Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter

F. ZHANG

Department of Atmospheric Sciences, Texas A&M University, College Station, Texas

CHRIS SNYDER AND JUANZHEN SUN

National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 12 December 2002, in final form 10 October 2003)

ABSTRACT

The ensemble Kalman filter (EnKF) uses an ensemble of short-range forecasts to estimate the flow-dependent background error covariances required in data assimilation. The feasibility of the EnKF for convective-scale data assimilation has been previously demonstrated in perfect-model experiments using simulated observations of radial velocity from a supercell storm. The present study further explores the potential and behavior of the EnKF at convective scales by considering more realistic initial analyses and variations in the availability and quality of the radar observations. Assimilation of simulated radial-velocity observations every 5 min where there is significant reflectivity using 20 ensemble members proves to be successful in most realistic observational scenarios for simulated supercell thunderstorms, although the same degree of success may not be readily expected with real observations and an imperfect model, at least with the present EnKF implementation. Even though the filter converges toward the truth simulation faster from a better initial estimate, an experiment with the initial estimate of the supercell displaced by 10 km still yields an accurate estimate of the storm for both observed and unobserved variables within 40 min. Similarly, radial-velocity observations below 2 km are certainly beneficial to capturing the storm (especially the detailed cold pool structure), but in their absence the assimilation scheme can still achieve a comparably accurate estimate of the state of the storm given a slightly longer assimilation period. An experiment with radar observations only above 4 km fails to assimilate the storm properly, but, with the addition of a hypothetical surface mesonet taking wind and temperature observations, the EnKF can again provide a good estimate of the storm. The supercell can also be successfully assimilated in the case of radar observations only below 4 km (such as those from the ground-based mobile radars). More frequent observations can help the storm assimilation initially, but the benefit diminishes after half an hour. Results presented here indicate that the vertical resolution and the uncertainty of observations, for the typical range of most of the observational radars, would have little impact on the overall performance of the EnKF in assimilating the storm.

1. Introduction

Radial-velocity and reflectivity observations from various radars are our primary tools in probing the structure and dynamics of convective-scale phenomena. How best to estimate the atmospheric state at this scale given limited radar observations has been the pursuit of many researchers since meteorological radars were first deployed. Until recently, such estimation from single-Doppler observations relied primarily on various retrieval algorithms (e.g., Qiu and Xu 1992; Laroche and Zawadzki 1994; Shapiro et al. 1995; Xu et al. 2001; Weygandt et al. 2002) that typically combine the radial velocity observations with the continuity equation to first estimate the wind fields and then derive (or retrieve) the thermodynamic variables through other physical constraints and assumptions. Three-dimensional and four-dimensional variational data assimilation methods, namely 3DVAR and 4DVAR, have also shown significant potential in convective-scale data assimilation using both simulated and real observations (Sun and Crook 1997, 1998; Gao et al. 1999).

Snyder and Zhang (2003, hereafter SZ) explored for the first time the use of an ensemble Kalman filter (EnKF) to assimilate single-Doppler radar observations in a cloud-scale model. The EnKF, which was first proposed in the geophysical literature by Evensen (1994), is an approximation to the optimal linear assimilation

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Dr. Fuqing Zhang, Dept. of Atmospheric Sciences, Texas A&M University, 3150 TAMU, College Station, TX 77843-3150.
E-mail: fzhang@tamu.edu

technique, the Kalman (1960) filter. The Kalman filter is designed to use all available information in order to produce the most accurate possible description of the state of the flow. It also provides the uncertainty in the state of the flow resulting from the uncertainties in various sources of information.

In the state estimation context described below, the update equation of the Kalman filter can be formulated as

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^f), \tag{1}$$

where $\mathbf{x}^f$represents the prior estimate or first guess, $\mathbf{x}^a$ is the posterior estimate or analysis, $\mathbf{y}$ is the observation vector, $\mathbf{H}$ is the observation operator that returns observed variables given the state, and the $\mathbf{K}$ is the so-called Kalman gain matrix defined as

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{2}$$

where $\mathbf{P}^f$ and $\mathbf{R}$ represent the background and observational error covariance, respectively. In the EnKF, the flow-dependent $\mathbf{P}^f$ is estimated through an ensemble of short-range forecasts. Further background on the EnKF can be found in SZ and references therein. A direct comparison of the EnKF with variational methods is underway (A. Caya 2003, personal communication) and will be reported elsewhere.

In the experiments of SZ, simulated observations of radial velocity from a reference supercell storm were assimilated using the EnKF, and the resulting analyses were shown to capture the reference storm accurately. Those experiments assumed that observations were taken beginning with the first radar echo and were available thereafter every 5 min at all model grid points with significant reflectivity. In reality, observations may only become available (and thus assimilation could start) at different stages of the convective storm with initial estimates of various qualities; moreover, portions of the storm, especially in the boundary layer, may be unobserved by the radar. In addition, the temporal and spatial resolution as well as the quality of the observations can vary from case to case. The present study thus seeks to explore further the potential and behavior of the EnKF at convective scales by considering more realistic initial analyses and variations in the availability and quality of the radar observations.

A general description of the implementation of the EnKF used in this study is presented in SZ. Section 2 introduces the forecast model and the reference simulation. The control experiment is presented in section 3, together with further details of the EnKF used here, which differs from that of SZ in its use of ''relaxation'' or averaging of the background and analysis covariances. The influence of the initial estimate on the analysis quality is presented in section 4, and the sensitivity experiments to data coverage, spatial and time resolution, frequency, and uncertainty of observations is presented in section 5. Summary and discussions are presented in section 6.

2. Forecast model and reference simulation

The numerical model used is that of Sun and Crook (1997) and is documented in detail there. The model was selected because of the availability of a 4DVAR system, which is used to compare against the EnKF developed herein in a separate study (A. Caya 2003, personal communication). Briefly, the model has a total of six prognostic variables including three Cartesian velocity components (u, v, w), rainwater mixing ratio (q_r), and total liquid-water mixing ratio ($q_t = q_r + q_c$, q_c is the cloud water mixing ratio), and liquid-water potential temperature (θ_l), which is conserved during condensation and evaporation processes and is defined as (Tripoli and Cotton 1981)

$$\theta_l = \theta\left[1 - \frac{L_v}{C_pT}(q_c + q_r)\right], \tag{3}$$

where L_v is the latent heat of vaporization, C_p is the heat capacity, $q_c(q_r)$ is the mixing ratio of cloud (rain) water, and θ and T are the potential temperature and temperature, respectively. The model solves the nonhydrostatic equations using θ_l as the thermodynamic variable and including only warm-rain microphysics. The equations are discretized using second-order centered differences in space and a second-order Adams–Bashforth time step.

The forecast model employs a domain of 70 km by 70 km with grid resolution of 2 km in the horizontal directions and 35 vertical layers with a grid resolution of 500 m. The model time step used in this study is 10 s. With six prognostic variables and a grid dimension of $36 \times 36 \times 35$, the state-vector dimension in the model is thus 272 160. The origin of the Cartesian coordinates is taken to lie at the lower left or southwest corner of the domain. The configuration of the model domain is shown in Fig. 1.

The reference simulation begins with a warm, moist bubble in a horizontally uniform environment; that is, u, v, θ_l, and q_r vary only with height outside the bubble, and w is zero. The perturbation liquid-water potential temperature inside the bubble is formulated as

$$\theta_l' = 4A_0\frac{z - z_b}{3}\left(1 - \frac{z - z_b}{3}\right)\cos\left[\frac{(y - y_c)\pi}{16}\right] \times \cos\left[\frac{(x - x_c)\pi}{16}\right]$$
$$\text{for} \quad -8 < x - x_c < 8 \text{ km},$$
$$-8 < y - y_c < 8 \text{ km}, \quad \text{and}$$
$$0 < z - z_b < 3 \text{ km}, \tag{4}$$

where the initial amplitude of the bubble $A_0 = 2$ K, the height of the lowest model level $z_b = 0.25$ km, and the location of the initial warm bubble $(x_c, y_c) = (22$ km, 28 km), as shown in Fig. 1. The reference simulation

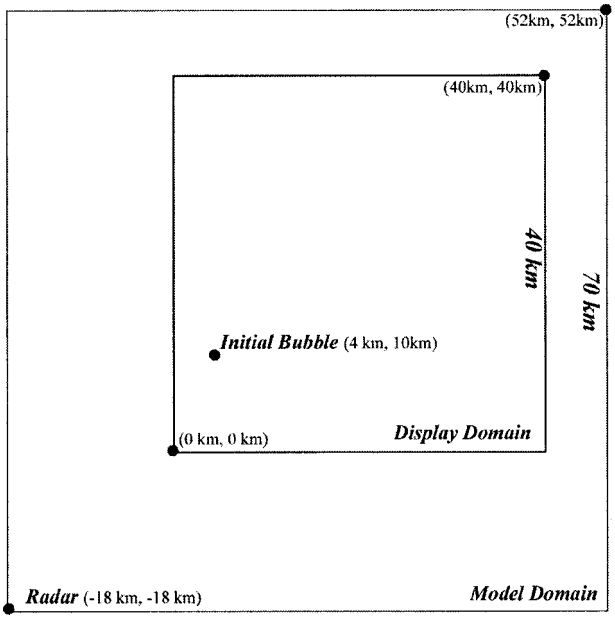

FIG. 1. The model domain (outer box) and the relative locations of the display domain in subsequent figures (inner box), the radar site, and the initial bubble.

then proceeds as a sequence of 5-min integrations as a necessary part of a perfect-model design (Mitchell et al. 2002).

The environmental sounding is that used in SZ (see their Fig. 2). It is based on the Oklahoma City, Oklahoma, sounding from 0000 UTC 25 July 1997, but 7 m s^{-1} is subtracted from the zonal wind at all levels in order to minimize the movement of the right-moving supercell through the domain.

In the control EnKF experiment described below, the assimilation cycle starts at time t_0 = 40 min after the initialization of the warm bubble. Figures 2a–e show the vertical velocity (w) at 5 km above ground level (AGL) at t_0 and every 20 min after t_0 in the reference simulation. The corresponding 5-km-AGL rainwater mixing ratio (q_r) is shaded in Figs. 3a–e. Also shown in these figures are the 5-km-AGL relative vorticity (contours in Fig. 3) and the surface liquid-water potential temperature (contours in Fig. 2), which are the primary indicators of the midlevel vortex and low-level cold pool structure of the supercell, respectively. These quantities from the reference simulations will be used to evaluate the assimilation results at these times in the following sections.

The warm bubble initiates a convective cell, which strengthens over the first 20 min and begins to rain at 30 min. As is typical for soundings such as in Fig. 2 of SZ, the initial cell forms a downdraft and then splits into a strong primary cell that moves to the right of the environmental shear and a weaker, secondary cell moving to the left of the shear. Here, the initial storm splits around 60 min after the initialization of the warm bubble.

3. The control experiment

In the control assimilation experiment (CNTL), simulated Doppler radar wind observations are taken from the reference simulation. The radar is located at the

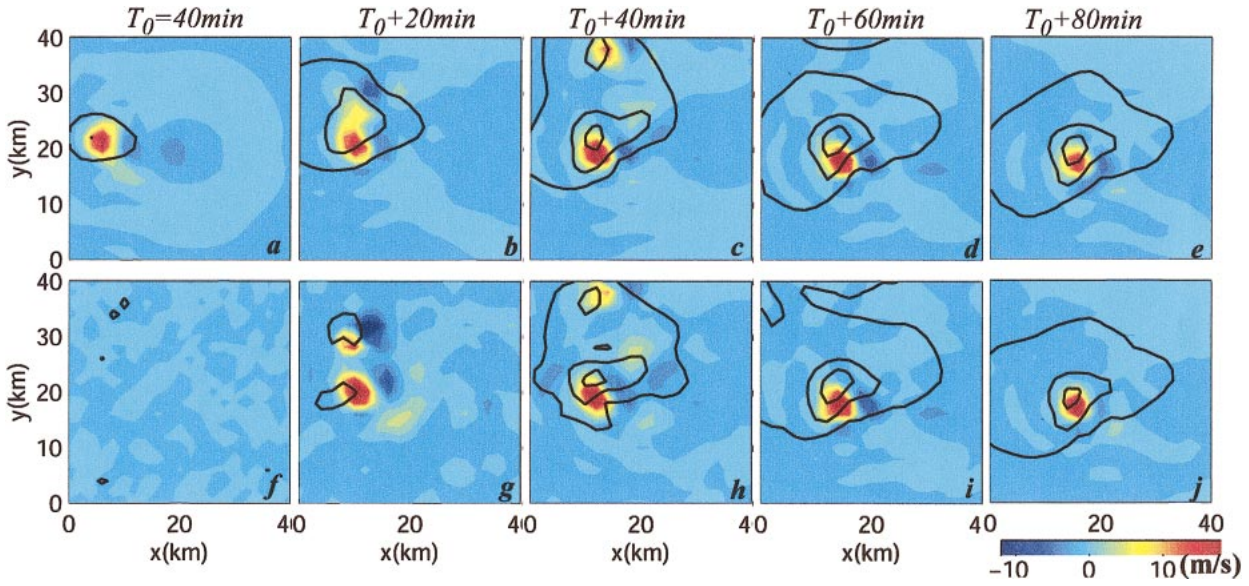

FIG. 2. The 5-km-AGL vertical velocity (colored) and surface perturbation liquid-water potential temperature (black lines, every 2 K) for (a)–(e) the reference simulation and (f)–(j) the EnKF analysis in the CNTL experiment. Shades of red and blue indicate upward and downward motion, respectively, with gradations of color every 2 m s^{-1}. Fields are shown at t_0, t_0 + 20, t_0 + 40, t_0 + 60, and t_0 + 80 min, labeled on top as *T_0=40min*, etc. (t_0 = 40 min).

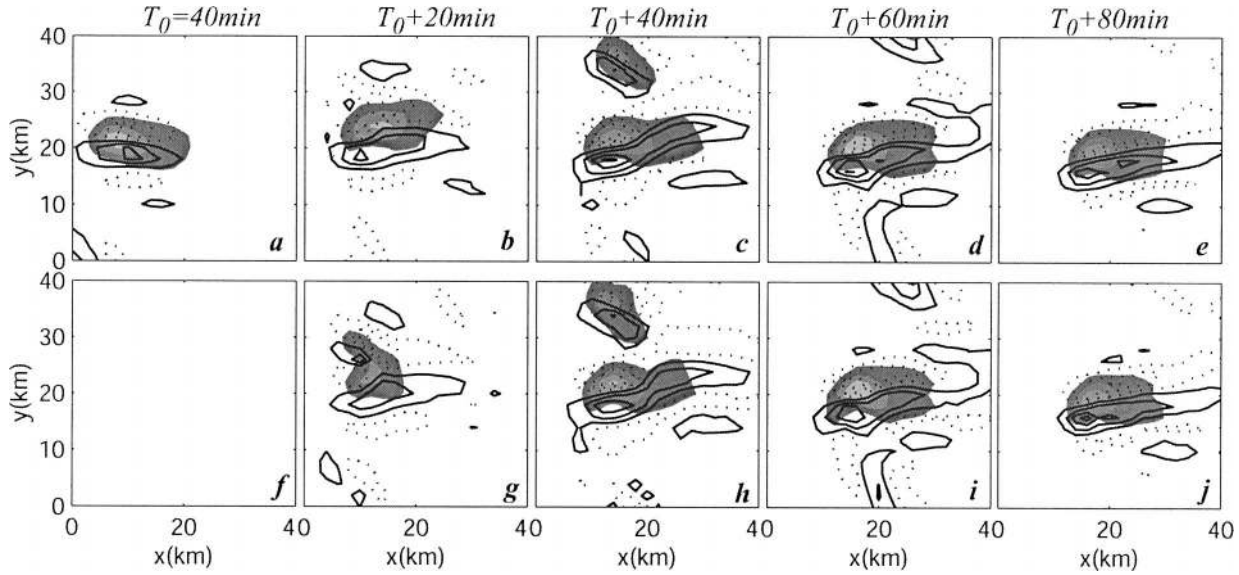

FIG. 3. The 5-km-AGL rainwater mixing ratio (shaded, every 2 g kg^{-1}) and relative vorticity (every 0.001 s^{-1}; solid lines are positive and dotted lines are negative) in (a)–(e) the reference simulation and (f)–(j) the EnKF analysis for the CNTL experiment. Fields are shown at t_0, t_0 + 20, t_0 + 40, t_0 + 60, and t_0 + 80 min, labeled on top as *T_0=40min*, etc. (t_0 = 40 min).

southwest corner of the computational domain as indicated in Fig. 1 and measures the radial velocity in a spherical coordinate system centered on the radar. As in SZ, we assume that (a) the observations have independent, Gaussian random errors of zero mean and variance of 1 m s^{-1}, and (b) radial velocity is observed only at grid points where there is significant reflectivity ($q_r > 0.13$ g kg^{-1}). Since the convective cell in the reference simulation covers only a small part of the domain (Figs. 3a–e), a typical observation set consists of $O(10^3)$ individual observations (refer to Table 1 in section 5). No radar reflectivity observations are assimilated in the current study.

Given velocities on the computational grid, the radial velocity v_r is calculated by first averaging u, v, and w, which are staggered, to a central point and then using the averaged velocities: $v_r = (x/r)u + (y/r)v + (z/r)w + \varepsilon$, where (x, y, z) denotes the location of the observation, r is its distance from the radar, and ε is the random error with zero mean and standard deviation of 1 m s^{-1}. The dependence of v_r on the fall speed of rain has been neglected for simplicity.

Observation sets, consisting of v_r at all points at which q_r exceeds the threshold given above, are available at t_0 and every 5 min thereafter. The forecast model is assumed to be perfect; that is, the same numerical model produces the forecasts and the reference simulation from which observations are taken. In addition, each observation is allowed to influence the analysis only within a sphere of radius 3 km, and 20 ensemble members are used, which differs from a radius of 4 km and 50 members used in SZ.

The EnKF used here also differs from that of SZ in the updating, given new observations, of the deviations of each member from the ensemble mean. Let deviations from the mean be denoted by primes. The EnKF of SZ provides an algorithm for updating the forecast deviation $(\mathbf{x}^f)'$, for each member to produce an analysis deviation $(\mathbf{x}^a)'$. [Typically $(\mathbf{x}^a)'$, is smaller than $(\mathbf{x}^f)'$, reflecting the reduction of uncertainty after assimilating observations.] We have chosen to obtain a modified analysis deviation by ''relaxing'' or weighting $(\mathbf{x}^f)'$, and $(\mathbf{x}^a)'$:

$$(\mathbf{x}^a_{\text{new}})' = (1 - \alpha)(\mathbf{x}^a)' + \alpha(\mathbf{x}^f)', \qquad (5)$$

where $\alpha = 0.5$. The modified analysis deviations are then used to initialize the ensemble members for forecasts to the next assimilation time.

Our use of (5), which overestimates the uncertainty in the analysis, is motivated by the tendency for the EnKF to underestimate the uncertainty of the analysis (Burgers et al. 1998; van Leeuwen 1999) if the ensemble size is small or if the initial estimate is poor (i.e., ensemble mean error is greater than the spread). Other approaches to dealing with this problem are the covariance inflation of Anderson (2001) or the use of a ''double'' ensemble in which covariances estimated from one-half of the ensemble are used in updating the other half (Houtekamer and Mitchell 1998). The use of (5) has the advantage relative to covariance inflation that variance of the analysis is artificially increased only where observations exist. However, both the relaxation method and the covariance inflation give up the important property of the EnKF that it converges to the Kalman filter as the ensemble size becomes large and thus are only of practical importance.

Although qualitatively similar assimilation results can be achieved even without the relaxation (as shown by

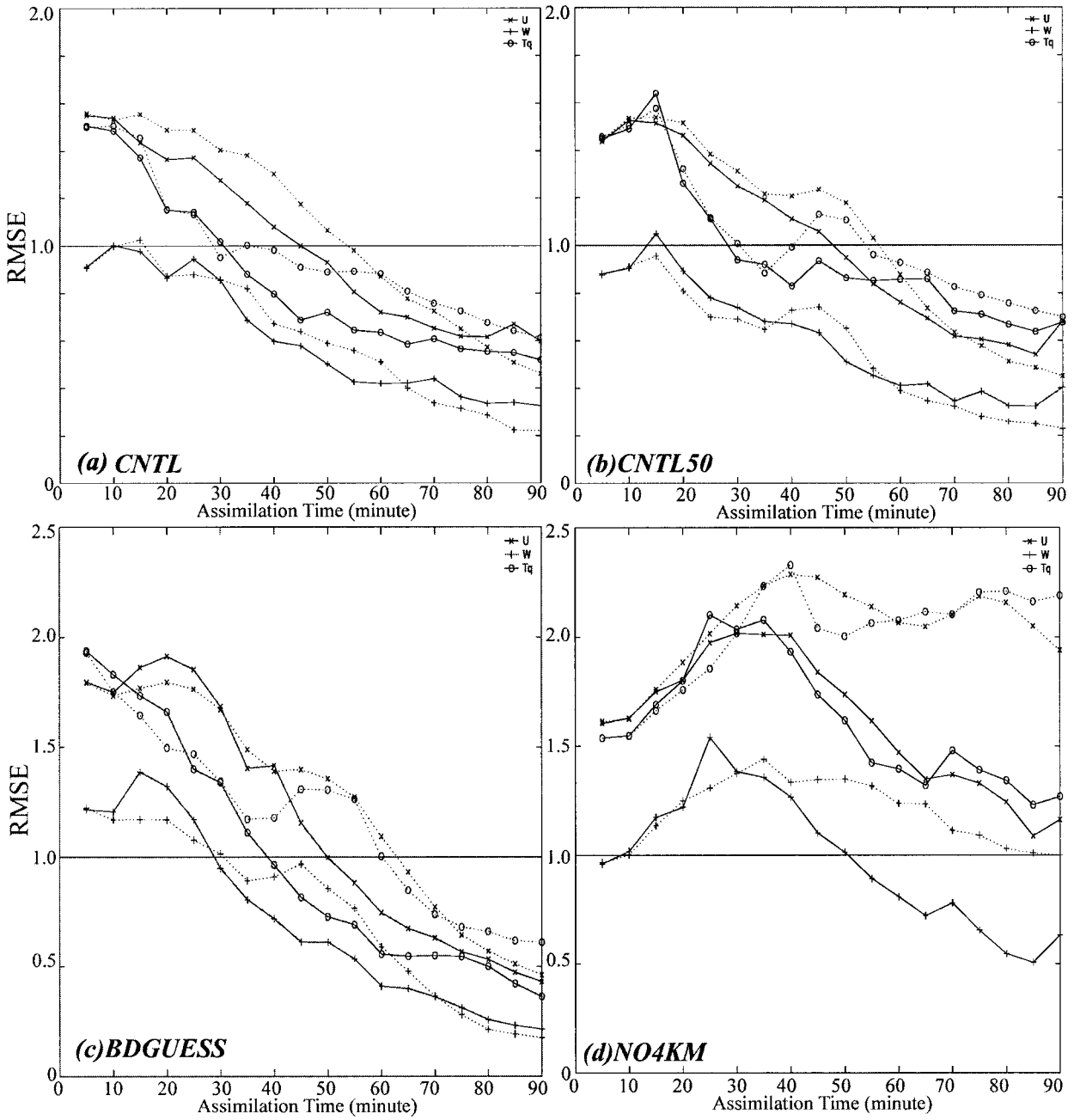

FIG. 4. Root-mean-square error (RMSE) of the analysis ensemble mean from the reference simulation for zonal wind (m s^{-1}, marked with "+"s), vertical velocity (m s^{-1}, with "×"s), and liquid-water temperature (K, with "○"s) in the EnKF expts with (solid lines) and without (dotted lines) relaxation of posterior covariances for (a) CNTL, (b) CNTL50, (c) BDGUESS, and (d) NO4KM.

the dotted curves in Fig. 4a), we found that the relaxation method improves quantitatively the assimilation efficiency, especially over the first 40–50 min. Similar improvements can also be found in the experiment using 50 ensemble members (CNTL50) in Fig. 4b. The use of (5) also becomes particularly helpful when the initial estimate is poor (e.g., "BDGUESS," to be detailed in section 4b; Fig. 4c) and fewer observations are assimilated (e.g., "NO4KM," to be detailed in section 5a; Fig. 4d). The root-mean-square error results (RMSE) presented in Fig. 4 and all subsequent figures are averaged over the entire computational domain in both horizontal and vertical directions. Since there are noticeable variations from realization to realization, the advantage of using covariance relaxation is further demonstrated in the filter performance averaged over eight additional realizations of both CNTL and CNTL50, with and without the use of (5), respectively (Figs. 5a,b).

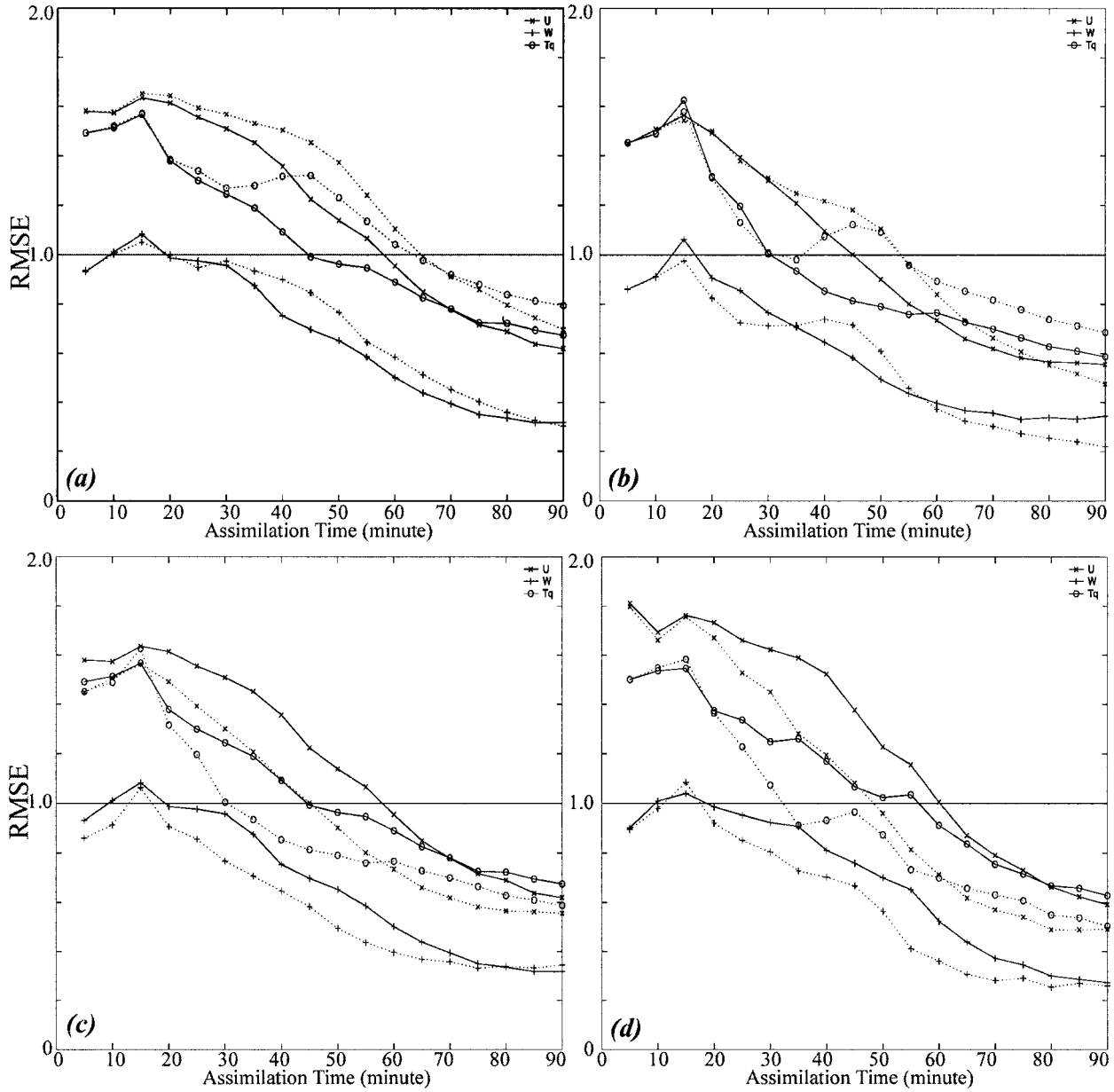

FIG. 5. (a) As in Fig. 4a, except averaged over eight realizations. (b) As in Fig. 4b, except averaged over eight realizations. (c) RMSE of ensemble analysis mean from the reference simulation using 20 (solid) and 50 (dotted) members, respectively. (d) As in (c), but for rmse of ensemble forecast mean.

Thus (5) will be used in all of the subsequent sensitivity experiments. Figures 5c and 5d show the direct comparisons of the filter performance [in terms of ensemble forecast/analysis mean (prior/posterior mean estimate of EnKF)] with 20 and 50 members averaged over the same sets of realizations in Figs. 5a and 5b. Though, on average, there are noticeable improvements by using the EnKF with a larger ensemble size, the EnKF performs reasonably well with only 20 ensemble members (Figs. 5c,d). All experiments hereafter will use a 20-member ensemble forecast for efficiency.

We begin assimilating observations at $t_0 = 40$ min, by which time the initial cell has developed in the reference simulation (Figs. 2a, 3a). The ensemble is initialized 5 min prior to the first observations. The ensemble mean is taken to be the environmental sounding shown in Fig. 2 of SZ (which is assumed known) together with the warm bubble used to initialize the reference simulation 35 min earlier, at $t = 0$. Each ensemble member is then initialized by adding realizations of Gaussian noise to the ensemble mean. This noise is independent at each grid point, has zero mean, and has

standard deviation of 3 m s^{-1} for each component of velocity and 3 K for liquid-water potential temperature. Water vapor and cloud water are initialized using the environmental sounding at each level.

We choose to generate the ensemble initial conditions with this approach because of its simplicity. One drawback, however, is that even at later times the members are noisy when compared to the reference simulation. For example, individual members may contain spurious convective cells (see Fig. 6 of SZ), and time series at a point indicate more variability in the members than in the reference simulation. This additional noise in the members is loosely analogous to the generation of spurious gravity waves in global or synoptic-scale data assimilation. As discussed in SZ, more sophisticated ensemble initiation methods can improve the performance of the EnKF in convective-scale assimilation, and these methods remain an important area for research.

The initialization of the ensemble differs from that in SZ by the use of the warm-bubble perturbation in the initial ensemble mean. The addition of the initial warm bubble in the ensemble mean expedites the development of useful covariance structures and thus leads to quicker assimilation convergence. Without the bubble, the filter yields similar results given slightly longer assimilation time (~15 min; not shown). The assimilation also becomes less robust, in that variance of the RMSE from one realization of the experiment to another is larger.

Results of CNTL are shown in Figs. 2 and 3. The initial and assimilated ensemble mean of the 5-km-AGL vertical velocity and rainwater mixing ratio every 20 min are shown in Figs. 2f–j and 3f–j, respectively. After four assimilation cycles (or 20 min), the storm is clearly indicated not only in the partially observed vertical velocity field (Fig. 2g) but also in the unobserved rainwater field (Fig. 3g). The 5-km-AGL vorticity (midlevel vortex) derived from the assimilated ensemble mean also resembles closely that from the reference simulation. At other levels, however, the assimilated storm still differs significantly from the reference storm at this time. In particular, since there are very few observations taken near the lower boundary, the reference surface cold pool (Fig. 2b) is mostly missed in the assimilated ensemble mean (Fig. 2g). Nevertheless, right below the convective cells, small pockets of cold outflows begin to appear in the assimilated ensemble mean as a result of the developing covariance (correlation) between the low-level radial velocity and liquid-water temperature fields as well as a dynamic/thermodynamic consequence of the assimilated storm(s).

At $t = t_0 + 40$ min, after another four cycles of assimilation, the assimilated ensemble mean has successfully captured most characteristic features of the reference state with two completely separated convective cells. The location, structure, and strength of the retrieved supercells, as indicated by the 5-km-AGL vertical velocity, relative humidity, and rainwater mixing ratio (Figs. 2h and 3h), match well with the true storm at this time. The difference between the analysis and reference simulation in the shallower, faster left-moving cell (which is on the northern edge of Figs. 2c and 3c) is slightly larger than in the dominant right-moving cell. Again, the retrieval of the surface outflow boundaries is less successful, although promising cold pools with similar structure to those in the reference state begin to form in the right location. As more observations are assimilated over time, the surface cold pools in the analysis mean gradually catch up and closely match the reference storm, and the actual storm becomes nearly indistinguishable from the reference run (Figs. 2i–j, 3i–j).

The performance of the control EnKF experiment can be best summarized by examining the domain-averaged RMSE between the assimilated (analysis) ensemble mean and the reference (truth) run. The RMSE of u, w, and the liquid-water potential temperature over the 90-min assimilation period are shown in Fig. 4a (solid curves). Except for the first several cycles, the RMSE drops consistently and rapidly for all variables. The RMSE of u drops below the observational error (1.0 m s^{-1}) after ~40 min (or a total of eight assimilation cycles) and continues to drop below 0.7 m s^{-1} after 60 min and appears to be stabilized afterward. The RMSE of v follows that of u closely (not shown), while the RMSE of w is approximately two-thirds of those of the horizontal winds and stabilizes at a smaller value (Fig. 4a). The RMSE of the unobserved prognostic variables (e.g., the θ_l also shown in Fig. 4a) decreases in a similar manner to the partially observed velocity variables.

4. Impacts of initial estimate

The initial estimate of most data assimilation systems usually comes from the previous short-term forecast or larger-scale forecast models. In either case, there is a strong possibility that a key meteorological feature (e.g., the supercell in this case) may exist in the initial estimate but occur in the wrong location. The dependence of the EnKF on the initial estimate is tested in this section under different realistic scenarios for the convective-scale data assimilation. We begin with an assimilation experiment (''GDGUESS'') in which the initial ensemble mean is a good initial estimate of the reference state at $t_0 = 40$ min.

a. EnKF with good initial estimate (GDGUESS)

In GDGUESS, there exists a supercell storm in the initial ensemble mean (Figs. 6a, 7a). The initial ensemble mean was formed by adding to the reference storm (at $t_0 = 40$ min) 3 m s^{-1} random error in each component of the wind and 3-K random error in the liquid-water potential temperature field, throughout the model domain. Random errors of 3 m s^{-1} or 3 K were then added to the perturbed ensemble mean to initiate 20 ensemble members. The GDGUESS experiment thus differs from

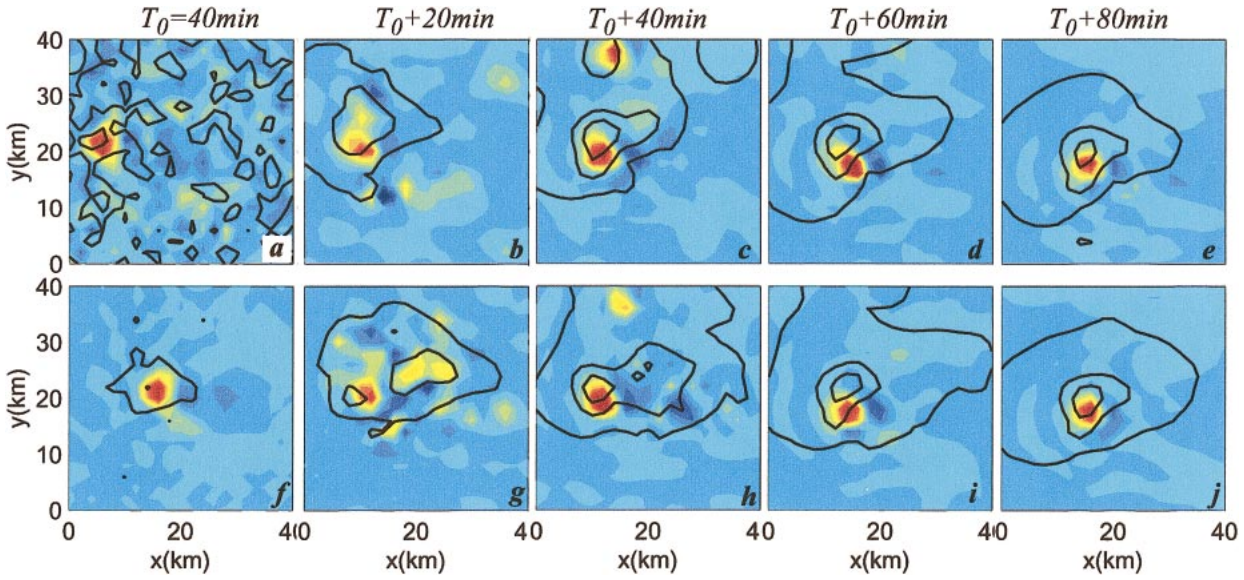

FIG. 6. The 5-km-AGL vertical velocity (colored) and surface perturbation liquid-water temperature (black lines every 2 K) in (a)–(e) GDGUESS and (f)–(j) BDGUESS. Shades of red and blue indicate upward and downward motion, respectively, with gradations of color every 2 m s^{-1}. Fields are shown at t_0, t_0 + 20, t_0 + 40, t_0 + 60, and t_0 + 80 min, labeled on top as *T_0=40min*, etc. (t_0 = 40 min).

CNTL, in which there is no prior information of the supercell except for the background sounding and a warm bubble.

As we can see from Figs. 6b and 7b, in GDGUESS the ensemble mean will be quickly drawn closer to the reference state after 20 min or four assimilation cycles. The surface cold pool, which exists in the initial ensemble mean, albeit greatly distorted (Fig. 6a), will be mostly recovered in the analysis ensemble mean at this time (Fig. 6b), much earlier than in the CNTL experiment (Fig. 2g). The ensemble mean is continuously improving to a higher accuracy after 40 min (Figs. 6c, 7c). The advantage of having a better initial estimate in GDGUESS over that in CNTL vanishes after an hour of assimilating observations (Figs. 6d–e, 7d–e) because both analyses are nearly errorless compared to the reference simulation at this time (Figs. 2d,e). The better performance of the filter with the presence of a ''noisy'' storm in the right place can be seen even more clearly in the evolution of RMSE compared to CNTL for the first 40 min (Fig. 8a).

We have also performed several forecast experiments, similar to those described in SZ, to examine the error growth characteristics at various stages using the analysis from the assimilation as the initial condition but without subsequent assimilation. We found that, even with a ''good'' initial estimate in GDGUESS, the RMSE of the pure ensemble forecast mean relative to the reference simulation can grow to a magnitude of 3 m s^{-1} for velocities and 3 K for liquid-water temperature within an hour because of the random errors in both the ensemble mean and individual members (refer to Fig. 9 in SZ). There is no doubt that the filter is adequately and continuously drawing information from the observations to reduce the RMSE and keep the analysis from diverging from the reference solution.

b. EnKF with poor initial estimate (BDGUESS)

In experiment ''BDGUESS,'' the initial ensemble mean is formed by displacing the reference (truth) simulation at T_0 = 40 min by a distance of 10 km so that the initial mean has a cell of the correct form but in the wrong location (Figs. 6f, 7f). Random errors of 3 m s^{-1} or 3 K again are then added to the ensemble mean with dislocated supercell to initiate 20 ensemble members. In this case, the assimilation process must develop a cell in the correct location and suppress the incorrectly located initial storm.

After 20 min, both the newly assimilated storm and the ''false'' initial storm coexist in the ensemble mean (Figs. 6g, 7g); the RMSE of each prognostic variable is thus considerably larger than in CNTL (Fig. 8b). After 40 min of assimilation of v_r observations, however, the ensemble mean represents the two cells from the reference simulation with fair success (Figs. 6h, 7h). The remnants of the preexisting cell can still be vaguely seen in the analysis mean, especially on the intensity and orientation of surface cold outflow boundary (Figs. 6g, 7g). After another 20 min with another four cycles of observations assimilated, the filter again achieves nearly the same accuracy (Figs. 6i, 7i) as in CNTL (Figs. 2i, 3i) and GDGUESS (Figs. 6d–e, 7d–e). Some small differences persist in the surface cold pool at this time (Fig. 6i), but these largely disappear at later times (Figs. 6j, 7j). The convergence of the RMSE (Fig. 8b) to values comparable to that of CNTL or GDGUESS (Fig. 8a) demonstrates the resilience of the EnKF under different

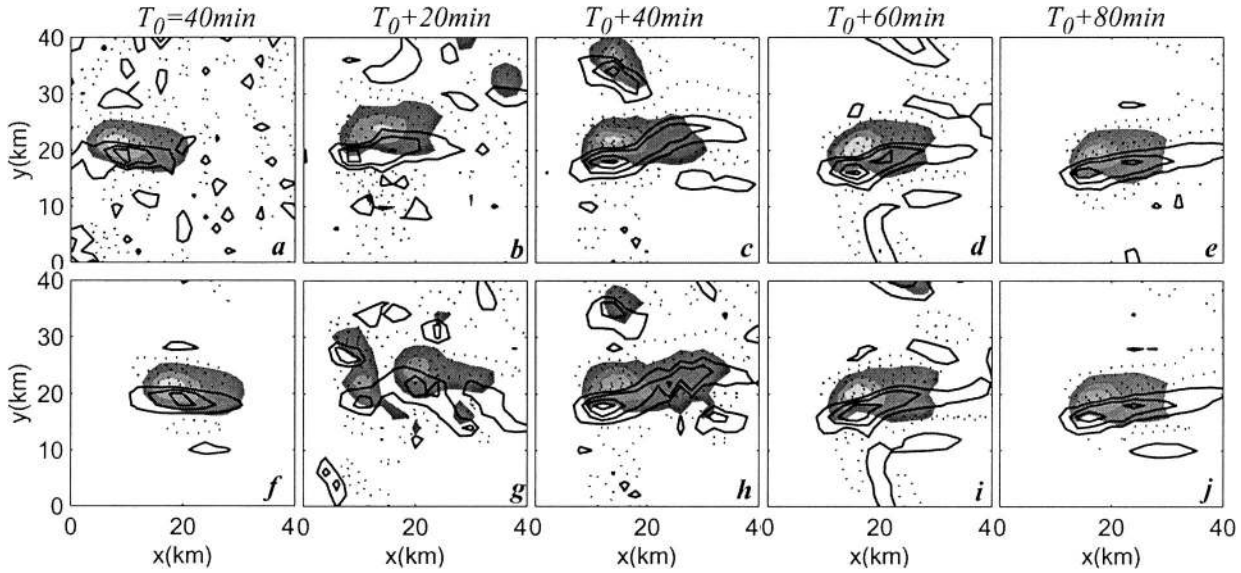

FIG. 7. The 5-km-AGL rainwater mixing ratio (shaded, every 2 g kg^{-1}) and relative vorticity (every 0.001 s^{-1}; solid lines are positive and dotted lines are negative) in (a)–(e) GDGUESS and (f)–(j) BDGUESS. Fields are shown at t_0, t_0 + 20, t_0 + 40, t_0 + 60, and t_0 + 80 min, labeled on top as *T_0=40min*, etc. (t_0 = 40 min).

initial estimates. It is worth mentioning that the ensemble filter will be less efficient in suppressing the preexisting, incorrectly located storms (especially their associated outflow boundaries) if they are beyond the radius of influence of any observations.

c. EnKF after the supercell splitting (LATE1 and LATE2)

The effect of the initial estimate can be further tested by beginning the assimilation after the supercell splits in the reference run. Experiments ''LATE1'' and ''LATE2'' both use the same reference state as in CNTL but start when the truth storm has just gone through splitting at 70 min after the warm-bubble initiation (Figs. 9a–e). These two experiments differ in their initial estimates. The initial ensembles in LATE1 are exactly the same as in CNTL and thus include no information about the initial storm(s) except for the environmental sounding and the warm bubble (Fig. 9f). In LATE2, the initial ensemble mean is taken to be the reference storm 30 min earlier, and the 20 ensemble members are initialized by adding random errors (again, of 3 m s^{-1} or 3 K) to the mean.

The ensemble filter in LATE1 behaves very similarly to CNTL. The dominant right-moving cell in the assimilated ensemble mean is present after 20 min of assimilation, while the assimilation of the surface outflow boundaries again lags behind (Fig. 9g). Unlike BDGUESS, the existence of an earlier storm in the initial ensemble mean turns out to be helpful in LATE2 (Fig. 9l), at least during the first 40 min of assimilation compared to LATE1. After 40–60 min of assimilation, the ensemble mean estimates in both experiments asymptote to similar levels of analysis error (Figs. 9c–d,h–i,m–n), and there are hardly any differences at all after 80 min of assimilation (Figs. 9e,j,l).

In summary, the filter is robust under the various initial conditions tested in this section, even though a better estimate of the initial ensemble mean can be helpful in shortening the number of assimilation cycles required to reach an analysis of a given accuracy. Typically, all initial estimates result in analyses of similar quality after 60 min of assimilation, or 12 cycles.

5. Observational coverage, resolution, frequency, and accuracy

The ensemble Kalman filter combines information from the initial estimate, the dynamics of the forecast model, and the observations to get the best estimate and the associated uncertainty. The quality (coverage, resolution, frequency, and accuracy) of radar observations of different convective storms varies from case to case, and this variation will influence our ability to estimate the true state. In this section, some more realistic observational scenarios are tested using the EnKF. These observing system simulation experiments should at least give an indication of the EnKF performance. We will use the same reference simulation and the same initial ensembles as those in CNTL for all the sensitivity experiments investigated in this section. We begin with sensitivities to the boundary layer observations.

a. Missing boundary layer observations (NO2KM and NO4KM)

In real cases, convective storms are often at a distance from the radar site, and thus the lowest level of the

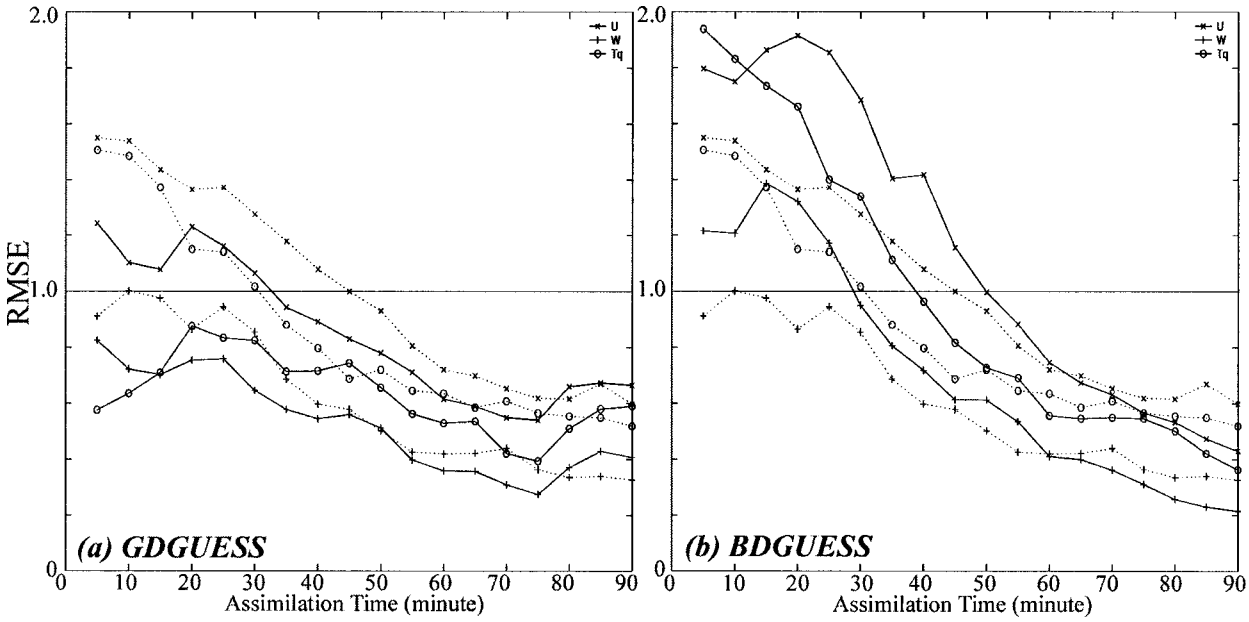

FIG. 8. As in Fig. 4, but for the RMSE in (a) GDGUESS and (b) BDGUESS. RMSE of CNTL is also shown as dotted lines in (a) and (b) for easy comparison.

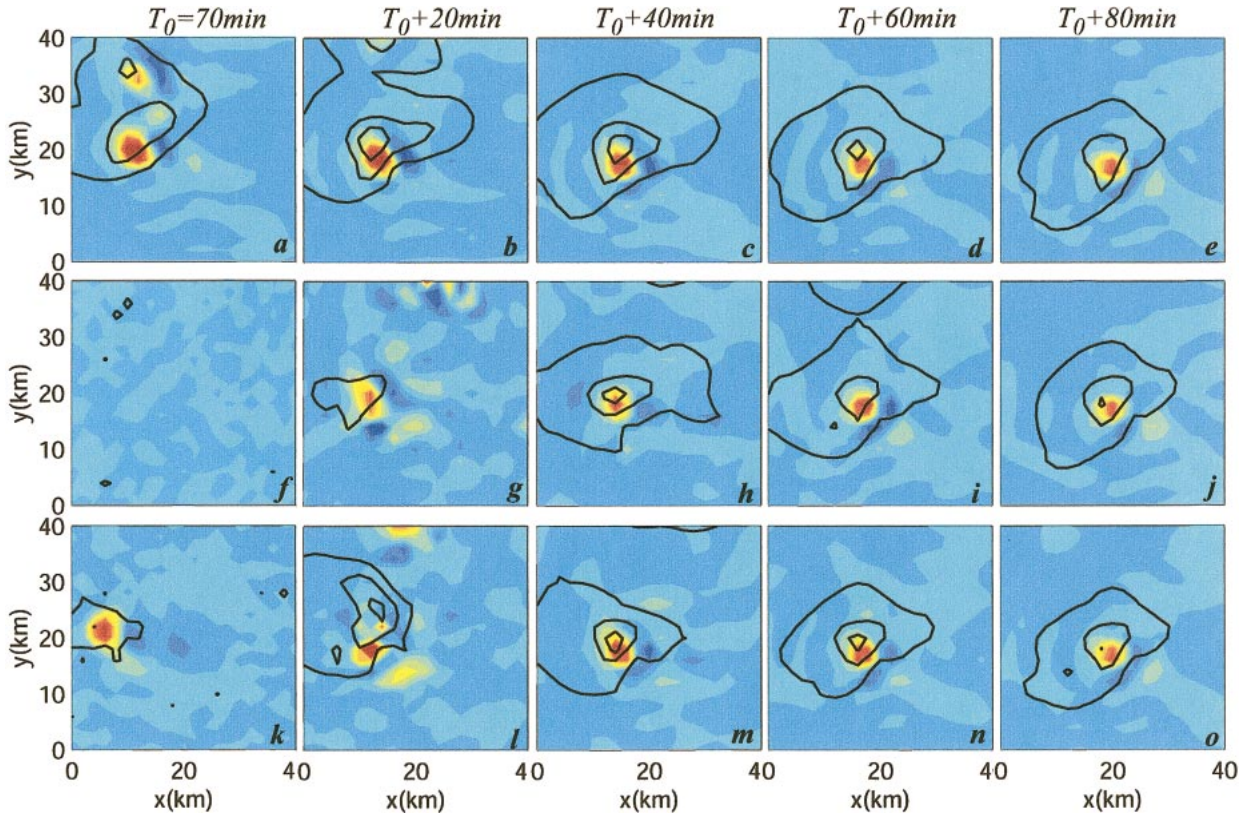

FIG. 9. The 5-km-AGL vertical velocity (colored) and surface perturbation liquid-water temperature (black lines every 2 K) for (a)–(e) the reference simulation and the EnKF analyses in (f)–(j) LATE1 and (k)–(o) LATE2. Shades of red and blue indicate upward and downward motion, respectively, with gradations of color every 2 m s^{-1}. Fields are shown at t_0, t_0 + 20, t_0 + 40, t_0 + 60, and t_0 + 80 min, labeled on top as *T_0=70min*, etc. (t_0 = 70 min).

TABLE 1. Number of observations for experiments CNTL, NO2KM, NO4KM, and ONLY4KM.

Expts	CNTL	NO2KM	NO4KM	ONLY4KM
t_0 + 20 min	1669	1470	1162	667
t_0 + 40 min	2458	2173	1716	965
t_0 + 60 min	2923	2573	2021	1176
t_0 + 80 min	2166	1898	1466	916

storm is likely missing from the radar observations. Two different experiments have been designed. Experiments ''NO2KM'' and ''NO4KM'' are identical to CNTL, except that there are no observations taken below 2 and 4 km AGL, respectively. The number of available observations at 20-min intervals for experiments CNTL, NO2KM, and NO4KM is shown in Table 1.

As might be expected, the analysis errors of the ensemble mean decrease more slowly in NO2KM than in CNTL. Without observations below 2 km, the analysis retains the signature of the right-moving storm and a midlevel vortex after 20 min of assimilation, but the reference cold pool is virtually missing in the analyzed ensemble mean (Figs. 10a, 11a). After another 20 min or four cycles of assimilation with the filter, the ensemble mean analysis begins to capture the basic structure and characteristics of the reference splitting storms while the surface cold pool catches up quickly (Figs. 10b, 11b). Nevertheless, the errors in NO2KM are still significant and larger than in CNTL; the vertical velocity in the shallower left-moving storm is overestimated, and that in the right mover is too weak at 5 km AGL. The analysis in NO2KM captures the reference storm ac-

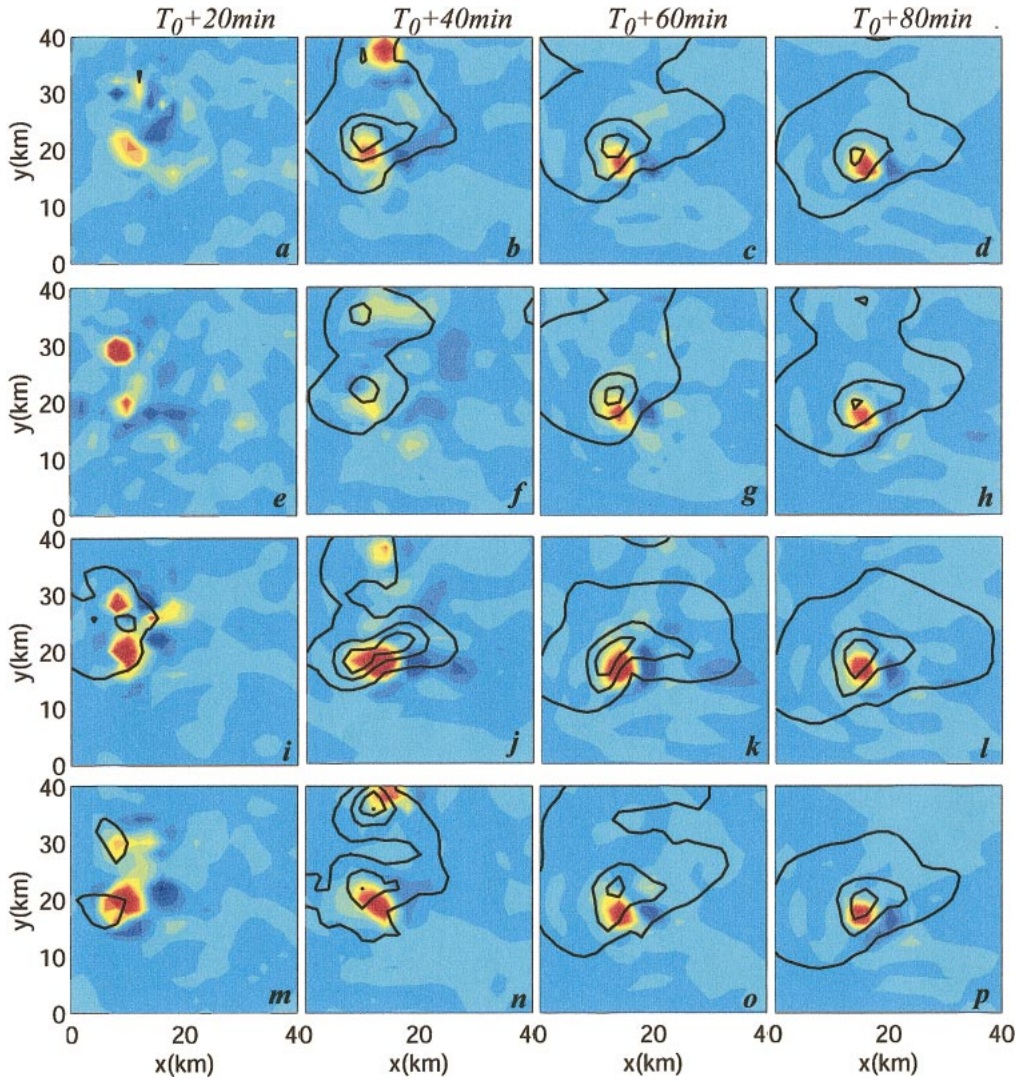

FIG. 10. The 5-km-AGL vertical velocity (colored) and surface perturbation liquid-water temperature (black lines every 2 K) in (a)–(d) NO2KM, (e)–(h) NO4KM, (i)–(l) SFCOBS, and (m)–(p) ONLY4KM. Shades of red and blue indicate upward and downward motion, respectively, with gradations of color every 2 m s^{-1}. Fields are shown at t_0 + 20, t_0 + 40, t_0 + 60, and t_0 + 80 min, labeled on top as T_0+*20min*, etc. (t_0 = 40 min).

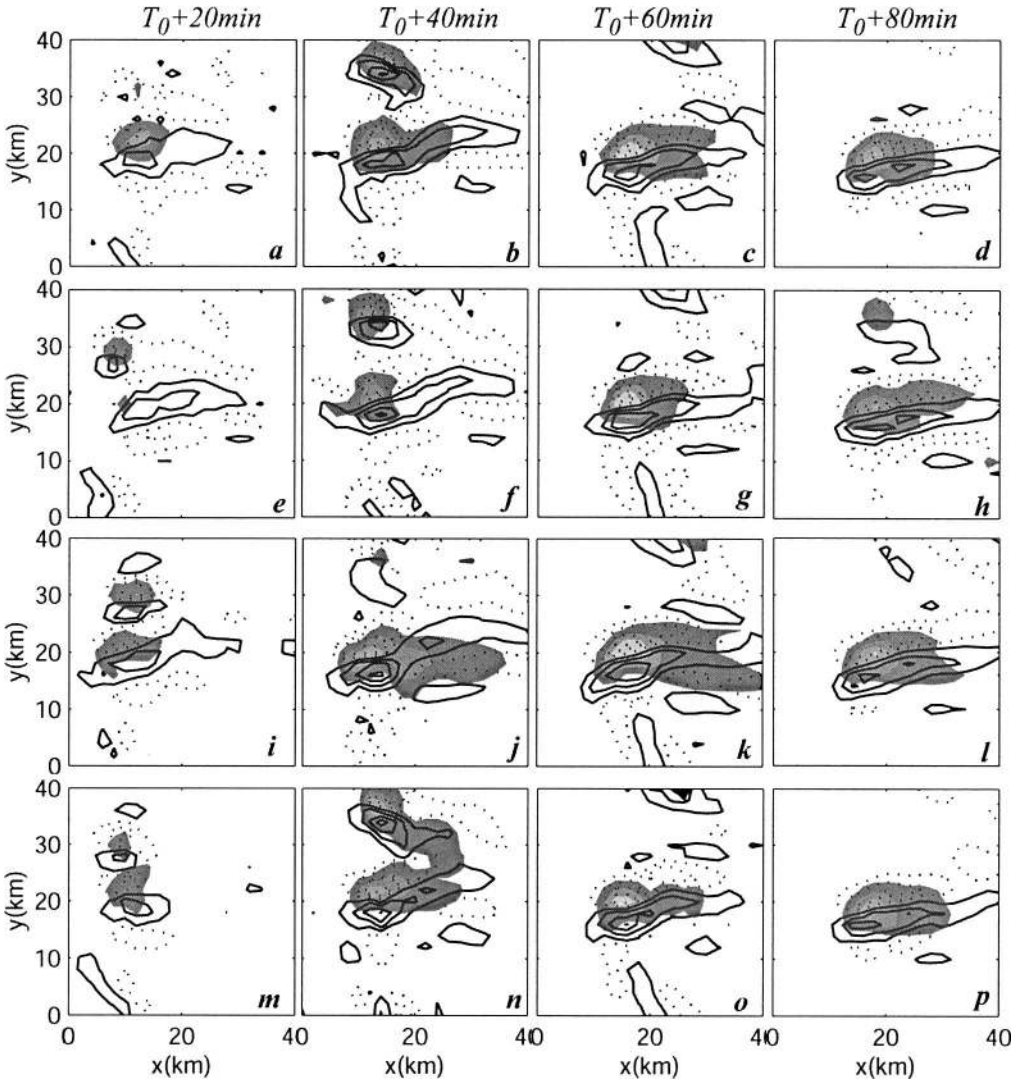

FIG. 11. The 5-km-AGL rainwater mixing ratio (shaded, every 2 g kg^{-1}) and relative vorticity (every 0.001 s^{-1}; solid lines are positive and dotted lines are negative) in (a)–(d) NO2KM, (e)–(h) NO4KM, (i)–(l) SFCOBS, and (m)–(p) ONLY4KM. Fields are shown at t_0 + 20, t_0 + 40, t_0 + 60, and t_0 + 80 min, labeled on top as *T_0+20min*, etc. (t_0 = 40 min).

curately after 60 to 80 min of total assimilation time (Figs. 10c,d), though the RMSE of each variable remains slightly larger than in CNTL throughout the entire 90-min assimilation period (Fig. 12a).

Results are markedly different in the experiment without observations below 4 km AGL (NO4KM). After 20 min, the reference storm has little resemblance to the analysis of the 5-km-AGL vertical velocity and rainwater or the surface cold pool except for some weak signal in the midlevel vortex (Figs. 10e, 11e). After 40 min, the analysis captures two cells in approximately the right locations, but each is much too weak (Figs. 10f, 11f). The assimilated ensemble mean differs noticeably from the reference storm even after 60–80 min (Figs. 10g–h, 11g–h). In contrast to previous experiments, the RMSE of the analyzed winds in NO4KM are higher than the observational errors throughout the 90-min assimilation (Fig. 12b). The results of NO4KM suggest that, given the limited lifetime of supercell thunderstorms, assimilation of observations only available at very high altitude (e.g., above 4 km AGL) may not be sufficient to estimate the phenomena accurately even with a perfect forecast model.

b. Impacts of surface observations (SFCOBS)

Experiment ''SFCOBS'' is similar to NO4KM, but a hypothetical 10 km by 10 km surface mesonet is made available to take observations of horizontal winds and liquid-water potential temperature at the lowest model level (250 m AGL). The mesonet stations are deployed on a regular grid, starting from (−8 km, −8 km) in Fig.

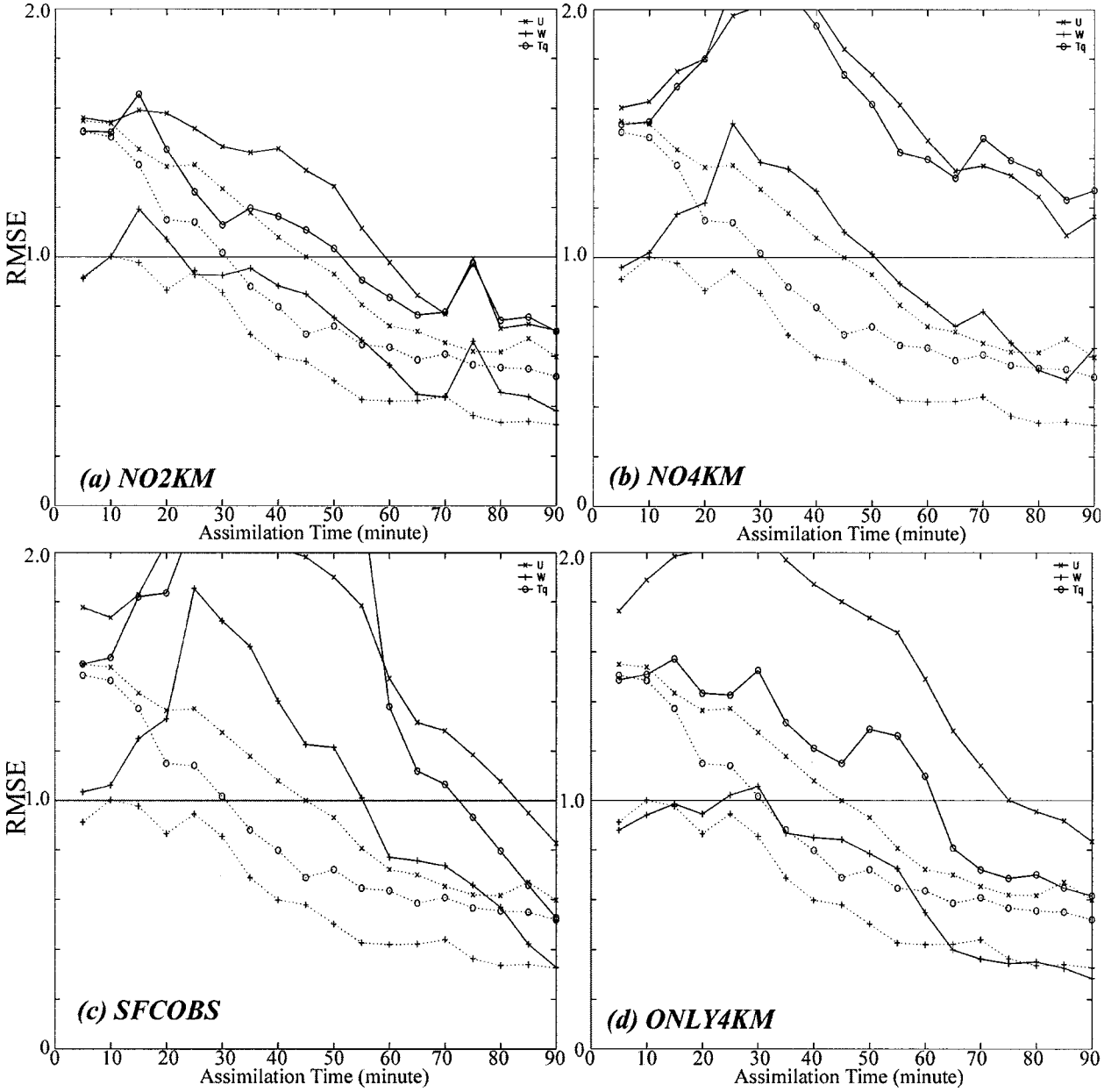

FIG. 12. As in Fig. 4, but for the RMSE in (a) NO2KM, (b) NO4KM, (c) SFCOBS, and (d) ONLY4KM. RMSE of CNTL is also shown as dotted lines in (a)–(d) for easy comparison.

1 extending northward and eastward with 10-km spacing, giving a total of 36 surface stations across the model domain. The standard deviation of the surface observational error is 1 m s^{-1} for horizontal winds and 1 K for liquid-water potential temperature. The radius of influence of each surface observation is set to 5 km.

Significant improvement of the EnKF performance can be found when the surface mesonet of wind and temperature observations are added (Figs. 10i–l, 11i–l). Even though the filter still has difficulty in capturing the weaker left-moving storm (which has relatively little signature above 4 km), it accurately reproduces the stronger, right-moving cell after 40–60 min of assimilation. The domain-averaged RMSE in SFCOBS falls steadily throughout the experiment and is close to that of CNTL after 90 min (Fig. 12c).

The improvement with the addition of the surface observations appears to come largely from the assimilation of surface thermal fields and thus better retrieval of the low-level cold pools (Figs. 10i–l). An additional experiment was performed exactly as SFCOBS except that only liquid-water potential temperature observations were taken from the 10 km by 10 km surface mesonet, and qualitatively similar results were achieved

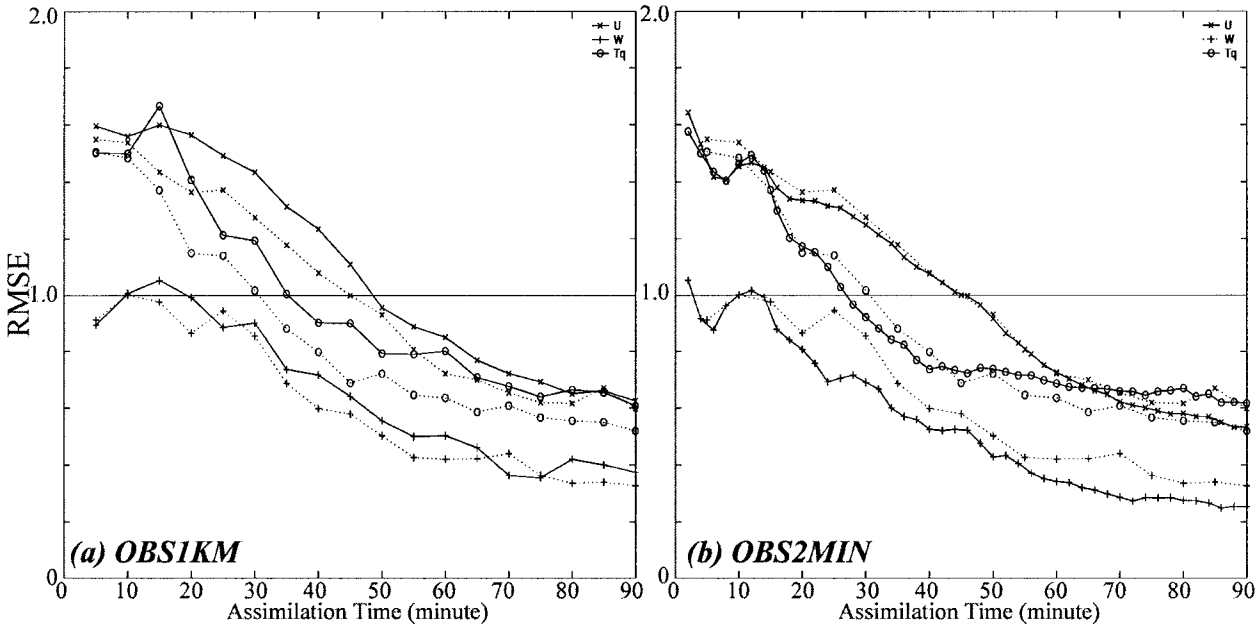

FIG. 13. As in Fig. 4, but for the RMSE in (a) OBS1KM and (b) OBS2MIN. RMSE of CNTL is also shown as dotted lines in (a) and (b) for easy comparison.

(not shown). Both experiments suggest the addition of a hypothetical surface mesonet, especially thermodynamic observations, can be beneficial to convective-scale data assimilation.

c. Mobile Doppler radar observations (ONLY4KM)

Mobile Doppler radar (or Doppler on Wheels; e.g., Wurman et al. 1996) is a promising tool to observe and study convective-scale phenomena. Typically, these mobile radars only scan the lower part of storms in order to obtain more accurate and frequent observations. Experiment ''ONLY4KM'' is designed to test the assimilation of radial velocity observations confined at and below 4 km.

With only the lower-level observations, the domain-averaged RMSE of all prognostic variables is significantly larger than in CNTL (Fig. 12d), even though the filter begins to develop storms in the ensemble mean after 20 min (Figs. 10m, 11m). Most of the differences exist in the upper troposphere where no observations are available and no updating of the state variables takes place (not shown). After 40 min, the filter begins to capture the splitting storms in the right locations with comparable strength to the reference simulation (Figs. 10n, 11n) though the zonal wind still contains significant errors (Fig. 12d). After 60 min, the filter has drawn the analyzed storm close to the reference solution (Figs. 10o–p, 11o–p) while continuously reducing the RMSE (Fig. 12d).

d. Sensitivity to vertical resolution (OBS1KM) and frequency (OBS2MIN) of observations

Especially far from the radars, radar observations have lower vertical resolution than the 0.5 km assumed in CNTL. Experiment ''OBS1KM'' differs from CNTL only in that the radial-velocity observations are taken at every other model level (that is, with 1-km vertical resolution). In this case, the ensemble filter turns out to have no difficulty in assimilating the storm at all: the rmse in OBS1KM is less than 10% higher than that of the CNTL throughout the assimilation (Fig. 13a).

In experiment ''OBS2MIN,'' the observations are taken every 2 min instead of every 5 min as in the control experiment. The reference simulation for OBS2MIN consists of a sequence of 2-min integrations as a necessary part of a perfect-model design (Mitchell et al. 2002). With the short assimilation cycle and more frequent observations, the filter performs marginally better for the first few observational cycles (Fig. 13b), leading to smaller RMSE initially, but beyond that the benefit of more frequent observations is minimal.

e. Sensitivity to observational error (OBSERR)

Experiment ''OBSERR'' differs from CNTL in that the observational error of the radial velocity has been doubled to 2.0 m s^{-1}. The RMSE curves, albeit slightly larger, show similar convergence toward levels similar to that of CNTL (Fig. 14). This experiment suggests that the accuracy of the analysis is not limited by observation errors but rather by errors arising from the

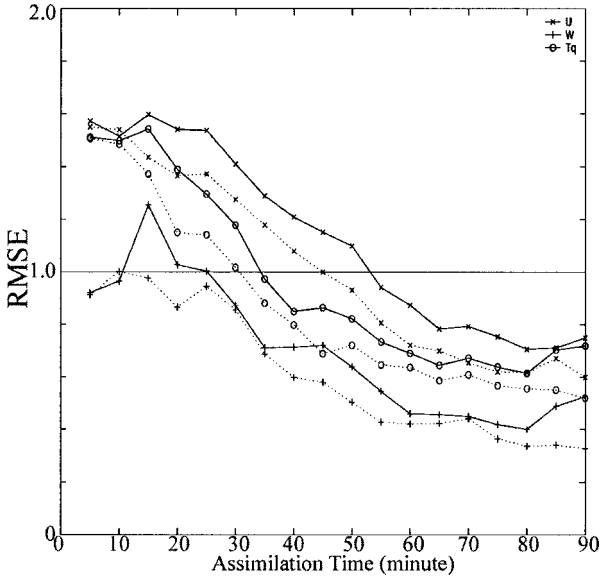

FIG. 14. As in Fig. 4 but for the RMSE in OBSERR. RMSE of CNTL is also shown as dotted lines for easy comparison.

use of a small ensemble size and the fact that only a portion of the state vector is observed.

6. Summary and discussion

This study explores the potential behavior of the ensemble Kalman filter at convective scales by considering realistic initial analyses and variations in the availability, resolution, and quality of the radar observations. We have demonstrated that the EnKF assimilations using radial-velocity observations are successful in most realistic observational scenarios for supercell thunderstorms. Even though the filter converges toward the truth simulation faster from a better initial estimate, an experiment with the initial estimate of the supercell displaced by 10 km still yields an accurate estimate of the storm for both observed and unobserved variables within an hour. Similarly, radial-velocity observations below 2 km are certainly beneficial to capturing the storm, especially the detailed cold pool structure, but in their absence the assimilation scheme can still achieve a comparably accurate estimate of the state of the storm given a slightly longer assimilation period. An experiment with radar observations only above 4 km fails to capture the storm properly but, with the addition of a hypothetical network taking surface wind and temperature observations, the EnKF can again provide an accurate estimate of the storm. The supercell can also be successfully assimilated in the case of radar observations only below 4 km, such as those from the ground-based mobile radars.

The results reported are likely optimistic in that they are based on assumptions of a perfect forecast model and known observation error statistics. In practice, error in forecast models, particularly from microphysical parameterizations, is likely considerable. Sources and statistics of errors in real observations are not well understood at present; these error sources include uncertainties in radar measurement and quality control as well as in forward operators for both radial velocity and reflectivity. In addition, correlated observational errors will make results here, with fully independent errors, overly optimistic.

On the other hand, a more sophisticated implementation of the filter than its current form and better ensemble initiation techniques might well further improve the EnKF performance at convective scales. We would also expect better filter performance if radar reflectivity observations are assimilated.

Acknowledgments. We acknowledge helpful discussions with Jeff Anderson, Howard Bluestein, Alain Caya, Andrew Crook, David Dowell, John Nielsen-Gammon, Bill Skamarock, Lou Wicker, Qin Xu, and Ming Xue. We also benefited greatly from two anonymous reviewers' comments on earlier versions of the manuscript. This research was supported by NSF Grants 0205599 and 0205655 and the U.S. Weather Research Program.

REFERENCES

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129,** 2884–2903.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126,** 1719–1724.

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.

Gao, J., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the analysis of three-dimensional wind fields from two Doppler radars. *Mon. Wea. Rev.,* **127,** 2128–2142.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

Kalman, R., 1960: A new approach to linear filtering and prediction theory. *Trans. ASME: J. Basic Eng.,* **83,** 95–108.

Laroche, S., and I. Zawadzki, 1994: A variational analysis method for the retrieval of three-dimensional wind field from single-Doppler radar data. *J. Atmos. Sci.,* **51,** 2664–2682.

Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130,** 2791–2808.

Qiu, C., and Q. Xu, 1992: A simple adjoint method of wind analysis for single-Doppler data. *J. Atmos. Oceanic Technol.*, **9,** 588–598.

Shapiro, A., S. Ellis, and J. Shaw, 1995: Single-Doppler velocity retrievals with Phoenix II data: Clear air and microburst wind retrievals in the planetary boundary layer. *J. Atmos. Sci.,* **52,** 1265–1287.

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.

——, and ——, 1998: Dynamical and microphysical retrieval from

Doppler radar observations using a cloud model and its adjoint. Part II: Retrieval experiment of an observed Florida convective storm. *J. Atmos. Sci.*, **55,** 835–852.

Tripoli, G. J., and W. R. Cotton, 1981: The use of ice-liquid water potential temperature as a thermodynamic variable in deep atmospheric models. *Mon. Wea. Rev.*, **109,** 1094–1102.

van Leeuwen, P. J., 1999: Comments on ''Data assimilation using an ensemble Kalman filter technique.'' *Mon. Wea. Rev.*, **127,** 1374–1377.

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.

Wurman, J., J. M. Straka, and E. N. Rasmussen, 1996: Fine-scale radar observations of tornadic storms. *Science*, **272,** 1774–1777.

Xu, Q., H. D. Gu, and S. Yang, 2001: Simple adjoint method for three-dimensional wind retrievals from single-Doppler data. *Quart. J. Roy. Meteor. Soc.*, **127,** 1053–1067.