RMetS
Royal Meteorological Society

An iterative ensemble square root filter and tests with simulated radar data for storm-scale data assimilation

Shizhang Wang,[a,b] Ming Xue,[b,c]* Alexander D. Schenkman[b,c] and Jinzhong Min[a]

[a]*Jiangsu Key Laboratory of Meteorological Disaster, Nanjing University of Information Science and Technology, China*
[b]*Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK, USA*
[c]*School of Meteorology, University of Oklahoma, Norman, OK, USA*

*Correspondence to: M. Xue, Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David L. Boren Blvd, Norman, OK 73072, USA. E-mail: mxue@ou.edu

An iterative procedure is designed to accelerate the 'spin-up' of ensemble square-root filter (EnSRF) data-assimilation cycles when starting from a poor initial ensemble. Referred to as the iterative EnSRF (iEnSRF), this iterative procedure follows the 'running in place' (RIP) concept developed for the local ensemble transform Kalman filter (LETKF) but because of algorithm differences is implemented differently. The iEnSRF is a three-step procedure: first, a backward EnSRF analysis is performed that updates the ensemble model states at an earlier time. Second, an ensemble of forecasts is run from these updated model states to the analysis time. These two steps are then repeated a prespecified number of times. The backward analysis is performed via an asynchronous ensemble Kalman filter (EnKF), which is capable of assimilating observations collected at times different than the analysis time. Like RIP, the iEnSRF uses the same observations repeatedly during the initial assimilation cycles, allowing for the extraction of additional information from observations when estimated ensemble mean state and ensemble covariance are poor. The iEnSRF algorithm is tested using simulated radar data for an idealized supercell storm. In experiments with a perfect model and the correct storm environment, as well as in the presence of model and environmental errors, the iEnSRF reduces the analysis error in the first few cycles more quickly than the regular EnSRF, leading to improved subsequent short-range forecasts. After the first few analysis cycles, continued use of iterations does not lead to further improvement. The better performance of the iEnSRF appears to be the result of improved background error covariance estimation as well as improved state estimation in the first few cycles, especially for correlations between observed and unobserved variables. Through iterations, the iEnSRF is also able to reach a steady level of state estimation error more quickly than the corresponding non-iterated version.

Key Words: radar data assimilation; iterative procedure; ensemble Kalman filter

Received 11 January 2012; Revised 25 September 2012; Accepted 17 October 2012; Published online in Wiley Online Library 4 January 2013

Citation: Wang S, Xue M, Schenkman AD, Min J. 2013. An iterative ensemble square root filter and tests with simulated radar data for storm-scale data assimilation. *Q. J. R. Meteorol. Soc.* **139**: 1888–1903. DOI:10.1002/qj.2077

1. Introduction

The ability of an ensemble Kalman filter (EnKF) (Evensen, 1994, 2003) to initialize convective storms from simulated Doppler radar radial velocity data was first demonstrated by Snyder and Zhang (2003) and Zhang *et al.* (2004) within an anelastic cloud model with simplified warm rain microphysics. Subsequent studies by Tong and Xue (2005) and Xue *et al.* (2006) demonstrated the capability of an EnKF to assimilate radar radial velocity as well as reflectivity data into a compressible model with complex ice microphysics. More recent studies have found that the ability to handle the complex nonlinear physics important for convective storms is an advantage that the EnKF method has over other data assimilation procedures (Jung *et al.*, 2008; Xue *et al.*, 2010; Snook *et al.*, 2011).

The EnKF is based on Monte Carlo sampling in which an ensemble is used to sample the model state uncertainty and to evolve the model state error covariance (Evensen, 1994). An optimal analysis can be obtained only if two independent requirements are satisfied: (i) the ensemble mean is close enough to the truth and (ii) the ensemble perturbations are representative of the true error characteristics (Kalnay and Yang, 2010).

For convective storms, Doppler weather radar is generally the only observation platform capable of providing detailed observations of the wind and precipitation structures within the storms. Due to sensitivity limitations, Doppler radar data are usually available only after precipitation-sized particles form, posing a challenge for the initialization of rapidly developing storms, especially for cycled data-assimilation (DA) methods that require the assimilation of many volume scans of radar data before a reasonably accurate estimate of the state of the atmosphere can be obtained.

The EnKF DA cycles for convective storms typically start from a first-guess field that contains information only about the broader-scale storm environment. Observing system simulation experiment (OSSE) studies have shown that it typically takes more than 10 assimilation cycles to obtain accurate analyses of convective storms when assimilating Doppler radar data via an EnKF (Caya *et al.*, 2005; Xue *et al.*, 2006; Yussouf and Stensrud, 2010). Caya *et al.* (2005) showed that because of the iterative nature of a four-dimensional variational (4D-Var) method, 4D-Var is able to establish a reasonably accurate analysis of a convective storm by assimilating only one to two radar volume scans. EnKF takes substantially longer to achieve this level accuracy, but is able to a produce a more accurate analysis than 4D-Var after several additional assimilation cycles. With the standard precipitation-mode volume scan interval of about 5 min for the operational WSR-88D weather radars of the U.S., Xue *et al.* (2006) showed that it took about 50 min for an EnKF to establish a well-defined storm, which was confirmed by Yussouf and Stensrud (2010). However, given a rapidly developing severe storm, as noted in Kalnay and Yang (2010), such a spin up time is likely to result in the EnKF analysis being less useful for severe storm forecast than 4D-Var (Caya *et al.*, 2005).

The desire to accelerate the 'convergence' (i.e., reduce the 'spin up time') of EnKF DA cycles, so that reasonably accurate state and covariance estimations can be established more rapidly, motivated the development of an iterative procedure called 'running in place (RIP)' by Kalnay and Yang (2010), in a local ensemble transform Kalman filter (LETKF) (Hunt *et al.*, 2007) framework. RIP is a three-step procedure: (i) a 'no-cost' ensemble Kalman smoother (EnKS) (Kalnay *et al.*, 2007; Yang *et al.*, 2009) is used to go backward in time within an assimilation cycle to improve the background state at the previous time; (ii) an advance with the standard LETKF procedure from the updated background; and (iii) a repeat of steps (i) and (ii) until a desired state is reached. Because it is based on the LETKF algorithm, the 'no-cost' EnKS uses LETKF weights obtained with observations at the later time to update ensemble states at the earlier time, hence improving both the state and covariance at the earlier time and allowing for the repetitive use of the observations within an assimilation cycle. Kalnay and Yang (2010) found that within an idealized global primitive equation model the RIP procedure improves the quality of the initial ensemble mean and background error covariance, so that fewer cycles are required for the filter to approach its optimal analysis. This result is especially true when initial ensemble perturbations are Gaussian noise and/or when the initial ensemble mean is far from the truth. The latter situation is often realized in thunderstorm initialization. Yang *et al.* (2012a) further studied the impact of RIP on nonlinearity and non-Gaussian noise with the Lorenz 63 model (Lorenz, 1963) and pointed out that RIP can significantly improve the analysis accuracy even in the presence of high nonlinearity.

Most recently, RIP has been applied to a typhoon case within OSSEs (Yang *et al.*, 2012b) based on the Weather Research and Forecast (WRF) model with an LETKF assimilation system, and RIP was able to accelerate the spin-up of dynamic structures of a typhoon. However, for radar, radial velocity and reflectivity observations are not model state variables. The impact of RIP with indirect observations or with storm-scale DA has yet to be examined. Testing the RIP idea for storm-scale radar DA is the main purpose of this article.

To date, almost all storm-scale radar DA studies have used either the ensemble square root filter (EnSRF) (Whitaker and Hamill, 2002) algorithm or the original EnKF algorithm with perturbed observations (Tong and Xue, 2005); they are serial algorithms that process observations one at a time, whereas the LETKF algorithm when RIP is applied is an algorithm that processes all observations (within the local domain) simultaneously. To apply the RIP concept to the serial EnSRF algorithm, we have developed a new procedure called iterative EnSRF (iEnSRF).

As an initial evaluation of the iEnSRF, we perform a set of OSSEs with simulated radar observations for a supercell storm. The OSSEs are conducted first for a perfect model experiment with no storm environment error and then in the presence of prediction-model and storm-environment errors. Environment error is generated by adding perturbations to the environmental sounding and model error is simulated by using different physical parametrization schemes in the ensemble forecasts. The benefits of the iEnSRF as compared to the regular EnSRF are assessed for these different situations.

The remainder of this article is organized as follows. In section 2, the iEnSRF algorithm is described and contrasted with RIP. Section 3 provides the configurations of the prediction model and data assimilation, and describes the simulation of radar observations and the design of OSSEs. The results of the OSSEs are discussed in section 4 and a summary is given in section 5.

2. The EnSRF and iEnSRF algorithms

2.1. The EnSRF algorithm

We first give a brief summary of the standard EnSRF algorithm on which the iEnSRF is based. Following Whitaker and Hamill (2002), the serial EnSRF algorithm analyzes uncorrelated observations one after another. Therefore, the observation error covariance $\mathbf{R}$ and background error covariance mapped to the observation space, $\mathbf{HP}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}$, reduce to scalars. The analysis equations for updating the ensemble mean model state $\overline{\mathbf{x}}$ and the ith ensemble deviation from the mean, $\mathbf{x}'_i$, are,

$$\overline{\mathbf{x}}^{\mathrm{a}} = \overline{\mathbf{x}}^{\mathrm{b}} + \mathbf{K}(y_j^{\mathrm{o}} - \overline{H(\mathbf{x}^{\mathrm{b}})}), \tag{1}$$

$$\mathbf{x}'^{\mathrm{a}}_i = \mathbf{x}'^{\mathrm{b}}_i - \alpha\mathbf{K}[H(\mathbf{x}_i^{\mathrm{b}}) - \overline{H(\mathbf{x}^{\mathrm{b}})}], \tag{2}$$

where

$$\mathbf{K} = \mathbf{P}^{\mathrm{b}}\mathbf{P}^{\mathrm{T}}(\mathbf{HP}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{3}$$

is the Kalman gain matrix, $\mathbf{P}^{\mathrm{b}}$ is the background error covariance, H represents the observation operator mapping variables from model state space to observation space, and $\mathbf{H}$ is a linearized version of H. Here, superscripts a, b, and o denote the analysis, background and observation, respectively. Subscripts i and j denote the ith ensemble member and the jth observation, respectively. In Eq. (2),

$$\alpha = [1 + \sqrt{\mathbf{R}(\mathbf{HP}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}}]^{-1} \tag{4}$$

is a coefficient derived by Whitaker and Hamill (2002) for the EnSRF algorithm. Equation (4) is valid only for single observation analysis and therefore both the numerator and denominator inside the square root are scalars. The background error covariance terms in $\mathbf{K}$ are estimated using the ensemble members, according to

$$\mathbf{P}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_i^{\mathrm{b}} - \overline{\mathbf{x}}^{\mathrm{b}})[H(\mathbf{x}_i^{\mathrm{b}}) - \overline{H(\mathbf{x}^{\mathrm{b}})}]^{T}, \tag{5}$$

$$\mathbf{HP}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[H(\mathbf{x}_i^{\mathrm{b}} - \overline{H(\mathbf{x}^{\mathrm{b}})}][H(\mathbf{x}_i^{\mathrm{b}} - \overline{H(\mathbf{x}^{\mathrm{b}})}]^{\mathrm{T}}, \tag{6}$$

where N is the ensemble size. For each observation, $\mathbf{P}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}$ is a vector and $\mathbf{HP}^{\mathrm{b}}\mathbf{H}^{\mathrm{T}}$ is a scalar. Observations are analyzed by using Eq. (1) and Eq. (2) sequentially, one after the other. After all observations at a given time are analyzed, an ensemble of forecasts proceeds from the analysis ensemble until new observations are available; the analysis cycles are then repeated.

2.2. The iEnSRF algorithm

As mentioned earlier, our iterative EnSRF procedure follows the RIP concept and also involves three steps in each analysis cycle, but its implementation differs from that of RIP due to differences between the LETKF and EnSRF algorithms on which they are based. This is particularly true for the first step where an earlier state is updated using later observations. This subsection describes the details of the iEnSRF while pointing out its differences to RIP.

In LETKF, the final ensemble analyses are obtained as linear combinations of the background ensemble forecasts, using transformation weights determined by the LETKF algorithm (Hunt *et al.*, 2007). In the first step of RIP (Kalnay and Yang, 2010), the 'weights' obtained at the current time, t_n, are used to transform the ensemble forecasts at a previous time, t_{n-1}, into updated states or ensemble analyses valid at t_{n-1}. This procedure is referred to as 'no-cost' EnKS (Kalnay *et al.*, 2007; Yang *et al.*, 2009). An EnSRF has no such weights, at least not explicitly, thus requiring a different procedure to update the previous time.

Specifically, in the EnSRF, an asynchronous implementation can be done so that observations at the current time can be used to update states at an earlier time. An algorithm to implement such asynchronicity was developed by Sakov *et al.* (2010) and can be applied to different variants of EnKF. However, due to the serial nature of the EnSRF, the implementation of the asynchronous algorithm has additional complications. In a typical asynchronous algorithm, observation priors computed at t_n, the time of observations, are used to update the model state at t_{n-1}. This usually requires the precalculation and simultaneous use of all observation priors, as described in Sakov *et al.* (2010); this is not possible for the serial EnSRF.

To address this complication, we treat the state vector $\mathbf{X}$ at t_{n-1} as an extension of the state vector at t_n and form an extended stated vector $(\mathbf{x}_{t_n}^{\mathrm{T}}, \mathbf{x}_{t_{n-1}}^{\mathrm{T}})^{\mathrm{T}}$. This extended state vector is treated like a regular vector in the EnSRF, and both state vectors are updated by the filter simultaneously. It is then straightforward to apply the asynchronous algorithm. This is similar to the joint state-observation vector used by Anderson (2001). Through this new vector, the update of observation priors at t_n follows the standard EnSRF algorithm and it is possible to analyze the model state at t_{n-1} from observations available at t_n. Here, the update of state at t_{n-1} by data available at t_n is achieved through asynchronous covariance, which involves model state samples distributed in time. The asynchronous covariance for updating the model state at t_{n-1} using observations t_n are

$$\mathbf{P}^{\mathrm{b}}_{t_{n-1}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}^{\mathrm{b}}_{i,t_{n-1}} - \overline{\mathbf{x}}^{\mathrm{b}}_{t_{n-1}})[H(\mathbf{x}^{\mathrm{b}}_{i,t_n}) - \overline{H(\mathbf{x}^{\mathrm{b}}_{t_n})}]^{\mathrm{T}}, \tag{7}$$

$$\mathbf{HP}^{\mathrm{b}}_{t_n}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}[H(\mathbf{x}^{\mathrm{b}}_{i,t_n}) - \overline{H(\mathbf{x}^{\mathrm{b}}_{t_n})}][H(\mathbf{x}^{\mathrm{b}}_{i,t_n}) - \overline{H(\mathbf{x}^{\mathrm{b}}_{t_n})}]^{\mathrm{T}}, \tag{8}$$

where $\mathbf{P}^{\mathrm{b}}_{t_{n-1}}\mathbf{H}^{\mathrm{T}}$ is the covariance between model states at t_{n-1} and observation priors at t_n, $\mathbf{HP}^{\mathrm{b}}_{t_n}\mathbf{H}^{\mathrm{T}}$ is the same as that in Eq. (6) because the observations used to update $\mathbf{x}_{t_{n-1}}$ are valid at t_n. Therefore, the Kalman gain matrix $\mathbf{K}_{t_{n-1}}$ for updating model state $\mathbf{x}_{t_{n-1}}$ is calculated according to

$$\mathbf{K}_{t_{n-1}} = \mathbf{P}^{\mathrm{b}}_{t_{n-1}}\mathbf{H}^{\mathrm{T}}(\mathbf{HP}^{\mathrm{b}}_{t_n}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}. \tag{9}$$

Then, the analysis equations for ensemble mean ($\overline{x}_{t_{n-1}}$) and ensemble deviations ($\mathbf{x}'_{i,t_{n-1}}$) are, respectively,

$$\overline{\mathbf{x}}^{\mathrm{a}}_{t_{n-1}} = \overline{\mathbf{x}}^{\mathrm{b}}_{t_{n-1}} + \mathbf{K}_{t_{n-1}}(y_j^{\mathrm{o}} - \overline{H(\mathbf{x}^{\mathrm{b}}_{t_n})}), \tag{10}$$

$$\mathbf{x}'^{\mathrm{a}}_{i,t_{n-1}} = \mathbf{x}'^{\mathrm{b}}_{i,t_{n-1}} - \alpha\mathbf{K}_{t_{n-1}}(H(\mathbf{x}^{\mathrm{b}}_{i,t_n}) - \overline{H(\mathbf{x}^{\mathrm{b}}_{t_n})}), \tag{11}$$

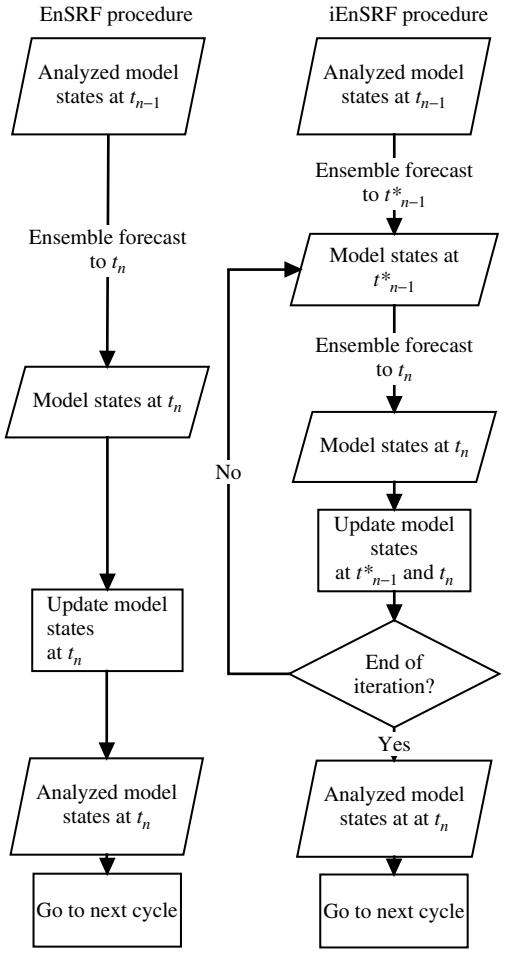

Figure 1. The flow chart for the EnSRF and iEnSRF procedures in each cycle, where t^*_{n-1} is an arbitrarily intermediate time between t_{n-1} and t_n.

Through Eqs (10) and (11), the model states at t_{n-1} can be updated by observations taken at t_n. Because Eqs (10) and (11) are similar to Eqs (1) and (2), except for the Kalman gain, it is easy to implement within an existing EnSRF framework. Once the ensemble states at t_{n-1} are updated, ensemble forecasts from t_{n-1} are produced to arrive at new forecast states at t_n and such new forecast states are used in the next iteration of asynchronous filter updating. The iterations are repeated in a similar manner as in Kalnay and Yang (2010). Because of the use of an updated state at t_{n-1}, the updated forecast states at t_n should be improved over that of the earlier iteration.

Our iEnSRF procedure is illustrated in Figure 1. It should be pointed out that, as part of an extended state, $\mathbf{x}_{i,t_n}$ and $\mathbf{x}_{i,t_{n-1}}$ must be simultaneously updated (note that in the figure, model states at t^*_{n-1} instead of t_{n-1} are updated, this will be discussed later in this section). It is noted that only the updated $\mathbf{x}_{i,t_n}$ in the final iteration are carried into the next assimilation cycle, while those in the intermediate iterations are needed only for calculating covariance for the next observation. Obviously, this is computationally inefficient. A possible solution to this problem follows the scalable implementation of an ensemble filter proposed by Anderson and Collins (2007), which precalculates and updates observation priors as part of an extended state. By updating the observation priors at t_n instead of the full state, the number of calculations can be greatly reduced; $\mathbf{x}_{i,t_n}$ will then need to be updated only in the final iteration. This approach is equivalent to state updating only when the observation operator is linear. In practice, differences due to such an approximation are probably no larger than other sources of uncertainty within a filter implementation (Anderson and Collins, 2007). Although this approach is attractive, computational efficiency is not the focus in this proof of concept article, and we chose to keep the implementation simple based on an existing EnSRF code.

Asynchronous algorithms are subject to temporal sampling error, in addition to spatial sampling error. Similar to the spatial sampling error, temporal sampling error can cause erroneous correlations when the time interval between samples is long. A temporal localization is introduced to help minimize such effects. In this study, the fifth-order correlation function (Gaspari and Cohn, 1999), typically used for spatial localization, is employed for temporal localization. Spatial and temporal localization are applied simultaneously to all state variables. With the spatiotemporal localization, analysis equations for updating ensemble mean $\overline{\mathbf{x}}_{t_{n-1}}$ and ensemble deviations $\mathbf{x}'_{i,t_{n-1}}$ are modified from those in (10) and (11),

$$\overline{\mathbf{x}}^{\mathrm{a}}_{t_{n-1}} = \overline{\mathbf{x}}^{\mathrm{b}}_{t_{n-1}} + (\rho_{tl}\boldsymbol{\rho}_{sl}) \circ \mathbf{K}_{t_{n-1}}(y^{\mathrm{o}}_j - \overline{H(\mathbf{x}^{\mathrm{b}}_{t_n})}), \tag{12}$$

$$\mathbf{x}'^{\mathrm{a}}_{i,t_{n-1}} = \mathbf{x}'^{\mathrm{b}}_{i,t_{n-1}} - \alpha(\rho_{tl}\rho_{sl}) \circ \mathbf{K}_{t_{n-1}}(H(\mathbf{x}^{\mathrm{b}}_{i,t_n}) - \overline{H(\mathbf{x}^{\mathrm{b}}_{t_n})}), \tag{13}$$

where scalar ρ_{tl} is the temporal localization coefficient; vector $\boldsymbol{\rho}_{sl}$ contains the spatial localization coefficients and '$\circ$' represents the Schur (element-wise) product (Houtekamer and Mitchell, 2001). Scalar ρ_{tl} is a function of the time interval between t_{n-1} and t_n while for vector $\boldsymbol{\rho}_{sl}$ the distance between y^{o}_j and model grid points can vary.

Another important consideration in our iterative procedure is how to determine the time interval of iteration between t_{n-1} and t_n. According to Kalnay and Yang (2010), RIP updates the state at t_{n-1} even though observations are available at t_n. They justified such a procedure on the grounds that the EnKS updated ensemble at t_{n-1} is more accurate than the analysis ensemble using data only at t_{n-1} (Yang *et al.*, 2009; Kalnay and Yang, 2010). However, similar to what was pointed out by Yang *et al.* (2012b), in our case, nonlinear growth of the perturbations becomes larger in the case of a fast developing and splitting storm, thus the linear assumption made in asynchronous EnKF that perturbations linearly evolve within the assimilation window may not be valid. This situation implies that updating the model state using future data far away from the model state in time may not yield the optimal analysis. In other words, if observations are available at t_{n-1}, they, instead of observations valid at t_n, are usually used to update the state at t_{n-1}. To allow later observations to update an earlier state, in our implementation, iterations are carried out between t_n and t^*_{n-1}, where t^*_{n-1} is an intermediate time between t_{n-1} and t_n when no observation is available. The other consideration in using t^*_{n-1} rather than t_{n-1} is to ensure that the observations at t_n have impacts on the model state at an earlier time; this means that this earlier update time should be within the cut off radius of temporal location; the asynchronous updating should occur within the time localization window. A longer time interval of iteration also costs more computationally because of the longer, repetitive, ensemble forecasts involved; tuning of an iterative procedure would probably need to balance analysis quality and computational cost.

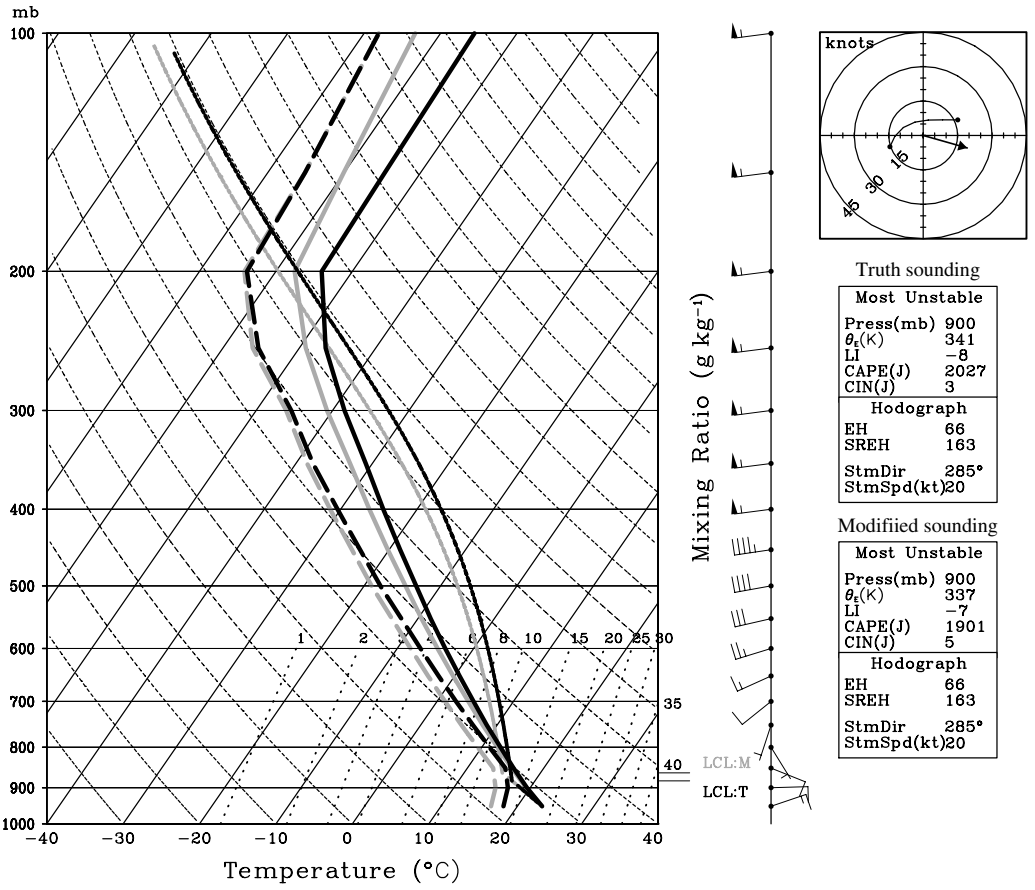

Figure 2. The original (black) and modified (grey) environmental soundings used by the truth simulation, and by the error-containing observing system simulation experiments, respectively. Solid and dashed lines are for temperature and dewpoint temperature, respectively. LCL:M indicates the lifting condensation level of the modified sounding while LCL:T indicates the original lifting condensation level.

The total computational cost of the iEnSRF increases linearly with the number of iterations. Assuming that the computational costs of the EnSRF analysis and ensemble forecast from t^*_{n-1} to t_n are A and B respectively, the cost for our current iEnSRF implementation with L iterations is about $L(2A + B)$. The computational cost of the EnSRF analysis has a factor of two because the model states at both t^*_{n-1} and t_n have to be updated in the current iEnSRF implementation. The cost of the EnSRF analyses can be reduced by 50% by updating the observation priors instead of the state vector, in a procedure similar to that of Anderson and Collins (2007); then the model state at t^*_{n-1} is updated in all except the final iteration, whereas the model state at t_n is updated only in the final iteration.

In addition, we note here the assumption that errors in the observations and the newly updated background are uncorrelated is no longer valid within the iterative procedure due to the influence of the same observations on the analysis states. Even so, because the quality of the initial background ensemble is poor, the benefit of the iteration procedure in improving the ensemble and hence the analysis apparently out-weighs the potential negative impact of the assumption.

3. The observing system simulation experiments

3.1. The prediction model settings and truth simulation

The non-hydrostatic and fully compressible WRF Version 2.2.1 model is used to produce the truth simulation of an idealized supercell storm and for the OSSEs. For all experiments, the physical domain used is 60 km × 60 km × 20 km. The model domain has a horizontal grid spacing of 2 km and a vertical grid spacing of 0.5 km. The truth simulation is initialized from the classic Weisman and Klemp (1982) analytic sounding provided in the WRF package (plotted in Figure 2 along with a modified sounding to be discussed later). The convective available potential energy (CAPE) of the sounding is about 2000 J kg^{-1} with directional (clockwise) vertical wind shear that favours right-moving cells following storm splitting (Klemp, 1987). The storm is triggered by an ellipsoidal thermal bubble centred at $x = 14$, $y = 28$ and $z = 1.5$ km, with a 10 km horizontal radius and a 1.5 km vertical radius and a 3 K maximum temperature perturbation. Other model configurations include: a Runge–Kutta third-order time integration scheme with a time-step of 12 s, WRF single-moment six-class (WSM6) microphysics parametrization scheme, and the Rapid Radiative Transfer Model (RRTM) and Dudhia schemes for long- and short-wave radiation. No cumulus parametrization is included. A 1.5-order turbulent kinetic energy (TKE) closure scheme is used to parametrize subgrid-scale turbulence and a positive definite scheme is used for the advection of moisture and water variables. Open conditions are used at the lateral boundaries. More details with regard to schemes in the WRF Version 2 model are described in Skamarock *et al.* (2005). The length of the truth simulation is 90 min.

We now briefly describe the truth simulation. In this article, all times are relative to the initial time (0 min) of the truth simulation. At 40 min (Figure 3(a)), two updraft cores

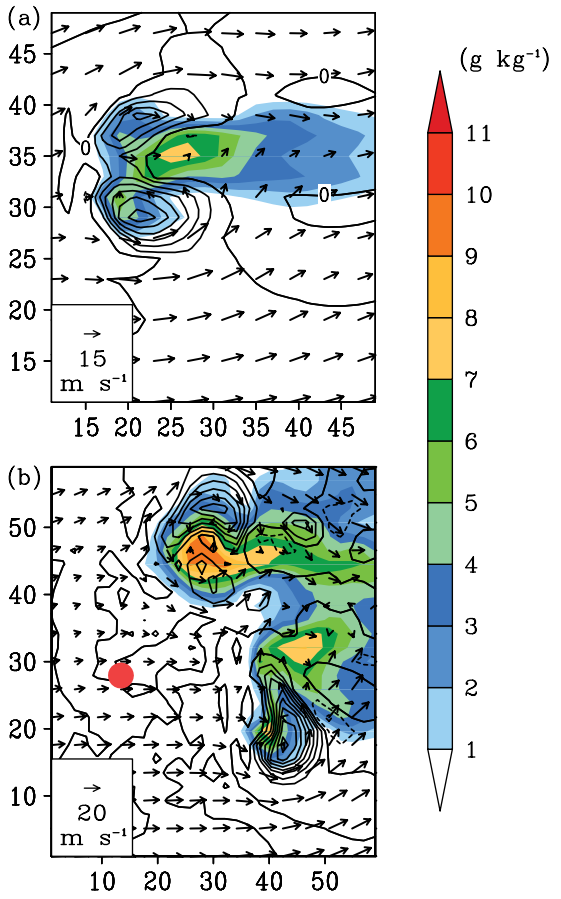

Figure 3. The graupel mix ratio (shaded), vertical velocity (contour at interval of 4 m s^{-1}) and horizontal wind (vector) at 5 km AGL for the truth simulation at model times of (a) 40 min and (b) 90 min. The red dot indicates the location of the initial thermal bubble.

are found at mid-levels, resulting from storm splitting. One core is located in the southern part of the rear of storm with the other located to the north. The southern cell (hereafter, SC) is stronger than the northern cell (hereafter, NC). The maximum updraft in the SC at this time reaches 25 m s^{-1} at 5 km above ground level (AGL). At 90 min (Figure 3(b)), the SC is located in the southeastern portion of the model domain, about 28 km from the position of initial thermal bubble (shown with a red dot in Figure 3(b)). The maximum updraft is greater than 30 m s^{-1} at 5 km AGL. Meanwhile, in the northern part of the model domain, the region of graupel associated with the NC becomes larger than that with the SC even though its updraft is still weaker. Plots at higher altitudes (not shown) feature a larger region of hydrometeors for the SC, indicative of a deeper updraft for the SC than the NC. Throughout the simulation period, the storm top of the SC reaches about 15 km and its maximum updraft reaches 40 m s^{-1}. This classic splitting supercell storm serves as the truth storm from which simulated radar data will be created for OSSEs.

3.2. *Simulation of radar observations*

A simulated WSR-88D type of radar is placed at $x = -22$ km and $y = 78$ km. This location is outside the model domain and about 60 km to the northwest of the initial thermal bubble. Following the recent OSSE articles of Xue *et al.* (2006) and Yussouf and Stensrud (2010), the simulated observations are sampled on radar elevations rather than at model grid points (Snyder and Zhang, 2003; Tong and Xue, 2005). Radar observations are simulated from model variables interpolated to the model scalar points in the horizontal direction and radar elevation in the vertical direction, as is common practice (Crook *et al.*, 2004). The radar operates in the standard U.S. operational WSR-88D radar precipitation scan mode (VCP 11), with 14 elevation levels and 5 min volume scan intervals. Following Yussouf and Stensrud (2010), the lower 12 sweeps of observations are generated at a rate of three sweeps per minute and the upper two sweeps are generated in the final minute of each volume scan. To take into account measurement and sampling error of radial velocity (V_r) and reflectivity observations (Z), random errors of zero mean and standard deviations of 2 m s^{-1} and 2 dBZ are added to V_r and Z, respectively. The operator for V_r is

$$V_r = u_g \cos\alpha \sin\beta + v_g \cos\alpha \cos\beta + (w_g - w_t)\sin\alpha, \quad (14)$$

where u_g, v_g and w_g are model-simulated velocities interpolated from the staggered model grid points to the observation location using tri-linear interpolation; w_t is the mean terminal fall speed of hydrometeors calculated according to Tong and Xue (2008); α and β are the elevation and azimuth angles of the radar beam, respectively. The simulated reflectivity, Z (in dBZ) is calculated from the mixing ratios of rainwater, snow and graupel using the formulations of Tong and Xue (2005) and Xue *et al.* (2006). The operator for Z satisfies the relationship

$$Z = 10\log_{10}\left(\frac{Z_r + Z_s + Z_h}{1\ mm^6 m^{-3}}\right), \quad (15)$$

where Z_r, Z_s and Z_h are equivalent reflectivity factors for rainwater, snow and graupel, respectively. Observations used in the assimilation experiments are calculated from Eqs (14) and (15) with added random errors. Radial velocity and reflectivity are only assimilated where reflectivity exceeds 10 dBZ.

3.3. *Data assimilation settings*

Forty ensemble members are used in all experiments. Similar to Tong and Xue (2005), ensemble forecasts begin 20 min into the simulation when the storm cell first develops from the thermal bubble. Random perturbations are added to a first-guess state to create an ensemble of initial conditions. These random perturbations have a Gaussian distribution with zero mean and a standard deviation of 3 K for perturbation potential temperature θ' (defined as total potential temperature θ minus 300 K) and 0.5 g kg^{-1} for water vapour mixing ratio q_v. The wind field is not perturbed. This configuration works reasonably well because wind perturbations can be induced effectively by thermal perturbations, somewhat like the triggering of storm cells by thermal bubbles. Perturbations are added at the grid points only where reflectivity > 10 dBZ is observed within 2 km in order to avoid spurious convection outside the observed area. After all the observations at an analysis time are analyzed, the relaxation inflation method of Zhang *et al.* (2004) is used to prevent ensemble underdispersion, and the inflation is applied to model states at both t^*_{n-1} and t_n. Following Zhang *et al.* (2004), the formula for relaxation inflation is

$$(\mathbf{x}^a_{new})' = (1-\gamma)(\mathbf{x}^a)' + \gamma(\mathbf{x}^b)', \quad (16)$$

where γ is the weight of the background ensemble perturbation and is set to 0.5 following Zhang *et al.* (2004). This parameter is used at both t^*_{n-1} and t_n. We applied additional inflation every 5 min to further increase the ensemble spread. This additional inflation is realized by scaling the spread of θ to 2 K in the areas influenced by observational data in the filter updating. This procedure has an effect of increasing the spread of thermal variables and limits added inflation to the areas impacted by observations; it has a similar effect as the additive noise method of Dowell and Wicker (2009) where temperature perturbations are added only in areas of precipitation.

For spatial localization a fifth-order correlation function (Gaspari and Cohn, 1999) is used. The horizontal radius is set to 8 km and the vertical radius is 4 km. The model variables updated by the DA system include the grid-point values of wind components u, v, w, potential temperature θ, perturbation geopotential ϕ, water vapour mixing ratio q_v, and mixing ratios of microphysical variables q_r, q_i, q_s and q_g for rain, ice, snow and graupel. Both radial velocity V_r and radar reflectivity Z are assimilated.

For the iEnSRF, the time interval between t^*_{n-1} and t_n is set to 4 min for the first analysis (at 25 min) and 3 min for all subsequent analysis cycles. Using a longer interval for the first analysis cycle allows for the development of flow-dependent covariance structures, which can be beneficial when starting from the poor initial ensemble. In RIP, the number of iterations is automatically adjusted according to a criterion based on the mean squares observation minus forecast or the forecast innovation. A similar criterion is not appropriate in our case, because none of the model state variables are directly observed, and a good fit of the model state to observations does not guarantee accurate state estimation. We choose instead to set the number of iterations to three for all analysis cycles with iterations, which is consistent with the experience of Kalnay and Yang (2010). In our iterative procedure, spatial covariance inflation and localization configurations for updating the state at t^*_{n-1} are the same as those for updating state at t_n. Temporal localization is performed using a fifth-order correlation function (Gaspari and Cohn, 1999). The cut-off radius of the temporal localization is set to 6 min (slightly longer than a radar volume scan time).

3.4. Design of the OSSE

As in many earlier OSSE studies with EnKF (e.g. Tong and Xue, 2005; Zhang *et al.*, 2006), we first examine the iEnSRF with a perfect model and a storm environment defined by the correct sounding. In this perfect model situation, two experiments are performed. One is named EnSRF_ne, using the EnSRF, while the other is named iEnSRF_ne, using the iEnSRF. The 'ne' in the names indicates no model or environmental error. Following Tong and Xue (2005) and Xue *et al.* (2006) we perform analyses every 5 min, from 25 min to 90 min for both experiments. Short-range deterministic forecasts are launched from the ensemble mean analyses at 40 min and 60 min, respectively. The forecasts launched at 40 min are designed to determine the impact of the iEnSRF on forecasts with a short assimilation window, whereas the forecasts launched at 60 min examine the impact of the iEnSRF after more data are assimilated through additional cycles.

In a second set of experiments, both the EnSRF and the iEnSRF are subject to a more realistic condition that includes error in both the prediction model and storm environment. Two experiments are again performed, the EnSRF_e and iEnSRF_e, where 'e' refers to the fact that error is included in the model and environmental sounding. As in the error-free experiments, cycled analyses are performed from 25 min to 90 min at 5 min intervals and two deterministic forecasts are launched from ensemble mean analyses at 40 min and 60 min. Similar to Zhang and Meng (2007) and Snook *et al.* (2011), the model error or uncertainty is modelled by using different physical parametrization schemes from those used in the truth simulation. Specifically, the Smagorinsky first-order closure scheme (SFOC) instead of the 1.5-order TKE scheme is used for the subgrid-scale turbulence parametrization; the Purdue Lin (Lin *et al.*, 1983; Rutledge and Hobbs, 1984) instead of the WSM6 scheme is used for microphysics, and a second-order instead of a third-order Runge–Kutta scheme is used for model time integration. The environmental error is introduced by adjusting the profile of the water vapour mixing ratio (q_v) and lapse rate of temperature of the sounding. A coefficient of 0.9 is multiplied to the q_v profile to represent an underestimate of the environmental moisture. For temperature, the amplitude of perturbation from the original profile mean is first reduced by 10%, a constant is then added to the profile so that its surface value remains unchanged. The adjusted sounding is shown in Figure 2 (grey lines). The largest differences in this modified sounding are the smaller value of CAPE, higher lifting condensation level (LCL), and larger dew point depression at the lower levels, indicating a less convectively unstable environment.

The impact of DA using the iEnSRF is then examined in terms of both analysis and forecast errors. In order to isolate the effects of the difference in the initial condition, we use the perfect model for the free-forecast period of both error-free and error-containing experiments. The configurations of the model and for all four experiments discussed above are listed in Table 1, including the parametrization schemes, assimilation window length, and the localization and inflation configurations.

4. Results and discussion

The root-mean-square (RMS) errors of the analyzed model state variables are used to quantify and compare the performance of the iEnSRF and the standard EnSRF. Following Tong and Xue (2005) and Xue *et al.* (2006), the RMS errors are calculated against the truth at grid points where the truth reflectivity is greater than 10 dBZ. The mass field used for verification is the perturbation geopotential height in metres. Meanwhile, for quantitative comparison of the iEnSRF and EnSRF, we define the improvement produced by the iEnSRF, *Imp*, as the percentage difference,

$$Imp = 100 \times (e_{\mathrm{ensrf}} - e_{\mathrm{iensrf}})/e_{\mathrm{ensrf}},$$

where e is the RMS error while the subscript denotes the analysis scheme used. Note that a positive value of *Imp* represents an improvement in the analysis of the iEnSRF.

4.1. Results with no model or environmental error

Figure 4 compares the RMS errors from the EnSRF_ne and iEnSRF_ne. When used from 25 min to 90 min, iEnSRF

Table 1. The model and data assimilation configurations of the observing system simulation experiments.

Parameters	Experiment name			
	EnSRF_ne	iEnSRF_ne	EnSRF_e	iEnSRF_e
Analysis scheme	EnSRF	iEnSRF	EnSRF	iEnSRF
Iteration interval time (min)	N/A	Cycle one: 4 Other cycles: 3	N/A	Cycle one: 4 Other cycles: 3
Iteration number for each cycle	N/A	3	N/A	3
Temporal localization (min)	N/A	6	N/A	6
Parametrization schemes for analysis	RK3/WSM6/TKE		RK2/Lin/SFOC	
Parametrization schemes for forecast	RK3/WSM6/TKE			
Environmental sounding	Truth	Truth	Adjusted	Adjusted
Spatial localization radius (km)	Horizontal: 8 Vertical: 4			
Inflation scheme/coefficient	Relaxation/0.5			
Additional inflation	Rescaling spread to 2 K for θ			

generally performs better than the regular EnSRF in the first four to five cycles, but its errors gradually become comparable with or even worse than those of the EnSRF in later cycles. It can be seen in Figure 4 that the iEnSRF reduces the initial background error for all variables except q_v more than the EnSRF does at the end of the first analysis cycle at 25 min. Table 2 shows that the *Imp* for the three wind components, u, v and w, is between 8.6% and 14.4% at this time. For variables that are not related directly to radial velocity, such as geopotential height ϕ and potential temperature θ, the *Imp* is 34.4% and 14.6%, respectively. For the hydrometeor variables, the *Imp* is somewhat less. For q_v, the *Imp* is -62% in the first cycle (Table 2). This suggests that the ensemble covariances used to update the moisture and microphysical variables are not reliable at the end of the first cycle even with the iterations of the iEnSRF.

At 40 min, after four assimilation cycles, the errors in both the EnSRF_ne and iEnSRF_ne are greatly reduced from their initial levels. The analysis errors for the wind components in the EnSRF_e are about 2.5 m s^{-1}, whereas those in the iEnSRF_e are reduced to about 2.0 m s^{-1} (Figure 4), representing an *Imp* of more than 20% for u and v and 30% for w by the iEnSRF. Although the analysis error of ϕ in the iEnSRF_ne is not much smaller than that in the EnSRF_ne, the analysis error of θ in the former is about 31.16% smaller than that in the latter. Moreover, substantial improvement can be seen for hydrometeor variables after applying the iEnSRF for four cycles. The *Imp* for the analysis error of q_v is about 8.5%, whereas for q_r, q_s and q_g *Imp* is over 30%. These results suggest that the iterations improved not only state estimation but also the covariance estimation; the updating of the mass and temperature fields relies on the cross-variable covariance when assimilating the radial velocity and reflectivity observations.

However, the positive impact of the iterative procedure did not last through the remainder of the assimilation cycles. This is seen most clearly in the analysis and background forecast errors of θ in the later cycles (Figure 4). From 75 through to 90 min, the errors of the iEnSRF are noticeably larger. The same degradation is seen for the wind components. This result agrees with Kalnay and Yang (2010), who also found that performing additional iterations after the filter reaches its asymptotic level could lead to larger analysis errors. This is due to overfitting of the analysis to observations, when the observations are used repeatedly after reasonable state and covariance estimations have already been established.

It is therefore recommended that the iterative procedure is stopped after the DA system stabilizes, which can be at the time when the change of the innovation falls below a specified value (Kalnay and Yang, 2010). Through further tuning and sensitivity experiments it may be possible to automatically determine when this stabilization has occurred; the extra iteration step should then be halted.

To further understand the filter behaviour, the ensemble spread is examined, in the same region where the RMS error is calculated. Figure 5 shows the forecast and analysis spreads for different variables for the EnSRF_ne and iEnSRF_ne. It can be seen that the ensemble spread is much smaller than the RMS error for all variables except for θ and q_v at the beginning of assimilation in both experiments. This is because only θ and q_v are perturbed in the initial ensemble. However, after about three cycles, the ensemble spread becomes comparable to the RMS error for most variables (c.f., Figure 4). The significant increase in the analysis ensemble spread for θ is due to our additional inflation that rescales its spread to 2 K after each analysis cycle (or each iteration in the iEnSRF_ne) if it is less than 2 K. This procedure actually causes overdispersion of the ensemble in terms of the potential temperature, especially with the analyses. The reason that the analysis spread is sometimes slightly larger than 2 K is because the region used for calculating the inflation differs slightly from that used for calculating the RMS error. Overall, the spreads of other variables are generally comparable to their RMSEs. We also examined the consistency ratios (e.g. Dowell *et al.*, 2004) in observation space; they convey similar information on the ensemble spread as Figures 4 and 5 and are therefore not shown. In addition, it can be seen from Figure 5 that the analysis ensemble spreads for u, v, w and ϕ in the iEnSRF_ne are larger than those in the EnSRF_ne after three iterations in the first cycle (25 min), even though the same observations were used several times in the iEnSRF_ne, indicating that our inflation procedure is effective in maintaining the ensemble spread. Starting from the third analysis cycle (35 min), the analysis and forecast spreads in the iEnSRF_ne become smaller for most variables than those in the EnSRF_ne and are actually closer to the corresponding RMS errors shown in Figure 4, giving a better consistency between the spread and error.

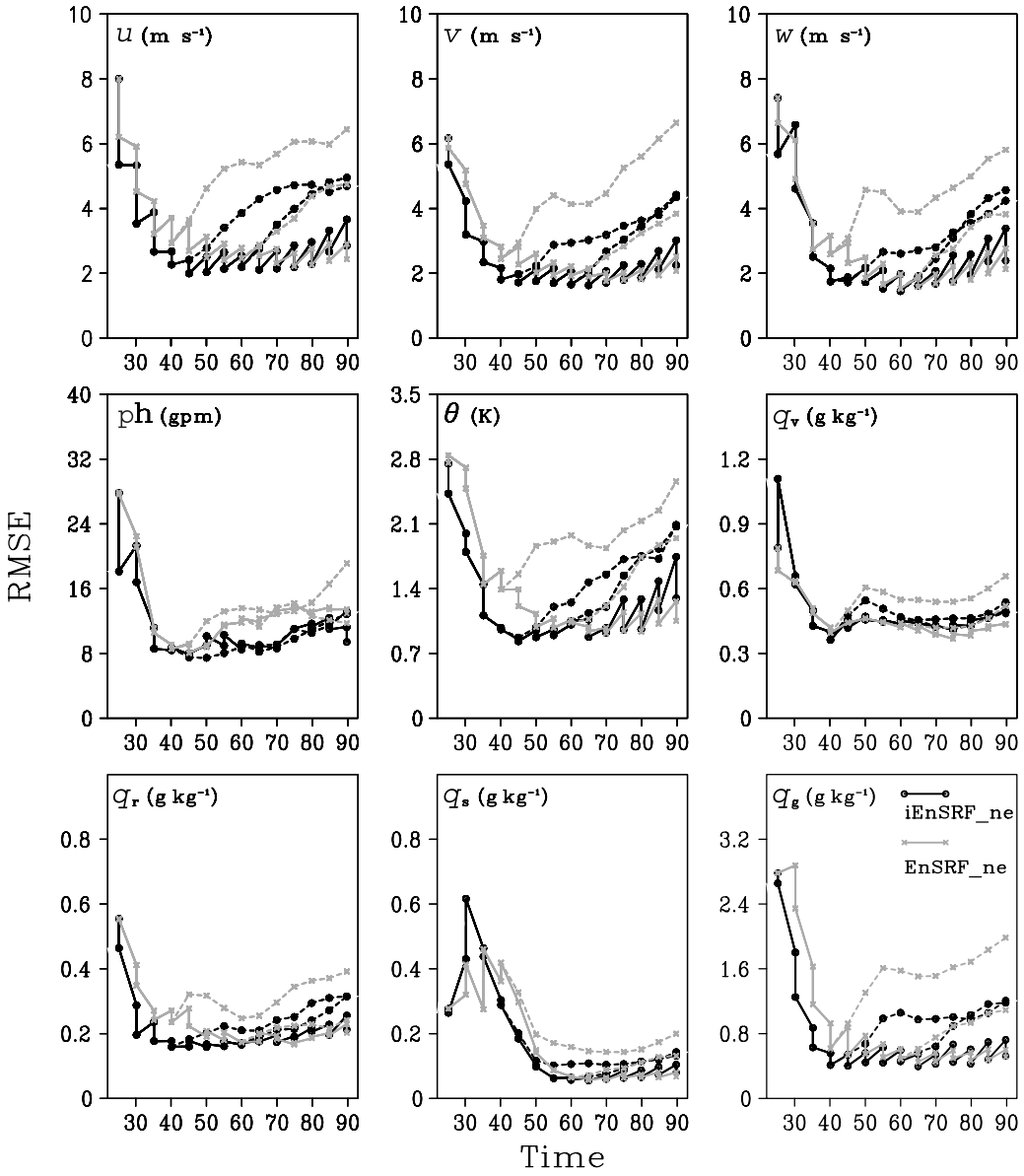

Figure 4. The evolution of RMS errors for experiments EnSRF_ne (grey line) and iEnSRF_ne (black line). Solid lines represent errors within the assimilation window while dashed lines represent errors of deterministic forecasts launched from the ensemble mean analyses at 40 and 60 min.

Table 2. Percentage differences between RMS errors of the iEnSRF_ne and EnSRF_ne at 25, 40 and 90 min of model time. At 25 and 40 min, the errors are for the analyses, while at 90 min they are for forecasts.

Time	u	v	w	ϕ	θ	q_v	q_r	q_s	q_g
25 min	13.80	8.65	14.42	34.41	14.60	−62.00	16.11	4.65	4.57
40 min	22.87	26.53	32.82	−1.48	31.16	8.51	32.20	31.21	33.21
90minF40	27.32	34.63	27.02	31.29	18.35	25.40	19.56	28.40	39.15
90minF60	−4.44	−15.17	−19.50	−9.50	−6.47	−3.03	−30.02	−0.20	−8.16

Positive values mean that errors are smaller in the iEnSRF_ne while negative ones mean errors are larger in the iEnSRF_ne.
F40 represents forecast proceeding from 40 min while F60 represents forecast proceeding from 60 min.

Launching from the analyses at 40 min, the forecast errors in the EnSRF_ne (grey dashed lines in Figure 4) increase rapidly and became about twice as large by 90 min for most variables. In contrast, the forecast errors in the iEnSRF_ne (black dashed lines in Figure 4) grow more slowly and remain substantially lower than those of the EnSRF_ne throughout the forecast period. For instance, the improvement over the EnSRF is between 27 and 35% for u, v and w at 90 min, while improvement for the other variables is at least 18% (Table 2). This is true even for variables that are not related directly to radial velocity or reflectivity, such as ϕ, θ and q_v. It is interesting to note that even though the improvement to the analysis at 40 min is slightly negative for ϕ, the forecast error in ϕ starting from this analysis is improved by close to 32% by 90 min, due to a more accurate analysis of most other state variables positively impacting the evolution of the forecast.

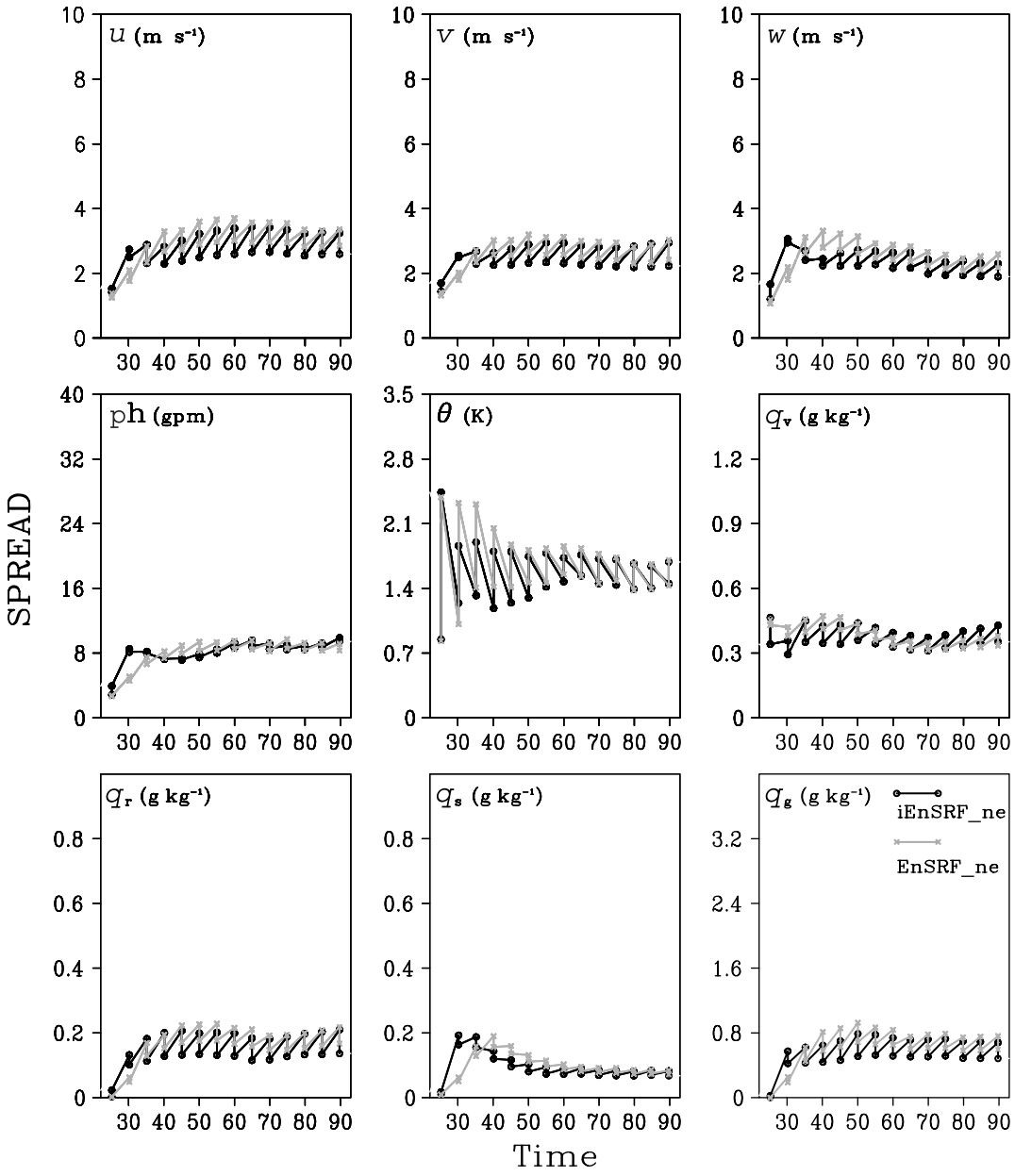

Figure 5. The forecast and analysis ensemble spread values for experiments EnSRF_ne (grey line) and iEnSRF_ne (black line). The spread values in the iEnSRF_ne are calculated for the final iteration of each cycle.

As mentioned earlier, we also launched deterministic forecasts from ensemble mean analyses at 60 min. By this time, both the EnSRF and iEnSRF had reached their asymptotic error levels and the analysis errors of all variables in the EnSRF_ne were comparable with those in the iEnSRF_ne. It can be seen in Figure 4 that the forecast error curves (dash lines) for these two experiments were nearly identical after being launched at 60 min, indicating that applying iterations after the filter reaches its asymptotic error level (shortly after 40 min in this case) does not improve the analysis and subsequent forecast. At 90 min, the forecast errors for nearly all variables in the iEnSRF_ne become slightly larger than those in the EnSRF_ne. Meanwhile, at 90 min, the error level of the forecast in the iEnSRF_ne starting from 40 min is comparable with that of the forecast starting from 60 min. This suggests that applying the iEnSRF effectively increases the forecast lead time by 20 min in the current case; in other words, the assimilation window can be 20 min shorter to reach a similar quality of state estimation for the current supercell storm, so that forecasts can be issued 20 min earlier.

4.2. *Results with model and environmental errors*

A perfect prediction model and perfect environmental conditions are not possible for real DA problems. In this section we examine the results of the second set of experiments; those that include prediction-model and storm-environment errors. Figure 6 shows the RMS errors from experiments EnSRF_e and iEnSRF_e. In this more realistic scenario, the iEnSRF still produces more accurate analyses than the EnSRF in the first several cycles. Similar to the no error case, the RMS errors in the later cycles become comparable to those of the EnSRF. Overall, the analysis errors in the error-containing experiments are comparable with the corresponding error-free experiments (compare Figure 6 with Figure 4).

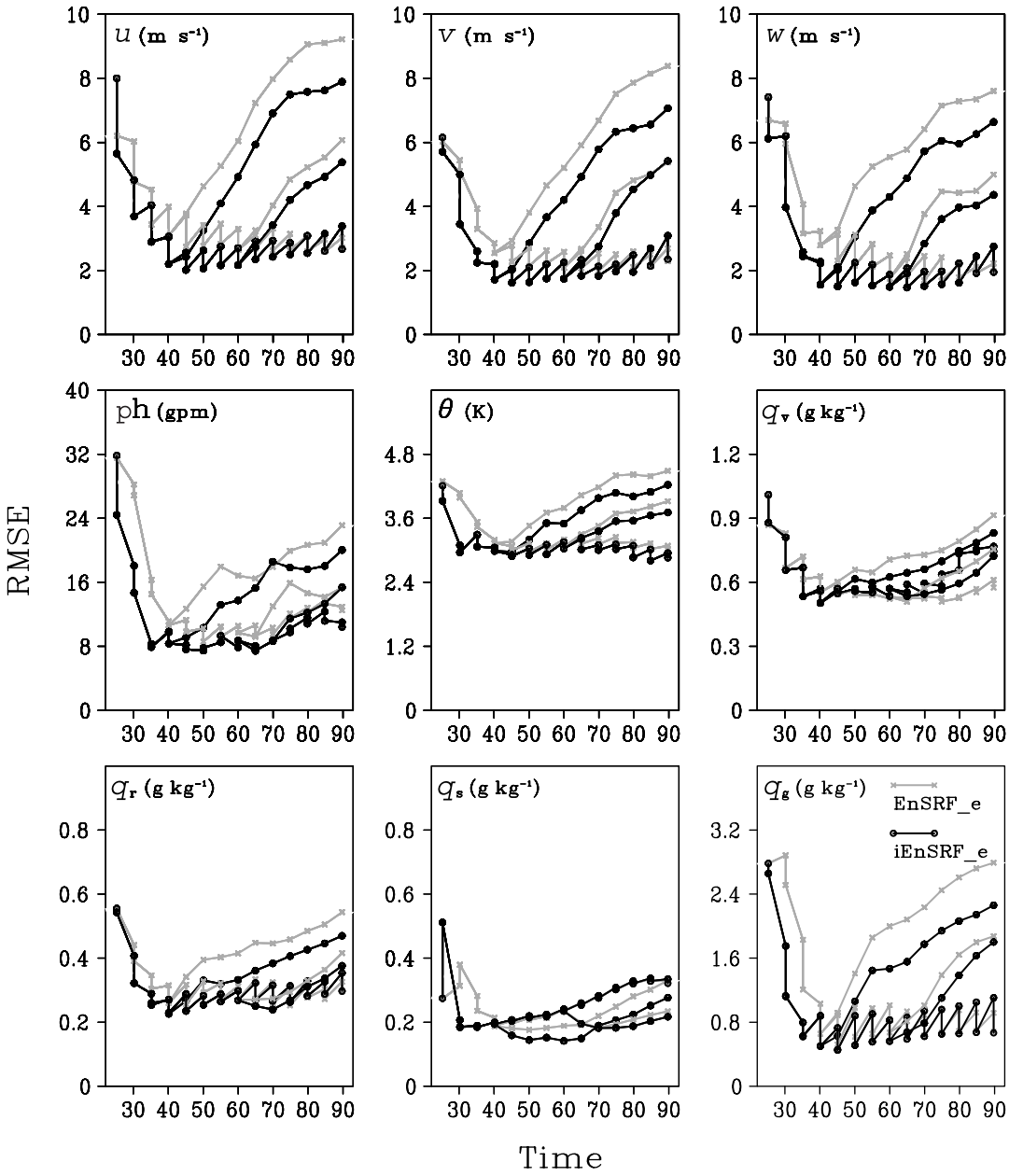

Figure 6. The same as Figure 4, but for experiments EnSRF_e (grey line) and iEnSRF_e (black line).

As can be seen in Table 3, the iteration procedure in the iEnSRF has positive impacts on the wind analysis at the end of the first cycle but the relative improvement is less than 10%, which is about 22–32% smaller than that in no-error experiments. The *Imp* for ϕ and θ are also smaller, achieving 22.66% and 8.52%, respectively, compared with the 34% and 14% of the no-error case. Among water vapour and hydrometeors, the iEnSRF produces smaller errors for q_r and q_g but it produces a large degradation in q_s verus q_v as in the no-error case. Similar to the no-error case, this larger analysis error for q_s here is probably due to unreliable covariance associated with the microphysical variables, and such a problem can be enlarged by use of the wrong microphysical scheme.

By 40 min or after four assimilation cycles, the improvements by the iEnSRF are again evident. At this time, the improvements to u, v and w are between 28% and 44% and the improvement to ϕ is also substantial (21.47%). The improvement to θ is smaller in the iEnSRF_e than in iEnSRF_ne (4.87 *vs* 31.16%), which is believed to be related to the error in the environment. The improvements to the moisture and hydrometeor variables are clearly evident, except for a small degradation in q_s. On average, the improvements to the hydrometeors are smaller than the no-error case. Similar to the no-error case, the analysis errors of the iEnSRF in later cycles become close to those of the EnSRF analyses, again suggesting that further iterations are unnecessary and may be undesirable after an asymptotic error level is reached.

The ensemble spreads in these experiments with imperfect conditions are also examined (not shown); the evolutions of ensemble spread in the iEnSRF_e and EnSRF_e are similar to their counterparts with perfect conditions. However, due to the environmental error, the ensemble spreads for θ and q_v are actually smaller than their RMSEs, unlike the error-free case.

When forecasts are launched at 40 min, the forecast errors in the EnSRF_e and iEnSRF_e initially grow at similar rates

Table 3. Percentage differences between RMS errors of the iEnSRF_e and EnSRF_e at 25, 40 and 90 min of model time. At 25 and 40 min, the errors are for the analyses while at 90 min they are for forecasts.

	u	v	w	ϕ	θ	q_v	q_r	q_s	q_g
25 min	9.07	5.33	8.50	22.66	8.52	−1.16	2.08	−86.79	4.53
40 min	28.63	32.79	44.26	21.47	4.87	9.12	13.09	−2.72	23.40
90minF40	14.39	15.75	12.71	13.36	5.81	9.05	13.55	16.16	19.04
90minF60	11.36	−0.73	12.64	−0.20	5.39	4.16	9.54	7.36	3.90

F40 represents forecast proceeding from 40 min while F60 represents forecast proceeding from 60 min.

as the corresponding no-error runs (compare Figure 6 with Figure 4), but after 10–15 min the errors grow much faster, reaching substantially higher levels by 90 min. These large forecast errors can be attributed to errors in the storm environment that cannot be easily corrected by the radar data. Throughout the forecast period, the RMS errors of the iEnSRF_e remain lower than those of the EnSRF_e for most variables. For a few variables, for example, u, v and q_s, the *Imp* is larger than it was at 40 min, indicating that the benefit of the iterative procedure, in this error-containing case (as well as the error-free case, Figure 4), is retained or even amplified for at least 50 min in the forecast period.

By 60 min, the analysis errors in both the EnSRF_e and iEnSRF_e have levelled off and the differences between them, although still identifiable, have become rather small. As a result, forecasts launched from analyses at this time produce similar forecast errors between the iEnSRF_e and EnSRF_e (Figure 6), although those of the iEnSRF_e are still slightly smaller. Compared with the error-free case, it seems to be beneficial to apply the iteration procedure for a couple more cycles in the error-containing case, given that *Imp* is mostly positive for 90minF60 in the error case (Table 3) whereas the corresponding *Imp* in the error-free case is mostly negative (Table 2). It is observed that the analysis errors for u and w at 60 min are still clearly lower in the iEnSRF_e than those in EnSRF_e, while the difference is already very small between the iEnSRF_ne and EnSRF_ne by this time; further iterations in the latter case would start to hurt. Overall, the benefit of using the iterative procedure to reduce spin up time in the iEnSRF is clear in both the error-free and the error-containing experiments.

4.3. *Error correlation structures, and analysis and forecast fields*

To help understand how the iterations produce the positive impacts noted in the previous subsection, we examine the background error correlation coefficients within the first assimilation cycle of the EnSRF_e and iEnSRF_e (Figure 7). The correlation between w and θ is first examined. A positive correlation between w and θ perturbations due to latent heat release within the updraft regions is physically consistent with, and important for, deep moist convection. In a vertical cross-section through the main updraft, one sample point marked by a red triangle is selected, where a hypothetical radial velocity (V_r) observation is assumed. At this point, the estimated w is smaller than the truth in both Figures 6(a) and 6(b). Therefore, if a V_r observation from a nearby radar can correct w at the sampled point, the increment of w at this point is expected to be positive. Surrounding this point, the initial estimate of θ was also smaller than the truth (Figure 7(a)), implying that a positive correlation between θ and w around the point is necessary in order for the assimilation of V_r to produce a positive θ increment, thereby reducing its error through variable cross-correlation. However, the correlation coefficients calculated from the ensemble background forecasts in the EnSRF_e are negative within the red ellipse (a region within the localization radius of the V_r observation) at this time (Figure 7(a)). Therefore, using the covariance between θ and the w component in radial wind prior to updating θ will lead to a negative impact on the θ analysis. Conversely, in Figure 7(b) for the iEnSRF, the correlations calculated from the background ensemble forecasts after three iterations are generally positive within the ellipse. This indicates that the ensemble correctly captured positive error correlations between θ and w after three iterations in the iEnSRF, but not in the EnSRF.

Similar behaviour can be observed with the q_g and w fields. Near the hypothetical observation point (marked by triangle), the forecast q_g and w in both the EnSRF_e and iEnSRF_e are smaller than the truth values (Figure 7(c) and (d)). Within the ellipse, positive correlations between q_g and w can be correctly estimated from the forecast ensembles in both the EnSRF_e and iEnSRF_e, but the maximum correlation coefficient in the former is small, reaching only 0.2, whereas in the latter, the value reaches 0.6. Given the larger correlation coefficient in the iEnSRF_e, the filter can produce a larger correction to q_g through the assimilation of V_r observations. In addition to the magnitude of error correlation, its spatial distribution is more reasonable in the iEnSRF_e. The error maximum of q_g in iEnSRF_e is located near the bottom boundary of the ellipse, implying that the error correlation should be smaller near the triangle but larger near the error maximum. This tendency is well captured in the iEnSRF_e. In the EnSRF_e, this spatial distribution is incorrectly estimated, with the error correlation maximum located near the triangle.

To further explore the impact of the iteration procedure on the analyzed storm, we examine the wind and q_g fields at 40 min, the time when we launch the first set of forecasts (Figure 8). As mentioned in section 3.1, the storm in the truth simulation begins to split into two cells (the NC and SC mentioned in section 3.1) at 40 min. The differences between the vertical profiles of analysis errors (not shown) of the iEnSRF_e and EnSRF_e were found to be substantial in the mid- and upper troposphere, especially for w and q_g. For example, at 5 km AGL, it is clear that there are two updraft maxima in the truth simulation (Figure 8(a)). This structure is captured in both the EnSRF_e and iEnSRF_e. However, the updrafts and precipitation core are stronger in the iEnSRF_e (Figure 8c), closer to the truth, than in the EnSRF_e (Figure 8(b)). The updraft of the northern cell in the EnSRF_e is particularly weak. In the truth simulation, the region of q_g greater than 2 g kg^{-1} extends to $x = 45$ km, which is not well captured by the EnSRF_e but better in the

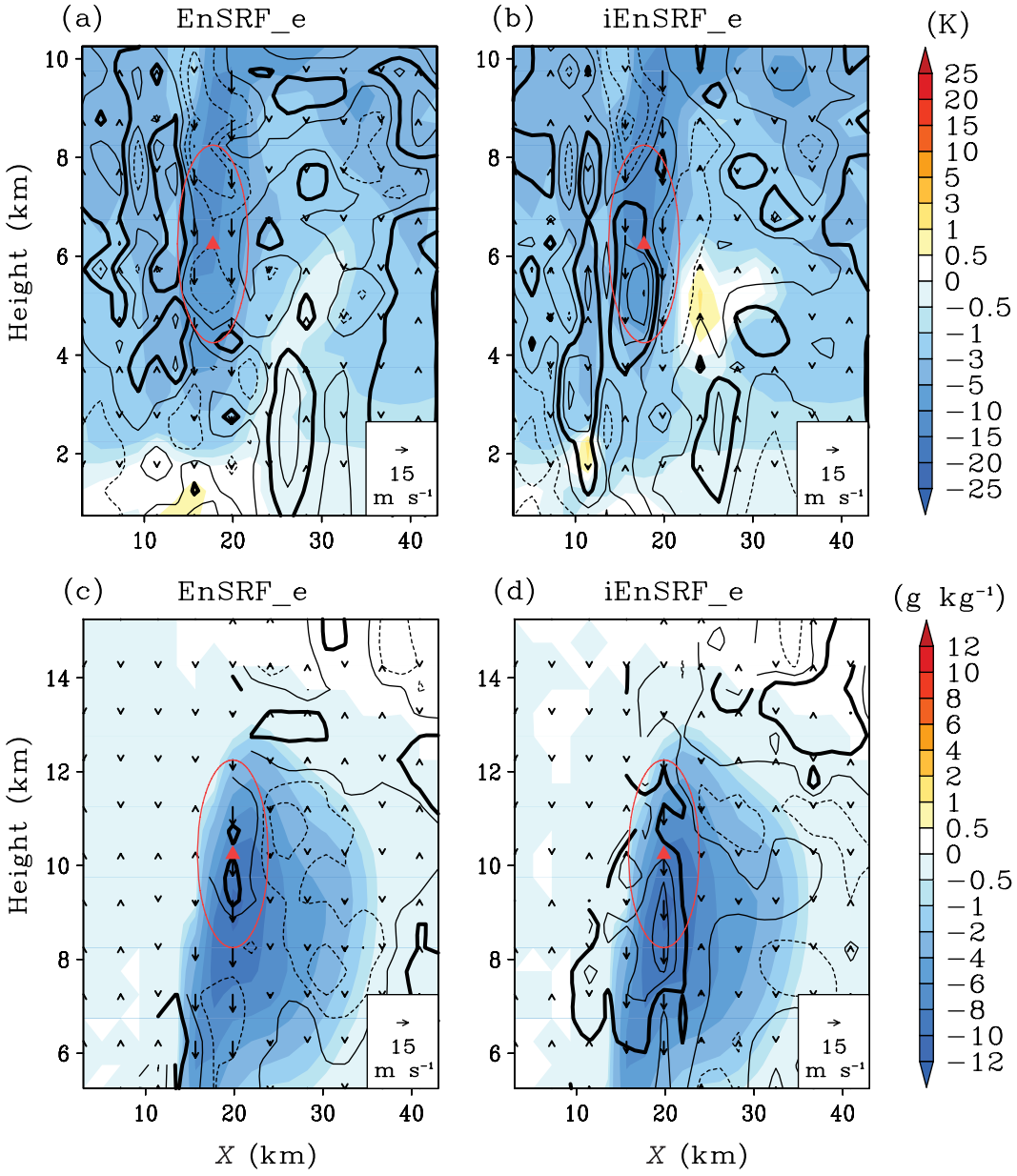

Figure 7. Vertical cross-sections of forecast errors, calculated as the ensemble mean minus the truth (shading for θ or q_g and vertical vectors for vertical velocity), and the correlation coefficients of forecast errors (contours at intervals of 0.2, with 0.2 correlation contours are in bold), at 25 min, the time of first analysis, in an x–y plane at $y = 30$ km. The red ellipse is the 0.2 contour for the localization coefficient for the observation located at the red triangle. For the iEnSRF_e the forecasts are at the end of the third iteration. The upper panels show the forecast errors of θ (shaded) and w (vertical vectors) and forecast error correlation coefficient between w at the point marked by the red triangle and θ at the grid point, for the EnSRF_e (left) and iEnSRF_e (right). The lower panels are the same as the upper panels except θ is replaced by q_g.

iEnSRF_e. The absolute errors of w shown in red contours clearly indicate much smaller errors in the iEnSRF_e, where the maximum error is about 4 m s^{-1} but that in the EnSRF_e reaches 12 m s^{-1}. These findings are consistent with the fact that the storm in the EnSRF_e is spun up less well because of the limited number of assimilation cycles without iterations.

Looking higher in the storm, at 10 km AGL, there is an indication of cell splitting in the structure of the updraft in the truth simulation (Figure 8(d)). In the EnSRF_e (Figure 8(e)), only a single updraft core is identifiable, whereas the iEnSRF_e produces an updraft structure (Figure 8(f)) that is closer to the truth. Such visual differences are consistent with the calculated absolute errors of w in the iEnSRF_e and EnSRF_e; their maximum errors are about 4 and 16 m s^{-1}, respectively. These results again show the benefit of iterations in accelerating the storm spin-up.

At the end of the 50 min forecast, the forecast errors of both the iEnSRF_e and EnSRF_e have become large (Figure 6), but the errors of the iEnSRF_e are still smaller than those of the EnSRF_e (Table 3). To see the visual differences in the forecasts, Figure 9 shows the forecast fields at 90 min from the EnSRF_e and iEnSRF_e at 5 and 10 km AGL. At this time, the strength of the storm in both the EnSRF_e and iEnSRF_e is weaker than the truth simulation, especially in terms of the graupel mixing ratio (q_g). As suggested earlier, the less unstable sounding that is used to define the storm environment probably contributed to such forecast errors. Nonetheless, the forecast storm in the iEnSRF_e is somewhat stronger than that in the EnSRF_e. For example, at 5 km AGL, the areal extent and maximum value of q_g associated with the SC are larger in the iEnSRF_e than those in the EnSRF_e. We see similar trends at 10 km AGL, with a larger updraft

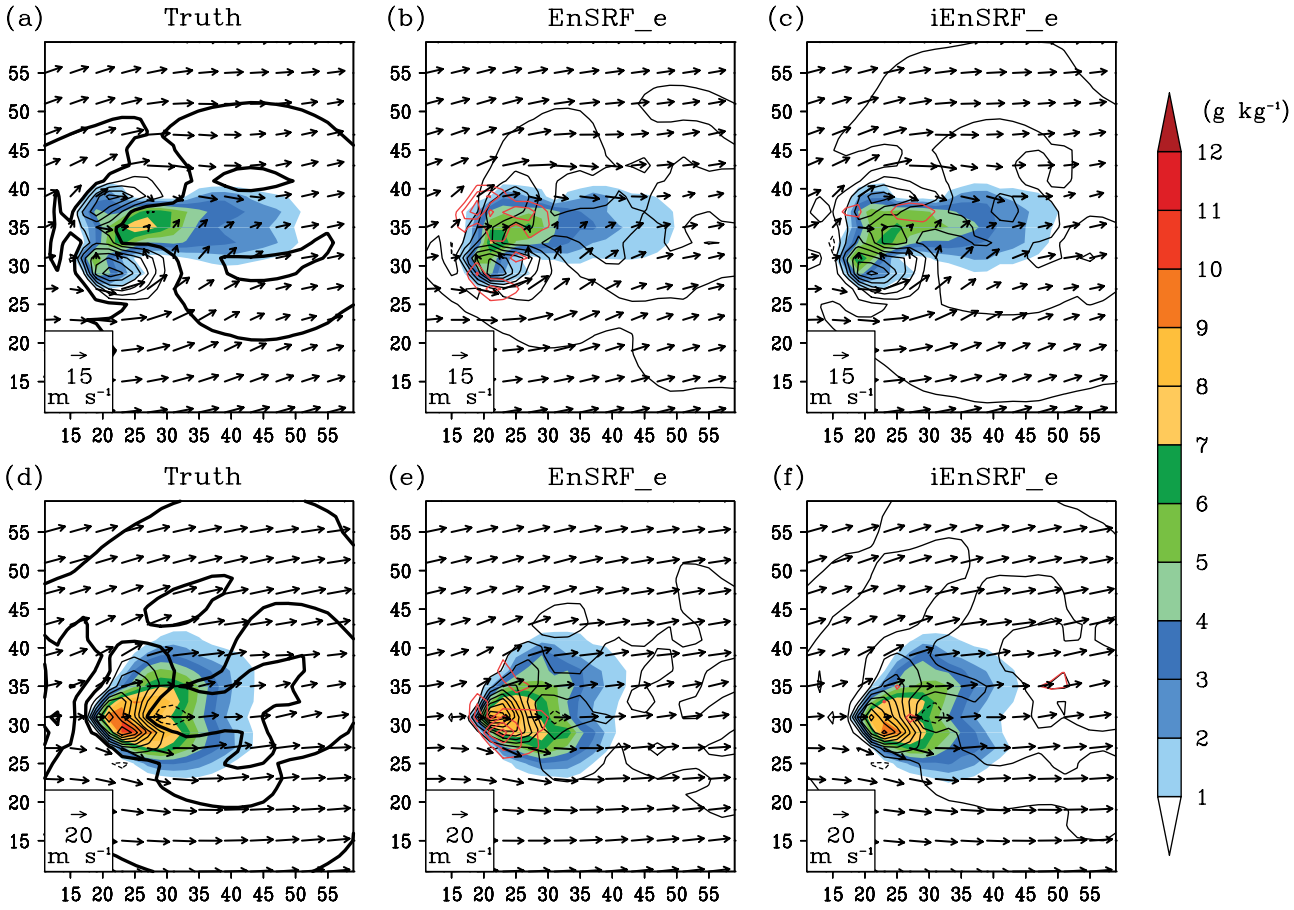

Figure 8. The graupel mix ratio (shaded), vertical velocity (contoured at intervals of 4 m s^{-1}) and horizontal wind vectors at 40 min for the truth simulation (left), EnSRF analysis (middle) and iEnSRF_e analysis (right), at 5 km (upper panels) and 10 km (lower panels) AGL. The red contours (at intervals of 4 m s^{-1}) in (b) (c) (e) and (f) are for the absolute w forecast errors.

and area of q_g in the iEnSRF_e than in EnSRF_e. Comparison with the truth simulation shows that the stronger and larger updraft of the iEnSRF_e is more accurate than that of the EnSRF_e. Therefore, the smaller error in the iEnSRF_e corresponds to a more accurate forecast of storm strength. In Figure 9, the maximum absolute error of w in the EnSRF_e is about 10 m s^{-1} larger than that in the iEnSRF_e, at both the middle and upper levels. Overall, although the relative improvement yielded by the iEnSRF in the presence of both model and environment error is not as large as that with perfect conditions, the benefits of the iteration procedure in the iEnSRF short-range storm forecasting through accelerated error reduction in the state estimation are clear.

5. Summary and conclusions

An iterative procedure based on the EnSRF, which we refer to as the iEnSRF, is designed with the goal of accelerating the 'spin-up' of the ensemble Kalman filter state estimation. The procedure is designed for use in those situations where the first-guess ensemble has a poor mean state estimate and poor ensemble error covariances. This procedure is similar to the 'running-in-place' (RIP) procedure proposed by Kalnay and Yang (2010), but differs substantially in its implementation due to differences between the local ensemble transform Kalman filter (LETKF) and the EnSRF algorithms used in the respective systems. In the iEnSRF, the background states at the analysis time and at an earlier time are combined into a new extended state vector. Using this extended vector the states at both times are updated by the filter, through using the asynchronous ensemble Kalman filter (Sakov *et al.*, 2010). By launching ensemble forecasts from the earlier updated states, and using the forecasts as the new background ensemble in Kalman filter updating in subsequent iterations within the same cycles, additional information can be extracted from the observations to reach a more accurate state estimate more rapidly.

Specifically, the iEnSRF contains three steps: first, a backward EnSRF analysis is performed that updates the ensemble model states at an earlier time. Second, an ensemble of forecasts is run from these updated model states to the analysis time. These two steps are then repeated a predetermined number of times. The backward analysis is performed via asynchronous EnKF, which can use observations not collected at the state updating time.

We test the iEnSRF algorithm using simulated radar data for an idealized supercell storm. Two sets of experiments are performed: one employs a perfect prediction model and a storm environment defined by the true environmental sounding. The other experiment set includes a combination of prediction model error and environmental error. In the error-free case, the iEnSRF accelerates the rate of error reduction and reaches a lower error level within the first four cycles. Continued application of the iterative procedure in later cycles leads to larger errors than in the non-iterative case, suggesting overfitting of the analysis to observations. In this idealized scenario, the error of the deterministic forecast

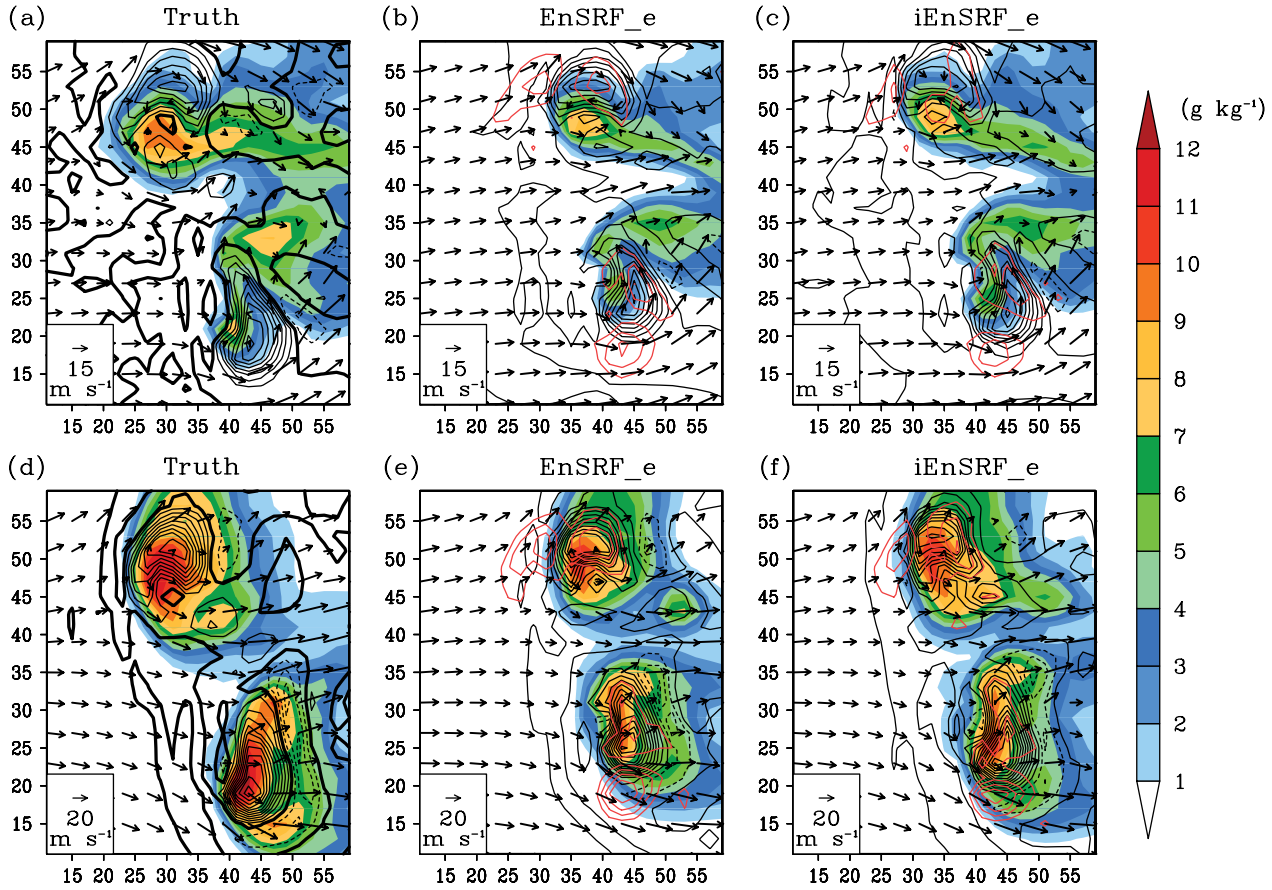

Figure 9. The same as Figure 8 but for deterministic forecasts at 90 min starting from the ensemble mean analyses at 40 min in EnSRF_e (middle panels) and iEnSRF_e (right panels), as compared to the truth (left panels). The red contours (at intervals of 10 m s^{-1}) in (b) (c) (e) and (f) are for the absolute *w* forecast errors.

starting from the iEnSRF analysis obtained after four cycles is found to be substantially smaller than that starting from the EnSRF analysis valid at the same time. After both the iEnSRF and EnSRF reach their asymptotic error level (about eight cycles), the difference between analyses and subsequent forecasts yielded by these two methods became small.

When model error and environmental error are present, the results are very similar. The iEnSRF performs better than the EnSRF in the first four cycles and becomes comparable in later cycles. The relative improvement over the non-iterative case is somewhat less than the error-free case but is still clearly evident. Examination of the analyzed storm at 40 min, or after four assimilation cycles, indicates that the iEnSRF is able to capture the storm splitting, intensity, and structure better than EnSRF. The benefit of the improved analysis is maintained throughout the 50 min forecast launched from the ensemble mean analysis at 40 min.

To better understand how the iEnSRF is able to produce more accurate analyses for unobserved variables in the first few cycles, we examined the background error correlations between different variables in the first cycle. Results showed that the error correlations calculated from the ensemble after three iterations with the iEnSRF are more physically consistent than those obtained without iterations. The improved ensemble error covariance obtained by the iterative procedure not only helps with the analysis of the wind and precipitating hydrometeor fields that are more directly observed, but also with the estimation of state variables, such as temperature, humidity and geopotential height and humidity that are not directly linked (through the observation operator) to the radar observations. Through iterations, the iEnSRF is able to produce more balanced analyses and reach a stable level of state estimation error in a fewer number of cycles than the corresponding non-iterated version. Because the iterative procedure not only improves the state estimation but also the error covariance, especially in the earlier cycles when these estimates tend to be poor, the method appears to be more beneficial for assimilating observations that are indirectly related to the state variables, as in the case of radar observations.

Similar to the findings of Kalnay and Yang (2010), the iterative procedure is found to dramatically improve state estimation in the first few assimilation cycles when starting from a poor initial ensemble. Such a poor initial ensemble is generally the case in thunderstorm initialization, where useful radar observations are typically not available before precipitation is present within a developing thunderstorm, or when DA is not running continuously in time so that ongoing thunderstorms have to be spun up starting from a coarse resolution operational model background that has little or no knowledge about the ongoing thunderstorms. After the cycles stabilize and the analysis error levels off, it is neither necessary nor desirable to continue the iterations, because overfitting to observations can occur and iterations also incur additional computational costs. For the problem of assimilating radar volume scan data at about 5 min intervals, it is recommended to use the iEnSRF only in the first four to five cycles. For practical implementations, we expect some level of tuning to reach optimal configurations.

Future work will examine the application of the iEnSRF for real data cases.

Acknowledgement

This research was primarily supported by NSF grants AGS-0802888 and OCI-0905040, the NOAA Warn-on-Forecast Project under NA080AR4320904, and by grant KLME060203 from NUIST. The second author was also supported by NSF grants AGS-0941491, AGS-1046171, and AGS-1046081. J. Min was also supported by Meteorology Commonwealth Project of China (GYHY200806029) and the Incubation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (NO708051). Computations were carried out at the University of Oklahoma Supercomputer Center for Education and Research (OSCER), and on the CAPS Linux Cluster machines.

References

Anderson JL. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Weather Rev.* **129**: 2884–2903.

Anderson JL, Collins N. 2007. Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.* **24**: 1452–1463.

Caya A, Sun J, Snyder C. 2005. A comparison between the 4D-VAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Weather Rev.* **133**: 3081–3094.

Crook NA, Dowell D, Sun J, Zhang Y. 2004. Assimilation of radar observations of a supercell storm using 4DVar: Parameter retrieval experiments. *22nd Conference on Severe Local Storms*, Hyannis, Massachusetts, American Meteorology Society, CDROM 8A.2.

Dowell DC, Wicker LJ. 2009. Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.* **26**: 911–927.

Dowell D, Zhang F, Wicker LJ, Snyder C, Crook NA. 2004. Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: Ensemble Kalman filter experiments. *Mon. Weather Rev.* **132**: 1982–2005.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**: 10143–10162.

Evensen G. 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynam.* **53**: 343–367.

Gaspari G, Cohn SE. 1999. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* **125**: 723–757.

Houtekamer PL, Mitchell HL. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Weather Rev.* **129**: 123–137.

Hunt BR, Kostelich EJ, Szunyogh I. 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D* **230**: 112–126.

Jung Y, Xue M, Zhang G, Straka J. 2008. Assimilation of simulated polarimetric radar data for a convective storm using ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Mon. Weather Rev.* **136**: 2246–2260.

Kalnay E, Yang SC. 2010. Notes and correspondence accelerating the spin-up of ensemble Kalman filtering. *Q. J. R. Meteorol. Soc.* **136**: 1644–1651.

Kalnay E, Li H, Miyoshi T, Yang SC, Ballabrera-Poy J. 2007. Response to the discussion on '4-D-Var or EnKF?' by Nils Gustafsson. *Tellus*, **59**: 778–780.

Klemp JB. 1987. Dynamics of tornadic thunderstorms. *Ann. Rev. Fluid Mech* **19**: 369–402.

Lin YL, Farley RD, Orville HD. 1983. Bulk parameterization of the snow field in a cloud model. *J. Clim. Appl. Meteorol.* **22**: 1065–1092.

Lorenz EN. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**: 130–141.

Rutledge SA, Hobbs PV. 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.

Sakov P, Evensen G, Bertino L. 2010. Asynchronous data assimilation with the EnKF. *Tellus*, **62**: 24–29.

Skamarock WC, Klemp JB, Dudhia J, Gill D, Barker D, Wang W, Powers JG. 2005. *A Description of the Advanced Research WRF Version 2*. NCAR Technical Note NCAR/TN-468+STR, National Center for Atmospheric Research: Boulder, CO.

Snook N, Xue M, Jung YS. 2011. Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-Band and WSR-88D radar data. *Mon. Weather Rev.* **139**: 3446–3468.

Snyder C, Zhang F. 2003. Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Weather Rev.* **131**: 1663–1677.

Tong M, Xue M. 2005. Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS Experiments. *Mon. Weather Rev.* **133**: 1789–1807.

Tong M, Xue M. 2008. Simultaneous estimation of microphysical parameters and atmospheric state with radar data and ensemble square-root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Weather Rev.* **136**: 1630–1648.

Weisman ML, Klemp JB. 1982. The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Weather Rev.* **110**: 504–520.

Whitaker JS, Hamill TM. 2002. Ensemble data assimilation without perturbed observations. *Mon. Weather Rev.* **130**: 1913–1924.

Xue M, Tong M, Droegemeier KK. 2006. 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.* **23**: 46–66.

Xue M, Jung YS, Zhang GF. 2010. State estimation of convective storms with a two-moment microphysics scheme and an ensemble Kalman filter: Experiments with simulated radar data. *Q. J. R. Meteorol. Soc.* **136**: 685–700.

Yang SC, Corazza M, Carrassi A, Kalnay E, Miyoshi T. 2009. Comparison of local ensemble transform Kalman filter, 3DVAR, and 4DVAR in a quasigeostrophic model. *Mon. Weather Rev.* **137**: 693–709.

Yang SC, Kalnay E, Hunt B. 2012a. Handling nonlinearity and non-Gaussianity in ensemble Kalman filter. Special collection 'Intercomparisons of 4DVariational Assimilation and the Ensemble Kalman Filter'. *Mon. Weather Rev.* **140**: 2628–2646.

Yang SC, Kalnay E, Miyoshi T. 2012b. Accelerating the EnKF Spinup for Typhoon Assimilation and Prediction. *Weather Forecast.* **27**: 878–897.

Yussouf N, Stensrud DJ. 2010. Impact of Phased-Array Radar Observations over a Short Assimilation Period: Observing System Simulation Experiments Using an Ensemble Kalman Filter. *Mon. Weather Rev.* **138**: 517–538.

Zhang FQ, Meng ZY. 2007. Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Weather Rev.* **135**: 1403–1423.

Zhang F, Snyder C, Sun J. 2004. Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filter. *Mon. Weather Rev.* **132**: 1238–1253.

Zhang F, Meng Z, Aksoy A. 2006. Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect model experiments. *Mon. Weather Rev.* **134**: 722–736.

