UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

APPLICATIONS OF ENSEMBLE KALMAN FILTER DATA ASSIMILATION: FROM CONVECTIVE THUNDERSTORMS TO HURRICANES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

JILI DONG
Norman, Oklahoma
2010

APPLICATIONS OF ENSEMBLE KALMAN FILTER DATA ASSIMILATION: FROM CONVECTIVE THUNDERSTORMS TO HURRICANES

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEROLOGY

BY

Dr. Ming Xue, Chair

Dr. Frederick Carr

Dr. Kelvin Droegemeier

Dr. Lance Leslie

Dr. Sivaramakrishnan Lakshmivarahan

ACKNOWLEDGMENTS

First and foremost I wish to thank my advisor, Dr. Ming Xue, director of Center for Analysis and Prediction of Storms (CAPS). I would like to sincerely thank Dr. Xue, for his guidance, teaching, discussion and support during my study and research at the University of Oklahoma (OU). Dr. Xue not only gave me this opportunity to work at CAPS, but also trained me for many skills at all levels. He gave me many good and important suggestions on my EnKF data assimilation research.

I would also like to gratefully acknowledge all the other members of my dissertation committee: Dr. Kelvin Droegemeier, Dr. Frederick Carr, Dr. Lance Leslie, and Dr. Sivaramakrishnan Lakshmivarahan, who have spent their time to read my dissertation and offered valuable advices. Without their important comments and suggestions on my research, this dissertation would not be completed.

Many thanks also go to Dr. Mingjing Tong for her suggestions and discussions on the ensemble Kalman filre (EnKF) system developed by her. Thanks to Dr. Tong's great works on the EnKF data assimilation system. I am grateful to Dr. Ming Hu for his help on Linux and FORTRAN programming. I thank Dr. Yunheng Wang for helping me on many Advanced Regional Prediction System (ARPS) issues patiently.

I would like to thank Dr. Youngsun Jung for her many suggestions and discussions on my study and research. Thanks to my office mate Guoqing Ge and Nate Snook for their friendship and help on my research. My research is also benefited from discussions with all members of the research group led by Dr. Xue.

Thanks also extends to staffs at school of meteorology and CAPS, particularly, Celia Jones, Marcia Pallutto, Eileen Hasselwander and Scott Hill for their kind help.

Outside the academic community, I shared with my family members. I want to thank my parents for their endless love, encouragement and support. Many thanks go to my fiancée, Fan Yang, for her support and love. She always encourages me on my research and careers with her enthusiasm. Without their constant support and encouragement, the research would not have been possible.

This research was supported by NSF Grants ATM-0331594, AGS-080288, a DOC-NOAA Grant NA17RJ1227 and a US Department of Defense Grants EPSCoR N00014-10-1-0133. The experiments were performed on the supercomputers of OSCER, University of Oklahoma, supercomputers at Pittsburg Supercomputing Center, supercomputers at Texas Advanced Computing Center and supercomputers at National Institute of Computational Science, University of Tennessee.

Table of Contents

Chapter 4 Assimilation of Radial Velocity and Reflectivity Data from Coastal WSR-88D Radars Using EnKF for Hurricane Ike (2008) ……98

List of Tables

List of Figures

Abstract

For the numerical prediction of severe thunderstorm and hurricane, data assimilation is one of the necessary tools to obtain accurate initial conditions. Ensemble Kalman filter (EnKF) is a state of the art data assimilation algorithm, with the advantage of using flow-dependent error covariance information and retrieving unobserved model quantities. In this dissertation, EnKF is first employed to assimilate additional surface observation in the presence of radar data for severe thunderstorm analysis and prediction with Observing System Simulation Experiments (OSSEs). The EnKF is then used to assimilate real coastal WSR-88D radar observations for Hurricane Ike (2008), and the impact of radar data on the analysis and forecast is investigated.

Due to the earth curvature effect and the non-zero elevation of the lowest scan of ground-based radars, low-level coverage of radar data decreases as distance from the radar increases, causing loss of coverage for important low-level features including the cold pool and gust front. Observations from surface networks are expected to help fill such low-level data gaps. To investigate the impact of additional surface observations on the analysis and forecast of convective storms, a series of OSSEs are performed using the ARPS model and its EnKF data assimilation system.

When the radar is located at a significant distance (e.g., the 115 and 185 km distances considered) from the main convective storm, a clear positive impact on the storm analysis and forecast is achieved by assimilating surface observations with a spacing of about 20 km. When the radar is located just 45 km from the storm center, a network spacing of 6 km is needed to achieve any noticeable positive impact. The impact of surface data in terms of relative error reduction increases linearly with decreased surface network spacing until the spacing is close to the grid interval of truth simulation. Assimilating observations from a coarser network over a longer period of time helps to achieve a similar level of impact as would be seen from a network of higher density.

The error correlation fields derived from the forecast ensemble exhibit dynamically consistent structures. Through flow-dependent error covariance and dynamical interactions in the prediction model, the surface observations not only correct the surface errors, but also improve analyses of state variables at the mid- and upper levels. Given typical observation error, surface wind observations produce the largest positive impact, followed by temperature measurements. Pressure measurements produce the least impact. Assimilating all surface observation variables together yields the largest impact.

The impact of surface data is sustained or even amplified during subsequent forecasts when their impact on the analysis is significant.

In the second part of this dissertation, EnKF assimilation and forecasting experiments are performed for the case of Hurricane Ike (2008), the third most destructive hurricane hitting the United States. Data from two coastal WSR-88D radars were carefully quality controlled, including automatic and manual velocity dealiasing. For the control experiment, 32 ensemble members are used in the EnKF system, and reflectivity (Z) and radial velocity (Vr) data from the two coastal radars are assimilated at 10-minute intervals over a 2-hour period shortly before Ike made landfall.

Compared to the corresponding NCEP GFS analysis, the assimilation resulted in a much improved vortex intensity at the final analysis time, although it is still weaker than observed. Compared to the forecast starting from GFS analysis at the same initial time, the forecast intensity, track and structure of Ike over a 12 hour period are improved in both deterministic and ensemble mean forecasts. The ensemble spread is well maintained with the help of multiplicative covariance inflation and posterior additive perturbations during the assimilation cycles. Assimilation of either Vr or Z alone leads to improvement in hurricane intensity, track and quantitative precipitation forecast. Vr leads to more improvement in intensity and track forecast, emphasizing more importance of Vr data.

Ensemble forecast has shown uncertainty growth in track forecast but not in intensity forecast. 30-minute assimilation interval has the similar results with 10-minute assimilation interval and 60-minute assimilation interval shows weaker intensity forecast.

Assimilation of additional minimum mean sea level pressure from best track data together with Z leads to further improvement in intensity and track forecast compared to assimilating Z alone. Assimilating minimum sea level pressure in addition to Vr leads to track forecast improvement but only small improvement in intensity analysis and forecast.

Chapter 1

Introduction and Overview

1.1 Background and motivation

1.1.1 Background

Numerical Weather Prediction (NWP) demands accurate initial conditions to produce weather forecasts with complex numerical weather models. Owing to the chaotic nature of the atmosphere, small errors in the initial conditions will amplify with time and lead to huge differences between the prediction and the truth state (Lorenz 1963). To reduce the uncertainty on initial conditions, various data assimilation methods are designed to combine the observations and background model states into an optimal initial state.

Most of the modern data assimilation methods provide a maximum likelihood estimation with a least square approach. Among these data assimilation methods, the optimal interpolation (OI) and the three-dimensional variational (3DVAR) methods have an assumption of constant background error statistics. This flow-independent assumption, however, is not realistic when the evolution of weather patterns is considered. The change of the background error statistics may be dramatic with time, like in convective-scale thunderstorms. Ensemble Kalman filter (EnKF) and the four-dimensional variational methods (4DVAR) both include flow-dependent background error statistics. Although more computationally expensive than OI and 3DVAR, both 4DVAR and EnKF still show advantage in theory for the inclusion of the forecast error variability.

In recent years, 4DVAR has been implemented at several operational centers, including European Centre for Medium-Range Weather Forecasts (ECMWF), MeteoFrance, and Japan Meteorological Agency (JMA). However, 4DVAR requires much effort in developing and maintaining a complicated adjoint model. The cost of developing this adjoint code is more expensive for convective scales than for large scales since more details in the physical processes have to be included in the code. Also, 4DVAR is still a deterministic method (Lewis et al, 2006). 4DVAR does not update the background error covariance for the next assimilation cycle, providing no uncertainty information about the forecast.

Since it was first proposed by Evensen (1994), EnKF has gained popularity among the research community. Different from 4DVAR, EnKF can explicitly estimate the background error covariance and carry the information through assimilation cycles. The ensemble analyses provided by EnKF are the natural choice for initializing ensemble forecasts. Without the need for developing and maintaining the adjoint code, the implementation of EnKF is relatively simple. Given a proper interface, a general EnKF system can be linked to multiple prediction models. The Data Assimilation Research Testbed (DART) developed at the National Center for Atmospheric Research (NCAR) is such an example, which applies a general EnKF code to various numerical models. Still a method in development, EnKF is already applied operationally at the Meteorological Service of Canada and has shown very promising results (Houtekamer and Mitchell 2005; Houtekamer et al. 2005; Houtekamer et al. 2009; Buehner et al. 2010a; Buehner et al. 2010b).

EnKF has bred an array of variants for the implementation in atmospheric science after it was initially proposed. Houtekamer and Mitchell (1998) perturbed the observations to increase ensemble spread and alleviate filter divergence problem. This perturbed observation approach was further justified by Burgers et al. (1998), who proved that without perturbed observations, the analysis covariance is always underestimated. However, as Whitaker and Hamill (2002) pointed out, the perturbed observations will introduce new sampling error. Another category of EnKF named square root filter (SRF) does not require perturbing the observations and is “deterministic” in this sense. Most of the current leading EnKF methods belong to this SRF category, such as the ensemble square-root filter (EnSRF) of Whitaker and Hamill (2002), the ensemble transform Kalman filter (ETKF) of Bishop et al. (2001), the ensemble adjustment Kalman filter (EAKF) of Anderson (2001), and the local ensemble Kalman filter (LEKF) of Ott et al. (2004) and local ensemble transform Kalman filter (LETKF) of Szunyogh et al. (2008). Some of these methods will be briefly introduced and discussed in section 2.2.

1.1.2 Motivation

The EnKF algorithms have been widely applied in meteorology. The application of EnKF ranges from the global (Houtekamer et al. 2009; Whitaker et al. 2008), synoptic (Zhang et al. 2006; Meng and Zhang 2007; Meng and Zhang 2008a; Meng and Zhang 2008b) to the convective or storm scale (Zhang et al. 2004; Tong and Xue 2005; Xue et al. 2006). Among these applications, EnKF is employed to assimilate observations from various platforms, including conventional (Fujita et al. 2007) and radar observations (Zhang et al. 2004; Tong and Xue 2005; Aksoy et al. 2009; Jung et al. 2010). Radar observations are able to provide helpful information with high temporal and spatial resolutions on convective storms. Assimilating radar observations with EnKF for better thunderstorm forecasts has been investigated since Zhang et al. (2004). Most of the studies on convective storms with EnKF, however, are focused on assimilating radar observations only. When a radar is far away from the storm and fails to provide low level information for the storm, conventional observations including those from surface networks can help to analyze the low-level storm structure. Surface observations measure the wind, temperature, moisture and pressure at the surface. With the flow-dependent error covariance from EnKF, surface observations can not only provide storm information at low levels, such as the cold pool structure, but also are able to spread the impact upward to higher levels. Zhang et al. (2004) first examined the impact of assimilating both radar and surface observations on thunderstorm analysis with Observing System Simulation Experiments (OSSEs). But their simplifying assumptions about the simulated observations were not very realistic. Their investigation was also limited to a single set of experiments. No other study, to our knowledge, has examined the assimilation of both radar and surface observations systematically using the EnKF method.

To investigate the impact of surface observations in addition to radar data, we perform a series of OSSE experiments with realistic observational networks and numerical model setups. This is the first time to examine the impact from assimilating additional surface observations with EnKF in the presence of radar data in a relatively comprehensive way. When the radar is far away from the thunderstorm, surface observations are expected to retrieve the low level storm information missed by the radar and improve the storm analysis and forecasts. This is our major hypothesis in Chapter 3. Besides this hypothesis, several questions remain unanswered before. How far does the radar have to be away from the storms for the surface data to have significant impact? Does surface data help much when radar is very close to the storm? Among the typical surface measurements of wind, temperature, moisture and pressure, which ones have the largest and smallest impacts? Which surface observation network density is enough to provide significant additional improvement? The answers to these questions can help to design surface networks and use surface observations efficiently. To address these questions, sensitivity experiments will be conducted by putting radar to different locations, assimilating various surface measurement types and changing the surface observation network densities. We will also aim to answer the following questions: How will the impact from surface observations change if imperfect forecast models are used and imperfect storm environments are considered? Model error and storm environment error will be included to examine whether the impact from the surface observations will increase or decrease. A relatively complete perspective on how the assimilation of the surface observations with EnKF will influence the thunderstorm analysis and forecasts in the presence of radar data will be given in Chapter 3.

To make accurate predictions for hurricanes, data assimilation is also needed to initialize the forecast. Numerical prediction of hurricane tracks has greatly improved recently (Houze et al. 2007) but improvement to hurricane intensity forecasting has been limited. It is believed that the internal meso- or convective scale structures of the wind, cloud and precipitation have direct or indirect impact on hurricane’s intensity and track forecast (Houze 2007; Wang 2009; Fovell et al. 2009; Fovell et al. 2010). Radar observations can detect cloud-resolving structures within hurricanes and have the potential to produce better hurricane predictions. For an evolving hurricane system, EnKF can use radar observations to provide valuable flow-dependent error statistics information, rather than the static error covariance in 3DVAR, to update model states. In one of the few published papers on assimilating radar observations for hurricane analysis and forecast with EnKF, it is shown that the assimilation of radar radial velocity with EnKF can improve hurricane track and intensity predictions (Zhang et al. 2009). Many issues remain with this problem, such as the impact of radar data on quantitative precipitation forecasting and the impact of assimilating radar reflectivity data.

The second part of this dissertation focuses on the assimilation of both radial velocity and reflectivity observations from two coastal WSR-88D radars with EnKF for Hurricane Ike (2008). The radial velocity and reflectivity observations assimilation with EnKF is expected to improve meso-scale hurricane vortex structure, thus improve the track, intensity and quantitative precipitation forecasts of Hurricane Ike (2008). This is our major hypothesis in Chapter 4. We will also plan to answer the following questions in Chapter 4: What will be the individual and combined impacts from assimilating radial velocity and reflectivity on Hurricane Ike’s track and intensity forecasts? Which assimilation interval is sufficient to give significant improvement? How does the uncertainty of track and intensity forecasts grow in ensemble forecasts? This is the first time to assimilate the reflectivity with EnKF for the cloud-resolving hurricane analysis and forecast. Aside from examining the impact of radar observations on the track and intensity forecasts, the impact of EnKF assimilation of radar data on precipitation forecast of a hurricane is investigated, for the first time according to our knowledge.

The assimilation of mean sea level pressure (MSLP) from the best track at the vortex center with EnKF is shown to be able to improve the meso-scale hurricane analysis and forecast (Hamill et al. 2010). In their experiments with relatively coarse 30-45 km resolutions, however, the improvement on the intensity analysis is difficult to maintain for the subsequent forecasts. For our cloud-resolving hurricane forecasts with a 4 km resolution, it is still an open question if there are any additional or individual improvements on the intensity and track forecast if MSLP is assimilated with radar observations or alone with EnKF. Our questions include: How much additional impact MSLP data might provide when radar observations are already used? Is assimilation of MSLP data effective in analyzing accurately the intensity of hurricane vortex, and if so, how long can the impact last in the forecast? These questions have not been addressed before to the author’s knowledge. We perform experiments assimilating MSLP with radar data or alone with EnKF to investigate its impact, aiming to give some insight to these issues mentioned above.

The experiments mentioned above will be discussed in Chapter 4. It is the first time to assimilate reflectivity and MSLP with EnKF for the cloud-resolving hurricane analysis and forecast in a real case.

1.2 Outline of dissertation

This dissertation is organized as follows. The EnKF algorithm will be briefly introduced and discussed in Chapter 2. The impact of EnKF assimilation of additional surface observations in the presence of radar data on thunderstorm analysis and forecast will be discussed in Chapter 3 with a series of OSSEs. Sets of sensitivity experiments are also conducted to investigate the impacts of surface observation variables and surface network density when the radar is located at different distances from the storm. Such data impact was investigated in perfect and imperfect model scenarios, with and without error in the storm environment, as defined by a single sounding. The model error is introduced by using different microphysical parameterization schemes from the truth simulation, and the storm environment error is through the environmental sounding.

The impact of assimilating radar reflectivity and radial velocity observations as well as minimum mean-sea-level-pressure from best track data on the track, intensity and precipitation forecast of Hurricane Ike's (2008) is examined in Chapter 4. Besides deterministic forecasts starting from the ensemble mean analyses, ensemble forecasts starting from the EnKF ensemble analyses are also performed to examine uncertainty growth in the forecast. Summary and future plans are provided in Chapter 5.

Chapter 2

Ensemble Kalman Filters

All data assimilation methods can be approximately categorized into deterministic methods and statistical methods. Variational methods belong to the deterministic one. Statistical methods include statistical least squares, maximum likelihood method, Bayesian framework and minimum variance methods (or Gauss-Markov theorem) (Lewis et al. 2006).

Different from deterministic variational methods, Kalman or Kalman-Bucy filter is related to stochastic dynamic systems and Bayesian estimation theory. However, the gold standard of data assimilation methods, extended Kalman filter (EKF), has several drawbacks. First the computation of EKF is extremely expensive when the order of the number of degrees of freedom of the model is large. EKF also has serious problems on closure scheme and boundary conditions (Evensen 1992; Evensen 1993). To overcome these problems, Evensen (1994) introduced the Monte Carlo method into Kalman filter, proposing EnKF. As one of the most promising data assimilation methods, EnKF combines ensemble-based assimilation approaches with the traditional Kalman filter to assimilate observations into numerical models. The ensemble forecast in the data assimilation system is used to estimate the forecast error covariance while the analysis is based on the Kalman filter.

In this section the background theory for EnKF, Bayesian estimation and Monte Carlo approach will be first briefly introduced. Then EnKF and its variants will be described. Finally some issues associated with EnKF will be discussed.

2.1 Theoretical background of EnKF

Much of the material in section 2 follows Jazwinski (1970), Anderson and Anderson (1999), Anderson (2003), Evensen (2003), Kalnay (2003), Miyoshi (2004), Lewis et al. (2006), Hamill (2006), Ehrendorfer (2007) and Anderson (2009).

2.1.1 Stochastic dynamic models

Following Lewis et al. (2006), a stochastic dynamic system can be described as

$$x_{t+1} = M(x_t) + w_{t+1} \tag{2.1}$$

where x_t is the state variable x at discrete time t; M denotes a mapping of state space into itself; and w_t is the random model uncertainty or error as an external forcing term. An observation in a stochastic dynamic model can be described as

$$y_t = H(x_t) + v_t \tag{2.2}$$

where y_t is the observation at time t; H is a mapping from model space into observation space; and v_t denotes an additive observation error at time t, which is always assumed as a white noise sequence.

2.1.2 Bayesian estimation

The data assimilation problem in atmospheric science can be seen as an estimation problem in an uncertain world. The true states of atmosphere need to be estimated as accurately as possible given an imperfect numerical model and a set of imperfect observations. Since the atmosphere is far from a deterministic system, statistical estimation theory is introduced to investigate the data assimilation problem.

Let x represent the unknown quantity to be estimated, or true state of the atmosphere. $\hat{x}$ is the estimate of x. The goal to obtain the best estimate can be achieved by minimizing the error in the estimate:

$$x_{error} = \hat{x} - x \tag{2.3}$$

There are two stochastic approaches on how to obtain the optimal estimate. Fisher's framework treated x as an unknown constant μ and developed the maximum likelihood method. Bayes framework, on the other hand, assumed x is a random variable with the *a priori* distribution $p(x)$ which is known. $p(x)$ can also be defined as a multivariate probability density function as

$$\Pr(a \le x \le b) = \int_a^b p(x)dx \tag{2.4}$$

where probability density integrates to 1.0 over the entire phase space. Consider the observation model we introduced before. Define Y_t as the set of all the observations that are taken at and before time t, e.g.

$$Y_t = (y_t, Y_{t-1}) \tag{2.5}$$

where Y_0=y_0. The conditional probability density function of x_t can be indicated as

$$p(x_t \mid Y_t)\,. \tag{2.6}$$

The data assimilation problem now is to estimate this probability density function or the current atmosphere state as accurately as possible.

This probability density function to be estimated is re-expressed with Bayes' rule as

$$p(x_t \mid Y_t) = p(Y_t \mid x_t)\, p(x_t) / p(Y_t)\,. \tag{2.7}$$

The denominator can be seen as a normalization term and can be dropped for simplicity. This normalization term will guarantee the total probability after the integration is 1.

It is assumed that the observation additive noise v_t is uncorrelated for different observation times. The observation error distribution at one time is independent from the observation error at a previous time. This assumption leads to

$$p(Y_t \mid x_t) = p(y_t \mid x_t)\, p(Y_{t-1} \mid x_t)\,. \tag{2.8}$$

Substituting (2.8) into (2.7) gives

$$p(x_t \mid Y_t) = p(y_t \mid x_t)\, p(Y_{t-1} \mid x_t)\, p(x_t) / p(Y_t)\,. \tag{2.9}$$

Bayes' rule is used again as

$$p(Y_{t-1} \mid x_t)\, p(x_t) / p(Y_{t-1}) = p(x_t \mid Y_{t-1})\,. \tag{2.10}$$

The final expression is given when combing (2.9) and (2.10)

$$p(x_t \mid Y_t) \propto p(y_t \mid x_t)\, p(x_t \mid Y_{t-1}) \tag{2.11}$$

where the normalization terms are dropped.

This equation 2.11 indicates that the posterior probability distribution function is a product of two terms. The first term $p(y_t \mid x_t)$ represents the new information at the current time *t*. The second term $p(x_t \mid Y_{t-1})$ is a prior or background probability distribution, which utilizes all the information from all the previous observations. Generally, the assimilation process will make the posterior distribution narrower than the prior distribution, leading to a reduction of uncertainty. This procedure, combined with the forecast, can be repeated recursively until the time of the latest observation.

Another forecast model form different from (2.1) is

$$dx_t = M(x_t)dt + G(x_t)dq \tag{2.12}$$

where dq represents a Brownian motion process with covariance $Q_t dt$ and the G term denotes the model-error forcing. The time evolution of probability density function can be described by the Fokker-Planck or Kolmogorov equation for the vector

$$\frac{\partial p(x_t)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial [p(x_t)M_i]}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 [p(x_t)(GQ_tG^T)_{ij}]}{\partial x_i x_j}. \tag{2.13}$$

The derivation can be found in section 4.9 of Jawinski (1970) with the formula 4.119. If GQG^T is zero, or there is no model error, the second term vanishes and the Fokker-Planck equation reduces to the Liouville equation. The resulting continuity equation indicates the conservation of probability. It is also clear that the second term of model error on the right hand side can lead to the probability diffusion with time due to the model uncertainty.

Even with an elegant and simple expression of probability density function as in the update or forecast step, it is not practical to compute it directly. The computational cost will be extremely expensive for real-world numerical models. From Hamill (2006), a 100-dimensional model state requires evaluating and modifying 100^{100} density estimates in the update/data assimilation step. For the typical numerical atmospheric model with dimensions of O(10^7), the computation will never be accomplished with the direct calculation of the exact probability density function. To solve this “curse of dimensionality” problem, the Monte Carlo method and Kalman filter come as two remedies for the approximation to Bayesian estimation.

2.1.3 The Monte Carlo method

The Monte Carlo method provides the basis for ensemble-based data assimilation methods. The term “Monte Carlo method” was first proposed in the 1940s by physicists working on nuclear weapon projects in the Los Alamos National Laboratory. It is always applied when the deterministic solution is difficult to obtain due to the computation cost. The basic idea is to use random sampling to represent the actual distribution. The explicit calculation of distribution is replaced by the simulation of random samples or ensemble members.

First the ensemble members are generated as random samples from the initial probability distribution. Then the error growth can be simulated by adding random noise or other ensemble forecast techniques. By the law of large numbers, the Monte Carlo method will display a $1/\sqrt{N}$ convergence rate, i.e., the error will approach zero at a rate proportional to $1/\sqrt{N}$.

2.2 Kalman filter and EnKF

2.2.1 Kalman filter and EKF

The Kalman filter is a linear approximation to Bayesian state estimation with linear assumptions on error growth, model and observation operator. Using a linearized form of forecast model and observation operator

$$x_{t+1} = \mathbf{M}x_t + w_{t+1} \tag{2.14}$$

$$y_t = \mathbf{H}x_t + v_t , \tag{2.15}$$

the Kalman filter can be summarized as below. In forecast step,

$$x_t^f = \mathbf{M}x_{t-1}^a \tag{2.16}$$

$$P_t^f = \mathbf{M}P_{t-1}^a\mathbf{M}^T + Q_t . \tag{2.17}$$

In assimilation step,

$$x_t^a = x_t^f + K(y_t - \mathbf{H}x_t^f) \tag{2.18}$$

$$K = P_t^f\mathbf{H}^T[\mathbf{H}P_t^f\mathbf{H}^T + R]^{-1} \tag{2.19}$$

$$P_t^a = [I - K\mathbf{H}]P_t^f . \tag{2.20}$$

In these two steps, subscript "*t*" denotes the time level. The superscript "*f*" and "*a*" represents forecast and analysis respectively. Superscript "*T*" means the transpose of the matrix. *K* is the Kalman gain. *P* denotes the model state covariance. *Q* is the model error covariance cov(w_t) and *R* is the observation covariance cov(v_t). **M** and **H** are the linearized form of the forecast model *M* and the observation operator *H*, respectively.

The EKF removes the linear assumption on the model and observation operators of the Kalman filter. As indicated in the name, it is an extension of the Kalman filter from linear version to nonlinearity. It is also a first order approximation of the exact moment dynamics of the nonlinear filter. The exact moment of nonlinear filter can be found in section 29.3 of Lewis et al. (2006) and will not be described here. Most of the difficulty of applying the exact moment of the nonlinear filter is the calculation of the conditional mean $E[M(x_t)\,|\,Y_t]$ and $E[h(x_t)\,|\,Y_t]$ since the conditional probability density is unknown. $M(x_t)$ is expanded with a *r*th-order Taylor series expansion around x_t^a, the current estimate. Then $M(x_t^a)$ and the first r moments are used to calculate an approximation to $E[M(x_t)\,|\,Y_t]$. To summarize, the steps of EKF or the first order approximation is shown below. In forecast step,

$$x_t^f = M(x_{t-1}^a) + w_t \tag{2.21}$$

$$P_t^f = \mathbf{M}P_{t-1}^a\mathbf{M}^T + Q_t\,. \tag{2.22}$$

In assimilation step,

$$x_t^a = x_t^f + K[y_t - H(x_t^f)] \tag{2.23}$$

$$K = P_t^f\mathbf{H}^T[\mathbf{H}P_t^f\mathbf{H}^T + R]^{-1} \tag{2.24}$$

$$P_t^a = [I - K\mathbf{H}]P_t^f\,. \tag{2.25}$$

These two steps are similar to the linear Kalman filter except that M and H in the Kalman gain and covariance calculations are replaced by $\mathbf{M} = \frac{\partial M}{\partial x}$ and $\mathbf{H} = \frac{\partial H}{\partial x}$, the Jacobian (or linearized form) of M(x) and H(x) at x_t. Both the Kalman filter and EKF assume a Gaussian or normal distribution of background and observation error covariance. In the EKF analysis scheme, the analysis state variable x_t^a is a result of correcting the forecast value x_t^f by the observation increment $y_t - H(x_t^f)$ weighted by Kalman gain K. The background error covariance is also updated with a reduction of $K\mathbf{H}P_t^f$, reflecting the loss of uncertainty by assimilating new observations. In the next forecast step, both the state variable and the error covariance are propagated with time until next observation is available. In this discrete filter, the estimator can provide the best linear unbiased estimate (BLUE).

From Hamill (2006), there are three major limitations when EKF is used. First is the assumption of linear error growth and normal distribution of errors. When moisture or cloud cover observations are assimilated, the predictability time scale might be small, the error growth might be nonlinear and the distribution non-normal. The second limitation is the requirement of accurately estimating background error and model error covariance. The estimate of model error covariance Q is particularly difficult. The third limitation is computation cost. Although the EKF reduce the computation cost from the Bayesian estimate greatly, the computational cost in EKF is still extremely high. For a state vector of dimension n, the computation cost for K is at least in an order of n^2. It is not practical to implement EKF in a typical atmospheric model with the dimension over $O(10^6)$.

2.2.2 Ensemble Kalman filter

Almost all of the variants of the ensemble-based data assimilation methods share some general properties. They all use an ensemble of forecasts to estimate the background error covariance. Then in the data assimilation step, all the ensemble members are updated to an ensemble analysis. Short-term ensemble forecasts follows until the next set of observations is available. When the error growth is linear and the error distribution is normal, the state estimate from the ensemble-based assimilation methods converges to EKF.

To include the additional dimension from the ensemble, the notation for the forecast state variables at time *t* is changed to $x_{t,i}^f$, where $i=1,\ldots,N$, the *i*th member of the ensemble and *N* is the ensemble size. The state variable ensemble represents the random samples from the actual distribution of model space. For each ensemble member of the state variable, the dynamical model and observation operator (or observation model) (2.1) and (2.2) still apply.

The ensemble mean for the forecast is defined as

$$\overline{x}_t^f = \frac{1}{N}\sum_{i=1}^{N} x_{t,i}^f \, . \tag{2.26}$$

The perturbation is defined as $x'^{f}_{t,i} = x^{f}_{t,i} - \overline{x}^{f}_{t}$ where i=1,…,N. Also define

$$X'^{f}_{t} = (x'^{f}_{t,1}, ...x'^{f}_{t,i}) \tag{2.27}$$

as the matrix form for the perturbations. The forecast error covariance is estimated with the ensemble as

$$\hat{P}^{f}_{t} = \frac{1}{N-1} X'^{f}_{t} (X'^{f}_{t})^{T} . \tag{2.28}$$

The analysis counterparts for ensemble mean, perturbation and error covariance are similar to the forecast ones and only the superscript "*f*" is replaced by "*a*". The accent "^" means the estimate.

The use of ensemble forecasts to estimate the background error covariance can mitigate the problem of linear assumption in EKF. In EKF, the nonlinear model is expanded with a Taylor expansion series and truncated to the first order approximation, which assumes the error growth from model advancing is linear. In ensemble-based data assimilation methods, the nonlinear forecast model is used for the ensemble forecasts. This non-linear error growth in ensemble data assimilation methods provides a more accurate estimate of the background error statistics.

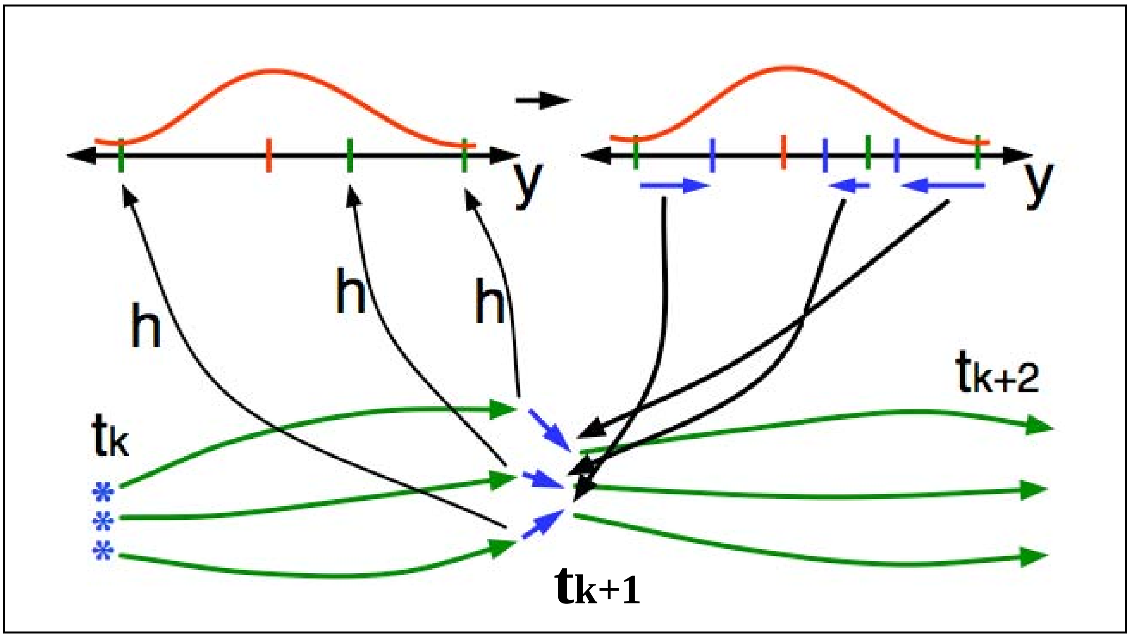

Fig. 2.1. Schematic plot of EnKF algorithm from Anderson (DART tutorial). The star denotes the ensemble state variables at the time level t_k. y means the observations and h is the observation operator. Green line denotes the ensemble forecast. The blue arrow on the lower part means the update of the state variable. The explanation for the figure can be found in the text.

The ensemble data assimilation procedure is briefly outlined in Fig. 2.1. The state variable ensemble evolves with time simultaneously from time t_k (or t) to t_{k+1} (or $t+1$). The uncertainty also grows during the forecast as indicated by the separation of the green arrows at time t_{k+1}. At time t_{k+1}, the observation operator h, is applied to the state variable ensemble to project them into the observation space (the black arrow with h). The ensemble increments in the observation space are calculated and transferred back into the state variable space (the black arrows at the right hand side). Then the forecast state variable ensemble is updated with the weighted analysis increment (the small blue arrows in the lower middle). After the update step, the new updated state variable ensemble is integrated with the dynamic model until the next set of observations is available again and this assimilation cycle is repeated.

The two major classes of ensemble filters include the stochastic (fully Monte Carlo) and deterministic schemes. The stochastic method adds random noise on the observations to simulate the uncertainty in the observations and is also called “perturbed observation” method. The deterministic method avoids the observation perturbations and the analysis ensemble perturbations are calculated by a linear combination of forecast ensemble perturbations. Since in these deterministic methods, the square root of the error covariance matrix is always used for the error covariance estimation, these deterministic methods are also called square root filters.

In the following sections, both stochastic and deterministic methods will be introduced and discussed first. A description of the major variants of the deterministic methods, such as EAKF, ETKF and EnSRF will also be provided. Then the issues associated with the EnKF, like the sampling error and the model error are discussed. The algorithms to reduce the sampling error, such as covariance inflation and covariance localization, are also introduced.

2.2.3 Stochastic methods

The stochastic method is also known as the perturbed observation method. As the name indicates, the observations are perturbed with random noise. The random noise has a normal distribution of zero mean and error covariance R:

$$\frac{1}{N}\sum_{i=1}^{N} y'_i = 0 \tag{2.29}$$

The observation perturbation vector is defined as $y' = (y'_1, \ldots\ldots, y'_N)$ and

$$\frac{1}{N-1} y'(y')^T = R_e \tag{2.30}$$

where R_e is the observation error covariance estimation. The perturbation is added to the observation as

$$y_i = y + y'_i, i = 1, \ldots\ldots, N. \tag{2.31}$$

For simplicity, the subscript indicating the time level is dropped in the remaining part of the section since all of the discussion about the update algorithm is only related to the current time level t. The individual state variable member is updated as below:

$$x_i^a = x_i^f + K[y_i - H(x_i^f)] \tag{2.32}$$

where

$$K = \hat{P}^f \mathbf{H}^T (\mathbf{H}\hat{P}^f \mathbf{H}^T + R)^{-1} \tag{2.33}$$

In the Kalman gain expression, **H** is still the Jacobian or linearized form of the observation operator *H*. The Kalman gain expression is similar to EKF. Only in EnKF the error covariance $\hat{P}^f$ is estimated with the ensemble, leading to a reduction of the computational cost. Burgers et al. (1998) shows that if the unperturbed observations are assimilated and the ensemble size approaches infinity, the estimated analysis error covariance $\hat{P}^a$ will be

$$(I - K\mathbf{H})\hat{P}^f (I - K\mathbf{H})^T, \tag{2.34}$$

which underestimates the error covariance in EKF: $(I - K\mathbf{H})P^f$. When the observation perturbation is added, the final $\hat{P}^a$ will be

$$(I - K\mathbf{H})\hat{P}^f (I - K\mathbf{H})^T + KRK^T. \tag{2.35}$$

The last term above is related to the observation perturbations. This new error covariance estimation will converge to the EKF error covariance when the infinite ensemble is assumed. However, the introduction of observation perturbation can lead to the spurious observation-background correlation and a bias in analysis error covariance (Whitaker and Hamill 2002). To correct the bias, several authors proposed their solution. To avoid the “inbreeding” problem of underestimating the error covariance, Houtekamer and Mitchell (1998) built a pair of EnKFs, using one set’s error covariance to update the other set. Pham’s (2001) second order exact EnKF adds perturbation to the background forecast members instead of observations.

2.2.4 Evaluation of error covariance in the practical implementation

For the practical implementation of EnKF, explicitly calculating the forecast error covariance matrix is still expensive. Instead, the covariance matrix components $\hat{P}^f \mathbf{H}^T$ and $\mathbf{H}\hat{P}^f \mathbf{H}^T$ are computed. Define

$$\overline{H(x^f)} = \frac{1}{N}\sum_{i=1}^{N} H(x^f)\,. \tag{2.36}$$

Then

$$\hat{P}^f \mathbf{H}^T = \frac{1}{N-1}\sum_{i=1}^{N}(x_i^f - \overline{x^f})[H(x_i^f) - \overline{H(x^f)}] \tag{2.37}$$

and

$$\mathbf{H}\hat{P}^f \mathbf{H}^T = \frac{1}{N-1}\sum_{i=1}^{N}[H(x_i^f) - \overline{H(x^f)}][H(x_i^f) - \overline{H(x^f)}] \tag{2.38}$$

where $\hat{P}^f$ denotes the estimation of the forecast error covariance. The first equation evaluates the covariance between the forecast state variables on the model grids and the observation points. The second equation evaluates the covariance between the forecast state variables projected in the observation space. The nonlinear observation operator in the covariance calculation is opposite to the linearization of the observation operator in EKF and is more realistic.

2.2.5 Sequential data assimilation

This reduction of computational cost in the last section will be lost if the dimension of the observation is as large as the model space dimension. In that case, the size of $\hat{P}^f\mathbf{H}^T$ and $\mathbf{H}\hat{P}^f\mathbf{H}^T$ will be the same as $\hat{P}^f$. To avoid the expensive cost of manipulating and storing the large matrix, sequential data assimilation is always employed in EnKF implementation. Bishop et al. (2001) demonstrated that if the observation errors are independent from each other, the effect on error covariance estimation of serially assimilating these observations is equivalent to assimilating all of the observations simultaneously. The updated ensemble state variables after the first observation is assimilated will be used as the background for assimilating the second observation. For each observation assimilated, $\mathbf{H}\hat{P}^f\mathbf{H}^T$ and R reduce to scalars. Thus the inverse of ($\mathbf{H}\hat{P}^f\mathbf{H}^T + R$) in the Kalman gain formula (eq. 2.33) will involve no matrix manipulation and is trivial to compute.

The only problem is that in the real observation set, horizontal or vertical correlation may exist between the serially assimilated observations. In this case, the observations can be organized into small batches so that the correlation among these distinct batches is small enough for the independent correlation assumption (Houtekamer and Mitchell, 2001). Then these batches are assimilated sequentially while the observations in each batch will be analyzed simultaneously. The computational cost will not be too expensive in each batch for manipulating the matrix ($\mathbf{H}\hat{P}^f\mathbf{H}^T + R$).

2.2.6 Parallelization for EnKF

Sequential EnKF algorithm still requires expensive computational cost. The high memory requirement for storing and analyzing the ensemble model states makes code parallelization necessary. Also, to speed up the assimilation, code optimization for parallelization is always welcome.

In a shared-memory system, all of the computer processors access and address the system memory uniformly. OpenMP directives can be inserted into the code to parallelize the update step. For each serially assimilated observation, multiple model state variables can be updated simultaneously with OpenMP parallelization. OpenMP is easy to implement but this parallelization is limited to shared-memory systems or computer nodes with large memory.

In a distributed-memory system, each processor has its own memory and other processors can access its memory only via network communications. Several authors proposed various parallelization algorithms. Keppnne and Rienecker (2002) used domain decomposition to realize parallelization. Each private sub-domain (or PE, processing element) contains its own set of model states of each ensemble member and the observations. The analysis can be conducted on each PE independently. And the analysis on each PE can be executed simultaneously. Houtekamer and Mitchell (2001) also used this domain decomposition method but their implementation still assimilated each batch of observations sequentially. Anderson’s (2007) scalable implementation of EnKF also discussed parallelization strategy. The model states are distributed over the PEs with certain loading balance criterion and need to be transposed over the PEs several times during the analysis. The observations are still assimilated sequentially.

2.2.7 Deterministic methods

The observation perturbations added in the stochastic method introduce a new source of spurious observation-background error correlation if a small ensemble is used (Whitaker and Hamill, 2002). The correlation terms between $y^{'o}$ and $x^{'b}$ terms in (7) of Whitaker and Hamill's (2002) are not negligible and can change the distribution shape of analysis error statistics. The analysis error covariance may be underestimated in the stochastic method. Several methods avoid the explicit calculation of the background error covariance and manipulate the square root of the error covariance matrix. They are called "deterministic" by avoiding random observation perturbations and "square root filter" by using the square root matrix. Define the square root of the forecast and analysis error covariance estimate as E^f and E^a:

$$\hat{P}^f = E^f (E^f)^T \text{ and } \hat{P}^a = E^a (E^a)^T . \tag{2.39}$$

which are also the scaled forecast or analysis ensemble perturbations:

$$E^f = \frac{1}{\sqrt{N-1}} X'^f . \tag{2.40}$$

Assume the scaled analysis ensemble perturbation is a linear combination of the forecast ensemble perturbation:

$$E^a = E^f U . \tag{2.41}$$

where U is the unknown matrix to be determined. Then the covariance update expression (2.25) can be rewritten as

$$E^f U U^T (E^f)^T = [I - K\mathbf{H}] E^f (E^f)^T . \tag{2.42}$$

U can be solved with matrix factorization and used in (2.41) for the update scheme. For the covariance $P=AA^T$, it also can be written as $P=(AS)(AS)^T$, S representing any orthonormal matrix such that $SS^T = S^T S = I$. Therefore the solution of the square root is not unique. Three square root filters with different solutions but the same analysis error covariance estimate are introduced below as the examples of the deterministic update algorithms.

2.2.8 Ensemble square root filter (EnSRF)

Whitaker and Hamill (2002) proposed this serial square root filter, or EnSRF. Assume the square root analysis matrix has the form of

$$E^a = (I - K_{rd}\mathbf{H})E^f . \tag{2.43}$$

The solution of the “reduced” gain K_{rd} is

$$K_{rd} = \hat{P}^f \mathbf{H}^T [(\mathbf{H}\hat{P}^f \mathbf{H}^T + R)^{-\frac{1}{2}}]^T [(\mathbf{H}\hat{P}^f \mathbf{H}^T + R)^{\frac{1}{2}} + R^{\frac{1}{2}}]^{-1} \tag{2.44}$$

(Andrews 1968), cf. (10) of Whitaker and Hamill 2002). When uncorrelated observations are assimilated sequentially, $\mathbf{H}\hat{P}^f \mathbf{H}^T$ and R are scalars. The expression of K_{rd} can be simplified as

$$K_{rd} = (1 + \sqrt{R(\mathbf{H}\hat{P}^f \mathbf{H}^T + R)^{-1}})^{-1} K \tag{2.45}$$

where K is the traditional Kalman gain in (2.24). Define

$$\alpha = (1 + \sqrt{R(\mathbf{H}\hat{P}^f \mathbf{H}^T + R)^{-1}})^{-1} \tag{2.46}$$

and $K_{rd} = \alpha K$. α is between 0 and 1, indicating a reduced Kalman gain relative to the traditional K. This expression is first derived by Potter (1964).

Ensemble data assimilation cycles in EnSRF are similar to EnKF. For convenience, the update step is divided into two parts with updating the ensemble mean and the ensemble perturbation or deviation separately:

$$\overline{x^a} = \overline{x^f} + K[y - H(\overline{x^f})], \tag{2.47}$$

$$x'^a_i = x'^b_i - K_{rd} H(x'^b_i). \tag{2.48}$$

The analysis ensemble is less modified by the observations than in the stochastic EnKF. This reduced weighting coefficient K_{rd} compensates for the over-reduction of variance by using the traditional K in the stochastic method. The computation cost of EnSRF is the same as the stochastic EnKF. As Whitaker and Hamill (2002) showed, with moderate ensemble size and Gaussian assumption, EnSRF produces better results than the perturbed observation method.

2.2.9 Ensemble transform Kalman filter (ETKF)

Bishop et al. (2001) proposed ETKF. Define

$$E^a = E^f T \tag{2.49}$$

where T is the transformation matrix. The Sherman-Morrison-Woodbury identity matrix formula

$$(I_m + cd^T)^{-1} = I_m - \frac{cd^T}{1 + d^T c}, \tag{2.50}$$

where c and d are m-dimension vectors, can be used to rewrite the Potter's formula. The solution for T is

$$T = C(\Gamma + I)^{-1/2} \tag{2.51}$$

where C is a matrix composed of eigenvectors of $(E^f)^T \mathbf{H}^T (R^{-\frac{1}{2}})^T R^{-\frac{1}{2}} \mathbf{H} E^f$ and Γ is a diagonal matrix composed of eigenvalues of the same matrix.

ETKF has faster computation speed than EnSRF but the covariance localization (section 2.3.3) cannot be applied in ETKF, inducing large sampling error unless large ensembles are employed.

2.2.10 Ensemble adjustment Kalman filter (EAKF)

EAKF is proposed by Anderson (2001) in a joint observation-state space with Bayesian estimation theory, assuming error statistics are Gaussian. EAKF uses rotation and scaling to transform the forecast error covariance into an identity matrix in a reference frame (appendix A of Anderson 2001). The adjustment A is used to update the forecast states as

$$E^a = AE^f \,. \tag{2.52}$$

A singular value decomposition gives $P^f = FDF^T$ where F is unitary (or orthonormal for real counterpart) and D is a diagonal matrix with the singular value μ_p of P^f on the diagonal. The adjustment A is given by

$$A = FD^{1/2}U[I + D]^{-1/2} D^{-1/2} F^T \tag{2.53}$$

cf. (3.3.1) of Miyoshi 2005, where U is an orthonormal matrix. As Tippet (2003) pointed out, this update scheme is equivalent to

$$E^a = E^f C[I + \Gamma]^{-1/2} D^{-1/2} F^T E^f \tag{2.54}$$

where C and Γ are the same denotation as in section 2.2.9. This scheme is the same as applying a transformation $D^{-1/2}F^T E^f$ to the ETKF analysis scheme.

2.3 Issues related to EnKF implementation

2.3.1 Filter divergence

Filter divergence always occurs when the analysis error covariance is underestimated. In the follow-up forecast cycles, the forecast errors will also be underestimated. The background will be falsely considered to be too accurate by the filter. This underestimation of background error will lead to a small impact from the observations or even disregarding the observations in the worst case when this effect accumulates. As more observations are ignored in cycles as the result of a feedback mechanism, the filter will diverge from the truth. Filter divergence can be caused by sampling error and model error. These two major sources of filter divergence will be briefly discussed below.

2.3.2 Sampling error and covariance inflation

As van Leeuwen (1999) has shown, the finite and small ensemble size in EnKF is an essential source of underestimating the analysis error covariance. This underestimation is a systematic effect due to the finite ensemble size and the nonlinearity in Kalman gain. In EnKF, modeling the background error covariance as accurately as possible is one of the most important requirements to reduce sampling error and avoid filter divergence (Hamill 2006). There are at least two aspects in correctly modeling the background error statistics: variance magnitude and spatial covariance structure. To keep the correct background error variance, covariance inflation is always used to maintain the spread among the ensemble members.

The idea behind covariance inflation is simple. It is assumed that the background correlation structure is correctly modeled but the variance of model state variables is too small. The error variance of the model state variables needs to be inflated by multiplying a constant or adding more perturbations with zero mean.

Anderson and Anderson (1999) proposed the simple multiplicative covariance inflation scheme as

$$x_{i(new)}^{f} = r(x_i^f - \overline{x^f}) + \overline{x^f} \tag{2.55}$$

where r is a constant slightly larger than 1. The deviation of each member from ensemble mean is increased before the observations are assimilated. Although it is simple to implement, this constant inflation may not be suitable for heterogeneously distributed observation network. Tuning the inflation coefficient r is also a problem for large geophysical models. In the practical implementation, Tong and Xue (2005) choose to inflate the state variables within the observation coverage. Anderson (2009) designed a spatially and temporally varying adaptive inflation strategy with the Bayesian approach. This adaptive method showed promising results compared to the constant inflation when it is used in low-order model. Another simple adaptive algorithm from Whitaker and Hamill (2010 EnKF workshop presentation) is applied to the posterior state variables after the analysis as

$$x'^{a}_{i(new)} = x'^{a}_{i} \sqrt{r \frac{\sigma_f^2 - \sigma_a^2}{\sigma_f^2} + 1} \tag{2.56}$$

where σ_f^2 and σ_a^2 are the variance of the model state variables before and after the analysis, and *r* is a tuning coefficient. This adaptive method also adaptively considers the impact of the increments and the density of observation network. However, the sharp spatial gradient of the inflation may occur and destroy the weather structures.

Zhang et al. (2004) used a relaxation method as

$$x'^{a}_{i(new)} = rx'^{f}_{i} + (1-r)x'^{a}_{i} \tag{2.57}$$

where *r* is a weighting coefficient to combine x'^{f}_{i} and x'^{a}_{i} linearly. This scheme modifies the analysis deviation by "relaxing" or weighting the forecast and analysis deviation.

Besides the multiplicative inflation, additive noise or additive inflation is another method to increase the spread artificially (Mitchell and Houtekamer 2000; Houtekamer and Mitchel 2005; Houtekamer et al. 2005; Hamill and Whitaker 2005). The additive noise with the same dimension as the model state will be added to each ensemble member before or after analysis. As Hamill and Whitaker (2005) pointed out, additive error can account for the model error and should sample the statistics of the accumulated model error. The additive error will be discussed more in section 2.3.4.

A combination of multiplicative and additive inflation can provide better results than single schemes (Whitaker and Hamill 2010 EnKF workshop presentation).

2.3.3 Covariance localization

The sampling error also has negative impact on the spatial structure of background error covariance estimate. Covariance localization is based on the fact that the spatially distant error covariance estimates are much less accurate than the local estimates. This measure limits the impact of observations to the local grid points of model state and is expected to remove the spurious background error covariance estimates at long distances. On the other hand, localization can reduce the computational cost and makes more efficient parallelization schemes possible.

A widely used localization is to modify the background error covariance by building a Schur product between the background error covariance by a correlation function. The correlation function always decreases monotonically as the distance between the observation assimilated and the model state variable increases. Gaspari and Cohn (1999) constructed such a Gaussian-shaped function with a fifth order piece-wise polynomial. This distance-dependent correlation function is 1 at the observation point and decreases gradually as the distance increases. The correlation turns to 0 beyond some pre-specified distance (Hamill et al. 2001). Covariance localization can also increase the rank of the forecast covariance estimate to the dimension of the state vector thereby introducing the extra degrees of freedom (section 5.3 of Hamill 2001).

The choice of localization cutoff radius requires efforts on tuning. If the radius is too large, spurious covariance will not be removed effectively. If the radius is too small, the observation will not have enough impact on the model states. Anderson (2007) developed an adaptive method to estimate the localization functions. A hierarchical ensemble filter, or an ensemble of ensemble filters, is used to estimate the impact of spurious correlations between the observations and model state variables, avoiding a prior specification of localization functions.

Other issues include the imbalance caused by severe localization (i.e. small localization cutoff radius). The spatial coherence can be disrupted by localization and some physical balance can be disturbed (Mitchel et al. 2002). For non-local observation operators, such as satellite radiance data or attenuation in the radar observations, it is still a problem to apply the localization and further study is needed.

2.3.4 Model error

During the whole EnKF discussion above, the model error term w in (2.1) is ignored. The numerical model is supposed to carry additional uncertainty due to the unresolved processes. To discuss the model error, another assumption is made that the model is unbiased. Ignoring the model error term leads to an evolved forecast error covariance as

$$\hat{P}_t^f \cong \mathbf{M}\hat{P}_{t-1}^a\mathbf{M}^T \, . \tag{2.58}$$

The comparison with (2.22) reveals that during the ensemble forecast, EnKF will have small forecast error covariance estimate without the model error covariance Q. Filter will diverge from the truth due to the model error.

There is several ways to deal with model error. The most ideal method is to use a stochastic dynamic model instead of a deterministic model. The error forcing can be added on the forecast model and integrated forward. Buizza et al. (1999) added noise to the parameterized physical process in the forecast model by multiplying a random number with the total parameterized tendency terms in the prognostic equations. Penland (2003) discussed several approaches for including the stochastic forcing in numerical models. However, the application to incorporate the stochastic forcing in real-time weather prediction model is still limited.

The second approach to reduce the impact from the model error is adding additive error to the background before data assimilation to increase the forecast error variance:

$$x_{i(new)}^{f} = x_i^f + \eta_i \tag{2.59}$$

where η_i satisfies $<\eta_i(\eta_i)^T>=Q$. Hamill and Whitaker (2005) tried several methods to generate the noise, such as using the structured difference between high and low resolution model forecasts as the sample, or using scaled time series perturbations obtained from subtracting a climatology mean. The sample from the resolution difference has the best result. Mitchell and Houtekamer (2000) used innovation information to estimate the model error with the maximum likelihood approach. The additive inflation method can span a different subspace from the original one spanned by the ensemble (Hamill and Whitaker 2005). It therefore can reduce both sampling error and model error. However, multiplicative inflation will not modify the subspace spanned by the ensemble and may not be effective in reducing the model error.

The last method discussed here to reduce model error is the use of multiple models to generate ensemble forecasts (Houtekamer et al. 1996; Hou et al. 2001; Hansen 2002). Even with the simplicity of implementation, there are still uncertainties involved in multi-model method. The balance and the representation of model error are some issues associated with multi-model ensemble (Hamill 2006).

2.4 The EnKF flowchart

The procedure for EnKF experiments in my PhD research is briefly listed step by step below:

1. Random perturbations for state variables are generated.

2. Gaussian-shaped function or recursive filter are applied on the perturbations to smooth the perturbations and create some spatial correlations.
3. Smoothed perturbations are added to the background filed to generate ensemble members.
4. Ensemble forecasts are started.
5. EnKF analysis is conducted when the observations are available.
6. Steps 4-5 are repeated until all the analysis cycles end.
7. The deterministic forecast or ensemble forecasts are started.

A flowchart is given in Fig. 2.2 to illustrate the above steps.

During step 5, EnSRF is used in Chapter 3 and Chapter 4. Eq. 2.47 and 2.48 are used to update ensemble mean and ensemble perturbations respectively. Eq. 2.33 and 2.45 are used to calculate the traditional and reduced Kalman gains respectively. Eq. 2.37 and 2.38 are used to evaluate the error covariance. The EnKF experiments in Chapter 3 and 4 will follow this procedure generally. Some steps may vary and will be described in detail in Chapter 3 and 4.

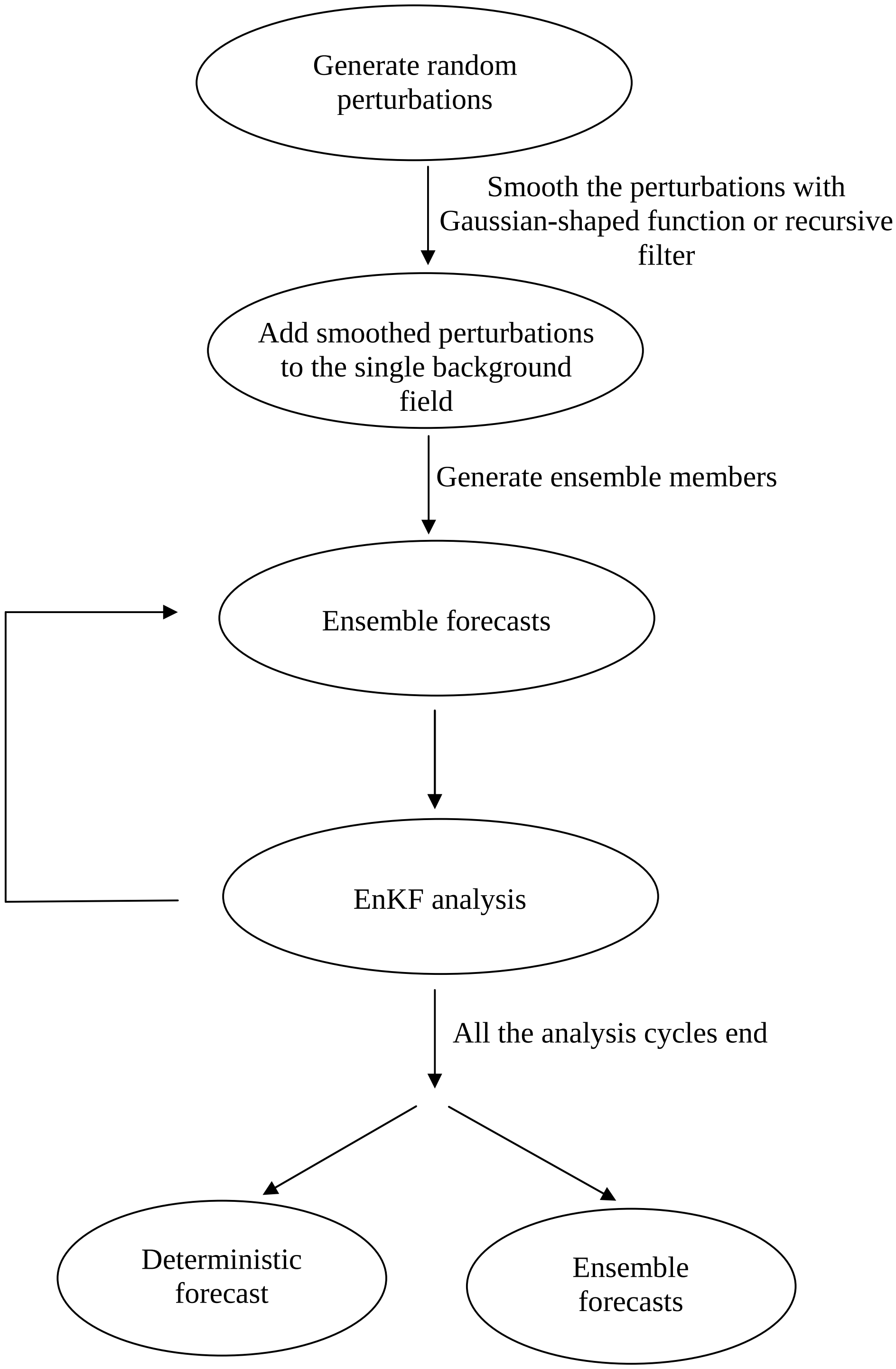

Fig. 2.2. Flow chart of EnKF algorithm in this dissertation.

Chapter 3

The Analysis and Impact of Simulated High-Resolution Surface Observations for Convective Storms with Ensemble Kalman Filter

3.1 Introduction

Successful numerical weather prediction (NWP) depends greatly on accurate initial conditions obtained with data assimilation. For convective storms, radar is the primary observational platform used. However, radars usually do not observe down to ground level, because of the non-zero elevation of the lowest scans and due to the earth curvature effect. This problem becomes worse when the storm is located far from the radar. For example, when a storm is 100 km away, the center of the lowest 0.5° elevation beam of a WSR-88D radar is more than 2 km above the ground. Even the lower edge of the half power beam does not reach ground (see, e.g., Fig. 2 of Xue et al. 2006, XTD06 hereafter). Yet, for convective systems, the low-level flows and cold pool are critically important in storm development and evolution. At ground level, observations from automated meteorological stations and sometimes from mesoscale observing networks are available. Effective assimilation of these observations, in combination with radar data, has the potential to significantly improve storm analysis and forecasts. These data are also valuable for defining the low-level environment of the convective storms, where radar has very limited observing capabilities. The resolution of the surface network, however, is usually low compared to the scales of convective features; thus their quantitative impact when radar is also present is not necessarily clear. Further, it is not clear if the information contained in surface observations can be effectively retained by the storm when observations are taken in the downdraft region, which is, however, primarily responsible for the cold pool.

The assimilation of the surface observations into NWP models is not as easy as it might appear. One issue is differing elevation of surface observations from the model ground level. Complex terrain will complicate the issue of spatial representativeness of the observations, which affects the proper spatial spreading of observation increments (e.g., Lazarus et al. 2002; Deng and Stull 2007). A more general issue is the vertical spread of observation information; this is non-trivial because of the typically large spatial and temporal variations of the boundary layer error structures.

For the prediction of convective initiation and its later evolution, Liu and Xue (2008) showed that the strength of analyzed cold pool is sensitive to the vertical correlation length scale specified in the ARPS (Xue et al. 2000) Data Analysis System (ADAS, Brewster 1996), when analyzing high-resolution surface observations. Similar sensitivity was found by Dawson and Xue (2006). Within ADAS, the vertical correlation length can be specified in terms of geometric height or potential temperature. The latter allows more vertical spreading of observation information in less stable condition, but the correlation scale is still empirical. Using ADAS and observations from both routine and special mesoscale surface networks that were gathered by the 2002 International H_2O Project (IHOP_2002, Weckwerth et al. 2004), Liu and Xue (2008) demonstrated significant positive impact of hourly surface data over a 6-hour period on convective initiation forecast in the ARPS model. Similar results were obtained by Xue and Martin (2006) for another convective initiation case from IHOP_2002.

To optimally assimilate surface observations, flow-dependent background error covariance is needed. The ensemble Kalman filter method (EnKF, Evensen 2003) is one method that estimates and evolves flow-dependent background error covariance using ensemble forecasts through assimilation cycles.

The assimilation of surface observations for the planetary boundary layer using EnKF has been examined recently in simple column model settings with simulated and real observations (Hacker and Snyder 2005; Hacker and Rostkier-Edelstein 2007). Simulated surface observations have also been tested with a mesoscale model for the case of a synoptic-scale winter cyclone by Zhang et al. (2006) and Meng and Zhang (2007). Fujita et al. (2007) examined the performance of EnKF for synoptic to mesoscale flows with real surface observations. These studies generally address the situations where the atmospheric boundary layer is strongly influenced by the land surface processes; they do not address the specific situation where the thunderstorm downdraft and cold pool play an important role.

Assimilation of high-resolution surface observations, from e.g., a high-density mesonet, for the initialization of explicit convective storms, however, is limited to the study of Zhang et al. (2004), which found positive impact of simulated surface observations at 10 km spacing when radar data were artificially limited to levels above 4 km. Their study also made the simplifying assumptions that the radar observations are available at model grid points; one of the surface observation types considered is the liquid-water potential temperature which is not directly measured. Only warm rain microphysics was used. These simplifying assumptions make it impossible to address the issue of radar distance from the storm, which in many cases is the most important factor determining the low-level data coverage. The assumption that the radar data are available at the model grid points also makes it impossible to include the effect of beam spreading which affects data coverage. The low-level data coverage is a key problem of the National Science Foundation (NSF) Engineering Research Center (ERC) for Collaborative Adaptive Sensing of the Atmosphere (CASA) project, which seeks to develop low-cost high-density radar networks (XTD06, McLaughlin et al. 2007) to improve lower atmospheric sensing. The lack of low-level radar data coverage is also suspected to be an important source of error in the EnKF analysis of thunderstorms in the studies of Dowell et al. (2004) and Tong (2006).

We set out to perform a systematic OSSE (Observing System Simulation Experiment, Lord et al. 1997) study on the impact of assimilating high-resolution observations from hypothetical surface networks, in addition to observations from a radar located at different distances from the main storm. A realistic radar simulator is used, providing realistic data coverage. Surface observing networks of different mean spacing are examined, and the impact of surface data is interpreted with the help of ensemble error covariance structures and our understanding of convective storm dynamics. Both perfect and imperfect forecasting model scenarios are considered in the OSSEs; in the latter case model error is introduced by using wrong microphysics schemes. It should be noted that for simplicity, only OSSEs are conducted in this study. No calibration or validation procedures with real observations sensitivity experiments are performed. In a more complete OSSE framework, the error statistics between “existing” simulated observations and real observations, which is radar data in this study, and the data impact from the existing observations between OSSE and Observing System Experiments (OSE) could be compared. Based on the guideline from the above validation, the impact of the "new" simulated observations from OSSEs could be better interpreted.

The rest of this chapter is organized as follows. Section 3.2 describes the numerical model used, the design of the observational network, and the EnKF algorithm used. Experiments assimilating surface observations with EnKF are discussed in section 3 and sensitivity tests are presented and analyzed in section 3.4. The impact of this assimilation on subsequent forecasts is discussed in section 3.5. Section 3.6 further examines the surface data impact in the presence of model error, and section 3.7 discusses the impact from surface observations under imperfect storm environment. Section 3.8 gives a summary and conclusions.

3.2 Model and experiment settings

We take the OSSE approach in this study, partly because it is impossible or impractical to observe the same real storm many different times using the same radar from different distances, and the real surface observing networks often have limited density. In the case of real thunderstorms, full measurement on the complete atmospheric state is not available, making the quality of analysis difficult to judge. With OSSEs, a model simulation serves as the 'truth' or 'nature run' for all experiments. Realistic observations can be simulated from this model atmosphere using a radar simulator and assuming certain surface network characteristics. With OSSEs, observation configurations that are currently unavailable can be tested. This study builds upon and extends the earlier OSSE studies of Tong and Xue (2005a, TX05 hereafter), XTD06, and Tong and Xue (2008a), which focused on radar data.

3.2.1 Description of simulation and assimilating model

The Advanced Regional Prediction System (ARPS, Xue et al. 2000; Xue et al. 2001; Xue et al. 2003) is used as the simulation and prediction model in this study. ARPS is a compressible nonhydrostatic model that predicts velocity components u, v, w, potential temperature θ, pressure p, and mixing ratios for water vapor, cloud water, rainwater, cloud ice, snow, and hail/graupel (q_v, q_c, q_r, q_i, q_s, and q_h, respectively) associated with ice microphysics schemes in the model. These variables make up the state vector that is estimated or analyzed using EnKF. The model also predicts the turbulence kinetic energy (TKE), which is used in a 1.5-order subgrid-scale turbulence closure scheme. TKE is not updated by the EnKF during assimilation as in our previous studies (e.g., TX05). The ARPS includes several single-moment ice microphysics packages. The Lin et al. (1983, LFO83 hereafter) scheme is the default and is used in the truth simulation and perfect model OSSEs. A modified version of the LFO83 scheme (Gilmore et al. 2004a, LFO04 hereafter), the Schultz simplified ice scheme (Schultz 1995, Schultz hereafter), the WRF 6-category Single-Moment Microphysics scheme (Hong and Lim 2006, WSM6 hereafter), and their combinations are used in the imperfect-model OSSEs.

3.2.2 Truth storm simulation

As in TX05 and XTD06, the 20 May 1977 Del City, Oklahoma, supercell storm (Ray et al. 1981) simulated by the ARPS is used as the truth for the OSSEs. A horizontal resolution of 2 km is chosen with a grid of $67 \times 67 \times 35$ points in the x, y and z directions, respectively, giving a *physical* domain of 128 × 128 × 16 km. Radar and surface observations are assumed to be available only within a 64×64 km^2 subdomain located at the center of the model grid. The main storm cells remained within this subdomain throughout assimilation and prediction, and this observation domain has the same size as that used by TX05. The larger domain used here is to avoid occasional problems arising from perturbations created by the open lateral boundary condition. In the vertical, a grid stretching scheme based on a hyperbolic tangent function is used and the vertical grid spacing is 200 m near the ground and increases to 800 m at the model top.

An initial thermal bubble with a maximum potential perturbation of 4 K is centered at $x = 80$ km, $y = 48$ km and $z = 1.5$ km to initialize convection in the truth simulation. The radii of the bubble are 10 km in the horizontal and 1.5 km in the vertical. Open lateral and free-slip top and bottom boundary conditions are used in both simulation and assimilation. A constant wind of $u = 3$ m s^{-1} and $v = 14$ m s^{-1} is subtracted from the original sounding to keep the main storm (right mover) near the center of the domain. These configurations are the same as those used in TX05, except for the larger computational domain and the use of vertical grid stretching; the stretched grid gives better vertical resolution near the surface.

The bubble-triggered storm updraft reaches its full updraft intensity within 30 min. At around 60 min, the supercell starts to split into two, with one right mover and one left mover (Fig. 1a). By 120 min, the left mover exists the northwest corner of the central subdomain (Fig. 2a). Additional details on the general evolution of the simulated storm can be found in TX05. The simulated model state is output every 5 min for observation simulation and for analysis/forecast verification.

3.2.3 Simulation of observations

A hypothetical WSR-88D radar with 1 degree beamwidth is positioned to the southwest of the storm. For the first set of experiments, the radar is located at x = - 68 km and y = - 68 km (the coordinate origin is located at the southwest corner of the 128 × 128 km model domain); this is about 185 km from the domain center; approximately where the right-moving cell is located. At this distance, the earth curvature effect combined with beam bending based on the 4/3 earth radius model (Doviak and Zrnic 1993) places the lower edge and the center of the half-power beam of 0.5 degree elevation at 1.98 and 3.60 km above ground, respectively. In another word, there is no direct radar data coverage below 1.98 km level at all at this distance. The vertical beamwidth as well as the spacing between two consecutive beams is about 3.8 km at this distance; for this reason, proper vertical beam pattern weighting is important. Two other sets of experiments assume that the radar is located 115 or 45 km southwest of the main storm (see Table 1).

As in XTD06, radar data are assumed to be available on the elevation levels in the vertical and already interpolated to model grid columns in the horizontal. The standard WSR-88D precipitation scanning mode is assumed (see Fig. 2 of XTD06). Radial velocity (V_r) is simulated as in XTD06, and the same V_r observation operator is used in the EnKF assimilation. The reflectivity formula described in TX05 is used for simulation and assimilation. This formula returns the reflectivity (Z) in dBZ from the mixing ratios of rain, snow and hail/graupel.

Gaussian-distributed random errors with zero mean and standard deviations of 1 m s^{-1} and 3 dBZ are added to simulated V_r and Z, respectively. The 3 dBZ standard deviation of Z error is smaller than the 5 dBZ used in TX05 and XTD06, but is suggested to be more appropriate by Xue et al. (2007); the value is also used by Tong and Xue (2008a). Using the larger error value results in slightly worse analyses.

Surface stations of a hypothetical mesoscale observing network are located inside the 64 × 64 km interior subdomain with station spacings of about 20 km in the control and other directly-related experiments for a total of 9 stations. To simulate a network whose stations are not on the grid point while keeping the network more or less uniform, the stations are placed randomly within 400 × 400 m square boxes that are centered on the grid points 20 km apart. The EnKF code is general; it does not require a uniform distribution of stations.

The observed variables at these stations include the horizontal wind components, the temperature, pressure, and water vapor mixing ratio. The standard deviation of the zero-mean Gaussian errors added to the simulated surface observations are: 1 m s^{-1} for wind components, 1 K for temperature, 1 hPa for pressure and 1 g kg^{-1} for water vapor mixing ratio. In sensitivity experiments, different network densities are tested (see Table 1).

Both radar and surface observations are assumed to be available every 5 minutes. The latter is actually true with the Oklahoma Mesonet while the WSR-88D radar volume scan interval in precipitation mode is typically 5 min.

3.2.4 The EnKF algorithm

The EnKF algorithm used in this study is based on the serial ensemble square root filter (EnSRF) of Whitaker and Hamil (2002). In this study, we refer to the algorithm as EnKF in general. The implementation follows XTD06 exactly, except for the addition of surface observations, which are assimilated after the radar data (results obtained assimilating surface observations first are very similar).

The state variables analyzed include u, v, w, θ, p, q_v, q_c, q_r, q_i, q_s, and q_h. To initialize the first ensemble forecast cycle, random initial perturbations as used in XTD06 are added to a horizontally homogenous first guess defined using the May 20 sounding of the truth simulation. The random perturbations are drawn from a Gaussian distribution with zero mean and standard deviations of 2 m s^{-1} for u, v, w, 2 K for θ, and 0.6 g kg^{-1} for q_v and all microphysical variables. Perturbations for all except for the microphysical variables are added in the entire subdomain. The perturbations for the latter are added only in the region where radar echo is present at 20 min, the start time of the first assimilation cycle. Reflectivity data in both the precipitation and clear air regions (negative Z values are set to zero) are used. Radial velocity data are only used in regions where Z is greater than or equal to 10 dBZ. We note here that because the storm environment is initialized with a perfect model sounding and there is no land surface process in the model, there is no error in the storm environment except for that introduced by the initial ensemble perturbations. The benefit of the surface observations is expected only where storm-induced disturbance exists.

Covariance localization (Houtekamer and Mitchell 2001) is used to limit the spatial influence of observations and reduce sampling error. A Schur product is applied with a smooth 5th-order distance-dependent function (Eq.(4.10) of Gaspari and Cohn, 1999) multiplying the calculated background error covariance. For radar observations, a 6 km localization radius is chosen in all directions to ensure the best results when only radar data is used (XTD06). For surface observations, when the mean station spacing (station spacing hereafter for short) is 20 km, a horizontal localization radius of 30 km and a vertical radius of 6 km are found to be optimal through experimentation. When the station spacing changes from 2 km through 32 km in sensitivity experiments, different optimal horizontal localization radii ranging between 6 km and 36 km are chosen, again based on experimentation. Further discussions can be found in section 3.4.2. Forty ensemble members are used for all experiments, as in XTD06.

To avoid the filter divergence problem caused by underestimation of covariance due to small sample size and/or model error, covariance inflation is used in all experiments following the procedure of TX05. For cases using radar observations only, covariance inflation is only applied in and near (within 6 km of) the region where observed reflectivity exceeds 10 dBZ. In experiments with only radar observations, a 5% inflation factor is applied for perfect-model experiments and 15% is used for imperfect-model experiments (see Section 6). For experiments including surface observations, additional covariance inflation is applied at the lower levels within the cutoff radius of the surface observations, in the entire subdomain, using the same inflation factor (at points influenced by both radar and surface data, covariance inflation is done only once). Therefore, covariance inflation in regions not directly reachable by observations is avoided. This avoids excessive increase in the ensemble spread where the ensemble is not constrained by any observations.

3.2.5 Assimilation experiments

A complete list of experiments can be found in Table 1. The control experiment, Cntl or Ra (for radar data only), uses only data from a radar located 185 km from the domain center where the main storm is located; this case serves as the baseline. Other experiments add data from a surface network of different station spacing, and test radar distances of 185, 115 or 45 km respectively. Some experiments assimilate a particular surface observation variable. In the experiment names, ‘Ra’ denotes the use of radar data, ‘Sfc’ indicates the use of surface observations, ‘D’ followed by a number indicates the radar distance, and ‘S’ followed by a number indicates the station spacing. For example, RaSfcD115S6 means that both radar and surface data are used and the radar is located 115 km from the domain center and the station spacing is 6 km. Additional characters, such as T in RaSfcT, indicate which variables in the surface data are assimilated. The results of these experiments are reported in the next three sections, focusing on the impact of additional surface observations. The experiments with the default radar distance of 185 km and default mean station spacing of 20 km have abbreviated names. Experiments discussed in section 3.6 include model error. In all experiments, the assimilation starts at 20 min and the first EnKF analysis occurs at 25 min.

3.3 Impact of surface observations

3.3.1 Results of experiments Ra and RaSfc

The control experiment Ra (Table 3.1) is first examined, together with RaSfc. Figures 3.1 and 3.2 show the surface fields of the truth simulation and the ensemble mean analyses of experiments Ra and RaSfc, at 60 and 120 min of model time, respectively. RaSfc adds data from a surface network of 20 km mean spacing with all observed variables (u, v, T, p, and q_v) assimilated. The ensemble mean analysis fields are plotted. At 60 min, or after eight 5-min analysis cycles, precipitation and cold pool associated with the storm are established in both Ra and RaSfc in general (Fig. 3.1b, 3.1c) but Ra does not capture the cell splitting or the left moving cell at this time. The cold pool is too broad in Ra based on the -1 K perturbation potential temperature, θ', contour (Fig. 3.1b) but is much closer to the truth in RaSfc (Fig. 3.1c). Ra completely misses the surface convergence center associated with the main updraft while the divergence center underneath the precipitation core is too weak and does not have the right pattern (compare Fig. 3.1e with Fig. 3.1d). In comparison, RaSfc does a much better job capturing the convergence pattern along the gust fronts on the south and north side, and reproduces the magnitude and pattern of the main divergence center much better (Fig. 3.1f). It does significantly underestimate the convergence associated with the main updraft. In both cases, the analyzed fields are not very accurate quantitatively at this time.

Experiment	Obs. Type	Sfc. Obs. Spacing (km)	Radar Dist. (km)	Remark/Aspects tested
Ra/Cntl	Radar only	N.A.	185	Control Exp.
RaSfc	Radar+Sfc All	20	185	Cntl sfc obs spacing
RaSfcUV	Radar+Sfc V	20	185	Individual sfc obs
RaSfcT	Radar+Sfc T	20	185	Individual sfc obs
RaSfcP	Radar+Sfc Pres	20	185	Individual sfc obs
RaSfcQv	Radar+Sfc Qv	20	185	Individual sfc obs
RaSfcS32	Radar+Sfc All	32	185	Sfc obs spacing
RaSfcS16	Radar+Sfc All	16	185	Sfc obs spacing
RaSfcS12	Radar+Sfc All	12	185	Sfc obs spacing
RaSfcS6	Radar+Sfc All	6	185	Sfc obs spacing
RaSfcS4	Radar+Sfc All	4	185	Sfc obs spacing
RaSfcS2	Radar+Sfc All	2	185	Sfc obs spacing
RaD115	Radar only	N.A.	115	Radar Dist.
RaSfcD115	Radar+Sfc All	20	115	Radar Dist.
RaSfcD115S6	Radar+Sfc All	6	115	Radar Dist. and Sfc Obs spacing
RaD45	Radar only	N.A.	45	Radar Dist.
RaSfcD45	Radar+Sfc All	20	45	Radar Dist.
RaSfcD45S6	Radar+Sfc All	6	45	Radar Dist. and Sfc. Obs. spacing
RaNr0	Radar only	20	185	Imperfect model with 10×Nr_0
RaLFO	Radar only	20	185	Imperfect model with LFO84
RaSchultz	Radar only	20	185	Imperfect model with Schultz
RaWSM6	Radar only	20	185	Imperfect model with WSM6
RaSfcNr0	Radar+Sfc All	20	185	Imperfect model with 10×Nr_0
RaSfcLFO	Radar+Sfc All	20	185	Imperfect model with LFO84
RaSfcSchultz	Radar+Sfc All	20	185	Imperfect model with Schultz
RaSfcWSM6	Radar+Sfc All	20	185	Imperfect model with WSM6
RaSfcNr0	Radar+Sfc All	6	185	Imperfect model with 10×Nr_0
RaSfcS6LFO	Radar+Sfc All	6	185	Imperfect model with LFO84
RaSfcS6Schultz	Radar+Sfc All	6	185	Imperfect model with Schultz
RaSfcS6WSM6	Radar+Sfc All	6	185	Imperfect model with WSM6

Table 3.1: List of OSSE experiments examining the impact of surface observation data of different spacings and for different radar distances under the scenario of perfect and imperfect model.

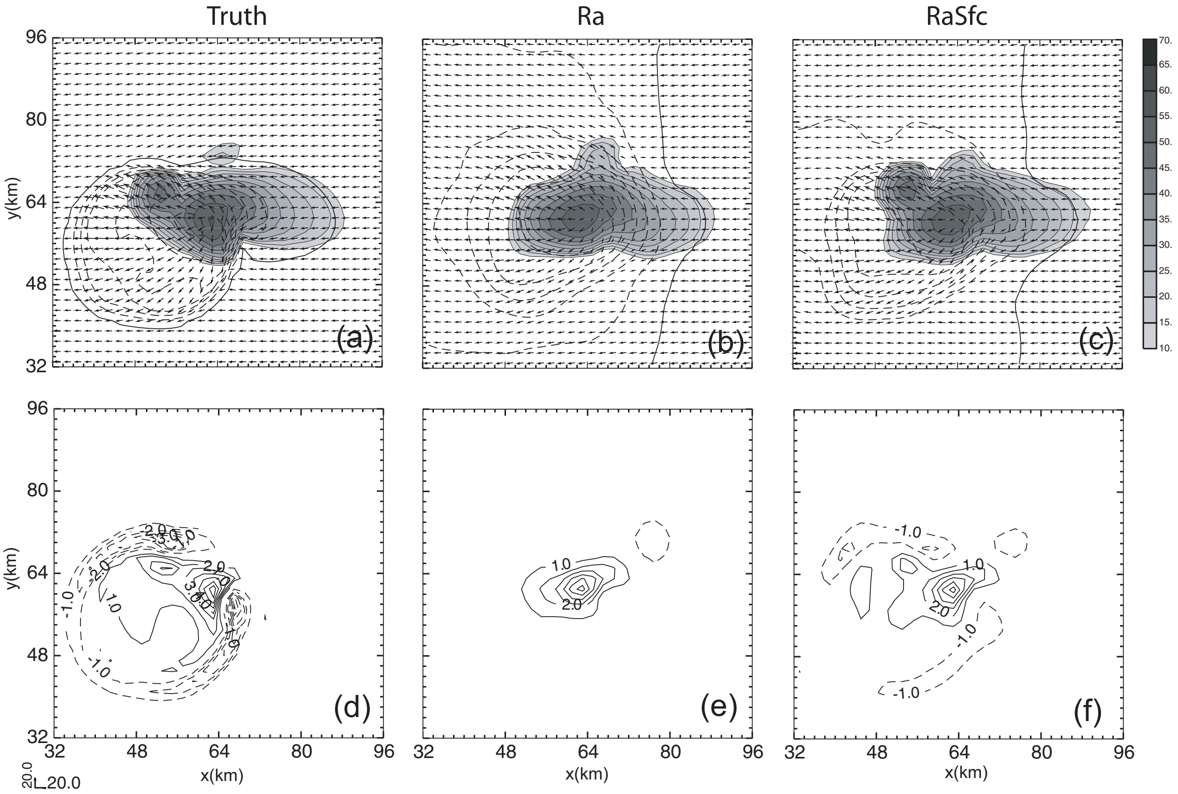

Fig. 3.1. Perturbation wind vectors, simulated reflectivity Z (dBZ) and perturbation potential temperature θ' (K) (upper panel), and divergence fields (× 1000 s^{-1}) (lower panel) at z = 100 m (first model level above ground) for truth (a, d), and experiments Ra (b, e) and RaSfc (c, f) at 60 min of model time.

At 90 min (not shown), the right and left moving cells as seen from the low-level *Z* are now reasonably well captured in both cases. However, the convergence field is still not accurately analyzed, especially along the gust fronts and in Ra. By 120 min (Fig. 3.2), the hook echo structure of the major storm is reproduced well in both Ra and RaSfc (Fig. 3.2b,c) but we see more differences in the analyzed surface divergence fields. Both the intensity and pattern of the divergence in RaSfc are reproduced accurately (Fig. 3.2f) but the divergence in Ra is generally broader and weaker. The tail of the divergence band at the southwest end of the gust front is missing in Ra. Also mostly missing is the convergence center associated with the left mover near the northwest corner of the plotted domain (Fig. 3.2e).

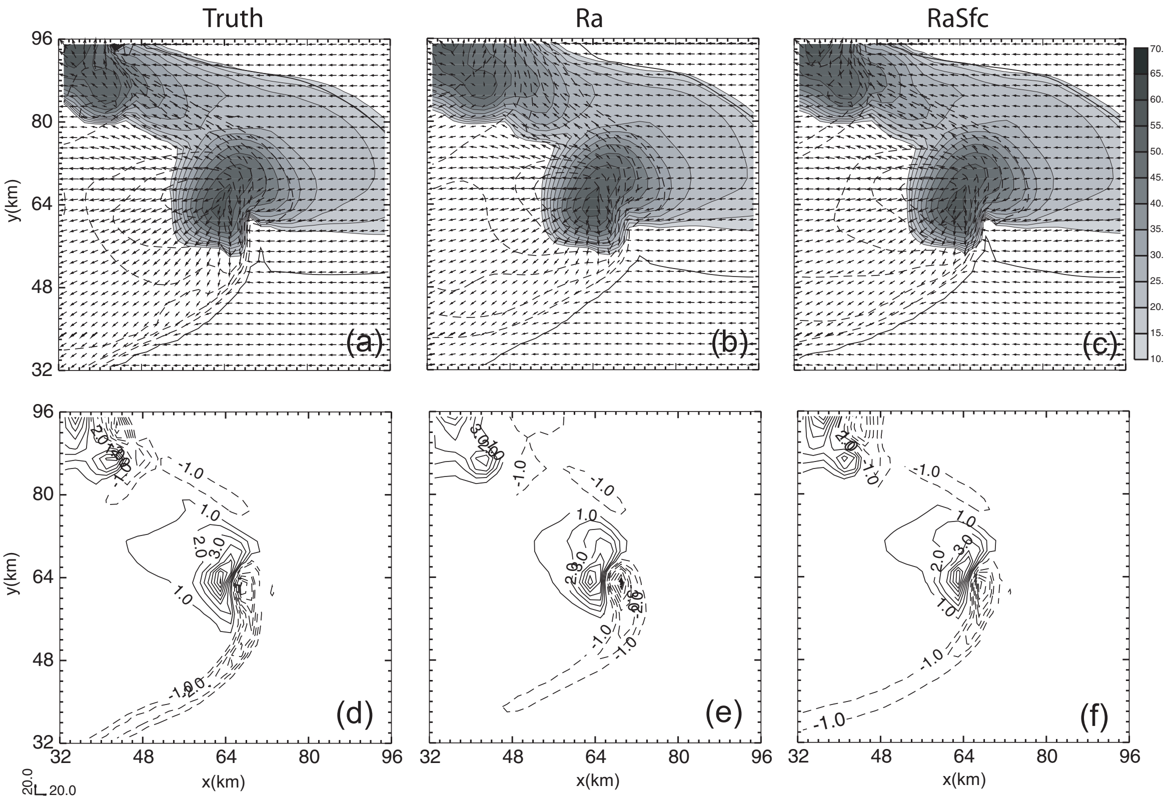

Fig. 3.2. As Fig. 3.1, but at 120 min.

There are also clear differences between the analyses of gust front. The temperature gradients along the gust fronts are significantly weaker in Ra (Fig. 3.2b) than in RaSfc (Fig. 3.2c) and there is also larger position error of the gust fronts in Ra, especially near the western and northern domain boundaries. In general, the low-level analysis of RaSfc can be considered very good at this time.

The subjective comparison between Ra and RaSfc indicates that the surface observations have a noticeable positive impact on the analysis of low-level features, including the cold pool and associated flows, even though the network has a much coarser resolution (~20 km) than would be needed to resolve the sharp gradient along the gust fronts. The impact on the precipitation field seems smaller, presumably because of that the hydrometeors that descend to the ground level are well captured by the radar data above. Without surface data, we have to rely on the model alone to establish the cold pool through microphysical and dynamic processes. As the cold air spreads away from the precipitation region, no more radar data (except for the clear air information contained in the zero reflectivity values) is available.

The root mean square (*rms*) error of the ensemble mean analysis calculated against the truth is used to further quantify EnKF performance and the impact of surface data. The evolution of *rms* error with time is shown in Fig. 3.3 for Ra and RaSfc. As in TX05, the errors were calculated at grid points where observed Z exceeds 10 dBZ. The errors decrease during the first few cycles rapidly, from the initially very high level associated with the poor sounding-based initial guess. Errors starting from 40 min, or after 4 analysis cycles, are shown here. The relative error ratio (RER) averaged over the last 10 analysis cycles ending at 120 min is also calculated to evaluate the impact of surface observations. It is defined as

$$\mathrm{RER} = \frac{1}{10}\sum_{i=n}^{n-9}\frac{E^{i}_{Ra+Sfc}}{E^{i}_{Ra}}, \tag{3.1}$$

where $E^{i}{}_{Ra+Sfc}$ is the analysis *rms* error of a given state variable at the i^{th} cycle when using both radar and surface observations and $E^{i}{}_{Ra}$ denotes the corresponding error using radar data only. The averaging over the last ten analysis cycles removes temporal fluctuations and provides a more reliable measure of the analysis accuracy. The RERs of various state variables can be combined further to form RERs for several major categories, such as that for the wind, which averages the RERs of u, v and w, and that for microphysical variables, which averages the RERs of 5 microphysics variables. The total RER (TRER) averages the RERs of all 11 state variables and is used as the main indicator quantifying the improvement by surface observations. The RERs or TRERs of all experiments are summarized in Table 3.2. A similarly defined error ratio has been used in Xu et al. (2008).

At the end of assimilation, or 120 min, the analysis errors in most fields are clearly lower in RaSfc than in Ra (Fig. 3.3). The percent improvement due to surface data is largest in u and θ, with the errors in RaSfc being only about 50-60% those of Ra. The absolute analysis *rms* error reduction in RaSfc is 0.8 ms^{-1} for u and 0.27 K for θ. Other variables also display improvement to various extents. The *rms* error in RaSfc is 65% of Ra for v, 61% for w, 73% for p, and 60% for q_v. Even for microphysical variables which are not directly observed by surface observations, significant improvements are also found: the *rms* error is 60% of Ra for q_c, 62% for q_r, 48% for q_i, 62% for q_s and 54% for q_h. This amounts to a 30% to 50% reduction in errors in various model fields due to the assimilation of surface observations. We note that there are temporary increases in the *rms* errors near 65 min; this also occurs in our earlier studies for this storm also (TX05 and XTD06) and we believe it is related to the cell splitting at this time. The TRER of RaSfc also shows evident improvement in the last 10 analysis cycles (Table 3.2). The reduction in the total relative error (1 - TRER) is 32% with the help of surface observations.

Experiment	*uvw*	Micro	θ	q_v	Total
RaSfcS32	0.85	0.83	0.78	0.80	0.83
RaSfc(S20)	0.72	0.67	0.67	0.67	0.68
RaSfcS16	0.66	0.64	0.62	0.61	0.64
RaSfcS12	0.61	0.62	0.55	0.56	0.60
RaSfcS6	0.50	0.55	0.48	0.43	0.51
RaSfcS4	0.47	0.53	0.45	0.39	0.48
RaSfcS2	0.47	0.56	0.45	0.38	0.49
RaSfcUV	0.80	0.76	0.73	0.76	0.77
RaSfcT	0.84	0.75	0.80	0.86	0.80
RaSfcQv	0.86	0.79	0.77	0.80	0.81
RaSfcP	0.93	0.90	0.91	0.93	0.91
RaSfcD115	0.71	0.71	0.72	0.68	0.70
RaSfcD115S6	0.52	0.61	0.56	0.50	0.56
RaSfcD45	0.97	1.04	0.96	0.95	0.99
RaSfcD45S6	0.85	0.96	0.89	0.82	0.89
RaSfcNr0	0.75	0.88	0.67	0.72	0.80
RaSfcNr0S6	0.47	0.76	0.48	0.44	0.60
RaSfcLFO	0.95	1.01	0.96	0.80	0.96
RaSfcLFOS6	0.70	1.00	0.84	0.55	0.84
RaSfcStz	0.88	1.03	0.91	0.79	0.94
RaSfcStzS6	0.76	1.01	0.85	0.61	0.87
RaSfcWSM6	0.92	0.98	0.96	0.86	0.94
RaSfcWSM6S6	0.78	0.95	0.91	0.62	0.86
RaSfcMulti	0.81	0.95	0.85	0.77	0.88
RaSfcMultiS6	0.58	0.95	0.68	0.49	0.75
RaNr0	1.19	1.19	1.56	1.49	1.27
RaLFO	1.27	2.47	1.25	1.40	1.81
RaSchultz	1.37	4.36	1.47	1.33	2.71
RaWSM6	1.95	4.17	1.95	1.56	2.91
RaMulti	1.35	2.95	1.27	1.33	2.04

Table 3.2: Relative *rms* error ratio (RER) of *u, v* and *w,* for microphysical variables, θ, q_v and all variables (total) for listed experiments. The errors are relative to experiment Ra, except for those of different radar distance where the error is relative to the radar-data-only experiment of the same distance.

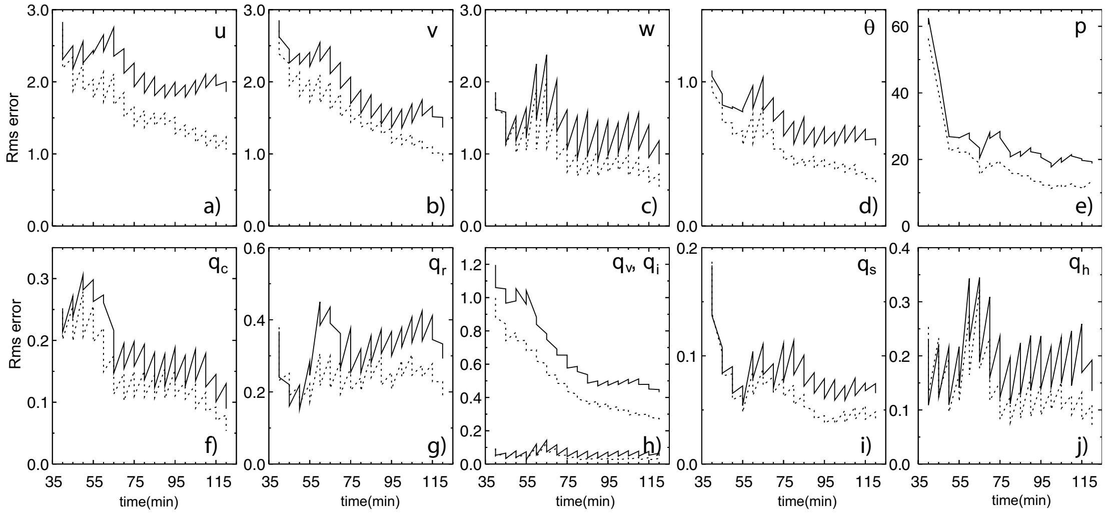

Fig. 3.3. The *rms* error of ensemble mean forecasts and analyses plotted against time for Ra (solid) and RaSfc (dotted) for a) u (m s^{-1}), b) v (m s^{-1}), c) w (m s^{-1}), d) potential temperature θ (K), e) pressure p (Pa), f) q_c ($\mathrm{g\,kg^{-1}}$), g) q_r ($\mathrm{g\,kg^{-1}}$), h) (upper curves) q_v and (lower curves) q_i ($\mathrm{g\,kg^{-1}}$), i) q_s ($\mathrm{g\,kg^{-1}}$), and j) q_h ($\mathrm{g\,kg^{-1}}$). The sharp reductions in the error at the analysis times are due to analysis updates. See TX05 for further explanations of this type of plots.

To examine the impact of surface data at different levels, vertical profiles of *rms* errors obtained by horizontal averaging over the precipitation regions are plotted in Fig. 3.4 for 120 min. The *rms* errors for u, v, θ, p, q_r, and q_v are largest at the low levels and near the surface in Ra, primarily because of the lack of low-level radar data. The largest improvements in RaSfc for these variables are also found at the surface. For example, the *rms* error reductions are 2.6 ms^{-1} for u, 0.8 K for θ and 12 Pa for p at the surface, reflecting the better analysis of surface cold pool and the associated fields with the use of surface observations. For w and all four microphysical variables shown (q_c, q_r, q_s and q_h), the largest *rms* error reduction is generally found at the levels where the corresponding Ra errors are largest. For example, q_h error is reduced from about 0.28 g kg^{-1} to about 0.14 g kg^{-1} near the 4 km level, and the maximum reduction in q_s error, about 0.08 g kg^{-1}, occurs close to 12 km. The error reduction in w is between 0.3 and 0.7 m s^{-1} between the levels of 2 and 13 km. This general improvement in the storm analysis through the addition of surface data, even at the upper levels away from the ground, is very encouraging.

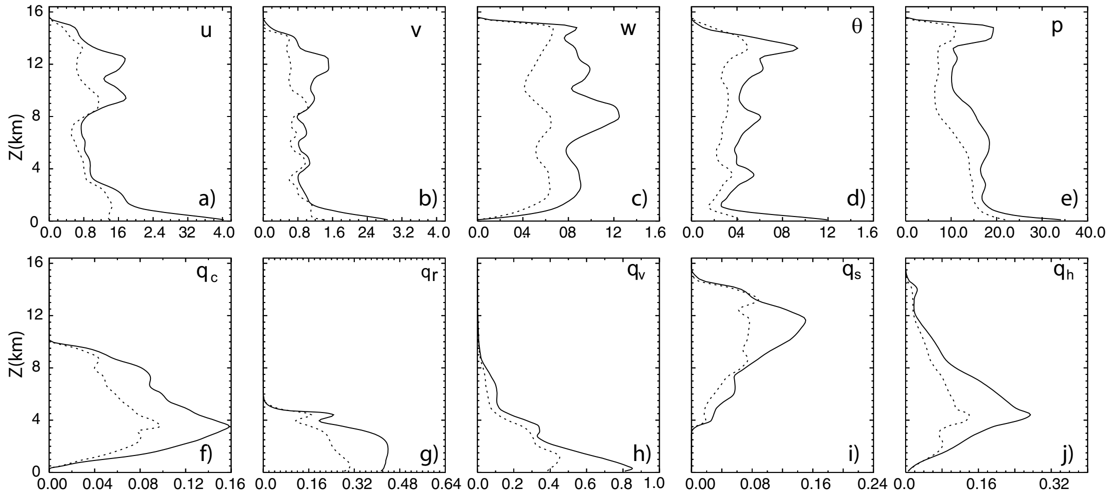

Fig. 3.4. The *rms* error profiles of the ensemble mean analyses of Ra (solid) and RaSfc (dotted), for a) u (m s^{-1}), b) v (m s^{-1}), c) w (m s^{-1}), d) θ (K), e) p (Pa), f) $q_c\,(\mathrm{g\,kg^{-1}})$, g) q_r $(\mathrm{g\,kg^{-1}})$, h) q_v $(\mathrm{g\,kg^{-1}})$ and q_i $(\mathrm{g\,kg^{-1}})$, i) $q_s\,(\mathrm{g\,kg^{-1}})$, and j) $q_h\,(\mathrm{g\,kg^{-1}})$ at 120 min.

3.3.2 Background error correlation structure

To gain insight on how the surface data affect the model state, we examine a few ensemble-derived background error correlation (ρ) fields. The background error covariance estimated from the ensemble members determines how observation information is spread in space and helps improve the analysis of variables not directly observed. Accurate error correlations should reflect the physical structure of the storm.

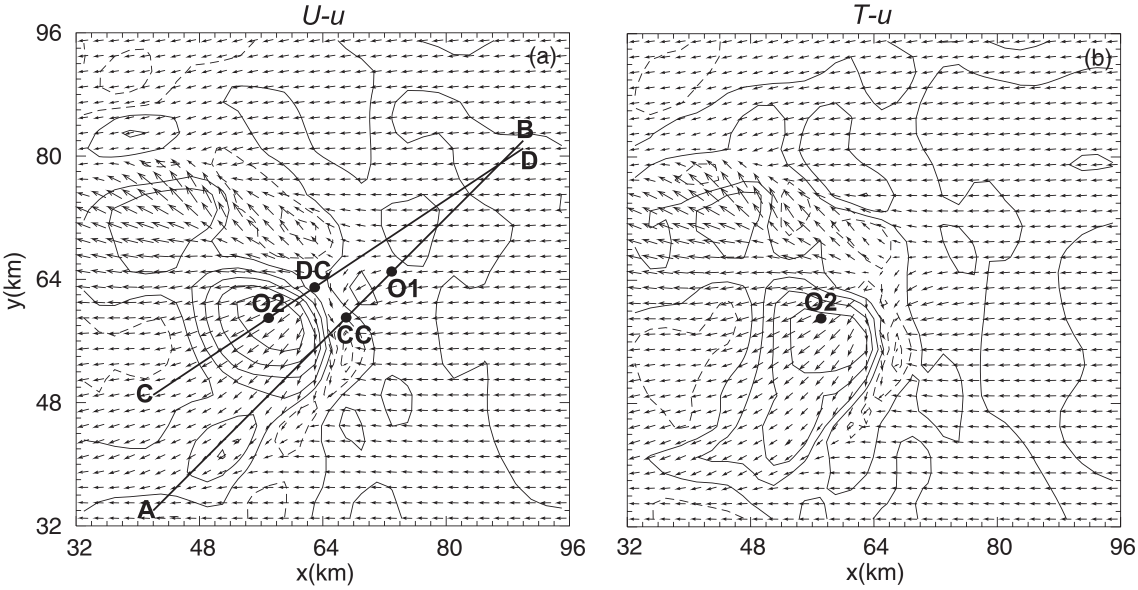

Fig. 3.5. The surface truth wind vectors, together with forecast error correlation coefficients estimated from the forecast ensemble at 75 min for experiment RaSfc. Error correlation between (a) surface wind observation U (wind along C-D line) at station O2 (x = 57 km, y = 59 km) and u at the grid points, and (b) between surface temperature at station O2 and u. Solid (dashed) contours represent positive (negative) correlations at intervals 0.2; zero contours are omitted. CC and DC in (a) mark the low-level convergence and divergence centers, respectively, and lines A-B and C-D indicate location of the vertical cross sections shown in Fig. 3.6.

We choose two locations of possible surface observations at 75 min (Fig. 3.5) to calculate the spatial forecast error correlations. The first location, O1, is in the inflow region of the storm, at (73, 65) km, and the second location, O2, is in the outflow region within the cold pool, at (57, 59) km (Fig. 3.5a). The surface inflow through O1 feeds the convergence center marked as CC and the outflow through O2 originates from the surface divergence center marked as DC in Fig. 3.5a. As in TX05, we calculate the correlation coefficients between a variable at a given observation location, and another (possibly different) variable at all other grid points. When these two variables are the same, we are calculating the spatial auto-correlation coefficient, and when they are different, we obtain a spatial cross-correlation. Such correlation information, after being localized, is used directly in the EnKF algorithm for analysis update.

We choose two vertical cross sections, along lines A-B and C-D, as shown in Fig. 3.5a. The horizontal wind vector is projected onto each of these two lines (positive towards B and D, respectively), and we calculate the correlation between this wind component (referred to as U) with other state variables.

We first present the ρ field at the surface, between U at location O2 (wind along line C-D) and u at the grid points (Fig. 3.5a), and between surface temperature at O2 and u (Fig. 3.5b). Positive U-u correlations are found to occupy a large part of the cold pool outflow region and extend out to 20 to 30 km from O2 (Fig. 3.5a), and more so in the generally westward direction. This positive correlation indicates that when U is smaller (more negative) at O2, u is smaller (more negative), which is physically reasonable because these U and u are located within the same outflow region; a stronger outflow observed at O2 suggests stronger outflows at locations with positive correlations. To the northeast of O2, correlation contours are more closely packed, and the correlation coefficient becomes negative immediately beyond the divergence center (but before going into the inflow region). This demonstrates that the outflow on the other side of DC is also stronger when that on this side of divergence center near O2 is stronger.

The correlation between surface temperature T at O2 and u shows a similar pattern as that between U and u (Fig. 3.5b), which shows that a stronger colder cold pool corresponds to stronger outflows, with more negative u west of DC and less positive u east of DC. When such cross-correlation is estimated with reasonable accuracy, surface temperature observations can be used effectively to update the flow field, and vice versa.

The generally well-defined correlation patterns up to 30 km from the observation location suggest that the localization radius of 30 km used for the current surface network is appropriate.

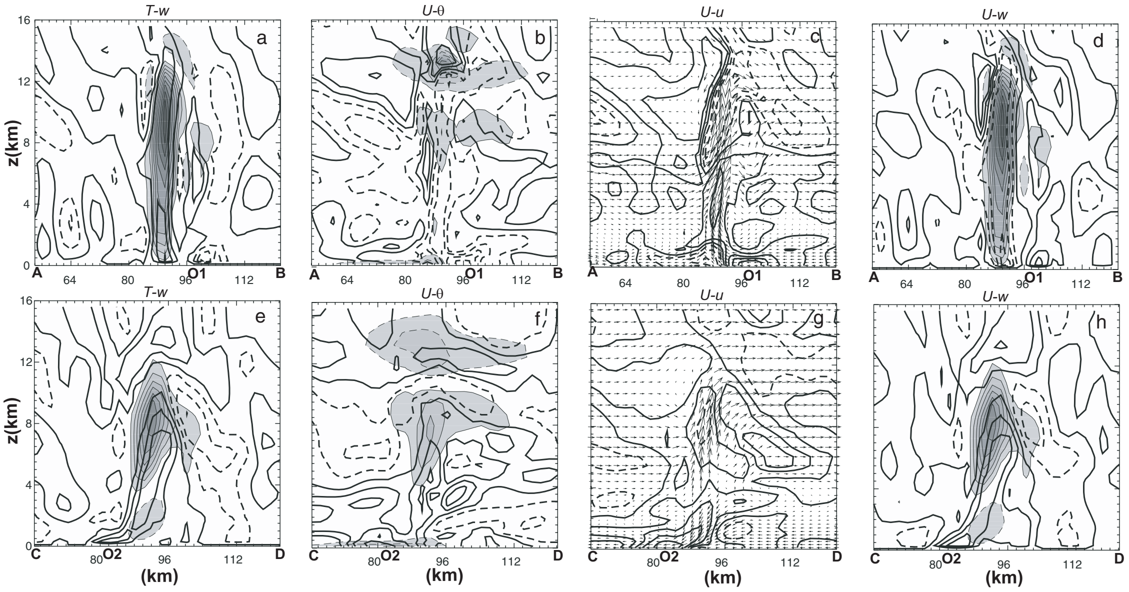

Fig. 3.6. Forecast error correlation coefficients estimated from the forecast ensemble at 75 min for experiment RaSfc, in the vertical cross section along line A-B (upper panels) and along C-D (lower panels) in Fig. 3.5. Error correlations (upper panels) between (a) surface temperature T at O1 (x = 73 km, y = 65 km) and w, (b) surface wind U at station O1 and θ, (c) U and u, (d) U and w. Error correlations (lower panels) between, (e) surface T at station O2 (x = 57 km, y = 59 km) and w, (f) U at O2 and θ, g) U and u, and (h) U and w. Thick solid (dashed) contours represent positive (negative) correlations at intervals 0.2. Shaded with thin contours shows the truth field of w in (a), (d), (e) and (h) with interval 2.5 m s^{-1}, and perturbation θ' in (b) and (f) with interval 2 K. Wind vectors in (c) and (g) show the truth perturbation wind field.

In the upper panels of Fig. 3.6, the correlations between surface observations at O1 and other state variables in the vertical cross-section along line A-B, through O1 and CC, are shown (in contours), together with the corresponding state variable fields (shaded contours). For the correlation between T at O1 and vertical velocity w, the most prominent feature is a deep column of positive correlation coinciding with the main updraft (Fig. 3.6a), suggesting the presence of a stronger updraft when surface T in the low-level inflow region is higher. A similar pattern is found between U at O1 and w, suggesting a stronger updraft when surface inflow towards CC is stronger (Fig. 3.6d). Again, this is physically consistent with the expected storm dynamics. The negative correlation with potential temperature, θ, in the updraft region indicates enhancement of the warm core structure by low-level inflow (Fig. 3.6b). The correlation of U with u is negative in the lower part of the troposphere and positive above (Fig. 3.6c), suggesting that a stronger easterly component at the surface contributes to more westward tilt of the updraft at the low levels while at the upper levels this enhanced easterly component may turn into an enhanced westerly component as the updraft air parcels rise and overturn.

The corresponding correlation fields in the C-D cross section are plotted in the lower panels of Fig. 3.6. In this case, the cross section passes through the surface divergence center (DC) with the observation located inside the cold pool (O2). Again, we find similar correlation patterns between surface T and U with w (Fig. 3.6e and Fig. 3.6h). The strong positive correlations between U, T in the downdraft outflow region and w in the updraft region, extending all the way from the surface to about 9 km, indicate a stronger downdraft (or weaker updraft) when temperature at O2 inside the cold pool is lower, or when the outflow near the surface is stronger. The positive correlation between U and θ in the downdraft region (Fig. 3.6f) also reflects the relation between outflow strength and the downdraft temperature. The correlation between U and u in this cross section changes sign at the middle troposphere (Fig. 3.6g), similar to what is observed in the A-B cross section (Fig. 3.6c).

The physically meaningful and dynamically consistent correlations of surface temperature and wind with the state variables presented here, taken in both the inflow and outflow regions, indicate that the ensemble system is able to properly estimate the spatial auto- and cross-correlations. Such flow-dependent error correlation information is valuable for optimally utilizing surface observations. The 6 km covariance localization radius that we use in the vertical allows the surface observations to directly influence state variables up to 6 km above the surface. Of course, the impact of surface observations is not limited to the lowest 6 km (see Fig. 3.4), because of information propagation in time and space by the prediction model.

3.4. Sensitivity experiments

In this section, we will further examine the impact of different observation variables or types, and the impact of surface data for different network densities and/or radar distances.

3.4.1 Impact of surface measurement type

In this sub-section, the surface measurement types or variables, including wind, temperature, pressure and water vapor, are assessed individually for their impact. The experiments are listed as RaSfcUV, RaSfcT, RaSfcP and RaSfcQv in Table 1, and are the same as RaSfc except that only the listed surface observation variables are included.

The relative analysis error plots for these four experiments are shown in Fig. 3.7. The relative error is defined as the ratio of the *rms* error of the individual sensitivity experiment to that of the chosen reference, which in this case is Ra (c.f., Fig. 3.3). The plots show that the surface wind observations have the largest impact while pressure observations have the least impact; the relative *rms* errors of RaSfcP are close to 1 at the end of assimilation cycles. In this case, the RERs or TRER are 93% to 90% (Table 3.2), indicating only 7% to 10% error reduction when only additional pressure observations are assimilated. For microphysical variables, temperature observations result in the smallest RER (Table 3.2). None of these individual measurements were able to produce as large an impact as RaSfc which assimilates all surface variables; it achieves a TRER of 68%.

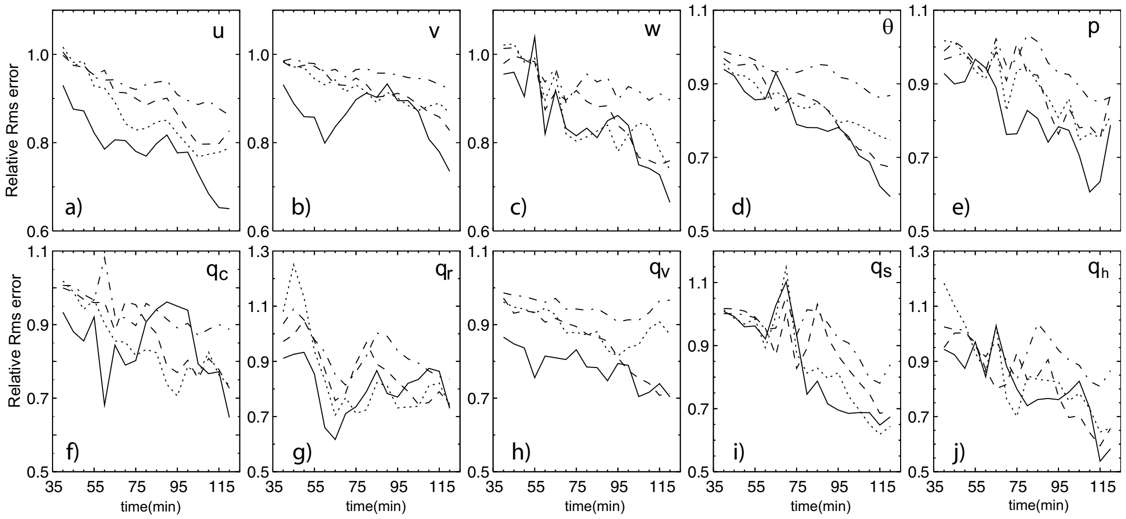

Fig. 3.7. Analysis *rms* errors relative to those of Ra for RaSfcUV (solid), RaSfcT (dotted), RaSfcP (dash-dotted) and RaSfcQv (dashed), as a function of analysis time.

It is noted that the largest improvements for state variables θ, p and q_v are also obtained when assimilating surface winds rather than direct measurements of these variables themselves. This seems to be at least partly because *rms* errors for θ, p and q_v in Ra are already much smaller than the standard deviations of the corresponding surface observation errors (Fig. 3.3), which are 1 K for T, 100 Pa for p and 1 g kg^{-1} for q_v. Such error-containing observations have a limited ability to further reduce the errors of the corresponding state variables (however, keep in mind that local errors in θ, p and q_v, especially underneath the active storms, are larger than those of precipitation-domain mean shown in Fig. 3.3, so the benefit of analyzing surface data should still be achievable). For u and v, the analysis errors from Ra are still relatively large (Fig. 3.3) compared to the surface wind observation errors; therefore there is more room for direct positive impact by the wind observations. When wind observations are improved, other variables benefit too. In general, θ benefits most from surface observations (73% RER) among all state variables; this is consistent with the fact the most important features at the surface are related to the cold pool.

That the impact of surface observations is largest when all of the measured variables are assimilated is expected, since in this case every individual measurement shows benefit. The total impact using all measurements together is smaller than the sum of individual impacts, in terms of the percentage error reduction.

Zhang et al. (2004) show that liquid-water potential temperature observations have a larger impact than winds in retrieving cold pool. Since liquid-water potential temperature is typically not measured, we cannot directly compare our results with theirs. Simplistic radar data coverage assumption, the use of V_r data only, and simple model microphysics used there further hinder a direct comparison.

In summary, the surface wind observations are found to have the largest impact on the analysis of wind and temperature fields for the chosen supercell and for the typical observation errors assumed. The improvement in analyzed cold pool temperature comes not only from thermodynamic observations, but even more from surface wind observations, which directly improve the analysis of convergence and divergence patterns that are closely linked to updraft and downdraft intensities as well as gust front positions. Such inter-connections are reflected in the background error covariance derived from the forecast ensemble, which drives the multivariate EnKF analysis. Wind observations are used to update all state variables, not just wind components themselves.

3.4.2 Varying surface network spacing

A series of experiments are conducted to examine the impact of surface network spacing on the analysis. Additional experiments with network spacings ranging from 32 km to 2 km are performed (Tables 3.1 and 3.2). In these experiments, surface observations are still approximately uniformly distributed in the central subdomain and there are 4, 16, 25, 100, 256 and 961 surface stations in RaSfcS32, RaSfcS16, RaSfcS12, RaSfcS6, RaSfcS4 and RaSfcS2, respectively.

For each network spacing, experiments are conducted to determine the optimal horizontal covariance localization radius. The optimal localization radius is plotted in Fig. 3.8a against the network spacing. As the spacing decreases, the optimal radius also decreases. When a denser network is employed to resolve smaller-scale features, reducing the radius has the effect of keeping the influence of observations more localized. The sampling error in the ensemble filter can also be decreased when a smaller localization radius is used (Anderson et al. 2005).

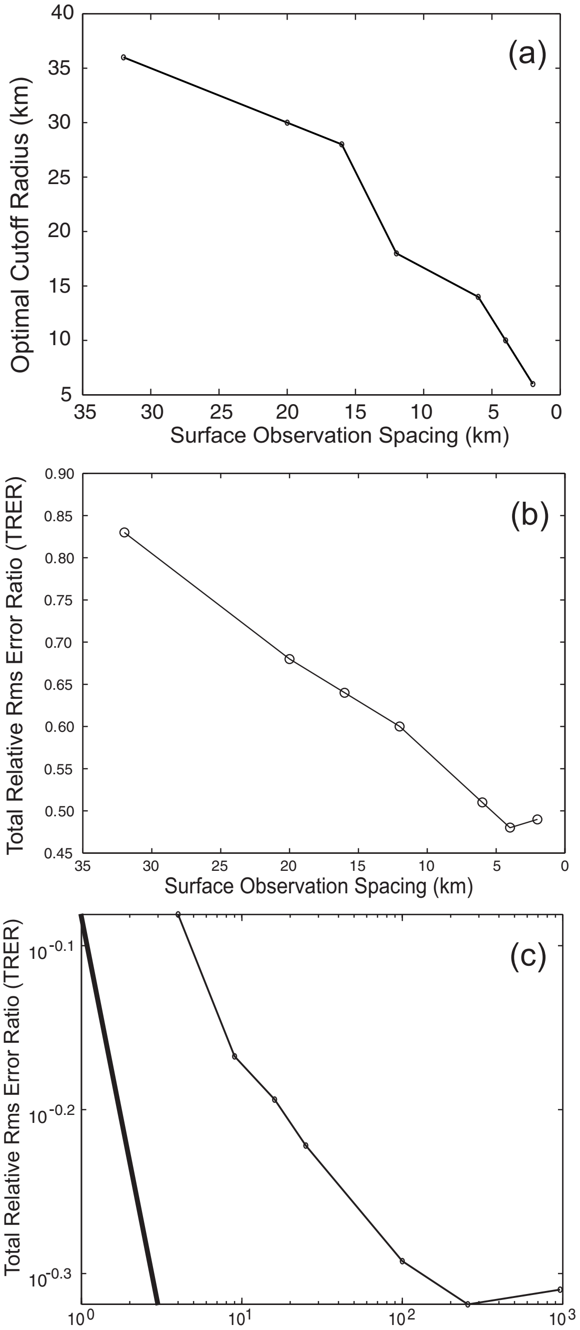

Fig. 3.8. (a) The optimal horizontal covariance localization radius for surface data as a function of mean surface network spacing, (b) the total *rms* error ratio (TRER) as a function of the mean surface network spacing, and (c) the TRER as a function of the number of surface observations, plotted in a logarithmic space. The thick straight line in (c) represent the -1/2 slope.

We first examine the cases of medium and high-density networks of 12 and 6 km spacing. In RaSfcS12, for u, v, θ', p', q_v, q_c, q_r and q_s, the relative analysis *rms* errors decrease faster and reach lower levels than in RaSfc by 90 min (Fig. 3.9). After 90 min, as the split left-moving cell propagates quickly towards the northwest corner of the subdomain, the *rms* errors start to grow in RaSfcS12 and become close to those of RaSfc for some variables at the end of the assimilation period; that is when the number of surface observations (which are confined to the subdomain) covering the split cell decreases.

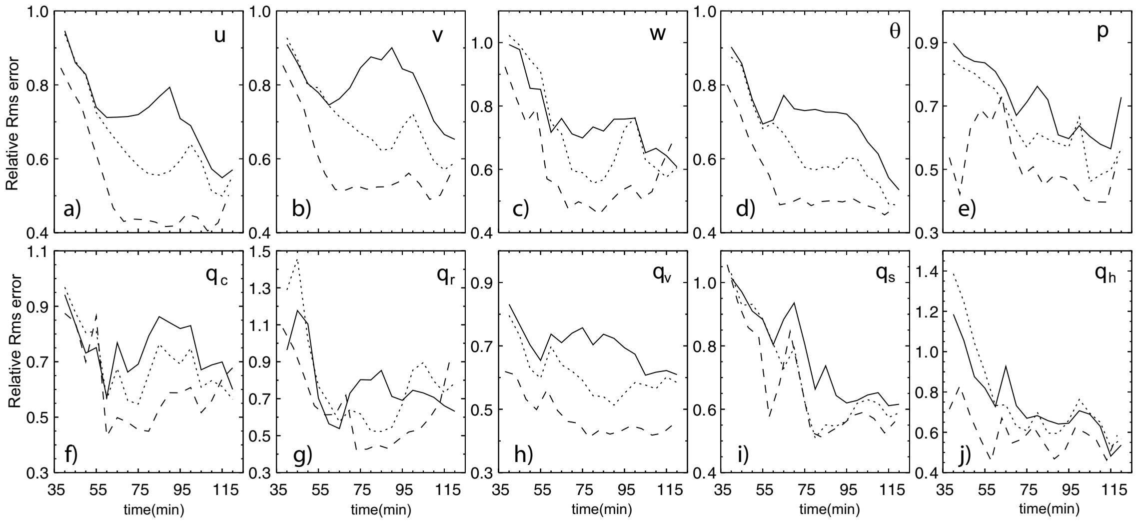

Fig. 3.9. As Fig. 3.7 but for RaSfc (solid), RaSfcS12 (dotted) and RaSfcS6 (dashed). All are errors relative to those of Ra.

When the number of surface stations is increased to 100, with a 6 km spacing, in RaSfcS6, the *rms* errors of the variables directly observed, i.e., of u, v, θ, p, and q_v, remain consistently lower than in RaSfc and RaSfcS12. For q_c, q_h, q_r and w, this smaller 6-km spacing leads to an improvement at most of the analysis times but a slight degradation compared to the coarser-resolution cases in the last couple of cycles. Even so, the improvement over Ra is still significant at the end of last cycle, with 10% to 45% error reductions. The degradation of analysis with increasing network density in some later cycles for some indirectly observed variables reflects the nonlinear nature of the storm-scale data assimilation problem. Relative to the generally large improvement in other variables and during other cycles, the amount and extent of the degradation are small.

The TRERs of the above experiments are plotted in Fig. 3.8b against the network spacing. It is interesting that the TRER decreases roughly linearly with the decrease of network spacing, or with the increase of network resolution, until the spacing is close to the grid interval of truth simulation. Given that the number of surface observations is much smaller than the number of radar observations, the number of which is kept fixed, such a strong dependency of analysis accuracy on surface network density is very interesting; it suggests increasingly large impact of surface networks on the analysis of convective-scale features as the network density increases.

Since Fig. 3.8b shows that TRER $\propto d$, where d is the mean network spacing, then TRER $\propto A\ n^{-1/2}$ and $\log(\mathrm{TRER}) \propto -1/\ 2\log(n)$, where A is the area covered by the network and n is the number of observation stations. Morss et al. (2001) point out that such a power-law behavior is expected of a network whose number of observations is much greater than the number of degrees of freedom in the analysis. This condition is obviously not met in our case, even if we include radar observations.

In Fig. 3.8c, we plot the TRERs against the number of surface observations in logarithmic space where the thick straight line represents a slope of -1/2. It can be seen that for relatively coarse resolutions, the error reduction rate is close to the -1/2 power law and for higher resolutions, the rate of error reduction is slower than the power law suggests. Morss et al. (2001) and Anderson et al. (2005) also found steep error reduction when the observations are relatively sparse (but not too sparse); for relatively dense network, the error reduction is close to or somewhat shallower than the power law suggests. While in our case, the error curve is never steeper than what the -1/2 power law suggests, the general trend is consistent with those two studies. Our case is more complicated because of the presence of much denser radar observations all the time.

The above experiments clearly demonstrate the benefit of higher surface network densities for storm-scale analysis. For real storms that may contain even more small-scale structures than the simulated truth storm does, more benefit is potentially achievable with higher-density networks.

3.4.3 Varying radar distance

In the next set of experiments, the radar is moved closer to the storm, at a distance of 115 km in RaD115, RaSfcD115 and RaSfcD115S6, and a distance of 45 km in RaSfcD45 and RaSfcD45S6 (see Table 3.1). Fig. 3.10 shows the analysis *rms* errors of RaD115 relative to those of Ra, also those of RaSfcD115 and RaSfcD115S6 relative to those of RaD115. With the radar at a closer distance of 115 km, RaD115 is consistently able to produce better analyses than Ra, especially for microphysical variables (which are directly linked to Z measurements). The final levels of error are lower for all variables, and are even smaller during the intermediate cycles. Relative error reductions reach 10% - 20% compared to Ra for most variables, with the largest error reduction around 75% at 65 min for q_r. In fact, the error peaks found in many variables in Ra (Fig. 3.3) and other experiments of the same radar distance (not directly shown) at around 60 min are mostly gone, resulting in significantly smaller relative errors around this time (Fig. 3.10). The closer location of radar provides more low-level data coverage and better vertical resolution at the location of storm. When the radar is located at 115 km from the main storm, the impact of additional surface observations taken at 20 km mean spacing is smaller during the early assimilation cycles (RaSfcD115 v.s. RaD115), compared to the 185 km case (RaSfc v.s. Ra), but the impact increases during later cycles (compare the solid curves in Fig. 3.10 and Fig. 3.9 noting the difference in vertical axis scales), partly because of the error increase in the case using no surface data (RaD115). The TRER of 70% is slightly larger but still close to that of RaSfc (68%).

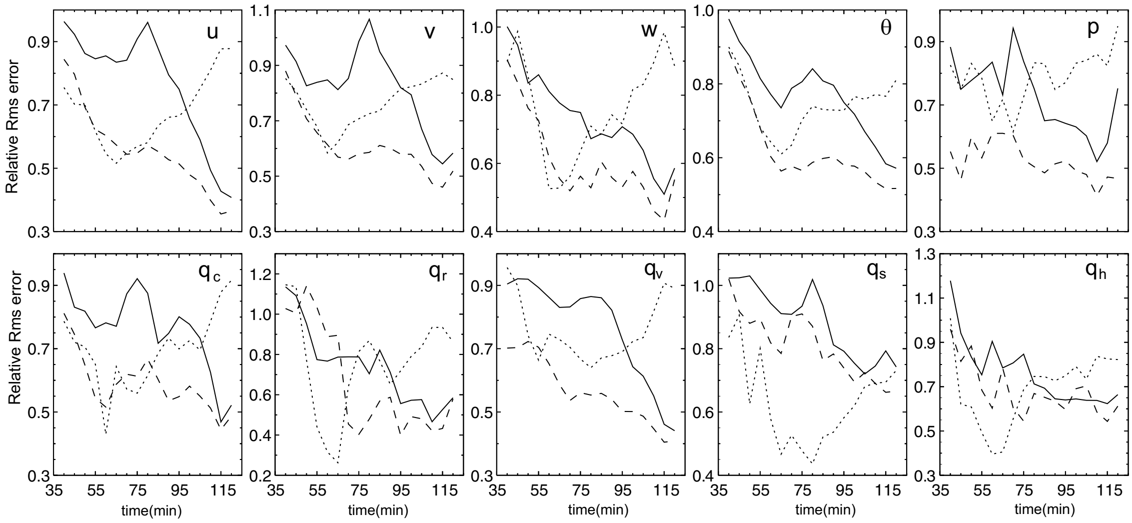

Fig. 3.10. Analysis *rms* errors of RaD115 relative to those of Ra (dotted), of RaSfcD115 relative to those RaD115 (solid), and of RaSfcD115S6 relative to those of RaD115 (dashed).

RaSfcD115S6, with a 6-km network spacing, shows a much more pronounced impact of surface data, starting from the early cycles (Fig. 3.10). After 90 min, both 12 and 6 km networks provide significant positive impacts, with the 6 km network providing the most. Unlike the 20-km spacing case (see section 3.4b), we observe no degradation in the analyses of microphysical variables as the spacing decreases in this case.

When the radar is located much closer to the storm center, at a distance of about 45 km, radar observations alone give very good analyses; there is not much room left for further improvement by surface observations of 20 km spacing, with the achieved TRER being 99%. The assimilation of surface observations at a spacing of 6 km in RaSfcD45S6 improves the results over those of RaD45, with an error reduction of 11% (TRER = 0.89, Table 3.2).

3.5 Impact of surface observations on forecast

To examine the impact of surface observations on the subsequent forecast, we perform one-hour forecasts from the ensemble mean analyses of Ra, RaSfc and RaSfcS6, at 60, 90 and 120 min, respectively. We plot the average relative error ratios for groups of state variables between the forecasts with and without surface data, where the station spacing is either 20 or 6 km. The groups of variables are wind components u, v and w, potential temperature and pressure, and microphysical variables.

Fig. 3.11 shows that when the forecasts start at 60 min of model time and when the station spacing is 20 km, the RER (thick black lines) of the wind components grows quickly from the initial value of close to 0.7 to about 1 at 100 min and exceeds 1 after 100 min. When the station spacing is 6 km, the corresponding RER (thick dashed lines) grows from about 0.56 at 60 min to about 0.75 by 90 min and remains about level afterwards. The RERs for the microphysical variable have very similar trends (Fig. 3.11c) while those for θ and p show that the benefit of surface data is sustained for the entire period for the 20 km case and even increases with time in the 6 km case (Fig. 3.11b).

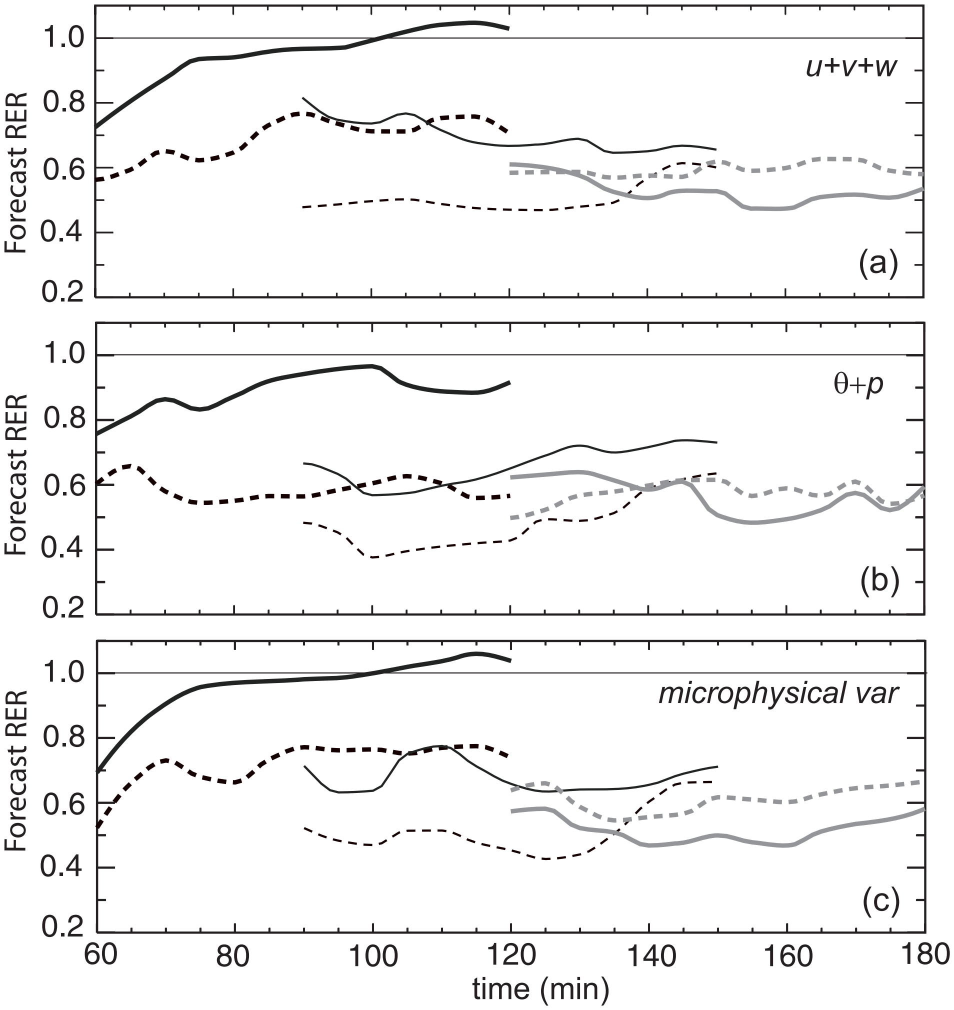

Fig. 3.11. The average relative *rms* error ratios of the *u*, *v* and *w* components (upper panel), θ and *p* (middle panel), and moisture and microphysical variables (lower panel), for 60-min-long forecasts starting from ensemble-mean analyses at 60 min (thick black lines), 90 min (thin black lines), and 120 min (thick gray lines). The solid curves are for the forecast errors starting from the analyses of RaSfc (with 20 km station spacing) relative to the corresponding errors of Ra (radar only), and the dashed lines are for the forecast errors of RaSfcS6 (6 km station spacing) relative to those of Ra.

Fig. 3.12 shows that the 60-min forecast of RaSfc valid at 120 min is rather poor (Fig. 3.12c) and is of similarly poor quality as that of Ra (Fig. 3.12b); the reflectivity associated with the main cell near the center is completely missing in both cases. The corresponding forecast of RaSfcS6 is much closer to the truth (Fig. 3.12d v.s. Fig. 3.12a). It is clear that when the surface data are assimilated for a relatively short period of time (8 cycles over 40 min), the benefit of a denser network is much greater. The benefit of the 20 km network is more or less lost after 40 min of forecast while for a denser 6 km network the benefit is sustained for the entire hour of forecast and even increases slightly with time for thermodynamic variables (Fig. 3.11b).

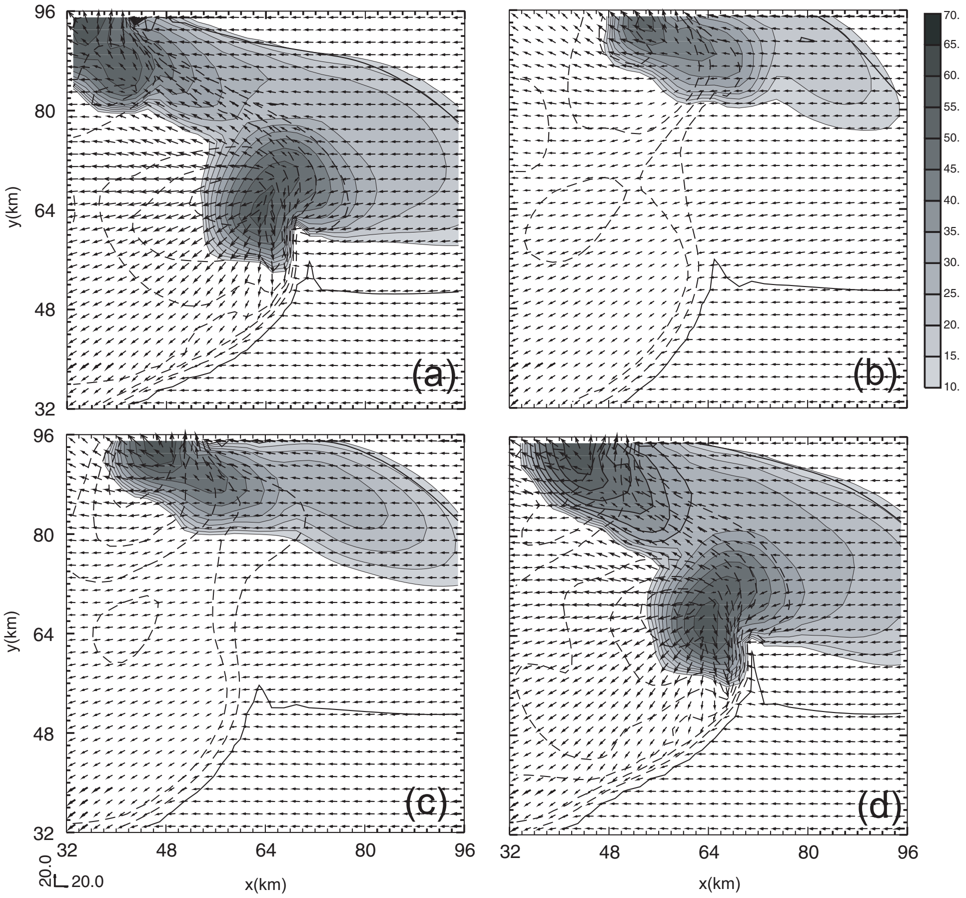

Fig. 3.12. Perturbation wind vectors, Z (dBZ, shaded) and θ' (K, contours) fields, valid at 120 min, from the truth (a), and 1-hour forecasts starting from 60-min ensemble-mean analyses of Ra (b), RaSfc (c) and RaSfcS6 (d).

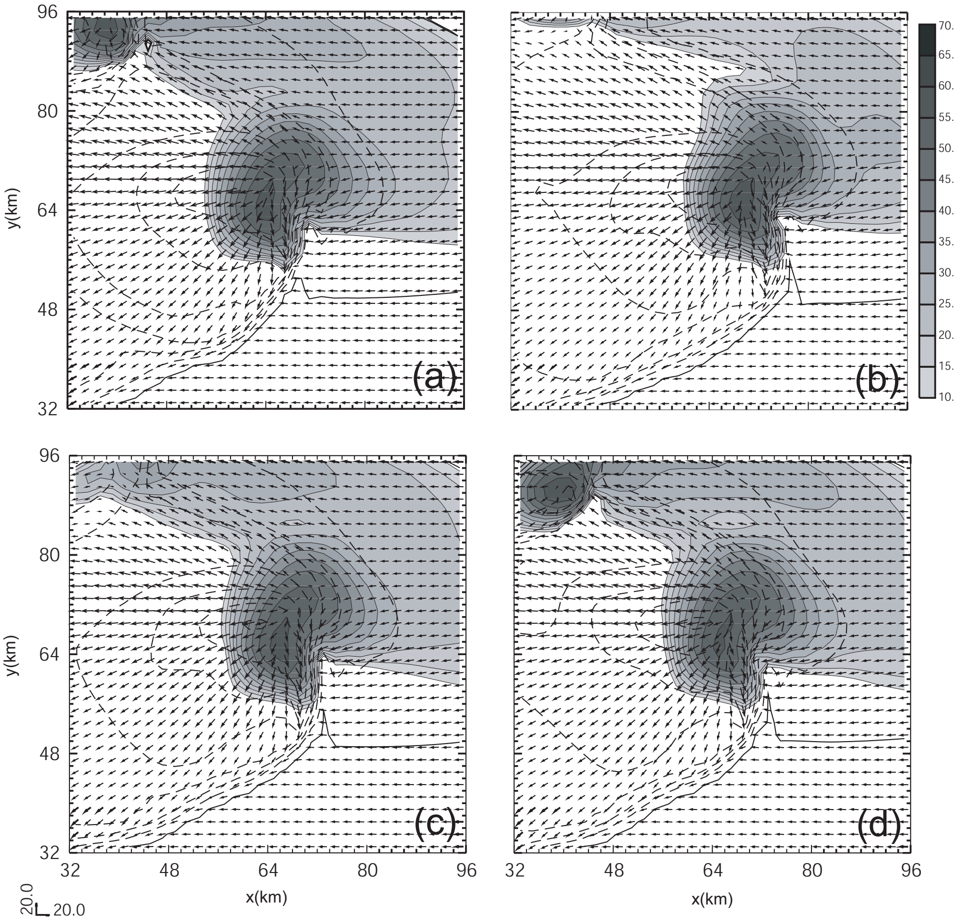

Fig. 3.13. As Fig. 3.12, but for fields valid at 150 min, of truth (a) and 1-hour forecasts starting from 90-min ensemble-mean analyses of Ra (b), RaSfc (c) and RaSfcS6 (d).

When surface observations are assimilated until 90 min of model time, significant benefit of surface observations is sustained for the entire hour of forecast (thin black lines in Fig. 3.11), for both network densities. In this case, the RERs for the wind components and the microphysical variables decrease with time for much of the time period. Overall, the RERs are lower than in the previous case of shorter assimilation period, and the RERs for the 6 km network is lower than those of 20 km network in both cases throughout the forecast, but the difference due to network density is generally larger when the forecasts are initialized at the earlier 60 min. This says that when surface observations are assimilated within fewer cycles, there is more benefit having a denser network. These are also supported by the plots of low-level forecast fields at 150 min (Fig. 3.13). The forecast of Ra (Fig. 3.13b) is poorest among the three and that of RsSfcS6 (Fig. 3.13d) is the best. The most noticeable differences are with the left-mover near the northwestern corner of the plotting domain. For the main cell near the center, the difference is smaller, especially when compared to the previous case where forecasts start from 60 min (c.f., Fig. 3.12).

When the assimilation cycles continue until 120 min, the analysis RERs become closer between the 6 km (RaSfcS6) and 20 km (RaSfc) cases (gray lines in Fig. 3.11). In fact, the analysis RERs for the wind components (Fig. 3.11a) are so close that after 10 min of forecast the RER of the 20 km case (thick gray curve) becomes lower than that of the 6 km case (this behavior is related to nonlinear error growth). The actual analysis errors at this time are already very small even with the 20 km spacing (c.f., Fig. 3.3) therefore further improvement due to higher network density becomes negligible. For temperature and pressure, the benefit of the denser network lasts a little longer (Fig. 3.11b) while the benefit for microphysical variables is sustained throughout the hour of forecast (Fig. 3.11c).

3.6 Impact of surface data in the presence of model error

Model error can significantly impact the behavior EnKF analysis. For real world problems, model error is inevitable. Model error can originate from physical parameterizations, model numerics and resolution, and simplifying assumptions made to the dynamic equations. For short-range (~ a few hours) simulation/prediction of convective storms, the microphysical parameterization (MP) is often the largest source of error or uncertainty (e.g., Gilmore et al. 2004b; Tong and Xue 2008b) and efforts to alleviate such uncertainty can be made through parameter estimation (Tong and Xue 2008a). In this section, we simulate the potential MP error by using MP scheme(s) in the assimilating model different from the one used in the truth simulation. Such an approach had been used with cumulus parameterization schemes in the OSSEs of Meng and Zhang (2007) for larger-scale applications. Our main goal here is to assess the impact of surface data in the presence of a dominant form of model error.

We first assess the impact of model error in the radar-only case. Experiments RaLFO04, RaSchultz and RaWSM6 are the same as Ra except for the use of “wrong” LFO04, Schultz and WSM6 (c.f., section 3.2.1) MP schemes, respectively. Experiment RaMulti uses LFO04, Schultz and WSM6 in 13, 13 and 14 of the 40 ensemble members and such a multi-physics approach has been found effective by Meng and Zhang (2007) and Fujita et al. (2007) in increasing the ensemble spread and reducing analysis error. Another experiment, RaNr0, uses the truth MP scheme (LFO83) but with the rain water intercept parameter, N_{r0}, increased by a factor of 10. This increases the number of rain drops and reduces their average sizes, leading to more rain evaporation and stronger cold pools. In the presence of model error, the covariance inflation coefficient is increased to 15%.

Fig. 3.14a shows the TRERs of the radar-only experiments with MP error, relative to perfect-model experiment Ra. In all cases, the TRER is larger than 1, indicating degradation of analysis due to model error. Among them, RaWSM6 has the largest TRER and RaNr0 has the smallest TRER. For RaNr0, the error is about 27% more while for

RaWSM6 the error is about a factor of 3 larger than the perfect-model case. For real cases, we do not know which MP scheme is more accurate and one way to account for such uncertainty is to use multiple MP schemes in the ensemble, which will at least help better sample the uncertainty. Fig. 3.14a shows that the TRER of RaMulti is smaller than those of RaWSM6 and RaSchultz but slightly larger than that of RaLFO04. Considering that LFO04 is a ‘close relative’ of LFO83 that is used in the truth simulation, the fact that RaMulti performs close to RaLFO04 suggests that the use of multiple MP schemes is rather effective.

In the presence of MP error, it is shown in Fig. 3.14b that with a network of 20 km spacing, surface observations still show positive impact on the storm analyses, but the magnitude of impact is reduced. The error reduction is largest for RaSfcNr0 with the TRER being about 80%. RaSfcNr0 has the smallest MP error since only the value of the rain intercept parameter is in error. The TRERs of all 20-km experiments using a single wrong microphysics scheme are close to 95%, indicating rather small positive impact from surface data. When multiple MP schemes are used in RaSfcMulti, the TRER is about 88%, smaller than the cases of single wrong scheme. We also notice that the microphysical state variables are improved least by the surface observations among these experiments (Table 3.1). A denser surface network of 6-km spacing is shown to produce larger positive impacts (Fig. 3.14b), with the error reduction in the multi-scheme case being about 25%, while that with wrong rain intercept parameter value is about 40%. The relative impacts among different schemes are similar to the 20-km case.

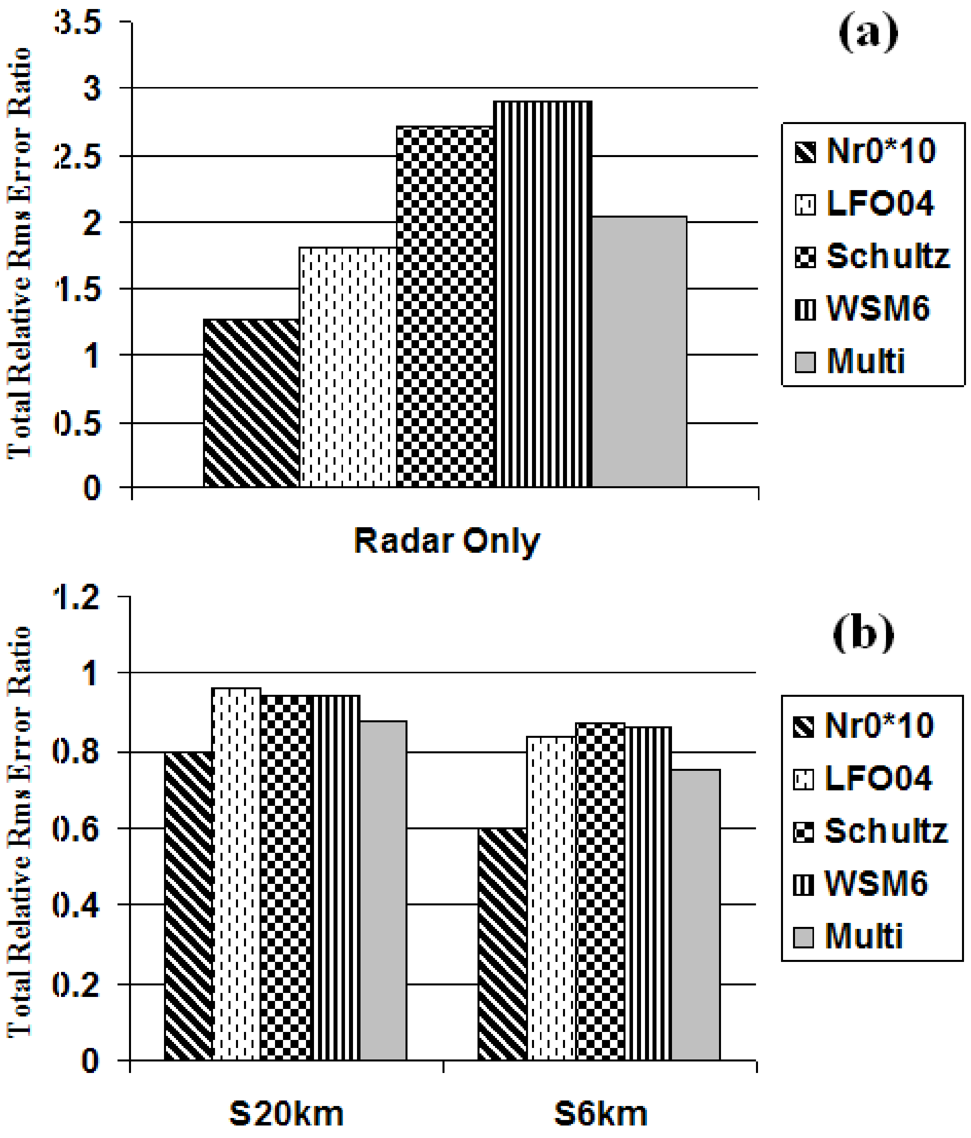

Fig. 3.14. (a) Total relative *rms* error ratios (TRERs) between imperfect and perfect model experiments for radar data only experiments RaNr0, RaLFO04, RaSchultz, and RaMulti, and (b) total relative *rms* error ratios (TRERs) for experiments with and without surface data when the model is imperfect.

We note here that in this study, the environment of the convective storms is assumed perfect (apart from the noise introduced when initializing the initial forecast ensemble), being based on a perfect sounding. The surface data are to help analyze the perturbations associated with the convective storms, and near the surface the cold pool and associated features are profoundly influenced by the microphysics (Snook and Xue 2008); for this reason, microphysics error significantly reduces the ability of surface observations in improving state estimation.

3.7 The case of imperfect storm environment

Most prior OSSE studies based on EnKF with convective scale analysis make the assumption of a perfect storm environment while in reality the storm environment always contain error which cannot be neglected. The source of these errors might result from previous inaccurate forecast or/and erroneous observations. Numerical simulations have shown that the environmental factor can affect storm initialization and development in terms of storm structure and propagation (Weisman and Klemp 1982; Weisman and Klemp 1984; Crook 1996; McCaul and Weisman 2001). Crook (1996) showed that moist convection forced by boundary layer process is sensitive to low-level thermodynamic environmental fields, e.g. moisture and temperature, with typical observational variability. Typical boundary moisture variability turns out to have larger impacts on storm strength for the well-developed convection than temperature. It is also found that low-level environment parameter could also exert influence on storm structure even if the land surface process is ignored (Park and Droegemeier 2000; Richardson and Droegemier 1999; Richardson et al. 2007). Park and Droegemeier (2000) examined the effect of prescribed subcloud water vapor error on storm evolution and suggested that the perturbation in ambient environment could affect both main and secondary storm. Richardson and Droegemier (1999) and Richardson et al. (2007) demonstrated that an inhomogeneous environment with horizontally variability of wind and low level moisture could also influence storm structure as well as propagation significantly. Compared with the studies examining the environment impact on storm development, there are almost no researches discussing the impact to assimilate radar observations for storm analysis in the existence of low-level environmental error. It is noted that the environment error might be involved in the real case data assimilation. However, it is hard to isolate these errors in complicated real cases when they are blended with other error sources. One of the advantages of OSSE is to provide insight into this specific problem by quantitatively adding prescribed errors on the ambient storm environment.

In this study, we try to investigate the impact of assimilating radar and surface observations in the existence of storm environment errors with the aid of OSSE. Various storm environment errors, e.g. moisture, temperature and wind, will be added to the background fields. Simulated radar and surface observations will be assimilated to study their individual and combined impact on storm analysis. We seek to answer the following questions in this study: What is the impact of assimilating radar observations on convective scale storm analysis in presence of various ambient storm environment errors? Could the assimilation of additional mesonet surface observations helps to correct the environment error and achieve improvement on analyses and forecasts?

Since the focus of this study is the development of the supercell storm related to low level cold pool, rather than an investigation of convection initialization etc., there is no boundary layer process (surface physics) involved in all the simulations and analyses for simplicity. It is also known that the cold pool produced by convection would suppress the boundary layer activity significantly so it appears to us reasonable to ignore this process in our OSSE setting.

3.7.1 Experiment setup

The horizontal grid resolution used in the model is 2 km and the vertical grid spacing stretches from 200 m near the ground to 800 m at the model top with the stretching scheme using a hyperbolic tangent function. A grid of 51×51×35 points in x, y and z direction is chosen, which gives a 96×96×16 km physical domain for all the simulations and analyses. The sounding extracted from the 20 May 1977 Del City, Oklahoma, supercell storm (Ray et al. 1981) is modified in low levels with a well mixed boundary layer (constant moisture and potential temperature) to initialize the truth storm simulation (Fig. 3.15). We refer this perfect sounding as PS thereafter. The low level environment moisture, potential temperature and wind error will be added to this sounding later to create an imperfect storm environment, with these erroneous soundings referred as MES, TES and WES thereafter respectively (Fig. 3.15). The environment error is -3 m/s for both u and v, 2 K for potential temperature and 2 g/kg for water vapor mixing ratio, respectively.

3.7.1.1 Truth simulation

An initial thermal bubble with a maximum potential perturbation of 4 K is centered at x = 48 km, y = 16 km and z = 1.5 km to trigger convection in the truth simulation. The radii of the bubble are 10 km in the horizontal and 1.5 km in the vertical. Open lateral and free-slip top and bottom boundary conditions are applied in both simulation and assimilation.

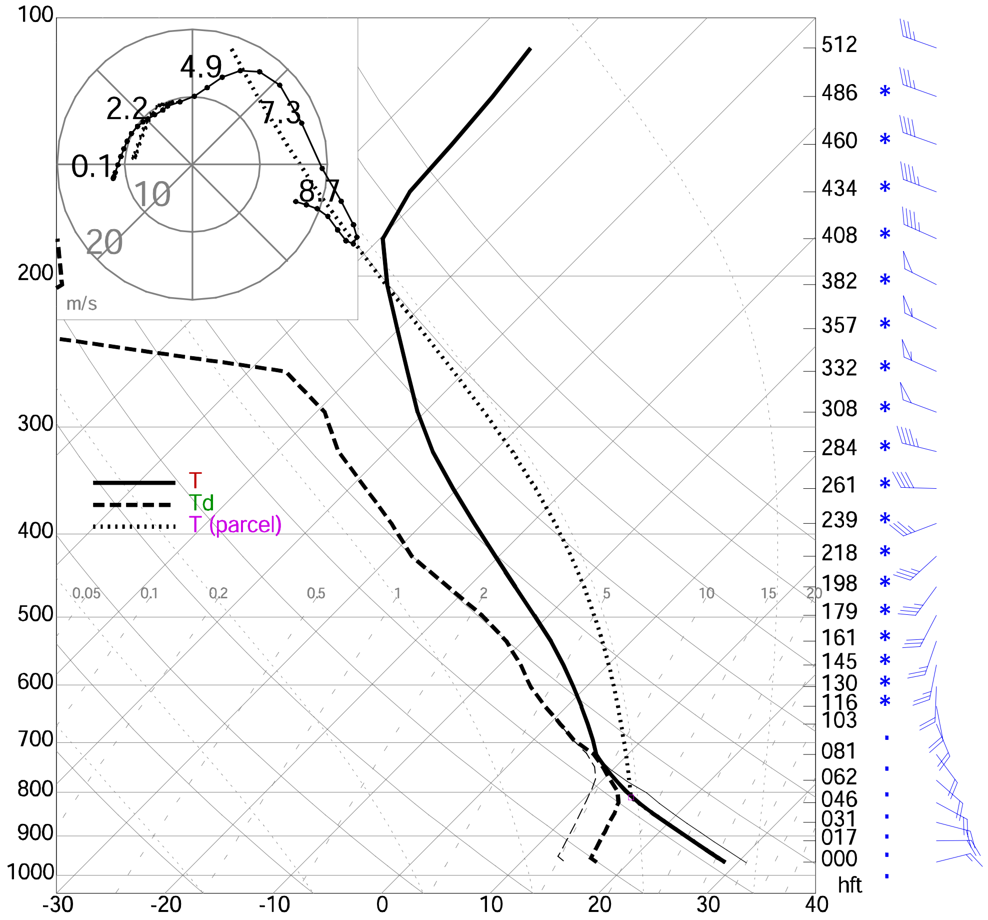

Fig. 3.15. Sounding to initialize the storm simulations. The gray lines below 700 hpa denote the environment errors added to the original sounding.

The bubble-triggered storm reaches its full updraft intensity within 20 min. The supercell starts to split into two, with one right mover and one left mover (Fig. 3.16a) at around 60 min. By 120 min, the left mover propagates to northwest of the physical domain (Fig. 3.16b) and the third cell splits from the left mover. After another hour, the major part of the secondary and third cell both moves out of the physical domain on the north boundary. The right-moving cell stays around the centre of the physical domain during the whole simulation time and the updraft continues to strengthen to a peak of 44 m/s at 180 min, the end of truth simulation, after a temporary weakening owing to the splitting at about 60 min. The simulated model state is output every 5 min for observation simulation and for analysis/forecast verification.

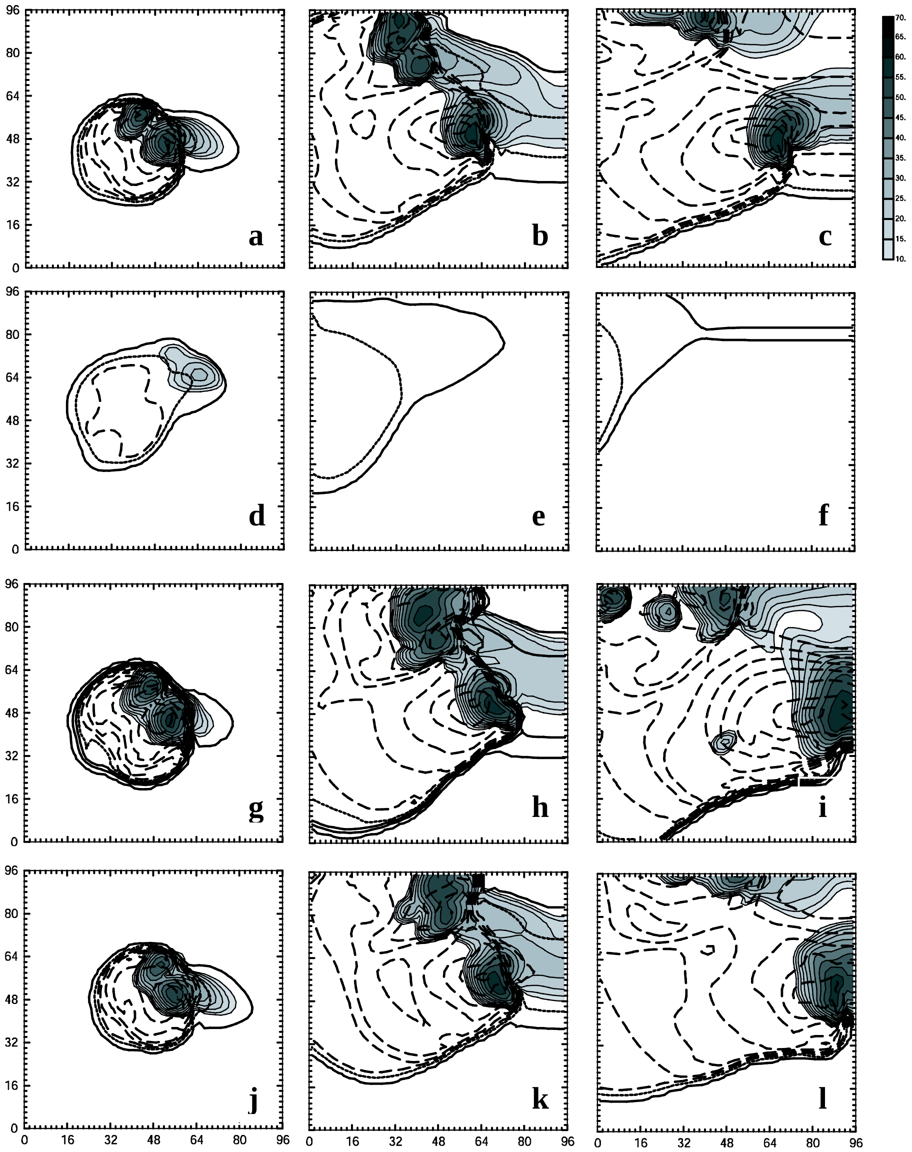

Fig. 3.16. Simulated reflectivity Z (dBZ) and perturbation potential temperature θ' (K) z = 100 m (first model level above ground) for perfect environment simulation (a-c), simulation with moisture error in the environment (d-f), simulation with potential temperature error in the environment (g-i) and simulation with moisture error in the environment (j-l) at 60 min (a, d, g, j), 120 min (b, e, h, k) and 180 min (c, f, i, l) of model time.

Before examining the impact of assimilating radar and surface observations, we set out to perform a series of simulation using the erroneous soundings with no assimilations conducted, aiming to gain some insight into the impact of the environment error on the storm development. With the low level -2 g/kg moisture error in the environment and a drier sounding, the storm is located at the northeast of the domain center at 60 min and too weak without cell splitting (Fig. 3.16d). It dissipates quickly after 60 min (Fig. 3.16e and f). With the 2 K potential temperature error, the storm development during the first two hours is similar to the perfect environment simulation (Fig. 3.16g and h). At 180 min, the right mover propagates to the east faster and has a larger storm area than the perfect environment simulation (Fig. 3.16i). More small cells split from the left mover at the north boundary of the physical domain. With the 3 m/s wind error, storm evolution is also similar with to the perfect environment simulation in the first two hours (Fig. 3.16 j and k). Owing to the weaker inflow, the right mover also exhibits a faster propagation to the east at 180 min (Fig. 3.16l).

3.7.1.2 Simulated observations

In most sets of the experiments, a hypothetical WSR-88D radar with 1 degree beamwidth is positioned at x = - 34 km and y = - 34 km (the coordinate origin is located at the southwest corner of the physical domain) which gives about 115 km from the domain center; approximately where the right-moving cell is located. At this distance, the earth curvature effect combined with beam bending based on the ¾ earth radius model (Doviak and Zrnic 1993) places the lower edge and the center of the half-power beam of

0.5 degree elevation at 1.98 and 3.60 km above ground, respectively. In another word, there is no direct radar data coverage below 1.98 km level at all at this distance.

Radar observations to be assimilated include radial velocity (V_r) and reflectivity (Z). The observation operators used in both the simulator and analysis are the same as XTD06 (V_r) and TX05 (Z). As in XTD06, radar data are assumed to be available on the elevation levels in the vertical and already interpolated to model grid columns in the horizontal. The standard WSR-88D precipitation scanning mode is assumed (see Fig. 2 of XTD06). Gaussian-distributed random errors with zero mean and standard deviations of 1 m s^{-1} and 3 dBZ are added to simulated V_r and Z, respectively.

The hypothetical mesoscale observing network has station spacings of about 20 km in the assimilation experiments. To simulate a network whose stations are not on the grid point while keeping the network more or less uniform, the stations are located randomly within 400 × 400 m square boxes that are centered on the grid points 20 km apart.

The horizontal wind components, the temperature, pressure, and water vapor mixing ratio consist of the assimilated surface measurements. The standard deviations of the zero-mean Gaussian errors added to the simulated surface observations are: 1 m s^{-1} for wind components, 1 K for temperature, 1 hPa for pressure and 1 g kg^{-1} for water vapor mixing ratio.

Both radar and surface observations are assumed to be available every 5 minutes, which is typical data update interval with the Oklahoma Mesonet and the WSR-88D radar volume scan in precipitation mode.

3.7.1.3 EnKF algorithm

EnSRF is applied as the analysis algorithm. The implementation follows the perfect environment experiments in section 3.2 exactly: surface observations are assimilated after the radar data.

The state variables analyzed include u, v, w, θ, p, q_v, q_c, q_r, q_i, q_s, and q_h. To handle the low level environment error, the initial sounding is perturbed with uniform random Gaussian noises in the vertical. These Gaussian noises have zero means and standard deviations of 3 m/s for horizontal wind components, 2 K for potential temperature and 2 g/kg for water vapor mixing ratio. To initialize the first ensemble forecast cycle, another set of random perturbations are drawn from a Gaussian distribution with zero mean and standard deviations of 2 m s^{-1} for u, v, w, 2 K for θ, and 0.6 g kg^{-1} for q_v and all microphysical variables. Perturbations for all except for the microphysical variables are added in the entire domain. The perturbations for the latter are added only in the region where radar echo is present at 20 min, the start time of the first assimilation cycle. Reflectivity data in both the precipitation and clear air regions (negative Z values are set to zero) are used. Radial velocity data are only used in regions where Z is greater than or equal to 10 dBZ.

Covariance localization (Houtekamer and Mitchell 2001) is used to limit the spatial influence of observations and reduce sampling error. The same scheme as used in section 3.2 is applied with a smooth 5th-order distance-dependent function (Eq.(4.10) of Gaspari and Cohn, 1999) multiplying the calculated background error covariance. For radar observations, a 6 km localization radius is chosen in all directions to ensure the best results when only radar data is used (XTD06). For surface observations, when the mean station spacing (station spacing hereafter for short) is 20 km, a horizontal localization radius of 30 km and a vertical radius of 6 km are found to be optimal through experimentation. Forty ensemble members are used for all experiments, as in XTD06.

To avoid the filter divergence problem caused by underestimation of covariance due to small sample size and/or model error, covariance inflation is used in all experiments following the procedure of section 3.2. A 15% inflation factor is used for all the experiments with storm environment error.

The data assimilation experiments are summarized in Table 3.3:

Data Type Environment Error	Radar alone	Radar + surface
Wind error	RaUVE	RaSfcUVE
Potential temperature error	RaPTE	RaSfcPTE
Moisture error	RaME	RaSfcME

Table 3.3: List of the experiments investigating the radar and surface impact under imperfect environment assumption.

3.7.2 The impact of surface observations

3.7.2.1 The impact on environment

The profiles at (86 km, 10 km) from the truth simulation and data assimilation experiments at 60 min, representing the storm environment, are extracted to investigate the impact of data assimilation on the environment (Fig. 3.17). For various environment errors, the radar data alone cannot correct the errors while the addition of the surface observations improves the storm environment significantly. The potential temperature profile from RaSfcPTE is close to the truth (Fig. 3.17b). The u profile from RaSfcUVE and the water vapor mixing ratio profile from RaSfcME are both close to the truth in the lowest 1 km, and the improvements are also evident in the higher levels (Fig. 3.17a and c).

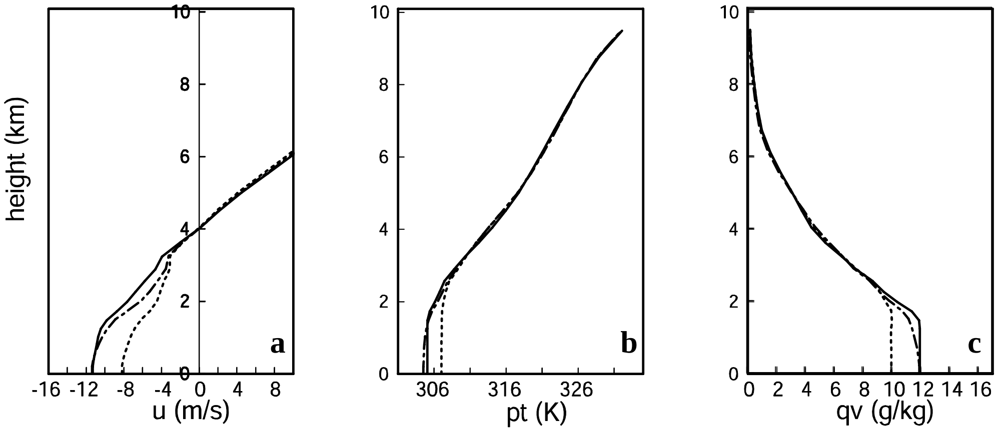

Fig. 3.17. The a) u, b) potential temperature and c) water vapor mixing ratio profiles at (86, 10) km from a) RaUVE (dotted) and RaSfcUVE (dot-dashed), b) RaPTE (dotted) and RaSfcPTE (dot-dashed), c) RaME (dotted) and RaSfcME (dot-dashed) and the truth simulation (solid). The profiles are all at 60 min.

3.7.2.2 The impact on storm analysis

The analysis rms error of RaME and RaSfcME in the precipitation region against time are plotted in Fig. 3.17. There is significant improvement on all the state variables from the assimilation of surface observations and the improvement increases with time during the assimilation cycles. For most of the state variables the relative improvement are over or around 50% at the end of the analysis cycles. Although with radar data assimilation, the storm develops and splits, there is no hook echo structure in the right mover (not shown). The surface observations retrieve the hook echo structure successfully (not sown).

The improvement of adding surface data over assimilating radar observations alone is still noticeable in terms of the rms error reduction when wind or potential temperature error exists but not as large as the moisture error experiment (Fig. 3.21 C1 and C2).

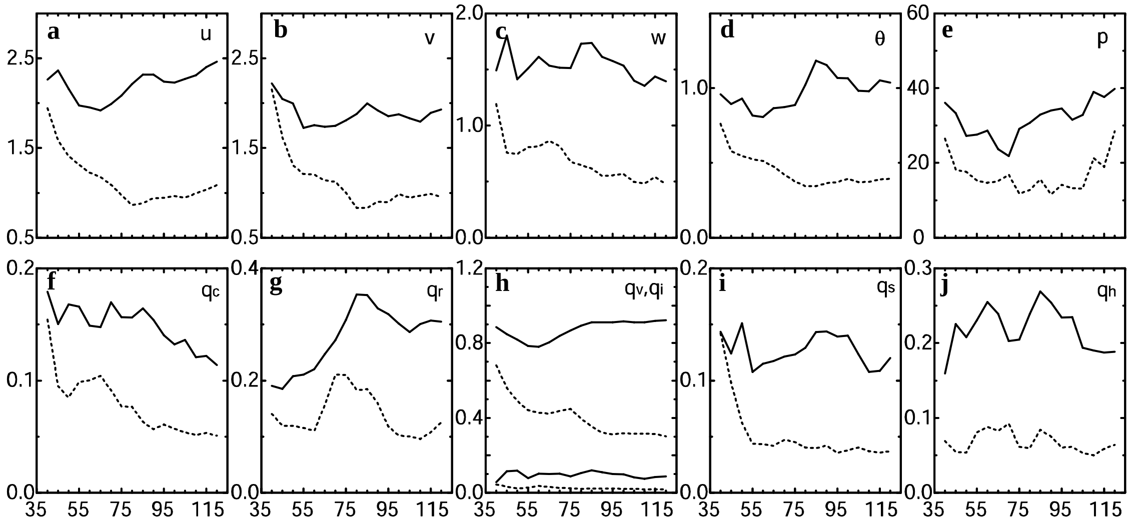

Fig. 3.18. The *rms* error of ensemble mean analyses plotted against time for RaME (solid) and RaSfcME (dotted) for a) u (m s^{-1}), b) v (m s^{-1}), c) w (m s^{-1}), d) potential temperature θ (K), e) pressure p (Pa), f) q_c (g kg^{-1}), g) q_r (g kg^{-1}), h) (upper curves) q_v and (lower curves) q_i (g kg^{-1}), i) q_s (g kg^{-1}), and j) q_h (g kg^{-1}).

3.7.2.3 The impact on forecast

Two sets of forecasts start from 60 min analysis and 120 min analysis respectively. The forecasts at 120 min from the analyses at 60 min are plotted in Fig. 3.18. For RaME, the storm cell is too weak due to the relatively dry environment (Fig. 3.18a). RaPTE and RaUVE both have the right and left mover but both with another spurious cell splitting from the right mover (Fig. 3.18b and c).

With the additional surface observations assimilated, the storm development and splitting are correctly forecasted in RaSfcME (Fig. 3.18d). In RaSfcPTE and RaSfcUVE, the spurious cell in RaPTE and RaUVE is removed (Fig. 3.18e and f). Generally, the assimilation of both radar and surface observations helps to capture the position and the strength of the storms accurately.

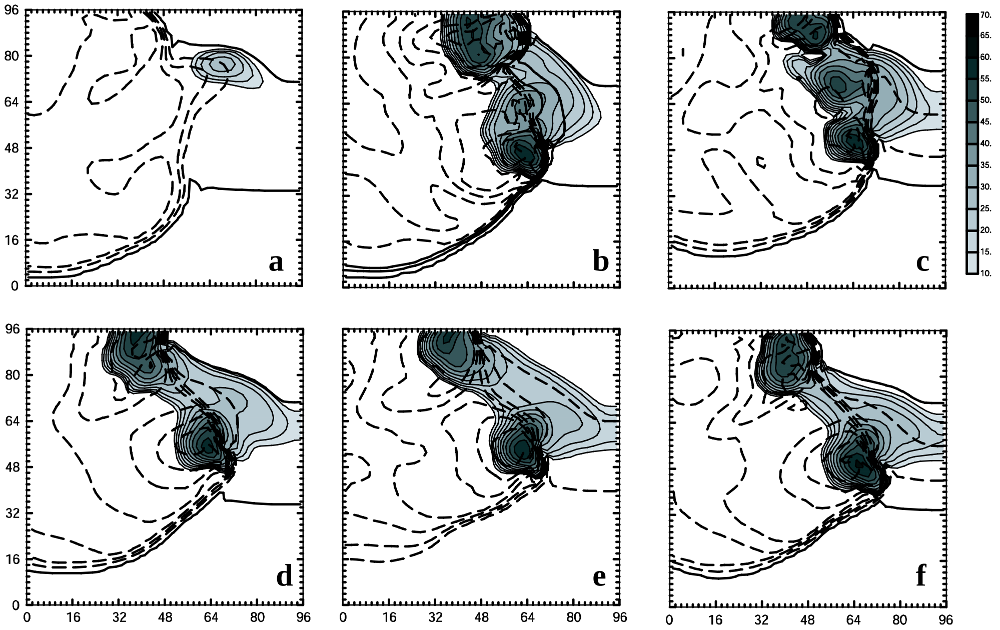

Fig. 3.19. As Fig. 3.16 but at 120 min forecast from the analysis at 60 min and for a) RaME, b) RaPTE, c) RaUVE, d) RaSfcME, e) RaSfcPTE, f) RaSfcUVE.

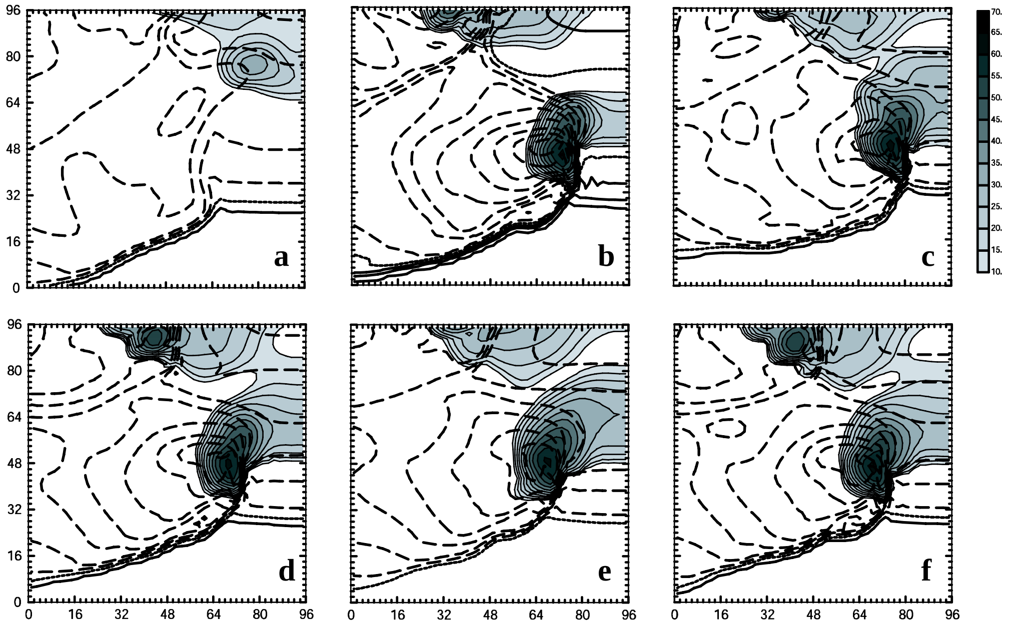

Fig. 3.20. As Fig. 3.19 but at 180 min forecast from the analysis at 120 min.

When the forecast starts from the RaME analysis at 120 min, the storms of the analysis at 120 min have the similar strength with the truth simulation (not shown). However, they both dissipate very quickly in one hour, with the storm remnant left (Fig. 3.19a). The RaPTE and RaUVE both have better forecast than RaME in terms of the storm strength and structure but with the right mover propagating faster to the east than the truth (Fig. 3.19b and c). There is a dramatic improvement from RaSfcME over RaME with the addition of surface observations with the structure, strength and the position of the storms much closer to the truth (Fig. 3.19e). The RaSfcPTE and RaSfcUVE both improve the position forecast of the right mover by around 10 km over RaPTE and RaUVE respectively.

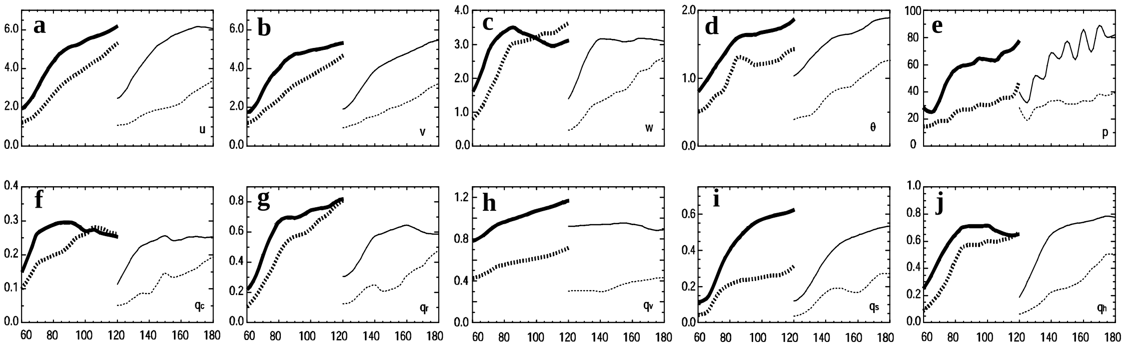

Fig. 3.21. The *rms* error of a) u (m s^{-1}), b) v (m s^{-1}), c) w (m s^{-1}), d) potential temperature θ (K), e) pressure p (Pa), f) q_c (g kg^{-1}), g) q_r (g kg^{-1}), h) (upper curves) q_v (g kg^{-1}), i) q_s (g kg^{-1}), and j) q_h (g kg^{-1}) for 60-min-long forecasts starting from ensemble-mean analyses at 60 min (thick black lines) and 120 min (thin black lines). The solid curves are for the forecast errors starting from the analyses of RaME and the dashed lines are for the forecast errors of RaSfcME.

To quantitatively explore the impact from the surface observations on the forecast, rms error in the precipitation region of the forecast is plotted for RaME and RaSfcME (Fig. 3.20). For the forecasts starting from 60 min, the improvement from the additional surface observations is clear during the first 40 min. The improvement over RaME increases with time during the first 40 min for wind, pressure and some of the microphysical variables (q_c, q_r and q_s). However, there is a degradation from the addition of the surface observations during the last 20 min in the one hour forecast, leading to a larger rms error than RaME for w and q_c, or similar rms error with RaME for q_r and q_h at 120 min. A careful investigation reveals that this degradation is from the phase error in RaSfcME. There is a 6 km forecast position error of the right mover in RaSfcME at 120 min, reducing the improvement. Despite the rms error degradation for some state variables, it is still clear that the storm structure in RaSfcME is much better than RaME at 120 min (Fig. 3.18a and d).

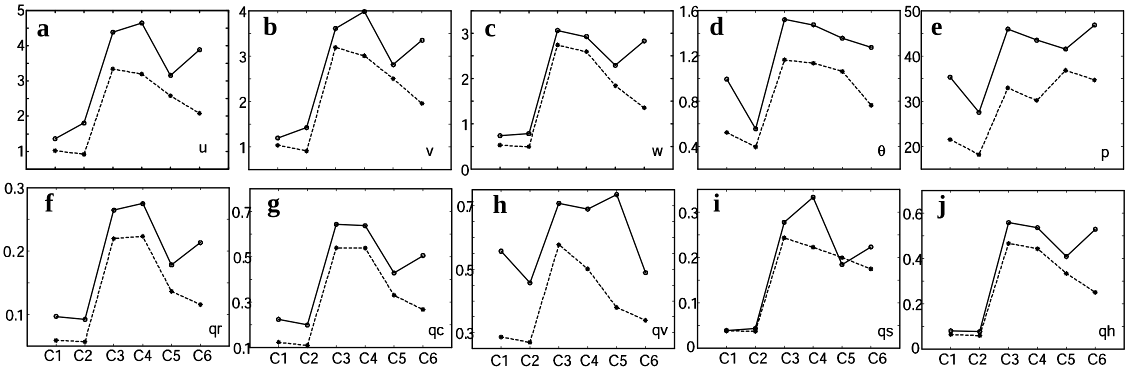

Fig. 3.22. The average rms error a) u (m s^{-1}), b) v (m s^{-1}), c) w (m s^{-1}), d) potential temperature θ (K), e) pressure p (Pa), f) q_c ($\mathrm{g\,kg^{-1}}$), g) q_r ($\mathrm{g\,kg^{-1}}$), h) (upper curves) q_v ($\mathrm{g\,kg^{-1}}$), i) q_s ($\mathrm{g\,kg^{-1}}$), and j) q_h ($\mathrm{g\,kg^{-1}}$) for RaPTE analysis (C1 solid), RaSfcPTE analysis (C1 dotted), RaUVE analysis (C2 solid), RaSfcUVE analysis (C2 dotted), one hour RaPTE forecast from the analysis at 60 min (C3 solid), one hour RaSfcPTE forecast from the analysis at 60 min (C3 dotted), one hour RaUVE forecast from the analysis at 60 min (C4 solid), one hour RaSfcUVE forecast from the analysis at 60 min (C4 dotted), one hour RaPTE forecast from the analysis at 120 min (C5 solid), one hour RaSfcPTE forecast from the analysis at 120 min (C5 dotted), one hour RaUVE forecast from the analysis at 120 min (C6 solid) and one hour RaSfcUVE forecast from the analysis at 120 min (C6 dotted).

The forecast starting from the analysis at 120 min of RaSfcME shows positive impact on every state variables over RaME during all the forecast times, indicating an improvement of assimilating additional surface observations not only on moisture but all other model fields (Fig. 3.20).

The average rms error is used to quantify the impact from data assimilation when the potential temperature or wind error exists in the storm environment. The average rms error is calculated by averaging the rms error over the specified analysis or forecast times and plotted for RaPTE, RaUVE, RaSfcPTE and RaSfcUVE in Fig. 3.21. The average rms error of assimilating additional surface observations in the analysis shows noticeable improvement over assimilating radar alone for most of the state variables (Fig. 3.21 C1 and C2). For the one hour forecasts starting from the analyses at 60 min or 120 min, the improvements from the additional surface observations are also evident (Fig. 3.21 C3-C6).

3.8 Summary and conclusions

In this chapter, an ensemble square root Kalman filter is used to assimilate simulated radar and surface network observations together for a supercell storm, for the purpose of examining the impact of additional surface observations on the storm analysis and forecasting. Realistic low-level radar data coverage, or the lack thereof, is simulated with a radar emulator using realistic beam pattern weighting.

It is shown that the assimilation of mesonet-like surface observations can significantly improve convective storm analysis when compared to the cases using radar observations only, particularly when the radar is at a sufficient distance from the storm to have a poor low-level coverage. Surface observations help fill the low-level data gap and improve the storm analysis even when the resolution of these observations is much coarser (e.g., at a 20 km spacing) compared to the radar data. The *rms* error reductions due to surface observations reach 30% to 50% in such a case when the assimilation model is perfect. Surface observations improve analyses of the cold pool and associated convergence and divergence in terms of pattern, magnitude and position while precipitation fields are improved less, because hydrometeors at higher levels are only indirectly related to surface observations and are better captured by radar.

It is shown that the surface observations not only help correct the near surface errors, but also improve the analyses of state variables at the mid- and upper levels. The greatest improvement from surface observations usually occurs at levels where analysis errors without the use of surface data are largest. The flow-dependent background error covariance estimated from the ensemble and the dynamical interactions realized through the prediction model are believed to play an important role. The background error correlations estimated from the ensemble exhibit physically reasonable structures and confirm the ability of the surface observations to properly influence all state variables at other grid points.

In general, a better analysis obtained using both radar and surface observations also improves the subsequent storm forecast. When the forecast starts at an early time when fewer assimilation cycles have been taken and when cold pool occupies a smaller area, a higher-density surface network shows a much greater positive impact. When many analysis cycles are performed, the 20-km spacing surface network is able to produce a similar level of positive impact as the 6-km network and the impact can last for at least 1 hour in both cases during forecast. In some cases, the improvement in the forecast even increases with forecast time.

Among the four surface observation variables with typical observation errors, wind observations show the largest positive impact, reflecting the important role of low level flows in the storm development, maintenance and propagation. Temperature observations are found to have the second largest impact and pressure observations are found to have the least impact. Part of the reason for the small impact of pressure observations may be due to the relatively large error specified for the pressure observations (1 hPa standard deviation) as compared to the error level of analyzed pressure achieved with radar data alone. The largest positive impact is achieved when assimilating all surface observation variables together.

Different covariance localization radii are used for radar and surface observations in the EnKF analysis. The optimal radii found through experimentation for surface observations generally decrease with the decrease in network spacing, and those of surface data are generally larger than those of radar data; such a 'multi-scale' analysis technique seems to be effective when dealing with data from networks of very different resolutions. An argument in support of such an approach is that observation networks are often designed to capture their respective target flow scales (e.g., a rawinsonde network is designed to primarily capture synoptic-scale structures). A practical consideration is the desire to produce relatively smooth analyses given the observation networks, and to avoid 'bull's eyes' in the analyses. Similar multi-scale technique has been used in the context of 3DVAR (e.g., Hu et al. 2006) and successive correction method (e.g., Xue and Martin 2006).

The *rms* error relative to the radar-only case is found to decrease linearly with the decrease in the mean surface network spacing until the spacing is close to the grid interval of truth simulation. When plotted against the total number of surface stations, this relative error is approximately proportional to the inverse square root of the number of stations when the network is relatively coarse, similar to the behavior found by Morss et al. (2001).

When the radar is located at a closer distance of 115 km, the radar-only analysis is significantly improved over that of 185 km case. In this case, the surface observations still show consistent positive impact and the impact increases with the network density. When the radar is only 45 km from the storm, a 6-km station spacing is needed to achieve a noticeable impact because the radar-only analysis is already very good.

The impact of surface observations is also examined in the presence of microphysics-related model error. Such model error is found to reduce the relative impact of surface observations. With the use of multiple microphysics schemes in the forecast ensemble, a surface network of 20-km (6-km) spacing produces error reduction of about 10% (25%).

When there is low level error in the storm environment, surface observations could help to correct the environment error, improving storm analysis and forecast. When low level moisture error (being too dry) exists in the environment, the storms dissipate quickly in the forecast when radar observations alone are assimilated. The addition of surface observations allowed correct analysis of the storm development, including the splitting process. Surface observations also improve the storm movement forecasts and remove spurious storm cells developing when low-level potential temperature or wind error exists.

The impact of surface observations does depend on the quality of analysis using radar data only. It also depends on the general quality of state estimation that affects the estimated error covariance between the surface observations and unobserved state variables. Attempts to analyze the convective storms using surface data only were unsuccessful. Finally we note that while this study may share some of the common limitations of OSSEs, the results obtained do give us valuable insights that are difficult, if at all possible, to obtain with real data.

Chapter 4

Assimilation of Radial Velocity and Reflectivity Data from Coastal WSR-88D Radars Using EnKF for Hurricane Ike (2008)

4.1 Introduction

Landfalling hurricanes can pose deadly threats to many lives and cause loss of billions of dollars. Accurate prediction of hurricane track, intensity and rainfall can save lives and properties. The official forecasts of the hurricane track from National Hurricane Center have improved greatly in recent years by the overall trend (Rappaport et al. 2009; National Hurricane Center 2010). However, the improvement in hurricane intensity forecasting has been limited (Houze et al. 2007). The accuracy of the current hurricane intensity forecasts are almost in the same level as in 1990 (Rappaport et al. 2009).

Many efforts have been made to identify the factors affecting intensity prediction and to improve intensity forecasting (Anthes 1982; Emanuel 2005; Krishnamurti et al. 2005; Rogers et al. 2006; Li and Pu 2009; Khain et al. 2010). These factors include the surface heat and moisture fluxes, sea surface temperature, and environment wind shear, etc. Among these factors, intensity forecasting is believed to be closely associated with the internal meso-scale or smaller scale structures of the wind, cloud and precipitation (Houze 2007; Wang 2009). On the other hand, while the hurricane track is primarily decided by the large scale steering flow, some of the cyclone scale factors and features, such as beta drift effect, convection distribution, vertical structure, microphysics processes and other asymmetries, can also influence the track forecasting (Kimberlain 2007; Fovell et al. 2009; Fovell et al. 2010). Radar is one of the most effective observation platforms to provide essential information on hurricane structures at high temporal and spatial resolutions. Radial velocity and reflectivity observed by Doppler radars can provide essential information about the wind and microphysics field of hurricanes at the meso-scale and convective scale.

High-resolution numerical models have proven to be one of the most promising tools to generate reliable hurricane forecasts (Chen et al. 2007; Davies et al. 2008). To improve the ability of the numerical model to accurately predict hurricane track and intensity, some of the researches focused on the sensitivity of model resolution and parameterization on hurricane structure and forecast (McFarquhar et al. 2006; Fierro et al. 2009). Other studies aimed to improve the initial condition with data assimilation methods. With bogus vortex data assimilated into numerical models, the hurricane forecast can be improved (Zou and Xiao 2000; Pu and Braun 2001; Kwon and Cheon 2010; Xiao et al. 2009a; Zou et al. 2010). Even without vortex bogussing, radar observations can help to create relatively accurate initial conditions and lead to improved intensity and track forecasts when assimilated into high-resolution numerical models (e.g., Xiao et al. 2007; Zhao and Jin 2008; Pu et al. 2009; Xiao et al. 2009b; Zhang et al. 2009; Zhao and Xue 2009).

Three-dimensional variational data assimilation method (3DVAR) is widely employed in most of the previous radar data assimilation researches for hurricane analysis and forecast. However, the static background error statistics assumption is not appropriate in the rapid evolving hurricanes. The use of independent background error covariance may lead to ill-suited problems for the initialization (Zhang et al. 2009). Among other advanced data assimilation methods, the ensemble Kalman filter (EnKF) employs ensemble forecasts to estimate flow-dependent background error covariance, thus showing advantages over 3DVAR. It was first proposed by Evensen (1994) and has bred an array of variants since then. Its simple formulation and easy implementation compared to four-dimensional variational data assimilation method, allow it to enjoy popularity within the research community as a state of the art assimilation algorithm. EnKF also generates initial conditions for ensemble forecasts in a natural way, providing uncertainty information on hurricane analysis and forecasts. It also can be used for predictability research (Sippel and Zhang 2010).

Previous studies have demonstrated that EnKF is a powerful tool for radar data assimilation, including radial velocity and reflectivity, to improve convective thunderstorm analysis (Dowell et al. 2004; Zhang et al. 2004; Tong and Xue 2005; Xue et al. 2006; Aksoy et al. 2009). Zhang et al. (2009) assimilated radial velocity from three coastal radars for the cloud-resolving hurricane analysis and prediction with EnKF for the first time. With the radial wind assimilation into Hurricane Humberto (2007) using EnKF, the analysis captures the best track position and intensity closely. The deterministic forecasts show improvement in track and intensity over operational forecasts and the forecasts initialized with 3DVAR. While radial velocity provides the wind structure from a hurricane, reflectivity includes information from the hydrometeors and microphysical fields. Assimilating reflectivity into numerical models with 3DVAR shows positive impact on hurricane quantitative precipitation forecasting (QPF) (Xiao et al. 2007). However, the assimilation of reflectivity for hurricane analysis and forecast with EnKF is still limited.

This study investigates for the first time the assimilation of both radar radial wind and radar reflectivity observations with EnKF for cloud-resolving hurricane analysis and forecasting. As one of the variants of EnKF, the serial ensemble square root filter (EnSRF, Whitaker and Hamill 2002), which is used extensively for thunderstorm radar assimilation, is chosen to assimilate radar observations for a hurricane in this study.

Mean sea level pressure (MSLP) from the best track at the vortex center has proven to be able to improve the hurricane track forecast when assimilated with EnKF and global models with 30-50 km resolutions (Hamill et al. 2010). In this study, "MSLP" hereafter always refers to the mean sea level pressure from the best track at the vortex center. The intensity forecast in Hamill et al's study, however, is problematic probably because their relatively coarse resolution is not sufficient to resolve the realistic hurricane vortex structure. It remains unknown if MSLP can further improve the cloud-resolving hurricane analysis and forecast when it is assimilated with radar observations using EnKF. We will assimilate MSLP from the best track dataset with radar data to examine its additional impact for the first time.

The intensity and track uncertainty in hurricane forecast can results from both model and initial condition uncertainties. It is still not clear which one of these two uncertainty sources has a larger contribution to the spread growth in intensity and track forecast. Experiments will be conducted to briefly investigate this issue.

This chapter is organized as follows. Section 4.2 will briefly introduce Hurricane Ike. The forecast model, the radar observations and the EnKF experiment setup are described in section 4.3. Ensemble spreads and innovation statistics during the analysis cycles are discussed in section 4.4. Section 4.5 discusses the impact of radar data on the deterministic forecasting of hurricane intensity, track and precipitation while ensemble forecasting results are shown in section 4.6, with a comparison with the deterministic forecasts. Sections 4.7, 4.8 and 4.9 examine the sensitivity to cutoff radius, data assimilation intervals, and use of single radars. The assimilation of MSLP from the best track will be presented and discussed in section 4.10. Section 4.11 will discuss the uncertainty growth issue briefly. A summary is given in section 4.12.

4.2 Hurricane Ike (2008)

Hurricane Ike (2008) studied here is the most intense hurricane during the 2008 Atlantic hurricane season. It started as a tropical disturbance near Africa at the end of August. On Sept. 1, 2008, it became a tropical storm west of the Cape Verde islands. During Sept. 3 and 4, Ike is in an area that lacked strong wind shear. It developed into a category 4 hurricane on the Saffir-Simpson Scale during its strongest stage of 145 mph (230 kmh^{-1}) with an estimated pressure of 935 hPa by the early morning at Sept. 4. The wind and pressure both are from satellite estimate since it was still too far from land for reconnaissance aircraft to reach (Berg 2009). It made landfall twice in Cuba before entering the Gulf of Mexico in the afternoon of September 9. The interaction with Cuba leaded to a disruption of Ike's inner core and an expansion of Ike's wind field. The intensity weakened to a category 1 hurricane. An eyewall replacement also happened when Ike moved into the southeastern Gulf when the outer rainband started to move inward and replaced the inner eyewall.

During the night of Sept. 10, Ike showed a rapid drop of minimum sea level pressure, falling from 963 hPa to 944 hPa as it passed over the Loop Current in the Gulf of Mexico. This drop was not reflected by the wind speed, however, which only increased to 100 mph (160 kmh^{-1}) from 85 mph (140 kmh^{-1}). Two well-defined wind maxima of nearly equal strength were reported by the National Hurricane Center. This unusually broad distribution of strong winds indicates the structure was absorbing and distributing energy over a large area, rather than concentrating it near the center. During this intensification, Ike strengthened to a category 2 from category 1.

During Sept. 11 and 12, Ike maintained its intensity and the broad wind field. When Ike moved toward Texas coast, the inner structure and eyewall became more organized. With a diameter of 900 mile, Ike is the largest Atlantic tropical cyclone in the recorded history. It made landfall in Galveston, Texas at 0700 UTC 13 September as a category 2 hurricane. After the first landfall, Ike passed over San Leon, Texas and made its final landfall near Baytown, Texas around 0900 UTC. Around and after landfall, Ike took a northwest, north, then northeast path (Fig. 4.1). It weakened to a tropical storm around 1800 UTC September 13 in the eastern Texas. During Ike's inland path, it brought heavy precipitation and damages to 11 U.S. states: Arkansas, Illinois, Indiana, Kentucky, Michigan, Missouri, New York, Ohio, Pennsylvania, Tennessee and West Virginia. Before and after Ike's landfall in the U.S., it caused 29.6 billion dollars of damage and 112 deaths in the U.S., making it the third costliest hurricane in U.S. history, only after Hurricane Katrina (2005) and Hurricane Andrew (1992). This study investigates the effect of assimilating radar reflectivity and radial velocity data from coastal WSR-88D radars using EnKF on the analysis and forecast of Ike.

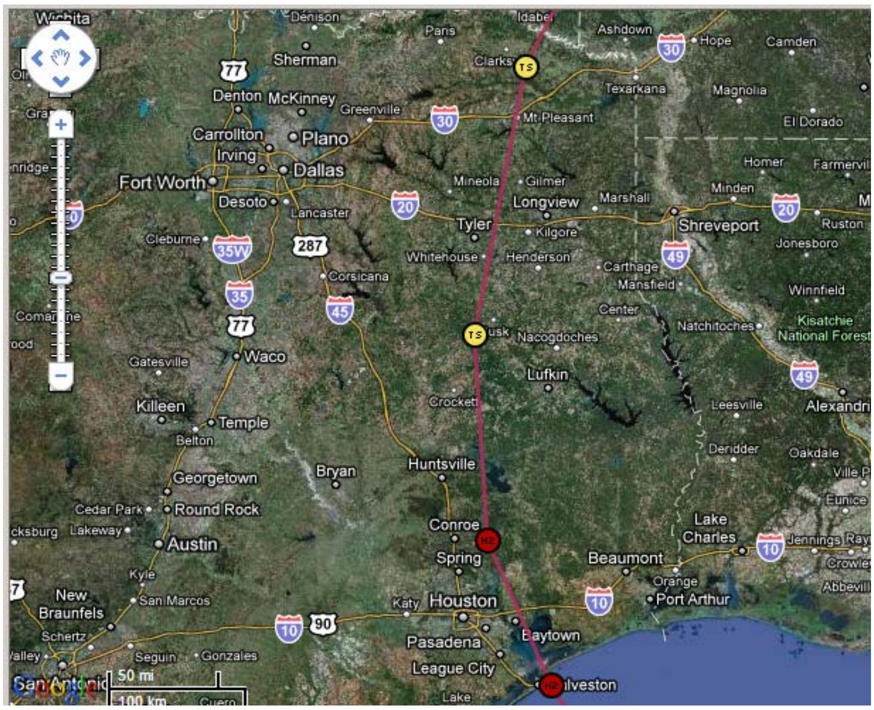

Fig. 4.1 Ike's path in Texas from 0600 UTC Sept. 13 to 0000 UTC Sept. 14, plotted every 6 hours. (from http://stormadvisory.org/map/atlantic/)

4.3 The prediction model and EnKF configurations

4.3.1 The prediction model

The Advanced Regional Prediction System (ARPS, Xue et al. 2000) is used in this study as the prediction model. A 515x515x53 grid with a horizontal resolution of 4 km defines the whole physical domain (Fig. 4.2). Mean vertical grid spacing is 625 m with a vertical grid stretching scheme having a grid spacing of 50 m at the surface. The Lin microphysical scheme (Lin et al. 1983) is used along with the 1.5 TKE-based sub-grid scale turbulence and PBL parameterizations. Details on these physics options can be found in Xue et al. (2001; 2003).

4.3.2 Radar observations

Coastal WSR-88D radars at Houston-Gavelston, Texas (KHGX) and Lake Charles, Louisiana (KLCH) provide the coverage for Ike when it approached to the east Texas coast (Fig. 4.2). From 2008 March to August, WSR-88D (Weather Surveillance Radar, 1988, Doppler) is upgraded with the super resolution capability. The radar data after August of 2008 can provide reflectivity data with a sample size of 0.25 km by 0.5 degree, and increases the range of Doppler velocity data to 300 km in the lower scan elevations, compared with the 230 km of the legacy range. Reflectivity (Z) and radial velocity (Vr) data from these two radars are assimilated with observation errors specified as 2 dBZ and 1 ms^{-1}, respectively.

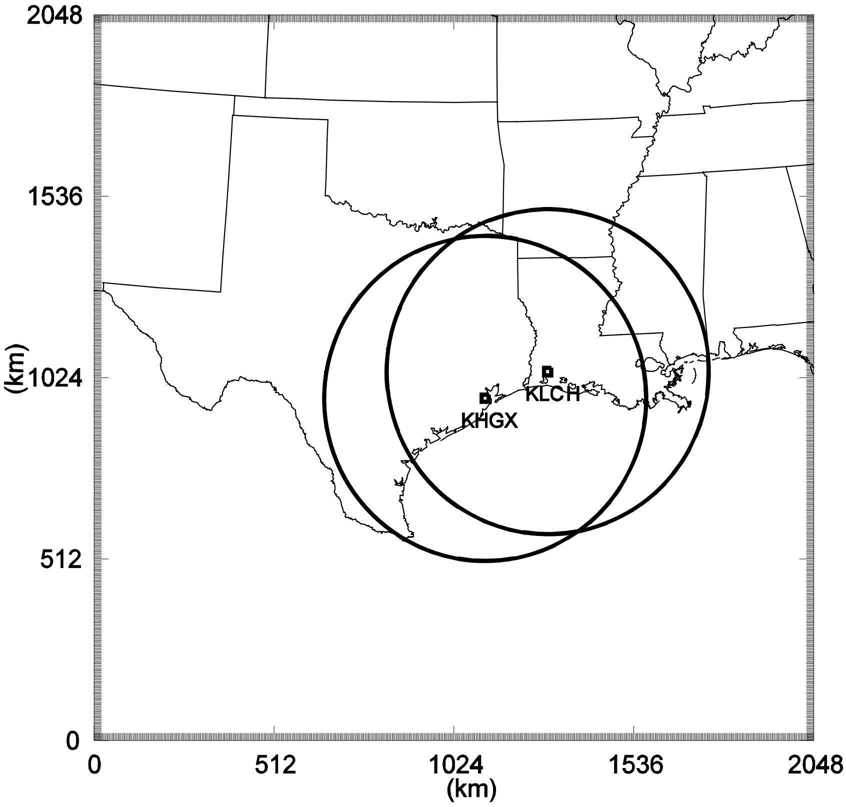

Fig. 4.2. The physical domain and radar coverage for Ike. The circles of KHGX and KLCH both have a maximum range of 460 km.

One problem associated with the Doppler radar data processing is the velocity aliasing. The maximum Doppler velocity measuring interval or Nyquist cointerval or Nyquist velocity, is decided by pulse repetition frequency (PRF) and the radar wavelength λ (Battan 1973):

$$V_{\max} = \pm PRF \times \lambda / 4. \tag{4.1}$$

$V_{\max}$ is proportional to PRF if λ is a constant for the specified radar. For the pulse Doppler radar, the Doppler dilemma is described as

$$V_{\max} \times r_{\max} = \pm c \times \lambda / 8 \tag{4.2}$$

where $r_{\max}$ is the maximum range covered by Doppler radar and c is the speed of electromagnetic wave or light. This expression says that increasing $V_{\max}$ will lead to the decrease of $r_{\max}$ and vice versa. For the typical WSR-88D radar PRF of 1000 Hz, the maximum unambiguous velocity is 25 ms^{-1}. So the environment wind speed with the magnitude more than 25 ms^{-1} will be aliased as $V_{aliased} = V_{real} \pm 2n \times V_{\max}$ where n represents the times the real velocity will be aliased. For the Ike dataset, the Nyquist velocity varies from 23.77 ms^{-1} to 35.55 ms^{-1} for different tilt scans.

In the ARPS program 88D2ARPS, an automatic dealiasing process is provided based on the algorithm proposed by Eilts and Smith (1990). This local environmental dealiasing (LED) technique performs a gate-to-gate and a following 9-point average velocity continuity check (Fig. 4.3) to choose the dealiased velocity value correctly.

However, this technique fails in some of Ike's radar database. The radial velocity cannot be dealiased correctly with this automatic algorithm in 88D2ARPS. NCAR SOLO software package provides the access to peruse and edit the sweep files. The Vr from

Level-II data is further manually quality-controlled for each tilt with SOLO to ensure correct velocity dealising (Fig. 4.4). Although the automatic quality control correct some velocity aliasing (Fig. 4.4a and Fig. 4.4b), some of the area with large negative radial velocities is still aliased after 88D2ARPS. Manual quality control provides the better dealiasing than 88D2ARPS (Fig. 4.4c).

There are other automatic algorithms to dealiase the radial velocity except for the LED technique used in 88D2ARPS. One of them is the variational method by Gao and Droegemeier (2004) which dealises the radial velocity through minimizing a cost function. They proposed this method based on the assumption that the gradients of radial velocity in both radial and azimuthal directions are not aliased. Although more complicated and computationally expensive than the LED technique, this variational method should be more effective in dealiasing than LED method and can save many efforts by otherwise using the manual dealiasing.

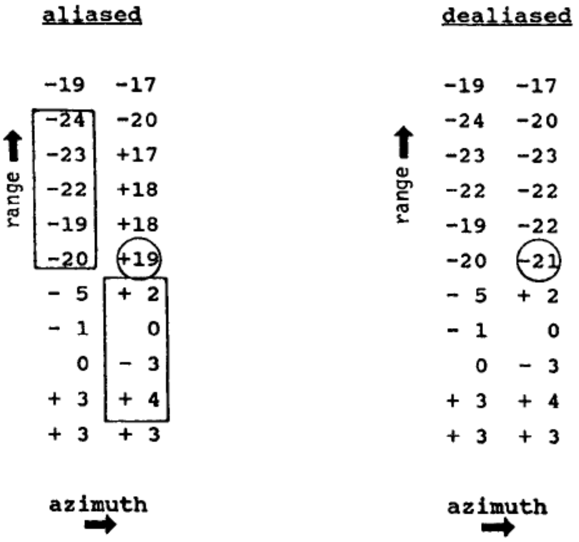

Fig. 4.3. LED technique used in 88D2ARPS. 9-point average in the box is calculated to check the spatial velocity continuity. If the difference between the current velocity (the circled one) and the average falls out of a threshold value, it will be dealiased. (From Eilts and Smith 1990)

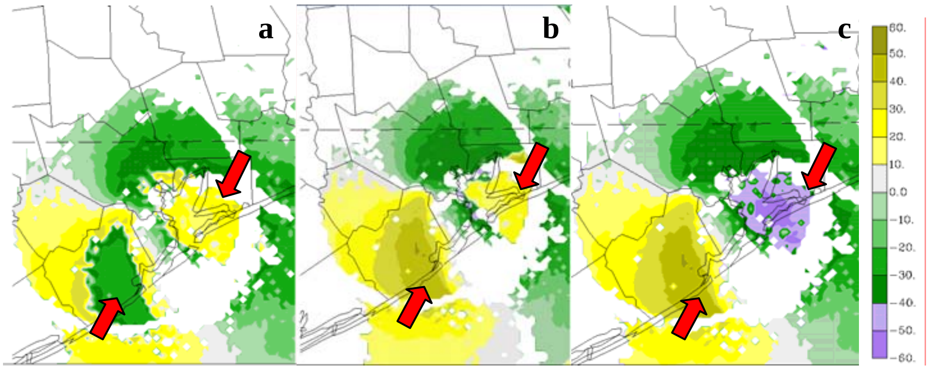

Fig. 4.4. Velocity field at 0.5° elevation angle and at 0410 UTC Sept. 13 for (a) no quality control, (b) with automatic 88D2ARPS quality control and (c) with manual quality control.

After the manual quality controlling, the Level-II data are bilinear-interpolated horizontally on the ARPS model grids. Since the horizontal resolution of the Level-II data (250m) is much smaller than the model resolution of 4 km, only the data closest to the model grid points are chosen and interpolated. No other data thinning schemes are applied in this study. The radar data to be assimilated are on the model grid horizontally and on the tilt vertically after the interpolation. No time interpolation is performed on the data. The radar observations which time is the closest to the regular model output time is used. For example, at 0410 UTC of the model output time, the radar observation with the first tilt scan at 0411 is chosen to be assimilated. How to assimilate the streaming data tilt by tilt is still under further study. In sensitivity experiments, using reflectivity to update the wind component, potential temperature and water vapor mixing ratio brings negative impact on intensity analysis. Hence the reflectivity only updates all the five microphysical variables: mixing ratio for snow, rain, hail, cloud and ice when Z is used. The radial velocity is used to update all of the eleven state variables.

4.3.3 Observation operators for radar observations

The forward observation operator used for Vr is from Tong and Xue (2005b). First the radial velocity on the scalar point of the model is calculated:

$$V_{rg} = u\cos\phi\sin\gamma + v\cos\phi\cos\gamma + (w - w_t)\sin\phi, \tag{4.3}$$

where ϕ is the elevation angle, γ is the azimuth angle, u, v and w are model wind components and w_t is the fall speed of hydrometers. w_t is calculated from

$$w_t = \frac{w_{tr}Z_{er} + w_{ts}Z_{es} + w_{th}Z_{eh}}{Z_{er} + Z_{es} + Z_{eh}}, \tag{4.4}$$

where w_{tr}, w_{ts} and w_{th} are the mass-weighted mean terminal velocities of rain, snow and hail, and Z_{er}, Z_{es} and Z_{eh} are the equivalent reflectivity factors. The calculation of mass-weighted mean terminal velocities follows Eqs (11)-(13) of Lin et al. (1983). After the radial velocity on the scalar point is calculated, the simplified radar emulator with power-gain-based sampling assumption is used to convert the radial winds from the model vertical levels to the radar elevation levels.

Reflectivity is also first calculated on the model vertical levels and then projected on the radar elevation levels using (4.3) and (4.4). The reflectivity formula in TX05 is applied as the Z observation operator to obtain the reflectivity in dBZ on the model levels from the mixing ratios of rain, snow and hail/graupel. First the equivalent reflectivity factor Z_e is calculated as

$$Z_e = Z_{er} + Z_{es} + Z_{eh}, \tag{4.5}$$

where Z_{er}, Z_{es} and Z_{eh} are contributions from rainwater, snow and hail with Eqs (4)-(6) in TX05. Then the logarithmic reflectivity factor is calculated as

$$Z = 10\log_{10}\left(\frac{Z_{er} + Z_{es} + Z_{eh}}{1mm^6m^{-3}}\right). \tag{4.6}$$

4.3.4 Generation of initial ensemble members

The baseline control forecast without radar data assimilation (NoDA) is run from 0600 UTC 13 September, initialized with NCEP GFS analysis. In all other experiments with radar data assimilation, there are two steps for generating and adding perturbations on the GFS analysis. In the first step, Gaussian random perturbations with zero mean smoothed with a 100 km horizontal de-correlation scale are added in the whole domain to initialize the 32-member ensemble at 2200 UTC 12 September. The perturbed state variables include horizontal wind component u, v, potential temperature θ, pressure p and mixing ratio for water vapor q_v. The standard deviations for these perturbations are 2 ms^{-1} for wind, 1 K for θ and 1 hPa for p. For q_v, 10% of the unperturbed value on the model grid is used as the magnitude of the perturbation. Six-hour-long ensemble forecasts were conducted from these perturbed initial conditions to allow evolved background error covariance to develop. The purpose of this procedure is to create reasonable meso-scale environment covariance for Hurricane Ike. In the second step, at 0400 UTC 13 September, another set of perturbations with a smaller horizontal de-correlation scale of 12 km and a vertical de-correlation of 4 km is added to the forecast fields in the observed precipitation regions (Z > 10 dBZ) only to introduce storm-scale perturbations. These storm-scale perturbations also have the Gaussian distribution with zero mean for u, v, vertical wind component w, θ, mixing ratio for water vapor q_v, rain water q_r, cloud water q_c, ice water q_i, snow water q_s and hail water q_h. The standard deviations are 2 ms^{-1} for the wind components, 2 K for θ, 1 gkg^{-1} for all the water content variables except for q_v.

q_v perturbations have the 10% relative perturbations smoothed with the similar small de-correlation scale.

The lateral boundary conditions are from National Center for Environmental Prediction's (NCEP) Global Forecast System (GFS) analysis. They are also perturbed with the 100 km horizontal de-correlation scale to produce 32 ensemble lateral boundary conditions every 6 hours using the perturbations with the same standard deviations as the first step.

	Observation type	Radar assimilation interval	Single or dual radars	Radar localization radius (km)
ExpVr	Vr	10 min.	Dual	12
ExpZ	Z	10 min.	Dual	12
ExpAll	Vr+Z	10 min.	Dual	12
Exp30Min	Vr+Z	30 min.	Dual	12
Exp60Min	Vr+Z	60 min.	Dual	12
ExpKHGX	Vr+Z	10 min.	KHGX only	12
ExpKLCH	Vr+Z	10 min.	KLCH only	12
ExpVrR6	Vr	10 min.	Dual	6
ExpVrR24	Vr	10 min.	Dual	24
ExpZR6	Z	10 min.	Dual	6
ExpZR24	Z	10 min.	Dual	24

Table 4.1. Summary of radar data assimilation experiments for different observation types, radar assimilation intervals and the number of radars used.

The first set of experiments employs 10-minute-long assimilation cycles. Experiments ExpVr, ExpZ and ExpAll assimilate Vr alone, Z alone and both Vr and Z, respectively. In these experiments, the first EnKF analysis of radar data occurs at 0410 UTC, and the assimilation cycles end at 0600 UTC 13 September. Two additional experiments, named Exp30Min and Exp60Min, are performed which has the same assimilation window length but only assimilate both Vr and Z data every 30 and 60 minutes, respectively. Instead of using both Doppler radars together, single radars KHGX and KLCH are also used separately in ExpKHGX and ExpKLCH. To test the sensitivity of horizontal localization radius, 4 additional experiments with various localization radii are also conducted. The experiments are summarized in Table 4.1. At the end of the assimilation window, at 0600 UTC, 18-hour-long deterministic forecast and an ensemble of forecasts are performed until 0000 UTC 14 September. The data assimilation schemes and NoDA control simulation are plotted in Fig. 4.5.

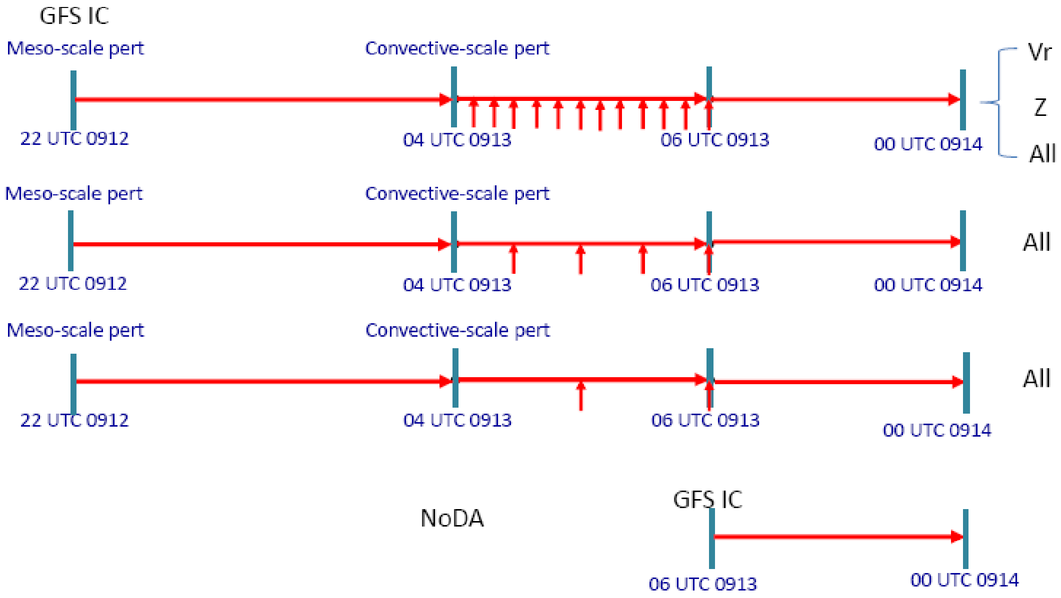

Fig. 4.5. The data assimilation and control simulation NoDA schemes. From top: 10-minute assimilation interval, 30-minute assimilation interval, 60-minute assimilation interval and NoDA simulation.

4.3.5 Covariance inflation and localization

To reduce sampling error caused by the small ensemble, a prior multiplicative covariance inflation of 5% (Tong and Xue 2005; Xue et al. 2006) before the analysis and a posterior additive error after analysis are used on the model grids influenced by radar observations. The multiplicative inflation is added to all the eleven state variables. The additive errors have a Gaussian distribution with zero mean. The standard deviations are 1 ms^{-1} for wind components u, v, w and 1 K for θ. Before added to the posterior analysis, the additive errors are smoothed with a horizontal de-correlation scale of 9 km and a vertical scale of 4 km. Besides increasing the spread artificially, the additive error can also alleviate the model error problem through introducing uncertainty in the model state space. Whitaker and Hamill (2010) found the combination of multiplicative and “evolved” additive inflation has the best results in their study. For most of the experiments, a covariance localization to limit the spatial impact of the observations has cutoff radii of 12 km in the horizontal and 4 km in the vertical. The inflation and localization parameters were chosen based on a number of sensitivity experiments.

4.4 Ensemble spread, observation innovation statistics and analysis increments

The EnKF relies on a sufficiently accurate estimate of the background error covariance to update the state variables. Due to sampling error and the lack of explicit representation of model errors, the forecast ensemble tend to be underdispersive. Maintaining adequate ensemble spread is necessary to prevent filter divergence. An examination of the spread of the state variables in the precipitation region during the analysis cycles of ExpAll (Fig. 4.6) reveals that for the horizontal wind components and pressure, the largest spread reduction by the EnKF analysis occurs in the first two cycles, suggesting more observation impact during those cycles. Despite the gradual reduction in the forecast spread in the subsequent cycles, the variances of the state variables remain at a reasonable level, which was helped by multiplicative and additive inflation.

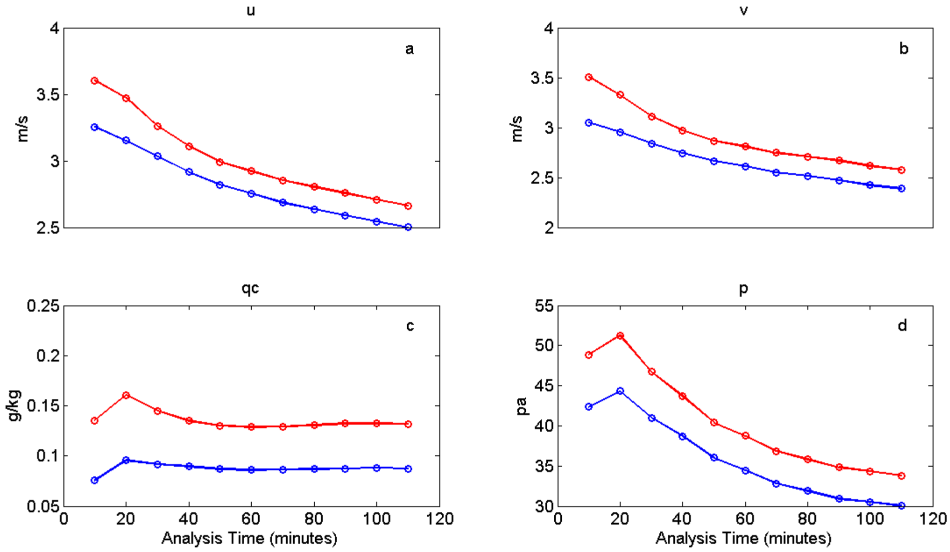

Fig. 4.6. Time evolution of ensemble forecast and analysis spread during the EnKF analysis cycles, spatially averaged in precipitation region (Z > 10 dBZ) for (a) u, (b) v, (c) cloud water mixing ratio (qc) and (d) pressure, from experiment ExpAll. Those for the background forecast are in red and those for analysis are in blue.

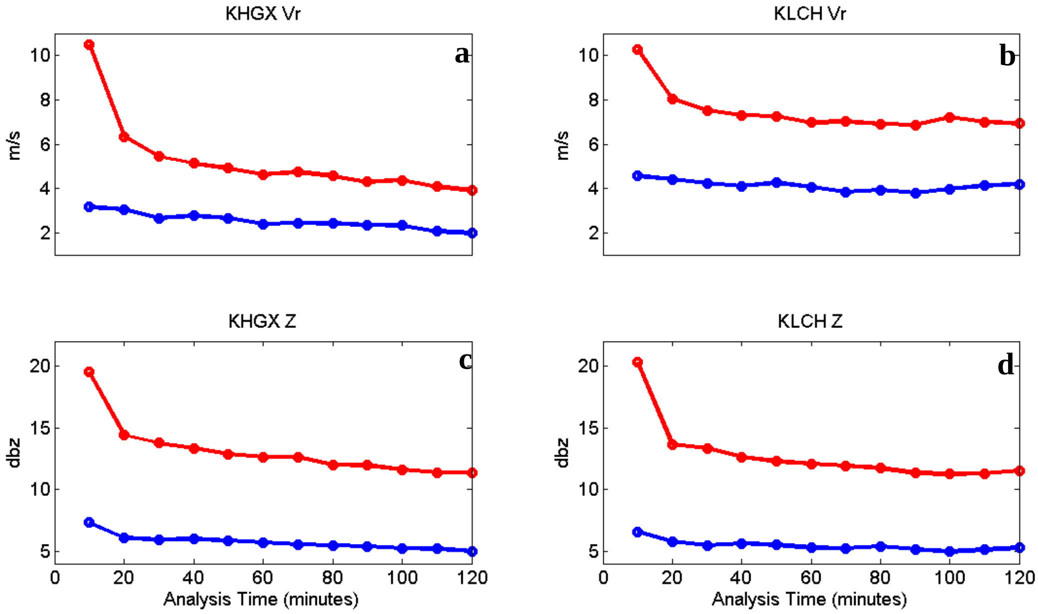

Fig. 4.7. Time evolution of innovation rms during the analysis cycles, averaged in precipitation region (Z > 10 dBZ) for (a) Vr of KHGX and (b) KLCH, (c) Z of KHGX and (d) KLCH from experiment ExpAll. Those for the background forecast are in red and those for analysis are in blue.

The observation innovations, or the root-mean-square (rms hereafter) difference between observations and the model states projected to the observation space, denoted as y-H(x), measure how well the model state fit the observations. The rms innovations of ExpAll averaged in precipitation region are shown in Fig. 4.7. for the background forecasts and analyses. With respect to both radars, the rms innovations of both Z and Vr have the largest reduction in the first two assimilation cycles. After 10 to 20 minutes of forecast and analysis, the innovation reductions remain and continue until the end of the analysis cycles. At the end, the rms innovations of Vr and Z are 2 to 4 ms^{-1} and 5 dBZ, respectively, which are much smaller than the initial values of about 10 ms^{-1} and 20 dBZ, respectively. This says that both the forecast and analysis states are significantly improved by the EnKF data assimilation in terms of the fit to observations.

To better understand the behavior of radar data analysis, the increments of horizontal wind components in the first and last analysis cycles are plotted in Fig. 4.8. It is found that during the first analysis at 0410 UTC, the horizontal wind increments appear to systematically enhance the hurricane vortex, with the increments having a well-organized structure of cyclonic rotation (Fig. 4.8a). At the end of the analysis cycles, the error in the overall vortex of the background forecast has been significantly reduced by this time. The wind increments are much less organized, indicating that most of the corrections now correspond to storm-scale structures at the sub-vortex scale (Fig. 4.8b).

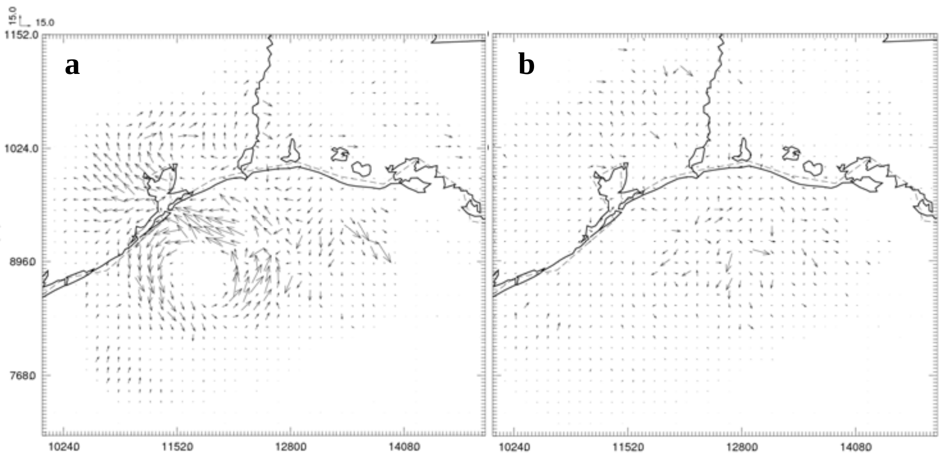

Fig. 4.8. Horizontal wind component increment at z=3km for (a) the first analysis and (b) the last analysis of ExpAll.

4.5 Data assimilation impact on the analysis and deterministic forecast

4.5.1 Impact on analyzed and forecasted hurricane structures

The composite reflectivity and horizontal wind vectors at the 3 km height from NoDA, ExpVr, ExpZ and ExpAll are presented in Fig. 4.9, together with the observed composite reflectivity (OBS). The stronger and tighter inner cores in the final analyses are identified with all radar data assimilation experiments, compared that in the GFS analysis at 0600 UTC (Fig. 4.9a-e). The reflectivity field in ExpVr shows a broader and stronger rainband than observations (Fig. 4.9c). There is no reflectivity in the GFS analysis for comparison. ExpAll and ExpZ have similar rainband structures and are closer to the observations than ExpVr (Fig. 4.9d-e). This difference is also reflected in rms innovations of Z at this time (not shown), where for both radars, rms innovation of Z is about 15 dBZ in ExpVr and only about 5 dBZ in ExpAll and ExpZ.

In the 6-hour forecast, the center of Ike is over the land north of Houston. Generally, all experiments with radar data assimilation display a more tightly wrapped rainband than NoDA. NoDA also has a spiral rainband on the north-west of the vortex center, which is too strong compared with the observation (Fig. 4.9f-j).

During 6 additional hours of forecast, the rainband in Ike moves further inland and an axis-asymmetric structure is seen on Fig.4.9k. Two major precipitation regions covering eastern Texas develops in the north-west and south-east quadrants around the vortex center. A clear-air hole without precipitation is visible in the vortex center in NoDA (Fig. 4.9l). With the radar data assimilated, the precipitation patterns are closer to the observations (Fig. 4.9m-o). Amid the radar data assimilation experiments, ExpZ has a broader precipitation region in the south-east quadrant and a tighter inner core, more similar to the observations (Fig. 4.9n).

At the final forecast time of 0000 UTC 14 September, most of the precipitation is out of Texas. The interaction with the cold front system to the north and the moisture transport from the Gulf leads to rainfall in Oklahoma and Arkansas and a more axis-asymmetric structure (Fig. 4.9p). The clear-air hole in the vortex center is still identifiable in NoDA (Fig. 4.9q). The rainbands in experiments with radar data assimilated are still more tightly wrapped. Like the observations, the hurricane eyes in ExpZ and ExpAll are filled with precipitation, and precipitation patterns in these two experiments are the closest to the observations (Fig. 4.9s and t).

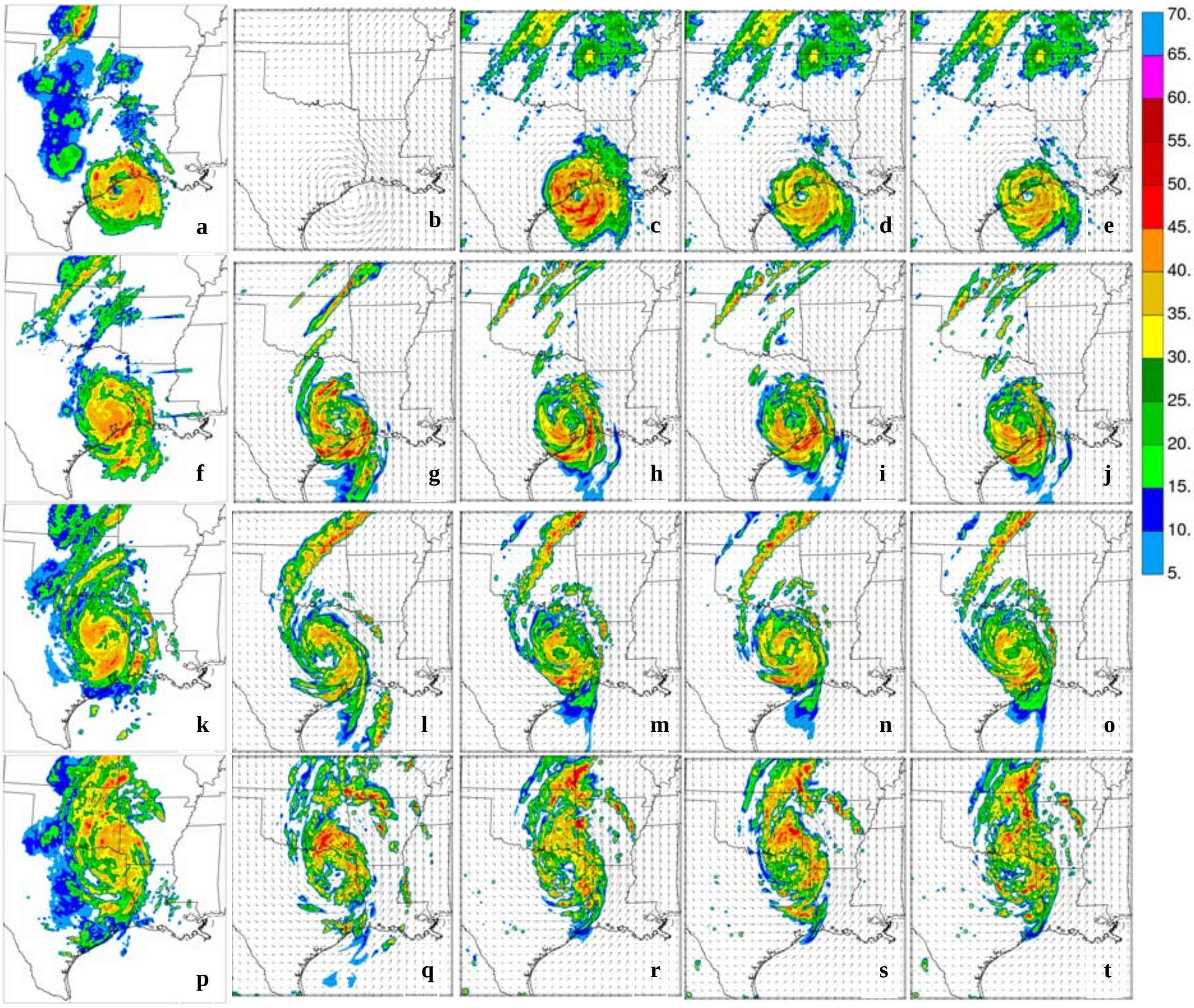

Fig. 4.9. Composite reflectivity (color shaded) and wind vectors at 3 km height analyzed and predicted by experiments (b, g, l and q) NoDA, (c, h, m and r) ExpVr, (d, i, n and s) ExpZ, and (e, j, o, t) ExpAll, as compared with (a, f, k and p) corresponding observations. The times shown are 0600, 1200, 1800 UTC, September 13 and 0000 UTC September 14, 2008.

4.5.2 Intensity and track forecasting

The minimum sea level pressure every 3 hours during the 18 hours of forecast from all experiments are plotted in Fig. 4.10, along with the best track minimum sea level pressure from the National Hurricane Center. All experiments with radar data exhibit solid improvement to NoDA during the first 12 hours in intensity forecasts. The analyzed intensities of 955 hPa in ExpVr and ExpAll at 0600 UTC are significantly lower than the 975 hPa of NoDA, although still somewhat higher than the best track value of 951 hPa. Assimilation of Z alone leads to a mild improvement of 9 hPa over NoDA at 0600 UTC, resulting in a weaker vortex than assimilating Vr or Vr plus Z. The relative improvement in intensity is defined as

$$(\text{intensity\_error}_{\text{NoDA}} - \text{intensity\_error}_{\text{EnKF}}) / \text{intensity\_error}_{\text{NoDA}}. \quad (4.7)$$

The relative improvements over NoDA are 80% for ExpVr, ExpAll and 55% for ExpZ in intensity at 0600 UTC.

The intensity of NoDA is too low and does not change dramatically during the first 12 hours of forecast while Ike in the best track data keeps weakening until 2100 UTC. ExpVr and ExpAll both capture the pressure rise at similar rates as the best track data before 1500 UTC. Between 1500 to 2100 UTC, the best track data show faster weakening than earlier, which is not reflected in any of the radar data assimilation experiments. The prediction model error may contribute to this discrepancy, in addition to possible initial condition error. At 1800 UTC, the fast weakening best track catches up with ExpVr and ExpAll in intensity, leading to an almost zero intensity error for these two experiments. After 1800 UTC, ExpVr and ExpAll both forecast stronger vortices than the best track due to the slower weakening in these two experiments. When both reflectivity and radial velocity are assimilated (ExpAll), the intensity forecasts are very close to the case assimilating radial wind alone (ExpVr), with a less than 0.5 hPa improvement in ExpAll over ExpVr at 0600 and 0900 UTC.

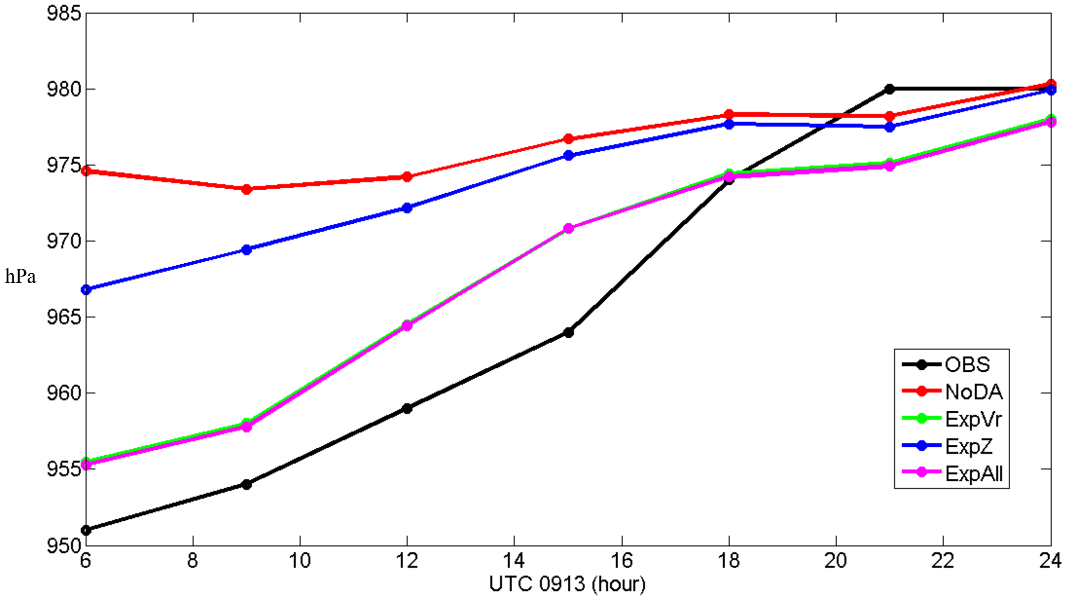

Fig. 4.10. The predicted minimum sea level pressure for Hurricane Ike, plotted every three hours from 0600 UTC September 13 to 0000 UTC September 14.

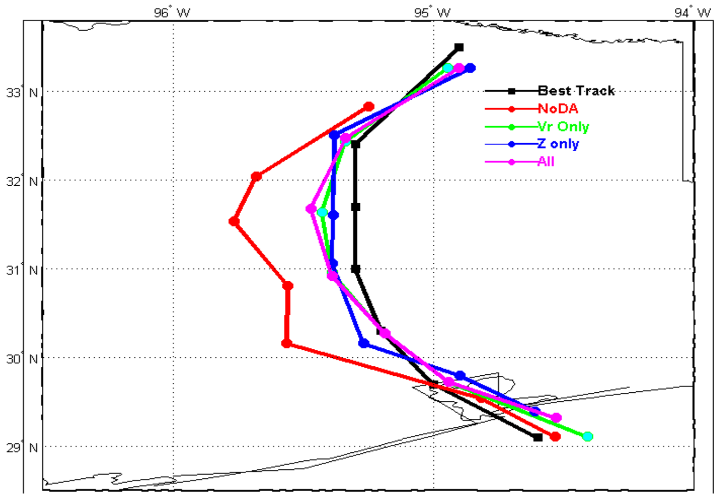

Fig. 4.11. The predicted track for Hurricane Ike, plotted every three hours from 0600 UTC September 13 to 0000 UTC September 14.

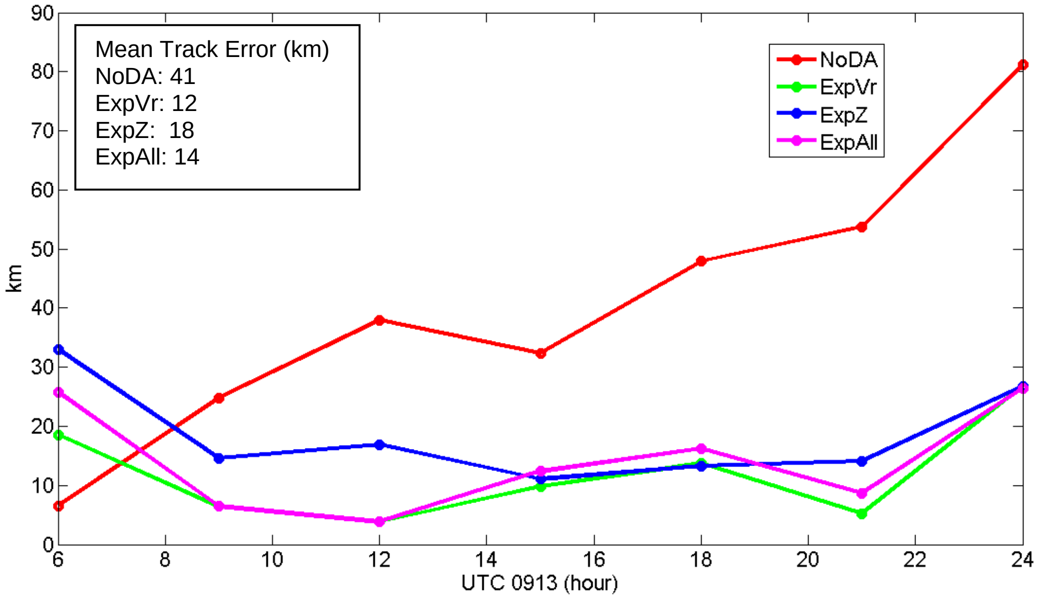

Fig. 4.12. The predicted track error for Hurricane Ike, plotted every three hours from 0600 UTC September 13 to 0000 UTC September 14.

The predicted tracks from all experiments are plotted in Fig. 4.11, along with the best track. Even with a quite accurate initial position of only 7 km track error from the GFS analysis at 0600 UTC 13 September, NoDA takes a west-most path in the 18-hour forecast. In the first 3-hour forecast from 0600 to 0900 UTC during and after the landfall, NoDA moves slower than all data assimilation experiments. Starting from 0600 UTC, NoDA takes a more west path than the best track and all the data assimilation experiments. The track error of NoDA increases with the forecast time and reaches 80 km at 0000 UTC 14 September (Fig. 4.12). With radar data assimilation, the track errors at 0600 UTC are all larger than NoDA. One problem related to the larger track error at the initial time is identified. The initial track of the deterministic forecast is determined by finding the minimum sea level pressure center in the mean field of the 32 member ensemble. After averaging the members, the mean field exhibits an elongated vortex owing to the hurricane position spread in the ensemble members. This creates some uncertainty with the vortex center estimation. An average of the tracks among all individual members provides a better result with an initial track error reduction of 10 km for ExpAll (Fig. 4.17 and Fig. 4.18). At 0900 UTC after the hurricane landfall, all predicted tracks in the data assimilation experiments are closer to the best track than in NoDA, with the track errors all being less than 20 km. The track errors in ExpVr and ExpAll are both less than 10 km. From 1200 to 2100 UTC, the best track took a north path toward Tyler, Texas. ExpZ captures the moving direction very well during the 9 hours but with a west shift of 15 to 18 km. ExpVr and ExpAll both have a minimum track error of less than 5 km at 1200 UTC but take a more curved path than ExpZ from 1200 to 2100 UTC. All of the three assimilation experiments are in the south of the best track at 0000 UTC Sept. 14 with a track error of 27 km. Although ExpZ has a larger mean track error of 18 km compared to ExpVr and ExpAll, it is encouraging to see that assimilating Z alone still results in 56% improvement in track forecasting on average over NoDA.

Pu et al. (2009) observed a marginal improvement in Hurricane Dennis (2005) intensity forecast when assimilating reflectivity data alone with 3DVAR. The impact from reflectivity alone on intensity forecast is smaller than assimilating radial wind alone or both radial wind and reflectivity. They attributed part of the reason to the small impact of reflectivity data alone on the track forecast. In our study, the track of ExpZ at 0600 UTC is too north to the coast, and has a larger track error than in ExpVr and ExpAll (Fig. 4.11), suggesting a similar reason at work. This track error of ExpZ at 0600 UTC can lead to a moderate impact on intensity forecast from assimilating reflectivity alone. The better vortex wind structure from assimilating radial wind is another possible reason for more accurate intensity forecasts from ExpVr and ExpAll. Radial velocity observes wind structure directly and wind field of the vortex appears to be dominant for intensity analysis and forecasts in this case. The microphysics fields observed by reflectivity have a direct impact on rainband structure analysis but mild influence on intensity analysis.

4.5.3 Precipitation forecasting

Flooding caused by Ike was one of the major culprits of deaths and economy loss, highlighting the importance of precipitation forecast. Fig. 4.13 shows the 18-hour accumulated precipitation for all experiments along with the Stage IV precipitation data. The observations show that the maximum accumulated rainfall is positioned around Huntsville and Conroe, Texas, north of Houston (Fig. 4.13a). NoDA fails to predict this strong rainfall region completely (Fig. 4.13b). Assimilation of radar data helps to capture this intense precipitation area in the three data assimilation experiments although the strength and area coverage are under-predicted (Fig. 4.13c-e). For lighter or stratiform precipitation, it is not easy to tell which experiment has a better prediction.

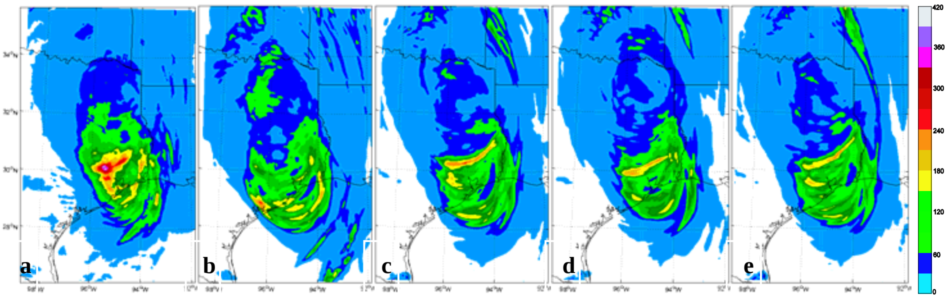

Fig. 4.13. 18-hour accumulated precipitation forecast from 0600 UTC September 13 to 0000 UTC September 14 for (a) observations, (b) NoDA, (c) ExpVr, (d) ExpZ and (e) ExpAll.

To quantify the precipitation forecast skills, equitable threat scores (ETS hereafter) against the Stage IV data for 3-hourly accumulated precipitation are calculated and plotted for all the experiments in Fig. 4.14. A threshold of 30 mm is chosen to present convective rainfall. In the first 6 hours of forecast, all the experiments with radar data assimilated have higher ETSs than NoDA. From 1200 UTC to 1800 UTC, the score is still higher in ExpZ while those of ExpAll and ExpVr are close to that of NoDA. There are increases in ETS scores for ExpAll and ExpVr from 1800 to 2100 UTC, which needs further examination. At the end of forecast, all the experiments have their scores below 0.1, partly due to the shrinking area of convective precipitation at this time; in this situation, small position errors can lead to very low scores.

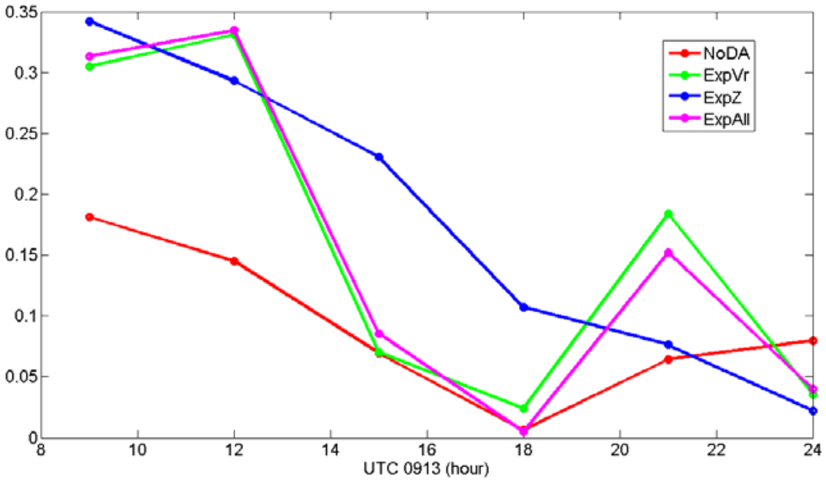

Fig. 4.14. ETS of 3-hour accumulated precipitation at the 30 mm threshold for NoDA, ExpVr, ExpZ and ExpAll.

The ETS of 18-hour accumulated precipitation is also calculated (Fig. 4.15) for four thresholds ranging from 30 mm to 120 mm. Lower thresholds represent more stratiform precipitations while higher thresholds more convective precipitations. It is noted that for all the thresholds, radar data assimilation helps to improve the quantitative precipitation forecast. The larger the threshold is, the stronger the relative improvement is, implying more importance in improving convective precipitation forecast than stratiform precipitation forecast. For the 120 mm threshold, the relative improvements of three radar data assimilation experiments over NoDA are all around 300%.

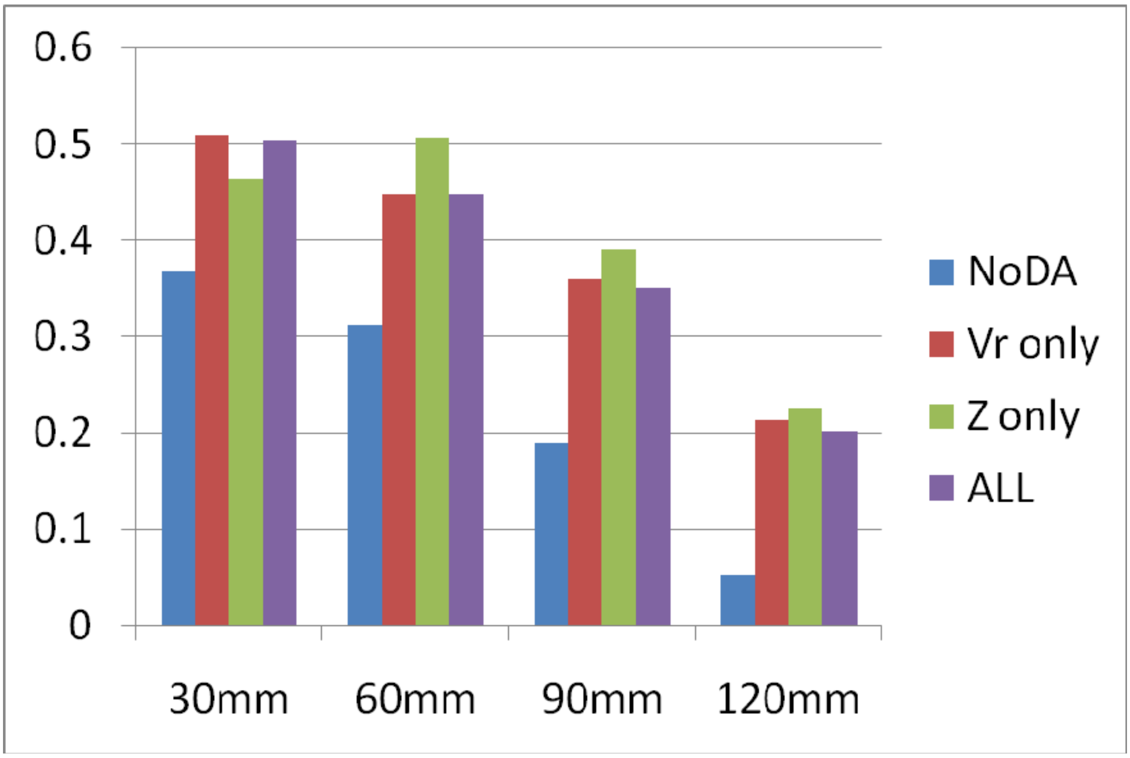

Fig. 4.15. ETS of 18-hour accumulated precipitation 0600 UTC September 13 to 0000 UTC September 14 at the threshold of 30 mm, 60 mm, 90 mm and 120 mm for NoDA, ExpVr, ExpZ and ExpAll.

4.6 Ensemble forecasts

4.6.1 Ensemble forecasts of intensity and track

The EnKF provides an ensemble of analyses which can be used to initialize an ensemble of forecasts. Ensemble forecasts of 32 members are therefore carried out from the 0600 UTC analyses of ExpAll. In this section 4.6, all the ensemble forecasts are started from ExpAll. The intensities, tracks and track errors of the ensemble forecasts are plotted in Fig. 4.16 to Fig. 4.18. The results of the deterministic forecast starting from the ensemble mean analysis and the mean of ensemble forecasts are also shown for comparison. The intensity spread of 0.7 hPa (indicated by minimum sea level pressure) at the end of the analysis is relatively small compared with the intensity error of 4 hPa. One of the reasons is that there is no enough synoptic scale variance in the environment. The addition of synoptic scale perturbation from global ensemble forecast error statistics in the hurricane environment may help to alleviate this underestimation problem. Because Ike is in a weakening stage during this forecast period, the spread of the ensemble forecasts in intensity did not increase noticeably with time (Fig. 4.16). It is also found in other studies that the intensity error growth of a decaying hurricane system is not as strong as an intensifying stage and the ensemble spread tends to decrease with time (see Fig. 12. of Zhang et al. 2009). The mean of ensemble minimum sea level pressures is similar with that of the deterministic forecast.

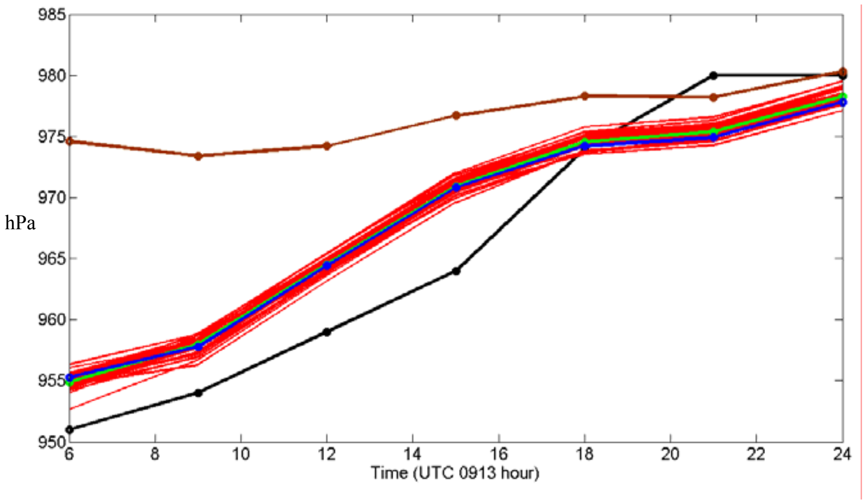

Fig. 4.16. The predicted ensemble minimum SLP of ExpAll (red), compared with the best track (black), NoDA (brown), ensemble average (green) and the deterministic forecast (blue).

An increase in uncertainty is observed in the ensemble track forecasts (Fig. 4.17). The vortex centers are more separated at the end of forecast than at the start. The spread of the ensemble track is defines as

$$\sqrt{\frac{1}{N-1}\sum_{i=1}^{n} d_i^2} \tag{4.8}$$

where N is the ensemble size and d_i is the distance between the track of the *i* th member and the mean track. The track spread increases during most of the forecast time except for a temporary reduction in the first 3-hour forecast (Fig. 4.19). The track spread at the end of 18-hour forecast increases by 75% over the spread at the start. This trend is also reflected in the predicted track error. It should be noted that the calculation of track error spread may conceal the actual large track spread among the members where two widely separated vortex center can have similar track errors. As mentioned before, the average position of ensemble members is closer to the best track than the single ensemble mean at 0600 UTC. The predicted track of the ensemble average also has certain improvement over the deterministic forecast at 1500 and 1800 UTC (Fig. 4.18), suggesting the potential benefit of using ensemble mean rather than the deterministic forecast for hurricane track prediction.

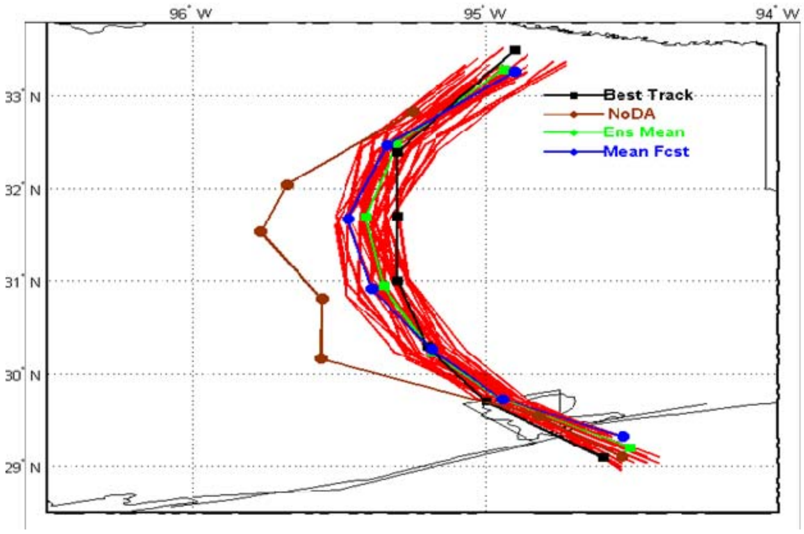

Fig. 4.17. The predicted ensemble track of ExpAll (red), compared with the best track (black), NoDA (brown), ensemble average (green) and the deterministic forecast (blue).

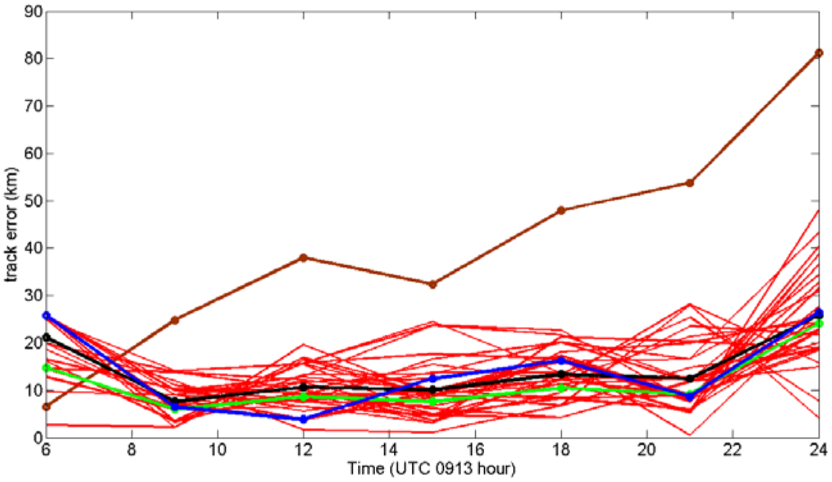

Fig. 4.18. The predicted ensemble track error of ExpAll (red), compared with the best track (black), NoDA (brown), ensemble average (green) and the deterministic forecast (blue).

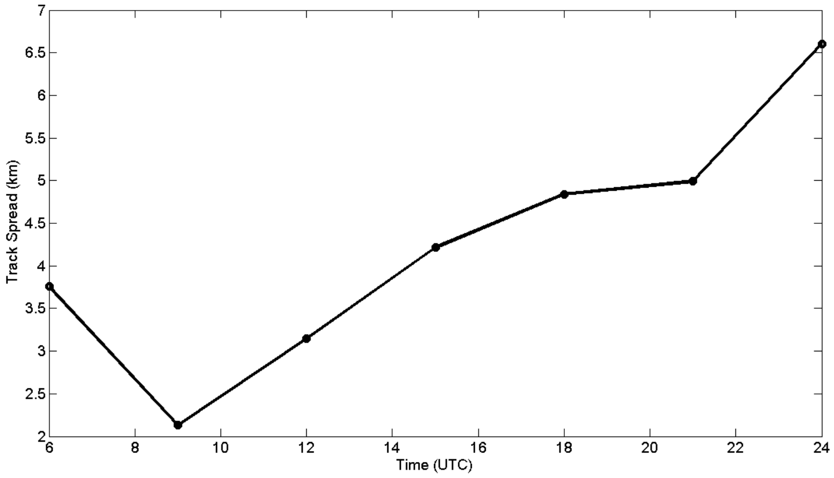

Fig. 4.19. The spread of ensemble tracks for ExpAll.

4.6.2 Correlation of intensity and track forecast

The ensemble forecasts provide an approach for the uncertainty and sensitivity study. The correlation coefficients between the initial intensity error and the intensity error of different forecast times are plotted in Fig. 4.20. There is a moderate correlation greater than 0.5 at the first 3-hour forecast. The correlation decreases gradually to 0.2 after 12 hour forecast, indicating the nonlinearity in intensity forecasts after hurricane landfall. The negative correlation appears at 2100 UTC 13 September and 0000 UTC 14 September when the best track has a weaker vortex than the forecast. The stronger vortices have smaller errors at early times but have larger errors at the end since the best track is weaker than the forecasts, resulting in the negative correlation.

The similar figure is also plotted for the track forecasts (Fig. 4.21). The correlations with the initial track error during most of the forecast time are around zero, reflecting a very small sensitivity in the track forecast. Also, the correlations between the intensity error and the track error with time are shown in Fig. 4.22. The small correlations during the 18 hours of forecast suggest that the improvement in intensity forecast does not guarantee an improvement in track forecast.

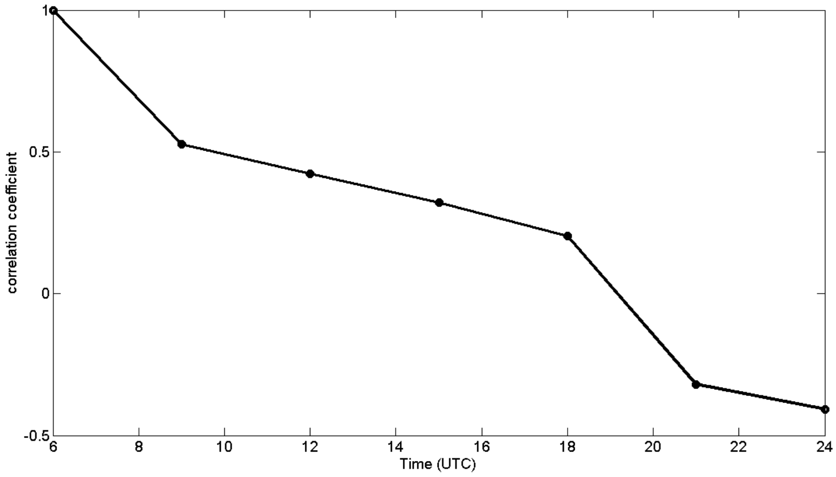

Fig. 4.20. The correlation coefficients between the initial intensity error and the intensity error of different forecast times for ExpAll.

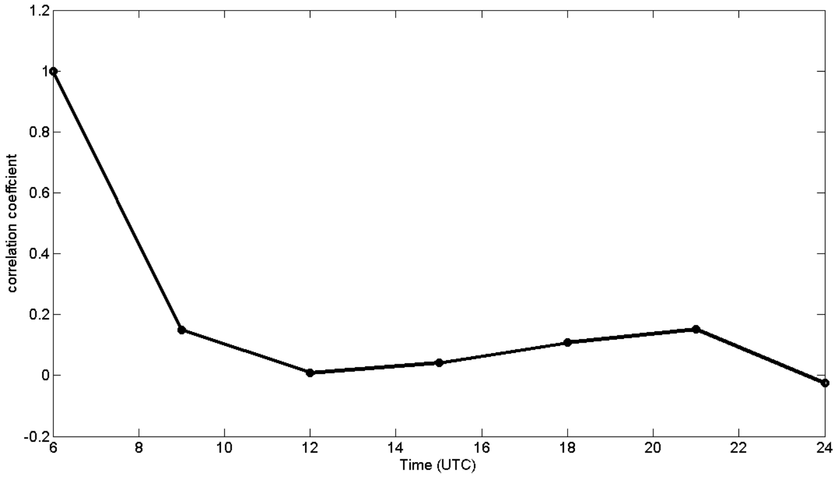

Fig. 4.21. The correlation coefficients between the initial track error and the track error of different forecast times for ExpAll.

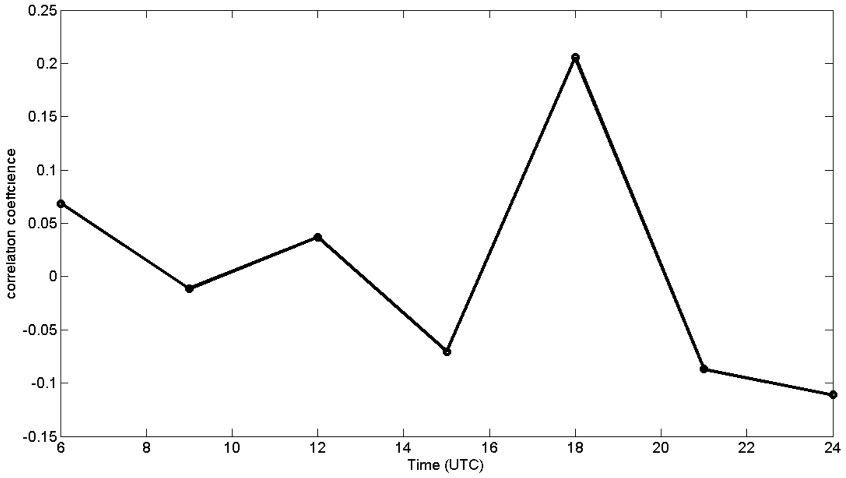

Fig. 4.22. The correlation coefficients between the track error and the intensity error of different forecast times for ExpAll.

4.6.3 Ensemble Forecasts of precipitation

ETSs are also calculated for the ensemble forecasts (Fig. 4.23). At most forecast times, most of the ensemble members have higher scores than NoDA. Similar to using ensemble track errors, we should be very cautious when using ETSs to estimate the precipitation forecast uncertainty as two highly different precipitation forecasts can have very similar ETSs. The mean of the ensemble ETSs is also shown in Fig. 4.23 (black). The mean ETS trend follows most of the ensemble members with no surprise. The ETS of ensemble mean is also calculated and plotted (green) in Fig 4.23. The ETS of mean precipitations is close to mean ETS but with lower scores at 1500 UTC and 1800 UTC and a higher score at 2100 UTC.

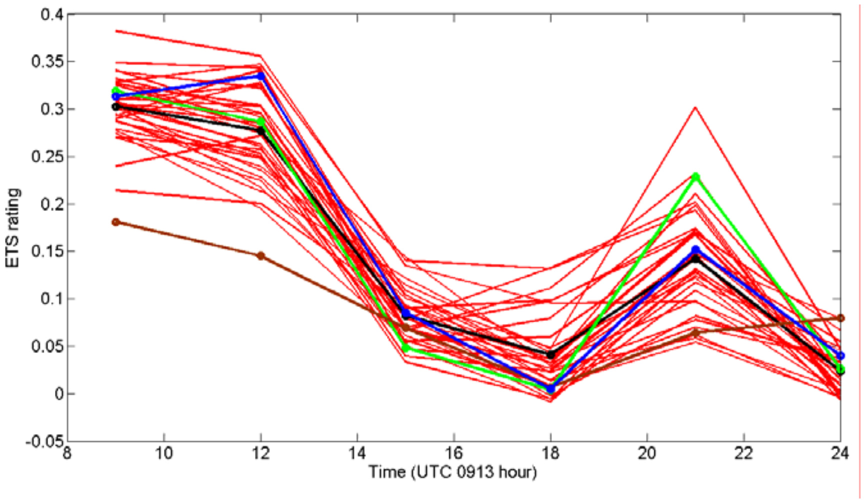

Fig. 4.23. The ensemble ETSs of 3-hour accumulated precipitation of ExpAll (red), against NoDA (brown) and the deterministic forecast (blue). The mean ETS (black) and the ETS of ensemble mean precipitation (green) is also plotted.

In high resolution numerical model simulations, the precipitation field tends to have high spatial and temporal variances among ensemble members. One problem related with the ensemble mean of precipitation is the excessively broad rainfall areas and weak rainfall magnitudes owing to averaging (Kong et al. 2009). In our study, the 18 hour accumulated ETS of the ensemble mean is larger than that of the deterministic forecast for 30 mm and 60 mm threshold (Fig. 4.25), indicating a better precipitation forecast for lighter rainfalls since the spatial difference among ensemble members leads to a broad rainfall region with light precipitations. For 90 mm and 120 mm threshold, however, the ETS of the mean is smaller than the deterministic forecast (Fig. 4.25), a result from the smaller convective rainfall amount caused by averaging among ensemble members.

Probability matching (PM, hereafter) provides a useful way to re-organize the precipitation fields among ensemble members and improve the quantitative precipitation forecast (QPF). Probability matching sets the probability distribution function (PDF) of the less accurate dataset to be the same as that of the more accurate dataset (Ebert 2001). The PM method in this study follows Ebert (2001) by assuming the best spatial distribution of rainfall field is represented by the ensemble mean while the best frequency distribution is given by the ensemble QPFs. PM algorithm first combines the accumulated precipitations from all the ensemble members on the selected domain and ranks the precipitations from the greatest to the smallest to obtain the PDF of the accumulated precipitations. This ensemble of precipitations is saved as array 1. Similarly, the precipitations of the ensemble mean over the same domain are also ranked with the highest to the lowest order, saved as array 2. Then the grid point with the highest value in array 2 is reassigned to the highest precipitation in array 1, and so on. The grid point with the n th value in array 2 is reassigned to the n×m th value in array 1, where m is the ensemble size.

With PM, the 3-hour accumulated rainfall ETS at a 30 mm threshold has slight but solid improvement consistently over the regular ensemble mean for all of the forecast time (Fig. 4.24). For the 18-hour accumulated precipitation ETS (Fig. 4.25), PM has higher scores at 60 mm, 90 mm and 120 mm threshold than the regular mean. At 120 mm threshold which represents strong rainfalls, the experiment with PM has the largest relative improvement of 49% to the experiment without PM. Overall, PM has the best ETS result for all four thresholds except for 120 mm, where the deterministic forecast is slightly better than PM.

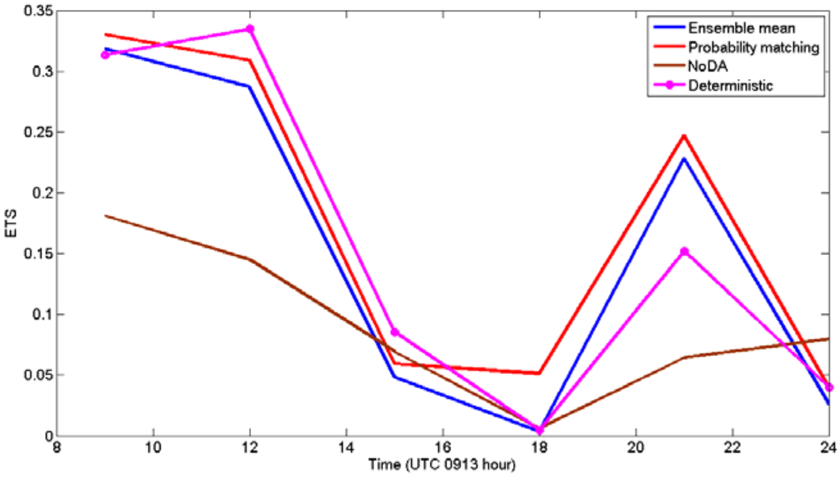

Fig. 4.24. ETS of 3-hour accumulated precipitation at the 30 mm threshold for NoDA (brown), deterministic forecast (magenta), mean of ensemble forecast (blue) and probability matching of ensemble forecast (red).

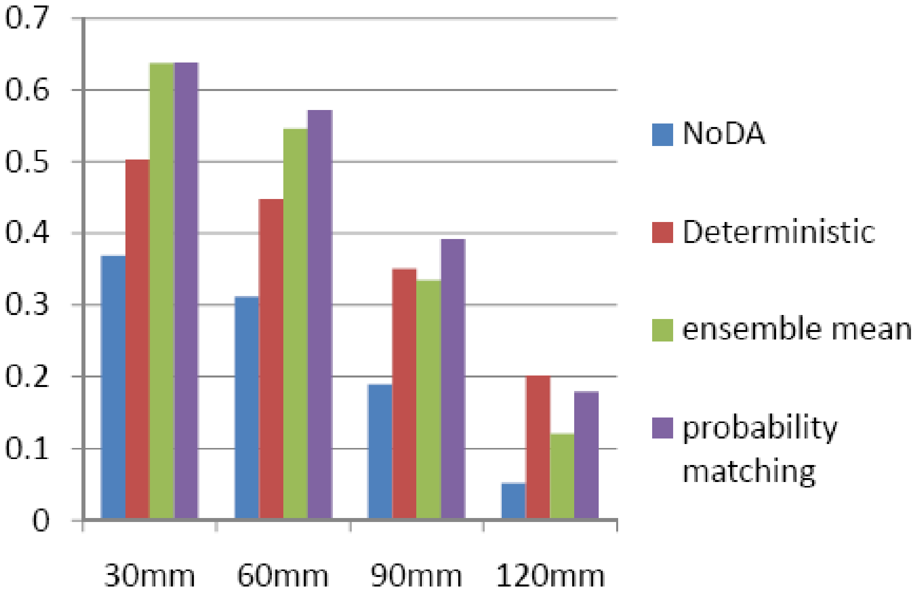

Fig. 4.25. The ETS of 18-hour accumulated precipitation for NoDA (blue), deterministic forecast of ExpAll (red), mean of ExpAll ensemble forecast (green) and probability matching of ExpAll ensemble forecast (purple) at various thresholds.

4.7 Sensitivity to localization cutoff radius

One important issue with EnKF assimilation of radar data for hurricanes is the optimal choice of the covariance localization radius. Usually, the localization radius has to be large enough to avoid the "bull's eyes" in the analysis, and should be small enough so as to avoid the use of unreliable covariance due to sampling error. The localization radius used in the previous experiments is mostly based on experiments assimilating radar data for individual thunderstorms at similar resolutions. It is not necessarily optimal for hurricanes which include flows of multiple scales. In our current study, the experiment assimilating reflectivity alone (ExpZ) has a smaller positive impact on the intensity and track forecasts than assimilating radial velocity alone (ExpVr). It is not clear if a more optimal localization radius exists that lead to a larger impact of reflectivity data. Sensitivity experiments with localization radius are conducted to find this out.

The localization radius is changed from the 12 km used in the control experiment ExpZ to 24 km and 6 km in ExpZR24 and ExpZR6, respectively (Table 4.1). Correspondingly, ExpVrR24 and ExpVrR6 are performed which are the same as ExpVr but using 24 and 6 km localization radius respectively. The predicted minimum sea-level pressures from these experiments are plotted in Fig. 4.26 and Fig. 4.27.

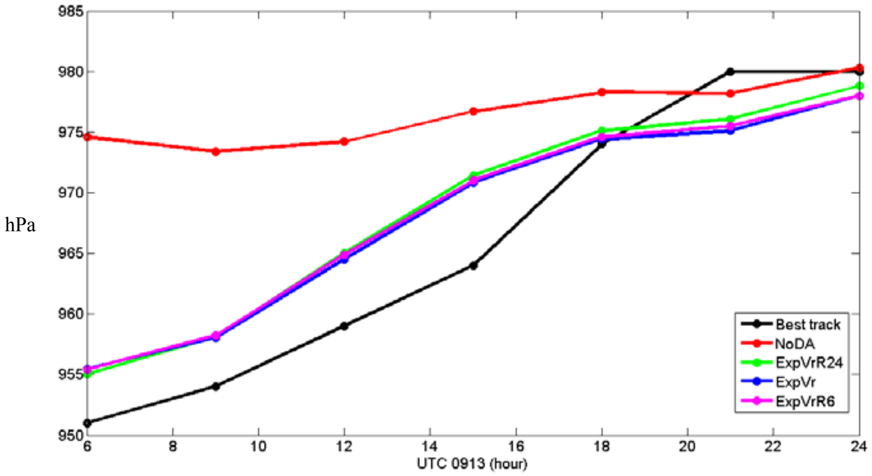

Fig. 4.26. The predicted minimum SLP of NoDA (red), ExpVrR24 (green), ExpVr (blue) and ExpVrR6 (also Exp10Min; magenta), compared with the best track (black).

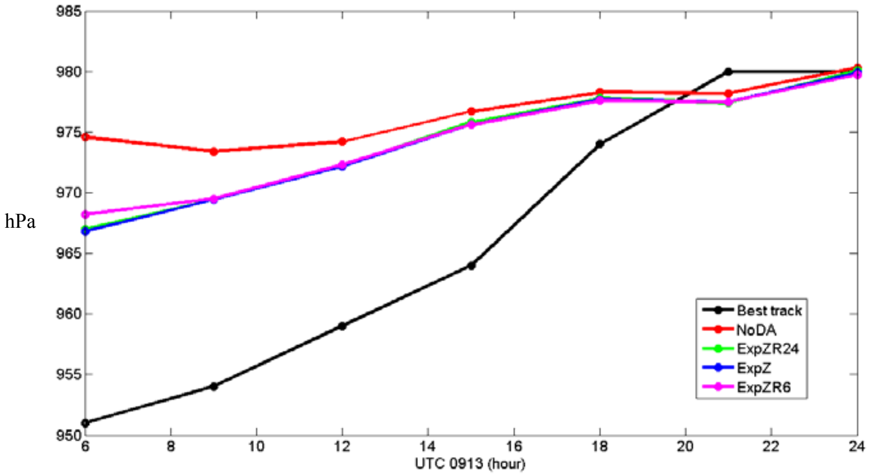

Fig. 4.27. The predicted minimum SLP of NoDA (red), ExpZR24 (green), ExpZ (blue) and ExpZR6 (also Exp10Min; magenta), compared with the best track (black).

Among the experiments assimilating Vr only, ExpVrR24 and ExpVrR6 predict almost the same minimum sea level pressures as ExpVr during the first 6 hours of forecast (Fig. 4.26). After 1200 UTC September 13, ExpVrR24 predicts a vortex that is nearly 1 hPa weaker than those in ExpVr and ExpVrR6. Among the experiments assimilating Z only, the analyzed minimum sea level pressure in ExpZR6 is slightly higher than in ExpZ and ExpZR24 at 0600 UTC. The forecast intensities in these three experiments are similar during the entire 18 hours of forecast.

The predicted tracks and track errors for Vr-assimilating experiments are plotted in Fig. 4.28 and Fig. 4.29. The analyzed vortex center in ExpVrR24 at 0600 UTC is to the northeast of the centers of ExpVr and ExpVr6, and is closer to the coast (Fig. 4.28). The track error of ExpVrR24 at 0600 UTC is a little larger than ExpVr and ExpVr6 while ExpVrR6 has the smallest track error (Fig. 4.29). During the 18 hours of forecast, these three experiments have similar tracks and track errors, with the mean track errors being 12.7 km, 12.1 km and 11.4 km for ExpVrR24, ExpVr and ExpVrR6, respectively.

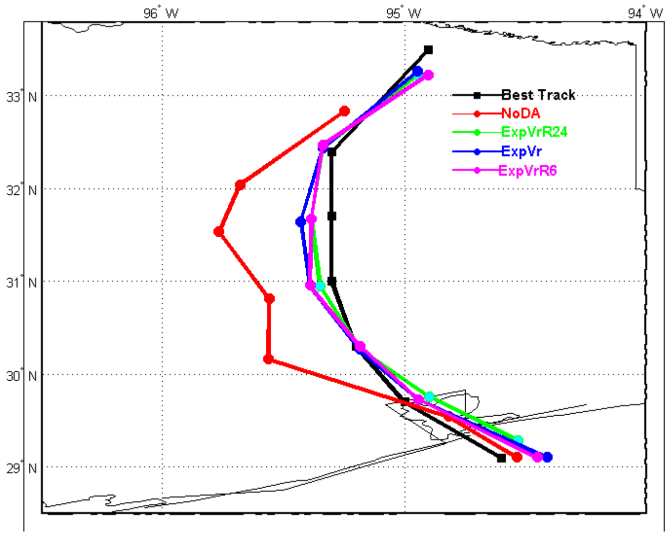

Fig. 4.28. The predicted track of NoDA (red), ExpVrR24 (green), ExpVr (blue) and ExpVrR6 (magenta), compared with the best track (black).

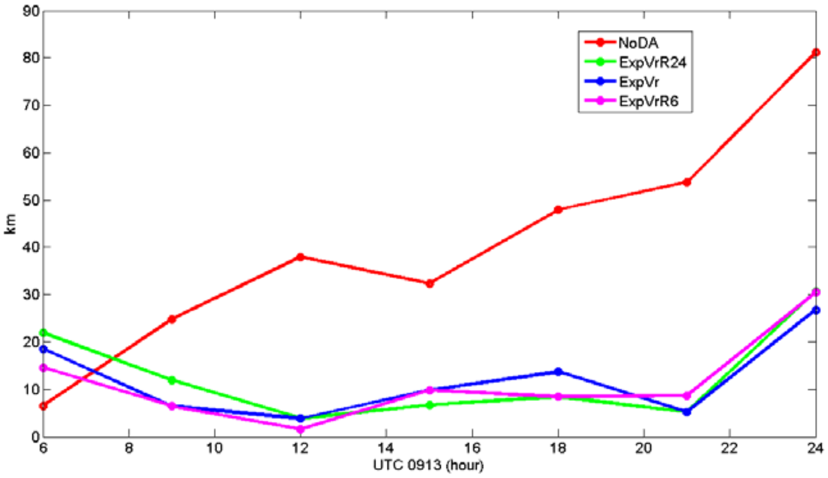

Fig. 4.29. The predicted track error of NoDA (red), ExpVrR24 (green), ExpVr (blue) and ExpVrR6 (magenta).

The tracks and track errors for Z-assimilating experiments are plotted in Fig. 4.30 and Fig. 4.31. Different from the ExpZ and NoDA cases, the analyzed positions at 0600 UTC in ExpZR24 and ExpZR6 are on the west side of the best track position (Fig. 4.30). ExpZR24 has a smaller track error than ExpZ and ExpZR6 (Fig. 4.31). During the first 3 hours of forecast, ExpZR24 and ExpZR6 take north-northwest paths, different from the northwest paths of the best track, NoDA and ExpZ. At 0900 UTC, it is apparent that the vortices in ExpZR24 and ExpZR6 change their moving directions, different from the best track, NoDA and ExpZ. After 0900 UTC, the tracks of the three experiments are similar. The mean track errors of ExpZR24, ExpZ and ExpZR6 are 16.6 km, 18.5 km and 17.6 km, respectively.

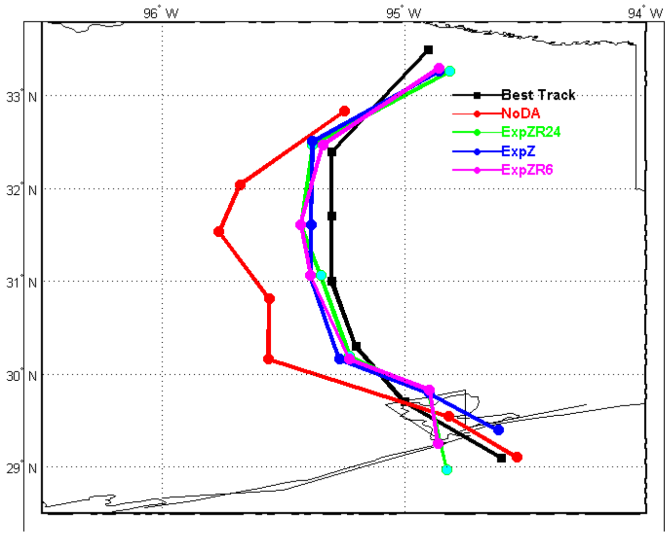

Fig. 4.30. The predicted track of NoDA (red), ExpZR24 (green), ExpZ (blue) and ExpZR6 (magenta), compared with the best track (black).

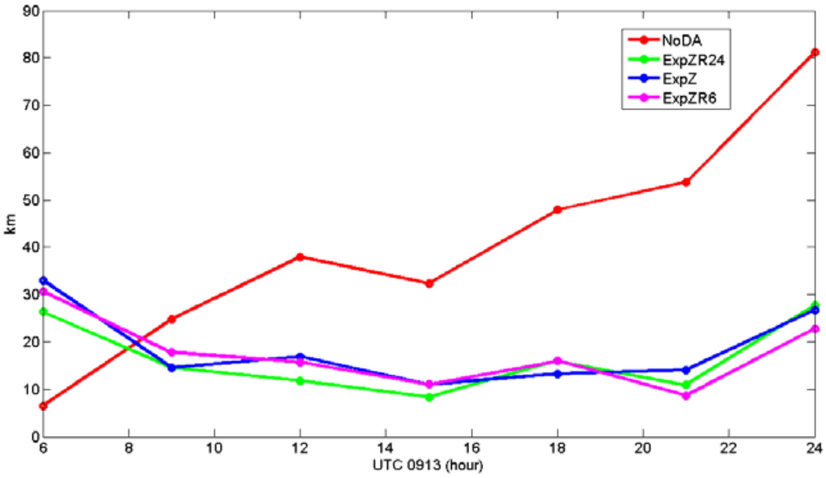

Fig. 4.31. The predicted track error of NoDA (red), ExpZR24 (green), ExpZ (blue) and ExpZR6 (magenta).

In summary, the change of the covariance localization radius has relatively small impacts on the intensity analysis and forecast among the experiments assimilating Z or Vr only. Except for some degradation in the final intensity analysis in ExpZR6, the localization radius has negligible effect on the intensity forecasts from up to 1500 UTC.

For the track forecast, among the experiments assimilating Vr alone, the impact of the localization radius is marginal. Among the experiments assimilating Z alone, although the 24 km radius shows the smallest mean track error, the predicted track with the 12 km radius is closer to the best track in terms of the general movement direction. In this specific hurricane case, a 12 km covariance localization radius appears to be optimal for the assimilation of radar data. This radius was used our control and most other experiments.

4.8 Sensitivity to assimilation interval

The data assimilation interval is changed to test the assimilation frequency's impact on intensity, track and precipitation forecasts in Exp30Min and Exp60Min (Fig. 4.32 to Fig. 4.34). With 10-minute intervals, ExpAll is also referred as Exp10Min in this section. The innovations of Vr and Z from KHGX during the assimilation cycles for three different intervals are plotted in Fig. 4.35. The difference of error growth between Exp10Min and Exp30Min are very close for both Vr and Z. There is a slightly larger error growth of Vr in Exp60Min from 60 minutes to 70 minutes.

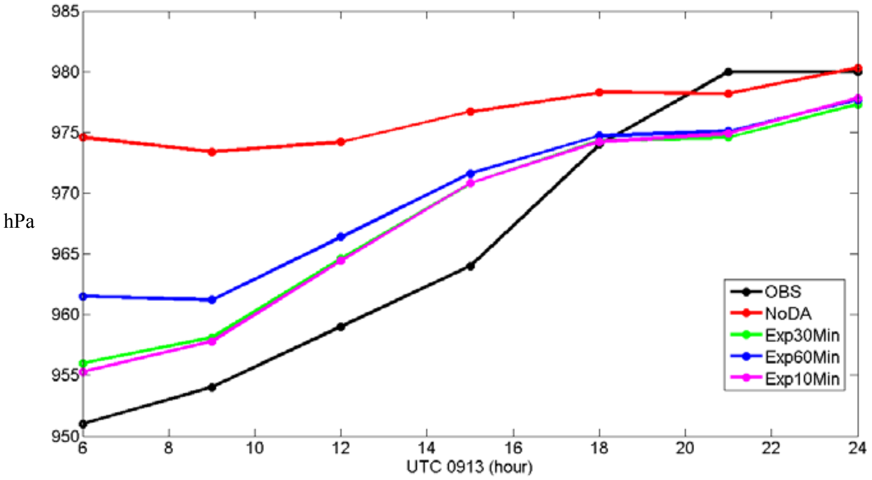

Fig. 4.32. The predicted minimum SLP of NoDA (red), Exp30Min (green), Exp60Min (blue) and ExpAll (also Exp10Min; magenta), compared with the best track (black).

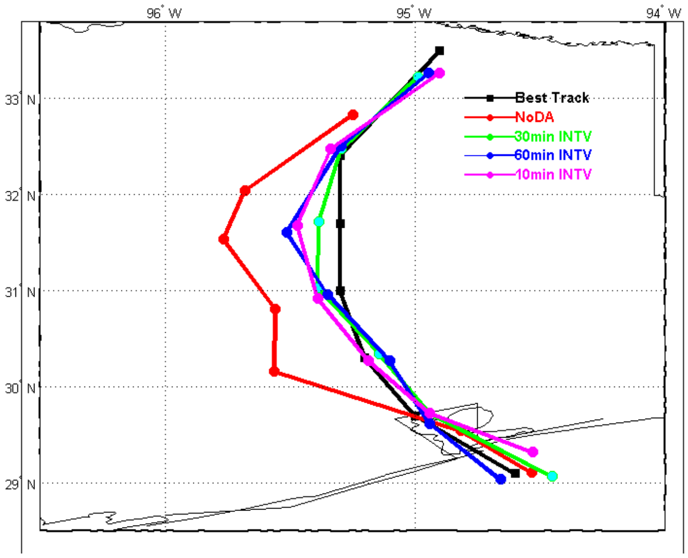

Fig. 4.33. The predicted track of NoDA (red), Exp30Min (green), Exp60Min (blue) and ExpAll (also Exp10Min; magenta), compared with the best track (black).

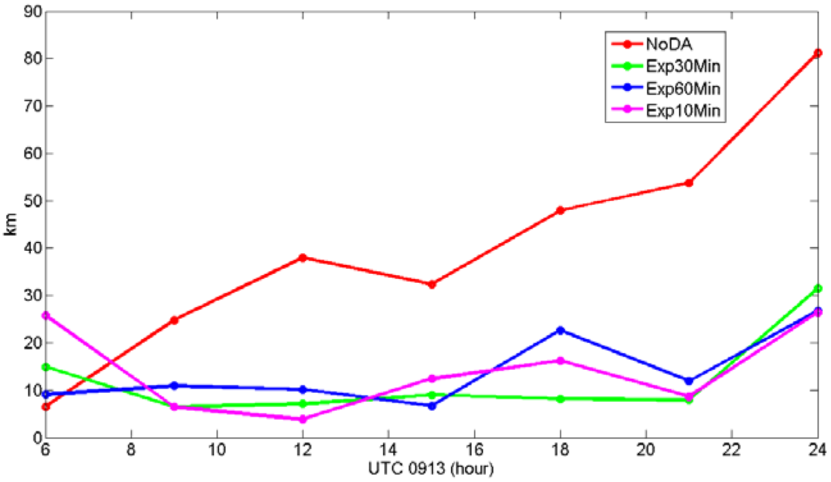

Fig. 4.34. The predicted track error of NoDA (red), Exp30Min (green), Exp60Min (blue) and ExpAll (also Exp10Min; magenta).

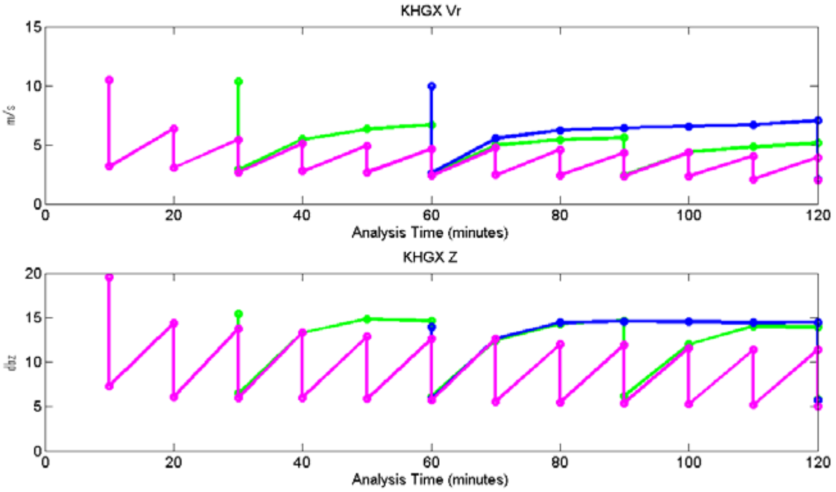

Fig. 4.35. The innovation for the radar KHGX (a) Vr and (b) Z during the analysis cycle for ExpAll (magenta), Exp30Min (green) and Exp60Min (blue).

The intensity forecasts indicate the 60-minute interval is not enough to predict a vortex as strong as 10-minure and 30-minute intervals in terms of minimum seal level pressure (Fig. 4.32). The latter two have similar intensity forecasts while minimum seal level pressures in 10-minute intervals is closer to the best track during the first 3 hour forecast. All the three intervals show similar track forecasts while Exp30Min has the smallest mean track error (Fig. 4.33 and Fig. 4.34). The ETSs in all of the three intervals share the similar accuracy of quantitative precipitation forecast (not shown).

4.9 Sensitivity to single radars

Instead of assimilating both of the coastal radars, the observations from the single radar KHGX and KLCH are also separately assimilated in ExpKHGX and ExpKLCH to investigate their individual impacts. Both Vr and Z are assimilated in these two experiments. It is interesting to notice that the assimilation of KHGX alone has the best intensity analysis with the strongest vortex in terms of minimum sea level pressure at 0600 UTC (Fig. 4.36) among all the three data assimilation experiments. The intensity difference with the best track for ExpKHGX , ExpKLCH and ExpAll is 2 hPa, 11 hPa and 4 hPa at 0600 UTC, respectively. During all the forecast times, ExpKHGX has the stronger predicted vortex than ExpKLCH and ExpAll. ExpKLCH also has evident improvement to NoDA but with a much weaker intensity forecast than the other two data assimilation experiments.

In the track forecast, ExpKLCH is closer to the best track from 1200 UTC to 1800 UTC with a more east path than ExpKHGX and ExpAll (Fig. 4.37). Assimilating KLCH alone always has the smallest or the second smallest track error during the 18 hours of forecast (Fig. 4.38). The mean track errors for ExpKHGX , ExpKLCH and ExpAll are 17 km, 9 km and 14 km, respectively.

It appears that KLCH has more contributions to the track forecast than KHGX while KHGX has a larger improvement in the intensity forecast than KLCH. Since the track and intensity forecasts of a hurricane is closely related to the axis-symmetric and axis-asymmetric circulation, it is possible that KLCH has more impact on the axis-asymmetric structure while KHGX exerts more influence on the axis-symmetric circulation. Generally, assimilating both radars has the better intensity forecast than assimilating KLCH alone and the better track forecast than assimilating KHGX alone, showing certain advantages of using multiple Doppler radars.

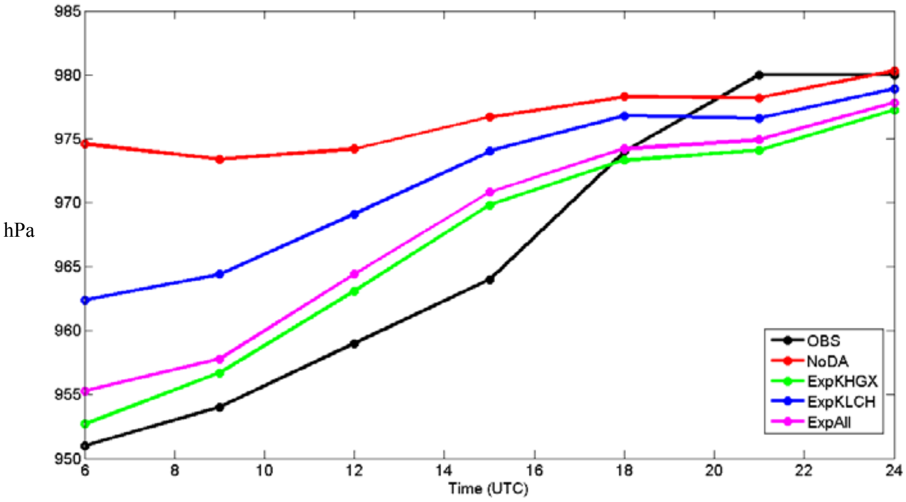

Fig. 4.36. The minimum seal level pressure of NoDA (red), ExpKHGX (green), ExpKLCH (blue) and ExpAll (magenta), compared with the best track (black).

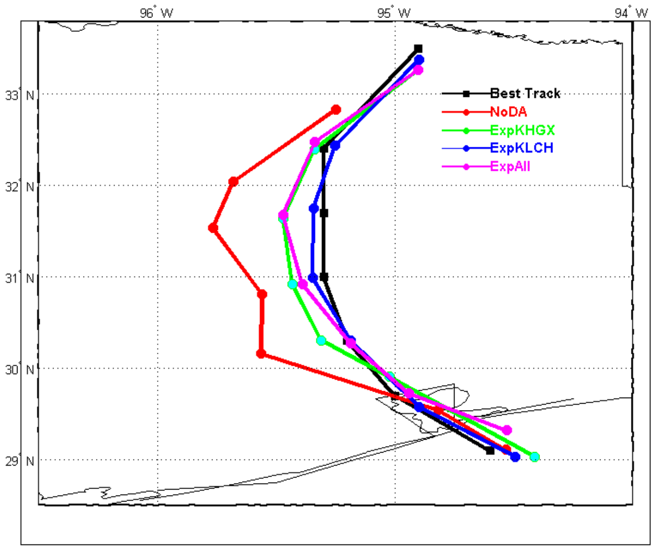

Fig. 4.37. The forecasted tracks of NoDA (red), ExpKHGX (green), ExpKLCH (blue) and ExpAll (magenta), compared with the best track (black).

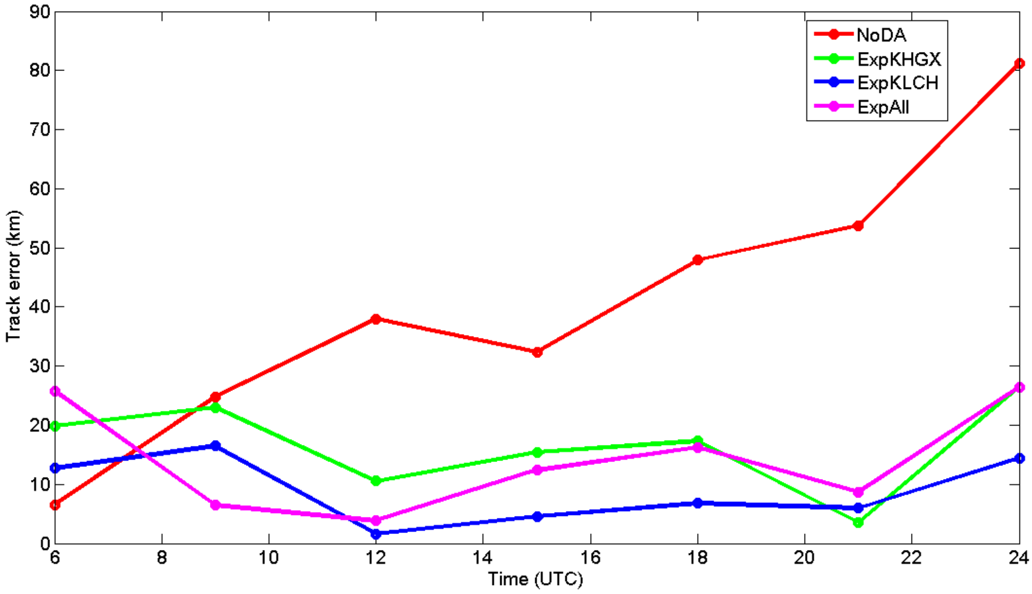

Fig. 4.38. The track error of NoDA (red), ExpKHGX (green), ExpKLCH (blue) and ExpAll (magenta).

4.10 Assimilation of MSLP

Besides radar observations, the assimilation of the minimum sea level pressure representing vortex intensity with EnKF has shown positive impacts on hurricane intensity and track analyses (Hamill et al's (2010) Fig. 15). Chen and Snyder (2007) assimilated the vortex intensity (vorticity magnitude in their study) with EnKF in a simple 2-dimensional barotropic model and also found the improvement in intensity and track analyses and forecasts. In our current experiments, there is about a 4-hPa error in the final intensity analysis in ExpAll, our best case, while ExpZ has a larger final analysis error of 16 hPa. The assimilation of best track MSLP is expected to directly impact Ike's intensity analysis and forecast and a larger improvement can be expected in ExpZ than in ExpVr and ExpAll experiments. The impact of assimilating MSLP on the track forecast will also be examined in this section.

4.10.1 Experiment setup

The MSLP from the best track are assimilated with radar observations together at 0500 and 0600 UTC September 13 in the control experiments. To test the sensitivity of MSLP assimilation intervals, the best track intensity are also interpolated with time with 10-minute interval between 0500 and 0600 UTC and assimilated with radar data. The radar observations are always assimilated with 10-minute intervals. The MSLP assimilation schemes for these two intervals of 10- and 60-minute are plotted in Fig. 4.39. Two additional experiments are conducted with MSLP assimilated alone with 10-minute and 60-minute intervals. The only difference between assimilating radar plus MSLP and MSLP alone is the removal of radar data in the latter two experiments.

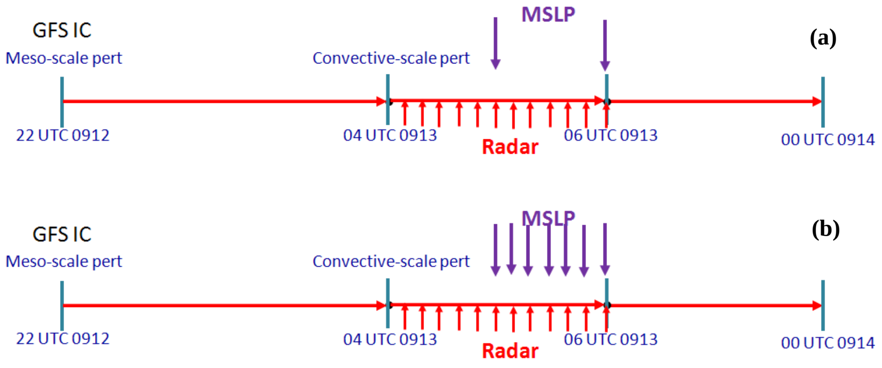

Fig. 4.39. The assimilations schemes for MSLP assimilation with (a) 60 min. interval and (b) 10 min. interval. The red upward arrows denote radar data assimilation. The purple downward arrows denote MSLP assimilation.

The positions of the MSLP are also extracted from the best track estimate. Chen and Snyder (2007) assimilated the vortex position into a 2D barotropic model and found the track forecast is improved. Different from their method, in this study MSLP from the best track is treated as a regular sea level pressure observation at the best track position.

The simple reduction equation (eq. (1) of Benjamin and Miller 1990) is applied as the observation operator H to obtain the sea level pressure from the pressure and the temperature of the model at the first vertical level above the surface:

$$P_{SL} = P_0\left(\frac{T_0 + \gamma z}{T_0}\right)^{g/R\gamma} \tag{4.9}$$

where P_{SL} is the sea level pressure, P_0 and T_0 are the pressure and temperature of the first level above the surface respectively at the model grid, z is the height of the first level, γ is the environmental temperature lapse rate, g is the gravity and R is the gas constant. This equation is based on the hydrostatic and hypsometric equations. A similar equation is derived by Wallace and Hobbs (1977). After the sea level pressure is obtained at the horizontal model grid, a bi-linear horizontal interpolation is used to project the sea level pressures to the best track position, which is not always on the model grid.

	Obs. Type	MSLP assimilation Interval from 0500 to 0600 UTC	MSLP obs. Error
ExpVrMSLP	Vr+MSLP	60 min.	1 hPa
ExpZMSLP	Z+MSLP	60 min.	1 hPa
ExpAllMSLP	Vr+Z+MSLP	60 min.	1 hPa
ExpVrMSLP2MB	Vr+MSLP	60 min.	2 hPa
ExpZMSLP2MB	Z+MSLP	60 min.	2 hPa
ExpAllMSLP2MB	Vr+Z+MSLP	60 min.	2 hPa
ExpVrMSLP10MIN	Vr+MSLP	10 min.	1 hPa
ExpZMSLP10MIN	Z+MSLP	10 min.	1 hPa
ExpAllMSLP10MIN	Vr+Z+MSLP	10 min.	1 hPa
ExpMSLP	MSLP	60 min.	1 hPa
ExpMSLP10MIN	MSLP	10 min.	1 hPa

Table 4.2. Experiments of assimilating MSLP observations.

The observation error of MSLP from human synthesized TcVital data set can range from 0.75 to 2 hPa (Tong 2010, personal communication). In this study, 1 hPa error is used in the control experiments and 2 hPa error is also tested. The MSLP are assimilated with Z or Vr or both Z and Vr. These experiments are summarized in Table 4.2.

The horizontal covariance localization radius of MSLP is 300 km, based on the size of the background vortex. Since there is only one MSLP observation assimilated at the analysis time and MSLP is expected to represent meso- or synoptic scale features, a large horizontal localization radius is preferred for this single MSLP observation to cover a large area of the vortex. The prior multiplicative covariance inflation of 5% is used to inflate the state variables covered by the radar or MSLP observations. The posterior additive covariance inflation same as section 4.3.4 is used only for the state variables covered by radar data.

MSLP is assimilated after radar observations during the analysis. The model state variables of wind components, potential temperature and pressure are updated by MSLP.

4.10.2 Impact of MSLP on hurricane analysis

The increment of wind and potential temperature at 1 km height from assimilating MSLP are plotted for ExpZMSLP in Fig 4.40. The increment in Fig. 4.40 is defined as the difference of the state variables before and after the MSLP analysis.

With the single MSLP observation analyzed, the increment field of wind shows the strong cyclonic circulation around the MSLP observation (Fig. 4.40a), indicating the enhancement of the vortex by the MSLP observation. The covariance between the pressure observation and the model wind fields allows the MSLP observation to update the wind field properly, which reflects the benefit of multivariate analysis. The pressure reduction from this MSLP observation is also evident in Fig. 4.40a. The pressure reduction is over 10 hPa at the vortex center and increases outward. The well-shaped circles of pressure increments around the MSLP observation is partly from the localization scheme which reduces the impact of the observation with distances.

The increment of potential temperature is plotted in Fig. 4.40b. While the increment of potential temperature is not as well-shaped as pressure, there is still noticeable positive increments at the center of the vortex with the maximum value of 5 K, suggesting the strengthening of the warm core structure by assimilating MSLP.

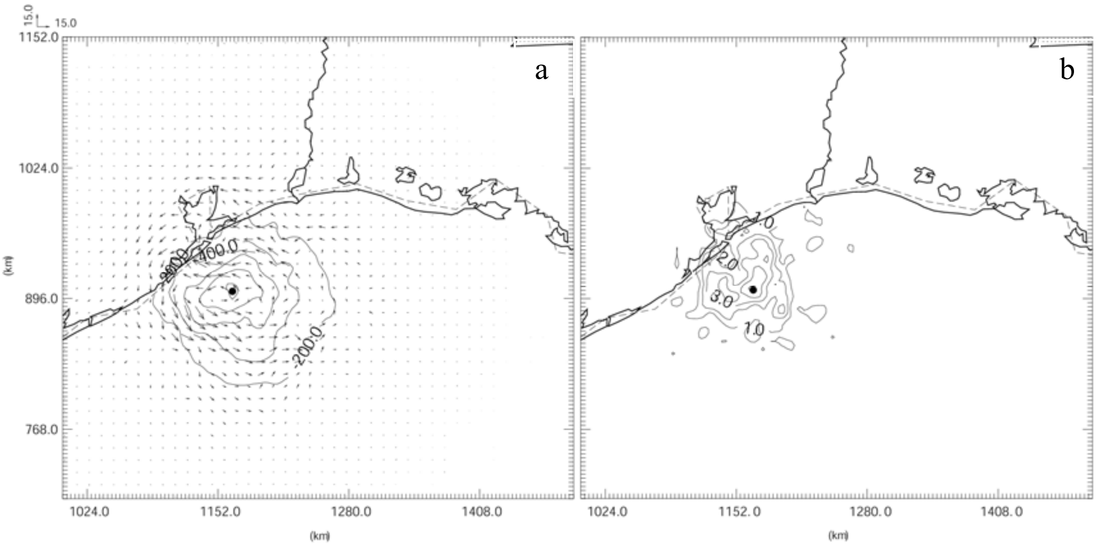

Fig. 4.40. Increment fields from assimilating MSLP at z=1km for (a) horizontal wind component and pressure (every 200 Pa), and (b) potential temperature (every 1 K) at 0500 UTC 0913 of ExpZMSL. The black dot denotes the position of the MSLP observation.

To investigate the impact of assimilating MSLP on intensity analysis, the minimum sea level pressures before and after each analysis cycle are plotted in Fig. 4.41 and Fig. 4.42. While the best track intensity is only available at 0500 of 952 hPa and 0600 UTC of 951 hPa, it may be reasonable to assume the intensity of the best track between these two times is close to 952 or 951 hPa.

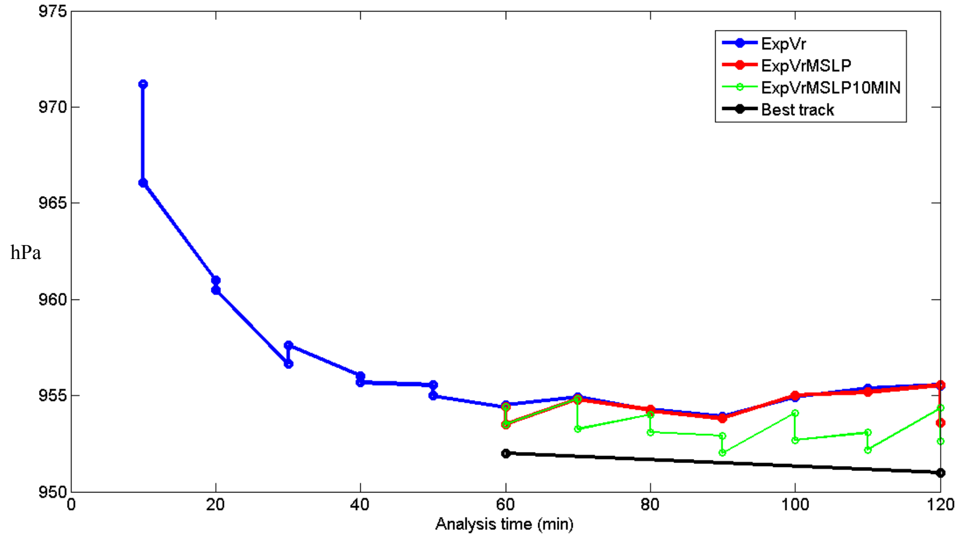

Fig. 4.41. The minimum sea level pressure during the assimilation cycles of ExpVr (blue), ExpVrMSLP (red), ExpVrMSLP10MIN (thin green) and the best track (black). Analysis time 10 min. corresponds to 0410 UTC September 13 and 120 min corresponds to 0600 UTC September 13.

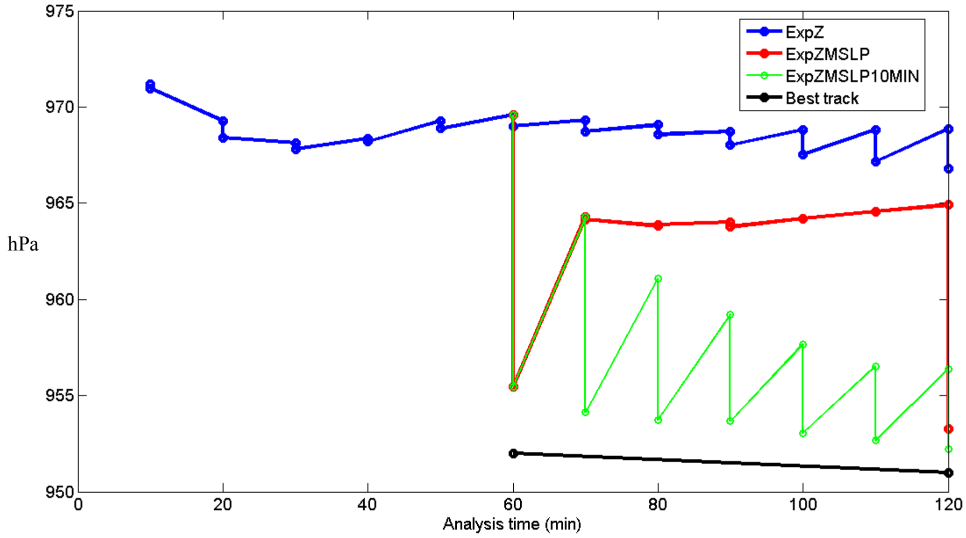

Fig. 4.42. The minimum sea level pressure during the assimilation cycles of ExpZ (blue), ExpZMSLP (red), ExpZMSLP10MIN (thin green) and the best track (black). Analysis time 10 min. corresponds to 0410 UTC September 13 and 120 min corresponds to 0600 UTC September 13.

When Vr is assimilated alone in ExpVr, the intensity does not change greatly from 0500 UTC to 0600 UTC (analysis time 60 to 120 min., Fig. 4.41) with the minimum sea level pressure always around 955 hPa. In ExpVrMSLP, there is about 1 hPa increase in intensity due to assimilating MSLP at 0510 UTC. Then this small improvement in analysis is lost quickly in the next forecast cycle, leading to the similar intensity level with ExpVr in the following assimilation cycles until the next MSLP observation is available at 0600 UTC. The improvement to ExpVr is about 2 hPa in the final intensity analysis at 0600 UTC.

When Z is assimilated alone in ExpZ, there is a slight intensification for the first 4 assimilation cycles starting from 0500 UTC (Fig. 4.42). Later, the impact from Z on the analysis is larger during the last 3 assimilation cycles with 1 to 2 hPa reductions of minimum sea level pressure, probably resulting from the better covariance structure between the microphysical fields and the pressure fields. In ExpZMSLP, the improvement due to assimilating the additional MSLP at 0500 UTC is as large as 14 hPa. While the vortex weakens quickly in the next 10 minutes to 964 hPa, the intensity is still 5 hPa stronger than ExpZ at 0510 UTC. From 0510 UTC to 0550 UTC, without MSLP assimilated in ExpZMSLP, there is no noticeable intensification in analysis-forecast cycles until the next MSLP observation is available at 0600 UTC. There is a considerable intensity increase of 12 hPa due to assimilating MSLP at 0600 UTC.

The impacts on analysis from more frequent MSLP assimilation cycles of 10 minutes intervals are also plotted in Fig. 4.41 and Fig. 4.42. They will be discussed later in section 4.10.5.

4.10.3 The impact on hurricane forecast

The predicted minimum sea level pressures are plotted in Fig. 4.43 and Fig. 4.44 to examine the impact of MSLP assimilation on Ike's forecast.

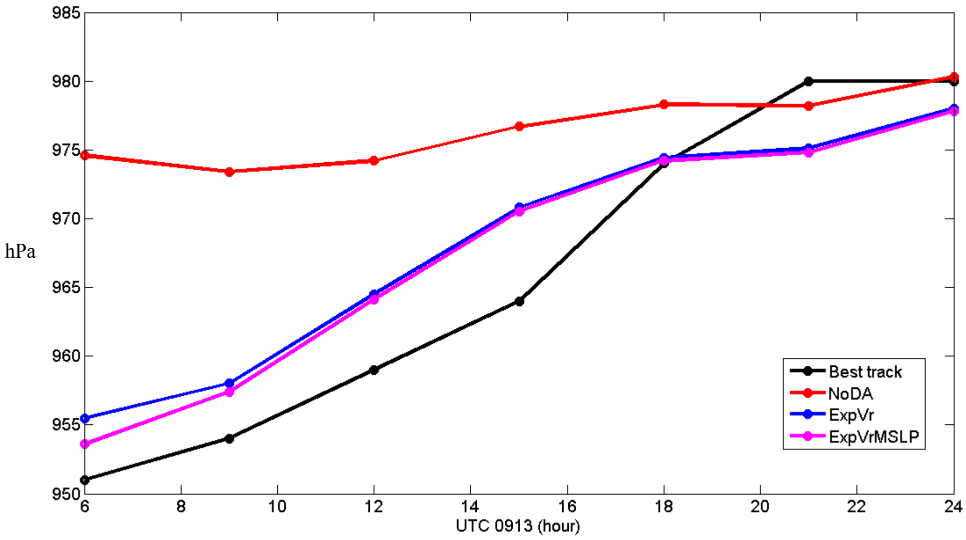

Fig. 4.43. The minimum sea level pressure of NoDA (red), ExpVr (blue), and ExpVrMSLP (magenta), compared with the best track (black).

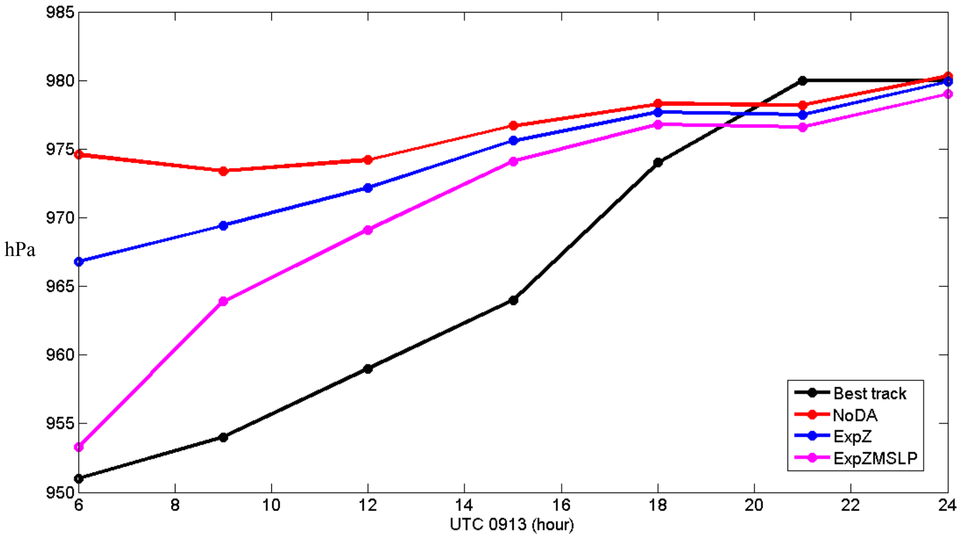

Fig. 4.44. The minimum sea level pressure of NoDA (red), ExpZ (blue), and ExpZMSLP (magenta), compared with the best track (black).

When Vr and MSLP are assimilated together, the intensity of the final analysis of ExpVrMSLP is 953.6 hPa at 0600 UTC, with a 2 hPa improvement to ExpVr (Fig. 4.43). The improvement, however, is reduced to less than 1 hPa after 3 hours of forecast at 0900 UTC. The difference becomes even smaller from 0900 UTC September 13 to 0000 UTC September 14.

When Z and MSLP are assimilated together in ExpZMSLP, the improvement due to assimilating the additional MSLP is significant with around a 14 hPa intensity increase over ExpZ (Fig. 4.44). The minimum sea level pressure of ExpZMSLP at 0600 UTC is 953.3 hPa, with only 2 hPa weaker than the best track. During the first 3 hours of forecast, the vortex weakens faster than the best track, suggesting some adjustment processes among the model state variables. From 0900 to 1500 UTC, the intensity weakening rate is similar to the best track estimate. The improvement in intensity to ExpZ is recognizable until 1800 UTC.

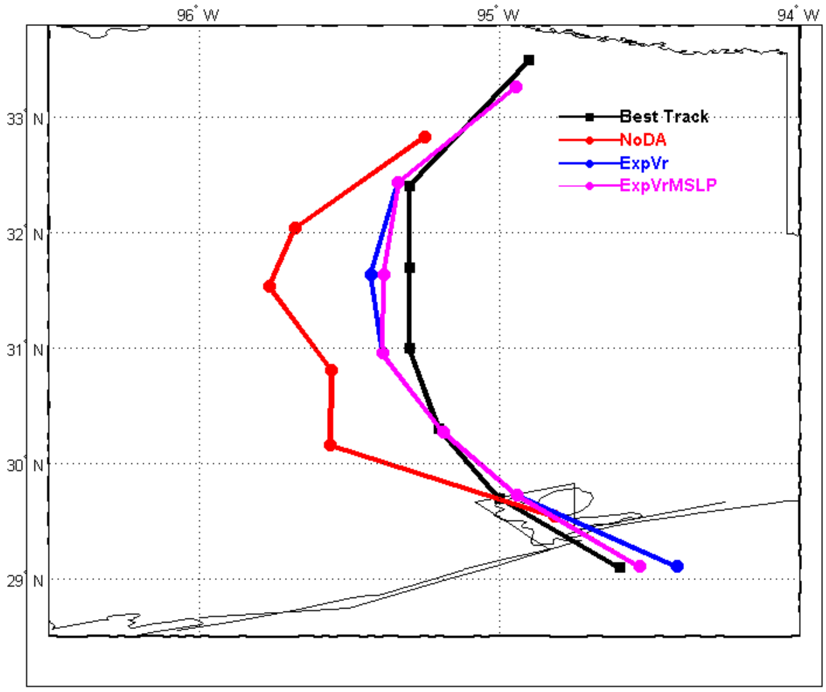

Fig. 4.45. The forecasted tracks of NoDA (red), ExpVr (blue) and ExpVrMSLP (magenta), compared with the best track (black).

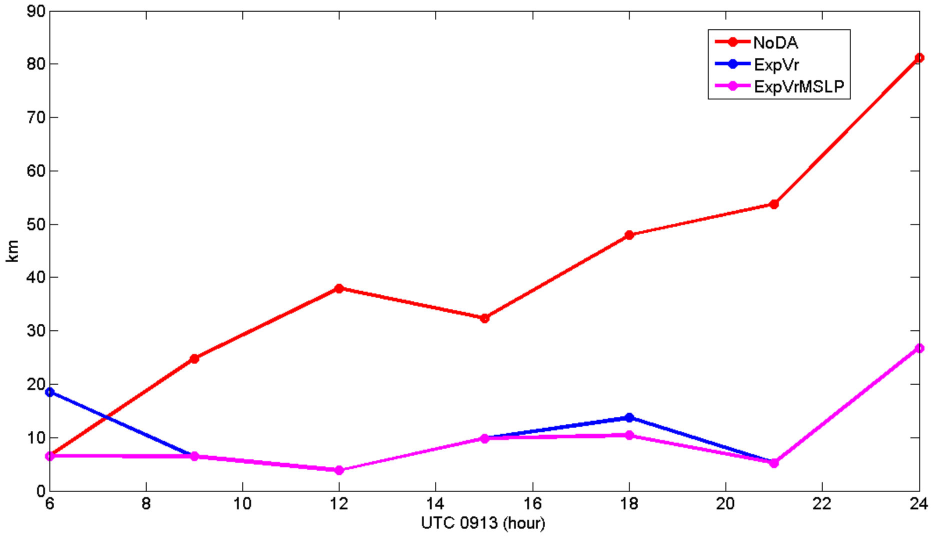

Fig. 4.46. The track error of NoDA (red), ExpVr (blue) and ExpVrMSLP (magenta).

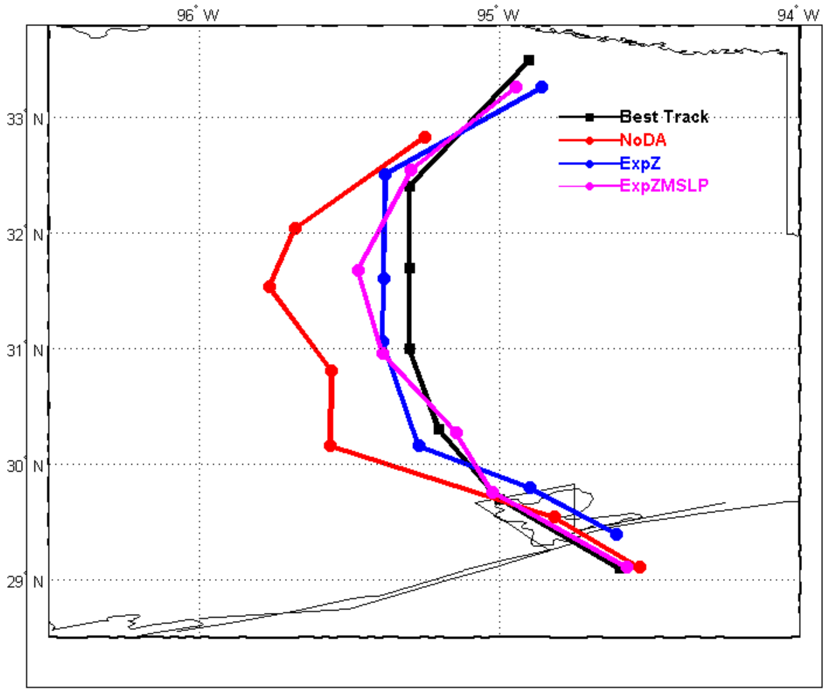

Fig. 4.47. The forecasted tracks of NoDA (red), ExpZ (blue) and ExpZMSLP (magenta), compared with the best track (black).

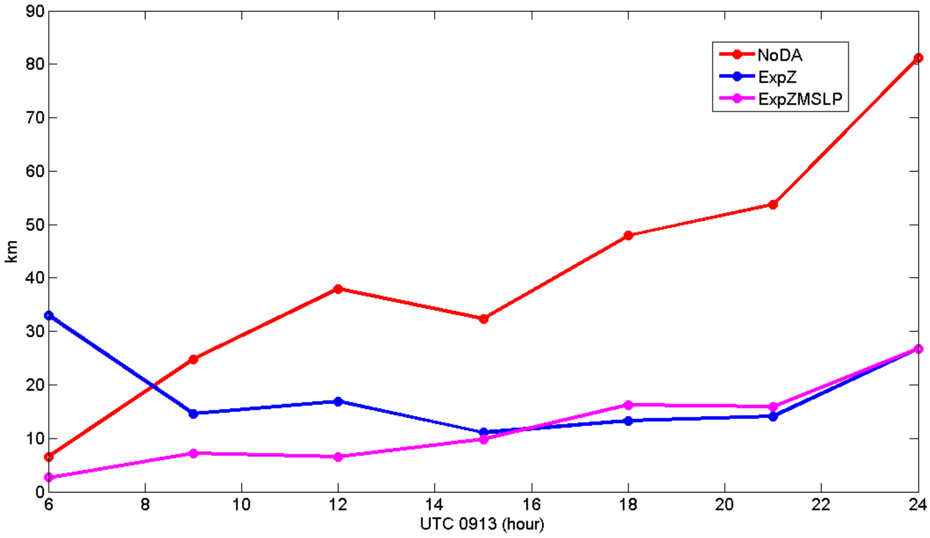

Fig. 4.48. The track error of NoDA (red), ExpZ (blue) and ExpZMSLP (magenta).

The forecasted tracks and track errors of ExpVr and ExpVrMSLP are plotted in Fig. 4.45 and Fig. 4.46. With the additional MSLP assimilated, the vortex center is closer to the best track at 0600 UTC with the track error of 6.6 km and the relative improvement of 65% in terms of track error reduction (Fig. 4.45 and Fig. 4.46). The improvement in the final analysis does not last long during the following forecast. The vortex movement and track errors of ExpVrMSLP are very similar to ExpVr during most of the forecast times.

In ExpZMSLP at 0600 UTC, the improvement in the track forecast due to assimilating additional MSLP is even larger than assimilating MSLP with Vr, with the track error of only 2.6 km and the relative improvement of 92% over ExpZ (Fig. 4.48). The movement of the hurricane in ExpZMSLP follows the best track closely from 0600 to 1200 UTC (Fig. 4.47). The improvements in track forecast are noticeable until 1500 UTC with the relative improvements in terms of track error reduction always over 50% during the first 6-hour forecast. After 1500 UTC, the track errors of ExpZMSLP are similar to ExpZ.

When MSLP is assimilated with both Vr and Z altogether in ExpAllMSLP, the impact due to assimilating additional MSLP on the intensity and track forecasts is very similar to the impact of assimilating MSLP in addition to Vr and will not be discussed in detail.

4.10.4 Sensitivity to MSLP observation errors

In this section, the observation error of MSLP is changed from 1 hPa in control experiments to 2 hPa to examine the impact on track and intensity forecasts. The intensity and track errors during the 18 hours of forecast are plotted in Fig. 4.49 to Fig. 4.52. Intensity error is defined as the difference of the minimum sea level pressures between the model forecasts and the best track estimates. It should be noticed that it is not a strictly defined "real" error since it is not the difference between the model simulation and the truth. This definition is chosen just for convenience.

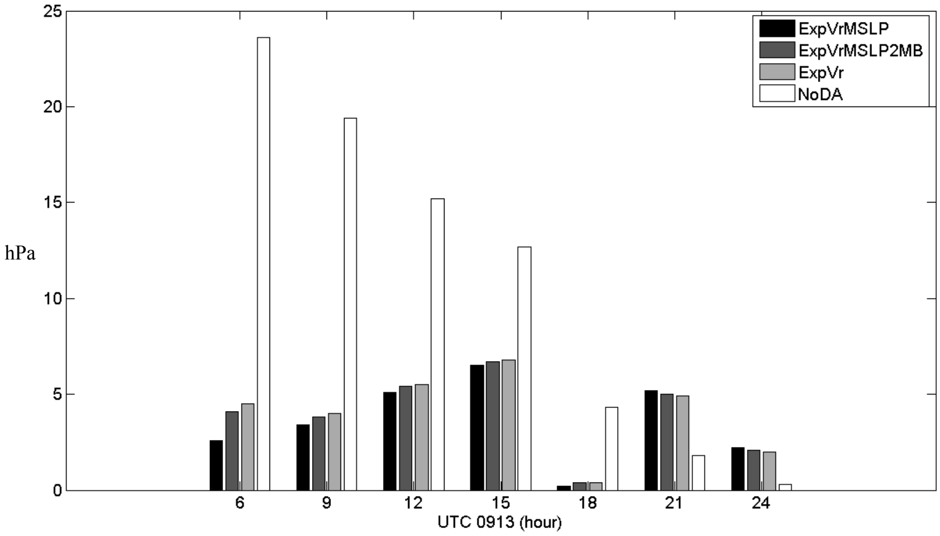

Fig. 4.49. The intensity errors of ExpVrMSLP, ExpVrMSLP2MB, ExpVr and NoDA experiments from 0600 UTC September 13 to 0000 UTC September 14. See text for the definition of intensity error.

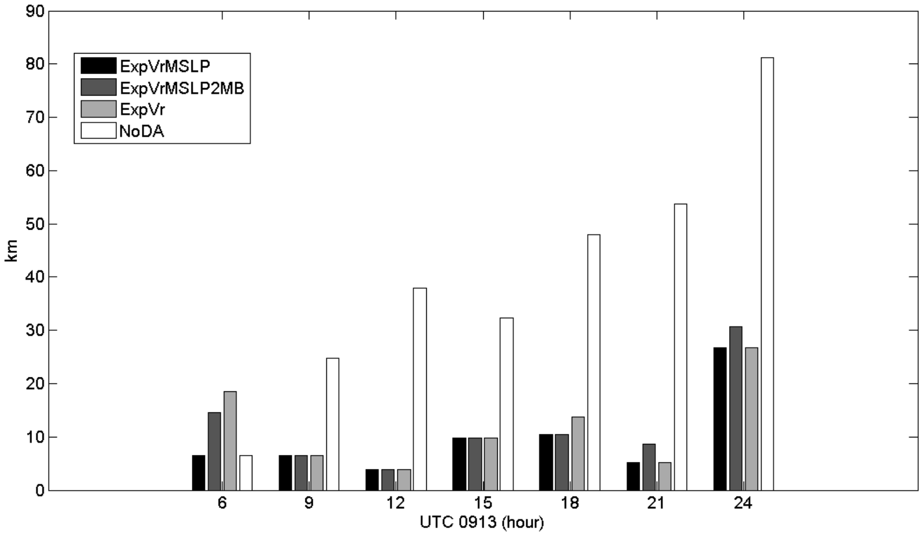

Fig. 4.50. The track errors of ExpVrMSLP, ExpVrMSLP2MB, ExpVr and NoDA experiments from 0600 UTC September 13 to 0000 UTC September 14.

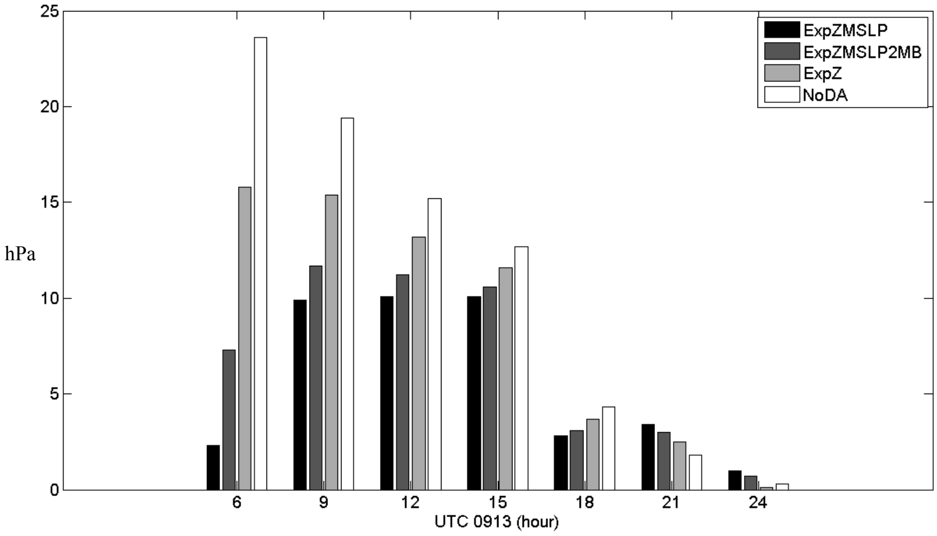

Fig. 4.51. The intensity errors of ExpZMSLP, ExpZMSLP2MB, ExpZ and NoDA experiments from 0600 UTC September 13 to 0000 UTC September 14.

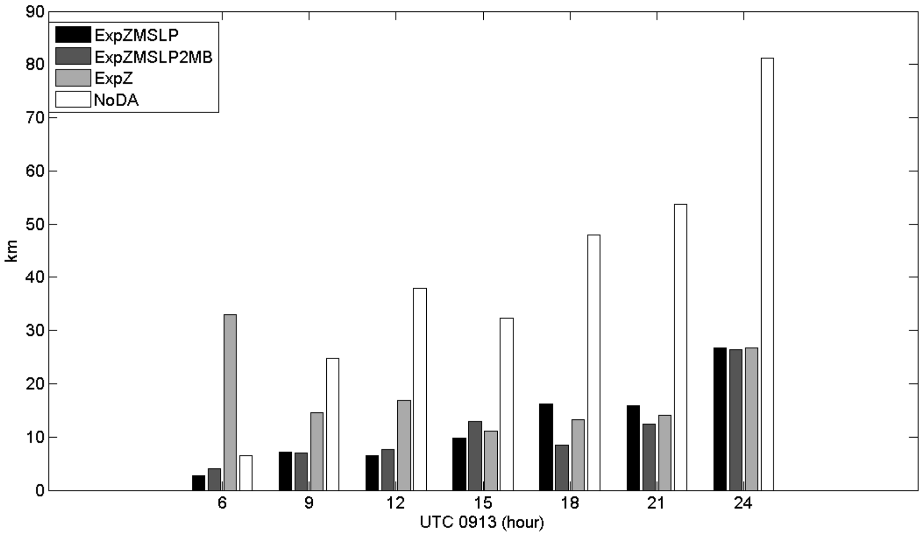

Fig. 4.52. The track errors of ExpZMSLP, ExpZMSLP2MB, ExpZ and NoDA experiments from 0600 UTC September 13 to 0000 UTC September 14.

When MSLP is assimilated with Vr, ExpVrMSLP has a better intensity analysis at 0600UTC than ExpVrMSLP2MB with a 1.5 hPa improvement (Fig. 4.49). The 1.5 hPa difference between ExpVrMSLP and ExpVrMSLP2MB on the final analysis of minimum sea level pressure is close to the 1 hPa difference between the MSLP observation errors in these two experiments. This difference decreases gradually in the following forecast. The time evolution of the difference in track errors between ExpVrMSLP and ExpVrMSLP2MB is similar to those in the intensity forecast (Fig. 4.50). The largest difference of track errors between these two experiments during 18 hours of forecast is 8 km at 0600 UTC with the smaller track error in ExpVrMSLP. The track errors of these two experiments are close to each other from 0900 UTC to 1800 UTC.

When MSLP is assimilated with Z, the difference of intensity between ExpZMSLP and ExpZMSLP2MB in the final analysis is as large as 5 hPa (Fig. 4.51) with a smaller intensity error in ExpZMSLP. The difference between ExpZMSLP and ExpZMSLP2MB decreases during the forecast, with 2 hPa at 0900 UTC and less than 1 hPa after 1500 UTC. In the track forecast, the difference between ExpZMSLP and ExpZMSLP2MB is small during most of the forecast hours (Fig. 4.52), except for 1800 UTC when the track error of ExpZMSLP is 8 km larger than ExpZMSLP2MB.

In this specific hurricane case, a smaller observation error of MSLP leads to better intensity and track forecast generally. In practice, a realistic observations error needs to be carefully chosen and should reflect the observation properties, such as the representative and instrument errors.

4.10.5 Sensitivity to MSLP assimilation intervals

In this section, the MSLPs is linearly interpolated with time from 0500 to 0600 UTC with 10-minute interval and assimilated with radar observations together. In operational applications, there are no such frequent MSLP estimates with 10-minute interval updates. In this study, the MSLP estimates at 0500 (952 hPa) and 0600 UTC (951 hPa) are so close that it is reasonable to assume the time interpolation does not introduce large errors. If there are abrupt changes with time in MSLP datasets, it may not be appropriate to use linear time interpolations.

The predicted intensities for ExpVrMSLP10MIN and ExpZMSLP10MIN are plotted in Fig. 4.53 and Fig. 4.54.

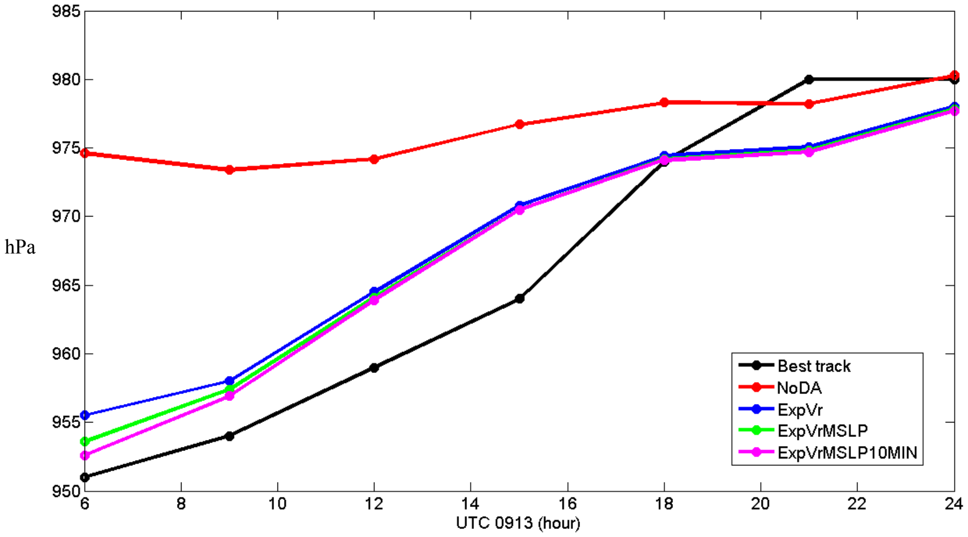

Fig. 4.53. The minimum sea level pressure of NoDA (red), ExpVr (blue), ExpVrMSLP (green), and ExpVrMSLP10MIN (magenta) compared with the best track (black).

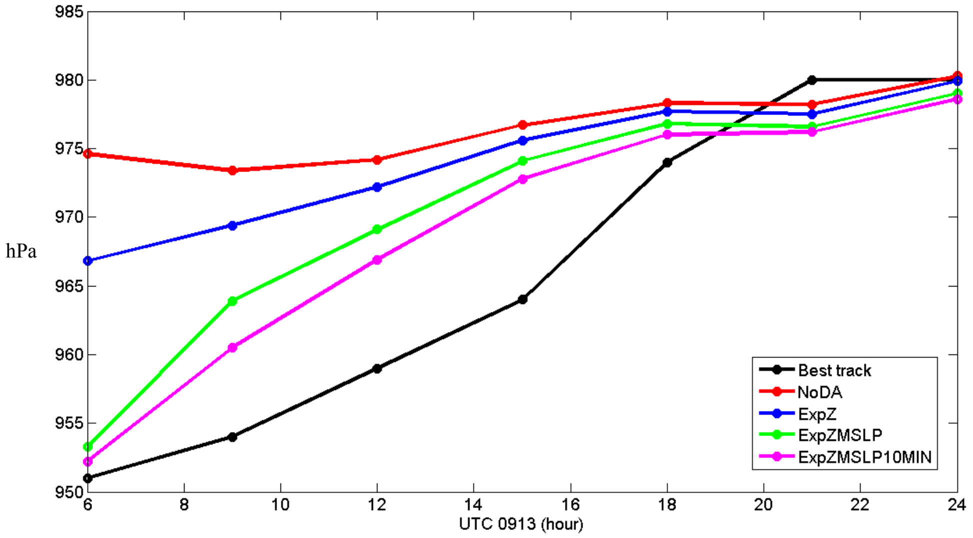

Fig. 4.54. The minimum sea level pressure of NoDA (red), ExpZ (blue), ExpZMSLP (green), and ExpZMSLP10MIN (magenta) compared with the best track (black).

With 10-minute intervals, the improvement from the frequent assimilation of MSLP to ExpVrMSLP during the assimilation cycles is distinguishable (Fig. 4.41 and Fig. 4.42). The forecasted minimum sea level pressures of ExpVrMSLP10MIN from 0530 to 0600 UTC are always smaller than ExpVrMSLP (Fig. 4.41). At the final analysis time, the intensity in ExpVrMSLP10MIN is 1 hPa stronger than ExpVrMSLP (Fig. 4.53). The difference decreases to less than 1 hPa in the first 3 hours of forecast at 0900 UTC and then almost disappears in the last 12 hours of forecast.

When MSLP is assimilated with Z, the more frequent assimilation of MSLP leads to larger additional impact than assimilating MSLP with Vr. The forecasted minimum sea level pressure in ExpZMSLP10MIN is evidently smaller than ExpZMSLP from 0520 UTC to the end of all assimilation cycles (Fig. 4.54). The intensity of the forecasted vortex increases gradually in each assimilation cycle with 10-minute intervals. In the first 2-3 cycles of forecast, the forecasted minimum sea level pressure increases very quickly after each analysis but increases slower during the last 2-3 cycles, suggesting a slower error growth in the later cycles. At the end of all the assimilation cycles, the intensity in ExpZMSLP10MIN is 1 hPa stronger than ExpZMSLP (Fig. 4.54). Different from the experiments assimilating Vr and MSLP together, the difference between ExpZMSLP and ExpZMSLP10MIN at the initial forecast time amplifies in the first 3 hours of forecast, reaching to 3.4 hPa at 0900 UTC. This noticeable improvement of using 10-minute intervals to 60-minute intervals lasts until 1800 UTC with a 0.8 hPa difference between ExpZMSLP and ExpZMSLP10MIN at 1800 UTC.

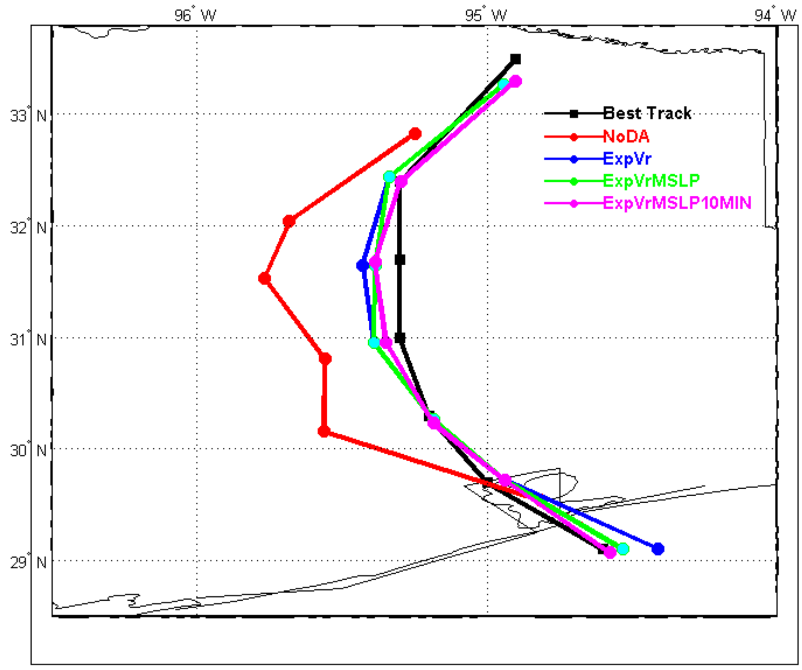

Fig. 4.55. The forecasted tracks of NoDA (red), ExpVr (blue), ExpVrMSLP (green) and ExpVrMSLP10MIN (magenta), compared with the best track (black).

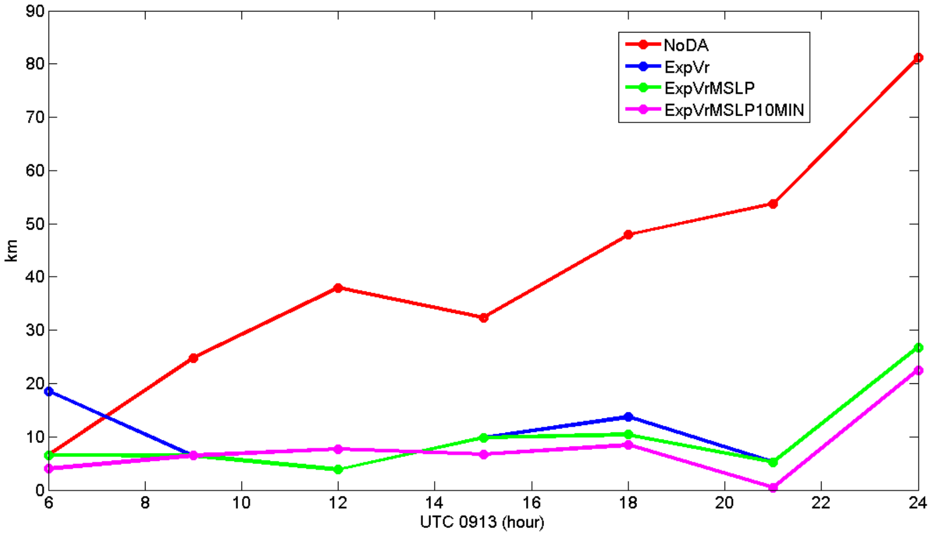

Fig. 4.56. The track error of NoDA (red), ExpVr (blue), ExpVrMSLP (green) and ExpVrMSLP10MIN (magenta).

The forecasted track and track error of ExpVrMSLP10MIN are plotted in Fig. 4.55 and Fig. 4.56, respectively. During most of the forecast hours, the track errors in ExpVrMSLP10MIN are smaller than ExpVrMSLP. The mean track error of ExpVrMSLP10MIN averaged in 18 hours of forecast is 8.0 km with a relative 19% improvement to ExpVrMSLP.

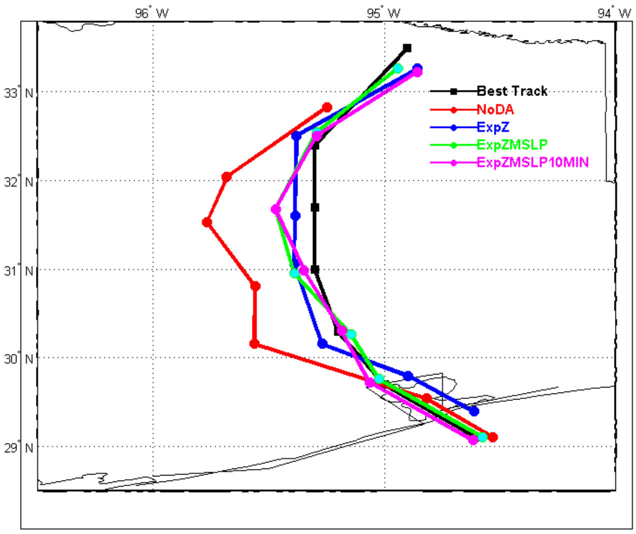

Fig. 4.57 The forecasted tracks of NoDA (red), ExpZ (blue), ExpZMSLP (green) and ExpZMSLP10MIN (magenta), compared with the best track (black).

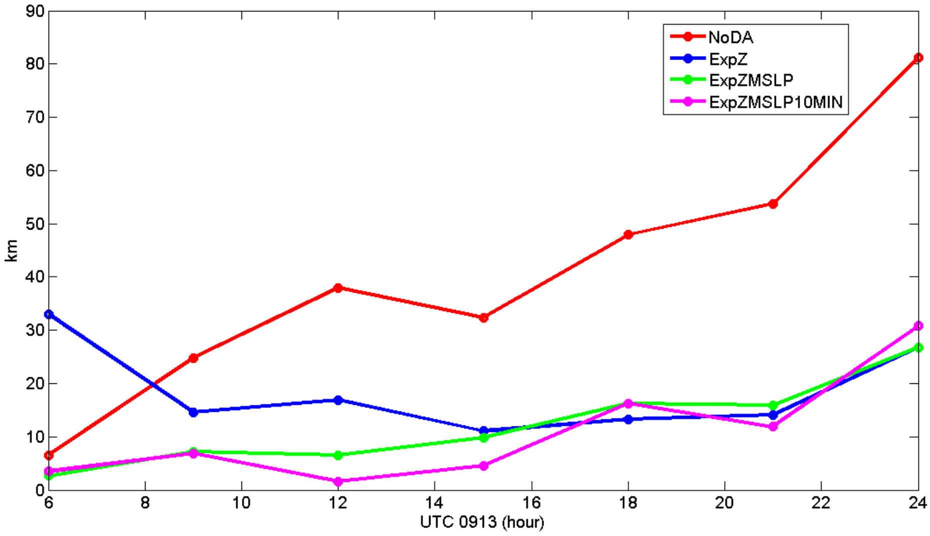

Fig. 4.58. The track error forecast of NoDA (red), ExpZ (blue), ExpZMSLP (green) and ExpZMSLP10MIN (magenta).

When MSLP and Z are assimilated together with 10-minute intervals, the track forecast is improved at 1200, 1400 and 2100 UTC due to using more frequent assimilation cycles (Fig. 4.57 and Fig. 4.58). The mean track error of ExpZMSLP10MIN is 10.8 km, with a relative improvement of 11% to ExpZMSLP.

4.10.6 Assimilation of MSLP alone

In the previous sections, MSLP is always assimilated with radar observations. It is of our interest to examine the individual impact of assimilating MSLP and radar observations separately. To investigate the impact of assimilating MSLP alone, ExpMSLP and ExpMSLP10MIN are conducted without radar data assimilation. In ExpMSLP, MSLP is assimilated at 0500 and 0600 UTC with the assimilation interval of 60 minutes as in ExpVrMSLP or ExpZMSLP. In ExpMSLP10MIN, MSLP is assimilated with the assimilation interval of 10 minutes between 0500 and 0600 UTC as in

ExpVrMSLP10MIN or ExpZMSLP10MIN. Covariance inflation of 5% is applied only for the state variables covered by the single MSLP observation and no additive error is used. The experiment setup for ExpMSLP is similar to ExpVrMSLP or ExpZMSLP except for the difference mentioned above. The experiment setup for ExpMSLP10MIN is similar to ExpVrMSLP10MIN or ExpZMSLP10MIN.

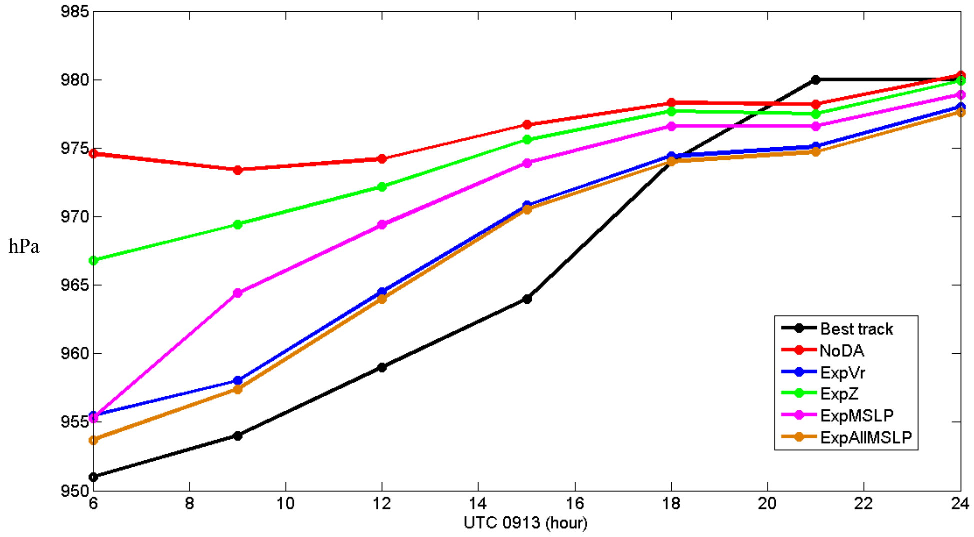

Fig. 4.59. The minimum sea level pressure of NoDA (red), ExpVr (blue), ExpZ (green), ExpMSLP (magenta) and ExpAllMSLP (brown) compared with the best track (black).

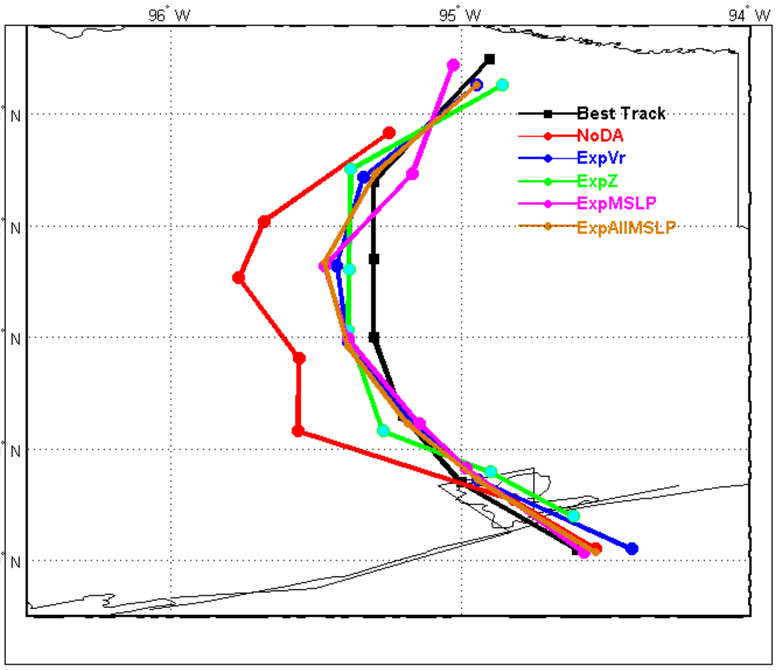

Fig. 4.60. The forecasted tracks of NoDA (red), ExpVr (blue), ExpZ (green), ExpMSLP (magenta) and ExpAllMSLP (brown) compared with the best track (black).

The forecasted minimum sea level pressures of ExpMSLP are plotted in Fig. 4.59 together with ExpVr, ExpZ and ExpAllMSLP. At 0600 UTC, ExpMSLP has an intensity of 955.3 hPa which is very close to the intensity of ExpVr and much stronger than ExpZ. When the forecast starts, the vortex in ExpMSLP weakens faster than ExpVr in the first 3 hours. The intensity at 0900 UTC is 964.4 hPa, 6.4 hPa weaker than ExpVr. Since there are much fewer observation numbers (only one MSLP observation assimilated) and less frequent assimilations cycles of MSLP assimilation than Vr data, it is not surprising to observe this fast weakening behavior in the first 3 hours of forecast. While the MSLP helps to improve the mesoscale intensity analysis, the storm or convective scale features in the hurricane system may not be improved as much as in ExpVr with radar data assimilation. When the forecast starts, the numerical model spends some time for the adjustment among various scales and a faster error growth than ExpVr is inevitable in the first several hours. From 0900 UTC, the vortex of ExpMSLP weakens more slowly than ExpVr, with the difference between ExpMSLP and ExpVr after 0900 UTC always smaller than that at 0900 UTC. It is also noticed that the intensity forecast of ExpMSLP is always better than ExpZ until 1800 UTC, indicating a larger impact from MSLP than Z on the intensity forecast.

When Vr, Z and MSLP are assimilated altogether in ExpAllMSLP, there is a 2 hPa improvement in intensity to ExpVr at the start of the forecast. This difference, however, becomes insignificant at and after 0900 UTC.

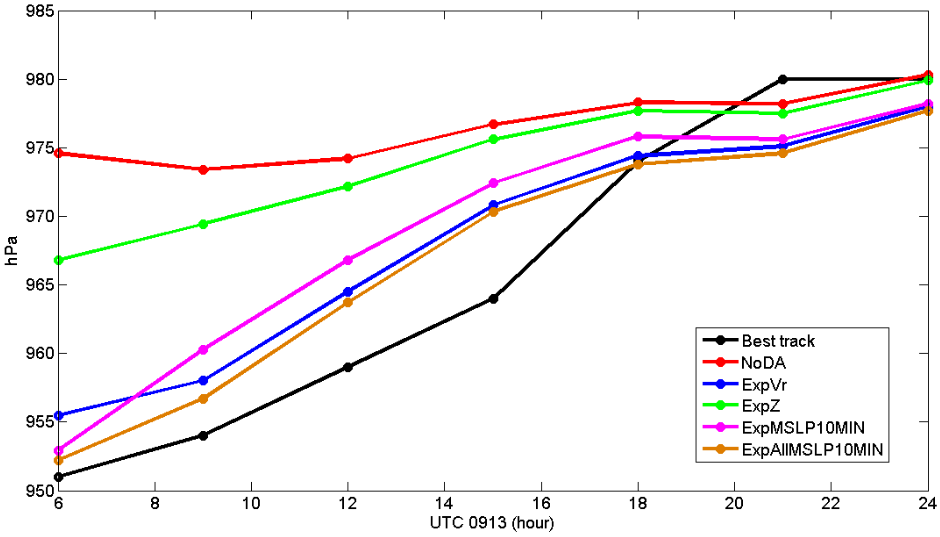

Fig. 4.61. The minimum sea level pressure of NoDA (red), ExpVr (blue), ExpZ (green), ExpMSLP10MIN (magenta) and ExpAllMSLP10MIN (brown) compared with the best track (black).

When MSLP is assimilated alone with 10-minute intervals in ExpMSLP10MIN, the intensity at 0600 UTC is closer to the best track than ExpVr and ExpZ (Fig. 4.61). The minimum sea level pressure of ExpMSLP10MIN is 2.6 hPa stronger than ExpVr and 13.9 hPa stronger than ExpZ at 0600 UTC. With a faster weakening rate than ExpVr in the first 3-hour forecast, the vortex in ExpMSLP10MIN is weaker than ExpVr from 0900

UTC but still much stronger than ExpZ during the 18 hours of forecast. With Vr, Z and MSLP at 10-minute intervals are assimilated together in ExpAllMSLP10MIN, the better intensity forecast is obtained than assimilating each observation individually.

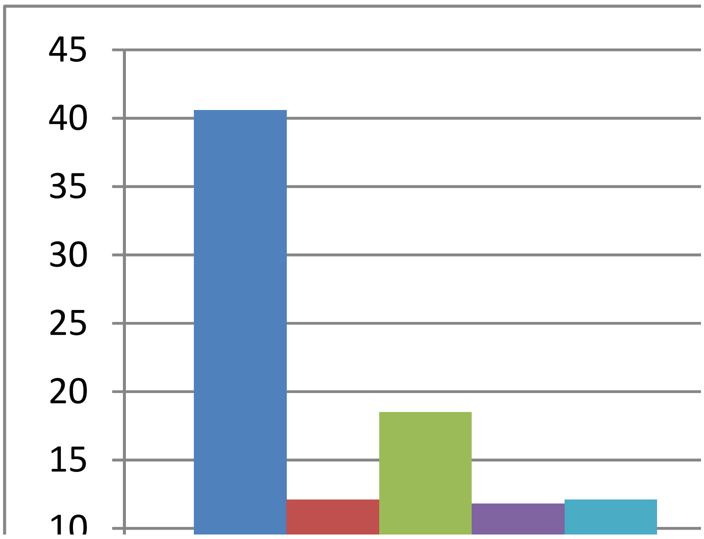

Fig. 4.62. Mean track error in 18-hour forecast of NoDA, ExpVr, ExpZ, ExpMSLP and ExpAllMSLP.

The predicted track of ExpMSLP is plotted in Fig. 4.60. The position of the vortex center in ExpMSLP is closer to the best track than ExpVr and ExpZ at 0600 UTC. From 0900 to 1800 UTC, the track of ExpMSLP is overlapped with ExpVr. At 2100 UTC, the vortex center position of ExpMSLP is on the east side of the best track, opposite to ExpVr and ExpZ, which are both on the west side of the best track. The track error of ExpMSLP at the end of the forecast is smaller than ExpVr and ExpZ (not shown). The mean track error of ExpMSLP during the 18 hours of forecast is similar to ExpVr and ExpAllMSLP, all showing relative improvements of 33% over ExpZ.

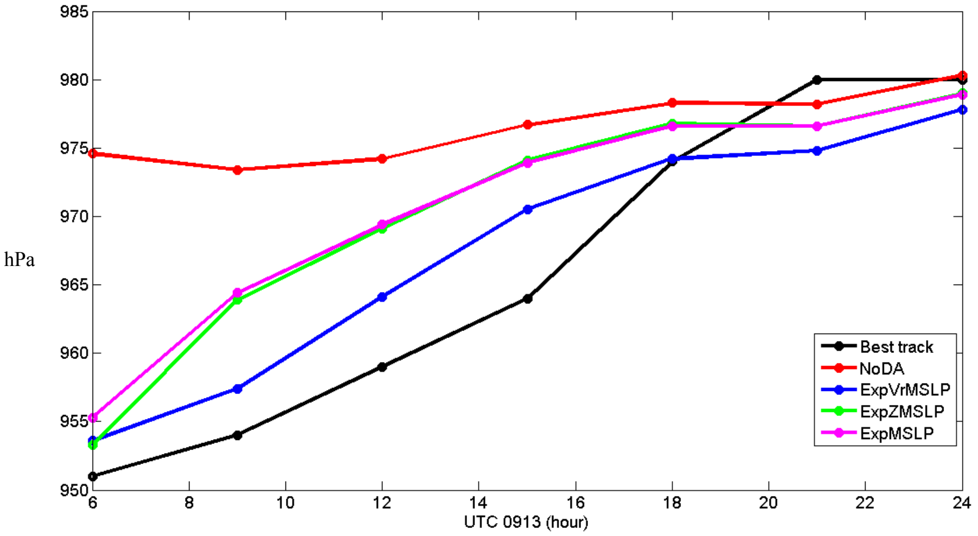

Fig. 4.63. The minimum sea level pressure of NoDA (red), ExpVrMSLP (blue), ExpZMSLP (green), and ExpMSLP (magenta), compared with the best track (black).

The MSLPs of ExpMSLP are also compared with ExpVrMSLP and ExpZMSLP (Fig. 4.62). It is interesting to notice that there are further improvements from both ExpVrMSLP and ExpZMSLP to ExpMSLP, therefore combining the radar data with MSLP information gives the best results.

4.11 Uncertainty growth from initial conditions and microphysical schemes

EnKF can help predictability studies by providing initial conditions that properly sample initial condition uncertainties. Sippel and Zhang (2010) used EnKF and short-range ensemble forecasts to examine factors affecting the predictability of Hurricane Humberto (2007). With the "hot-start" EnKF analysis, the artificial over-reaction of initial convection to the environment convective instability during the spin-up stage can be removed (Sippel and Zhang 2010). Given limited time available, we will focus on a different aspect of predictability. The development of hurricane track and intensity uncertainty results partly from the uncertainty in the initial conditions and partly from the uncertainty in the prediction model. The forecast uncertainty growth owing to initial condition uncertainties will be compared with the uncertainty growth due to model physics uncertainties, especially in microphysics. We will include the effect of physics uncertainty by using different microphysics parameterization schemes in the ensemble forecasts and examine its relative impact on uncertainty growth.

4.11.1 Motivation and past work

Different from the deterministic forecast predicting the single atmospheric state, ensemble forecasts aim to predict the weather system's probability as completely as possible. Due to the fact that errors in initial conditions and forecast models are unavoidable, ensemble forecasts can provide useful uncertainty informations on the future PDF of weather systems and help forecasters to better understand the predictability issues.

There are several studies on comparing uncertainty growth from initial conditions and model errors. Stensrud et al. (2000) investigated the uncertainty growth from the initial conditions and the microphysical schemes for two mesoscale convective systems (MCS). They found a faster uncertainty growth from microphysical schemes than from initial conditions for MCS cases during the first 12-hour forecasts. Kong et al. (2007) compared the spread development from the initial conditions and microphysical schemes for convective thunderstorm predictions. They found different contributions from the two uncertainty sources to the spreads of different model fields.

While it is known that the forecast of the hurricane system is both sensitive to the microphysical process (Wang 2009; Fovell et al. 2009; Fovell et al. 2010) and initial conditions, there has been no research to investigate the contributions of these two factors on the uncertainty growth of cloud-resolving hurricane forecasts. The intent of this study is to examine the role of microphysical parameterization schemes on the uncertainty growth of Hurricane Ike's intensity and track forecasts. The individual contributions of initial conditions and microphysical schemes will also be compared and discussed. This study can provide helpful perspectives for the probabilistic cloud-resolving hurricane forecasts and the predictability research of hurricane systems.

4.11.2 Methodology

The first sets of experiments are to examine the additional contribution of microphysical schemes on the uncertainty growth. The 32-member ensemble forecasts similar to ExpAll in section 4.6.1 are conducted except that the multiple microphysical schemes are applied in the ensemble to add forecast model uncertainty in ExpAllMulti. Among these 32 members, ExpAllMulti uses Lin, LFO04, Schultz and WSM6 in every 8 members. In the ensemble forecasts in ExpAll, the only uncertainty source is from the initial conditions at 0600 UTC 0913 while ExpAllMulti includes both initial conditions and microphysical schemes uncertainty. The intensity and track forecasts from ExpAll and ExpAllMulti will be compared.

The second sets of experiments are to investigate the individual and combined contributions of the initial conditions and microphysical schemes on the uncertainty growth. Three 4-member ensembles are designed for this purpose and listed in Table 4.3 to 4.5. In these three ensembles, Exp4PERT includes only initial conditions uncertainty, Exp4PHYS includes only microphysical schemes uncertainty and Exp4FULL includes both initial conditions and microphysical schemes uncertainty. There are 24 possible combinations to include both initial conditions and microphysical schemes perturbations in the 4-member ensemble. In this study, only one combination is used for simplification.

	Initial condition	Microphysical Schemes
MEM1	Ensemble mean	LFO83
MEM2	IC1	LFO83
MEM3	IC2	LFO83
MEM4	IC3	LFO83

Table 4.3: The members in the ensemble forecasts of Exp4PERT. The initial condition of ensemble mean is from the ensemble mean analysis of ExpAll at 0600 UTC 0913. Initial conditions of IC1-IC3 are from 3 members of ExpAll final ensemble analysis at 0600 UTC. The acronyms of microphysical schemes can be found in Chapter 3 section 3.2.1.

	Initial condition	Microphysical Schemes
MEM1	Ensemble mean	LFO83
MEM2	Ensemble mean	LFO04
MEM3	Ensemble mean	Schultz
MEM4	Ensemble mean	WSM6

Table 4.4: The members in the ensemble forecasts of Exp4PHYS. The initial condition of ensemble mean is from the ensemble mean analysis of ExpAll at 0600 UTC 0913. The acronyms of microphysical schemes can be found in Chapter 3 section 3.2.1.

	Initial condition	Microphysical Schemes
MEM1	Ensemble mean	LFO83
MEM2	IC1	LFO04
MEM3	IC2	Schultz
MEM4	IC3	WSM6

Table 4.5: The members in the ensemble forecasts of Exp4FULL. The acronyms have the same meanings as in Table 4.3 and Table 4.4.

4.11.3 Single microphysical scheme vs. multiple microphysical schemes

The ensemble intensities of ExpAllMulti are plotted in Fig 4.63. The intensity spreads of ExpAll and ExpAllMulti are plotted in Fig 4.64.

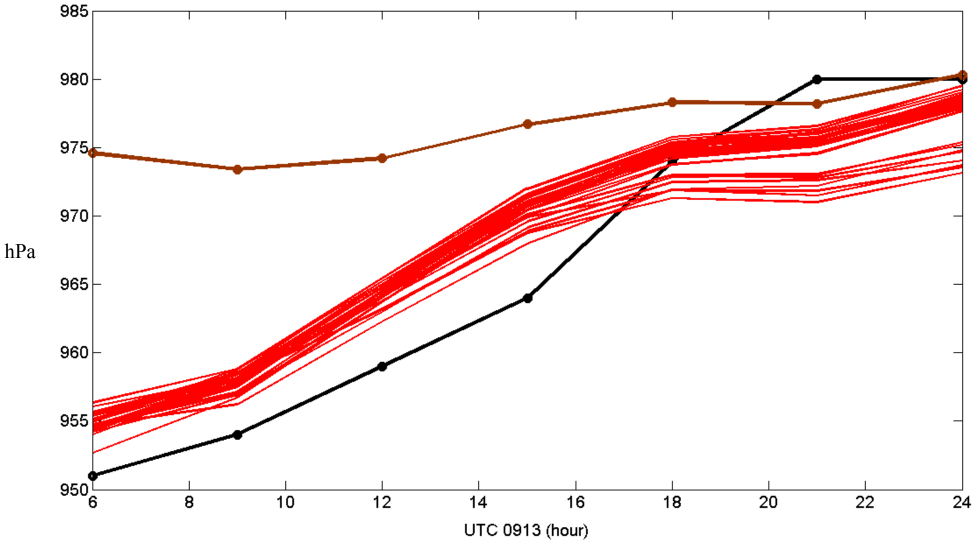

Fig. 4.64. The predicted ensemble minimum SLP of ExpAllMulti (red), compared with the best track (black) and NoDA (brown).

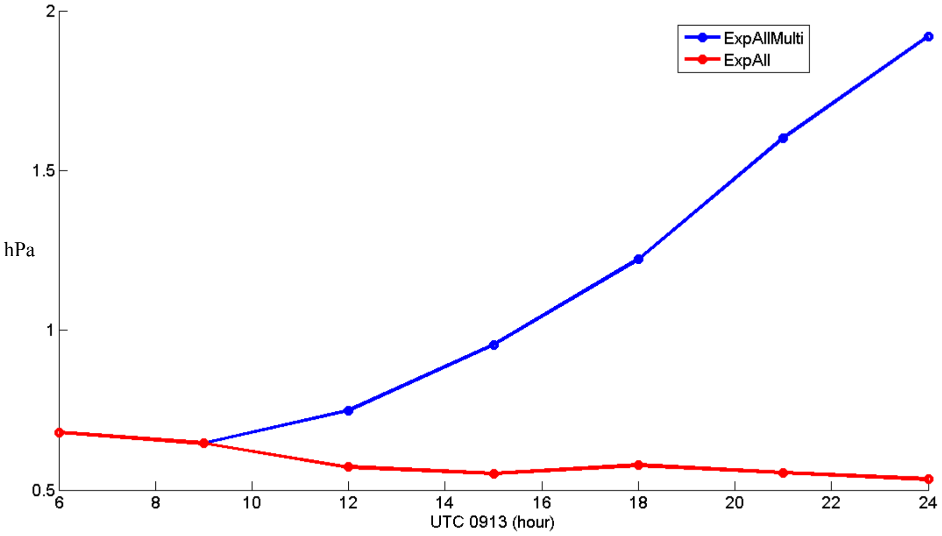

Fig. 4.65. The spread of intensity forecasts of ExpAll (red) and ExpAllMulti (blue).

The ensemble intensities in ExpAllMulti are similar to ExpAll during the first 3 hours of forecast starting from 0600 UTC (Fig. 4.16 and Fig. 4.63). After 0900 UTC, the spread of ExpAllMulti starts to increase monotonically with time while the spread of

ExpAll does not change greatly until the end of forecast (Fig. 4.64). The growth rate of the spread of ExpAllMulti also increases with time, showing a faster spread growth in the last 2-3 cycles than earlier. The final intensity spread of ExpAllMulti is about three times of the initial spread and much larger than the final spread of ExpAll.

It is interesting to notice that there is a bifurcation point at 1800 UTC in the ensemble intensity forecasts of ExpAllMulti (Fig. 4.63). 8 members using the Schultz scheme diverge from the other members with stronger vortex forecasts and contribute to the spread increase in the last 6 hours of forecast.

The ensemble mean intensity of ExpAllMulti is very close to the mean of ExpAll (not shown), indicating there is no forecast skill improvement in the intensity forecast from the multiple microphysical schemes. On the other hand, the model microphysical parameterizations uncertainty increases the intensity spread of ensemble forecasts in this specific case.

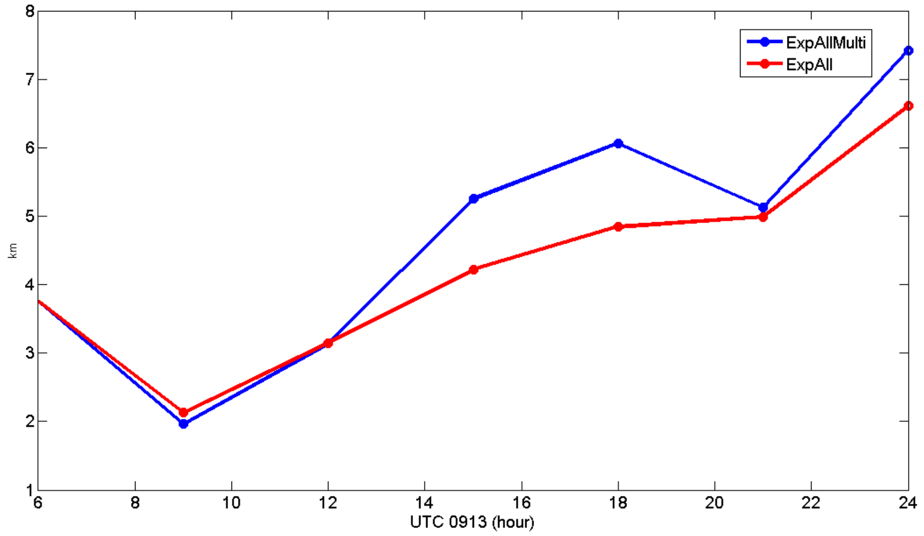

Fig. 4.66. The track spread of ensemble forecasts of ExpAll (red) and ExpAllMulti (blue).

The track spreads of ensemble forecasts ExpAll and ExpAllMulti are plotted in Fig. 4.65. The track spread of ExpAllMulti is similar to ExpAll generally. At 1500, 1800 UTC 13 September and 0000 UTC 14 September, there are 1-1.5 km increases of track spread from ExpAllMulti. Since the mean track error of ExpAll is as large as 14 km, the difference of track spreads between ExpAll and ExpAllMulti are generally negligible during most of the forecast times.

4.11.4 Comparison of contributions on uncertainty growth from IC and microphysical scheme perturbations

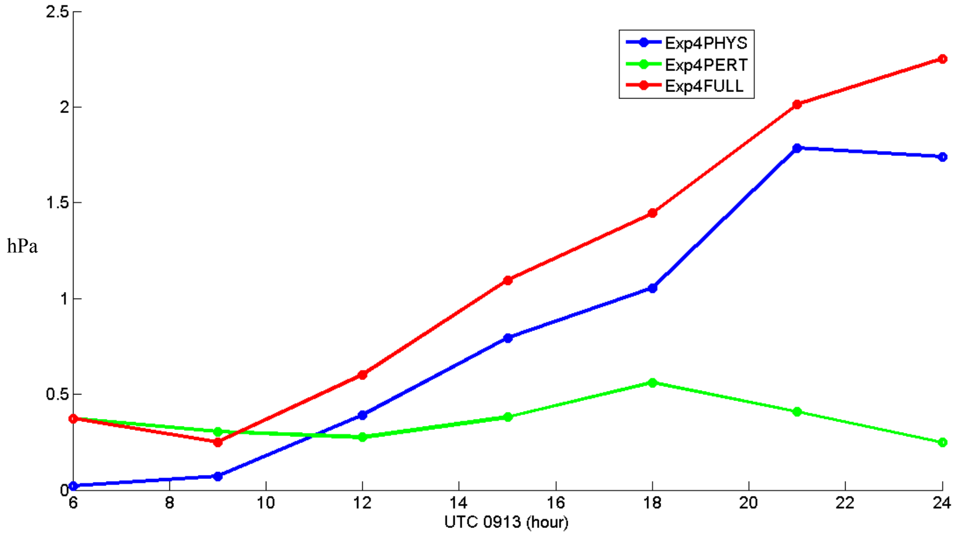

Fig. 4.67. The spread of intensity forecasts of Exp4PHYS (blue), Exp4PERT (green) and Exp4FULL (red).

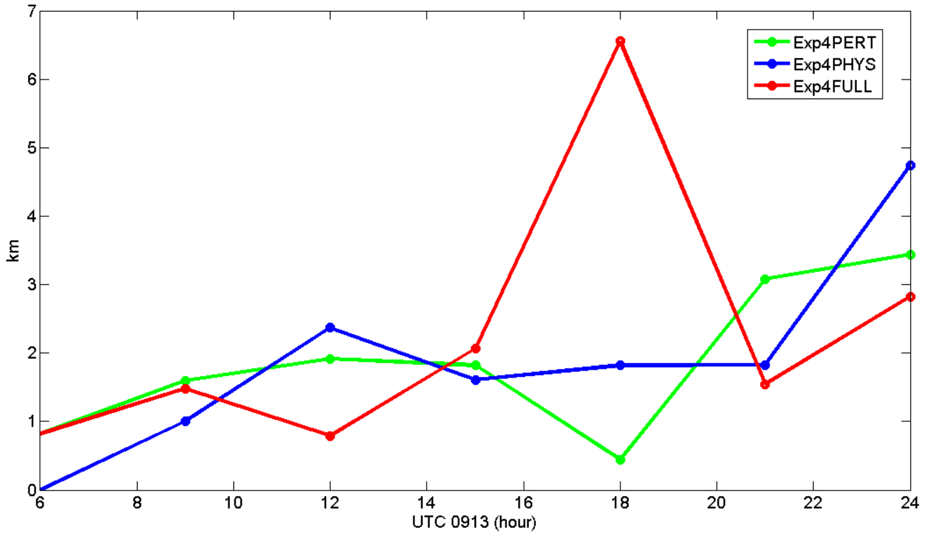

Fig. 4.68. The track spreads of Exp4PHYS (blue), Exp4PERT (blue) and Exp4FULL (red).

The intensity and track spread of Exp4PHYS, Exp4PERT and Exp4FULL are plotted in Fig. 4.66 and Fig. 4.67. Starting from a zero spread with the same IC in each ensemble member at 0600 UTC, the intensity spread of Exp4PHYS increases with forecast time, reaching to 1.6 hPa at the end of the 18-hour forecasts (Fig. 4.66). The intensity spread of Exp4PERT, on the other hand, does not change dramatically during the 18 hours of forecasts. When the initial condition perturbations and the microphysical schemes perturbations are combined together in Exp4FULL, the spread is larger than Exp4PERT and Exp4PHYS after 1200 UTC. The growth rate of intensity spread in Exp4FULL is close to Exp4PHYS, demonstrating the dominate role of microphysical schemes perturbations in intensity uncertainty growth in this case. For most of the forecast times, the spread of Exp4FULL is always smaller than the linear combination of the spreads of Exp4PERT and Exp4PHYS due to the nonlinear effect.

For the track forecasts, the track spreads in all of the three ensembles are comparable except for Exp4FULL at 1800 UTC with a higher track spread than other two ensembles (Fig. 4.67).

4.12 Summary

The impact of radar data assimilation on the analysis and forecast of Hurricane Ike's (2008) intensity, track and precipitation is investigated with the cloud-resolving ARPS model and the ARPS EnKF data assimilation system in this study. Radial velocity (Vr) and reflectivity (Z) observations from two coastal Doppler radars are assimilated within 2 hours. With prior multiplicative and posterior additive covariance inflations, the ensemble spread is well maintained and large impacts from the observations are obtained on the analyzed wind and microphysical fields. The assimilation of data in the first one to two cycles clearly strengthens the hurricane vortex.

The assimilation of radar observations is also found to evidently improve the structure, intensity, track and precipitation forecasts of Ike. Assimilating Vr alone leads to a much greater improvement in the intensity forecast than assimilating Z alone. For the track forecast, Vr alone produces a slightly better forecast than Z alone. Z alone results in a precipitation forecast improvement that lasted longer than using Vr alone. Assimilating both Vr and Z has similar results as assimilating Vr alone, indicating a dominant role of Vr data when analyzed using EnKF.

The similarity in the track and intensity forecasts between the cases assimilating radial wind alone and assimilating both radial wind and reflectivity is also observed by Pu et al. (2009) in their air-borne radar data assimilation with WRF 3DVAR. They also observed small impacts on hurricane intensity and track forecasting from reflectivity data and attributed the smaller improvement to the simple warm rain microphysics scheme used in their study. Zhao and Jin (2008) observed a marginal or negligible impact on the intensity analysis and forecast from assimilating reflectivity alone with 3DVAR. Zhao and Xue (2009) noticed significant improvements from reflectivity assimilation in Ike intensity forecast when using the ARPS 3DVAR combined with a complex cloud analysis package employing an ice-microphysics scheme. In our experiments with EnKF, which use the same ice microphysics scheme as Zhao and Xue (2009), there is evident improvements from reflectivity data in the track forecast but only a mild impact on the intensity forecast. Existing studies on the impact from radar reflectivity observations on the hurricane intensity and track forecast are few and more future studies are needed to investigate this topic in depth.

Ensemble forecasts starting from our EnKF analyses exhibit some uncertainty growth in track, but not much growth in intensity spread. The latter is most likely due to the weakening of the hurricane itself during the forecast period. The ensemble mean precipitation forecast has better performance than the deterministic forecast in stratiform precipitation regions, giving higher ETSs owing to the averaging effect among ensemble members. The probability matching technique improves the ETS of the ensemble precipitation forecast for heavy rainfall thresholds. Overall, the ETS with probability matching appears to be the highest or nearly the highest for all the thresholds evaluated.

The correlation between the initial intensity error and the forecast intensity error decreases with forecast time. There is no strong correlation between the track error and intensity error, suggesting different mechanisms controlling the track and intensity forecasts in this case.

The experiment with 30-minute assimilation cycles shows similar results to the 10-minute cycles, while assimilating radar data at 60-minute intervals fails to obtain a strong enough vortex in either analysis or forecast.

The experiment with both coastal radars has better track forecasts than using KHGX radar alone and a better intensity forecast than using KLCH alone, indicating, not surprisingly, advantages of using data from multiple Doppler radars.

When MSLP from the best track data is assimilated with radial velocity together, there is a 2-3 hPa improvement in intensity at the end of the analysis cycles when assimilated at 60-minute or 10-minute intervals. The impacts from MSLP last for only 3-6 hours during the forecast, however. When only reflectivity is assimilated from radar, the improvement due to MSLP is as large as 14-15 hPa in the final analysis. The improvement decreases with time but is still noticeable 12 hours into the forecast. The assimilation of MSLP in addition to radar data also improves the track analysis. The relative improvements in the track error due to assimilating additional MSLPs are always over 50% in the first 9 hours of forecast when reflectivity and MSLP is assimilated.

It appears to us that for Z assimilation experiments, the analyses on both intensity and track are not good enough, leaving more space for the improvement from additional MSLP assimilation, while for the Vr assimilation experiments with better analyses, the additional improvement from MSLP is much smaller.

With frequent assimilations of MSLP with 10-minute intervals, there is noticeable further improvement in the intensity forecast to 60-minute intervals for 12 hours when MSLP is assimilated with reflectivity.

The assimilation of MSLP alone at both 60-minute and 10-minute intervals shows better intensity forecast than assimilating Z alone but not as good as assimilating Vr alone. The assimilation of Z, Vr and MSLP together has a larger improvement in the intensity forecast than assimilating each observation type individually.

The addition of multiple microphysical schemes in the ensemble forecasts increases the intensity spread but not much the track spread. The uncertainty created by the microphysical scheme perturbations, if added, contributes more to the intensity spread than the initial condition perturbations.

Radar data assimilated in this study are mostly at the convective and sub-vortex scales; their impact on the forecast may therefore be limited to a relatively short period of time. Other observational sources such as surface observations, soundings, profilers and satellite observations can help improve the larger scale environment and improve the forecast at longer ranges. There are many other issues associated with the predictability of hurricanes, the EnKF data assimilation and the subsequent ensemble forecasting can be helpful when investigating these issues (Sippel and Zhang 2010).

Chapter 5

Summary and Future Plans

5.1 Summary

Atmospheric data assimilation combines observations and model states to produce the best estimate of the atmosphere that can serve as the initial condition for numerical weather prediction models. The accuracy of the initial condition depends greatly on the data assimilation algorithm. Traditional 3DVAR method uses static background error covariance which is not appropriate for the highly flow-dependent background error statistics of thunderstorm or hurricane systems. Ensemble Kalman filter (EnKF) uses ensemble forecasts to estimate and evolve flow-dependent error covariance statistics and is one of the most popular assimilation methods among the research community today. Compared with the 4DVAR method, the implementation of EnKF does not require the development of an adjoint model and its code can be easily linked with different prediction models.

Amid the applications of EnKF, cloud-resolving numerical analysis and prediction of thunderstorms and hurricanes have received much attention in recent studies, since both thunderstorms and hurricanes have significant societal impacts. This research focuses on the EnKF applications to the analysis and prediction of these two weather phenomena.

Radar is one of the most important observation platforms for thunderstorm analysis and forecast. However, when the radar is far away from a convective storm, it cannot observe the low levels due to the earth curvature effect. Surface observations can provide the low-level coverage for thunderstorm analysis, especially those from dense networks. Chapter 3 investigates, systematically, the impact of surface observations on thunderstorm analysis and forecast with a series of Observing System Simulation Experiments (OSSEs) for the first time. When the radar is located at a significant distance (e.g., the 115 and 185 km distances considered) from the main convective storm, clear positive impacts on the storm analysis and forecast are achieved by assimilating surface observations with a spacing of about 20 km. When the radar is located just 45 km from the storm center, a network spacing of 6 km is needed to achieve any noticeable positive impact. The impact of surface data in terms of relative error reduction increases linearly with decreased surface network spacing until the spacing is close to the grid interval of truth simulation. Assimilating observations from a coarser network over a longer period of time helps to achieve a similar level of impact as would be seen from a network of higher density. Given the typical observation errors, surface wind observations produce the largest positive impact, followed by temperature measurements. Pressure measurements produce the least impact. Assimilating all surface observation variables together yields the largest impact. The impact of surface data is sustained or even amplified during subsequent forecasts when their impact on the analysis is significant.

Chapter 4 of this study investigates the impact of assimilating coastal WSR-88D radar radial velocity and reflectivity data on Hurricane Ike (2008), on the analysis and forecast of its intensity, track and precipitation. The analysis and prediction were carried out on a large uniform resolution grid with a 4 km grid spacing and the ARPS was used as the prediction model. Radial wind and reflectivity data from two coastal radars were assimilated with the EnSRF algorithm at 10-minute, 30-minute and 60-minute intervals for a duration of two hours. This study represents the first time that assimilates reflectivity and/or MSLP together with radial velocity for hurricane forecast using an EnKF method.

With radar data assimilation, forecasts of hurricane's structure, track, intensity and precipitation are generally improved. The assimilation of radial velocity alone has a much greater improvement on the intensity forecast and a slightly larger impact on the track forecast than assimilating reflectivity alone. The quantitative precipitation forecast (QPF) initialized by assimilating reflectivity alone shows an improvement that lasts longer than assimilating radial wind alone, suggesting better analyzed microphysical fields from reflectivity observations. Assimilating both radial wind and reflectivity produces similar results as assimilating radial wind alone, suggesting the dominant role of radial velocity observations for hurricanes.

The ensemble forecasts starting from the ensemble analyses show some uncertainty growth in track forecasts. Probability matching processing for the ensemble improves the QPF scores over the simple ensemble mean for most precipitation thresholds.

The assimilation of additional MSLP from the best track data with 10-minute and 60-minute assimilation intervals both show solid improvement in both intensity and track forecasting when radial velocity data are not assimilated. When radial velocity data are also assimilated, the intensity forecast improvement is small but still recognizable. When time-interpolated MSLP values are assimilated without radar data at all at 60-minute or 10-minute interval, the improvement in the intensity forecast is smaller than radial velocity assimilation but larger than reflectivity assimilation. Assimilating all of Vr, Z and MSLP altogether shows better intensity forecast than assimilating them individually, suggesting the potential benefit of assimilating observations of multiple platforms.

Using multiple microphysical parameterization schemes increases the uncertainty growth of intensity. In a 4-member ensemble, the contribution to the uncertainty growth of intensity from the microphysical scheme perturbations is found to be larger than that from initial condition perturbations.

5.2 Future plans

As mentioned in Chapter 3, one of the limitations of the OSSEs for surface observation impact is that calibration and validation of the OSSE results were not performed therefore the data impact seen may not be reliable. Before the additional impact of surface observations is investigated, the results of the “existing” radar observations impact from OSSEs can be compared with the results from the Observing System Experiments (OSEs) using real radar observations. The calibration includes comparison of error statistics between simulated and real radar observations, comparison of data impacts on the analysis and forecast between OSEs and OSSEs. Through the calibration steps, the OSSE system can be validated. This calibration and validation provide the guidelines to interpret the impact of the “future” additional simulated surface observations in OSSE.

However, to practice the calibration effectively, a full set of observations needs to be included. Also, the impact of the radar observations can vary greatly for different convective thunderstorm cases, posing an additional challenge for the calibration. These issues need to be considered carefully in the future.

For EnKF analysis of hurricanes, more cases need to be examined in the future, beyond this Ph.D. program, to establish robust results on the radar data impact on tropical storm forecasting. Before working with the Ike case, we performed EnKF radar data assimilation experiments using data from multiple ground-based radars for Tropical Storm Erin (2007), which made landfall along Texas coast, re-intensified in Oklahoma and caused heavy flooding. Despite its re-intensification, the vortex structure of Erin was not as well defined in the radar observations as Ike was, and there was significant axis-asymmetry in the radar data coverage in the vortex region of Erin. The track forecast from the control simulation without radar data assimilation was poor. Assimilating radar data did not result in much improvement to the track forecasting in this case in our earlier experiments. The forecast was found to be sensitive to the initial background analysis, which was either from the operational NAM model or from the RUC. This suggests that there was significant uncertainty or error in the large scale environment that may have to be reduced before the track forecast can be improved. In such a case, assimilating additional observations together with the radar data may be most beneficial.

Hamill et al. (2010) assimilated conventional and non-conventional observations, including surface observations and rawinsondes, cloud track winds, aircraft observations, human-synthesized MSLP, satellite-based Global Positioning System (GPS) radio occultation and satellite radiance with EnKF for global models at 45 km and 30 km resolutions. During the 2009 Northern Hemisphere summer Tropical Cyclones (TCs) forecasts, they found generally better or competitive track forecasts from EnKF compared to operational global forecasts. They suggested that their increased resolutions over the National Centers for Environmental Predictions' (NCEP) operational global model with

137 km resolution and the improvement in the environmental steering wind in the EnKF analyses may have contributed to the better track forecast. However, the intensity forecasts in their study are problematic. They suspected that the possible reason is that their resolutions of around 40 km are not fine enough to produce realistic vortex structures.

A parallel MPI version of the ARPS EnKF that can handle both radar and conventional observations at the same time has been recently developed at CAPS, which would enable experiments including a full set of observations. These observations may include, but not limited to, the surface observations, the rawinsondes, the profilers, MSLP from the best track estimate dataset, the radar observations and the satellite radiance. Also, realistic synoptic environmental perturbations based on, e.g., global ensemble forecasts, should be introduced to properly sample uncertainties in the storm environment. Will the assimilation of data covering a variety of scales improve the hurricane analysis and forecast of different cases? This is our question for the future work. We plan to pursue such research for the Erin and other cases in the future.

References

Aksoy, A., D. C. Dowell, C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805-1824.

Aksoy, A., D. Dowell and C. Snyder, 2010: A multi-case comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Wea. Rev.*, **138**, 1273-1292.

Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741-2758.

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

Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634-642.

Anderson, J. L., B. Wyman, S. Zhang, and T. Hoar, 2005: Assimilation of surface pressure observations using an ensemble filter in an idealized global atmospheric prediction system. *J. Atmos. Sci.*, **62,** 2925–2938.

Anderson, J. L., N. Collins, 2007: Scalable implementations of ensemble filtering algorithms for data assimilation. *J. Atmos. Ocean Tech.*, **24**, 1452-1463.

Anderson, J. L., 2007: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99-111.

Anderson, J. L., 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus*, **61A**, 72-83.

Anderson, J. L., 2009: Ensemble Kalman filters for large geophysical applications. *IEEE Cntrl. Sys. Mag.*, **29**, 66-82.

Andrews, A., 1968: A square root formulation of the Kalman covariance equations. *AIAA J.*, 6, 1165-1168.

Anthes, R. A., 1982: Tropical cyclones—their evolution, structure, and effects. Meteor. Monogr., No. 41, Amer. Meteor. Soc., 208 pp.

Battan, L. J., 1973: Radar observation of the atmosphere. Univ. of Chicago Press, Chicago, 324pp.

Benjamin, S. G. and P. A. Miller, 1990: An alternative sea level pressure reduction and a statistical comparison of geostrophic wind estimates with observed surface winds. *Mon. Wea. Rev,* **118**, 2099–2116.

Berg, R., 2009: Hurricane Ike tropical cyclone report. National Hurricane Center.

Bishop, C. H., B. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transformation Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.,* **129**, 420-436.

Brewster, K., 1996: Application of a Bratseth analysis scheme including Doppler radar data. *Preprints, 15th Conf. Wea. Anal. Forecasting*, Norfolk, VA, Amer. Meteor. Soc., 92-95.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, 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. Wea. Rev.*, **138**, 1550-1566.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, B. He., 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev.*, **138**, 1567-1586.

Buizza, R., M. Miller and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887-908.

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

Chen, S. S., J. F. Price, W. Zhao, M. A. Donelan, and E. J. Walsh, 2007: The CBLAST-Hurricane Program and the nextgeneration fully coupled atmosphere–wave–ocean models for hurricane research and prediction. *Bull. Amer. Meteor. Soc.*, **88**, 311–317.

Chen, Y., and C. Snyder, 2007: Assimilating vortex position with an ensemble Kalman filter. *Mon. Wea. Rev.,* **135**, 1828-1845.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.,* **124**, 1767–1785.

Davis, C. A., and Coauthors, 2008: Prediction of landfalling hurricanes with the advanced hurricane WRF model. *Mon. Wea. Rev.,* **136**, 1990–2005.

Dawson, D. T., II and M. Xue, 2006: Numerical forecasts of the 15-16 June 2002 Southern Plains severe MCS: Impact of mesoscale data and cloud analysis. *Mon. Wea. Rev.*, **134,** 1607-1629.

Deng, X. and R. Stull, 2007: Assimilating surface weather observations from complex terrain into a high-resolution numerical weather prediction model. *Mon. Wea. Rev.*, **135,** 1037–1054.

Doviak, R. and D. Zrnic, 1993: Doppler radar and weather observations. 2nd ed. Academic Press, 562 pp.

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

Ebert, E.E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461-2480.

Ehrendorfer, M., 2007: A review of issues in ensemble-based Kalman filtering. *Meteor. Z.*, **16**, 795-818.

Eilts, M. D., and S. D. Smith, 1990: Efficient dealiasing of Doppler velocities using local environment constraints. *J. Atmos. Ocean Tech.*, **7**, 118–128.

Emanuel, K., 2005: Divine wind: The history and science of hurricanes. Oxford University Press, 296 pp.

Evensen, G., 1992: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *J. Geophys. Res*., **97**(C11): 17 905-17 924.

Evensen, G., 1993: Open boundary conditions for the extended Kalman filter with a quasi-geostrophic ocean model. *J. Geophys. Res.*, **98**(C9): 16529-16546.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10143-10162.

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

Fierro, A. O., R. F. Rogers, F. D. Marks, D. S. Nolan, 2009: The impact of horizontal grid spacing on the microphysical and kinematic structures of strong tropical cyclones simulated with the WRF-ARW model. *Mon. Wea. Rev.,* **137**, 3717-3743.

Fovell, R. G., K. L. Corbosiero, and H.C. Kuo, 2009: Cloud microphysics impact on hurricane track as revealed in idealized experiments. *J. Atmos. Sci.*, **66**, 1764-1778.

Fovell, R. G., K. L. Corbosiero, A. Seifert, and K.-N. Liou, 2010: Impact of cloud-radiative processes on hurricane track. *Geoph. Res. Lett.*, **37**, L07808, doi:10.1029/2010GL042691.

Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846-1868.

Gao, J. and K. K. Droegemeier, 2004: A variational technique for dealising Doppler radial velocity data. *J. Appl. Meteor.*, **42,** 934-940.

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

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004a: Precipitation and evolution sensitivity in simulated deep convective storms: Comparisons between liquid-only and simple ice and liquid phase microphysics. *Mon. Wea. Rev.*, **132,** 1897-1916.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004b: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Mon. Wea. Rev.*, **132,** 2610-2627.

Hacker, J. P. and C. Snyder, 2005: Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133,** 3260-3275.

Hacker, J. P. and D. Rostkier-Edelstein, 2007: PBL state estimation with surface observations, a column model, and an ensemble filter. *Mon. Wea. Rev.*, **135,** 2958-2972.

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

Hamill T.M., and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation: a comparison of different approaches. *Mon. Wea. Rev.*, **133**, 3132-3147.

Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. Chapter 6 of *Predictability of Weather and Climate*, Cambridge Press, 124-156.

Hamill T.M. and J.S.Whitaker, 2010: What constrains spread growth in forecasts initialized from ensemble Kalman filters? *Mon. Wea. Rev.*, conditionally accepted.

Hamill T.M., J.S.Whitaker, M. Fiorino and S. G. Benjamin, 2010: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.,* conditionally accepted.

Hansen, J. A., 2002: Accounting for model error in ensemble-based state estimation and forecasting. *Mon. Wea. Rev.*, **130**, 2373-2391.

Hong, S.-Y. and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42,** 129-151.

Hou, D., E. Kalnay and K. K. Droegemeier, 2001: Objective verification of the SAMEX-98 ensemble forecasts. *Mon. Wea. Rev.*, **129**, 73-91.

Houtekamer, P. L., J. Derome, H. Ritchie and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225-1242.

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

Houtekamer, P. L. and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129,** 123-137.

Houtekamer, P. L., and H. L. Mitchell, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126-2143.

Houtekamer, P. L., and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **131**, 3269-3289.

Houtekamer, P. L., and H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604-620.

Houze, R. A., S. S. Chen, B. F. Smull, W.-C. Lee, and M. M. Bell, 2007: Hurricane intensity and eyewall replacement. *Science*, **315**, 1235–1238.

Hu, M., M. Xue, J. Gao, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part II: Impact of radial velocity analysis via 3DVAR. *Mon. Wea. Rev.*, **134,** 699-721.

Jazwinski, A. H., 1970: Stochastic processes and filtering theory. Academic Press, 376 pp.

Jung, Y., M. Xue and G. Zhang, 2010: Simultaneous estimation of microphysical parameters and the atmospheric state using simulated polarimetric radar data and an ensemble Kalman filter in the presence of an observation operator error. *Mon. Wea. Rev.*, **138**, 539-562.

Kalnay, E., 2002: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, 364 pp.

Keppenne, C. L., and M. M. Rienecker, 2002: Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.*, **130**, 2951-2965.

Khain, A., B. Lynn, J. Dudhia, 2010: Aerosol effects on intensity of landfalling hurricanes as seen from simulations with the WRF model with spectral bin microphysics. *J. Atmos. Sci.*, **67**, 365-384.

Kimberlain, T., Tropical cyclone motion and intensity talk, 2007.

Krishnamurti, T. N., S. Pattnaik, L. Stefanova, T. S. V. Vijaya Kumar, B. P. Mackey, A. J. O'Shay, and R. J. Pasch, 2005: The hurricane intensity issue. *Mon. Wea. Rev.*, **133**, 1886– 1912.

Kong, F., M. Xue, K. K. Droegemeier, D. Bright, M. C. Coniglio, K. W. Thomas, Y. Wang, D. Weber, J. S. Kain, S. J. Weiss, and J. Du, 2007: Preliminary analysis on the real-time storm-scale ensemble forecasts produced as a part of the NOAA hazardous weather testbed 2007 spring experiment. 22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Pred., Salt Lake City, Utah, Amer. Meteor. Soc., CDROM 3B.2.

Kong, F. M. Xue, K. W. Thomas, J. Gao, Y. Wang, K. Brewster, K. K. Droegemeier, J. Kain, S. Weiss, D. Bright, M. Coniglio, and J. Du, 2009: A realtime storm-scale ensemble forecast system: 2009 spring experiment. 23nd Conf. Wea. Anal. Forecasting/19th Conf. Num. Wea. Pred., Omaha, NB, Amer. Meteor. Soc., Paper 16A.3.

Kwon, I. H. and H. B. Cheong, 2010: Tropical cyclone initialization with a spherical high-order filter and an idealized three-dimensional bogus vortex. *Mon. Wea. Rev.,* **138**, 1344-1367.

Lazarus, S., C. M. Ciliberti, J. D. Horel, and K. Brewster, 2002: Near-real-time applications of a mesoscale analysis system to complex terrain. *Wea. Forecast.*, **17,** 971-1000.

Lewis, J. M., S. Lakshmivarahan, S. Dhall, 2006: Dynamic data assimilation: a least squares approach, Cambridge University Press, 680 pp.

Li, X., and Z. Pu, 2008: Sensitivity of numerical simulation of early rapid intensification of Hurricane Emily (2005) to cloud microphysical and planetary boundary layer parameterizations. *Mon. Wea. Rev.*, **136**, 4819–4838.

Li, X. and Z. Pu, 2009: Sensitivity of numerical simulation of early rapid intensification of Hurricane Emily (2005) to cumulus parameterization schemes in different model horizontal resolutions, *J. Meteor. Soc. Japan.*, **87**, 403-421.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22,** 1065-1092.

Liu, H. and M. Xue, 2008: Prediction of convective initiation and storm evolution on 12 June 2002 during IHOP. Part I: Control simulation and sensitivity experiments. *Mon. Wea. Rev.***,** **136**, 2261-2283.

Lord, S. J., E. Kalnay, R. Daley, G. D. Emmitt, and R. Atlas, 1997: Using OSSEs in the design of the future generation of integrated observing systems. *Preprint volume,* *1st Symposium on Integrated Observation Systems*, Long Beach, CA, Amer. Meteor. Soc., 45-47.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183-3203.

McCaul, E. W., Jr., and M. L. Weisman, 2001: The sensitivity of simulated supercell structure and intensity to variations in the shapes of environmental buoyancy and shear profiles. *Mon. Wea. Rev.*, **129**, 664-687.

McFarquhar, G. M., H. Zhang, G. Heymsfield, R. Hood, J. Dudhia, J. B. Halverson, and F. Marks, 2006: Factors affecting the evolution of Hurricane Erin (2001) and the distributions of hydrometeors: Role of microphysical processes. *J. Atmos. Sci.*, **63**, 127–150.

McLaughlin, D., E. Knapp, Y. Wang, and V. Chandrasakar, 2007: Short wavelength technology and the potential for distributed networks of small radar systems. *IEEE Radar 2007 Conference Digest*.

Meng, Z., and F. Zhang, 2007: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403-1423.

Meng, Z., and F. Zhang, 2008a: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVar in a real-data case study. *Mon. Wea. Rev.*, **136**, 522-540.

Meng, Z., and F. Zhang, 2008b: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Comparison with 3DVar in a month-long experiment. *Mon. Wea. Rev.*, **136**, 3671-3682.

Mitchell, H. L., and P. L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Wea. Rev.*, **128**, 416-433.

Mitchell, H. L., P. L. Houtekamer and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791-808.

Miyoshi, T., A review of the deterministic ensemble Kalman filtering methods and related techniques. Department of Meteorology, University of Maryland, 21.

Morss, R. E., K. A. Emanuel, and C. Snyder, 2001: Idealized adaptive observation strategies for improving numerical weather prediction. *J. Atmos. Sci.*, **58,** 210-232.

National Hurricane Center, 2010: National Hurricane Center forecast verification. http://www.nhc.noaa.gov/verification/verify5.shtml.

Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus,* **56A**, 415-428.

Park, S.K. and K.K. Droegemeier, 2000: Sensitivity analysis of a 3-D convective storm: Implications for variational data assimilation and forecast error. *Mon. Wea. Rev.*, **128**, 140-159.

Penland, C., 2003: A stochastic approach to nonlinear dynamics: a review. (Electronic supplement to 'Noise out of chaos and why it won't go away'), *Bull. Am. Meteorol. Soc.*, **84**, 921-925.

Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.*, **129**, 1194-1207.

Potter, J., 1964: W matrix augmentation. M.I.T. Instrumentation Laboratory Memo SGA 5–64, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Pu, Z.-X., and S. Braun, 2001: Evaluation of bogus vortex techniques with four - dimensional variational data assimilation. *Mon. Wea. Rev.*, **129**, 2023-2039.

Pu, Z., X. Li and J. Sun., 2009: Impact of airborne Doppler radar data assimilation on the numerical simulation of intensity changes of Hurricane Dennis near a landfall. *J. Atmos. Sci.*, **66**, 3351-3365.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419.

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

Richardson, Y. P., 1999: The influence of horizontal variations in vertical shear and low-level moisture on numerically simulated convective storms, School of Meteorology, University of Oklahoma, 236.

Richardson, Y. P., K. K. Droegemeier, and R. P. Davies-Jones, 2007: The influence of horizontal environmental variability on numerically simulated convective storms, Part II: Variations in low-level moisture. To be submitted to *Mon. Wea. Rev.*

Rogers, R., and Coauthors, 2006: The Intensity Forecasting Experiment (IFEX): A NOAA multiyear field program for improving tropical cyclone intensity forecasts. *Bull. Amer. Meteor. Soc.*, **87**, 1523–1537.

Schultz, P., 1995: An explicit cloud physics parameterization for operational numerical weather prediction. *Mon. Wea. Rev.*, **123,** 3331-3343.

Sippel, J. A. and F. Zhang, 2010: Factors affecting the predictability of hurricane Humberto (2007). *J. Atmos. Sci.*, **63**, 1759-1778.

Snook, N. and M. Xue, 2008: Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms. *Geophy. Res. Letters,* **35**, L24803, doi:10.1029/2008GL035866.

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. Wea. Rev.*, **128**, 2077–2107.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield, J. A. Yorke, 2008: A local ensemble transform Kalman Filter data assimilation system for the NCEP global model. Tellus A, **60**, 113-130.

Tong, M. and M. Xue, 2005a: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSSE Experiments. *Mon. Wea. Rev.*, **133**, 1789-1800.

Tong, M. and M. Xue, 2005b: Simutaneous retrieval of microphysical parameters and atmospheric state variables with radar data and ensemble Kalman filter method. Extended abstract, 17th Conf. Num. Wea. Pred., Washington DC, Amer. Meteor. Soc., P1.30.

Tong, M., 2006: Ensemble Kalman filer assimilation of Doppler radar data for the initialization and prediction of convective storms, School of Meteorology, University of Oklahoma, 243.

Tong, M. and M. Xue, 2008a: Simultaneous estimation of microphysical parameters and atmospheric state with radar data and ensemble square-root Kalman filter. Part II: Parameter estimation experiments. *Mon. Wea. Rev.*, **136**, 1649–1668.

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

Wallace, J. M., and P.V. Hobbs, 1977: Atmospheric Science, Academic Press, 464 pp.

Wang, Y., 2009: How do outer spiral rainbands affect tropical cyclone structure and intensity? *J. Atmos. Sci.*, **66**, 1250-1273.

Weckwerth, T. M., D. B. Parsons, S. E. Koch, J. A. Moore, M. A. LeMone, B. B. Demoz, C. Flamant, B. Geerts, J. Wang, and W. F. Feltz, 2004: An overview of the International H2O Project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.*, **85,** 253-277.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520.

Weisman, M. L., and J. B. Klemp, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, **112**, 2479–2498.

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

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2007: Ensemble data assimilation with the NCEP global forecast system, *Mon. Wea. Rev.*, **136**, 463-482.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecast system. *Mon. Wea. Rev.*, **136**, 463-482.

Whitaker, J. S., and T. M. Hamill, 2010: A simple state-dependent multiplicative inflation algorithm, the 4th Ensemble Kalman Filter Workshop.

Wikipedia: Hurricane Ike.

Xiao, Q., Y. H. Kuo, J. 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. Meteor. Climatol.*, **46**, 14–22.

Xiao, Q., L. Chen and X. Zhang, 2009a: Evaluations of BDA scheme using the Advanced Research WRF (ARW) model. *J. Appl. Meteor., Climat.*, **48**, 680-689.

Xiao, Q., X. Zhang, C. Davis J. Turlle, G. Holland and P. J. Fitxpatrick, 2009b: Experiments of hurricane initialization with airborne Doppler radar data for the Advanced-research Hurricane WRE (AHW) model. *Mon. Wea. Rev.*, **137**, 2758-2777.

Xu, Q., H. Lu, S. Gao, M. Xue, and M. Tong, 2008: Time-expanded sampling for ensemble Kalman filter: Assimilation experiments with simulated radar observations. *Mon. Wea. Rev.*, **136**, 2651-2667.

Xue, M. and W. J. Martin, 2006: A high-resolution modeling study of the 24 May 2002 case during IHOP. Part I: Numerical simulation and general evolution of the dryline and convection. *Mon. Wea. Rev.*, **134,** 149–171.

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

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

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

Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Ocean Tech.*, **23**, 46-66.

Xue, M., Y. Jung, and G. Zhang, 2007: Error modeling of simulated reflectivity observations for ensemble Kalman filter data assimilation of convective storms. *Geophys. Res. Letters*, **34,** L10802, doi:10.1029/2007GL029945.

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

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

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

Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter: Humberto (2007). *Mon. Wea. Rev.*, **137**, 2105-2125.

Zhao, Q. Y., and Y. Jin, 2008: High-resolution radar data assimilation for Hurricane Isabel (2003) at landfall. *Bull. Am. Meteorol. Soc.*, **89**, 1355–1372.

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). *Geophy. Res. Letters*, **36**, L12803.

Zou, X., and Q. Xiao, 2000: Studies on the initialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.*, **57**, 836–860.

Zou, X., Y. Wu and P. S. Ray, 2010: Verification of a high-resolution model forecast using airborne Doppler radar analysis during the rapid intensification of Hurricane Guillermo. *J. Appl. Meteor., Climat.*, **49**, 807-820.