

An Efficient Dual-Resolution Approach for Ensemble Data Assimilation and Tests with Simulated Doppler Radar Data

JIDONG GAO

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

(Manuscript received 14 December 2006, in final form 13 July 2007)

ABSTRACT

A new efficient dual-resolution (DR) data assimilation algorithm is developed based on the ensemble Kalman filter (EnKF) method and tested using simulated radar radial velocity data for a supercell storm. Radar observations are assimilated on both high-resolution and lower-resolution grids using the EnKF algorithm with flow-dependent background error covariances estimated from the lower-resolution ensemble. It is shown that the flow-dependent and dynamically evolved background error covariances thus estimated are effective in producing quality analyses on the high-resolution grid.

The DR method has the advantage of being able to significantly reduce the computational cost of the EnKF analysis. In the system, the lower-resolution ensemble provides the flow-dependent background error covariance, while the single-high-resolution forecast and analysis provides the benefit of higher resolution, which is important for resolving the internal structures of thunderstorms. The relative smoothness of the covariance obtained from the lower 4-km-resolution ensemble does not appear to significantly degrade the quality of analysis. This is because the cross covariance among different variables is of first-order importance for “retrieving” unobserved variables from the radar radial velocity data.

For the DR analysis, an ensemble size of 40 appears to be a reasonable choice with the use of a 4-km horizontal resolution in the ensemble and a 1-km resolution in the high-resolution analysis. Several sensitivity tests show that the DR EnKF system is quite robust to different observation errors. A 4-km thinned data resolution is a compromise that is acceptable under the constraint of real-time applications. A data density of 8 km leads to a significant degradation in the analysis.

1. Introduction

To gain an understanding of convective storm dynamics and to initialize storm-scale numerical weather prediction (NWP) models, a complete description of the three-dimensional (3D) wind, thermodynamic, and microphysical fields associated with convective storms is needed. Doppler weather radar, as a platform that provides volumetric information at the convective storm scale, only observes radial velocity (V_r) and reflectivity (Z). The assimilation of such data into storm-scale NWP models requires the estimation or retrieval

of many variables that are not directly observed (hence the term retrieval), and it remains a significant challenge. Effective assimilation of Doppler radar data into operational NWP models is of increasing importance as nonhydrostatic models are being adopted at the National Centers for Environmental Prediction (NCEP) and the full-volume data from the entire Weather Surveillance Radar-1988 Doppler (WSR-88D) radar network have recently become available in real time at NCEP.

Over the past two decades, various techniques have been developed for analyzing and retrieving the atmospheric state at the convective scale from Doppler radar data. These techniques range from relatively simple single-Doppler velocity retrieval methods (e.g., Lin et al. 1993; Crook and Tuttle 1994; Xu et al. 1995; Gao et al. 2001; Lazarus et al. 2001; Weygandt et al. 2002),

Corresponding author address: Dr. Jidong Gao, Center for Analysis and Prediction of Storms, National Weather Center, Suite 2500, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: jdgao@ou.edu

three- and four-dimensional variational data assimilation (3DVAR and 4DVAR) methods (e.g., Sun and Crook 1997; Gao et al. 1998, 2004; Xue et al. 2003; Hu et al. 2006a,b), to the emerging ensemble Kalman filter (EnKF) techniques (e.g., Snyder and Zhang 2003; Dowell et al. 2004; Tong and Xue 2005; Xue et al. 2006).

For the purpose of initializing NWP models, single-Doppler velocity retrieval methods are insufficient because they only provide the wind analysis. Although they can be combined with thermodynamic retrievals (Gal-Chen 1978), such a stepwise approach does not always result in dynamically consistent initial conditions for NWP models. The more advanced 4DVAR method promises to effectively use multiple radar volume scan data and provide an initial condition that is consistent with the prediction model. The need to develop and maintain the full model adjoint and the associated high computing costs have, however, limited 4DVAR assimilation of Doppler radar data to relatively simple applications and model settings.

More recently, the ensemble Kalman filter (Evensen 1994, 2003) method was applied to the assimilation of Doppler radar data for convective storms, using simulated data (e.g., Snyder and Zhang 2003; Tong and Xue 2005; Xue et al. 2006) and real radar data (Dowell et al. 2004; Tong 2006). Very encouraging results were obtained in these studies. For example, Tong and Xue (2005) reported on the application of the EnKF to the assimilation of both radial velocity and reflectivity data from single Doppler radar using a forecast model that involves multiphase ice microphysics. The ability of the EnKF scheme to “recover” the complete state of the model thunderstorms, including wind, temperature, pressure, and all water and ice fields, is shown to be excellent under the perfect model assumption; that is, the same model is used both in the assimilation and in the truth simulation.

The EnKF method has the advantage of being able to dynamically evolve the background error covariances through the Monte Carlo approach and through multiple assimilation cycles. Being ensemble-based, it can also provide valuable uncertainty information on both analysis and forecast. However, the computational cost of ensemble-based assimilation methods is significant because of the need to run an ensemble of forecast and analysis of significant sizes (usually from a few dozens to a few hundreds), especially when high-density data are involved and/or when the forecast ensemble is run at high resolution. With limited ensemble sizes, the sampling error associated with the Monte Carlo approach can be significant, which can lead to ensemble underdispersion and filter divergence (Anderson and Anderson 1999). A larger ensemble helps improve the

background error covariance estimation but incurs a higher computational cost.

The incremental 4DVAR method proposed by Courtier et al. (1994) enables the use of double iteration loops and the use of different resolutions of a forecast model. In this approach, the nonlinear prediction model is run at a high spatial resolution to define the nonlinear trajectory around which the linear tangent and the adjoint models are formulated. To reduce computational cost, the tangent linear and adjoint models are run at a reduced resolution within the inner cost-function minimization loop. The analysis increment obtained at this lower resolution (LR) is interpolated and added to the higher-resolution (HR) forecast background for an HR analysis. This analysis serves as the initial condition for the HR forecast. It was the cost saving associated with such an incremental procedure that made the initial operational implementations of 4DVAR practical, and such a procedure is employed in essentially all of current operational 4DVAR systems (e.g., Courtier et al. 1994; Rabier et al. 2000).

In this paper, we propose a dual-resolution (DR) ensemble Kalman filter data assimilation (DA) strategy, which is in a way analogous to the DR incremental 4DVAR approach. The primary goal is to reduce the computational cost of the overall EnKF analysis while trying to maintain the benefits of the EnKF algorithms. With this strategy, an ensemble of forecasts is run at an LR, which provides the background error covariance estimation for both an ensemble of LR analyses and a single HR analysis. For storm-scale applications where the grid resolution tends to be marginal at resolving convective storms, the benefit of having a high-resolution component within the DA system can be significant. At the same time, the cross covariances of analysis background errors between the observed quantities (e.g., radial velocity) and the state variables play a key role in “retrieving” the unobserved variables. We test this strategy for the assimilation of simulated radial velocity data, sampled from a simulated supercell storm. The covariance structure obtained at the lower resolution is analyzed to help us understand the behavior of the dual-resolution system.

In section 2 of this paper, we describe the proposed DR algorithm. In section 3, further details on the algorithm implementation, using the Advanced Regional Prediction System (ARPS; Xue et al. 2000, 2003) as the prediction model, and the design of the OSSEs (observing system simulation experiments; see, e.g., Lord et al. 1997) are given. The experiment results are presented and discussed in section 4, and a summary and conclusions are given in section 5.

2. Dual-resolution strategy for ensemble square root filter algorithm

The key to the ensemble-based filter algorithms is the estimation of the background error covariance and the calculation of the Kalman gain matrix using a forecast ensemble; this was first proposed by Evensen (1994). Since then, there have been a number of developments with the EnKF algorithm to improve the performance of the filter when the ensemble size is relatively small (Burgers et al. 1998; Houtekamer and Mitchell 1998). These methods generally require adding perturbations to the observations. Whitaker and Hamill (2002) proposed an ensemble square root filter algorithm that does not require perturbing observations. Further, the assumption that the observational errors are uncorrelated enables the processing of the observations serially, one by one, leading to a considerable simplification of the analysis scheme (Houtekamer and Mitchell 2001). This square root filter algorithm is used in most radar data assimilation studies, including those of Snyder and Zhang (2003), Dowell et al. (2004), and Xue et al. (2006).

We build our DR EnKF DA system based on the ensemble square root filter algorithm of Whitaker and Hamill (2002). The basic update equation of the original Kalman filter is

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}[\mathbf{y}^o - H(\mathbf{x}^b)]^T, \quad \text{where} \quad (1)$$

$$\mathbf{K} = \mathbf{P}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{P}\mathbf{H}^T)^{-1} \quad (2)$$

is the optimal weighting or Kalman gain matrix (Evensen 1994). Here \mathbf{x} is the state vector we seek to analyze or estimate; superscripts a and b refer to the analysis (posteriori estimate) and background forecast (prior estimate), respectively; and \mathbf{y}^o is the observation vector, following the standard notation of Ide et al. (1997). The forward observation operator that maps the model state to the observations is denoted by H , and \mathbf{H} is the linearized version of H . The covariance matrices for the observation and background errors are \mathbf{R} and \mathbf{P} , respectively.

The ensemble square root algorithm, as with all ensemble-based filters, proceeds in two steps: an analysis step and a forecast or propagation step. In the analysis step, the following equations are used to update the state vectors for the ensemble mean and individual ensemble members:

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}[\mathbf{y}^o - H(\bar{\mathbf{x}}^b)] \quad \text{and} \quad (3)$$

$$\mathbf{x}_n^a = \bar{\mathbf{x}}^a + \beta(\mathbf{I} - \alpha\mathbf{K}\mathbf{H})(\mathbf{x}_n^b - \bar{\mathbf{x}}^b), \quad (4)$$

where n represents the n th ensemble member and the overbar denotes the ensemble mean. Here β is a co-

variance inflation factor that is usually slightly larger than 1, and \mathbf{I} is an identity matrix. The α is given in the square root filter algorithm by Whitaker and Hamill (2002) as

$$\alpha = \left[1 + \sqrt{\mathbf{R}(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1}}\right]^{-1}. \quad (5)$$

This procedure produces an ensemble of analyses, as given by (4).

For completeness, the terms related to the calculation of background error covariance \mathbf{P} are also given here:

$$\mathbf{P}\mathbf{H}^T = \frac{1}{N-1} \sum_{n=1}^N \{(\mathbf{x}_n^b - \bar{\mathbf{x}}^b)[H(\mathbf{x}_n^b) - \overline{H(\mathbf{x}^b)}]^T\}, \quad \text{and} \quad (6)$$

$$\mathbf{H}\mathbf{P}\mathbf{H}^T = \frac{1}{N-1} \sum_{n=1}^N \{[H(\mathbf{x}_n^b) - \overline{H(\mathbf{x}^b)}][H(\mathbf{x}_n^b) - \overline{H(\mathbf{x}^b)}]^T\}, \quad (7)$$

where N is the number of ensemble members. As can be seen, the nonlinear observation operator H instead of its linearized version \mathbf{H} is used in these two equations, removing the need for linearizing the observation operators.

In the forecast step, forecasts are made from each ensemble analysis and are used as the prior estimate or background in the next analysis-forecast cycle; the algorithm continues as the analysis cycles are repeated.

As discussed earlier, EnKF algorithms have been successfully applied to the assimilation of Doppler radar observations into cloud-scale models. The algorithms are, however, computationally expensive. For this reason, a more efficient algorithm is proposed here that involves the use of forecast and analysis ensembles produced at a lower resolution. It also involves a single analysis and forecast produced at a higher resolution. In this case, the background error covariance is estimated from the lower-resolution ensemble. The specific steps for this DR algorithm are given as follows:

- 1) Integrate a single HR model and an ensemble of LR models forward for the length of the analysis cycle or until the next observation is available, so as to yield a single HR forecast \mathbf{x}_h^f and an ensemble of LR forecasts \mathbf{x}_{ln}^f , respectively.
- 2) Calculate the ensemble mean and the ensemble perturbations from the mean for the LR ensemble according to

$$\bar{\mathbf{x}}_l^f = N^{-1} \sum_{n=1}^N \mathbf{x}_{ln}^f \quad \text{and} \quad (8)$$

$$\mathbf{x}_{ln}' = \mathbf{x}_{ln}^f - \bar{\mathbf{x}}_l^f. \quad (9)$$

- 3) Use (1)–(7) to perform EnKF analyses. The analyses are done serially, one observation at a time, as in the regular single-high-resolution (SR) EnKF algorithms. That is, for a given observation, the background error covariance is calculated on the LR grid from the LR ensemble; it is then interpolated to the high-resolution grid and used by the analysis on the HR grid. This covariance is also used on the LR grid to update all ensemble members. This procedure is then repeated for the next available observation until all observations are analyzed. In this study, we assume that the entire volume scan of radar data is valid at the same analysis time.
- 4) The LR EnKF analyses and the single HR analysis are then used as the initial conditions for the LR ensemble forecasts and for the single HR forecast, respectively. The forecasts are carried out to the next analysis time, and the analysis cycle is then repeated.

As pointed out earlier in the introduction, our proposed DR procedure aims to obtain a quality analysis from which an HR prediction can be initialized by using background error covariance derived from a low-cost, low-resolution ensemble. An analogous strategy, in which a single HR forecast and a LR forecast ensemble are used together to improve the quality of the short-range NWP, was proposed by Du (2004). The performance of our proposed dual-resolution algorithm is evaluated in the next two sections through OSSEs.

3. Assimilation system and experimental design

This section describes the prediction model that is used to create the truth simulation and used for the EnKF data assimilation. A description on the OSSEs for assimilating radar radial velocity data follows. Additional sensitivity experiments will also be performed to examine aspects related to data density, observation error magnitude, and ensemble size.

a. Prediction model and truth simulation for OSSEs

We test our DR EnKF algorithm and compare its results with those of single HR and LR EnKFs, using simulated data from a classic supercell storm of 20 May 1977 in Del City, Oklahoma (Ray et al. 1981). 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

TABLE 1. Parameters of the model grids used in the experiments. NX, NY, and NZ are the number of grid points in the x , y , and z direction, respectively.

Model resolution	High-resolution grid	Low-resolution grids	
$\Delta x, \Delta y$	1 km	2 km	4 km
Δz	500 m	500 m	500 m
NX, NY	57	29	15
NZ	35	35	35

q_i , snow q_s , and hail q_h . The microphysical processes are parameterized using the three-category ice scheme of Lin et al. (1983). More details on the model can be found in (Xue et al. 2000, 2001).

For our experiments, the model domain is $57 \times 57 \times 16 \text{ km}^3$. The LR and HR models have a horizontal grid spacing (Δx and Δy) of 4 and 1 km, respectively, while the vertical grid spacing (Δz) is 500 m in both cases (Table 1). The truth simulation is created using the 1-km horizontal resolution and is initialized from a modified real sounding plus a 4-K ellipsoidal thermal bubble centered at $x = 48$, $y = 16$, and $z = 1.5 \text{ km}$, with radii of 10 km in x and y and 1.5 km in the z direction. Open conditions are used at the lateral boundaries. The length of simulation is two hours. A constant wind of $u = 3 \text{ m s}^{-1}$ and $v = 14 \text{ m s}^{-1}$ 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. (2001). 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 in two at around 55 min (Figs. 1a,b). The right-moving (relative to the storm motion vector, which is toward north-northeast) cell tends to dominate the system; the updraft reaches a peak value of over 40 m s^{-1} at 90 min (Fig. 1c). The left-moving cell starts to split again at 95 min. The initial cloud started to form at about 10 min, and rainwater formed at about 15 min. Ice phase fields appeared at about 20 min. A similar truth simulation was also used in Gao et al. (2001, 2004) and Tong and Xue (2005).

b. Simulation of radar observations

As in Snyder and Zhang (2003) and Tong and Xue (2005), the simulated observations are assumed to be available on the grid points. The simulated radial velocity, v_r , is calculated from

$$v_r = u \sin \phi \cos \mu + v \cos \phi \cos \mu + w \sin \mu, \quad (10)$$

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 stag-

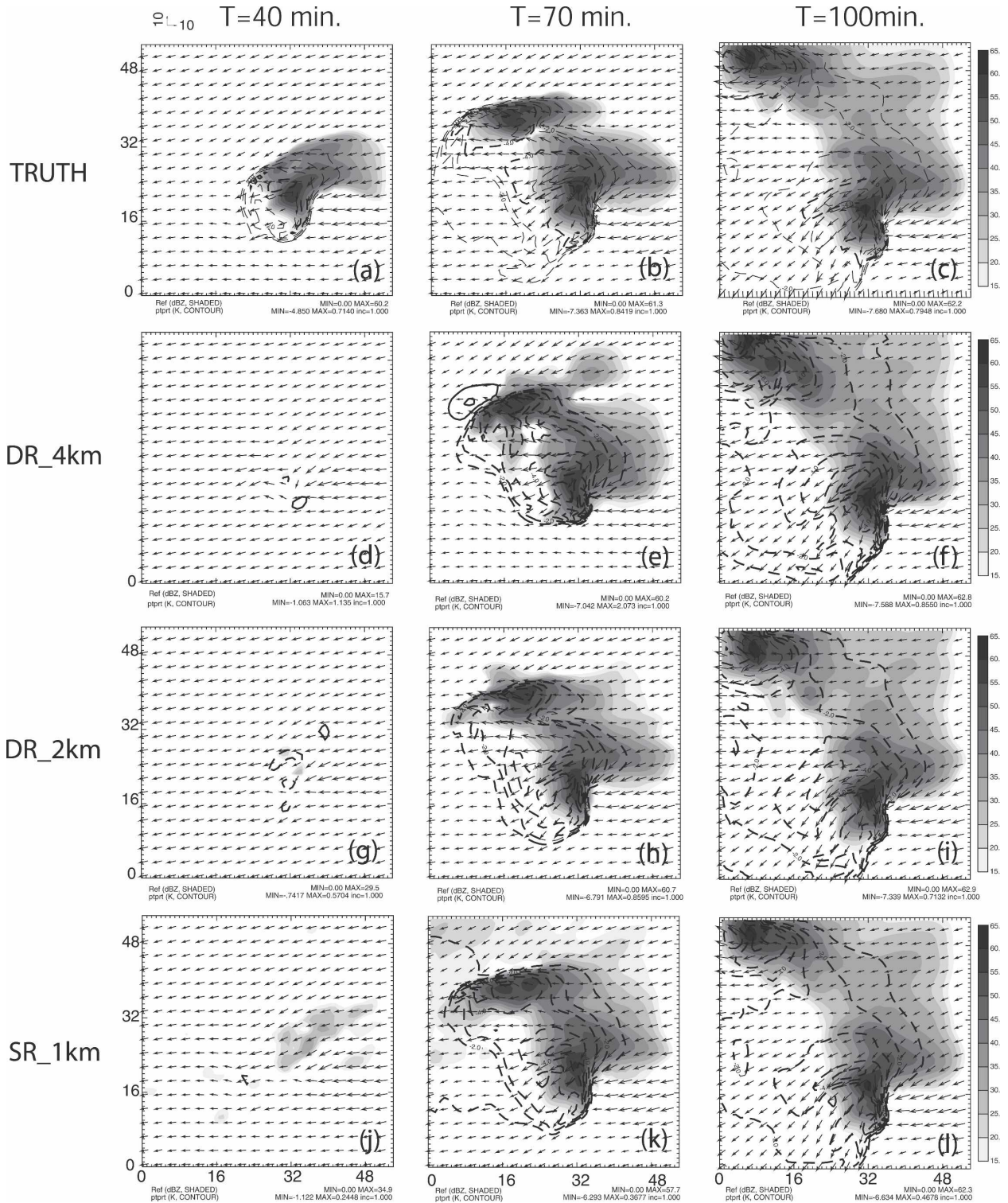


FIG. 1. Horizontal winds (vectors; m s^{-1}), perturbation potential temperature (contours at 1-K intervals), and simulated reflectivity (shaded contours; dBZ) at 250 m AGL for (a)–(c) the truth simulation; (d)–(f) single-high-resolution analyses from DR_4km; (g)–(i) single-high-resolution analyses from DR_2km; and (j)–(l) the ensemble mean analyses from SR_1km. The times shown are 40, 70, and 100 min. Wind vectors are shown every 2 km.

TABLE 2. List of all experiments and their assimilation configuration, data resolution, purpose, and the total wall clock time. See text for more details.

Experiment	Description	Purpose	Wall clock time (s)
DR_4km	DR analysis with 4-km ensemble and 1-km high resolution	Test the quality and efficiency of DR as	4642
DR_2km	DR analysis with 2-km ensemble and 1-km high resolution	compared to standard EnKF	14 044
SR_1km	Single-resolution EnKF analysis with $\Delta x = 1$ km	high-resolution analysis	65 295
SR_4km	Single-resolution EnKF analysis with $\Delta x = 4$ km	Test the performance of SR analyses at	3780
SR_2km	Single-resolution EnKF analysis with $\Delta x = 2$ km	different resolutions	12 041
DR_s10	As DR_4km, with 10 members	Test the sensitivity of DR analysis to	2027
DR_s400	As DR_4km, with 400 members	ensemble size	50 215
DR_ob04	As DR_4km, but observations at every 4 km	Test the sensitivity of DR analysis to	4167
DR_ob08	As DR_4km, but observations at every 8 km	data density	2875
DR_err03	As DR_4km, with obs error SD of 3 m s^{-1}	Test the sensitivity of DR analysis to	4688
DR_err05	As DR_4km, with obs error SD of 5 m s^{-1}	data error	4661

gered model grid. Random errors drawn from a normal distribution with zero mean and a standard deviation of 1 m s^{-1} (unless otherwise stated) 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. For the HR (1 km) assimilation, the assimilation model is perfect because it is identical to that used by the truth simulation. For the low-resolution ensemble, the assimilation model is imperfect because of the degraded resolution. In this study, only the radial velocity data are assimilated and they are only available where the truth reflectivity is greater than zero in the analysis domain. We also use the data at every other grid point of the 1-km truth simulation grid in all three directions unless otherwise stated.

c. Design of assimilation experiments

We start the initial ensemble forecast at 25 min of the model integration time when the storm cell reaches peak intensity. 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. (2006). The random noise is sampled from Gaussian distributions with zero mean and standard deviations of 5 m s^{-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 (2005) 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 observations are simulated and assimilated every 5 min. The first analysis is performed at 30 min and 40 ensemble members are used unless otherwise noted. To localize covariances during the analysis, Eq. (4.10) of Gaspari and Cohn (1999) is used when calculating the background error matrix $\mathbf{P}\mathbf{H}^T$, as suggested by Houtekamer and Mitchell (2001). A cutoff radius of 8 km is used in most of our experiments.

We perform a set of EnKF experiments listed in Table 2. For all of the experiments, the qualities of analysis are measured by rms error, and the follow-up forecast experiments are not performed for the purpose of brevity. The first group of experiments is performed to test the impact of model resolution. DR_4km is a DR experiment with ensemble forecast and analysis using $\Delta x = 4$ km, and the single-high-resolution forecast and analysis using $\Delta x = 1$ km. Experiment DR_2km differs from DR_4km in that the lower-resolution ensemble uses a 2-km horizontal resolution. For comparison purposes, we also perform single-high-resolution (SR_1km) EnKF analyses at the 1-km horizontal resolution. In the second group of experiments, two single-resolution analyses (SR_4km and SR_2km) are performed with 4- and 2-km horizontal resolutions, respectively.

The third group of experiments, DR_s10 and DR_s400, using 10 and 400 ensemble members, respectively, examine the impact of ensemble size. As another way of reducing the cost, data thinning is often performed when the data density is close to or higher than the model resolution. The fourth group of experiments, DR_4ob and DR_8ob, are the same as experiment DR_4km, except that the radial velocity observations at every fourth and eighth grid point, respectively, from the 1-km truth grid are used. In the vertical, the data are available at 1- and 2-km intervals, respectively. This group of experiments tests the sensitivity of the analysis to reduced data density, that is, to the degree of data

thinning. Finally, experiments DR_err03 and DR_err05 are performed to test the sensitivity of the analysis to different magnitudes of observation errors. We note here again that the standard data density for non-data thinning experiments is one data every other point of the 1-km truth grid.

4. Results of experiments

a. Dual- and single-high-resolution experiments

As stated above, the first two dual-resolution experiments, DR_4km, DR_2km, and a single-high-resolution experiment, SR_1km, are performed to test the effectiveness of our proposed DR algorithm. Figure 1 shows the horizontal winds, perturbation potential temperature, and reflectivity at 250 m AGL (first model level above surface) and Fig. 2 shows the horizontal wind and vertical velocity fields at 3.5 km AGL, at 40, 70, and 100 min of model time. They are shown for the truth simulation, the 10-km analyses of DR_4km and DR_2km, and the ensemble mean analysis of SR_1km.

For DR_4km, Fig. 1d shows that, after 2 cycles at $t = 40$ min, the analysis has only captured some weak perturbations in the wind field. No precipitation is present at the surface and a small positive temperature perturbation is found where there should be rain-cooled air. At the 3.5-km level (Fig. 2d), an updraft is established at about the right location but its structure differs from the truth (Fig. 2a). After six more analysis cycles at $t = 70$ min, the low-level flow immediately underneath the storm cells becomes similar to the truth (Fig. 1e versus Fig. 1b) but the outflow and cold pool on the southwest side are too weak. At this time, the minimum temperature of the cold pool is close to the truth (-7 K versus -7.4 K). There is a small positive temperature anomaly northwest of the left-moving cell in Fig. 1e that is spurious.

At midlevels, the perturbation (from the environmental wind profile) horizontal winds and the updrafts associated with the two split cells are well captured in DR_4km by $t = 70$ min (Fig. 2e). Cyclonic and anticyclonic rotation associated with the right and left movers are clearly evident. Compared to the truth (Fig. 2b), the main updrafts of both right and left movers are broader and slightly weaker. This is believed to be related to the relatively smooth spatial covariance that is derived from the 4-km ensemble and used in the analysis. In comparison, the updrafts from DR_2km and SR_1km are more intense and closer to the truth (Figs. 2h,k).

By $t = 100$ min, the analysis of the DR_4km is much improved. In fact, by this time, there is no longer much difference from the truth in either the low-level or the

midlevel fields (Figs. 1f and 2f). The midlevel updrafts are now stronger and close to the truth. However, the updrafts in DR_2km and SR_1km (Figs. 2i,l) are closer to the truth. The midlevel horizontal flow fields are very well analyzed in all three experiments by this time, including the mesocyclones. The results of the DR_4km analysis indicate that the reduced 4-km resolution used by the ensemble is still effective in providing error covariance information for the EnKF analysis on the 1-km resolution grid, even though it takes more assimilation cycles to achieve a good analysis.

Overall, the analyses of the DR_2km are better than those of the DR_4km. At $t = 40$ min, negative potential temperature perturbations are produced at the surface (Fig. 1g), and by 70 min the pattern of the low-level reflectivity matches the truth simulation better (Fig. 1h) than that of the DR_4km. The midlevel updrafts are more intense, as pointed out earlier (Fig. 2h). The differences become smaller by 100 min. Increasing the resolution of the ensemble from 4 to 2 km, therefore, does result in better analyses. The overall computational cost is, however, increased by about a factor of 3 (Table 2, 14 044 s versus 4642 s). Considering the large cost difference and the relatively small difference in the analyzed results at the end of the assimilation cycles, the use of a 4-km ensemble for the 1-km analysis appears rather attractive.

With SR_1km, the assimilation procedure reduces to the standard EnKF algorithm with the perfect model assumption, which is also assumed by Snyder and Zhang (2003) and Tong and Xue (2005). In this best-case scenario, the analysis converges much faster and the final analysis errors are generally very small (Figs. 1j–l and 2j–l). The final analysis looks almost identical to the truth at the end of the assimilation window. The analyzed fields have similar accuracies as reported in the referenced earlier studies.

The rms errors of the analyzed fields on the 1-km grid are shown in Fig. 3. As in Tong and Xue (2005), the rms error calculation is limited to the regions where the truth reflectivity exceeds 10 dBZ. Figure 3 shows that the rms errors of w , θ' , q_v , and reflectivity Z (derived from the hydrometeor mixing ratios) generally decrease with the cycles in all three experiments. The errors in DR_4km decrease more slowly and remain at a higher level at the end of assimilation cycles than those in DR_2km and SR_1km. The rms error of w is close to 4 m s^{-1} at 100 min, while that in SR_1km is close to 1 m s^{-1} . The θ' rms errors for DR_4km and SR_1km are over 1.5 K and below 0.5 K, respectively. While these differences are significant, we note here that the error levels of SR_1km are unrealistically low due to the perfect model assumption. For real data

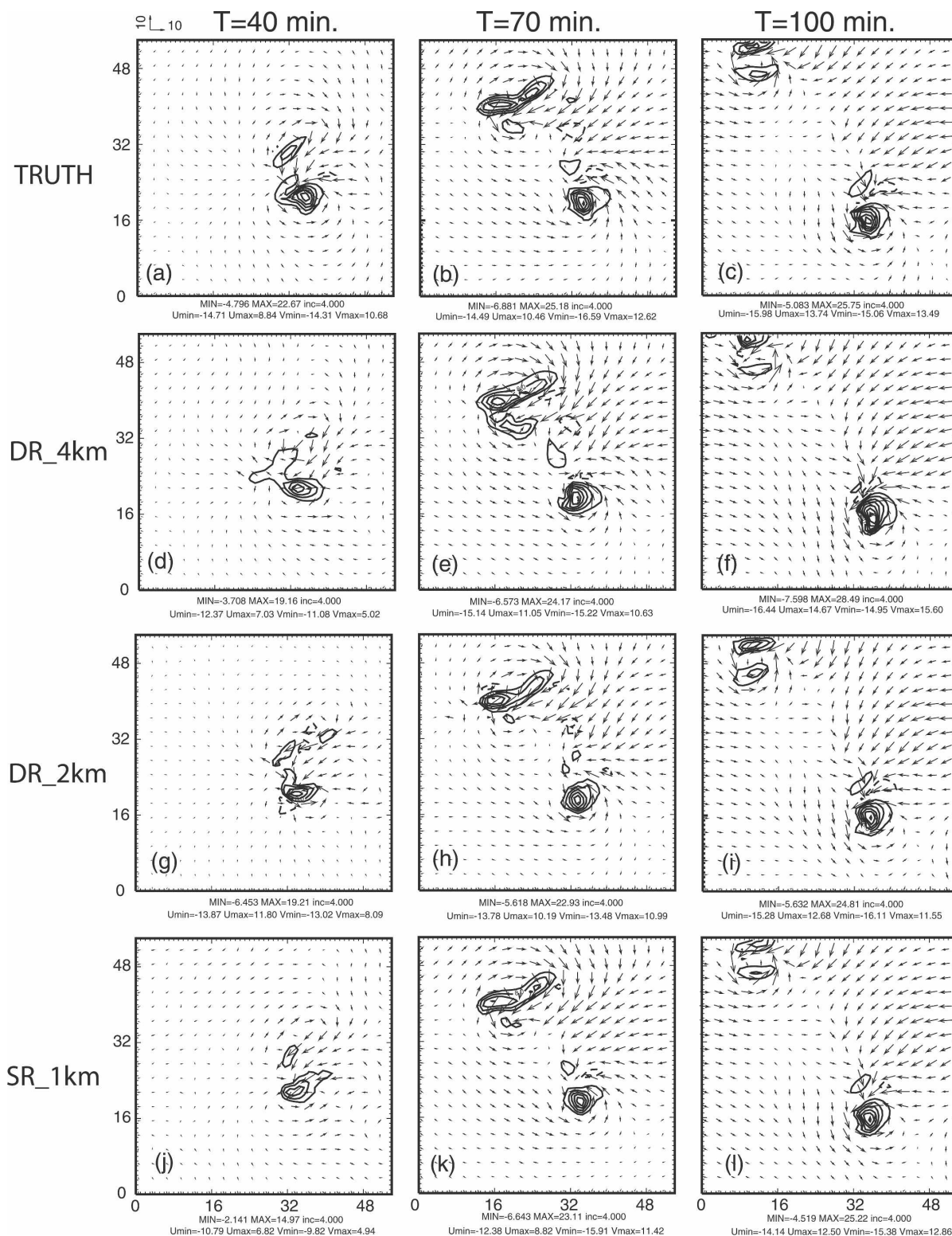


FIG. 2. Same as Fig. 1, but for perturbation horizontal winds from the environment and the vertical velocity w fields at 3.5 km AGL. The w contour interval is 4 m s^{-1} ; positive contours are solid and negative contours are dashed.

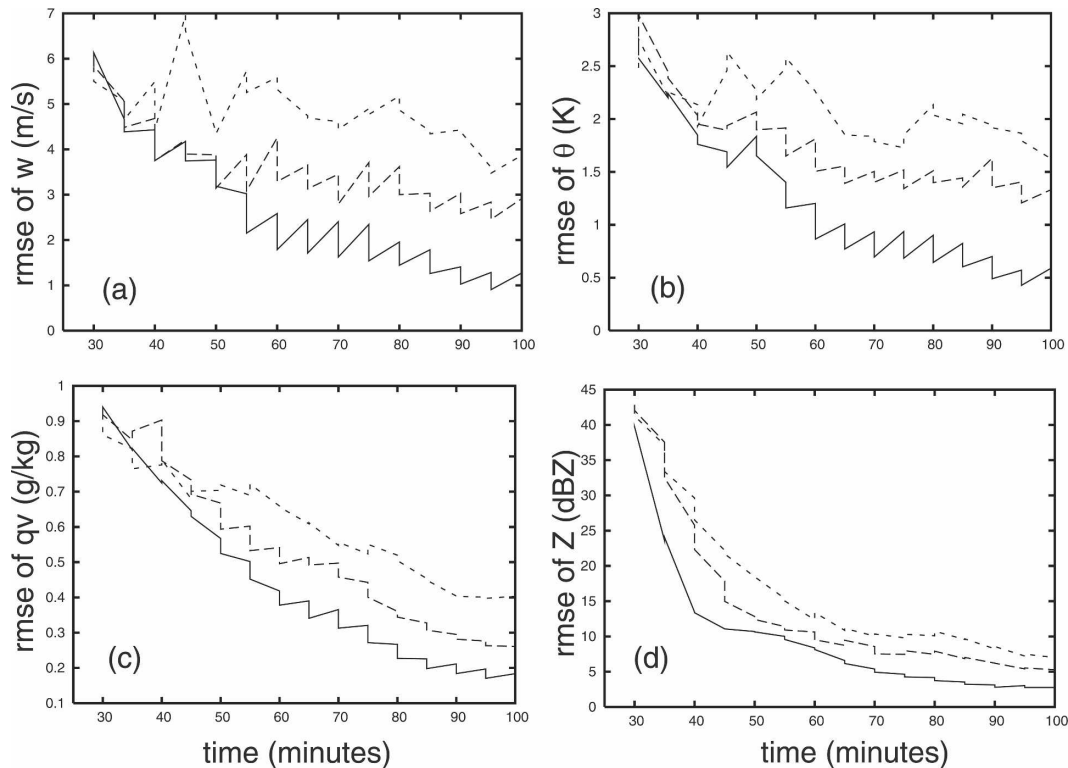


FIG. 3. The rms errors of the ensemble forecast and analyses for dual resolution and single high resolution, averaged over points at which the reflectivity is greater than 10 dBZ for (a) w , (b) θ' , (c) q_v , and (d) Z for experiments SR_1km (solid), DR_2km (long-dashed), and DR_4km (short-dashed) for different data densities.

cases where model error exists, the analysis errors are most likely to be much higher (see, e.g., Dowell et al. 2004). We note that, 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 Fig. 2. Further, the θ' error of about 1.5 K is comparable to typical errors of conventional temperature observations. Interestingly, the differences among the rms errors for q_v and Z for different experiments are smaller (Fig. 3). The q_v rms error is reduced to 0.4 g kg^{-1} in DR_4km by 100 min and that of Z is decreased to about 7 dBZ, both relatively low values. The rms error results again suggest that the dual-resolution approach is viable for storm-scale radar data assimilation.

b. Single-lower-resolution EnKF experiments

Experiments SR_4km and SR_2km use the conventional single-resolution approach. The ensemble forecast and analysis are performed at a lower, 4- or 2-km, horizontal resolution. No high-resolution grid is involved in the procedure. The truth is the same 1-km simulation used earlier. Because of the resolution difference between the analysis and truth, model error

exists. As before, the observations are available at every other point of the truth grid.

Figures 4 and 5 show the analyzed fields at the surface and at the 3.5-km level, respectively, from SR_4km and SR_2km. At the lower resolution, fewer low-level structures are captured at 40 min than in DR_4km and DR_2km (Figs. 4a,d). At the later times, the low-level reflectivity starts to show a hook echo pattern (Figs. 4c,f), with DR_2km being closer to the truth. For DR_4km, the overall fields are much smoother and the cold pool is too warm, however. At midlevels, the analyzed updrafts are much weaker in SR_4km (8 versus 24 m s^{-1}), with those of the left mover completely missing at 100 min (Fig. 5c). The pattern and location of the analyzed updrafts in SR_2km are much better (Figs. 5e,f) but their intensities are still weak compared to the truth (16 versus 24 m s^{-1}) or those of DR_2km (cf. Fig. 2). In both cases, cyclonic and anticyclonic circulations associated with the right and left movers are present in the midlevel analysis, but there are significant errors in their structure and location (Fig. 5). These results indicate that even though the coarser-resolution ensembles can provide valuable covariance information for the high-resolution EnKF analysis, the analyses at the

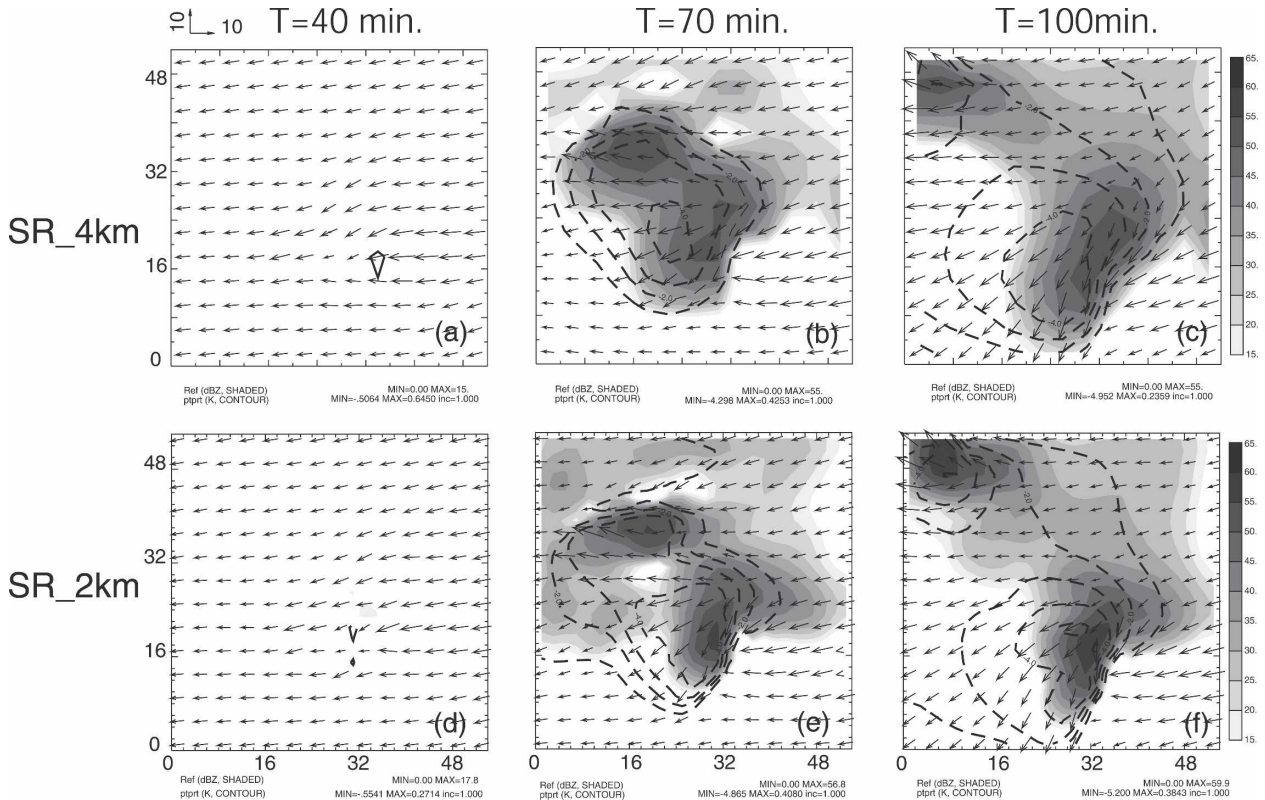


FIG. 4. Same as Fig. 1, but for experiments (a)–(c) SR_4km, and (d)–(f) SR_2km. Wind vectors are shown every 4 km.

coarser resolution contain significant error. It is beneficial to perform an additional analysis at the higher resolution, which takes advantage of the higher resolution and more accurate forecast background while utilizing the covariance information derived from the coarser-resolution ensemble.

c. The correlation structures and computational costs

As discussed by Snyder and Zhang (2003) and Tong and Xue (2005), estimating the unobserved model variables from radar radial velocity data in the EnKF system depends on reliable flow-dependent correlations or covariances between the background errors of state variables and radial velocity. In the EnKF, it is actually the covariances between the ensemble deviations from the mean that serve as the surrogate of the error covariances because the truth error is unknown. Such covariances are calculated according to (6) and (7).

Here, we are interested in finding out the quality of the estimated background error covariances when they are derived from the lower-resolution ensemble and how they compare to those of a full-resolution ensemble. For interpretation purposes we present in Fig. 6, for experiments SR_1km, SR_2km, SR_4km, and

DR_s400 (to be discussed later), example error (strictly the ensemble deviation) correlations which are normalized covariances. The correlations are calculated between the radial velocity (background velocity projected to the radial direction of radar beams) at $x = 34$ km and $z = 8$ km and state variables w , θ' , q_c , and q_i in an east–west vertical cross section through the updraft core ($y = 17$ km) of the truth simulation. The time of the calculations is 100 min. The correlation between radial velocity v_r at an observation point and scalar s at a particular grid point is given by

$$\langle v_r, s \rangle = \frac{\sum_{n=1}^N (v_{rn}^b - \bar{v}_r^b)(s_n^b - \bar{s}^b)}{\sqrt{\sum_{n=1}^N (v_{rn}^b - \bar{v}_r^b)^2 \sum_{n=1}^N (s_n^b - \bar{s}^b)^2}}, \quad (11)$$

where v_{rn}^b and s_n^b are the background (denoted by b) radial velocity and scalar, respectively, of the n th member of the ensemble, and \bar{v}_r^b and \bar{s}^b are the corresponding ensemble means. The statistics are calculated from N samples where N is the ensemble size. In (11), s is one of w , θ' , q_c , and q_i , variables for which correlations are examined.

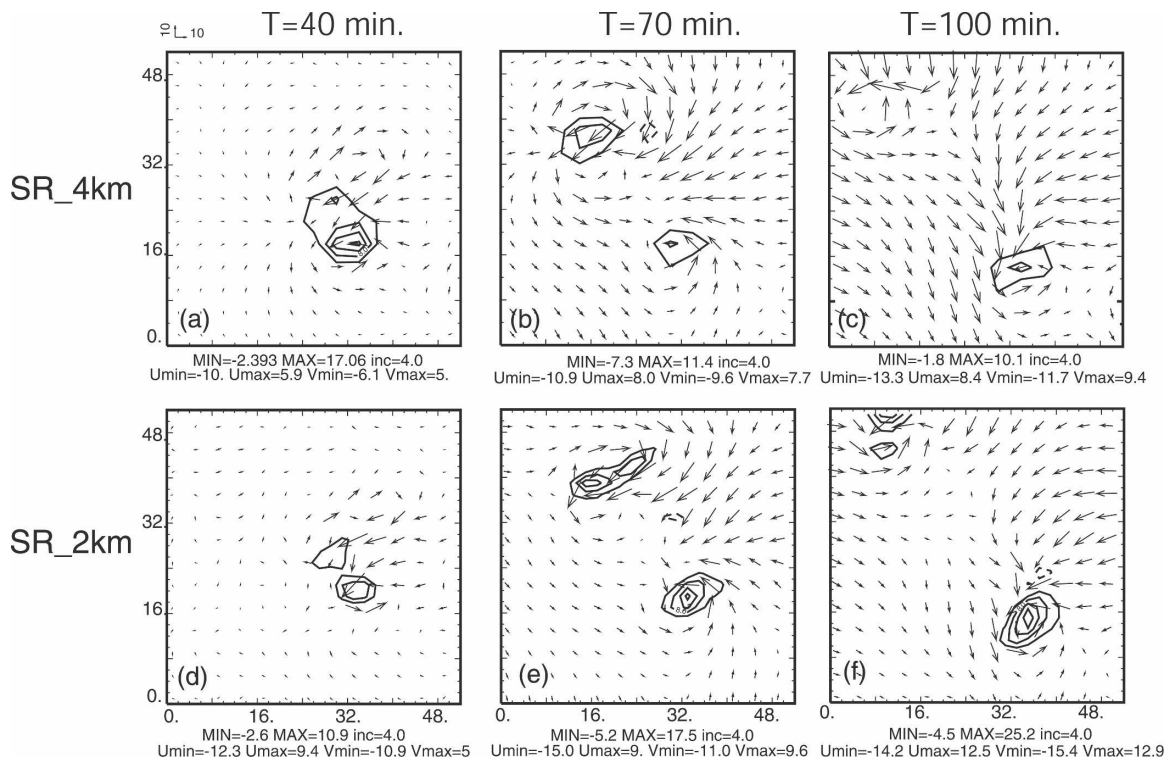


FIG. 5. Same as Fig. 2, but for experiments (a)–(c) SR_4km, and (d)–(f) SR_2km.

As expected, in all the experiments shown in Fig. 6, for cloud water and cloud ice fields, correlations are confined to the regions where their values are greater than zero (see the two rightmost columns of Fig. 6). Significant correlations extend through much of the troposphere depth for w and θ' , which is consistent with the nature of deep convection. Generally, positive correlations are found for w and θ' within the updraft core while negative correlations are found in the recirculation regions on the flanks of the updraft. The correlations for all variables from SR_1km contain many detailed structures (Figs. 6a–d), while those from SR_2km look similar except for being somewhat smoother (Figs. 6e–h). SR_4km produces the smoothest and broadest correlation patterns (Figs. 6i–l), but the general structures, such as the positive and negative correlation centers, remain similar to those of SR_1km. For example, the correlation between v_r and w is generally positive in the updraft region and is generally negative in the surrounding regions of return circulation in both cases. These regions of negative correlations are, however, broader and located further away from the updraft in SR_4km, reflecting the weaker and broader return circulations expected in the coarser-resolution forecasts.

For radar data assimilation, the cross correlations among the observed parameters (v_r in this case) and the

unobserved model state variables play the most important role in retrieving the unobserved quantities. The spatial covariance helps spread observation information to nearby grid points, but in the presence of dense radar data, the spatial covariance is of less importance than the cross covariances. Further, the spatial covariance localization typically employed in EnKF (Houtekamer and Mitchell 2001) truncates remote spatial correlations, lessening the impact of inaccurate estimation of spatial covariance structures. These appear to be the reasons why a rather good analysis can be obtained in the case where the covariances are estimated from a lower-resolution ensemble.

Figure 6 also shows that the correlation patterns for SR_2km are much closer to those of SR_1km, suggesting that the 2-km resolution is much better at representing the storm-scale circulations than the 4-km resolution. To further investigate the impact of ensemble size on the covariance estimation, we perform experiment DR_s400, which is the same as DR_4km except that the ensemble size is 400 instead (note that the covariances derived from the 4-km ensemble of DR_4km should be very similar to those of SR_4km). The sampling error with 400 members should be much smaller. With 400 members, the positive correlation centers become narrower (closer to those of SR_1km) for w and

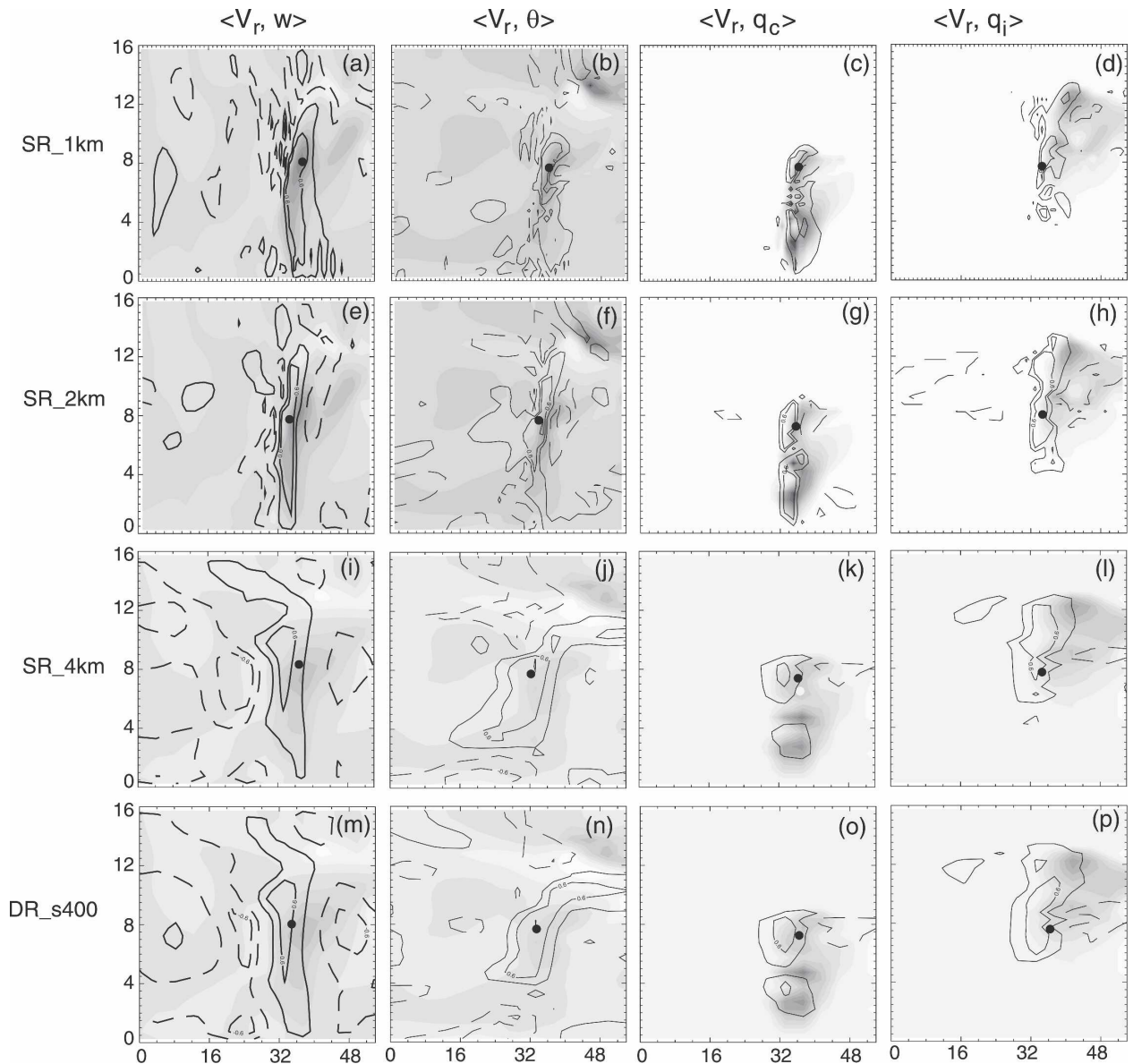


FIG. 6. Error correlation coefficients between the forecast radial velocity at $x = 34$ km and $z = 8$ km (indicated by a black dot) and the forecast vertical velocity w , perturbation potential temperature θ' , cloud water q_c , and ice q_i for experiments (a)–(d) SR_1km, (e)–(h) SR_2km, (i)–(l) SR_4km, and (m)–(p) DR_S400 prior to the analysis at $t = 100$ min, in an x – z cross section through $y = 17$ km, where the maximum updraft is found in the truth simulation. The contour interval is 0.3 and negative contours are dashed. The shaded contours show the corresponding model fields from the truth simulation and the contour intervals are 4 m s^{-1} , 2 K , 0.5 g kg^{-1} , and 0.5 g kg^{-1} for w , θ' , q_c , and q_i , respectively.

θ' (Figs. 6m,n versus Figs. 6i,l) and the negative correlation pattern for w west of the updraft is also narrower (Fig. 6m), but the overall structures are close to those of SR_4km. This suggests that the 40-member ensemble is capturing the most important correlation structures.

The main reason for using the lower-resolution ensemble in our dual-resolution algorithm is the reduced computational cost. Table 2 lists the total wall clock time used by each assimilation experiment on a dedi-

cated Linux workstation with a single 3.8-GHz Pentium 4 processor. We can see that experiment SR_4km only uses 3780 s of wall clock time. Compared to SR_4km, experiment DR_4km adds one high-resolution model run and the corresponding EnKF analysis. This adds only 862 s to the wall clock time, but the improvement to the quality of analysis is dramatic.

Experiment SR_1km produces the best analysis but uses 65 295 s of wall clock time, which is about 14 times

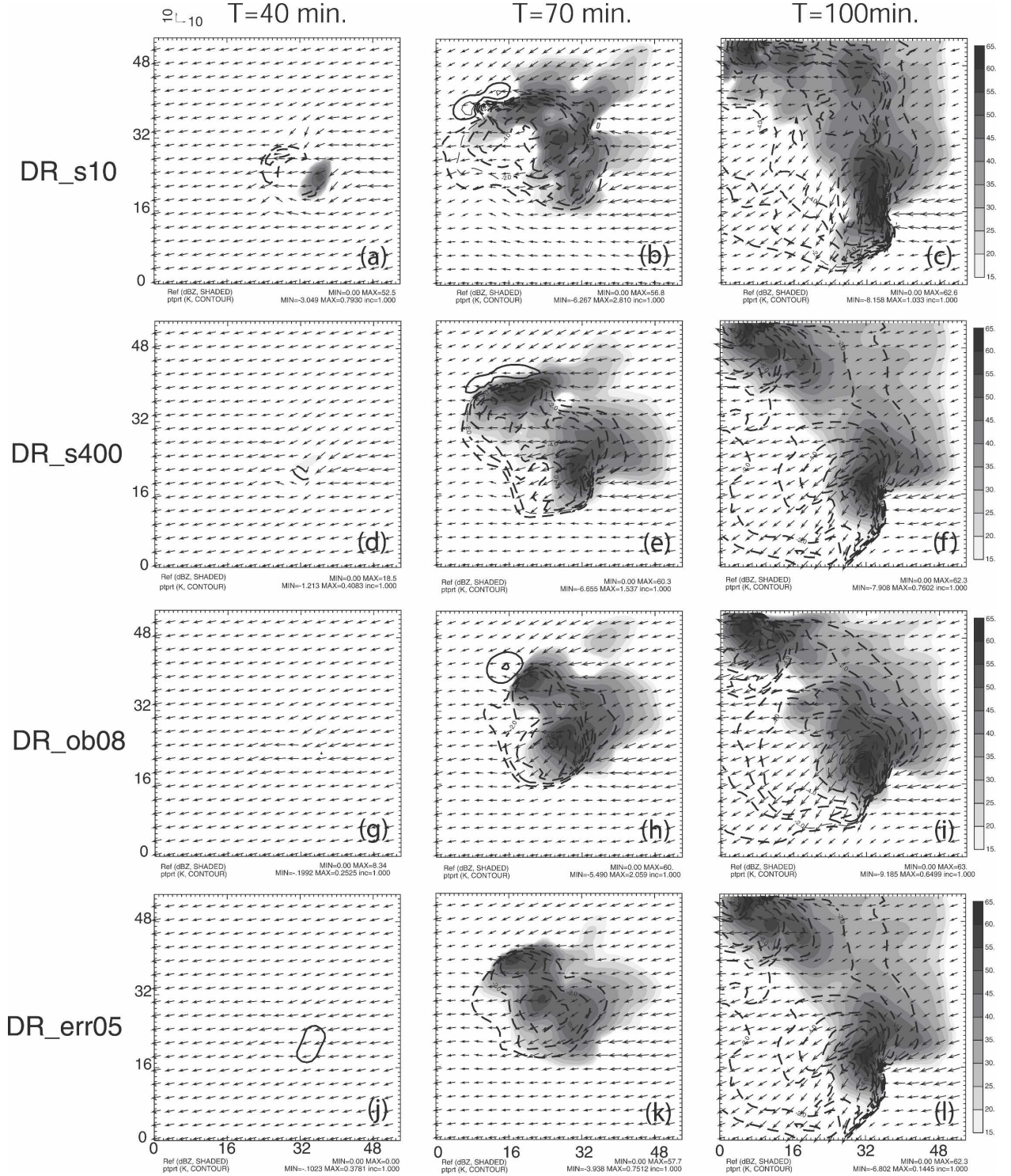


FIG. 7. Same as Fig. 1, but for experiments (a)–(c) DR_s10, (d)–(f) DR_s400, (g)–(i) DR_ob08, and (j)–(l) DR_err05.

that of DR_4km and about 17 times that of SR_4km. The other dual-resolution experiments listed in Table 2 all have significant computational cost savings. Such computational cost savings are very important for real-

time implementations of convective-scale ensemble-based data assimilation, as the cost in general is very high compared to available computational resources. We expect this methodology to be applicable to large-

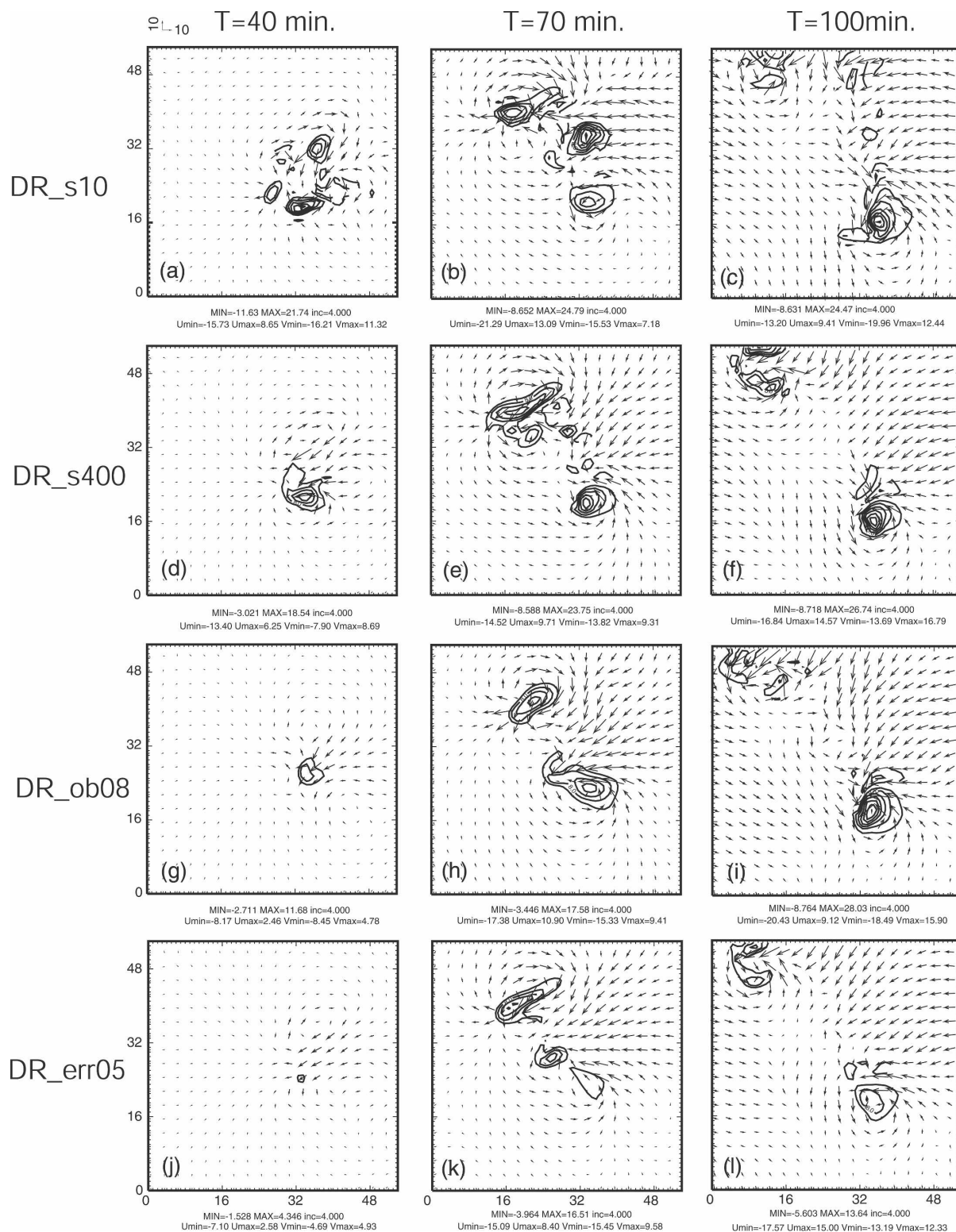


FIG. 8. Same as Fig. 2, but for experiments (a)–(c) DR_s10, (d)–(f) DR_s400, (g)–(i) DR_ob08, and (j)–(l) DR_err05.

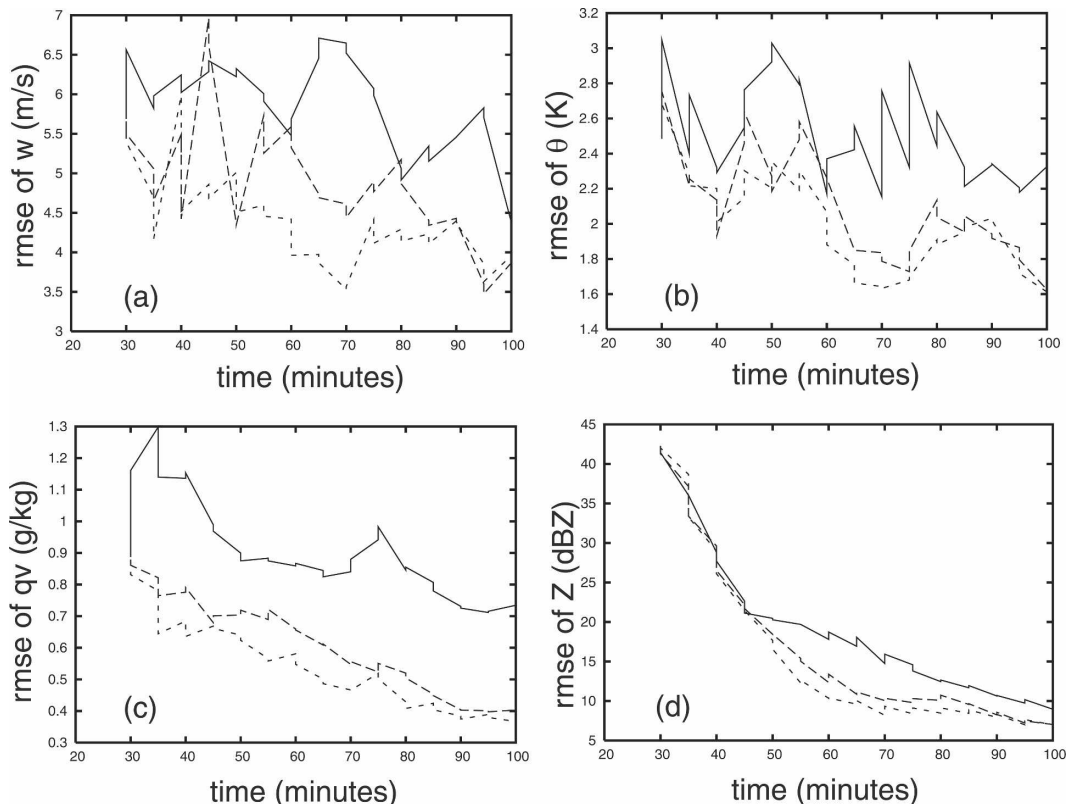


FIG. 9. Same as Fig. 3, but for experiments DR_s10 (solid), DR_4km (long-dashed), and DR_s400 (short-dashed), which employ ensemble sizes of 10, 40, and 400, respectively.

scale NWP also where the method may work even better, because the dominant flow features such as the baroclinic waves tend to be reasonably well resolved even at the coarser resolutions.

d. Ensemble size and covariance localization

Most ensemble Kalman filters are sensitive to ensemble size and covariance localization, but these sensitivities may be quite different for different cases and/or when different dynamics are involved or even for different types of observations. In the previous section, we showed that the forecast error covariances obtained from a 40-member ensemble is reasonably close to those obtained from 400 members. In this section, we discuss a dual-resolution experiment in which the lower-resolution ensemble has only 10 members. Figures 7a–c and 8a–c show that the analyses are significantly worse when only 10 ensemble members are used; several spurious cells are found in the analysis domain and the overall storm structures are poor. When 400 ensemble members are used in DR_s400 the analyses are very good. At the end of the analysis, the general shapes of the storms are quite similar to those of the truth (Figs. 7d–f and 8d–f versus Figs. 1a–c and 2a–c).

The low-level hook echo in DR_s400 is slightly sharper than in DR_4km (Fig. 7f versus Fig. 1f) and is certainly much better than that of DR_s10.

Figure 9 shows that the analyzed rms errors in DR_s400 are smaller than in DR_4km, while those of DR_s10 are much larger. In fact, the latter are larger throughout the assimilation period, and for variable θ' the analysis has more error most of the time. This behavior indicates that the cross covariance estimated from the 10 ensemble members is not reliable, and that the use of such information to update the related state variables can degrade the analysis. Such a behavior was also found in Tong and Xue (2005) for some state variables during the early assimilation cycles when random initial perturbations without smoothing were used to initialize the forecast ensemble. These experiments indicate that an ensemble size of 10 is too small, but 40 is a reasonable choice. Going to 400 members significantly increases the computational cost while the improvement in accuracy is moderate. This conclusion is also supported by the fact that the covariances obtained in DR_4km and DR_s400 are similar (Fig. 6).

We also tested different cutoff radii for covariance localization (not shown). We found that generally a

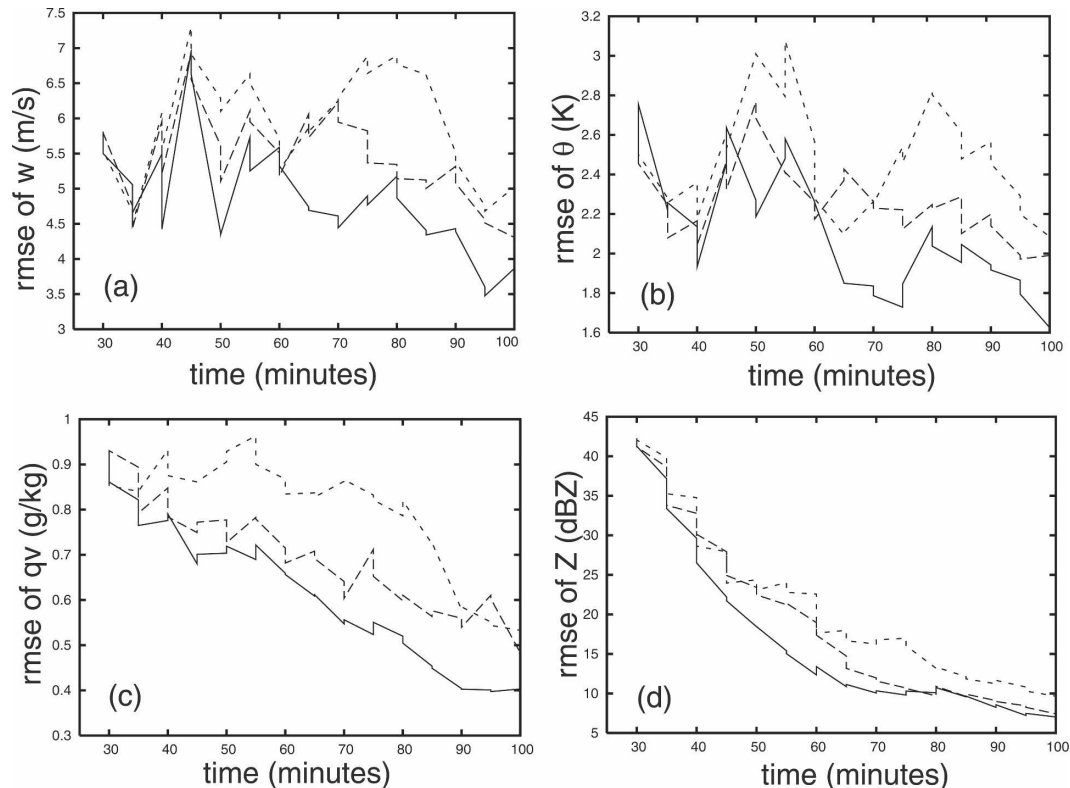


FIG. 10. Same as Fig. 3, but for experiments DR_4km (solid), DR_ob04 (long-dashed), and DR_ob08 (short-dashed) in which horizontal spacings of v_r data are 2, 4, and 8 km, respectively.

radius between 6 and 9 km works best for configurations presented in this paper. In practice, this parameter should be tuned based on the data density, ensemble size, and grid resolution. For sparse observations, a larger cutoff radius would generally be beneficial, while smaller or more local weather systems may also benefit from a smaller cutoff radius.

e. Thinning of observations

Another important factor that can affect the quality of analysis as well as the cost is the data density. The operational WSR-88D radars have a nominal range resolution of 250 m and an azimuthal resolution of 1° for radial velocity data while in the vertical there can be up to 14 elevations within the lowest 20° . For a model grid coarser than the radar data resolution, thinning the data can result in significant cost savings. In experiments DR_ob04 and DR_ob08, we thin the radar observations by assuming that they are available every fourth and eighth grid point in the horizontal on the 1-km truth grid (Table 2). The data resolution in the vertical is 2 km for both experiments. Other settings are the same as in DR_4km.

With the total number of observations reduced by a

factor of 8 in DR_ob04 compared to DR_4km, the rms errors of the analyzed fields decrease more slowly (Fig. 10). At the end of the assimilation cycles, the rms error in w is about 4.3 m s^{-1} compared to 3.8 m s^{-1} for DR_4km. For θ' , the corresponding values are about 2.0 and 1.6 K. A smaller difference is found for the rms errors of Z (Fig. 10d). In general, the analysis is still successful. At the end of the assimilation window, the analyzed midlevel flow and updraft and the surface cold pool and precipitation pattern still match the truth quite well, although the overall structures are smoother (not shown).

When the horizontal data resolution is reduced by another factor of 2 in DR_ob08, the analyzed low-level cold pool, gust front, and precipitation pattern at the end of the assimilation window differs more from those of the truth (Fig. 7i); the reflectivity core becomes broader and the hook echo is less well defined. At 70 min, the midlevel updrafts appear significantly broader and weaker (Fig. 8e versus Fig. 2b). At 100 min, there are some circulation features between the right and left movers at the midlevel that are not found in the truth (Figs. 8f and 2c). The rms errors of DR_ob08 are also higher (Fig. 10), especially during the intermediate

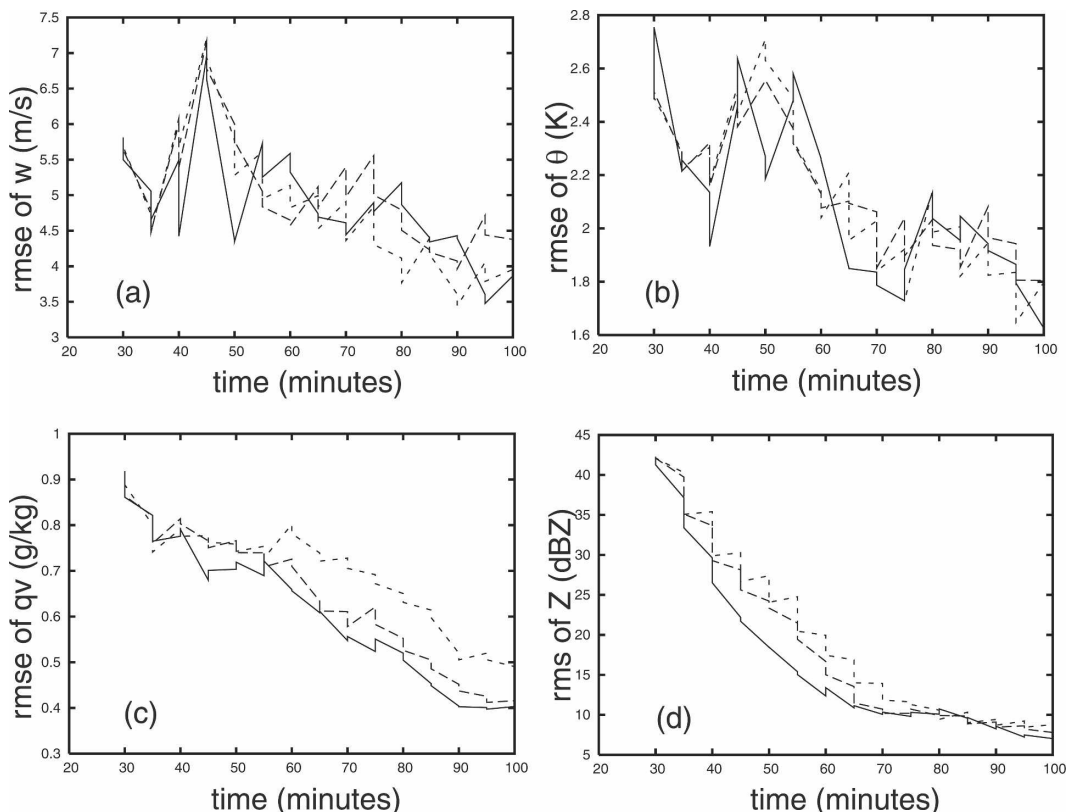


FIG. 11. Same as Fig. 3, but for experiments DR_4km (solid), DR_err03 (long-dashed), and DR_err05 (short-dashed), in which the standard deviations of radial velocity observation error are 1, 3, and 5 m s^{-1} , respectively.

cycles. The apparently larger sensitivity of the w analysis to the data density appears to be related to the decrease of horizontal divergence when the v_r resolution is reduced. Overall, the horizontal data density of 8 km appears too low for our dual-resolution analysis that uses a 4-km LR ensemble and a 1-km HR grid. A 4-km horizontal data resolution is a reasonable compromise between accuracy and cost.

f. Observation errors

In all cases above, we assumed that the standard deviation of radial velocity observation errors is 1 m s^{-1} . In reality, radial velocity observations can contain larger errors, especially when the data are contaminated by, for example, anomalous propagation, ground clutter, and/or improper velocity de-aliasing. For this reason, we test the sensitivity of our dual-resolution EnKF analysis to the observational error. In experiments DR_err03 and DR_err05 (Table 2), we assume error standard deviations of 3 and 5 m s^{-1} , respectively, and compare the results with those of DR_4km. Figure 11 shows that the rms errors for w and θ' are, on average, at similar levels for all three experiments (Figs.

11a,b), while the errors for q_v and Z are generally larger in DR_err05. The general pattern of the storm in DR_err03 is still close to the truth (not shown), but there exists significant degradation in the analysis of DR_err05, especially in terms of the midlevel updraft (the fourth row of Fig. 8). At 40 min, the updraft at the 3.5-km level is essentially absent (Fig. 8j), while at 70 min the main updraft of the right mover is too weak (only about 4 m s^{-1}) and a spurious updraft exists in between the right and left mover updrafts (Fig. 8k; cf. Fig. 2b). At the end of the assimilation (Fig. 8l), the pattern of the updrafts is closer to the truth simulation but the updrafts are still too weak (8 m s^{-1} versus 24 m s^{-1} for the right mover). Clearly, the 5 m s^{-1} error standard deviation in the radial velocity data is too large for obtaining accurate analyses of the internal structures of thunderstorms. Fortunately, the WSR-88D radial velocity errors are lower than 5 m s^{-1} when proper quality control is applied.

5. Summary and discussion

A new efficient dual-resolution (DR) data assimilation algorithm is developed based on the ensemble Kal-

man filter (EnKF) method and tested using simulated radar radial velocity data for a supercell storm. Within this algorithm, radar observations are assimilated on both high- and low-resolution grids using ensemble Kalman filter algorithm and the flow-dependent background error covariance estimated from the low-resolution ensemble. It is shown that the flow-dependent and dynamically evolved background error covariances thus estimated is effective in producing quality analysis on the high-resolution grid.

The DR method has the advantage of a much lower computational cost compared to standard full-resolution ensemble Kalman filter method. In this system, the low-resolution ensemble provides the flow-dependent background error covariance, while the single-high-resolution forecast and analysis provide the benefit of high resolution, which is important for resolving internal structures of thunderstorms. The background error covariances obtained from a 4-km 40-member ensemble are shown to be reasonably accurate, as compared to the 400-member ensemble at the same resolution. The smoothness of the covariance obtained from the 4-km resolution ensemble, as compared to those from a corresponding 1-km ensemble, does not appear to significantly degrade the quality of analysis. This is because the cross covariance among different variables is of first-order importance in radar data assimilation.

Several sensitivity experiments are conducted to test the performance of the DR algorithm, for different ensemble sizes, data densities, and data errors. For dual-resolution analysis, the use of a 4-km horizontal resolution in the ensemble and a 1-km resolution for the high-resolution analysis with an ensemble size of 40 appears to be a reasonable choice, and a 4-km horizontal resolution of thinned radial velocity data is an acceptable compromise between accuracy and cost. A data density of 8 km causes significant degradation in the analysis quality. Also, an error standard deviation of 5 m s^{-1} in the radial velocity data results in poor analyses, especially in terms of the midlevel updraft.

More experiments are needed to investigate the general tradeoffs between accuracy and cost with different resolution ratios and their impact on subsequent forecasts. The conclusions obtained in this paper may be somewhat different for atmospheric phenomena of different scales, where the model and data resolution requirements can be different. Still, the general methodology of our algorithm should be applicable to ensemble-based data assimilation at other scales. We note that, with our dual-resolution implementation, resolution-related model error does exist with the lower-resolution ensemble. Other types of model errors, such

as those related to model physics and numerics, should be considered in future studies.

Acknowledgments. The work was mainly supported by NSF Grants ATM-0331756 and ATM-0530814. The authors were also supported by NSF EEC-0313747 and a DOT-FAA grant via DOC-NOAA NA17RJ1227. Xue was further supported by NSF ATM-0331594, ATM-0608168, and Chinese NSF 40620120437. The computations were performed partly on the supercomputers of OSCER, University of Oklahoma. Three anonymous reviewers helped improve the paper. Dr. William Martin is acknowledged for proofreading the manuscript.

REFERENCES

- Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.
- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.
- Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387.
- Crook, A., and J. D. Tuttle, 1994: Numerical simulations initialized with radar-derived winds. Part II: Forecasts of three gust-front cases. *Mon. Wea. Rev.*, **122**, 1204–1217.
- Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005.
- Du, J., 2004: Hybrid ensemble prediction system: A new ensemble approach. Preprints, *Symp. 50th Anniversary Operational Numerical Weather Prediction*, College Park, MD, Amer. Meteor. Soc., P4.2.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162.
- , 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.
- Gal-Chen, T., 1978: A method for the initialization of the anelastic equations: Implications for matching models with observations. *Mon. Wea. Rev.*, **106**, 587–606.
- Gao, J., M. Xue, Z. Wang, and K. K. Droegemeier, 1998: The initial condition and explicit prediction of convection using ARPS adjoint and other retrieval methods with WSR-88D data. Preprints, *12th Conf. on Numerical Weather Prediction*, Phoenix, AZ, Amer. Meteor. Soc., 176–178.
- , —, A. Shapiro, Q. Xu, and K. K. Droegemeier, 2001: Three-dimensional simple adjoint velocity retrievals from single-Doppler radar. *J. Atmos. Oceanic Technol.*, **18**, 26–38.
- , —, 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.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.

- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
- , and —, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.
- Hu, M., M. Xue, and K. Brewster, 2006a: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698.
- , —, 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.
- Ide, K., P. Courtier, M. Ghil, and A. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189.
- Lazarus, S., A. Shapiro, and K. K. Droegemeier, 2001: Application of the Zhang–Gal–Chen single-Doppler velocity retrieval to a deep convective storm. *J. Atmos. Sci.*, **58**, 998–1016.
- Lin, Y., P. S. Ray, and K. W. Johnson, 1993: Initialization of a modeled convective storm using Doppler radar-derived fields. *Mon. Wea. Rev.*, **121**, 2757–2775.
- 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.
- Lord, S. J., E. Kalnay, R. Daley, G. D. Emmitt, and R. Atlas, 1997: Using OSSEs in the design of the future generation of integrated observing systems. Preprints, *First Symp. on Integrated Observing Systems*, Long Beach, CA, Amer. Meteor. Soc., 45–47.
- Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170.
- Ray, P. S., B. C. Johnson, K. W. Johnson, J. S. Bradberry, J. J. Stephens, K. K. Wagner, R. B. Wilhelmson, and J. B. Klemp, 1981: The morphology of several tornadic storms on 20 May 1977. *J. Atmos. Sci.*, **38**, 1643–1663.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677.
- 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.
- Tong, M., 2006: Ensemble Kalman filter assimilation of Doppler radar data for the initialization and prediction of convective storms. Ph.D. dissertation, University of Oklahoma, 243 pp.
- , 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.
- Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002: Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part I: Single-Doppler velocity retrieval. *Mon. Wea. Rev.*, **130**, 433–453.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.
- Xu, Q., C.-J. Qiu, H.-D. Gu, and J.-X. Yu, 1995: Simple adjoint retrievals of microburst winds from single-Doppler radar data. *Mon. Wea. Rev.*, **123**, 1822–1833.
- Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.*, **75**, 161–193.
- , and Coauthors, 2001: The Advanced Regional Prediction System (ARPS)—A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76**, 143–165.
- , D.-H. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170.
- , M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Oceanic Technol.*, **23**, 46–66.
- Zupanski, M., S. J. Fletcher, I. M. Navon, B. Uzunoglu, R. P. Heikes, D. A. Randall, T. D. Ringler, and D. Daescu, 2006: Initiation of ensemble data assimilation. *Tellus*, **58A**, 159–170.