UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

HYBRID EN3DVAR RADAR DATA ASSIMILATION AND COMPARISONS WITH 3DVAR AND ENKF WITH OSSES AND A REAL CASE

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

RONG KONG
Norman, Oklahoma
2017

HYBRID EN3DVAR RADAR DATA ASSIMILATION AND COMPARISONS WITH 3DVAR AND ENKF WITH OSSES AND A REAL CASE

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

_____
Dr. Ming Xue, Chair

_____
Dr. Alan Shapiro

_____
Dr. David Parsons

_____
Dr. Fanyou Kong

_____
Dr. Keith Brewster

_____
Dr. Xiangming Xiao

Acknowledgements

I would like to express my sincere gratitude to my advisor, Dr. Ming Xue, for his guidance and mentoring and for his encouragement and support at all levels throughout my Ph.D. research. He provided me with many valuable suggestions on my research at our regularly meetings, without which the reported research progresses would not have been possible.

I also would like to thank all other members of my dissertation committee: Drs. Alan Shaprio, David Parsons, Fanyou Kong, Keith Brewster, Lance Leslie, and Xiangming Xiao for their advices on my research, for reading this dissertation and their suggestions for improvement.

I especially thank Dr. Youngsun Jung for her suggestions and providing the initial and boundary condition for the experiments conducted in Chapter 5. I thank Jonathan Labriola for his help and suggestions on the IAU algorithms used in Chapter 5, and thank Shizhang Wang for providing the background and observation data for the initial tests of En3DVar based on a real case. Thanks also extend to staffs at school of meteorology and CAPS, particularly, Celia Jones, Marcia Pallutto, Eileen Hasselwander, and Debra Farmer for their kind help. My research is also benefited from discussions with members of the research group led by Dr. Xue, including Drs. Gang Zhao, Chong-Chi Tong, and others.

My deepest thanks go to my family for their unconditional love and support. I thank my husband, Chengsi, for sharing my frustrations and joys throughout my dissertation work, and for the discussions and suggestions on my research. I especially own much to my parents Fanyu Kong and Guifen Song, father and mother in law Shubo Liu and Xinfang Huang, stepfather and stepmother in law Biao Fu and Ping for taking

care of my son and my daughter for three years. Without their supports, the research would not have been possible.

This research was primarily supported by the NOAA Warn-on-Forecast (WoF) grant NA160AR4320115. Computational resources of the Oklahoma Supercomputing Center for Research and Education (OSCER), and the NSF XSEDE Supercomputing Centers were used.

Table of Contents

List of Tables

List of Figures

Abstract

Studies have shown advantages of the hybrid ensemble-variational data assimilation (DA) algorithms over pure ensemble or variational algorithms, although such advantages at the convective scale, in the presence of complex ice microphysics and for radar data assimilation, have not yet been clearly demonstrated, if the advantages do exist. A hybrid ensemble-3DVar (En3DVar) system is developed recently based on the ARPS 3DVar and EnKF systems at the Center for Analysis and Prediction of Storms (CAPS). In this dissertation, hybrid En3DVar is compared with 3DVar, EnKF, and pure En3DVar for radar DA through observing system simulation experiments (OSSEs) under both perfect and imperfect model assumptions. It is also applied to a real case including multiple tornadic supercells. For the real case, radar radial velocity and reflectivity data are assimilated every 5 minutes for 1 hour that is followed by short-term forecasts. DfEnKF that updates a single deterministic background forecast using the EnKF updating algorithm is introduced to have an algorithm-wise parallel comparison between EnKF and pure En3DVar.

In the perfect-model OSSEs, DfEnKF and pure En3DVar are compared and are found to perform differently when using the same localization radii. The serial (EnKF) versus global (pure En3DVar) nature of the algorithms, and direct filter update (EnKF) versus variational minimization (En3DVar) are the major reasons for the differences. Hybrid En3DVar for radar DA is also compared with 3DVar, EnKF, DfEnKF, and pure En3DVar. Experiments are conducted first to obtain the optimal configurations for different algorithms before they are compared; the optimal configurations include the optimal background decorrelation scales for 3DVar, optimal localization radii for EnKF, DfEnKF, and pure En3DVar, as well as the optimal hybrid weights for hybrid En3DVar.

When the algorithms are tuned optimally, hybrid En3DVar does not outperform EnKF or pure En3DVar, although their analyses are all much better than 3DVar. When ensemble background error covariance is a good estimation of the true error distribution, pure ensemble-based DA methods can do a good job, and the advantage of including static background error covariance **B** in hybrid DA is not obvious.

In the imperfect-model OSSEs, model errors are introduced by using different microphysical schemes in the truth run (Lin scheme) and in the ensemble forecasts (WSM6 scheme). Experiments are conducted to obtain the optimal configurations for different algorithms, similar to those in perfect-model OSSEs. Hybrid En3DVar is then found to outperform EnKF and pure En3DVar (3DVar) for better capturing the hail analyses below the freezing level (intensity of the storm). The advantage of hybrid En3DVar over pure ensemble-based methods is most obvious when ensemble background errors are systematically underestimated. In addition, the impact of adding a mass continuity constraint in 3DVar, pure and hybrid En3DVar is also examined. Overall, adding the mass continuity constraint improving the analyses by producing a little stronger vertical velocity analyses that are much closer to the truth and by smoothing noise present in the velocity and hydrometer fields.

Finally, the ARPS hybrid En3DVar system is applied to the assimilation of radar data for a real tornadic supercell storm. Hybrid En3DVar is compared with 3DVar, EnKF, DfEnKF, and pure En3DVar based on both objective verification and the analyses and forecasts of storm intensity and structures. Hybrid En3DVar with 75% weight of static **B** clearly outperforms 3DVar in better capturing the hook echo structure and rotating updraft in the forecasts, and outperforms EnKF and DfEnKF in better

capturing forecast reflectivity between 35 and 45 dBZ. The low-level mesocyclone is better forecast by hybrid En3DVar than by other methods, suggesting stronger rotations and a larger tornado threat in the forecast.

In perfect-model OSSEs, only precipitation reflectivity ($\geq$ 5 dBZ) is assimilated. In the imperfect-model OSSEs and the real data case, the clear-air reflectivity (< 5 dBZ) is also assimilated to help suppress spurious storms. The assimilation of clear-air reflectivity by the variational DA algorithms is found to seriously degrade the analyses in storm region, with the intensity of the reflectivity analysis being much weaker than that assimilating reflectivity larger than 5 dBZ only. When using hydrometeor mixing ratios as the control variables, the gradient of the cost function becomes extremely large when background reflectivity is small. In the imperfect-model OSSEs, a double-pass procedure is proposed and found to help alleviate this problem. In the real data case, an alternative way of using the logarithmic hydrometeor mixing ratios as the control variables is found to be a better solution to the problem. In such a case, the excessively large gradient of the cost function is avoided.

Chapter 1 Introduction and Overview

1.1 Background and Motivation

1.1.1 Background

Numerical weather prediction (NWP) is a method of weather forecasting that uses a set of equations to describe flows of the atmosphere. In a NWP model, an initial state is integrated forward to get the model state at a later time, making use of the boundary conditions. As pointed out by Kalnay (2002), NWP is an initial-boundary value problem. The more accurate the estimate of the initial condition, the better the quality of the forecasts. Data assimilation, by which the observations and short-range forecast are combined to obtain an initial condition for NWP, is found to be able to greatly improve the forecast skill in the past four decades (Bauer et al. 2015). The operational implementation of the four-dimensional variational (4D-Var) satellite data assimilation at the European Centre for Medium-Range Weather Forecasts (ECWMF) in 1997 marks a major milestone in operational global NWP.

A major task of NWP is to issue severe weather alerts and advisories to protect life and property. Severe weather phenomena include tornadoes, damaging winds or gust greater or equal to 25.8 ms-1 (mph), and hail with its size greater than or equal to 2.5 cm (1 inch). According to NOAA's National Centers For Environmental Information (NCEI), 15 billion-dollar weather disasters occurred in the U.S. during the first 9 months (Jan - Sep) of 2017. Damage from severe thunderstorms, including tornadoes, large hail and straight-line winds, accounts for 7 of the 15 weather disasters that have cost $1 billion or more in 2017.

Numerical predictions of convective storms date back to Lilly (1990). Steady progress has been made during the past two decades, aided by advances of the observation network (particularly in radars) and high-performance computing. The convection-resolving simulation is usually conducted at a horizontal resolution in the order of a few kilometers. Most research communities and operational forecast centers now use or plan to use models at kilometer scale (Davies et al. 2005; Lean et al. 2008; Weisman et al. 2008; Baldauf et al. 2011; Hirahara et al. 2011; Seity et al. 2011; Verrelle et al. 2015; Stratman and Brewster 2017). Storm simulations from even higher resolution (sub-kilometer) are reported from Warren et al. (2014), Xue et al. (2014), and Hanley et al. (2016) etc.

To improve the accuracy of numerical weather prediction (NWP) at the convective scale, active research has been carried out in the past two decades to assimilate radar observations into numerical models. The assimilation of radar data using traditional three-dimensional variational approach (3DVar) or ensemble Kalman filtering (EnKF) has been shown to improve the prediction of convective systems (Hu et al. 2006a; Hu et al. 2006b; Kain et al. 2010; Snook et al. 2011; Sun et al. 2014). More recently, the hybrid data assimilation (DA) approach that combines 3DVar and EnKF methods has been found to have some advantages over pure EnKF or pure 3DVar for large scale and mesoscale DA.

3DVar is attractive for radar DA because of its relatively low computational cost and the ability to include weak equation constraints (Gao et al. 1999) to the cost function. Because the radial velocity (Vr) observation operator is relatively simple when the hydrometeor terminal velocity effect is pre-corrected, it is straightforward to analyze

radial velocity data variationally. In Gao et al. (2004), a 3DVar system was developed within the ARPS model (Xue et al. 2000;(Xue et al. 2003)) framework for analyzing Vr data that utilizes recursive filters to model the background error covariance. Xiao et al. (2005) assimilated Vr data into the WRF 3DVar for a heavy rainfall case and obtained improvement in rainfall forecasts. In comparison, the observation operator of radar reflectivity (Z) data is complicated; it involves multiple hydrometeor species and is highly nonlinear. Without additional physical constraints, the problem of assimilating reflectivity data within a 3DVar framework is under-determined and/or non-unique.

For the 3DVar systems of MM5 and WRF models, Xiao et al. (2007) developed a procedure to directly assimilate Z data but with a warm-rain assumption in which Z is a function of rain water mixing ratio only. In Wang et al. (2013), it was proposed that rainwater is first retrieved from Z data before assimilation into the WRF 3DVar due to large linearization errors of the Z operator when the background rainwater mixing ratio is small. The needed assumption of warm rain microphysical processes is clearly a major limitation with above procedures. For real storms, neglecting ice microphysics processes is a major source of error.

When ice processes are included in the model, the Z operator is at least a function of the mixing ratios of several precipitating hydrometeors, typically the rainwater, snow and hail/graupel mixing ratios (Tong and Xue 2005). In a typical 3DVar framework, the analysis algorithm itself is incapable of properly partitioning the precipitation information contained in Z among the liquid and ice hydrometeors, unless additional physical or statistic constraints are introduced; the flow-dependent background error statistics for the hydrometeors provided by ensemble forecasts are one

of such examples. In Gao and Stensrud (2012), the standard reflectivity formula is modified so that Z is only dependent on the snow and hail mixing ratios above the freezing level while below the level Z is only dependent on rainwater mixing ratio. With such a modification; they call this a hydrometeor classification-based algorithm. While this approach avoids the creation of spurious liquid phases at the upper levels by the 3DVar, it also excludes the presence of hail below the freezing level, sometimes all the way to the ground.

For this reason, indirect assimilation of Z is most commonly employed so far when using the 3DVar method, such as the use of a complex cloud analysis scheme. For example, the ARPS 3DVar system is often used in combination with a complex cloud analysis package to assimilate Vr and Z data together. Xue et al. (2003), Hu et al. (2006a); Hu et al. (2006b) and many subsequent studies (e.g., Dawson and Xue 2006; Zhao and Xue 2009; Kain et al. 2010; Schenkman et al. 2011; Xue et al. 2014) have demonstrated the effectiveness of this system for analyzing and predicting supercell storms, mesoscale convective systems as well as tropical cyclones. Additional physical constraints are used in the ARPS cloud analysis system to partition among hydrometeors associated with ice microphysics when assimilating Z data, as well as to adjust in-cloud temperature and moisture. Current cloud analysis is very valuable to improve NWP systems. For example, thunderstorms that have increments in hydrometeors may imply that saturation has occurred. By examining the ambient profile, the amount of latent heating occurring can be obtained, and the positive temperature perturbations can be deduced. This buoyancy helps to get a stronger updraft in the subsequent forecast through buoyancy accelerations. More recent applications of the

ARPS 3DVar/cloud analysis system to assimilating radar radial velocity and/or reflectivity data include Du et al. (2012); Ge et al. (2012); Gasperoni et al. (2013); Ge et al. (2013), Xue et al. (2014), Brewster et al. (2016); Stratman and Brewster (2017). Such cloud analysis schemes are computationally fast, over a large range of scales, and efficient to handle a broad selection of data. However, cloud analysis does not consider the observation error so that its solution is not optimal. In addition, there is no model constraint used (such as the microphysical constraint), so that background error covariance cannot consistently adjust wind, temperature, and moistures fields. For reflectivity assimilation, the analysis increments assign to different hydrometeor fields are purely based on the distribution of background hydrometeor mixing ratios. When there exist severe errors in background hydrometeor fields, the hydrometeor analyses will have large errors.

Compared to 3DVar, 4DVar incorporates the NWP model as a strong constraint into the DA system, enabling physical linkages among the model state variables and the retrieval of unobserved state variables (Le Dimet and Talagrand 1986; Courtier 1997; Kalnay 2002). Sun et al. (1991) demonstrated 4DVar of wind and thermodynamic fields for Doppler radar data, and Sun and Crook (1994) further tested this technique for a gust front case. Sun and Crook (1997, 1998) found that assimilating rainwater-mixing ratio derived from reflectivity observations produced better retrievals than directly assimilating reflectivity. However, only rainwater mixing ratio is considered and ice hydrometeors are neglected in their studies. Wu et al. (2000) represents one of few studies that include ice processes within a 4DVar DA. In that study, a simple ice microphysics scheme was used that included only one precipitating ice hydrometeor

category called graupel/hail. Rainwater and ice mixing ratios were initially diagnosed first from Z and available differential reflectivity observations before being assimilated into the 4DVar. Here, the differential reflectivity observations helped partitioning the liquid and ice hydrometeors. Because of many issues associated with the high nonlinearity of complex yet important physical processes (such as ice microphysics) in mesoscale and convective scale models, 4DVar has not enjoyed as much success for convective-scale DA as with large scale NWP (Sun et al. 2014).

EnKF is an alternative method that has enjoyed increasing popularity in convective scale DA, since its first application to radar DA by Snyder and Zhang (2003). In EnKF, flow-dependent background error covariances are derived from an ensemble of forecasts and used to update the state variables. EnKF allows for direct use of nonlinear observation operators and is more suitable for a NWP model with complex parameterizations. The benefits of EnKF for convective-scale NWP and radar DA had been demonstrated in many studies (Tong and Xue 2005; Xue and Martin 2006; Xue et al. 2006; Jung et al. 2008; Tong and Xue 2008; Aksoy et al. 2009, 2010; Dowell et al. 2011; Snook et al. 2011; Jung et al. 2012; Snook et al. 2012; Stensrud et al. 2013; Snook et al. 2015). In Tong and Xue (2005), EnKF was shown for the first time to be able to successfully retrieve from radar reflectivity data multiple microphysical species associated with a complex ice microphysics scheme for a simulated storm. Significant positive impact is found when both precipitation and clear-air reflectivity are assimilated together. Jung et al. (2012) assimilated both reflectivity and radial velocity data into the ARPS EnKF system coupled with a two-moment microphysics scheme and

obtained analyses of polarimetric radar signatures in a supercell storm that reasonably matched real observations.

Snook et al. (2011, 2012) successfully demonstrated the ability of EnKF on producing skillful analyses and corresponding forecasts for a tornadic mesoscale convective system by assimilating real radar observations. Dowell et al. (2011) found that assimilating reflectivity data could improve the analysis of hydrometeor fields for a supercell storm case. However, bias errors in the background hydrometeor fields could be projected onto other model variables via the ensemble background error covariances, and led to unreliable temperature and cold pool analyses. More recent studies applying EnKF to radar DA for real cases include, e.g., Snook et al. (2015) and Supinie et al. (2016). Zhang et al. (2011) and Stensrud et al. (2013) review some of the progresses in recent years.

Although flow-dependent background error covariances can be approximated from ensemble forecasts in EnKF, the estimated covariance matrix is usually rank deficient because of the much smaller ensemble size compared to the degrees of freedom of NWP models (Houtekamer and Mitchell 1998; Hamill et al. 2000). Using much larger ensembles can help alleviate the problem but the computational cost can become prohibitively high. Covariance localization is commonly used to alleviate the rank-deficient problem (Hamill et al. 2001), but it has its own issues, such as preventing the use of distant correlations that are physically meaningful, or introducing imbalance into the background error cross-correlations (Lorenc 2003; Houtekamer and Mitchell 2005; Kepert 2009; Greybush et al. 2011).

An alternative approach that can help alleviate the rank deficiency problem is the so-called hybrid method; in which a weighted average of ensemble-derived flow-dependent background error covariance and the typical static flow-independent variational background error covariance is used. This strategy was first proposed by Hamill and Snyder (2000) while Lorenc (2003) proposed a computationally efficient implementation through introducing a set of extended control variables preconditioned on the ensemble covariance.

The advantages of the hybrid algorithms over traditional EnKF and 3DVar were first demonstrated using simulated observations and for simple model (Hamill and Snyder 2000; Etherton and Bishop 2004; Wang et al. 2007b), and the hybrid algorithm generally has more benefit when the ensemble size is small. The applications of hybrid algorithms to real cases are mostly at the large scales (Buehner et al. 2010a, 2010b; Buehner et al. 2013; Clayton et al. 2013) and mesoscales (Li et al. 2012; Zhang et al. 2013; Pan et al. 2014). For conventional observations and month-long experiments at a mesoscale resolution, Zhang et al. (2013) showed that 12–72 h forecasts initialized with a hybrid ensemble-3DVar (En3DVar) system had noticeably smaller root-mean-square errors (RMSEs) as opposed to EnKF. For a 40-km regional forecast system, Pan et al. (2014) compared the performance of the hybrid En3DVar with 3DVar and EnKF based on the NCEP operational GSI (Grid-point Statistical Interpolation) framework, and found better performance of the hybrid method than EnKF in most aspects of the ensuing forecasts and in all aspects than 3DVar. Li et al. (2012) compared hybrid DA to 3DVar for assimilating radial velocity data from coastal radars for a landfalling hurricane. Clear advantages were found of the hybrid method over 3DVar, although no

benefit was found to include any of the static covariance in the hybrid algorithm. Therefore in this particular study the hybrid algorithm performed the best when it was a pure En3DVar algorithm (when using 100% ensemble-derived covariance).

Compared to the global scale and mesoscale, the development and testing of hybrid ensemble-variational (EnVar) algorithm for the convective scales and for radar data assimilation have been more limited. Gao et al. (2013) first reported preliminary results comparing hybrid En3DVar, 3DVar, and EnKF for assimilating simulated radar data for a supercell storm. EnKF was found to outperform hybrid En3DVar with equal weights assigned to the static and ensemble covariances for the analyzed dynamic variables but underperform hybrid En3DVar for hydrometeor variables when assimilating data from single radar. When assimilating data from two radars, the hybrid method produced the best analyses for most model variables. Gao et al. (2004) further examined the dependency of the relative performance of hybrid En3DVar for different ensemble sizes and covariance weights, and they found that smaller ensemble sizes would benefit from a higher weight for the static covariance, consistent with earlier findings with larger scale applications. In Gao et al. (2016), similar to Li et al. (2012), an ensemble of the 3D variational DA approach was taken in which the En3DVar system is run multiple times to provide the ensemble perturbations rather than running a parallel EnKF system. The sensitivities of supercell analyses to the inclusion of a mass continuity constraint, microphysics errors, and reflectivity assimilation were examined, again in an observing system simulation experiments (OSSE) framework. The assimilation of reflectivity data was found to accelerate storm spin-up and have a small positive impact on wind analyses.

1.1.2 Motivation

While interesting results have been obtained with the previous studies through developing and testing hybrid algorithms for radar DA, further improvements and investigations are still needed. For example, the EnKF system used in Gao et al. (2013, 2014) was an experimental version of EnKF DA, and the 3DVar used constant background error variances for all state variables and empirical spatial correlation scales. If the background error covariances are further optimized for 3DVar and EnKF is optimally tuned in terms of covariance localization and inflation, would the relative performance of the 3DVar, EnKF and En3DVar change? Also, in principle, when En3DVar uses 100% ensemble-derived covariance, its analysis should be identical to that of EnKF under linearity and Gaussian error assumptions; will their analyses actually be very close? If not, what are the sources of differences? In what situations that the static covariance in the hybrid algorithm does help, if at all, for convective storms? Do the conclusions change when model errors are included in the OSSEs? Does hybrid En3DVar have advantages over the other methods for a real storm case?

In this study, the above questions will be investigated via perfect-model OSSE, imperfect-model OSSEs, and a real case experiments. In the OSSEs, the truth is known, enabling a quantitative assessment of different algorithms. Based on the OSSEs, any difference between EnKF and En3DVar algorithms can be easily investigated, and the potential benefits of hybrid En3DVar over traditional 3DVar and EnKF can be better understood. OSSEs are therefore necessary before applying the algorithms to real data cases where many possible, unknown sources of error can make understanding of the algorithms difficult.

1.1.3 Hypothesis and summary of study

Considering hybrid En3DVar has some advantages over 3DVar and EnKF for large scale and mesoscale DA applications in the literature. This dissertation evaluates the hybrid Ensemble-3DVar (En3DVar) DA scheme and compares it with 3DVar and EnKF for convective-scale radar DA. The study is conducted based on the following hypothesis*:* ***hybrid En3DVar radar data assimilation can produce better analyses and short-range (e.g., 1-hr) forecasts than stand alone 3DVar and EnKF for convective scale DA.***

The hybrid En3DVar algorithm is implemented within the ARPS 3DVar framework (Gao et al. 2004) based on the Lorenc (2003) extended control variable approach. Full ensemble covariance localization is implemented in all three directions via the correlation matrix in the extended control variable term of the cost function. The En3DVar system is coupled with a mature EnKF DA system that has been developed and tested for radar DA over the past decade at the Center for Analysis and Prediction of Storms (CAPS) to form a coupled EnKF-hybrid En3DVar system. The static background error covariances for the hydrometeors adopt temperature-dependent vertical profiles recently proposed to improve variational analyses of hydrometeors from reflectivity observations. To facilitate the most fair and direct comparison between EnKF and pure En3DVar, we formulate an alternative EnKF algorithm in which an additional deterministic forecast is produced each cycle that is updated in the same manner as the ensemble mean background in the EnKF, and we call this algorithm DfEnKF (because of the use of deterministic forecast), and DfEnKF will be directly compared with pure En3DVar. In OSSEs, sensitivity experiments are conducted to

obtain the optimal localization radii, optimal background error decorrelation scales, and optimal hybrid weights for EnKF/pure En3DVar, 3DVar, and hybrid En3DVar algorithms, respectively, so that different algorithms are fairly compared. We aim to answer some of the questions posed in the earlier section.

1.1.4 Uniqueness

This study has several unique aspects:

- For the first time, optimal configurations are obtained for 3DVar, EnKF, pure and hybrid En3DVar before they are compared, so that the comparisons among different algorithms can be more meaningful. The optimal configurations include optimal background error decorrelation scales for 3DVar, optimal localization radii for EnKF and pure En3DVar, and the optimal hybrid weights for hybrid En3DVar.
- A new DfEnKF algorithm that updates a single deterministic background field using the mean updating equation of EnKF is introduced so that EnKF is directly parallel to pure En3DVar (that uses 100% ensemble covariance).
- For the first time, hybrid En3DVar is compared with both 3DVar and EnKF using the imperfect-model OSSEs for convective-scale radar DA, which can provide more robust conclusions on the relative performance of various algorithms.
- For the first time, hybrid En3DVar is compared with both 3DVar and EnKF for a real convective storm case.

1.2 Dissertation Outline

The rest of this dissertation is organized as follows. In Chapter 2, we briefly introduce the theoretical background of current hybrid En3DVar algorithms. In Chapter 3, hybrid En3DVar for radar DA is compared with 3DVar, EnKF, and pure En3DVar using perfect-model OSSEs for a simulated supercell storm. In Chapter 4, model error is introduced into the OSSEs by using a different microphysical scheme in the truth run (Lin scheme) (Lin et al. 1983) from that in the forecasts (Hong and Lim 2006) within the DA cycles. Hybrid En3DVar is compared with 3DVar, EnKF, and pure En3DVar to check whether the conclusions obtained in Chapter 3 still hold. In Chapter 5, the hybrid En3DVar and other algorithms are applied to the May 10, 2010 Oklahoma tornado outbreak case and performances of the algorithms are compared. Chapter 6 summarizes the results of this dissertation research and discusses additional issues that can be examined in the future.

Chapter 2 The Hybrid Ensemble-Variational (EnVar) DA Scheme

2.1 Theoretical background of hybrid EnVar

The Hybrid EnVar Data assimilation (DA) scheme is developed based on both the variational and the Ensemble Kalman Filter (EnKF) DA schemes by utilizing a combination of the static and ensemble covariance in calculation of the optimal solution.

2.1.1 Variational data assimilation scheme

The formulation of variational methods was derived based on Bayesian theory and under the assumption of the Gaussian error distributions (Lorenc 1986). Its optimal solution is obtained by minimizing a cost function

$$J(x) = J_B + J_o = \frac{1}{2}\left(x - x_b\right)^T B^{-1}\left(x - x_b\right) + \frac{1}{2}\left[H\left(x\right) - y_o\right]^T R^{-1}\left[H\left(x\right) - y_o\right]. \qquad (2.1)$$

Here $J(x)$ is the cost function, which is the sum of two quadratic terms J_B and J_o. J_B and J_o measure the departure of the analysis $\mathbf{x}$ from the background $\mathbf{x}_b$ and from the observation $\mathbf{y}_o$ respectively. $H(\ )$ is the observation operator that may be nonlinear. **B** and **R** are the static background error covariance and the observation error covariance respectively. After minimization the cost function, the optimal analysis x satisfies

$$\nabla J(x) = \mathrm{B}^{-1}\left(x - x_b\right) + \frac{1}{2}\mathrm{H}^{\mathrm{T}}\mathrm{R}^{-1}\left[H\left(x\right) - y_o\right] = 0. \qquad (2.2)$$

Here **H** is the linearized observation operator, and $\mathbf{H}^{\mathbf{T}}$ is its adjont.

When dealing with the four-dimensional variational DA, the observation operator $H(\ )$ refers to both the observation operator and the forward model so that the departures of the analyses and the observations can be calculated at different times.

Correspondingly, $\mathbf{H}^{\mathrm{T}}$ is the adjoint of both the nonlinear observation operator and the forward model.

2.1.2 Ensemble Kalman filter

The original Kalman filter was developed for linear prediction model based on the Bayesian theory. When using the nonlinear model to advance the state variables, the traditional Kalman filter is adapted to Extended Kalman Filter (EKF) (Evensen 1992). However, it is too expensive to advance the error covariance evolution equation for EKF. As a variant to EKF, the ensemble Kalman filter (EnKF) was first introduced by Evensen (1994) to deal with the high computational cost of integrating the background error covariance. EnKF is a Monte Carlo approximation of the traditional extended Kalman filter (EKF) by estimating the background error covariance based on an ensemble of forecasts (Evensen 1992; Miller et al. 1994). EnKF DA cycle starts with an ensemble of initial first guess fields and mainly includes two repeated steps: the forecast step and the analysis step. At the forecast step, the ensemble models are integrated forward from the analyses until new observations are available, and at the analysis step, each model state is updated based on the following equations:

$$x^{a} = x^{b} + \mathrm{K}\left[\mathrm{y}\text{-}\, H\left(\mathrm{x}^{b}\right)\right], \tag{2.3}$$

$$\mathrm{K} = \hat{\mathrm{P}}^{\mathrm{b}}\mathrm{H}^{\mathrm{T}}\left(\mathrm{H}\hat{\mathrm{P}}^{\mathrm{b}}\mathrm{H}^{\mathrm{T}} + \mathrm{R}\right)^{-1}. \tag{2.4}$$

Here K is the traditional Kalman gain. x_k^b and x_k^a are the ensemble analysis and background for each ensemble member k respectively. y represents the observation, H and H are the nonlinear and tangent linear observation operator. $\hat{\mathrm{P}}^{\mathrm{b}}$ is the ensemble covariance estimated based on the N size of ensemble background field (Evensen 1994):

$$\hat{\mathbf{P}}_{\mathbf{b}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{k=1}^{N}\left(x_b^k - \bar{x}^b\right)\left(H\left(x_b^k\right) - \overline{H\left(x_b\right)}\right)^{\mathrm{T}}, \tag{2.5}$$

$$\mathbf{H}\hat{\mathbf{P}}_{\mathrm{b}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N-1}\sum_{k=1}^{N}\left(H\left(x_b^k\right) - \overline{H\left(x_b\right)}\right)\left(H\left(x_b^k\right) - \overline{H\left(x_b\right)}\right)^{\mathrm{T}}. \tag{2.6}$$

There are mainly two types of EnKF: the stochastic EnKF and the deterministic EnKF. Houtekamer and Mitchell (1998) found that in order for the EnKF to maintain sufficient spread in the ensemble and prevent filter divergence, the observations should be treated as random variables. They introduced the concept of using perturbed sets of observations to update each ensemble member. The perturbed observations consisted of the actual, or "control" observations plus random noise, with the noise randomly sampled from the observational-error distribution used in the data assimilation. Different ensemble members were updated using different sets of perturbed observations. Whitaker and Hamill (2002) put forward another ensemble filter formulation (EnSRF), which uses the traditional Kalman gain for updating the ensemble mean but uses a 'reduced' Kalman gain to update deviations from the ensemble mean. EnSRF is shown to have similar performance as the traditional EnKF that uses the 'perturbed observations' (Whitaker and Hamill 2002) . Other variants of deterministic EnKF that avoids perturbing the observations include Ensemble Adjustment Kalman Filter (EAKF) (Anderson 2001), Ensemble Transform Kalman Filter (ETKF) (Bishop et al. 2001), and Local Ensemble Transform Adjustment Kalman Filter (LETKF) (Hunt et al. 2007; Miyoshi and Yamane 2007). These methods are shown to belong to a broad class of square root filters (Tippett et al. 2003). In this dissertation, only EnSRF is used, whose formulation will be addressed in detail in Chapter 3.

2.1.3 Hybrid EnVar and its variants

There are mainly two types of approaches to realize the hybrid En3DVar Scheme, which are a) the hybrid scheme that directly combine the static and ensemble covariance (Hamill and Synder, 2000) and b) the hybrid scheme developed based on the extended control variables that are preconditioned upon the square root of the ensemble covariance (Lorenc 2003). They are proven to be equivalent in (Wang et al. 2007a)

a. Hybrid EnKF-3DVar scheme using the direct combination of the static and ensemble covariance

In Hamil and Synder (2000), a hybrid EnKF-3DVar analysis scheme was constructed based on a simple model, in which a weighted average of the static background errors from 3DVar ($\mathbf{B}$) and the ensemble based background error covariance from ensemble forecasts ($\mathbf{P}_\mathrm{b}$) are used, which is

$$\mathbf{P_H} = (1-\alpha)\mathbf{P}_\mathrm{b} + \alpha\mathbf{B}^T. \tag{2.7}$$

The sample covariance matrix $\mathbf{P}_b$ derived from the background ensemble forecasts is flow-dependent. The calculation of $\mathbf{P}_b$ for each ensemble member are given as below:

$$\mathbf{P}_b = \frac{1}{N-1}\sum_{k=1}^{N}\left(x_b^k - \bar{x}\right)\left(x_b^k - \bar{x}\right)^{\mathbf{T}}, \tag{2.8}$$

where x_b^k is the ensemble forecasts for N ensemble members (k=1, ...,N), and $\bar{x}$ is the mean of the ensemble forecasts averaged over all the ensemble members. The analysis increments for the hybrid scheme are calculated iteratively by minimization a cost function similar to 3DVar,

$$J(\mathrm{x}) = \frac{1}{2}(\mathrm{x}-\mathrm{x}_b)^{\mathrm{T}}\mathrm{P}_{\mathrm{H}}^{-1}(\mathrm{x}-\mathrm{x}_b) + (H(\mathrm{x})-\mathrm{y})^{\mathrm{T}}\mathrm{R}^{-1}(H(\mathrm{x})-\mathrm{y})\ . \tag{2.9}$$

With hybrid weights α changing from 0.0 to 1.0, the analysis hybrid changes from using stand-alone ensemble-based error covariances to using the stand-alone static background error covariances.

b. Hybrid EnVar scheme based on extended control variables

Lorenc (2003) described how to realize the hybrid scheme within the variational framework by introducing a set of extended control variable that are preconditioned upon the normalized ensemble background perturbations $\mathbf{X}_b'$.

$$\mathbf{x}_{\boldsymbol{a}} = \mathbf{x}_b + \mathbf{X}_b'\boldsymbol{\alpha}, \tag{2.10}$$

where $\boldsymbol{\alpha}_{\boldsymbol{r}}$ is the extended control variable that has the size of N*1. N is the ensemble size.

$$\mathbf{X}_b' = \frac{1}{\sqrt{N-1}}(x_b^1 - \overline{\mathbf{x}_b}, x_b^2 - \overline{\mathbf{x}_b}, \dots, x_b^N - \overline{\mathbf{x}_b}), \tag{2.11}$$

where

$$\overline{\mathbf{x}_b} = \frac{1}{N}(x_b^1 + x_b^2 + \cdots + x_b^N) \tag{2.12}$$

is the ensemble mean.

The optimal analysis is obtained by minimizing a cost function that measures the misfit between observational prior $H(\mathbf{x})$ and observations $\mathbf{y}$:

$$J(\boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha} + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^{\mathrm{T}}\mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}), \tag{2.13}$$

$$H(\mathbf{x}) = H(\mathbf{X}_b + \mathbf{X}_b'\boldsymbol{\alpha}). \tag{2.14}$$

Here R is the observational error covariance and $H(\quad)$ is the nonlinear observational operator. The above $\boldsymbol{\alpha}$ has a size of $N \times 1$, which is usually rank-deficient (i.e., the limited size of ensemble members for covariance estimation is much smaller than the number of degrees of freedom of the model (Kalnay 2002)). In their realization, Shur product was used to reduce the rank-deficient problem.

In order to match the covariance modified by the Shur product, they use as control variable a matrix $\boldsymbol{\alpha}$ whose columns are N fields $\boldsymbol{\alpha_i}$, each with the desired modifying covariance C. The variational problem is then to minimize

$$J(\boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\begin{pmatrix} \mathbf{C} & & 0 \\ & \ddots & \\ 0 & & \mathbf{C} \end{pmatrix}^{-\mathbf{1}} \boldsymbol{\alpha} + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^{\mathrm{T}}\mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}), \qquad (2.15)$$

$$\mathbf{y} = H\left\{\mathbf{x}_b + \mathbf{x}'_b \circ \boldsymbol{\alpha}\begin{pmatrix} \mathbf{1} \\ \vdots \\ \mathbf{1} \end{pmatrix}\right\}. \qquad (2.16)$$

The main advantage of this hybrid approach is that it is easily built on the existing variational framework with the incremental variational method being used. The corresponding cost function is:

$$J(\mathbf{v}, \boldsymbol{\alpha}) = \frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{v} + \frac{1}{2}\boldsymbol{\alpha}^{\mathrm{T}}\begin{pmatrix} \mathbf{C} & & 0 \\ & \ddots & \\ 0 & & \mathbf{C} \end{pmatrix}^{-\mathbf{1}} \boldsymbol{\alpha} + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^{\mathrm{T}}\mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}) \quad . \qquad (2.17)$$

$$\mathbf{y} = H\left\{\mathbf{x}_b + \sqrt{(1-\beta^2)}\mathbf{U}\mathbf{v} + \beta\mathbf{x}'_b \circ \boldsymbol{\alpha}\begin{pmatrix} \mathbf{1} \\ \vdots \\ \mathbf{1} \end{pmatrix}\right\}. \qquad (2.18)$$

where, as noted above, a further horizontal transform of $\boldsymbol{\alpha}$ removed the need to invert $\mathbf{C}$. The terms $1-\beta^2$ and β^2 are included because the two control variables are really alternatives; when adding the $\boldsymbol{\alpha}$ control variable one should reduce the use made of the traditional control variable. The ensemble perturbations can come from another source such as an error breeding cycle (Toth and Kalnay 1997).

As is shown above, the hybrid scheme can be realized either by 1) directly using the weighted average of the original background error covariance and the ensemble covariance or 2) or through modifying the cost function by introducing a set of extended control variables that are preconditioned upon the square root of the ensemble

covariance. In our study, the latter approach is adopted for its easier implementation based on the existing variational framework.

The hybrid method can be easily extended from three dimensions to four dimensions by introducing a forecast step to the variational cost function of the hybrid En3Dvar algorithm, similar to the implementation of 4D-Var. In this dissertation, we will only focus on the hybrid En3DVar scheme.

2.2 The ARPS Hybrid En3DVar implementation

The hybrid En3DVar DA algorithm is implemented within the existing ARPS 3DVar (Gao et al. 2004) framework. The hybrid En3DVar algorithm is based on the extended control variable method of Lorenc (2003) although the actual formulations follow Liu and Xue (2016) more closely. The combined use of the full-rank static background error covariance and the rank-deficient ensemble covariance in En3DVar is achieved through the extended state vector method. The analysis increment is given by

$$\Delta\mathbf{x}^a = \mathbf{x}^a - \mathbf{x}^b = \beta_1\Delta\mathbf{x}_1 + \beta_2\Delta\mathbf{x}_2\,, \tag{2.19}$$

$$\beta_1^2 + \beta_2^2 = 1\,. \tag{2.20}$$

Where $\Delta\mathbf{x}_1$ and $\Delta\mathbf{x}_2$ are the analysis increments related to the static and ensemble background error covariances, respectively. $1/\beta_1^2$ and $1/\beta_2^2$ are the weights given to the static and ensemble covariances, respectively.

$$\Delta\mathbf{x}_1 = \mathbf{U}\mathbf{v}\,, \tag{2.21}$$

$$\mathbf{B} = \mathbf{U}\mathbf{U}^T\,, \tag{2.22}$$

$$\Delta\mathbf{x}_2 = \sum_{i=1}^{N}\left(\mathbf{x}'_{bi}\circ\boldsymbol{\alpha}_i\right), \tag{2.23}$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_N \end{bmatrix}. \tag{2.24}$$

$\mathbf{v}$ is the traditional variational control variable (vector) preconditioned on the square root of the 3DVar static background error covariance $\mathbf{B}$. $\mathbf{U}$ is the square root of $\mathbf{B}$. $\mathbf{x}'_b$ is a matrix of ensemble perturbations normalized by $\sqrt{N-1}$, which has N state vectors, $\mathbf{x}'_{bi} = \mathbf{x}_{bi} - \overline{\mathbf{x}}^b$ whose dimension is n. N and n are the ensemble size and the dimension of state vector respectively. $\circ$ denotes the Schur product.

$$\boldsymbol{\alpha} = \begin{bmatrix} \mathbf{C}' & & \\ & \ddots & \\ & & \mathbf{C}' \end{bmatrix} \tilde{\boldsymbol{\alpha}}. \tag{2.25}$$

Here $\tilde{\boldsymbol{\alpha}}$ is a new control vector related to the extended control vector $\boldsymbol{\alpha}$, a $n \times n$ correlation matrix $\mathbf{C}$ is used to localize the ensemble covariance, and $\mathbf{C}'$ is the corresponding decomposed matrix that satisfies $\mathbf{C} = \mathbf{C}'\mathbf{C}'^T$. A hybrid solution is derived by setting β_1 or β_2 to values between but not equal to 0 and 1.

The analysis increment in Eq. (10) is obtained via minimizing the following cost function

$$\begin{aligned} J(\mathbf{v}, \tilde{\boldsymbol{\alpha}}) = & \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\tilde{\boldsymbol{\alpha}}^T\tilde{\boldsymbol{\alpha}} + \\ & \frac{1}{2}\left(\beta_1\mathbf{HUv} + \beta_2\mathbf{H}\sum_{i=1}^{N}\left(\mathbf{x}'_{bi} \circ \mathbf{C}'\tilde{\boldsymbol{\alpha}}_i\right) + \mathbf{d}\right)^T \mathbf{R}^{-1}\left(\beta_1\mathbf{HUv} + \beta_2\mathbf{H}\sum_{i=1}^{N}\left(\mathbf{x}'_{bi} \circ \mathbf{C}'\tilde{\boldsymbol{\alpha}}_i\right) + \mathbf{d}\right), \end{aligned} \tag{2.26}$$

where $\mathbf{d}$ is the observation innovation vector, $\mathbf{H}$ is the linearized observational operator, and $\mathbf{R}$ is the observational error covariance. Additional details can be found in Liu and Xue (2016).

In this study, as is typically done, the En3DVar is coupled with EnKF to form a coupled EnKF-En3DVar system, where the EnKF system provides ensemble perturbations to En3DVar for flow-dependent covariance, while En3DVar itself updates a single background forecast to obtain a single new analysis. With two-way coupling, the analysis of En3DVar is used to replace the ensemble mean analysis of EnKF, i.e., $\mathbf{x}^a$ from Eq. (10) is used to replace $\overline{\mathbf{x}}^a$ in Eq. (1). As mentioned earlier, in this paper, we choose to stick with one-way coupling to keep the algorithms more separate (so that they are not too mingled) for comparison purposes (for the relatively short period of DA cycles presented in this paper, the two coupling systems generally will not diverge much; for long periods of DA cycles, two-way coupling is recommended). When the ensemble-derived covariance is used at 100% without static covariance, the En3DVar is referred to as pure En3DVar. When static and ensemble covariances are used in combination, the algorithm is referred to as hybrid En3DVar. The readers are referred to e.g., Liu and Xue (2016) for discussions on related terminology.

2.3 Advantages and disadvantages of hybrid EnVar

2.3.1 Advantages of hybrid DA schemes

Hybrid DA schemes has a number of potential advantages than the stand-alone 3DVar, EnKF, and pure En3DVar:

1) Relative to pure 3DVar scheme, hybrid En3DVar could take into account the flow-dependent background error covariance, better than the static background error covariance when the ensemble background error covariance is a good estimation of the true error covariance.

2) Pure ensemble-based methods (EnKF and pure En3DVar) are subject to model errors so that the ensemble spread tends to be systematically underestimated and results in filter divergence. For example, the EnKF may fail to capture the storm systems in the analysis when storms in the background are missed and at the same time the ensemble spread is too small because of systematic model errors. Blending the ensemble covariance with the 3DVar statistics may ameliorate the sampling error problems relative to using stand-alone ensemble covariance.
3) The hybrid weight of the static and ensemble covariance can be adjusted based on reliabilities of ensemble and static background error covariance.
4) It is easily to add weak constraint (such as the mass-continuity constraint) in to hybrid DA cost function relative to EnKF.

2.3.2 Disadvantages of hybrid DA schemes

Considering hybrid DA using both the ensemble and the static covariance, it is subject to the disadvantages of both EnKF and 3DVar, relying on the hybrid weight specified. Overall, the disadvantages can be summarized as follows:

1) When hybrid weight of the static error covariance is too small or not big enough, the hybrid analysis may be subject to the filter divergence problem when severe model error exists.
2) When hybrid weight of ensemble part is too small, the analysis increment is not flow-dependent, similar to the problem of 3DVar.
3) An optimal weight needs to be determined to combine the static and the ensemble background error covariance.

4) The computational cost of hybrid En3DVar is more expensive than EnKF and orders of magnitude more expensive than 3DVar. Unlike EnKF, for which the localization is conducted in the observation space, the En3DVar needs the recursive filter to conduct the localization, which is more expensive.

2.4 Special treatments of ensemble covariance in implementation of hybrid DA

As mentioned above, hybrid En3DVar is prone to all the problems of EnKF, especially when ensemble weight of static background is small. The methods used in the implementation of EnKF: covariance inflation and covariance localization are also used in implementation of hybrid En3DVar to achieve better performance.

2.4.1 Covariance inflation

Similar to the implementation of EnKF, inflation algorithms that modify the prior or posterior estimates of the ensemble variance to reduce filter error and avoid filter divergence are also used on ensemble covariance of hybrid En3DVar.

There are mainly four types of methods to inflate the covariance:

1) The multiplicative inflation

In multiplicative inflation, the ensemble deviation from the ensemble mean is multiplied by a constant. As a result, a constant increase of the ensemble variance is made. The inflation can be made before or after filtering (calling prior or posterior inflation). A 20% multiplicative inflation (with an inflation coefficient of 1.2) is used in OSSEs in this dissertation for its simplicity.

2) Additive inflation

In order to simulate model errors, random noises are added to the background (prior) or analysis (posterior) ensembles so that the diagonal entries of the ensemble

covariance can be increased. Usually random noise is smoothed first to make it less noisy. By applying additive inflation, new directions are introduced into the error space so that the ensemble spread can be increased. However, it is difficult to obtain reasonable additive noise.

3) Covariance relaxation inflation

Zhang et al. (2004) proposed another covariance inflation method that relaxes posterior (analysis) perturbations back toward the prior (first guess) perturbations independently at each point through

$$x_i'^a \leftarrow (1-\alpha)x_i'^a + \alpha x_i'^b, \tag{2.27}$$

where $x_i'^a$ is the deviation from the posterior ensemble mean for the ith ensemble member, and $x_i'^b$ is the deviation from the prior ensemble mean for the ith ensemble member. Whitaker and Hamill (2012) referred this method as relaxation-to-prior perturbations (RTPP). Unlike the multiplicative inflation, relaxation inflation tries to increase the posterior ensemble variance in proportion to the amount of the prior variance that is reduced by assimilation of observations. A tunable parameter α is used to control the amount of weight given to the prior ensemble. When a is 0, only posterior ensemble is used; when a is 1.0, the posterior ensemble is completely replaced by the prior ensemble; for values in between, part of the posterior ensemble is replaced by the prior ensemble.

4) Adaptive covariance inflation

Whitaker and Hamill (2012) proposed an alternative approach to RTPP (they call it relaxation-to-prior (RTPS)) by relaxing the ensemble standard deviation back to

the prior instead of relaxing the posterior perturbations back to their prior values at each grid point as in RTPP. The perturbations are relaxed via,

$$\sigma^{a} \leftarrow (1-\alpha)\,\sigma^{a} + \alpha\,\sigma^{b}, \tag{2.28}$$

where $\sigma^{b} = \sqrt{[1/(n-1)]\sum_{i=1}^{n} x_{bi}^{\prime\,2}}$ and $\sigma^{a} = \sqrt{[1/(n-1)]\sum_{i=1}^{n} x_{ai}^{\prime\,2}}$ are the prior and posterior ensemble standard deviation (spread) at each analysis grid point, and n is the ensemble size. This formula can be rewritten as:

$$x_{i}^{\prime\,a} \leftarrow x_{i}^{\prime}\ \alpha\left(\alpha\,\frac{\sigma^{b}-\sigma^{a}}{\sigma^{a}}+1\right). \tag{2.29}$$

For a given value of α, the multiplicative inflation in RTPS is proportional to the amount the ensemble spread reduced by assimilation of the observations.

2.4.2 Covariance localization

Houtekamer and Mitchell (1998) show that the background error correlations between observation and distant analysis points usually exhibit a large amount of structure that does not have physical meaning with limited ensemble size. Without correction, an observation would produce anomalous analysis increments on distant state. Covariance localization is a modification of the covariance matrices that removes spurious long distance correlations due to sampling error of the model covariance from finite ensemble size (Houtekamer and Mitchell, 2001; Anderson, 2007).

Unlike EnKF, which uses a fifth-order polynomial (Gaspri and Cohn, 1999) to conduct the localization in most applications, En3DVar uses the recursive filter to localize the background error ensemble covariance. It re-uses the one or more of its output as an input to mimic a Gaussian convolution operator.

The localization of EnKF is an element-wise multiplication of the ensemble estimates (a) with a smooth correlation function via

$$\widehat{K} = \left(\rho_s \circ \widehat{P^b}\right) H^T \left[\left(\rho_s \circ \widehat{P^b}\right) H^T + R\right]^{-1}. \tag{2.30}$$

It can be approximately expressed as,

$$\widehat{K} = \left(\rho_s \circ \widehat{P^b}\right) \left[\rho_s \circ \left(H^T \widehat{P^b H^T}\right) + R\right]^{-1}. \tag{2.31}$$

$\widehat{P^b}$ is the ensemble covariance estimated from the limited ensemble perturbations. ρ_s is the localization function.

$$\rho_s = \begin{cases} -\frac{1}{4}\left(\frac{d}{c}\right)^5 + \frac{1}{2}\left(\frac{d}{c}\right)^4 + \frac{5}{8}\left(\frac{d}{c}\right)^3 - \frac{5}{3}\left(\frac{d}{c}\right)^2 + 1 & (0 \le \lceil d \rceil \le c) \\ \frac{1}{12}\left(\frac{d}{c}\right)^5 - \frac{1}{2}\left(\frac{d}{c}\right)^4 + \frac{5}{8}\left(\frac{d}{c}\right)^3 + \frac{5}{3}\left(\frac{d}{c}\right)^2 - 5\left(\frac{d}{c}\right) + 4 - \frac{2}{3\left(\frac{d}{c}\right)} & (0 \le \lceil d \rceil \le c) \\ 0 \quad \lceil d \rceil \ge 2c \end{cases}, \tag{2.32}$$

where d is the distance from observation to grid points, and c is the horizontal cutoff radius.

The recursive filter length scale (S_{RF}) of En3DVar and the cutoff radius of EnKF localization (c) can be roughly converted from each other based on Eq. (4) in Pan et al. (2013), which is

$$S_{RF} = \sqrt{0.15}c/\sqrt{2}. \tag{2.33}$$

Chapter 3 Evaluation of a Hybrid En3DVar Radar Data Assimilation System and Comparisons with 3DVar and EnKF for Radar Data Assimilation with Observing System Simulation Experiments under Perfect Model Assumptions[1]

3.1 Introduction

Convective scale NWP is vital to provide timely and accurate forecasts and watches for severe thunderstorms and tornadoes. Convective scale weather systems develop on very short timescales (typically from minutes to hours), making the analysis and forecasts more challenging than larger scale weather systems. Radar data that provides the wind and precipitation information with very high spatial (~1km) and temporal (5~10 min) resolution is very useful to extract the wind and precipitation information from the observations and merge into the analysis to improve the initial condition for NWP forecasts.

3DVar have been used for storm scale radar DA and was found to be able to improve the analysis and forecast of storm evolution and short-range QPF (Gao et al. 2004; Lindskog et al. 2004; Hu et al. 2006a; Hu et al. 2006b; Xiao et al. (2007)). However, it is commonly believed that 3DVAR is not suitable for convective scale DA because the static background error covariance fails to reflect the convective-scale balance.

Since last decade, the application of the EnKF to storm scale radar DA has become popular (e.g., Snyder and Zhang 2003; Dowell et al. 2004; Zhang et al. 2004;

[1] This chapter is accepted as: Kong, R., Xue, M., and Liu, C., 2017: Development of a Hybrid En3DVar Data Assimilation System and Comparisons with 3DVar and EnKF for Radar Data Assimilation with Observing System Simulation Experiments. *Mon. Wea. Rev.*.

Tong and Xue 2005, 2008; Xue et al. 2006; Aksoy et al. 2009, 2010; Dowell et al. 2011; Jung et al. 2008, 2012; Snook et al. 2011, 2012, 2015; Dowell et al. 2011; Wang et al. 2014). EnKF was found to help producing more skillful analyses and forecasts and bring positive impact to the analysis of hydrometeor fields. However, EnKF is usually rank-deficient when ensemble size is small, and is subject to filter divergence problem when there exists severe sampling or model error.

Recently, hybrid schemes (e.g., En3DVar, En4DVar, and 4DEnVar) that blend the static background error covariance with the flow-dependent ensemble covariance have been applied to radar DA and were found to be able to generate better analysis than both 3DVar and EnKF for smaller sized ensembles (Hamill and Snyder 2000; Lorenc 2003; Etherton and Bishop 2004; Wang et al. 2007b; Wang et al. 2007a). However, most of the studies are conducted using simple models and simulated observations. Applications of hybrid DA to real cases are mostly at the large scales (Buehner et al. 2010a, 2010b; Buehner et al. 2013; Clayton et al. 2013) and mesoscales (Li et al. 2012; Zhang et al. 2013; Pan et al. 2014).

For convective scale radar DA, the development and testing of hybrid ensemble-variational (EnVar) algorithm have been more limited. In Gao et al. (2013), hybrid En3DVar data assimilation scheme is found to help reducing the storm spin-up time. Comparisons made among 3DVar, 40-member EnKF, and corresponding hybrid En3DVar indicate, EnKF method outperform hybrid method for the model dynamic variables, and underperform hybrid method for hydrometeor related variables when assimilating data from single radar. When assimilating data from two radars, the RMSE for the hybrid method are the smallest for most of the model variables. Gao et al. (2014)

indicates, hybrid analysis is benefit from a weaker weight for a relative smaller ensemble size, and is benefit from a stronger weight for larger ensemble size when assimilating two radars at the same time. In Gao et al. (2016), a mass continuity equation is included into the cost function of the hybrid scheme as a weak constraint, the analyses is improved when radial velocity observations contain large errors. Sensitivity experiments indicate that the En3DVar analyses are sensitive to different microphysics schemes, suggesting that multiple-microphysical ensembles could be used to reduce the uncertainty of the model physical processes.

While interesting results have been obtained with the above studies through developing and testing hybrid algorithms for radar DA, further improvements and investigations are still needed. For example, the EnKF system used in Gao et al. (2013, 2014) was an experimental version of EnKF DA, and the 3DVar used constant background error variances for all state variables and empirical spatial correlation scales. If the background error covariances are further optimized for 3DVar and EnKF is optimally tuned in terms of covariance localization and inflation, would the relative performance of the 3DVar, EnKF and En3DVar change? Also, in principle, when En3DVar uses 100% ensemble-derived covariance, its analysis should be identical to that of EnKF under linearity and Gaussian error assumptions; will their analyses actually be very close? If not, what are the sources of differences? In what situations that the static covariance in the hybrid algorithm does help, if at all, for convective storms?

In this study, the above questions will also be addressed in an OSSE framework, under the assumption of a perfect prediction model. In the OSSEs, the truth is known,

enabling a quantitative assessment of different algorithms. Based on the OSSEs, any difference between EnKF and En3DVar algorithms can be more easily investigated. This is necessary before applying hybrid DA to the real cases where many possible, unknown sources of error make understanding of the results difficult. The hybrid En3DVar algorithm based on the Lorenc (2003) extended control variable approach is implemented within the ARPS 3DVar framework (Gao et al. 2004). Full ensemble covariance localization is implemented in all three directions via the correlation matrix in the extended control variable term in the cost function. The En3DVar system is coupled with a mature EnKF DA system that has been developed and tested for radar data DA over the past decade to form a coupled EnKF-hybrid En3DVar system. The static background error covariances for the hydrometeors adopt temperature-dependent vertical profiles we recently proposed to improve variational analyses of hydrometeors from reflectivity observations. To facilitate the most fair and direct comparison between EnKF and pure En3DVar, we formulate an alternative EnKF algorithm in which an additional deterministic forecast is produced each cycle which is updated in the same manner as the ensemble mean background in the EnKF, and we call this algorithm DfEnKF (because of the use of deterministic forecast), and DfEnKF will be directly compared with pure En3DVar. We aim to answer some of the questions posed in the earlier paragraph.

3.2 Assimilating System and Experimental Design

3.2.1. The EnKF system

The EnKF system used in the following study is the one initially developed for the ARPS model, as originally described in Tong and Xue (2005) and Xue et al. (2006). The ensemble squall root (EnSRF) algorithm, a flavor of EnKF without perturbing observations as developed by Whitaker and Hamill (2002), as available in the ARPS EnKF system is used. Following Whitaker and Hamill (2002) and Xue et al. (2006), the EnSRF analysis equations are briefly described in the following, which consist of update equations for the ensemble mean, and for ensemble perturbations.

The ensemble mean forecast is first updated by Eq. (3.1) below:

$$\overline{\mathbf{x}}^{a} = \overline{\mathbf{x}}^{b} + \mathbf{K}\left[\mathbf{y} - \overline{H\left(\mathbf{x}^{b}\right)}\right], \tag{3.1}$$

where

$$\mathbf{K} = \left[\rho_s \circ \left(\mathbf{P}^{b}\mathbf{H}^{\mathbf{T}}\right)\right]\left(\mathbf{H}\mathbf{P}^{b}\mathbf{H}^{\mathbf{T}} + \mathbf{R}\right)^{-1}, \tag{3.2}$$

$$\mathbf{P}^{b}\mathbf{H}^{\mathbf{T}} = \frac{1}{N-1}\sum_{k=1}^{N}\left[\mathbf{x}_k^{b} - \overline{\mathbf{x}}^{b}\right]\left[H(\mathbf{x}_k^{b}) - \overline{H(\mathbf{x}^{b})}\right]^{T}, \tag{3.3}$$

$$\mathbf{H}\mathbf{P}^{b}\mathbf{H}^{\mathbf{T}} = \frac{1}{N-1}\sum_{k=1}^{N}\left[H(\mathbf{x}_k^{b}) - \overline{H(\mathbf{x}^{b})}\right]\left[H(\mathbf{x}_k^{b}) - \overline{H(\mathbf{x}^{b})}\right]^{T}. \tag{3.4}$$

In the equations, $\overline{\mathbf{x}}^{a}$ and $\overline{\mathbf{x}}^{b}$ are the ensemble mean analysis and background forecast state vectors, respectively and $\mathbf{y}$ is the observation vector.

$$\overline{\mathbf{x}}^{a,\,b} = \frac{1}{N-1}\sum_{k=1}^{N}\mathbf{x}_k^{a,\,b}, \tag{3.5}$$

$$\overline{H\left(\mathbf{x}^{a,\,b}\right)} = \frac{1}{N-1}\sum_{k=1}^{N}H\left(\mathbf{x}_k^{a,\,b}\right). \tag{3.6}$$

H is the observation operator (which can be nonlinear) and $\mathbf{H}$ is tangent linear observation operator. $\mathbf{K}$ is the Kalman gain. $\mathbf{P}^b$ is the background error covariance. $\mathbf{P}^b\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}^b\mathbf{H}^T$ are the background error covariances defined between the grid points and observation points, and between the observation points, respectively, and are evaluated directly from the ensemble background states $\mathbf{x}_k^b$ (k is the index of ensemble member) and their observation counterparts, $H(\mathbf{x}_k^b)$, called observation priors. $\mathbf{R}$ is the observation error covariance matrix. A distance-dependent localization function (Gaspari and Cohn 1999) is used for the localization of the ensemble covariance; $\rho_s \circ$ denotes a Schur product of the correlation matrix ρ_s.

The ensemble perturbations, denoted by superscript prime, are updated by

$$\mathbf{x}_k'^a = \mathbf{x}_k'^b - \tilde{\mathbf{K}} H\left(\mathbf{x}_k^b\right)', \tag{3.7}$$

and

$$H\left(\mathbf{x}_k^b\right)' = H\left(\mathbf{x}_k^b\right) - \overline{H\left(\mathbf{x}^b\right)}. \tag{3.8}$$

$\tilde{\mathbf{K}}$ is the "reduced" Kalman gain that is equal to $\mathbf{K}$ amplified by a factor following Whitaker and Hamill (2002) and the actual formula can be found in Xue et al. (2006) also. The final analyses for the ensemble members are

$$\mathbf{x}_k^a = \overline{\mathbf{x}}^a + \mathbf{x}_k'^a. \tag{3.9}$$

For more detailed descriptions about EnKF algorithm, please refer to Xue et al. (2006).

3.2.2. The DfEnKF algorithm

For the convective scale, the ensemble forecast mean tends to smooth out detailed convective scale structures that are important. Therefore the mean of ensemble forecasts may not necessarily be the best representation of convective storms(Yussouf et al. 2013). In En3DVar cycles, the background forecast is a deterministic forecast starting from the En3DVar analysis of the previous cycle. To facilitate more direct comparison of the EnKF results with those of En3DVar, and to potentially improve the results of EnKF, we developed an alternative implementation of EnKF, in which a separate set of analysis and forecast is run that uses a deterministic forecast, $\mathbf{x}_d^b$, in place of the ensemble mean background forecast, $\overline{\mathbf{x}}^b$, in the ensemble mean update Eq. (3.1), to obtained analysis $\mathbf{x}_d^a$:

$$\mathbf{x}_d^a = \mathbf{x}_d^b + \mathbf{K}\left[\mathbf{y} - H\left(\mathbf{x}_d^b\right)\right]. \tag{3.10}$$

The deterministic forecast is started from this version of EnKF analysis, and we call this EnKF formulation DfEnKF, where Df denotes deterministic forecast.

Figure 3.1 shows the flow charts of the DfEnKF and En3DVar DA one-way coupled with EnKF. Similar to En3DVar, DfEnKF can be treated as an independent algorithm that borrows the ensemble covariance from the EnKF, and updates its single background field based on the EnKF update equation (the EnSRF algorithm is used here).

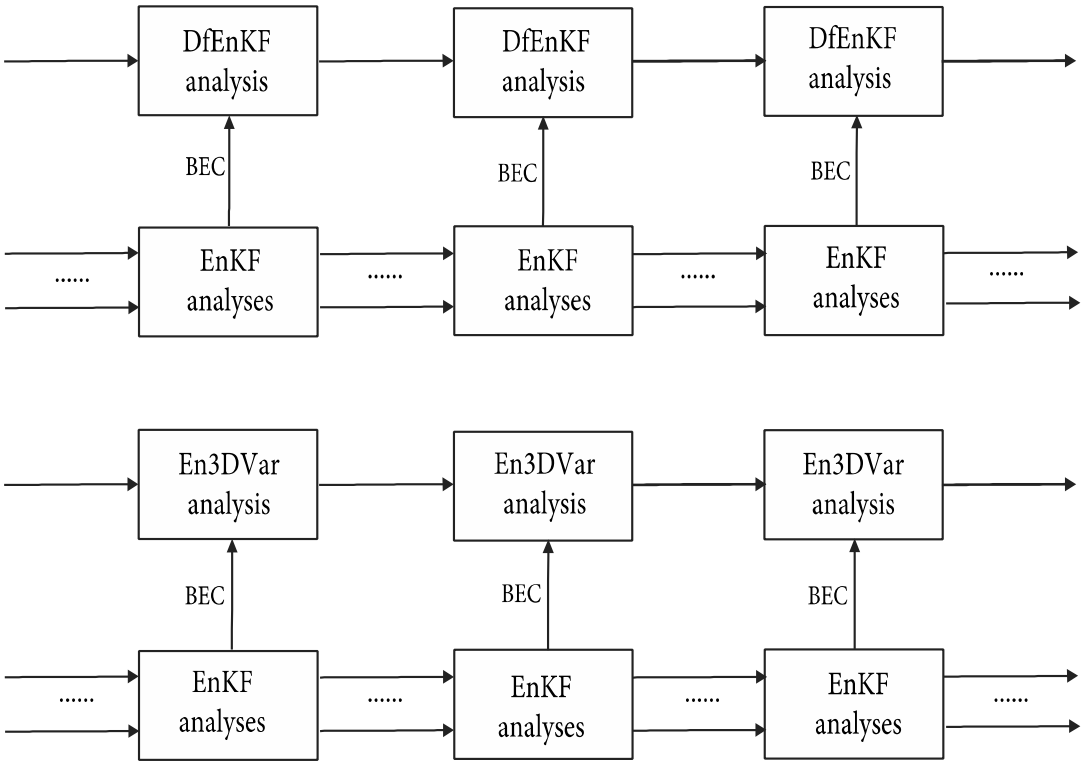

Fig. 3.1 Flow chart of the cycles in one-way coupling between EnKF and DfEnKF, EnKF and En3DVar analysis schemes, where BEC represents the background error covariance. The ensemble perturbations are updated by the EnKF.

It is worth noting that our DfEnKF algorithm is different from the so-called deterministic EnKF (DEnKF) proposed by Sakov and Oke (2008), which modifies the original EnKF perturbation updating equations by using only half of the Kalman gain matrix and avoids perturbing the observations. Given that we use the EnSRF algorithm that does not perturb observations, our DfEnKF algorithm is also a deterministic EnKF algorithm.

We point out here that one may choose to replace the ensemble mean analysis of the regular EnKF with the analysis of DfEnKF so that the ensemble perturbations are effectively defined around the DfEnKF analysis; doing so results in a "two-way-

coupled" DfEnKF and EnKF systems. Similar is done in "two-way-coupled" En3DVar and EnKF, as will be discussed later, but in this paper, we choose one-way coupling in both cases to keep the algorithms more independent so that performance of different algorithms can be more clearly compared.

3.2.3. The prediction model and truth simulation for OSSEs

In this study, the nonhydrostatic ARPS (Xue et al. 2000; Xue et al. 2001; Xue et al. 2003) is used as the prediction model for EnKF and En3DVar DA cycles. The En3DVar algorithm developed is compared with 3DVar and EnKF via perfect-model OSSEs. Simulated radial velocity and reflectivity data from a single Doppler radar are assimilated for a tornadic supercell storm. A sounding derived from that of the 20 May 1977 in Del City, Oklahoma (Ray et al. 1981) supercell storm case is used to define the storm environment. The Lin et al. (1983) ice microphysics parameterized is used.

The model configurations follow closely those of Tong and Xue (2005). Specifically, the model grid size is 35×35×35, and the grid resolutions are 2 km and 0.5 km in the horizontal and vertical directions respectively. The storm in the truth simulation is triggered by an elliptical thermal bubble with a 4 K maximum temperature perturbation, and the bubble is located at 24 km, 16 km, and 1.5 km in x, y, and z directions. The horizontal and vertical radii of the bubble are 10 km and 1.5 km respectively. A constant wind of (3, 14) m s^{-1} is subtracted from the original sounding so that the supercell could remain within the simulation domain. The storm is simulated up to 2 hours with an evolution similar to those documented in Tong and Xue (2005). For EnKF, a distance-dependent localization function following Gaspari and Cohn

(1999) is used to localize the ensemble covariance, while the covariance localization in En3DVar is realized through recursive filters applied to the extended control variable term of the cost function (corresponding to the second term on the right hand side of Eq. (2.26)), to realize similar localization effects in the En3DVar algorithm. When the vertical model grid is stretched, the vertical localization realized through the recursive filter in the En3DVar algorithm is usually asymmetric; to avoid this complication when comparing the EnKF and En3DVar algorithms, we choose to utilize a vertically uniform grid, which is also used in the OSSEs of Tong and Xue (2005).

3.2.4. Simulation of the radar observations

The radial velocity (v_r) and reflectivity (Z) data are simulated based on the model velocity and mixing ratios of hydrometeor fields, respectively, with the following formula:

$$v_r = u\sin\phi\cos\mu + v\cos\phi\cos\mu + w\sin\mu \,, \tag{3.11}$$

$$Z_e = Z_{er}\left(q_r\right) + Z_{es}\left(q_s\right) + Z_{eh}\left(q_h\right) \,, \tag{3.12}$$

$$Z = 10\log_{10}^{Z_e} \,. \tag{3.13}$$

Here, ϕ and μ are the azimuth and elevation angles of radar beams. u, v, and w are the model velocity components in the observational space that are interpolated from the truth simulation grid. Z_e is the equivalent radar reflectivity factor, which contains contributions from the mixing ratios of rain (q_r), snow (q_s), and hail (q_h), respectively. Z (in dBZ) is the reflectivity factor and is the quantity assimilated. The actual reflectivity formula for the individual components follow those given in Tong and Xue (2005). The default values of the intercept parameters of the size distributions of the

hydrometeors in the Lin et al. (1983) are used, as in Tong and Xue (2005). Equations (15) to (17) are also the observation operators for radial velocity and reflectivity data in the DA systems.

One WSR-88D radar is assumed to be located at the origin (i.e., the southwest corner) of the simulation domain, and its data are simulated based on the truth simulation. Unbiased and normal distributed random errors with standard deviations of 3 dBZ and 1 m s^{-1} are added to the simulated reflectivity and radial velocity data, respectively. The same error standard deviations are also specified in the DA. The simulated reflectivity and radial velocity observations are collected on 14 elevation levels ranging from 0.5 to 19.5 degrees based on volume coverage pattern 11 (VCP11) in precipitation mode. The maximum radar coverage is 230 km. Radial velocity data are assimilated only in regions where truth reflectivity is greater than 15 dBZ in the analysis domain, while reflectivity data are assimilated in both precipitation and clear air regions. The assimilation of clear air reflectivity data helps to suppress possible spurious precipitation in clear air regions as found by Tong and Xue (2005). These practices follow Tong and Xue (2005) closely.

3.2.5. Design of assimilation experiments

The simulated storm cell is well developed after 25 min of model integration. As in Tong and Xue (2005), the unbiased and Gaussian-distributed random perturbations are added to the sounding defined first guess field to generate the initial ensemble perturbations. The standard deviations are 2 K for potential temperature, 2 m s^{-1} for u, v, and w, and 0.6 g kg^{-1} for mixing ratios of water vapor and hydrometeor fields,

respectively. Previous studies have shown that adding random perturbations to the whole domain would introduce a lot of noise into the model fields (Snyder and Zhang 2003; Dowell et al. 2004). Different from Tong and Xue (2005), the perturbations of *u, v, w*, and potential temperature θ are confined to the regions with observed reflectivity higher than 0 dBZ, and hydrometeor perturbations are confined to regions with reflectivity higher than 15 dBZ. The radial velocity and reflectivity observations are assimilated every 5 min for one hour, with the first assimilation performed at 30 min of the truth simulation time. Experiments with different combinations of localization radii and recursive filter length scales in horizontal and vertical directions are first tested to obtain the optimal localization radii and recursive filter length scales for EnKF and En3DVar, respectively.

To obtain more physical hydrometeor analyses, height-dependent (or in a more general form temperature-dependent) static background error profiles for different hydrometeors are used in 3DVar, and in hybrid En3DVar (for the static background error). Specifically, the background error of snow at the low levels (< 4 km) is set to values two orders of magnitude smaller than the error at other levels (set to 0.6 g kg^{-1}), which ensures small analysis increments of snow at the lower levels (below freezing levels). The background error of rain-water at higher levels (>8 km) are set to values one order of magnitude smaller than default value of 0.6 g kg^{-1} to help prevent unphysical rainwater analysis increments at the upper levels. For levels between 4 km and 8 km, 1/5 of the default value (i.e, 0.1 g kg^{-1}) is used in this study to suppress analysis increment of rainwater but still allow for the existence of supercool water close

to and above the freezing level. The background error of hail is 0.6 g kg^{-1} at all levels; considering that hail can fall below the freezing level.

We perform five types of experiments that are named according to the methods they use (Table 1). The first three types are 3DVar, (using the stand-alone static background error covariance $\mathbf{B}$), EnKF, and DfEnKF, (using ensemble covariance $\hat{\mathbf{P}}^{b}$). The remaining two types are the pure En3Dvar that uses 100% ensemble covariance and the hybrid En3DVar that uses a combination of the static and ensemble covariance. Performance of DfEnKF and pure En3DVar are compared, and the reasons for differences are investigated though sensitivity experiments. Hybrid En3DVar with different combinations of ensemble members and ensemble covariance weights are tested to see the sensitivity of the hybrid analysis to these factors. Finally, Hybrid En3DVar is compared with 3DVar, EnKF, and pure En3DVar to see if hybrid En3DVar has any advantage over other methods.

Table 3.1 Descriptions of the assimilating methods

Method	Use of background error covariance	Background updating	Ensemble size
3DVar	Static background error covariance	Update background field deterministically	N/A
EnKF	Background error covariance derived from ensemble background forecasts	Updates ensemble mean background and ensemble perturbations using EnKF algorithm	40
DfEnKF	Using ensemble covariance from an EnKF system	Update a single deterministic background forecast using EnKF mean updating algorithm	40
Pure En3DVar	Using 100% ensemble covariance from an EnKF system	Update a single deterministic background forecast using variational algorithm	40
Hybrid En3DVar	Using weighted average of ensemble covariance and static 3DVar covariance	Update a single deterministic background forecast using variational algorithm	10, 20... 100

3.3 Results of assimilation experiments

To evaluate the performance of algorithms in different experiments, we define a mean error that is the average of the RMSEs of both the analysis and 5-min background forecast at all analysis times scaled by the background RMSE at the beginning of the assimilation window (at 25 min). The actual error formulation is

$$\overline{\varepsilon} = \frac{1}{2(N+1)\varepsilon_0} \sum_{t=0}^{N} (\varepsilon_{bt} + \varepsilon_{at}), \tag{3.14}$$

where $\overline{\varepsilon}$ is the mean error. ε_{bt} and ε_{at} are the grid-point average of the RMSEs at time t for a certain variable from the background and the analysis fields, respectively. ε_0 is

the grid-point average of the RMSEs at the beginning of the assimilation window, which are the same for all experiments. N is the total number of assimilation cycles. The RMSEs are calculated only over grid points where the truth reflectivity is higher than 15 dBZ. We refer to the mean error defined in (18) as the mean scaled RMSE for a specific variable. Further average can be performed over all variables to give an overall mean RMSE.

3.3.1. Comparisons between DfEnKF and pure En3DVar

a. Optimal localization radii for DfEnKF and EnKF

In EnKF, a Gaussian-like fifth-order piecewise polynomial of Gaspari and Cohn equation (1999) is used for covariance localization. Sensitivity experiments with different combinations of the horizontal and vertical cut-off radii of localization that range from 1 km to 15 km with a 1 km interval are conducted for DfEnKF. As indicated in Fig. 3.2, the optimal localization radii for different variables are similar except for hail mixing ratio, and on average the combination of 15 km and 6 km for horizontal and vertical cut-off radii yields the best result in terms of the minimal analysis and forecast errors. Similarly, Sobash and Stensrud (2013) found that 12~18 km horizontal and ~3km vertical localization cutoff radii were beneficial for their EnKF experiments with 50 ensemble members and a 3 km grid resolution. On the "optimal" localization radii, some discrepancies between this and previous studies are worth mentioning. For example, Tong and Xue (2005) noted that a localization radius of 6 km produced the best result with 40 ensemble members, but their "optimal" radii were obtained based on the assumption that the horizontal and vertical localization radii were the same.

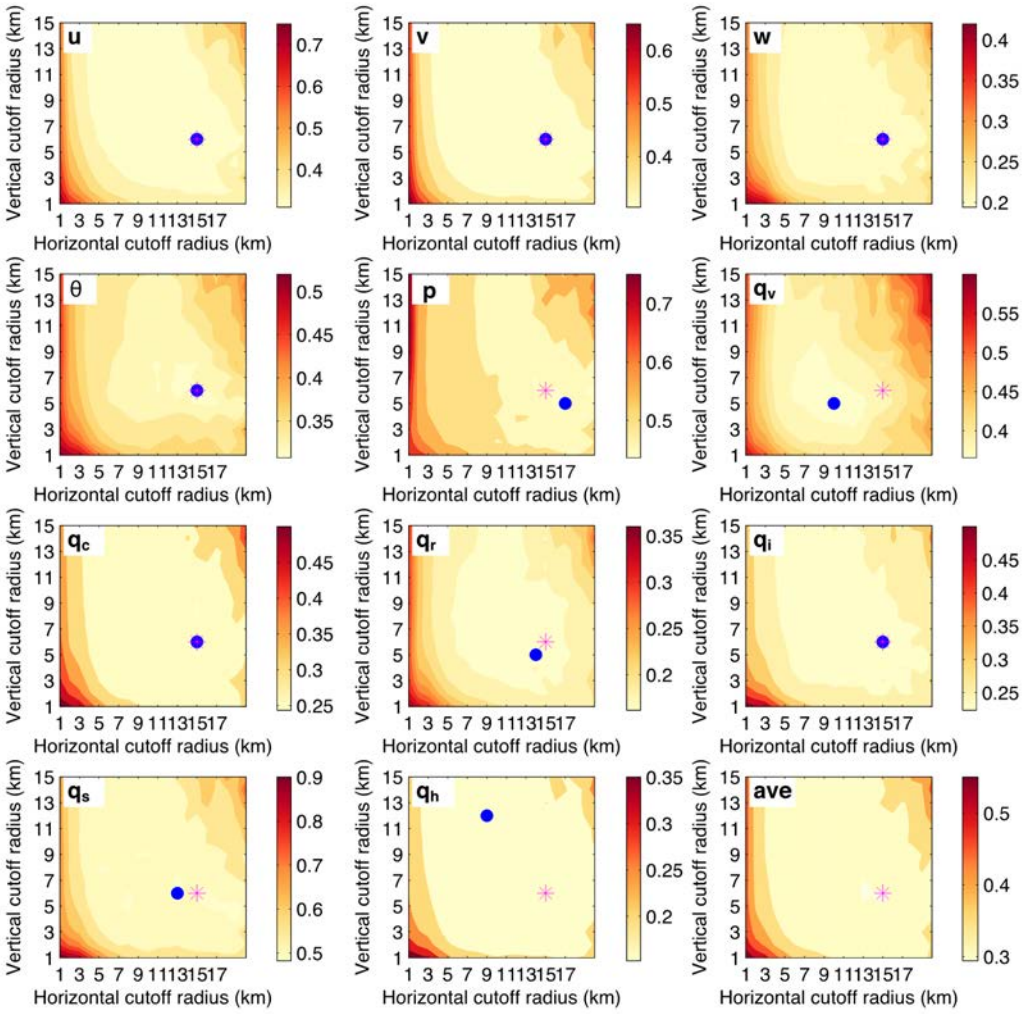

Fig. 3.2 The mean scaled RMSEs for different state variables as defined in Eq. (3.14) and averaged over all variables (ave), for DfEnKF experiments with different combinations of horizontal and vertical cutoff radius (km), with magenta asteroid indicating the minimum values over all values. The blue dot is the minimum for each individual variable.

Sensitivity experiments show that the optimal localization radii for EnKF are almost the same as those for DfEnKF, so the "optimal" localization radii of DfEnKF are adopted as the common optimal localization radii for both EnKF and DfEnKF in later experiments.

Unlike DfEnKF that uses a fifth-order Gaussian-like polynomial function and a Schur product (see Eq. (2.32)) to localize the ensemble covariance, En3DVar uses a recursive filter in each coordinate direction to mimic a second-order autoregressive function (that is close to Gaussian) for the localization. To investigate the source of difference between pure En3DVar and DfEnKF, the optimal localization radii of DfEnKF are also used by pure En3DVar in this section. The recursive filter length scale (S_{RF}) and the cutoff radius of DfEnKF localization (S_{GC}) can be roughly converted from each other based on the following equation,

$$S_{RF} = \sqrt{0.15} S_{GC} / \sqrt{2}, \tag{3.15}$$

which is given as Eq. (4) in Pan et al. (2014). The equivalent recursive-filter length scales for localization in En3DVar are therefore 4.11 km in the horizontal and 1.64 km in the vertical, corresponding to the 15 and 6 km cut off localization radii, respectively. In next section, the optimal localization radii of En3DVar will be determined by similar sensitivity experiments and En3DVar will be compared with other algorithms with their own optimal localization radii.

b. Single reflectivity or radial velocity observation assimilation

Single-observation experiments are often used to examine the correctness of newly developed DA systems, and the effects of spatial spreading of observation innovation by the background error covariance (Hu et al. 2006a; Zhu et al. 2013). The flow-dependent nature of the ensemble covariance in pure En3DVar and DfEnKF can be revealed by single-observation experiments. In our experiments, a single radial velocity observation is taken at a model grid point near the updraft region (as indicated

by the upward vector in Fig. 3.3). The ensemble forecasts from the 4th cycle of experiment EnKF to be discussed later are used in the single observation EnKF test. En3DVar uses the localization radii that are equivalent to the optimal localization configurations of DfEnKF.

For the single radial velocity DA experiment, the analysis increment of u wind is positive (negative) to the left (right) of the updraft for both pure En3DVar and DfEnKF (Fig. 3.3c and d), which is more consistent with the true increment (which is the truth minus background) in Fig. 3.3a, while for 3DVar, the analysis increment is a flow-independent ellipse that is not realistic (Fig. 3.3b). Similarly, for single reflectivity observation assimilation, the analysis increments of hail mixing ratio for En3DVar and DfEnkF are similar and show flow dependent structures while that of 3DVar is again elliptical (Fig. 3.4). An important point to note is that 3DVar creates hail mixing ratio increment symmetrically above and below the observation point, with the increment extending way below the freezing level, which in this case is unrealistic. In En3DVar and DfEnKF, the flow-dependent covariance is aware of the temperature dependency of hail distributions.

More differences between DfEnKF and pure En3DVar are found within the analysis increments of snow mixing ratio. Reasons for these differences will be addressed later in section 3b. Another point that worth mentioning is that radial velocity assimilation only updates the wind fields (u, v, w) and reflectivity assimilation only updates hydrometeors (q_r, q_s, q_h) in 3DVar. In EnKF and En3DVar, however, all model prognostic variables can be updated via background error cross-correlations; this is an

important advantage for algorithms utilizing the ensemble-derived background error covariances.

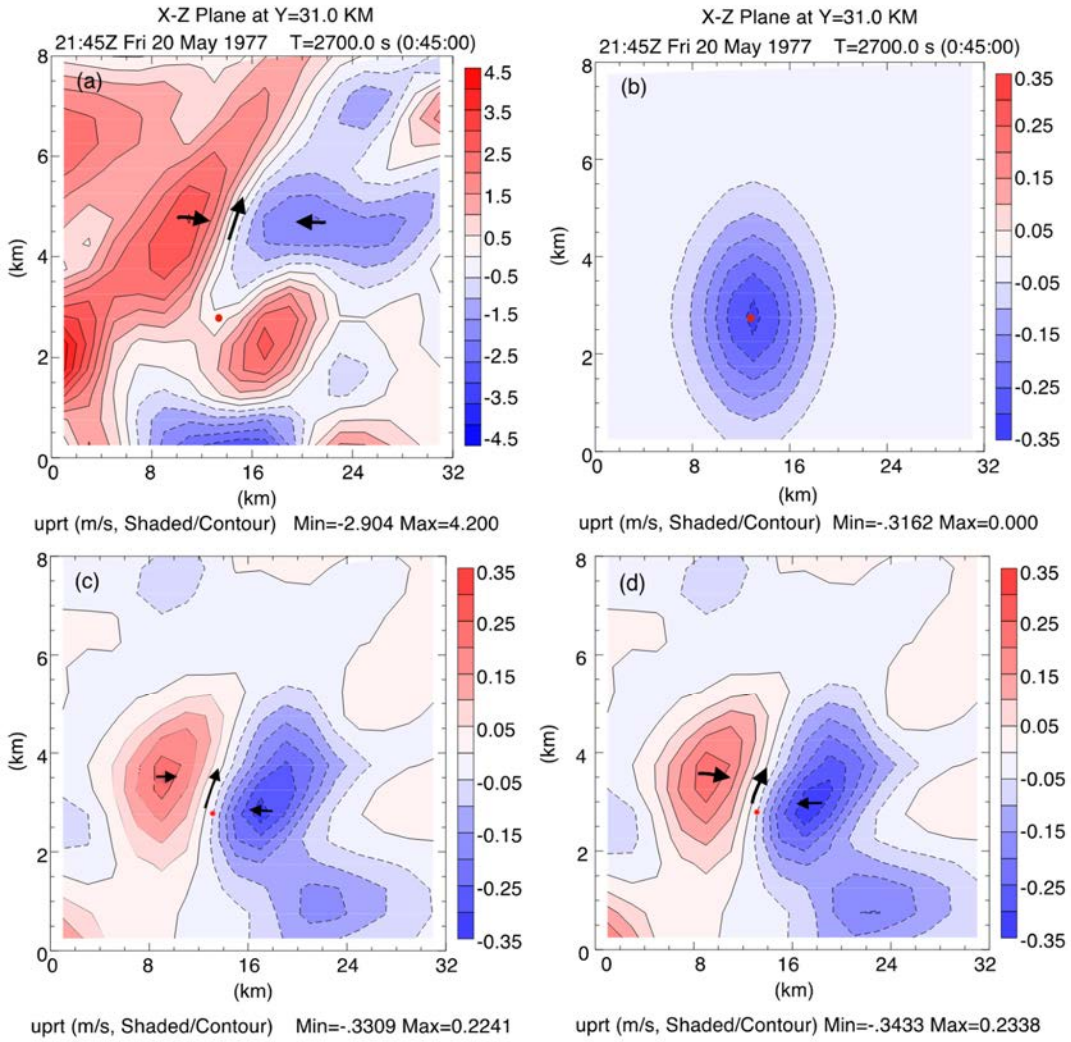

Fig. 3.3 Vertical cross-sections of (a) the truth minus background, and the analysis increment (analysis minus background) of u wind (m s^{-1}) when assimilating radial velocity at a single point, using (b) 3DVar, (c) pure En3DVar (c), and (d) DfEnKF (d), respectively. Horizontal convergence and upward motion are indicated by the horizontal and vertical black arrows, respectively.

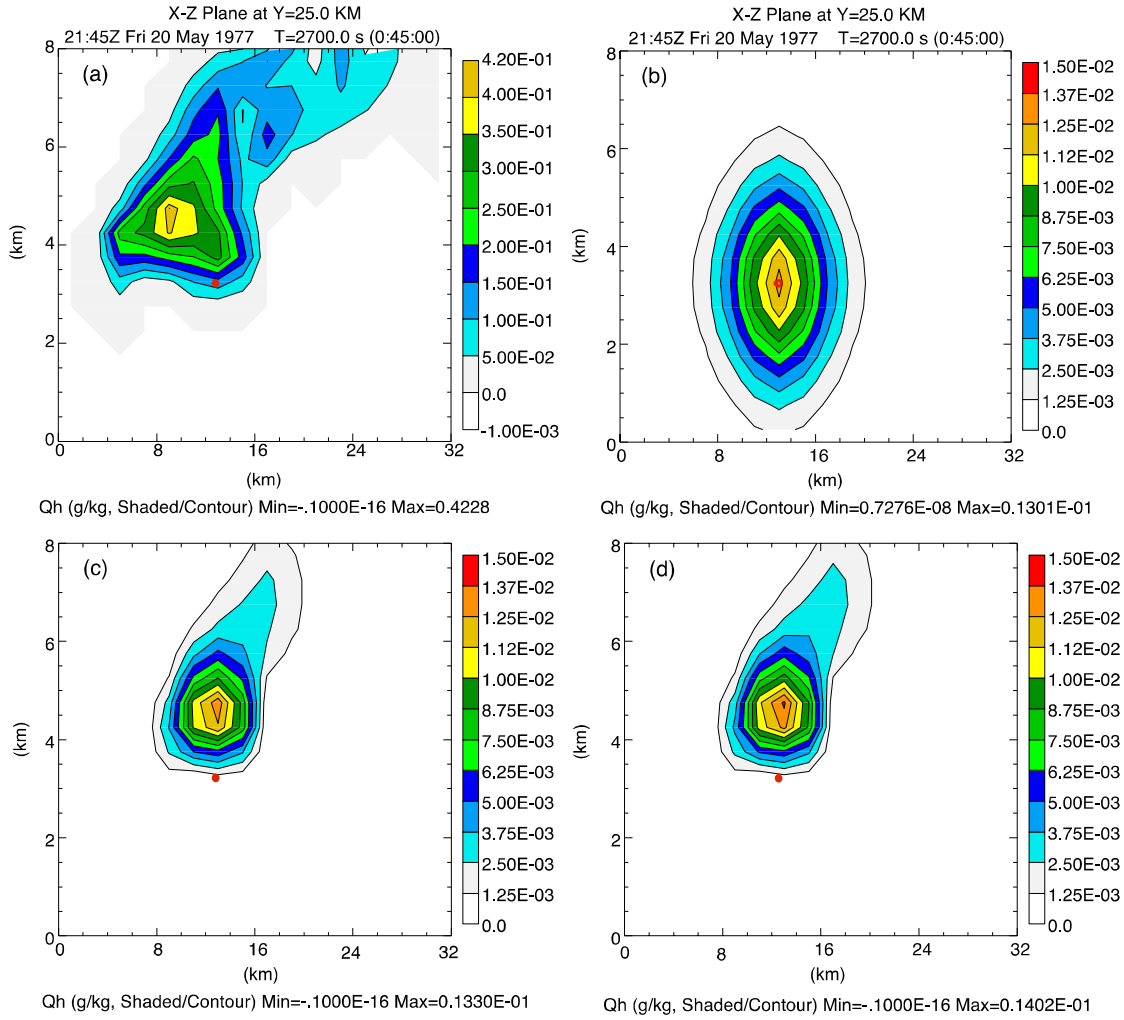

Fig. 3.4 Vertical cross-sections of the background error (truth minus background, a) and analysis increment (analysis minus background) for hail mixing ratio (g kg^{-1}) when assimilating a single reflectivity observation at the red dot location, using (b) 3DVar, (c) En3DVar, and (d) DfEnKF, respectively

c. OSSE comparisons between DfEnKF and pure En3DVar

In this section, the performance of EnKF, DfEnKF, and pure En3DVar are compared by assimilating both radial velocity and reflectivity data every 5 min for one hour. When using the optimal localization radii, DfEnKF clearly outperforms pure En3DVar in terms of RMSEs, especially for snow and some of the other hydrometeor

variables (Fig. 3.5). Only pressure has larger errors in some of the earlier cycles when using DfEnKF, but pressure is more sensitive to acoustic noise as pointed out by Tong and Xue (2005).

To see how well the precipitation fields are analyzed, the reflectivity components from individual hydrometeor mixing ratios are calculated according to

$$Z_x = 10log_{10}(Z_{ex}), \tag{3.16}$$

where Z_{ex} is the equivalent reflectivity, with *x* represent *r, s, or h* for rain, snow, and hail, respectively. Z_x is the corresponding reflectivity in dBZ and is introduced here for convenience only, and the total reflectivity $Z = 10log_{10}(Z_{er} + Z_{es} + Z_{eh})$ is not equal to the sum of Z_r, Z_s, and Z_h. Horizontal cross-sections of the analyzed reflectivity and contributions to reflectivity from snow and hail at 8 km AGL (the level where the differences between pure En3DVar and DfEnKF are most apparent at the 10th cycle, i.e., 2210 UTC) are depicted in Fig. 3.6. The spatial extent of the analyzed snow (hail) from pure En3DVar is somewhat underestimated (overestimated), while those from DfEnKF agree better with the truth. The intensities of updraft and downdraft from En3DVar at 8 km AGL are also underestimated while those from DfEnKF are much closer to the truth. Considering that DfEnKF and pure En3DVar use the same background error covariance derived from the EnKF ensemble (at 100% in pure En3DVar), both are based on deterministic background forecast configured the same way, and the background error covariance localization radii are configured equivalently, the two algorithms should theoretically yield very similar results. Why are the results of En3DVar worst, especially for snow, then? The reasons for the differences will be investigated in detail in the following section.

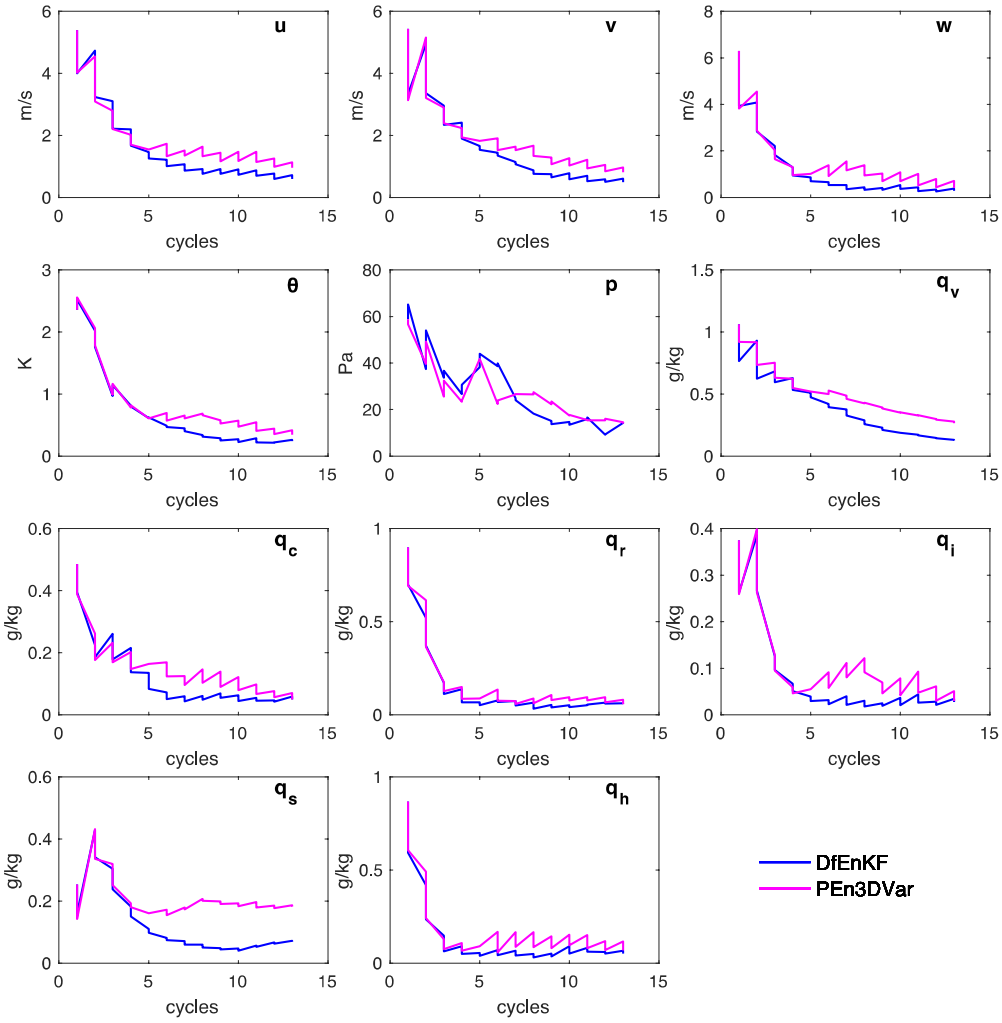

Fig. 3.5 RMSEs of the background forecasts and analyses of state variables verified in regions with true reflectivity higher than 15 dBZ (solid lines), assimilating both radial velocity and reflectivity data using EnKF (cyan lines), DfEnKF (blue lines), and pure En3DVar (magenta lines) algorithms.

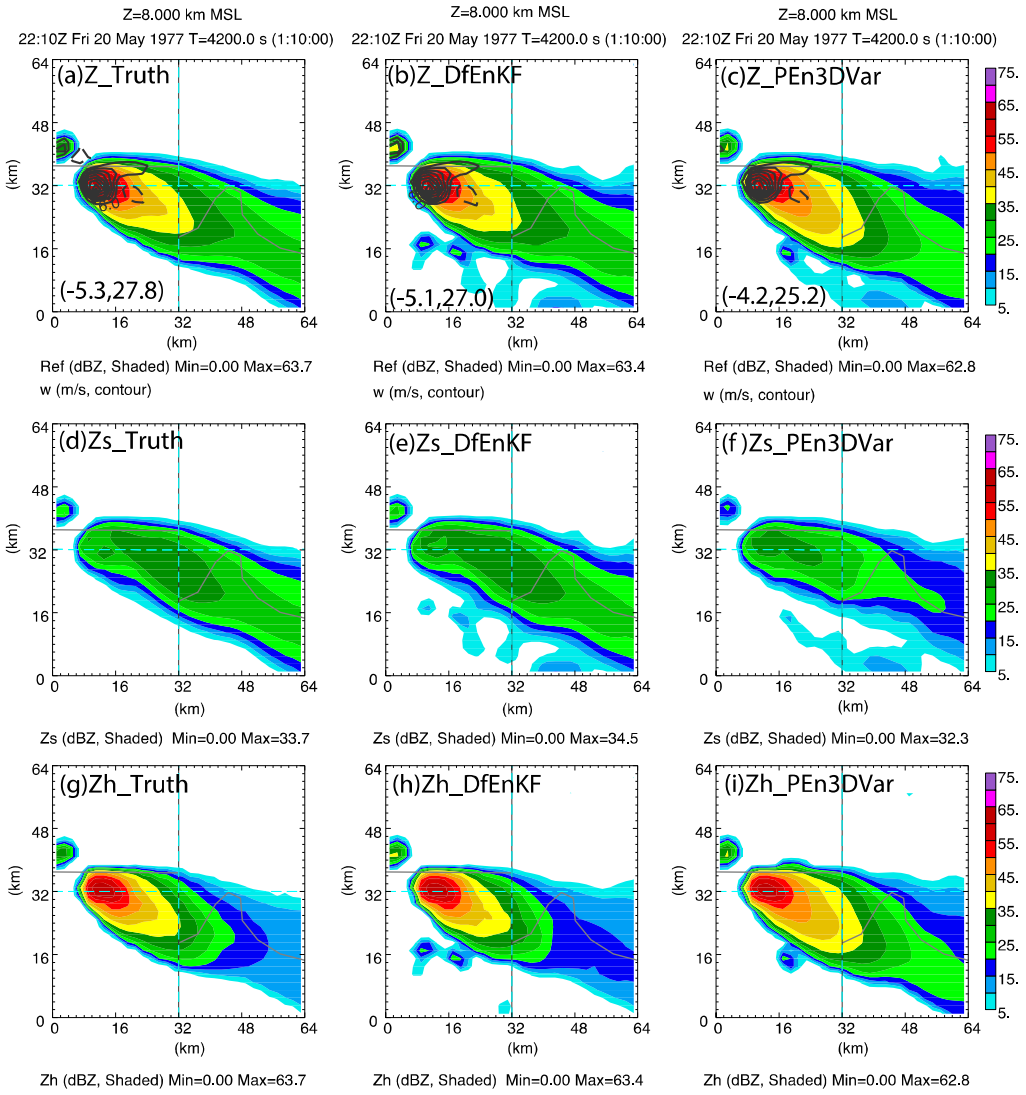

Fig. 3.6 Truth (left column) and analyzed fields at 8 km AGL of total reflectivity (Z, upper row) and reflectivity calculated from mixing ratio of snow (Zs, middle row) and of hail (Zh, low row) from DfEnKF (central column) and pure En3DVar (right column) at the 10th cycle (2210 UTC). Vertical velocity contours are overlaid in the upper panels together with the minimum and maximum values.

3.3.2 An analysis of the differences between DfEnKF and pure En3DVar

Possible reasons for the differences between DfEnKF and pure En3DVar include: covariance localization difference, serial versus global nature of the algorithms, and high nonlinearity of the observational operator of reflectivity that can cause

differences between filter update and variational minimization. Sensitivity experiments are conducted to investigate the impact of the above sources.

a. Covariance localization difference

The ARPS model state variables are defined on C grid. For ARPS En3DVar, the minimization procedure and the localization are implemented on A-grid, while for ARPS-EnKF, the localization is distance dependent so that it can be conducted between A and C grid directly for vector variables. This will cause some difference between ARPS-En3DVar and ARPS-EnKF for the localizations of vector variables. In order to check the localization difference between DfEnKF and En3DVar for both scalar and vector variables, temperature and horizontal wind observations on a model grid (x=8, y=15, z=7) are assimilated separately using the optimal cutoff of EnKF and the corresponding recursive filter length scale for En3DVar. Up to 10% difference is noticed between the localization curves for both scalar and vector variables (Fig. 3.7a, b). The difference is larger for vector variables (Fig. 3.7b) than for scalar variables (Fig. 3.7a) for horizontal localization due to the 5% difference of the peak values. Unlike EnKF, for which localization is directly conducted between A and C grid; for pure En3DVar, the localization is conducted on A grid, so that one more interpolation is needed to transform the control variables back to the C grids to get the analysis increments, this back-and-forth interpolations will introduce additional errors compared to that of EnKF. The difference of the localization schemes on z direction is small with higher resolution (500m). Overall, the impact of the localization difference between DfEnKF and pure En3DVar is fairly small and not likely to be significant.

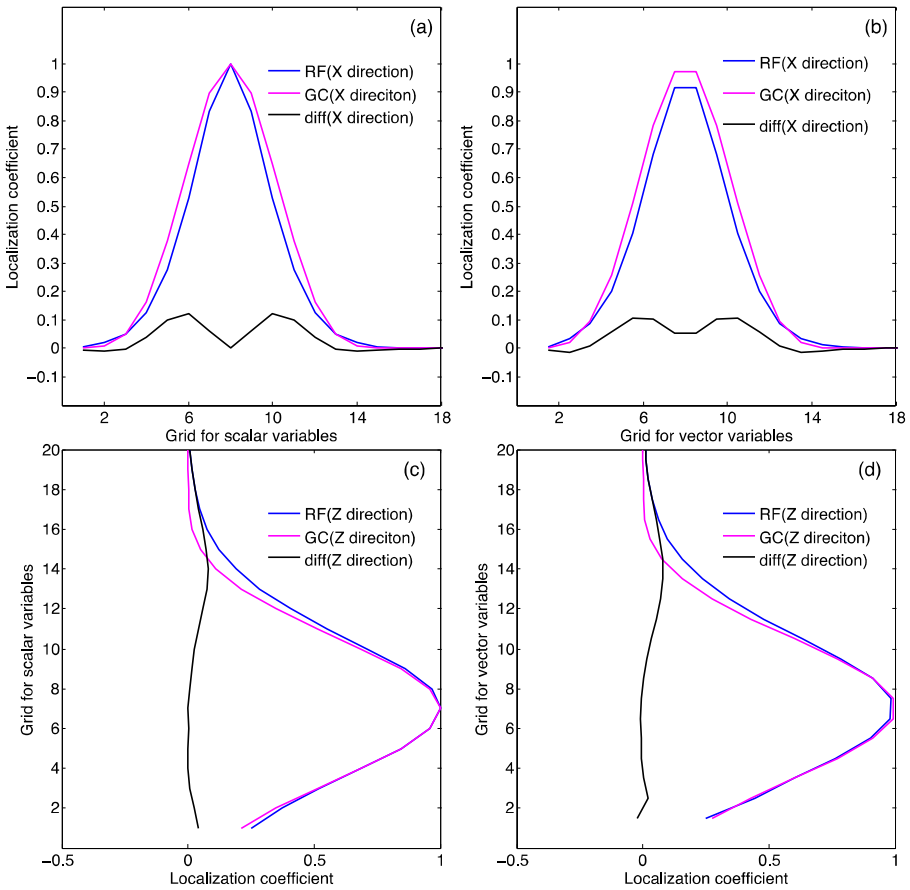

Fig. 3.7 Gaspri-Cohn functions (magenta line), the recursive-filter based localization curves (blue line), and their difference (dark line) for scalar variable (Fig. a, c) and vector variable under C grids (Fig. b, d) based on the corresponding optimal localization radii of DfEnKF (15km in the horizontal and 6km in vertical) and their counterparts based on Eq. (3.15)) in X coordinate (a), (b), and Z coordinate (c), (d) respectively.

b. Serial versus global nature of algorithms

To compare DfEnKF and pure En3DVar under a linear observation operator condition, only radial velocity data are assimilated. In such a case, DfEnKF clearly outperforms pure En3DVar for most of the state variables except for water vapor mixing ratio (q_v) when using the optimal localization radii of DfEnKF, i.e., 15 km in horizontal and 6 km in vertical (Fig. 3.8). The analyses and forecasts of snow mixing ratio in DfEnKF are again better than those of pure En3DVar.

In DfEnKF, the observations are assimilated serially, while observations in En3Dvar are assimilated globally (or simultaneously). In order to reduce the effect of the order in which data are assimilated, the horizontal (vertical) localization radius is decreased from 15 km (6 km) to 3 km (3 km). This significantly reduces the overlap of the influence regions from neighboring observations (given the 2 km interval of radar data in the horizontal and data on the radar elevations of about 1 degree apart). When no influence regions from different observations overlap, simultaneous assimilation of all observations should yield the same result as assimilating one observation at a time, because the assimilation of a prior observation does not affect the assimilation of the follow on observations.

After significantly decreasing the localization radii, the results of DfEnKF and pure En3DVar become very close (Fig. 3.8), suggesting the serial versus global nature of the two algorithms is a major cause of the differences in the results of DfEnKF and pure En3DVar when assimilating radial velocity data (for which the observation operator is linear). We point out here that with the reduced non-optimal localization radii, the error levels are clearly higher for all variables than when optimal radii are used (Fig. 3.8); despite that, data assimilation using either method is still able to decrease error in individual state variables, and the analyzed storm is reasonable (not shown).

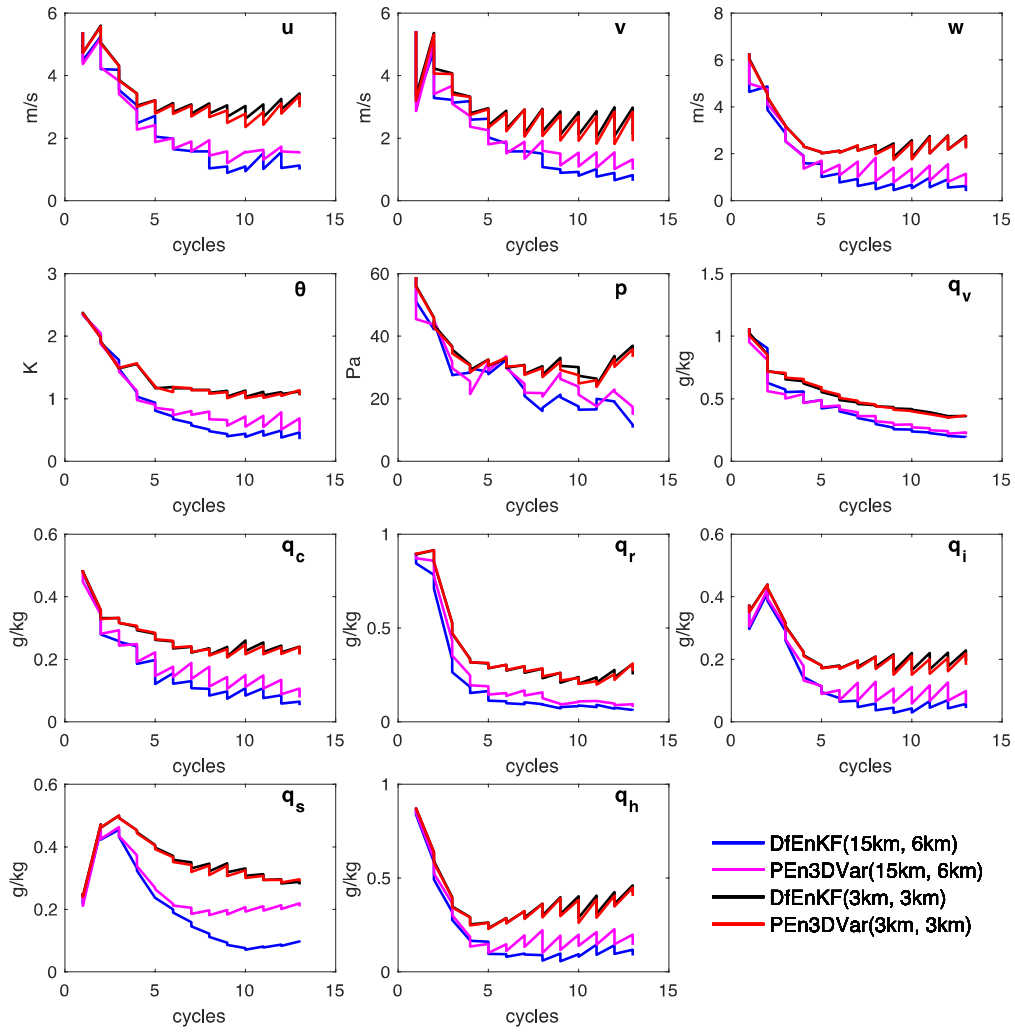

Fig. 3.8 RMSEs of the forecasts and analyses of state variables verified in regions with true reflectivity higher than 15 dBZ from DfEnKF (blue and black contours) and pure En3DVar (magenta and red contours) when assimilating only radial velocity data with the (horizontal, vertical) localization radii being (15 km, 6 km) (blue and magenta contours) and (3 km, 3 km) (black and red contours), respectively.

c. Influence of the nonlinear observational operator and its resultant difference between variational minimization and filter update

To isolate the effect of nonlinearity in the reflectivity observation operator, only reflectivity data are assimilated in the next set of experiments (Fig. 3.9). The differences between DfEnKF and pure En3DVar are more obvious compared to the analyses that

assimilate radial velocity data only. To reduce the impact of data assimilation sequence, localization radii are again decreased in the same way as earlier for radial velocity data. Unlike the radial velocity case, the differences between DfEnKF and pure En3DVar are still quite obvious even with the reduced localization radii (Fig. 3.9). This suggests that the nonlinear nature of the reflectivity observation operator is another significant contributor to the differences seen between DfEnKF and pure En3DVar.

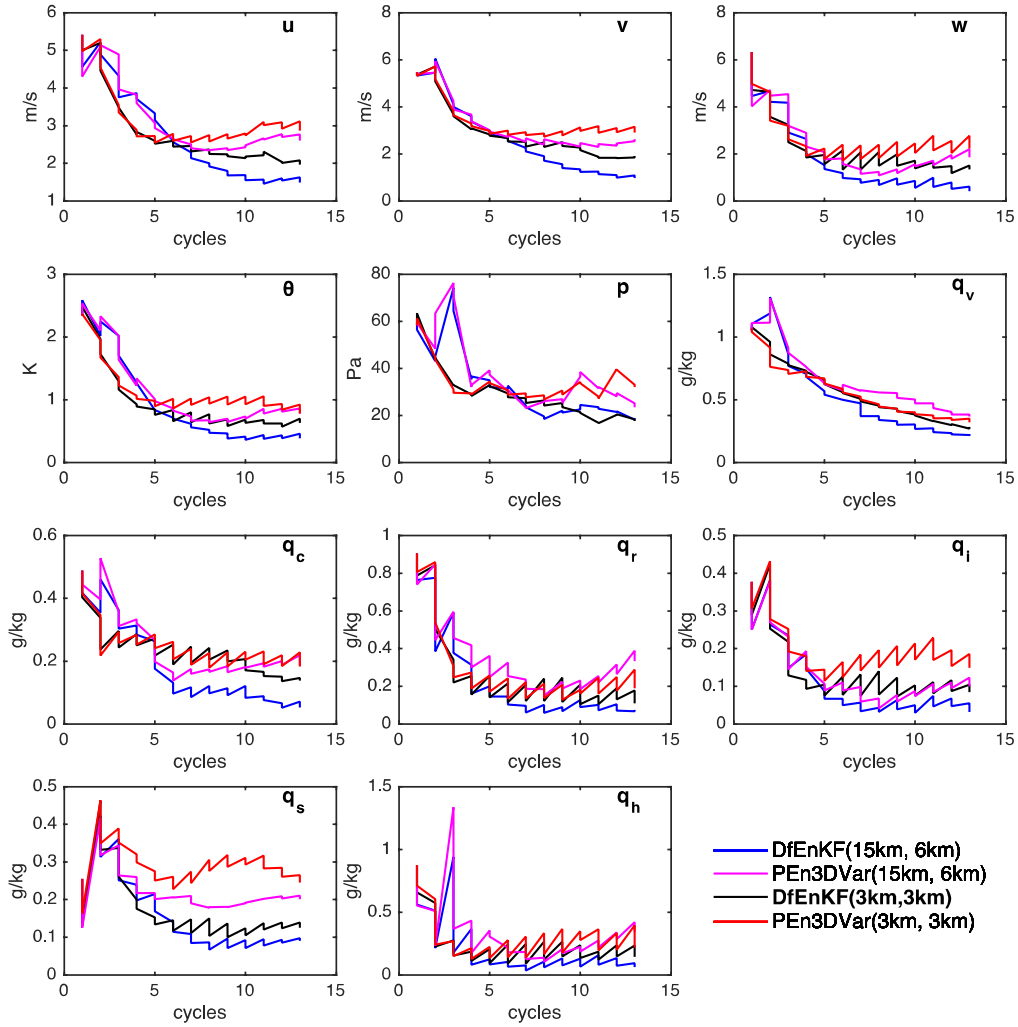

Fig. 3.9 Same as Fig. 3.8, when assimilating reflectivity data only.

To further examine this issue, another single observation test is performed at the time of the second analysis cycle when the analysis differences from the two experiments become large. The background of DfEnKF at the second cycle is used as the background of pure En3DVar also. A single reflectivity observation is placed at a grid point where the analysis difference between DfEnKF and pure En3DVar is most apparent for snow mixing ratio. To better capture the difference, relative error are calculated based on the background, analyzed, and true reflectivity (or equivalent reflectivity) at the given point. The relative error is defined as below,

$$\eta = \frac{Z^x - Z^t}{Z^t}, \tag{3.17}$$

where η is the relative error (%), and Z^x is the total reflectivity or equivalent reflectivity for analysis ($x = a$) and forecast ($x = b$), respectively. As is shown in Table 3.2, the relative error of the background reflectivity at the observation point (Z^b) is 35.2%, which is greatly reduced to 1.8% and 1.9%, respectively, after the assimilation of the single reflectivity data by pure En3DVar and DfEnKF. The relative errors of analyzed equivalent reflectivity (Z_e^a) from pure En3DVar and DfEnKF are 13.6 % and 14.3 %, respectively, with a small difference of 0.7 % only. However, when examining the analyzed equivalent reflectivity components from different hydrometeor species, the relative error differences for snow and hail from DfEnKF and pure En3DVar are much larger (37.0% for Z_e^a of snow and 9.0% for Z_e^a of hail), indicating that the differences in the algorithms are creating significant differences in the analyses of individual hydrometeor fields. Unlike EnKF, for which the analysis is obtained via direct filter update, analysis of En3Dvar is obtained through variational minimization, which tends

to adjust hail more than snow because the reflectivity is more sensitive to hail than to snow in the reflectivity observation operator. In variational algorithms, the sensitivity of the cost function to the control variables has a large effect on the amount of adjustment to the individual variables in the minimization process. Variables with small sensitivity may receive little adjustment before the minimization iteration is terminated.

Table 3.2 Comparison of single reflectivity observation assimilation between pure En3DVar and DfEnKF, where Z^t, Z^b, Z^a, indicate the true, background and analyzed reflectivity in dBZ, and Z_e^b, Z_e^a represent the background and analyzed equivalent reflectivity in mm^6m^{-3}, Z_{er}^a, Z_{es}^a, Z_{eh}^a, are the equivalent reflectivity components for rain, snow, and hail respectively.

	Background or analyzed reflectivity and equivalent reflectivity	**Relative Error (%) $(Z^x\text{-}Z^t)/Z^t$, x=a or b**		**Difference of relative error between pure En3DVar and DfEnKF**
		Pure En3DVar	**DfEnKF**	
Background	$Z^b = 10log_{10}Z_e^b$ (dBZ)	35.17%		0%
Analysis	$Z^a = 10log_{10}Z_e^a$ (dBZ)	1.78%	1.88%	0.1%
	$Z_e^a = Z_{er}^a + Z_{es}^a + Z_{eh}^a$ (mm^6/m^3)	13.59%	14.26%	0.67%
	Z_{er}^a (mm^6/m^3)	0%	0%	0%
	Z_{es}^a (mm^6/m^3)	54.73%	17.71%	37.02%
	Z_{eh}^a (mm^6/m^3)	4.53%	13.50%	8.97%

3.3.3 Comparisons of hybrid En3DVar with EnKF, DfEnKF, and 3DVar based on own optimal states

a. Optimal localization radii for En3DVar and optimal background error decorrelation scales for 3DVar

Because of the differences found in the results obtained with DfEnKF and pure En3DVar in the previous sections, the optimal localization radii of DfEnKF are not necessarily the optimal radii of pure En3DVar. To obtain optimal analysis with pure

En3DVar, 30 × 20 combinations of horizontal and vertical recursive filter length scales that range from 0.27 km to 8.10 km in the horizontal (corresponding to 1 km to 30 km cutoff radius of DfEnKF), and from 0.27 km to 5.40 km in the vertical (corresponding to 1 km to 20 km cutoff radius) are tested to determine the optimal radii in both horizontal and vertical directions. According to the error statistics shown in Fig. 3.10, the recursive filter length scale of 6.75 km (25 km cutoff) in horizontal and 3.24 km (12 km cutoff) in vertical yield the lowest mean error averaged across all variables. The optimal horizontal cut off radii in terms of the errors of individual fields are close to the overall optimal cut off radius, but in the vertical direction, the optimal radii for individual fields are either the same or larger than the overall cut off radii (Fig. 3.10), although further error decrease for vertical cut off radii beyond the overall optimal radius is generally small. In this study, the optimal radii are chosen based on the error averaged across all variables.

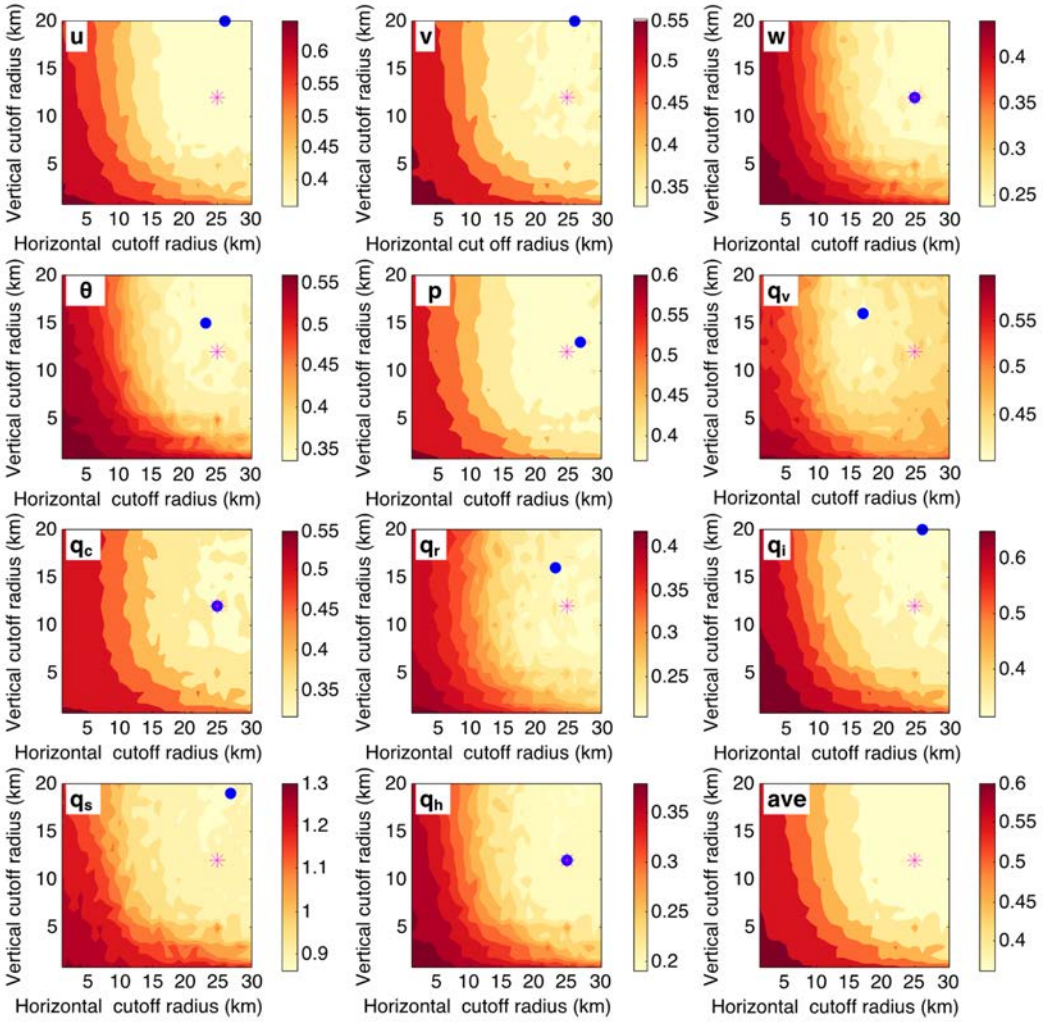

Fig. 3.10 The mean scaled RMSEs as defined in Eq. (3.14) for different state variables and averaged over all variables (ave), which are obtained based on pure En3DVar experiments with 30×20 combinations of horizontal and vertical length scales after being transformed to cutoff-radii of EnKF based on Eq. (3.15), with the magenta asteroid indicating the location of the minimum value over all variables. The blue dot is the minimum for each individual variable.

The optimal background error decorrelation scales for 3DVar are determined the same way. Background error decorrelation scales are first transformed to cutoff radii of EnKF based on Eq. (3.15) to keep the experiments of finding optimal cutoff radii (for EnKF) and length scales (for pure En3DVar and 3DVar) more consistent. According to Fig. 3.11, there are more scatters in the optimal scales based on errors of difference

variables, especially for the horizontal scale. Based on errors averaged across all variables, background error decorrelation scales of 4.05 km (equivalent to 15 km cutoff based on Eq. (3.15)) in the horizontal and 1.35 km (5 km cutoff) in the vertical yield the best results. These scales are therefore chosen as the optimal scales for 3DVar experiments.

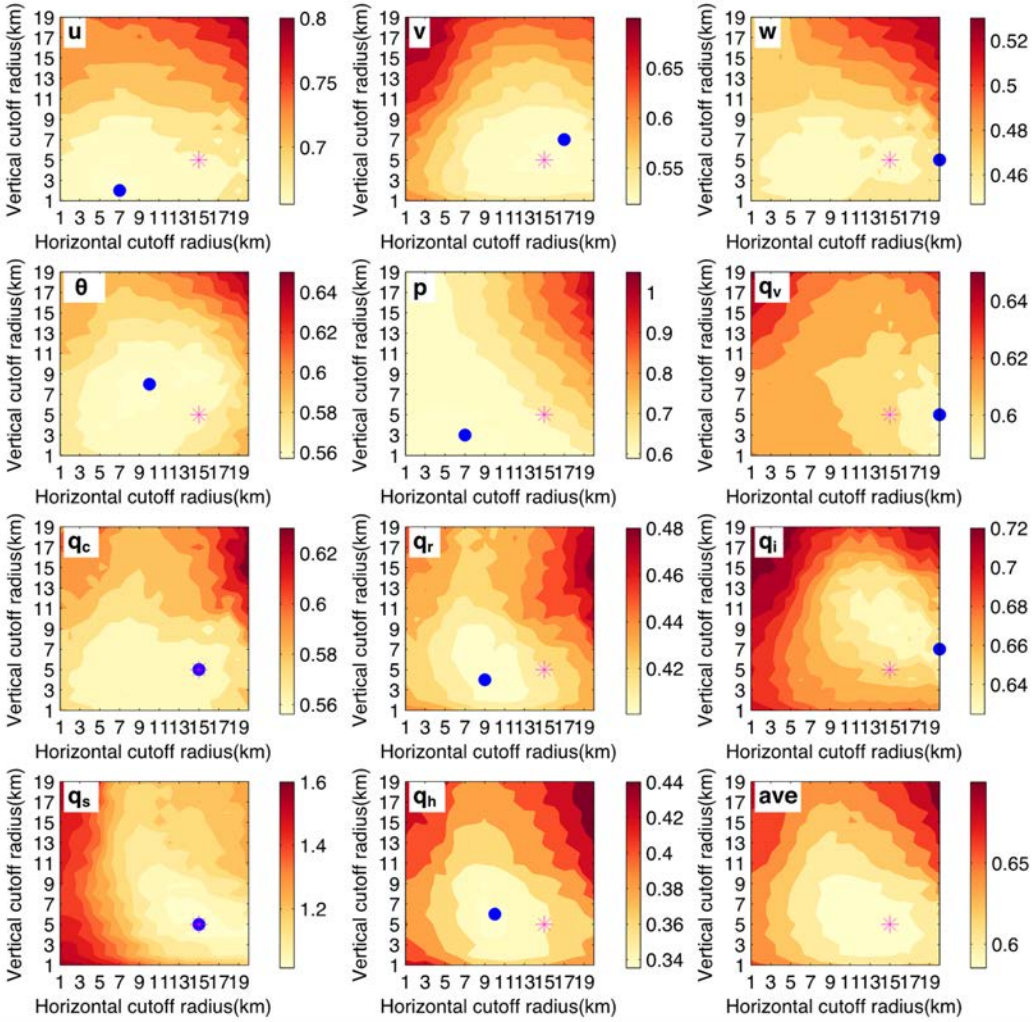

Fig. 3.11 The mean scaled RMSEs as defined in Eq. (3.14) for different state variables and averaged over all variables (ave), which are obtained based on 3DVar experiments with 20×20 combinations of horizontal and vertical decorrelation length scales after being transformed to EnKF's cutoff radii based on Eq. (3.15) (to be consistent with Fig. 3.9), with magenta asteroid indicating the location of minimum value over all values. The blue dot is the minimum for each individual variable

In the remaining hybrid En3DVar experiments, the optimal localization radii obtained for pure En3DVar and the optimal background error decorrelation scales obtained for 3DVar will be applied to the flow-dependent ensemble covariance and static covariance, respectively.

b. Optimal hybrid weights as function of ensemble size

To determine the optimal hybrid weights in hybrid En3DVar as a function of ensemble size, 10×10 experiments with different combinations of ensemble size (from 10 to 100 with increment of 10) and the weight given to the static background error covariance (from 0 to 1 with increment of 0.1) are performed. Pure En3DVar and 3DVar correspond to experiments with the static error weight set to 0 and 1, respectively.

As indicated in Fig. 3.12,. when ensemble size is 30 or larger, hybrid En3DVar with a 5% to 10% static B performs the best in analyses and forecasts of u, v, θ, p, and q_v (with a couple of exceptions for u and q_v). When ensemble size is 50 or larger, pure En3DVar (with 0% static B) performs the best in w and all hydrometeor variables. When the ensemble size is only 10 or 20, some variables, including v, θ, cloud water mixing ratio q_c, and q_v, require 50% or larger static weight to achieve the smallest error. For cloud ice mixing ratio (q_i), pure En3DVar always has the smallest error.

In terms of the average error across all variables, hybrid En3Dvar with a 5 ~ 10% (60%) static background error gives the best results when ensemble size is larger (smaller) than 20. Errors of pure and hybrid En3DVar increase quickly when the ensemble sizes falls below 40. These results indicate that an ensemble size of 30 or larger is needed for the hybrid system to benefit significantly from the ensemble error

covariances, and in that case only around 5% of static covariances is needed. For variables possessing convective and precipitation nature (*w* and hydrometeors), no static B is needed, especially when ensemble size is 50 or larger.

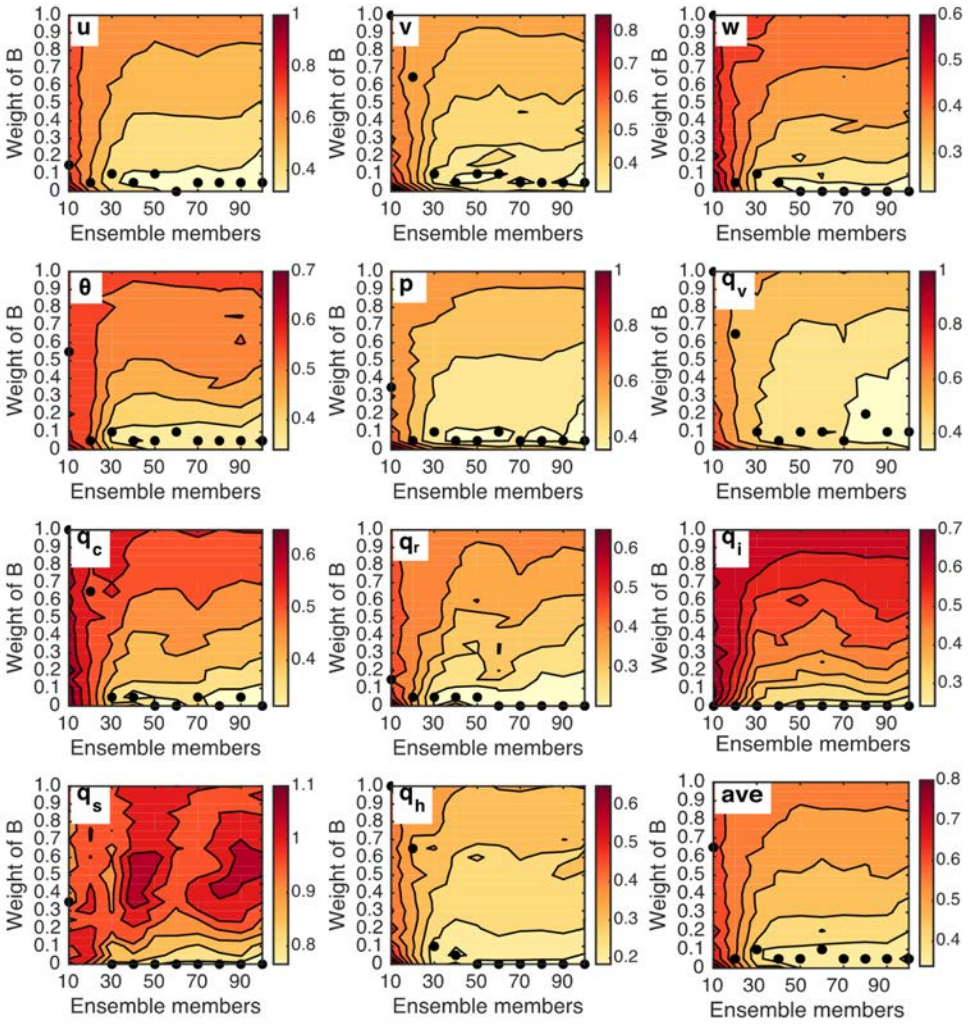

Fig. 3.12 The Grid-point average of the scaled RMSEs as defined in Eq. (3.14) for different variables and averaged over all variables (ave), which are obtained based on hybrid En3DVar experiments with different combinations of ensemble sizes and hybrid weights. The black dots indicate locations of the optimal weights for different ensemble sizes

c. Comparisons of hybrid En3DVar with EnKF, DfEnKF, and 3DVar

Comparisons are made among the analyses and forecasts from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar with their own optimal configurations (including localization/decorrelation scales and covariance weights). In all cases, the ensemble size used is 40, and this choice is guided by the results of the previous subsection, and is also consistent with many previous studies with EnKF for radar assimilation (e.g., Xue et al. 2006; Snook et al. 2011). Basically, we want to choose an ensemble size that can produce decent EnKF analyses without excessive costs, and ask the question if their analyses can be further improved by including some static background error via hybrid En3DVar.

The RMSEs are calculated only over grid points where the true reflectivity exceeds 15 dBZ, as have been done earlier. Fig. 3.13 shows that 3DVar performs the worst with its RMSEs being much larger than those of others. EnKF and DfEnKF have similar or the same error levels and both outperform pure and hybrid En3DVar for almost all variables in almost all cycles, with the exception being with p in the earlier cycles. As discussed earlier, the pressure analysis is more sensitive to acoustic noise that may affect the liability of the cross covariance between pressure and radar observations; we therefore give its error less weight in our performance assessment. Among the variables, q_v seems to benefit from the hybrid covariance the most as the error levels of hybrid En3DVar are lower than pure En3DVar and close to EnKF (and DfEnKF). The difference between EnKF (or DfEnKF) and pure (or hybrid) En3DVar is the largest for q_s. As discussed in section 3b, their differences are mainly caused by the serial versus global nature of the algorithms and the variational minimization versus

direct filter updating (which is sensitive to nonlinearity). In a variational framework, it is much more difficult to adjust mixing ratio of dry snow by reflectivity assimilation because the gradient of the reflectivity operator to dry snow is much smaller than those to hail and rain. In another word, when hail and dry snow coexist, the adjustment to hail tends to dominate in variational schemes, making correction to errors in dry snow more difficult.

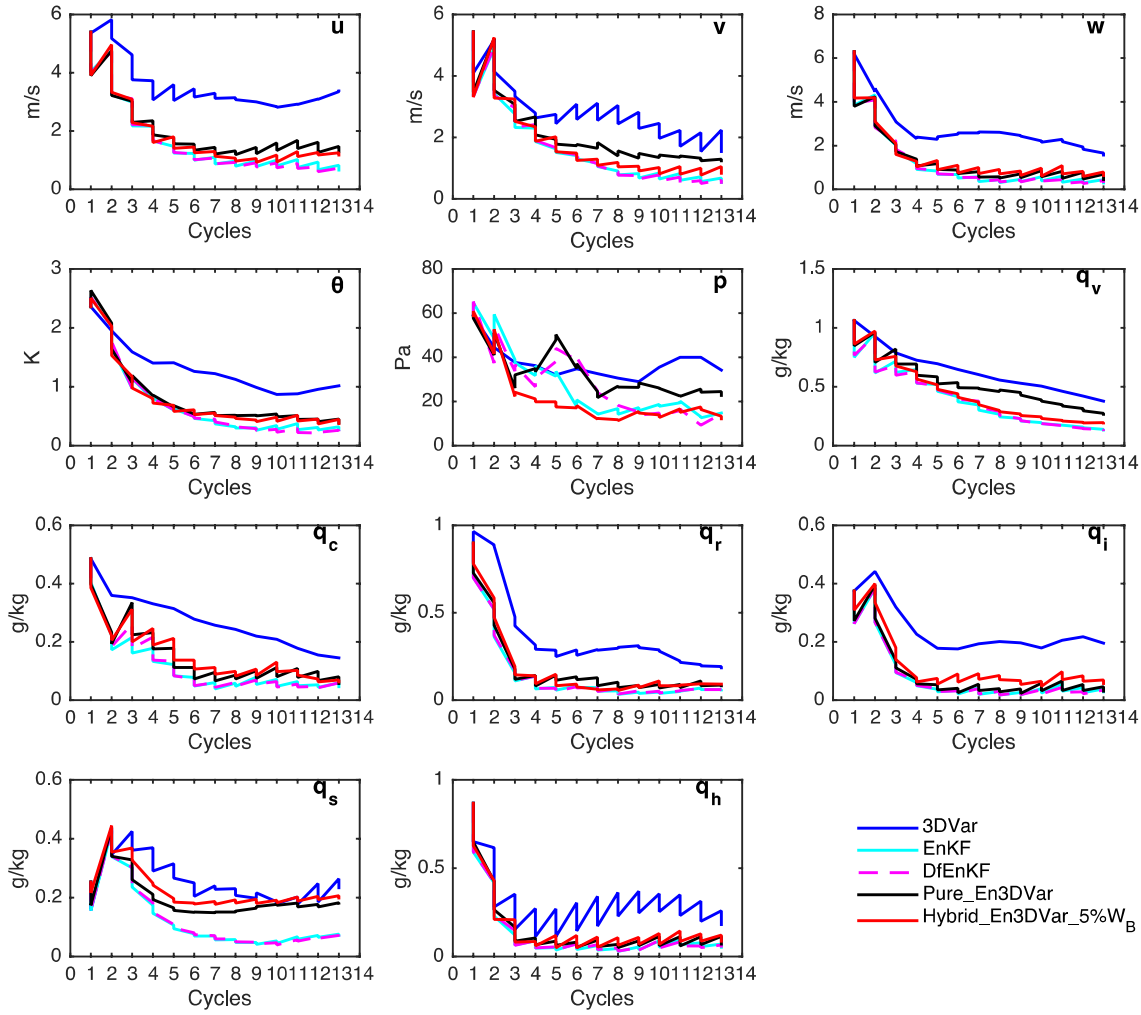

Fig. 3.13 The grid-point average of the RMSEs over the region where true reflectivity is higher than 15 dBZ for 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar (5% weight given to static B).

Vertical cross sections of analyzed reflectivity and winds from 3DVar, 40-member EnKF, DfEnKF, pure and hybrid En3DVar at the end of the one-hour DA cycles are shown in Fig. 3.14 together with the truth. The maximum reflectivity analyses from 3DVar and hybrid En3DVar are 67.7 dBZ and 70.3 dBZ, which are slightly less than those form the truth and EnKF (about 71 dBZ), indicating that both hybrid En3DVar and 3DVar underestimate the intensity of the reflectivity core (e.g., reflectivity exceeding 70 dBZ). The maximum vertical velocity from EnKF, DfEnKF, pure and hyrbrid En3DVar are over 25 m s^{-1}, which are much larger than that from 3DVar (18.8 m s^{-1}) and closer to that of the truth (27.1 m s^{-1}). Some spurious reflectivity is found outside the truth storm region in the analyses of 3DVar and hybrid En3DVar, which can be caused by the assimilation of error-containing reflectivity observations (whose errors can exceed 3 dBZ). In EnKF, DfEnKF, and pure En3DVar, spurious perturbations in the background forecasts can be effectively suppressed through flow-dependent background error correlations if zero reflectivity is assimilated, as was pointed out by Tong and Xue (2005).

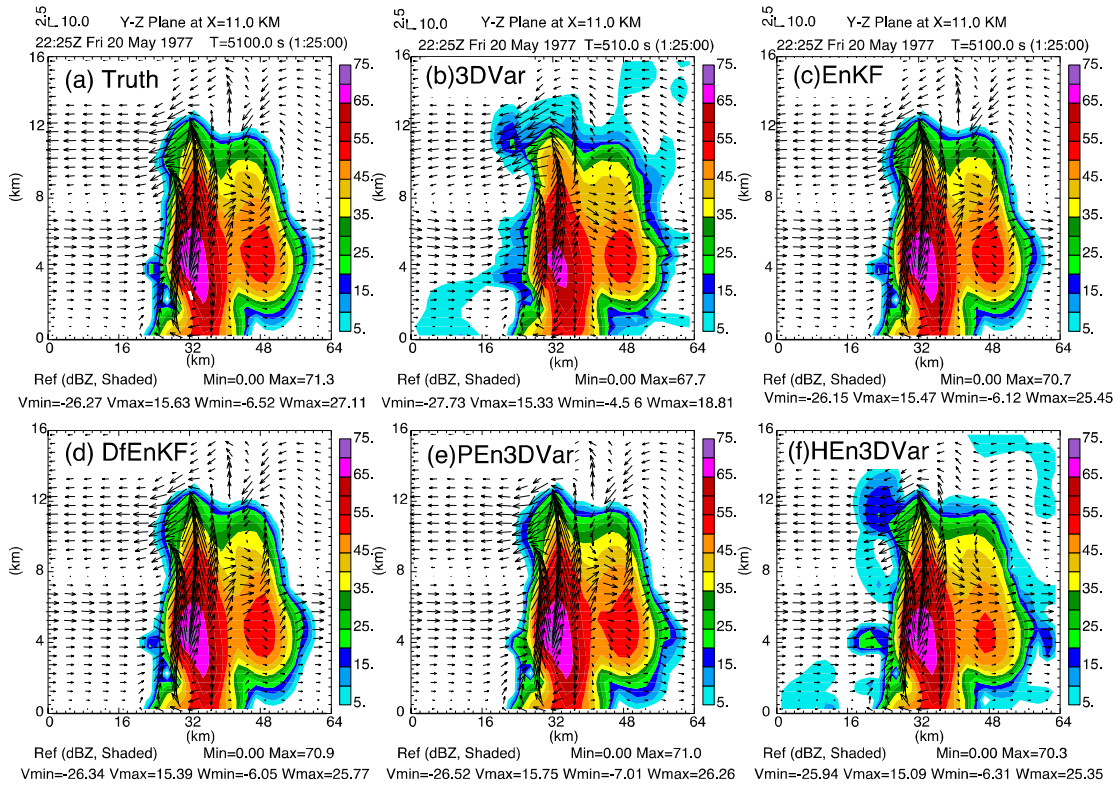

Fig. 3.14 Vertical cross-sections of the reflectivity field overlaid with the wind vector at the end of the one-hour DA window from (a) truth, and the analyses of (b) 3DVar, (c) pure En3DVar, (d) DfEnKF, (e) pure En3DVar, and (f) hybrid En3DVar with 5% static covariance.

To examine the quality of analyzed hydrometeor fields, we compare the reflectivity components from individual hydrometeor mixing ratios according to Eq. (3.17). According to Fig. 3.15, the contribution to equivalent reflectivity from analyzed rainwater mixing ratio q_r from different algorithms are very similar and close to that of truth. For hail, 3DVar and hybrid En3DVar underestimate the hail reflectivity at the reflectivity core region, contributing to the under-estimation of maximum analyzed total reflectivity noted earlier. For snow, EnKF and DfEnKF obtain better analyses than pure En3DVar, hybrid En3DVar, and 3DVar. The contribution of snow to total reflectivity is,

however, much smaller than that of hail, making correction to its error more difficult in a variational framework.

Overall, EnKF, DfEnKF and pure En3DVar give similar reflectivity or precipitating hydrometeor analyses that are very close to truth. There is some benefit to bring in 5% or so static error covariance into En3DVar to form a hybrid algorithm when verified within the truth precipitation regions, but some spurious reflectivity can appear in the clear air regions (Fig. 3.15) that can degrade the overall analyses, similar to the 3DVar case.

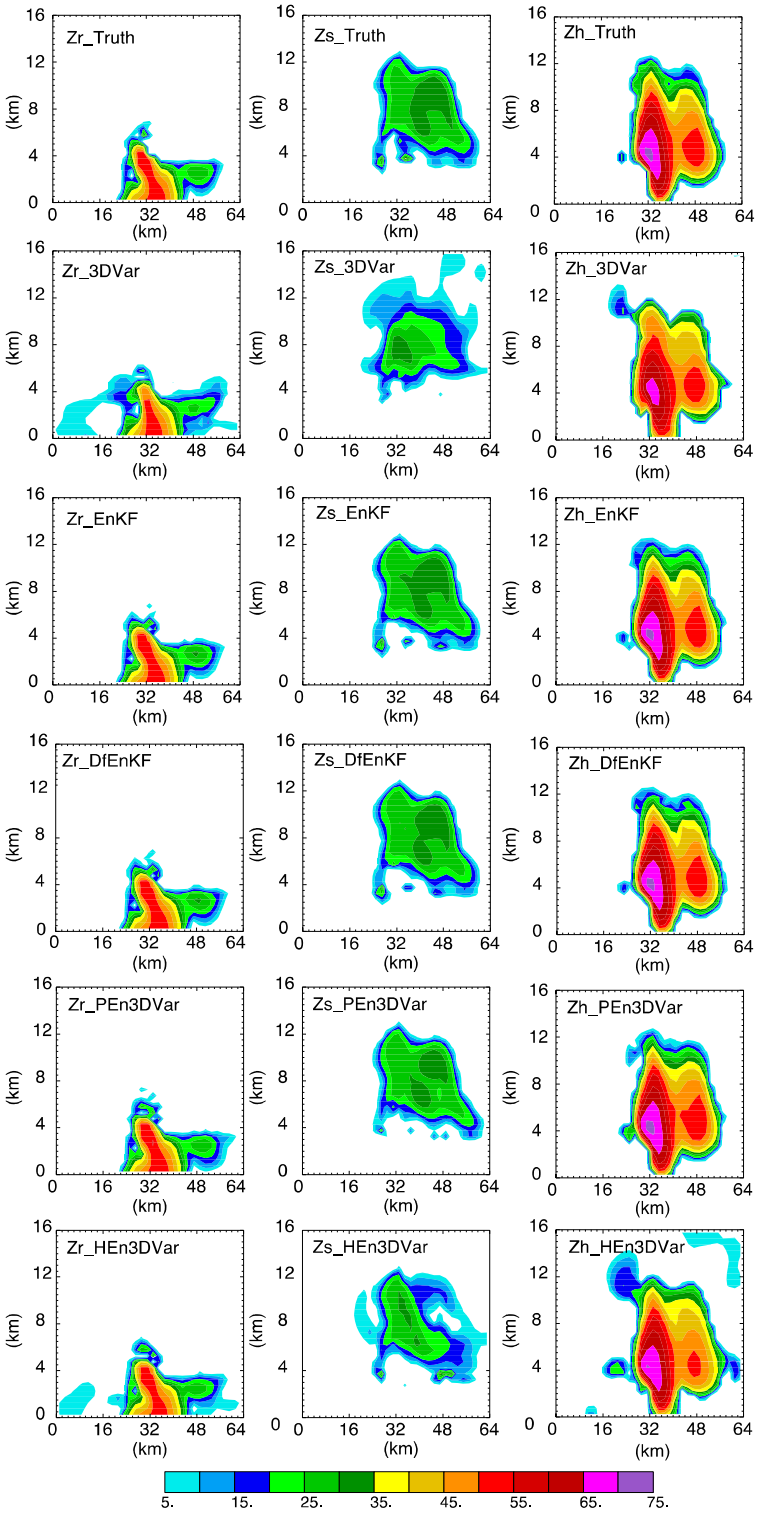

Fig. 3.15 Vertical cross-sections of the reflectivity (in dBZ) calculated based on the mixing ratio analyses of rain, snow, and hail from truth, analyses of 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar respectively.

3.4 Summary and discussion

In this chapter, a hybrid En3DVar system that was developed based on the ARPS variational framework is applied to the assimilation of simulated radar data for a supercell storm. Radar radial velocity and reflectivity data were assimilated every five minutes for one hour. Differences between EnKF and pure En3DVar are first investigated to better understand the difference among different algorithms. DfEnKF, which updates a single deterministic background forecast using the EnKF mean updating algorithm, is introduced to have an algorithm-wise parallel comparison between EnKF and pure En3DVar. 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar are tuned to obtain their optimal configurations before they are compared. The experiments performed and related conclusions are summarized as follows:

- Single-point radial velocity or reflectivity data assimilation experiments show that pure En3DVar and DfEnKF produce similar flow-dependent analysis increments that depict flow-dependent structures while 3DVar produces isotropic analysis increments that are not necessarily physical.
- Differences are found in the analyses of DfEnKF and pure En3DVar when both use the same effective localization scales even though they are supposed to be the same for linear and Gaussian problems. The serial versus global nature of the algorithm is shown to be responsible for the analysis differences in radar radial velocity DA, for which the observation operator is linear; variational minimization versus direct filter update is responsible for analysis difference in reflectivity DA, for which the observation operator is highly nonlinear. As a variational algorithm, pure En3DVar tends to adjust hail more than snow,

because the analysis results are controlled to a large extent by the sensitivity of reflectivity with respect to individual hydrometeor, i.e., the gradient of the reflectivity operator with respect to the hydrometeor state variables. The EnKF/DfEnKF algorithm relies on the background error covariances for analysis updating.

- For 3DVar, optimal de-correlation scales for static background error covariances are obtained via sensitivity experiments, and the optimal localization radii for ensemble background error covariances are similarly obtained for EnKF/DfEnKF and En3DVar. These optimal configurations are then used to construct hybrid En3DVar.
- The sensitivity of hybrid En3DVar to covariance weights and ensemble size is then examined. It is found that when the ensemble size is 30 or larger, a 5% to 10% weight for static covariance produces the smallest mean errors for u, v, θ, p, and q_v, fields. For w and hydrometeor fields, 0% static covariance performs the best when the ensemble size is 50 or larger. On average, when the ensemble size is 20 or larger, a 5 or 10% static covariance gives the best results while for smaller ensembles a larger static covariance with a ~50% weight produces somewhat better results. Using an ensemble size of 40, EnKF and DfEnKF perform similarly, and both are better than pure and hybrid En3DVar overall. Using 5% static error covariance, hybrid En3DVar outperforms pure En3DVar for most variables except the hydrometeor variables; the improvement is the largest for q_v and the degradation is most notable for q_s. In a sense, the extra static error covariance does not help for hydrometeor variables. Overall, in the

current perfect OSSE framework, EnKF and DfEnKF perform similarly and the best, 3DVar the worst. En3DVar with or without static background error covariance does not perform as well as EnKF/DfEnKF, and static covariance only helps slightly via hybrid En3DVar.

Finally, we note that the conclusions obtained here are based on the assumption of a perfect prediction model. In next chapter, we will compare hybrid En3DVar against EnKF and 3DVar under imperfect model conditions and will apply the algorithms to a real supercell storm case afterwards.

Chapter 4 Evaluation of the ARPS Hybrid En3DVar System for Radar Data Assimilation with Observing System Simulation Experiments (OSSEs) Under Imperfect Model Assumptions

4.1 Introduction

In the previous chapter, the newly developed ARPS En3DVar data assimilation system was tested using an OSSE under perfect model assumptions (i.e., the model is an exact representation of the true state). Radar data was assimilated for a simulated supercell storm case using EnKF, DfEnKF, 3DVar, pure and hybrid En3DVar algorithm. Optimal localization radii (background error de-correlation scales) for ensemble (static) background error covariance are first obtained through sensitivity experiments. A separate 'deterministic forecast' EnKF analysis, called DfEnKF, is performed to obtain an algorithm-wise parallel comparison between EnKF and pure En3DVar. Performance of hybrid En3DVar in assimilating radar data is compared with those of 3DVar, EnKF, DfEnKF, and pure En3DVar. On average, when ensemble size is 20 or larger, a 5 to 10% static covariance gives the best results, while for smaller ensembles, more static covariance is beneficial. Using an ensemble size of 40, EnKF/DfEnKF performs the best, 3DVar the worst, and static covariance only helps slightly via hybrid En3DVar. Under perfect model assumption, when the ensemble background error covariance is a good estimation of the truth, the impact of the static background error covariance in hybrid DA is not obvious.

Perfect-model OSSEs tend to underestimate the analysis and forecast errors. When severe model errors exist, the ensemble covariance may be systematically underestimated so that EnKF can suffer from filter divergence in the following-up

cycles (Maybeck 1982, Mitchell and Houtekamer 2000; Hamill et al. 2002). In this case, the static background error covariance that could take into account the experience and statistical factors, might be a good supplement and constraint to the pure ensemble-based data assimilation methods (Hamill and Snyder 2000; Lorenc 2003; Etherton and Bishop 2004; Wang et al. 2007b; Gao et al. 2013). Based on the above hypothesis, we raised the following two questions: 1) Do the conclusions obtained based on the perfect-model OSSEs change if the ensemble background error covariance is not a good estimation of the true error distribution? 2) In what situation that the static background error covariance could help, if any, for convective scale radar DA?

To answer the above questions, model errors need to be introduced. This is done first via OSSE. This is necessary before applying the algorithm to the real case, where many sources of errors make the understanding of the algorithm behaviors difficult. Microphysical error is one of the major sources of the model error (Houtekamer et al. 1996; Houtekamer and Lefaivre 1997; Stensrud et al. 2000). There are several ways to account for the microphysics errors, such as using multiple microphysics (Stensrud et al. 2000), or perturb the parameters within a single microphysics scheme (Stainforth et al. 2005; Lee et al. 2012; Baker et al. 2014). In this section, model errors were introduced by using different microphysical schemes in the truth run (Lin scheme, Lin 1983) and in the forecasts within the DA cycles (WSM6 scheme, Hong and Lim 2006). With this model error, hybrid En3DVar is compared with 3DVar, EnKF, DfEnKF, and pure En3DVar to examine the relative advantage of static background error covariance B in hybrid DA.

Similar to what was done in the perfect-model OSSEs, optimal localization radii (background error de-correlation scales) for ensemble (static) background error covariance are first obtained through sensitivity experiments before they are compared. DfEnKF was also introduced to make an algorithm-wise parallel comparison with pure En3DVar. Optimal hybrid weights are obtained as functions of the ensemble sizes. At the ensemble size of 40, hybrid En3DVar with the optimal weight of the static **B** will be compared with 3DVar, EnKF, DfEnKF, and pure En3DVar. So that the relative advantage of the static **B** in hybrid DA can be further examined, as a follow-up studies made to Chapter 3.

The rest of this chapter is organized as follows. In section 4.2, we briefly review DA schemes (EnKF, DfEnKF, and En3DVar) used as well as the experimental design. In section 4.3, sensitivity experiments are conducted to obtain optimal cutoff radii (recursive filter length scale) for EnKF, optimal background error decorrelation scale for 3DVar, and optimal weight for hybrid En3DVar. In the experiments, the reflectivity DA suffers from excessively large gradient due to small background reflectivity. A double-pass procedure, which assimilates precipitation and clear-air reflectivity separately in different passes, is proposed and found to be able to greatly alleviate the problem. It will also be discussed in section 4.4. Finally, hybrid En3DVar, 3DVar, and EnKF are compared based on both objective verifications and physical field analyses. A summary and outlook are then concluded in section 45?

4.2 Experiment design

4.2.1 The prediction model and truth simulation

As in the previous chapter, the ARPS En3DVar DA system tested in this chapter is developed based on the full version of the ARPS. In this chapter, the true fields are simulated in the same way as that did for the perfect-model OSSEs in Chapter 3, in which the classic supercell storm of 20 May 1977 in Del City, Oklahoma (Ray et al. 1981) is simulated based on the single-moment, three-category ice scheme of Lin et al. (1983). The domain size is 35×35×35, with a 2 km grid interval in the horizontal directions and 0.5 km in the vertical. Similar to the perfect-model OSSEs, non-stretching vertical grid is used. Unlike EnKF, for which the Shur-product based localization is distance dependent, the recursive filter based localization for pure En3DVar is grid-size dependent. If using stretched grid in vertical direction, the localization for hybrid En3DVar will be asymmetric in vertical. To reduce the potential difference caused by two different localization realizations when comparing the EnKF and En3DVar algorithms, non-stretched grid will be adopted in the OSSEs. However, in the following real case applications, a stretched grid will be used in the vertical direction to enable higher vertical resolution near the surface.

The truth simulation or nature run is initialized from a modified real sounding valid at 21:00 UTC, with a thermal bubble being used to initialize the storm. Details of the truth simulation are the same as that described in the previous chapter. Here model errors are introduced by using different microphysical schemes in nature run (Lin scheme, Lin et al. 1983) and forecasts (WSM6 scheme, Hong and Lim 2006) within the DA window.

4.2.2 Simulation of the radar observations

The radial velocity and reflectivity are simulated based on model velocity and mixing ratios of hydrometeor fields, respectively. For details of the operators, please refer to Chapter 3. The radar data is simulated in the same way as the perfect-OSSEs. Specifically, single WSR-88D radar, which is assumed to be located at the southwest corner of the simulation domain, is produced based on the truth simulation. Unbiased and normal distributed random errors with standard deviations of 3 dBZ and 1 m s^{-1} are added to the simulated reflectivity and radial velocity data, respectively. The same error levels are also specified in the following DA experiments. Radial velocity data are assimilated only in regions where truth reflectivity is greater than 15 dBZ in the analysis domain, while reflectivity data are assimilated in both precipitation and clear air regions.

4.2.3 Design of assimilation experiments

The initial ensemble members are generated at 22:20 UTC by adding random noise to the horizontally homogeneous first guess defined based on the environmental sounding. The noise is sampled from Gaussian distribution with zero mean and standard deviations of 2 ms^{-1} for u, v, and w, 2 K for potential temperature, and 0.6 g kg^{-1} for mixing ratios of water vapor and hydrometeor fields respectively.

The hydrometeor perturbations are confined to the region with reflectivity higher than 10 dBZ to suppress spurious storms in the analyses and forecasts. The radial velocity and reflectivity observations are assimilated every 5 min with the first one conducted after 5-minute integration of the initial perturbations (22:25 UTC). Experiments with different combinations of localization radii and recursive filter length

scales are tested to obtain the optimal localization radii and recursive filter length scales for EnKF and En3DVar respectively.

Seven sets of experiment that are named based on the DA algorithms being used are conducted in this study. In experiment 3DVar_SP (3DVar_DP), radar data are assimilated using 3DVar with the clear-air reflectivity being assimilated together (separately) with the precipitation reflectivity in a single (double pass). Here 'SP' and 'DP' denote single and double pass, respectively. The next two experiments are EnKF and DfEnKF. DfEnKF updates a single deterministic background forecast based on EnKF mean updating equation, it is algorithm-wise parallel to pure En3DVar. Experiments PEn3DVar and HEn3DVar both use the En3DVar algorithm with the stand-alone ensemble covariance and a combination of static and ensemble-covariance respectively. Different combinations of ensemble sizes and weights are tested in hybrid En3DVar to obtain the optimal weights as functions of the ensemble sizes. Finally, Hybrid En3DVar is compared with 3DVar, EnKF, and pure En3DVar with their own optimal configurations to examine the advantage of hybrid En3DVar over other methods. Detailed descriptions about each experiment can be found in Table 4.1.

Table 4.1 Descriptions of the assimilating methods

Experiment	Use of background error covariance	Background updating	Use of double-pass	Ensemble size
3DVar_SP	Static background error covariance	Update background field deterministically	Assimilate clear-air (<5 dBZ) and precipitation reflectivity (≥5 dBZ) together in a single pass	N/A
3DVar_DP			Assimilate clear-air and precipitation reflectivity separately in a double pass	
EnKF	Background error covariance derived from ensemble background forecasts	Updates ensemble mean background and ensemble perturbations using EnKF algorithm		40
DfEnKF	Using ensemble covariance from an EnKF system	Update a single deterministic background forecast using EnKF mean updating algorithm		40
PEn3DVar	Using 100% ensemble covariance from an EnKF system	Update a single deterministic background forecast using variational algorithm		40
HEn3DVar_SP	Using weighted average of ensemble covariance and static 3DVar covariance	Update a single deterministic background forecast using variational algorithm	Assimilate clear-air and precipitation reflectivity together in a single pass	10, 20... 100
HEn3DVar_DP			Assimilate clear-air and precipitation reflectivity separately in a double pass	

4.3 Results of assimilation experiments

4.3.1 Optimal localization radii for EnKF and En3DVar and optimal background error de-correlation scales for 3DVar

To obtain the optimal localization radii and background error correlation scales for imperfect-model OSSEs, the scaled root mean square error defined in Chapter 3 is adopted here. Please refer to section 3.3.1 for the definition. Similarly to what has been done in the perfect-model OSSEs, sensitivity experiments with different combinations

of the horizontal and vertical localization radii that range from 1 km to 15 km with 1 km interval are first conducted to obtain the optimal localization radii for EnKF and DfEnKF. According to Fig. 4.1, the optimal localization radii for DfEnKF are 6 km and 3 km in horizontal and vertical directions respectively, which are much smaller than that obtained under perfect model assumption (16 km and 5 km for horizontal and vertical directions respectively, see section 3.3.1), indicating smaller localization radii are needed to obtain better performance in the imperfect-model scenario. The optimal cutoff radii of EnKF are similar to those of DfEnKF based on our results. The optimal cutoff radii of DfEnKF will be used as the cutoff radii for EnKF.

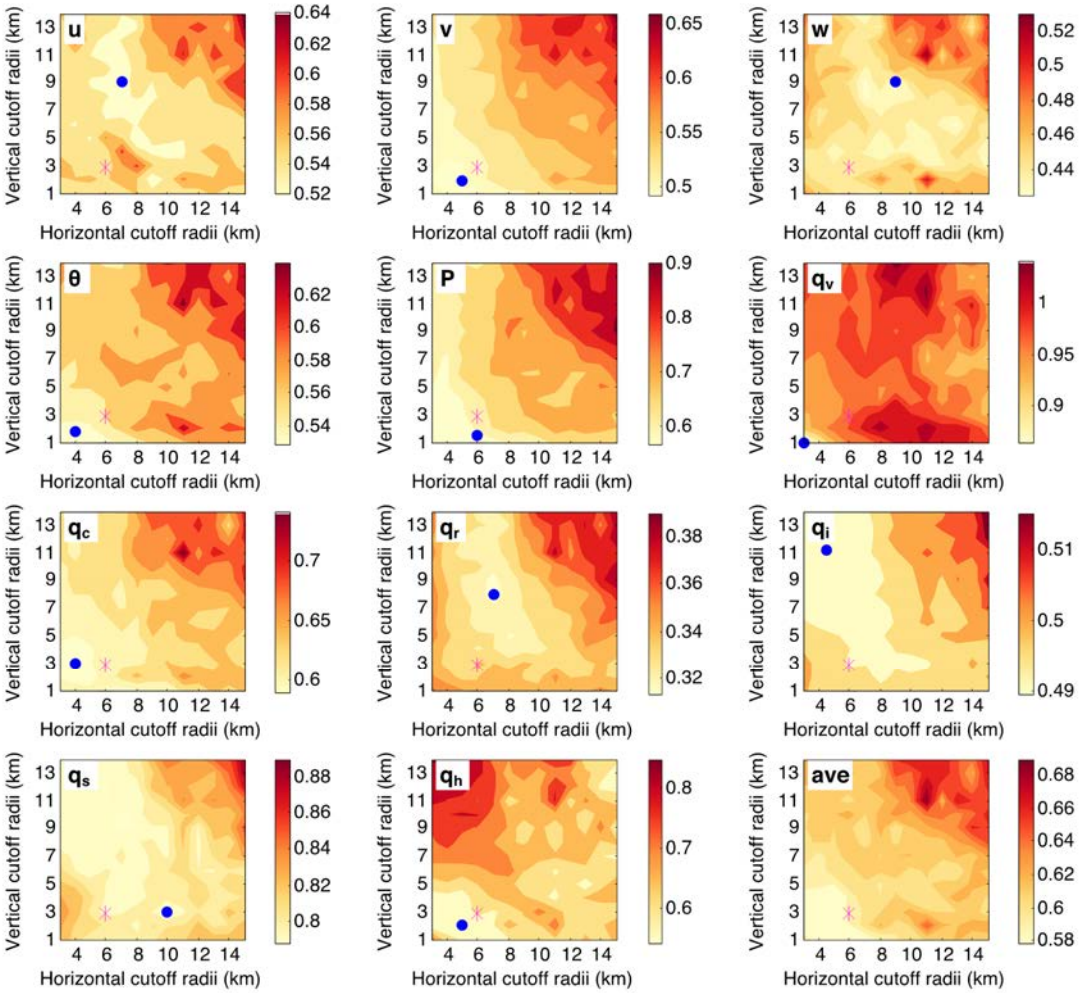

Fig. 4.1 The mean scaled RMSEs for different state variables as defined in Eq. (3.14) and averaged over all variables (ave), for DfEnKF experiments with different combinations of horizontal and vertical cutoff radius (km), with magenta asteroid indicating the minimum values

As indicated in the last chapter, the optimal localization radii of DfEnKF are not necessarily the optimal radii of pure En3DVar because of the differences between DfEnKF and pure En3DVar. Similarly, sensitivity experiments are conducted to obtain the optimal localization radii for pure En3DVar. According to the error statistics, the recursive filter length scale of 5.2 km (19 km cutoff) in horizontal and 2.2 km (8 km cutoff) in vertical yield the lowest error. In similar way, the optimal background error

de-correlation scales for 3DVar are also determined, which are 3.6 km and 1.1 km in horizontal and vertical respectively.

In hybrid En3DVar experiments, the optimal localization radii obtained for pure En3DVar and the optimal background error decorrelation scales obtained for 3DVar will be applied to the flow-dependent ensemble covariance and static covariance, respectively.

4.3.2 Optimal hybrid weights as function of ensemble size

Similar to the perfect-model OSSEs conducted in Chapter 3, the optimal hybrid weight in hybrid En3DVar is determined as functions of the ensemble sizes and the hybrid weights. 10×20 experiments with different combinations of the ensemble size (from 10 to 100 with increment of 10) and the hybrid weights (from 0 to 100% with increment of 5%) given to the static background error covariance are conducted. Pure En3DVar and 3DVar correspond to experiments with the weight of static background error being set to 0 and 1, respectively.

Similar to what have been done in the perfect-model OSSEs, the optimal weight of static background error covariance in hybrid En3DVar is also determined as functions of the ensemble sizes. 10 × 20 experiments with different combinations of the ensemble size (from 10 to 100 with increment of 10) and the hybrid weights (from 0% to 100% with increment of 5%) given to the static background error covariance are conducted. Pure En3DVar and 3DVar correspond to experiments with the weight of static background error being set to 0 and 1, respectively. Scaled RMSEs that are defined earlier are calculated for each experiment and the results are depicted in Fig. 4.2. Hybrid En3DVar performs much better than pure En3DVar and 3DVar with the optimal

weight of the static background error covariance ranged from 10% to 90% for most of the variables and for all the tested ensemble sizes, except for the analyses of cloud ice, and analyses of potential temperature and snow mixing ratio for small ensemble sizes (less than or equal to 30). On average, Hybrid En3DVar with weights ranged from 30% to 40% yield the best results for different ensemble sizes. When ensemble size is 40, the optimal weight is 30%.

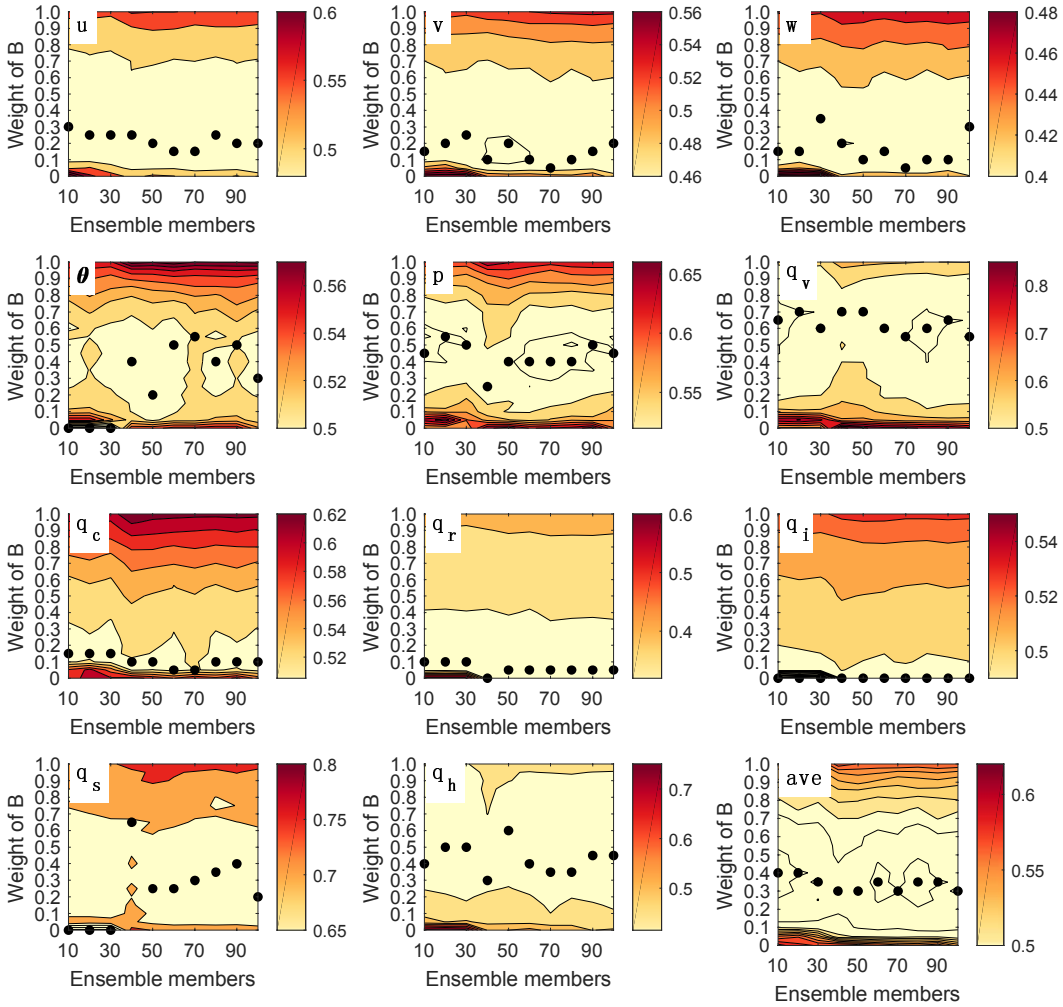

Fig. 4.2 The Grid-point average of the scaled RMSEs as defined in Eq. (3.14) for different variables and averaged over all variables (ave), which are obtained based on hybrid En3DVar experiments with different combinations of ensemble sizes and hybrid weights. The black dots indicate locations of the optimal weights for different ensemble sizes.

4.3.3 Assimilation of clear-air and precipitation reflectivity via double passes in variational DA

According to Tong and Xue (2005), clear-air reflectivity could help to suppress the spurious storms in EnKF analyses. Thomas et al. (2015) also demonstrated that clear-air reflectivity has the desired impact of reducing spurious convection early in the assimilation cycle of EnKF. Previous studies also indicate that when using hydrometeor-mixing ratios as the control variables, reflectivity DA by 3DVar may suffer from extremely large gradient of the cost function, and causes the assimilation of reflectivity and radial velocity data inefficient due to small background reflectivity (Sun and Crook, 1997; Wang and Wang, 2016). This problem is more obvious when including the clear-air reflectivity (< 5 *dBZ*) into reflectivity DA relative to those do not. Because the background reflectivity tends to be small in the area where the observed reflectivity is also small (e.g., <5dBZ), especially after one or two cycles when the location errors of convective storms are greatly decreased by reflectivity DA. Thus the assimilation of the clear-air reflectivity involves more observations, on which the background reflectivity is very small, being included in to the calculation of the gradient, making it excessively large and the assimilation of radial velocity and precipitation reflectivity ineffective. To solve the issue, a double-pass procedure was then proposed and found to be able to greatly alleviate the problem. In the double-pass procedure, the radial velocity and precipitation reflectivity are assimilated together in the first pass, and the clear-air reflectivity is assimilated separately in the second pass. The analyses and forecasts based on the double-pass DA experiments are compared with their single-

pass counterparts in both 3DVar and hybrid En3DVar DA. EnKF, and DfEnKF were then compared with the double-pass 3DVar and hybrid En3DVar.

a. Double-pass 3DVar

Performance of the single-pass and double-pass 3DVar analyses in radar data assimilation are compared based on the RMSEs of the analyses and forecasts at different analysis times (错误!未找到引用源。). The 3DVar analysis has been greatly improved in terms of smaller root mean squared error by applying the double-pass procedure to assimilation of clear-air and precipitation reflectivity. The improvements are most obvious for velocity component v, snow and hail mixing ratios (q_s and q_h). According to the reflectivity observation operator used in chapter 3, the reflectivity observation operator is expressed as follows,

$$Z = 10\log_{10}^{Z_e} = 10\log_{10}^{\left(a_r(\rho q_r)^{1.75} + a_s(\rho q_s)^{1.75} + a_h(\rho q_h)^{1.6625}\right)} . \quad (4.1)$$

The derivative of Z respect to individual hydrometeor mixing ratios q_x (x = r, s, h indicative of rain, snow, and hail), taking rain-water mixing ratio q_r as example, can be expressed as

$$\frac{\partial Z}{\partial q_r} = \frac{10}{Z_{eb}} \times 1.75 \times a_r \times \rho\left(\rho q_{rb}\right)^{0.75} . \quad (4.2)$$

Where

$$Z_{eb} = a_r\left(\rho q_{rb}\right)^{1.75} + a_s\left(\rho q_{sb}\right)^{1.75} + a_h\left(\rho q_{hb}\right)^{1.6625} . \quad (4.3)$$

Eq. (4.2) is not valid when Z_{eb} is zero. In the NWP models, the magnitude of hydrometeor mixing ratios is usually from 0 to 10^{-2} kg kg^{-1}. If all of the hydrometeors mixing ratios are equal to 0 kg kg^{-1}, Z_{eb} in the denominator is also zero. To avoid this

problem, the hydrometeor mixing ratios are usually given a lower limit in the applications (such as 10^{-8} kg kg^{-1} used in this study). As a result, the orders of magnitude for Z_{eb} is 10^{-3} mm^{6}m^{-3}, for $\partial Z/\partial q_r$ and $\partial Z/\partial q_s$ are 10^6 dBZ kg kg^{-1}, and for $\partial Z/\partial q_h$ is 10^7 dBZ kg kg^{-1} when hydrometeors mixing ratios are all as small as 10^{-8} kg kg^{-1} (to facilitate the demonstration, ρ is assumed to be 1.0 kg kg^{-1}), leading to excessively large gradient of the cost function. On the other hand, when hydrometeor-mixing ratios are as large as 10^{-2} kg kg^{-1}, the order of magnitude for Z_{eb} is 10^7 mm^{6}m^{-3}, for $\partial Z/\partial q_r$, $\partial Z/\partial q_s$, and $\partial Z/\partial q_h$ are all 10^1 dBZ kg kg^{-1}, which are much smaller than those of the small value hydrometeor mixing ratios (10^{-8} kg kg^{-1}). Thus the assimilation of reflectivity observation is ineffective in the area where background reflectivity is large. By applying a double-pass procedure in assimilation of radial velocity, precipitation reflectivity, and clear-air reflectivity, the assimilation of much more gradient-consuming clear-air reflectivity won't influence the assimilation of radial velocity and precipitation reflectivity. In this way, the spurious storms are suppressed by the clear-air reflectivity, and at the same time, the reflectivity analyses in the storm region won't be influenced by assimilation of the clear-air reflectivity.

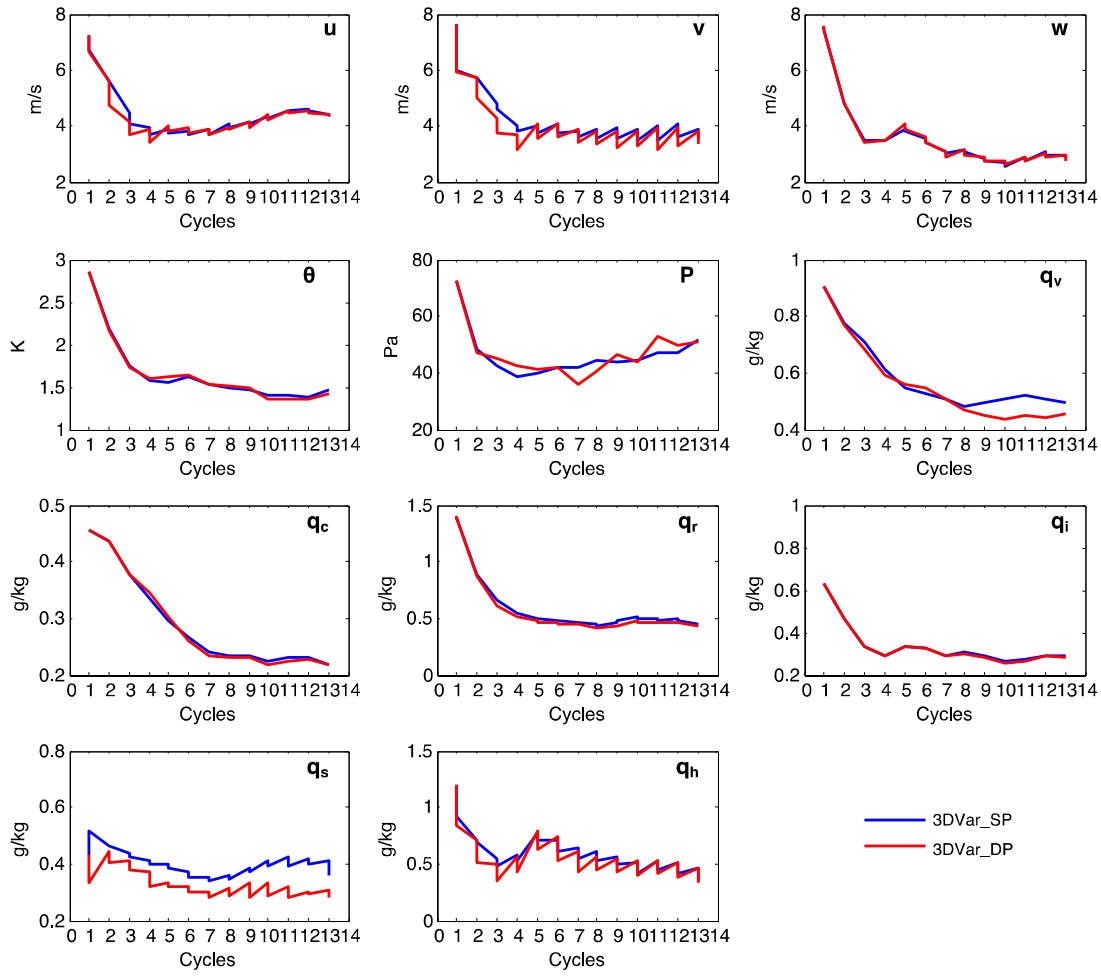

Fig. 4.3 Grid-point averages of the root mean squared error for single (blue contour) and double pass (red line) 3DVar analyses.

Vertical cross-sections of reflectivity analyses and forecasts at the end of 1-hour DA cycle for 3DVar are compared and the results are depicted in Fig. 4.4. By applying the double-pass procedure in assimilation of 'clear-air' and precipitation reflectivity, the maximum reflectivity analysis at the end of the DA cycle from 3DVar has been increased from 59.0 dBZ to 62.7 dBZ, which is much closer to the truth (72.2 dBZ). The maximum vertical velocity is also increased from 2.8 ms^{-1} to 7.0 m s^{-1}, though it is still much weaker than that of the truth (40.2 ms^{-1}). When examine the analyzed

reflectivity calculated based on different hydrometeors according to Eq. (3.16) in Chapter 3, the analyses of the rain, snow, and hail mixing ratio analyses and forecasts have been significantly improved. The analysis of rainwater, dry snow, and hail from the single-pass experiments are all much stronger than those of the single-pass, and are much closer to the truth.

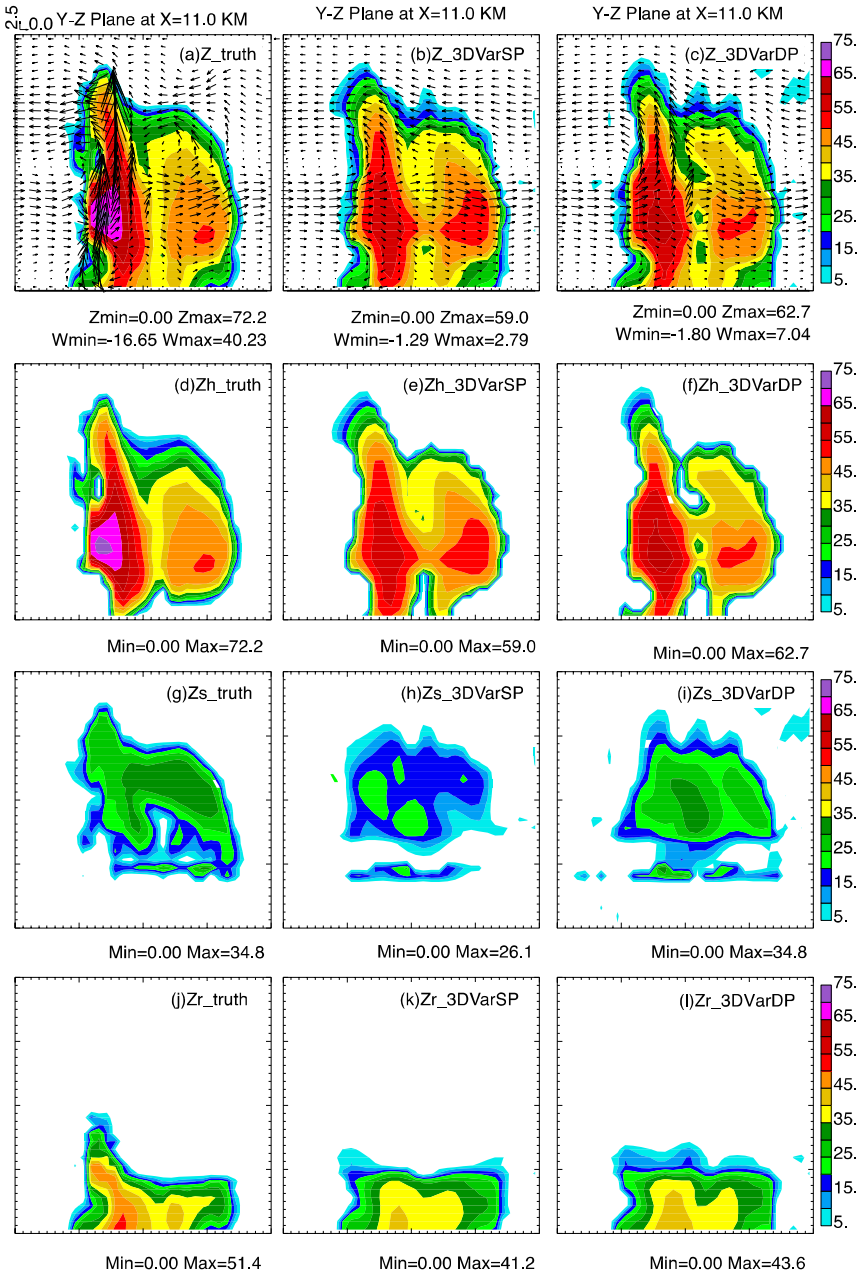

Fig. 4.4 Vertical cross-sections of the reflectivity through the maximum vertical velocity of truth overlaid with the wind vector at the end of the one-hour DA window (a, b, c), and reflectivity calculated based on hail (d, e, f), snow (g, h, i), and rain water mixing ratio (i, k, l) fields from truth (a, d, g, i), single-pass 3DVar analyses (b, e, h, k) and double-pass 3DVar analyses (c, f, i, l).

b. Double-pass hybrid En3DVar

Similar double-pass procedure is applied to hybrid En3DVar as that to 3DVar, with the radial velocity and precipitation reflectivity being assimilated in the first pass, and the clear-air reflectivity in the second. In this way, the influence of clear-air reflectivity assimilation to precipitation reflectivity due to its excessively large value of the gradient can be avoided.

Performances of the single-pass and double-pass hybrid En3DVar analyses in radar data assimilation are compared based on the RMSEs of the analyses and forecasts at different analysis times (Fig. 4.5). The 3DVar analysis has been greatly improved in terms of smaller root mean squared error by applying the double-pass procedure to assimilation of clear-air and precipitation reflectivity. The improvements are most obvious for analyses of snow and hail mixing ratios (q_s and q_h).

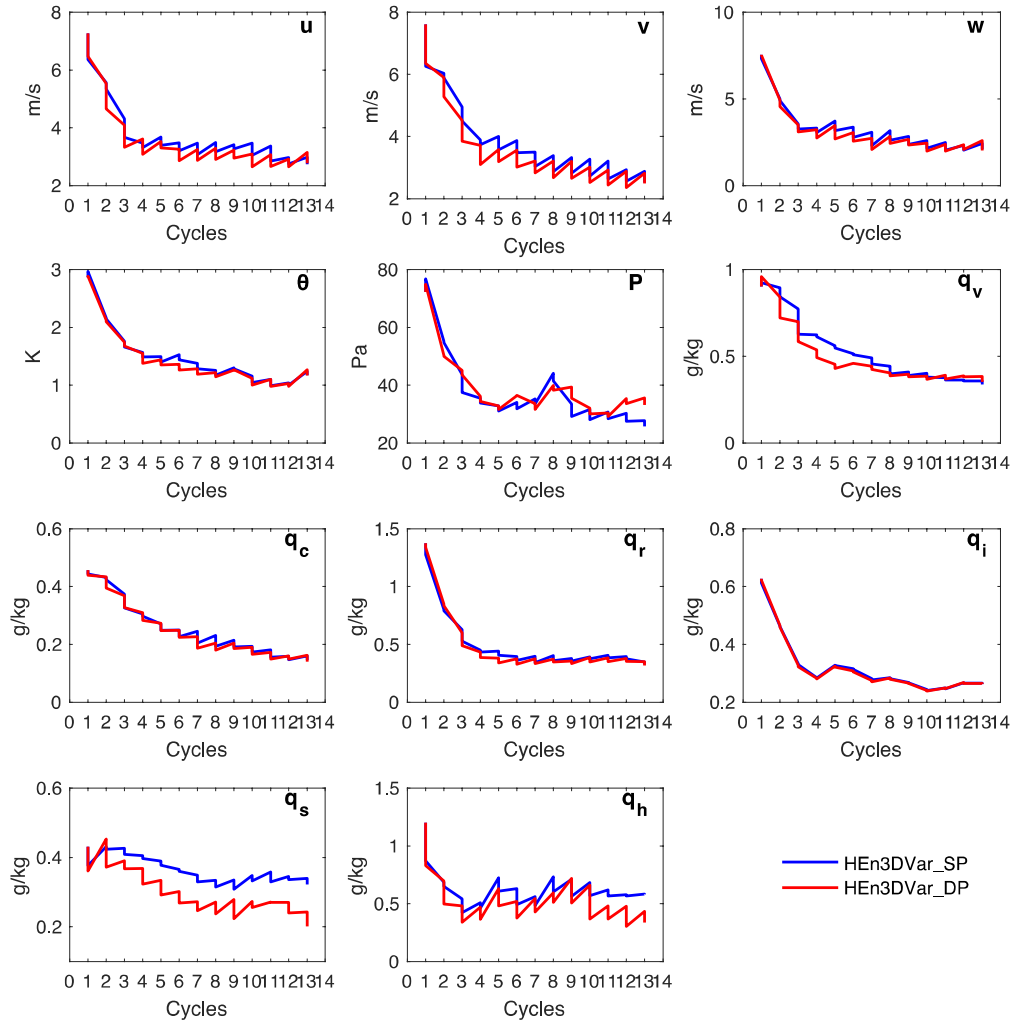

Fig. 4.5 The Grid-point averages of the root mean squared error for single (blue contour) and double pass (red line) hybrid En3DVar analyses.

Based on the comparison of the surface reflectivity between single and double pass experiments (Fig. 4.6), the spurious storms (5 ~ 20 dBZ) outside the storm region are suppressed, and at the same time the intensity of the storm center is much stronger and closer to the true condition after applying the double-pass procedure.

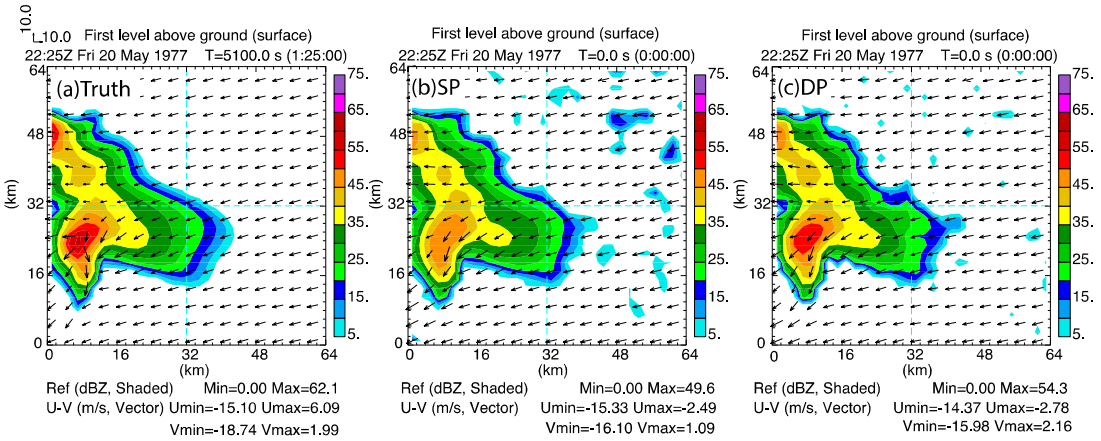

Fig. 4.6 Surface reflectivity (shaded contour, dBZ) overlaid with wind vector field (ms^{-1}) from (a) truth, (b) single-pass hybrid En3DVar analysis, and (c) double-pass hybrid En3DVar analysis.

The vertical cross-section of the analyzed reflectivity and the reflectivity calculated based on analyses of different hydrometeor mixing ratios are compared between single- and double- pass hybrid En3DVar (Fig. 4.7). The maximum analyzed reflectivity calculated based on snow mixing ratio is 31.6 dBZ for the single-pass hybrid En3DVar, which is increased to 33.8 dBZ after applying the double-pass procedure, and is much closer to the truth (34.8 dBZ). The distribution of snow mixing ratio from double-pass hybrid En3DVar is much closer to the truth as well. The hail analyses from single and double pass hybrid En3DVar are similar, except that the vertical extending of the hail from double-pass hybrid En3DVar is a little bit lower than that from the single-pass. The updraft intensity from double-pass hybrid En3DVar is 27.0 ms^{-1}, which is larger than that of the single-pass hybrid En3DVar (24.8 ms^{-1}), and they are both smaller than the true updraft (40.2 ms^{-1}). In the following comparisons, results from double-pass 3DVar, pure and hybrid En3DVar will be used to compare with EnKF and DfEnKF.

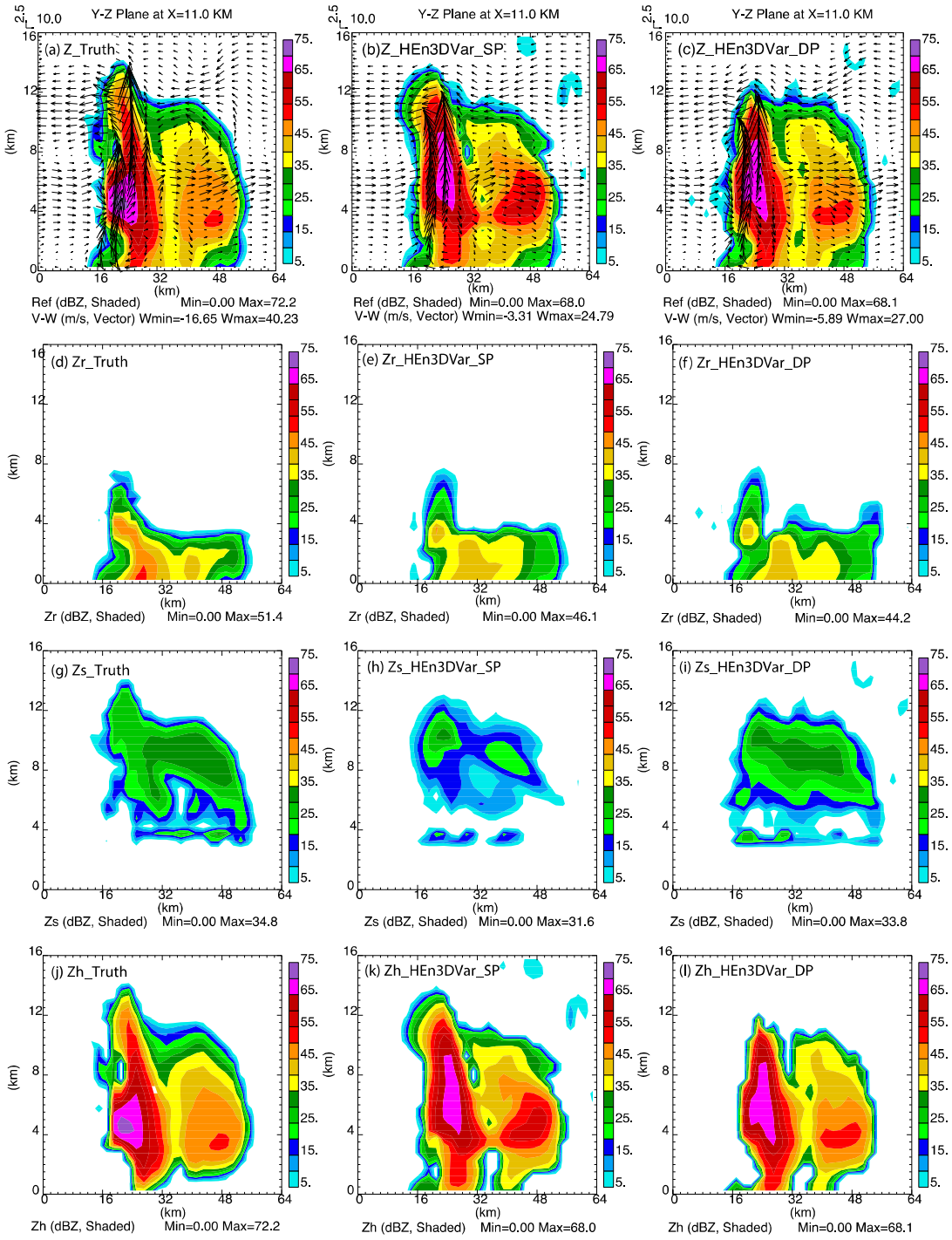

Fig. 4.7 Vertical cross-sections of the reflectivity through the maximum vertical velocity of truth overlaid with the wind vector at the end of the one-hour DA window (a, b, c), and the reflectivity calculated based on rainwater (d, e, f), snow (g, h, i), and hail mixing ratio (j, k, l) fields from truth (a, d, g, j), single-pass hybrid En3DVar analyses (b, e, h, k) and double-pass 3DVar analyses (c, f, i, l).

c. Problems of using the double-pass procedure

Although encouraging results have been obtained by using the double-pass procedure in reflectivity DA. There still exist some problems in the double-pass procedure. First, small background reflectivity can still be included into calculation of the gradient in the first pass, and thus influence the assimilation of radial velocity and precipitation reflectivity, especially when significant location errors exist. In this case, observed clear-air reflectivity can no longer represent the region of small background reflectivity. Second, the background error covariance needs to be updated after assimilating radial velocity and precipitation in the first pass, using the same background error covariance in the second pass as the first one will introduce errors into the DA process. Third, the double-pass procedure involves minimizing two different cost functions with different orders of magnitude separately, which is a lot more expensive. For real case study in Chapter 5, an alternative way of using logarithmic mixing ratios as the control variables was used to dealing with the excessive large gradient caused by small background reflectivity and at the same time avoided using of the double-pass procedure.

4.3.4 Comparisons of hybrid En3DVar with 3DVar, EnKF, DfEnKF, and pure En3DVar with optimal configurations

Comparisons are made among the RMSEs of the analyses and forecasts from 3DVar, 40-member EnKF, DfEnKF, pure and hybrid En3DVar with their own optimal configurations. RMSEs are calculated in the region where the true reflectivity exceeds 15 dBZ. As is indicated in Fig. 4.8, 3DVar performs the worst with its RMSE being the largest, except for mixing ratios of water vapor and hail. Hybrid En3DVar obviously

outperforms the other algorithms. The outperformance is most obvious for analyses of velocity, mixing ratios of water vapor, cloud ice, and hail. 3DVar and hybrid En3DVar outperform the pure ensemble based DA algorithms (EnKF, DfEnKF, and pure En3DVar) for analyses of water vapor mixing ratio, indicating the cross-correlation between mixing ratios of water vapor and hydrometeor variables is not reliable when severe microphysical error exists. In this case, updating water vapor mixing ratio field by reflectivity assimilation does not benefit the analyses and related forecasts. EnKF outperforms DfEnKF for velocity components of u, v, and cloud ice mixing ratio at later cycles. DfEnKF outperforms pure En3DVar for velocity analyses at later cycles, and under performs for analysis of hail mixing ratio.

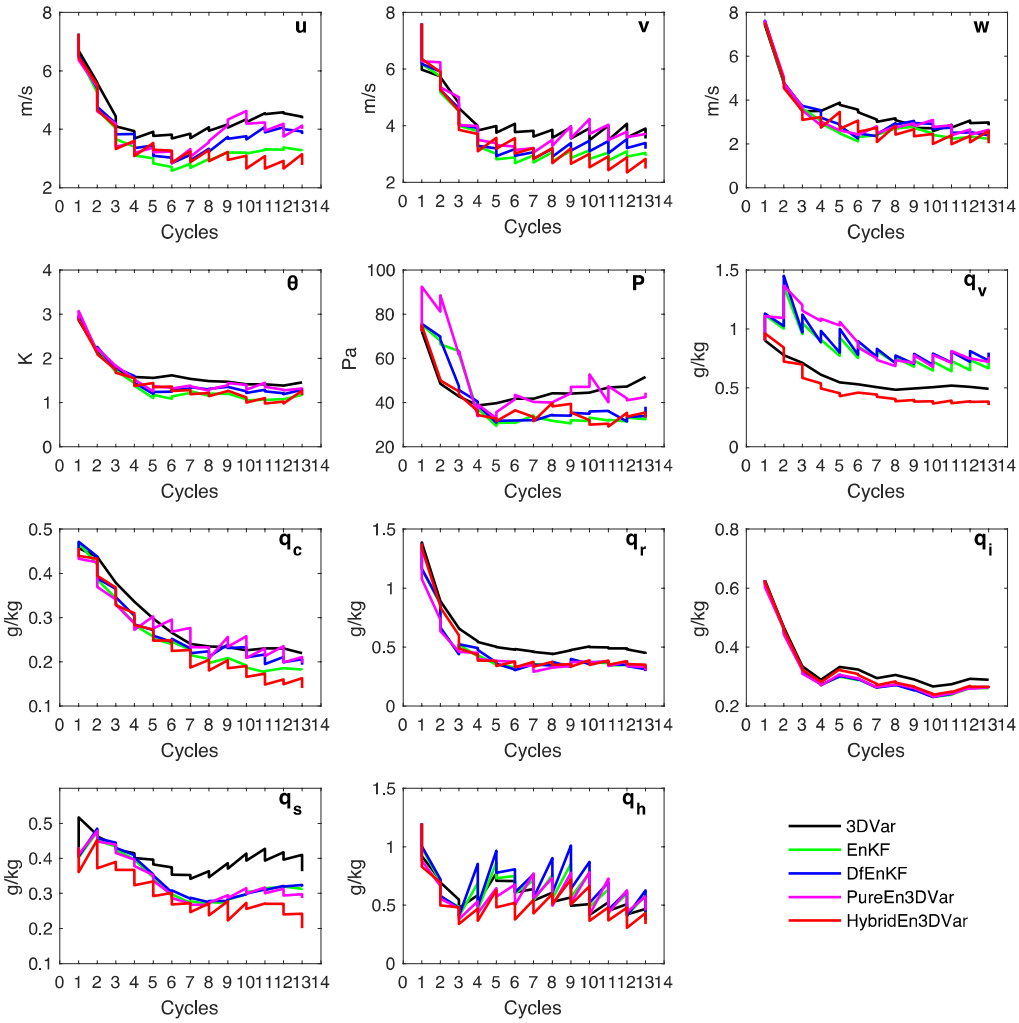

Fig. 4.8 The Grid-point averages of the root mean squared error for 3DVar, EnKF, DfEnKF, pure En3DVar, and hybrid En3DVar (30% weight given to static B).

Cross-sections of the reflectivity analyses at the end of one-hour DA window are compared among different experiments (Fig. 4.9). The total reflectivity analyses from EnKF, DfEnKF, and hybrid En3DVar are similar, which all outperform 3DVar for better capture the intensity of the storm center. The maximum reflectivity from EnKF, DfEnKF, pure and hybrid En3DVar are 67.2 dBZ, 68.1 dBZ, 67.9 dBZ, and 68.6 dBZ respectively, which are similar and much larger than that of 3DVar (58.6 dBZ), and is much closer to that of the truth (72.2 dBZ). The relative weaker storm center in the

downshear direction is better captured by 3DVar and hybrid En3DVar relative to EnKF, DfEnKF, and pure En3DVar for having more smoothed storm structure. The updraft intensity from EnKF, DfEnKF, pure and hybrid En3DVar are 31.6 ms-1, 31.6 m/s, 30.8 m/s, and 32.26 m/s respectively, which are all much larger than that of 3DVar (5.38 m/s), and closer to the truth (40.2 m/s).

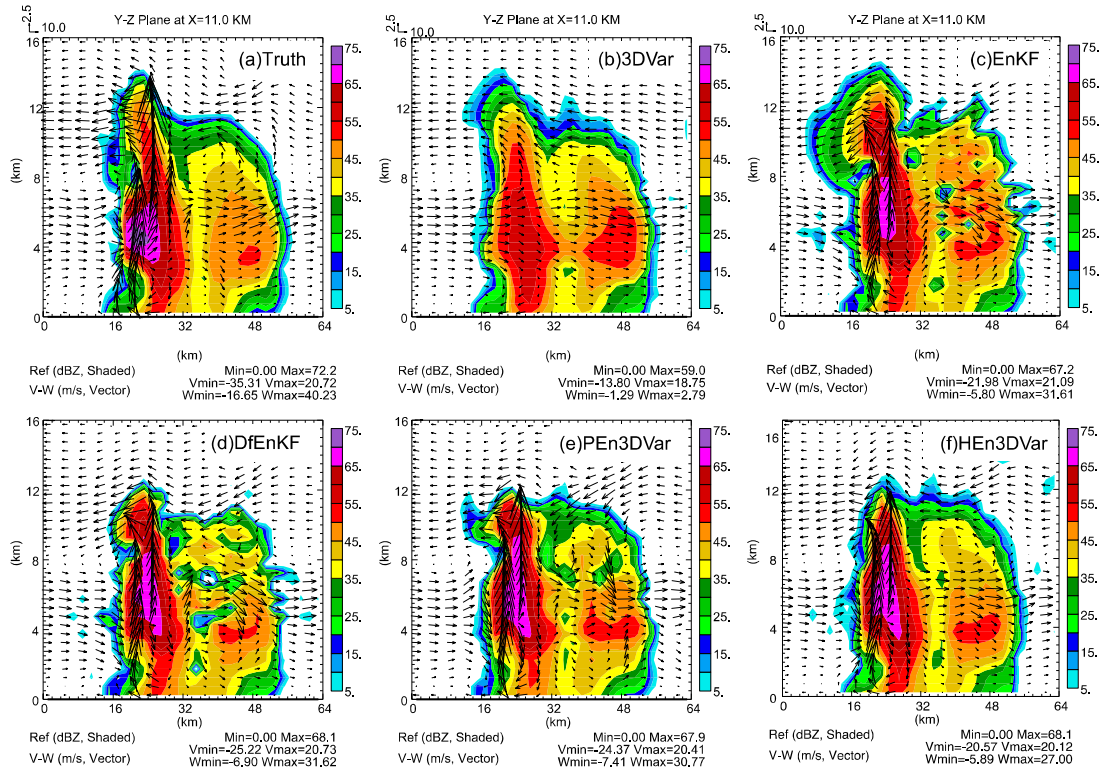

Fig. 4.9 Vertical cross-sections of the reflectivity field through the maximum vertical velocity of truth overlaid with the wind vector at the end of the one-hour DA window from (a) truth, (b) 3DVar, (c) EnKF, (d) DfEnKF, (e) pure En3DVar, and (f) hybrid En3DVar with 30% static covariance.

Considering the difference in order of magnitude in mixing ratio analyses for different hydrometeors, the hydrometer mixing ratio analyses are compared by calculating the reflectivity based on single type mixing ratio. According to Fig. 4.10, both EnKF and DfEnKF overestimated the rainwater mixing ratio analyses, while

3DVar and pure En3DVar underestimated the rainwater mixing ratio analyses. Pure En3DVar has reflectivity intensity more close to the truth for the left center, but with the right center still overestimated.

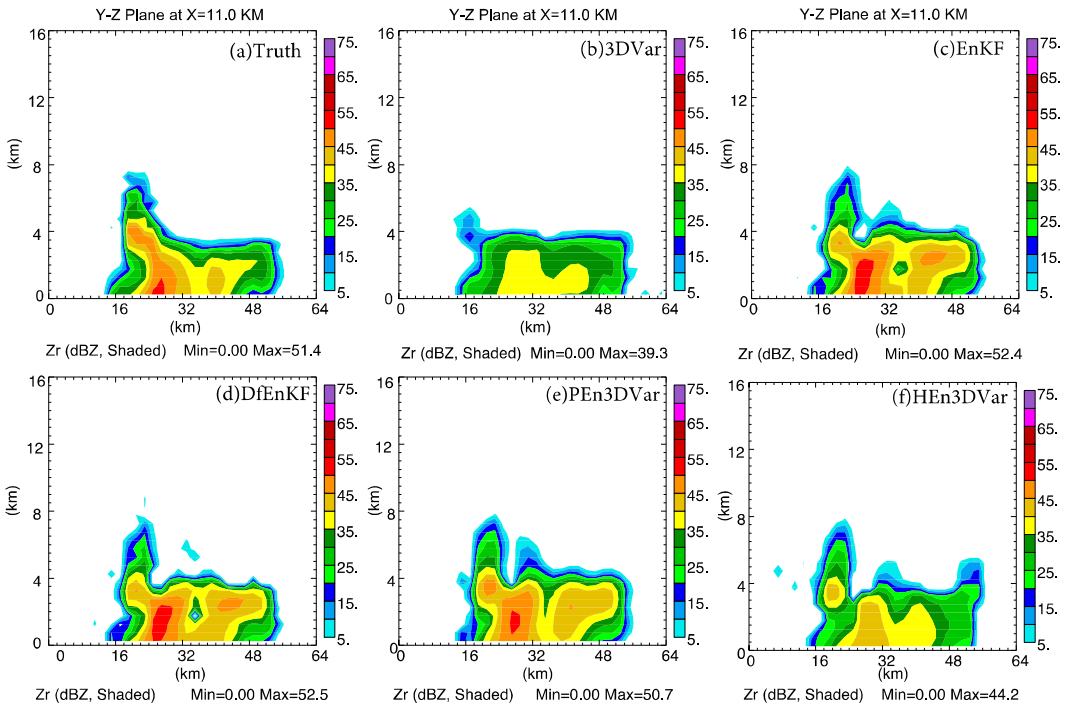

Fig. 4.10 Vertical cross-sections of the reflectivity calculated based on mixing ratio of rain water mixing ratio through the maximum vertical velocity of truth and at the end of the one-hour DA window from (a) truth, (b) 3DVar, (c) EnKF, (d) DfEnKF, (e) pure En3DVar, and (f) hybrid En3DVar with 30% static covariance.

When checking the analyses of mixing ratio of hail from each analysis, hybrid En3DVar algorithm outperforms the other methods (3DVar, EnKF, DfEnKF, and pure En3DVar) in producing better analysis in storm intensity and structure (Fig. 4.11). Low-level hail that is missed in the analysis of EnKF, DfEnKF, and pure En3DVar is captured by hybrid En3DVar at the end of one-hour data assimilation window. For WSM6 scheme, the hail/graupel melted too fast as it fell near the ground, so that there is no hail in the background field. When calculating the ensemble covariance based on

such ensemble background fields, the error variance and the ensemble spread (Fig. 4.12) for hail in the low level is severely under-estimated. That's why pure ensemble-based DA methods (EnKF, DfEnKF, and pure En3DVar) cannot generate hail analysis in the low levels. By including the static **B** into hybrid En3DVar, the low-level hail that is absent in the pure ensemble-based DA experiments are captured. In addition, hybrid En3DVar also outperforms pure 3DVar in better capturing the intensity of the hail core.

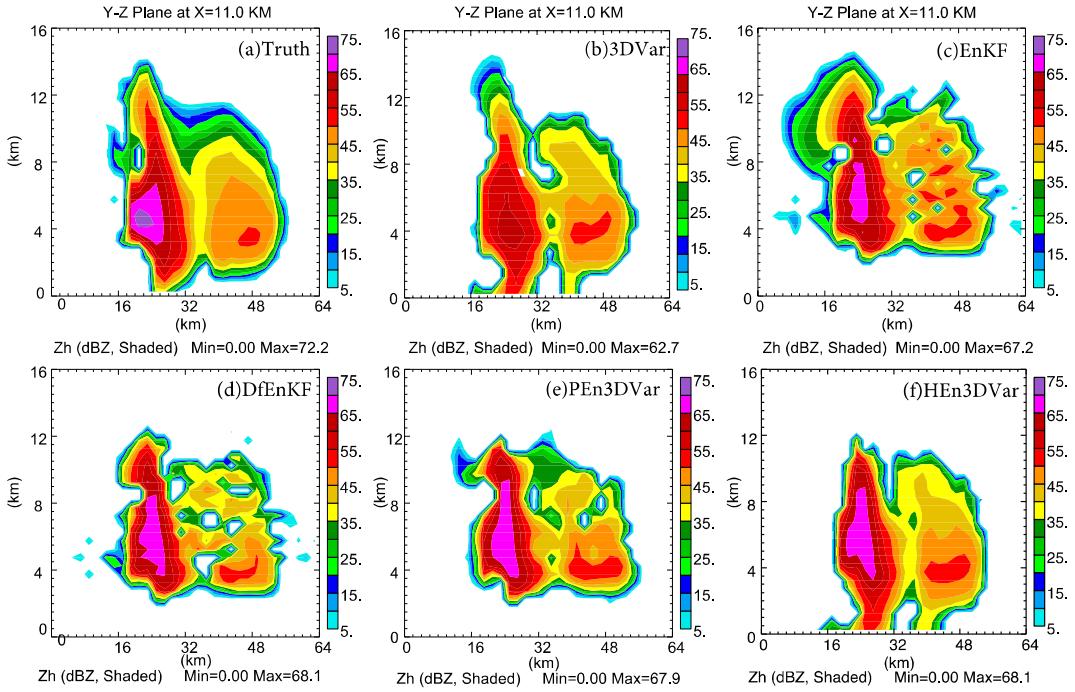

Fig. 4.11 Vertical cross-sections of the reflectivity calculated based on mixing ratio of hail through the maximum vertical velocity of truth at the end of the one-hour DA window from (a) truth, (b) 3DVar, (c) EnKF, (d) DfEnKF, (e) pure En3DVar, and (f) hybrid En3DVar with 30% static covariance.

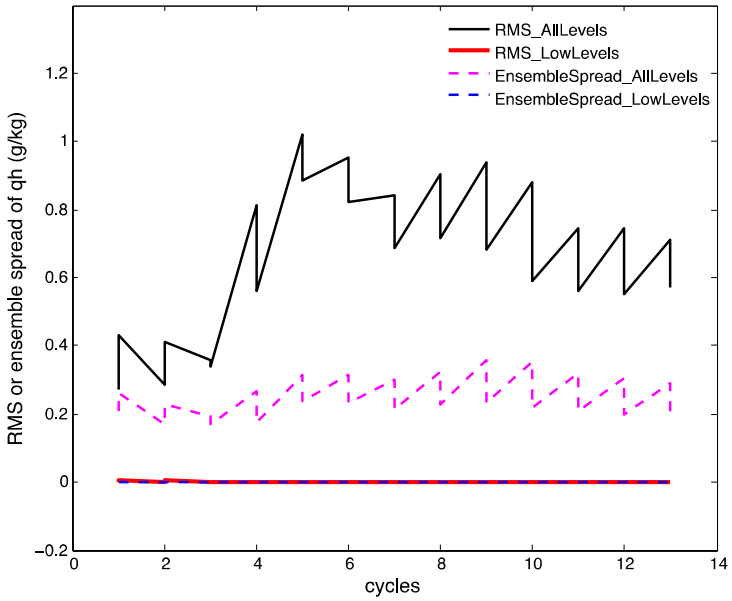

Fig. 4.12 Comparisons of the root mean squared value of mixing ratio of hail between all levels and low levels (<=2km), and ensemble spread at all levels and low levels (<=2km) respectively, unit: g/kg.

The analyzed reflectivity calculated based on mixing ratio of snow (denote the reflectivity as Z_s) from 3DVar and hybrid En3DVar are 34.8 and 33.8 dBZ, which are larger than those from EnKF (31.7 dBZ), DfEnKF (33.3 dBZ), and are closer to the truth (34.8 dBZ) (Fig. 4.13). Although the maximum Z_s from pure En3DVar (36.1 dBZ) is larger than those of the others, the Z_s from pure En3DVar is still underestimated, with the coverage of the reflectivity larger than 30 dBZ being much smaller than those of 3DVar, hybrid En3DVar, as well as the truth.

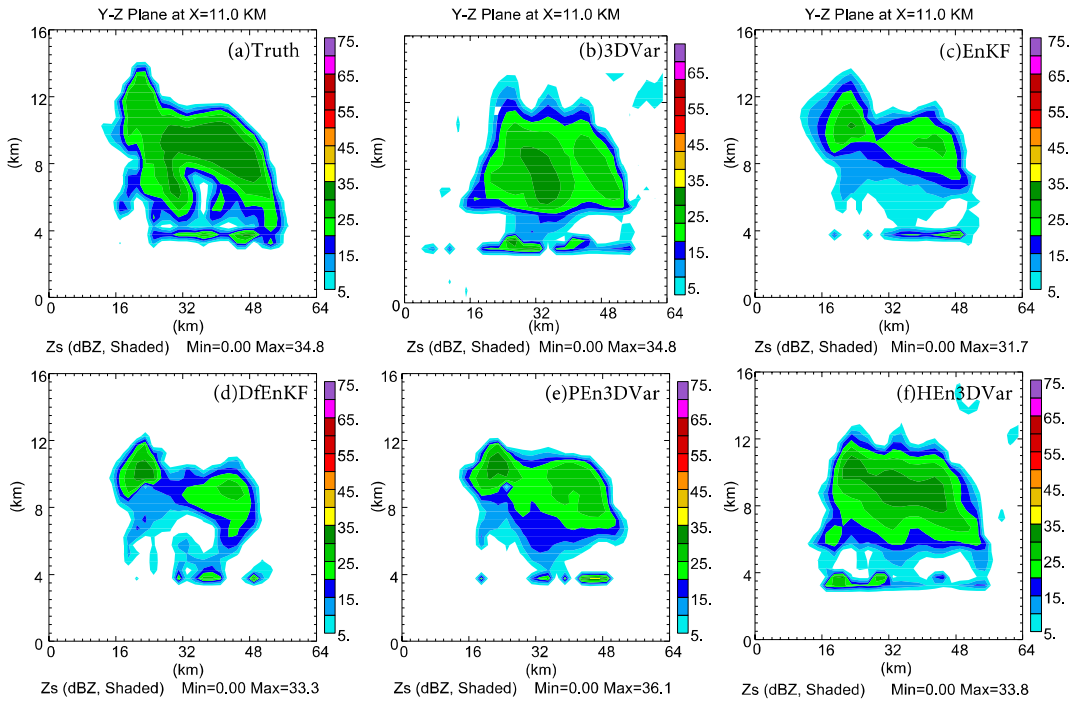

Fig. 4.13 Vertical cross-sections of the reflectivity calculated based on mixing ratio of snow through the maximum vertical velocity of truth at the end of the one-hour DA window from (a) truth, (b) 3DVar, (c) EnKF, (d) DfEnKF, (e) pure En3DVar, and (f) hybrid En3DVar with 30% static covariance.

4.3.5 Impacts of using the mass-continuity constraint

An advantage of the hybrid En3DVar over EnKF is that the hybrid DA algorithm allows the weak constraint (such as the mass continuity constraint) being included into the cost function. Gao et al. (1999) and Hu et al. (2006a) reported some benefits of using the mass continuity constraint in radar radial velocity DA based on 3DVar. For hybrid radar DA, Gao et al. (2016) shows that incorporating a mass continuity equation as a weak constraint into the En3DVar algorithm can improve the quality of the analyses when radial velocity observations contain large errors. To further investigate its impact, experiments with and without the mass-continuity constraint are compared among radar DA from 3DVar, pure and hybrid En3DVar. The corresponding

experiments are denoted as "3DVar_W/C", "3DVar_WC", "PEn3DVar_W/C", "PEn3DVar_WC", "HEn3DVar_W/C", and "HEn3DVar_WC" respectively, with the suffix "_W/C" and "_WC" representing "without" and "with" the mass continuity constraint. The mass continuity constraint can be achieved by adding a term J_{div} to the cost function, which is

$$J_{div} = \frac{1}{2}\left(w_h\left(\frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y}\right) + w_v\left(\frac{\partial \rho w}{\partial z}\right)\right)^2. \quad (4.1)$$

Here ρ is the air density; w_h (w_v) is the coefficient of the horizontal (vertical) divergent term. Experiments with different levels (10^1, 10^2, 10^3, 10^4, 10^5) of the coefficients are tested. Based on the experiments, $w_h = w_v = 10^3$ yields the smallest RMSEs for 3DVar (pure and hybrid En3DVar).

The RMSEs of the analyses and forecasts within the DA window are compared between 3DVar DA with and without adding the weak constraint. As is indicated in Fig. 4.14. The impacts of the weak constraint to 3DVar analyses and forecasts are not obvious. There is only small improvement made to *w* after including the weak constraint. The reflectivity analyses at the end of one-hour DA window also resemble form the experiments with and without adding the weak constraint (Fig. 4.15). The intensity of the updraft from 3DVar_W/C and 3DVar_WC are 7.42 ms^{-1} and 6.69 ms^{-1}, which are all much smaller than the truth (40.23 ms^{-1}). The benefit of the mass continuity to analysis and forecasts of vertical velocity is not obvious for convective scale radar DA using 3DVar.

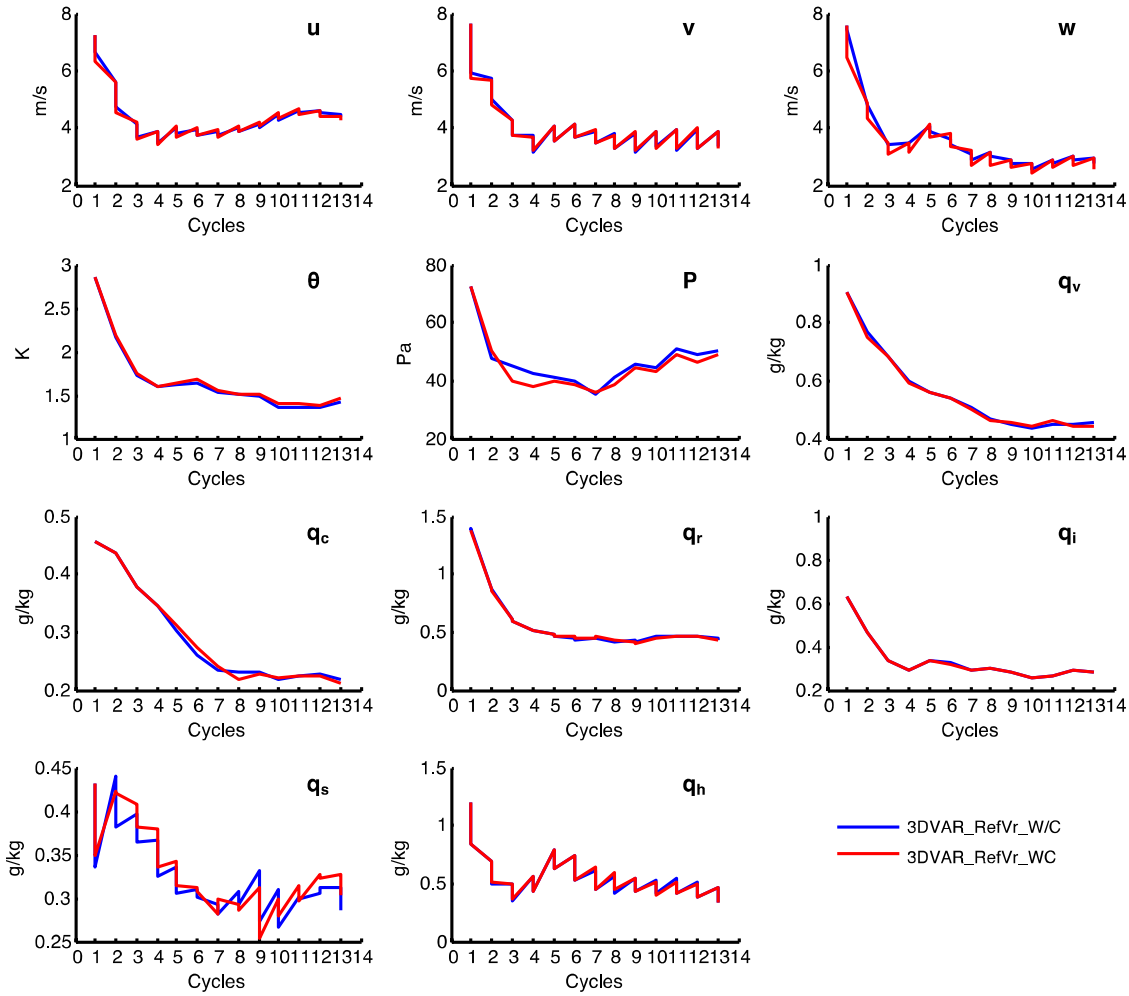

Fig. 4.14 The Grid-point averages of the root mean squared error for 3DVar without (blue contour) and with (red contour) adding the weak constraint.

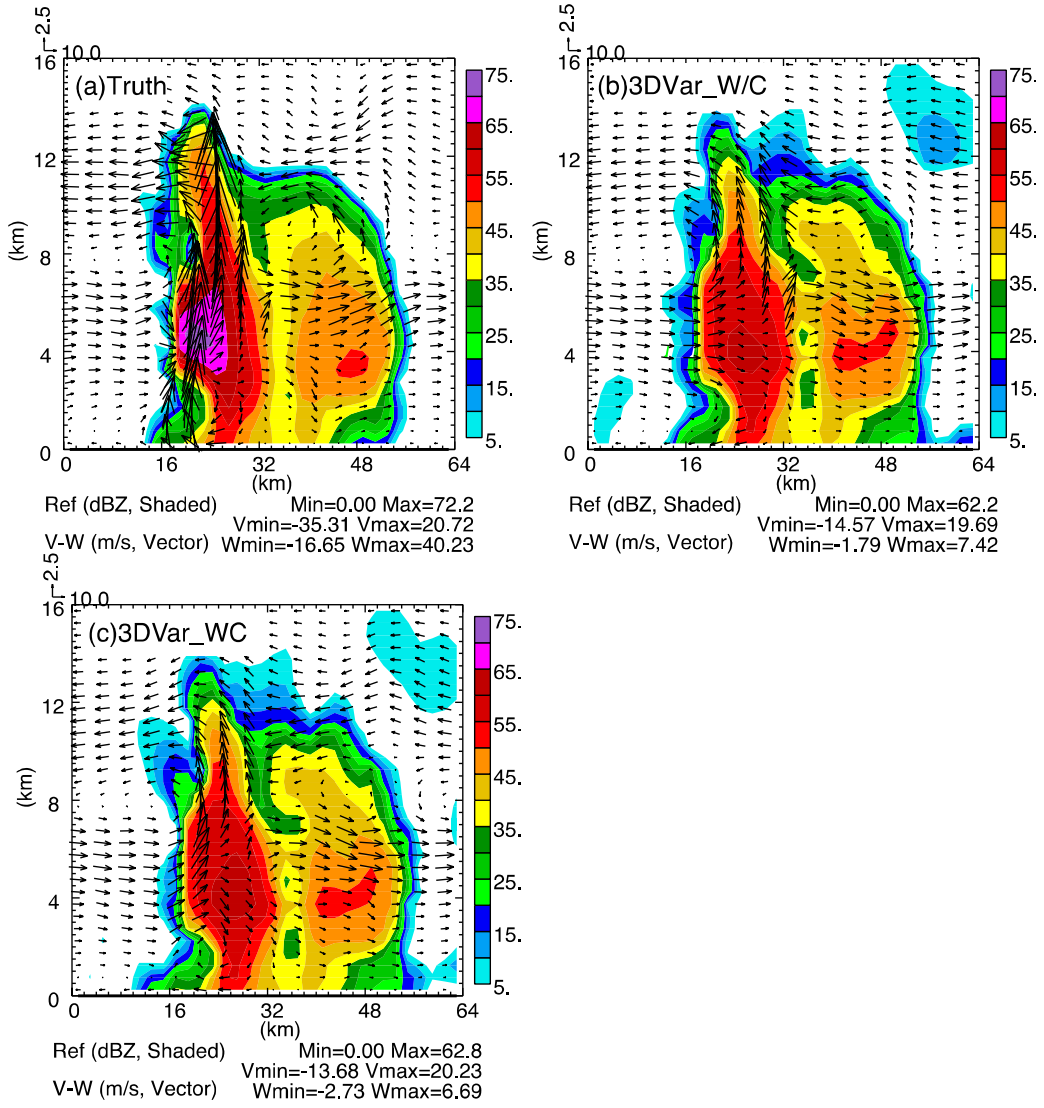

Fig. 4.15 Vertical cross-sections of the reflectivity field through the maximum vertical velocity of truth overlaid with the wind vector at the end of the one-hour DA window from (a) truth, 3DVar without (b), and with (c) the mass continuity constraint.

The RMSEs of the analyses and forecasts are compared between pure En3DVar DA with and without adding the weak constraint. As is indicated in Fig. 4.16. The impacts of the weak constraint to analyses (forecasts) of pure En3DVar are more obvious than those for 3DVar. By adding the mass continuity constraint to the cost function, the analyses and forecasts of pure En3DVar has been greatly improved,

especially for velocity fields at later cycles and water vapor mixing ratio field. The reflectivity analyses at the end of one-hour DA window indicate that both the velocity and the reflectivity analyses for the second storm center have been greatly improved (Fig. 4.17). By adding the mass continuity constraint, both the velocity and the reflectivity become more smoothed, and are much closer to the truth.

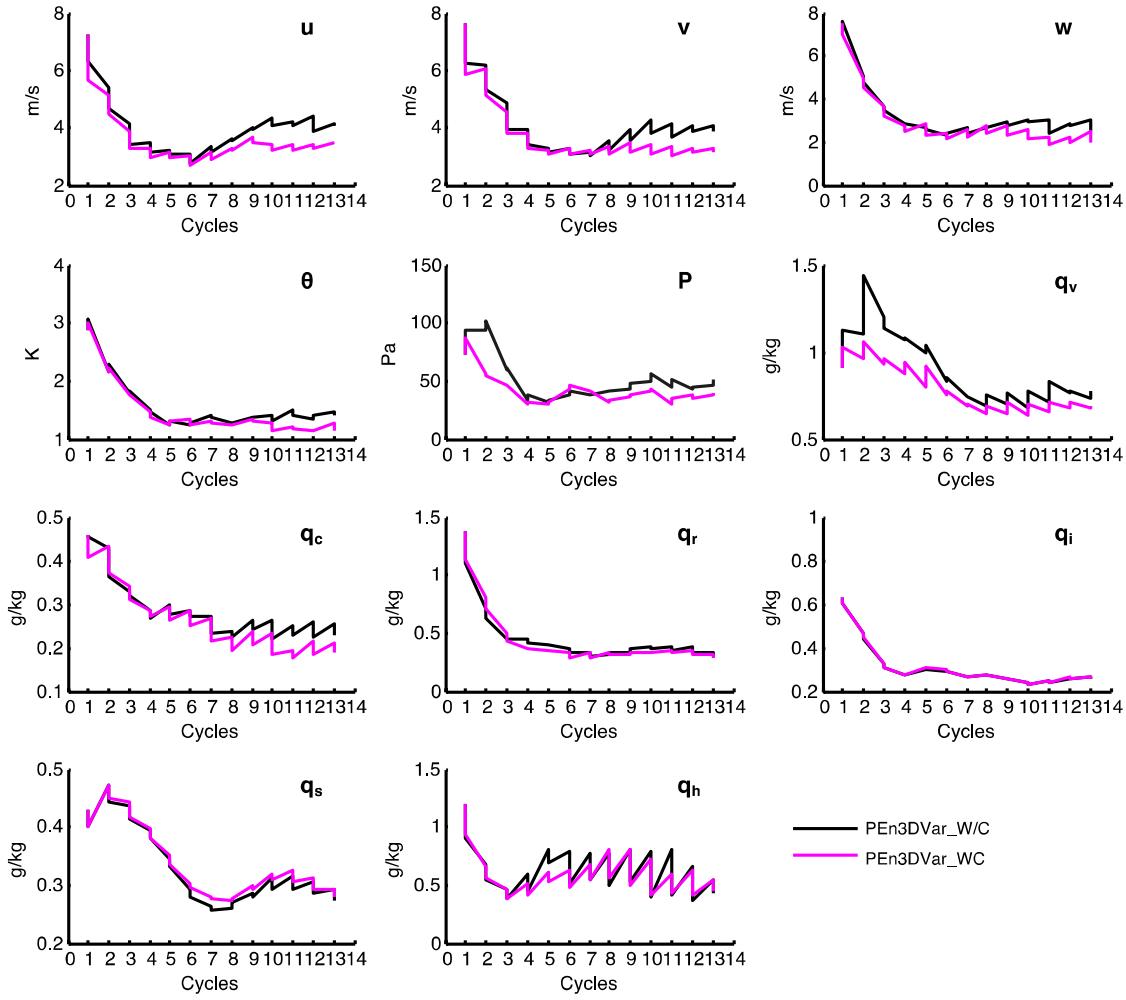

Fig. 4.16 The Grid-point averages of the root mean squared error for 3DVar without (blue contour) and with (red contour) adding the weak constraint.

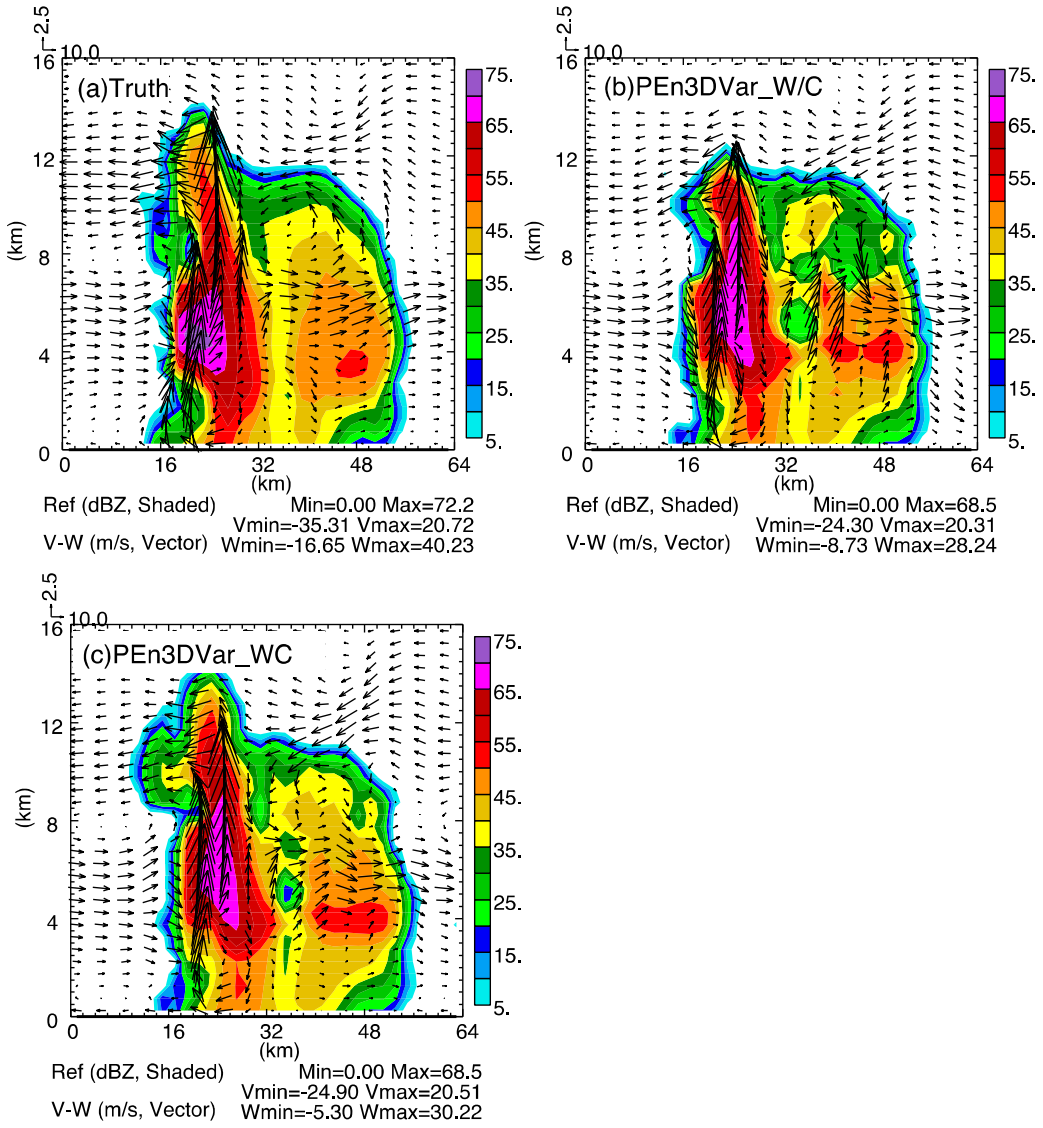

Fig. 4.17 Vertical cross-sections of the reflectivity field through the maximum vertical velocity of truth overlaid with the wind vector at the end of the one-hour DA window from (a) truth, PEn3DVar without (b), and with (c) the mass continuity constraint.

Similarly, the RMSEs of the analyses and forecasts are also compared between hybrid En3DVar DA with and without adding the weak constraint. As is indicated in Fig. 4.18. The impacts of the weak constraint to analyses (forecasts) of hybrid En3DVar are not obvious in terms of similar RMSEs. The reflectivity analyses at the end of one-hour DA window also resemble form the experiments with and without adding the weak

constraint (Fig. 4.15). The intensity of the updraft from HEn3DVar_W/C is 28.1 ms^{-1}, which is increased to 29.5 ms^{-1} by HEn3DVar_WC, and is a little bit closer to that of the truth (40.23 ms^{-1}). The maximum echo top of the storm from HEn3DVar_WC is a little bit higher than that from HEn3DVar_W/C, and is much closer to the truth.

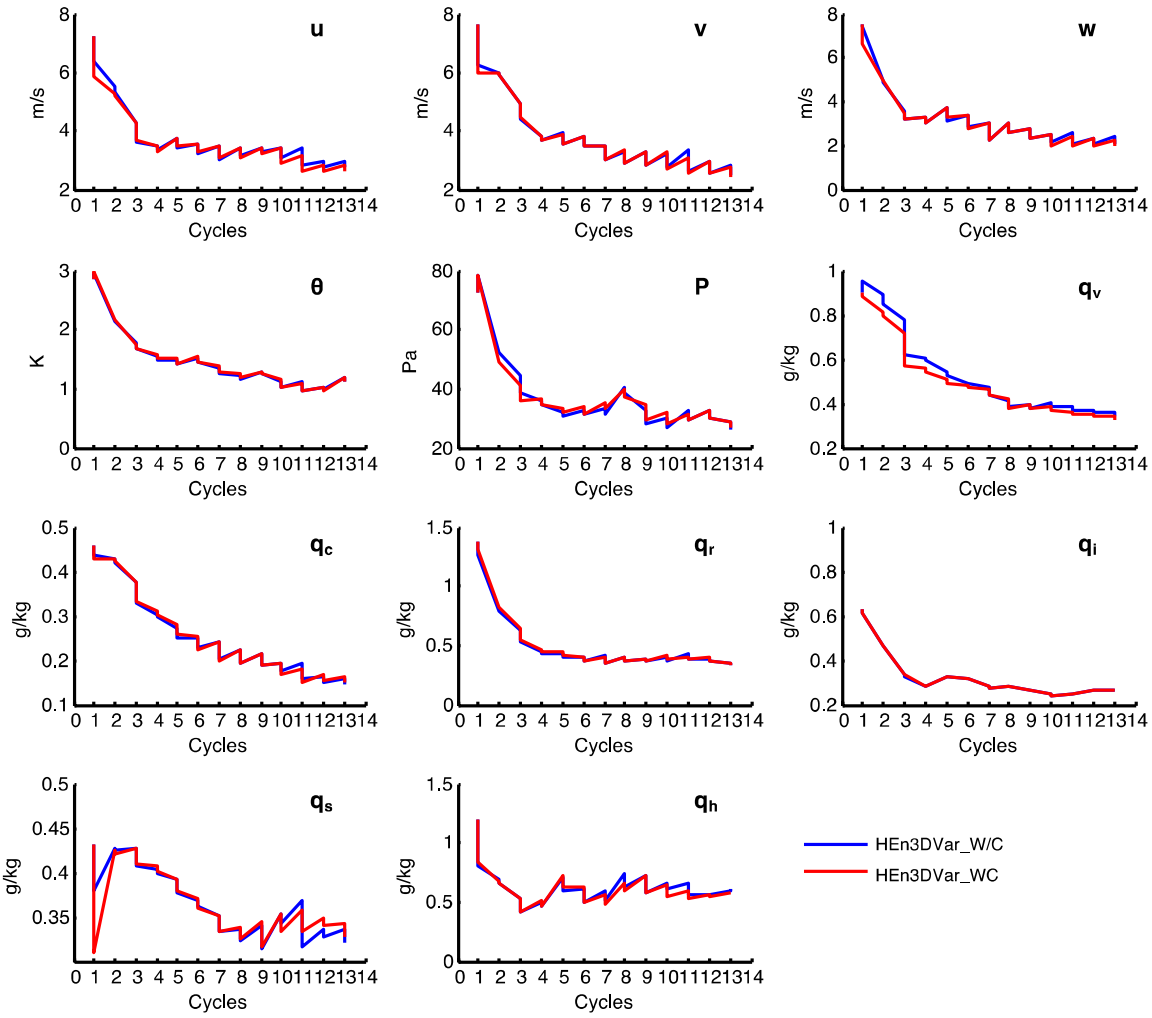

Fig. 4.18 The Grid-point averages of the root mean squared error for 3DVar without (blue contour) and with (red contour) adding the weak constraint.

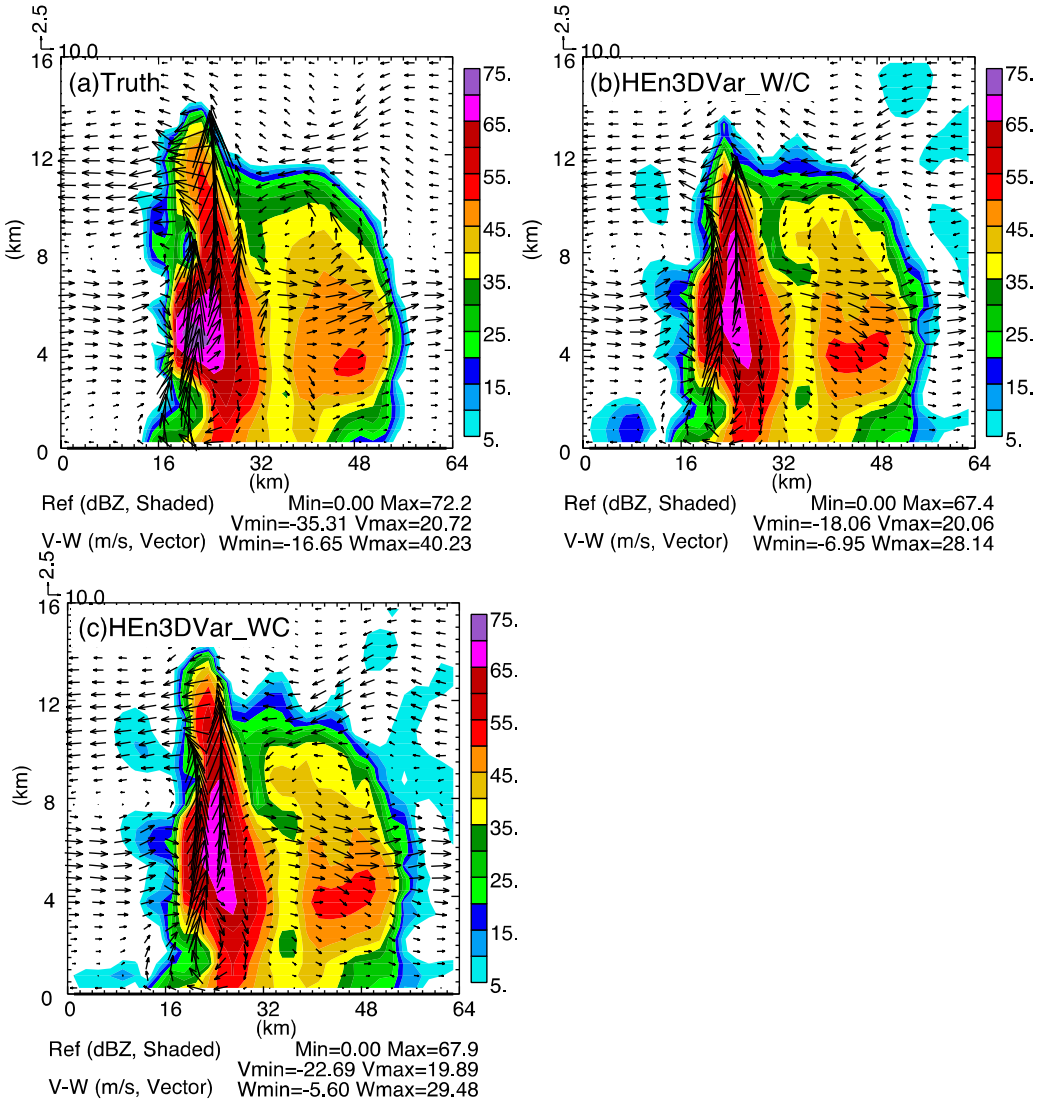

Fig. 4.19 Vertical cross-sections of the reflectivity field through the maximum vertical velocity of truth overlaid with the wind vector at the end of the one-hour DA window from (a) truth, HEn3DVar without (b), and with (c) the mass continuity constraint.

Overall, adding the mass continuity constraint still benefits the analyses in providing a little bit stronger vertical velocity analyses as well as smoothing the noises existed in the velocity and hydrometer fields.

4.4 Summary and discussion

In this chapter, the ARPS hybrid En3DVar DA system is applied to the assimilation of simulated radar data for a supercell storm using imperfect-model OSSEs. Model errors are introduced by using different microphysical schemes in the truth run (Lin scheme) and forecasts within the DA cycles (WSM6 scheme). Radar radial velocity and reflectivity data were assimilated every five minutes for one hour. A deterministic forecast EnKF algorithm (DfEnKF) introduced in Chapter 3 was also used here. It updates a single deterministic background forecast using the EnKF mean updating equation, so that EnKF can be compared with pure En3DVar more directly. 3DVar, EnKF, DfEnKF, pure and hybrid were tuned to obtain their optimal configurations before being compared. For reflectivity DA based on the variational framework, clear-air reflectivity (< 5 dBZ) was found to seriously degrade the analysis in the storm region, with the intensity of the analyzed storm being much weaker than that assimilates precipitation reflectivity only as well as the truth. A double-pass procedure, in which the radial velocity and precipitation reflectivity ($\geq$5dBZ) are assimilated in the first pass, and the clear-air reflectivity assimilated in the second, is found to greatly improve the reflectivity analyses. The influence of mass-continuity constraint to convective storm DA is also examined among 3DVar, pure and hybrid En3DVar. The experiments conducted and related conclusions are summarized as follows:

- Optimal localization radii (background error de-correlation scales) in terms of the smallest scaled root mean squared errors are obtained for EnKF, DfEnKF,

and pure En3DVar (3DVar) through sensitivity experiments, these configurations are then applied to hybrid En3DVar.

- The sensitivity of hybrid En3DVar to covariance weights and ensemble size is examined. Hybrid En3DVar performs much better than pure En3DVar and 3DVar with the weight of the static background error covariance ranged from 10% to 90% for most of the variables and for all the tested ensemble sizes. There are some exceptions for analyses of cloud ice, and potential temperature as well as snow mixing ratio for small ensemble sizes (less than or equal to 30), for which pure En3DVar performs better. On average, Hybrid En3DVar with weights ranged from 30% to 40% yield the best results for all the ensemble sizes. When ensemble size is 40, the optimal weight of static **B** is 30%.
- Using an ensemble size of 40, Hybrid En3DVar with 30% static **B** obviously outperforms the other algorithms (3DVar, EnKF, DfEnKF, and pure En3DVar). The outperformance is most obvious for analyses of velocity, mixing ratios of water vapor, cloud water, and hail. Hybrid En3DVar successfully capture the hail analyses below the freezing level, which are missed in the analyses of EnKF, DfEnKF, and pure En3DVar due to underestimation of the ensemble spread. Hybrid En3DVar also outperforms 3DVar for better capturing the intensity of the reflectivity core as well as the updraft. EnKF, DfEnKF, and pure En3DVar obviously underperforms 3DVar and hybrid En3DVar for analyses of water vapor mixing ratio fields, indicating the cross-correlations between mixing ratios of water vapor and hydrometeor variables are not very reliable when severe microphysical error exists. However, when applying 30% weight to static B, the

ensemble correlation between mixing ratio of the water vapor and hydrometeors can still benefit the analysis by generating small decrease of the errors from the analysis relative to the background in hybrid En3DVar DA. DfEnKF performs similarly as pure En3DVar, and both underperform EnKF for velocity components of u, v, and cloud ice mixing ratio at later cycles. Overall, in the current imperfect OSSE framework, hybrid En3DVar performs the best, 3DVar, EnKF, DfEnKF, and pure En3DVar the worst. With 30% weight given to the static background error covariance B, the rank deficiency problem happened in pure ensemble-based DA is greatly alleviated. The advantage of hybrid DA over pure ensemble-based DA is most obvious when ensemble background errors are systematically underestimated.

- In the imperfect-model OSSEs, clear-air reflectivity (reflectivity < 5 dBZ) was assimilated to help suppress the spurious storms. In the reflectivity DA based on the variational framework, the assimilation of clear-air reflectivity was found to severely degrade the analyses in storm region when it was assimilated together with the precipitation reflectivity (≥5dBZ). The degradation is mainly caused by the excessively large gradient related to small background reflectivity. The background equivalent reflectivity tends to be small in the area where the observed reflectivity is also small (<5dBZ), especially after one or two cycles when the location errors of convective storms are greatly decreased by reflectivity DA. The small (near zero) background equivalent reflectivity causes the gradient of the cost function excessively large and makes the assimilation of the radial velocity and the precipitation reflectivity ineffective. To solve this

problem, a double pass procedure is proposed, with the first pass assimilating radial velocity and precipitation reflectivity and the second pass assimilating clear-air reflectivity only. In this way, the spurious storms are suppressed by the clear-air reflectivity, and at the same time the influence of the clear-air reflectivity to assimilation of radial velocity and precipitation reflectivity is greatly suppressed.

- Experiments with and without the mass-continuity constraint are compared among 3DVar, pure and hybrid En3DVar. Overall, adding the mass continuity constraint benefits the analyses (forecasts) in providing a little bit stronger vertical velocity analyses as well as smoothing the noises existed in the velocity and hydrometer fields.

Finally, we note that the conclusions obtained here are still based on the OSSEs. In next chapter, we will compare hybrid En3DVar against EnKF and 3DVar using a real supercell storm case.

Chapter 5 Evaluation of ARPS Hybrid En3DVar DA System for Radar Data Assimilation with a Real Tornadic Storm Case

5.1 Introduction

In the previous two chapters, the newly developed ARPS hybrid ensemble-3DVar (three-dimensional variational) data assimilation system (En3DVar) was tested based on the Observing System Simulation Experiments (OSSEs) under both perfect and imperfect model assumptions. Radar data was assimilated for a simulated supercell storm case using EnKF, 3DVar, pure and hybrid En3DVar algorithm. Performance of hybrid En3DVar in assimilating radar data is compared with those of 3DVar, EnKF, and pure En3DVar. To better understand the differences found between pure En3DVar and EnKF, a separate 'deterministic forecast' EnKF analysis, called DfEnKF, is conducted so that EnKF is more comparable to En3DVar.

In perfect-model OSSEs, DfEnKF and pure En3DVar perform differently when using the same localization radii. The serial (EnKF) versus global (pure En3DVar) nature of algorithms, and direct filter update (EnKF) versus variational minimization (En3DVar) are the major reasons for the differences. When the algorithms are tuned optimally, hybrid En3DVar does not outperform EnKF and pure En3DVar in perfect-model OSSEs, though their analyses are all much better than 3DVar. When ensemble background error covariance is a good estimation of the true error distribution, pure ensemble-based DA methods can do a good job, and the advantage of static **B** is not obvious in hybrid DA. Imperfect-model OSSEs were conducted to further examine whether the conclusions obtained based on the perfect-model OSSEs change after considering the model errors. In the imperfect-model OSSEs, model errors are

introduced by using different microphysical schemes in the truth run (Lin scheme) and in the forecasts (WSM6 scheme). Hybrid En3DVar then outperforms EnKF and pure En3DVar (3DVar) for better capturing the hail analyses below the freezing level (intensity) of the storm in the analysis. The advantage of hybrid En3DVar over pure ensemble-based methods is most obvious when ensemble background errors are systematically underestimated.

However, the above conclusions are made based on the OSSEs, which tend to underestimate the errors of analyses and forecasts. In reality, the sources of the model errors are more complicated, which include the incomplete formulation of equations of motion, numerical approximation errors, time interpolation errors, boundary condition errors, convective parameterization errors, heat, moisture, and momentum flux approximation errors, etc. The imperfect-model OSSEs only considers a possibility of the microphysical error. In this chapter, the hybrid En3DVar algorithm was applied to a real storm case, so that any advantage of hybrid En3DVar over pure 3DVar and EnKF in storm-scale radar DA can be further examined.

The rest of this chapter is organized as follows. In section 5.2, we briefly introduce the real storm case, DA schemes (EnKF, DfEnKF, and En3DVar) used, and the experimental design. In section 5.3, logarithmic hydrometeor mixing ratios are used as the control variables and the experimental results are compared with those directly using hydrometeor mixing ratios as the control variables. Sensitivity experiments are conducted to obtain a relative optimal weight for hybrid En3DVar. Finally, hybrid En3DVar, 3DVar, and EnKF are compared based on both objective verifications and physical field analyses. A summary and outlook are then concluded in section 5.4.

5.2 Assimilating system and experimental design

5.2.1 The ARPS En3DVar DA scheme

In the ARPS En3DVar DA system, flow-dependent ensemble covariance and static background error covariance are combined within the existing ARPS 3DVar variational framework. The hybrid scheme follows (Lorenc 2003; Buehner 2005), in which the static and ensemble background error covariances are combined through introduction of a set of extended control variables. Detailed descriptions about the hybrid algorithm can be found in chapter 3.

5.2.2 The reflectivity observation operator

The reflectivity observation operator follows that used in chapter 3, in which the radar radial velocity and reflectivity data are calculated based on model velocity and hydrometeor mixing ratio fields. One thing to be noted is that the logarithmic hydrometeor mixing ratios for rain, snow, and hail were used as the control variables in this chapter to help alleviate the problem caused by small background equivalent reflectivity (Section 4.3.3). Results of using logarithmic hydrometeor mixing ratio as control variables will be compared with those directly using hydrometeor-mixing ratio as control variables.

5.2.2 Case overview

On 10 May 2010, a total of 56 tornadoes produced by storms triggered by an intense dry line occurred in the state of Oklahoma, including two EF4 (the Moore-Lake Draper-Harrah and Norman-Little Axe-Pink tornadoes) and four EF3 tornados that induced significant damage over other areas. Information from the Oklahoma

Department of Emergency Management indicates, 3 people were killed and over 450 people were injured in the state during the whole tornado outbreak. This was the second largest tornado outbreak documented in Oklahoma (with the largest outbreak occurring on May 3, 1999). From 5:00 PM to 7:41 pm CDT, there was at least one tornado occurring in Oklahoma every minute. A total of 42 different tornadoes occurred within this time. The tornadic storms move 50 to 60 mph, with the damage paths of tornadoes spreading over a north-south area of over 200 miles; softball size hail was also reported in several locations.

According to information from National Weather Services, there were 13 different storms that produced tornadoes. They are denoted as storms A to N (more information about this case may be found at*: http://www.srh.noaa.gov/oun/?n=events-20100510)*. Among these storms, storm J was the most productive, spawning 20 out of 55 tornadoes, including one EF4 (J1) and one EF3 (J4) tornado (Fig. 5.1), causing 1 fatality and several injuries. Considering tornado J4 has the largest damage width, the whole convective system with its development covers the lifetime of tornado J4 was selected for this study.

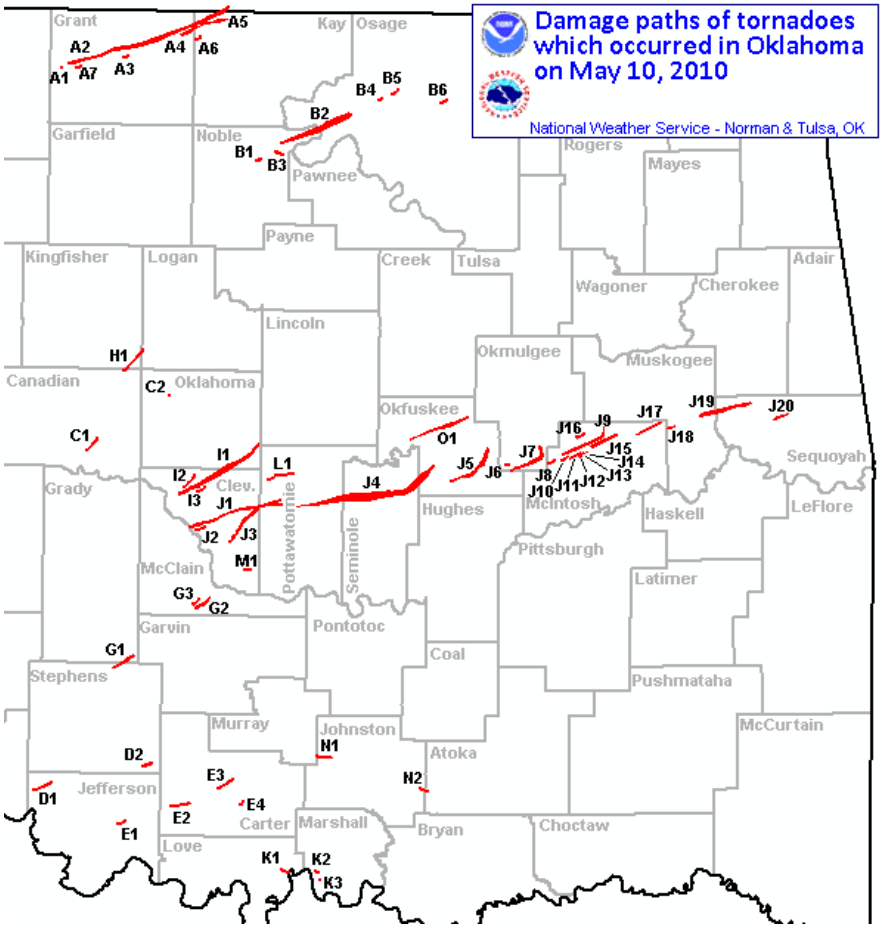

Fig. 5.1 The damage path of tornadoes occurred in Oklahoma on May 10, 2010 (Image courtesy of National Weather Service).

5.2.3 Radar data preprocessing and quality control

Radar data is interpolated onto the ARPS terrain-following Cartesian model grid using the ARPS-88d2ARPS module. The interpolation is only performed in horizontal based on a quadratic local least-square fit of the observations. In vertical direction, no interpolation is performed; the data is saved on the elevation angles as the raw data instead of the model grid. Prior to remapping the radar data are checked for velocity folding, transient echoes in clear air are removed (image is despeckled) by eliminating

any datum that has non-missing neighbors fewer than 3 in a 3 × 3 square template, and the data are checked for anomalous propagation (AP).

5.2.4 Experimental design

a. Model configuration

The Advanced Regional Prediction System (ARPS) model was used to simulate the tornadic storm case happened on 10 May 2010. Considering the long trajectory of the storms, the domain is set to cover the influencing area of the interested convective storms with its center located at 35.7°N and 97.5°W (Fig. 5.2). The domain size is 363 × 363 × 53, with a 1 km grid interval in the horizontal directions and 0.4 km on average in vertical stretching grid. The minimum vertical resolution in low levels is 20 m.

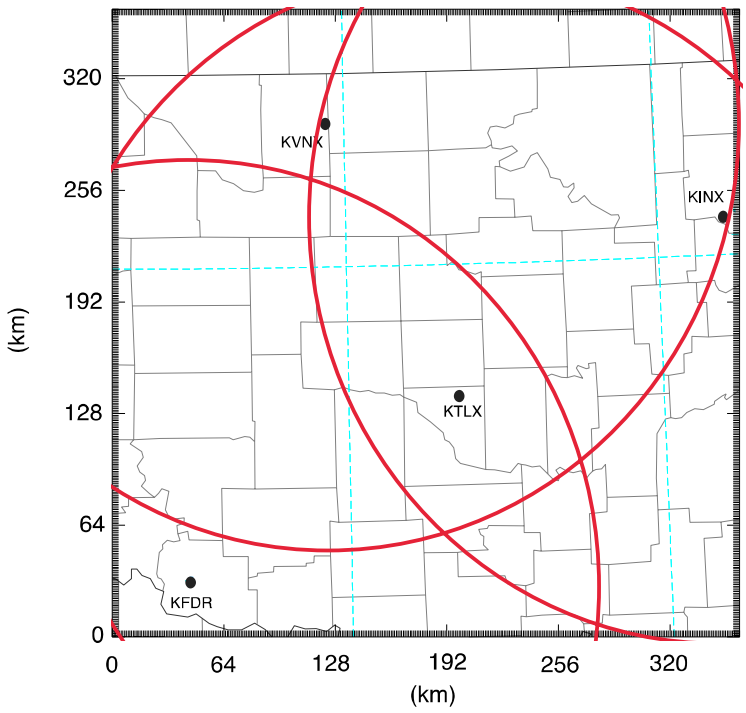

Fig. 5.2 The simulation domain and truncated radar coverage circle, radar locations are marked by the black dot.

Full model physics (Xue et al. 2001) were applied, which includes 1 1.5-order turbulent kinetic energy (TKE) based subgrid-scale turbulence parameterization scheme (Deardorff, 1980; Klemp and Wilhelmson, 1978; Moeng, 1984), a two-layer soil model, simplified surface radiation physics and an ice microphysics package (Lin et al. 1983). 4^{th} order monotonic computational mixing is used in model integration. The default values of the intercept parameter for rain ($8 \times 10^6 \mathrm{m}^{-4}$) and hail ($4 \times 10^4 \mathrm{m}^{-4}$) were used.

b. Experiment configuration

40 member ensemble analyses (that assimilate surface, upper-air, profiler, and radar observations) with a 40-km horizontal resolution and valid at 21:00 UTC May 10, 2010 and its corresponding hourly forecasts are interpolated to the testing domain to serve as the initial ensemble perturbations and boundary conditions. For more information about the initial ensemble fields, the reader is referred to Jung et al. (2012).

On the experimental domain, the Oklahoma Mesonet, radar radial velocity and reflectivity data are assimilated every 5 minutes using 3DVar, EnKF, pure and hybrid En3DVar DA systems (Fig. 5.3). To suppress the spin-up problem of the forecast model, the first analysis is conducted after 30-minute integration of the initial perturbations at 21:00 UTC to increase the spread of the initial background ensembles. The observational errors are taken as 2 m s^{-1} for u and v, 1.5 K for potential temperature, and 2 hpa for pressure. 20% multiplicative inflation was added to the background fields where the observed reflectivity is higher than 5 dBZ to maintain the ensemble spread. The localization for EnKF is conducted based on a fifth-order correlation function (Gaspri and Cohn 1999), with a cutoff radius of 6 km in the horizontal and 3 km in the

vertical for radar data and 30 km in the horizontal and 6 km in vertical for mesonet DA. For pure and hybrid En3DVar, the recursive filter length scales for localization are 1.64 km in horizontal and 0.82 km in vertical for radar DA, which are equivalent to 6 km and 3 km cutoff based on Eq. (4) in Pan et al. (2013), and 8.22 km in horizontal and 1.64 km in vertical for mesonet DA. For 3DVar, the background error decorrelation scales for 3DVar are 3.56 km in horizontal and 0.10 km in vertical for radar DA, and 8.22 km in horizontal and 1.64 km in vertical for mesonet DA.

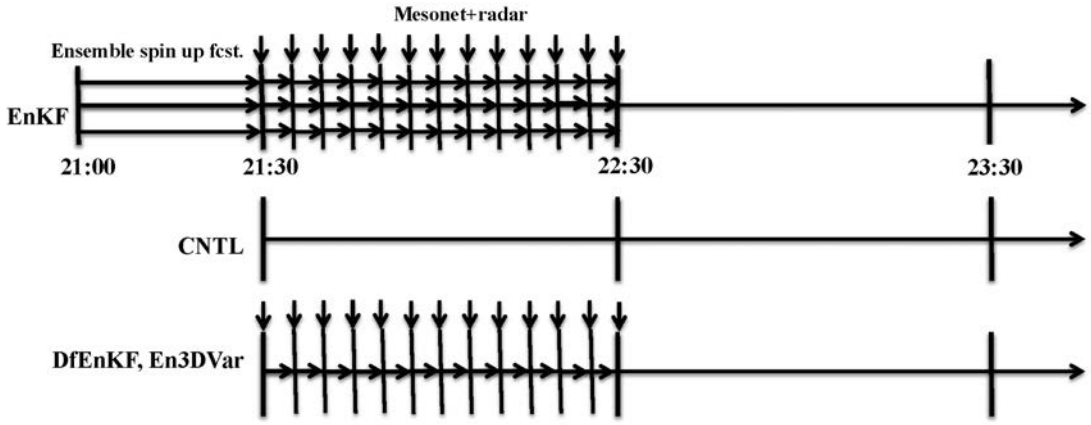

Fig. 5.3 The flow diagram of the experiments.

The radar radial velocity and reflectivity data (including 'clear-air' reflectivity, e.g., reflectivity <5 dBZ) form four S band WSR-88D radars (KTLX, KINX, KVNX, KFDR) available from 22:30 Z to 23:30 Z are assimilated and the performances are compared among hybrid En3DVar, 3DVar, EnKF, and pure En3DVar algorithms. Same as those did for the OSSEs, a deterministic forecast EnKF (DfEnKF), for which the initial ensemble mean is evolved deterministically within the data assimilation cycles, was also introduced to have an algorithm-wise parallel comparison between pure En3DVar and EnKF.

In EnKF and En3DVar DA, the updating of the velocity fields (u, v, w) by reflectivity DA is turned off because of the rapid growth of the errors in velocity analyses and forecasts. Hybrid En3DVar analyses with 5%, 25%, 50%, and 75% weights given to the static background error covariance (experiments names are suffixed with "%B") are compared to get a relative optimal weight among them. Detailed descriptions about each experiment can be found in table 5.1.

Table 5.1 Descriptions of the assimilating experiments

Experiment name. With suffix "_qx" indicating using q_x (x=r, s, h, indicative of rain, snow, and hail) as the control variables, otherwise using $log_{10}^{q_x}$ as the control variables.	**Use of background error covariance**	**Background updating**
CNTL	None	Update background field deterministically
3DVar_qx/3DVar	Static background error covariance	Update background field deterministically
EnKF	Background error covariance derived from ensemble background forecasts	Updates ensemble mean background and ensemble perturbations using EnKF algorithm
DfEnKF	Ensemble covariance from an EnKF system	Update a single deterministic background forecast using EnKF mean updating algorithm
PEn3DVar(_qx)	Pure En3DVar using 100% ensemble covariance from an EnKF system	Update a single deterministic background forecast using variational algorithm
HEn3DVar5%B(_qx)	Hybrid En3DVar using 5% static and 95% ensemble covariance	
HEn3DVar25%B(_qx)	Hybrid En3DVar using 25% static and 75% ensemble covariance	
HEn3DVar50%B(_qx)	Hybrid En3DVar using 50% static and 50% ensemble covariance	
HEn3DVar75%B(_qx)	Hybrid En3DVar using 75% static and 25% ensemble covariance	

c. Application of Incremental Analysis Updates (IAU)

In order to make the analysis increment better accepted by the model and reduce the shock of data insertion. IAU is applied in the DA cycles. In IAU, analysis increments are incorporated into model integration gradually (Bloom et al. 1996). Traditionally, the IAU assimilation is applied with a triangular distribution of the fractional increments in time (Brewster et al. 2016; Brewster and Derek 2015). In order to decrease the time error, a step-down time function is used in our experiments, with the largest increment set at the beginning and the smallest at ends. In our experiments, the IAU is applied to all the model prognostic variables with the interval of 100 s, indicating the analysis increment is added to the background forecast in 3 steps within a single data assimilation window. For example, for a time window from 0 s to 300 s, the analysis increment is added to the background forecasts at 100, 200, and 300 s with the time scales of 0.6633, 0.33, and 0.0033 respectively, Fig. 5.4).

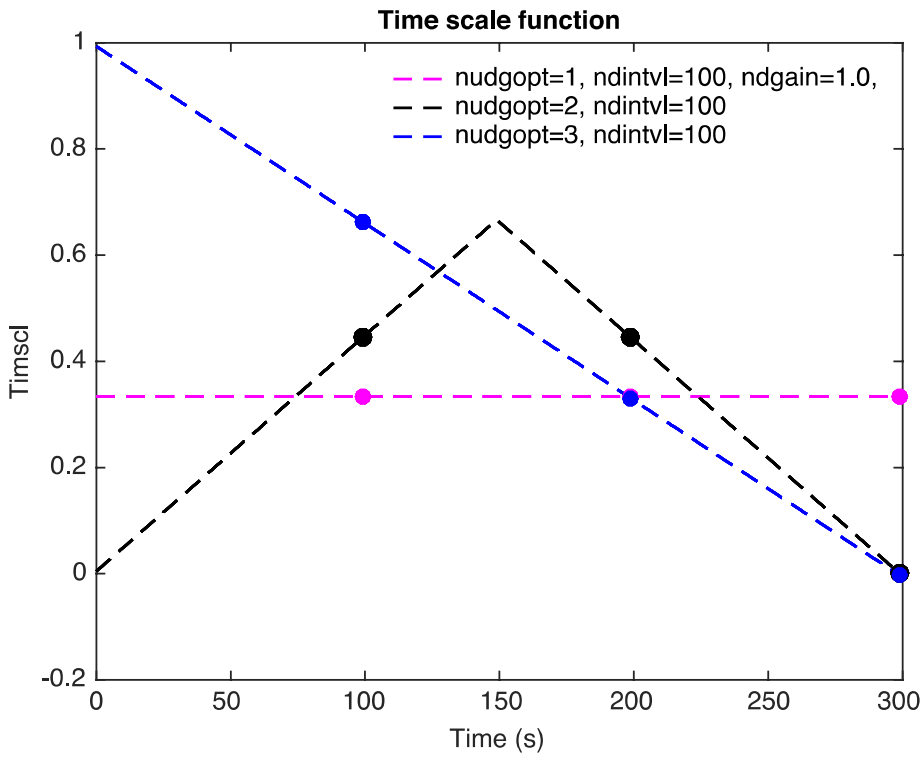

Fig. 5.4 Three types of time scales as a function of the integration interval for IAU.

5.3 Results of assimilation experiments

In this section, logarithmic hydrometeor mixing ratios are used as the control variables and the experimental results are compared with those directly using hydrometeor mixing ratios as the control variables. Sensitivity experiments are conducted to obtain a relative optimal weight for hybrid En3DVar. Performances of 3DVar, EnKF, pure and hybrid En3DVar are compared based on both the objective error statistics (such as the root mean square error: RMSE, the scatter and QQ plot, and the fractional skill score: FSS) and the analyses/forecasts of convective storms features, such as the analyses and forecasts of the reflectivity, radial velocity, updrafts and downdrafts, mesocyclone, and the cold pool, etc.).

5.3.1 Using hydrometeor-mixing ratios versus using logarithmic hydrometeor mixing ratios as the control variables

As mentioned in Chapter 4, when directly using hydrometeor-mixing ratios as the control variables, the reflectivity DA by 3DVar and En3DVar suffers from extremely large gradient of the cost function, and leads to ineffective assimilation of reflectivity and radial velocity data. The intensity of the storm core and the updraft in the storm region were apparently underestimated due to inclusion of small background reflectivity into calculation of the gradient, especially when clear-air reflectivity (< 5 dBZ) were assimilated. Because the background reflectivity tends to be small in the region where the observed reflectivity is also small (i.e., the clear-air region), especially after several DA cycles when the location errors of the storms have been greatly suppressed by reflectivity DA. To address this issue, a double-pass procedure is proposed in the imperfect-model OSSEs (Chapter 4), with the radial velocity and precipitation reflectivity (≥ 5 dBZ) assimilated together in the first pass, and the clear-air reflectivity in the second. Through the double-pass procedure, two cost functions with apparent different orders of magnitude are minimized separately, so that the minimization of the cost function based on the largest gradient won't influence that based on the smaller one. However, there still exist some problems in the double-pass procedure. First, small background reflectivity can still be included into calculation of the gradient in the first pass, and producing large gradient that hurts the radial velocity and precipitation reflectivity DA, especially when significant location errors exist. Second, the background error covariance needs to be updated after assimilating radial velocity and precipitation in the first pass, using the same background error covariance

in the second pass as the first one will introduce errors into the DA process. In addition, the double-pass procedure involves minimizing two different cost functions with different orders of magnitude separately, which is a lot more expensive.

An alternative way to avoid the extremely large gradient in reflectivity DA is to use the logarithmic hydrometeor mixing ratios, instead of the hydrometeor mixing ratios, as the control variables. The experiments that using (logarithmic) hydrometeor-mixing ratios as the control variables are denoted as CV_q (CV_logq)) to facilitate the discussion. CV_logq was implemented into the NCEP Grid-point Statistical Interpolation by Carley (2012). The hydrometeors mixing ratios are transformed into logarithmic mixing ratio, taking rain water-mixing ratio as example, based on

$$\hat{q}_r = \log q_r . \tag{5.1}$$

The corresponding equivalent reflectivity is,

$$Z_{er} = 3.63 \times 10^9 \times (\rho 10^{\hat{q}_r})^{1.75} . \tag{5.2}$$

Here ρ is the air density. By taking into account three types of hydrometeor fields, the total equivalent reflectivity is

$$Z_e = Z_{er}(\hat{q}_r) + Z_{es}(\hat{q}_s) + Z_{eh}(\hat{q}_h) . \tag{5.3}$$

To understand how CV_q and CV_logq affect the order of magnitude of the gradient in reflectivity DA, the tangent linear operators corresponding to CV_q and CV_logq are compared. For CV_q, the tangent linear of Z observation operator with respect to hydrometeor mixing ratio, taking rainwater mixing ratio as example, is

$$\frac{\partial Z}{\partial q_r} = \frac{10 \times 3.63 \times 10^9 \times \rho^{1.75} \times 1.75 \times q_r^{0.75}}{\ln 10 \times Z_e} . \tag{5.4}$$

Eq. (5.6) is not valid when Z_e is zero. In the NWP models, the magnitude of hydrometeor mixing ratios is usually from 0 to 10^{-2} kg kg^{-1}. If all of the hydrometeors mixing ratios are equal to 0 kg kg^{-1}, Z_e in the denominator is also zero. To avoid this problem, the hydrometeor mixing ratios are usually given a lower limit in the applications (such as 10^{-8} kg kg^{-1} used in this study). As a result, the orders of magnitude for Z is 10^{-3} mm^{6}m^{-3}, for $\partial Z/\partial q_r$ and $\partial Z/\partial q_s$ are 10^6 dBZ kg kg^{-1}, and for $\partial Z/\partial q_h$ is 10^7 dBZ kg kg^{-1} when hydrometeors mixing ratios are all as small as 10^{-8} kg kg^{-1} (to facilitate the demonstration, ρ is assumed to be 1.0 kg kg^{-1}), leading to excessively large gradient of the cost function. On the other hand, when hydrometeor-mixing ratios are as large as 10^{-2} kg kg^{-1}, the order of magnitude for Z_e is 10^7 mm^{6}m^{-3}, for $\partial Z/\partial q_r$, $\partial Z/\partial q_s$, and $\partial Z/\partial q_h$ are all 10^1 dBZ kg kg^{-1}, which are much smaller than those of the small value hydrometeor mixing ratios (10^{-8} kg kg^{-1}). Thus the assimilation of reflectivity observation is ineffective in the area where background reflectivity is large. In addition, the amount of adjustment made to individual hydrometeors for CV_q is controlled by the sensitivity of reflectivity to individual hydrometeor mixing ratios as discussed in Chapter 3. Variables with small sensitivity may receive little adjustment before the minimization iteration is terminated.

For CV_logq, the gradient of Z respect to logarithmic mixing ratios, taking rainwater-mixing ratio as example, can be written as,

$$\frac{\partial Z}{\partial \hat{q}_r} = \frac{10 \times 3.63 \times 10^9 \times 1.75 \times \left(\rho \times 10^{\hat{q}_r}\right)^{1.75}}{Z_e} . \qquad (5.5)$$

Similarly, the orders of magnitude for Z_e is still 10^{-3} mm^{6}m^{-3}, for $\partial Z/\partial q_r$ and $\partial Z/\partial q_s$ and $\partial Z/\partial q_h$ are 10^{-1} dBZ kg kg^{-1} when hydrometeors mixing ratios are all as

small as 10^{-8} kg kg^{-1}, which are much smaller than that for CV_q (10^6 kg^{-}mm^{-6} for q_r, q_s and 10^6 kg^{-}mm^{-6} for q_h). On the other hand, when hydrometeor-mixing ratios are as large as 10^{-2} kg kg^{-1}, the orders of magnitude for $\partial Z/\partial q_r$, $\partial Z/\partial q_s$, and $\partial Z/\partial q_h$ are still 10^{-1} dBZ kg kg^{-1}. In this way, the excessively large value of the gradient due to small background reflectivity can be avoided in reflectivity DA.

Comparisons are made between "CV_q" and "CV_logq" based on the RMSEs. As indicated in Fig. 5.5, the RMSEs of CV_q is much large than that of CV_logqx. However, the RMSEs of the one-hour reflectivity forecasts between experiments CV_q and CV_logq resemble. The radial velocity analyses and forecasts between the two experimental sets also resemble, indicating the errors in the hydrometeor fields for experiment set CV_q have not been projected into the velocity fields, because the updating of the velocity fields by reflectivity DA has been turned off.

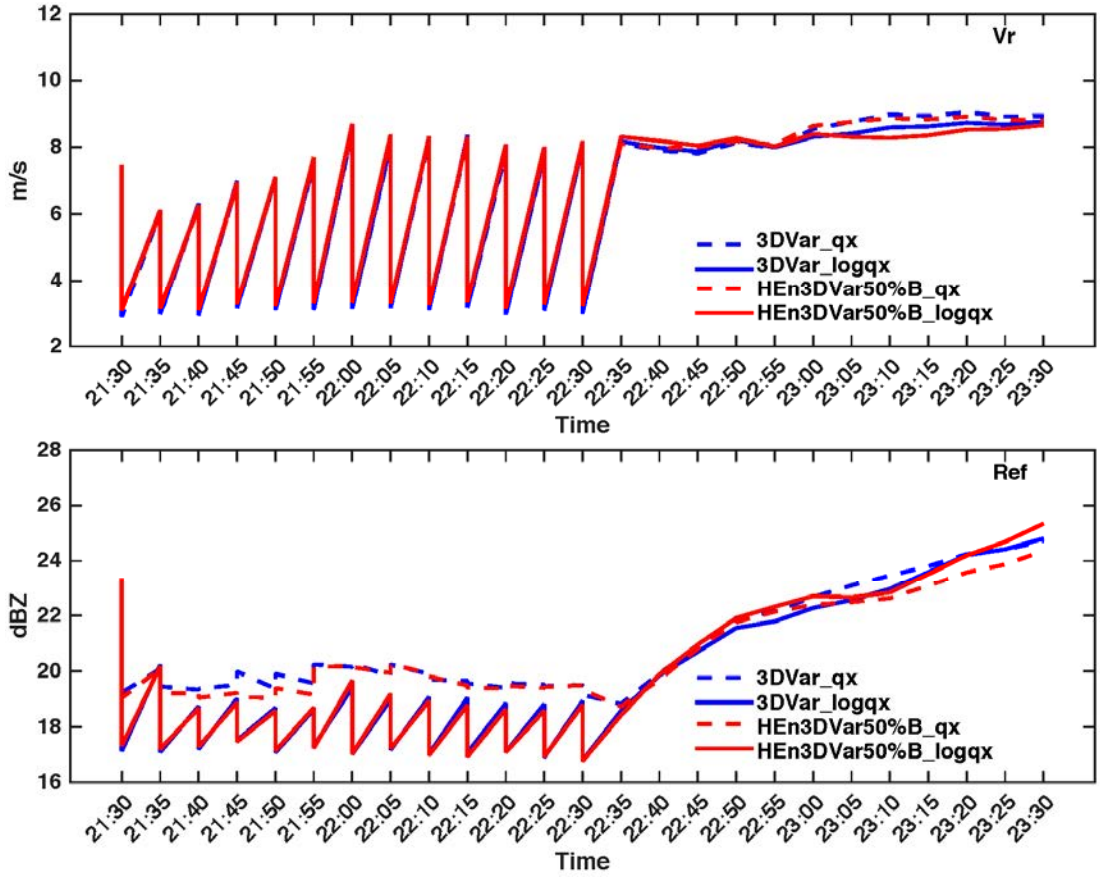

Fig. 5.5 Comparisons of the RMSEs for reflectivity and radial velocity analyses and forecasts between experiments CV_q (with dashed lines) and CV_logq (with solid lines) for 3DVar (blue contours) and hybrid En3DVar (red contours) methods with 50% weight given to the static background error covariance B.

When examining the composite reflectivity fields from CV_q (corresponding to experiment 3DVar_qx, HEn3DVar50%B_qx) and CV_logq (experiment 3DVar, HEn3DVar50%B) at the 2nd analysis time, the spurious reflectivity echoes outside of the storm region has been greatly suppressed for CV_logq relative to CV_q (Fig. 5.6). On the other hand, the area of the storm center with the analyzed reflectivity higher than 50 dBZ (depicted in red color) from experiments 3DVar, HEn3DVar50%B are larger than those from experiment 3DVar_qx and HEn3DVar50%B_qx, and are closer to the observation. The maximum reflectivity analyses form 3DVar and HEn3DVar are 68.1

dBZ and 69.0 dBZ respectively, which are much smaller than those from 3DVar_qx (73.4 dBZ) and HEn3DVar_qx (73.5 dBZ), and closer to that of the observation (64.5 dBZ). Considering experiment sets CV_logq has better performance than CV_q in reflectivity analysis. In the following sections, logarithmic hydrometeor mixing ratios will be used as the control variables.

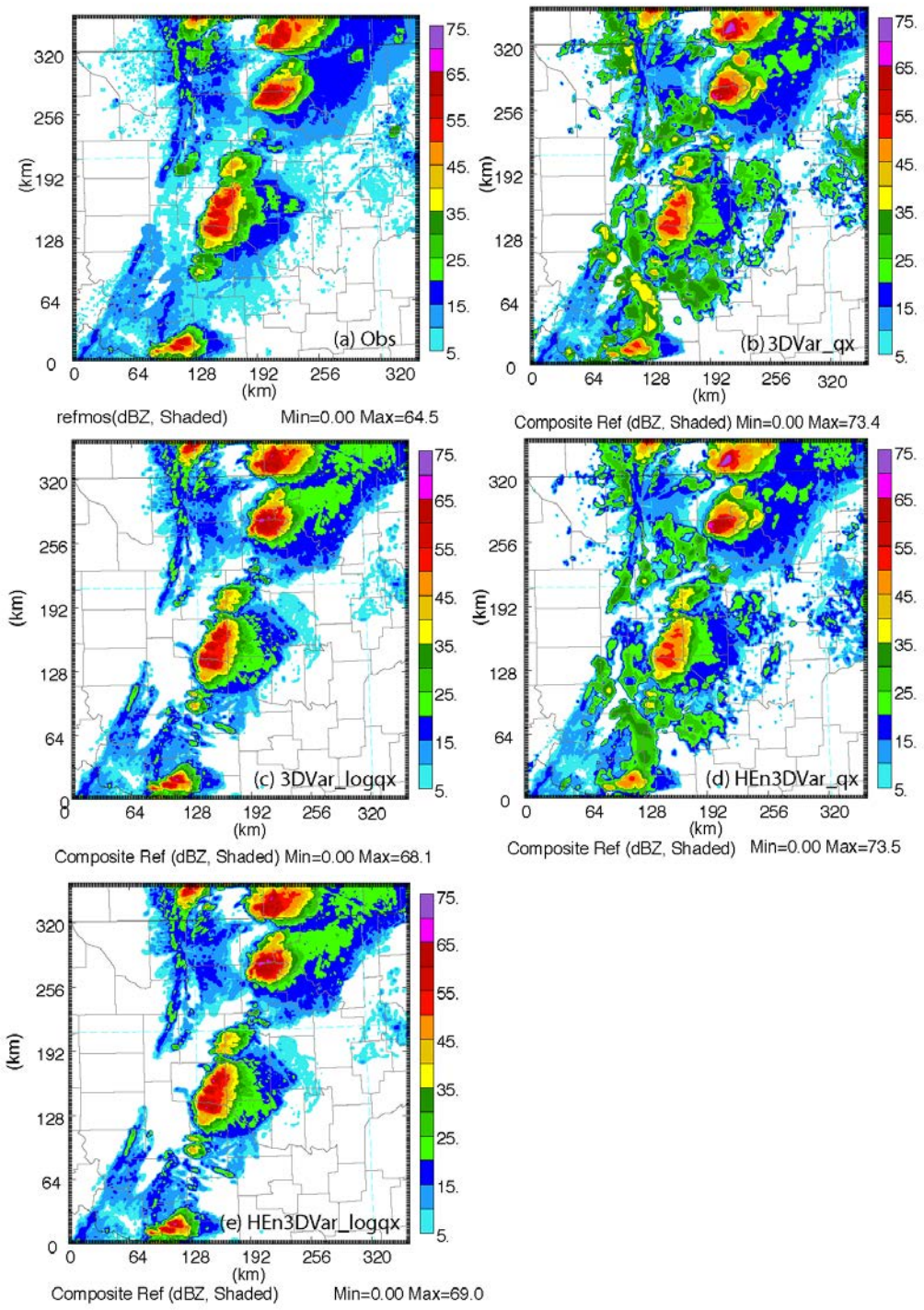

Fig. 5.6 Comparisons of (a) the composite reflectivity observation and composite reflectivity analyses at the 2nd cycle time (21:45 UTC) between experiments that directly using hydrometeor mixing ratios (Fig. b, d) and the logarithmic hydrometeor mixing ratios (Fig. c, e) as the control variables for analyses of 3DVar (Fig., a, b) and

hybrid En3DVar with 50% weight given to the static background error covariance B (Fig. c, e).

5.3.2 Tests on hybrid weights

To obtain a relative 'optimal' weight for hybrid En3DVar, four sets of experiments are conducted with 5%, 25%, 50%, and 75% weight given to the static background error covariance **B** respectively. The corresponding experiments are named as HEn3DVar5%B, HEn3DVar25%B, HEn3DVar50%B, and HEn3DVar75%B respectively. Their results are also compared with pure En3DVar and 3DVar that actually have 0% and 100% static **B**.

The RMSE is commonly used to quantify model analysis and forecast error and compare the performance among different DA algorithms. In this chapter, the RMSEs are calculated over regions where the observed or analyzed reflectivity exceeds 15 dBZ to take into account the possibility of overestimation. As indicated in Fig. 5.7, the radial velocity and reflectivity analyses and forecasts within the DA window from all experiments are quite similar based on the RMSEs. Hybrid En3DVar with 5% B performs the worst in terms of the largest RMSEs in the forecasts of radial velocity and reflectivity, though it is still better than the CNTL. For the following-on reflectivity forecasts, 3DVar, hybrid En3DVar with 50% and 75% weight given to the static **B** perform similar and are much better than those of the other two (5% and 25%).

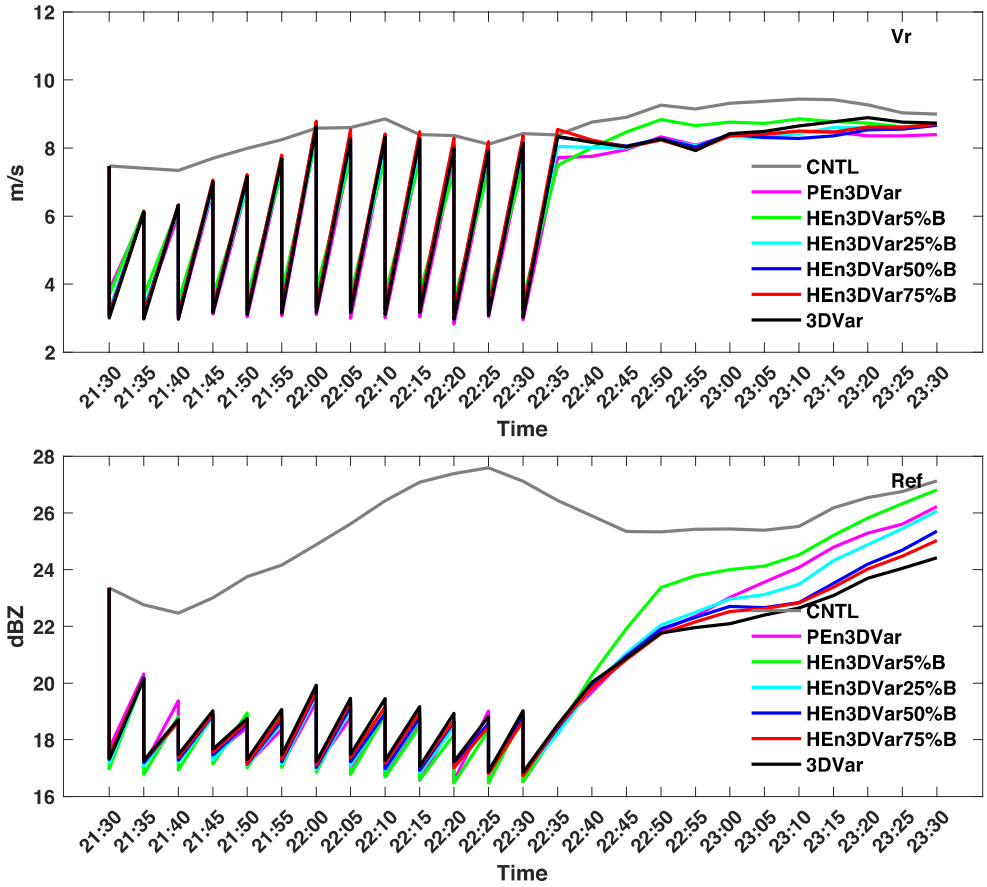

Fig. 5.7 Comparisons of the RMSEs for radial velocity and reflectivity from control and the experiments that using hybrid En3DVar algorithm with 0% (actually is pure En3DVar), 5%, 25%, 50%, and 75% and 100% (actually is 3Dvar) weight given to static background error covariance B respectively.

When comparing the reflectivity analyses at 1 km height above ground level (AGL) among different experiments, the reflectivity analyses at the end of the 1-hour DA window are similar in capturing the intensity and shape of the main supercell storm, and are close to that of the observation (Fig. 5.8).

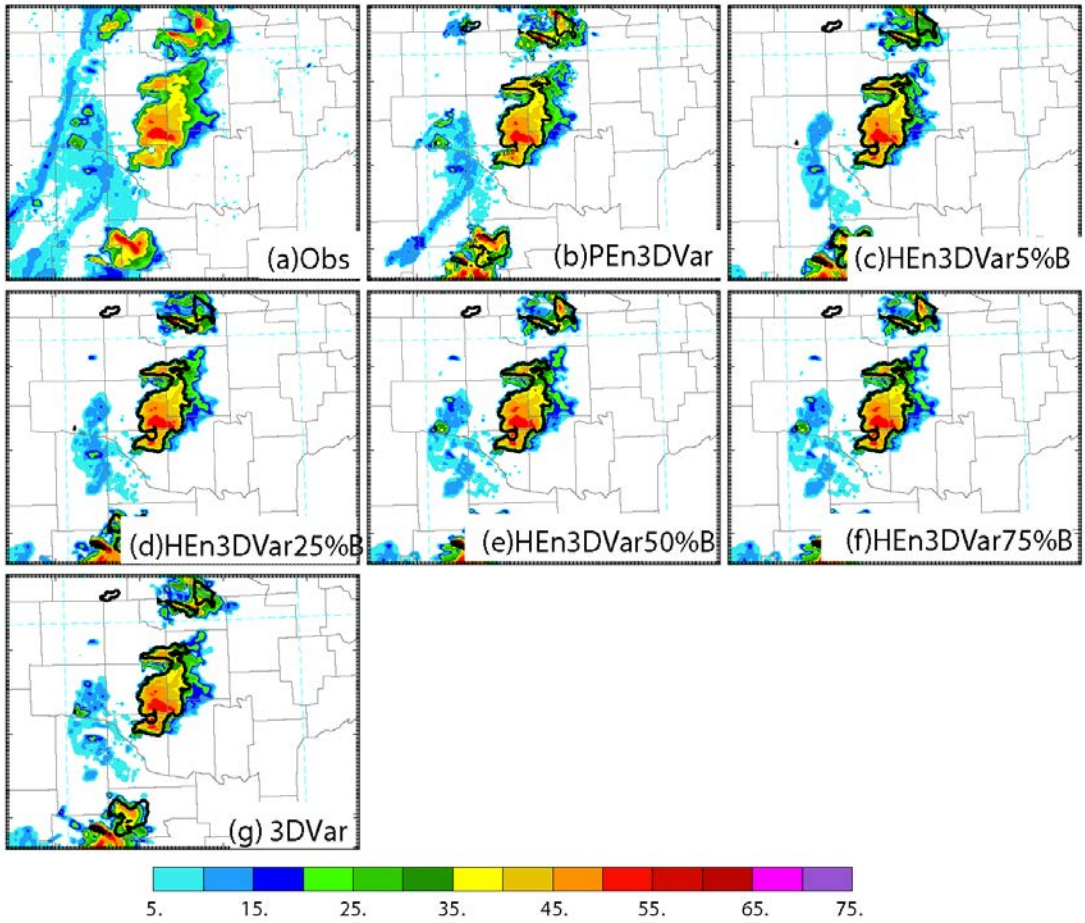

Fig. 5.8 Reflectivity observation at 1 km AGL (a) and reflectivity analyses at the end of one-hour DA wndow from (b) pure En3DVar, hybrid En3DVar with (c) 5%, (d) 25%, (e) 50%, (f) 75% weight given to static background error covariance B, and (g) 3DVar, overlaid with observed reflectivity that is higher than 35 dBZ.

The 45-min forecasts of the reflectivity at 1 km AGL are compared among different experiments as well as the observation. The forecasts of the storm intensity from different experiments are similar and close to that of the observation. The storm structure forecast from PEn3DVar and HEn3DVar5% are less organized than those of the others as well as the observation. For 3DVar, the storm structure is too linear. Hybrid En3DVar with 50% and 75% weight of **B** outperforms the other experiments in better capture the hook echo structure of the major tornadic supercell storm (Fig. 5.9).

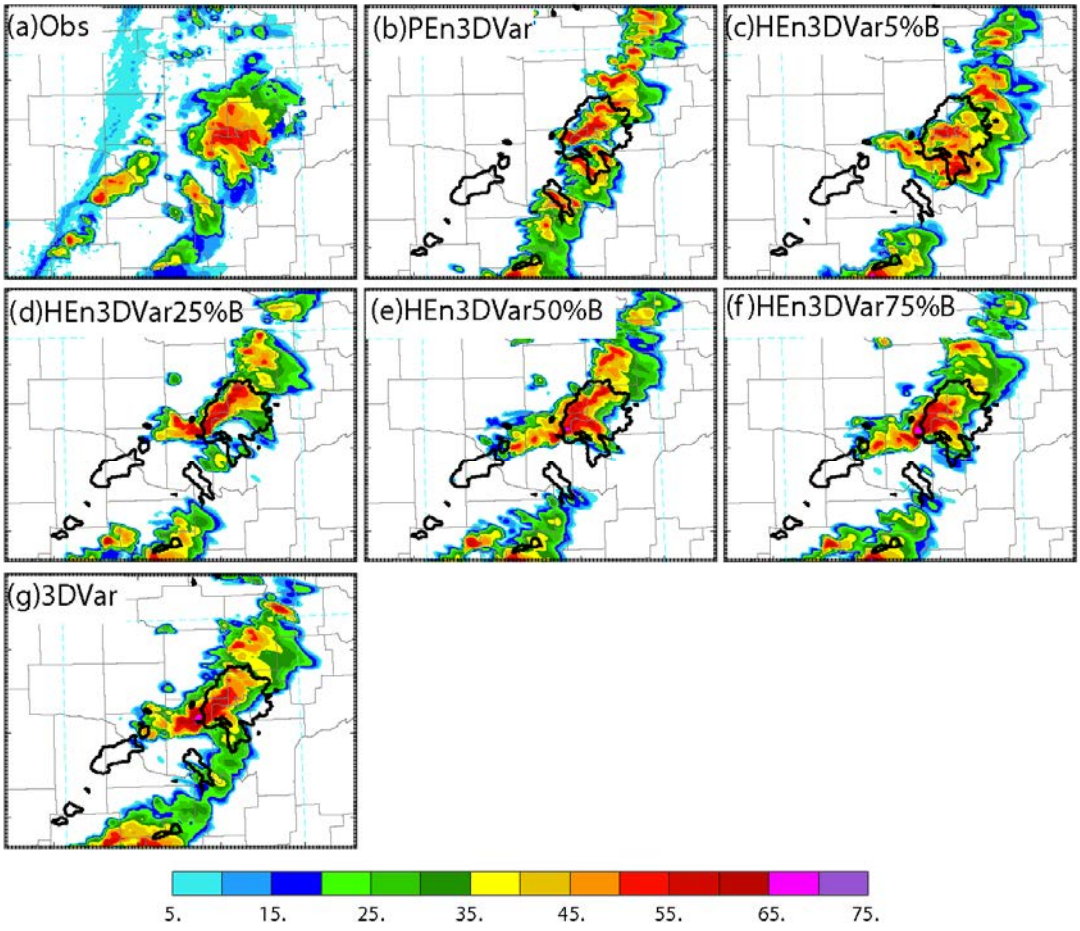

Fig. 5.9 Reflectivity observation at 1 km AGL (a) and 45-minute reflectivity forecasts from (b) pure En3DVar, hybrid En3DVar with (c) 5%, (d) 25%, (e) 50%, (f) 75% weight given to static background error covariance B, and (g) 3DVar, overlaid with observed reflectivity that is higher than 35 dBZ.

To obtain a hybrid En3DVar analysis that also better captures the low-level rotation of the tornadic supercell storm, radial velocity analyses and forecasts at the lowest elevation angle of the KTLX radar are compared among different experiments. The radial velocity analyses from different experiments are similar and close to the observations (Fig. 5.10). For the 45-minute radial velocity forecasts, experiments HEn3DVar25%B, HEn3DVar75%B, and 3DVar better capture the mesocylone that are indicated by the black circles (Fig. 5.11). For 1-hour radial velocity forecasts,

experiment HEn3DVar50%B and HEn3DVar75%B better capture the intensity and the location of the mesocyclone that are close to the observations (Fig. 5.12). The location error of the mesocyclone forecast from HEn3DVar with 5%weight given to static **B** is more obvious than those with the 50%B and 75%B. Based on the above results, 75% was chosen to be the optimal weight of static **B** for hybrid En3DVar, thus HEn3DVar75%B was used to compare with the other algorithms.

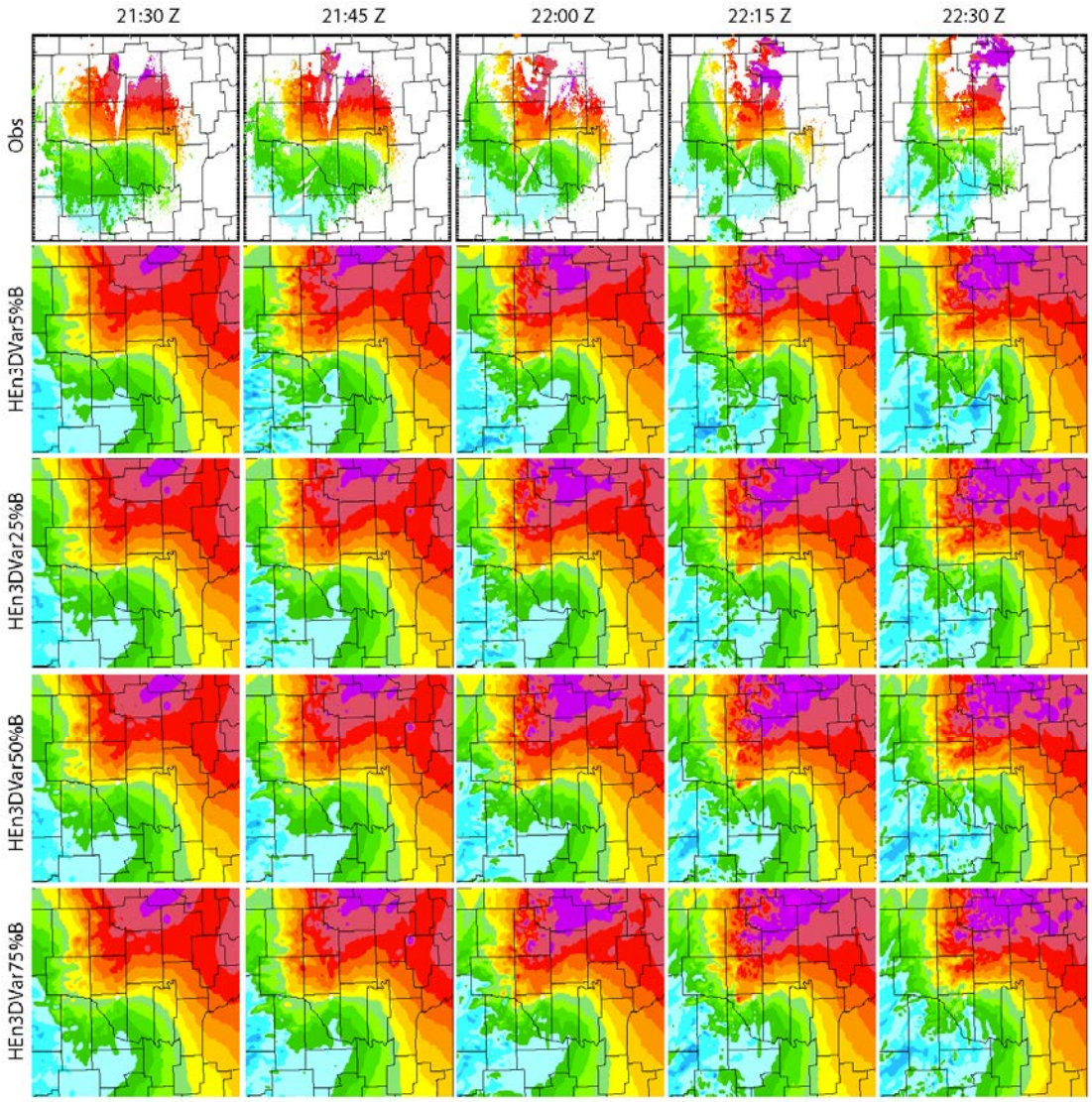

Fig. 5.10 Observed and analyzed radial velocity at the first elevation angle of radar KTLX from hybrid En3DVar with 5%, 25%, 50%, and 75% weight given to static background error covariance B with 15 min interval.

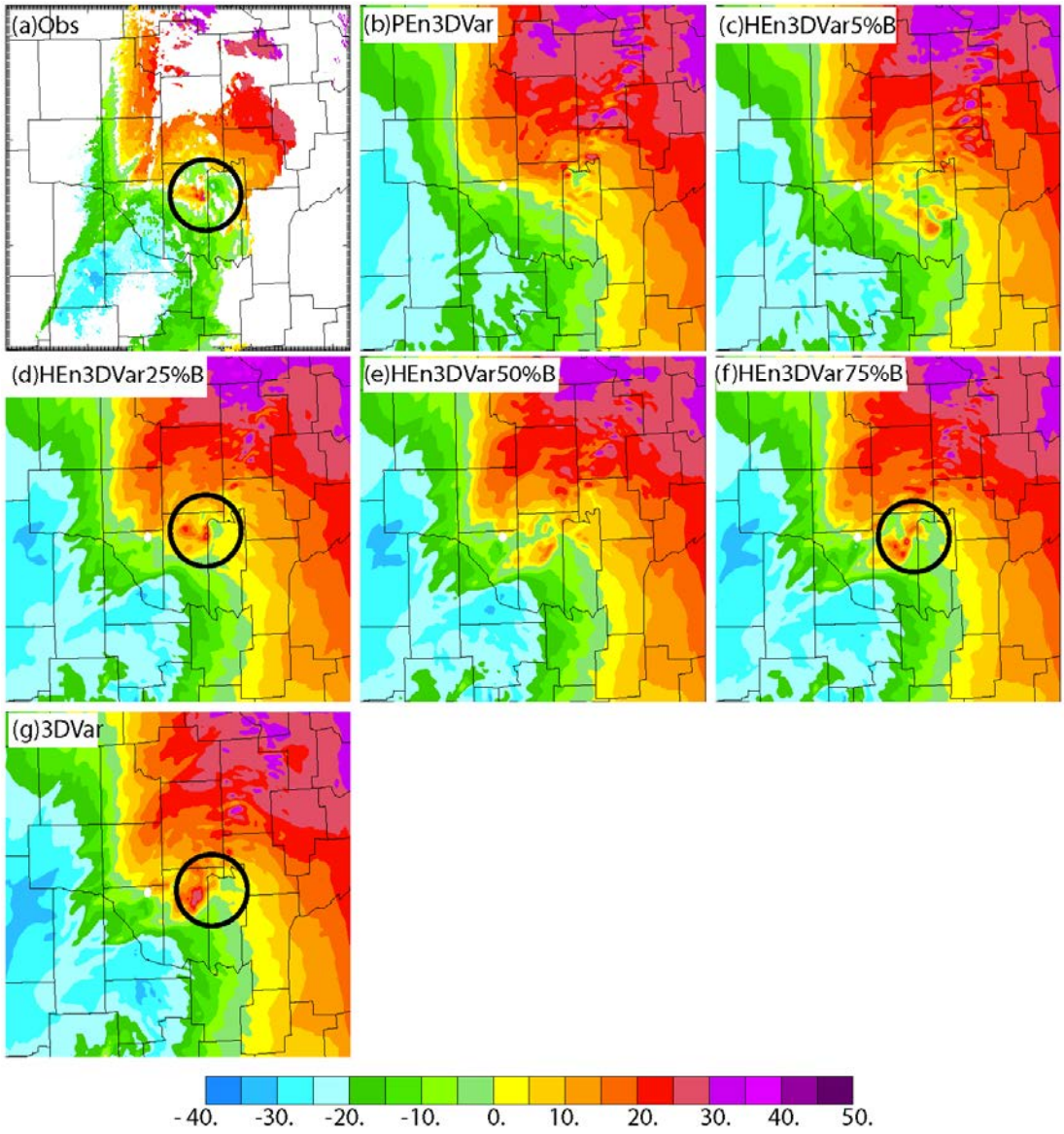

Fig. 5.11 (a) Observed and 45-minutes forecast radial velocity at the first elevation angle of radar KTLX from (b) pure En3DVar, hybrid En3DVar with (c) 5%, (d) 25%, (e) 50%, (f) 75% weight given to static background error covariance B, as well as (g) 3DVar.

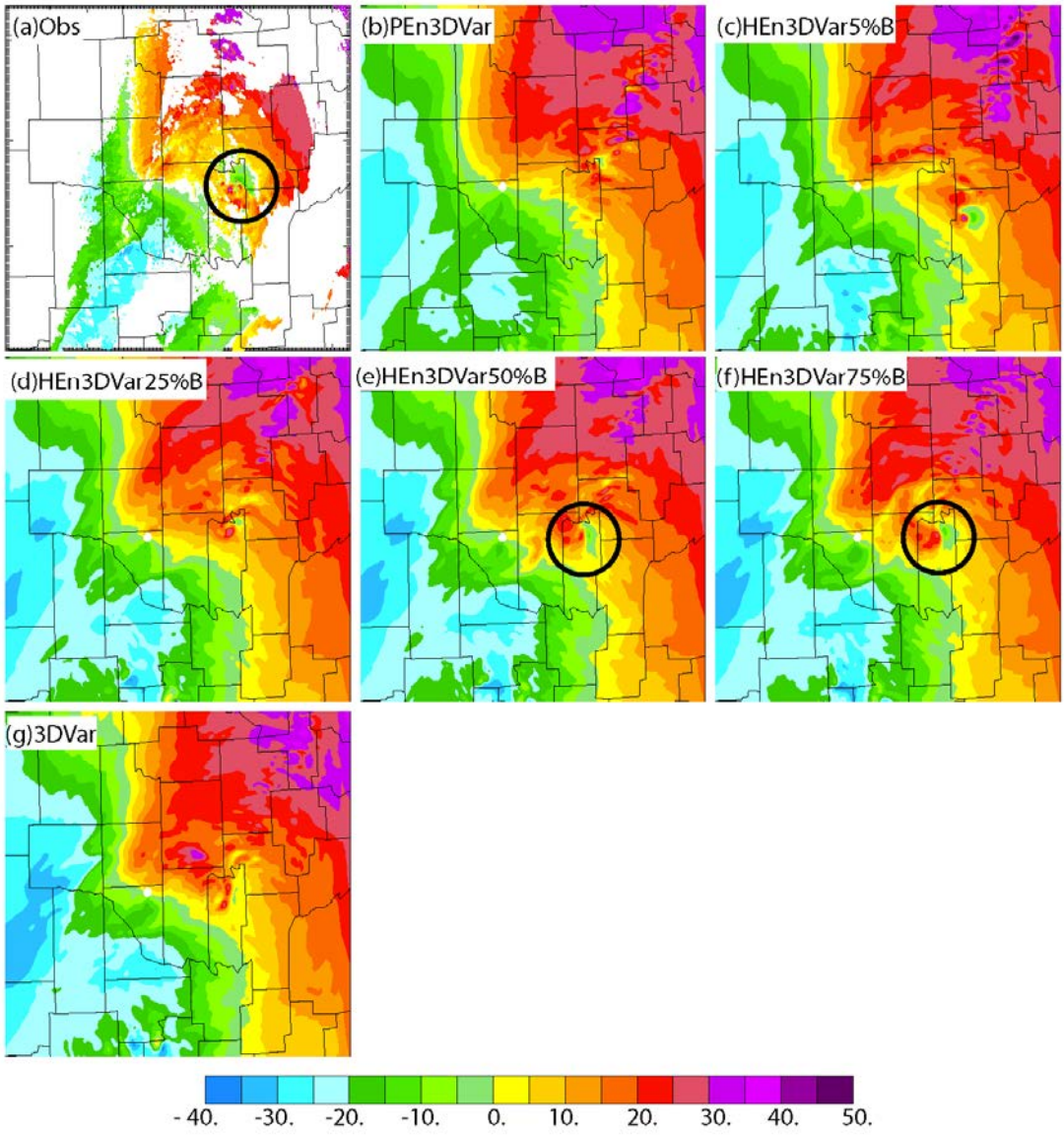

Fig. 5.12 (a) Observed and 60-minutes forecast radial velocity at the first elevation angle of radar KTLX from (b) pure En3DVar, hybrid En3DVar with (c) 5%, (d) 25%, (e) 50%, (f) 75% weight given to static background error covariance B, as well as (g) 3DVar.

5.3.3 Comparisons among 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar

a. Objective verification

1) The RMSE:

The analyses and forecasts of the radar radial velocity and reflectivity from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar are compared based on the RMSEs. For radial velocity analyses and forecasts within the one-hour DA window, EnKF outperforms the other algorithm in terms of the smallest RMSEs, suggesting averaging the velocity forecasts (for EnKF) can greatly decrease the errors relative to that uses a single deterministic background (for DfEnKF, 3DVar, pure and hybrid En3DVar). For the follow-on one-hour radial velocity forecasts, different algorithms perform similarly, and are all much better than the CNTL (Fig. 5.13). For reflectivity analyses and forecasts, hybrid En3DVar performs similarly as 3DVar and both are better than EnKF, DfEnKF, and pure En3DVar in terms of smaller RMSEs.

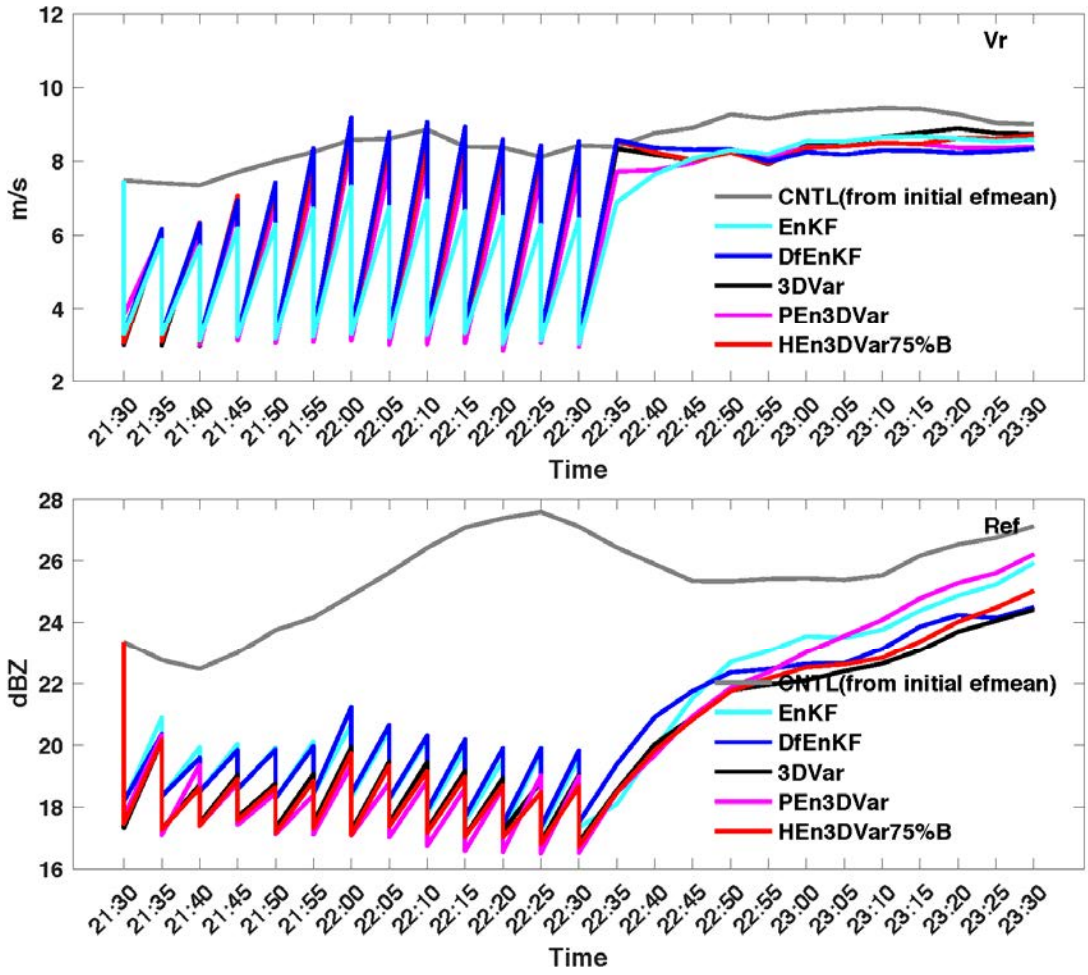

Fig. 5.13 RMSEs of the analyzed and forecasts radial velocity and reflectivity from CNTL, 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar.

2) The quantile-quantile verification (QQ-plot)

To have a better idea of the over and under performance presented in the RMSEs, QQ plot is used to further examine intensity distributions of the reflectivity analyses or forecasts against the observations. A QQ plot is a probability plot, which is the quantiles of the first data set (such as the analysis or forecast) against the quantiles of the second data set (such as the observations). If the two distributions being compared are similar, the points in the QQ plot will approximately lie on the 45-degree

diagonal line. A QQ plot can be used to compare the shapes of distributions, providing a graphic view of how properties such as intensity, locations, scales are similar or different in the two distributions. In this study, the QQ plot was used to examine the intensity distribution between the reflectivity analyses (forecasts) and the observations. To make the QQ plot, the reflectivity analyses (forecasts) and observations are sorted in increasing orders. The sorted analyses (forecasts) and the observations are then paired off and plotted using the corresponding values. The two data sets need to have the same size. Otherwise, interpolation is needed so that quantiles corresponding to the same underlying probability can be constructed.

To better understand the underperformance of EnKF and DfEnKF relative to the other algorithms (3DVar, pure and hybrid En3DVar) at the 15 minute DA window shown in Fig. 5.13, the forecast and analyzed reflectivity valid at 21:45 UTC were compared with the observations based on both the scatter plot (pair off the reflectivity analyses/forecasts and the observations without sorting) and the QQ plot (Fig. 5.14). According to the scatter plot (Fig. 5.14 a, b), the number of points located above the 45 degree diagonal line are a lot more than those below the diagonal line from all the experiments, indicative of overestimation of the forecasts (Fig. 5.14 a) and analyses (Fig 5.14 b) relative to the observations. The scatter points locate much closer to the diagonal line after reflectivity DA, indicating smaller overestimation from the analyses than those from the background forecasts. QQ plots more clearly capture the intensity distribution of the reflectivity forecasts and analyses (Fig. 5.14 c, d). The overestimation of the reflectivity intensity forecasts (analyses) is much stronger from EnKF and DfEnKF than those of the others (Fig. 5.14 c, d), which are consistent with

the larger RMSEs shown in Fig. 5.13. After the reflectivity DA, the overestimation for reflectivity above 20 dBZ is reduced, but still exists. For reflectivity less than 20 dBZ, different experiments perform similar, with the analyzed reflectivity a little bit smaller than the observations.

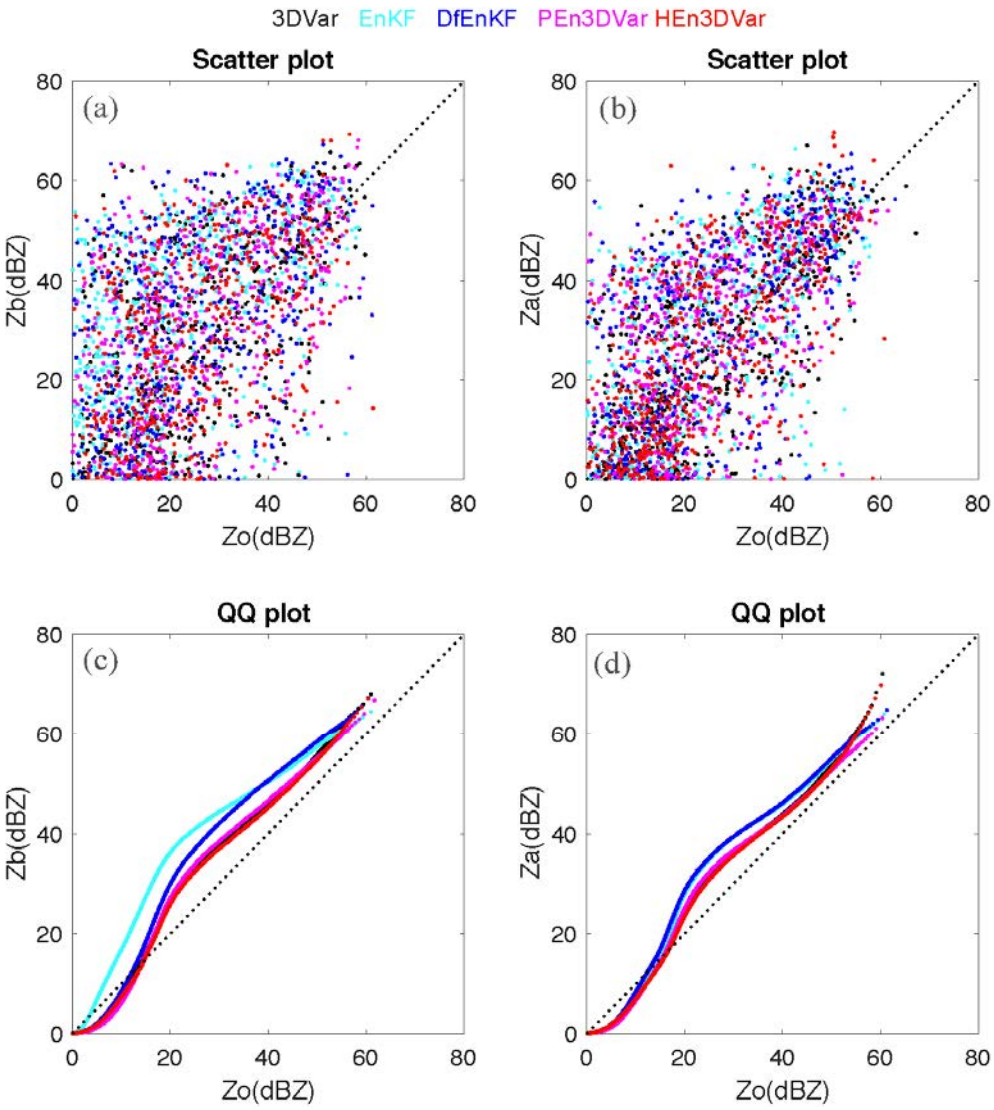

Fig. 5.14 The scatter-plot (a, b) and QQ-plot (c, d) for reflectivity forecasts (a, c) and analyses (b, d) at the 15-minute DA window (valid at 22:45 UTC).

After one-hour DA, the forecast/analyzed reflectivity was still overestimated relative to the observation, with the overestimation from the analysis a little bit smaller

than from the forecasts. The overestimation from EnKF and DfEnKF are still more obvious than those of the others (Fig. 5.15).

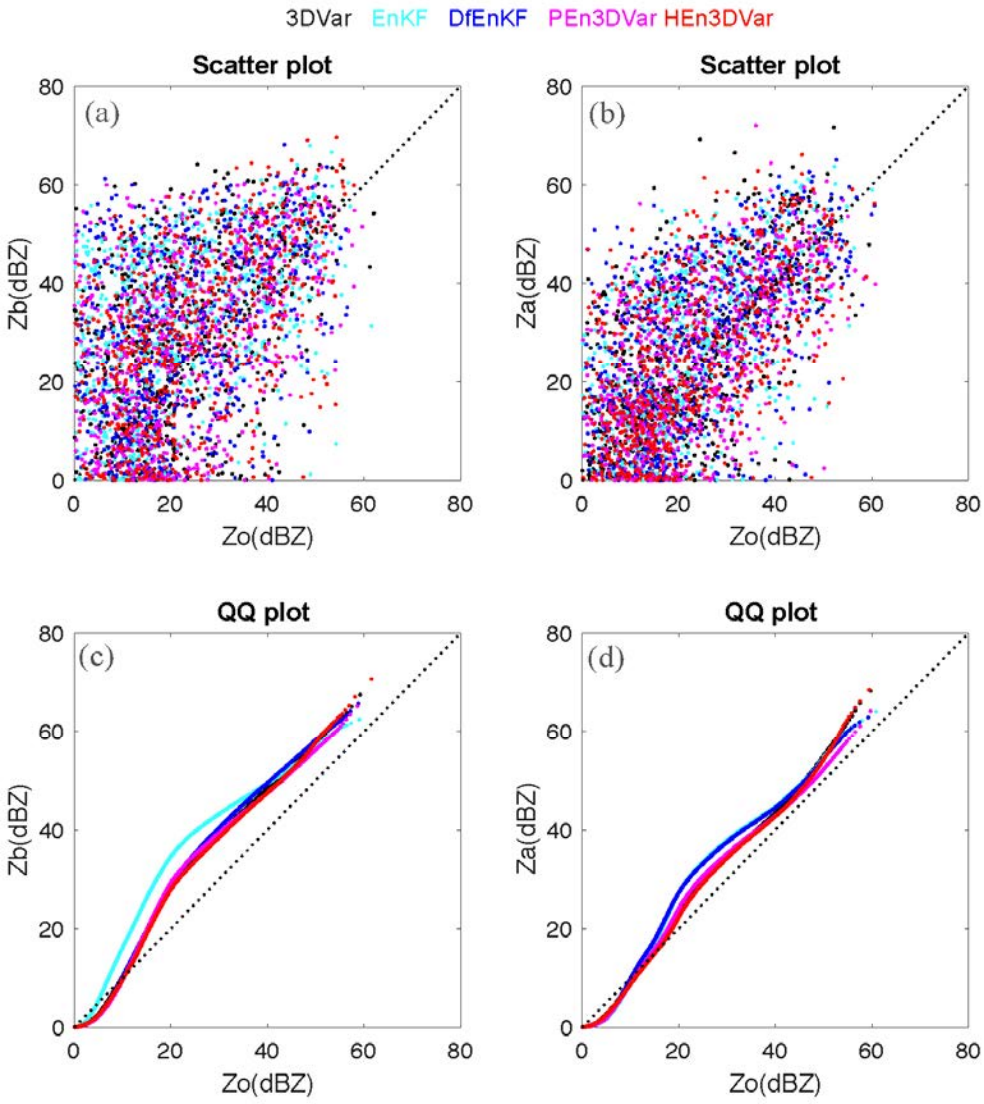

Fig. 5.15 The scatter-plot (a, b) and QQ-plot (c, d) for reflectivity forecasts (a, c) and analyses (b, d) at the end of one-hour DA window (valid at 22:30 UTC).

The reflectivity forecasts from different experiments all underestimate the 30-minute and 60-minute forecasts when the observed reflectivity is less than 20 dBZ, and

overestimates otherwise. The corresponding underestimation (overestimation) from EnKF is a little bit larger (smaller) than those of the others (Fig. 5.16).

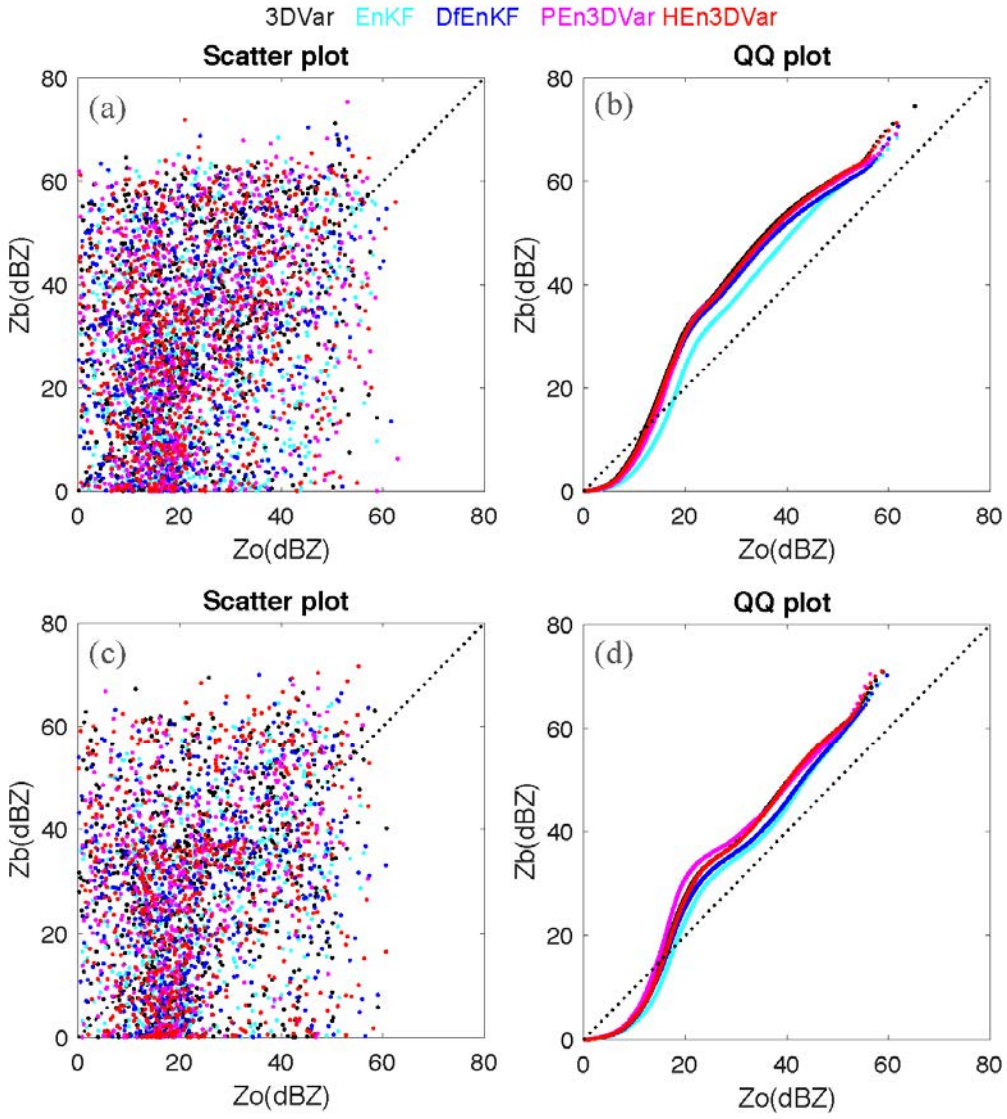

Fig. 5.16 The scatter-plot (a, c) and QQ-plot (b, d) for 30-minute (a, b) and 60 minute (c, d) reflectivity forecasts from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar.

3) Fractions Skill Score Verification

The purpose of the Fractions Skill Score (FSS) is to attain a measure of how analysis/forecast skill varies with spatial scale and intensity threshold. Radar data observations from four S-band WSR-88D radars (KTLX, KINX, KFDR, KVNX) are used for the comparison with the analyzed and forecast reflectivity. FSS belongs to the category of spatial neighborhood verification that enables the forecast being compared with the observations within specified neighborhood rather than strictly at the grid scale. The fractional coverage of the forecast or analysis within a neighborhood can be assessed for similarity to the observations. By varying the sizes of the neighborhood, the scales for which the forecast or analysis has sufficient skill can be determined and compared (Ebert 2008; Roberts and Lean 2008). The FSS is calculated based on a variation of the Brier Skill Score (FBS). According to Roberts and Lean (2008), the FSS can be calculated by:

$$\mathrm{FSS} = 1 - \frac{\mathrm{FBS}}{\mathrm{FBS}_{\mathrm{worst}}}, \tag{5.6}$$

where FBS is a variation of the Brier Score (Brier, 1950) that is given by:

$$\mathrm{FBS} = \frac{1}{N}\sum_{j=1}^{N}\left(O_j - M_j\right)^2 . \tag{5.7}$$

Here N is the total number of grid points. M_j and O_j are the forecast (or analyzed) and observed reflectivity fractions respectively at different grid point j, which are determined based on reflectivity thresholds with its values ranged between 0 and 1. A forecast with perfect skill has a score of 1; a score of 0 means zero skill. Skill is lowest at the grid scale, that is, when the neighborhood is only one grid point and the fractions are binary ones or zeros. Figure 5.17 gives a schematic example of how fractions are

computed over different neighborhood scales (1, 3, and 5 in this example). When the neighborhood box is 1 grid scale, i.e., only the central box is considered, the forecast/analysis fraction is 0/1=0, while the observation fraction is 1/1=1. When the neighborhood box is 3 times of the grid scale, the corresponding forecast and observation fractions are 4/9 and 3/9 respectively. When neighborhood box is 5 times of the grid scale, the observation and forecast (analysis) fractions are the same, which is 6/25. For more information about FSS, the readers are referred to Roberts and Lean (2008).

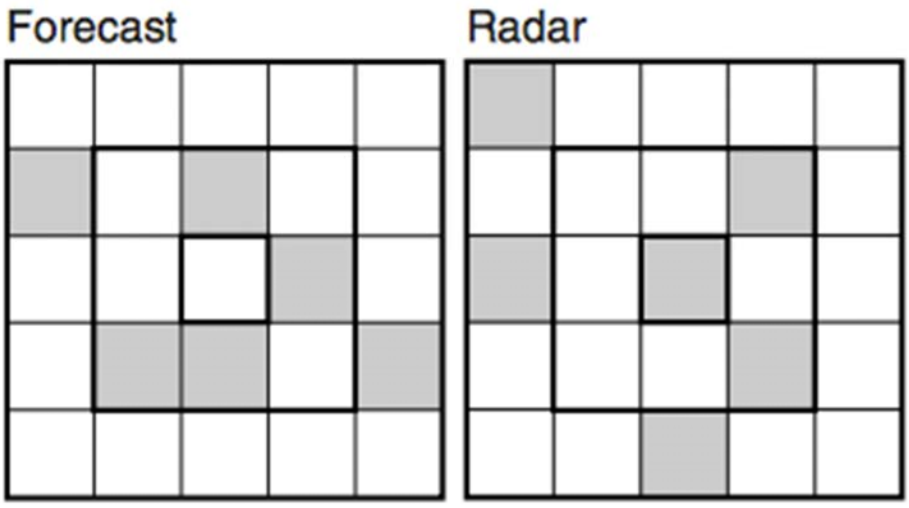

Fig. 5.17 A schematic comparison between forecast and observed reflectivity (from Roberts and Lean (2008)).

FSS from different experiments are calculated as functions of the reflectivity thresholds (5, 15, 35, 40, 60 dBZ) and sizes of the neighborhood scales (1, 5, 10, 20, 40, 80, 160, 320 km). FSS for reflectivity analysis at the end of one-hour DA window from hybrid En3DVar is compared with those of the others to see whether it is able to produce more skillful reflectivity analyses at different scales, and how the skill varies

with spatial scale and reflectivity threshold. The FSS for the reflectivity analyses and forecasts within the DA windows are plotted as functions of neighborhood scales for different reflectivity thresholds (Fig. 5.18). Unlike the 60-dBZ reflectivity analyses/forecasts, for which the FSS is not very sensitive to the neighborhood scales, the FSSs for 5-dBZ, 15-dBZ, 35-dBZ, and 45-dBZ reflectivity all increase with the neighborhood scales. This indicates that reflectivity analyses/forecasts on smaller neighborhood scales are less skillful than those on the larger scales. This is reasonable because the larger the neighborhood scale, the less the penalty given to typical location errors. When comparing the FSS among different algorithms, EnKF and DfEnKF are less skillful than 3DVar, pure and hybrid En3Dvar for the 35-dBZ and 45-dBZ reflectivity thresholds. EnKF is most skillful for the 60-dBZ reflectivity threshold. In addition, pure En3DVar (DfEnKF) performs the worst (best) for 5-dBZ (15-dBZ) reflectivity threshold in terms of the smallest (largest) FSS. Overall, 3DVar and hybrid En3DVar are most skillful for analyses of the 35 and 45 dBZ reflectivity, for higher intensity (e.g., 60 dBZ), EnKF performs much better than those of the others.

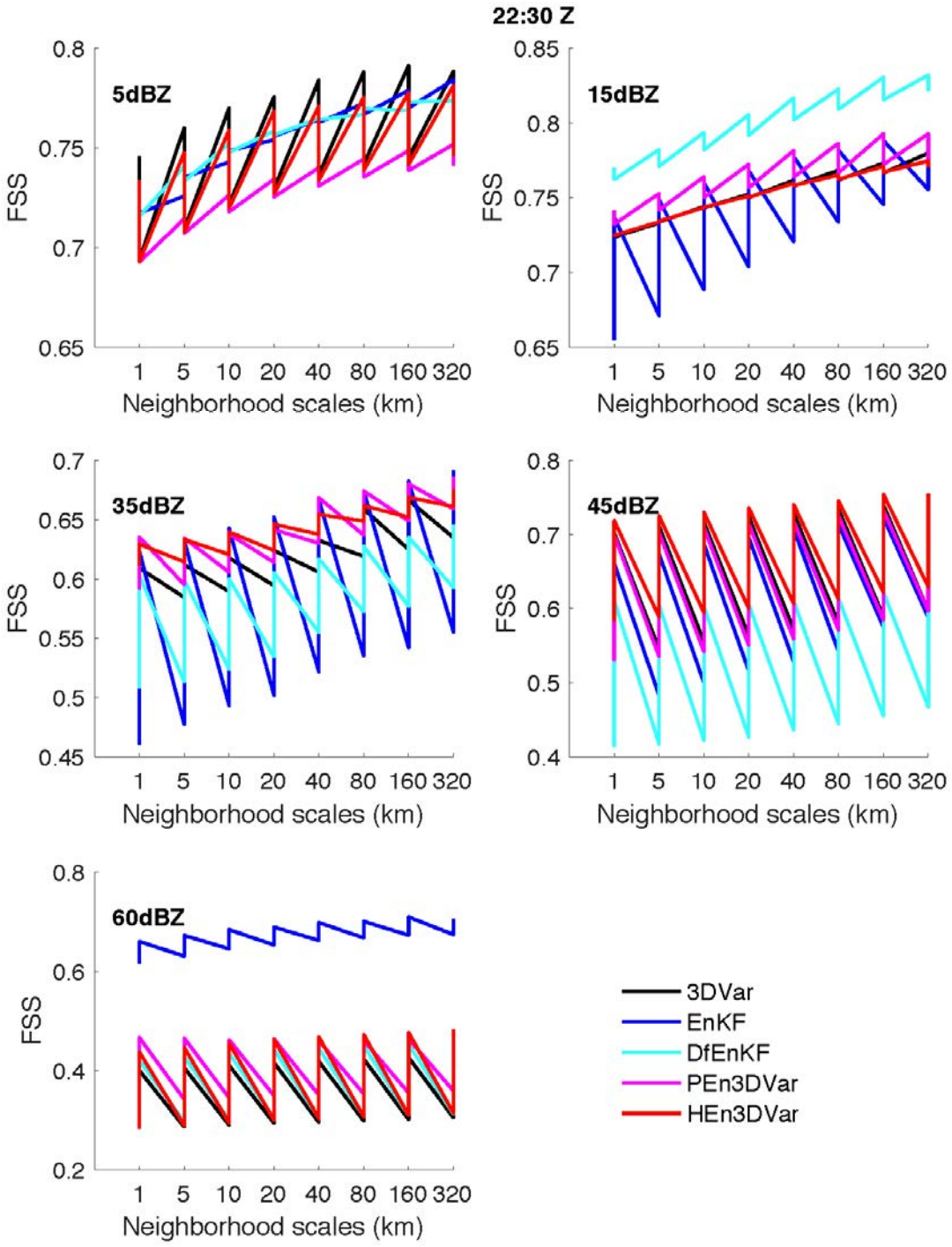

Fig. 5.18 The fractional skill scores for reflectivity analyses/forecasts at the end of one-hour DA window (22:30 UTC) from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar as functions of the neighborhood scales and for different thresholds.

Similar to the reflectivity analyses at the end of one-hour DA window, the FSSs for one-hour reflectivity forecasts also increase with the neighborhood scales (Fig. 5.19). Different from the conclusion obtained for the reflectivity analyses, for which lower-value reflectivity analyses are more sensitive to the neighborhood scales, the FSS for higher-value reflectivity (15, 35, 45, 60 dBZ) forecasts are more sensitive to the neighborhood scales relative to the lower-value reflectivity (5 dBZ), indicative of lower spatial accuracy in higher reflectivity intensity. Overall, 3DVar is more skillful than hybrid En3DVar, and both are more skillful than the others for small reflectivity thresholds (5dBZ and 15 dBZ). For higher reflectivity thresholds of 35 dBZ and 45 dBZ, pure En3Dvar and EnKF perform worse than the others, and the other algorithms perform similarly. For even higher reflectivity threshold (60 dBZ), EnKF is most skillful, and 3DVar has the poorest performance.

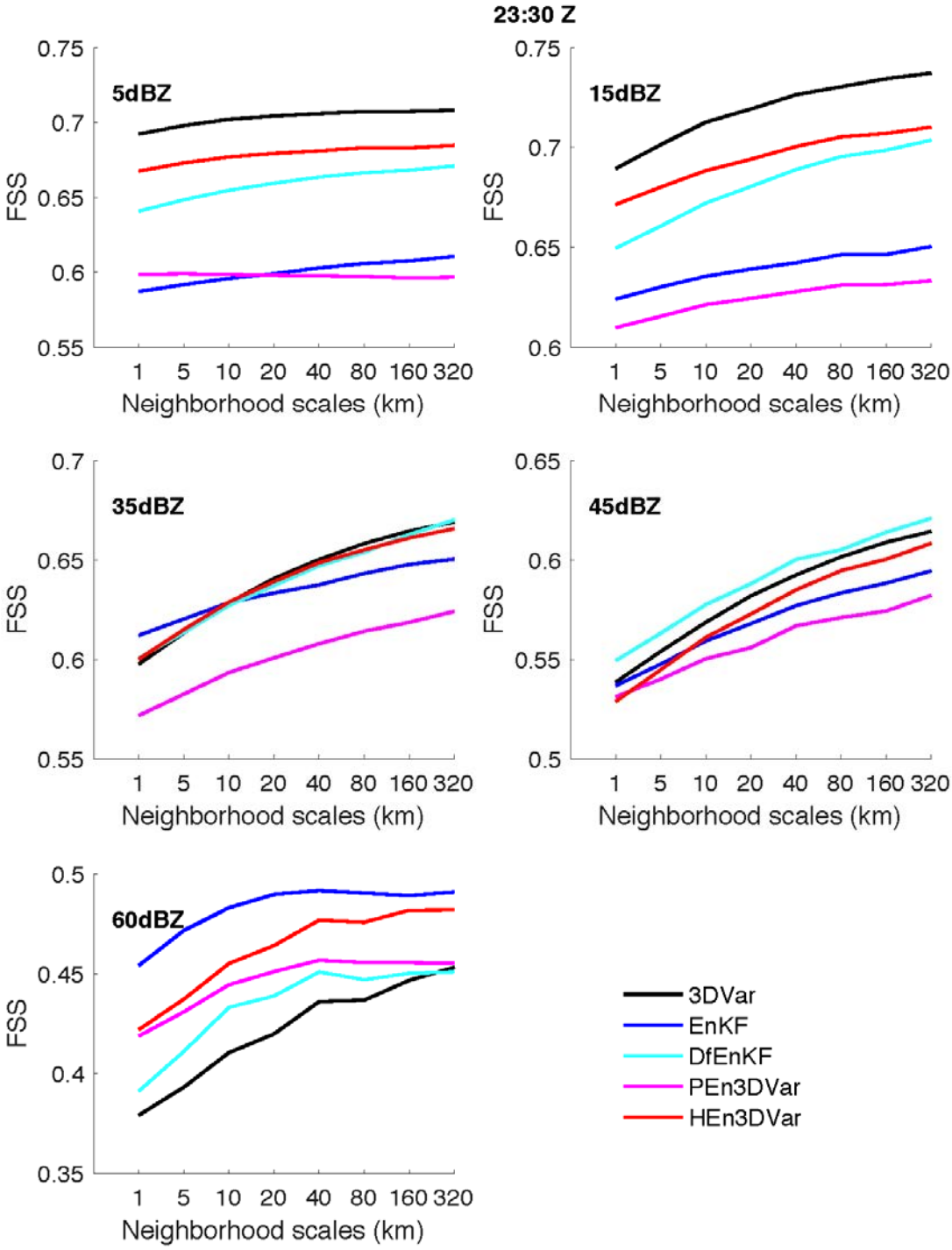

Fig. 5.19 The fractional skill scores for one-hour reflectivity forecasts (23:30 UTC) from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar as functions of the neighborhood scales and for different reflectivity thresholds.

The FSSs for reflectivity analyses/forecasts at the end of the one-hour DA window are also examined as functions of the reflectivity intensity (Fig. 5.20). On average, the FSS decreases as the reflectivity intensity increases, the higher the intensity, and the lower the FSS except EnKF. For EnKF, the 60-dBZ reflectivity analyses/forecasts are more skillful than those of the 35-dBZ and 45-dBZ. For reflectivity analyses at 35 and 45 dBZ, 3DVar, pure and hybrid En3DVar are more skillful. DfEnKF is most skillful for 15-dBZ reflectivity analyses/forecasts.

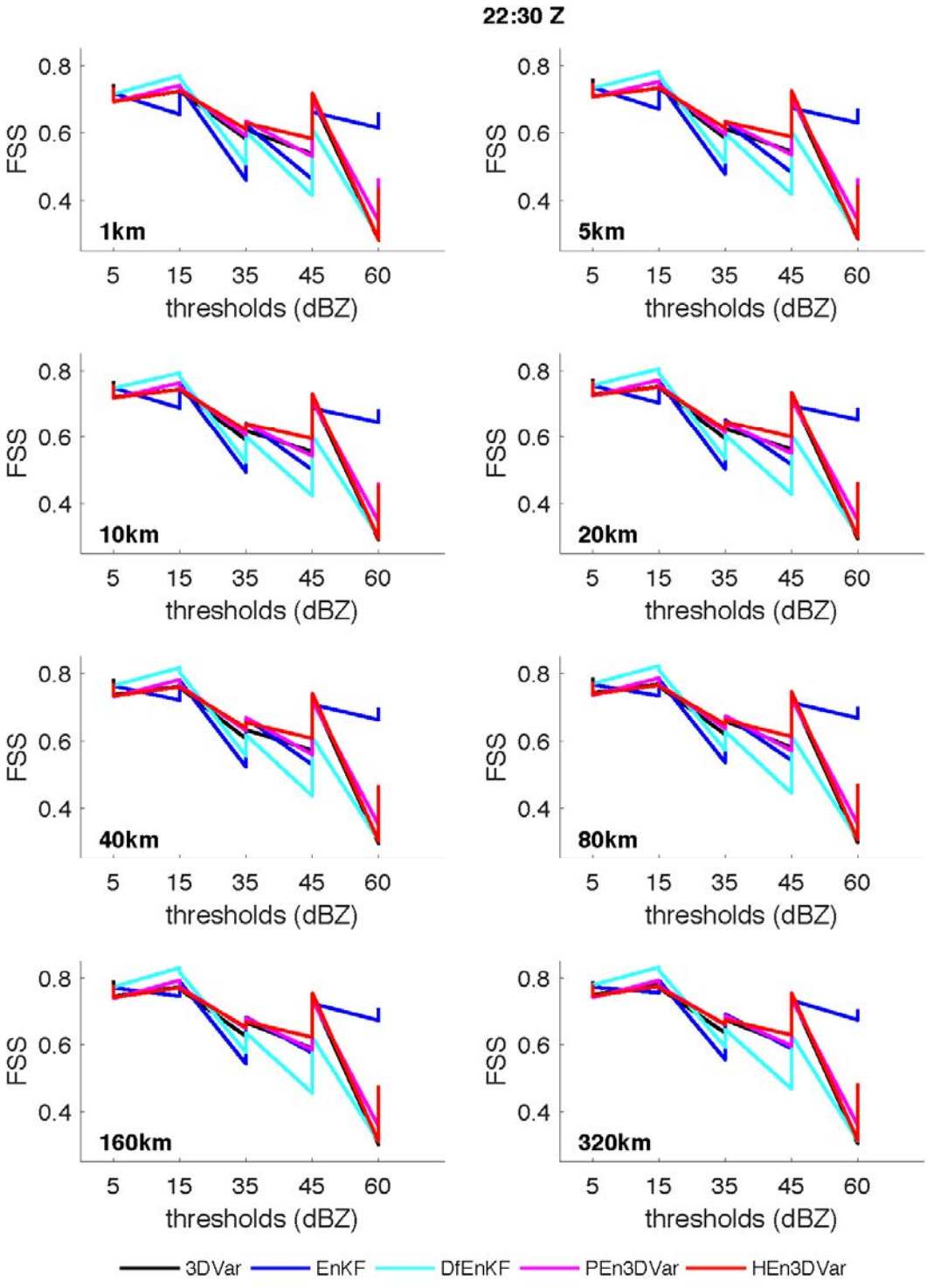

Fig. 5.20 The fractional skill scores for reflectivity analysis at the end of one-hour DA window (22:30 UTC) from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar as functions of the reflectivity thresholds for different neighborhood scales.

The FSSs for one-hour reflectivity forecasts are examined as functions of the reflectivity intensity (Fig. 5.21). On average, the FSS decreases as the reflectivity intensity increases, indicating more skillful forecast in smaller reflectivity intensity that has a broader spatial coverage. Overall, 3DVar and hybrid En3DVar outperform EnKF, DfEnKF, and pure En3DVar for small reflectivity intensity (< 45 dBZ), for higher intensity (60 dBZ), EnKF performs the best.

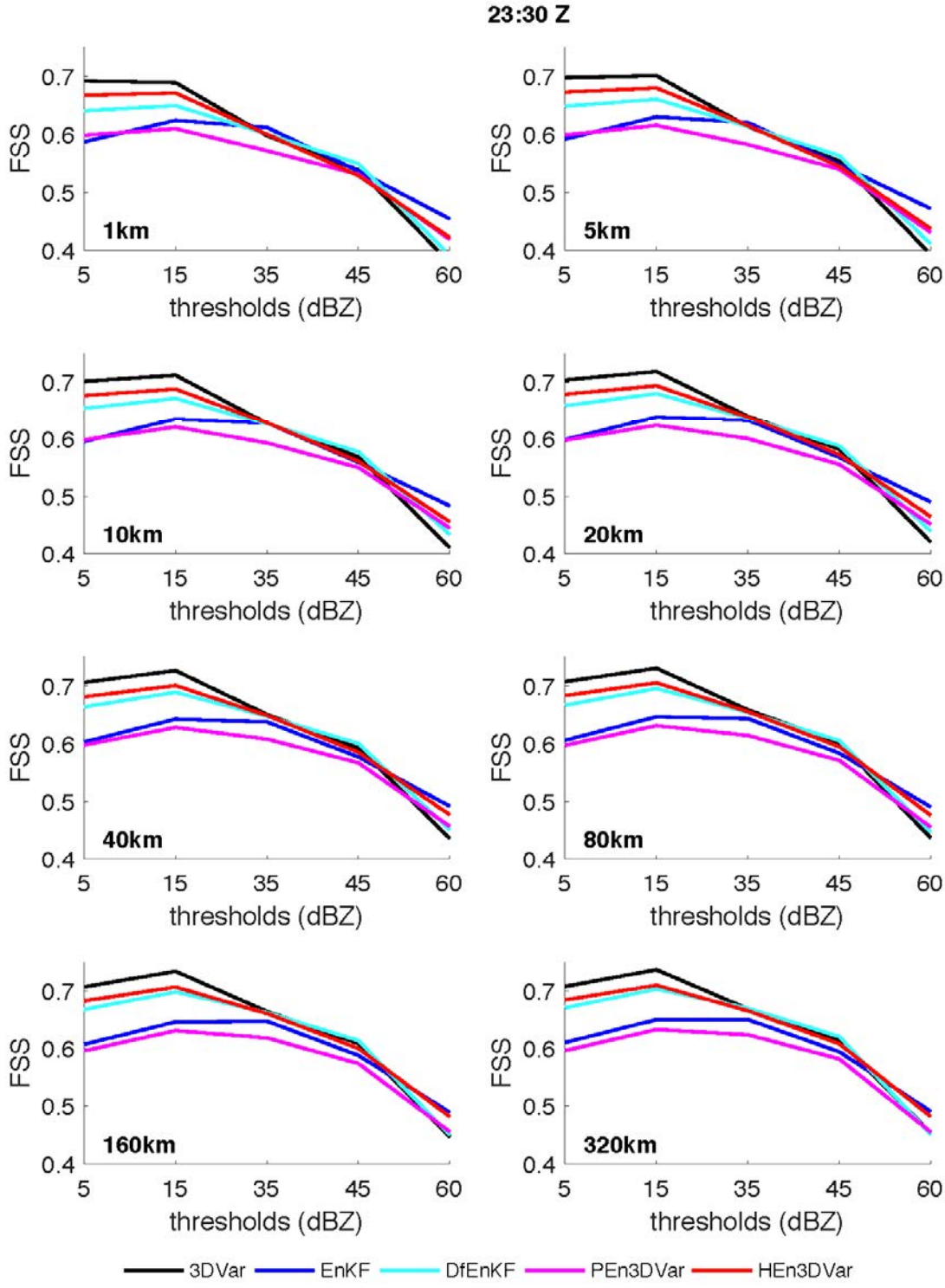

Fig. 5.21 The fractional skill scores for one-hour reflectivity forecasts (23:30 UTC) from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar as functions of the reflectivity thresholds for different neighborhood scales.

b. Analyses and forecasts of reflectivity

Comparisons of the reflectivity analyses and forecasts at 1 km above the ground level (AGL) are made among observation, 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar (Fig. 5.22). At the end of one-hour DA window, the reflectivity analyses from different algorithms have similar patterns that are close to the truth, with the maximum reflectivity analyses from EnKF (53.0 dBZ) being a little bit weaker than those of the observations (58.8 dBZ). While 3DVar (69.7 dBZ), pure En3DVar (64.3 dBZ), and hybrid En3DVar (64.3 dBZ) all overestimate the maximum intensity of the analyzed reflectivity, and DfEnKF has the maximum reflectivity analyses that is the same as the observation (58.8 dBZ).

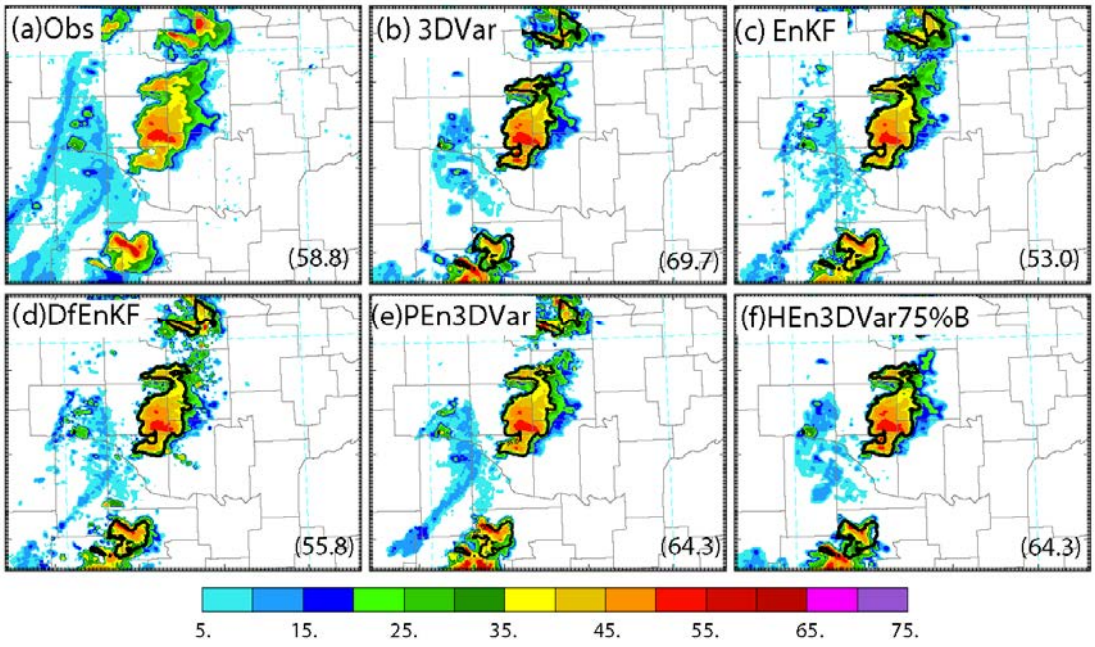

Fig. 5.22 The reflectivity observation at 1km AGL (a), and the reflectivity analyses at the end of one-hour DA from, (b) 3DVar, (c) EnKF, (d) DfEnKF, (e) pure and (f) hybrid En3DVar (75% weight given to **B**) overlaid with the observed reflectivity with its value higher than 35dBZ (black line). The maximum reflectivity is indicated by the value at the southeast corner.

30-minute reflectivity forecasts at 1-km AGL are compared among different experiments (Fig. 5.23). EnKF performs the worst for underestimating the intensity and coverage of the storms. DfEnKF outperform EnKF with the coverage and intensity of the storms being much closer to the observation. Pure En3DVar fails to capture the storm cell near the southwest corner. 3DVar and HEn3DVar performs similar in capturing the intensity and pattern of the storm. Overall, DfEnKF, 3DVar, and hybrid En3DVar provide very good 30-minute forecasts that are close to the observation.

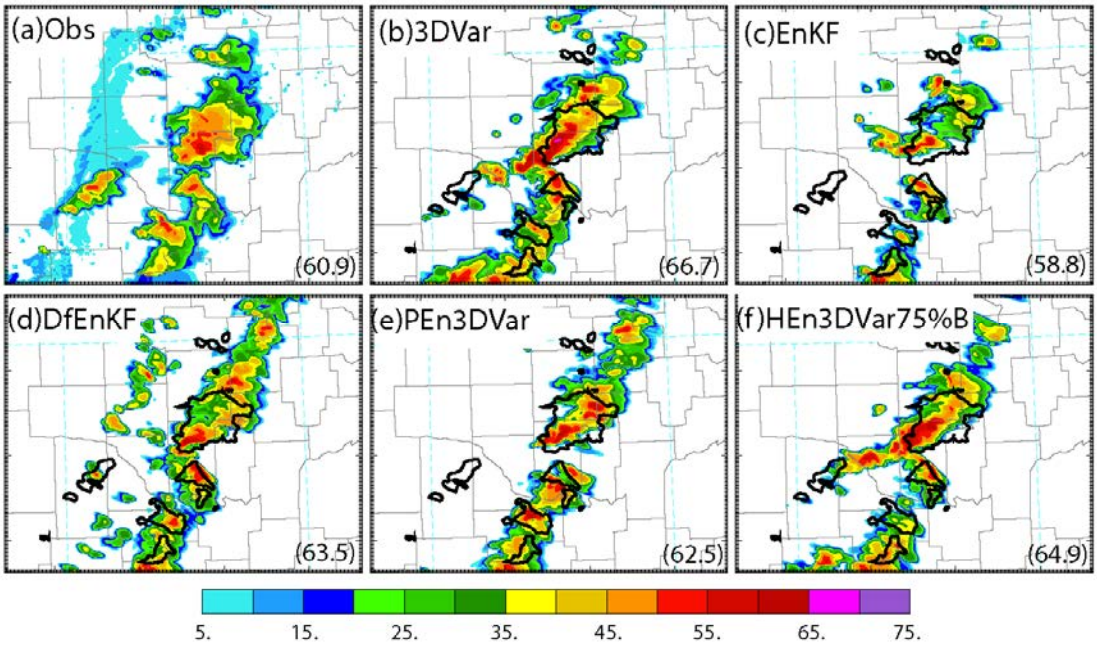

Fig. 5.23 The reflectivity observation at 1km AGL (a), and 30-min reflectivity forecasts based on analyses of (b) 3DVar, (c) EnKF, (d) DfEnKF, (e) pure and (f) hybrid En3DVar (75% weight given to **B**) overlaid with the observed reflectivity with its value higher than 35dBZ (black line). The maximum reflectivity is shown at the southeast corner.

45-minute reflectivity forecasts at 1-km AGL are compared among different experiments (Fig. 5.24). EnKF and DfEnKF perform the worst for underestimating coverage of the storms. 3DVar and HEn3DVar performs similar in capturing the

intensity of the storm. However, the storm structure from 3DVar tends to be too linear. The hook echo structure is more clearly captured by HEn3DVar compared to 3DVar.

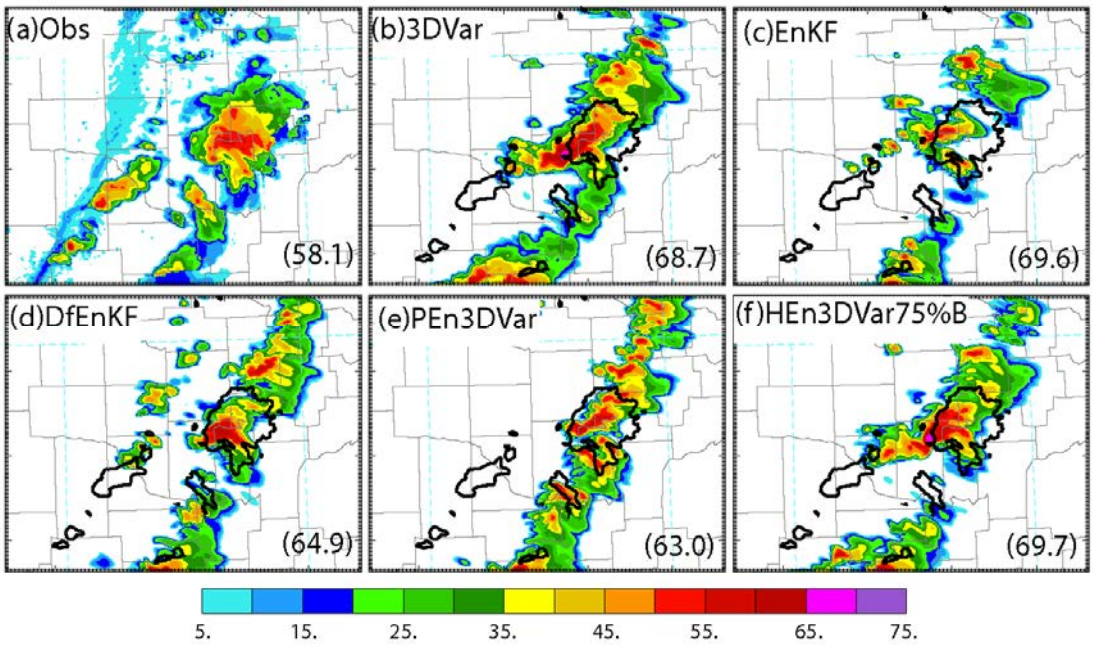

Fig. 5.24 Same as Fig. 5.23 except that it is for the 45-minute reflectivity forecasts.

c. Forecasts of updrafts and downdrafts

Strong updrafts and downdrafts occur in thunderstorms. The intensity of the vertical updraft/downdraft determines the degree of storm severity. Updraft is inevitable for storm's early development, during which warm air rises to the lifted condensation level. Downdrafts are usually caused by evaporational cooling, which originates at mid-levels with dry air that becomes cold, then the cold dense air sinks to the ground and spreads out laterally. The strong updraft and rear flank downdraft in the storm cause the weak echo region (WER) and the hook echo feature observed on radar. The hook is

caused by the interaction of the rear flank downdraft and updraft with the resultant precipitation wrapping around the southwest side of the updraft.

To compare the structure of the forecast updraft valid at 23:30 UTC, vertical cross-sections of the one-hour velocity forecasts are made through their own maximum values from experiments 3DVar, EnKF, DfEnKF, and HEn3DVar75%B (Fig. 5.25). The updrafts forecasts from experiments EnKF, DfEnKF, and HEn3DVar75%B are much stronger and more consistent than that of 3DVar. There is some tilting for updrafts form DfEnKF and HEn3DVar75%B, while the updraft from EnKF is more erect. The increasing of wind speed with altitude causes the updraft to tilt. Because it tilts, when precipitation falls, it does not fall into the updraft, thus allowing a more continuous source of moist warm air to fuel the storm for experiments DfEnKF and HEn3DVar75%B. The updraft strength from DfEnKF and HEn3DVar75%B are 44.2 m s^{-1}, and 46.9 m s^{-1} respectively, which are slightly larger than those from EnKF and 3DVar (39.9 m s^{-1}). In a mature storm, downdrafts are present alongside with updrafts. Compared to the updrafts, the intensity of the downdrafts, especially the rear flank downdrafts (RFD), are much weaker, which are -16.1 m s^{-1} for experiment 3DVar, -12.4 m s^{-1} for EnKF, -17.5 m s^{-1} for DfEnKF, and -10.4 m s^{-1} for HEn3DVar75%B respectively. Such kind of storms often forms in very moist atmosphere where there is little mid-level dry air and thus evaporational cooling to drive the downdrafts. The RFD from 3DVar is a little bit stronger than those from the other experiments, which is related to the much drier mid-level (700 hPa) atmosphere relative to those of the others (Fig. 5.26).

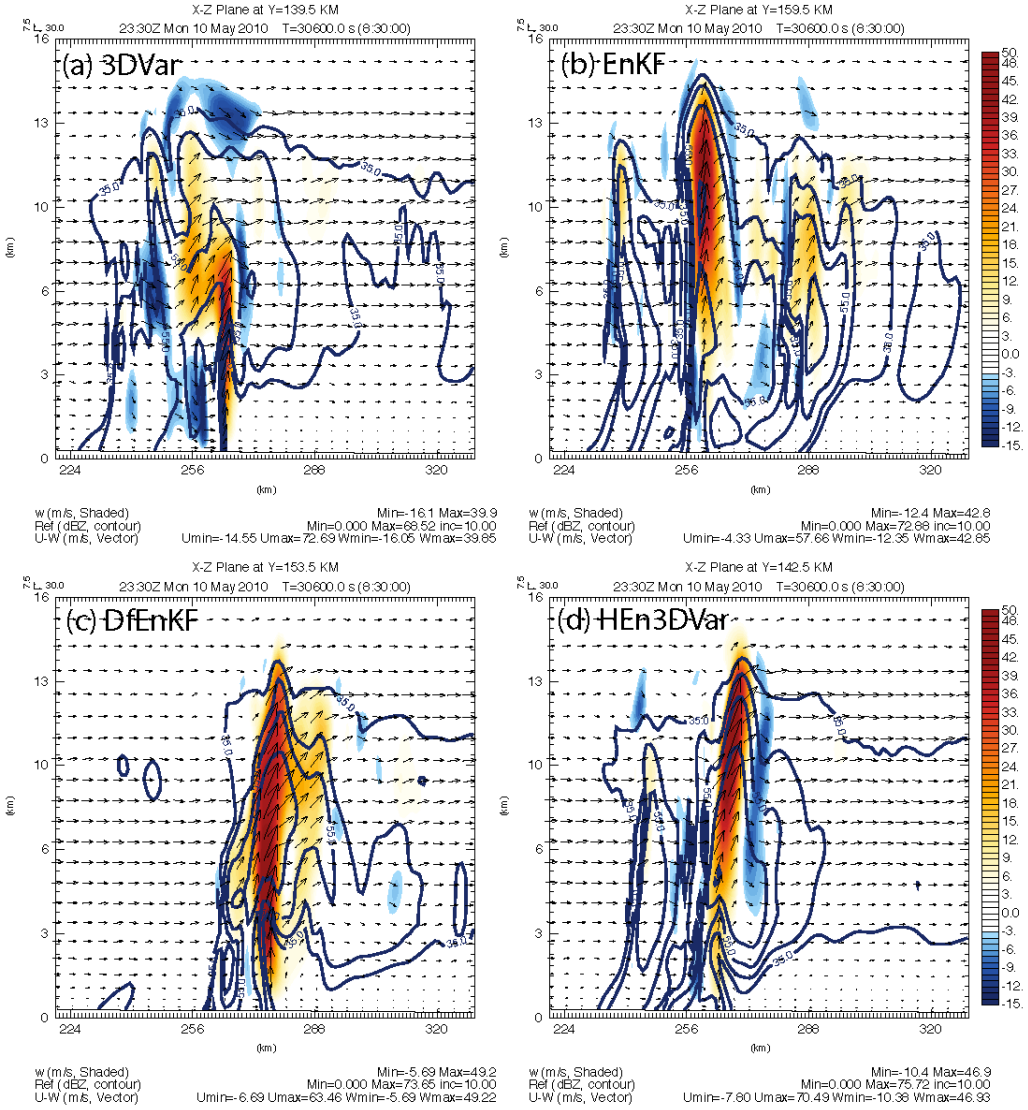

Fig. 5.25 Vertical cross-sections of the one-hour vertical velocity forecasts valid at 23:30 Z for experiments 3DVar, EnKF, DfEnKF, HEn3DVar75%B, overlaid with the wind vector field and one-hour reflectivity forecasts with their value higher than 35 dBZ, and with an interval of 10 dBZ.

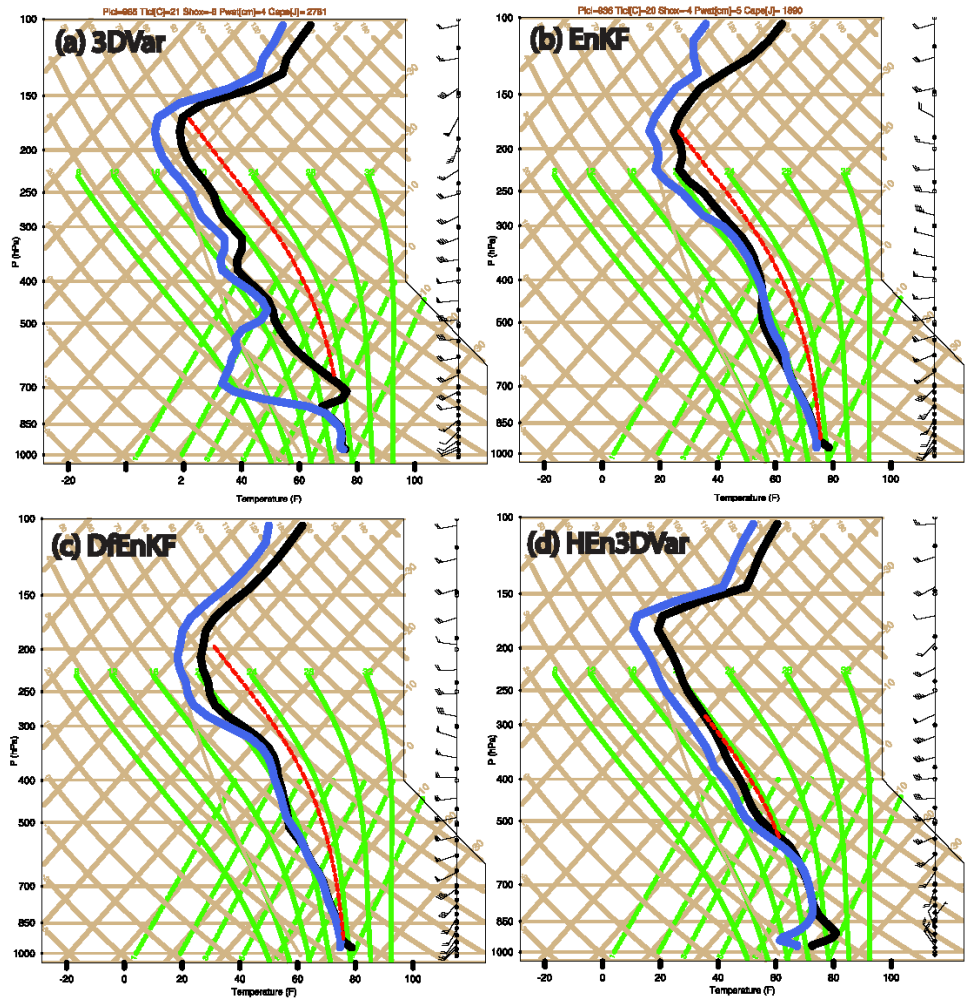

Fig. 5.26 Model soundings taken through the maximum one-hour vertical velocity forecasts of different experiments for experiments (a) 3DVar, (b) EnKF, (c) DfEnKF, and (d) HEn3DVar75%B.

d. Forecasts of the meso-cyclone

The rotating updraft is known as the mesocyclone, which is a key to the development of a supercell, and the potential for a tornado. Once the updraft is rotating and being fed by warm, moist air flowing in at ground level, conditions are ripe for tornado formation. To compare the intensity of the rotating updraft among different experiments, vertical cross-sections of the vertical vorticity forecasts through the maximum one-hour vertical velocity forecasts are made and compared among different experiments (Fig. 5.27). The vertical cyclonic vorticity indicative of rotations of the updrafts from experiments EnKF, DfEnKF, and HEn3DVar75%B are 2.18×10^{-1} s^{-1}, 1.98×10^{-1} s^{-1}, and 1.66×10^{-1} s^{-1} respectively, which are much stronger and more consistent than that form 3DVar (1.25×10^{-1} s^{-1}).

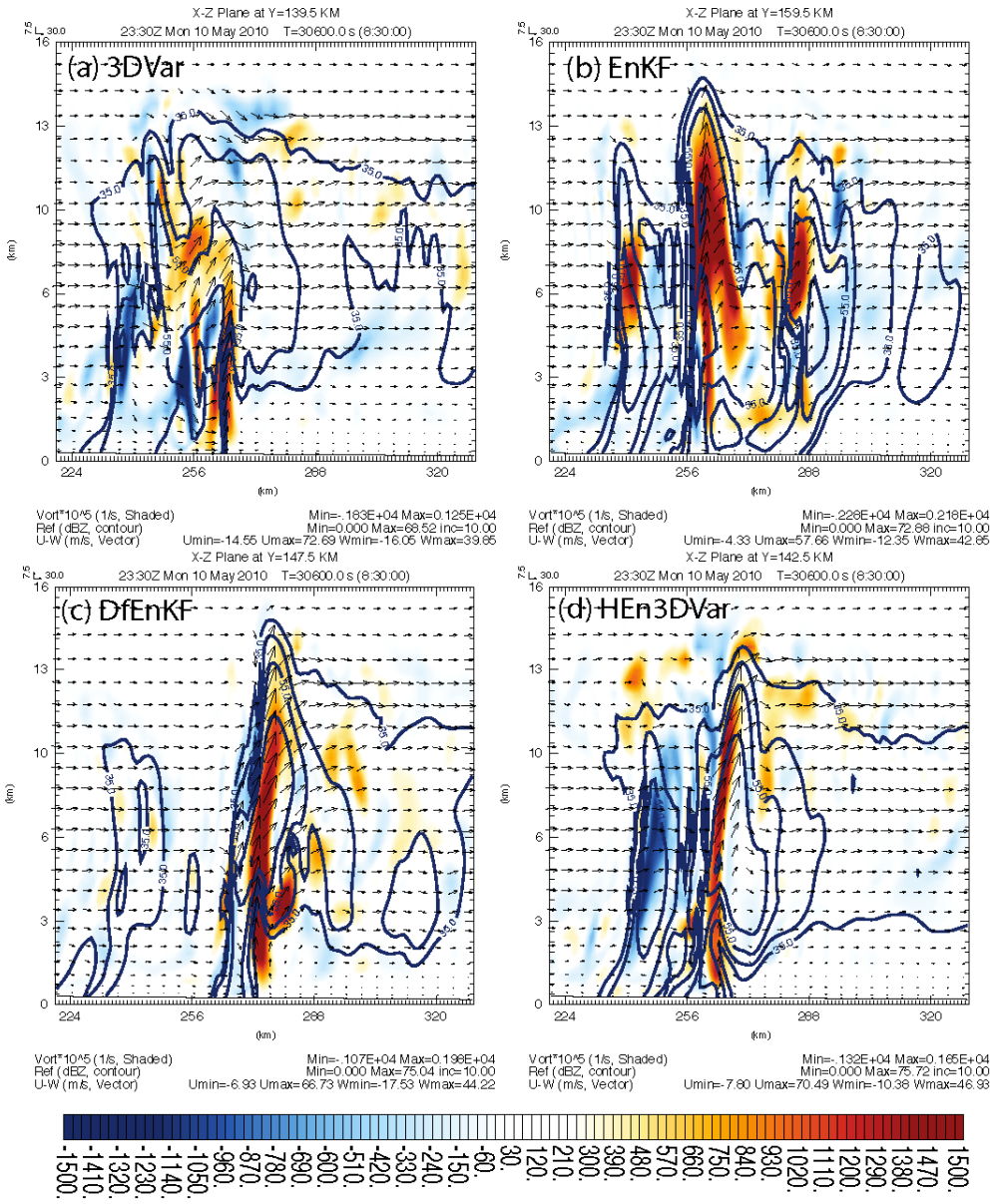

Fig. 5.27 Vertical cross sections of the vertical vorticity forecasts (units 10^{-5} s^{-1}) valid at 23:30 Z for experiments 3DVar_logqx, EnKF, DfEnKF, HEn3DVar75%B, overlaid with the wind vector field and one-hour reflectivity forecasts with their value higher than 35 dBZ, and with an interval of 10 dBZ.

A way to effectively detect and verify the presence of a mesocyclone is by Doppler weather radar. Nearby high values of radial velocity data with opposite sign (the velocity couplet) indicate the mesocyclone. The velocity couplet for which bright

reds (outbound) and bright greens (inbound) are next to each other indicates rotation in the storm. It is extremely important to determine if a supercell is strongly rotating and poses a tornado threat. The radial velocity analyses at the lowest elevation angle of radar KTLX are compared among experiments 3DVar, EnKF, DfEnKF, and HEn3DVar75%B (Fig. 5.28). All the analyses are able to capture the strengthening of velocity at the corner of "L-shaped" outbound velocity (red color), indicating stronger low-level convergence when diagnosing together with the inbound velocity (green color). The 0 ~ 60 min radial velocity forecasts at the lowest two elevation angles (0.59^o and 0.92^o) of radar KTLX from experiments 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar are shown in Fig. 5.29 and Fig. 5.30. The velocity couplet is more clearly captured by experiment HEn3DVar75%B than those from 3DVar, EnKF, and hybrid En3DVar, indicative of stronger convergence. The rotation of the velocity couplet from 0.92^o elevation angle (Fig. 5.30) is a little bit larger than that on the 0.59^o elevation angle (Fig. 5.29), and they are all smaller than that of the observation based on orientations of the velocity couplets..

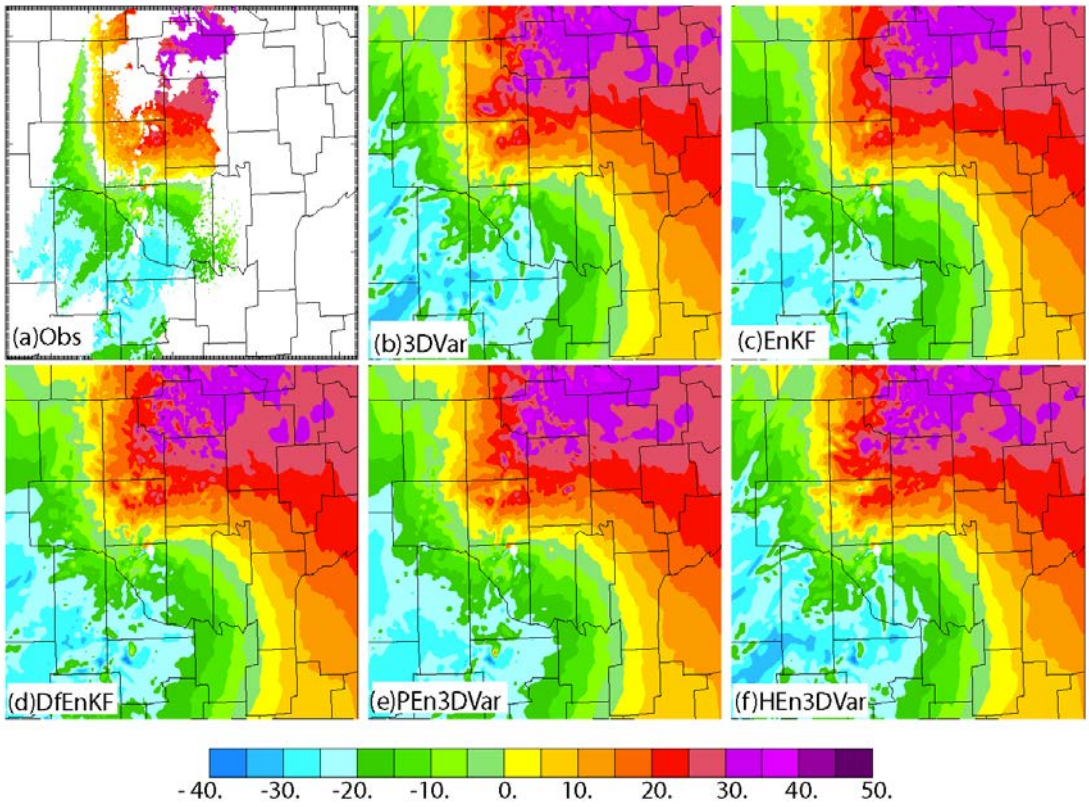

Fig. 5.28 The observed and analyzed radial velocity at the lowest elevation angle of KTLX from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar with 75% weight given to B at the end of one-hour DA window (22:30 UTC).

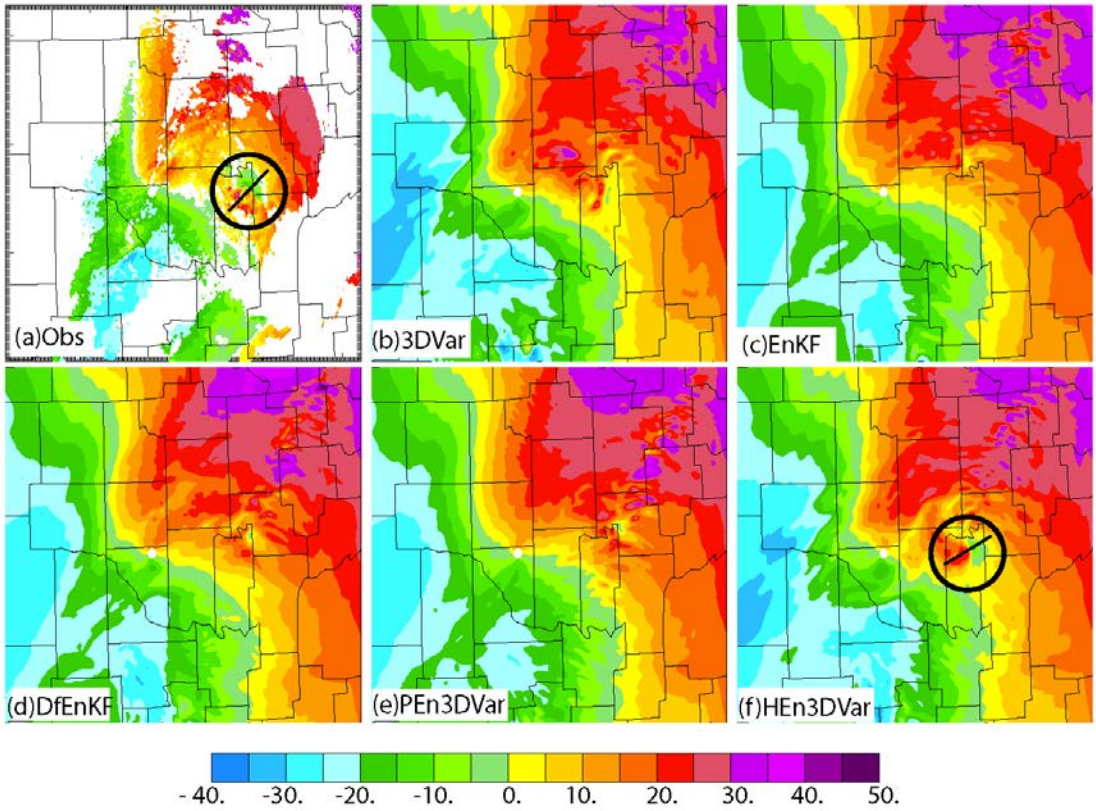

Fig. 5.29 The observed and forecast radial velocity at 23:30 UTC at the lowest elevation angle of KTLX radar from CNTL, 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar with 75% weight given to B. The black circle indicates location of the mesocyclone, and the short black line indicates orientation of the velocity couplet that passes through the maximum inbound and outbound velocity for the circled velocity couplet.

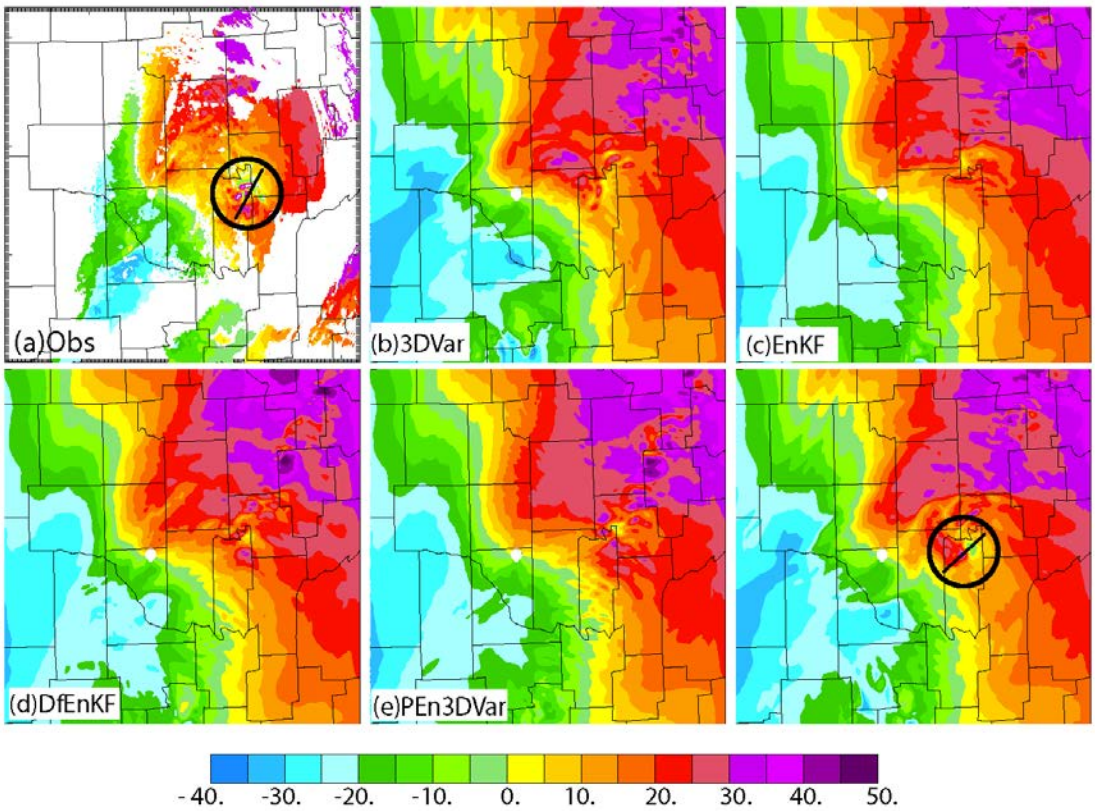

Fig. 5.30 The observed and forecast radial velocity at 23:30 UTC at the elevation angle of 0.92° for KTLX radar from CNTL, 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar with 75% weight given to B. The black circle indicates location of the mesocyclone, and the short black line indicates orientation of the velocity couplet that passes through the maximum inbound and outbound velocity for the circled velocity couplet.

e. Forecasts of the cold pool

Cold pool is storm-scale cold air masses, which are produced by both downdrafts from thunderstorms that bring cool air from aloft down to the surface and from evaporation of rain near the surface and the subsequent latent cooling. Cold pools represent sinking air but can provide lift due to convergence and density differences and is important for storm development. The circulation at the leading edge of the cold pool can pull the cold air outward at the surface and at the same time lift the warm environmental air up over the leading edge of the cold pool.

The characteristic of the cold pool has been found to be a significant factor impacting storm development. The baroclinic vorticity induced by the cold pool is an important source of low-level horizontal vorticity (Straka et al. 2007; Markowski et al. 2008; Markowski et al. 2012; Marquis et al. 2012; Kosiba et al. 2013). Tornadoes are found to be more probable as the environmental low-level horizontal vorticity increases (Rasmussen 2003; Thompson et al. 2003). Most low-level mesocyclones are attribute to baroclinic vorticity rather than the environmental vorticity (Rotunno and Klemp 1985; Markowski et al. 2012). However, tornadogenesis is more likely to be observed in supercells that have intermediate cold pools rather than too strong or too weak cold pools (Markowski 2002; Grzych et al. 2007; Hirth et al. 2008; Markowski and Richardson 2014). When cold pool is too strong, rapid acceleration of low-level air parcel is prohibited, making it difficult to be lifted to level of free convection (LFC). Weak cold pool may lead to suppression of significant near-ground rotation due to weaker surface convergence (Dawson et al. 2013). Cold pool also impacts updraft tilting. Cold pool causes the updraft to lean upshear, and the positive vorticity related to the vertical wind shear causes the updraft to lean downshear. With both cold-pool and shear, the two effects may negate each other and produces an erect updraft (Rotunno et al. 1988). When cold pool is too strong, the mid-level mesocyclone deviate from the low-level circulation centers near the gust front. Intermediate cold pool strength helps to maintain vertical updrafts near and above the low-level circulation centers, providing strong dynamic lifting and vertical stretching to the low-level parcels and favoring tornadogenesis (Snook and Xue 2008).

To check the strength of the heat sink (cold pool) in our study, one hour potential temperature perturbation forecasts overlaid with the reflectivity forecasts with their values higher than 35 dBZ on the surface are shown in Fig. 5.31. The potential temperature perturbations near the surface are positive and are too weak relative to the mesonet observations from all the experiments. As is shown in the figure, there are relative colder air masses around the storm area with respect to the surrounding warm air. The minimum potential temperature perturbation forecasts based on analyses from 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar are 1.59 K, 5.21 K, 4.70 K, 4.32 K, and 3.52 K, respectively. The intensity of the relative cold air mass from 3DVar is the strongest. The cold air mass intensity from pure and hybrid En3DVar are similar and are a little bit weaker than that from 3DVar, and they are all stronger than those from EnKF and DfEnKF. The area of the 35 dBZ contour of the reflectivity forecasts based on analyses from EnKF and DfEnKF are similar, and smaller than those from 3DVar, pure En3DVar, indicating that the stronger the cold pool strength, the stronger the intensity of the storm for this period.

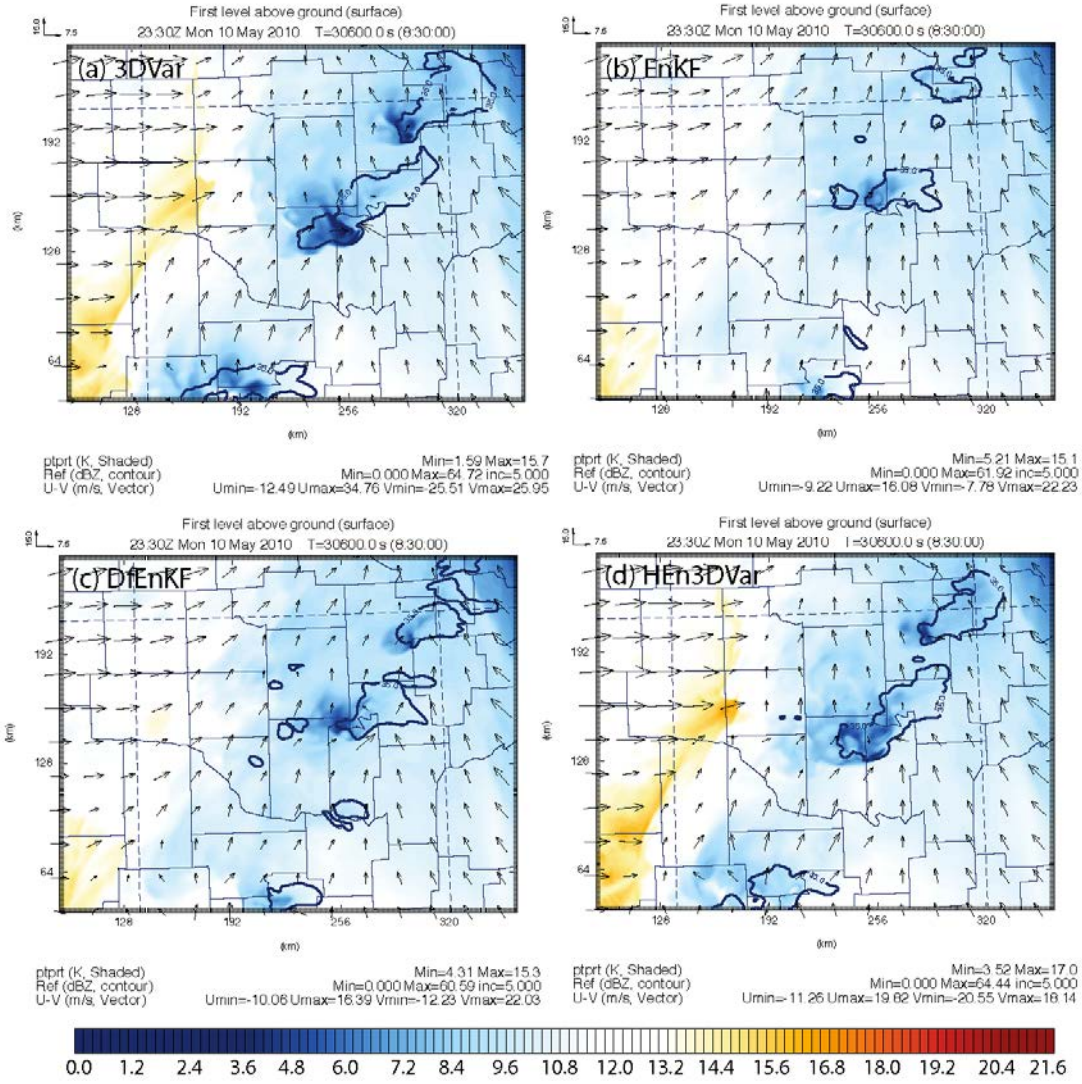

Fig. 5.31 The one-hour surface potential temperature forecasts (°C, shaded area) based on analyses at the end of DA window (22:30 Z) from 3DVar, EnKF, DfEnKF, pure En3DVar, and hybrid En3DVar, overlaid the reflectivity forecasts with their values higher than 35dBZ (contour).

To check the relation between the vertical distribution of the updraft and storm with the cold pool, vertical cross-sections of the one-hour potential temperature perturbation forecasts from different experiments are shown in Fig. 5.32. It shows that there is no strong cold pool generated in the low levels from all the experiments. Without a strong cold pool, the axis of the updraft and the storm from all the experiments all lean in the downshear direction due to the weak cold pool.

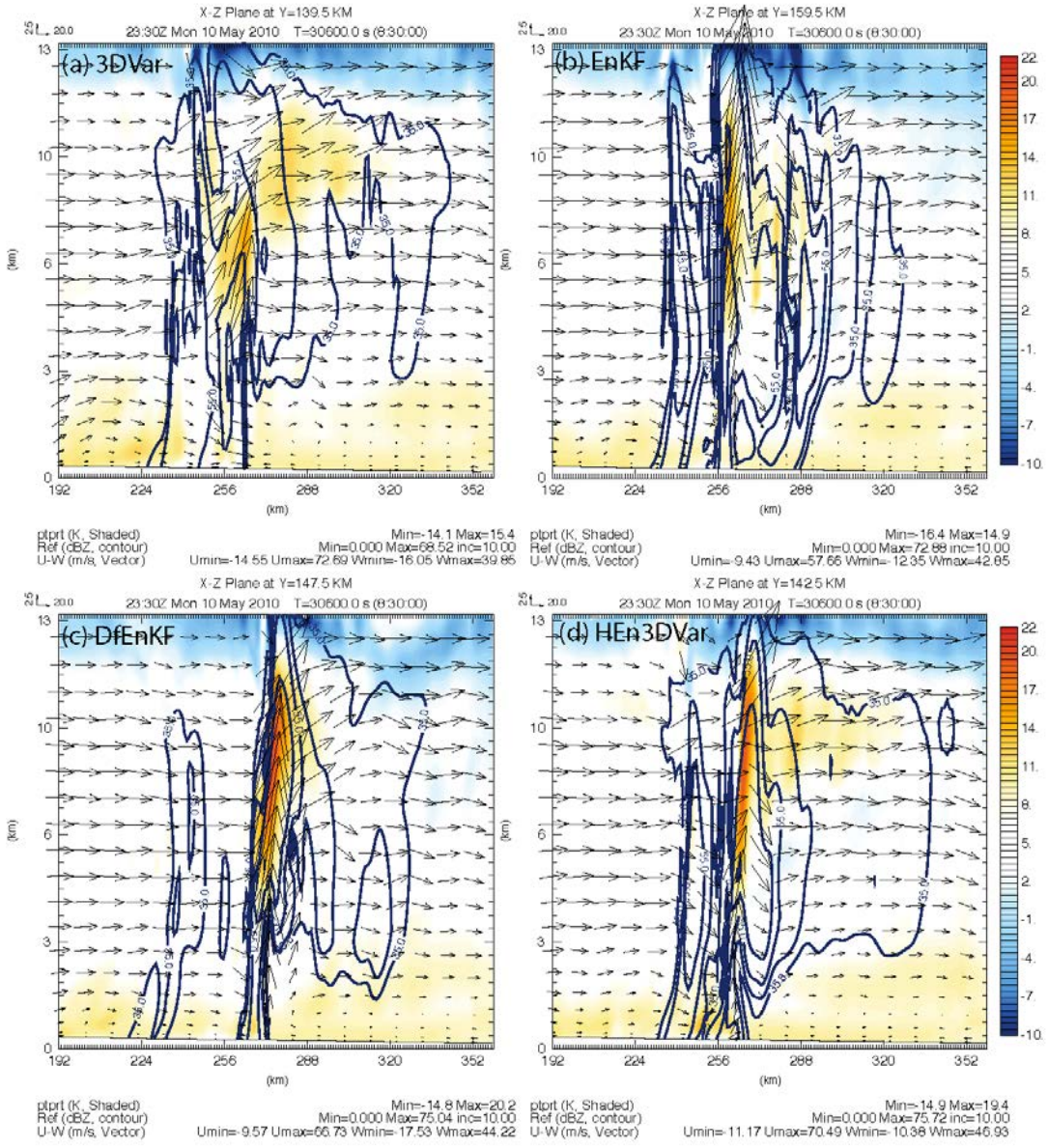

Fig. 5.32 Vertical cross-sections of one-hour potential temperature perturbation forecasts (ºC, shaded area) based on analyses at the end of DA window (22:30 Z) from 3DVar, EnKF, DfEnKF, PEn3DVar, and hybrid En3DVar, overlaid the reflectivity forecasts with their values higher than 35dBZ (contour) with intervals of 20 dBZ.

5.4 Summary and discussion

In this chapter, ARPS hybrid En3DVar system is applied to assimilation of radar data for a real (including tornadic supercell) storm case. Radar radial velocity and reflectivity data were assimilated every 5 minutes for one hour, and an hour following-on forecast was made as well. A deterministic forecast EnKF (called DfEnKF) used in the former two chapters are also applied here to get an algorithm wise parallel comparison between EnKF and pure En3DVar. Logarithm hydrometeor mixing ratios, instead of the hydrometeor mixing ratios, are used as the control variables so that excessive large gradient caused by small background mixing ratios can be avoided in reflectivity DA. Localization of EnKF, DfEnKF, pure and hybrid En3DVar are configured equivalently. 3DVar, EnKF, DfEnKF, pure and hybrid En3DVar are compared based on both the objective verifications and analyses (forecasts) of storm features. The experiments performed and related conclusions are summarized as follows:

- Comparisons are made between experiments that directly use hydrometeor-mixing ratio as the control variables with that use the logarithmic mixing ratios as the control variables. Results show that there are a lot of spurious reflectivity echoes in the analyses that directly uses hydrometer-mixing ratios as the control variables. The gradient of the cost function becomes extremely large when background hydrometeor mixing ratios are small, making the total cost function difficult to converge to the optimal solution. The problem can be avoided when using the logarithmic hydrometeor mixing ratios as the control variables.
- Hybrid En3DVar with four different weights (5%, 25%, 50%, 75%) given to the background error covariance B are compared based on RMSEs as well as

analyses/forecast of storm intensity and structure to obtain a relative optimal weight among them. They are also compared with pure En3DVar and 3DVar that actually have 0% and 100% weight of **B**. Hybrid En3DVar with 50% and 75% weight given to the static **B** performs similar and are much better than those with 5% and 25% in terms of smaller RMSEs for both radial velocity and reflectivity forecasts but much of the RMSE difference is lost during the subsequent forecast. Hybrid En3DVar with 75% weight given to **B** was also found to better capture the hook echo structure of the major tornadic supercell storm in 45 min forecast of the reflectivity forecast, and 45 and 60 min forecasts of the low-level mesocyclone than the other experiments. Thus 75% is chosen to be the optimal weight of static **B** for hybrid En3DVar.

- Hybrid En3DVar with 75% **B** was compared with 3DVar, EnKF, DfEnKF, and pure En3DVar based on both objective verification and analyses/forecasts of storm intensity/structures. Hybrid En3DVar with 75% weight of **B** obviously outperform 3DVar for better capturing the forecasts of the hook echo structure as well as the rotating updraft, and outperforms EnKF and DfEnKF in better capturing the reflectivity forecasts that are less than 20 dBZ. The low-level mesocyclone was better forecast by hybrid En3DVar than those of the others, indicative of stronger rotation and a larger tornado threat. The RMSE of radial velocity forecasts within the DA cycles from EnKF is much smaller than those of the others, indicating averaging the forecasts from different ensemble members could help to decrease the errors in velocity forecasts but the advantage is lost during the subsequent forecast.

In the above study, the advantage of hybrid En3DVar over 3DVar is most obvious for better capturing the storm structure, rotating updraft after including the flow-dependent ensemble covariance at the end of the DA period, though much of that advantage is lost in the subsequent forecast. The advantage of hybrid En3DVar over EnKF and DfEnKF is not as clear except that hybrid En3Dvar better capture the reflectivity forecast below 20 dBZ as well as the strength of the low-level mesoscyclone. In the future, more storm cases are needed to further examine the advantage of hybrid En3DVar over pure 3DVar and EnKF and to see what, if anything can be done to better retain advantages gained in the DA through the subsequent forecast.

Chapter 6 Summary and Future Work

6.1 Summary

Studies have shown benefits of hybrid ensemble-variational data assimilation (DA) over pure ensemble or variational algorithms. Such advantages, if any, for convective-scale radar DA have not been clearly demonstrated, however. A hybrid ensemble-3DVar (En3DVar) system is developed recently based on the ARPS 3DVar and EnKF systems at CAPS. In this dissertation, the impact of combing the flow-dependent background error covariance with the static one was examined for storm-scale radar DA. Hybrid En3DVar is compared with 3DVar, EnKF, and pure En3DVar for radar DA through observing system simulation experiments (OSSEs) under both perfect and imperfect model assumptions. It is also applied to a real case including multiple tornadic supercells. For the real case, radar radial velocity and reflectivity data are assimilated every 5 minutes for 1 hour that is followed by short-term forecasts. DfEnKF that updates a single deterministic background forecast using the EnKF updating algorithm is introduced to have an algorithm-wise parallel comparison between EnKF and pure En3DVar. In this study, we sought to answer the following four questions:

1) In principle, pure En3DVar and EnKF that use 100% ensemble covariance should be identical (under linearity and Gaussian error assumption), will their analyses actually be very close? If not, what are the sources of their difference?
2) Does the relative performance of 3DVar, EnKF, pure/hybrid En3DVar change, if algorithms are tuned to their optimal states?

3) Does the results change after considering model errors? In what situation that the static covariance in the hybrid algorithm does help, if any, for the convective storms?
4) What is relative performance of the algorithms for a real storm case?

The first two questions are answered via perfect-model OSSEs. In the perfect-model OSSEs, DfEnKF and pure En3DVar are compared and are found to perform differently when using the same localization radii. The serial (EnKF) versus global (pure En3DVar) nature of algorithms, and direct filter update (EnKF) versus variational minimization (En3DVar) are the major reasons for the differences. As a variational algorithm, pure En3DVar tends to adjust hail more than snow, because the analysis results are controlled to a large extent by the sensitivity of reflectivity with respect to individual hydrometeor, i.e., the gradient of the reflectivity operator with respect to the hydrometeor state variables. Hybrid En3DVar in radar data assimilation was also compared with 3DVar, EnKF, DfEnKF, and pure En3DVar. Sensitivity experiments were conducted first to obtain the optimal configurations for different algorithms before they are compared, which include the optimal background decorrelation scales for 3DVar, optimal localization radii for EnKF, DfEnKF, and pure En3DVar, as well as the optimal hybrid weight for hybrid En3DVar. When the algorithms are tuned optimally, hybrid En3DVar does not outperform EnKF and pure En3DVar, though their analyses are all much better than 3DVar. When ensemble background error covariance is a good estimation of the true error distribution, pure ensemble-based DA methods can do a good job, and the advantage of static B is not obvious in hybrid DA.

Imperfect-model OSSEs are then used to examine whether the conclusions obtained based on the perfect-model OSSEs change after considering model errors. In the imperfect-model OSSEs, model errors are introduced by using different microphysical schemes in the truth run (Lin scheme) and in the forecasts (WSM6 scheme). Sensitivity experiments are conducted to obtained the optimal configurations for different algorithms, similar to what have been done in the perfect-model OSSEs. Hybrid En3DVar then outperforms EnKF and pure En3DVar (3DVar) for better capturing the hail analyses below the freezing level (intensity) of the storm in the analysis. The advantage of hybrid En3DVar over pure ensemble-based methods is most obvious when ensemble background errors are systematically underestimated.

Finally, ARPS hybrid En3DVar system is applied to the assimilation of radar data for a real tornadic supercell storm. Hybrid En3DVar is compared with 3DVar, EnKF, DfEnKF, and pure En3DVar based on both objective verification and analyses/forecasts of storm intensity/structures. Hybrid En3DVar with 75% weight of B obviously outperforms 3DVar for better capturing the forecasts of the hook echo structure as well as the rotating updraft, and outperforms EnKF and DfEnKF in better capturing the 35dBZ to 45dBZ reflectivity forecasts. The low-level mesocyclone is better forecast by hybrid En3DVar than those of the others, indicating stronger rotation and a larger tornado threat.

In the process to answer the four major questions raised in the first paragraph, some issues arose and were investigated. For example, in the imperfect-model OSSEs, the assimilation of clear-air reflectivity (< 5 dBZ) by the variational DA algorithms was found to seriously degrade the analyses in storm region, with the intensity of the

reflectivity analysis being much weaker than that assimilates precipitation reflectivity (≥5 dBZ) only as well as the truth. When directly using hydrometeor-mixing ratios as the control variables, the gradient of the cost function becomes extremely large when background reflectivity is small. The background reflectivity tends to be small in the area where the observed reflectivity is also small (<5dBZ), especially after one or two cycles when the location errors of convective storms are greatly decreased by reflectivity data assimilation (DA). Thus the assimilation of the clear-air reflectivity involves more grids, on which the background reflectivity are very small, being included in to the calculation of the gradient, making the gradient excessively large and the assimilation of the precipitation reflectivity ineffective. To keep the storm analyses by precipitation reflectivity and at the same time suppress the spurious storms through clear-air reflectivity DA, a double-pass procedure was then proposed and found to be able to greatly alleviate the problem. In the double-pass procedure, the radial velocity and precipitation reflectivity are assimilated together in the first pass, and the clear-air reflectivity is assimilated separately in the second pass. In this way, two cost functions with apparently different orders of magnitude of gradient are minimized separately without influencing the other.

However, there still exist some problems with the double-pass procedure. First, small background reflectivity can still be included into calculation of the gradient in the first pass, and thus influences the assimilation of radial velocity and precipitation reflectivity, especially when significant location errors exist. Second, the background error covariance needs to be updated after assimilating radial velocity and precipitation in the first pass, using the same background error covariance in the second pass as the

first one will introduce errors into the DA process. In addition, the double-pass procedure involves minimizing two different cost functions with different orders of magnitude separately, which is a lot more expensive. So in the real case, an alternative way of using the logarithmic hydrometeor mixing ratios as the control variables are found to be able to solve the problem. By using the logarithmic hydrometeor mixing ratios as the control variables, the excessively large gradient of the cost function is avoided.

In the imperfect-model OSSEs, the impacts of adding the mass continuity constraint are also examined using 3DVar, pure and hybrid En3DVar. Overall, adding the mass continuity constraint benefits the analyses (forecasts) in providing a little bit stronger vertical velocity analyses that are closer to the truth and smoothing the noises present in the velocity and hydrometer fields.

6.2 Future work

Our current studies show that the hybrid En3DVar algorithm has some advantages over stand-alone 3DVar and EnKF, especially when ensemble background errors are systematically underestimated. Given that our findings in this study are primarily based on OSSEs and study of one real case. More research is needed to improve the application of these algorithms to real storm cases. Possible follow-up studies applying ensemble variational DA to storm cases include:

1) Applying the hybrid En3DVar to more real storm cases to better understand the relative performance of hybrid En3DVar for storm scale DA.

2) In our study, the nonlinear reflectivity observation operator is directly used to calculate the cost function. To get its gradient, the nonlinear reflectivity observation operator is linearized around the analysis of the former loop, instead of the initial first guess. Thus the use of the outer-loop is avoided in this study. Past studies indicate that using the linearized observation operator alongside with the outer-loop procedure can decrease the computational cost by increasing the convergence speed of the minimization. However, based on our preliminary tests, using the linearized reflectivity observation operator greatly increases the magnitude of the cost function. The performance of using the tangent linear reflectivity observation operator alongside the outer-loop procedure versus directly using the nonlinear observation operator needs investigation in terms of the computational cost and the performance in future.

3) Past studies indicated that when data timing errors are present with the regular EnSRF, 4DEnSRF worked better than EnSRF (Wang et al., 2012). The four-dimensional DA is expected to improve the analysis and related forecast when more radar volume scans are assimilated. Recently, the ARPS hybrid 4DEnVar has been developed at Center for Analysis and Prediction of Storms (CAPS). We will test the hybrid 4DEnVar algorithms (Liu and Xue 2016) based on the ARPS hybrid 4DEnVar system using OSSEs and real case studies to understand the potential benefits of hybrid En4DVar over En3DVar and 4DEnKF for storm-scale radar DA.

References

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part I: Storm-Scale Analyses. *Mon Weather Rev*, **137,** 1805-1824.

——, 2010: A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part Ii: Short-Range Ensemble Forecasts. *Mon Weather Rev*, **138,** 1273-1292.

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon Weather Rev*, **129,** 2884-2903.

Baker, L. H., A. C. Rudd, S. Migliorini, and R. N. Bannister, 2014: Representation of Model Error in a Convective-Scale Ensemble Prediction System. *Nonlinear Proc Geoph*, **21,** 19-39.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational Convective-Scale Numerical Weather Prediction with the Cosmo Model: Description and Sensitivities. *Mon Weather Rev*, **139,** 3887-3905.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The Quiet Revolution of Numerical Weather Prediction. *Nature*, **525,** 47-55.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon Weather Rev*, **129,** 420-436.

Bloom, S. C., L. L. Takacs, A. M. DaSilva, and D. Ledvina, 1996: Data Assimilation Using Incremental Analysis Updates. *Mon Weather Rev*, **124,** 1256-1271.

Brewster, K.A., D.R. Stratman, K.W. Thomas and F.H. Carr, 2016: Incremental analysis updating with variable-depending timing applied to a real-time high resolution forecast system. *28th Conf. on Severe Local Storms*, Portland, OR, Amer. Meteor. Soc., Paper 69.

Buehner, M., 2005: Ensemble-Derived Stationary and Flow-Dependent Background-Error Covariances: Evaluation in a Quasi-Operational Nwp Setting. *Q J Roy Meteor Soc*, **131,** 1013-1043.

Buehner, M., J. Morneau, and C. Charette, 2013: Four-Dimensional Ensemble-Variational Data Assimilation for Global Deterministic Weather Prediction. *Nonlinear Proc Geoph*, **20,** 669-682.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic Nwp. Part I: Description and Single-Observation Experiments. *Mon Weather Rev*, **138,** 1550-1566.

——, 2010b: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic Nwp. Part Ii: One-Month Experiments with Real Observations. *Mon Weather Rev*, **138,** 1567-1586.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational Implementation of a Hybrid Ensemble/4d-Var Global Data Assimilation System at the Met Office. *Q J Roy Meteor Soc*, **139,** 1445-1461.

Courtier, P., 1997: Variational Methods. *J Meteorol Soc Jpn*, **75,** 211-218.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A New Dynamical Core for the Met Office's Global and Regional Modelling of the Atmosphere. *Q J Roy Meteor Soc*, **131,** 1759-1782.

Dawson, D. T., and M. Xue, 2006: Numerical Forecasts of the 15-16 June 2002 Southern Plains Mesoscale Convective System: Impact of Mesoscale Data and Cloud Analysis. *Mon Weather Rev*, **134,** 1607-1629.

Dawson, D. T., L. J. Wicker, E. R. Mansell, Y. S. Jung, and M. Xue, 2013: Low-Level Polarimetric Radar Signatures in Enkf Analyses and Forecasts of the May 8, 2003 Oklahoma City Tornadic Supercell: Impact of Multimoment Microphysics and Comparisons with Observation. *Adv Meteorol*.

Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman Filter Assimilation of Radar Observations of the 8 May 2003 Oklahoma City Supercell: Influences of Reflectivity Observations on Storm-Scale Analyses. *Mon Weather Rev*, **139,** 272-294.

Dowell, D. C., F. Q. 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 Weather Rev*, **132,** 1982-2005.

Du, N. Z., M. Xue, K. Zhao, and J. Z. Min, 2012: Impact of Assimilating Airborne Doppler Radar Velocity Data Using the Arps 3dvar on the Analysis and Prediction of Hurricane Ike (2008). *J Geophys Res-Atmos*, **117**.

Ebert, E. E., 2008: Fuzzy Verification of High-Resolution Gridded Forecasts: A Review and Proposed Framework. *Meteorol Appl*, **15,** 51-64.

Etherton, B. J., and C. H. Bishop, 2004: Resilience of Hybrid Ensemble/3dvar Analysis Schemes to Model Error and Ensemble Covariance Error. *Mon Weather Rev*, **132,** 1065-1080.

Evensen, G., 1992: Using the Extended Kalman Filter with a Multilayer Quasi-Geostrophic Ocean Model. *J Geophys Res-Atmos*, **97,** 17905-17924.

——, 1994: Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte-Carlo Methods to Forecast Error Statistics. *J Geophys Res-Oceans*, **99,** 10143-10162.

Gao, J., M. Xue, K. Brewster, and K. Droegemeier, 2004: Three-Dimensional Variational Data Analysis Method with Recursive Filter for Doppler Radars. *J Atmos Ocean Tech*, **21,** 457-469.

Gao, J., C. Fu, D. J. Stensrud, and J. S. Kain, 2016: Osses for an Ensemble 3dvar Data Assimilation System with Radar Observations of Convective Storms. *J Atmos Sci*, **73,** 2403-2426.

Gao, J. D., and D. J. Stensrud, 2012: Assimilation of Reflectivity Data in a Convective-Scale, Cycled 3dvar Framework with Hydrometeor Classification. *J Atmos Sci*, **69,** 1054-1065.

Gao, J. D., M. Xue, and D. J. Stensrud, 2013: The Development of a Hybrid Enkf-3dvar Algorithm for Storm-Scale Data Assimilation. *Adv Meteorol*.

Gao, J. D., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A Variational Method for the Analysis of Three-Dimensional Wind Fields from Two Doppler Radars. *Mon Weather Rev*, **127,** 2128-2142.

Gaspari, G., and S. E. Cohn, 1999: Construction of Correlation Functions in Two and Three Dimensions. *Q J Roy Meteor Soc*, **125,** 723-757.

Gasperoni, N. A., M. Xue, R. D. Palmer, and J. D. Gao, 2013: Sensitivity of Convective Initiation Prediction to near-Surface Moisture When Assimilating Radar Refractivity: Impact Tests Using Osses. *J Atmos Ocean Tech*, **30,** 2281-2302.

Ge, G. Q., J. D. Gao, and M. Xue, 2012: Diagnostic Pressure Equation as a Weak Constraint in a Storm-Scale Three-Dimensional Variational Radar Data Assimilation System. *J Atmos Ocean Tech*, **29,** 1075-1092.

——, 2013: Impact of a Diagnostic Pressure Equation Constraint on Tornadic Supercell Thunderstorm Forecasts Initialized Using 3dvar Radar Data Assimilation. *Adv Meteorol*.

Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2011: Balance and Ensemble Kalman Filter Localization Techniques. *Mon Weather Rev*, **139,** 511-522.

Grzych, M. L., B. D. Lee, and C. A. Finley, 2007: Thermodynamic Analysis of Supercell Rear-Flank Downdrafts from Project Answers. *Mon Weather Rev*, **135,** 240-246.

Hamill, T. M., and C. Snyder, 2000: A Hybrid Ensemble Kalman Filter-3d Variational Analysis Scheme. *Mon Weather Rev*, **128,** 2905-2919.

Hamill, T. M., C. Snyder, and R. E. Morss, 2000: A Comparison of Probabilistic Forecasts from Bred, Singular-Vector, and Perturbed Observation Ensembles. *Mon Weather Rev*, **128,** 1835-1851.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Mon Weather Rev*, **129,** 2776-2790.

Hanley, K. E., A. I. Barrett, and H. W. Lean, 2016: Simulating the 20 May 2013 Moore, Oklahoma Tornado with a 100-Metre Grid-Length Nwp Model. *Atmos Sci Lett*, **17,** 453-461.

Hirahara, Y., J. Ishida, and T. Ishimizu, 2011: Trial Operation of the Local Forecast Model at Jma. *CAS/JSC WGNE Res. Act. Atmos. Ocea. Model,*, **41,** 11-12.

Hirth, B. D., J. L. Schroeder, and C. C. Weiss, 2008: Surface Analysis of the Rear-Flank Downdraft Outflow in Two Tornadic Supercells. *Mon Weather Rev*, **136,** 2344-2363.

Hong, S., and J. J. Lim, 2006: The Wrf Single-Moment 6-Class Microphysics Scheme (Wsm6). *Journal of the Korean Meteorological Society*, **42,** 129-151.

Houtekamer, P. L., and L. Lefaivre, 1997: Using Ensemble Forecasts for Model Validation. *Mon Weather Rev*, **125,** 2416-2426.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Mon Weather Rev*, **126,** 796-811.

——, 2005: Ensemble Kalman Filtering. *Q J Roy Meteor Soc*, **131,** 3269-3289.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A System Simulation Approach to Ensemble Prediction. *Mon Weather Rev*, **124,** 1225-1242.

Hu, M., M. Xue, and K. Brewster, 2006a: 3dvar and Cloud Analysis with Wsr-88d Level-Ii Data for the Prediction of the Fort Worth, Texas, Tornadic Thunderstorms. Part I: Cloud Analysis and Its Impact. *Mon Weather Rev*, **134,** 675-698.

Hu, M., M. Xue, J. D. Gao, and K. Brewster, 2006b: 3dvar and Cloud Analysis with Wsr-88d Level-Ii Data for the Prediction of the Fort Worth, Texas, Tornadic Thunderstorms. Part Ii: Impact of Radial Velocity Analysis Via 3dvar. *Mon Weather Rev*, **134,** 699-721.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient Data Assimilation for Spatiotemporal Chaos: A Local Ensemble Transform Kalman Filter. *Physica D-Nonlinear Phenomena*, **230,** 112-126.

Jung, Y. S., M. Xue, and M. J. Tong, 2012: Ensemble Kalman Filter Analyses of the 29-30 May 2004 Oklahoma Tornadic Thunderstorm Using One- and Two-Moment Bulk Microphysics Schemes, with Verification against Polarimetric Radar Data. *Mon Weather Rev*, **140,** 1457-1475.

Jung, Y. S., M. Xue, G. F. Zhang, and J. M. Straka, 2008: Assimilation of Simulated Polarimetric Radar Data for a Convective Storm Using the Ensemble Kalman Filter. Part Ii: Impact of Polarimetric Data on Storm Analysis. *Mon Weather Rev*, **136,** 2246-2260.

Kain, J. S., and Coauthors, 2010: Assessing Advances in the Assimilation of Radar Data and Other Mesoscale Observations within a Collaborative Forecasting-Research Environment. *Weather Forecast*, **25,** 1510-1521.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation, and Predictability.* Cambridge University Press, 341 pp.

Kepert, J. D., 2009: Covariance Localisation and Balance in an Ensemble Kalman Filter. *Q J Roy Meteor Soc*, **135,** 1157-1176.

Kosiba, K., J. Wurman, Y. Richardson, P. Markowski, P. Robinson, and J. Marquis, 2013: Genesis of the Goshen County, Wyoming, Tornado on 5 June 2009 During Vortex2. *Mon Weather Rev*, **141,** 1157-1181.

Le Dimet, F. X., and O. Talagrand, 1986: Variational Algorithms for Analysis and Assimilation of Meteorological Observations: Theoretical Aspects. *Tellus*, **38A,** 97-110.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of High-Resolution Versions of the Met Office Unified Model for Forecasting Convection over the United Kingdom. *Mon Weather Rev*, **136,** 3408-3424.

Lee, L. A., K. S. Carslaw, K. J. Pringle, and G. W. Mann, 2012: Mapping the Uncertainty in Global Ccn Using Emulation. *Atmos Chem Phys*, **12,** 9739-9751.

Li, Y. Z., X. H. Wang, and M. Xue, 2012: Assimilation of Radar Radial Velocity Data with the Wrf Hybrid Ensemble-3dvar System for the Prediction of Hurricane Ike (2008). *Mon Weather Rev*, **140,** 3507-3524.

Lilly, D. K., 1990: Numerical Prediction of Thunderstorms - Has Its Time Come. *Q J Roy Meteor Soc*, **116,** 779-798.

Lin, Y. L., R. D. Farley, and H. D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model. *J Clim Appl Meteorol*, **22,** 1065-1092.

Lindskog, M., K. Salonen, H. Jarvinen, and D. B. Michelson, 2004: Doppler Radar Wind Data Assimilation with Hirlam 3dvar. *Mon Weather Rev*, **132,** 1081-1092.

Liu, C. S., and M. Xue, 2016: Relationships among Four-Dimensional Hybrid Ensemble-Variational Data Assimilation Algorithms with Full and Approximate Ensemble Covariance Localization. *Mon Weather Rev*, **144,** 591-606.

Lorenc, A. C., 2003: The Potential of the Ensemble Kalman Filter for Nwp - a Comparison with 4d-Var. *Q J Roy Meteor Soc*, **129,** 3183-3203.

Markowski, P., E. Rasmussen, J. Straka, R. Davies-Jones, Y. Richardson, and R. J. Trapp, 2008: Vortex Lines within Low-Level Mesocyclones Obtained from Pseudo-Dual-Doppler Radar Observations. *Mon Weather Rev*, **136,** 3513-3535.

Markowski, P., and Coauthors, 2012: The Pretornadic Phase of the Goshen County, Wyoming, Supercell of 5 June 2009 Intercepted by Vortex2. Part Ii: Intensification of Low-Level Rotation. *Mon Weather Rev*, **140,** 2916-2938.

Markowski, P. M., 2002: Hook Echoes and Rear-Flank Downdrafts: A Review. *Mon Weather Rev*, **130,** 852-876.

Markowski, P. M., and Y. P. Richardson, 2014: The Influence of Environmental Low-Level Shear and Cold Pools on Tornadogenesis: Insights from Idealized Simulations. *J Atmos Sci*, **71,** 243-275.

Marquis, J., Y. Richardson, P. Markowski, J. Wurman, and D. Dowell, 2012: The Maintenance of Tornadoes Observed with High-Resolution Mobile Radars. *Mon Weather Rev***,** 3-27.

Miller, R. N., M. Chil, and F. Gauthiez, 1994: Advanced Data Assimilation in Strongly Nonlinear Dynamic Systems. *J Atmos Sci*, **51,** 1037-1056.

Miyoshi, T., and S. Yamane, 2007: Local Ensemble Transform Kalman Filtering with an Agcm at a T159/L48 Resolution. *Mon Weather Rev*, **135,** 3841-3861.

Pan, Y. J., and Coauthors, 2014: A Gsi-Based Coupled Ensrf-En3dvar Hybrid Data Assimilation System for the Operational Rapid Refresh Model: Tests at a Reduced Resolution. *Mon Weather Rev*, **142,** 3756-3780.

Rasmussen, E. N., 2003: Refined Supercell and Tornado Forecast Parameters. *Weather Forecast*, **18,** 530-535.

Ray, P. S., and Coauthors, 1981: The Morphology of Several Tornadic Storms on 20 May 1977. *J Atmos Sci*, **38,** 1643-1663.

Roberts, N. M., and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon Weather Rev*, **136,** 78-97.

Rotunno, R., and J. Klemp, 1985: On the Rotation and Propagation of Simulated Supercell Thunderstorms. *J Atmos Sci*, **42,** 271-292.

Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A Theory for Strong, Long-Lived Squall Lines. *J Atmos Sci*, **45,** 463-485.

Sakov, P., and P. R. Oke, 2008: A Deterministic Formulation of the Ensemble Kalman Filter: An Alternative to Ensemble Square Root Filters. *Tellus A*, **60,** 361-371.

Schenkman, A. D., M. Xue, A. Shapiro, K. Brewster, and J. D. Gao, 2011: The Analysis and Prediction of the 8-9 May 2007 Oklahoma Tornadic Mesoscale Convective System by Assimilating Wsr-88d and Casa Radar Data Using 3dvar. *Mon Weather Rev*, **139,** 224-246.

Seity, Y., and Coauthors, 2011: The Arome-France Convective-Scale Operational Model. *Mon Weather Rev*, **139,** 976-991.

Snook, N., and M. Xue, 2008: Effects of Microphysical Drop Size Distribution on Tornadogenesis in Supercell Thunderstorms. *Geophys Res Lett*, **35**.

Snook, N., M. Xue, and Y. S. Jung, 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.

——, 2012: Ensemble Probabilistic Forecasts of a Tornadic Mesoscale Convective System from Ensemble Kalman Filter Analyses Using WSR-88d and Casa Radar Data. *Mon Weather Rev*, **140,** 2126-2146.

——, 2015: Multiscale Enkf Assimilation of Radar and Conventional Observations and Ensemble Forecasting for a Tornadic Mesoscale Convective System. *Mon Weather Rev*, **143,** 1035-1057.

Snyder, C., and F. Q. Zhang, 2003: Assimilation of Simulated Doppler Radar Observations with an Ensemble Kalman Filter. *Mon Weather Rev*, **131,** 1663-1677.

Sobash, R. A., and D. J. Stensrud, 2013: The Impact of Covariance Localization for Radar Data on Enkf Analyses of a Developing Mcs: Observing System Simulation Experiments. *Mon Weather Rev*, **141,** 3691-3709.

Stainforth, D. A., and Coauthors, 2005: Uncertainty in Predictions of the Climate Response to Rising Levels of Greenhouse Gases. *Nature*, **433,** 403-406.

Stensrud, D. J., J. W. Bao, and T. T. Warner, 2000: Using Initial Condition and Model Physics Perturbations in Short-Range Ensemble Simulations of Mesoscale Convective Systems. *Mon Weather Rev*, **128,** 2077-2107.

Stensrud, D. J., and Coauthors, 2013: Progress and Challenges with Warn-on-Forecast. *Atmos Res*, **123,** 2-16.

Straka, J. M., E. Rasmussen, R. Davies-Jones, and P. Markowski, 2007: An Observational and Idealized Numerical Examination of Low-Level Counter-Rotating Vortices in the Rear Flank of Supercells. *Electronic J. Severe Storms Meteor.*, **2,** 1-22.

Stratman, D. R., and K. A. Brewster, 2017: Sensitivities of 1-Km Forecasts of 24 May 2011 Tornadic Supercells to Microphysics Parameterizations. *Mon Weather Rev*, **145,** 2697-2721.

Sun, J. Z., and A. Crook, 1994: Wind and Thermodynamic Retrieval from Single-Doppler Measurements of a Gust Front Observed During Phoenix-Ii. *Mon Weather Rev*, **122,** 1075-1091.

Sun, J. Z., and N. A. Crook, 1997: Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint .1. Model Development and Simulated Data Experiments. *J Atmos Sci*, **54,** 1642-1661.

——, 1998: Dynamical and Microphysical Retrieval from Doppler Radar Observations Using a Cloud Model and Its Adjoint. Part Ii: Retrieval Experiments of an Observed Florida Convective Storm. *J Atmos Sci*, **55,** 835-852.

Sun, J. Z., D. W. Flicker, and D. K. Lilly, 1991: Recovery of 3-Dimensional Wind and Temperature-Fields from Simulated Single-Doppler Radar Data. *J Atmos Sci*, **48,** 876-890.

Sun, J. Z., and Coauthors, 2014: Use of Nwp for Nowcasting Convective Precipitation. *B Am Meteorol Soc*, **95,** 409-426.

Supinie, T. A., Y. Jung, M. Xue, D. J. Stensrud, M. M. French, and H. B. Bluestein, 2016: Impact of Vortex2 Observations on Analyses and Forecasts of the 5 June 2009 Goshen County, Wyoming, Supercell. *Mon Weather Rev*, **144,** 429-449.

Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close Proximity Soundings within Supercell Environments Obtained from the Rapid Update Cycle. *Weather Forecast*, **18,** 1243-1261.

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

Tong, M. J., and M. Xue, 2005: Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: Oss Experiments. *Mon Weather Rev*, **133,** 1789-1807.

——, 2008: Simultaneous Estimation of Microphysical Parameters and Atmospheric State with Simulated Radar Data and Ensemble Square Root Kalman Filter. Part I: Sensitivity Analysis and Parameter Identifiability. *Mon Weather Rev*, **136,** 1630-1648.

Toth, Z., and E. Kalnay, 1997: Ensemble Forecasting at Ncep and the Breeding Method. *Mon Weather Rev*, **125,** 3297-3319.

Verrelle, A., D. Ricard, and C. Lac, 2015: Sensitivity of High-Resolution Idealized Simulations of Thunderstorms to Horizontal Resolution and Turbulence Parametrization. *Q J Roy Meteor Soc*, **141,** 433-448.

Wang, H. L., J. Z. Sun, S. Y. Fan, and X. Y. Huang, 2013: Indirect Assimilation of Radar Reflectivity with Wrf 3d-Var and Its Impact on Prediction of Four Summertime Convective Events. *J Appl Meteorol Clim*, **52,** 889-902.

Wang, X. G., C. Snyder, and T. M. Hamill, 2007a: On the Theoretical Equivalence of Differently Proposed Ensemhle-3dvar Hybrid Analysis Schemes. *Mon Weather Rev*, **135,** 222-227.

Wang, X. G., T. A. Hamill, J. S. Whitaker, and C. H. Bishop, 2007b: A Comparison of Hybrid Ensemble Transform Kalman Filter-Optimum Interpolation and Ensemble Square Root Filter Analysis Schemes. *Mon Weather Rev*, **135,** 1055-1076.

Warren, R. A., D. J. Kirshbaum, R. S. Plant, and H. W. Lean, 2014: A 'Boscastle-Type' Quasi-Stationary Convective System over the Uk Southwest Peninsula. *Q J Roy Meteor Soc*, **140,** 240-257.

Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-H Explicit Convective Forecasts with the Wrf-Arw Model. *Weather Forecast*, **23,** 407-437.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Mon Weather Rev*, **130,** 1913-1924.

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

Xiao, Q. N., Y. H. Kuo, J. Z. Sun, W. C. Lee, D. M. Barker, and E. Lim, 2007: An Approach of Radar Reflectivity Data Assimilation and Its Assessment with the Inland Qpf of Typhoon Rusa (2002) at Landfall. *J Appl Meteorol Clim*, **46,** 14-22.

Xue, M., and W. J. Martin, 2006: A High-Resolution Modeling Study of the 24 May 2002 Dryline Case During Ihop. Part I: Numerical Simulation and General Evolution of the Dryline and Convection. *Mon Weather Rev*, **134,** 149-171.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (Arps) - a Multi-Scale Nonhydrostatic Atmospheric Simulation and Prediction Model. Part I: Model Dynamics and Verification. *Meteorol Atmos Phys*, **75,** 161-193.

Xue, M., M. J. Tong, and K. K. Droegemeier, 2006: An Osse Framework Based on the Ensemble Square Root Kalman Filter for Evaluating the Impact of Data from Radar Networks on Thunderstorm Analysis and Forecasting. *J Atmos Ocean Tech*, **23,** 46-66.

Xue, M., M. Hu, and A. D. Schenkman, 2014: Numerical Prediction of the 8 May 2003 Oklahoma City Tornadic Supercell and Embedded Tornado Using Arps with the Assimilation of Wsr-88d Data. *Weather Forecast*, **29,** 39-62.

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. *Meteorol Atmos Phys*, **82,** 139-170.

Xue, M., and Coauthors, 2001: The Advanced Regional Prediction System (Arps) - a Multi-Scale Nonhydrostatic Atmospheric Simulation and Prediction Tool. Part Ii: Model Physics and Applications. *Meteorol Atmos Phys*, **76,** 143-165.

Yussouf, N., J. D. Gao, D. J. Stensrud, and G. Q. Ge, 2013: The Impact of Mesoscale Environmental Uncertainty on the Prediction of a Tornadic Supercell Storm Using Ensemble Data Assimilation Approach. *Adv Meteorol*.

Zhang, F. Q., M. Zhang, and J. Poterjoy, 2013: E3dvar: Coupling an Ensemble Kalman Filter with Three-Dimensional Variational Data Assimilation in a Limited-Area Weather Prediction Model and Comparison to E4dvar. *Mon Weather Rev*, **141,** 900-917.

Zhang, M., F. Q. Zhang, X. Y. Huang, and X. Zhang, 2011: Intercomparison of an Ensemble Kalman Filter with Three- and Four-Dimensional Variational Data Assimilation Methods in a Limited-Area Model over the Month of June 2003. *Mon Weather Rev*, **139,** 566-572.

Zhao, K., and M. Xue, 2009: Assimilation of Coastal Doppler Radar Data with the Arps 3dvar and Cloud Analysis for the Prediction of Hurricane Ike (2008). *Geophys Res Lett*, **36**.

Zhu, K. F., and Coauthors, 2013: A Regional Gsi-Based Ensemble Kalman Filter Data Assimilation System for the Rapid Refresh Configuration: Testing at Reduced Resolution. *Mon Weather Rev*, **141,** 4118-4139.