Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS Experiments

Mingjing Tong and Ming Xue*

School of Meteorology and Center for Analysis and Prediction of Storms
University of Oklahoma, Norman, OK 73019

June, 2004
Submitted to Monthly Weather Review

Revised August, 2004
Minor revisions September 2004

*Corresponding author address:

Dr. Ming Xue
School of Meteorology, University of Oklahoma
SEC 1310, 100 E. Boyd, Norman OK 73019
mxue@ou.edu

Abstract

A Doppler radar data assimilation system is developed based on ensemble Kalman filter (EnKF) method and tested with simulated radar data from a supercell storm. As a first implementation, we assume the forward models are perfect and radar data are sampled at the analysis grid points. A general purpose nonhydrostatic compressible model is used with the inclusion of complex multi-class ice microphysics. New aspects compared to previous studies include the demonstration of the ability of EnKF method in retrieving multiple microphysical species associated with a multi-class ice microphysics scheme, and in accurately retrieving the wind and thermodynamic variables. Also new are the inclusion of reflectivity observations and the determination of the relative role of radial velocity and reflectivity data as well as their spatial coverage in recovering the full flow and cloud fields. In general, the system is able to reestablish the model storm extremely well after a number of assimilation cycles, and best results are obtained when both radial velocity and reflectivity data, including reflectivity information outside precipitation regions, are used. Significant positive impact of the reflectivity assimilation is found even though the observation operator involved is nonlinear. The results also show that a compressible model that contains acoustic modes hence the associated error growth performs at least as well as an anelastic model used in previous EnKF studies at the cloud scale.

Flow-dependent and dynamically consistent background error covariances estimated from the forecast ensemble play a critical role in successful assimilation and retrieval. When the assimilation cycles start from random initial perturbations, better results are obtained when one withholds during the first few cycles the updating of the fields that are not directly related to radar reflectivity. In fact, during the first few cycles, updating of variables indirectly related to reflectivity hurts the analysis. This is so because the estimated background covariances are unreliable at this stage of data assimilation, which is related to the way the forecast ensemble is initialized. Forecasts of supercell storms starting from the best assimilated initial conditions are shown to remain very good for at least two hours.

1. Introduction

Since its first introduction by Evensen (1994), the ensemble Kalman filter (EnKF) technique for data assimilation has received much attention. A rapidly increasing number of studies are appearing that examine its performance for various applications. In the field of meteorology, EnKF was first applied to large-scale data assimilation problems and observations are treated as random variables that are subject to perturbations (Evensen 1994; Burgers et al. 1998; Houtekamer and Mitchell 1998; 2001; Evensen 2003). Deterministic methods were developed more recently to avoid sampling errors associated with the use of perturbed observations or to address the adaptive observational network design problem; these methods include the ensemble square-root filter (EnSRF, Whitaker and Hamill 2002; Tippett et al. 2003), ensemble adjustment filter (Anderson 2001) and ensemble transform Kalman filter (Bishop et al. 2001), all of the three belong to the broader class of square-root filters (Tippett et al. 2003).

In general, EnKF and related methods are designed to simplify or make possible the computation of flow-dependent error statistics. Rather than solving the equation for the time evolution of the probability density function of model state, EnKF methods apply the Monte Carlo method to estimate the forecast error statistics. A large ensemble of model states are integrated forward in time using the dynamic equations, the moments of the probability density function are then calculated from this ensemble for different times (Evensen 2003).

In recent years, various techniques have been developed for analyzing and retrieving atmospheric state at the convective scale from Doppler radar data. These methods range from purely kinematic to sophisticated 4D variational method (4DVAR) that employs a nonhydrostatic prediction model and its adjoint (e.g., Gal-Chen 1978; Sun et al. 1991; Qiu and Xu 1992; Shapiro et al. 1995; Sun and Crook 1997; Gao et al. 1999; Wu et al. 2000; Weygandt et

al. 2002). Most of the work deals with retrieval and assimilation of radial velocity and/or reflectivity data from single Doppler radar because dual or multiple-Doppler coverage is not generally available. For the purpose of initializing NWP models, the 4DVAR method (e.g., Sun and Crook 1997; Gao et al. 1998) promises to provide an initial condition that is consistent with the prediction model and is able to effectively use multiple volume scans from radar. However, the high cost of developing and maintaining an adjoint code, especially one that can run efficiently on distributed-memory parallel computer systems and the need to include in the adjoint detailed physical processes which are more important at the convective scales, have limited 4DVAR assimilations of Doppler radar and other high-resolution data to relatively simple applications and model settings.

Very recently, EnKF was applied to the assimilation of simulated Doppler radar data for a modeled convective storm (Snyder and Zhang 2003; Zhang et al. 2004) and of real radar data by Dowell et al. (2004). Very encouraging results are obtained in these studies in retrieving wind, temperature and moisture field for convective storms. The first two studies assimilated only radial velocity data, while in Dowell et al. (2004), the use of reflectivity data is limited to the update of rainwater mixing ratio only. Neither of these studies included ice microphysics processes, and the assimilation system was based on an anelastic cloud model. Warm-rain microphysics scheme is also used in all afore-quoted 4DVAR studies except for Wu et al (2000) in which a simplified ice microphysics scheme was used. Dual-polarization radar data were assimilated by Wu et al (2000) by first deriving hydrometeor mixing ratios from regular and differential reflectivities.

Because of the need to run an ensemble of forecast and analysis of nontrivial sizes (usually a few tens to a few hundreds), the overall computational cost of ensemble-based

assimilation methods is also significant. Fortunately, a significant portion of the assimilation procedure, including the forecast components, is easily parallelizable. There is no need in general for the adjoint of the forward observation operators, therefore indirect observations with complex observation operators, at least those that involve primarily local influences, can be easily included. Furthermore, the analysis code is more or less independent of the prediction model, just the opposite of 4DVAR method. In addition, the system provides valuable uncertainty information on both analysis and forecast, and when combined with an existing ensemble prediction system, the incremental cost can be small.

The EnKF method is, however, not as mature as 4DVAR. No ensemble Kalman filter assimilation system has been implemented so far for operational use; EnKF also shares some of the common problems with 4DVAR, including issues with model errors, linear assumption associated with the optimality of solution, and Gaussian error assumption. For observations, such as the integrated total water along the slant paths of GPS data, the data influence region tends to nonlocal. In such a case, the covariance localization will have to be done around the entire slant path instead of around local observation points. Additional issues may arise in such cases, with cost being one. For systems that involve error growth at very different temporal and spatial scales, more difficulties may arise. For these reasons, much research is still needed before reliable operational implementations of ensemble-based assimilation methods can be achieved. More recently, work has been done in comparing the relative performance of 4DVAR and EnKF methods when assimilating Doppler radar data (Crook et al. 2002; Dowell et al. 2002; Caya et al. 2004).

In this study, we report on the development of an EnKF system based on a general-purpose compressible nonhydrostatic model, and on the application of the system to the

assimilation of simulated radial velocity and/or reflectivity data from a single Doppler radar. The forecast model employs a complex multi-class ice microphysics scheme. The performance of the EnKF scheme in 'recovering' the complete state of the model thunderstorms, including wind, temperature, pressure and all water and ice fields are examined. The relative impact of radial velocity and reflectivity data as well as their spatial coverage on the analysis are also investigated. The use of a compressible model, the inclusion of three-category ice microphysics in addition to the liquid water species, and the retrieval of multiple microphysics species with and without reflectivity data are aspects that are new compared to previous studies.

Even though we also performed experiments using ensemble square-root filter, we report here only results using the perturbed observation method. The rest of the paper is outlined as follows. In section 2, we describe our EnKF assimilation system and the design of OSS (Observing System Simulation) experiments. In section 3 we present the experiment results. The impact of various analyses on the forecast is discussed in section 4. A concluding section is given at the end.

2. Assimilation System and Experimental Design

a. The prediction model and truth simulation

In this study, we test our EnKF assimilation system using simulated data from a classic May 20, 1977 Del City, Oklahoma supercell storm case (Ray et al. 1981). Such simulation experiments are commonly referred to as Observing System Simulation Experiments (OSSE, see, e.g., Lord et al. 1997). The forecast model used is the Advanced Regional Prediction System (Xue et al. 2000; 2001; 2003). In this study, the ARPS is used in a 3D cloud model mode and the prognostic variables include three velocity components u, v, w, potential temperature θ, pressure

p, and six categories of water substances, *i.e.*, water vapor specific humidity q_v, and mixing ratios for cloud water q_c, rainwater q_r, cloud ice q_i, snow q_s and hail q_h. In addition, turbulence kinetic energy is also predicted which is used to determine turbulent mixing coefficients based on a 1.5-order turbulence closure scheme. The microphysical processes are parameterized using the three-category ice scheme of Lin *et al.* (1983) and its implementation follows Tao and Simpson (1993). More details on the model can be found in (Xue et al. 2000; 2001).

For all experiments, the physical domain is $64\times64\times16\,\mathrm{km}^3$. The model grid comprises of $35\times35\times35$ grid points (including points that facillitate the specification of boundary conditions), with grid intervals of 2 km in the horizontal directions and 0.5 km in the vertical. The truth simulation or nature run is initialized from a modified real sounding as used in Xue et al (2001). The CAPE of the sounding is about 3300 J kg^{-1}. A 4 K ellipsoidal thermal bubble centered at $x = 48$, $y = 16$ and $z = 1.5$ km, with radii of 10 km in *x* and *y* and 1.5 km in *z* direction is used to initiate the storm. Open conditions are used at the lateral boundaries. A wave radiation condition is also applied at the top boundary. Free-slip conditions are applied to the bottom boundary. The length of simulation is up to three hours. A constant wind of $u = 3$ m s^{-1} and $v = 14\,\mathrm{m\ s}^{-1}$ is subtracted from the observed sounding to keep the primary storm cell near the center of model grid. Despite the differences in resolutions, the evolution of the simulated storms is very similar to those documented in Xue *et al.* (2001).

During the truth simulation, the initial convective cell strengthens over the first 20 minutes. The strength of the cell then decreases over the next 30 minutes or so, which is associated with the splitting of the cell into two at around 55 minutes (Fig. 1). The right moving (relative to the storm motion vector which is towards north-northeast) cell tends to dominate the system; the updraft reaches a peak value of 44 m s^{-1} at 90 minutes. The left moving cell starts to

split again at 95 minutes. The initial cloud started to form at about 10 minutes, and rainwater formed at about 15 minutes. Ice phase fields appeared at about 20 minutes.

b. Simulation of radar observations

As a first implementation, we assume that the simulated observations are available on the scalar grid points. Future work will assume the availability of data in radar coordinate. The simulated radial velocity, V_r, is calculated from

$$V_r = u\cos\alpha\sin\beta + v\cos\alpha\cos\beta + w\sin\alpha + \text{a random error}, \quad (1)$$

where α is the elevation angle and β the azimuth angle of radar beams, and *u*, *v* and *w* are the model-simulated velocities interpolated to the scalar points of staggered model grid. The random error is drawn from a normal distribution with zero mean and standard deviation of 1 m s^{-1}. Since V_r is sampled directly from velocity fields, the effect of hydrometeor sedimentation does not come into play.

The simulated logarithmic reflectivity factor (referred to simply as reflectivity in most parts of this paper), in dBZ, is estimated from equations as follows:

$$Z = 10\log_{10}\left(\frac{Z_e}{1mm^6m^{-3}}\right) + \text{a random error}. \quad (2)$$

The equivalent reflectivity factor, Z_e, is made up of three components,

$$Z_e = Z_{er} + Z_{es} + Z_{eh}, \quad (3)$$

where Z_{er}, Z_{es}, Z_{eh} are contributions from rain water, snow and hail. The random error is drawn from a normal distribution with zero mean and standard deviation of 5 dBZ. Reflectivity relations corresponding to the 10 cm wavelength of WSR-88D radars are given below and used in our experiments.

The rain component of the reflectivity is calculated, based on Smith *et al* (1975), from

$$Z_{er} = \frac{10^{18} \times 720(\rho q_r)^{1.75}}{\pi^{1.75} N_r^{0.75} \rho_r^{1.75}}, \tag{4}$$

where $\rho_r = 1000$ kg m^{-3} is the density of rainwater, ρ in kg m^{-3} the density of air. $N_r = 8.0 \times 10^6$ m^{-4} is the intercept parameter in the assumed Marshall-Palmer exponential rain drop size distribution.

If the temperature is less than 0 C°, then the component of reflectivity is, for dry snow,

$$Z_{es} = \frac{10^{18} \times 720 K_i^2 \rho_s^{0.25} (\rho q_s)^{1.75}}{\pi^{1.75} K_r^2 N_s^{0.75} \rho_i^{1.75}}. \tag{5}$$

Here $\rho_s = 100$ kg m^{-3} is the density of snow and $\rho_i = 917$ kg m^{-3} the density of ice. $N_s = 3.0 \times 10^6$ m^{-4} is the intercept parameter for snow. K_i^2 = 0.176 is the dielectric factor for ice and K_r^2 =0.93 the same for water. Wet snow, which occurs at temperature higher than 0 C°, is treated in a similar way as rain water, and the equivalent reflectivity factor-mixing ratio relation is

$$Z_{es} = \frac{10^{18} \times 720 (\rho q_s)^{1.75}}{\pi^{1.75} N_s^{0.75} \rho_s^{1.75}}. \tag{6}$$

For hail, the wet hail formulation of Smith *et al* (1975) is used, *i.e.*,

$$Z_{eh} = \left(\frac{10^{18} \times 720}{\pi^{1.75} N_h^{0.75} \rho_h^{1.75}} \right)^{0.95} (\rho q_h)^{1.6625}, \tag{7}$$

where $\rho_h = 913$ kg m^{-3} is the density of hail. Hail intercept parameter of $N_h = 4.0 \times 10^4$ m^{-4} is used, which is slightly larger than that used in Smith *et al* (1975), implying more hails of smaller sizes. The same value is used in the ARPS implementation of Lin *et al* (1983) microphysics scheme. SI units are used in all equations above. Relations similar to the above are used in, e.g., Ferrier (1994). The above equations define the observation operator denoted as *H* in the

following section. When creating simulated observations for our OSS experiments, we impose a lower limit of 1 mm^6m^{-3} on the equivalent reflectivity factor to yield a lower limit of 0 dBZ for the logarithmic reflectivity factor.

The ground-based radar is located at the southwest corner of the computational domain, *i.e.,* at the origin of *x*-*y* coordinate (c.f., Fig. 1). For data sampling and assimilation, we assume that the observation operators, given by the above equations, are perfect. As with most atmospheric data assimilation systems, the prediction model is also assumed to be perfect, i.e., no model error is explicitly taken into account. It is worth noting here that observation operators for reflectivity factors are nonlinear and because of that the error distribution of reflectivity data is likely non-Gaussian. Discussions on the non-Gaussian nature of error distributions associated with nonlinear observation operators can be found in, e.g., Lorenc (2003). One of the goals of this study is to assess the effectiveness of such data given these characteristics that violate the basic assumptions used in deriving the optimal EnKF solution.

c. The EnKF data assimilation procedure

Our EnKF implementation is primarily based on the algorithm described by Evensen (1994), Burgers *et al.* (1998) and Houtekamer and Mitchell (1998) which uses the perturbed-observation method. In this algorithm, the ensemble mean is supposed to be the best estimate of the true state and the spread of the ensemble around the mean is a good estimate of the error in the ensemble mean (Evensen 2003). The forecast and analysis error covariances, given by matrix **P**, are defined as

$$\mathbf{P}^f \cong \mathbf{P}_e^f = \overline{(x^f - \overline{x^f})(x^f - \overline{x^f})^T}, \tag{8}$$

$$\mathbf{P}^a \cong \mathbf{P}_e^a = \overline{(x^a - \overline{x^a})(x^a - \overline{x^a})^T}, \tag{9}$$

where the overbar denotes the ensemble mean and superscripts *a* and *f* denote analysis and forecast, respectively. Each ensemble member is updated by analysis equation

$$x_i^a = x_i^f + \mathbf{P}^f \mathbf{H}^T [\mathbf{H}\mathbf{P}^f \mathbf{H}^{\mathrm{T}} + \mathbf{R}]^{-1} (y_i^o - H x_i^f) , \tag{10}$$

where i represents the i^{th} ensemble member and x_i^f is the first guess obtained from the i^{th} ensemble forecast. H is the observation operator, which converts the model states to the observed parameters. **H** in its standard notion is the linearized version of H in a matrix form but will not be explicitly evaluated anywhere in the analysis procedure. The linearization is required to arrive the optimal analysis-covariance minimizing solution (see, e.g., Kalnay 2002). Following Evensen (2003) and Houtekamer and Mitchell (1998), observations are treated as random variables having a distribution with a mean equal to the first guess observations y^o and error covariances given by matrix **R**. Thus an ensemble of observations are obtained by

$$y_i^o = y^o + \varepsilon_i , \tag{11}$$

and the observation error covariance matrix is calculated from the random errors added to the observations, as given in Eq.(11):

$$\mathbf{R} \approx \mathbf{R}_e = \overline{\varepsilon \varepsilon^T} , \tag{12}$$

where ε is observation error. The above calculation of **R** is only possible with OSSE experiments where observations contaminated by errors with known variance and distribution are constructed. For real observations, certain estimate of **R** has to be made. Also, our assumption about the non-correlation among the observational errors determines that **R** is diagonal. Finally, in our current application, y^o includes radial velocity and reflectivity.

The forecast error covariances are calculated from

$$\mathbf{P}^f \mathbf{H}^{\mathrm{T}} \cong \frac{1}{N-1} \sum_i^N (x_i^f - \overline{x^f})(H(x_i^f) - \overline{H(x^f)})^{\mathrm{T}} , \tag{13}$$

and

$$\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} \cong \frac{1}{N-1}\sum_i^N (H(x_i^f) - \overline{H(x^f)})(H(x_i^f) - \overline{H(x^f)})^{\mathrm{T}}. \quad (14)$$

where N is the number of ensemble members. The use of the nonlinear observation operator in the evaluation of covariances in Eqs. (13) and (14) is an advantage of EnKF compared to traditional extended Kalman filter for nonlinear problems (Evensen 2003).

As can be seen, the nonlinear observation operator H is directly used in Eqs. (13-14), **H** and its transpose are not evaluated directly (the transpose of **H** or the adjoint of the observation operator is needed by variational assimilation methods). The underlying assumption of standard extended Kalman filter theory (which is the basis of EnKF) is that the forecast model and the observation operator are linear. The nonlinearity in the forecast model and the observation operator (such as that of reflectivity data) can make the algorithm suboptimal. Strong nonlinearity would also render initially Gaussian error distributions non-Gaussian. In this study, the performance of the EnKF in such a situation will be examined.

We start the initial ensemble forecast at the 20 minutes of model time when the storm cell developing out of an initial bubble reaches peak intensity. To initialize the ensemble members, random noises are added to the initially horizontally homogeneous background that is based on the environmental sounding. The random noises are sampled from Gaussian distributions with zero mean and standard deviation of 3 m s^{-1} for u, v, and w, and 3 K for potential temperature. The pressure, moisture, and microphysical variables are not perturbed. The technique of adding initial perturbation has been discussed by former studies of Snyder and Zhang (2003) and Dowell et al (2004). They found that adding initial perturbations to the entire computational domain would introduce spurious cells in individual members. Limiting the perturbation region around the observed storm area helps improve the assimilation result. However, for general applications

of EnKF, the ensemble forecast should cover all scales of motion. We chose to apply the initial perturbations to the entire domain except for the grid points at the lateral boundary to keep the system more general. The open lateral boundary condition is sensitive to boundary errors and initial perturbations introduced at the boundary were found to trigger spurious storm development within some ensemble members. Assimilating reflectivity in clear area region is found to help suppress spurious cells in the interior domain, hence allowing a more general application of initial perturbations.

The observations are assimilated every 5 minutes. The first analysis is performed at 25 minutes and one hundred ensemble members are used. As mentioned earlier, observations are perturbed by adding Gaussian noises, with a standard deviation of 1 m s^{-1} for radial velocity and 5 dBZ for reflectivity. The observation errors are assumed to be uncorrelated; therefore, observations can be and are analyzed sequentially one at a time, following Houtekamer and Mitchell (2001). When using a small (relatively to the degrees of freedom of the analysis system) ensemble to estimate the background error covariances, the estimated values between distant grid points are not reliable. Therefore covariance localization is necessary (Houtekamer and Mitchell 1998; Anderson 2001; Hamill et al. 2001; Houtekamer and Mitchell 2001).

In our early study (Tong and Xue 2004), we limited the influence region of each observation to a rectangular region with half width of two grid intervals in both horizontal and vertical directions, equivalent to 4 and 1 km in the horizontal and vertical, respectively. Some useful information beyond that region of influence was not utilized (though the computation was faster). To spatially smooth the analysis increments as well as to localize covariances, Houtekamer and Mitchell (2001) proposed a method that applies Schur (elementwise) product of the background error covariance calculated from the ensemble and a correlation function with

local support. This method was also used by Dowell et al. (2004). In the same way, we multiply each element of the background error matrix $\mathbf{PH}^T$ with a weight computed from correlation function given by Eq. (4.10) of Gaspari and Cohn (1999). The weight decreases gradually from 1 at the observation point to zero at an effective cutoff radius and remains zero beyond. Through assimilation experiments with cutoff radii ranging from 4 to 10 km, we found that 8 km worked the best with 100 ensemble members. This value is therefore used in all experiments presented in this paper. The use of a smooth filter function is found to produce significantly better analysis than the sharp cutoff function used earlier in Tong and Xue (2004).

The EnKF algorithm tends to underestimate the analysis uncertainty owning to the use of limited ensemble size. To solve this problem, we tested the double ensemble method (Houtekamer and Mitchell 1998), in which the statistics estimated from one ensemble is used to update the other. However, this approach overestimated the analysis uncertainty in our case and did not help improve our analysis. We also tested the covariance inflation approach of Anderson (2001), which did not work well at first. As found in Snyder and Zhang (2003), it enhances the spurious cell in ensemble members. A modification is then made by applying covariance inflation to the grid points that will be influenced directly during the analysis update by the observations found within the precipitation (where observed Z > 10 dBZ) regions. To do this, we check for each grid point to see if within 8 km radius (the same radius used by the covariance localization) there is at least one reflectivity data greater than 10 dBZ. If true, then covariance inflation is applied to this grid point. This is found to improve our analysis. The inflation factor we used is 1.07 or 7 percent when radial velocity or/and reflectivity data in the precipitation region are assimilated. When reflectivity data of complete coverage are used, the inflation factor is increased to 1.1 (10 percent). Zhang et al (2004) proposed another alternative to the traditional

covariance inflation that is to avoid enhancing spurious cells but Caya et al (2004) found that the use of the method degrades the quality of analysis.

3. The Assimilation Experiments

Table 1 lists eight experiments to be discussed in this paper. The assimilation scheme is first tested by assimilating radial velocity or reflectivity data alone or by assimilating both. Further, data coverage tests are performed in which each data type is available either in the entire domain or only in regions where reflectivity exceeds 10 dBZ (referred to as precipitation region). Additional variations have to do with the ways analysis variables are updated during the assimilation cycles. The results are discussed in the following sections.

a. Assimilations using radial velocity data only

In experiment VrP (see Table 1), we assume that the radial velocity data are available in precipitation regions where reflectivity is greater than 10 dBZ. After only one analysis cycle (at t = 25 min), the basic patterns of middle-level horizontal winds around the storm appear reasonable (not shown). After three more analysis cycles, *i.e.*, by 40 min., the basic structures of updraft and horizontal flow (Fig. 1*e*) as well as perturbation temperature (not shown) above the low-level cold pool are reasonably well retrieved. The low-level cold pool and the associated divergence are too weak at this time (Fig. 2*e*). Some hydrometeors can be retrieved at this time, but their amount and locations are not yet accurate. Two more assimilation cycles later, at 50 min., the strength of the updraft and the magnitude of the temperature perturbation (except that in the low-level cold pool) become reasonably good.

At 60 min., the retrieved microphysical fields also become rather close to the truth as seen from the vertical cross-sections (Fig. 3 (a2)-(f2)), so do most other fields (Fig. 1*f* and Fig.

2*f*). The largest differences at this time are found at the low levels (Fig. 2*f*) where precipitation and cooling in the otherwise unperturbed low-level inflow region exist, and near the western boundary. Evaporative cooling due to spurious precipitation in some ensemble members was the cause. Much of the error is corrected by subsequent analyses and by 100 min. (Fig. 2*h*), the precipitation in the inflow region is generally gone except for a small area ahead of the rear-flank gust front. The low-level flow and reflectivity patterns as well as the shape of the cold pool now agree quite well with the truth (Fig. 2*h* and *d*); although small differences are still found with the exact location of the boundary of rain-cooled regions (as indicated by the zero degree θ' contours) and in the areal coverage of leading precipitation region on the east-northeast side. The agreement at the 6 km level is even better (Fig. 1*h* and *d*). These results indicate that even with radial velocity data in precipitation region only, the EnKF system is able to rebuild the model storm remarkably well after a sufficient number of assimilation cycles. Analysis of such quality can only be expected of methods that make use of multiple radar volume scans effectively and in a way that is compatible with the hopefully correct model physics.

We use the root-mean-square (*rms*) error of ensemble mean analysis to judge quantitatively the quality of the analysis. The *rms* errors are averaged over those grid points where the reflectivity is greater than 10 dBZ. The *rms* errors of velocities, temperature, cloud and hydrometeor variables in experiment VrP are seen to decrease rapidly during the first four assimilation cycles (over 20 minutes) and the analysis tends to converge at about 70 minutes (Fig. 4, gray curves). When the analysis converges, the *rms* analysis error for horizontal wind components is generally less than 1 m s^{-1} and about 0.5 m s^{-1} for vertical velocity. For perturbation potential temperature θ', the analysis error decreases to less than 0.5 K at 60 minutes. Such velocity errors are similar to or less than the observational errors added to the radial

velocity. It means that after eight to ten assimilation cycles, the EnKF system is producing a very good estimate of the state of simulated storm.

In experiment VrF, in which V_r data cover the entire domain, the analysis converges more quickly and the analysis errors are smaller (black dotted lines in Fig. 4). The retrieved winds, potential temperature, pressure and water vapor content are significantly better than those of VrP case. The retrieved microphysical fields are also improved, but not as much as when full coverage reflectivity data are used (results to be presented later). For radial velocity data to be available outside the precipitation region, the radar has to be operating in high sensitivity mode, which is generally not the case with the WSR-88D network when precipitation is present. However, when such data are available, our experiment shows clear positive impact of the data.

In our earlier study (Tong and Xue 2004), we found that when pressure field is analyzed (updated by the analysis procedure), the pressure error increases instead of decreases after each analysis cycle. A careful examination of the results reveals that the main problem was associated with a domain-wide pressure drift in the forecast that often occurs in simulations using a small computational domain and an open boundary condition. The background error covariance between pressure and observations in the presence of pressure drift was apparently not correct, causing negative analysis impact on pressure field. In all experiments presented here, the pressure detrending option of ARPS model is turned on, which forces the domain-mean perturbation Exner function to zero after each forecast time step. The model solution, at least in the dry case, should be independent of the perturbation Exner function to an arbitrary constant.

After applying pressure detrending, pressure drift is much controlled and the update of pressure by analysis does reduce its error in general, though not as much as for other fields (compare, e.g., gray curves in Fig. 4*e* and Fig. 4*c*). At 40 min. in the current experiment (VrP),

the basic pattern of perturbation pressure can be reasonably reconstructed (not shown). Noticeable noises associated with acoustic oscillations remain at low levels at this time which are reduced by additional assimilation cycles. However, it can be seen from Fig. 4*e* (gray curve), the forecast error of the pressure perturbation starts to increase again at 65 min., and the error is not effectively reduced by further analysis. It is found that in this case some positive pressure drift still occurs in the ensemble-mean forecast and analysis below 4 km. The positive pressure perturbation associated with the cold pool is about 20 Pa higher than in the truth. The low pressure below the updraft is not as low as it should be. Such pressure biases appear to be preventing the EnKF scheme to work effectively in reducing the pressure errors, which is only indirectly related to V_r. It will be shown later that when reflectivity data are assimilated, the pressure drift problem is alleviated and the analysis of pressure is improved.

b. Impact of assimilating reflectivity data

Reflectivity is a measurement that is provided by all types of weather radar. In this section, we examine if reflectivity data alone is sufficient for the model to reproduce the true storm; we also study its value when used in combination with radial velocity data. We note that the observation operator for reflectivity is nonlinear and there exist more uncertainties with reflectivity operator (because of uncertainties with cloud microphysics, effects of attenuation, among others) than with radial velocity. In our model, the rainwater, snow and hail mixing ratios, q_r, q_s, and q_h, are directly related to observed reflectivity through Eqs. (3)-(7). Therefore, the experiments further test the performance of the EnKF scheme for nonlinear observations.

In experiment ZP (see Table 1), reflectivity (Z) data greater than 10 dBZ are assimilated but not radial velocity. Generally, the analysis obtained by assimilating reflectivity data only is not as good as the analysis assimilating redial velocity data alone. As can be seen in Fig. 4 (black

curves), the analysis acts to reduce the *rms* errors in q_r, q_h, q_s, q_c, q_i, w and θ' from the third cycle with the reduction of error in q_h being much larger. Significant reduction in errors of horizontal wind components and in q_v did not start until after four to five cycles. In fact, during the first two cycles, the update of most variables increases rather than decreases the error (see, *e.g.*, the black curves in Fig. 4*d* for θ' and Fig. 4*h* for q_v). Significant reduction in errors of horizontal wind components and in q_v did not start until after four to five cycles.

The delay in the reduction of or even increase in the *rms* errors in variables not directly related to Z is related to our initial perturbation method. When the ensemble members are initialized, cloud and hydrometeor fields are not perturbed. This is reasonable because the location of precipitation region is unknown before observational data are introduced. Adding hydrometeor perturbations everywhere in the model is undesirable because non-zero values in certain variables like cloud ice should only exist under certain conditions. If the initial background contains some information about the cloud, then random perturbations can be added to the related fields in a similar way we do the other fields. Without direct initial perturbations to them, it takes the model a couple of assimilation cycles to develop cloud and hydrometeors and to establish coherent covariance structures for a reliable estimation of error covariances. For these reasons, updating indirectly related variables in the first few cycles increases their errors. Enough ensemble spread is established starting from the third assimilation cycle; from then on, the analysis is able to correct forecast errors in fields including w, θ', q_c, q_r, q_i, q_s and q_h (Fig. 4). Reliable covariances between reflectivity and p, q_v and horizontal wind components are slower to establish and the analysis correction starts a couple of cycles later.

Because of the lack of reliable covariances, the retrieved model fields in ZP are not good at 40 min. (Fig. 1*i*). By 45 min., the storm updraft and temperature perturbation pattern become

reasonable. The locations of the hydrometeors are now close to the truth. However, horizontal wind patterns are still not very good at this time. From 60 minutes on, the basic structure and the evolution of the storm including the split storm cells become rather accurate. The errors in all fields are further reduced in the subsequent assimilation cycles (Fig. 4), but remain higher than the VrP case by the end of the assimilation period (100 min.).

Retrieving the model state from the observed variables in the EnKF system relies heavily on a good estimate of the flow-dependent multivariate background error covariances. Fig. 5 shows, for experiment ZP, the forecast background error correlations between reflectivity at point x = 32 km, z = 6 km and model variables at each grid point in an x-z vertical cross section through the maximum updraft (y = 29 km) at 80 min. (the correlation patterns for other experiments at this time are similar). As can be seen, for the water and ice fields, significant correlations are mostly confined to the regions where their non-zero values are found. For w and θ', the correlation patterns also match those of w and θ' themselves in general, and significant correlations extend through much of the troposphere depth, which is consistent with the nature of deep convection. The maximum correlations for w and θ' exceed 0.8, indicating that they are highly correlated with reflectivity. Generally, positive correlations are associated with updraft and high buoyancy; negative correlations are found in the recirculation regions. The correlation for θ' spreads above the tropopause in a wavy pattern to both upstream and downstream, which should be related to the gravity wave propagation in the stable stratification there.

Even for pressure, coherent structures of significant correlation exist. Maximum negative correlation is found at the mid levels where perturbation pressure has a minimum; maximum positive correlation exists near the cloud top. This correlation pattern can be explained by the rotating supercell storm dynamics. Stronger updraft rotation produces lower mid-level pressure

which promotes stronger updraft therefore produces more reflectivity. Stronger updraft, accompanied by larger reflectivity, will produce larger positive pressure perturbations at the cloud top due to both hydrostatic and Bernoulli effects. These suggest that the error correlations estimated from the forecast ensemble are dynamically consistent. It is these valuable correlations that act to spread the information from the observation points to the finite domain surrounding them and that retrieve the unobserved variables from the observed ones. The results also justify our choice of the influence region in covariance localization. Most significant correlations within 4 km from the center of the influence region are kept and smaller weights are given to the correlations extending up to 8 km.

c. Assimilation of both radial velocity and reflectivity

In the next set of experiments (VrZPa-c, VrPZF, see Table 1), we combine the radial velocity and reflectivity data into the assimilation process. Experiment VrZPa is a combination of VrP and ZP, in which both the radial velocity and reflectivity in precipitation regions (Z > 10 dBZ) are assimilated, using the same procedure as in VrP and ZP. Compared to VrP (Fig. 6), when additional reflectivity data are introduced, the *analyses* of q_r, q_s and q_h are generally improved, especially for q_r and q_h, as indicated by lower errors after each analysis. The forecast error growth in this experiment is faster, however, even for those three variables directly related to Z.

The analysis errors of q_v and p' remain higher than those in VrP until 60 and 80 min., respectively, but become noticeably smaller afterwards. For most part of the assimilation period, the wind components are not as accurate as those retrieved by assimilating only V_r, but can reach the same accuracy in the last a few cycles. For θ', q_c and q_i, the *analysis* errors from VrZPa are larger in the first 2-3 cycles, smaller in the next few then remain similar to those in VrP case.

The lack of significant positive impact by the inclusion of reflectivity data, especially during the first few assimilation cycles, is due to the reasons discussed earlier for ZP. In the first two to four cycles, background error covariances between reflectivity and the model variables that are not directly related to Z are not reliable. Updating these variables based on reflectivity data and the unreliable covariances hurts the analysis, as indicated by the increase of errors in w, θ, q_c and q_v by, e.g., the second analysis cycle (the error increase is actually rather bad for q_c and q_v, see Fig. 6). The increased analysis errors lead to less accurate forecast and faster error growth. Despite all this, the analyses towards the end of the assimilation period is similar or better for most fields in VrZPa than in VrP, indicating significant positive impact of Z when good error estimation becomes available. Interestingly, the degradation of the analysis in the early cycles by the assimilation of rainwater content is also noted by Caya et al (2004), which includes warm rain microphysics.

We note that in Dowell *et al.* (2004) which employs warm rain microphysics, it is found that updating q_r only when assimilating reflectivity observations help maintain a realistic structure of the precipitation core and improve velocity verification score. We performed a corresponding experiment, VrZPb (black curves in Fig. 7), in which only q_r, q_s, and q_h are updated when assimilating reflectivity. As was found in VrZPa, the *analyses* of q_r, q_s and q_h are improved over those of VrP. The analysis of q_v is also better. For most of the other variables, the improvement over VrP is generally small, if any (Fig. 7). There are also a few cycles in which the retrieved w and q_i are not as good as those of VrP. Since it was shown earlier that reflectivity does have significant correlations with indirectly related variables (Fig. 5), we do not believe it appropriate to completely exclude those variables from the analysis update when assimilating reflectivity data. A better solution needs to be found.

From previous experiments, it can be seen that the negative impact of using reflectivity to update variables rather than q_r, q_s, and q_h is mainly caused by the incorrect background covariances in the first two to three cycles. After that, the reflectivity becomes beneficial in retrieving those model fields. To see if we can further improve the analysis, in experiment VrZPc (Table 1), we apply the update due to reflectivity only to q_r, q_s, and q_h before the fourth cycle. Starting from the fourth cycle, all models variables are updated. It is shown in Fig. 8 that doing so improves the analysis overall. The retrieved q_v, potential temperature and pressure are improved over VrP more than those in experiment VrZPb. Significant improvement could also be found in q_c and q_s. For velocity components, the improvements are also slightly larger, especially in the later cycles.

The retrieved updraft and horizontal wind field in the middle level are shown in Fig. 1 in panels *m* though *p*. The detailed structures of the updraft of the split cells starting from 60 min (panels *n* through *p*) are retrieved better than those in either VrP (plots *f* through *h*) or ZP (plots *j* through *l*). At 80 and 100 min., the pattern and strength of the left mover, a less organized therefore more difficult one analyze, are clearly better retrieved by experiment VrZPc. Fig. 2 shows the retrieved low-level cold pool, gust front and precipitation pattern by different experiments, in terms of the low-level θ', wind vectors and model computed reflectivity. The benefit of using reflectivity data can be better seen from this figure. Without reflectivity data (VrP), the retrieved precipitation is much less, especially during the early cycles. Spurious echoes are found in the clear air region in both cases but much less so in VrZPc (Fig. 2). They are related to spurious cells in individual ensemble members. The assimilation of reflectivity data reduced the spurious echoes.

In the previous experiments, only reflectivity larger than 10 dBZ is assimilated. In reality,

reflectivity outside the precipitation regions also contains valid information. At least it indicates the absence of precipitating particles there. One can therefore assume that the reflectivity data cover the entire region within the radar range, which in our current case means the entire computational domain. In our next OSS experiment, named VrPZF (Table 1), reflectivity information is available everywhere while radial velocity data remain available in precipitation (Z > 10 dBZ) regions only. As in VrZPc, u, v, w, θ', q_v, q_c, and q_i are updated starting from the fourth cycle when assimilating reflectivity while q_r, q_s and q_h are updated from beginning.

The complete coverage of reflectivity data can remove spurious cells that would otherwise develop in the data void regions. The *rms* analysis errors in VrPZF (Fig. 9, thin black curves) are decreased more over those in VrP or VrZPc (Fig. 8). The retrieved horizontal wind fields, the strength of the updraft (Fig. 1), the low-level perturbation potential temperature (Fig. 2) and the microphysical fields (Fig. 3) are more accurate than those retrieved by any of the previous experiments. The reflectivity computed from the analysis (Fig. 1*m*-*p*) very closely resembles that in the truth simulation with no spurious echoes found in the clear air region. Towards the later period of assimilation, the strength of the low-level cold pool can be retrieved very well by all of the three experiments (Fig. 2). However, only when complete reflectivity data coverage is used, as in the case of VrPZF, can the extent of the low-level rain-cooled air, as indicated by the 0 K contour of perturbation potential temperature, be accurately determined. This is actually very important for the subsequent forecast as spurious cooling in the inflow region is found in some cases to reduce the CAPE in the inflow and adversely affects the evolution of ensuing forecast. VrPZF produces the best forecast (more later) among all assimilation experiments, for this and other reasons.

Finally, we briefly discuss the ensemble spread which is defined as the square root of the

ensemble variance and plotted in Fig. 9 for experiment VrPZF. Since EnKF estimates the forecast error using the ensemble spread, in ideal situation, the ratio of the spread of the ensemble to the error of the ensemble mean forecast is equal to $\sqrt{N_e/(N_e+1)}$ (Murphy 1988). As we can see from Fig. 9, the ensemble spread gradually approaches the ensemble mean forecast error for most variables, a favorable sign. During the early cycles, the ensemble spread is much smaller than the rms error and therefore provides a poor representation of forecast error, a behavior also observed by Snyder and Zhang (2003). The ratio increases rapidly during the first 6 to 8 cycles for most variables (mainly because of the rapid decrease in forecast errors) then becomes more stable. In the last four or so cycles, this ratio starts to decrease again, especially for *u, v, w,* θ' and p'. Such a decrease is typically observed as EnKF goes through successive assimilation cycles (e.g., Houtekamer and Mitchell 1998) owing to a systematic underestimation of the analysis variance. In our case, this decrease in the later cycles is mainly due to the faster error growth in the forecast rather than to change in the ensemble spread; in fact, the ensemble spread remains essentially constant in the later cycles. The faster forecast error growth is, we believe, due to the more transient nature of the flow after the storm went through more than two splits. Snyder and Zhang (2003) noted, however, a continued increase in the spread-to-error ratio, and attibuted the increase to the presence of spurious cells in individual memebers. A similar behavior is also observed in our experiment, VrP, in which spurious cells are not effectively supressed as they are in VrPZF. We further note that in the last few cycles of VrPZF, despite the increased forecast errors (Fig. 9), the analysis errors do not increase or increase very little, indicating that the filter is well behaved.

d. Retrieval of microphysical fields

The microphysics retrieval is an important aspect of convective-scale data assimilation.

Relatively few previous data assimilation studies have focused on this problem. Most of these studies used only a simple microphysical parameterization and the ice phase is usually excluded. The more recent attempt of Wu et al (2000) uses a 4DVAR system to assimilate dual-polarization radar data into a cloud model. Generally, a good bulk ice microphysics parameterization includes ice categories for cloud ice (individual crystals), snow (aggregates), and graupel and/or hail. As pointed out by Wu et al (2000), such a model will have a complex adjoint model involving many nonlinearities, and the 4DVAR system based on such adjoint tends to have poor convergence properties. For these reasons, Wu et al (2000) made simplifications by using a scheme without the snow category and by merging cloud water and cloud ice categories. Such simplifications reduce the degrees of freedom (or the number of control or analysis variables), and also reduce the number of, usually highly nonlinear, microphysical processes. In addition, in Wu et al (2000), the reflectivity and differential reflectivity data are converted first to rain and hail mixing ratios before assimilation, rather than being assimilated directly. The differential reflectivity data are necessary for such a conversion.

In our study, the original detailed ice microphysics parameterization of Lin et al (1983) is used and only standard reflectivity is assumed available. As a result, our problem is more difficult because more water and ice species have to be determined and no dual polarization information is available. Yet, the EnKF scheme does not seem to have any serious difficulty. Fig. 3 shows the distribution of the perturbation water vapor and five categories of water substances for the truth run and for selected assimilation experiments. It can be seen that the EnKF data assimilation system is able to establish detailed microphysical structures that have very high fidelity. The quality of actual analysis does depend on the usage and availability of data, as indicated quantitatively by the error plots discussed earlier.

To better understand the way the EnKF scheme works and the role of updating microphysical fields, we performed another experiment, named VrPnoIce, in which ice variables q_i, q_s, and q_h are not updated by the analysis and only radial velocity in precipitation regions are assimilated. The differences in error between VrPnoIce and VrP start to show from the third cycle and increase with successive cycles (Fig. 10). It can be seen that without the analysis update to these three variables, the *rms* error of q_s *analysis* is greater than even that of the ensemble-mean *forecast* of experiment VrP. The q_s field at 60 min (Fig. 3(*e*3)) exhibits larger spatial coverage in VrPnoIce and is less accurate than that of VrP (Fig. 3(*e*2)). The smaller difference in the earlier period reflects relatively weak link (through background error covariance) between V_r and these three variables. The link apparently becomes stronger and more effective in correcting errors in these fields at the later stage (as in the case of VrP). On the other hand, despite the lack of direct correction in VrPnoIce to q_i, q_s, and q_h, the errors in these fields are still reduced with time in general. Such reductions are achieved through model dynamics; when other model fields are improved, fields that are not directly updated have to adjust and become consistent with these fields. This points to the power of model-based dynamic data assimilation methods.

Further comparison of VrP and ZP helps us understand the interactions between the analyses of different variables. We see from Fig. 4, that the errors in the wind, temperature and pressure fields remain much higher in ZP case than in VrP for almost the entire assimilation period, with the differences being largest during the earlier cycles. During this early period, the assimilation of reflectivity data in ZP is very effectively in reducing errors in q_h, q_s as well as in q_r while the generally poor analysis of the other fields is causing rapid increases in the forecast error. We note here that the fact that q_r, q_s and q_h are directly related to Z does not guarantee a

good retrieval of them because without the help of error covariances offered by the EnKF scheme, the problem is underdetermined. When reflectivity data are combined with radial velocity data, very accurate microphysical fields can be retrieved (Fig. 2 and Fig. 3). With V_r data only (the case of VrP), the maximum value of q_h in Fig. 3(f 2) is 5.73 g kg^{-1} at 60 min while the true value is 6.63 g kg^{-1}. If the ice hydrometeors are not updated when assimilating V_r (experiment VrPnoIce), the maximum is larger than 1 g kg^{-1} smaller than the truth. The addition of full-coverage reflectivity data (Fig. 3(f5)) yields a maximum value of 6.37 g kg^{-1} that is closer to the truth.

4. Forecasts from ensemble-mean analyses

Since the goal of data assimilation is to provide a good initial condition for numerical weather prediction, in this section, we look at the quality of forecasts produced from the analyses. Fig. 11 shows the truth and two forecasts over a 130 minute period that are initialized from the ensemble-mean analyses from experiments VrP and VrPZF at 80 min. Within the first 20 minutes (from 80 to 100 min), the forecast of VrPZF maintains most details of the storm very well, including the strength, structure and location of the updraft and the further splitting of the left mover. Within the first hour of forecast, the position of the right mover is forecast very accurately by experiment VrPZF. After that, the center of the right moving cell is displaced about 4 to 8 km eastward during its southeastward propagation. The forecast beginning from VrP roughly captures the evolution of the storm but not as accurately as in VrPZF. The right mover propagates southeastwards even faster than in VrPZF and reaches the southern boundary earlier; at 210 min., the structure of forecast storm is rather different from that of true one.

Figure 12 shows the *rms* errors (averaged over the entire domain) of forecasts beginning from the ensemble-mean analyses of different times from experiments VrP and VrPZF.

Generally, more accurate estimate of the initial condition yields better forecast. It can be seen that for all these forecasts, within the first 80 minutes, both the forecast error and the forecast error growth rate of VrPZF (black thick curves) is smaller than those of VrP (gray thin curves). For the forecasts starting from 80 min., a large difference is seen in rainwater at around 120 min between the forecast errors of these two experiments (Fig. 12c). Within the first 40 minutes of forecast, the q_r forecast error in VrP grows much faster than in VrPZF. The low-level rainwater was over-forecast by experiment VrP. For example, at 105 min, the maximum value of the forecast rainwater at 250 m AGL is 8.54 g kg^{-1} for experiment VrP, while that of the truth is only 5.91 g kg^{-1} (not shown). The low-level rainwater was also over-forecast by experiment VrPZF, but not as much. The forecast distribution of the rainwater in VrPZF is always closer to the truth than in VrP. The comparison of the forecast errors of these two experiments further verifies the benefit of assimilating reflectivity data in both precipitation and clear air regions. For forecasts starting from the ensemble-mean analysis of VrF (not shown here), the errors are smaller than those of VrP, but generally bigger than those of VrPZF within the 80 minutes of forecast.

We note that for the forecasts starting from 100 min, the errors of VrPZF grow faster after 175 min and eventually become larger than those of VrP. This is found to be mainly associated with the slower southeastward propagation of the forecast storm in VrPZF than the true storm. Usually, phase error increases with the length of forecast. In this case, the center of the updraft forecast by experiment VrPZF is located 2 to 8 km, or 1 to 4 grid intervals, north of the true storm during the last 50 minutes. The center of the updraft in VrP is closer to the true storm during the later forecast period. However, within the last 20 minutes phase error also occurs within VrP. The forecast storm moves faster than the true storm and the center of the updraft becomes displaced by 2 to 4 km to the south of the true storm. Still, the storm structure in

VrPZF is better than that in VrP at this stage.

The success of the forecast depends not only on the accuracy of the analysis of storm, but also on the analysis of the storm environment. In our initial tests, in which the boundary zone is perturbed at the initial time of ensemble forecast, we obtained similarly good analyses. However, the forecasts from some of these analyses deteriorated quickly with time. In a forecast starting from analysis at 100 min, the right mover quickly decays and moves northeastward. We found in that case the low-level air was drier and the CAPE was reduced in the storm inflow region. Spurious cells triggered near the boundary in individual ensemble members contributed to this. Not perturbing the boundary at the initial time and the use of zero reflectivity information outside the precipitation regions helped alleviate the problem. For general applications, other means and data should be employed to ensure a good analysis of the storm environment, which for prediction purpose appears to be at least as important as the analysis of the storms themselves.

5. Summary and discussion

In this study we applied the ensemble Kalman filter technique to the assimilation of simulated radar radial velocity and reflectivity data, using a compressible model with a complex multi-class ice microphysics scheme. The inclusion of a complex multi-class ice microphysics has not been done before in either EnKF or 4DVAR assimilation of radar data. Doing so introduces additional model state (ice microphysics) variables that impose additional challenges to the retrieval problem. Previously published EnKF research did not carefully examine the impact of assimilating reflectivity data either, nor has attention been paid to the analysis quality of microphysical variables. Our work also represents the first time that a compressible model is used in the context of EnKF assimilation at the cloud scale.

The EnKF method is shown to have great potentials for the assimilation of such data and

for thunderstorm prediction. Although the observation operator for the reflectivity data is nonlinear, which may lead to non-Gaussian error distribution and violate the basic assumptions of the Kalman filter algorithm, direct assimilation of reflectivity data is shown to provide positive impact overall on the analysis and subsequent forecast.

Flow-dependent forecast error covariances estimated from the ensemble states play a critical role in the assimilation process. Through them not only can the wind and thermodynamic fields be retrieved accurately, all five categories of cloud and hydrometeor variables can also be retrieved successfully. Compared to 4DVAR, the EnKF is a much easier yet effective method to obtain microphysical fields that are compatible with the model dynamics, thermodynamics and microphysics. Compared to early efforts that insert radar observations directly into numerical models without the use of covariance information (e.g., Liou et al. 1991), the EnKF method enjoys a much greater success. The importance of spatial covariance information is further supported by that fact that worse analysis is obtained when the effective radius of covariance localization is too small (results not shown).

Reliable and dynamically consistent multivariate covariances between the observed quantities and the state variables not directly related to them can be obtained after a few assimilation cycles, even when the ensemble is started from initial guesses made of an environmental sounding plus random perturbations. Spatial covariance structures are shown to be dynamically consistent. After the initial two to three cycles, useful observational information can be spread in space and to indirectly related variables. Delaying the update of indirectly related variables until after first few cycles when assimilating reflectivity data produces the best analysis. Using reflectivity information in clear air regions is very beneficial in suppressing spurious storms (assuming that the radar network provides a complete spatial coverage). The forecast

initiated from the ensemble-mean analysis using V_r and full coverage reflectivity data is shown to be the best within about 80 minutes of all three forecasts presented. We should point out here if the EnKF assimilation system is run in a continuously-cycled mode, as an operational system should be, or if the ensemble is initalized in a more physical way, such as using the breeding method (Toth and Kalnay 1997), then the delay in updating indirectly related variables may not be necessary.

The EnKF data assimilation results reported here used 100 ensemble members. We found that 40 ensemble members are enough to produce good analysis result, except that the filter convergence rate is a little slower. With 40 members, the radius of the data influence region needs to be smaller. A radius of 6 km is found to produce the best result in this case.

Snyder and Zhang (2003) noted that the results of their EnKF experiments were subject to some random variability associated with the specific realizations of the initial ensemble and observation errors. In this study, we have repeated some of the experiments with different realizations and we did not notice any signifnicant deviation from the results reported here in terms of the error characteristics. The larger ensemble size used here may account for some of the reduced variability.

For larger applications, the analysis algorithm needs to be parallelized for distributed-memory parallel platforms. For our current application, we used shared-memory parallelization via OpenMP for the computation of the background error covariance $\mathbf{P}\mathbf{H}^{\mathrm{T}}$, which is the most expensive part of the algorithm. Four processors of IBM Regatta (model p-690) give a speedup of 3.5 for the analysis. The ensemble forecasts were distributed among a number of processors.

Finally, we note that caution should be used when interpreting OSSE results. Both forecast model and forward observation operators are assumed perfect in the current work. The

effects of model error need to be studied in the future. Much work is still needed in moving us in the direction of real case and real data. We also mention that similar set of assimilations using square-root EnKF, which does not perturb the observations (as used in e.g., Snyder and Zhang 2003), has also been conducted with generally similar conclusions.

Acknowledgement

This work was supported by NSF grant ATM0129892. The second author was also supported by NSF ATM-9909007, ATM-0331594 and EEC-0313747, a DOT-FAA grant via DOC-NOAA NA17RJ1227 and a grant from Chinese Natural Science Foundation No. 40028504. Dr. William Martin contributed to the initial development of the analysis code. The authors also benefited from discussions with Drs. Fuqing Zhang and David Dowell and the review of Chris Snyder. The IBM Regatta is operated by OSCER of University of Oklahoma.

References

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

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the Ensemble transform Kalman rilter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129,** 420.

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

Caya, A., J. Sun, and C. Snyder, 2004: A comparison between the 4D-VAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.***,** In review.

Crook, N. A., D. Dowell, and J. Sun, 2002: Wind and thermodynamic retrievals in a supercell

thunderstorm: 4DVar results. *19th Conf. Wea. Anal. Forecasting/15th Conf. Num. Wea. Pred.*, Amer. Meteor. Soc., 120-123.

Dowell, D., 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.***,** 1982-2005.

Dowell, D., F. Zhang, L. Wicker, C. Snyder, B. Skamarock, and A. Crook, 2002: Wind and thermodynamic retrievals in a supercell thunderstorm: Ensemble Kalman filter results. *19th Conf. Weather Anal. Forecasting/15th Conf. Num. Wea. Pred.*, Amer. Meteor. Soc., 375-378.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99(C5),** 10 143-10 162.

——, 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics,* **53,** 343-367.

Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.*, **51,** 249-280.

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 retrievals methods with WSR-88D data. *12th Conf. Num. Wea. Pred.*, Phoenix AZ, Amer. Meteor. Soc., 176-178.

Gao, J.-D., 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.

Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125,** 723-757.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129,** 2776-2790.

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

——, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129,** 123-137.

Kalnay, E., 2002: *Atmospheric modeling, data assimilation, and predictability*. Cambridge University Press, 341 pp.

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.

Liou, Y. C., T. Cal-Chen, and D. K. Lilly, 1991: Retrievals of wind, temperature and pressure from single-Doppler radar and a numerical model. *Preprint, 25th Int. Conf. on Radar Meteor.*, Paris, France, Amer. Meteor. Soc., 151-154.

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. *Preprint volume, 1st Symposium on Integrated Observation Systems*, Long Beach, CA, Amer. Meteor. Soc.

Lorenc, A., 2003: The potential of the ensemble Kalman filter for NWP - a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129,** 3183-3204.

Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor.*

Soc., **114,** 463-493.

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

Ray, P. S., B. Johnson, K. W. Johnson, J. S. Bradberry, J. J. Stephens, K. K. Wagner, R. B. Wilhelmson, and J. B. Klemp, 1981: The morphology of severe tornadic storms on 20 May 1977. *J. Atmos. Sci.*, **38,** 1643-1663.

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

Smith, P. L., Jr., C. G. Myers, and H. D. Orville, 1975: Radar reflectivity factor calculations in numerical cloud models using bulk parameterization of precipitation processes. *J. Appl. Meteor.*, **14,** 1156-1165.

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.

Sun, J., D. W. Flicker, and D. K. Lilly, 1991: Recovery of three-dimensional wind and temperature fields from simulated single-Doppler radar data. *J. Atmos. Sci.*, **48,** 876-890.

Tao, W.-K. and J. Simpson, 1993: Goddard cumulus ensemble model. Part I: Model description. *Terres. Atmos. Ocean Sci.*, **4,** 35-72.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble Square Root Filters{ast}. *Mon. Wea. Rev.*, **131,** 1485-1490.

Tong, M. and M. Xue, 2004: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model. *Extended Abstract, 20th Conf. Wea. Analy. Forecasting/16th Conf. Num. Wea. Pred.*, Seattle, WA, Amer. Meteor. Soc.

Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125,** 3297-3319.

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.

Wu, B., J. Verlinde, and J. Sun, 2000: Dynamical and Microphysical Retrievals from Doppler Radar Observations of a Deep Convective Cloud. *J. Atmos. Sci.*, **57,** 262-283.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Physics*, **75,** 161-193.

Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Physics*, **82,** 139-170.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phy.*, **76,** 143-165.

Zhang, F., C. , Snyder, and J. Sun, 2004: Impacts of initial estimate and observations on the

convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132,** 1238-1253.

List of Tables

Table 1. List of Data Assimilation Experiments

Experiment	Observation: Radial velocity (Vr) and/or Reflectivity (Z)	Update q_i, q_s, q_h	Update u, v, w, q_v, q_c, q_i when assimilating reflectivity
VrP	V_r (Z > 10 dBZ)	yes	
VrF	V_r	yes	
ZP	Z (Z > 10 dBZ)	yes	yes
VrZPa	V_r & Z (Z > 10 dBZ)	yes	yes
VrZPb	V_r & Z (Z > 10 dBZ)	yes	no
VrZPc	V_r & Z (Z > 10 dBZ)	yes	yes, start from 4th cycle
VrPZF	V_r (Z > 10dBZ) & Z	yes	yes, start from 4th cycle
VrPnoIce	V_r (Z > 10 dBZ)	no	

List of Figures

Fig. 1. Vertical velocity (contours and shading at intervals of 4 m s^{-1}) and horizontal perturbation wind (vectors, plotted every other grid point; m s^{-1}), at level z = 6 km: (a)-(d) truth simulation; analyses from VrP (e)-(h), ZP (i)-(l); VrZPc (m)-(p); VrPZF (q)-(t), at t = 40, 60, 80 and 100 min during the assimilation period.

Fig. 2. Horizontal perturbation wind (vectors; m s^{-1}), perturbation potential temperature (thick black lines for 0 K and thin dashed contours at 0.5 K intervals) and computed reflectivity (thick solid contours and shading at intervals of 5 dBZ, starting from 15 dBZ) at z = 250 m: truth (a)-(d), EnKF analyses from VrP (e)-(h), VrZPc (i)-(l) and VrPZF (m)-(p).

Fig. 3. The retrieved (a) q_v perturbation (solid (dash) contours represent positive (negative) values) (b) q_c, (c) q_r, (d) q_i, (e) q_s and (f) q_h. (contours and shading at intervals of 0.5 g kg^{-1} for q_v perturbation, q_c, q_i and q_s, and of 1.0 g kg^{-1} for q_r and q_h.) in the x-z plane at y = 29 km that pass through the maximum updraft at t = 60 min: (a1)-(f1) truth simulation, (a2)-(f2) VrP, (a3)-(f3) VrPnoIce, (a4)-(f4) ZP and (a5)-(f5) VrPZF.

Fig. 4. The rms errors of ensemble-mean forecast and analysis, averaged over points at which the reflectivity is greater than 10dBZ for: a) u (ms^{-1}), b) v, c) w and d) perturbation potential temperature θ', e) perturbation pressure p', f) q_c, g) q_r, h) q_v (the curves with larger values), q_i (the curves with lower values), i) q_s and j) q_h, for experiment VrP (gray), experiment ZP (black) and experiment VrF (black dotted lines). Units are shown in the plots. The drop of the error curves at specific times corresponds to the reduction of error by analysis.

Fig. 5. Forecast error correlations estimated from an ensemble at t = 80 min for experiment ZP in the x-z plane at y=29 km, which passes through the maximum updraft. The error correlation [thick solid (dash) contours represent positive (negative) correlations at intervals of 0.2]

between forecast reflectivity Z at x=32 km and z=6 km (indicated by a black dot) and (a) u, (b) w, (c) θ', (d) p', (e) q_c , (f) q_r, (g) q_i, and (h) q_h. The shading and thin solid (thin dashed) contours in (a) - (d) indicate positive (negative) values of model fields from the truth simulation with increment of 4 ms^{-1} for u and w; 2 K for θ' and 40Pa for p'. The shading and thin contours in (e)-(h) indicate the values of mixing ratio of water and ice fields from the truth simulation with increment of 0.5 g kg^{-1} for q_c, and q_i and 1 g kg^{-1} for q_r and q_h.

Fig. 6. As in Fig. 4, but for experiment VrZPa (black) and experiment VrP (gray).

Fig. 7. As in Fig. 4, but for experiment VrZPb (black) and experiment VrP (gray).

Fig. 8. As in Fig. 4, but for experiment VrZPc (black) and experiment VrP (gray).

Fig. 9. As in Fig. 4, but for experiment VrPZF (thin black curves) and experiment VrP (thin gray curves). The additional thick black curves are for the analysis and forecast ensemble spread of VrPZF, and the spread is calculated only at the points where reflectivity is greater than 10 dBZ, as the errors are.

Fig. 10. RMS error of (a) q_i, (b) q_s, (c) q_h for VrP (gray) and VrPnoIce (black).

Fig. 11. Vertical velocity (contours and shading at intervals of 4 ms^{-1}) and horizontal wind perturbation (vectors; ms^{-1}) at 6 km level for (a)-(e) truth simulation and the forecast beginning from the ensemble-mean analysis at t = 80 min: (f)-(j) forecast from ensemble-mean analysis of VrPZF, and (k)-(o) forecast from ensemble-mean analysis of VrP. Note the difference in the plotting domains before and after 150 min.

Fig. 12. The *rms* errors of forecasts averaged over the entire domain for: (a) w (m s^{-1}), (b) θ' (K) and (c) q_r (g kg^{-1}). The forecasts begin from ensemble-mean analysis at t = 60 min (dot), t =

80 min (dash) and t = 100 min (solid) of experiment VrP (gray thin curves), VrPZF (black thick curves).

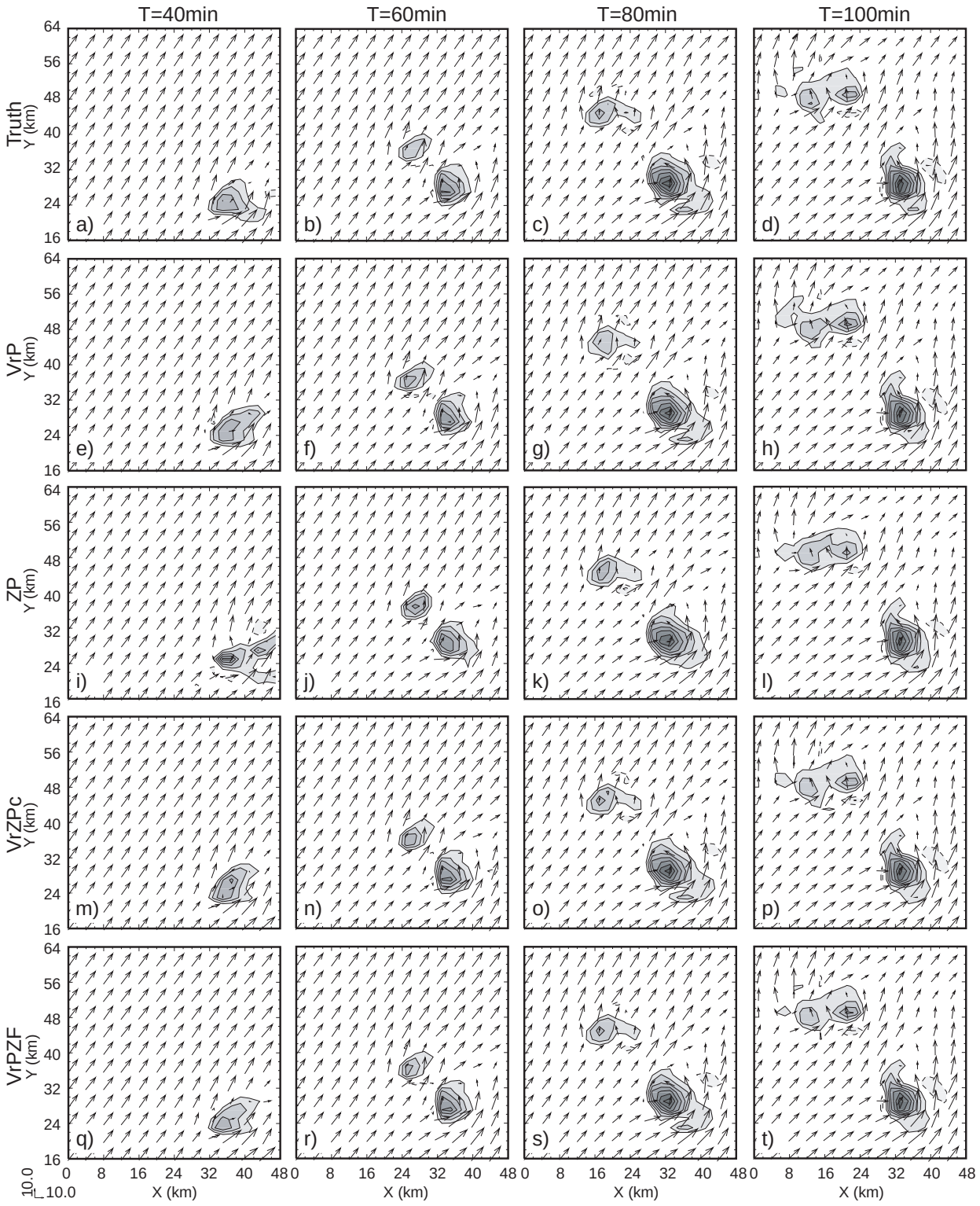

Fig. 1. Vertical velocity (contours and shading at intervals of 4 m s^{-1}) and horizontal perturbation wind (vectors, plotted every other grid point; m s^{-1}), at level z = 6 km: (a)-(d) truth simulation; analyses from VrP (e)-(h), ZP (i)-(l); VrZPc (m)-(p); VrPZF (q)-(t), at t = 40, 60, 80 and 100 min during the assimilation period.

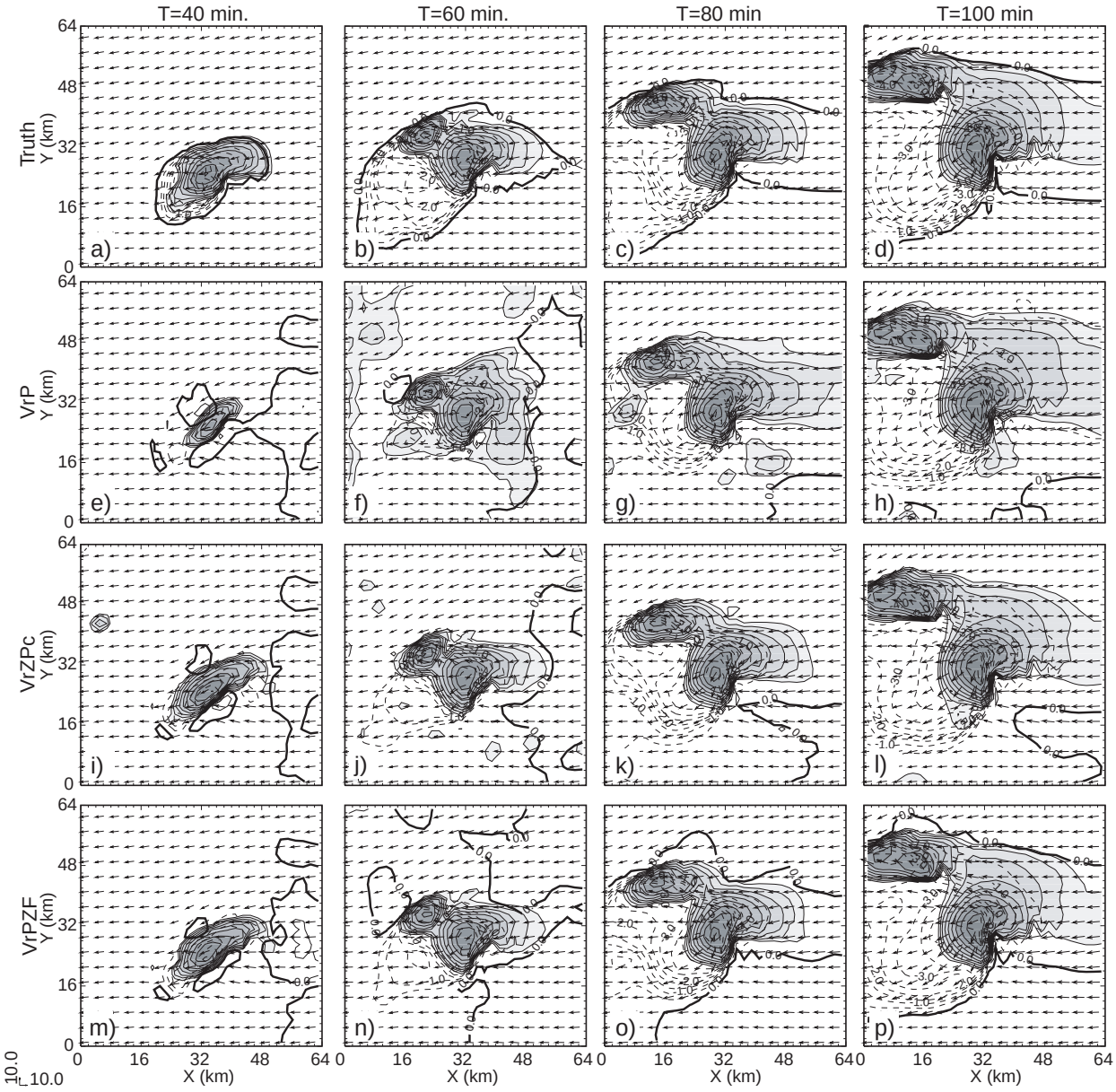

Fig. 2. Horizontal perturbation wind (vectors; m s^{-1}), perturbation potential temperature (thick black lines for 0 K and thin dashed contours at 0.5 K intervals) and computed reflectivity (thick solid contours and shading at intervals of 5 dBZ, starting from 15 dBZ) at z = 250 m: truth (a)-(d), EnKF analyses from VrP (e)-(h), VrZPc (i)-(l) and VrPZF (m)-(p).

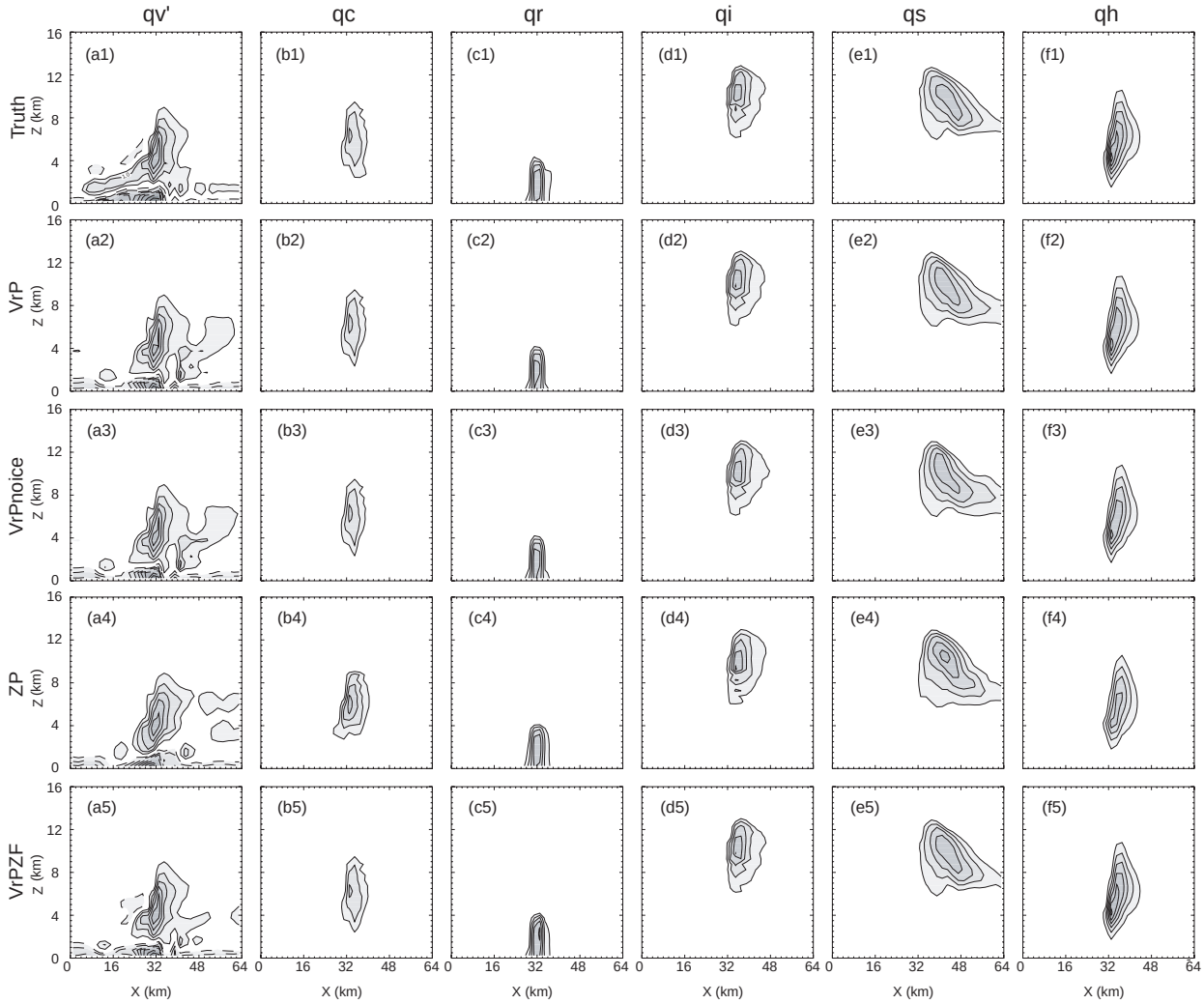

Fig. 3. The retrieved (a) q_v perturbation (solid (dash) contours represent positive (negative) values) (b) q_c, (c) q_r, (d) q_i, (e) q_s and (f) q_h. (contours and shading at intervals of 0.5 g kg^{-1} for q_v perturbation, q_c, q_i and q_s, and of 1.0 g kg^{-1} for q_r and q_h.) in the *x*-*z* plane at y = 29 km that pass through the maximum updraft at t = 60 min: (a1)-(f1) truth simulation, (a2)-(f2) VrP, (a3)-(f3) VrPnoIce, (a4)-(f4) ZP and (a5)-(f5) VrPZF.

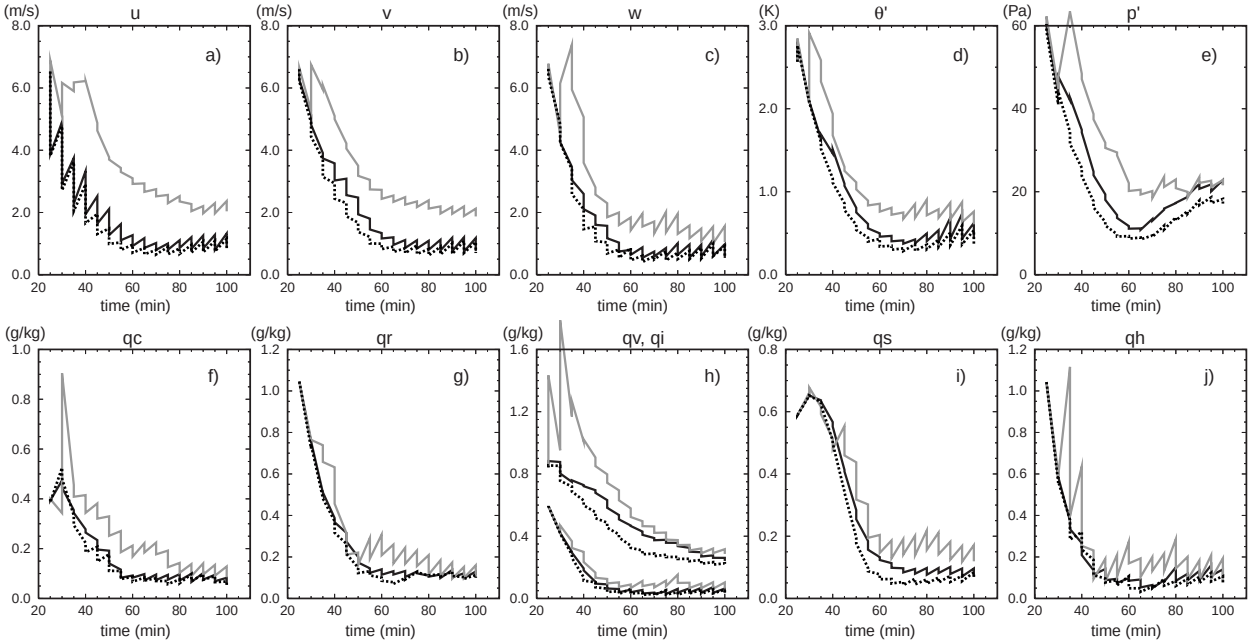

Fig. 4. The rms errors of ensemble-mean forecast and analysis, averaged over points at which the reflectivity is greater than 10dBZ for: a) u (ms^{-1}), b) v, c) w and d) perturbation potential temperature θ', e) perturbation pressure p', f) q_c, g) q_r, h) q_v (the curves with larger values), q_i (the curves with lower values), i) q_s and j) q_h, for experiment VrP (gray), experiment ZP (black) and experiment VrF (black dotted lines). Units are shown in the plots. The drop of the error curves at specific times corresponds to the reduction of error by analysis.

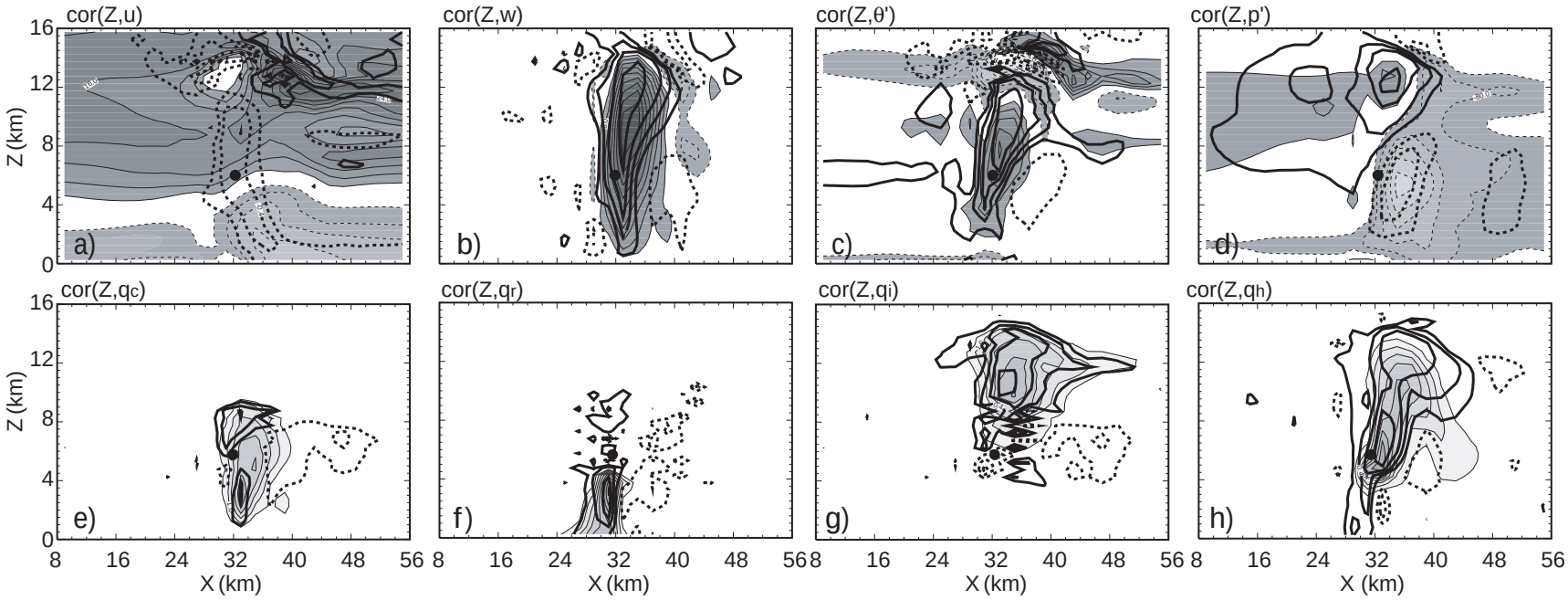

Fig. 5. Forecast error correlations estimated from an ensemble at t = 80 min for experiment ZP in the x-z plane at y=29 km, which passes through the maximum updraft. The error correlation [thick solid (dash) contours represent positive (negative) correlations at intervals of 0.2] between forecast reflectivity Z at x=32 km and z=6 km (indicated by a black dot) and (a) u, (b) w, (c) θ', (d) p', (e) q_c , (f) q_r, (g) q_i, and (h) q_h. The shading and thin solid (thin dashed) contours in (a) - (d) indicate positive (negative) values of model fields from the truth simulation with increment of 4 ms^{-1} for u and w; 2 K for θ' and 40Pa for p'. The shading and thin contours in (e)-(h) indicate the values of mixing ratio of water and ice fields from the truth simulation with increment of 0.5 g kg^{-1} for q_c, and q_i and 1 g kg^{-1} for q_r and q_h.

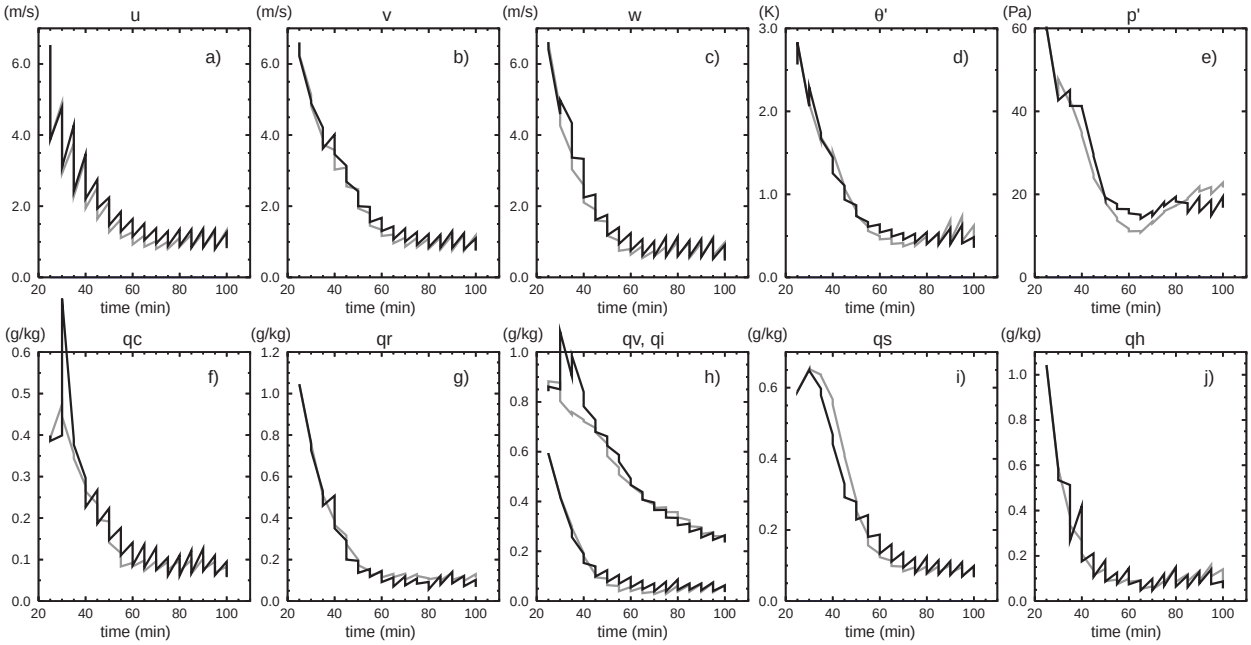

Fig. 6. As in Fig. 4, but for experiment VrZPa (black) and experiment VrP (gray).

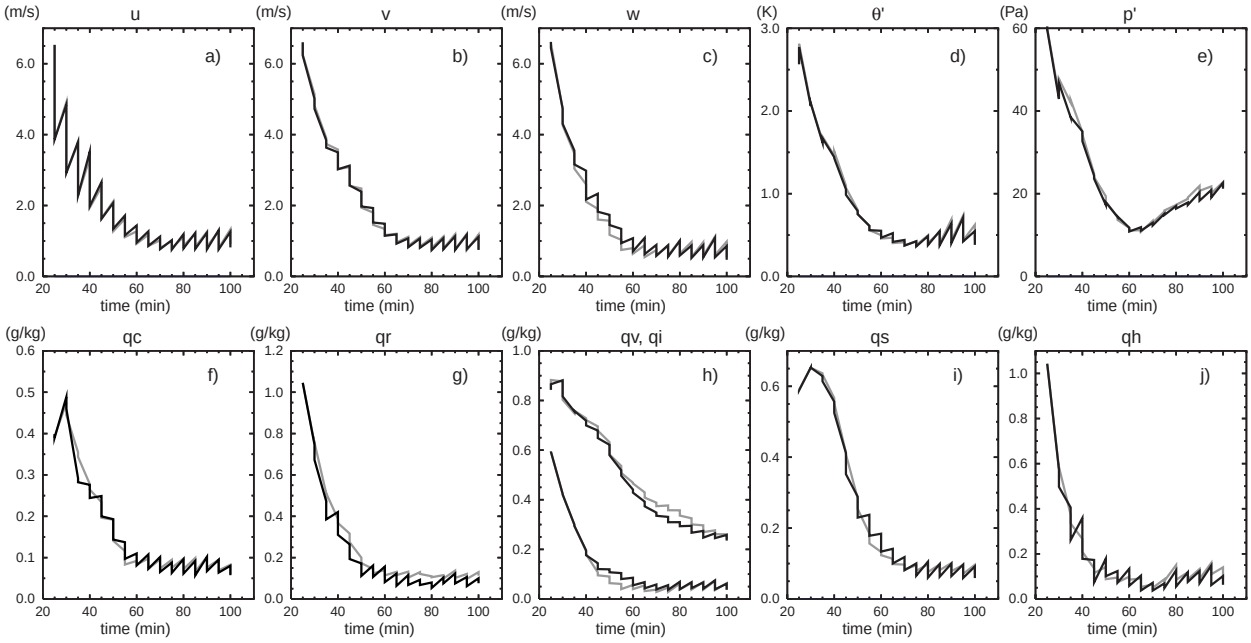

Fig. 7. As in Fig. 4, but for experiment VrZPb (black) and experiment VrP (gray).

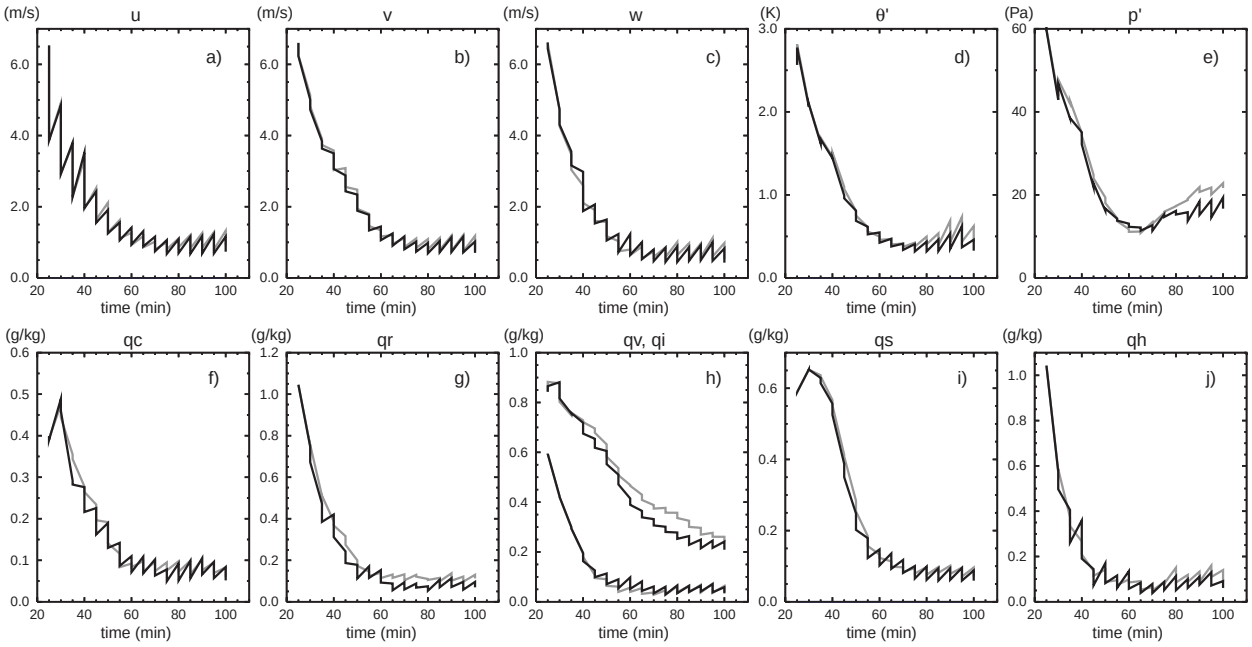

Fig. 8. As in Fig. 4, but for experiment VrZPc (black) and experiment VrP (gray).

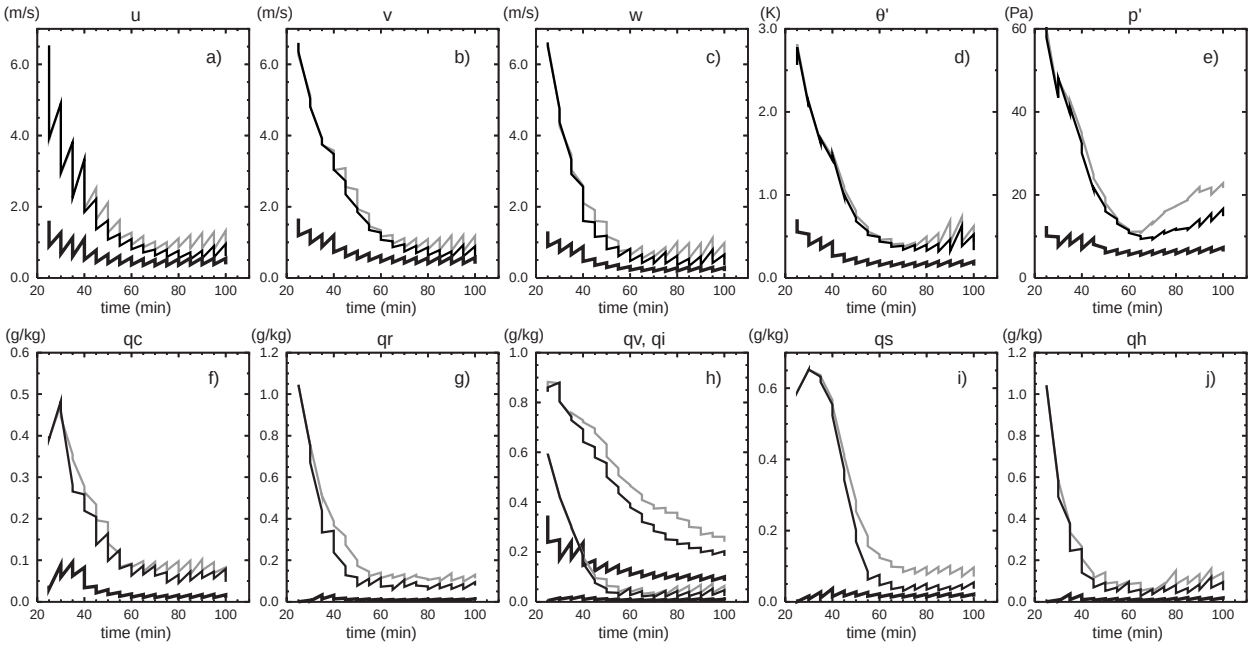

Fig. 9. As in Fig. 4, but for experiment VrPZF (thin black curves) and experiment VrP (thin gray curves). The additional thick black curves are for the analysis and forecast ensemble spread of VrPZF, and the spread is calculated only at the points where reflectivity is greater than 10 dBZ, as the errors are.

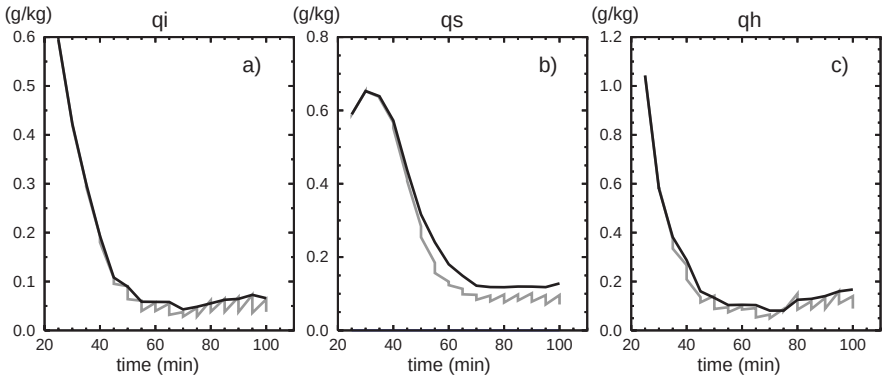

Fig. 10. RMS error of (a) q_i, (b) q_s, (c) q_h for VrP (gray) and VrPnoIce (black).

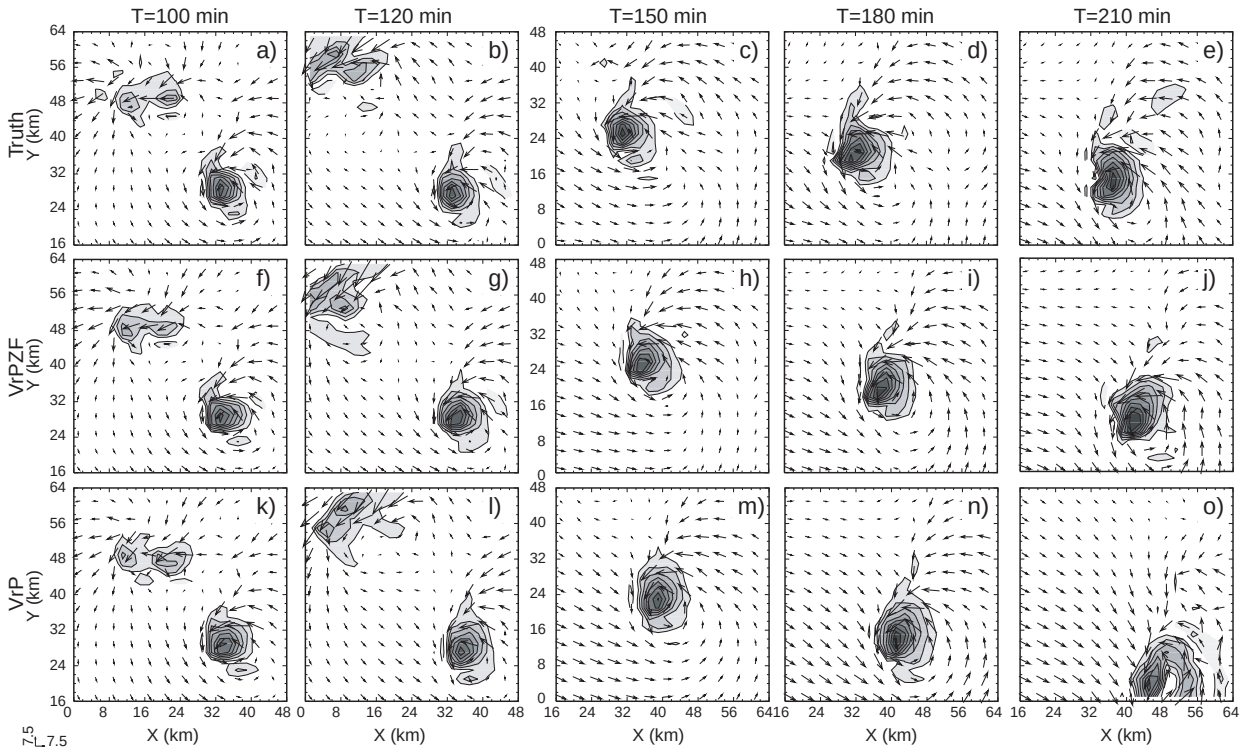

Fig. 11. Vertical velocity (contours and shading at intervals of 4 ms^{-1}) and horizontal wind perturbation (vectors; ms^{-1}) at 6 km level for (a)-(e) truth simulation and the forecast beginning from the ensemble-mean analysis at t = 80 min: (f)-(j) forecast from ensemble-mean analysis of VrPZF, and (k)-(o) forecast from ensemble-mean analysis of VrP. Note the difference in the plotting domains before and after 150 min.

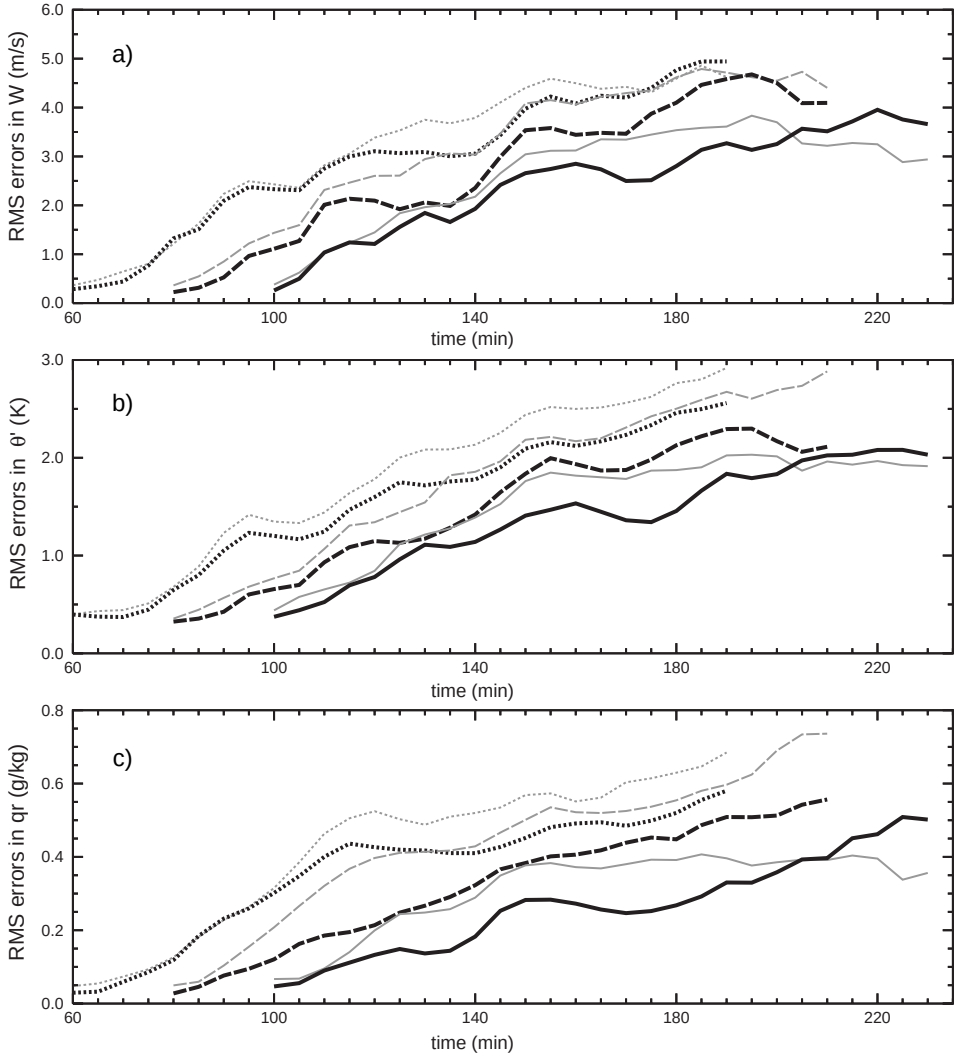

Fig. 12. The *rms* errors of forecasts averaged over the entire domain for: (a) w (m s^{-1}), (b) θ' (K) and (c) q_r (g kg^{-1}). The forecasts begin from ensemble-mean analysis at t = 60 min (dot), t = 80 min (dash) and t = 100 min (solid) of experiment VrP (gray thin curves), VrPZF (black thick curves).