pp. 272–294, 2011.

[24] N. Yussouf and D. J. Stensrud, "Impact of phased-array radar observations over a short assimilation period: observing system simulation experiments using an ensemble Kalman filter," *Monthly Weather Review*, vol. 138, no. 2, pp. 517–538, 2010.

[25] N. Snook, M. Xue, and Y. Jung, "Analysis of a tornadic mesoscale convective vortex based on ensemble kalman filter assimilation of CASA X-band and WSR-88D radar data," *Monthly Weather Review*, vol. 139, no. 11, pp. 3446–3468, 2011.

[26] N. Snook, M. Xue, and Y. Jung, "Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data," *Monthly Weather Review*, vol. 140, pp. 2126–2146, 2012.

[27] P. L. Houtekamer and H. L. Mitchell, "A sequential ensemble Kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 129, no. 1, pp. 123–137, 2001.

[28] H. L. Mitchell, P. L. Houtekamer, and G. Pellerin, "Ensemble size, balance, and model-error representation in an ensemble Kalman filter," *Monthly Weather Review*, vol. 130, no. 11, pp. 2791–2808, 2002.

[29] T. M. Hamill and C. Snyder, "A hybrid ensemble Kalman filter-3D variational analysis scheme," *Monthly Weather Review*, vol. 128, no. 8, pp. 2905–2919, 2000.

[30] A. C. Lorenc, "The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var," *Quarterly Journal of the Royal Meteorological Society*, vol. 129, no. 595, pp. 3183–3203, 2003.

[31] M. Buehner, "Ensemble-derived stationary and flow-dependent background-error covariances: evaluation in a quasi-operational NWP setting," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 607, pp. 1013–1043, 2005.

[32] M. Zupanski, "Maximum likelihood ensemble filter: theoretical aspects," *Monthly Weather Review*, vol. 133, no. 6, pp. 1710–1726, 2005.

[33] X. Tian, A. Dai, D. Yang, and Z. Xie, "Effects of precipitation-bias corrections on surface hydrology over northern latitudes," *Journal of Geophysical Research*, vol. 112, no. 14, pp. 1984–2012, 2007.

[34] C. Liu, Q. Xiao, and B. Wang, "An ensemble-based four-dimensional variational data assimilation scheme. Part I: technical formulation and preliminary test," *Monthly Weather Review*, vol. 136, no. 9, pp. 3363–3373, 2008.

[35] C. Liu, Q. Xiao, and B. Wang, "An ensemble-based four-dimensional variational data assimilation scheme. Part II: observing system simulation experiments with advanced research WRF (ARW)," *Monthly Weather Review*, vol. 137, no. 5, pp. 1687–1704, 2009.

[36] N. E. Bowler, J. Flowerdew, and S. R. Pring:, "Tests of different flavours of EnKF on a simple model," *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 675, pp. 1505–1519, 2013.

[37] X. Wang, C. Snyder, and T. M. Hamill, "On the theoretical equivalence of differently proposed ensemble—3DVAR hybrid analysis schemes," *Monthly Weather Review*, vol. 135, no. 1, pp. 222–227, 2007.

[38] D. M. Barker and Coauthors, "The Weather Research and Forecasting model's community variational/ensemble data assimilation system: WRFDA," *Bulletin of the American Meteorological Society*, vol. 93, pp. 831–843, 2012.

[39] Y. Li, X. Wang, and M. Xue, "Assimilation of radar radial velocity data with the WRF hybrid ensemble-3DVAR system for the prediction of hurricane Ike (2008)," *Monthly Weather Review*, vol. 140, pp. 3507–3524, 2012.

[40] F. Zhang, M. Zhang, and J. Poterjoy, "E4DVar: coupling an ensemble kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar," *Monthly Weather Review*, vol. 141, pp. 900–917, 2013.

[41] X. Wang, D. M. Barker, C. Snyder, and T. M. Hamill, "A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: observation system simulation experiment," *Monthly Weather Review*, vol. 136, pp. 5116–5131, 2008.

[42] X. Wang, D. M. Barker, C. Snyder, and T. M. Hamill, "A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II: real observation experiments," *Monthly Weather Review*, vol. 136, pp. 5116–5131, 2008.

[43] M. Buehner, P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, "Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: description and single-observation experiments," *Monthly Weather Review*, vol. 138, no. 5, pp. 1550–1566, 2010.

[44] M. Buehner, P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, "Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: one-month experiments with real observations," *Monthly Weather Review*, vol. 138, no. 5, pp. 1567–1586, 2010.

[45] J. C. Derber and A. Rosati, "A global oceanic data assimilation system," *Journal of Physical Oceanography*, vol. 19, pp. 1333–1347, 1989.

[46] J. Gao and M. Xue, "An efficient dual-resolution approach for ensemble data assimilation and tests with simulated Doppler radar data," *Monthly Weather Review*, vol. 136, no. 3, pp. 945–963, 2008.

[47] P. S. Ray, "The morphology of several tornadic storms on 20 May 1977," *Journal of the Atmospheric Sciences*, vol. 38, no. 8, pp. 1643–1663, 1981.

[48] Y. L. Ying Lin, P. S. Ray, and K. W. Johnson, "Initialization of a modeled convective storm using Doppler radar-derived fields," *Monthly Weather Review*, vol. 121, no. 10, pp. 2757–2775, 1993.

[49] M. Xue, K. K. Droegemeier, and V. Wong, "The advanced regional prediction system (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: model dynamics and verification," *Meteorology and Atmospheric Physics*, vol. 75, no. 3-4, pp. 161–193, 2000.

[50] M. Xue, K. K. Droegemeier, V. Wong et al., "The advanced regional prediction system (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: model physics and applications," *Meteorology and Atmospheric Physics*, vol. 76, no. 3-4, pp. 143–165, 2001.

[51] J. Gao, M. Xue, A. Shapiro, Q. Xu, and K. K. Droegemeier, "Three-dimensional simple adjoint velocity retrievals from single-Doppler radar," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 1, pp. 26–38, 2001.

[52] P. L. Jr. Smith, C. G. Myers, and H. D. Orville, "Radar reflectivity factor calculations in numerical cloud models using bulk parameterization of precipitation processes," *Journal of Applied Meteorology*, vol. 14, pp. 1156–1165, 1975.

[53] B. S. Ferrier, "A double-moment multiple-phase four-class bulk ice scheme. Part I: description," *Journal of the Atmospheric Sciences*, vol. 51, no. 2, pp. 249–280, 1994.

[54] M. Zupanski, S. J. Fletcher, I. M. Navon et al., "Initiation of ensemble data assimilation," *Tellus A*, vol. 58, no. 2, pp. 159–170, 2006.

[55] S. Penny, "The hybrid local ensemble transform kalman filter," *Monthly Weather Review*. Accepted.