P1.30 SIMULTANEOUS RETRIEVAL OF MICROPHYSICAL PARAMETERS AND ATMOSPHERIC STATE VARIABLES WITH RADAR DATA AND ENSEMBLE KALMAN FILTER METHOD

Mingjing Tong* and Ming Xue
School of Meteorology and Center for Analysis and Prediction of Storms
University of Oklahoma Norman, Oklahoma, 73019

1 . INTRODUCTION

After several years' implementation, the ensemble Kalman filter (Burgers *et al.* 1998; Evensen 1994; 2003; Houtekamer and Mitchell 1998; Whitaker and Hamill 2002) was found to be a potential atmospheric data assimilation method for both large scale (Houtekamer and Mitchell 2001) and small scale (Dowell *et al.* 2004; Snyder and Zhang 2003; Tong and Xue 2005; Zhang *et al.* 2004) applications. The most important advantage of the ensemble-based data assimilation scheme (Anderson 2001; Bishop *et al.* 2001; Houtekamer and Mitchell 1998; Tippett *et al.* 2003; Whitaker and Hamill 2002) is that it provides a practical way to calculate and evolve the error statistics by using an ensemble to represent the probability density function (PDF) of the error and propagating the PDF through ensemble forecast (Evensen 1994). In the context of convective scale data assimilation, the flow-dependent multivariate background error covariances, provided by the ensemble, play an essential role, with which dynamically consistent wind, thermodynamic and microphysical fields can be retrieved accurately from simulated radar radial velocity and reflectivity observations (Snyder and Zhang 2003; Tong and Xue 2005; Zhang *et al.* 2004). Encouraged by mostly Observing System Simulation Experiment (OSSE) results, people have been moving toward using real data (Dowell *et al.* 2004; Dowell and Wicker 2004; Houtekamer *et al.* 2005) and exploring the possibility of operational implementation of this data assimilation method.

In most OSSE studies, only forecast error due to uncertain initial conditions was taken into account, while forecast error due to model deficiencies was neglected. However, in real-world applications, the first challenge might be encountered by the ensemble Kalman filter (hereafter EnKF) is the model error. The essential part of the EnKF is the error covariance, which determines the accuracy of the analysis. Convective-scale data assimilation, which mainly depends on the observations from the radar, is more of a retrieval problem, since the model state is not directly observed. Therefore, multivariate error covariance is more important for convective-scale data assimilation. The flow-dependent and multivariate covariacnes are built up through model evolution of each ensemble member. Whether they can be determined correctly depends on whether the model evolutions are correct. The systematic model error can cause the ensemble member not be drawn from the distribution that produces truth because the model attractor and the system attractor differ (Hansen 2002).

Systematic errors can result from uncertain parameters used in the model. One way to account for the model error of this type is through parameter estimation, so that the parameters can be more adequately constrained by available observations. Different techniques, e.g. maximum likelihood method (Dee 1995); extended Kalman filter (Hao *et al.* 1995); variational method (Derber 1989; Yu and O'Brien 1991; Zou *et al.* 1992) have been applied for parameter estimation purpose in meteorology. Navon (1992) reviewed the variational approach via an adjoint model for parameter estimation and discussed the issue of parameter identifiability. Recently, Crook et al. (2004) applied the 4DVAR method to estimate a coefficient of the terminal velocity formulation used in their cloud model.

Anderson (2001) first suggested that the EnKF can be used for parameter estimation by including the model parameters as part of the model state and being estimated simultaneously with the model state. Annan et al. (2005a) applied the EnKF method to simultaneously estimate 12 parameters in a low-resolution coupled atmosphere-ocean model with steady-state dynamics, which works successfully with identical twin testing. Annan and Hargreavers (2004) also successfully applied this method to perform multivariate parameter estimation in the presence of chaotic dynamic with the Lorenz model. More recently, they have extended their results to a realistic intermediate

* Corresponding author address:
Mingjing Tong, SOM, SEC 1337, 100 E. Boyd,
Norman OK 73019,mingjing@ou.edu.

complexity atmospheric GCM with both an identical twin experiment and reanalysis data (Annan *et al.* 2005b). However, in contrast to weather prediction, the climate forecasts depend strongly on parameterizations rather than initial conditions (Annan *et al.* 2005a). Kivman (2003) found that the EnKF performed poorly when applied to simultaneous state and parameter estimation in the Lorenz model. He attributed this to utilizing only two statistical moments in the analysis step by Kalman filter based method, which is unable to deal with highly non-Gaussian probability distribution in the parameter space. Aksoy *et al.* (2005) applied the EnKF method to simultaneous estimation of up to 6 parameters and model state with a two-dimensional, hydrostatic, non-rotating, and incompressible sea-breeze model. They found that the estimation of single imperfect parameters with the EnKF is successful, while the results were declined when estimated parameter increased.

For convective-scale numerical weather prediction, explicit microphysics schemes are used to predict the evolution of cloud and precipitation. Most cloud model utilize bulk microphysics scheme, in which the particle distributions are parameterized in functional forms. McCumber *et al.* (1991) tested the sensitivity of tropical convective system simulation to the change of size distribution parameters. Ferrier *et al.* (1995) also did some microphysical parameter sensitivity study when simulating squall systems in midlatitude continental and tropical environment. More recently, Gilmore *et al.* (2004) examined the precipitation uncertainty of simulated midlatitude multicell and supercell due to variations in particle parameters. All these studies demonstrated that the structure and the evolution of simulated convective systems are sensitivity to microphysical parameterization. Variations in microphysical parameters, such as collection coefficients, drop size distribution parameters and particle densities, have profound effect upon the characteristics of precipitation systems and their associated dynamical process.

As indicated by those simulation sensitivity studies, microphysical parameterization could be an important source of model error for convective scale data assimilation and prediction. The purpose of this study is to examine the impact of the errors in some of these microphysical parameters on the retrieved model state and to determine the ability of the EnKF method in correcting these errors through parameter estimation.

The remainder of this paper is organized as follows. Section 2 outlines our ensemble square root filter (EnSRF) data assimilation configurations. In section 3, we will describe our parameter estimation process. The results of the parameter retrieval experiments are discussed in section 4. Conclusions and discussions are given in section 5.

2 . DATA ASSIMILATION ENVIRONMENT

2.1. The Model and the Natural Simulation

The forecast model and the truth simulation are inherited from Tong and Xue (2005, hereafter TX05). Briefly, the Advanced Regional Prediction System (Xue *et al.* 2000; Xue *et al.* 2003; Xue *et al.* 2001), a fully compressible and nonhydrostatic atmospheric prediction system is used. The ARPS contains 12 prognostic state variables, including three velocity components u, v, w, potential temperature θ, pressure p, the mixing ratios for water vapor q_v, cloud water q_c, rainwater q_r, cloud ice q_i, snow q_s and hail q_h, plus the turbulence kinetic energy used by the 1.5-order subgrid-scale turbulence closure scheme.

The truth simulation or the nature run is for the May 20, 1977 Del City, Oklahoma supercell storm case (Ray et al. 1981). The physical domain is 64×64×16 km^3. The grid spacing is 2 km in the horizontal directions and 0.5 km in the vertical. A sounding of 3300 J kg^{-1} CAPE (Xue et al. 2001)is used to define the environmental condition and a 4 K ellipsoidal thermal bubble is used to initiate the storm. Open conditions are used at the lateral boundaries. Free-slip conditions are applied to the top and the bottom boundaries. A constant wind of $u = 3$ m s^{-1} and $v = 14$ m s^{-1} is subtracted from the observed sounding to keep the primary storm cell near the center of model grid. More detailed information about the natural run can be found in TX05. The actual sounding used by the truth simulation can be found in Xue et al. (2001) and the general evolution of the storm is similar to that documented in the same paper.

2.2. Simulated Radar Data

We assume that the radial velocity and reflectivity data are available from a WSR-88D radar located at the south-west corner of the model domain. The WSR-88D radar is assumed to operate in standard precipitation-mode, having 14 elevations with the lowest elevation at 0.5º and highest at 19.5º. The maximum range is 230 km. The effects of the curvature of the earth and the beam bending due to vertical change of refractivity are taken into account by using the simply effective earth radius model discussed in (Doviak and Zrnic 1993).

Following Xue *et al.* (2005, hereafter XTD05), the simulated observations are assumed to be available on the original radar elevation levels, i.e., the data are on the radar plan position indicator (PPI) rather than at the model vertical levels. We do assume that on each elevation level, radar observations are already interpolated from the radar polar coordinate to the Cartesian coordinate; in another word, the observations are found in the vertical columns through the model scalar points. A simplified radar emulator that does power-gain-based sampling in the vertical direction is used to interpolate the data from the model vertical levels to the radar elevation levels:

$$\eta_e = \frac{\sum G\eta_g \Delta z}{\sum G\Delta z}, \tag{1}$$

where η_e and η_g are respectively the elevation level and grid point values of either radial velocity (V_r) or reflectivity factor (Z in mm^6 m^{-3}). Δz is the depth of the layer in which grid point value η_g is found. The two-way power weighting function is assumed to be Gaussian and has the form of

$$G = \exp\left[-4\ln 4\left(\frac{\phi_g - \phi_0}{\phi_w}\right)^2\right], \tag{2}$$

following (Wood and Brown 1997), where ϕ_w is the 1 degree beam width. ϕ_g is the elevation angle for the grid point value and ϕ_0 the elevation at the beam center.

The grid point values of radial velocity involved in the numerator of are calculated from

$$V_{rg} = u\cos\alpha\sin\beta + v\cos\alpha\cos\beta + (w - w_t)\sin\alpha, \tag{3}$$

where α is the *local* elevation angle and β the azimuth angle of the radar beam that goes through the given grid point. *u, v* and *w* are the model air velocity components interpolated to the scalar point of a staggered model grid. w_t is the fall velocity calculated on scalar point. The fall velocity is calculated from

$$w_t = \frac{w_{tr}Z_{er} + w_{ts}Z_{es} + w_{th}Z_{eh}}{Z_{er} + Z_{es} + Z_{eh}}, \tag{4}$$

where Z_R, Z_S and Z_H are the reflectivity factors (in mm^6 m^{-3}) of rain, snow and hail respectively. w_{tr}, w_{ts} and w_{th} are the mass-weighted mean terminal velocities of rain, snow and hail, which are calculated using Eq. (11)-(13) of Lin et al. (1983) and are consistent with those used in the model's microphysics scheme. After V_r is sampled from the grid, random error drawn from a normal distribution with zero mean and standard deviation of 1 m s^{-1} are added as the simulated observation errors.

The simulated reflectivity factor at grid point is calculated from the mixing ratio of rain, snow and hail. The same formulations of reflectivity factor for rain and snow are used as those in TX05. The reflectivity equation of dry hail is also included and follows Smith *et al.* (1975). The transition zone from dry to wet hail is defined to be -2.5°C to 2.5°C. After the values of equivalent reflectivity on elevation levels are obtained, they are transformed into the commonly used reflectivity, *Z*, in dBZ (Z=10 log$_{10}$(Z_e)). In our system, reflectivity *Z*, in dBZ, is directly assimilated.

2.3. The EnSRF Data Assimilation Method

An ensemble square root filter (EnSRF, Whitaker and Hamill 2002) is used in this study and the actual implementation of the filter for radar data assimilation is described in Xue et al. (2005) with noted differences. For the control OSS experiment (CNTL), the model is assumed to be perfect, i.e., the default values of the microphysical parameters in the ARPS are assumed to be the true values. The procedure of initializing the ensemble is different from that in TX05. Spatially smoothed perturbations are added to the first guess of the initial condition, which is a homogeneous environment defined by the sounding. For each model variable at grid point *(l, m, n)*, the spatially smoothed perturbation is calculated as

$$\varepsilon(l,m,n) = E\sum_{(i,j,k)\in S} r(i,j,k)W(i,j,k), \tag{5}$$

where $r(i,j,k)$ is a random number sampled from a normal distribution with zero mean and standard deviation of 1. $W(i,j,k)$ is a 3D distance-dependent weighting function. E is a scaling parameter to obtain the right standard deviation of the perturbation field. The fifth-order correlation function (Eq. (4.10) of Gaspari and Cohn 1999) is used here to calculate *W*. The sum is over grid points, which are located within the 3D radius of influence. The radius of influence is 6km, which was chosen based on the correlation length scale of errors obtained in TX05. The same correlation function and cut off radius are used by the covariance localization.

After the smoothed perturbations are obtained, they are rescaled, i.e., E in Eq. 1 is determined, so that the standard deviation equals to that of the specified value for each variable. The standard deviations of the perturbations are, respectively, 2 m/s for velocity components, 2K for perturbation potential temperature, 0.6 g kg^{-1} for q_v, q_r and q_h,

0.4 g kg^{-1} for q_i and q_s, and 0.2 g kg^{-1} for q_c. For the mixing ratio of hydrometeors, the perturbations are only added within 6km distance from the first observed precipitation region in horizontal. They are further limited to the vertical levels, where the hydrometeors are expected. Negative values of perturbed mixing ratios are set to zero. The perturbations for velocity components, potential temperature and specific humidity are added to the entire domain except the boundary zone. Some spurious cells triggered by adding perturbations in non-precipitation region can be suppressed by assimilating reflectivity everywhere. We found that by using the spatially smoothed initial perturbations, the ensemble spread of most model variables grows within the first 5 min forecast. Perturbing microphysical fields also contributes to larger ensemble spread for microphysical variables. Larger initial ensemble spread results in smaller ensemble mean rms errors in early assimilation period (see green curves Fig. 6). We also found that with this new initial perturbation method, updating model variables that are indirectly related to reflectivity via observation operation does not hurt the analysis at early assimilation period. Therefore, with our current assimilation configuration, we do not hold the update of those variables when assimilating reflectivity data as what we did in TX05.

The same covariance localization procedure as that in TX05 and XTD05 are applied here to avoid the influence of unreliable covariances at large distance from the observations. No covariance inflation is applied here, because the difference of the analysis rms (root-mean-square) errors caused by covariance inflation is not as large as that caused by different realization of the initial perturbations, i.e., using different set of initial ensemble members. We do find some sensitivity of the analysis to the realization of the initial perturbations and mostly in the first few cycles, but the sensitivity is not as large as that found in Snyder and Zhang (2003).

Forty ensemble members are used. The first ensemble forecast start at 20 minutes of model time. Ensemble members are integrated for 5 minutes before the first analysis. Both radial velocity and reflectivity, including reflectivity in non-precipitation region, are assimilated in all experiments.

3 . THE DESIGN OF THE PARAMETER ESTIMATION EXPERIMENTS

3.1. Microphysics Scheme and Selection of Parameters to be Estimated

The ice microphysics scheme in the ARPS is a 5-class (cloud water, rain, cloud ice, snow and hail/graupel) single moment scheme after Lin et al. (1983). The scheme assumes that the particle size distribution functions for rain, snow and hail/graupel have an exponential form:

$$n_x(D) = n_{0x} \exp(-\lambda_x D_x), \tag{6}$$

where *x* represents *r* (rain), *s* (snow) or *h* (hail). $n_x(D)\delta D$ is the number of drops per unit volume between diameters *D* and *D*+δ*D*. n_{0x} is the intercept parameter, which is the value of n_x for *D*=0. The slope parameter, which equals to the inverse of the mean size diameter of each distribution, is diagnosed as:

$$\lambda_x = \left(\frac{\pi \rho_x n_{0x}}{\rho q_x} \right)^{0.25}, \tag{7}$$

where ρ_x is the constant particle density, ρ is the air density and q_x is the hydrometeor mixing ratio.

The limitation of the single moment bulk microphysics scheme is that the intercept n_{0x} and the density ρ_x are prescribed constants. It can be seen from Eqs. (6) and (7) that for a given mixing ratio q_x, the larger the intercept or the density, the more the hydrometeor spectrum are weighted toward small particles. When performing model simulation, adjusting these constant parameters can directly impact the bulk terminal velocity and the number concentration of species, which result in the change of the trajectory of the hydrometeors within the cloud and the particle growth rate. These changes in the microphysical processes will affect the water budgets within the cloud and hence the latent heating and hydrometer loading, which lead to the change of the buoyancy and subsequent updraft and downdraft pattern.

The limitation of the 3-ice microphysics scheme is that the parameterization can not represent the convective clouds in a variety of large-scale environments and are not necessarily suitable for different precipitation systems. For example, the parameterization of LFO83 microphysical scheme is formulated for the intense continental storms with the presence of high-density hail. The 3-ice scheme of (Rutledge and Hobbs 1983; 1984) is more suitable for oceanic systems, because it represents the large precipitating ice in the form of graupel. The differences come from either the treatment of the microphysical process or using different parameters, such as hydrometeor densities and intercept parameters.

The way to treat the microphysical processes is not the focus of this study. What we are interested in is how much our storm-scale data assimilated can be influenced by prescribing those adjustable parameters and whether we can correct the incorrectly specified parameters by using the data, i.e., retrieving the parameters as well as the model state with radar observations. The parameters selected for this study are the intercepts of rain, snow and hail size distributions, the density of snow and the density of hail. Observation and sensitivity studies (e.g. Ferrier *et al.* 1995; Passarelli 1978) indicate that the coefficients associated with the formula for hydrometeor fall speeds and the collection efficiency parameters are also uncertain parameters, which affect the microphysical processes significantly. In this study, we only focus on the density and the intercept parameters, partly because they are easier be controlled in the model code. Estimate fall speed and collection efficiency parameters could be our future work.

3.2. Parameter Estimation with EnSRF

With the single-moment bulk microphysics scheme, the intercept parameters and the bulk density of snow and hail are spatially and temporally constants. The default values of the intercept parameters for rain, snow and hail size distributions in the ARPS are 8×10^{-2} cm^{-4}, 3×10^{-2} cm^{-4} and 4×10^{-2} cm^{-4}, respectively, following LFO83. The density of rainwater, snow and hail are specified to be 1.0 g cm^{-3}, 0.1 g cm^{-3} and 0.913 g cm^{-3}, respectively (see Table 1).

Table 1 List of experiments and the first guess values of microphysical parameters

Experiments*	n_{0r} (cm^{-4})	n_{0s} (cm^{-4})	n_{0h} (cm^{-4})	ρ_s (g cm^{-3})	ρ_h (g cm^{-3})
CNTL	0.08	0.03	4×10^{-4}	0.1	0.913
Experiments retrieving or with wrong hail parameters In experiment names HN*a*ρ*b* or HN*a*ρ*b*NE, *a* is the exponent without the minus sign in the intercept parameter, e.g., 4×10^{-a}, and *b* is the first digit after the decimal point in density.					
HN6ρ9/HN6ρ9NE	0.08	0.03	**4×10^{-6}**	0.1	0.913
HN3ρ9/HN3ρ9NE	0.08	0.03	**4×10^{-3}**	0.1	0.913
HN4ρ4/ HN4ρ4NE	0.08	0.03	4×10^{-4}	0.1	**0.4**
HN3ρ4/ HN3ρ4NE	0.08	0.03	**4×10^{-3}**	0.1	**0.4**
Experiments retrieving or with wrong snow parameters In experiment names SN*a*ρ*b* or SN*a*ρ*b*NE, *a* and *b* represent the digits after the decimal point in the intercept parameter and density, respectively.					
SN007ρ1/SN007ρ1NE	0.08	**0.007**	4×10^{-4}	0.1	0.913
SN1ρ1/ SN1ρ1NE	0.08	**0.1**	4×10^{-4}	0.1	0.913
SN3ρ1/ SN3ρ1NE	0.08	**0.3**	4×10^{-4}	0.1	0.913
SN03ρ4/ SN03ρ4NE	0.08	0.03	4×10^{-4}	**0.4**	0.913
SN007ρ4/ SN007ρ4NE	0.08	**0.007**	4×10^{-4}	**0.4**	0.913
Experiments retrieving or with wrong rain parameters In the experiment names, RN*a* and RN*a*NE, *a* represents the digits after the decimal point in the intercept parameter.					
RN8/RN8NE	**0.8**	0.03	4×10^{-4}	0.1	0.913
RN3/RN3NE	**0.3**	0.03	4×10^{-4}	0.1	0.913
RN03/RN03NE	**0.03**	0.03	4×10^{-4}	0.1	0.913
Experiments retrieving or with three wrong intercept parameters					
RSHNa/ RSHn_0aNE	**0.30**	**0.1**	**4×10^{-3}**	0.1	0.913
RSHNb/ RSHn_0bNE	**0.30**	**0.007**	**4×10^{-3}**	0.1	0.913

* In the experiment names, H, S, R, N and ρ denote hail, snow, rain, intercept parameter and density, respectively. The experiment with the name ending with 'NE' means that the parameter estimation or retrieval is not performed, and the wrong initial guesses of the parameters are kept throughout the assimilation cycles. The numbers in bold represent initial guesses that deviate from the true values.

A number of observational studies indicate that the intercept parameters of hydrometeor distribution can vary widely among precipitation systems occur in different large-scale environment. Also, within a same precipitation system the intercept parameters can vary spatially and with the evolution of the system, although in this study we do not consider the spatial and temporal variations of the microphysical parameters. The hail/graupel intercept n_{0h}, reviewed by Gilmore *et al.* (2004) (Gilmore et al. 2004), ranges from 10^{-6} to greater than 10^{0} cm^{-4}. Observed snow intercept parameter n_{0s} varies from 10^{-3} cm^{-4} to 10^{0} cm^{-4} (Braham 1990; Houze *et al.* 1979; Houze *et al.* 1980; Lo and Jr. 1982; Mitchell 1988; Passarelli 1978). For raindrop spectra, several studies (Joss and Waldvogel 1969; Sekhon and Srivastava 1971; Srivastava 1971; Waldvogel 1974) demonstrated that n_{0r} cannot be a constant as defined by Marshall and Palmer (1948), but is a function of rainfall rate. Joss et al. (1968) found that n_{0r} varies between 3×10^{-2} cm^{-4} and 10^{0} cm^{-4} (Pruppacher and Klett 1978). Sudden change in the raindrop spectra, recognized as 'n_{0r} jump', were observed by Waldvogel (1974), when the precipitation changed from one type to another and even when the precipitation type remained the same. The changes of n_{0r} were attributed to the changes in the microphysical processes occurring in the cloud system (Pruppacher and Klett 1978).

In the LFO83 scheme, the term hail is used loosely to represent high density graupel, ice pellets, frozen rain and hailstones. According to Pruppacher and Klett (1978), the bulk density of hail has been found to vary between 0.7 and 0.9 g cm^{-3} and the observed density of graupel ranges from 0.05 g cm^{-3} to 0.89 g cm^{-3}. The term snow in the LFO83 scheme is used to represent snow crystals, snowflakes and low-density graupel particles. The bulk densities of snow particles ranges from 0.05 to 0.89 g cm^{-3} (Lin *et al.* 1983; Pruppacher and Klett 1978).

Based on the above observation studies of hydrometeor size spectrum, we designed several parameter retrieval experiments (Table 1). In this study, we still assume that the true values of those microphysical parameters, which are used in the CNTL experiment, do not change with space and time. No parameter estimation is performed for the CNTL experiment. In the parameter retrieval experiments, we picked the first guesses of one or some of the five parameters to be different from that in the CNTL experiment, but within the range of the observed values. For the three intercept parameters, we first choose their initial values to be commonly an order different from the true value, so that the sensitivity and the retrieval results can be compared directly among them. We picked the first guesses of the hail intercept and the hail density based on the sensitivity study of Gilmore *et al.* (2004). For snow and rain intercept parameters, we also chose other two first guesses (a larger and a smaller value than the 'true' value), which are closer to the 'true' value and are more often observed. We randomly picked a larger value of snow density.

As mentioned before, the parameter estimation with the EnKF is realized by considering the parameters as part of the model state. For each ensemble member, the perturbations of the parameter to be estimated is sampled from a normal prior distribution with mean equals to zero. Ideally, the initial ensemble spread, i.e. the standard deviation of the prior distribution, should represent the error of the first guess. Therefore, the standard deviation of the prior distribution is specified to be the initial error of each parameter. However, the initial error is usually unknown in reality. Therefore, we tested a narrower prior distribution with experiment HN6ρ9, in which the hail intercept of ensemble members are sampled from a prior distribution with the one standard deviation width equals to ½ of its initial error.

The uncertainty associated with the intercept parameters can be more than an order of magnitude. Initial sampling the parameter ensemble from a broad prior distribution can easily result in unphysical negative values. Any inaccuracy in the analysis can also result in negative intercept and density. Therefore, the five microphysical parameters are logarithmically transformed and multiplied by 10 before the analysis, and then transformed back during the forecast step.

At each analysis step, the covariances between the parameters and the observations are calculated and are used in the update equation of the EnSRF. For parameter retrieval purpose, only reflectivity data greater than 10 dBZ are used. These data are assimilated sequentially.

An important issue we confronted with the parameter estimation is the 'filter divergence'. The tendency of filter divergence is much more pronounced with parameter retrieval than model state retrieval because of two reasons. First, at each analysis step, 400 (the first cycle) to more than 4000 (the last cycle) reflectivity observations are used to update the parameters, while the data used to update the model variables at a certain grid point are limited through covariance localization. The parameter ensemble narrows quickly by the repeated use of the data. Another reason

causes the continuous narrowing of the parameter ensemble is that they remain constant during the forecast step, while errors and ensemble spread associated with model state variables grow during the forecast step.

To compensate the infinitely shrinking of the parameter ensemble, a minimum standard deviation is pre-specified, so that when the posterior standard deviation becomes smaller than the minimum value, the parameter ensemble spread is adjusted to restore the minimum. A similar variance inflation procedure was applied in Aksoy *et al.* (2005). However, we found in our case that a minimum standard deviation, as small as less than 5% of the initial standard deviation is required. The ensemble spread can decrease to 5% of its initial amount within the first two assimilation cycles, because a large amount of data is used to estimate the parameters. The sensitivity of the parameter retrieval to the variance inflation procedure will be discussed.

4 . RESULTS

4.1. Sensitivity of EnSRF analyses to microphysical parameters

Whether the chosen parameters can be retrieved from the observational data depends on whether the model output, in the forms of observations, is sensitive to these parameters. To test the sensitivity of the analyzed model state to the five parameters that we want to retrieve, we calculate the difference between the observations and the corresponding analysis projected to the observations:

$$d\eta = \sum_n (\eta^a - \eta^{obs})^2 \ , \qquad (8)$$

where η represents the radial velocity or reflectivity. The sum is over the data points with observed reflectivity greater than 10dBZ, which are the data used for parameter estimation. Without the superscript '*a*' in the equation, it is actually the observation term without the weighing coefficient of the cost function used in typical variational analysis.

The sensitivities to the five individual parameters are calculated from the output of five assimilation experiments, that are, respectively, HN3ρ9NE, SN3ρ1NE, RN8NE, HN4ρ4NE and SN03ρ4NE listed in Table 1. In the five experiments, the wrong initial guesses of the individual parameters are kept the same throughout the assimilation cycles and the parameter retrieval is not performed. The sensitivity measures how different the results of analysis are from the experiment with correct parameter values. The relative sensitivities of the analyzed reflectivity to the five parameters are plotted in Fig. 1. The relative sensitivity is defined as $d\eta$ normalized by that of experiment CNTL. For all cases, the first analysis is performed at 25 minutes of the truth storm and the analysis cycles end at 100 minutes.

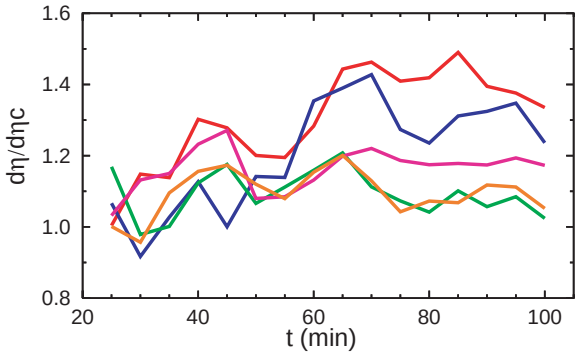

Fig. 1. The sensitivity of the analyzed reflectivity to the hail intercept n_{0h} (red, corresponding to HN3ρ9 but without parameter retrieval), snow intercept n_{0s} (blue, SN3ρ1), rain intercept n_{0r} (green, RN8), hail density ρ_h (pink, HN4ρ4) and snow density ρ_s (orange, SN03ρ4) from data assimilation experiments with wrong parameters and without parameter retrieval. dη defined by Eq. (8) is normalized by that of experiment CNTL.

It can be seen from Fig. 1 that among all five parameters, the analyzed reflectivity is most sensitive to n_{0h}, the intercept parameter of hail. The analyzed reflectivity is least sensitive rain intercept parameter, n_{0r}, and the sensitivity to snow intercept parameter n_{0s} is in-between those of hail and rain. Generally, the sensitivity increases with assimilation cycles for n_{0h} and n_{0s}. In early period of assimilation, except for the first analysis cycle, the retrieved reflectivity shows larger sensitivities to n_{0h} and the density of hail ρ_h but smaller sensitivities to n_{0s} and n_{0r}. At the later stage, the reflectivity shows larger sensitivities to n_{0h} and n_{0s} and is relatively insensitive to n_{0r} and ρ_s. The retrieved radial velocity (not shown) is much less sensitive to these five microphysical parameters than the retrieved microphysical fields are, although the sensitivity is still highest for n_{0h}. For this reason, that we will not use radial velocity data to retrieve (or update) these microphysical parameters. They are used, however, to update the state variables.

When an incorrect value of n_{0r}. or $\boldsymbol{\rho_h}$ is used in the model to perform the data assimilation, the ensemble mean forecast and analysis rms (root-mean squre) errors of all model state variables increase (blue curves in Fig. 6, not shown for ρ_h). The q_s is most sensitive to the change of hail pa-

rameters. Less q_s in the anvil and more q_h aloft are found in the ensemble mean analysis when larger hail intercept or smaller hail density is used in the model (Fig. 2b and 2d) and more q_s in the anvil and less q_h aloft are found in the ensemble mean analysis when smaller hail intercept is used (Fig. 2d). The sensitivity of the precipitation structure to the hail intercept and the hail density is similar to what was found in Gilmore *et al.* (2004). Larger hail intercept results in higher number concentrations, smaller mean particles sizes of hail and smaller terminal velocity and vise versa. The terminal velocity of hail can also be greatly reduced by decreasing the particle density. Increasing the number of smaller drops and reducing the terminal velocity of hail both result in enhanced upward fluxes of q_h, higher-altitude transport of q_h, longer residence time of q_h aloft, more collection of snow in the updraft region and lower amounts of snow transported to the anvil region. Larger upward flux of q_h also leads to smaller q_r and q_h at the low levels.

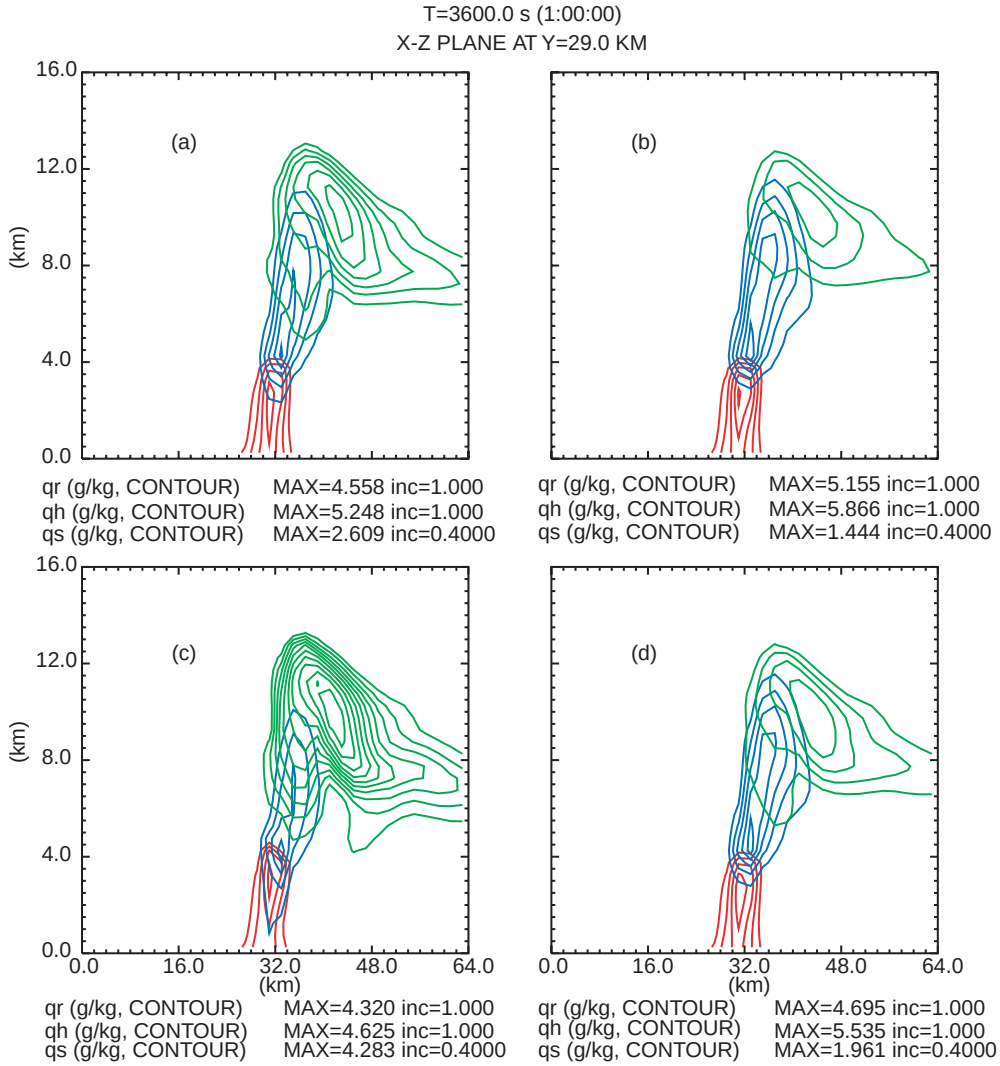

Fig. 2. Vertical cross section of q_r (red, with intervals of 1 g kg^{-1}), q_h (blue, with intervals of 1 g kg^{-1}) and q_s (green, with intervals of 0.2 g kg^{-1}) through the point of domain-wide maximum vertical velocity of the ensemble mean analysis at 60 min for (a) CNTL, (b) HN3ρ9NE, (c) HN6ρ9NE, and (d) HN4ρ4NE. The last three experiments use incorrect microphysical parameters without parameter estimation.

Snow, cloud ice and cloud water are the three species that are most sensitive to the changes in the intercept parameter and density of snow. Other model state variables, including q_r and q_h are relatively insensitive to these parameters (c.f., blue curves in Fig. 7). We can see from Fig. 3 that

the model storm has systematically more/less snow and less/more cloud water and ice (Fig. 3b and 3c) when larger/smaller snow intercept parameter is used. The snow distribution becomes more heavily weighted toward smaller particles and the bulk terminal velocity decreases when the intercept parameter of snow is increased. In the model, the accretional growth of snow through the interaction of snow with cloud water and ice are proportional to the snow intercept parameter. Therefore, the snow content increases by accreting more cloud water and ice when snow intercept parameter increases. With smaller terminal velocity due to larger n_{0s}, the centroid of q_s is located at higher altitudes (see Fig. 3b and 3c). The two accretion terms in the model are inversely proportional to the snow density, therefore, larger snow density leads to less snow content (Fig. 3d).

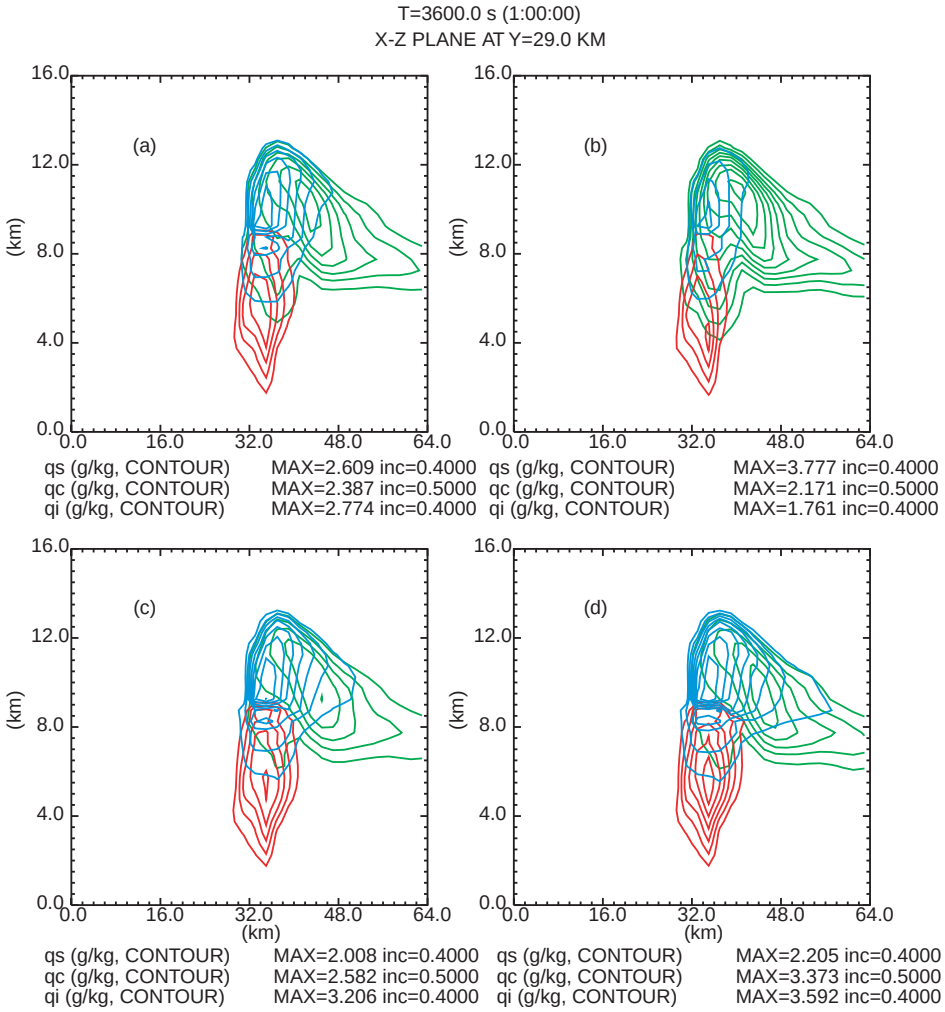

Fig. 3. Vertical cross section of q_c (red, with intervals of 0.5 g kg^{-1}), q_i (blue, with intervals of 0.4 g kg^{-1}) and q_s(green, with intervals of 0.4 g kg^{-1}) through the maximum vertical velocity of the ensemble mean analysis at 60 min for (a) CNTL, (b) SN3ρ1NE, (c) SN007ρ1NE and (d) SN03ρ4NE at 60 min. Panels (b) – (d) are for the experiments with incorrect parameters without parameter estimation.

Larger/smaller rain intercept parameter results more/less rain (compare Fig. 4 with Fig. 2a). Temperature and vertical velocity are more sensitive to the changes in the rain intercept parameter than to the changes in the snow parameters (blue curves of Fig. 7c and 7d, and Fig. 8c and 7d). Larger intercept parameter of rain results in more raindrops with smaller sizes, which contributes to stronger downdraft and low-level cold pool via evaporation.

4.2. Results of Experiments Retrieving Single Parameters

First, we performed ten experiments, in which the five parameters are initialized with wrong values and estimated individually. In each of these experiments, only the parameter that is to be estimated is perturbed among the ensemble members around its first guess value and updated using the reflectivity data; all other parameters are kept at their true values. Fig. 5 shows the time series of the ensemble mean (red curves) and the standard deviation (blue curves) of the estimated parameters alone with their true values (black lines) during the 80 minutes of assimilation. The values at 20 min indicate the prior distributions of these parameters.

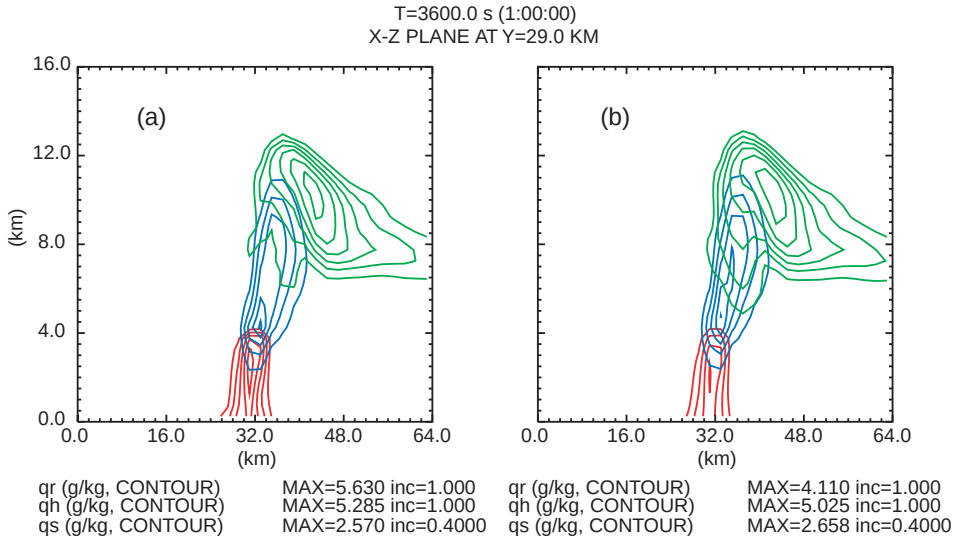

Fig. 4. As Fig. 2, but for experiments (a) RN8NE and (b) RN03NE, that have incorrect values of rain intercept parameter.

As can be seen from Fig. 5a, the estimated hail intercept of experiment HN3ρ9 reaches the true value at the end of the third assimilation cycle and then oscillates around the true value in the rest of the cycles. The estimated hail intercept approaches the true value slower in experiment HN6ρ9 and even more so when the initial prior distribution is narrower. For the hail intercept parameter retrieval experiments, the minimum standard deviation specified here is 0.35 in logarithmic units. The rate of convergence towards the truth for this parameter is comparable to those of the ensemble mean rms errors of most model state variables. The errors of the estimated hail intercept from these three experiments are around 0.0001 cm^{-4} when converted back to its original units, where the true hail intercept parameter is 0.0004 cm^{-4}. The estimated model state is generally insensitive to this amount of parameter error, as shown by Fig. 6 that the rms errors of most state variables (red curves) are almost indistinguishable from those of CNTL experiment (black curves) that uses true parameters. Ideally, the posterior ensemble spread of the parameter provides a good measure of the parameter error. Obviously, the error here is underestimated by the posterior ensemble spread. A larger pre-specified minimum spread will result in larger oscillations around the truth in the estimated parameter, which will be discussed in more detail later.

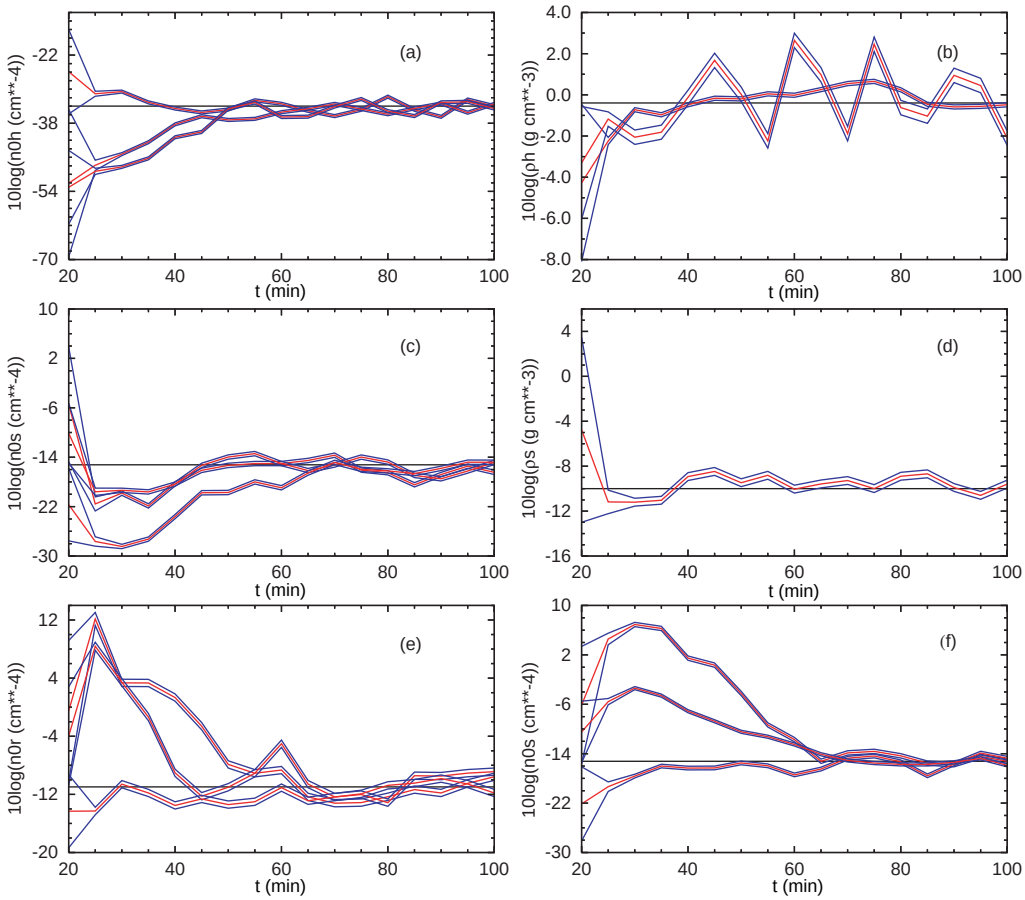

Fig. 5. Time series of ensemble mean (red curves) and standard deviation (blue curves) of the ensemble of estimated parameters from single-parameter retrieval experiments (a) 10log(n_{0h}) from HN3ρ9 (the one has the larger initial guess), HN6ρ9 (the one has the smaller initial guess) and HN6ρ9b (the same as HN6ρ9, but with smaller initial spread) (b) 10log(ρ_h) from HN4ρ4 and HN4ρ4b (the same as HN4ρ4, but with larger pre-specified minimum standard deviation) , (c) 10log(n_{0s}) from SN3ρ1 (the one has the largest initial guess), SN1ρ1 and SN007ρ1 (the one has the smallest initial guess), (d) 10log(ρ_s) from SN03ρ4, (e) 10log(n_{0r}) from RN8 (the one has the largest initial guess), RN3, RN03 (the one has the smallest initial guess) and (f) the same as (c) but with different realizations of the initial random perturbations to the parameter.. The true values of the parameters are shown as straight black lines,

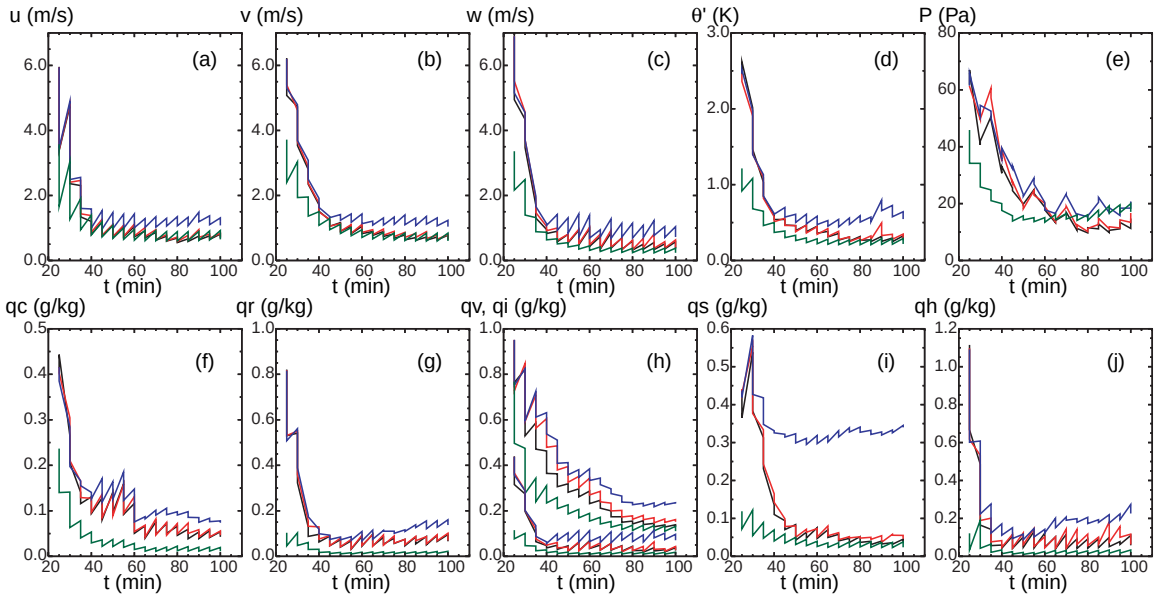

Fig. 6. The *rms* errors of ensemble-mean forecast and analysis, averaged over points at which the true reflectivity is greater than 10dBZ for (a) u, (b) v, (c) w, (d) θ', (e) p', (f) q_c, (g) q_r, (h) q_v (the curves with larger values) and q_i (the curves with lower values), (i) q_s and (j) q_h, for experiment CNTL (black), HN3ρ9 (red), HN3ρ9NE (blue) and the ensemble spread of CNTL (green). Units are shown in the plots. The drop of the error curves at specific times corresponds to the reduction of error by analysis.

We performed three parameter retrieval experiments with three different initial guess values for snow and rain intercept parameters. One of the experiments, SN3ρ1 or RN8, has an initial guess one order of magnitude different from the true value, so that it can be easily compared with the retrieval experiment for the hail intercept parameter HN3ρ9, but this value is rarely observed for snow or rain. The other two experiments, SN1ρ1 and SN007ρ1 for snow intercept parameter or RN3 and RN03 for rain intercept parameter, have their first guesses set to more realistic values. The pre-specified minimum standard deviation is 0.35 for the snow intercept and 0.5 for the rain intercept in logarithmic units. It can be seen from Fig. 5c, e that with the same magnitude of initial errors, the retrieved snow and rain intercept parameters converge more slowly than the hail intercept parameter. This is consistent with our earlier sensitivity analysis; initially, the retrieved model state is not very sensitive to n_{0s} and n_{0r} therefore the retrieval of them is difficult We can also see that the retrieved n_{0s} in all three experiments first arrives at values that are lower than their first guesses (Fig. 5c) while the retrieved value of n_{0r} first rise above their first guess values (Fig. 5e). Such behaviors are found to be related to the initial sampling errors. A repeat of the same set of experiments for n_{0s}, i.e., experiments SN3ρ1, SN1ρ1 and SN007ρ1, but with different realizations of the initial random perturbations to the parameter can result in different evolutions of the retrieved values, as seen in Fig. 5f; the retrieved parameter eventually converges to a value that is close to the truth in all cases, however.

The error of the retrieved n_{0s} is around 0.01 cm^{-4} after the estimation converges, while the true value is 0.03 cm^{-4}. The model state analysis is not sensitive to this amount of error, as can be seen in Fig. 7; the retrieved model variables are as good as those of CNTL starting from 45 min or the fifth cycle. The error of the retrieved n_{0r} is around 0.03 cm^{-3} for experiment RN8 while the truth value is 0.08 cm^{-3}. The errors of most retrieved state variables are comparable to those of CNTL starting from 70 min, except for q_v whose errors are still smaller than those when no parameter estimation is performed (Fig. 8). When the initial errors in the snow or rain intercept parameters are less than an order of magnitude, the impact of the parameter estimation on the model state retrieval is limited, because the analysis is relatively insensitive to small errors in these parameters. Even though all retrieved parameters are closer to their true values than their initial guesses, whether the retrieved model state is improved at the end of assimilation cycles depends on the rate of convergence of these retrieved parameters towards their true values. The convergence rate, on the other hand, is

found to be very sensitive to the realization of the initial random perturbations to the parameters (see, Fig. 5c and 5f).

Both retrieved hail and snow densities, ρ_h and ρ_s, approach their true values very quickly (Fig. 5b and 5d) in the single-parameter retrieval experiments. The pre-specified minimum spread for snow density is 0.3 in logarithmic unit. The error of the retrieved ρ_s after convergence is less than 0.015 g kg^{-1}, which is very low. The rms error of the retrieved model state is indistinguishable from that of CNTL (not shown). For ρ_h, we tried two minimum standard deviations, a smaller one of 0.1 and a larger one of 0.35 in logarithmic units. It can be seen from Fig. 5b that with the larger minimum, the retrieved ρ_h shows large oscillations around the true value. This kind of oscillations was also seen in the retrieval experiments for other parameters. The model state retrieval with the smaller minimum is also as good as the CNTL (not shown). With the larger minimum, the retrieved model state is still very good except for q_v, q_s and q_h (Fig. 9). A large increase in ensemble mean rms forecast error in q_h can be seen at 50, 65, 80 and 95 minutes (Fig. 9j), which is caused by the retrieved ρ_h being much larger than the truth at 45, 60, 75 and 90 minutes (Fig. 5b). Since our standard deviation restoration procedure for the parameters is somewhat artificial, the specification of the ensemble spread thresholds require some care. A larger threshold can result in significant oscillations of the retrieved parameter. Too smaller a threshold can lead to slow convergence. Often, numerical experiments are required to arrive at the most suitable thresholds.

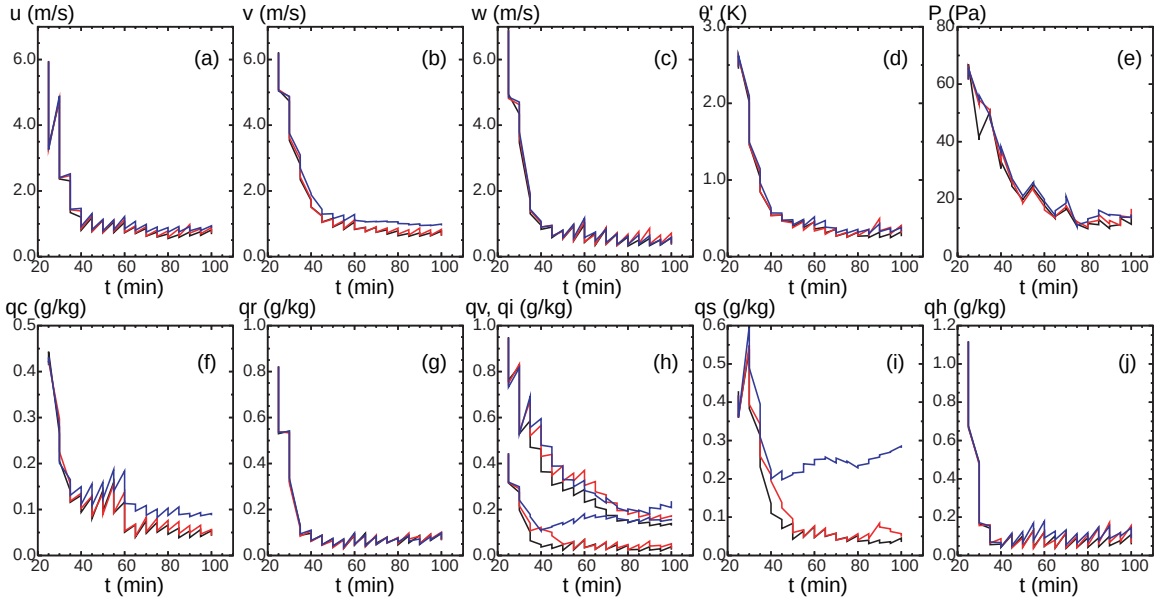

Fig. 7. As Fig. 6 but for experiments CNTL (black), SN3ρ1 (red) and SN3ρ1NE (blue). The ensemble spread curves for CNTL are not plotted.

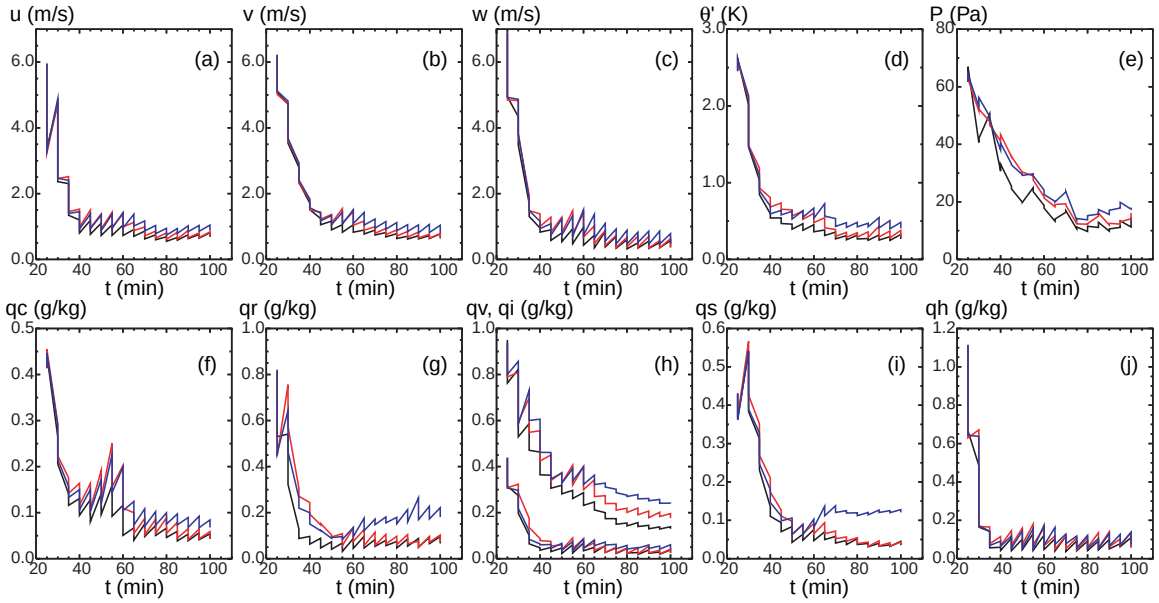

Fig. 8. As Fig. 7 but for experiments CNTL (black), RN8 (red) and RN8NE (blue).

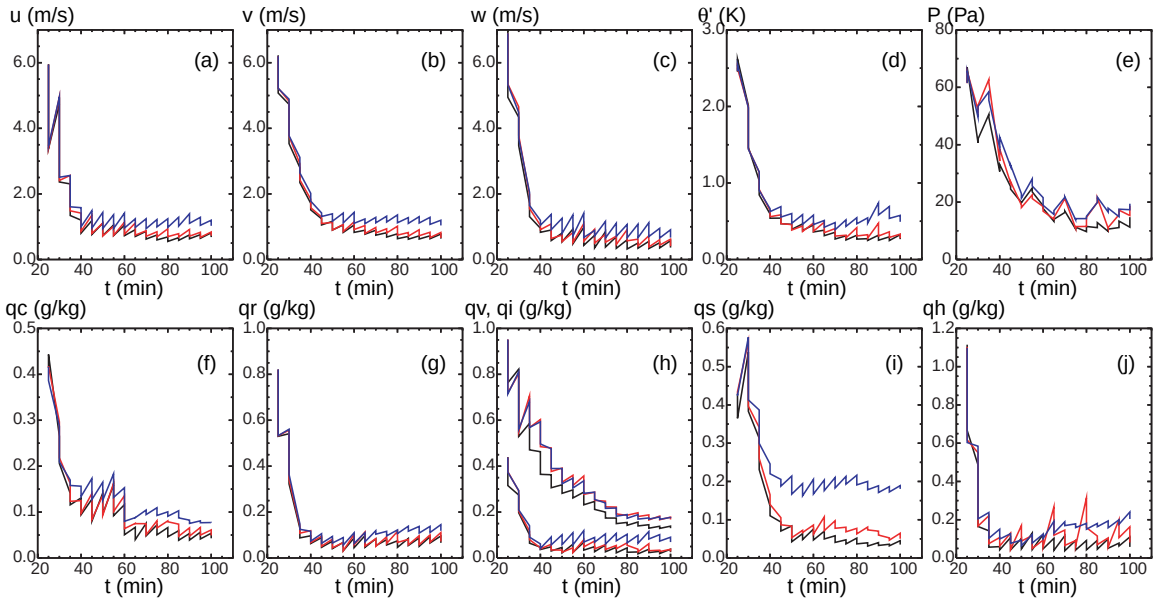

Fig. 9, As Fig. 7 but for experiments CNTL (black), HN4ρ4 (red) with pre-specified minimum standard deviation of 0.35 in logarithmic unit, and HN4ρ4NE (blue)

To understand how the microphysical parameters are retrieved from reflectivity data, i.e., to examine the issue of parameter identifiability (Navon 1997), we calculate the error correlation coefficients between each individual parameter and the forecast reflectivity. It is the cross correlation between the parameter and the forecast reflectivity that determines how the parameter can be adjusted from the reflectivity observations.

Significant and meaningful correlation coefficient can be found between the forecast reflectivity and n_{0h}, ρ_h, n_{0s} and ρ_s (Fig. 10). The color shades in Fig. 10a and c show the observed reflectivity at the first radar elevation. The echoes of the right cell are located 40 to 70 km away from the radar. Based on the radar geometry, the reflectivity at the first elevation level represents the precipitation between surface and 1.5 km AGL. The reflectivity at the seventh elevation level (Fig. 10b and 10d)

represents the precipitation between 4 and 9 km AGL. It can be seen from Fig. 10a and 10c that a negative/positive correlation coefficient between reflectivity and n_{0h}/ρ_h is found at low level heavy precipitation region. A positive/negative correlation coefficient center between forecast reflectivity and n_{0h}/ρ_h is located at the transient region between the updraft and anvil (Fig. 10b and 10d). This is consistent with the sensitivity of the analysis to these two parameters. Larger n_{0h} or smaller ρ_h results in less q_r and q_h at low levels, leading to higher reflectivity. The reflectivity is therefore negative correlated with n_{0h} and positively correlated with ρ_h. At the middle levels, larger n_{0h} or smaller q_h results in more q_h being transported to high altitudes, leading to higher reflectivity at the updraft-to-anvil transient region. The reflectivity in the transient region is therefore positively correlated with n_{0h} and negatively correlated with q_h. At the higher altitude anvil region, the reflectivity and n_{0h} are negatively correlated (Fig. 10b), because larger hail intercept parameter leads to smaller q_s and therefore smaller reflectivity there. Significant correlation coefficients between reflectivity and the snow intercept parameter and snow density are also found in the anvil region (Fig. 10e and 10f).

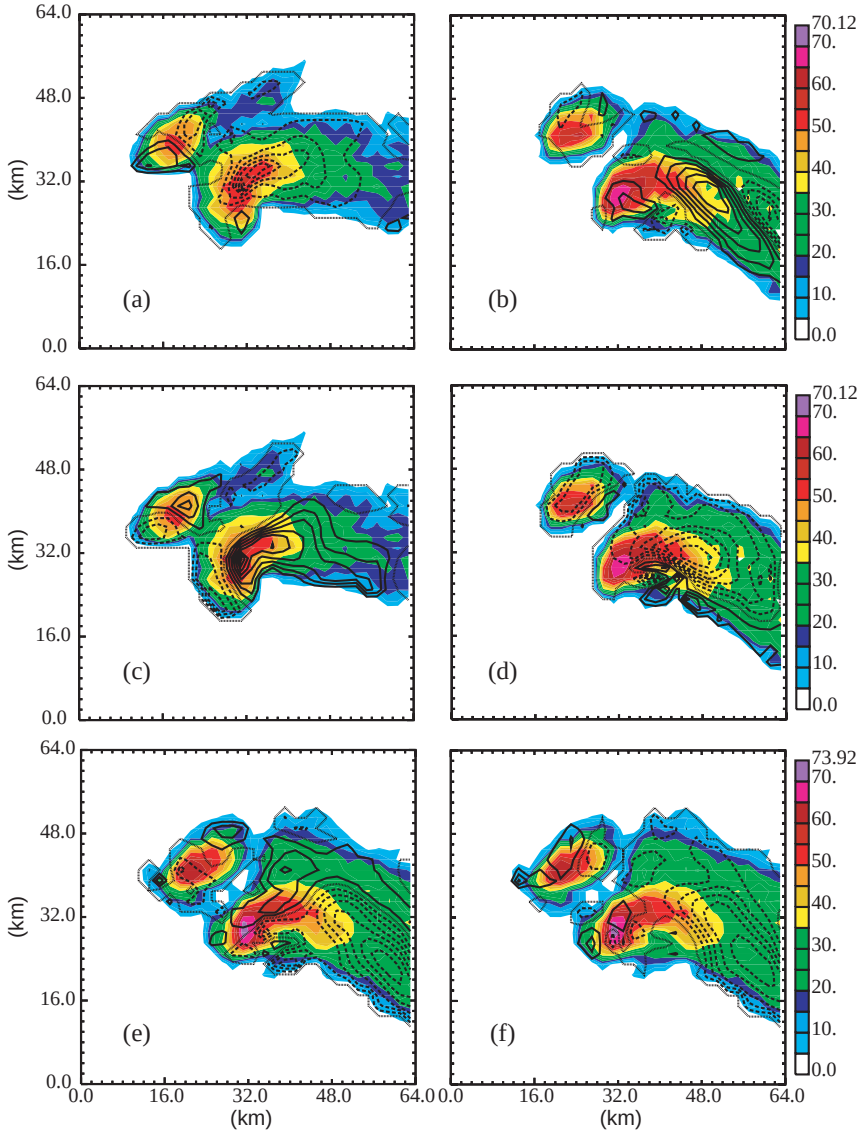

Fig. 10. Forecast error correlation coefficients at intervals of 0.1, between reflectivity and (a) n_{0h} on the 1st elevation level from HN3ρ9, (b) n_{0h} on the 7th elevation level from HN3ρ9, (c) ρ_h on the 1st elevation level from HN4ρ4, (d) ρ_h on the 7th elevation level from HN4ρ4, (e) n_{0s} on the 6th elevation from SN3ρ1, (f) ρ_s on the 6th elevation level from SN03ρ4, calculated from the ensemble at t = 70 min.

4.3. Results of Multiple Parameter Retrievals

It was shown in the last section that the five microphysical parameters can be retrieved reasonably well individually from the radar reflectivity data together with the model state variables. In this section, we further explore the ability of the EnKF in simultaneously retrieving more than one parameter. Four experiments are presented here for this purpose. First, the two parameters defining the hail size distribution (HN3ρ4) then the two parameters defining the snow size distribution (SN007ρ4) are estimated, in separate experiments. Then, the three intercept parameters are estimated simultaneously, starting from two different combinations of their first guesses

For the two-parameter case HN3ρ4 , the combination of the first guesses of the hail intercept and the hail density is for a storm with the presence of moderate-density graupel rather than high density hail (see Table 1). The 'true' storm is a typical mid-latitude continental storm producing mainly high density hail. It can be seen from Fig. 11a and 11b for experiment HN3ρ4 that the retrieved hail intercept parameter approaches the true value at the eighth assimilation cycle. The estimated hail density first deviates to values much higher than the true value but is gradually drawn back and reaches the true value during the last two assimilation cycles. The parameters approach their true values much slower than when they are retrieved individually. The retrieved model state is much better than the case without parameter estimation (see the red and blue curves in Fig. 12), but is not as good as that from CNTL or the corresponding experiments retrieving single parameters. Similarly, the intercept parameter and density of snow can also be simultaneously retrieved from reflectivity data (Fig. 11c and 11d) and the retrieved model state in this case is generally better than the case with incorrect parameters.

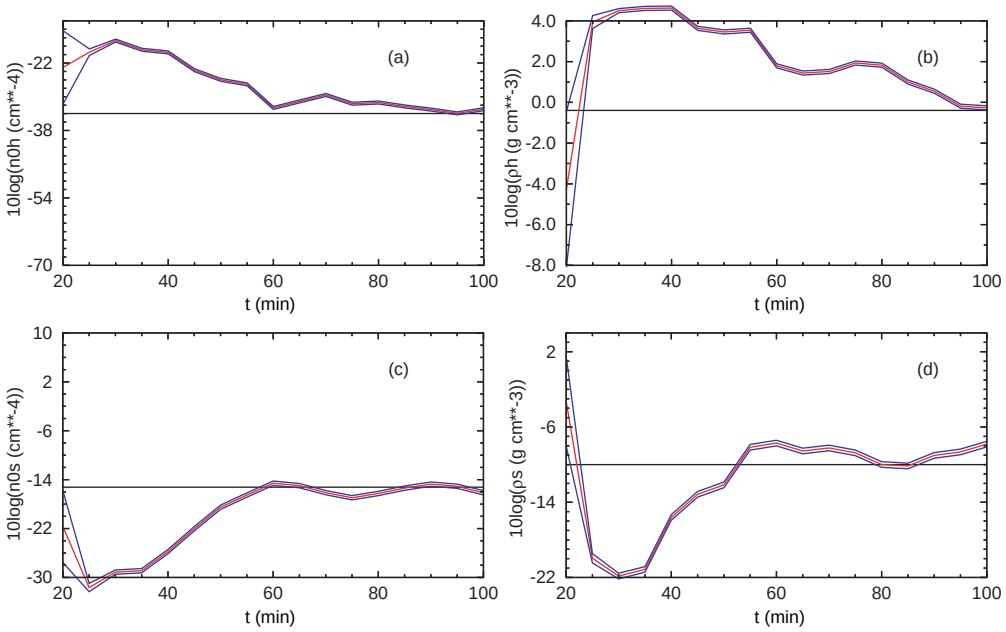

Fig. 11. As Fig. 5, but (upper panel) for experiment HN3ρ4 (a) 10log(n_{0h}) , (b)10log(ρ_h) and (lower panel) for experiment SN007ρ4 (c) 10log(n_{0s}) and (d)10log(ρ_s).

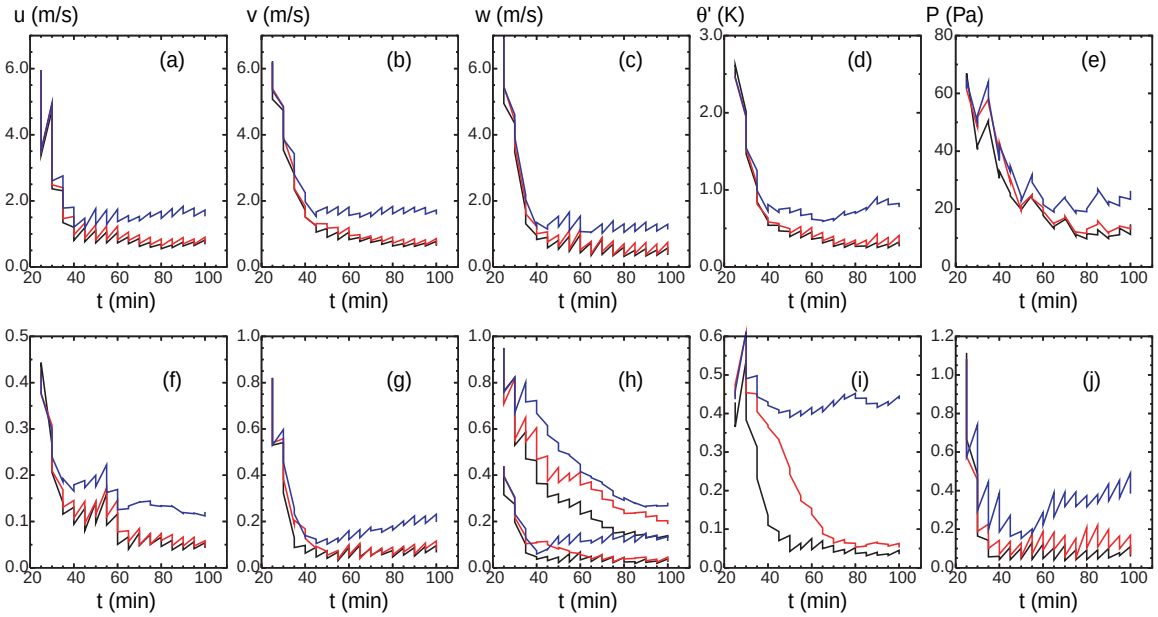

Fig. 12. As Fig. 6, but for experiments CNTL (black), HN3ρ4 (red) and HN3ρ4NE (blue).

In the first three-parameter retrieval experiment, the three intercept parameters start from first guesses that are larger than their true values (RSHNa, Table 1). The retrieval results are shown in Fig. 13. It can be seen that the intercept parameters of both hail and rain can be estimated correctly, but at a slower rate. The retrieval of the snow intercept parameter is, however, not successful. The structure of correlation coefficient between hail intercept parameter and forecast reflectivity is similar to that of single-parameter retrieval experiment HN3ρ9, but the maximum value of the correlation coefficient is smaller (not shown). No significant correlation is found for the intercept parameter of snow, explaining the failure of retrieval. Despite this problem, the retrieved model state is still improved over the case where none of the three intercept parameters are corrected via retrieval, but it is worse than that of CNTL.

In the other experiment that tries to retrieve the three intercept parameters (RSHNb, Table 1), the initial guess of the snow intercept parameter is set to 0.007 cm^{-4}, smaller than the true value of 0.03 cm^{-4}. It is otherwise the same as RSHNa. With this set of initial guesses, none of the three parameters can be correctly retrieved. Their values occasionally approach the level of true value and but stay away from it most of the other times (not shown). The reason for this failure will be investigated further.

5 . SUMMARY AND DISCUSSIONS

In this study, we first examined the impact of errors in several uncertain parameters in the ice microphysics scheme used in the ARPS model, on the retrieval of the model state for a simulated supercell thunderstorm. We then explored the possibility of correcting this type of model errors through simultaneous retrieval of these uncertain parameters as well as the model state using an ensemble square-root Kalman filter (EnSRF). Radar reflectivity data are used for the parameter retrievals. The microphysical parameters examined include the intercept parameters of assumed drop size distributions (DSDs) for rain, snow and hail, and the densities of hail and snow.

The EnSRF performs very well when retrieving a single error-containing parameter. The times for the parameters to approach their true values depend on the specific parameter to be retrieved. Generally, a parameter is easier to retrieve when the model output is more sensitive to the parameter. The analyzed model state is most sensitive to the parameters of hail DSD; hence the hail intercept parameter and hail density are the most retrievable among the five parameters. The analyzed model state, when the hail intercept parameter and density are properly retrieved, is as good as that of the experiment with perfect parameters. The snow density, and the snow and rain intercept parameters can also be successfully retrieved from reflectivity data, even though the model output is less sensitive to these parameters. Significant and

physically meaningful spatial error correlations were shown to exist between the retrieved microphysical parameters and the model reflectivity output, and such correlations play a key role in successful parameter estimation.

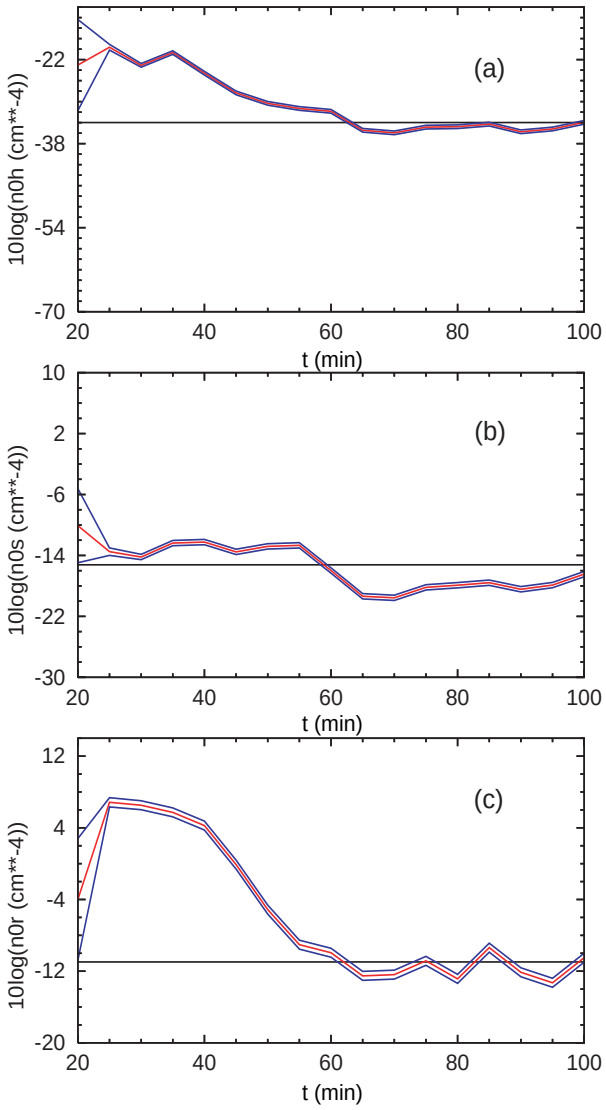

Fig. 13. As Fig. 6, but for experiment RSHNa. (a) 10log(n_{0h}), (b) 10log(n_{0s}), and (c) 10log(n_{0r}).

It is found that the parameter retrieval is very sensitive to the realization of initial perturbations sampled for the initial parameter ensemble. This is mainly caused by the limited ensemble size used here. The random variations in the initial sampling can directly affect the rate at which the parameter converges to the true value, and in some cases even the success or failure of the retrieval. The amount that the retrieved model state can be improved by the parameter estimation highly depends on the rate of convergence of the retrieved parameter, because when the convergence is slow, most of the analysis cycles are then performed with the wrong values of the parameters. This is especially important for the snow and the rain intercept parameters, because the realistic ranges of these two parameters are not as wide as that of hail intercept; even with the same magnitude of errors in the parameters, the model output is less sensitive to the intercept parameters of snow and rain DSDs.

It is also found that the results of parameter retrieval are very sensitive to the variance-inflation procedure applied to the retrieved parameters. The specification of both the width of initial prior distribution and the imposed minimum variance threshold can affect the convergence rate and the accuracy of the final parameter estimate. In our case, rather small minimum variance thresholds are needed and larger variances can cause large oscillations around the true value in the retrieved parameters. The need for parameter variance inflation is because that a huge number of reflectivity data, taken from the entire storm body, is used to estimate single or few parameters. The correlation information determines the direction to which the parameter should be adjusted, while the variance determines the amount of adjustment. If a major of the data provides the same direction of adjustment, the application of adjustments due to many data quickly diminishes the ensemble spread of the retrieved parameter. On the other hand, constantly inflating the variance can cause overadjustment, leading to large oscillations around the truth. A better, optimal, procedure is needed to make the evolutions of the error and ensemble spread consistent during the analysis. This will become more important when we start to deal with real data, where the truth is unknown.

The results of multiple parameter retrievals are generally not as good as those of single-parameter retrieval, not surprisingly because of the added errors in the initial guesses of the parameters. The two parameters defining the hail or the snow distribution, i.e., the intercept parameter and density, can be simultaneously retrieved sucessfully from the reflectivity data. The retrieval of the hail parameters may be especially important for storm-scale NWP and data assimilation, because hail/graupel can significantly influence the quantity and type of precipitation in many types of mid-latitude storms (Gilmore *et al.* 2004). When all three intercept parameters contain errors and are simultaneously estimated, the retrieval results are very sensitive to the initial guesses of the parameters.

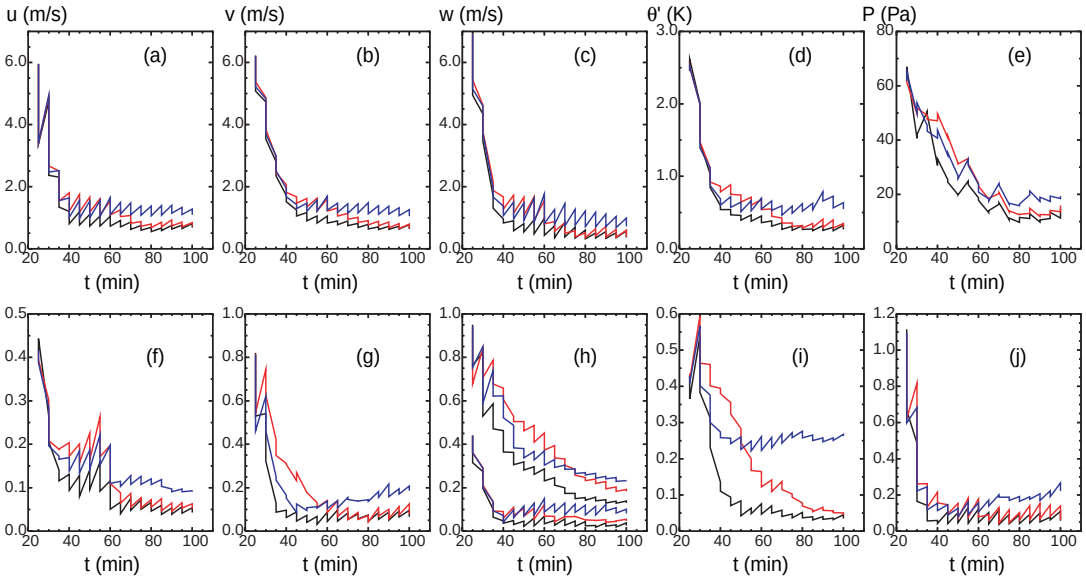

Fig. 14. As Fig. 6, but for experiments CNTL (black), RSHNa (red) and RSHNaNE (blue).

The sensitivities of the analysis and forecast to the microphysical parameters, and therefore, the identifiability of these parameters, are probably case dependent and may differ for different types of convective systems. In this study, we applied the parameter estimation to a supercell storm only. Some parameters may be more retrievable with other convective systems, such as the squall line systems that contain both vigorous convection and the stratiform precipitation regions. Other, such as the dual-polarization, data, may be very helpful for microphysical parameter retrieval. Work in assimilating polarimetric Dopple radar data using EnKF is in progress (Jung et al. 2005).

We should point out that currently we considered only the uncertainties in the microphysical parameters in the model. The five parameters retrieved are actually also involved in the observation operators of reflectivity. In this study, this involvement is not considered; true values of these parameters are used in the observational operators. Preliminary tests showed that the analysis is very sensitive to parameter errors in the observation operators because they directly affect what the model thinks the observational data are. We have not obtained any successful retrieval so far, based on a limited number of experiments, when these uncertainties with the observation operators are taken into account. Future work should investigate how to correct the errors in both prediction model and the observation operators.

Acknowledgement: This work was primarily supported by NSF grant ATM-0129892. Ming Xue was also supported by NSF grants ATM-0331594, ATM-0331756, EEC-0313747, DOT-FAA grant NA17RJ1227, and the "Outstanding Overseas Scholars" Award No. 2004-2-7 from Chinese Academy of Sciences. The computations were performed at the Pittsburgh Supercomputing Center supported by NSF.

REFERENCES

Aksoy, A., F. Zhang, and J. W. Nielsen-Gammon, 2005: Treatment of model error in the thermally driven circulations through ensemble-based simulaneous state and parameter estimation. *Mon. Wea. Rev.*, in review.

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

Annan, J. D. and J. C. Hargreaves, 2004: Efficient parameter estimation for a highly chaotic system. *Tellus*, **56A,** 520-526.

Annan, J. D., J. C. Hargreaves, N. R. Edwards, and R. Marsh, 2005a: Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter. *Ocean Modelling*, **8,** 135-154.

Annan, J. D., D. J. Lunt, J. C. Hargreaves, and P. J. Valdes, 2005b: Parameter estimation in an atmos-

pheric GCM using the ensemble Kalman filter. *Nonlinear Processes in Geophysics*, **12,** 363-371.
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.
Braham, R. R. J., 1990: Snow Particle Size Spectra in Lake Effect Snows. *Journal of Applied Meteorology*, **29,** 200-207.
Burgers, G., P. J. v. Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126,** 1719-1724.
Crook, N. A., D. Dowell, J. Sun, and Y. Zhang, 2004: Assimilation of radar observations of a supercell storm using 4DVar: Parameter retrieval experiments. *Preprints, 22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., CDROM 8A.2.
Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123,** 112-81145.
Derber, J. C., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev.*, **117,** 2347-2446.
Doviak, R. and D. Zrnic, 1993: *Doppler Radar and Weather Observations*. 2nd ed. Academic Press, 562 pp.
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. C. and L. J. Wicker, 2004: High-resolution analyses of the 8 May 2003 Oklahoma City storm. Part II: EnKF data assimilation and forecast experiments. *Preprints, 22nd Conf. on Severe Local Storms*, Hyannis, MA,, Amer. Meteor. Soc., CDROM, 12.5.
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., W. K. Tao, and J. Simpson, 1995: A double-moment multiple-phase four-class ice scheme. Part II: Simulations of convective storms in different large-scale environments and comparisons with other bulk parameterizations. *J. Atmos. Sci.*, **52,** 1001-1033.
Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125,** 723-757.
Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Monthly Weather Review*, **132,** 2610-2627.
Hansen, J. A., 2002: Accounting for Model Error in Ensemble-Based State Estimation and Forecasting. *Mon. Wea. Rev.*, **130,** 2373-2391.
Hao, Z., M. Ghil, and J. D. Neelin, 1995: Sequential parameter estimation for a coupled ocean-atmosphere model. *The Second International Symposium on Assimilation of Observations in Meteorology and Oceanography*, Tokyo, Japan, World Meteorological Organization, Geneva, Switzerland, 181-186.
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.
Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Monthly Weather Review*, **133,** 604-620.
Houze, R. A. J., P. V. Hobbs, P. H. Herzegh, and D. B. Parsons, 1979: Size distributions of precipitation particles in frontal clouds. *J. Atmos. Sci.*, **36,** 156-162.
Houze, R. A. J., P. V. Hobbs, D. B. Parsons, and P. H. Herzegh, 1980: Reply. *Journal of the Atmospheric Sciences*, **37,** 699-700.
Joss, J. and A. Waldvogel, 1969: Raindrop Size Distribution and Sampling Size Errors. *Journal of the Atmospheric Sciences*, **26,** 566-569.
Jung, Y., M. Xue, and J. M. Straka, 2005: Assimilation of polarimetric radar data using ensemble Kalman filter: Experiments with simulated data. *Extended abstract, 17th Conf. Num. Wea. Pred.*, Washington DC, Amer. Meteor. Soc., 13A.3.
Kivman, G. A., 2003: Sequential parameter estimation for stochastic systems. *Nonlinear Processes in Geophysics*, 253-259.
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.
Lo, K. K. and R. E. P. Jr., 1982: The growth of snow in winter storms: An airborne observational study. *J. Atmos. Sci.*, **39,** 697-706.
Marshall, J. S. and W. M. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5,** 165-166.
McCumber, M., W.-K. Tao, and J. Simpson, 1991: Comparison of ice-phase microphysical parameterization schemes using numerical simulations of tropical convection. *J. Appl. Meteor.*, **30,** 985-1004.

Mitchell, D. L., 1988: Evolution of snow-size spectra in cyclonic storms. Part I: Snow growth by vapor deposition and aggregation. *J. Atmos. Sci.*, **45,** 3431-3451.

Navon, I. M., 1997: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dynamics of Atmospheres and Oceans*, **27,** 55-79.

Passarelli, R. E. J., 1978: Theoretical and Observational Study of Snow-Size Spectra and Snowflake Aggregation Efficiencies. *Journal of the Atmospheric Sciences*, **35,** 882-889.

Pruppacher, H. R. and J. D. Klett, 1978: *Microphysics of Clouds and Precipitation*. D. Reidel, 714 pp.

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.

Rutledge, S. A. and P. V. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. Part VIII: A model for the feeder-seeder process in warm frontal rainbands. *J. Atmos. Sci.*, **40,** 1185-1206.

——, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. Part Xll: A diagnostic modeling study of precipitation development in narrow cold-frontal rainbands. *J. Atmos. Sci.*, **41,** 2949-2972.

Sekhon, R. S. and R. C. Srivastava, 1971: Doppler Radar Observations of Drop-Size Distributions in a Thunderstorm. *Journal of the Atmospheric Sciences*, **28,** 983-994.

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.

Srivastava, R. C., 1971: Size distribution of raindrops generated by their breakup and coalescence. *J. Atmos. Sci.*, **28,** 410-415.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Monthly Weather Review*, **131,** 1485-1490.

Tong, M. and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS Experiments. *Mon. Wea. Rev.***,** 1789-1807.

Waldvogel, A., 1974: The N0-jump of raindrop spectra. *J. Atmos. Sci.*, **31,** 1067-1078.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130,** 1913-1924.

Wood, V. T. and R. A. Brown, 1997: Effects of radar sampling on single-Doppler velocity signatures of mesocyclones and tornadoes. *Wea. Forecast.*, **12,** 928-938.

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., M. Tong, and K. K. Droegemeier, 2005: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Ocean Tech.***,** Accepted.

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. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76,** 143-166.

Yu, L. and J. J. O'Brien, 1991: Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile. *J. Phys. Oceanogr.*, **21,** 709-719.

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.

Zou, X., I. M. Navon, and F. X. Le Dimet, 1992: An optimal nudging data assimilation scheme using parameter estimation. *Quart. J. Roy. Meteor. Soc.*, **118,** 1163-1186.