UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ON THE FURTHER STUDIES OF SUITABLE STORM-SCALE 3DVAR DATA ASSIMILATION FOR THE PREDICTION OF TORNADIC THUNDERSTORMS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

GUOQING GE
Norman, Oklahoma
2011

ON THE FURTHER STUDIES OF SUITABLE STORM-SCALE 3DVAR DATA ASSIMILATION FOR THE PREDICTION OF TORNADIC THUNDERSTORMS

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEROLOGY

BY

Dr. Kelvin Droegemeier, Chair

Dr. Jidong Gao, Co-Chair

Dr. Alan Shapiro

Dr. Luther White

Dr. Qin Xu

Dr. Ming Xue

Acknowledgement

First of all, I would like to thank my advisor Dr. Kelvin Droegemeier and co-advisor, Dr. Jidong Gao, for their teaching, guidance, advice and encouragement during my Ph.D. study at University of Oklahoma. I am especially grateful to Dr. Jidong Gao for his support, kindness and patience in mentoring me through my dissertation research and in editing the dissertation. His broad knowledge and expertise on the storm-scale data assimilation and numerical weather prediction help me greatly.

I also owe my deeply thanks to all other members of my doctoral committee: Dr. Alan Shapiro, Dr. Luther White, Dr. Qin Xu and Dr. Ming Xue for their reading through my dissertation and providing valuable comments. Without their important suggestions, this dissertation would not be finished.

Dr. Ming Xue is deeply appreciated for his critical support and important advices on my research. Dr. Xu is gratefully thanked for very valuable discussions with him. Dr. Keith Brewster is thanked for his helpful comments and editing the Chapter 3.

CAPS scientists and staffs are acknowledged for their providing powerful computing capability and maintaining a research friendly atmosphere. I benefit greatly from the CAPS weekly science meeting and discussion with other CAPS scientists and graduate students. Dr. Yunheng Wang is thanked for his professional support and generous help on ARPS model and ARPS code development. Scott Hill is thanked for computer system support. Dr. Ming Hu is thanked for his useful suggestions. Dr. Jili Dong is thanked for his help and friendship. Thanks also go to my office mate Nate

Snook for his kindness and help on my research. Dr. Yungsun Jung and Dr. Ting Lei are acknowledged for interesting discussions with them.

Celia Jones, Marcia Pallutto, Eileen Hasselwander and Debra Farmer are thanked for their kind helps that make me studying here more smoothly.

I would like to dedicate this dissertation to my wife, Lesheng Hua, my son, Justin Ge, my parents and my parents-in-law. My wife and my son accompany me through the tough time and many challenges in my study here. My wife takes good care of the family while I am busy at my dissertation research. My parents and parents-in-law are always ready to help when we need. They understand my choice and support me whenever they can. My sisters, my brothers-in-law, Xuhui Pan, De Gu and Chongxiu Dou are also gratefully thanked for their continuous help and encouragement during my Ph.D. study.

This research was supported by NSF Grants ATM-0331756, ATM-0738370, ATM-0608168, ATM-0802888, ATM-0530814 and EEC-031347. The experiments in this dissertation were performed at Pittsburgh Supercomputing Center (PSC) and Oklahoma Supercomputing Center for Research and Education (OSCER) at University of Oklahoma.

Table of Contents

List of Tables

List of Figures

Abstract

Storm-scale 3DVAR data assimilation and NWP for the prediction of tornadic supercell thunderstorms still faces many challenges. Some fundamental issues are still not thoroughly (or explicitly) investigated. To name a few: what data field(s) plays the most important role in storm-scale data assimilation? How much information is required to get a quality data assimilation results? What is the model's first response to different types of observations? How will the neglecting of beam broadening and earth curvature factors in radar forward observation operator affect the data assimilation results? How to build a more dynamic consistent analysis by imposing weak constraints in cost function that is aimed to couple different model variables? This dissertation tries to address some of these questions.

The impacts of different data fields are firstly investigated. OSS Experiments are conducted under a simplified 3DVAR framework. The model's first responses at storm scale to the assimilation of different types of observations are thoroughly examined. It is also demonstrated that the horizontal wind fields have the greatest impact on the storm-scale data assimilation. In addition to the horizontal wind fields, extra observations from other model variables will improve the quality of data assimilation. Among these "other model variables", the water vapor field exerts the largest impact. A follow-on real case study confirms the important role of wind fields.

The impact of beam broadening or earth curvature on storm-scale 3DVAR data assimilation is also examined using OSS experiments. It is shown that the effect of beam broadening can be generally overlooked in storm-scale radar data assimilation without noticeable degradation of assimilation results. However, the effect of earth curvature can

only be neglected when the radar is near the storm (within 60 km as demonstrated by this study). The impact of refractive index gradient is also tested and shown to be small.

To help boost dynamic consistency among model variables, the storm-scale diagnostic pressure equation is incorporated into the storm-scale 3DVAR cost function in the form of a weak constraint. The impact of the constraint has been examined by applying it to case studies of one idealized tornadic supercell thunderstorm and two real-world tornadic supercell thunderstorms. It is demonstrated in the idealized case study that at single analysis step, the use of the constraint can help slightly improve the analysis of wind fields and pressure field. After a given period of intermittent data assimilation, the use of the constraint can evidently improve the quality of the data assimilation results. For the 8 May 2003 OKC tornadic supercell thunderstorm case, it is shown that the use of the constraint help improve the forecast in term of the general evolution and the mesocyclone rotation of the major tornadic supercell thunderstorm. For the 5 May 2007 Greensburg tornadic supercell thunderstorm case, two different assimilation configurations are introduced to examine the impact of the constraint under different situations. It is shown that assimilating wind data alone produces reasonable forecast and the use of the diagnostic pressure equation constraint evidently improve the forecast. When assimilating reflectivity data in addition to wind data, the impact of the constraint is also positive. Overall, it is demonstrated that the constraint can improve the quality of radar data assimilation and the subsequent forecast.

Chapter 1 Introduction and Overview

1.1 Background and motivation

The convective thunderstorm is one of hazardous weather that endangers people's life and properties. The implement of storm-scale numerical weather prediction (NWP) model is a great advance in improving the forecast of convective thunderstorms (Droegemeier 1997). The NWP is an initial-boundary value problem, therefore, providing an accurate initial condition for the storm-scale NWP model becomes a great concern (using appropriate boundary conditions is another concern but beyond the scope of this research).

There are generally two ways to get a better initial condition. One is to develop and improve NWP models and data assimilation techniques (such as 3DVAR, 4DVAR, and EnKF etc.) to make best use of available observations and background information. The other is to design and implement more high resolution observing systems to provide more types of observations. Currently, one component of three-dimensional wind fields can be observed by single Doppler radar. The horizontal wind fields can be retrieved from multiple NEXRAD radar velocity observations to certain accuracy if a good multiple radar coverage can be obtained. The rainfall information (rain water mixing ratio, snow water mixing ratio, hail mixing ratio) can be assumed to be derived from radar observed reflectivity (including dual-pol information), satellite imagery data and surface cloud reports. The storm-scale observing systems for other model variables are being advanced by the community. In the future, it will be expected that the vertical velocity field can be observed in high resolution by spaceborne/airborne high frequency Doppler radar. The

water vapor field can be derived in high resolution from observations by next generation GOES (Geostationary Operational Environmental Satellite), observations by dense ground-based GPS receiver network and radar refractivity observations. The temperature field can also be profiled in high resolution by next generation GOES.

As more and more data assimilation and observing system studies are devoted into this area, some general questions are raised naturally across the community: What are the impacts of different data fields for successful data assimilation and the following forecasts? How much observation information is required to get a quality initial condition? Will a more frequent assimilation (rapid update) naturally yield good results? How a storm-scale NWP model responds to the ingestion of different types of observations? Many studies (e.g. Weygandt et al. 1999; Park and Droegemeier 2000; Weygandt et al. 2002b, 2002a; Sun 2005a; Tong and Xue 2005; Hu et al. 2006a; Hu et al. 2006b; Nascimento and Droegemeier 2006; Hu and Xue 2007; Fabry 2010; Fabry and Sun 2010) have been made to examine the impact of different data fields on storm-scale NWP. However, because of the different context and different focus, there are some differences and conflicts among the conclusions of these studies (details will be provided in Chapter 2). It calls for further studies on this area. On the other hand, the model's first response to the ingestion of different types of observations at storm scale is still not thoroughly investigated. Our research will revisit the above issues. This kind of research will have practical implications for the design of the storm-scale observing systems and data assimilation experiments.

To successfully assimilate the radar data into NWP models, it is necessary to accurately simulate the radar measurements from the model data. This requires good

radar forward operator that computes model counterparts in radar observation space from related model variables on the three-dimensional grids. The radial velocity observation operator projects the three components of wind fields on model grids in a Cartesian coordinate to radial velocities in a spherical coordinate. A full radial velocity observation operator should consider the propagation path of the electromagnetic waves that can be affected by the earth curvature and the refractivity of the atmosphere, the radar beam pattern, the terminal velocity and even the signal processing algorithm used to determine the radial wind observation (Fabry 2010). This full equation is often hard to be applied in data assimilation system. Some kinds of simplifications are often introduced. For example, some early studies neglect the earth curvature and beam broadening effect. Recently, the factors of earth curvature and beam broadening are already included in many applications. However, a detailed study of the impact of beam broadening and earth curvature on storm-scale data assimilation has not yet been carried out explicitly.

Various data assimilation techniques have been developed in order to extract maximum amount of information from WSR-88D radar observations and to get a better initial condition for storm-scale NWP model. These techniques include (but not limited to) the three-dimensional variational (3DVAR) technique, the four-dimensional variational (4DVAR) technique and the Ensemble Kalman Filter (EnKF) technique. These three techniques are briefed as follows.

The 4DVAR technique uses the NWP model as a strong constraint and fits the model to observations at different time levels during a time window. By doing so, the best representation of the observations in the initial condition can be achieved and the initial condition is naturally dynamic consistent. In addition, the 4DVAR technique implicitly

includes the effects of time-evolving flow-dependent forecast error covariance. In recent years, the 4DVAR technique has enjoyed significant success at several operational NWP centers, including ECMWF, Meteo-France, Meteorological Service of Canada , and Japan Meteorological Agency (JMA), but mainly in global NWP systems (Rabier 2006). The application of the 4DVAR technique to storm-scale data assimilation has also been pioneered by some studies (Sun and Crook 1997, 1998; Sun and Crook 2001; Crook and Sun 2002; Sun 2005a). In spite of their inspiring results, the storm-scale application of 4DVAR has so far been limited to simple microphysics. The strong nonlinearity in complex microphysics schemes is difficult to handle in the minimization process of 4DVAR cost function. Honda and Koizumi (2006) reported difficulties, including slow convergence, in including complex ice microphysics within their 4DVAR system for a non-hydrostatic model at JMA. There are also other drawbacks for 4DVAR. The computational cost of 4DVAR is high, which limits its high-resolution applications. The development and maintenance of the adjoint model required by 4DVAR is not trivial. In real time applications, the 4DVAR assimilation procedure will not start unless the observations during the whole assimilation window are already available. This inevitably delays the dispatch of analysis results.

The EnKF technique is an emerging advanced data assimilation method, which promises to produce similar assimilation quality with 4DVAR but avoids the derivation of the adjoint model and appears to be less sensitive to nonlinearity. The method generates an ensemble of model forecasts with different model settings and initial disturbances. The forecast error covariance can then be estimated from the ensemble members and it naturally evolves with time. Since first proposed by Evensen (1994), the

EnKF technique is gained considerable attentions in recent years in meteorology and oceanography (e.g. Burgers et al. 1998; Houtekamer and Mitchell 1998; Hamill and Snyder 2000; Bishop et al. 2001; Whitaker and Hamill 2002). Recently, the EnKF method has been applied to the assimilation of Doppler radar data (e.g. Snyder and Zhang 2003; Zhang et al. 2004; Caya et al. 2005; Tong and Xue 2005; Aksoy et al. 2009; Zhang et al. 2009; Aksoy et al. 2010; Dowell et al. 2011). Although these experiments have shown a very good potential of the EnKF method for high-resolution applications, the EnKF technique remains a rather new, relative immature method. Its application to storm-scale NWP of thunderstorms remains a challenge. In the meantime, the EnKF method is also computationally intensive, generally requiring dozens of parallel analyses and forecasts.

The 3DVAR technique is, although theoretically sub-optimal, much more computationally efficient and easy to be implemented compared to the 4DVAR, EnKF methods. It has been reaching a considerable maturity at operational NWP centers, but mainly in the context of large-scale hydrostatic flows (Derber et al. 1991; Parrish and Derber 1992; Courtier et al. 1998; Wu et al. 2002; Rabier et al. 2006).

There are also some mesoscale applications (their model resolutions cannot resolve thunderstorms) of 3DVAR technique. Lindskog et al. (2004) reports the development of 3DVAR scheme for the High Resolution Limited Area Model (HIRLAM). Barker et al. (2004) implements the 3DVAR scheme for the MM5 system and Xiao et al. (2005) further develops the MM5 3DVAR. Zhao et al. (2006, 2008) and Xu et al. (2010) applies the so-called 3.5DVAR method (Gu et al. 2001; Xu et al. 2001b; Xu et al. 2001a) to the Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS, Hodur 1997) (The

3.5DVAR is an extension of 3DVAR, which uses two time levels of radar volume scans to retrieve one time level of wind fields). Xiao and Sun (2007) demonstrated the ability of WRF 3DVAR to assimilate multiple radar data to improve short-range quantitative precipitation forecast.

The application of 3DVAR method to storm-scale data assimilation is now an active research area. Gao et al. (1999; 2002; 2004) incorporated a 3DVAR system for the ARPS model (Xue et al. 2000; Xue et al. 2001; Xue et al. 2003) and Hu et al. (2006a; 2006b) applied it to assimilate radar data for the prediction of tornadic supercell storms. The ARPS 3DVAR system and its cloud analysis package have also been used to produce continental-US real-time weather predictions at a high 1km resolution (CAPS news, 5/2009, http://www.caps.ou.edu). However, despite its successful application, the 3DVAR scheme is often challenged by its sub-optimum due to its use of static isotropic background covariance structure and the lack of suitable balances among model variables in analysis. Efforts have been made to alleviate the negative impact of these drawbacks. Liu and Xue (2006; 2007) reports the effort to build a flow-dependent background error covariance for a 3DVAR system using an anisotropic recursive filters (Purser et al. 2003a, 2003b) and demonstrates the improvement from this method in the retrieval of moisture from GPS slant-path water vapor observations. Hamill and Snyder (2000) and Wang et al. (2008a, 2008b) illustrates another direction to provide more reasonable flow-dependent time-evolving background covariance for a 3DVAR system from an EnKF method. This immerging technique is called the hybrid data assimilation and is still at its early stage of development.

Another alternative to mitigate the imbalance among model variables in a 3DVAR analysis is to develop suitable weak constraints to help spread the disturbance from ingested observations to other model variables that is not directly linked with observations. Gao et al.(1999, 2001, 2004), Hu et al. (2006a, 2006b), and Hu and Xue (2007) incorporated the continuity equation into the cost function and found that this weak constraint can effectively help build more reasonably wind fields. This is a good progress. However, there is still no suitable balance among the dynamic and thermodynamic fields. Further research is needed to investigate this issue. We propose in this research updating the ARPS 3DVAR system with a weak constraint based on the diagnostic pressure equation, which is derived from the full ARPS model momentum equations. This weak constraint couples the dynamic field and the thermodynamic fields. It is expected that this weak constraint can help properly balance different model variables and therefore, the analysis and the subsequent forecast can be improved.

1.2 Outline of dissertation

This dissertation is organized as follows. Chapter 2 will investigate the impact of different model variables on storm-scale 3D variational data assimilation using simulated radar data and further evaluate the role of wind fields in storm-scale NWP using a real data case study. Chapter 3 will discuss the impacts of beam broadening and earth curvature on storm-scale 3D variational data assimilation of radial velocity with two Doppler radars. Chapter 4 will report the development of a weak constraint derived from ARPS model momentum equations for storm-scale 3DVAR and its applications to tornadic supercell thunderstorms observed by multiple radars. Summary and future plans will be provided in Chapter 5. All the following Chapters will be presented in a

manuscript form, which allows each chapter has its own detailed background, introduction and methodology sections.

Chapter 2 Impact of Different Model Variables on Storm-scale Three-dimensional Variational Data Assimilation

2.1 Background

The numerical weather prediction of thunderstorms is very important for the saving of people's lives, properties. To get a good prediction of thunderstorms, the initial condition where a forecast starts from is expected to be as accurate as possible. During the past 20 years, lots of research has been done in order to get a better initial condition for storm-scale numerical weather prediction. There are generally two ways to do so. One is to develop and improve data assimilation techniques (such as 3DVAR, 4DVAR, and EnKF etc.) to make best use of available observations and background information. The other is to design and implement more high resolution observing systems to provide more types of observations.

Currently, one component of three-dimensional wind fields can be observed by single Doppler radar. The horizontal wind fields can be retrieved from multiple NEXRAD radar velocity observations to certain accuracy if a good multiple radar coverage can be obtained. The rainfall information (rain water mixing ratio, snow water mixing ratio, hail mixing ratio) can be assumed to be derived from radar observed reflectivity (including dual-pol information), satellite imagery data and surface cloud reports. The storm-scale observing systems for other model variables are being advanced by the community. In the future, it will be expected that the vertical velocity field can be observed in high resolution by spaceborne/airborne high frequency Doppler radar. The water vapor field can be derived in high resolution from observations by next generation

GOES (Geostationary Operational Environmental Satellite), observations by dense ground-based GPS receiver network and radar refractivity observations. The temperature field can also be profiled in high resolution by next generation GOES.

As more and more data assimilation and observing system studies are devoted into this area, some general questions are raised naturally across the community: What are the impacts of different data fields for successful data assimilation and the following forecasts? How much observation information is required to get a quality initial condition? Will a more frequent assimilation (rapid update) naturally yield good results?

Weygandt et al. (1999) performed some experiments to study the relative importance of different data fields in a numerically simulated convective storm by withdrawing information about each model variable and then rerunning the simulation. It is found that the perturbation horizontal velocity has the greatest influence on the evolution of the simulated convective storm. Park and Droegemeier (2000) examined the sensitivities of a supercell storm to errors in model fields in the context of four-dimensional variational data assimilation. They concluded that the forecast error is most sensitive to the inaccuracy of temperature, followed by pressure and water vapor. Weygandt et al. (2002b, 2002a) conducted several sensitivity tests and found that the supercell storm simulation was greatly dependent on initial moisture fields, especially water vapor field. Sun (2005a) studied the relative importance of different initial fields on the forecast of an observed supercell storm by resetting a given initial field to its base state. The result is that wind, water vapor and temperature perturbations showed largest sensitivities. Nascimento and Droegemeier (2006) examined, using an idealized bow echo convective system, the nature of dynamic adjustment that occurred after resetting a given model data field to its

base state. They found that horizontal wind fields are crucial for the correct evolution of the simulation. Fabry and Sun (2010) and Fabry (2010) studied the propagation of initial condition errors in mesoscale convections under 4DVAR context and found that error in midlevel moisture (humidity) has the greatest impact on the quality of the forecast.

All the above researches contribute to our understanding of the relative importance of different data fields on the mesoscale/storm-scale data assimilation and prediction. However, because of their different context and different focus, there are some differences and conflicts among the conclusions of these researches. Some argued that the horizontal wind fields exert the greatest impact. Some showed that the moisture fields (especially water vapor / humidity) have the largest sensitivity. Some suggested that the error in temperature field was the most important factor that would affect the forecast. The differences among these studies call for more research on this issue. In the next section (Section 2.2), we will perform over dozen of idealized experiments to study the impact of different model variables on the accuracy of storm-scale data assimilation within a three-dimensional variational framework. In Section 2.3, we will further examine the role of wind fields in storm-scale NWP using a real data case study.

2.2 Impact of different model variables - Experiments using simulated data

2.2.1 Introduction

In this part, we will try to re-investigate the impact of different data fields thoroughly in the context of a simplified 3D variational data assimilation (3DVAR) system. Unlike the “remove” method or sensitivity method used in most previous studies (e.g.Weygandt et al. 1999; Sun 2005a; Nascimento and Droegemeier 2006), we will try to examine the impact of different data fields in a direct way by assimilating them into the model. We

will perform a series of OSS data assimilation experiments, which assimilate different combinations of observed data fields, and then check whether each of them can successfully reproduces the storm structures (both dynamic and thermodynamic structures) and how long it will take to get such a successful assimilation. In the meantime, the impact of assimilation frequency will also be examined.

This part is organized as follows. Section 2.2.2 will cover the methodology and the design of experiments, Section 2.2.3 will discuss the results from these experiments. Summary and future plan will be provided in Section 2.2.4.

2.2.2 Methodology and experimental design

2.2.2.1 The 3D variational formulation

The standard formulation of variational methods is derived from first principles by Lorenc using Bayesian probabilities and assuming Gaussian error distributions (Lorenc 1986).The concept of a variational method is to determine the analysis by direct minimization of a cost function. The cost function may be written as:

$$J(x) = J_B + J_O = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}[H(x) - y_o]^T R^{-1}[H(x) - y_o] \quad (2.1)$$

where J_B measures the departure of the analysis x from the background, and is weighted by the inverse of the background error covariance matrix B; J_O measures the departure $H(x)$, which is the projection of the analysis x in observational space, from the observations y_o and is weighted by the inverse of the observational error covariance matrix R. B and R are uncorrelated, and both are symmetric and positive definite.

In this study, the observations are directly drawn from the model variables, so no projection or interpolation is needed. Hence, we get a simplest form of the cost function:

$$J(x) = J_B + J_O = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(x - x_o)^T R^{-1}(x - x_o) \qquad (2.2)$$

The goal of an analysis is to find state x_a, for which J is minimized. At the minimum, the derivative of J vanishes, and x_a satisfies $\nabla J(x) = B^{-1}(x - x_b) + R^{-1}(x - x_o) = 0$. The construction of background error covariance B^{-1} is similar to Gao et al. (1999). The standard deviation for each model variable will be provided in next section.

In Eq. (2.1), the penalty term, usually represented by J_c, which can be used to build linkages among model variables by using some kind of equation constraints (e.g. mass continuity equation, buoyancy compensation equation, diagnostic pressure equation, etc) is not included. The reason is that although the equation constraints can help spread the observation information to some unobserved model variables; it complicates the data impact problem here. This study will particularly focus on the data impact of individual model fields, especially the first response of the model to the ingestion of data observations. The impact of equation constraints will be examined in Chapter 4.

2.2.2.2 The prediction model and truth simulation

In this study, we use simulated data from a classic May 20, 1977 Del City, Oklahoma supercell storm case (Ray et al. 1981). The Advanced Regional Prediction System (ARPS) is used to simulate such a deep convective storm within a 64 x 64 x 16 km physical domain. The model grid comprises of 67 x 67 x 35 grid points. Horizontal resolution of 1km and vertical resolution of 0.5km are used. The truth simulation is initialized from a

modified real sounding plus a +4K ellipsoidal thermal bubble centered at x=48, y=16 and z=1.5km, with radii of 10km in x and y directions while 1.5km in z direction. The warm rain microphysical scheme is used together with a 1.5-order turbulent kinetic energy subgrid parameterization. Open conditions are used at the lateral boundaries. A wave radiation condition is also applied at the top boundary. Free-slip conditions are applied to the bottom boundary. The length of simulation is up to three hours. A constant wind of $u=3ms^{-1}$and $v= 14ms^{-1}$ is subtracted from the observed sounding to keep the primary storm cell near the center of model grid. The evolution of the simulated storms is similar to those documented in Xue et al. (2001).

Fig. 2.1 shows the horizontal winds, perturbation potential temperature and reflectivity pattern from the control run at z=250m every twenty minutes from t=50 minutes into the simulation run to t=110 minutes. During the control run, the supercell strengthens over the first 20 minutes. The strength of the cell then decreases thereafter. At around 55 minutes, the cell splits into two. The north-northeastward moving cell tends to dominate the system. Another cell moves northwestward and splits again at 95 minutes.

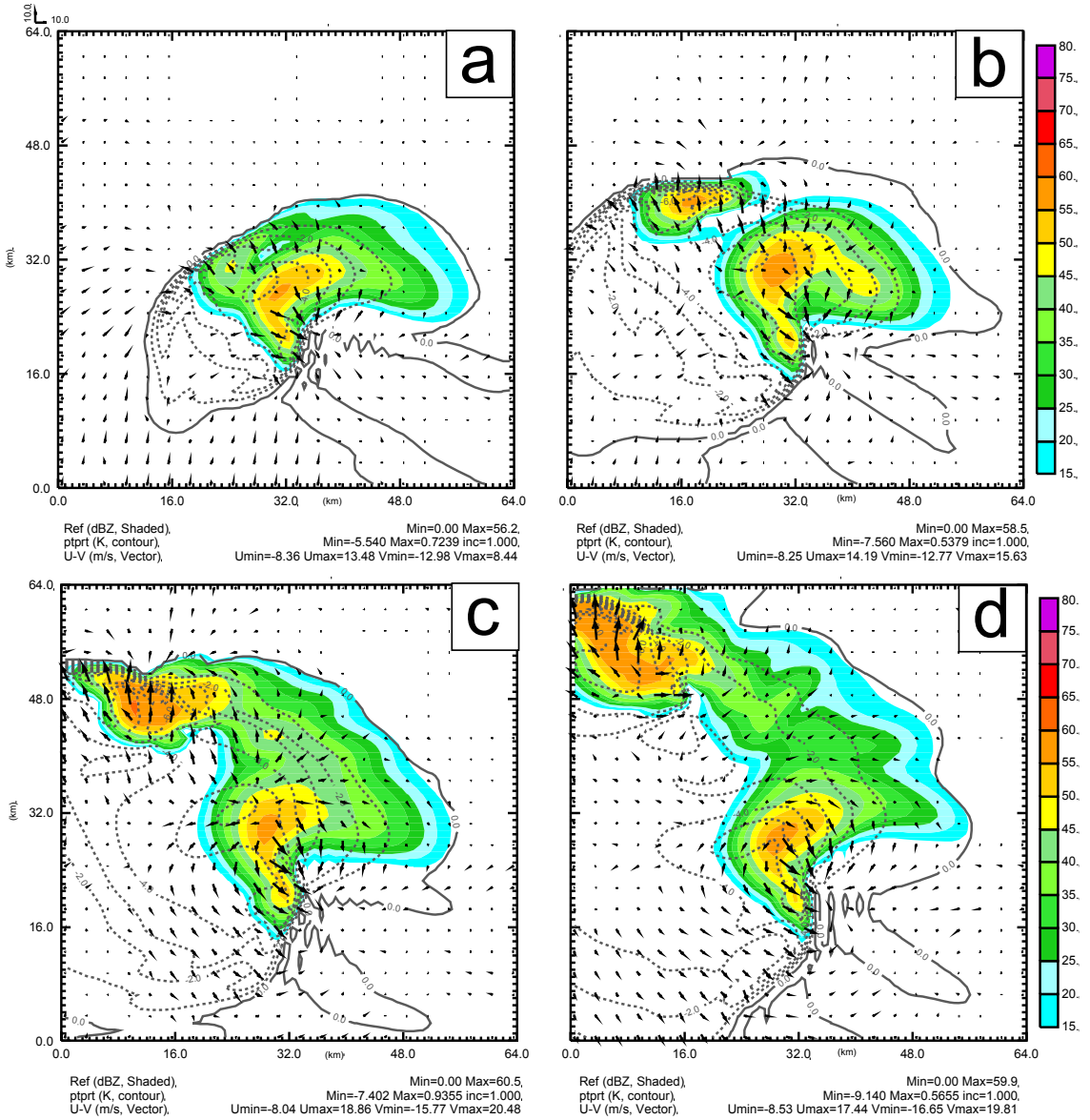

Fig. 2.1. The perturbation horizontal winds, perturbation potential temperature and reflectivity pattern from the control run at z=250m every twenty minutes from t=50 minutes into the simulation run to t=110 minutes. (a) t=50min, (b) t=70min, (c) t=90min, (d) t=110min

2.2.2.3 Experimental design

After creating the truth (control) simulation of the tornadic thunderstorm, pseudo observations are generated by directly taken from the evolution of corresponding model variables. These pseudo observations are:

- Perturbation horizontal wind in x direction (u’)
- Perturbation horizontal wind in y direction (v’)
- Perturbation Vertical velocity (w)
- Perturbation potential temperature (θ')
- Perturbation water vapor mixing ration (q_v')
- Rain water mixing ratio (q_r)

The pseudo observations are assumed nearly perfect, exist at every model grid point, and thus the assimilation experiments are performed using the simple 3D variational method as mentioned in Section 2.2.2.1. Gaussian noises are then added to the above data fields to model the observation error. The standard deviation for each data field is list in Table 2.1 (σ_o). In the meantime, Table 2.1 (σ_b) also lists the standard deviation of model error for each corresponding model variable.

Table 2.1. Standard deviation of observation error (σ_o) and background error (σ_b)

	σ_o	σ_b
u’	1m/s	3m/s
v’	1m/s	3m/s
w	0.667m/s	2m/s
θ'	0.667K	2K
q_r	0.1g/kg	0.3g/kg
q_v'	0.25g/kg	0.75g/kg

The pseudo observations are taken from t=30 minutes into the control run to t=120 minutes. The assimilation experiments start with a horizontally homogeneous background whose vertical variation is given by the same sounding as in the initial condition for the simulation run. The available pseudo observations are then assimilated into the model. The model runs for a given time span according to the selected data assimilation frequency. After the given time span elapsed, the observations are assimilated into the model again. This process repeats until the assimilation runs for ninety minutes.

With the consideration of testing the impacts of different data fields under the context that these data fields can be observed by potential observing systems mentioned before, we assimilated sixteen combinations of different data fields respectively. These combinations are list in Table 2.2. Each combination is named by the data fields used in the assimilation. For example, "UVQv" represents assimilating perturbation horizontal winds in x direction u', perturbation horizontal winds in y direction v' and perturbation water vapor qv into the model at the same time. Note that for simplicity, the ' sign is omitted, the first letter is capitalized, θ' is replaced by "Pt". For every combination of observations, three assimilation experiments are preformed which are designed to use three different assimilation frequencies, i.e. every 1 minute, every 5 minutes , every 10 minutes. Hence, it is a natural choice to refer an individual data assimilation experiment as the combination of observations it used plus an underscore "_" followed by the assimilation frequency. As an example, "UVQv_5" refers to the experiment that assimilate u', v' and qv' every 5 minutes.

Table 2.2. The list of experiments and their corresponding "Successful-recovery time" (see text for the definition "Successful-recovery time")

Assimilation frequencies (per mins) / Observation combinations	1	5	10
U			
W			
Pt			
Qv			
Qr			
UV	30	68	
UPt			
UQv		51	
UQr			
UVPt	41	32	
UVQv	19	23	31
UVQr	28	48	81
UVW	26	38	71
UVWPt	30	26	42
UVWQv	18	20	22
UVWQr	24	27	61

At the beginning of the assimilation experiments, there is no storm signature in the model. As the observations are assimilated into the model, the storm structures are gradually recovered. It is expected that the data assimilation could finally recover both the dynamic and thermodynamic fields of the thunderstorm and these fields should be very close to that of the control run. Since these are much idealized experiments, we set a strict criterion to evaluate whether an experiment does successfully recover the storm structures. The criterion goes as follows: First we compute the RMS error of the three components of wind fields (u', v', w'), the perturbation potential temperature(θ'), the simulated reflectivity(ref) and the perturbation water vapor mixing ratio (qv') between the assimilation run and the control run every one minute. As the assimilation cycles forward, the RMS error is expected to decrease. When the RMS error of the three components of wind fields (u', v', w') is below 2.0m/s, the RMS error of the perturbation potential temperature(θ') is below 1.0K, the RMS error of the simulated reflectivity(ref) is below 10dBz, and the RMS error of the perturbation water vapor mixing ratio (qv') is below 0.3g/kg, the storm can be described as successfully recovered in this assimilation run. A "successful-recovery" is therefore defined as the status when the storm structures are successfully recovered under the above criterion. A "successful-recovery time" is defined as a time duration that an assimilation run takes to reach a "successful-recovery".

It should be noted that when calculating the RMS error, only grids which are located in the cloudy region (here we refer to the region where simulated reflectivity $\geqslant$10dBz) are taken into account.

2.2.3 The Results of assimilation experiments

As mentioned in Section 2.2.2.3, there are sixteen different combinations of observations; each is assimilated with three different frequencies (every one minute, five minute, ten minutes). Hence, there are a total of forty-eight data assimilation experiments. For every experiment, the RMS error evolution is computed and then is used to decide the so-called "successful-recovery time" as defined in Section 2.2.2.3. Table 2.2 lists all experiments and their corresponding "successful-recovery time" in unit of minutes. A blank will be left for those experiments that do not reach a "successful-recovery" at the end of the assimilation run.

2.2.3.1 Experiments with observations from only one model variable

The experiments assimilating observations from single model variable will be investigated here to examine the model's first response to the ingestion of different observations and to what extent the storm dynamic and thermodynamic structure can be recovered during the assimilation runs. For simplicity, the following discussion will be based on experiments that assimilate observations every five minutes.

2.2.3.1.1 Assimilating U component of wind fields

Fig. 2.2 shows the evolution of U component of perturbation wind fields, vertical velocity, perturbation potential temperature and water vapor mixing ratio from the "U_5" experiment during the first 5-minutes assimilation cycle. It can be seen that at the first data assimilation cycle, the U component of perturbation wind fields is updated by the analysis step. The model is then integrated forward and the convergence/divergence in the U component of wind fields leads to the adjustments in the vertical velocity field. The

established vertical air movement then perturbs the potential temperature field and the water vapor field, which are horizontally homogeneous before. As the model integrates further forward, the magnitude of U component of perturbation wind fields and vertical velocity field decreases while those of water vapor field and perturbation potential temperature field increases.

This kind of data impact will be reinforced as new observations are ingested into the model at the subsequent data assimilation cycles. The adjustments in dynamic fields and thermodynamic fields will eventually induce rainfall at some time. Fig. 2.3 shows the perturbation horizontal winds, perturbation potential temperature and reflectivity pattern at z=250m MSL every twenty minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). It is clear that after four data assimilation cycles (20 minutes into the assimilation run), a small area of weak rainfall has been produced (see Fig. 2.3a). As the assimilation run goes on, the rainfall becomes stronger and spreads to wider area. After 80 minutes of assimilation run, the recovered storm at the center of domain can be even comparable to that in control run. However, considerable discrepancies in the perturbation potential temperature field and the perturbation wind fields still exist. The left-moving storm near the up-left corner of the domain is still barely recovered.

To investigate further the assimilation results, Fig. 2.4 shows the evolution of RMS error for the “U_5” experiment. It can be seen that although the RMS error of simulated reflectivity decreases steadily with time, the RMS errors of other model variables do not show evident decrease during ninety minutes of data assimilation. Bearing in mind that this is a much-idealized experiment and under our definition of "successful-recovery",

this experiment is regarded that it fails to successfully recover the simulated thunderstorms.

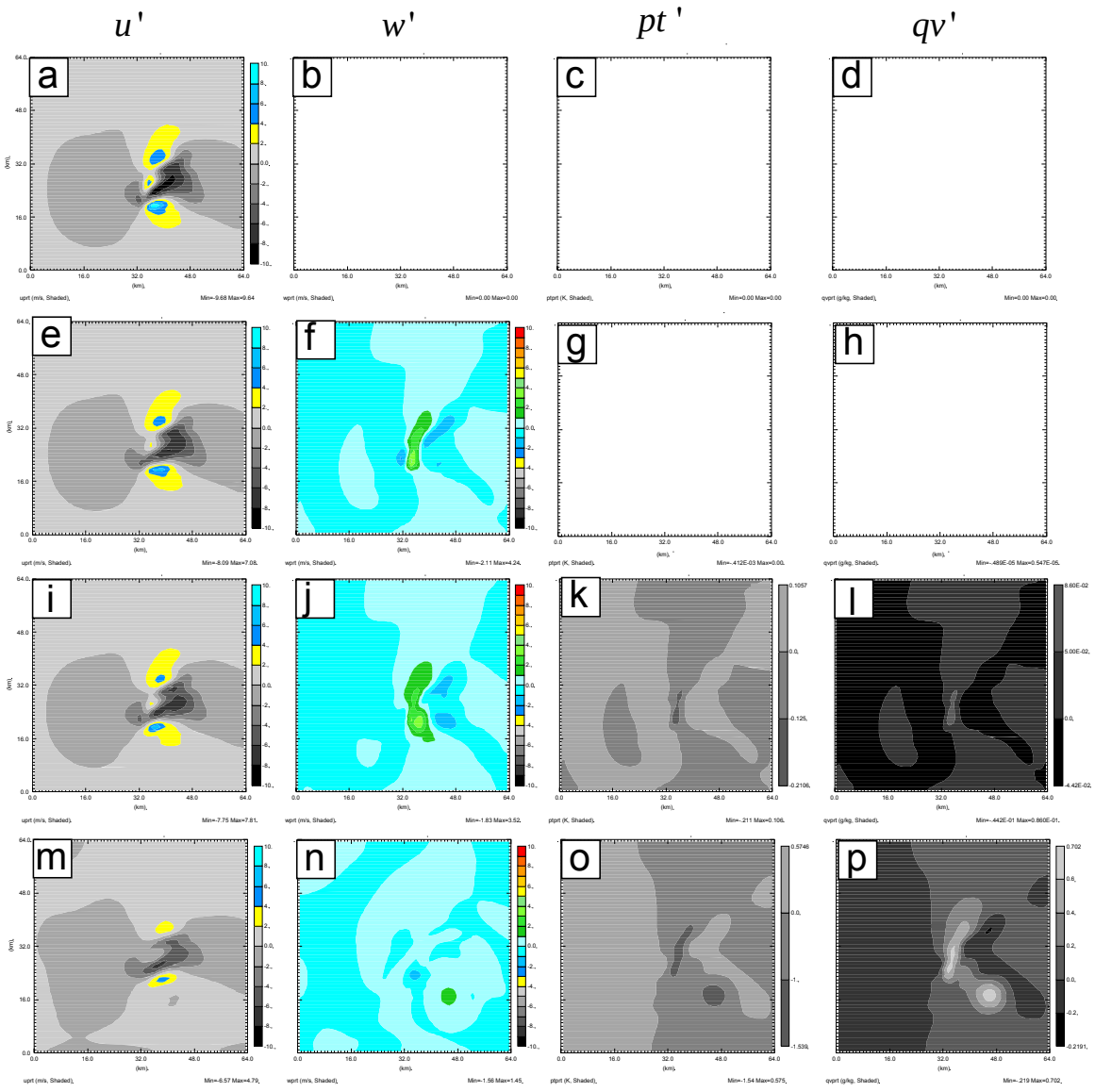

Fig. 2.2. The U component of perturbation wind fields, perturbation vertical velocity, perturbation potential temperature, water vapor mixing ratio for the "U_5" experiment at z=4km MSL. (a),(b),(c),(d) are at t=0s into the assimilation run; (e),(f),(g),(h) are at t=6s into the assimilation run, (i),(j),(k),(l) are at t=12s into the assimilation run; (m),(n),(o),(p) are at t=300s into the assimilation run.

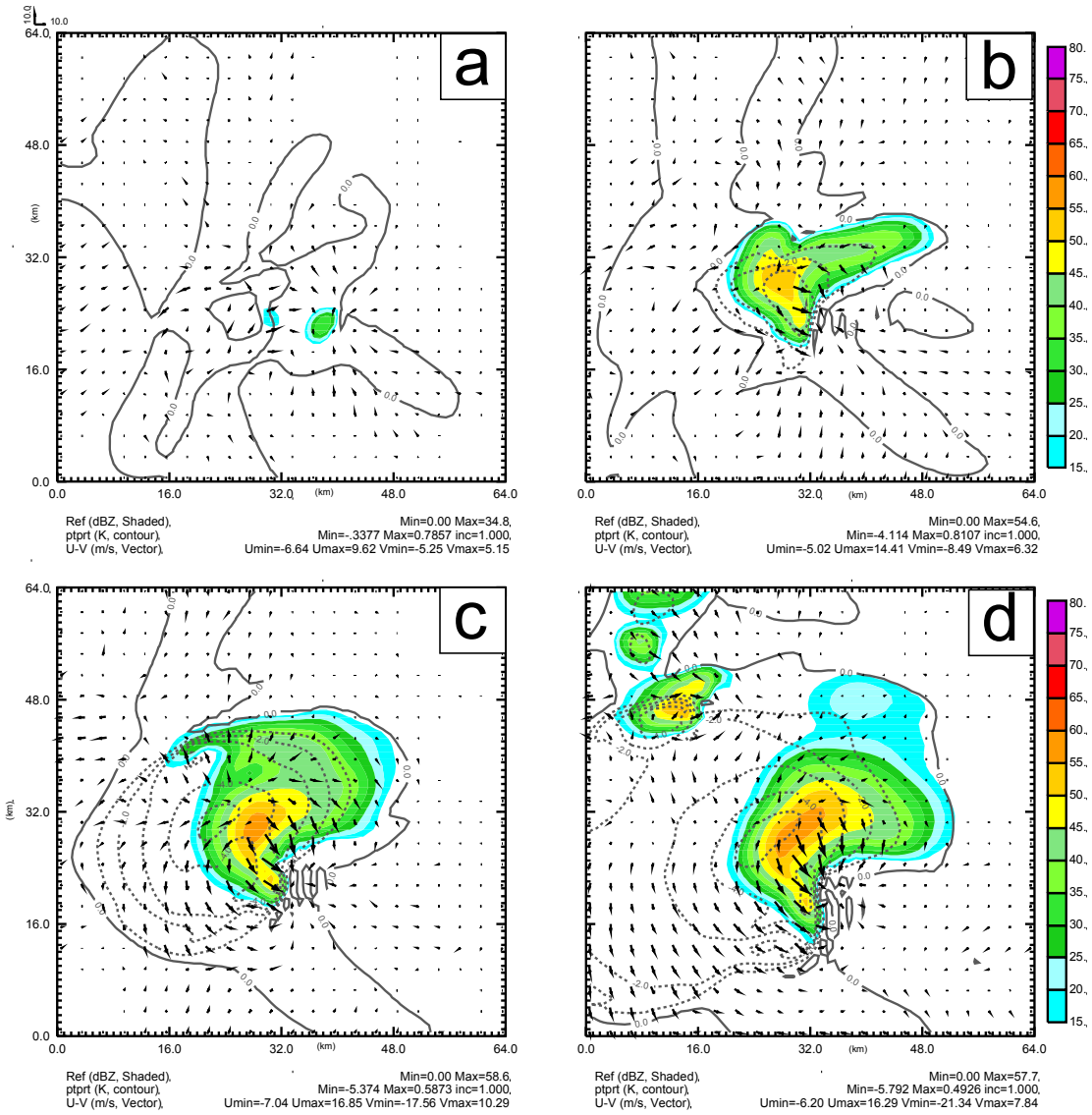

Fig. 2.3. The perturbation horizontal winds, perturbation potential temperature and reflectivity pattern for the "U_5" experiment at z=250m every 20 minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). (a) t=20min, (b) t=40min, (c) t=60m, (d)t=80m.

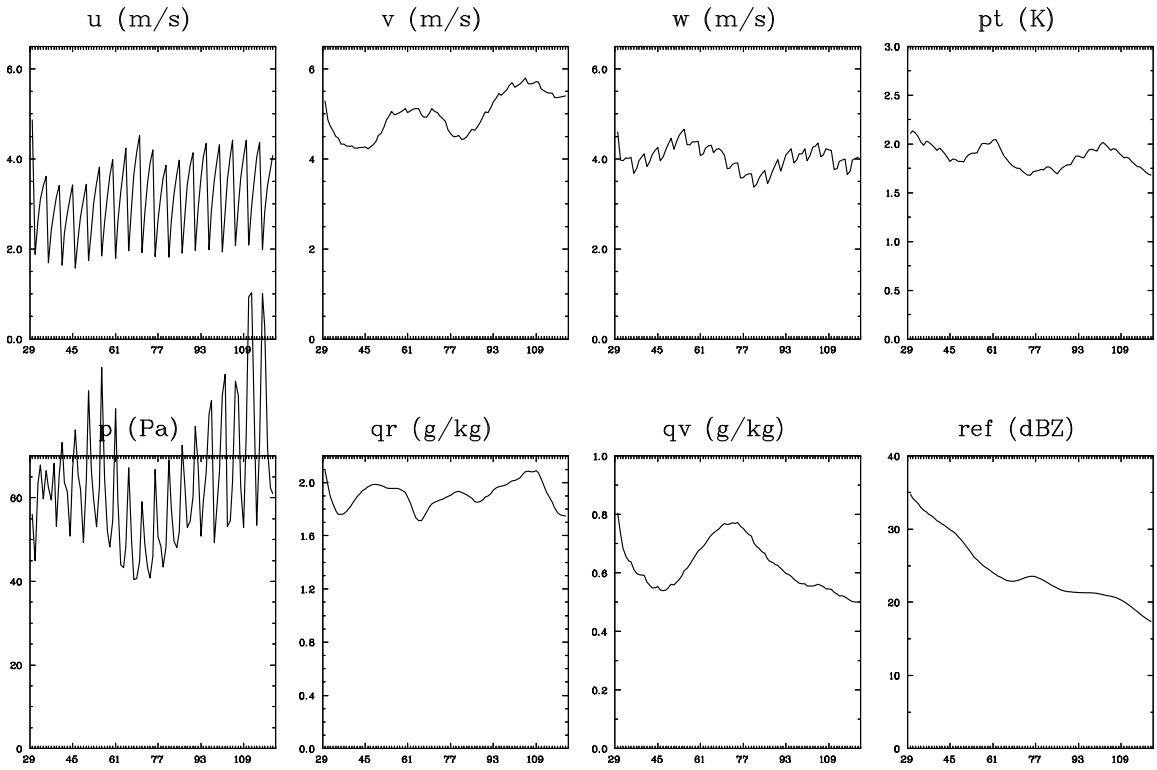

Fig. 2.4. The RMS error evolution every one minute for the "U_5" experiment

2.2.3.1.2 Assimilating perturbation vertical velocity

When the perturbation vertical velocity is assimilated into the model at the first cycle, other model fields will be perturbed by the upward or downward advection. Warm moist air in the low level is then brought upward. When it gets saturated, condensation begins. Model dynamic and thermodynamic fields will then adjust accordingly. The data impact is reinforced through the subsequent intermittent data assimilation cycles. Fig. 2.5 shows the perturbation horizontal winds, perturbation potential temperature and reflectivity pattern at z=250m MSL every twenty minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). It can be seen that in terms of rainfall pattern, the storm cell near the center of the domain is essentially recovered at nearly the end of the

assimilation run (Fig. 2.5d). The storm cell at the upper-left corner is also partially rebuilt. Overall, the "W_5" experiment recovers a better rainfall pattern than the "U_5" experiment.

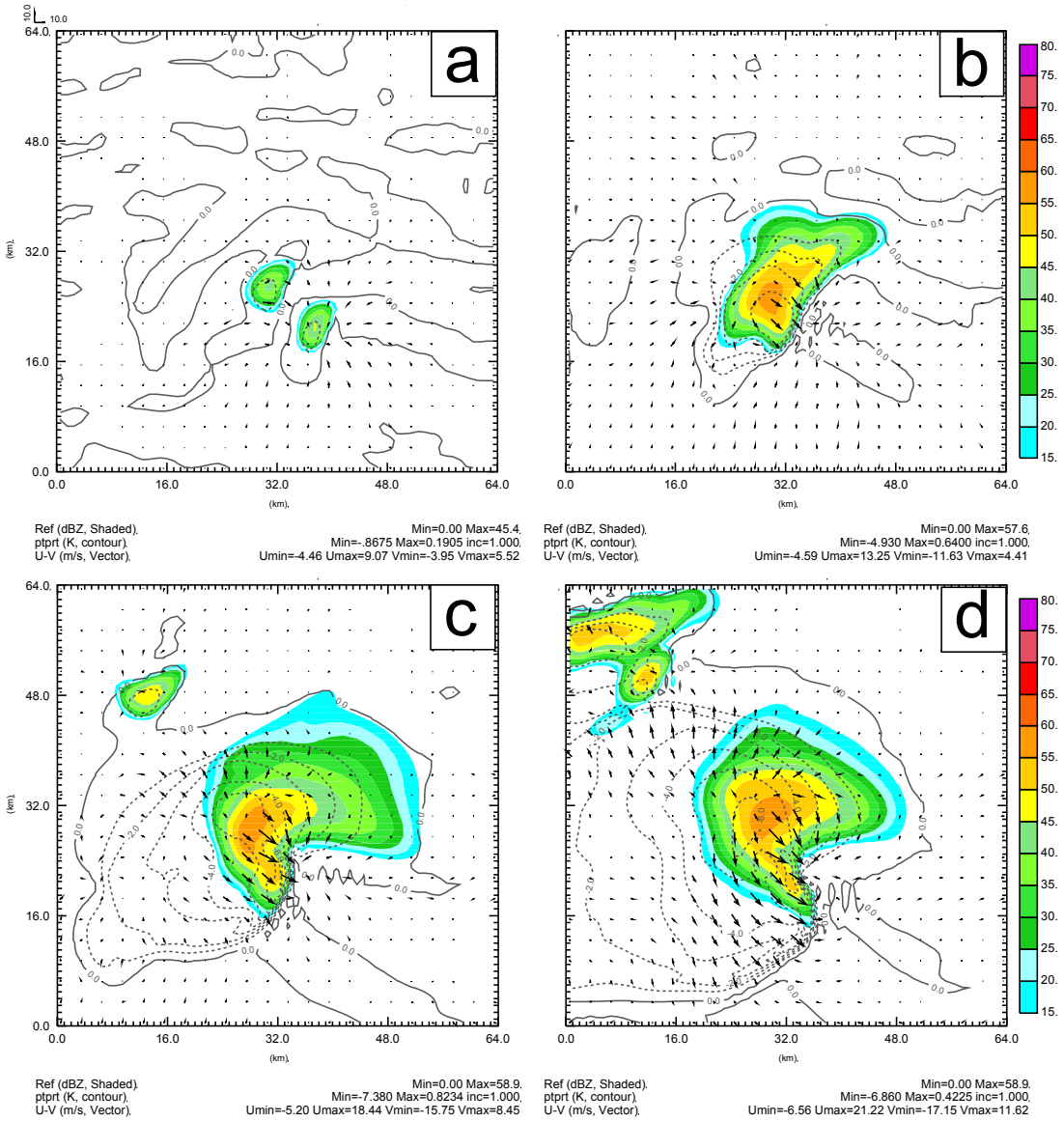

Fig. 2.5. The perturbation horizontal winds, perturbation potential temperature and reflectivity pattern for the "W_5" experiment at z=250m every 20 minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). (a) t=20min, (b) t=40min, (c) t=60m, (d)t=80m.

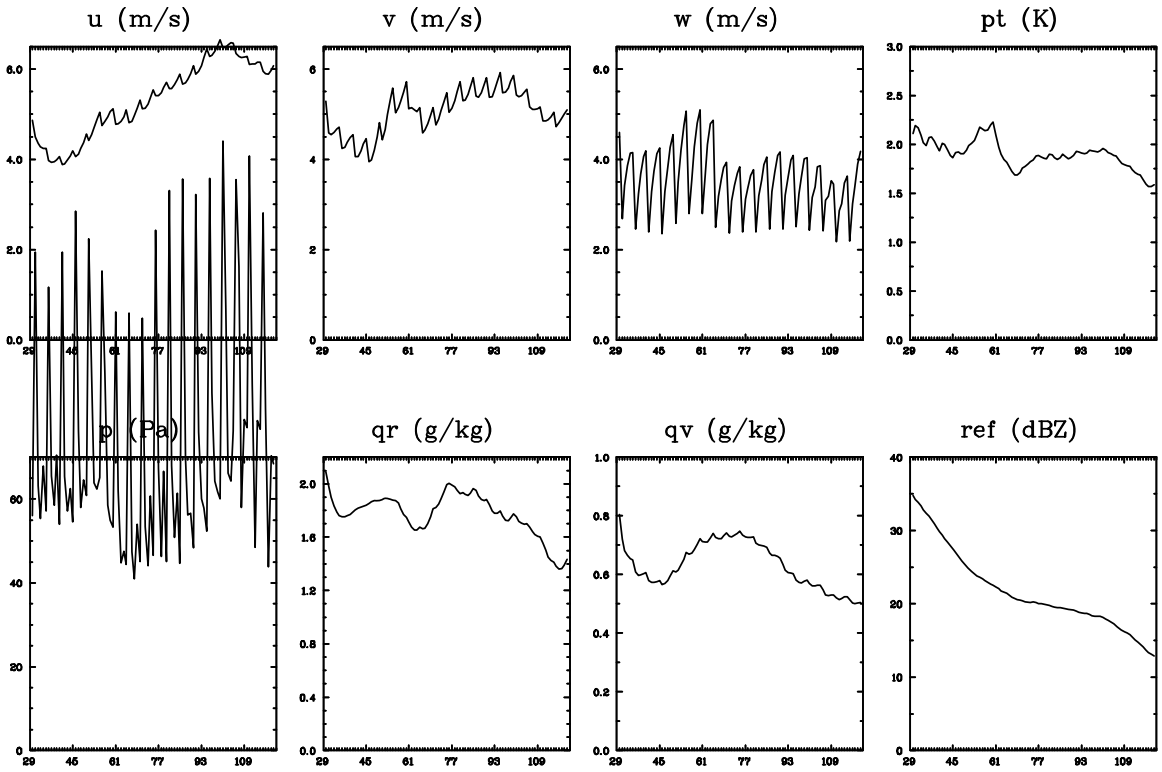

Fig. 2.6. The RMS error evolution every one minute for the “W_5” experiment

On the other hand, noticeable discrepancies still exist in the horizontal wind fields, perturbation potential temperature fields. Further, the rainfall pattern is not close enough to the truth either. The evolution of RMS error (Fig. 2.6) confirms this conclusion. Under our definition of "successful-recovery", this experiment is also regarded that it fails to successfully recover the simulated thunderstorms.

2.2.3.1.3 Assimilating perturbation potential temperature

When the perturbation potential temperature is assimilated into the model, the direct impact on the model is the change of air buoyancy, which in turn to promote vertical air motion. The horizontal wind field and water vapor field then change accordingly. The rainfall is produced gradually with continued data assimilation cycles.

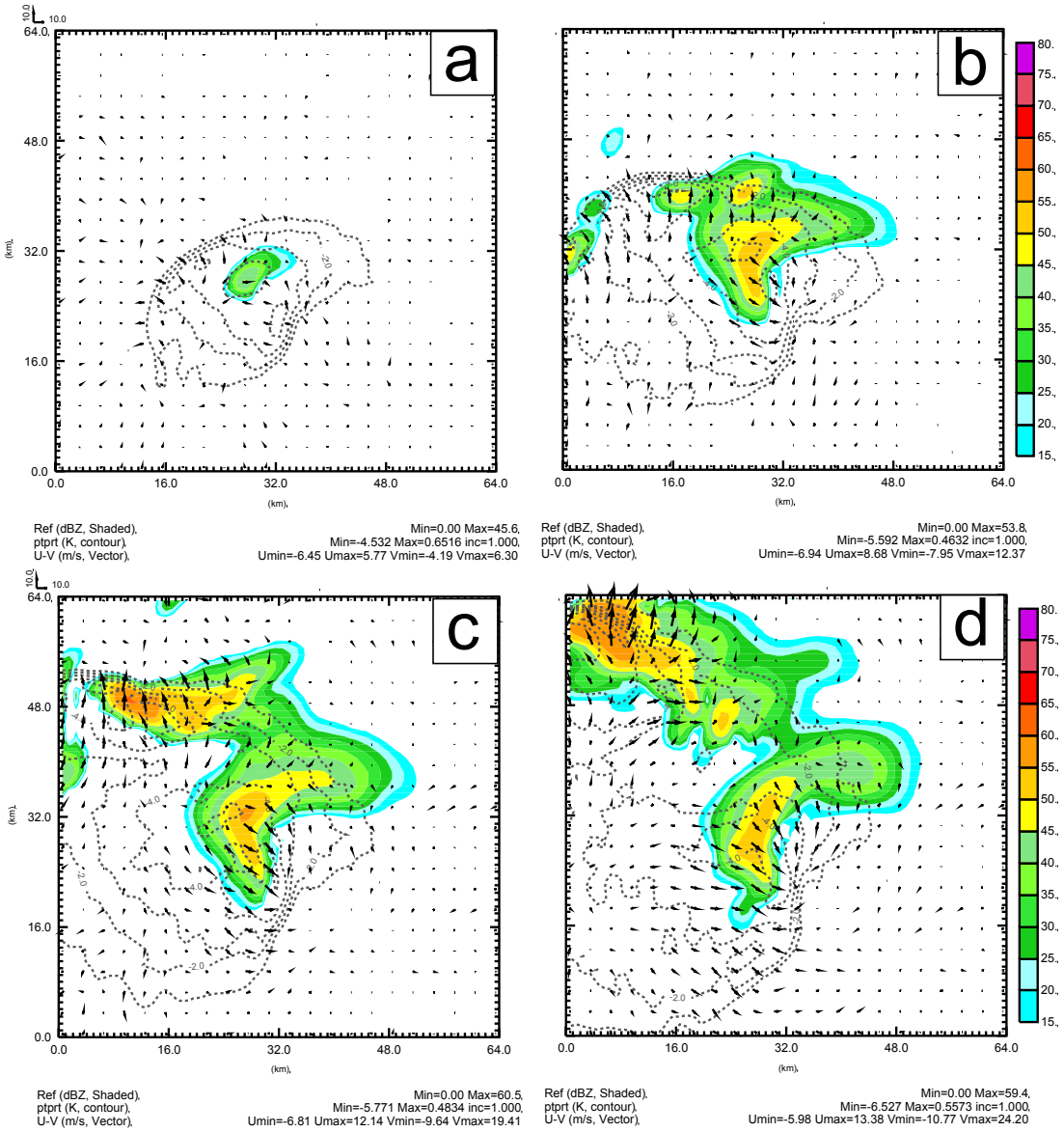

Fig. 2.7. The perturbation horizontal winds, perturbation potential temperature and reflectivity pattern for the "Pt_5" experiment at z=250m every 20 minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). (a) t=20min, (b) t=40min, (c) t=60m, (d)t=80m.

Fig. 2.7 shows the perturbation horizontal winds, perturbation potential temperature and reflectivity pattern at z=250m MSL every twenty minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). It is shown that near the end of the assimilation

run, the recovered rainfall pattern (Fig. 2.7d) is comparable to those in the truth run (Fig. 2.1d). Two storm cells are located at the correct location and with similar strength. However, the area with reflectivity value in-between 15dBZ-25dBZ is still evidently different from those in the truth run.

Fig. 2.8 presents the evolution of RMS error from the “Pt_5” experiment. It clearly shows that the assimilation of temperature observations have greater impact on the recovery of rainfall pattern than other model structures. The vertical velocity field is only partially rebuilt. The horizontal wind field and the water vapor field are poorly retrieved. On the whole, under our definition of "successful-recovery", this experiment is regarded that it fails to successfully recover the simulated thunderstorms.

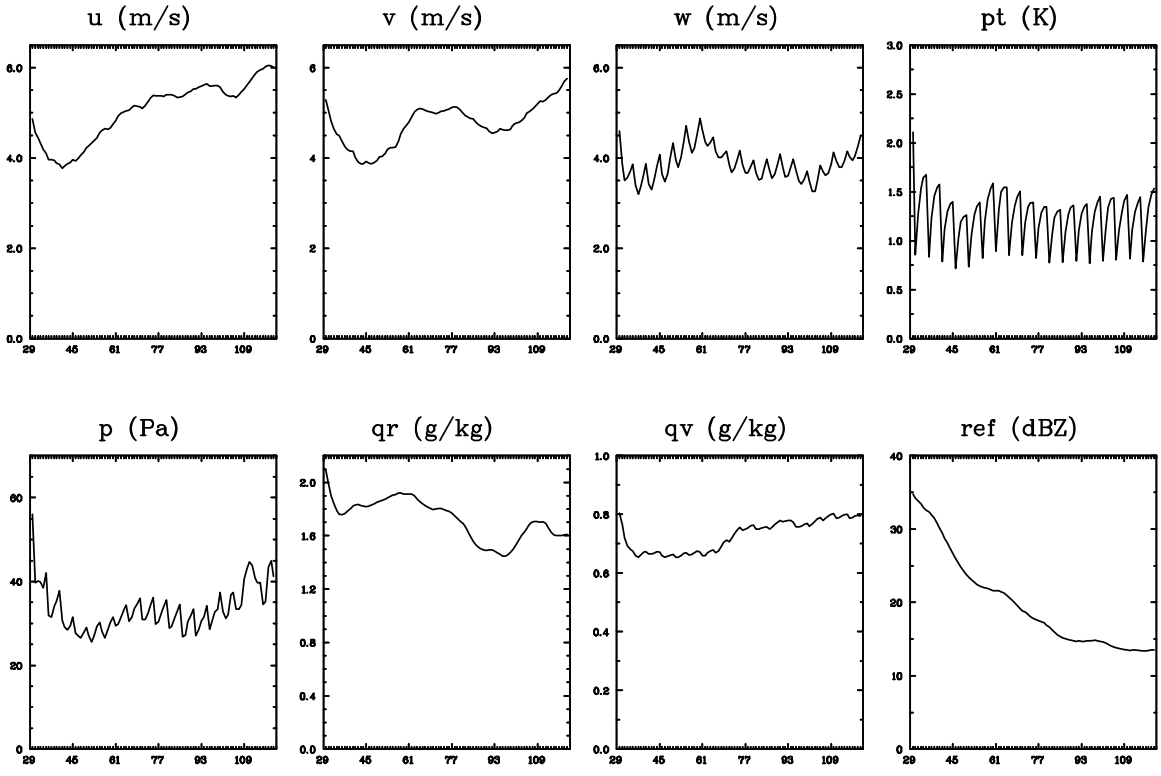

Fig. 2.8. The RMS error evolution every one minute for the “Pt_5” experiment

2.2.3.1.4 Assimilating perturbation water vapor

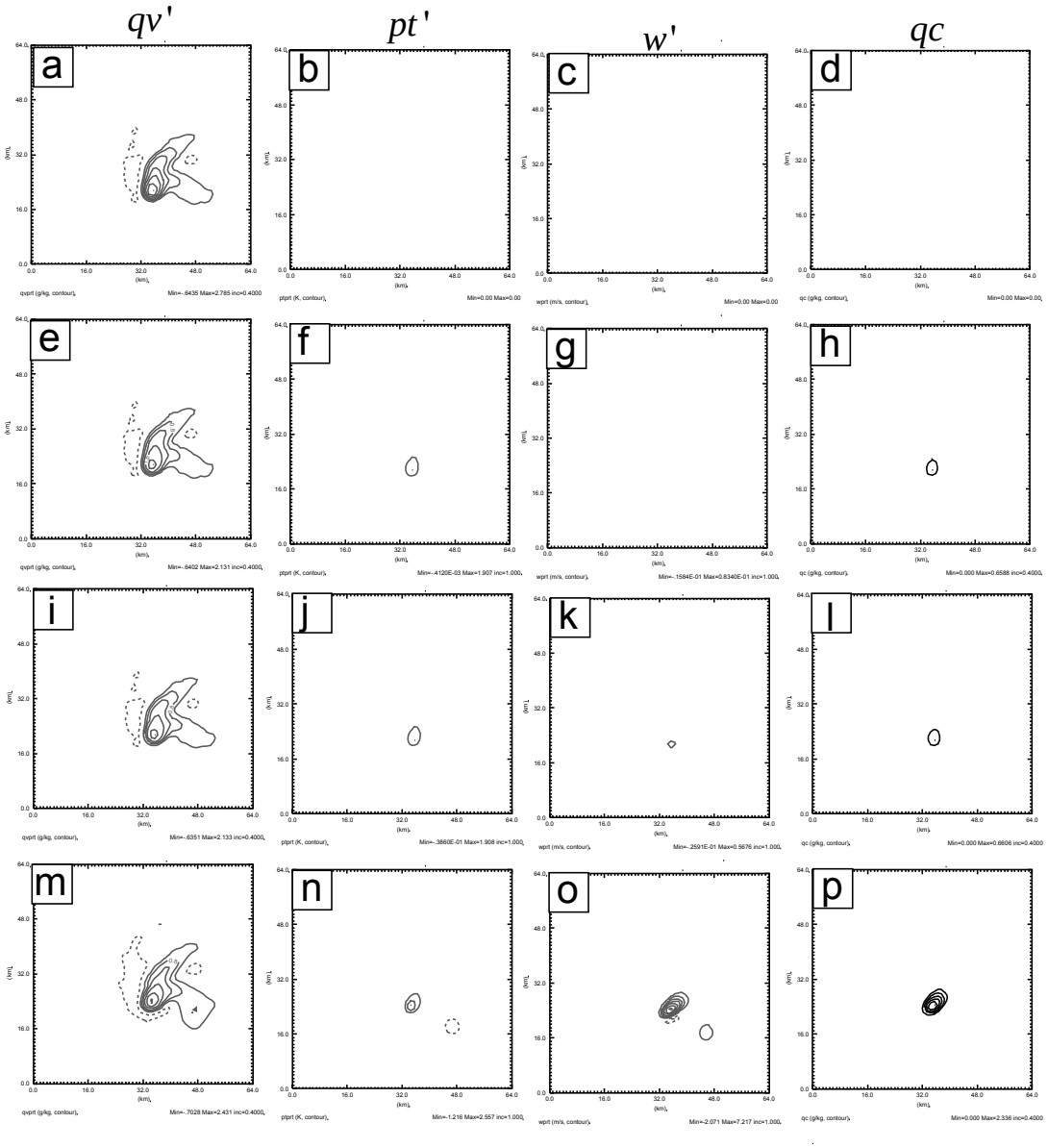

Fig. 2.9. The perturbation water vapor mixing ratio, perturbation potential temperature, perturbation vertical velocity, cloud water mixing ratio for the "Qv_5" experiment at z=4km MSL. (a),(b),(c),(d) are at t=0s into the assimilation run; (e),(f),(g),(h) are at t=6s into the assimilation run, (i),(j),(k),(l) are at t=12s into the assimilation run; (m),(n),(o),(p) are at t=300s into the assimilation run.

When the perturbation water vapor observations are assimilated into the model at the first data assimilation cycle, the major response from the model is to produce cloud water

through condensation (Fig. 2.9h) and heat the air through latent heating released from the condensation (Fig. 2.9f). The change of buoyancy due to the contribution of perturbation water vapor is another response from the model; however, the impact is very small. This can be confirmed from Fig. 2.9g, which shows that the vertical velocity barely changes initially. Detailed scale analysis (not shown here) for buoyancy terms indicates that the contribution of perturbation water vapor to the buoyancy is at the order of 0.01 m s^{-2} near storm center while the contribution of perturbation potential temperature is at the order of 0.1m s^{-2}. Therefore, the buoyancy response from the assimilation of water vapor observations is rather small compared to the impact of condensation and latent heating, at least in this case.

The recovery of rainfall pattern goes well in this experiment as shown in Fig. 2.10. As early as at t=40 minutes (Fig. 2.10b) into the assimilation run, the rainfall pattern has already been comparable to those in the truth run. At nearly the end of the assimilation run (Fig. 2.10d), the rainfall pattern is closer to the truth. The final RMS error for simulated reflectivity at t=90 minutes into the assimilation run is 7.35dBZ, already below our "successful-recovery" criterion 10dBZ. The recovery of temperature field also goes okay. Fig. 2.10d shows that the cold pool is re-established around both the two storm cells. The location and coverage are very acceptable as compared to those in the truth simulation, although noticeable differences still exist in term of the distribution of temperature field near the two storm centers. The final RMS error for perturbation potential temperature is 1.2 K, which is very close to our "successful-recovery" criterion 1K. However, the wind fields, especially horizontal wind fields are poorly recovered. This can be seen from Fig. 2.11, which presents the evolution of RMS errors for the

"Qv_5" experiment. The decreases of RMS errors for U, V components of wind fields are very limited. Therefore, on the whole, under our definition of "successful-recovery", this experiment is still considered that it fails to successfully recover the simulated thunderstorms.

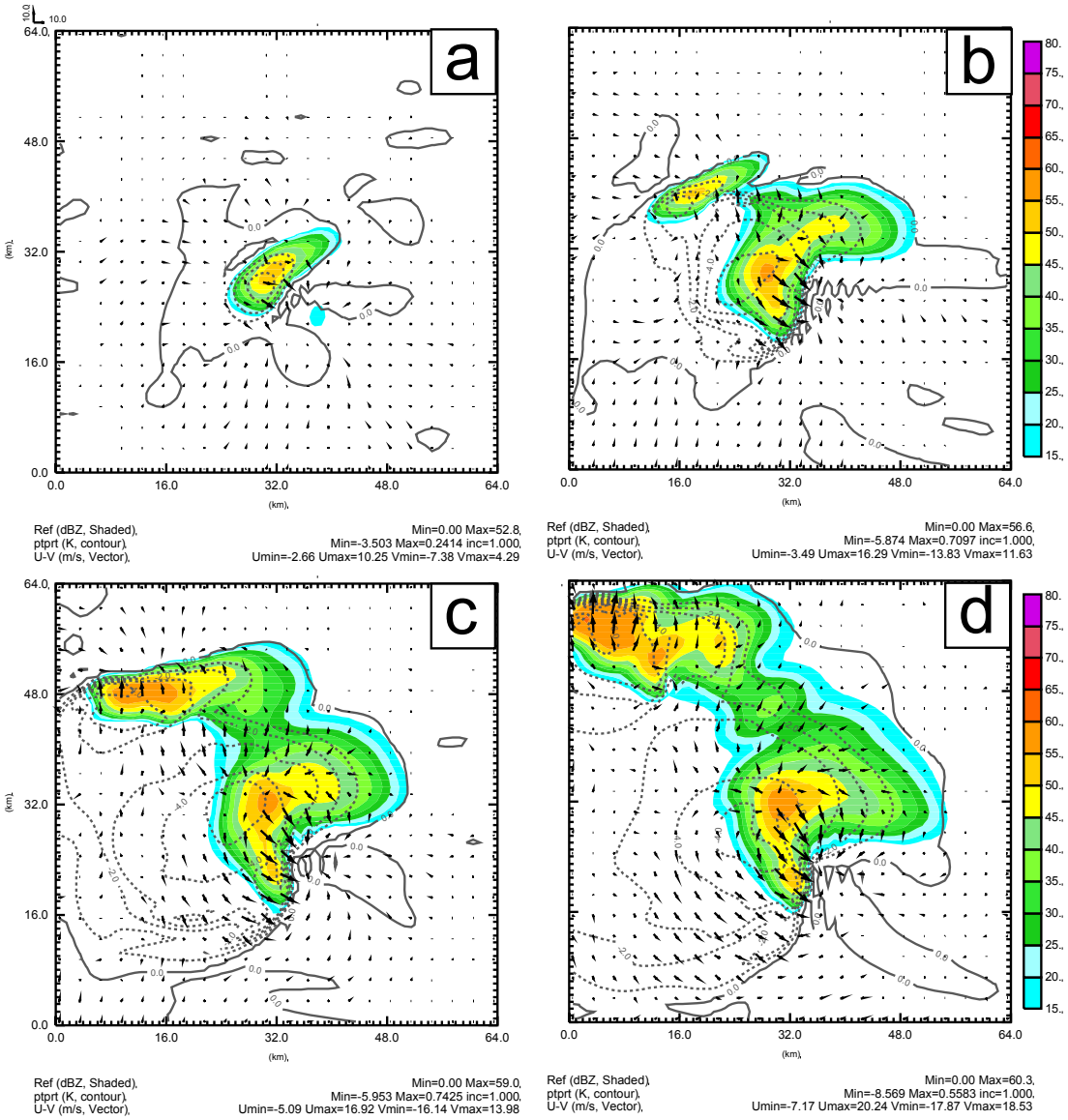

Fig. 2.10. The perturbation horizontal winds, perturbation potential temperature and reflectivity pattern for the "Qv_5" experiment at z=250m every 20 minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). (a) t=20min, (b) t=40min, (c) t=60m, (d)t=80m.

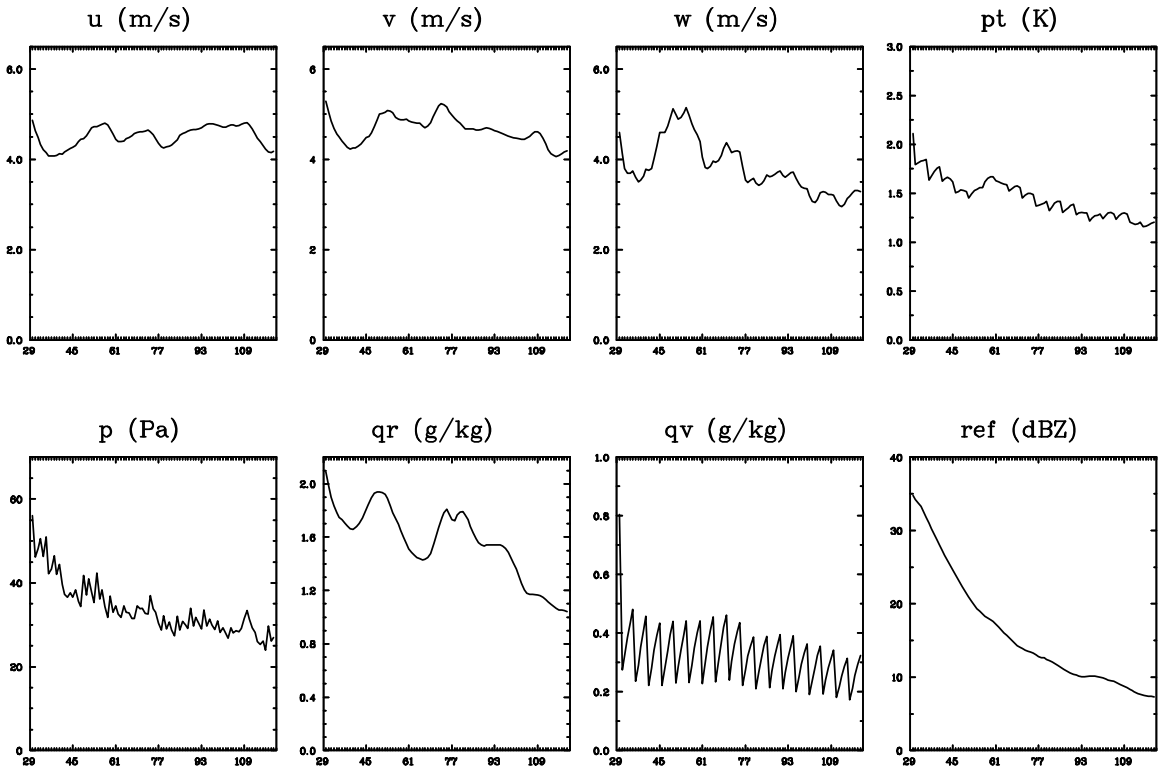

Fig. 2.11. The RMS error evolution every one minute for the "Qv_5" experiment

2.2.3.1.5 Assimilating rain water

When the rainwater pseudo observations are assimilated into the model, its major impact is to change the buoyancy through water loading effect and therefore to produce downward vertical motion (Fig. 2.12e, h). The evaporative cooling (Fig. 2.12f, i) is another noticeable impact. Wind fields and water vapor fields will then adjust accordingly.

With continued intermittent data assimilation, the cold pool is rebuilt well (Fig. 2.13) although there still exists noticeable difference in the strength and distribution. On the other hand, the mid-upper level temperature field is not-so-well retrieved (not shown). The wind fields and water vapor field are barely recovered as shown in Fig. 2.14, which

presents the evolution of RMS error for the "Qr_5" experiment. This means that this experiment fails to successfully recover the simulated thunderstorms.

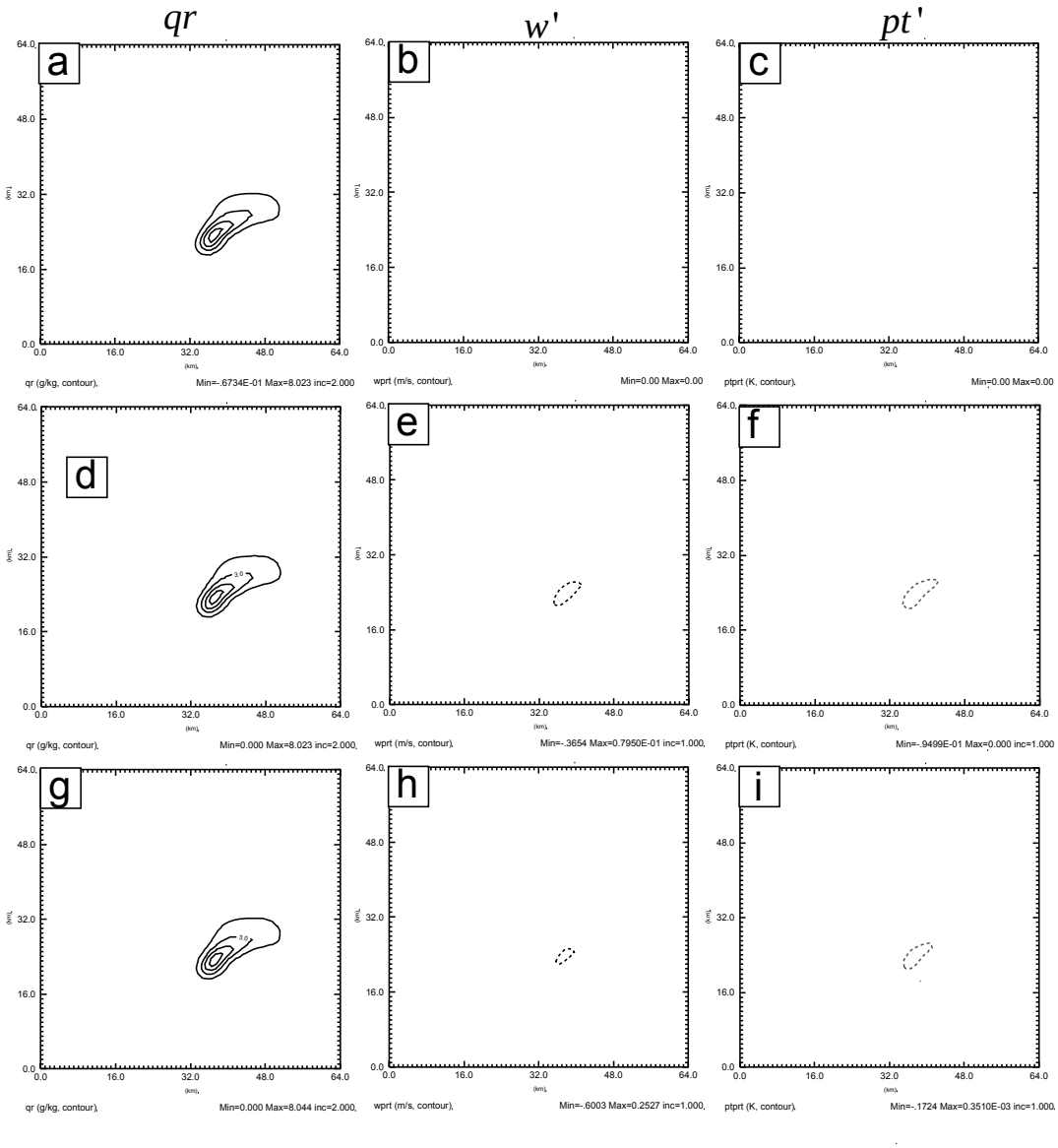

Fig. 2.12. The rainwater mixing ratio, perturbation vertical velocity, perturbation potential temperature from the "Qr_5" experiment, at z=4km MSL. (a),(b),(c) are at t=0s into the assimilation run; (d),(e),(f) are at t=6s into the assimilation run, (g),(h),(i) are at t=12s into the assimilation run.

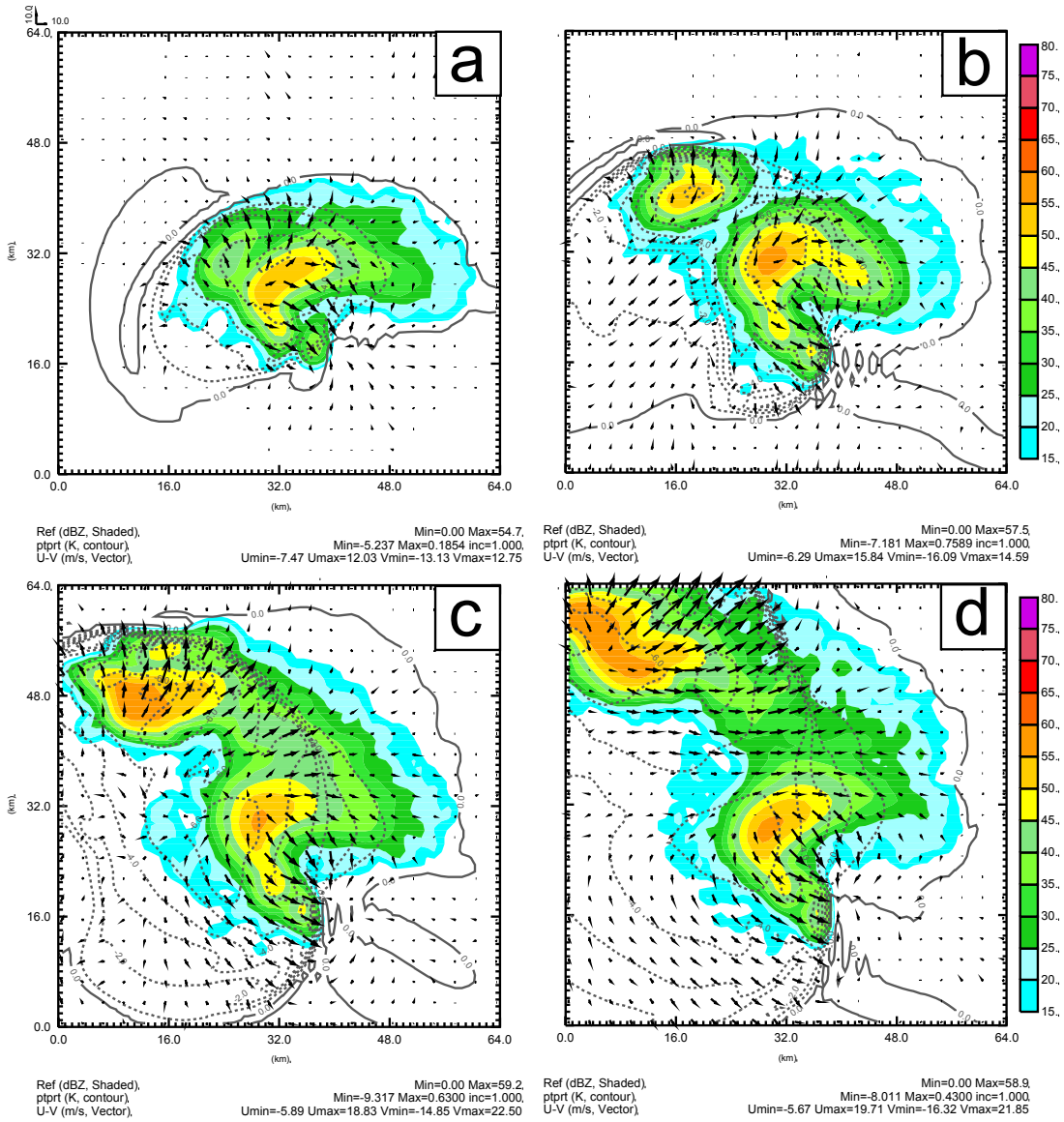

Fig. 2.13. The perturbation horizontal winds, perturbation potential temperature and reflectivity pattern for the "Qr_5" experiment at z=250m every 20 minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). (a) t=20min, (b) t=40min, (c) t=60m, (d)t=80m.

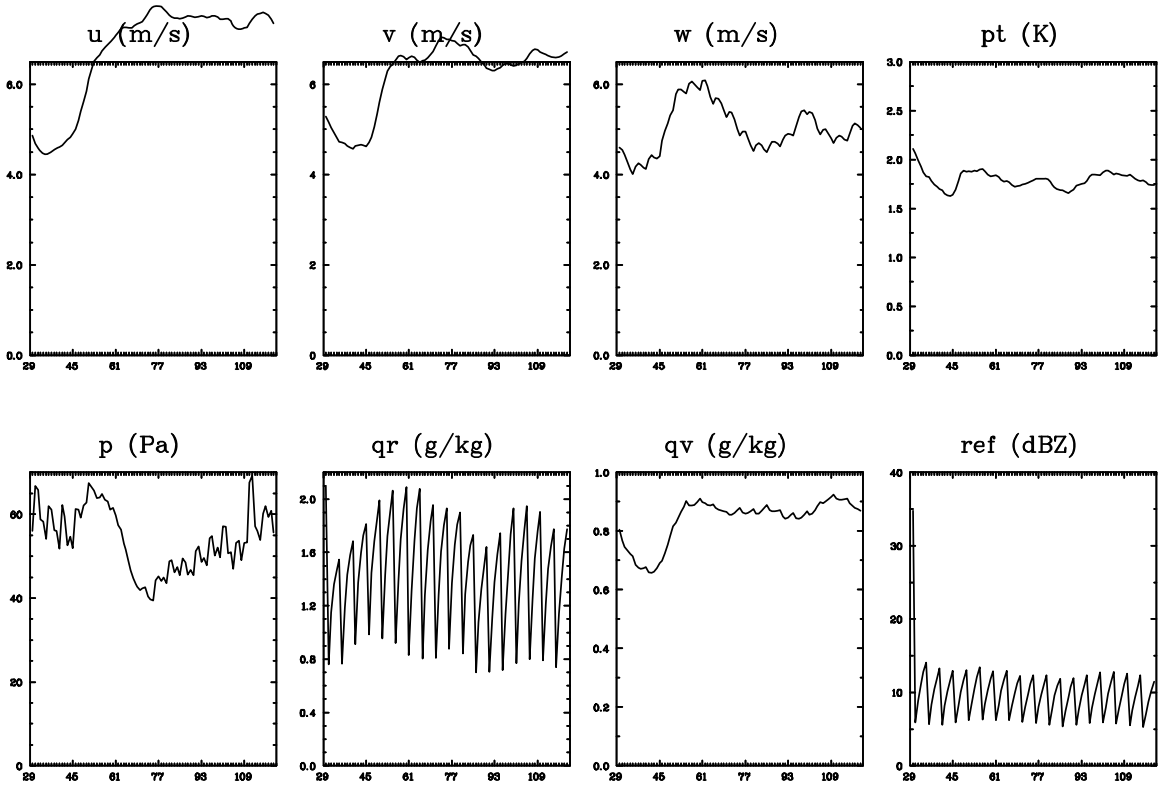

Fig. 2.14. The RMS error evolution every one minute for the "Qr_5" experiment

2.2.3.2 Experiments with observations from two model variables

In this section, we will mainly discuss the experiments "UV_5", "UPt_5", "UQv_5" and "UQr_5", which are performed with an assimilation frequency of every five minutes. Other experiments with same amount of model field combinations but with different assimilation frequencies (every one minute, every ten minutes) will be investigated in Section 2.2.3.5.

The experiments "UPt_5" and "UQr_5" do not reach a "successful-recovery" while the experiment "UV_5" and "UQv_5" successfully recover the storm structures at t=68 minutes and t=51minutes into the assimilation run respectively. This implies that the horizontal wind field and the water vapor field have much larger impact on the storm-

scale data assimilation. As an example, Fig. 2.15 shows the evolution of the RMS error for the experiment "UQv" and Fig. 2.16 shows the plots of perturbation horizontal winds, perturbation potential temperature and reflectivity pattern at z=250m every twenty minutes from t=20 minutes into the assimilation run to t=80 minutes (which corresponds to the time period of the simulation run from t=50 minutes to t=110minutes). In Fig. 2.15, the RMS errors in every model data fields decreased quickly as the U component of the perturbation horizontal wind fields and the perturbation water vapor are assimilated into the model. After sixty minutes of assimilation, the storm structures have already been successfully recovered, and the storm looks nearly exactly the same as that in the control run (Fig. 2.16c).

The experiment "UPt_5" does not successfully recover the storm structures even after 90 minutes of assimilation. This behavior is a little bit unexpected since previous studies (Park and Droegemeier 2000; Sun 2005a) demonstrated the importance of the temperature field. A detailed investigation of this experiment reveals that at the end of the assimilation run, the RMS errors of the wind fields and water vapor fields remain above 2.0m/s and 0.4g/kg respectively. Under the criterion of "successful-recovery" in this study, this assimilation run cannot be regarded as an experiment with a "successful-recovery". It can be concluded that the perturbation potential temperature has less impact on the data assimilation than horizontal winds and moisture field. However, on the other hand, the rainfall pattern (not shown) is recovered very well in this experiment (The RMS error of reflectivity reaches 5dBz at the end of the assimilation).

The experiment "UQr_5" also fails to reach a "successful-recovery" after ninety minutes of assimilation. This implies that directly assimilating the radial velocity (U) and

reflectivity (Qr) from single radar alone (without the proper adjustment in other hydrometers and temperature) may not be good enough to make a successful storm-scale data assimilation.

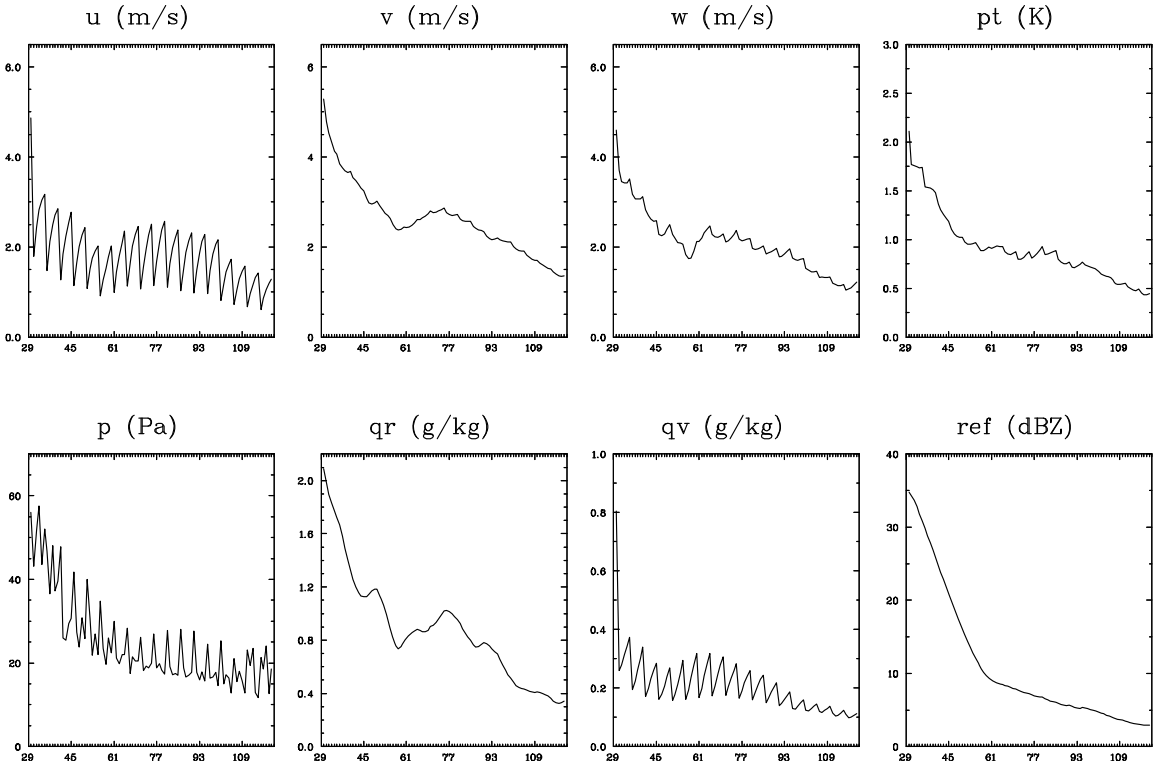

Fig. 2.15. The RMS error evolution every one minute for the “UQv_5” experiment

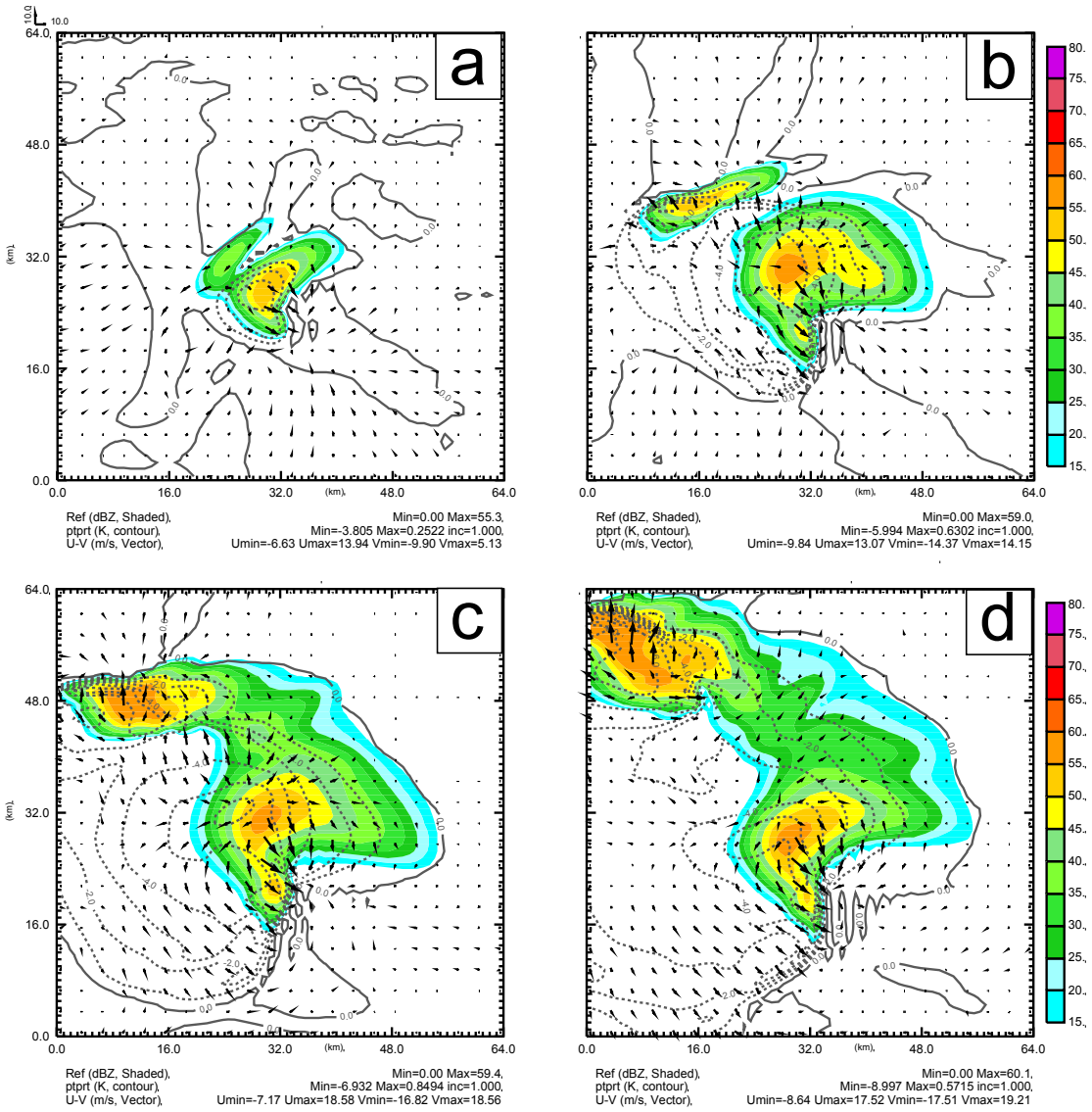

Fig. 2.16. The perturbation horizontal winds, perturbation potential temperature and reflectivity pattern for the experiment "UQv_5" at z=250m every 20 minutes from t=20 minutes into the simulation run to t=80 minutes (which corresponds the time period from t=50 minutes into the simulation run to t=110minutes). (a) t=20min, (b) t=40min, (c) t=60min, (d) t=80min.

2.2.3.3 Experiments with observations from horizontal winds plus one more model variable

The experiment "UV_5" yields a "Successful-recovery time" of sixty-eight minutes (Table 2.2). It indicates that with observed horizontal winds, the assimilation run can

successfully recover the storm structures. This has good practical implication for dual-radar or multiple-radar data assimilation where the horizontal wind fields can be revealed in a better quality than single radar data assimilation. On the other hand, the experiment "UV_5" still takes sixty-eight minutes to reach a "successful-recovery". That is a little bit long considering the operational need of quick delivery of storm-scale predictions.

Assimilating observation from one more model variable may mitigate this problem. As seen in Table 2.2, the experiment "UVPt_5" has a "successful-recovery time" of thirty-two minutes, the experiment "UVQv_5" has a "successful-recovery time" of twenty-three minutes and the experiment "UVQr_5" with a "successful-recovery time" of forty-eight minutes. All these "successful-recovery time" are much shorter than the "sixty-eight" minutes in the "UV_5" experiment.

It is worthy to note that the shortest "successful-recovery" time in "UVQv_5" experiment confirms the importance of water vapor field, which exerts much larger impact than the perturbation potential temperature field and rain water mixing ratio field. In the meantime, the perturbation potential temperature field has evidently positive contribution to the data assimilation at the presence of a quality horizontal wind fields. The rain water mixing ratio filed can also help accelerate the "successful-recovery" time. It implies that directly assimilating the radar reflectivity is helpful to the storm-scale data assimilation under the context that there exists sufficient horizontal wind information.

2.2.3.4 Experiments with observations from 3D winds plus one more model variable

The experiment assimilating the completely 3D wind fields yields a "successful-recovery time" of thirty-eight (the "UVW_5" experiment in Table 2.2). This is quite acceptable for storm-scale data assimilation and forecast. It is much shorter than the

experiment assimilating only horizontal wind fields (the "UV_5" experiment in Table 2.2). Along with the fact that the "UV_5" experiment performed much better than the "U_5" experiment, it is suggested that for storm-scale data assimilation, most efforts should be made on getting a complete 3D wind fields as accurate as possible. In practice, this implies that multiple radar data assimilation can benefit storm-scale NWP. It is also confirmed by other research. Schenkman et al (2011) has shown that assimilating extra CASA (Collaborative and Adaptive Sensing of the Atmosphere) radar data in addition to WSR-88D radar data can improve the forecast of convective storms(A CASA radar observes the lower part of the atmosphere where a WSR-88D radar generally cannot observe). In Chapter 4, we will demonstrate the benefits from four WSR-88D radars compared to single radar for the prediction of a tornadic supercell thunderstorm. Another implication from the above finding is to further develop advanced velocity retrieval schemes in order to get better wind analysis. For example, Shapiro et al (2009) reported that including a vorticity equation constraint into a variational framework could improve dual-Doppler wind analysis.

The experiments "UVWPt_5", "UVWQv_5" and "UVWQr_5" all performs better than the experiment "UVW_5" since all of them have much smaller "successful-recovery time" than the experiment "UVW_5". It indicates that one more observed model variable can improve the data assimilation further even the full dynamic fields are observed very accurately. When the full 3D wind fields are assimilated, the model will experience an adjustment period, during which the perturbations in the 3D wind fields spread to other variables to re-construct dynamic balance among different model fields. One more

observed model variable can help accelerate the adjustment process and then improve the performance of the data assimilation.

Among the experiments “UVWPt”, “UVWQv”, “UVWQr”, the water vapor field again shows larger impact than the potential temperature field and the rain water field.

2.2.3.5 The impact of assimilation frequency

In Table 2.2, the experiments for assimilating observations every five minutes all have a small “successful-recovery time” than the corresponding experiments assimilating the same amount of observations but every ten minutes. The better performance for the five-minute-interval experiments over the ten-minute-interval experiments is due to the benefit from a high assimilation frequency. If observations are assimilated into the model and no more new data are ingested, the impact of the observations will gradually decreased as the assimilation model is integrated forward in time. However, when the model variables are updated more frequently in time, the data impact can be consolidated and sustained longer. In this way, the high data updating frequency improve the quality of the data assimilation.

The above statements naturally lead to a question: since a higher assimilation frequency can produce better results, can we assimilate observations as frequent as possible in order to get a better quality data analysis? The answer is no. In Table 2.2, assimilating observations every one minute does not generally show an evident improvement over assimilating observations every five minutes. The “successful-recovery time” of the experiments “UVQv_1”, “UVWQv_1”, “UVWQr_1” are generally only 2~4 minutes smaller than that of the experiments “UVQv_5”, “UVWQv_5”, “UVWQr_5”. This kind of slight improvement of one-minute-interval assimilation over

five-minute-interval assimilation is very negligible compared to the much more burden it may bring to the computation resources.

The experiments of "UQv_1", "UVPt_1", and "UVWPt_1" have a longer "successful-recovery time" than the experiments of "UQv_5", "UVPt_5", and "UVWPt_5". This demonstrates that in some situation, such as the above combinations of observations, high time frequency of assimilation may have negative impact on the analysis. There are also some exceptions where more frequent data assimilation yields better results.

The above findings suggest that for different observation combinations, the optimal assimilation frequencies may be different. Hu and Xue (2007) investigated the impact of the assimilation settings on the data assimilation and its following forecast. They found that a ten-minute-interval assimilation scheme is the best choice for their data assimilation and forecast of the 8 May 2003 Oklahoma City Tornadic Thunderstorm. However, in their research, a five-minute-interval assimilation scheme, if properly configured, could also make a good prediction although slightly worse than the best run. This finding contributes to our understanding of the impact of storm-scale data assimilation frequency. It can be therefore concluded that the data assimilation frequency has important impact on the quality of data assimilation and the subsequent forecast. An assimilation frequency of every five or ten minutes is now a common practice in storm-scale data assimilation.

2.2.4 Summary and future plan

Nowadays, there is a great need for quality storm-scale NWP (Numerical Weather Prediction) of thunderstorms. Various storm-scale data assimilation schemes are

developed in order to provide a more accurate initial condition for the storm-scale NWP model in order to deliver a better forecast. Despite many studies focus on this area, a clear understanding of the impacts of different data fields and data assimilation frequency is still lacking. Some researchers have put their attentions to this issue. However, the complexity of this topic and the difference in the findings of previous research call for further investigation on this topic.

In this study, under the context of simplified 3D variational data assimilation, we examined the impact of different data fields and assimilation frequencies through a series of data assimilation experiments that ingested different combinations of observations taken from model variables. A term of "successful-recovery" is defined using the RMS error of model variables between the assimilation run and the control run. It describes a kind of criteria when both the dynamic and thermodynamic structures of the storm in the assimilation run are recovered to be very close to the simulated storm in the control run. This method is then used to evaluate the performance of different data assimilation experiments so that the impacts of different data fields and assimilation frequencies are disclosed.

It is found that observations from only one model variable are not sufficient to make a "successful-recovery" after ninety minutes of data assimilation. It indicates that the ability of one observed model variable to make a good storm-scale data assimilation is really limited. The observations from the vertical velocity, the water vapor mixing ratio, the potential temperature are good at recovering thermodynamic fields and spinning-up the rainfall pattern to an acceptable accuracy but perform poorly in rebuilding the three dimensional wind fields. The rain water mixing ratio observations are very helpful to

reduce the rainfall "spin-up" problem and build a cold pool but fail to re-establish other dynamic and thermodynamic fields. The observations from one component of wind fields can produce some kind of recovery at temperature, moisture, vertical velocity fields and rainfall pattern. However, the recovered structure, especially the three-dimensional wind structure, is still noticeable different from the truth.

Another important finding is that horizontal wind fields have the most important impact on the storm-scale 3D data assimilation. It is not only because the assimilation of the horizontal wind fields alone can successfully recover all other model fields, but also due to the fact that the horizontal wind fields are a must and a good sustain to express/enhance the impact of other model variables such as perturbation water vapor, perturbation potential temperature, rain water mixing ratio. In practice, great efforts should be made to get as much wind information as possible and as accurate as possible. This calls for dual-radar data assimilation, multiple radar data assimilation and the development of advanced wind analysis techniques.

When a good picture of the horizontal wind fields can be obtained, extra observations from other model variables will improve the data assimilation and the subsequent forecast. Among these "other model variables", the perturbation water vapor field exerts the greatest impact. In practice, to get storm-scale water vapor observations is a very difficult task. However, some of water vapor information can be derived from the refractivity data observed by radars. These derived data can then be assimilated into the model. There are already some research focusing on this issue (Fabry et al. 1997; Bodine et al. 2010). In the future, water vapor information may be available in high resolution due to the advances in the observing systems such as next generation GOES (Geostationary Operational

Environmental Satellite), dense ground-based GPS receiver network. At that time, our NWP for storm-scale phenomena may be greatly improved.

The last finding is about the impact of data assimilation frequency. Generally, the data assimilation frequency will exert important effect on the quality of the data assimilation. Despite of the difficulties to find an optimal data assimilation frequency for a data assimilation configuration, it is now a common practice to assimilate observations every five or ten minutes for storm-scale NWP.

2.3 The role of wind fields in storm-scale NWP - a real data case study

2.3.1 Introduction

One of the most important findings in Section 2.2 is that the horizontal wind fields exert the greatest impact on the storm-scale 3D variational data assimilation. This conclusion is drawn from an idealized case study. Although it agrees with some previous studies (e.g. Weygandt et al. 1999), it is still in question whether the same conclusion can be drawn under the context of real world data assimilation and numerical forecast. Hu et al. (2006b) reported that assimilating radial velocity from single radar alone failed to predict the Fort Worth, Texas, tornadic thunderstorms. In the meantime, it is also mentioned in Hu and Xue (2006b) that the small impact of radial velocity data is partly due to the limited data availability from single radar. What will happen if we can assimilate the radial velocity data from multiple radars? This calls for more real case studies. Considering the current operational WSR-88D radars in NEXRAD network are densely deployed in some areas of the Nation, we may easily find thunderstorms that fall into the coverage of several radars. The 4-5 May 2007 Greensburg tornadic thunderstorm

is such a case. In the following sections, we will seek to investigate the impact of the wind fields, observed by six radars, on the storm-scale analysis and forecast.

This part is organized as follows. Section 2.3.2 briefs the ARPS prediction model, the ARPS 3DVAR system and the included cloud analysis package. Section 2.3.3 is a general description of the 4-5 May 2007 Greensburg, Kansas (KS) tornadic thunderstorms. In Section 2.3.4, the design of experiments is discussed and Section 2.3.5 presents the results from the experiments. Summary and future plan is provided in Section 2.3.6.

2.3.2 The ARPS model, 3DVAR system and cloud analysis scheme

The ARPS (Advanced Regional Prediction System) is used as the prediction model in this study. It is a general-purpose three-dimensional, non-hydrostatic and compressible atmospheric model that is well documented in several early publications (Xue et al. 2000; Xue et al. 2001; Xue et al. 2003). In this section, we will only briefly review ARPS 3DVAR data assimilation system and its cloud analysis scheme. Following Gao et al. (2004), the standard cost function of 3DVAR can be written as,

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) + \frac{1}{2}\left[H(\mathbf{x})-\mathbf{y}^o\right]^T \mathbf{R}^{-1}\left[H(\mathbf{x})-\mathbf{y}^o\right] + J_c(\mathbf{x}) \qquad (2.3)$$

where the first term on the right hand side measures the departure of the analysis vector, $\boldsymbol{x}$, from the background vector, $\boldsymbol{x}^b$, weighted by the inverse of the background error covariance matrix **B**. In the current ARPS 3DVAR system, the analysis vector $\boldsymbol{x}$ contains the three wind components (u, v, and w), potential temperature (θ), pressure (p) and water vapor mixing ratio (q_v). The second term, observation term, measures the departure of the analysis vector, projected into observation space, from the observation vector, $\mathbf{y}^o$. In this study, $\mathbf{y}^o$ only includes radar radial velocity data. The analysis is

projected to the observation space by the forward operator H that is defined by radar radial wind equation and interpolation operator from model grid points to radar observation locations. The observation term is weighted by the inverse of observation error covariance matrix **R** that includes both the instrument and representativeness errors. Term $J_c(\mathbf{x})$ in Equation (2.3) represents dynamic or equation constraints.

By defining $\sqrt{\mathbf{B}}\mathbf{v} = \left(\mathbf{x} - \mathbf{x}^b\right)$, the cost function is changed into an incremental form:

$$J_{inc}(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\left(\mathbf{H}\mathbf{B}^{1/2}\mathbf{v} - \mathbf{d}\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{B}^{1/2}\mathbf{v} - \mathbf{d}\right) + J_c(\mathbf{v}) \tag{2.4}$$

where **H** is the linearized version of H and $\mathbf{d} \equiv \mathbf{y}^o - H\left(\mathbf{x}^b\right)$. In the current version of the ARPS 3DVAR system, the cross-correlations between variables are not included in the background error covariances. The spatial covariances for background error are modeled by a recursive filter (Purser et al. 2003b, 2003a). The corresponding covariance matrix, **R**, is diagonal, and its diagonal elements are specified according to the estimated observation errors.

In ARPS 3DVAR, the mass continuity equation is imposed as a weak constraint. This constraint builds up the relationship between different wind components. Gao et al. (1999; 2004) and Hu et al. (2006b) found that this constraint is very effective in producing reasonable analyses of vertical velocity. When a stretching grid strategy is used in vertical direction, a special treatment (Hu et al. 2006b), which assigns different weighting coefficients in the horizontal and vertical directions, is required to apply this weak constraint.

The cloud analysis is based on the Local Analysis and Prediction system (LAPS, Albers et al. 1996) with significant modifications by Zhang (1998), Brewster (2002), and

Hu et al (2006a) and used to assimilate reflectivity data into the ARPS. The purpose of assimilating reflectivity is to decrease the "spin up" time of storm development in numerical models.

2.3.3 The Greensburg Kansas tornadic thunderstorm

We chose the 4-5 May 2007 Greensburg, Kansas (KS), tornadic thunderstorm case for our test because it is well documented, falling in the coverage of six WSR-88D radars. The storm complex produced 18 tornadoes in the Dodge City forecast area and 47 tornado reports in Kansas, Nebraska and Missouri. One of them is the strongest tornadoes in recent years. The tornado started moving through Greensburg at 0245 UTC 5 May 2007 (21:45 CDT 4 May) and destroyed over 90 % of the town. The tornado damage was rated at EF5 - the highest rating on the Enhanced Fujita scale (McCarthy et al. 2007).

The synoptic setting for this event consisted of a deep long-wave trough over the western U.S., a surface low over eastern Colorado, and a quasi-stationary front extending from the low across northwest Kansas into northeast Nebraska (Fig. 2.17). A dryline stretched generally southward across western Kansas, Oklahoma, and into west Texas. A very moist and unstable air mass was found east of the dryline, where values of surface-based convective available potential energy (CAPE) were above 4000 J kg^{-1} across central Oklahoma and south of central Kansas. Values of 0-3 km storm-relative environmental helicity (SREH) were in excess of 150 m^2 s^{-2} throughout much of Oklahoma and Kansas, providing an environment favorable for supercell thunderstorms.

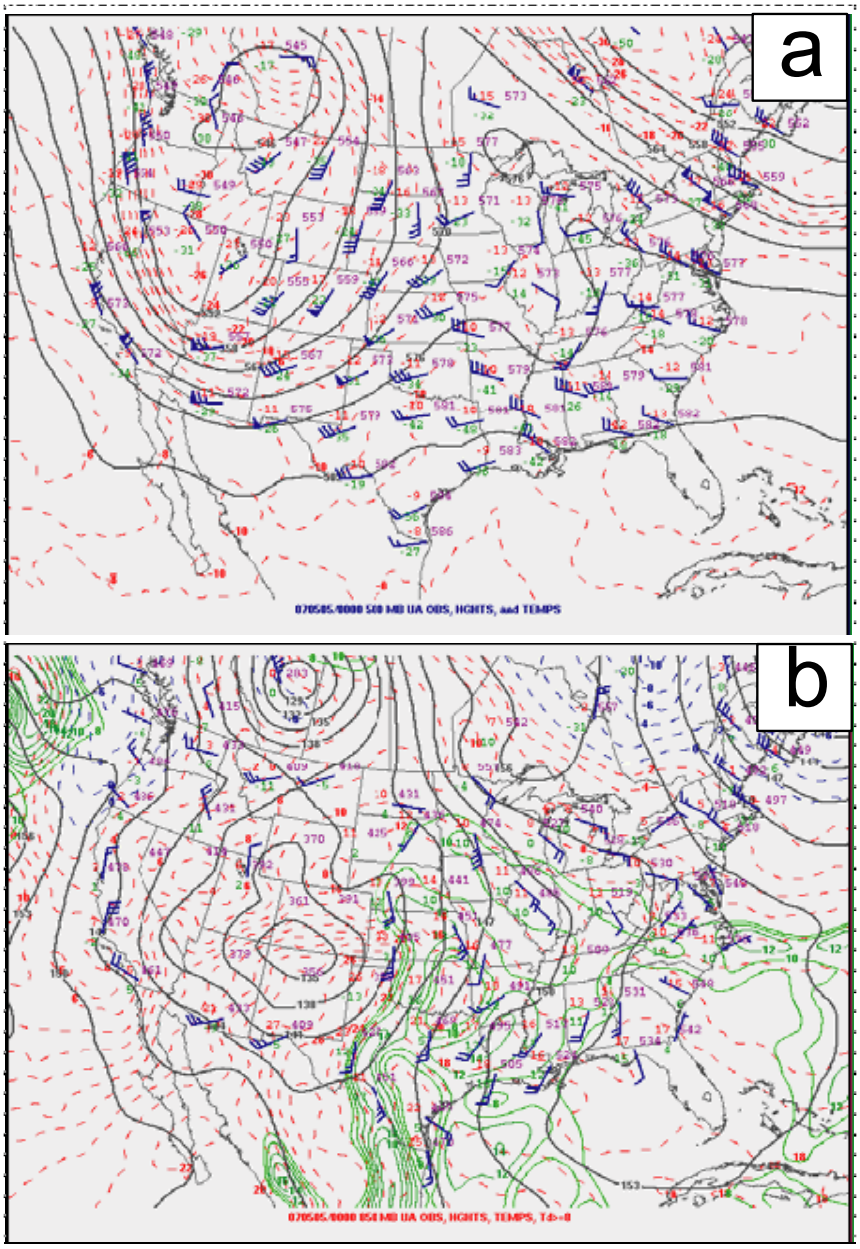

Fig. 2.17. NCEP NAM analysis valid at 00 UTC 5 May 2007 at (a) 500 hPa, (b) 850 hPa. Heights are shown as black contours (in decameters); and temperatures are shown by dashed red contours.

Initial storm development occurred over the northern Texas panhandle/Oklahoma border around 2210 UTC on 4 May 2007. A complex cell evolution ensued in which several storm splits were observed in succession over the next 2 h. As one of the storms

crossed the border into Kansas near 0040 UTC, it split with the right-moving storm evolving into the tornadic supercell thunderstorm that passed over Greensburg. This storm moved from 212°(the direction with the north as 0° and clockwise turn, hereafter) at 13 ms^{-1} and developed its hook echo signature by 0106 UTC. Between 0130 UTC and 0148 UTC, a strong middle-level mesocyclone was very clear and persistent in the data of Dodge City WSR-88D radar (not shown). The supercell was observed to take a classic hook echo shape by 0230 UTC as the strength of its rotation increased dramatically. The tornado that eventually produced the violent EF-5 damage at Greensburg was first observed near 0200 UTC (Lemon and Umscheid 2007). Forecasters at the National Weather Service Dodge City Weather Forecast Office issued a tornado warning with 30 minutes lead-time for this event.

Over the next hour from 0230 to 0330 UTC, this tornadic supercell thunderstorm (which we call the dominant storm) turned a bit more to the right, moving from 219° as the storm motion slowly decreased from 10 m s^{-1} to near 8 m s^{-1} (Lemon and Umscheid 2007). In comparison, the group of non-supercell thunderstorms to the northwest of the dominant storm moved much faster at 23 m s^{-1} from 206°. While the violent EF-5 tornado that hit Greensburg dissipated near 0305 UTC, a second strong EF-3 tornado developed near 0303 UTC, lasted for 65 min, and had a path length of over 43 km. This 1-h period from 0230 to 0330 UTC is selected for study. During this period, the storm motion is fairly steady and strong tornadoes are observed throughout the period.

2.3.4 Experimental design

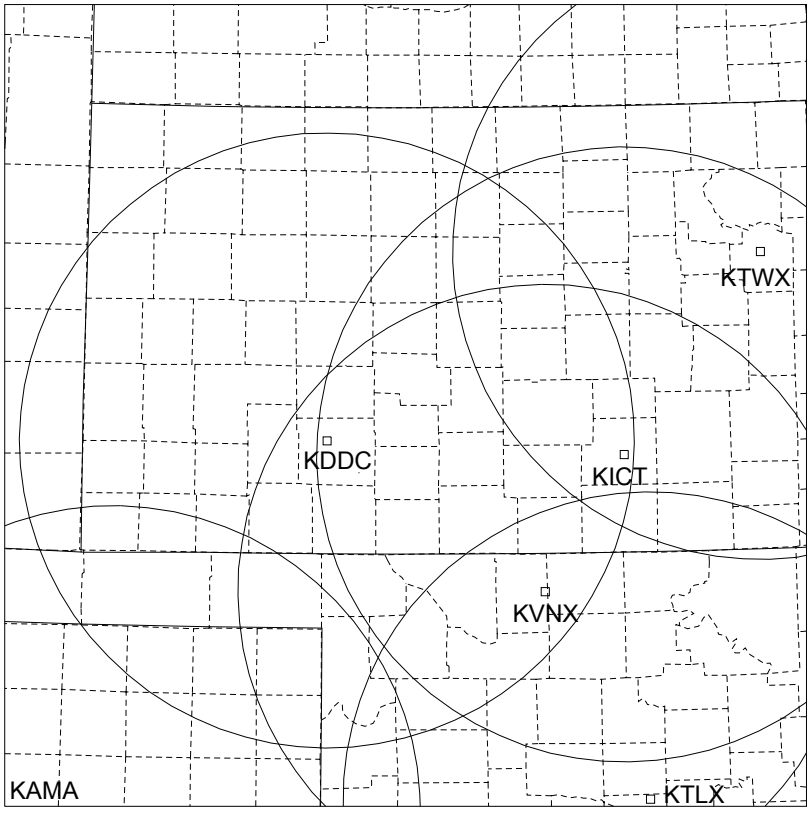

Fig. 2.18. The model domain with county boundaries. The six radars as well as their coverage circles are also shown.

For this real data case experiment, we use 3-km grid spacing with 200x200 grid points in the horizontal. The model domain is shown in Fig. 2.18. The domain is selected with sufficient coverage to contain the principal features of interest while maintaining some distance between primary storms and the lateral boundaries. The model uses 47 terrain-following vertical layers, with nonlinear stretching, via a hyperbolic tangent function, that yields a spacing of 100 m at the ground and expands to approximately 800 m at the top of the domain. The ARPS 3DVAR technique is used to create rapid analysis cycles and the cloud analysis scheme follows the 3DVAR analysis step to assimilate the

radar reflectivity data. In addition, the mixing ratio of precipitation (including rain water, snow, and hail) and potential temperature are adjusted within the cloud analysis procedure based on reflectivity measurements. The other hydrometeor variables are not adjusted.

In the first experiment (named as experiment VrOnly), only radial velocity observations from the six radars are used. The second experiment Vr&RF uses both radial velocity and reflectivity data. For both experiments, data from six radars at Dodge City (KDDC), (Vance AFB, OK (KVNX), Wichita Kansas (KICT), Oklahoma City (KTLX), Amarillo TX (KAMA) and Topeka Kansas (KTWX) are used in the 3DVAR and cloud analysis system. Each experiment consisted of a 1-h assimilation period (from 0130-0230 UTC) and a 1-h forecast period (0230-0330 UTC). The background and boundary condition came from an analysis from a mesoscale ensemble assimilation system (Stensrud and Gao 2010). While Stensrud and Gao (2010) performed a 3DVAR analysis only at one time level before the launch of the forecast, the present study uses an assimilation period that consists of thirteen analysis cycles at 5-min interval. A five minutes ARPS forecast follows each analysis. This process is repeated until the end of the 1-h assimilation period. From the final analysis, a 1-h forecast is launched.

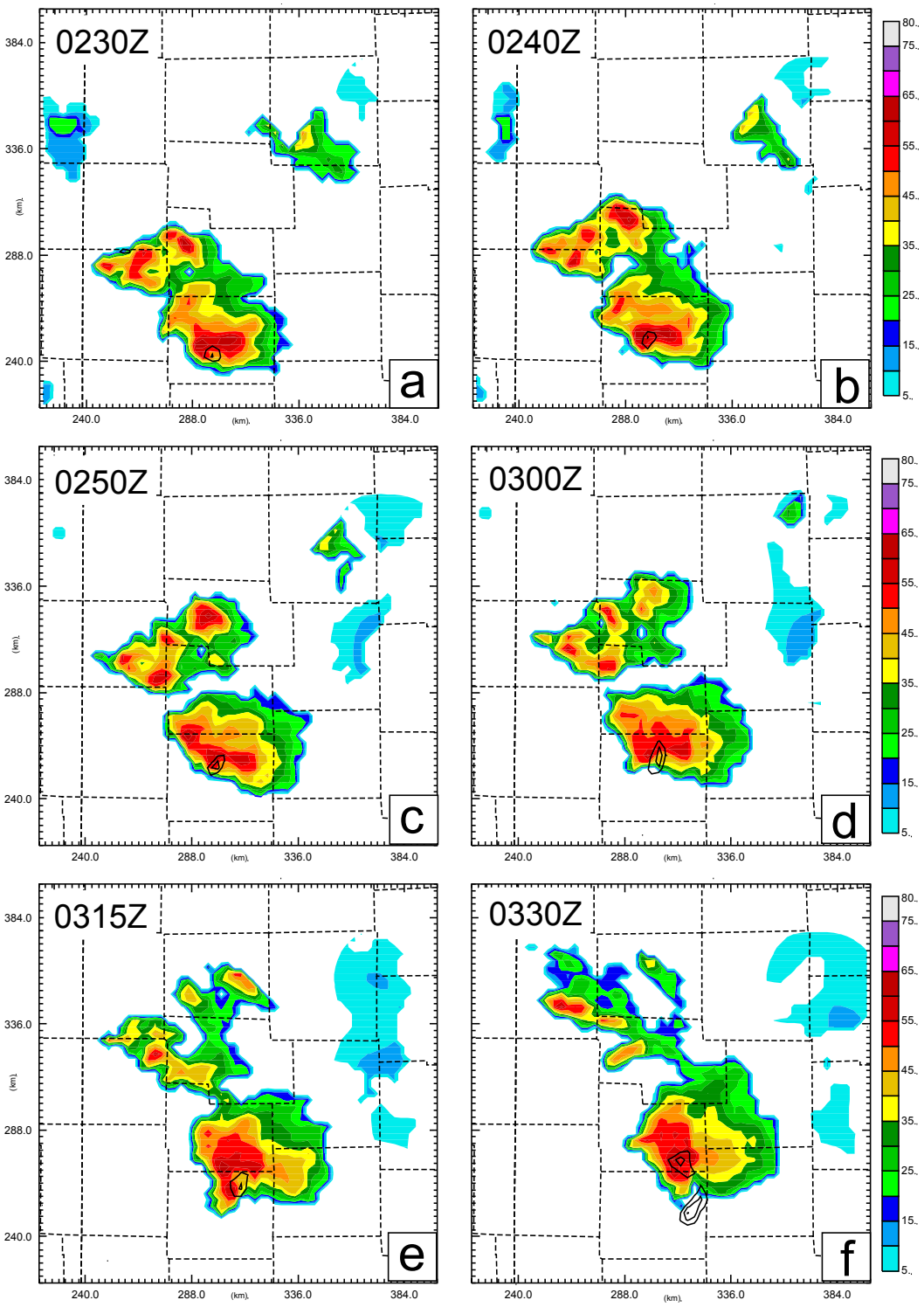

Fig. 2.19. Observed radar reflectivity mosaic (dBZ) at 2km MSL from the KDDC, KICT, KVNX Doppler radars valid at (a) 0230, (b) 0240, (c) 0250, (d) 0300, (e) 0315, and (f) 0330 UTC 5 May 2007 over western Kansas. Solid lines indicate the locations of strong cyclonic rotations (see text for details).

The radar reflectivity mosaic from the aforementioned six WSR-88D radars is used for forecast verification. The evolution of the storm as indicated by the radar reflectivity mosaic at the 2 km MSL is shown in Fig. 2.19 from 0230 to 0330 UTC. The major supercell thunderstorm at southernmost side in Fig. 2.19 is the focus of this study. It produces the EF-5 tornado hitting the Greensburg area between 0245UTC~0305UTC. It bears a hook echo sign at 0230 UTC (Fig. 2.19a, NOTE: the hook echo is not as clear as shown in original radar PPI display due to data interpolation and multiple radar mosaic, however, it is still distinguishable). As the major storm reaches Greensburg, the hook echo signature becomes less prominent (Fig. 2.19c,d, it is also not easy to tell the hook echo in radar PPI display during this period) due to reflectivity wrapping up. During this period, the radar velocity observations (not shown here) show strong cyclonic rotation, indicating the location of the strong tornado. This strong rotation is illustrated by circles of solid lines in Fig 2.18. These solid lines are drawn based on the contour of the analyzed vertical vorticity using data from the six Doppler radars, which is of $0.003s^{-1}$ and $0.006s^{-1}$ respectively. The major storm moves gradually towards northeast. After passing the town Greensburg, the storm maintains a very strong circulation and continues to move to the northeast. The second EF-3 tornado develops at the end of Greensburg tornado just northeast of the town (McCarthy et al. 2007).

2.3.5 Results of experiments

2.3.5.1 Experiment assimilating radial velocity data alone

The experiment VrOnly assimilates alone the radial velocity observations from the aforementioned six NEXRAD radars. Fig. 2.20 presents for this experiment the

predicted radar reflectivity, horizontal wind vector and vertical vorticity as 2km MSL from 0230UTC to 0330UTC. It can be seen that at the end of 1-h data assimilation (Fig. 2.20a), the major storm has been successfully developed although it is much weaker than observed reflectivity (Fig. 2.19a) at z=2km MSL. It is accompanied with a strong rotation as indicated by the contour of large vertical vorticity. The predicted major storm develops quickly in the following one hour. By 0250UTC (Fig. 2.20c), the reflectivity field has already shown a hook echo sign (although it is not very well defined), which is collated with a strong mesocyclone with vertical vorticity already over 0.01 s^{-1}. The hook echo signature and the very strong rotation maintain until at least 0300 UTC in the forecast. After 0315UTC, the hook echo sign can barely be distinguished and the cyclonic rotation weakens gradually.

During the whole 1-h forecast period, the predicted major storm moves slowly towards northeast. The location, moving path, and timing for the major storm are predicted very well. The storms northwest of the major storm is also reasonably captured during the period 0305 UTC~0330UTC.

On the whole, this VrOnly experiment is able to capture the general evolution of the dominant storm that produces the Greensburg tornado during the 1-h forecast (Fig. 2.20). This is quite inspiring. It means that with much more wind information observed by multiple radars we can reproduce a storm after 1-h data assimilation and then make a reasonable forecast. On the other hand, as indicated by Fig. 2.20a vs. Fig. 2.19a, it is also clear that assimilating only wind fields will delay the spin-up process of rainfall. If additional observations, such as reflectivity data, can be added into the assimilation run, the spin-up problem will be reduced. This will be discussed in the next section.

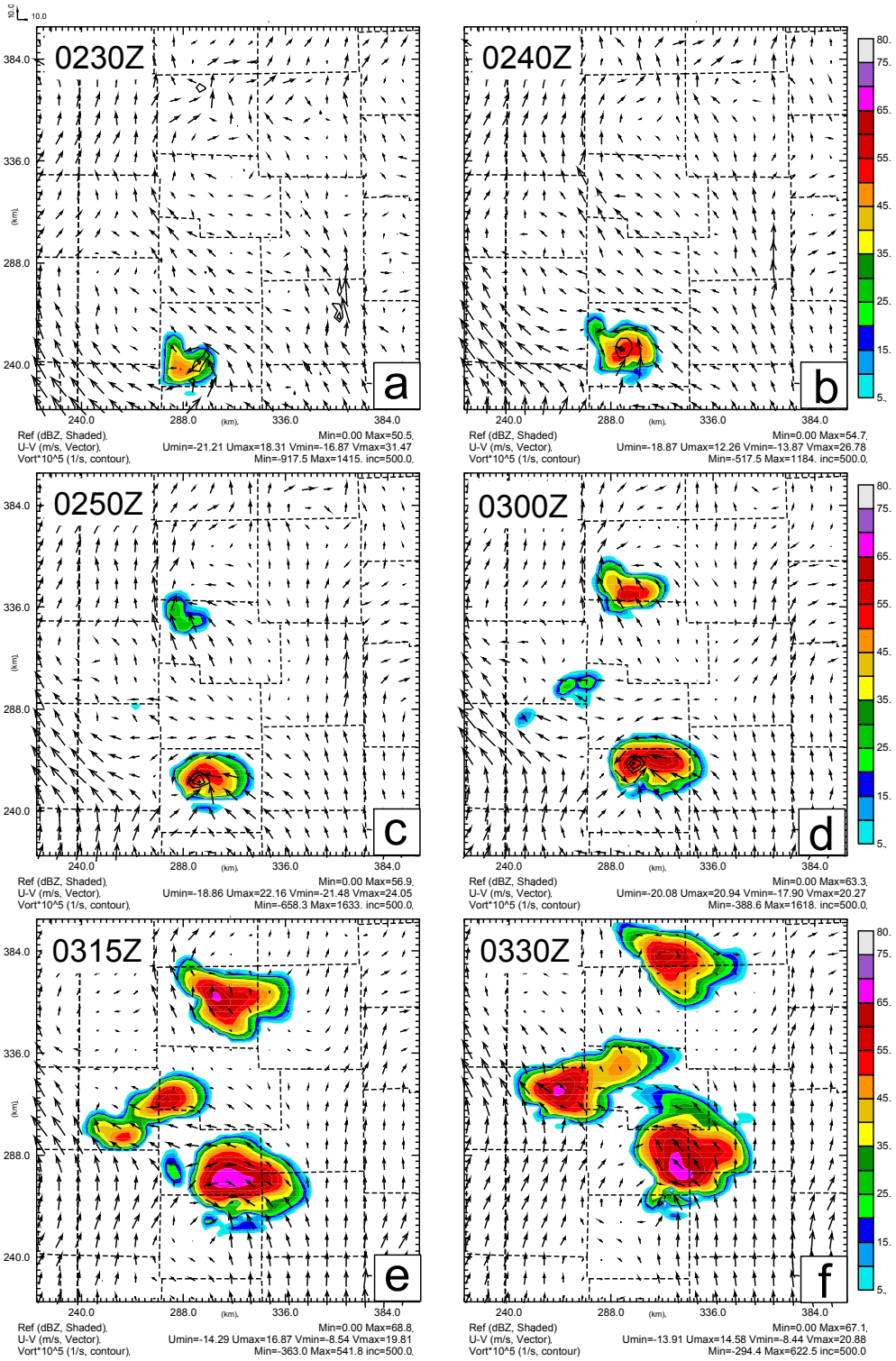

Fig. 2.20. Radar reflectivity (dBZ), horizontal winds, and vertical vorticity (contours staring at $0.005s^{-1}$ with an interval of $0.005s^{-1}$) at 2 km MSL from the VrOnly experiment during 0230~0330UTC 5 May 2007 over western Kansas.

2.3.5.2 Experiment assimilating both wind data and reflectivity data

The Vr&Rf experiment assimilates both the wind data and reflectivity data. The wind data helps establish dynamic fields and recover other model fields. The reflectivity data helps reduce the rainfall spin-up problem. It is expected that this experiment will produce better results than the VrOnly experiment.

Fig. 2.21 shows the simulated reflectivity, the horizontal wind vector, the vertical vorticity at 2km MSL for the 1-h forecast from this experiment. It can be seen that the reflectivity field at the end of this assimilation run looks very close to the observed reflectivity mosaic (Fig. 2.21a vs. Fig. 2.19a) while the VrOnly experiment only produce weak storm echo at the same time (Fig. 2.20a). Therefore, it is very clear that adding the reflectivity data in addition to the wind data greatly speeds up the rainfall spin-up. During this first half hour forecast period, the major storm moves slowly to the northeast with similar timing and location as demonstrated by the observation (Fig. 2.19). For the last half hour forecast period, the major storm moves a little bit faster than observation but still acceptable. Overall, the "Vr&Rf" experiment produces a reasonable forecast after one hour data assimilation.

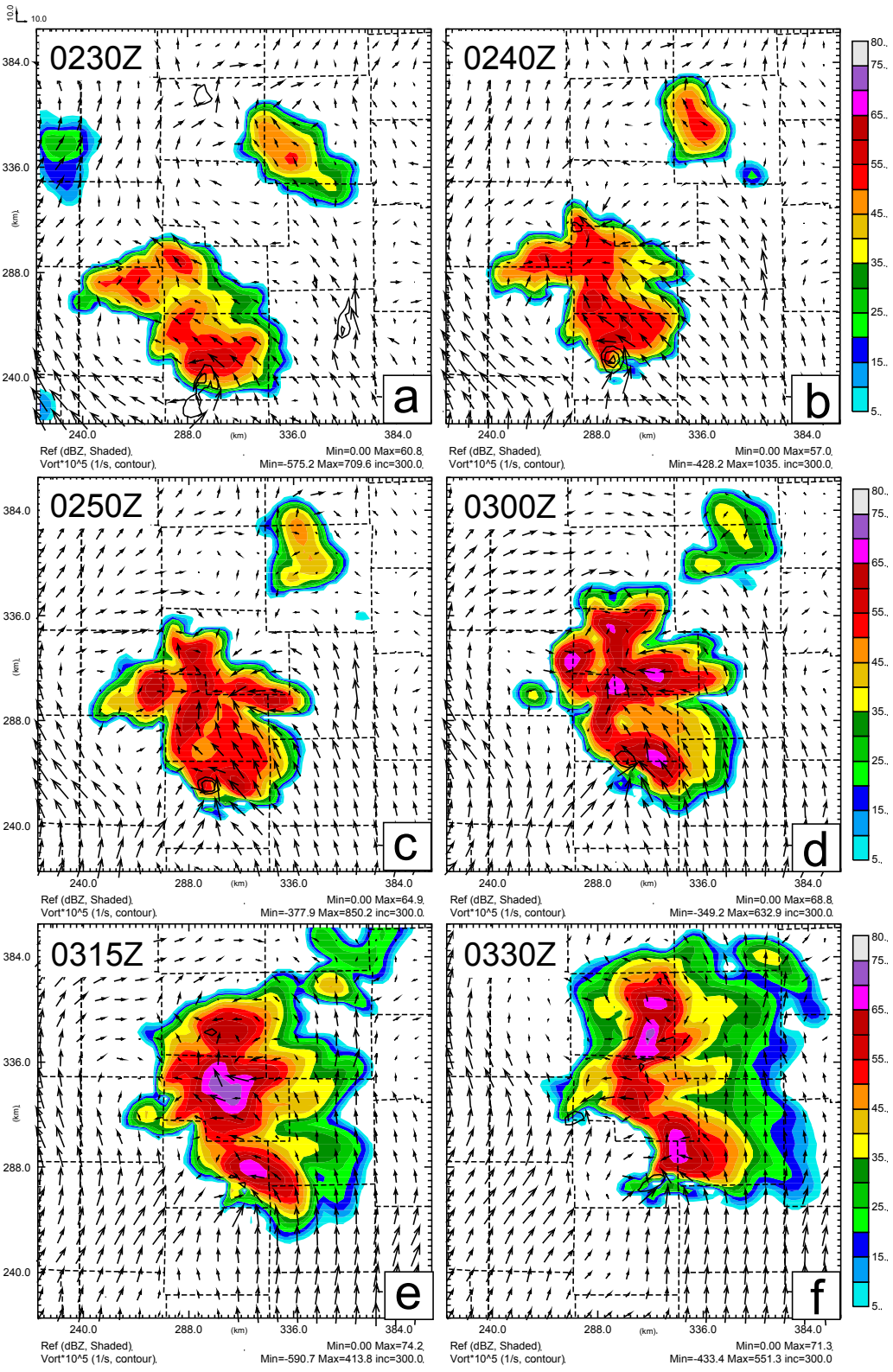

Fig. 2.21. Similar as Fig. 2.20, but for the Vr&Rf experiment.

2.3.6 Summary and future plan

Inspired by the findings in Section 2.2 of this study, the 4-5 May 2007 Greensburg, Kansas (KS), tornadic thunderstorm case is selected to study whether the wind fields can exert the greatest impact on storm-scale 3D variational data assimilation. Two data assimilation experiments are conducted. Their corresponding 1-h long forecasts are evaluated by comparing them with reflectivity mosaic derived from radars.

It is shown that assimilating wind fields observed by six radars can rebuild the storm after sixty minutes of intermittent assimilation and reasonably predict the general evolution of the dominant storm cell that produced the EF-5 Greensburg tornado. This success is much encouraging considering the general failure of single radar data assimilation of radial velocity alone in previous studies. This confirms that more accurate wind fields can lead to better performance in data assimilation and the following forecast. On the other hand, assimilating reflectivity data in addition to wind data can help reduce the spin-up problem.

The above conclusions are mainly based on single real case study. While these findings can provide a guide in the design/improvement of storm-scale observing systems and storm-scale data assimilation configurations, it should be cautious to apply them in general. More case studies on this issue will be explored in the future.

Chapter 3 * Impacts of Beam Broadening and Earth Curvature on Storm-scale 3D Variational Data Assimilation of Radial Velocity with Two Doppler Radars

3.1 Introduction

The operational WSR-88D Doppler radar network (NEXRAD) is an important tool for real-time detection and warning of hazardous weather (Crum and Alberty 1993; Crum et al. 1998; Serafin and Wilson 2000). It is also an essential observing system for initializing non-hydrostatic, storm-resolving (i.e., horizontal grid spacing on the order of 1 km) numerical weather prediction (NWP) models (e.g., Droegemeier 1990; Lilly 1990; Droegemeier 1997). To assimilate these radar data into NWP models, it is necessary to accurately determine the spatial locations of individual radar measurements. Because the propagation path of the electromagnetic waves can be affected by the refractivity of the atmosphere, the propagation path or the ray path is usually not a straight line. A suitable ray path equation is therefore needed. The local direction of the ray path also affects the radial velocity forward operator that projects the Cartesian velocity components on the model grid to the local radial direction in data assimilation systems.

Most early radar data assimilation studies used relatively simple ray path equations in the forward operator formulation which are based on the Cartesian geometry, essentially assuming a flat earth (e.g., Sun et al. 1991; Qiu and Xu 1992; Xu et al. 1995; Qiu and Xu

* This Chapter is published as: Ge, G., J. Gao, K. Brewster, and M. Xue, 2010: Impacts of Beam Broadening and Earth Curvature on Storm-Scale 3D Variational Data Assimilation of Radial Velocity with Two Doppler Radars. *Journal of Atmospheric and Oceanic Technology*, **27,** 617-636.

1996; Sun and Crook 1997; Gao et al. 1998; Sun and Crook 1998; Xu et al. 2001a; Weygandt et al. 2002b, 2002a; Shapiro et al. 2003; Gao et al. 2004). Brewster (2003) applied complete ray path equations into the ARPS (Advanced Regional Prediction system, Xue et al. 2000; Xue et al. 2001; Xue et al. 2003) Data Assimilation System (ADAS) and phase correction technique. Similar ray path equations were applied into the 3.5dVar radar data assimilation system (Gu et al. 2001; Zhao et al. 2006) developed for the Coupled Ocean/Atmosphere Mesoscale Prediction System (Hodur 1997). However, these previous studies did not investigate the impact of the ray path equations on the radar data assimilation systems. Gao et al. (2006, hereafter Gao06) have shown that using simplified radar ray path equations introduces errors that are significant for ranges beyond 30 km. In that paper, a set of four-thirds earth-radius ray path equations is recommended, especially at low elevation angles. However, Gao06 mainly addresses the error in physical location of individual radar measurement. It is also of interest to study how, and to what extent, the neglecting of earth curvature will affect the results of storm-scale radar data assimilation.

In order to compute most accurately the model counterpart of radial wind, one must integrate over all possible model grid points within the radar beam main lobe, which broadens with range. Most radar data assimilation studies do not consider this beam broadening effect. Wood and Brown (1997) introduced a power gain weighted average in the radar forward observation operator in their study on the effects of radar sampling on velocity signatures of mesocylones and tornadoes. Sun and Crook (2001) incorporated a similar beam broadening equation in their 4DVAR radar analysis system. Salonen (2002) approximated the beam broadening effect with a Gaussian function (Probert-Jones 1962)

in the vertical direction and demonstrated slightly positive impact on radar analysis using the High Resolution Limited Area Model (HIRLAM) 3DVAR system. Xue et al. (2006) and Tong (2006) used a power-gain-based sampling in vertical direction to compute the model counterpart of radial velocity in their EnKF work. All these treatments are more reasonable since they are more close to the nature of the radar measurement. Caumont and Durocq (2008) showed that neglecting the beam broadening could cause large errors at distant gates in the simulation of radar data. However, a detailed study of the effect of beam broadening in storm-scale data analysis and assimilation has not yet been investigated.

In this study, the effect of earth curvature and beam broadening in radar data assimilation is investigated using an idealized supercell tornadic thunderstorm. The ARPS 3DVAR system, described in Gao et al. (2002; 2004) and Hu et al. (2006b) is used for this purpose. The ARPS 3DVAR system is capable of analyzing radar radial velocity data along with conventional observations. It is usually used together with the cloud analysis system to initialize hydrometeor related variables and provide a latent heating adjustment. For simplicity in studying the radial velocity effects, in this paper only the simulated radial winds derived from an idealized thunderstorm are used and the cloud analysis is not used. In the ARPS 3DVAR system, the mass continuity weak constraint is included in the cost function that serves to link three wind components together and helps improve wind analysis.

This paper is organized as follows. In Sections 3.2 and 3.3, we will briefly introduce the radar forward observation operator and the ARPS 3DVAR system respectively. In

Section 3.4, the model configuration and experiment design are discussed. The results are presented in Section 3.5, and summary and discussion in Section 3.6.

3.2 The radar forward observation operator

Under the assumption that the refractivity is a function only of height above mean sea level, Doviak and Zrnic (1993) present a formulation that expresses the ray path in terms of a path following a curve of a sphere of radius,

$$a_e = \frac{a}{1 + a\left(dn/dh\right)} = k_e a \tag{3.1}$$

where a is the earth's radius, k_e is a multiplier which is dependent on the vertical gradient of refractive index of air dn/dh, h is the height above the radar altitude, n is the refractive index of air. The assumptions under which the eq. 3.1 is reached also include: 1) The radar ray is launched at a low elevation angle, which is usually the case with weather radars; 2) The refractive index n is close to 1; 3) h<<a; 4) dh/ds <<1, where s is the surface range (distance along the earth's surface).

The refractive index of air, *n*, is a function of its temperature, pressure and humidity. It is convenient to use the quantity *N* called radio refractivity instead of *n*. *N* represents the departure of *n* from unity in parts per million and its variations can be considered more conveniently. *N* has a value of about 300 (at the surface). N is usually taken, subject to certain assumptions, as (Bean and Dutton 1968),

$$N = (n-1)\times 10^6 = 77.6P/T + 3.73\times 10^5 eT^{-2} \tag{3.2}$$

where P is air pressure in hPa (including water vapor pressure), e is water vapor pressure in hPa, and T is air temperature in degrees K. In the above equation, the first term on the right hand side is known as the dry term, the second term is the moist term. The value of

N can be computed from measurement of *P, T, and e*. If h is limited to the lowest 20 km of the atmosphere and $^{dn}/_{dh}$ is -1/(4a) in the lower atmosphere, k_e will equal to 4/3 (Doviak and Zrnic 1993). This is often referred to as the "four-thirds earth radius model".

The following two equations relate *h* and the surface range (distance along the earth's surface), *s*, to radar-measurable parameters, the slant path, *r* and radar elevation angle, θ_e (Doviak and Zrnic 1993),

$$s = k_e a \sin^{-1}\left(\frac{r\cos\theta_e}{k_e a + h}\right) \tag{3.3}$$

$$h = \left[r^2 + (k_e a)^2 + 2rk_e a\sin\theta_e\right]^{\frac{1}{2}} - k_e a \tag{3.4}$$

To consider the curvature of the Earth, the radar forward observation operator can be written as the following equation:

$$v_r^{\ *} = u\cos\theta_e^{'}\sin\phi + v\cos\theta_e^{'}\cos\phi + (w - w_t)\sin\theta_e^{'} \tag{3.5}$$

where ϕ is radar azimuth angle, w_t is the terminal velocity of precipitation, and $\theta_e^{'}$ includes the effect of the curvature of the earth as the following:

$$\theta_e^{'} = \theta_e + \tan^{-1}[(r\cos\theta_e/(k_e a + r\sin\theta_e)] \tag{3.6}$$

In this study, only the effect of beam broadening in the vertical direction is considered. The reason is as the following. In storm-scale NWP, the horizontal resolution is normally between 1 km and 3 km and a 1° half-power beam width will measure about 3490 m at a surface range of 200 km. So a beam lobe at a surface range of 200 km and gate spacing less than 1-km will enclose only 1 to 3 horizontal grid points, even at 1-km grid spacing, which we judge to be too few to have a material difference. However, the vertical resolution of NWP models typically ranges from 20 to 500 m and a beam lobe at

a range of 200 km can span more than seven vertical grid points, much greater than the two grid points that might be used to compute the model counterpart of radial wind with linear interpolation.

At the same time the height of the lowest ray above the ground will increase rapidly with range (Gao et al. 2006). At a surface range of 100 km, the height of the center of a 0.5-degree ray above the ground is about 1.5 km and at 200 km it is about 4 km. So there may be little information observed of the boundary layer, especially far from the radar. Considering beam broadening in the radar forward observation operator may also spread information below the center of the lowest ray.

Following Rihan et al. (2008), the observation operator for mapping data from multiple vertical model levels onto elevation angles is formulated as:

$$V_{r,e} = H_e(V_r) = \left(\sum G V_r \Delta z\right) / \left(\sum G \Delta z\right) \tag{3.7}$$

where $V_{r,e}$ is the radial velocity on an elevation angle, H_e is the radar forward observation operator, V_r is the model counterpart of radial velocity, Δz is the vertical model grid spacing. G describes the two-way power gain distribution within the radar beam and is formulated as $G=e^{-4\ln 4\alpha^2/\beta^2}$ (Wood and Brown 1997) with α as the distance from the center of the radar beam in radians and β as the one degree beam width. The summation is over vertical model grid points enclosed by the half-power beam lobe.

3.3 The ARPS 3DVAR system

Following Gao et al. (2004), the standard cost function of 3DVAR can be written as,

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) + \frac{1}{2}\left[H(\mathbf{x})-\mathbf{y}^o\right]^T \mathbf{R}^{-1}\left[H(\mathbf{x})-\mathbf{y}^o\right] + J_c(\mathbf{x}) \tag{3.8}$$

where the first term on the right hand side measures the departure of the analysis vector, $\boldsymbol{x}$, from the background vector, $\boldsymbol{x}^b$, weighted by the inverse of the background error covariance matrix **B**. In the current ARPS 3DVAR system, the analysis vector $\boldsymbol{x}$ contains the three wind components (u, v, and w), potential temperature (θ), pressure (p) and water vapor mixing ratio (q_v). The second, observation term, measures the departure of the analysis from the observation vector, $\mathbf{y}^o$. In this study, $\mathbf{y}^o$ only includes radar radial velocity data. The analysis is projected to the observation space by the forward operator H which is defined by equations (3.1) ~ (3.7) and an interpolation operator from model grid points to radar observation locations. The observation term is weighted by the inverse of observation error covariance matrix **R** that includes both instrument and representativeness errors. Because only radial velocity data are used in the analysis system and there are no cross-correlations between variables in the B matrix, only wind components will be updated during the minimization process. Term $J_c(\mathbf{x})$ in Eq. (3.8) represents dynamic or equation constraints.

By defining $\sqrt{\mathbf{B}}\mathbf{v} = \left(\mathbf{x} - \mathbf{x}^b\right)$, the cost function is changed into incremental form:

$$J_{inc}(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\left(\mathbf{H}\mathbf{B}^{1/2}\mathbf{v} - \mathbf{d}\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{B}^{1/2}\mathbf{v} - \mathbf{d}\right) + J_c(\mathbf{v}) \tag{3.9}$$

where **H** is the linearized version of H and $\mathbf{d} \equiv \mathbf{y}^o - H\left(\mathbf{x}^b\right)$. In the current version of ARPS 3DVAR system, the spatial covariances for background error are modeled by a recursive filter (Purser et al. 2003b, 2003a). The corresponding covariance matrix, **R**, is diagonal, and its diagonal elements are specified according to the estimated observation errors (1m s^{-1} in this study).

In the ARPS 3DVAR, the mass continuity equation is imposed as a weak constraint. This constraint builds up the relationship among the three wind components. Gao et al. (1999; 2004) found that this constraint is very effective in producing suitable analyses of vertical velocity. When a stretched grid strategy is used in the vertical direction, a special treatment (Hu et al. 2006b), which assigns different weighting coefficients in horizontal and vertical direction, is needed to apply this constraint. More recently, the modified ARPS model equations are included as weak constraints in the 3DVAR scheme. These newly introduced constraints couple the wind components with thermodynamic variables (Ge and Gao 2007). In this study, for simplicity, only the mass continuity constraint is included.

3.4 Experimental design

In this study, we evaluate the impact of beam broadening and earth curvature on data assimilation system using simulated data. Such simulation experiments are usually referred to as observing system simulation experiments (OSSEs). The ARPS model is used in a 3D cloud model mode. The 20 May 1977 Del City, Oklahoma tornadic supercell storm is used to conduct several series of experiments. This storm has been thoroughly studied by multiple Doppler analysis and numerical simulation (Klemp et al. 1981; Ray et al. 1981; Klemp and Rotunno 1983).

The model is configured as the following: $67 \times 67 \times 35$ grid points and $1\text{km} \times 1\text{km} \times 0.5\text{km}$ grid intervals for the x, y, and z directions, respectively, so as to establish a physical domain of $64 \times 64 \times 16$ km. The simulation starts with a modified sounding (as in Klemp et al. 1981) which favors the development of a supercell thunderstorm. The thermal bubble has a 4 K perturbation, and is centered at x=48 km,

y=16 km and z=1.5 km with the lower-left corner of the domain as the origin. The radius of the bubble is 10 km in the x and y directions and 1.5 km in the z direction. The three-category ice microphysical scheme of Lin et al. (1983) is used together with a 1.5-order turbulent kinetic energy subgrid parameterization. Open boundary conditions are used for the lateral boundaries and rigid wall conditions for the top and bottom boundaries. An upper-level Rayleigh damping layer is also included to inhibit wave reflection from the top of the model.

The simulation runs for 3h. The initial convective cell strengthens over the first 20 min and begins to split into two cells at around 1h. To keep the right-moving storm near the center of the model domain, a mean storm speed (U=3 m s^{-1}, V=14 m s^{-1}) is subtracted from the sounding. At about 2h into the simulation, the right mover is still near the center of the domain as expected and the left mover is located at the northwest corner. Fig. 3.1a and Fig. 3.2a show horizontal and vertical cross sections of simulated wind, vertical velocity at 2h respectively (vertical cross section is plotted through line A-B in Fig. 3.1a). A strong rotating updraft (with maximum vertical velocity exceeding 29 m s^{-1}) and associated low-level downdraft are evident near the center of the domain. The updraft tilts eastward in the upper part of the troposphere. The evolution of the simulated storm is qualitatively similar to that described by Klemp and Wilhelmson (1981). After 2 h, the major storm gradually moves a little bit toward the southeastern corner of the model domain, and remains a very strong supercell structure until the end of simulation at 3 h (Fig. 3.7a-c).

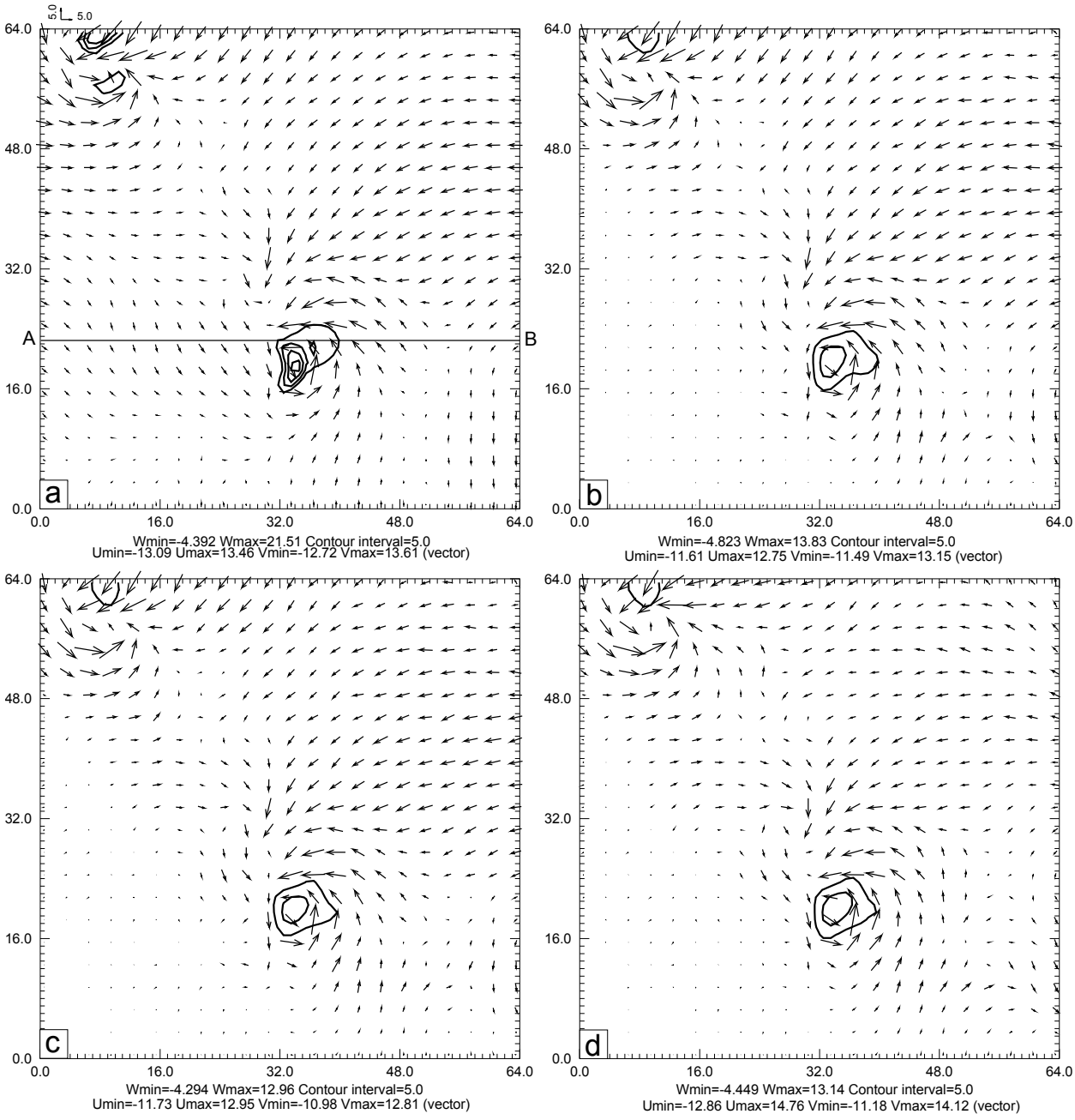

Fig. 3.1. Perturbation horizontal winds (vectors, m s^{-1}) and vertical velocity *w* (contours, m s^{-1}) at t=120 min and 3.5 km AGL for (a) truth simulation; (b) CNTL1_60; (c) NoBB1_60; (d) NoCV1_60. The *w* contour starting from 5 m s^{-1} with an interval of 5 m s^{-1}.

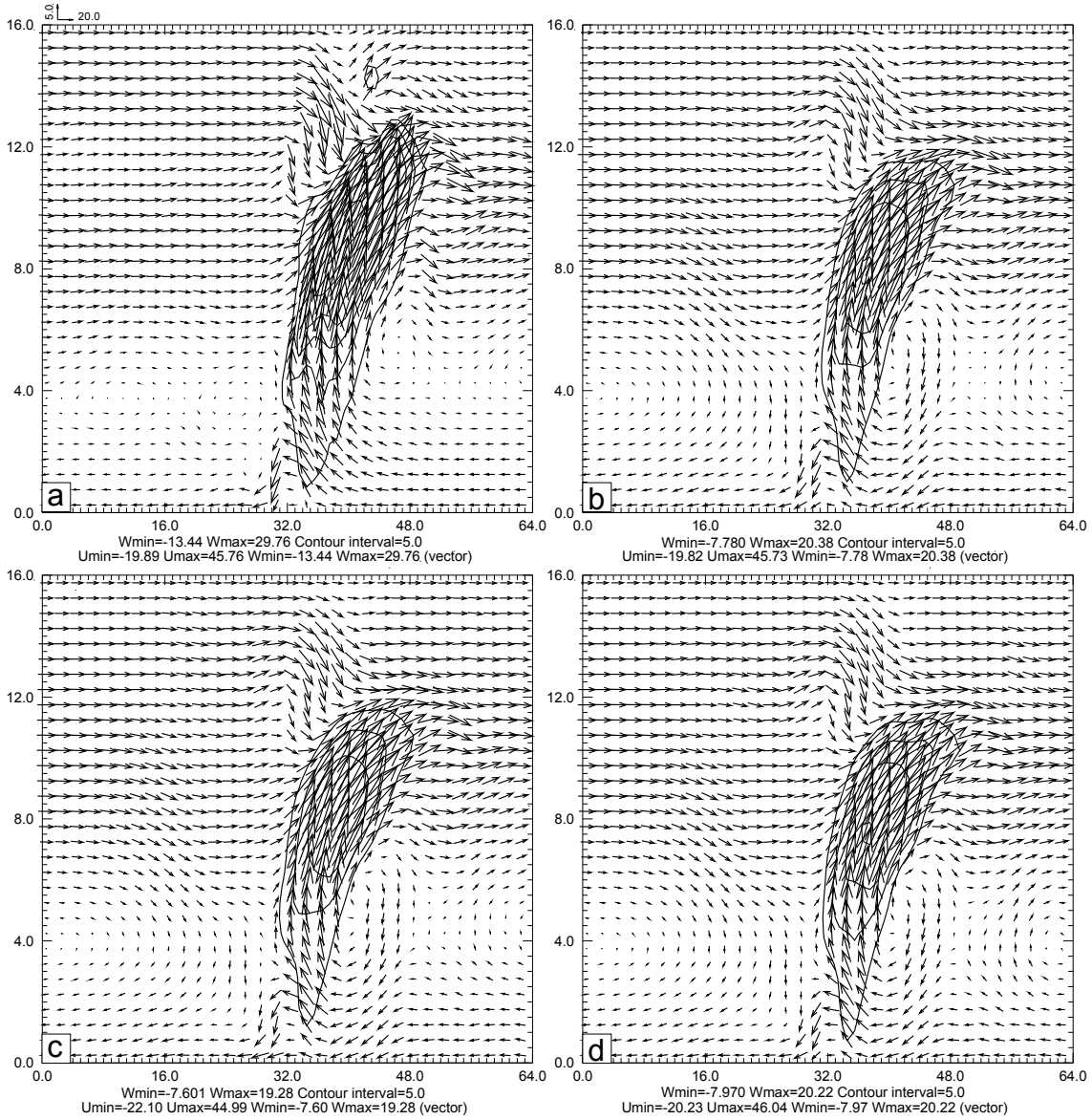

Fig. 3.2. Total u-w wind vectors and vertical velocity (contours) of the 20 May 1977 supercell storm at t=120 min and y=22.5 km (along the line A-B in Fig. 3.2a) for (a) truth simulation; (b) CNTL1_60; (c) NoBB1_60; (d) NoCV1_60.

Four series of pseudo radar radial observations from two Doppler radars are obtained by sampling the evolution of this simulated storm every 5 min from 2h to 3h using radar

forward operators expressed in eq. (3.1)-(3.7). The first series of simulated data are obtained from the simulated wind field fixed at t=2 h, as a function of various radar locations. Of the two radars, one is put at x= 33 km relative to the origin of model domain (lower left corner), while its y coordinate is varied in increments of 10 km from y=-190 km to y=10 km. A second radar is set at position y=25 km while its x coordinate is varied from x=0 km to x= -200km in intervals of 10km. In this way, we are able to test the impact of the beam broadening and the earth curvature as a function of distance from the center of the storm ranging from about 20 km to 220 km. The center of the storm is estimated to be (32.5km, 22.5km). The second series of pseudo observations are sampled in a similar way to the first one, except that the refractive index gradient dn/dh is no longer -1/(4a) (about -39.2×10^{-6} km^{-1}) in Eq. (3.1) for the "four-thirds earth radius model". Instead, the dn/dh takes the value of -10×10^{-6} km^{-1}, -70×10^{-6} km^{-1}, -100×10^{-6} km^{-1}, -130×10^{-6} km^{-1} respectively, representing most possible cases in the atmosphere.

The third series of radial velocity observations are obtained every 5 minutes from model simulation between 2h and 3 h using the same forward operator, but two radars are at fixed locations (33km, -40km) and (-30 km, 25km) respectively. In this case, the surface range between the storm center and either of the radars is about 60 km. The fourth series of pseudo observations are sampled in a similar way to the third series, for two radars at fixed locations (33km, -130km) and (-120km, 25km). In this case, the distance between the storm center and either of the radars is about 150 km.

The elapsed times for the radars to obtain the volume scans are neglected, and thus we assume that the radial wind observations are simultaneous. For simplicity, the two radars will cover the entire horizontal physical grids (i.e. 64×64 km) which assumes that

the radars sweep almost continuously in horizontal direction. The elevation angles are $0.5^{\circ}, 0.9^{\circ}, 1.3^{\circ}, 2.4^{\circ}, 3.1^{\circ}, 4.0^{\circ}, 5.1^{\circ}, 6.4^{\circ}, 7.5^{\circ}, 8.7^{\circ}, 10.0^{\circ}, 12.0^{\circ}, 16.7^{\circ}, 19.5^{\circ}$ (same as the WSR-88D convective precipitation volume coverage pattern (VCP) 11). The simulated data are only specified in precipitation regions (where reflectivity is greater than zero dBZ). In order to simulate the radar measurement statistical error, 1m s^{-1} random error (white noise) is added to the radial velocities in the pseudo observation data.

Corresponding to the first series of radial wind observations, three categories, 21 experiments each category, of data analysis experiments (see Table 3.1, which lists all experiments) will be conducted at t=2 h with varied surface ranges between radar location and storm center. In the first category of experiments, both the effect of beam broadening and the effects of earth curvature are considered using the radar forward observation operator as defined in Eqs. 1-7. They will be referred as CNTL1 experiments (label 1 means at single time level). In the second category of experiments, the effect of beam broadening is not considered and Eq. (3.7) will be replaced with a simple tri-linear interpolation scheme. It will be referred as NoBB1 experiments. In the third category of experiments, the effect of earth curvature will not be considered and Eq. (3.3) ~ (3.6) will be replaced with the commonly used Cartesian radar forward operator (Gao et al. 1999). It will be referred as NoCV1 experiments. The distance between the storm and the radar varies from 20 km to 220 km at an interval of 10 km for both radars. Therefore, each individual experiment will be referred by its category name followed by the distance in km, as described above, e.g. CNTL1_60, NoBB1_60, NoCV1_60, etc. Corresponding to the second series of pseudo observations, four categories, twenty-one experiments each category, of data analysis experiments are performed (see Table 3.1). The settings are

similar to that in CNTL1 experiments except that the refractive index gradient dn/dh is no longer -1/(4a). The four categories of experiments are named DnDh-10, DnDh-70, DnDh-100, and DnDh-130, respectively, according to the value of dn/dh used.

Table 3.1. List of data analysis/assimilation experiments

<table>
<tr><th>Name[a]</th><th>Radar distance</th><th>Description</th></tr>
<tr><td>CNTL1_xxx</td><td rowspan="3">20km~220km at an interval of 10km
(xxx is the radar distance in km)</td><td rowspan="3">one-time analyses at t=2h
(21 experiments for each type)</td></tr>
<tr><td>NoBB1_xxx</td></tr>
<tr><td>NoCV1_xxx</td></tr>
<tr><td>CNTLM_60</td><td rowspan="3">60km</td><td rowspan="6">One hour assimilation from t=2h~3h
at an interval of 5min</td></tr>
<tr><td>NoBBM_60</td></tr>
<tr><td>NoCVM_60</td></tr>
<tr><td>CNTLM_150</td><td rowspan="3">150km</td></tr>
<tr><td>NoBBM_150</td></tr>
<tr><td>NoCVM_150</td></tr>
<tr><td>DnDhxxx</td><td>20km~220km at an interval of 10km
(xxx is the dn/dh value)</td><td>one-time analyses at t=2h
(21 experiments for each dndhxxx experiment)</td></tr>
</table>

[a]CNTL means both the effects of beam broadening and earth curvature are considered;
NoBB means the effect of beam broadening are neglected;
NoCV means the effect of earth curvature are neglected.

Corresponding to the third series of pseudo observations, three intermittent data assimilation experiments (see Table 3.1) are performed with an interval of 5 minutes and a window covering t=2 h to t=3 h of the model simulation. For these three experiments, the distance from the radar to the storm center is about 60 km when the data assimilation experiments begin. These three experiments are referred as CNTLM_60, NoBBM_60, NoCVM_60 experiments with similar literal meaning as the above (where the label M is added to denote multiple time levels). Corresponding to the fourth series of pseudo

observations, three more intermittent data assimilation experiments (see Table 3.1) are performed. The setting is same as above, but the distance between radar location and storm center is changed to 150 km at the beginning of data assimilation. Similarly, these three more experiments are named CNTLM_150, NoBBM_150, and NoCVM_150. These six experiments are designed to assess the impact of the beam broadening and the earth curvature on radar data assimilation over a data assimilation window while radar sites are near, or far from a storm. There are 13 assimilation cycles with 5 minute interval in these 6 experiments. The ARPS 3DVAR system is used to obtain the model initial condition first, and then the ARPS system runs for a five-minute forecast starting from this initial analysis. This intermittent assimilation cycle is applied every five minutes until the end of assimilation period.

To compare the accuracy of the analysis from different experiments, the RMS error statistics of the horizontal wind components (V_h) and scalar model variables (s) between the experiments and the truth simulation run are computed using the following equations:

$$RMS\_Vh = \sqrt{\frac{\sum_{i=1}^{N}(u-u_{simu})_i^2 + \sum_{i=1}^{N}(v-v_{simu})_i^2}{2N}} \tag{3.10}$$

$$RMS\_s = \sqrt{\frac{\sum_{i=1}^{N}(s-s_{simu})_i^2}{N}} \tag{3.11}$$

where N is the total number of 3-dimensional grid points used in the calculation, and the subscript $simu$ stands for the data from the simulation run. The computation of the RMS error statistics is only done over model grid points where the reflectivity (estimated from the local hydrometeor mixing ratios) of the simulation run is greater than 5 dBZ.

3.5 Results of experiments

3.5.1 The impact on 3DVAR wind analysis at t=2h time level

As stated above, the purpose of first series of experiments is to test the impact of beam broadening and earth curvature on 3DVAR wind analysis at a single time level. The variations of RMS errors for NoBB1 and NoCV1 are plotted in Fig. 3.3 along with that for CNTL1. The horizontal section at z=3.5km AGL and the vertical cross section at y=22.5km of wind fields for the truth simulation, CNTL1_60, NoBB1_60, NoCV1_60 and CNTL1_150, NoBB1_150, NoCV1_150 are plotted in Fig. 3.1, 3.2, 3.4 and 3.5.

We first discuss the impact of beam broadening. The RMS error of the horizontal winds and the vertical velocities plotted as a function of the distance for both CNTL1 (solid lines) and NoBB1 (dashed lines) experiments are shown in Fig. 3.3. It is found that the RMS error differences for both horizontal winds and vertical velocities between these 21 CNTL1 experiments and their corresponding NoBB1 experiments gradually increase as the distance between the storm center and radar locations increase. These differences are less than 0.35 m s^{-1} for horizontal winds and less than 0.1 m s^{-1} for vertical velocities within the range of 60 km. Beyond 60 km, the differences for horizontal winds become more noticeable as the range increases, reaching over 1 m s^{-1} at the range of 220 km, while the difference for vertical velocity shows little change. This means that additional errors due to the neglect of beam broadening are gradually introduced in NoBB1 experiments.

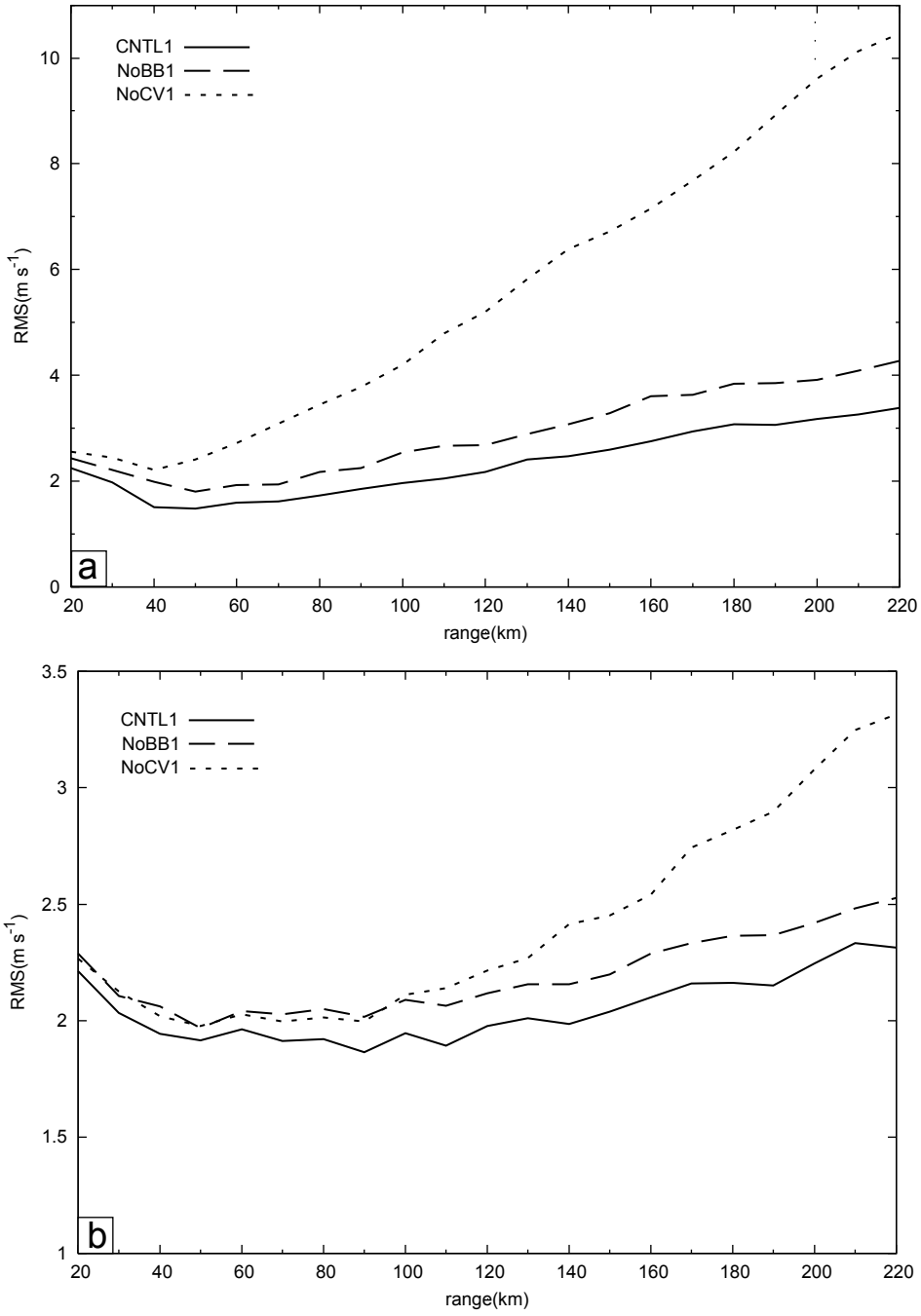

Fig. 3.3. The variation of RMS errors with the distance between the center of the storm and radar locations, for (a) horizontal wind components, and (b) vertical velocity. The solid lines are for CNTL1 experiments, the dashed lines are for the NoBB1 experiments, and the dotted lines are for the NoCV1 experiments.

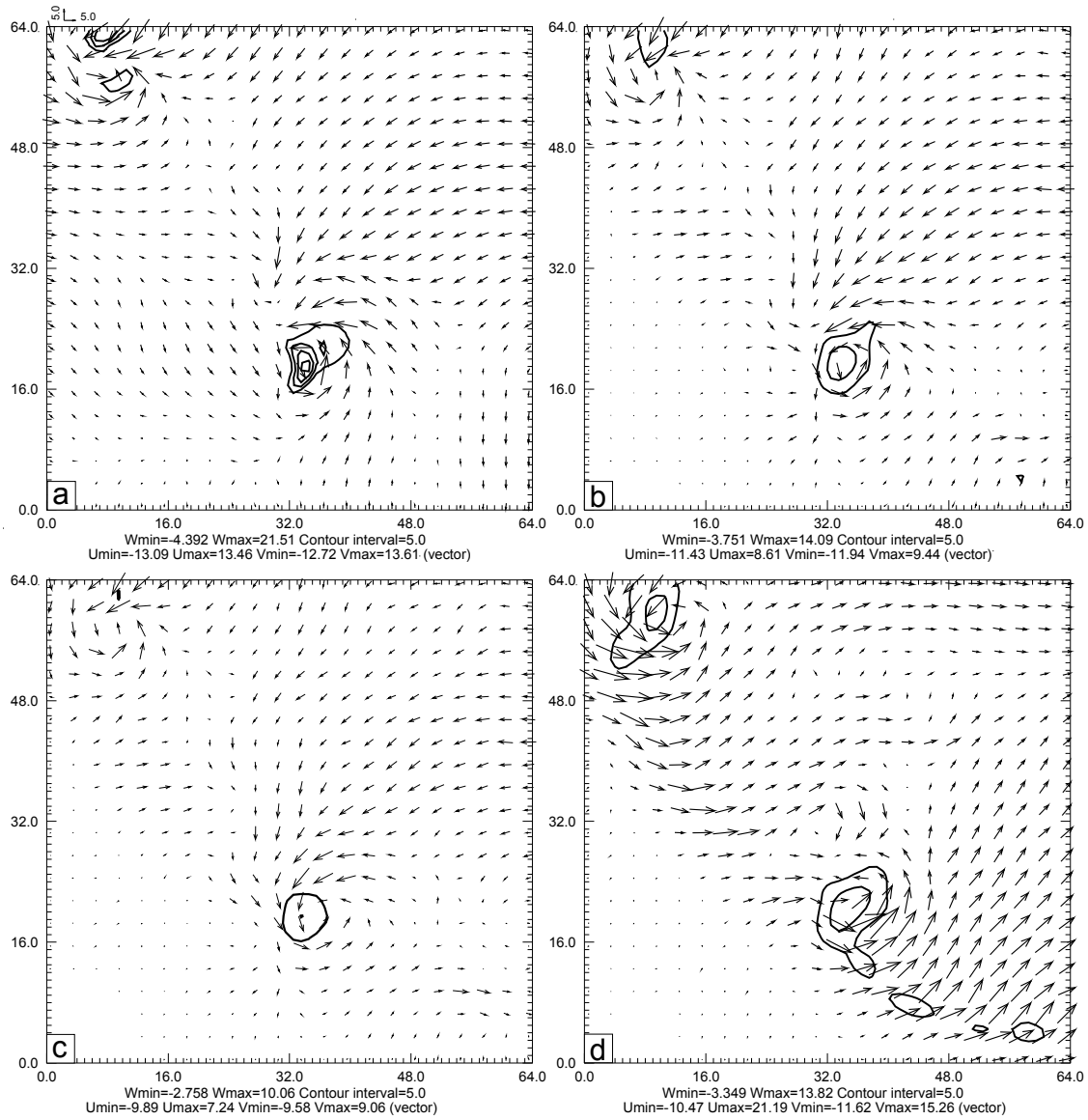

Fig. 3.4. Same as Fig. 3.1, but for (a) truth simulation; (b) CNTL1_150; (c) NoBB1_150; (d) NoCV1_150.

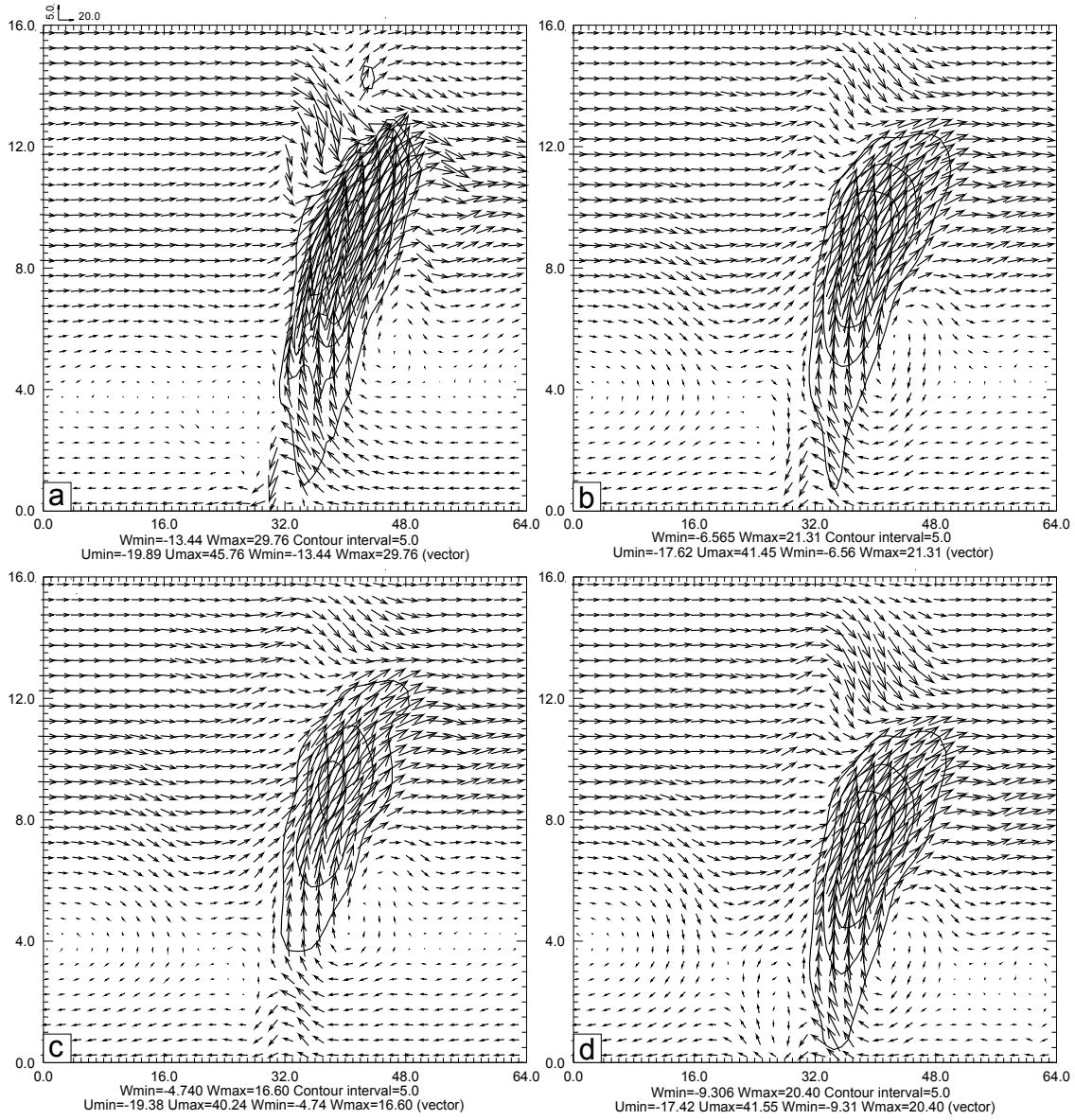

Fig. 3.5. Same as Fig 3.2, but for (a) truth simulation; (b) CNTL1_150; (c) NoBB1_150; (d) NoCV1_150.

The variation in the RMS errors for horizontal winds and vertical velocities as a function of distance for experiment NoCV1 is also plotted in Fig. 3.3 in dotted lines. It is easily identified that the neglecting of the earth curvature can lead to very large RMS errors in the analysis of horizontal winds, especially beyond 60 km. It exhibits an additional 7.1 m s^{-1} RMS error of horizontal winds compared to CNTL1 experiment at the range of 220 km (Fig. 3.3a). The RMS error differences for vertical velocities between CNTL1 and NoCV1 experiments are evident when the surface range is over 150 km (Fig. 3.3b). Therefore, in the sense of the evolution of RMS errors, we can conclude that overlooking the earth curvature has a much greater negative impact on variational wind analysis than the neglect of beam broadening.

As the RMS statistics suggest, the differences in the 3-D wind fields among all three categories of experiments CNTL1, NoBB1 and NoCV1 should be very small when the distance between the storm and radars is less than 60 km. Fig. 3.1 and Fig. 3.2 confirm this conclusion. Fig. 3.1 shows that the horizontal wind and vertical velocity fields at 3.5 km AGL for the truth simulation and the three experiments, CNTL1_60, NoBB1_60, and NoCV1_60, where the radar is 60km from the storm. Though the 3DVAR analysis is not perfect, the horizontal cyclonic rotation associated with the right and left movers are evident in all three experiments (Fig. 3.1b, c, d). They are all pretty close to the truth simulation (Fig. 3.1a). The analyzed maximum vertical velocities (Fig. 3.2b, c, d) for all three categories of experiments are generally several meters per second weaker than the truth simulation, but the pattern is nearly the same for all three experiments. So the error from neglecting both beam broadening and earth curvature at this range is pretty small.

When the distance between the storm and radar location is 150 km or greater, the differences among these experiments become larger and can no longer be ignored. As an example, horizontal cross sections at z=3.5 km and vertical cross sections are plotted as in Fig. 3.4 and Fig. 3.5 for the surface range of 150 km (the truth simulation is re-plotted for ease of comparison). It is clearly evident that the rotation signature near the center of the storm in Fig. 3.4b for CNTL1_150 is stronger than that in Fig. 3.4c for NoBB1_150. In addition, Fig. 3.5b shows a much stronger and deeper rotation updraft than Fig. 3.5c. The maximum vertical velocity in Fig. 3.5b is 21.31 m s^{-1}, much closer to the simulation result (as shown in Fig. 3.5a) than that in Fig. 3.5b which is only 16.60 m s^{-1}. Apparently, CNTL1_150 experiment does a better job for the wind analysis than NoBB1_150 in which no effect of beam broadening is considered.

For experiment NoCV1_150 in which the influence of the earth's curvature is not considered, Fig. 3.4d shows that the perturbation horizontal winds are unexpectedly strong and quite noisy. The signatures of cyclonic rotation within each of the cells are not so well analyzed. Although the strength of the major updraft in Fig. 3.5d is well captured, just as in Fig. 3.5b of CNTL1_150, the updraft in Fig. 3.5d is incorrectly positioned in the vertical direction, about 1 km below than that in Fig. 3.5a. All these distorted features are evidently caused by the neglect of the effect of the earth curvature in the radar forward observation operator.

It should be noted that the wind analysis generally becomes worse even in CNTL1_150km experiment because of the poorer resolution in the data at that distance.

It is demonstrated that the impacts of both the beam broadening and earth curvature are dependent on the surface range between the center of the storm and the radar location.

It appears that within a range of 60 km, both the impacts of beam broadening and earth curvature can be neglected. As the distance increases beyond 60 km, more and more additional errors are introduced into the wind analysis from both earth curvature and beam broadening effects. Specifically, the neglect of the earth curvature exhibits much more negative impact than the neglect of the beam broadening. When the distance to the storm exceeds 150 km, overlooking the earth curvature and the beam broadening will both bring much more obvious negative impact on the 3-dimensinal wind analysis. So the Cartesian ray path equation and a simple interpolation are not recommended when the distance to the storm is greater than 150 km.

The "four –thirds earth radius model" assumes that dn/dh equals to -1/(4a), about -39.2×10^{-6} km^{-1}, for the standard atmosphere. However, dn/dh can deviate from this value more than 100×10^{-6} km^{-1} in storm favoring environments (Gao et al., 2006). The impact of refractivity gradient with different dn/dh is investigated now through four categories of data analysis experiments as described in previous section (Section 3.4). The RMS error of horizontal wind for the experiments CNTL1, DnDh-10, DnDh-70, DnDh-100, and DnDh-130 are plotted in Fig. 3.6. It is shown that the impact of using different values of dn/dh instead of -1/(4a) for standard atmosphere is rather small. The additional RMS error due to the use of the "fourth-thirds earth radius model" is generally less than 0.42 m s^{-1} within a range of 100 km. The impact will gradually increase as the range increases. When the radar is very far from the storm (beyond 190km) and the absolute value of dn/dh is very large (more than 130×10^{-6} km^{-1}), the additional RMS error is over 1 m s^{-1}. The additional RMS error of vertical velocity (not shown) introduced by the use of dn/dh=-1/(4a) are all less than 0.2 m s^{-1}. The 3D wind plots (not shown) also confirmed

that above statements. It is concluded that the impact of refractive index can be neglected for most applications. Because of the impact is so small, this effect will not be discussed in the following intermittent data assimilation experiments.

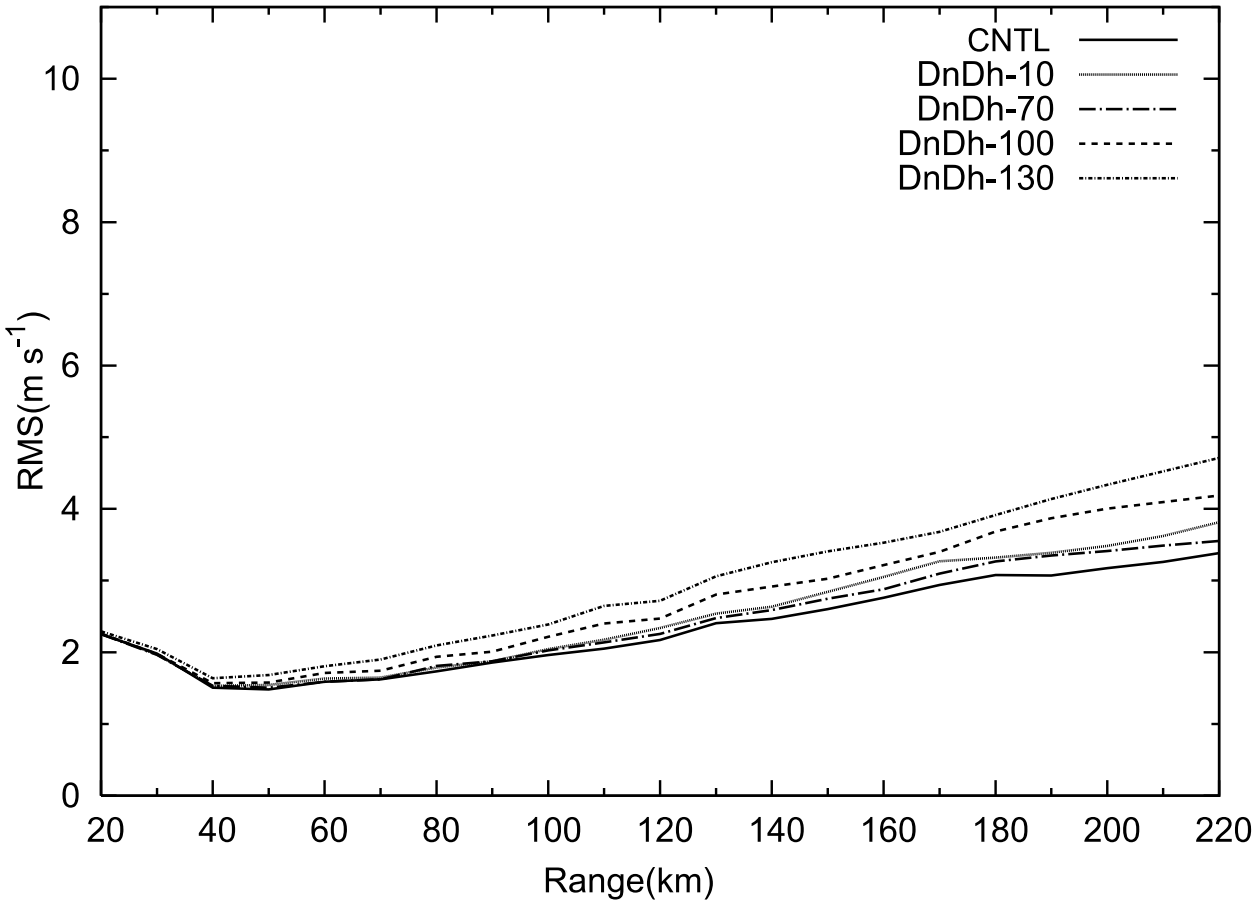

Fig. 3.6. The variation of RMS errors with the distance between the center of the storm and radar locations for horizontal wind. The solid lines are for CNTL1 experiments, the dot lines are for DnDh-10 experiments, the dot-dashed lines are for DnDh-70 experiments, the dashed lines are for DnDh-100 experiments, the short dot-dashed lines are for the DnDh-130 experiments.

3.5.2 The impact on radar data assimilation cycles

To investigate how the errors introduced by neglecting of the beam broadening and the earth curvature are accumulated during an intermittent data assimilation and

investigate how the retrievals of other model variables, such as potential temperature, moisture are impacted, two time series of data assimilation with 5 minutes interval are performed during a one-hour-long data assimilation period. As discussed in Section 3.4, first three intermittent data assimilation experiments referred as CNTLM_60, NoBBM_60, NoCVM_60 are conducted using data sampled from t=120 min to t=180 min of model simulation with a radar distance of 60 km when the data assimilation experiments begin. Three more experiments CNTLM_150, NoBBM_150, NoCVM_150 are conducted for the radar distance of 150 km at the beginning of data assimilation. The results from these six experiments are discussed in the following.

Fig. 3.7 shows the horizontal winds, perturbation potential temperature and reflectivity at 250 m AGL (first model level above surface) and Fig. 3.8 shows the horizontal wind and vertical velocity fields at 3.5 km AGL, at 140, 155 and 170 min of model time. Recall that the model assimilation begins at t=120 min. They are shown for the truth simulation, cycled 3DVAR assimilation for experiments CNTLM_60, NoBBM_60 and No_CVM_60, as described in above. For all three experiments, Fig. 3.7d, g, j show that after 4 cycles at t = 140 min, the assimilation has retrieved some weak potential temperature perturbations. Though no reflectivity is assimilated, the model established the reflectivity pattern quite similar to the truth simulation, although covering a smaller area after 20 minutes of assimilation. A small positive temperature perturbation is found where there should be cooling (Fig. 3.7d, g, j). At the 3.5 km level (Fig. 3.8d, g, j), an updraft is established well at the correct location, and its strength and structure are quite similar to the truth (Fig. 3.8a). After three more

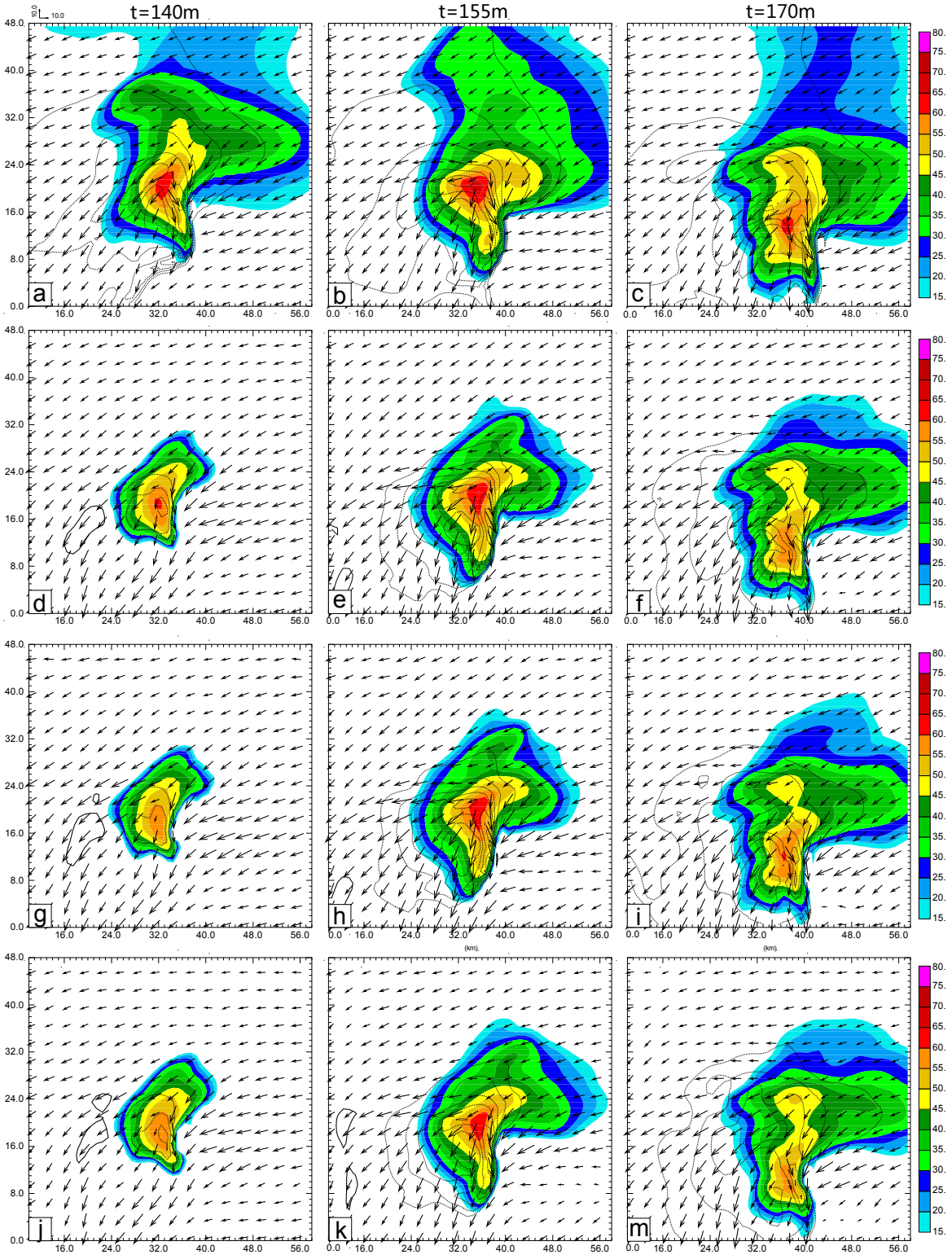

Fig. 3.7. The total u-v wind vector, perturbation potential temperature (contour at every 1K) and reflectivity (colored) at z=250m AGL and t=140min, 155min, 170min respectively. (a), (b), (c) are for truth simulation, (d), (e), (f) are for CNTLM_60, (g), (h), (i) are for NoBBM_60, (j), (k), (m) are for NoCVM_60. Solid contour for positive, and dashed contour for negative.

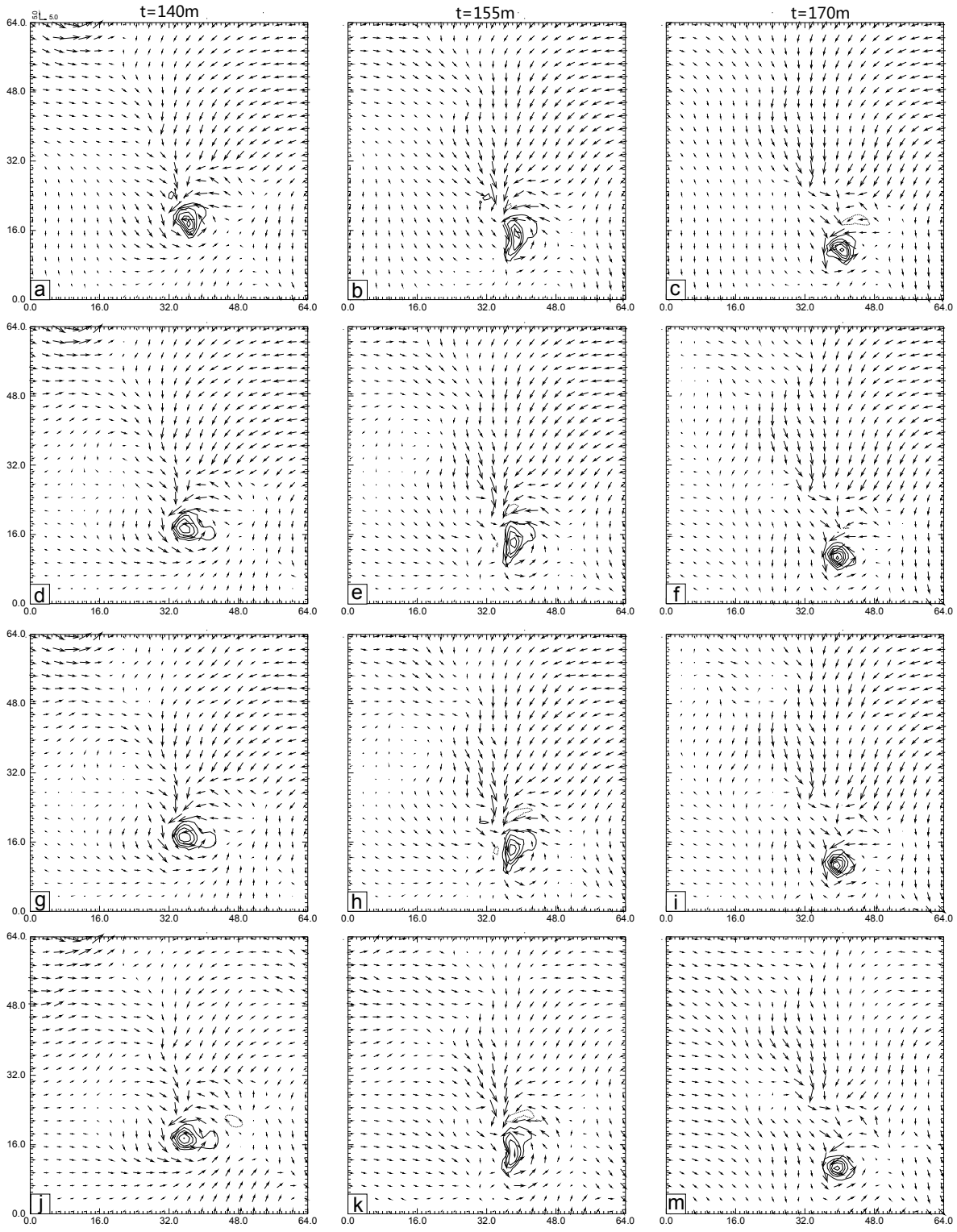

Fig. 3.8. The perturbation u-v wind vector, vertical velocity (contour at every 5 m s^{-1}) at z=3.5km AGL and t=140min, 155min, 170min respectively. (a), (b), (c) are for truth simulation; (d), (e), (f) are for CNTLM_60; (g), (h), (i) are for NoBBM_60; (j), (k), (m) are for NoCVM_60. Solid contour for positive, and dashed contour for negative.

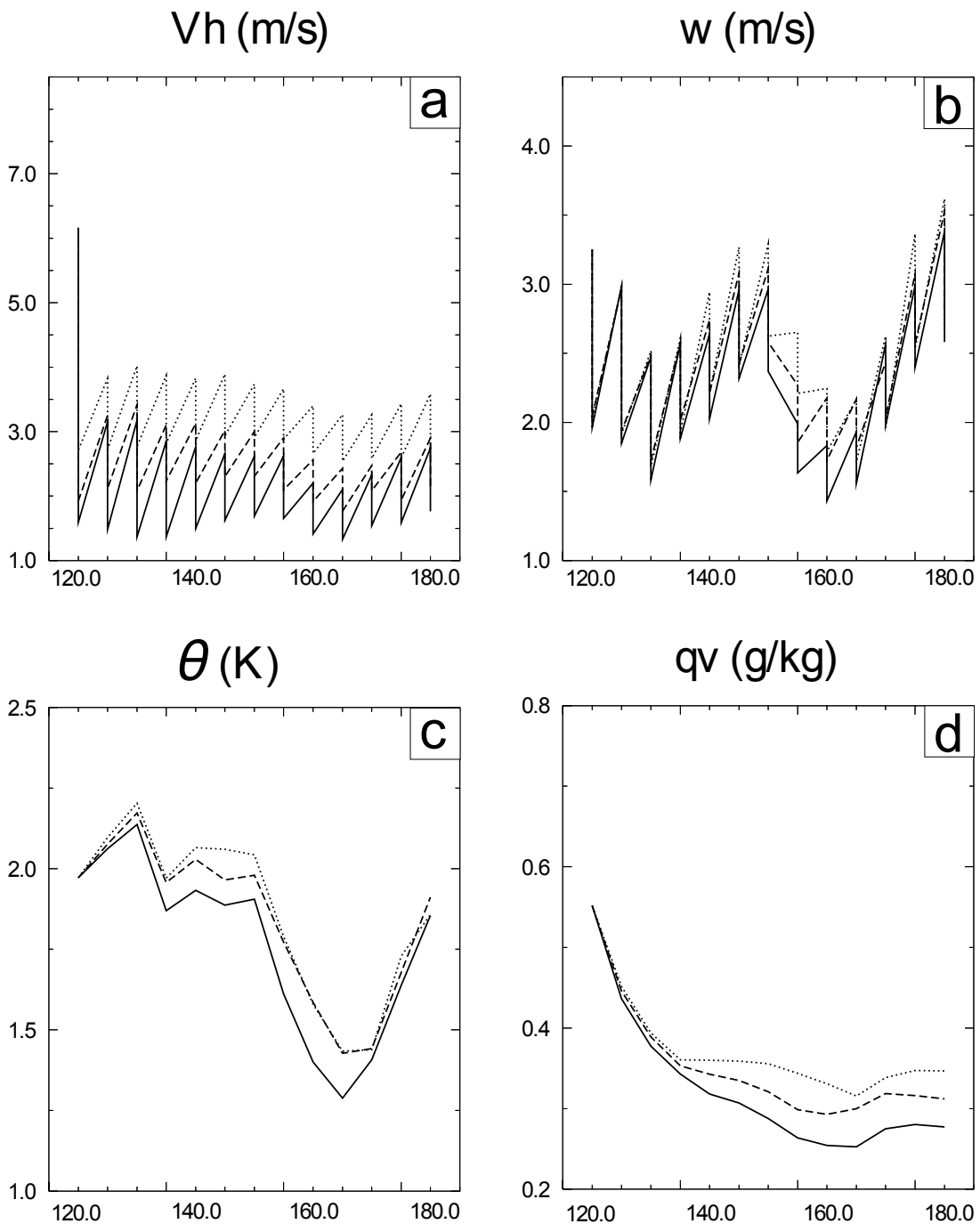

Fig. 3.9. The evolution of the RMS errors with time for different model variables. The solid lines are for CNTLM_60, the dashed lines are for NoBBM_60, and the dotted lines are for NoCVM_60.

analysis cycles at t = 155 min, the low-level flow immediately underneath the storm cells becomes closer to the truth (Fig. 3.7e, h, k vs. Fig. 3.7b) but the area of outflow and cold pool on the southwest side remain smaller than the truth. At the 3.5 km level, the perturbation horizontal winds and the updrafts are well captured in all three experiments by t= 155 min (Fig. 3.8e, h, k vs. Fig. 3.8b).

By t = 170 min, the analysis is further improved. In fact, by this time, there are no significant differences from the truth in either the low-level and mid-level fields (Fig. 3.7f, i, m and Fig. 3.8f, i, m). General storm structures including the precipitation pattern are well retrieved during this 1h data assimilation in all three experiments though the results from NoCVM_60 are not quite as good. This reinforces that the impacts of beam broadening and earth curvature on radar data assimilation cycles for retrieving other model variables from the radial wind of two radars are generally small when the storm is not far from two radars.

Although the RMS error is generally not well suited as a verification metric for storm-scale phenomena, we use it here for comparison among different experiments while also visually comparing plotted fields to verify the result. The RMS errors for several analyzed fields are shown in Fig. 3.9. The RMS errors for V_h components decrease with time, but very slowly. The variations of RMS errors for w are not stable, possibly because of small phase or position errors. The RMS errors for θ' decrease for the first 40 minutes of assimilation, then increase with time again. Only the errors for q_v decrease nearly monotonically with time. The q_v RMS error is reduced to 0.28 g kg^{-1} in CNTLM_60, and to 0.31, 0.35 g kg^{-1} in NoBBM_60 and NoCVM_60 respectively. Fig. 3.9 generally shows that the RMS errors of V_h, w, θ', and q_v stay very close for all three

experiments though NoCVM_60 has slightly larger errors in V_h. The RMS errors again suggest that the effect of beam broadening and earth curvature is generally small when the storm is not far from radar.

We now turn to the results for experiments CNTLM_150, NoBBM_150 and NoCVM_150. Fig. 3.10 and Fig. 3.11 show that, in general, the results are significantly worse in all three experiments than the prior 60 km experiments. The overall storm structures are poorly resolved compared to CNTLM_60, NoBBM_60 and NoCVM_60. However, among the three experiments for the range of 150 km, the overall structure of the storm for CNTLM_150 is the best and quite similar to those of the truth toward the end of the assimilation.

In experiment NoBBM_150, the precipitation area is pretty small and the cold pool is very weak at 140 min, i.e. after 20 min of assimilation (Fig. 3.10g), but the pattern of horizontal winds and strength of updraft at the 3.5 km level is similar to the truth (Fig. 3.11g vs. Fig. 3.11a). At 155 min, the analysis looks better, but both the horizontal wind and vertical velocity field look noisy, and there exist several small centers for positive, or negative contours that are not supported by the truth simulation (Fig. 3.11h vs. Fig. 3.11b). At the end of the assimilation (Fig. 3.10i and Fig. 3.11i), the reflectivity and updraft patterns look much closer to the truth. Clearly, at this very large radar distance, the neglect of beam broadening worsens the assimilation results. But the impact is limited and the internal structures of thunderstorms can still be obtained well by the end of 1h assimilation.

When the effect of earth curvature is not considered at the range of 150 km, the analyzed low-level cold pool, gust front, and precipitation pattern differ markedly from

those of the truth (Fig. 3.10j,k,m vs. Fig. 3.10a,b,c) and from the control assimilation at the same radar distance (vs. Fig. 3.10d,e,f). At t=155 min, the mid-level updraft appears broader and the pattern of horizontal flow is significantly different from the truth. At t=170 min, the reflectivity core becomes distorted and the hook echo is poorly defined after 50 min assimilation (Fig. 3.10m vs. Fig. 3.10c). Also at this time, there are a few spurious updrafts within the analysis domain (Fig. 3.11m). Overall, the analysis is significantly worsened when the effect of earth curvature is not considered at a radar distance of 150 km.

The variations of the RMS error in horizontal wind components (V_h), vertical velocity (w), perturbation potential temperature (θ) and perturbation water vapor mixing ratio (q_v) are plotted in Fig. 3.12. It is demonstrated that the RMS errors in NoBBM_150 are generally larger than that in CNTLM_150 but do not deviate much. The NoCVM_150 experiment yields the worst results with the largest RMS errors during the 1h long assimilation period among all three experiments, especially for variables V_h and qv. These error statistics also indicate that when a storm is 150 km from the radar, neglecting beam broadening worsens the results slightly while overlooking earth curvature produces significantly worse results for retrieved model variables.

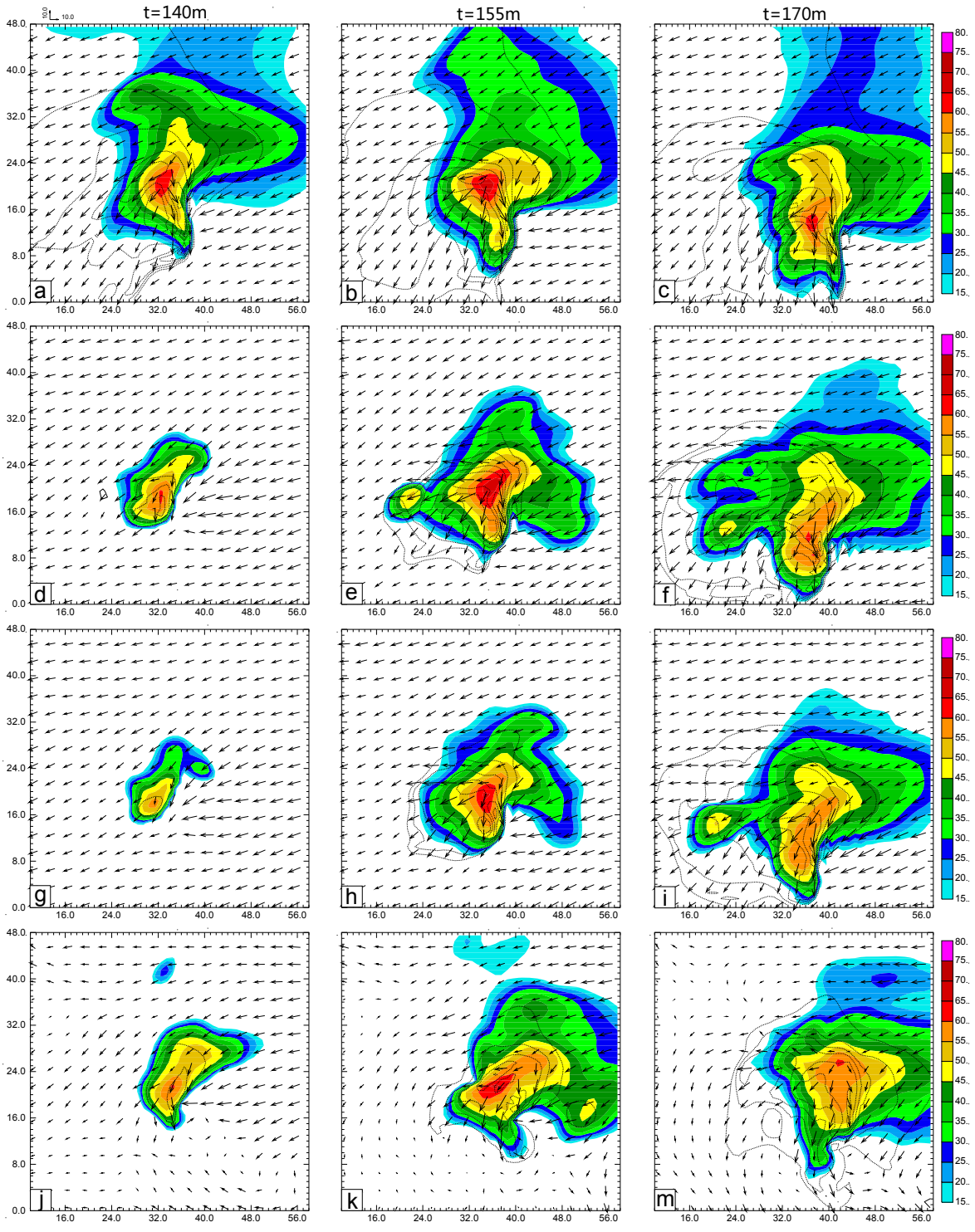

Fig. 3.10. Same as Fig. 3.7, but (a), (b), (c) are for truth simulation; (d), (e), (f) are for CNTLM_150; (g), (h), (i) are for NoBBM_150; (j), (k), (m) are for NoCVM_150. Solid contour for positive, and dashed contour for negative.

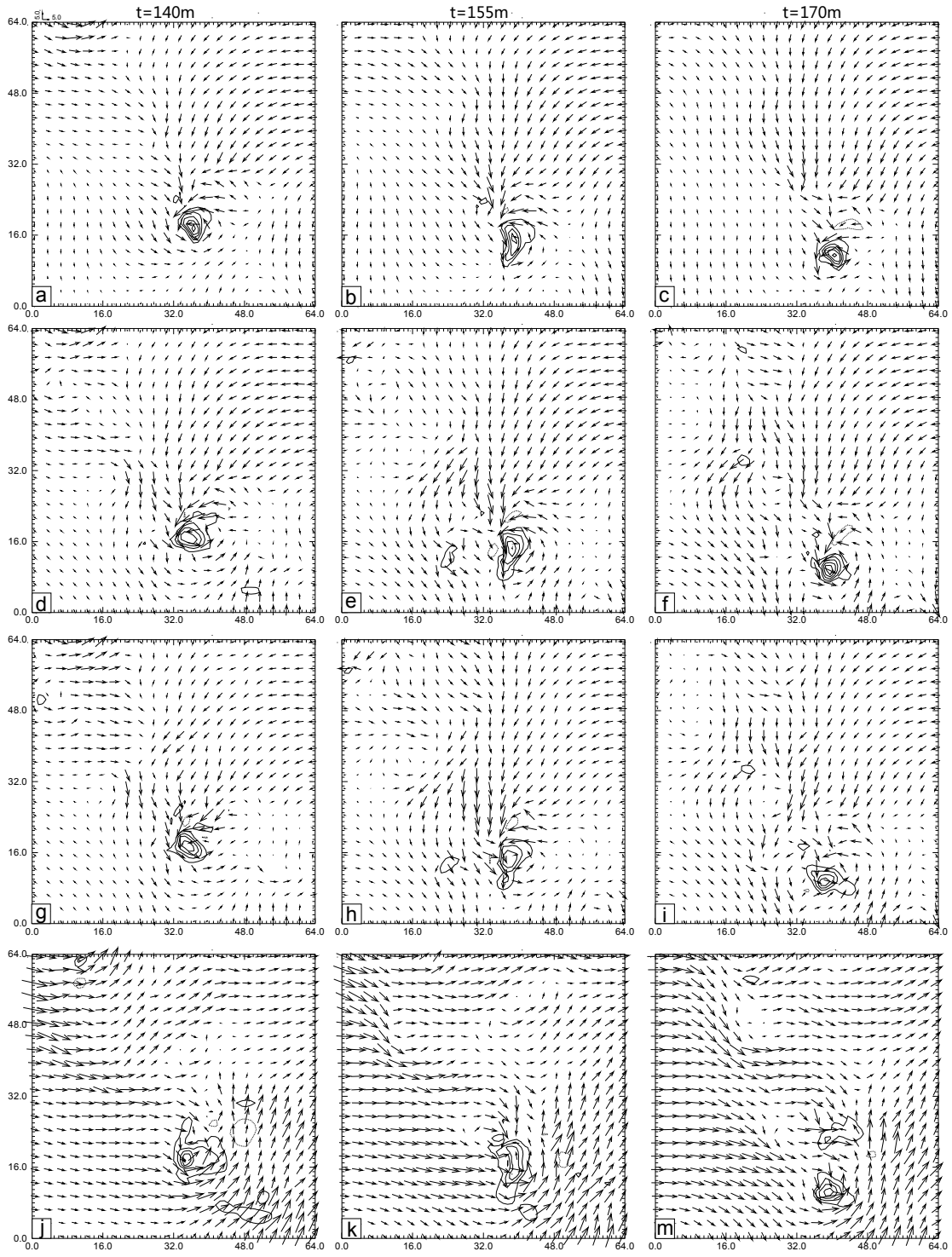

Fig. 3.11. Same as Fig. 3.8, but (a), (b), (c) are for truth simulation; (d), (e), (f) are for CNTLM_150; (g), (h), (i) are for NoBBM_150; (j), (k), (m) are for NoCVM_150. Solid contour for positive, and dashed contour for negative.

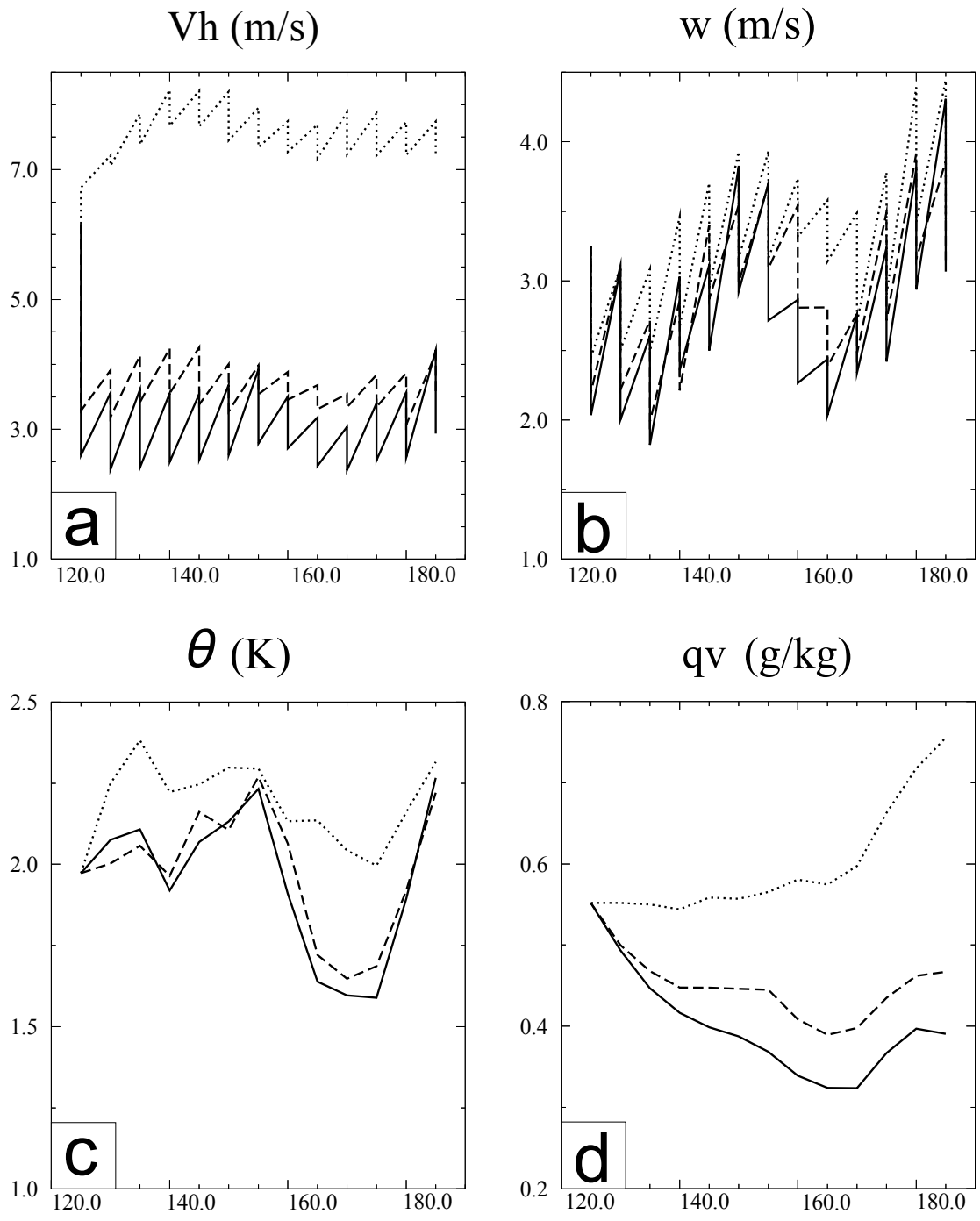

Fig. 3.12. The evolution of the RMS error for different model variables. The solid lines are for CNTLM_150, the dashed lines are for NoBBM_150, and the dotted lines are for NoCVM_150.

3.6 Summary and discussion

To utilize high-resolution radar radial velocity data in storm-scale data assimilation, it is necessary to compute the model counterpart of radial winds by converting *u, v, w* winds on model grids into radial velocity in radar coordinates. This is called the radar forward observation operator. The most accurate forward observation operator includes considering the effect of beam broadening and the earth curvature. However, this may lead to higher computational cost that could impact the lead time of a forecast system or require additional computational resources. So some past research used a very simple form of radar observation operator by neglecting the two effects mentioned above, or where better formulations were used, but the impact of that choice was not explicitly measured. In this study, we studied the effects of these assumptions on assimilating data from an idealized simulated supercell storm. It is shown that both the effects of beam broadening and earth curvature can only be neglected when the radar is near the storm, within 60 km, as demonstrated by this study.

For wind analysis at a single time, as the surface range increases, more and more additional error will be introduced into the analysis by the neglect of the two effects. The effect of beam broadening becomes evident and can cause relatively large errors for ranges at and beyond 150 km. The effect of earth curvature is very significant when the surface range is beyond 60 km due to vertical location errors in the data. The impact of refractive index gradient is also tested. It is shown that the variation of refractive index gradient has a very small impact on the wind analysis. It is acknowledged there are extreme cases where ducting and other effects can occur causing false echoes and ground

or sea clutter, but it is assumed such data will be properly screened from use in data assimilation.

In two series of one-hour-long data assimilation experiments it is shown that the impact of both effects is not significant for retrieving all model variables when the radars are relatively close to the storm (generally within 60 km). When the radars are far from the storm, not accounting for beam broadening has a rather small effect on the accuracy of assimilation results after one-hour assimilation. So the effect of beam broadening can be generally overlooked in radar data assimilation. On the other hand, ignoring the earth's curvature leads to significant errors (especially beyond 150 km) for retrieved model variables and reflectivity due to vertical location error in the data.

The results of this study may provide useful guidance for application of radar radial velocity data to storm scale diagnostic studies as well as numerical weather prediction.

Chapter 4 Storm-scale 3DVAR with diagnostic pressure equation as a weak constraint and its applications to the prediction of tornadic supercell thunderstorms

4.1 Introduction

There exist many challenges in forecasting convective storms due to their high localization. One of them is how to produce a dynamic consistent initial condition for storm-scale Numerical Weather Prediction (NWP) model. Currently, the Weather Surveillance Radar-1998 Doppler (WSR-88D) radar network is the only source to routinely provide observations that can resolve storm-scale features in a high temporal and spatial resolution. Therefore, in recent years, many studies are focused on the assimilation of the WSR-88D radar level II data into a NWP model to provide better initial condition so as to improve the forecast of the storms.

Many assimilation methods have been developed for this purpose, such as 4DVAR, 3DVAR, EnKF etc. The 4DVAR method is considered theoretically the best one. It uses the NWP model as a strong constraint and fits the model to observations at different time levels during a time window. By doing so, the best representation of the observations in the initial condition can be achieved and the initial condition is naturally dynamic consistent. Sun and Crook (1997, 1998), and Sun (2005a) had shown some encouraging results by using the 4DVAR method. However, in spite of their inspiring results, the 4DVAR method is currently in limit usage for storm-scale NWP because the tedious work to derive and maintain adjoint model and the expensive computational cost. The EnKF Method is an emerging technique, which promises to produce the similar

assimilation quality with 4DVAR but avoids the derivation of the adjoint model. Lots of experiments have been conducted using EnKF (e.g.Snyder and Zhang 2003; Zhang et al. 2004; Caya et al. 2005; Tong and Xue 2005; Meng and Zhang 2008a, 2008b; Aksoy et al. 2009; Zhang et al. 2009; Aksoy et al. 2010; Torn 2010; Dowell et al. 2011). These experiments have shown a very good potential of the EnKF method. On the other hand, EnKF is not as mature as 4DVAR, 3DVAR and its computational cost is still very intensive. Its realtime application to storm-scale NWP remains a big challenge.

The 3DVAR method is much more efficient, compared to the 4DVAR and EnKF methods, and easy to be applied to operational storm-scale data assimilation. Some studies (e.g. Bishop et al. 2001; Hu et al. 2006a; Hu et al. 2006b; Zhao et al. 2006; Hu and Xue 2007; Stensrud and Gao 2010) have successfully demonstrated the ability of the 3DVAR to assimilate radar data to predict tornadic supercell storms. The ARPS 3DVAR system and its cloud analysis package had been used to produce continental-US-scale realtime weather predictions at a high 1km resolution (CAPS news, 5/2009, http://www.caps.ou.edu). However, despite its successful application, the 3DVAR scheme is often challenged by its sub-optimum theoretically due to its use of static isotropic background covariance structure and the lack of suitable balances among model variables. Efforts have been made to alleviate the negative impact of this drawback. Liu and Xue (2006; 2007) reported the effort to build a flow-dependent background error covariance for a 3DVAR system using an anisotropic recursive filter and demonstrated the improvement from this method in the retrieval of moisture from GPS slant-path water vapor observations. Hamill and Snyder (2000) and Wang et al. (2008b, 2008a) illustrated a direction to provide more reasonable flow-dependent time-evolving background

covariance for a 3DVAR system from an EnKF method. This immerging technique is called the hybrid data assimilation and is still at its early stage of development.

Another alternative to overcome the imbalance among model variables in a 3DVAR analysis is to develop suitable weak constraints to help spread the information from ingested observations to other model variables that is not directly linked with observations. Gao et al. (1999; 2001; 2004), Hu et al. (2006a; 2006b) and Hu and Xue (2007) incorporated the mass continuity equation into the cost function and found that this weak constraint can effectively help build more reasonable balance among the three components of wind fields. This is a good progress. However, there is still no suitable balance among the dynamic and thermodynamic fields in an analysis. Xiao et al. (2005) reported their efforts to build suitable balance between the wind fields and the thermodynamic fields in the MM5 3DVAR system using a constraint based on the Richardson equation, which combines the continuity equation, adiabatic thermodynamic equation and hydrostatic equation. This is a good attempt, however, the hydrostatic and adiabatic assumption are not applicable to storm-scale data assimilation. Some other 3DVAR radar data analysis and data assimilation studies (Protat and Zawadzki 2000; Liou 2001; Protat et al. 2001; Weygandt et al. 2002a, 2002b; Liou et al. 2003; Zhao et al. 2006, 2008; Liou and Chang 2009) turned to adopt a two-step thermodynamic retrieval technique to derive the temperature and pressure fields from already recovered wind fields. It is expected that this method can improve the balance among model fields and hence lead to a better forecast. This technique were pioneered by Gal-chen (1978) and Hane and Scott (1978). Since then, lots of researches (e.g. Gal-chen and Hane 1981; Hane et al. 1981; Brandes 1984; Gal-Chen and Kropfli 1984; Hane and Ray 1985; Roux

1985; Lin et al. 1986; Roux 1988; Roux and Sun 1990; Sun and Houze 1992; Lin et al. 1993; Roux et al. 1993; Shapiro and Lazarus 1993; Crook 1994; Crook and Tuttle 1994; Sun and Crook 1996; Protat and Zawadzki 2000; Liou 2001; Protat et al. 2001; Yu et al. 2001; Liou et al. 2003; Liu et al. 2005; Liou and Chang 2009) contributed to test, understand and improve this technique. However, the application of this method into storm-scale data assimilation and numerical weather prediction remains a big problem. The difficulty lies mainly in how to estimate wind tendency terms in model momentum equations. The different time differential schemes and different elapsed times between two wind observations significantly affect the accuracy of the wind tendency calculation(Crook 1994).

In this study, the calculation of the wind tendency term will be avoided by applying the divergence operator to the three model momentum equations. The derived equation is called the diagnostic pressure equation. This diagnostic pressure equation will be incorporated into the 3DVAR cost function in the form of a weak constraint in addition to the aforementioned mass continuity equation constraint. The main goal of this diagnostic divergence constraint is to help improve dynamic consistency between the dynamic model fields and thermodynamic model fields. Xu et al. (2001a) tried to include a similar diagnostic pressure equation constraint in their simple adjoint system for three-dimensional wind retrieval from single-Doppler radar by treating the radial velocity as a tracer. They found that the diagnostic pressure equation constraint can help improve the retrieval of wind fields in single time data analysis. The results are encouraging. However, the impact of the diagnostic pressure equation constraint on intermittent data assimilation and the following forecast has not been investigated.

In this Chapter, we will discuss the development of the diagnostic pressure equation constraint under ARPS 3DVAR framework and its applications to tornadic supercell thunderstorms. The following part is organized as follows. Section 4.2 will discuss the schemes adopted by the 3DVAR system and focus on the development of the diagnostic pressure equation constraint. Section 4.4 examines the impact of the diagnostic pressure equation constraint on storm-scale data assimilation using an idealized supercell thunderstorm. Section 4.4 and Section 4.5 presents the results from the application of the updated 3DVAR scheme to the 8 May 2003 Oklahoma City tornadic supercell thunderstorm case and the 5 May 2007 Greensburg tornadic supercell thunderstorm case. Summary and future work will be discussed in Section 4.6.

4.2 The scheme for the 3DVAR system

A 3DVAR system within the ARPS framework (Xue et al. 2000; Xue et al. 2001; Xue et al. 2003) has been developed and applied to the assimilation of WSR-88D radar data and other data (Gao et al. 1999; Gao et al. 2004; Hu et al. 2006a; Hu et al. 2006b; Hu and Xue 2007; Stensrud and Gao 2010). The system consists of two components, one is the 3DVAR subsystem, which is to assimilate radar radial velocity data as well as other conventional observations in a three dimensional variational framework; the other one is the cloud analysis subsystem which is to assimilate the radar reflectivity data, satellite data, etc, based mainly on semi-empirical rules. The cloud analysis system not only updates the hydrometeor fields, but also adjusts in-cloud temperature field and water vapor field according to users' parameter setting.

4.2.1 The 3DVAR subsystem

In the 3DVAR subsystem, the cost function, J, is written as the sum of the background term and the observational term plus a penalty or equation constraint term (J_c):

$$J(x) = J_b + J_o + J_c = \frac{1}{2}(x - x^b)^T B^{-1}((x - x^b) + \frac{1}{2}\left[H(x) - x^o\right]^T R^{-1}\left[H(x) - y^o\right] + J_c \quad (4.1)$$

Following the standard notion of Ide et al (1997), x and x^b are the analysis and background state vectors, and y^o is the observation vector. B and R are the background and observation error covariance matrices respectively. H(x) is the observation operator. To improve the conditioning of the cost function minimization and avoid the need for the inverse of B, a new control variable v is introduced, which is related to the analysis increment according to

$$\mathbf{x} - \mathbf{x}^b = \mathbf{B}^{1/2}\mathbf{v} \quad (4.2)$$

In terms of v, the background term becomes,

$$J_b = (1/2)\mathbf{v}^T\mathbf{v} \quad (4.3)$$

Consequently, the minimization is performed in the space of v. The recursive filter proposed by Purser et al. (2003a, 2003b) is used to model the effect of the background error covariance, or more precisely the square root of B. Currently in our 3DVAR subsystem, the background is provided by a previous ARPS model forecast, or other large scale models' forecast. The observations include Doppler radar radial velocity, single-level surface data (such as Mesonet), and multiple-level observations (such as rawinsondes and wind profilers).

Term J_c in (1) includes any penalty or equation constraint terms that serve the important role of correlating the desired analysis variables. Currently it includes two terms as defined in the following,

$$J_c = P(x)^T A_P^{-1} P(x) + Q(x)^T A_Q^{-1} Q(x) \tag{4.4}$$

The first term on the right hand side (R.H.S) of equation (4.4) is the diagnostic pressure equation constraint in which,

$$\begin{aligned} P &\equiv \nabla \cdot \vec{E} \equiv -\nabla^2 p' - \nabla \cdot (\bar{\rho}\vec{V} \cdot \nabla \vec{V}) \\ &+ g \frac{\partial}{\partial z}\left(\bar{\rho}\left[\frac{\theta'}{\bar{\theta}} - \frac{p'}{\bar{\rho} c_s^2} + \frac{q_v'}{\varepsilon + \bar{q}_v} - \frac{q_v' + q_{liquid+ice}}{1 + \bar{q}_v} \right] \right) + \nabla \cdot \vec{C} + \nabla \cdot \vec{D} \end{aligned} \tag{4.5}$$

Where,

$$\vec{E} = \frac{\partial\left(\overline{\rho}\overline{V}\right)}{\partial t} = \vec{i}\frac{\partial\left(\overline{\rho}u\right)}{\partial t} + \vec{j}\frac{\partial\left(\overline{\rho}v\right)}{\partial t} + \vec{k}\frac{\partial\left(\overline{\rho}w\right)}{\partial t} \tag{4.6}$$

$$\vec{V} = \vec{i}u + \vec{j}v + \vec{k}w \tag{4.7}$$

$$\vec{C} = \vec{i}\left(\overline{\rho} f v - \overline{\rho}\widetilde{f} w\right) + \vec{j}\left(\overline{\rho} f u\right) + \vec{k}\left(\overline{\rho}\widetilde{f} u\right) \tag{4.8}$$

$$\vec{D} = \vec{i}D_u + \vec{j}D_v + \vec{k}D_w \tag{4.9}$$

vector $\vec{E}$ is the forcing term of the vector Euclidian momentum equation. The primed variables are perturbations from a base state, c_s is the acoustic wave speed, and ε is the ratio of the gas constants for dry air and water vapor. The Coriolis coefficients $f = 2\Omega\sin(\phi)$ and $\widetilde{f} = 2\Omega\cos(\phi)$, where Ω is the angular velocity of the earth and ϕ is latitude. The terms, D_u, D_v and D_w contain the subgrid scale turbulence and computational mixing terms. Other symbols follow conventions.

Equation (4.5) is derived by applying the diverging operator to the three ARPS model momentum equations:

$$\bar{\rho}\frac{\partial u}{\partial t} = -\bar{\rho}\vec{V}\bullet\nabla u - \frac{\partial p'}{\partial x} + \left(\bar{\rho} fv - \bar{\rho}\tilde{f} w\right) + D_u \tag{4.10}$$

$$\bar{\rho}\frac{\partial v}{\partial t} = -\bar{\rho}\vec{V}\bullet\nabla v - \frac{\partial p'}{\partial y} + \bar{\rho} fu + D_v \tag{4.11}$$

$$\begin{aligned}\bar{\rho}\frac{\partial w}{\partial t} &= -\bar{\rho}\vec{V}\bullet\nabla w - \frac{\partial p'}{\partial z} \\ &+\bar{\rho}g\left[\frac{\theta'}{\bar{\theta}} - \frac{p'}{\bar{\rho}c_s^2} + \frac{q_v'}{\varepsilon + \bar{q}_v} - \frac{q_v' + q_{liquid+ice}}{1+\bar{q}_v}\right] + \bar{\rho}\tilde{f} u + D_w\end{aligned} \tag{4.12}$$

Equations (4.10)-(4.12) are the basis of the thermodynamic retrieval technique mentioned in previous section (Section 4.1). To get a reasonable storm-scale thermodynamic retrieval, it is required to get a very good estimation of the three wind tendency terms on the left hand side of equations (4.10)-(4.12), i.e. $\bar{\rho}\frac{\partial u}{\partial t}$, $\bar{\rho}\frac{\partial v}{\partial t}$, $\bar{\rho}\frac{\partial w}{\partial t}$. However, this task is often very difficult since the storm scale features change rapidly in time and the radar observations are usually taken in a relative slow pace (every five to six minutes in practice). The inaccuracy and incompleteness of the wind observations worsens the scenario. To overcome this problem and help establish some kind of balance among model variables, we incorporate the diagnostic pressure equation (4.5), which is derived from the three momentum equations (4.10)-(4.12), into the ARPS 3DVAR system in the form of a weak constraint (named as P in equation (4.4)). In this way, the calculation of wind tendency terms can be avoided.

The second term on R.H.S of equation (4.4) is intended to minimize the 3D anelastic mass divergence so as to provide the key coupling among the three wind

components. The definition and the impact of this constraint had been thoroughly investigated by Gao et al. (1999; 2004) and Hu et al. (2006b).

The two A's in Eq. (4.4) are the error covariance associated with the corresponding constraints, which are assumed to be diagonal matrices with empirically defined constant diagonal elements. They determine the relative importance of each constraint and their optimal values can be determined through many numerical experiments in a trial-and-error fashion suggested by Sun and Crook (2001) and usually should not far from the order of each term.

4.2.2 The cloud analysis subsystem

The cloud analysis subsystem is based on semi-empirical physical laws and used to derive the hydrometeor information from radar reflectivity data, satellite infrared and visible imagery data, METARs and cloud reports from surface observations from Global Observing System (GOS) of the World Meteorological Organization. The in-cloud temperature field and/or water vapor field can also be adjusted according to user's choice. More details about the ARPS cloud analysis package can be found in Zhang et al (1998), Brewster (2002), Hu et al (2006a).

4.2.3 Connection between the two subsystems

Under the context of ingesting radar data alone (radial wind and reflectivity), the analysis variables in the 3DVAR subsystem are the three wind components u, v, w; and the analysis variables in the cloud analysis subsystem can be potential temperature θ', water vapor mixing ratio qv, rain water mixing ratio qr, snow water mixing ratio qs, hail mixing ratio qh, cloud water mixing ratio qc and ice mixing ratio qi. Currently, the cloud

analysis subsystem is a follow-on step after the finish of running the 3DVAR subsystem. These two subsystems are separated from each other and there is no suitable coupling between the wind fields and the thermodynamic fields. Therefore, there may be inconsistency between the different model variables in a single analysis step. This inconsistency may harm the quality of subsequent data assimilation cycles and the ensuing forecast.

To alleviate this kind of inconsistency, we propose that the cloud analysis subsystem is done first when it is used in the assimilation runs. The results from the cloud analysis package will then be treated as pseudo observations and be ingested, as well as radar radial velocity data, by the 3DVAR subsystem. The diagnostic pressure equation constraint incorporated into the cost function will then act to help improve the balance between the dynamic and thermodynamic fields. In this way, it is expected that a more dynamically consistent analysis will be achieved and therefore, the following data assimilation cycles and the subsequent forecast will be improved.

4.2.4 The verification of adjoint codes and the behavior of cost function

The diagnostic pressure equation constraint involves the use of the three full ARPS momentum equations (excluding the wind tendency terms, hereafter). When computing the gradient part of the diagnostic pressure equation constraint, adjoint codes for the three full ARPS momentum equations is required. Developing adjoint codes is not a trivial work. However, it is lucky that the adjoint code here is only for the three full ARPS momentum equations. There is no need to considering complex microphysics processes in the model that contain strong nonlinearity. Therefore, the maintenance of the adjoint code here is relative easier than that of 4DVAR data assimilation. In practice, the adjoint

coding work is done according to the general rules proposed by Giering and Kaminski (1998).

It is very important to make sure the gradient of the cost function is computed correctly otherwise the minimization is erroneous. Similar to Wang (1993), let Z be the control vector, J(Z) be the cost function. To expand $J\left(Z+\alpha\nabla_z J\right)$ at the direction $\nabla_z J$ using the Taylor series, it can be derived that:

$$\phi(\alpha)=\frac{J\left(Z+\alpha\nabla_z J\right)-J(Z)}{\alpha\nabla_z J\cdot\nabla_z J}=1.0+\mathrm{O}\left(\alpha\left\|\nabla_z J\right\|\right) \tag{4.13}$$

For a very small α, if the gradient is computed correctly, the $\phi(\alpha)$ should take the value of one. The updated ARPS 3DVAR with the diagnostic pressure equation constraint has been verified using the above method. Fig. 4.1 shows that when α takes a small value from 10^{-5} to 10^{-15}, $\phi(\alpha)$ takes the value of one. This justifies that the gradient calculation is correct.

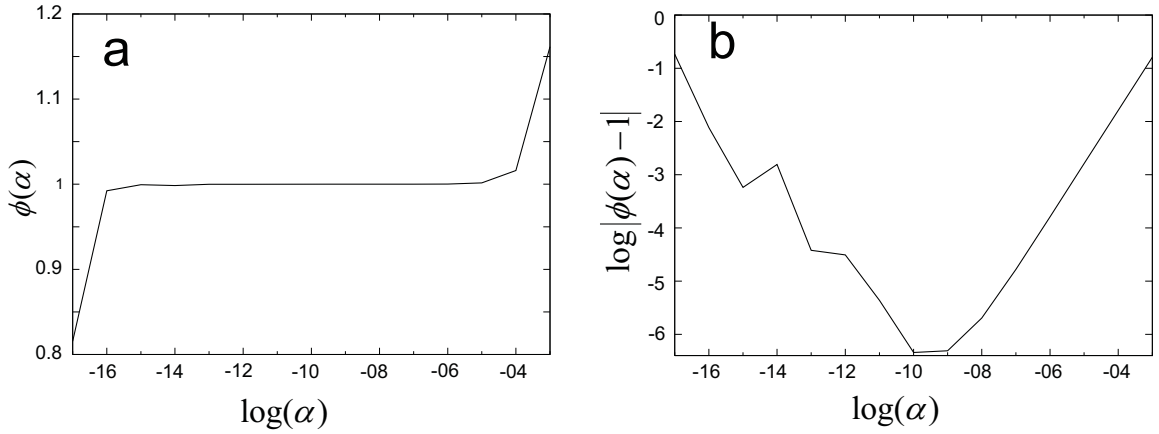

Fig. 4.1. The verification of the gradient calculation (a) variation of $\phi(\alpha)$ with $\log(\alpha)$ and (b) variation of $\log\left|\phi(\alpha)-1\right|$ with $\log(\alpha)$.

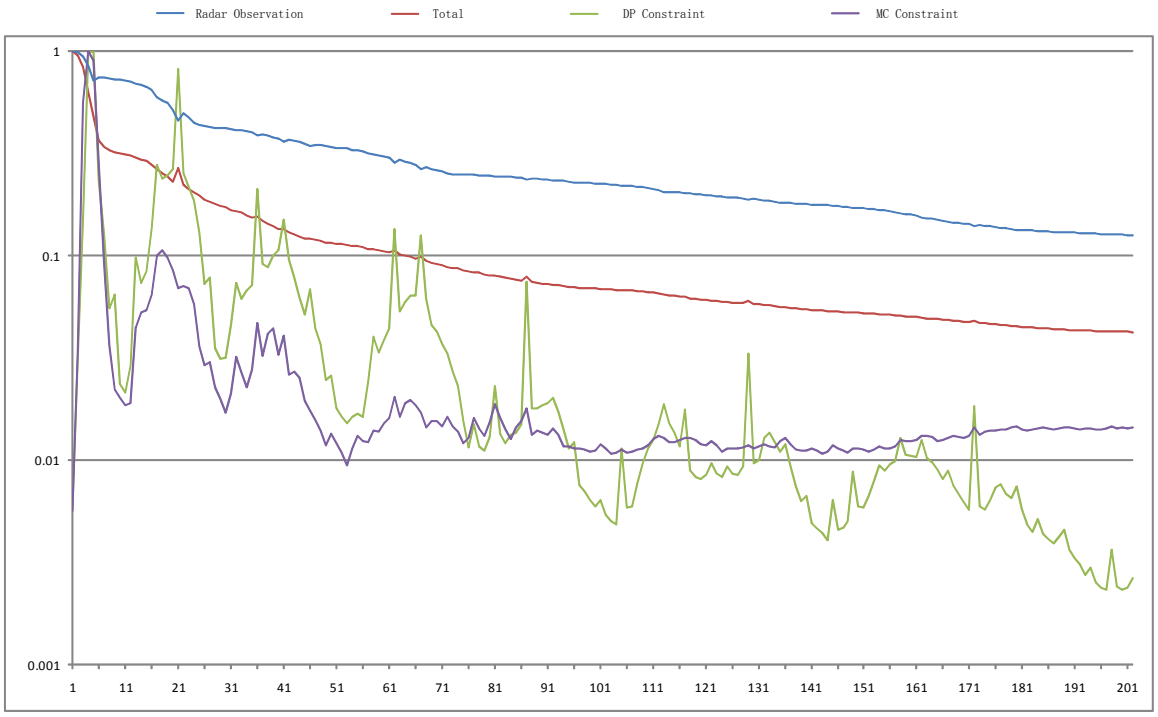

Fig. 4.2 The scaled cost function (J_k/J_{max}) as a function of the number of iterations. The red line is for the total cost function, the blue line is the part contributed by the radar observation term, the olive green line is the part contributed by the diagnostic pressure equation constraint and the purple line is the part contributed by the mass continuity equation constraint.

To check whether the minimization process goes well after including the diagnostic pressure equation constraint in the 3DVAR system, the behavior of the cost function is examined by plotting the evolution of the scaled cost function with the number of iterations in a similar way as Gao et al. (2001). The individual parts in the total cost function contributed respectively by the radar observation term, the mass continuity equation constraint and the constraint are also plotted for investigation. As an example, Fig. 4.2 shows the evolution of scaled cost function for analyses valid at the beginning of the data assimilation run for the 8 May 2003 Oklahoma City tornadic supercell storm in Section 4.4. The weighting coefficient, or the diagonal elements a_{pi} of matrix A_p for the diagnostic pressure equation constraint used for this test is 7.0E-8. It can be seen from

Fig. 4.2 that the total cost function and the three parts (from radar observation, DP constraint, MC constraint respectively) all decrease well in the minimization process. Other different weighting coefficients (a_{pi}*10, a_{pi}/10) for the diagnostic pressure equation constraint are also tested and the cost function decreases similarly. These tests confirm that the updated ARPS 3DVAR system with one more weak constraint (the diagnostic pressure equation constraint) works correctly and is ready for the following idealized testing and real case studies.

4.2.5 The measurement of dynamic consistency in single analysis step

It is expected that the dynamic consistency among model variables in single analysis step can be improved by the use of diagnostic pressure equation constraint. However, it is a challenge to directly measure the consistency among model variables of an analysis.

In previous thermodynamic retrieval researches, a so-called "momentum checking" method (Gal-chen and Hane 1981) is mainly used to check the quality of retrieved thermodynamic fields. The 3D momentum checking (Er) quantity is defined based on equations (4.10)-(4.12) as follows (similar to Liou et al. 2003):

$$E_r = \frac{\iiint \left[\left(\frac{\partial p'}{\partial x} - F \right)^2 + \left(\frac{\partial p'}{\partial y} - G \right)^2 + \left(\frac{\partial p'}{\partial z} - \bar{\rho} g \left[\frac{\theta'}{\bar{\theta}} - \frac{p'}{\bar{\rho} c_s^2} + \frac{q_v'}{\varepsilon + \bar{q}_v} - \frac{q_v' + q_{liquid+ice}}{1 + \bar{q}_v} \right] - H \right)^2 \right] dxdydz}{\iiint (F^2 + G^2 + H^2) dxdydz} \tag{4.14}$$

where

$$F = -\bar{\rho} \frac{\partial u}{\partial t} - \bar{\rho} \vec{\bar{V}} \bullet \nabla u + \left(\bar{\rho} f v - \bar{\rho} \tilde{f} w \right) + D_u \tag{4.15}$$

$$G = -\bar{\rho} \frac{\partial v}{\partial t} - \bar{\rho} \vec{\bar{V}} \bullet \nabla v + \bar{\rho} f u + D_v \tag{4.16}$$

$$H = -\overline{\rho}\frac{\partial w}{\partial t} - \overline{\rho}\overline{\vec{V}} \bullet \nabla w + \overline{\rho}\widetilde{f}u + D_w \tag{4.17}$$

The symbols in equations (4.14)-(4.17) follow the same meanings as in equations (4.5) - (4.12).

Although the momentum checking method is useful for the evaluation of the quality of the thermodynamic retrievals, it cannot be used directly for our study. As can be seen in equations (4.15)-(4.17), in order to compute the Er quantity, the wind tendency terms in the three momentum equations should be known in advance. In some previous thermodynamic retrieval studies, the tendency terms is calculated by the difference between two or among three successive radar volume scans, which are collected every 5 or 6 minutes. In some mesoscale studies, the tendency terms are just approximated by zero values. Both of these approximations are not acceptable for storm-scale data assimilation.

An alternate approach is adopted in this research. The tendency terms is estimated from the previous ARPS model run. For example, for an analysis at time *t* (except the beginning of the data assimilation), the background is the five-minute ARPS model forecast starting at time *t-5min*. The model run also computes the wind tendency terms in the forecast every integral time step. The wind tendency terms at time *t* are then dumped out to be treated as a rough estimate of the true tendency terms. This is not ideal, but may provide a general view how the dynamic consistency will be improved by including the diagnostic pressure equation constraint in the 3DVAR cost function. Generally, a smaller Er value means more dynamic consistent analysis, however there still exist exceptions as discussed in Hane and Ray (1985). Therefore, it is the best to consider the momentum checking method as a partial or relative measurement.

We now use again the 8 May 2003 Oklahoma City tornadic supercell storm in Section 4.4 as an example to demonstrate the decrease of Er value when the diagnostic pressure equation constraint is used in the analysis step. A first analysis if performed at t=2140UTC. Starting from this analysis, a five-minutes (2140UTC~2145UTC) forecast is produced and the wind tendency terms at t=2145UTC are dumped from the model. At t=2145UTC, four analysis experiments using different weighting coefficients for the diagnostic pressure equation constraint are conducted and the Er values for each experiments are computed. It should be noted that all these analysis experiments include the mass continuity equation constraint. Table 4.1 shows the calculation results. It can be seen that the Er Value for the analysis without the use of the diagnostic pressure equation constraint is the largest. As the weighting coefficient decreases (means more part of the cost function is contributed by the diagnostic pressure equation constraint), the Er Value get smaller. The Er checking is also conducted for other case studies in this paper and at other analysis times, the behavior is very similar. That is, the use of the diagnostic pressure equation constraint yields smaller Er Values. Therefore, it can be concluded that the diagnostic pressure equation constraint does help improve the dynamic consistency among model variables in single analysis step in terms of the momentum checking quantity.

Table 4.1 List of Er Value with different weighting coefficients for the diagnostic pressure equation constraint

DP weighting coefficient	n/a	7E-7	7E-8	7E-9
Er Value	0.0011958	0.0010777	0.0007783	0.0006086

4.3 The impact of the constraint in idealized experiments

4.3.1 Experimental design

4.3.1.1 The prediction model and truth simulation

To examine the impact of the constraint on the storm-scale 3DVAR, a series of OSS experiments are conducted using simulated data from the May 20, 1977 Del City, Oklahoma supercell storm case (Ray *et al.* 1981). The Advanced Regional Prediction System (ARPS) is used to simulate such a deep convective storm within a 54 x 54 x 16 km^3 physical domain. The model grid comprises of 57 x 57 x 35 grid points. Horizontal resolution of 1km and vertical resolution of 0.5km are used. The truth simulation is initialized from a modified real sounding plus a +4K ellipsoidal thermal bubble centered at x=48, y=16 and z=1.5km, with radii of 10km in x and y directions while 1.5km in z direction. The Lin three categories ice microphysical scheme is used together with a 1.5-order turbulent kinetic energy subgrid parameterization. Open conditions are used at the lateral boundaries. A wave radiation condition is also applied at the top boundary. Free-slip conditions are applied to the bottom boundary. The length of simulation is up to two hours. A constant wind of u=3$ms^{-1}$and v= 14ms^{-1} is subtracted from the observed sounding to keep the primary storm cell near the center of model grid. The evolution of the simulated storms is similar to those documented in Xue et al. (2001).

During the control run, the supercell strengthens over the first 20 minutes. The strength of the cell then decreases thereafter. At around 55 minutes, the cell splits into two. The north-northeastward moving cell tends to dominate the system. Another cell moves northwestward and splits again at 95 minutes.

4.3.1.2 Simulation of radar observations

The experiments assimilates pseudo radial velocity observations from two radars, which are located at the southwest corner (i.e. x=0km, y=0km) and the southeast corner (i.e. x=54km, y=0km) of the model domain. The pseudo radar radial velocity (Vr) observations are assumed available on the grid points and calculated as follows.

$$V_r = u\sin\phi\cos\theta + v\cos\phi\cos\theta + w\sin\theta \tag{4.18}$$

where θ is the elevation angle, ϕ is the azimuth angle, and *u, v* and *w* are the three components of wind fields taken from the truth simulation. Random Gaussian noise with a standard deviation of 1m s^{-1} is added to the pseudo radial velocity. Terminal velocity is not considered.

4.3.1.3 Experimental design

The first 3DVAR analysis is started at t=30 minutes into truth simulation. From this analysis, the ARPS model runs for a 5-minute forecast. Then new radar data at the end of the 5-minute forecast is ingested into the model through the 3DVAR analysis again. This intermittent assimilation repeats until t=90 minutes into the truth simulation. Therefore, the data assimilation covers a sixty-minute period.

Table 4.1 lists all the data assimilation experiments. First four experiments (CNTL, onlyMC, onlyDP, NOEC) are designed mainly to investigate the impact of the diagnostic pressure equation constraint on the storm-scale data assimilation. For comparison purpose, the impact of the mass continuity equation constraint is also examined here. The "CNTL" experiment uses both the diagnostic pressure equation constraint and the mass continuity equation constraint. The "onlyMC" experiment uses only the mass continuity equation constraint and "MC" is an abbreviation representation of "Mass Continuity

equation" (hereafter). The "onlyDP" uses only the diagnostic pressure equation constraint and "DP" is an abbreviation of "Diagnostic Pressure equation" (hereafter). The "NOEC" means "NO Equation Constraint" is used.

Table 4.2. List of data assimilation experiments ("DP" stands for "Diagnostic Pressure equation", "MC" stands for "Mass Continuity equation")

	DP weighting coefficient if used	MC weighting coefficient if used
CNTL	2.5E-7	2.5E-7
onlyDP	2.5E-7	
onlyMC		2.5E-7
NOEC		
CNTL_DP*5	1.25E-6	2.5E-7
CNTL_DP/5	5.0E-8	2.5E-7
CNTL_DP*25	6.25E-6	2.5E-7
CNTL_DP/25	1.0E-8	2.5E-7
CNTL_MC*5	2.5E-7	1.25E-6
CNTL_MC/5	2.5E-7	5.0E-8
CNTL_MC*25	2.5E-7	6.25E-6
CNTL_MC/25	2.5E-7	1.0E-8

The next four experiments (CNTL_DP*5, CNTL_DP/5, CNTL_DP*25, CNTL_DP/25) are designed to test the sensitivity of the data assimilation to the DP weighting coefficient. The "CNTL_DP*5" experiment is the same as the "CNTL" experiment except that the DP weighting coefficient is increased by five times, which is

indicated by the symbols “DP*5”. Similarly, the “CNTL_DP*5” experiment is the same as the “CNTL” experiment except that the DP weighting coefficient is decreased by five times. Similar explanations go to the experiment “CNTL_DP*25”, “CNTL_DP/25” which change the weighting coefficient by twenty-five times.

Finally, the last four experiments (CNTL_MC*5, CNTL_MC/5, CNTL_MC*25, CNTL_MC/25) are designed to test the sensitivity of the data assimilation to the MC weighting coefficient. Similar symbols are used as the previous four experiments.

To compare the accuracy of the data assimilation results from different experiments, the RMS error statistics of model variables between the experiments and the truth simulation are computed using the following equation:

$$RMS\_s = \sqrt{\frac{\sum_{i=1}^{N} (s - s_{simu})_i^2}{N}} \tag{4.19}$$

where N is the total number of three dimensional grid points used in the calculation, and the subscript $simu$ stands for the data from the truth simulation. The computation of the RMS error statistics is only done over model grid points where the reflectivity (estimated from the local hydrometeor mixing ratios) of the truth simulation is greater than 10 dBZ.

4.3.2 Results of experiments

4.3.2.1 The impact of the diagnostic pressure equation constraint

To investigate the impact of the diagnostic pressure equation constraint, the RMS error statistics are calculated during the whole assimilation period for all model variables and are shown in Fig. 4.3. For easy display and without harm to the conclusions, the RMS errors of the rain/snow/hail mixing ratios are substituted by the RMS error of

simulated reflectivity, and the RMS error for V component of wind field is not shown as it evolves similar to the RMS error for U component of wind fields.

First, look at the results from the "NOEC" experiments (the red solid line), which excludes both constraints, with that from the "onlyDP" experiments (the blue dashed line), which imposes the diagnostic pressure equation constraint alone. After fifteen minutes of data assimilation (i.e. after three data assimilation cycles, at t=45 minutes of truth simulation), the improvement on the rainfall/cloud-water RMS error can already be distinguished although the improvement is still slight. After thirty minutes of data assimilation (i.e. after six data assimilation cycles, at t=60 minutes of truth simulation), the data assimilation results from the "onlyDP" experiment are further improved and much better than that of the "NOEC" experiment (Fig. 4.3). After that, the RMS errors from the "onlyDP" experiment are generally smaller than that from the "NOEC" experiment.

Similar behaviors can also been seen by comparing the "CNTL" experiments (the magenta solid line) and the "onlyMC" experiments (the green dashed line) which is the same as the "CNTL" experiment but without the diagnostic pressure equation constraint. After fifteen minutes data assimilation (i.e. t=45 minutes into truth simulation), the "CNTL" experiment produces smaller RMS error in the reflectivity field. After thirty-five minutes data assimilation (i.e. t=65 minutes into truth simulation), the "CNTL" experiment produces smaller RMS errors in all model fields than the "onlyMC" experiment.

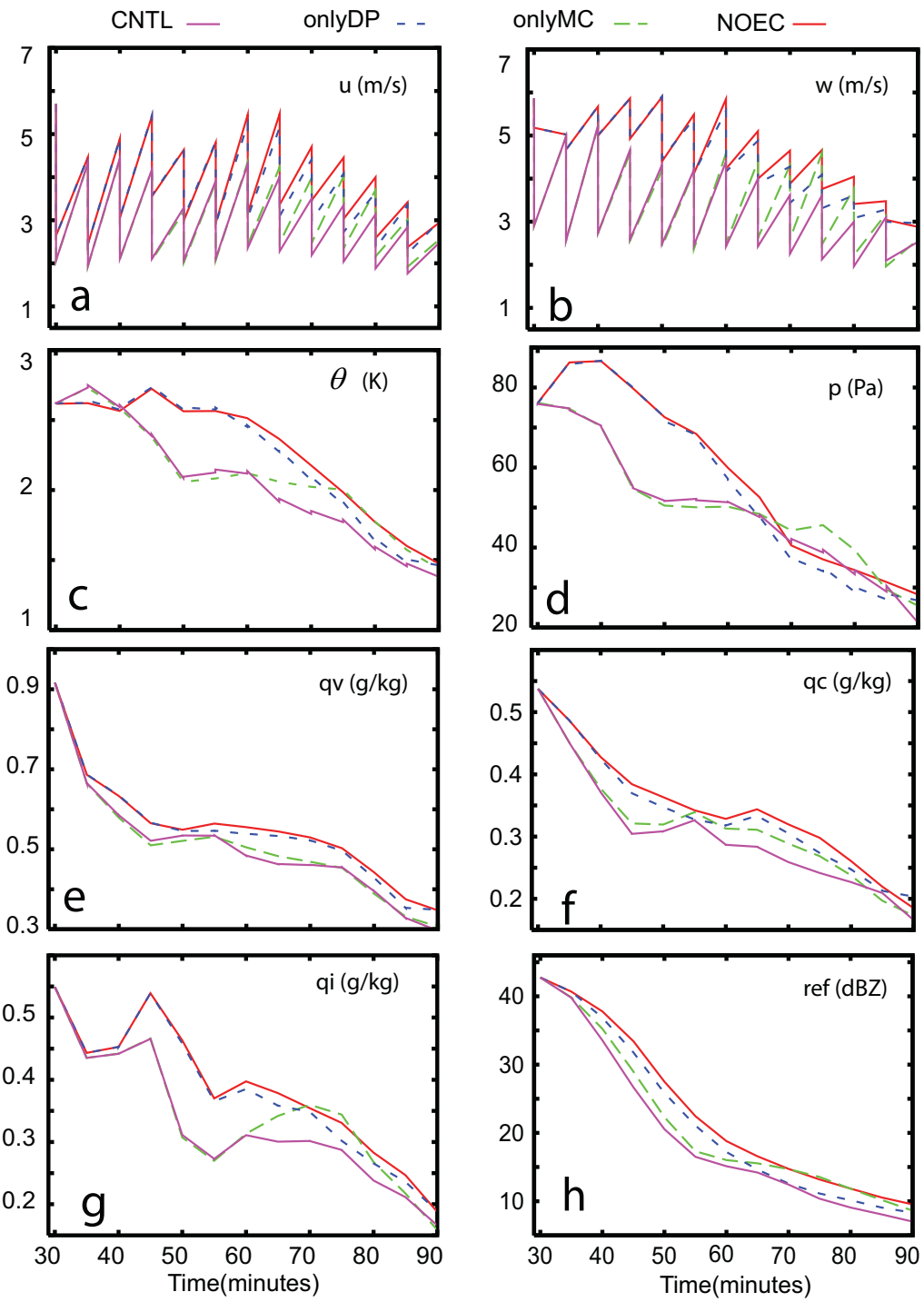

Fig. 4.3. The evolution of RMS error of model fields during the 1-h assimilation period for (a) U component of wind fields, (b) vertical velocity, (c) perturbation potential temperature, (d) pressure, (e) water vapor mixing ratio, (f) cloud water mixing ratio, (g) cloud ice mixing ratio, (h) simulated reflectivity from model rain/snow/hail mixing ratio. The solid magenta line is for the "CNTL" experiment, the blue dashed line is for the "onlyDP" experiment, the green dashed line is for the "onlyMC" experiment and the solid red line is for the "NOEC" experiment.

From the above discussions, it can be concluded that the diagnostic pressure equation constraint does improve the data assimilation results after several data assimilation cycles. It is so no matter whether the mass continuity equation is included or not.

In practice, the mass continuity equation is often included in the storm-scale 3DVAR data assimilation since it can help retrieve better vertical velocity and horizontal winds in single analysis step. This can be clearly seen from Fig. 4.3a and Fig. 4.3b (comparing the red solid line and the green dashed line), where the mass continuity equation constraint helps reduce the RMS error of U component of wind fields analysis by about 0.5 m s^{-1} and vertical velocity analysis by about 2 m s^{-1} for first data analysis cycle. The better-analyzed wind fields then promote an improved recovery of other model fields.

The impact of the diagnostic pressure equation constraint on single time analysis is not easy to be demonstrated as that of the mass continuity equation constraint discussed in the above. In the first analysis, the RMS errors from the "onlyDP" experiment are almost the same as that from the "NOEC" experiment. Further investigation on the RMS errors in different vertical model levels does show some differences. The RMS errors of U component of wind fields between roughly z=7km MSL and z=12km MSL are very slightly smaller in the "onlyDP" experiment than that in the "NOEC" experiment (decreased about 0.001 m s^{-1}). The RMS errors of vertical velocity in the low levels (below 2.5km MSL) and mid-upper levels (7km MSL ~ 9km MSL) are also very slightly smaller in the "onlyDP" experiment than that in the "NOEC" experiment (decreased about 0.004 m s^{-1}). The RMS errors of perturbation pressure in vertical levels of 2km ~

8km MSL are again very slightly smaller in the "onlyDP" experiment than that in the "NOEC" experiment (decreased about 0.1 Pa). The RMS errors of perturbation potential temperature in the vertical levels of 3.5km MSL~9km MSL are very slight larger in the "onlyDP" experiment than that in the "NOEC" experiment (increased about 0.004K). The RMS errors of water vapor field in the "onlyDP" experiment are essentially the same as that in the "NOEC" experiment, which is expectable since the buoyancy contribution from the perturbation water vapor is normally at least one order of magnitude less than that from other buoyancy terms in the vertical momentum equation. Although the above impact at single time analysis is very slight, it actually gives out a picture that how the model variables are adjusted by the diagnostic pressure equation constraint. It can be assumed that the diagnostic pressure equation constraint helps improve the analysis of wind fields and pressure fields. At the same time, the temperature field is slightly adjusted to be dynamically consistent with the updated wind fields and pressure fields. The adjustment in the temperature field does not necessary mean to reduce the RMS error of it at single time analysis. Instead, the RMS error of potential temperature field may actually increases very slightly. However, this adjustment is necessary and good to boost dynamic consistency among model variables and therefore to produce improved data assimilation results in the following cycles, which has been demonstrated in the above discussions about Fig. 4.3.

Overall, it is shown in Fig. 4.3 that the diagnostic pressure equation constraint and the mass continuity equation constraint both can improve the data assimilation results. The impact of the mass continuity equation constraint is more direct at the first several data assimilation cycles as it greatly promotes a better analysis of vertical velocity. The

impact of the diagnostic pressure equation constraint is not obvious in single time analysis although it does slightly improve the analysis of the wind fields (also demonstrated by Xu et al. 2001a) and the pressure field. For the entire intermittent data assimilation cycles, the impact of the diagnostic pressure equation constraint is evidently positive. Including both the constraints at the same time yields the best data assimilation results with least RMS error statistics in all model variables.

All the above discussions are based on the RMS error statistics. How these RMS error differences are related to changes in individual model fields? As examples, here we will show the recovery results after forty-five minutes data assimilation (i.e. at t=75 minutes into truth run) for the reflectivity field, the potential temperature field and the vertical velocity field. Fig. 4.4 shows the distribution of the RMS error at different vertical levels from z=0km MSL to z=12km MSL. The solid magenta line is for the "CNTL" experiment, the blue dashed line is for the "onlyDP" experiment, the green dashed line is for the "onlyMC" experiment and the solid red line is for the "NOEC" experiment. From Fig. 4.4a, it can be seen that the RMS error of the reflectivity field is evidently reduced in the "onlyDP" and "CNTL" experiments as compared to that in the "NOEC" and "onlyMC" experiment respectively. The most noticeable decreases (by about 4dBZ) are at the vertical levels from about z=6km MSL to z=8km MSL. Fig. 4.5 shows the simulated reflectivity fields at z=6.5km MSL. For easy comparing, the difference of the reflectivity field between each data assimilation experiment and the truth simulation are also plotted in Fig. 4.5. It can be seen that at this time (after forty-five minutes of data assimilation), the rainfall pattern for the domain-centered storm cell is recovered well in all the four data assimilation experiments. The most noticeable

difference lies in the recovery of the rainfall pattern of the upper-left-cornered storm cell and the region between these two storm cells. It is very clear that the "onlyDP" and the "CNTL" experiments produce better results in the above area over the "NOEC" and the "onlyMC" experiments. The "CNTL" experiment yields the best data assimilation results in terms of the recovery of the rainfall pattern in the whole domain.

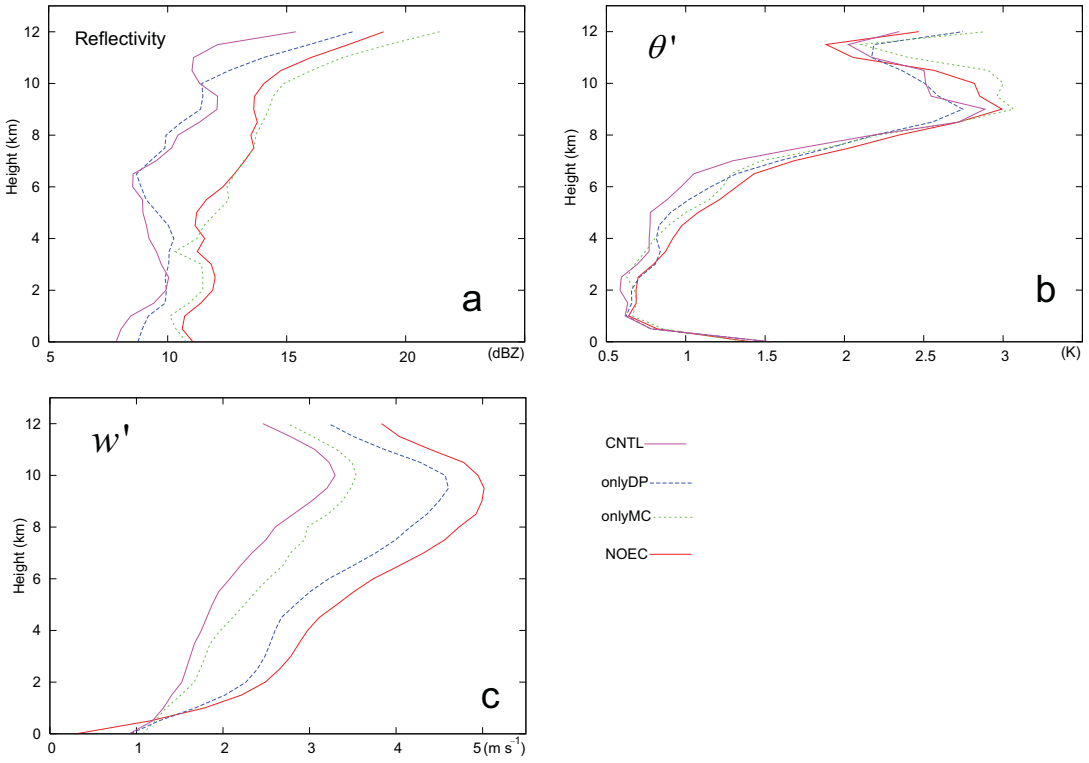

Fig. 4.4. The RMS error at vertical levels from z=0km MSL to z=12km MSL at t=75minutes into truth simulation (after forty-five minutes of data assimilation) for (a) reflectivity field, (b) perturbation potential temperature, and (c) vertical velocity. The solid magenta line is for the "CNTL" experiment, the blue dashed line is for the "onlyDP" experiment, the green dashed line is for the "onlyMC" experiment and the solid red line is for the "NOEC" experiment.

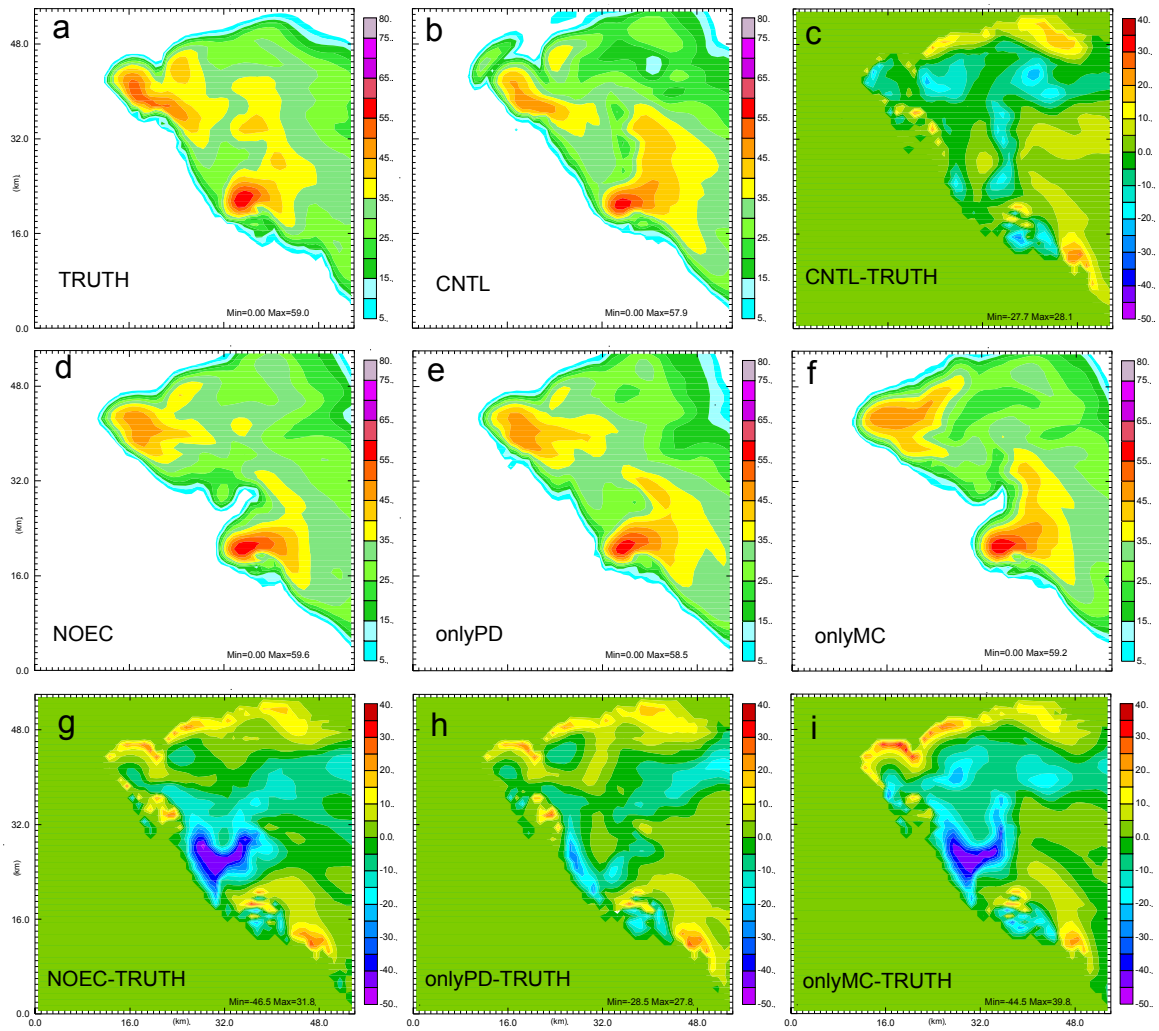

Fig. 4.5. The simulated reflectivity field at z=6.5km MSL for (a) the truth simulation, (b) the "CNTL" experiment, (d) the "NOEC" experiment, (e) the "onlyDP" experiment, (f) the "onlyMC" experiment; and the difference of the reflectivity field between each data assimilation experiment and the truth simulation at z=6.5km MSL for (c) the "CNTL" experiment, (g) the "NOEC" experiment, (h) the "onlyDP" experiment, (i) the "onlyMC" experiment . All the above plots are available at t=75 minutes into truth simulation (i.e. after forty-five minutes data assimilation).

From Fig. 4.4b, it can be seen that the decrease of the RMS error of the perturbation potential temperature field by the use of equation constraints is mainly for the vertical levels from z=2km MSL to z=8km MSL. The most noticeable reductions (by about 0.3~0.4K) are at the vertical levels from about z=5km MSL to z=7.5km MSL. Fig. 4.6 shows the perturbation potential temperature fields at z=6.5km MSL. For easy comparing,

the differences of the perturbation potential temperature field between each data assimilation experiment and the truth simulation are also plotted in Fig. 4.6. It can be seen that at this time (after forty-five minutes of data assimilation), the potential temperature field is recovered very well in all the four data assimilation experiments and there is no significant pattern difference among the results from the four experiments at first glance. However, from the difference fields (Fig. 4.6c, g, h, i), it can be more clearly seen that the mass continuity equation constraint mainly helps improve the recovery of the temperature structure for the domain-centered storm cell (Fig. 4.6i vs. Fig. 4.6g) and the diagnostic pressure equation constraint mainly helps improve the recovery of the temperature structure for the upper-left-cornered storm cell (Fig. 4.6h vs. Fig. 4.6g). Similar as before, the "CNTL" experiment yields the best data assimilation results in terms of the recovery of the temperature structure of the whole storm systems.

From Fig. 4.4b, it can be seen that after forty-five minutes of data assimilation, the use of the diagnostic pressure equation constraint (the "onlyDP" experiment) can help reduce the RMS error of the vertical velocity at roughly an order of 0.4 m s^{-1}. The use of the mass continuity equation constraint can reduce it more at roughly an order of 1.0 m s^{-1}. The use of both equation constraints produces the most reduction with roughly an order of 1.4 m s^{-1}. The decrease of the RMS error of the vertical velocity field by the use of equation constraints is for most vertical levels (above z=1.5km MSL). The most noticeable decreases are at the vertical levels above z=4.5km MSL. Fig. 4.7 shows the vertical velocity fields at z=6.0km MSL. For easy comparing, the difference of the vertical velocity field between each data assimilation experiment and the truth simulation are also plotted in Fig. 4.6. It can seen that at this time (after forty-five minutes of data

assimilation), the vertical velocity field is already recovered fairly well in all the four data assimilation experiments for the domain-centered storm cell although the use of the mass continuity equation constraint can still make some slightly improvements in this region. The most improvements are for the upper-left-cornered storm cell. Similar as before, the "CNTL" experiment again yields the best data assimilation results in terms of the recovery of vertical velocity field of the whole storm systems.

To summary for this section, both the diagnostic pressure equation constraint and the mass continuity equation constraint have positive impact on the intermittent data assimilation. In single time analysis or first one/two data assimilation cycles, the impact of the diagnostic pressure equation constraint is not as large as that of the mass continuity equation constraint. The mass continuity equation constraint can evidently improve the retrieval of wind fields, especially the vertical velocity, in single analysis step while the diagnostic pressure equation constraint only slightly improve the retrieval of the wind fields and the pressure field. The temperature field is also slightly adjusted by the diagnostic pressure equation constraint during the analysis process but this adjustment is assumed mainly to boost dynamic consistency among model variables and not necessary to produce a more close resemblance to the truth. After fifteen minutes of data assimilation (three cycles), the positive impact of the diagnostic pressure equation appears and becomes more evident after thirty-five minutes of data assimilation (seven cycles). The mass continuity equation constraint can also improve the intermittent data assimilation results; however, its impact becomes a little bit less prominent after forty minutes data assimilation (eight cycles) than its direct impact on single time data analysis.

Including both the diagnostic pressure equation constraint and the mass continuity equation constraint yields the best data assimilation results.

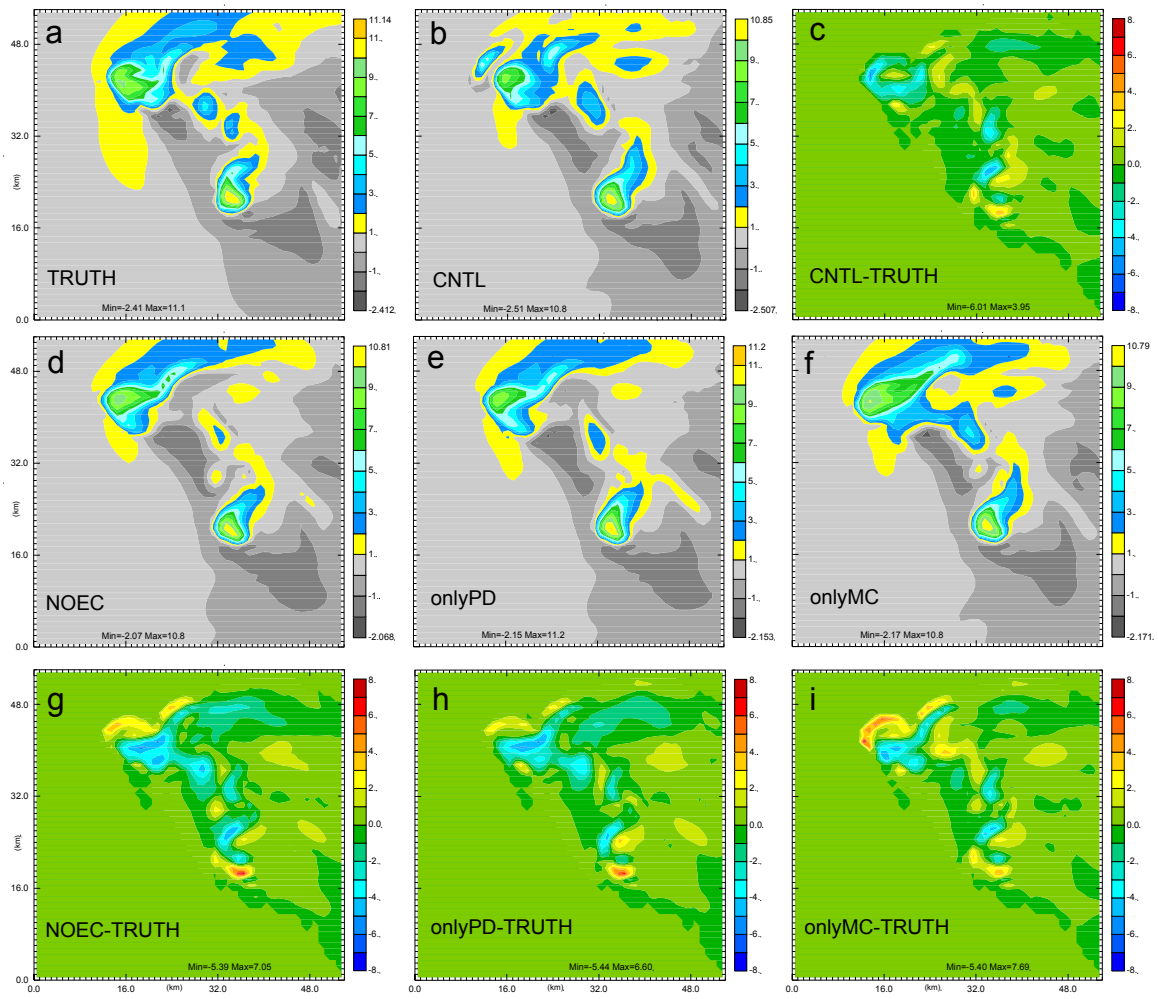

Fig. 4.6. The perturbation potential temperature field at z=6.5km MSL for (a) the truth simulation, (b) the "CNTL" experiment, (d) the "NOEC" experiment, (e) the "onlyDP" experiment, (f) the "onlyMC" experiment; and the difference of the perturbation potential temperature field between each data assimilation experiment and the truth simulation at z=6.5km MSL for (c) the "CNTL" experiment, (g) the "NOEC" experiment, (h) the "onlyDP" experiment, (i) the "onlyMC" experiment . All the above plots are available at t=75 minutes into truth simulation (i.e. after forty-five minutes data assimilation).

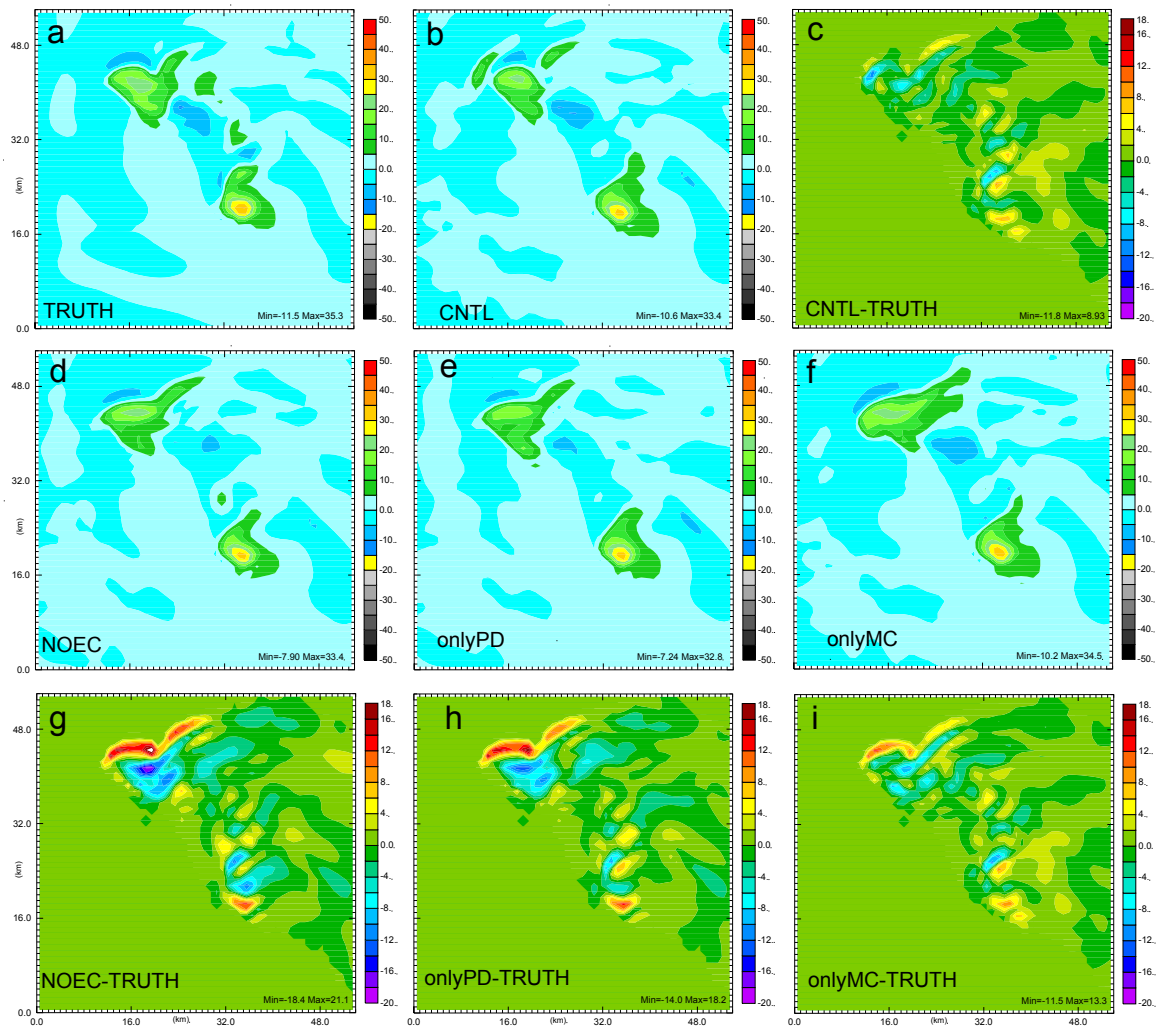

Fig. 4.7. The vertical velocity field at z=6.0km MSL for (a) the truth simulation, (b) the "CNTL" experiment, (d) the "NOEC" experiment, (e) the "onlyDP" experiment, (f) the "onlyMC" experiment; and the difference of the vertical velocity field between each data assimilation experiment and the truth simulation at z=6.0km MSL for (c) the "CNTL" experiment, (g) the "NOEC" experiment, (h) the "onlyDP" experiment, (i) the "onlyMC" experiment . All the above plots are available at t=75 minutes into truth simulation (i.e. after forty-five minutes data assimilation).

4.3.2.2 Sensitivities to the weighting coefficients

In this section, eight more data assimilation experiments (as listed in Table 4.2 and described in Section 4.3.1.3 "Experimental design") are conducted to test the sensitivity of the weighting coefficients of the diagnostic pressure equation constraint and the mass continuity equation constraint. These experiments are based on the "CNTL" experiment

that uses both the two constraints and yields the best data assimilation results as discussed in the above.

Fig. 4.8 and Fig. 4.9 show the evolution of the RMS errors of all these experiments during the 1-h data assimilation period for all model fields. Following the practices in previous section, the RMS errors of the rain/snow/hail mixing ratio are substituted by the RMS error of simulated reflectivity, and the RMS error for V component of wind field is not shown as it evolves similar to the RMS error for U component of wind fields.

Fig. 4.8 presents the sensitivities of the data assimilation to the weighting coefficient for the DP constraint. It can be seen that when the weighting coefficient for the DP constraint varies by five times, the assimilation results are very close to that produced by the "CNTL" experiment. When the DP weighting coefficient is increased by twenty-five times, the data assimilation results are still acceptable and not far away from that of the "CNTL" experiment. When the DP weighting coefficient is decreased by twenty-five times, the data assimilation results are much worse than that from the "CNTL" experiment. This indicates that it should be cautious in choosing the DP weighting coefficient. Very small DP weighting coefficient should be avoided, as it might give the diagnostic pressure equation constraint much more share of the total cost function and then degrade the quality of the data assimilation.

Fig. 4.9 presents the sensitivities of the data assimilation to the weighting coefficient for the MC constraint. It can be seen that when the weighting coefficient of the MC constraint varies by five times, the data assimilation results are very close to that of the "CNTL" experiment. When the MC weighting coefficient varies by twenty-five times, the data assimilation results are still comparable to that of the "CNTL" experiment.

Therefore, it can be concluded that the mass continuity equation constraint is not very sensitive to the weighting coefficient.

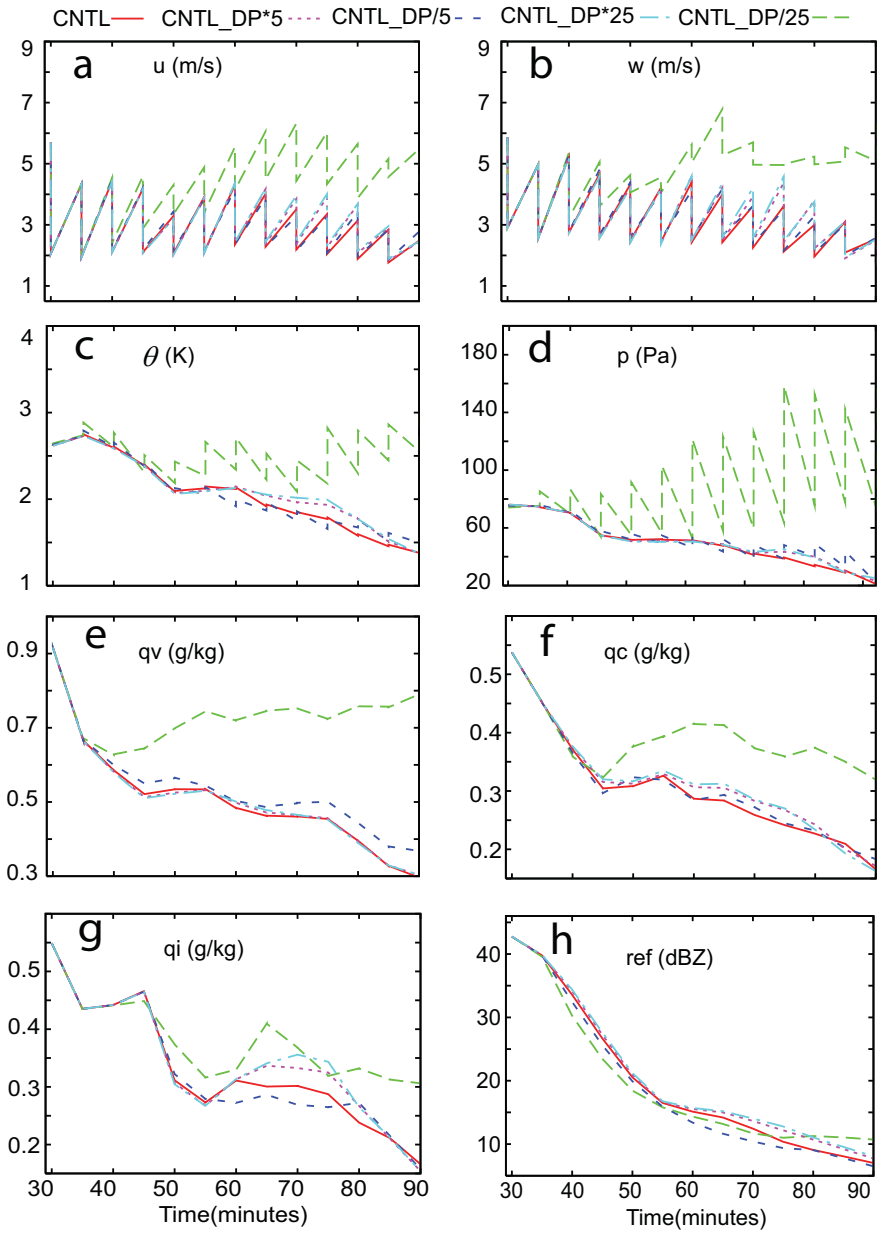

Fig. 4.8. The evolution of RMS error of model fields during the 1-h assimilation period for (a) U component of wind fields, (b) vertical velocity, (c) perturbation potential temperature, (d) pressure, (e) water vapor mixing ratio, (f) cloud water mixing ratio, (g) cloud ice mixing ratio, (h) simulated reflectivity from model rain/snow/hail mixing ratio. The solid red line is for the "CNTL" experiment, the magenta dashed line is for the "CNTL_DP*5" experiment, the blue dashed line is for the "CNTL_DP/5" experiment, the cyan dashed line is for the "CNTL_DP*25" experiment and the green dashed line is for the "CNTL_DP/25" experiment.

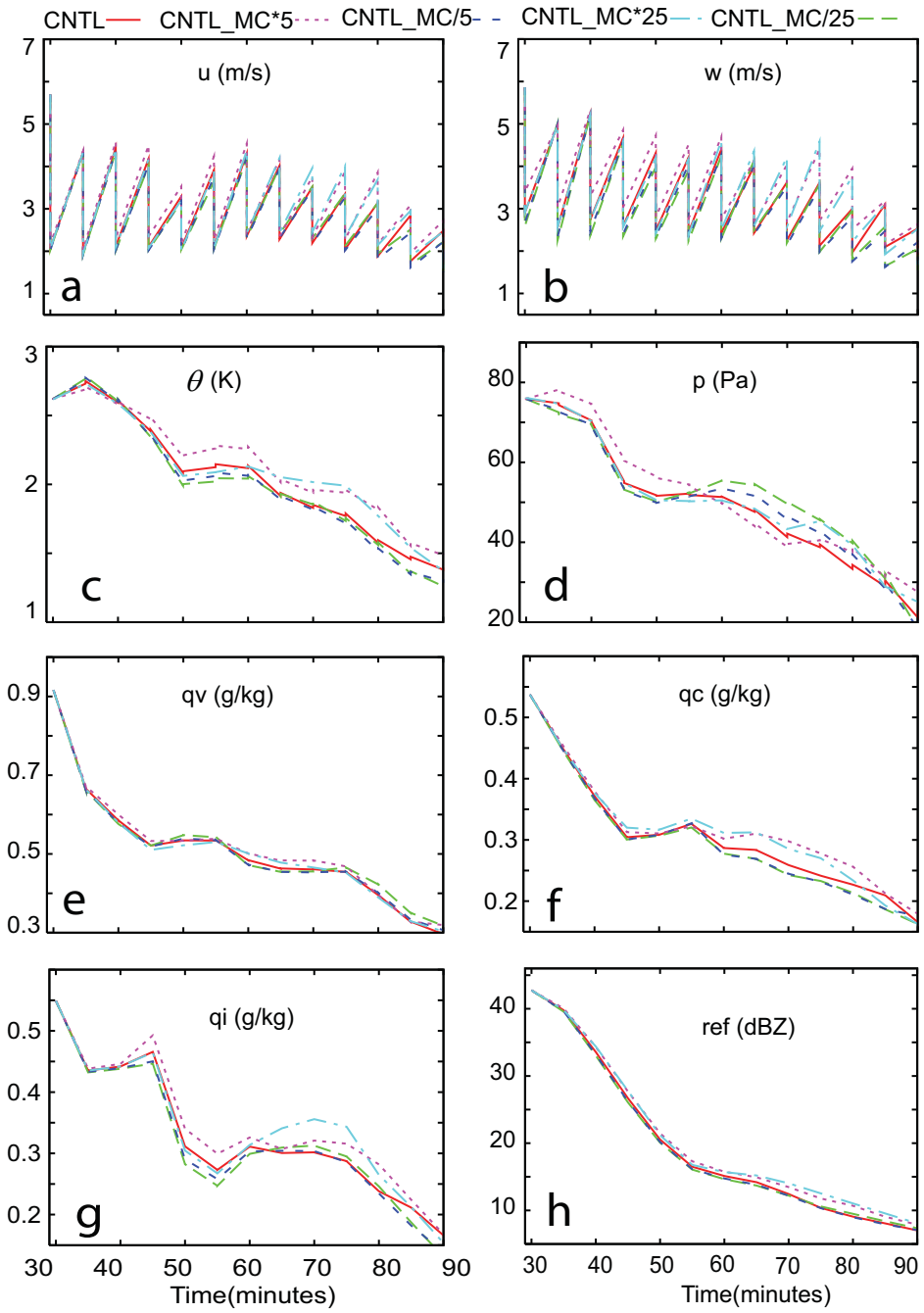

Fig. 4.9. The evolution of RMS error of model fields during the 1-h assimilation period for (a) U component of wind fields, (b) vertical velocity, (c) perturbation potential temperature, (d) pressure, (e) water vapor mixing ratio, (f) cloud water mixing ratio, (g) cloud ice mixing ratio, (h) simulated reflectivity from model rain/snow/hail mixing ratio. The solid red line is for the "CNTL" experiment, the magenta dashed line is for the "CNTL_MC*5" experiment, the blue dashed line is for the "CNTL_MC/5" experiment, the cyan dashed line is for the "CNTL_MC*25" experiment and the green dashed line is for the "CNTL_MC/25" experiment.

4.3.3 Conclusions

The impact of the diagnostic pressure equation constraint is investigated using an idealized tornadic supercell thunderstorm case. A 1-h data assimilation period and radar velocity observations from two radars are used for all the experiments. For comparison purpose, the impact of the mass continuity equation constraint is also examined.

It is demonstrated that the use of the diagnostic pressure equation constraint can improve the data assimilation results after a period of data assimilation (in this idealized case, it takes about thirty-five minutes, i.e. seven cycles, to get better recovery of all model variables). The impact of the diagnostic pressure equation constraint at single analysis step or first one/two assimilation cycles is not as large as that of the mass continuity equation constraint. While in single analysis step, the mass continuity equation constraint can evidently improve the retrieval of wind fields, especially the vertical velocity field, the diagnostic pressure equation constraint slightly improves the retrieval of the wind fields and the pressure field. The temperature field is also slightly adjusted by the diagnostic pressure equation constraint during the analysis process. This adjustment is assumed to boost dynamic consistency among model variables. These dynamic-balance-improved analyses help produce a better data assimilation results after several assimilation cycles. On the other hand, the impact of the mass continuity equation constraint becomes less prominent at the later stage of the assimilation period. Including both the diagnostic pressure equation constraint and the mass continuity equation constraint yields the best data assimilation results.

The sensitivity tests show that the mass continuity equation constraint is less sensitive to the choice of weighting coefficient while the diagnostic pressure equation

constraint is more sensitive and very small weighting coefficient of it can produce poor data assimilation results. Therefore, it should be cautious to determine the weighting coefficient of the diagnostic pressure equation constraint and very small weighting coefficient should be avoided. On the other hand, in practice, the weighting coefficient for the diagnostic pressure equation constraint might be a little bit different from case to case because the weighting coefficient is determined according to the magnitude of the diagnostic pressure equation constraint in cost function and this magnitude varies from case to case. A general rule is that this weighting coefficient should be within two order of magnitude of those used in the CNTL experiments.

4.4 The 8 May 2003 Oklahoma City tornadic supercell storm case

4.4.1 The case

On the late afternoon of 8 May 2003, a major tornado hit the southern Oklahoma City metropolitan area (Fig. 4.10). It first touched down at Moore, a suburban city close to and south of Oklahoma City, then traveled east north-east through south of Oklahoma City to Choctaw. The life span of the tornado is about 28 minutes from 2210UTC (1610 CDT, Central Daylight-saving Time) to 2238UTC. It caused up to F4 (Fujita scale) damages but no death. The tornado is thereafter named as the OKC tornado and the parent storm as the OKC tornadic thunderstorm.

The synoptic environment on 8 May 2003 over Oklahoma is very favorable for the development of supercell storms and even tornados, as has been discussed by Hu and Xue (2007), Romine et al. (2008). The low-level southerly wind flowed over Oklahoma all the day. An evident north-south-oriented dryline moved eastward approaching Moore, Oklahoma. A large instability with a 4004 J kg^{-1} CAPE (convective available potential

energy) , a 1 J kg^{-1} CIN (convective inhibition) and about 25 m s^{-1} vertical shear over the lowest 6km presented in the 1800 UTC 8 May Norman, Oklahoma (OUN) sounding. All these conditions indicate that there is a high possibility for tornadic supercell thunderstorms to develop.

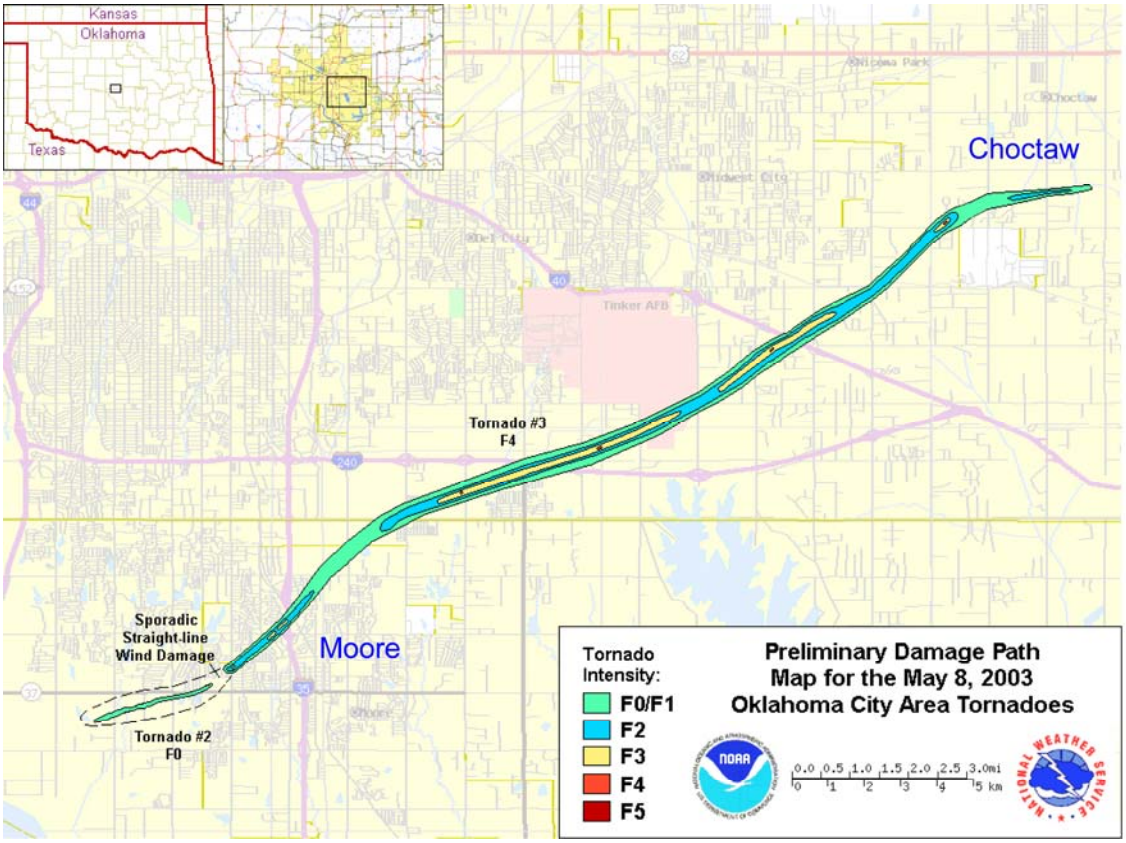

Fig. 4.10. The Damage Path Map for the 8 May 2003 Oklahoma City Area Tornadoes (National Weather Service, Norman).

At about 2030UTC, a first sign of the OKC tornadic storm showed up as a weak echo at the KTLX radar reflectivity field. By 2101UTC, the storm developed into a strong cell. In the following one hour, the storm grew rapidly and moved northeastward. By 2201UTC (see Fig. 4.11), the storm bore an obvious hook echo signature at its southwestern end. The hook echo was then located at northwest of Moore, just several

miles away. The pronounced hook echo signature sustained until at least 2235UTC while the parent supercell storm propagated east northeastward. The storm weakened since 2240UTC and dissipated by 0020UTC 9 May. In addition to OKC tornadic thunderstorm, there are three other short-lived storms (not shown). Here we will just focus on the major thunderstorm. Fig. 4.2 shows the general evolution of the major thunderstorm roughly every thirty minutes as observed by the KTLX radar at the 1.45 elevation angle.

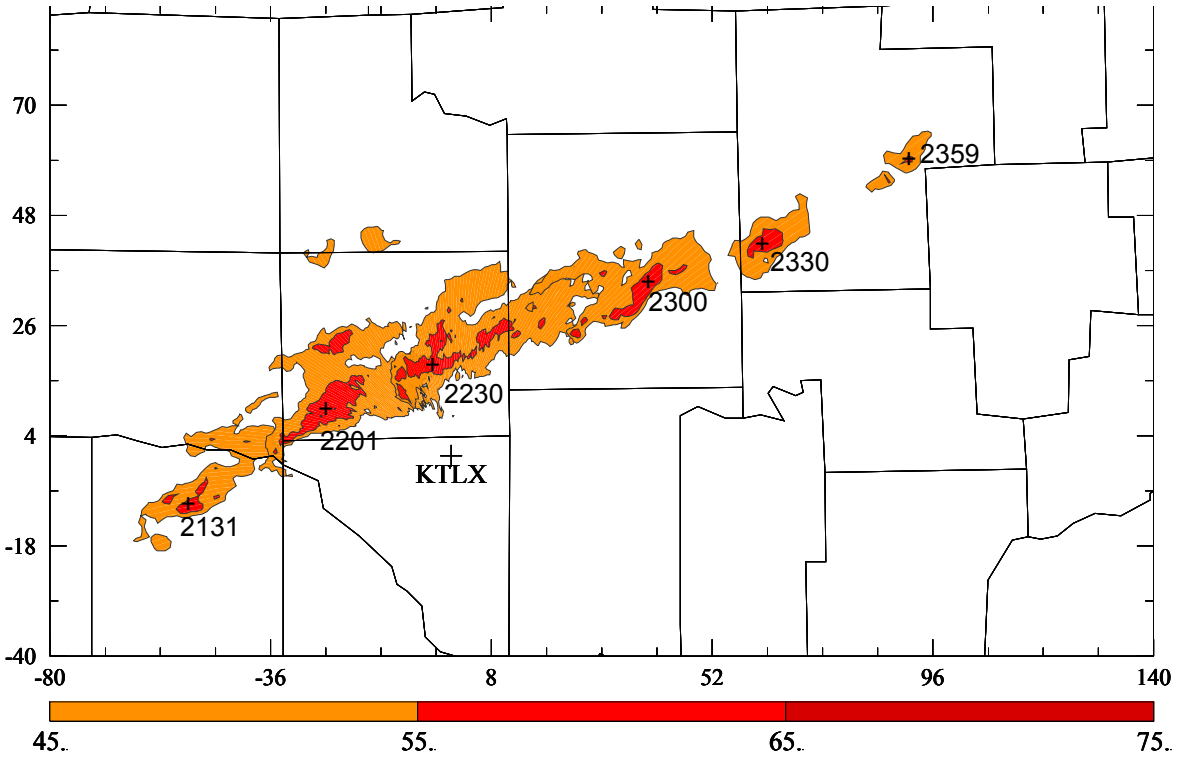

Fig. 4.11. Regions of radar echoes observed by the KTLX radar at the 1.45 elevation angle, from 2131 to 2159 UTC 8 May 2003. The contours are plotted every 10dBz, starting from 45dBz. The plus sign indicates the reflectivity center, which is followed by its corresponding time in UTC.

4.4.2 Experimental design

All experiments are conducted with a horizontal resolution of 3km. There are 195 grid points in both x and y directions. In the vertical direction, a stretched grid scheme is used. It contains 53 layers with an average grid spacing of 400m, stretching from about 20 m at the surface to 770m at the model top. The model domain is shown in Fig. 4.12. It covers nearly the whole Oklahoma. The evolution of the 8 May 2003 Oklahoma City tornadic supercell thunderstorm is roughly at the center of the domain. This big domain configuration will help alleviate the negative impact of boundary problems. The four WSR-88D radars KTLX, KVNX, KINX, KFDR and their associated coverage region are also shown in Fig. 4.12. The wavelet line near the KTLX radar is the damage path of the 8 May 2003 OKC tornado.

The ARPS system is used as the prediction model. Lin 3-category ice microphysics scheme as well as a 1.5-order turbulent kinetic energy subgrid parameterization is adopted for the model run. A wave radiation condition is applied at the top boundary and rigid-wall conditions are applied to the bottom boundary. The lateral boundaries are forced externally by the forecast from a 9-km data assimilation experiment. This 9-km data assimilation experiment is done in the same way as in Hu and Xue (2007). It assimilates rawinsonde data and wind profiler data every 1hr for a total of six hours. The Eta model analysis and forecast provide the background and lateral boundaries for the 9-km experiment.

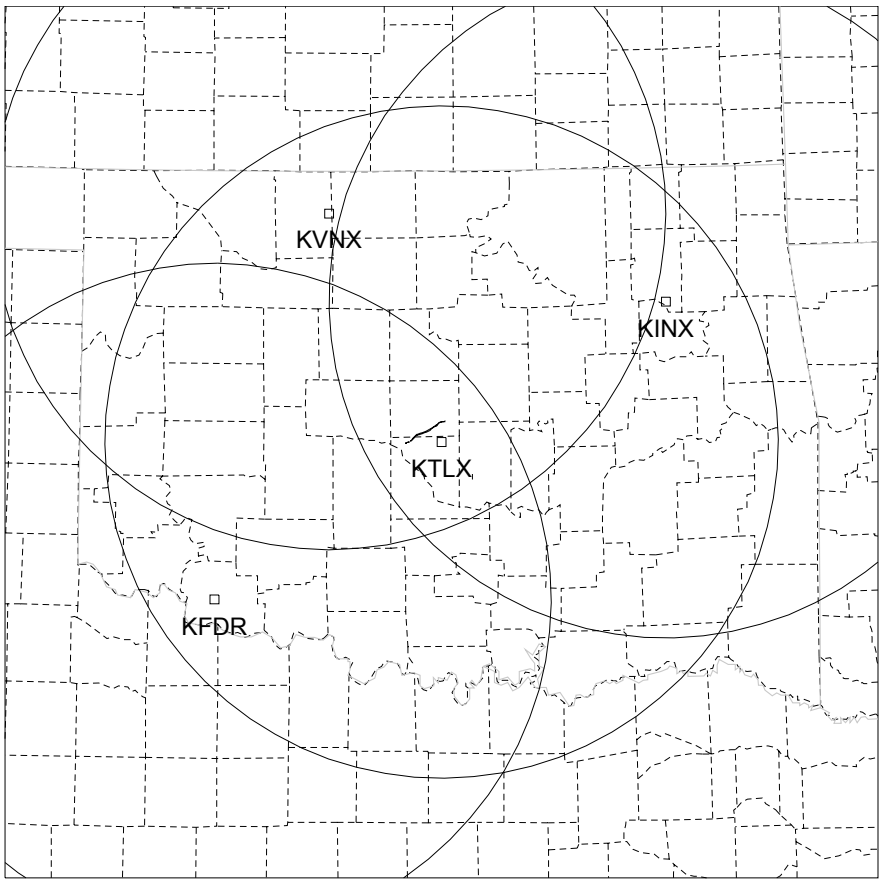

Fig. 4.12. The model domain with county boundaries. The four radars as well as their coverage circle are also shown. The wavelet line near KTLX radar shows the damage path of the 8 May 2003 OKC tornado.

The 8 May 2003 Oklahoma City tornadic thunderstorm case is observed by four WSR-88D radars (KTLX, KVNX, KINX, KFDR, Fig. 4.12) in NEXRAD network. The KTLX radar is close to the storm, it observes the lower to middle parts of the storm, and the other three radars KINX, KVNX, KFDR observe the mid to higher part of the storm. Observations from all these four radars will help produce a more complete picture of the thunderstorm as discussed in Chapter 2. A quality control procedure will first be applied to the radar data before it is ingested into the data assimilation step. This includes clutter removal, velocity dealiasing, etc. After the quality check and processing, the radar data will be projected into the model grid space in the form of a series of column observations.

This will be referred as "gridded" radar data. When ingesting the radial velocity data into model, the abnormal large values of absolute difference between the gridded radar observations with the background counterparts will be denied by the ARPS 3DVAR subsystem. For the reflectivity data from multiple radars, a mosaic is made before starting the cloud analysis subsystem.

The radial velocity data and reflectivity data are both assimilated from all the four radars mentioned in the above. In order to mitigate the negative impact of small spurious cells, the noisy data in the radar observation can be discarded according to a user-specified reflectivity threshold. In our research, only the data where the observed reflectivity is larger than 25 dBZ is used for the data assimilation experiments. The CNTL experiment started at 2100UTC when a strong cell is observed by the KTLX radar. After assimilating the radial velocity data and the reflectivity data, a 5-minute forecast is made using the ARPS model and new radar data is ingested into the model again. This process is repeated until a final analysis is made at 2140UTC after a total of forty minutes data assimilation with a frequency of every five minutes. At the end of the assimilation, there are thirty minutes left before the OKC tornado first touches down near the Moore area. The final analysis is used to launch a 2h-20min forecast. The forecast stops at 000UTC 9 May 2003. At that time, the storm is in its later dissipating stage.

Four different data assimilation experiments (see Table 4.3) are conducted to demonstrate the impact of the diagnostic pressure equation constraint and the sensitivity to its associated weighting coefficient. The aforementioned ARPS 3DVAR system and assimilation configurations/procedures are applied to all the four experiments except that the first experiment, the "NODP" experiment, does not use the diagnostic pressure

equation constraint while the other three use the constraint. The mass continuity equation constraint is included for all experiments. The last three experiment, the "DP7E-7", "DP7E-8" and "DP7E-9" experiments differ only on the weighting coefficients they used for the DP constraint. The specific coefficient values are lists in Table 4.3 and also contained within the experiment names.

Table 4.3. List of experiments

Expr. Name	NODP	DP7E-7	DP7E-8	DP7E-9
Description	DP constraint is not used	DP weighting coefficient is 7.0E-7	DP weighting coefficient is 7.0E-8	DP weighting coefficient is 7.0E-9

Since no observations other than radar data can resolve the storm-scale features in a high spatial and temporal resolution, the reflectivity observations at 1.45° tilt from KTLX radar are used to assist the evaluation of the forecast quality of the three experiments.

4.4.3 Results of experiments

4.4.3.1 The analyses of the experiments

For the idealized case study in Section 4.3, we can compare the data assimilation results to the truth simulation to examine the impact of the diagnostic pressure equation constraint. However, for real data experiments here, the truth is unknown so it is difficult to evaluate directly the quality of analyses from the four different data assimilation experiments. Some methods are introduced to check the dynamic consistency among model variables in an analysis. First, the momentum checking method is used and it does yield smaller Er value for experiments using the diagnostic pressure equation constraint (the DP7E-7, DP7E-8 and DP7E-9 experiments) than the experiment not using it (the

NODP experiment). This partially indicates that a more dynamic consistent analysis is achieved by the use of the diagnostic pressure equation constraint. The "acoustic wave checking" method similar to Hu et al (2006b) is also used. It shows that the acoustic oscillation amplitude, presented in the time series of the pressure field (not shown), is reduced in the "DP7E-7", "DP7E-8" and "DP7E-9" experiments over the "NODP" experiment. This partially justifies in a different way that the dynamic consistency is improved by the including of the constraint.

However, in spite of the above checking, it is still interesting to examine how the model fields are affected in single analysis by the diagnostic pressure equation constraint. Fig. 4.13 shows the vertical distribution of the RMS errors between the analyses from the "NODP" experiment and each of the "DP7E-7", "DP7E-8", "DP7E-9" experiments valid at the beginning of the data assimilation (2100 UTC 8 May 2003) when the backgrounds for all the four experiments are the same. The statistics is calculated only in rainy area (where the observed reflectivity mosaic is larger than 5dBZ). It can be seen from Fig. 4.13 that the use of the diagnostic pressure equation constraint leads to evident difference in the analysis. The difference in the horizontal wind fields is noticeable in nearly entire vertical direction. The difference in the vertical velocity, the potential temperature and water vapor fields mainly lie in mid-upper levels (5km~10km, 3km~8km and 2.5km~7.5km respectively). Fig. 4.14 shows the U component of perturbation wind fields, vertical velocity, perturbation potential temperature and water vapor mixing ratio at Z=7km MSL for all the four experiments. It can be seen that when the diagnostic pressure equation constraint plays more important role in the cost function (i.e. with smaller weighting coefficient), more analysis differences present between the experiment

that imposes the constraint and the "NODP" experiment that does not use the constraint. The impact of the diagnostic pressure constraint on analysis includes controlling noise introduced by the ingestion of radar observations (Fig. 4.14f,g,h vs. Fig. 4.14e) and applying some kind of smoothness (Fig. 4.14j,k,l vs. Fig. 4.14i and Fig. 4.14n,o,p vs. Fig. 4.14m). It is also clear that when the diagnostic pressure equation constraint plays more important role in the cost function (from Fig. 4.14b to Fig. 4.14d), smaller perturbation of U component of wind fields is analyzed (Fig. 4.14b,c,d vs. Fig. 4.14a). This indicates that the observation impact might be traded off with dynamic consistency among model variables.

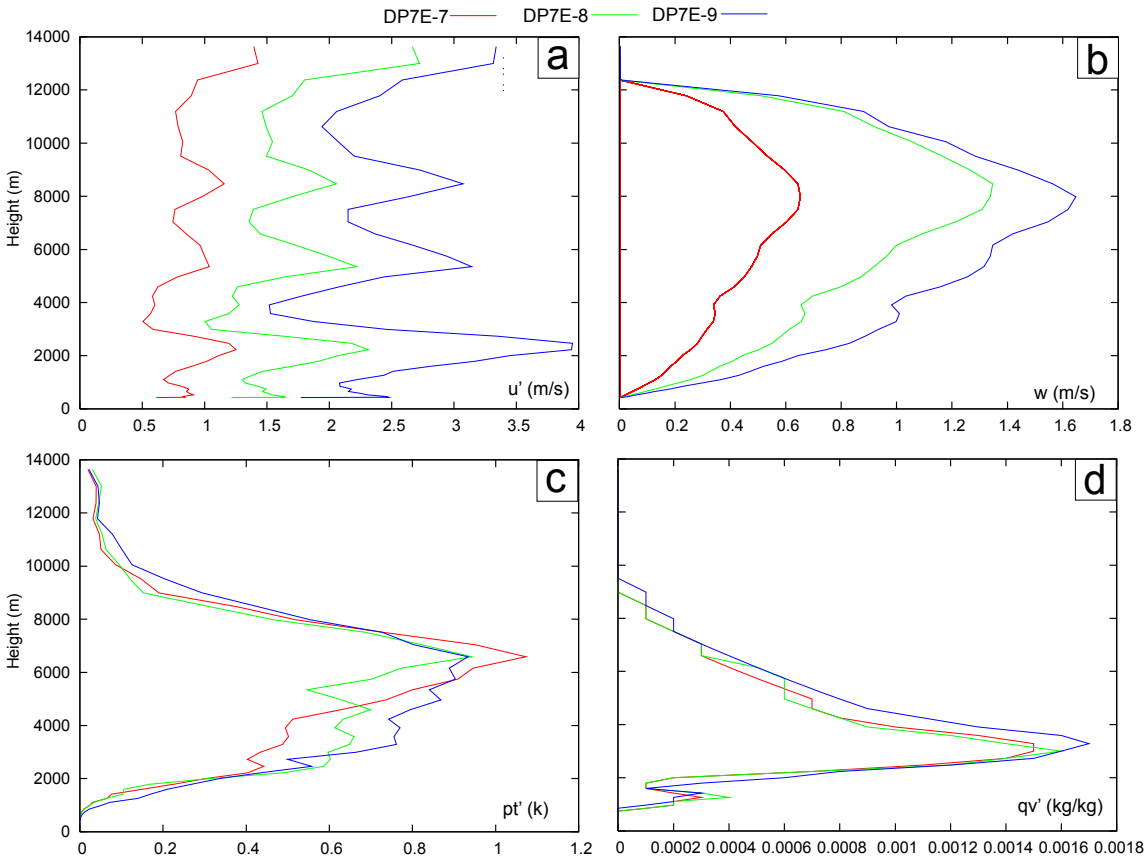

Fig. 4.13. The vertical distribution of the difference, in terms of the RMS errors, between the analyses from the "NODP" experiment and each of the "DP7E-7"(red line) , "DP7E8"(green line), "DP7E-9"(blue line) experiments valid at 2100 UTC 8 May 2003

(the beginning of the data assimilation). The statistics is calculated only in rainy area (where the observed reflectivity mosaic is larger than 5dBZ).

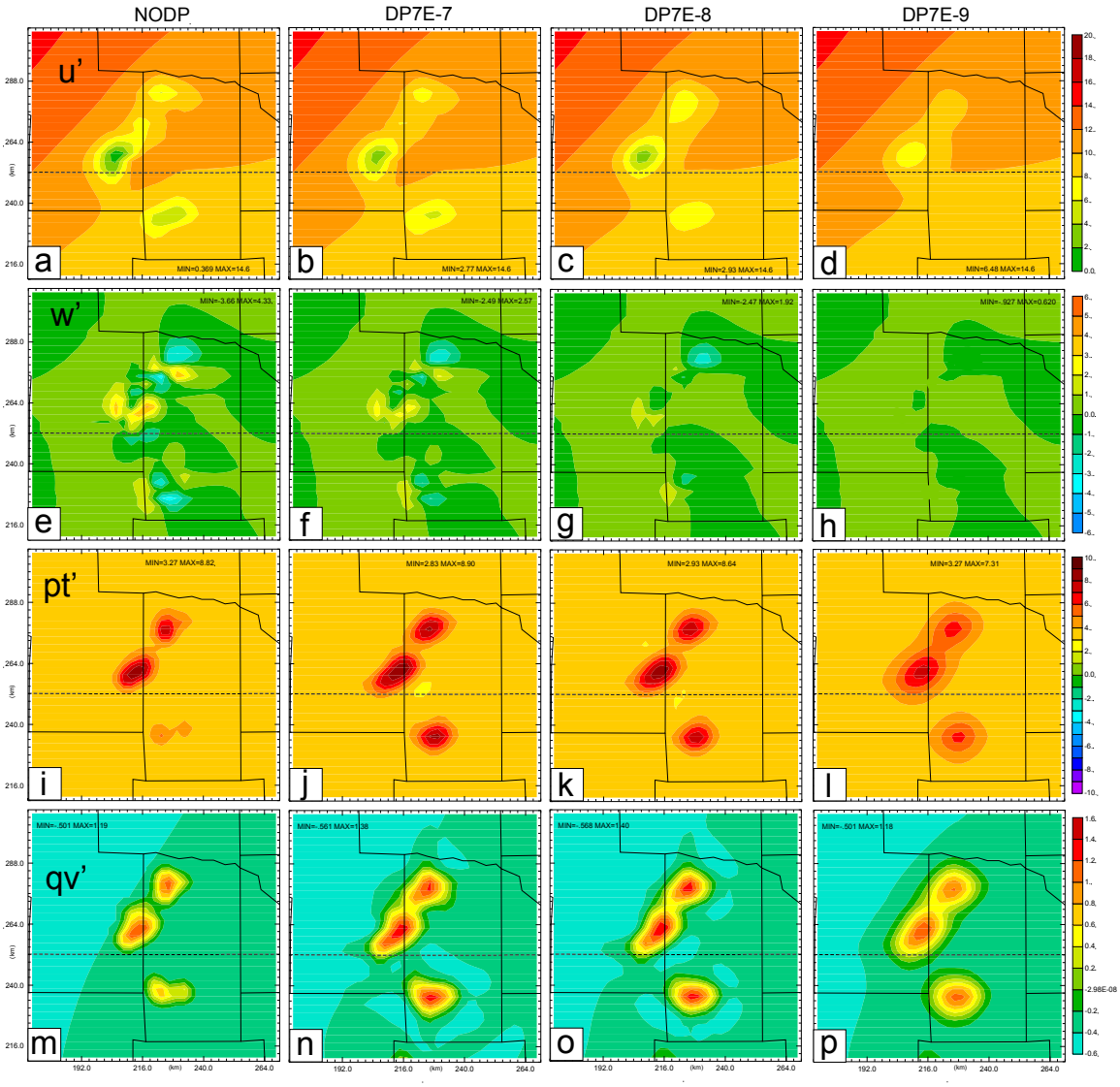

Fig. 4.14. The U component of perturbation wind fields, vertical velocity, perturbation potential temperature and water vapor mixing ratio at z=7km MSL at 2100 UTC 8 May 2003. (a),(e),(i),(m) for the "NODP" experiment, (b),(f),(j),(n) for the "DP7E-7" experiment, (c),(g),(k),(o) for the "DP7E-8" experiment, (d),(h),(l),(p) for the "DP7E-9" experiment.

After forty minutes of data assimilation at 2140UTC 8 May 2003, the analysis differences among the four experiments become more noticeable as Fig. 4.15 shows

much larger RMS errors in all vertical levels. The most noticeable difference in the vertical velocity still lies in the mid-upper levels (about 5km~10km). The difference in the potential temperature field increases with altitude. Fig. 4.16 further shows the U component of perturbation wind fields, vertical velocity, perturbation potential temperature and water vapor mixing ratio at z=7km MSL. The most noticeable difference is that the southeast and southwest updraft centers (A and B in Fig. 4.16e) get stronger in the data assimilation experiments that use the diagnostic pressure equation constraint (Fig. 4.16f,g,h vs. Fig. 4.16a). The southeast warm center (D in Fig. 4.16i) and the warm tongue (E in Fig. 4.16i) also get warmer. The difference in other fields is not so evident, but still distinguishable.

Because there is no simple balance for storm-scale phenomena, it is difficult to explain what kind of balance has been build up by the use of the diagnostic pressure equation constraint. However, the results show that the impact of the constraint on the analysis is very noticeable. The direct verification of whether this impact is positive is limited since there is no reliable high-resolution analysis of the storm. However, From the idealized case study in Section 4.3, where the truth is known, it has been demonstrated that the diagnostic pressure equation constraint does have positive impact on the intermittent data assimilation.

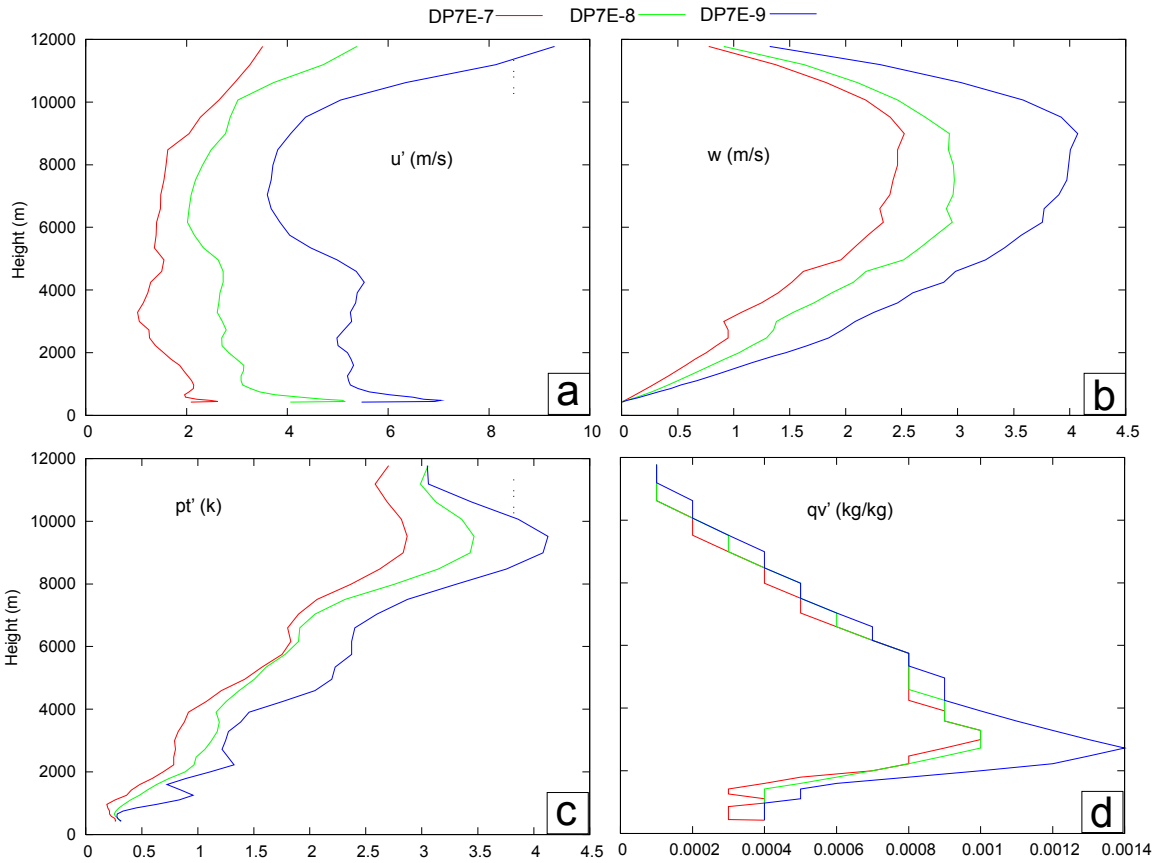

Fig. 4.15. The vertical distribution of the RMS errors between the analyses from the "NODP" experiment and each of the "DP7E-7"(red line) , "DP7E8"(green line), "DP7E-9"(blue line) experiments valid at 2140 UTC 8 May 2003 (the beginning of the data assimilation). The statistics is calculated only in rainy area (where the observed reflectivity mosaic is larger than 5dBZ).

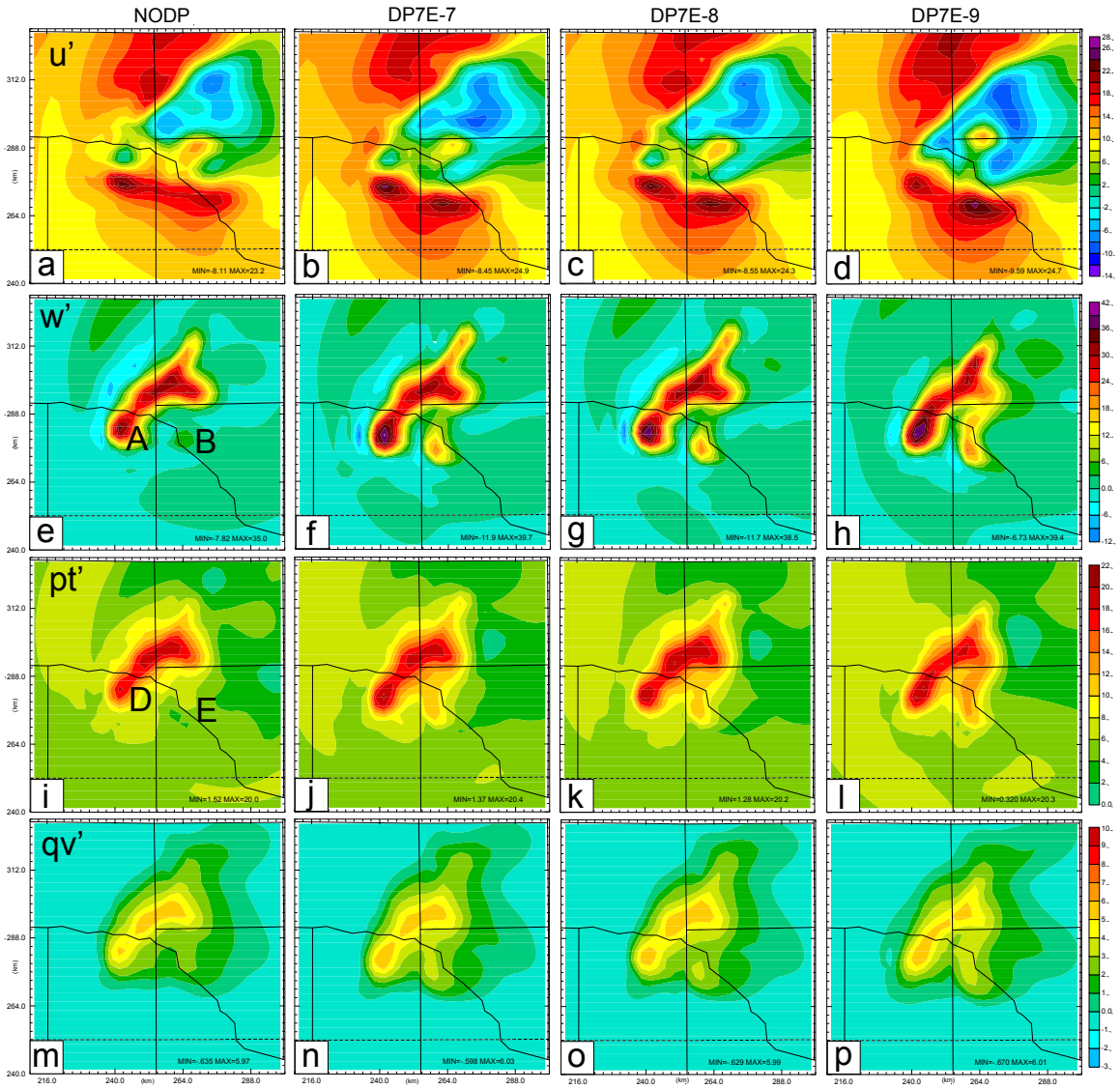

Fig. 4.16. The U component of perturbation wind fields, vertical velocity, perturbation potential temperature and water vapor mixing ratio at z=7km MSL at 2140 UTC 8 May 2003. (a),(e),(i),(m) for the "NODP" experiment, (b),(f),(j),(n) for the "DP7E-7" experiment, (c),(g),(k),(o) for the "DP7E-8" experiment, (d),(h),(l),(p) for the "DP7E-9" experiment.

4.4.3.2 The forecasts of the experiments

From the final analysis at 2140UTC, a 2h 20min forecast is made for all the four experiments. To evaluate the quality of the forecast, simulated reflectivity, using reflectivity forward operator as in Tong and Xue (2005), is produced from the forecasts

and then projected to the 1.45 elevation angle in the KTLX radar observation space. These emulated radar echoes are then compared with the observed reflectivity from KTLX radar at the same elevation angle.

Fig. 4.17 shows the evolution of the emulated radar echoes every thirty minutes from all the four experiments. It can be seen the general evolution of the OKC tornadic thunderstorm is predicted very well by the "DP7E-7", "DP7E-8" and "DP7E-9" experiments (Fig. 4.17b,c,d). The storm intensifies into its mature stage by 2200UTC. The mature stage maintains until 2235UTC. Since 2240UTC, the thunderstorm starts to weaken but in a slow pace. After 2330UTC, the storm enters its later dissipating stage and weakens quickly. The "NODP" experiment also makes a reasonable forecast. Similar to the observed evolution, the storm also experiences the intensifying stage, the mature stage and the dissipating stage during the whole 2h 20min forecast period. However, Fig. 4.17a shows a very quick weakening process since 2300UTC. After 2340UTC (not shown in the figure), there is no more high reflectivity area (>45dBZ). Therefore, regarding the prediction of the general evolution of the OKC tornado thunderstorm, the use of the diagnostic pressure equation constraint produces better results. The experiments using different weighting coefficients yield little difference in the forecast except that the "DP7E-9" experiment predicts a strong spurious storm cells since 2340UTC.

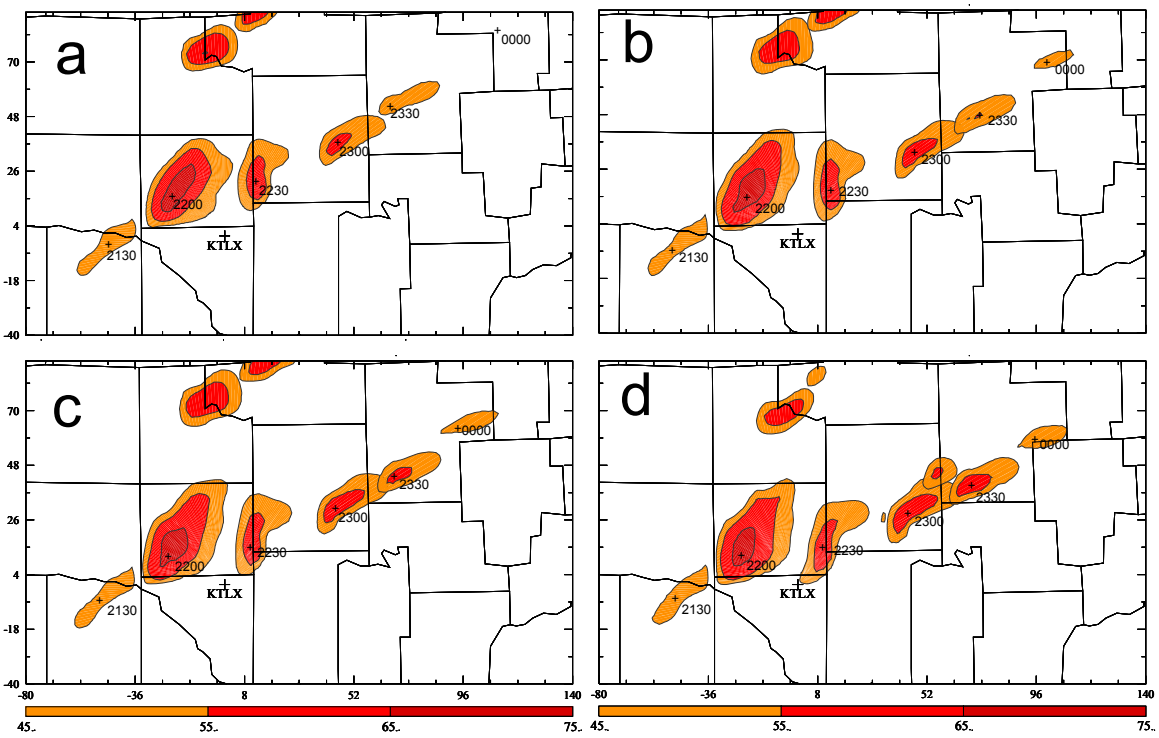

Fig. 4.17. Regions of predicted radar echoes at the 1.45 elevation angle, from 2130UTC 8 May 2003 to 0000 UTC 9 May 2003 every 30 minutes, by the experiments: (a) NODP, (b) DP7E-7, (c) DP7E-8, (d) DP7E-9. The contours are plotted every 10dBz, starting from 45dBz. The plus sign indicates the reflectivity center and is followed by the corresponding time in UTC.

To quantitatively evaluate the quality of the forecast, the ETS (Equitable Threat Score, Schaefer 1990) of composite reflectivity against reflectivity mosaic from the four radars is calculated for all the experiments. Fig. 4.18 shows the results for the 5-, 15-, 30- and 45-dBZ thresholds. It can be seen that for the 5-dBZ threshold, the ETS values of the experiments "DP7E-7", "DP7E-8", "DP7E-9" are generally better than the "NODP" experiment except during the period from about 2220UTC to 2300UTC, when there is almost no forecast score in all experiments due to relative large storm location errors. For the 15-and 30-dBZ threshold, the ETS values of the experiments "DP7E-7", "DP7E-8", "DP7E-9" are much better than the "NODP" experiment after 2255UTC. For the 45-dBZ threshold, a better score is produced after 2330UTC for the experiments "DP7E-7",

"DP7E-8", "DP7E-9" over the "NODP" experiment, which corresponds to a better forecast of the storm dissipating stage. Therefore, it can be confirmed that the use of the diagnostic pressure equation constraint helps improve the forecast in term of the prediction of the general evolution of the major thunderstorm. The experiments adopting different weight coefficients produce similar forecasts. The experiment using the smallest weight coefficient (i.e. 7.0E-9) produces a small strong spurious cell at the later dissipating stage of the major storm, but in spite of that, the forecast is still reasonable and not far from that of the "DP7E-7" and "DP7E-8" experiments.

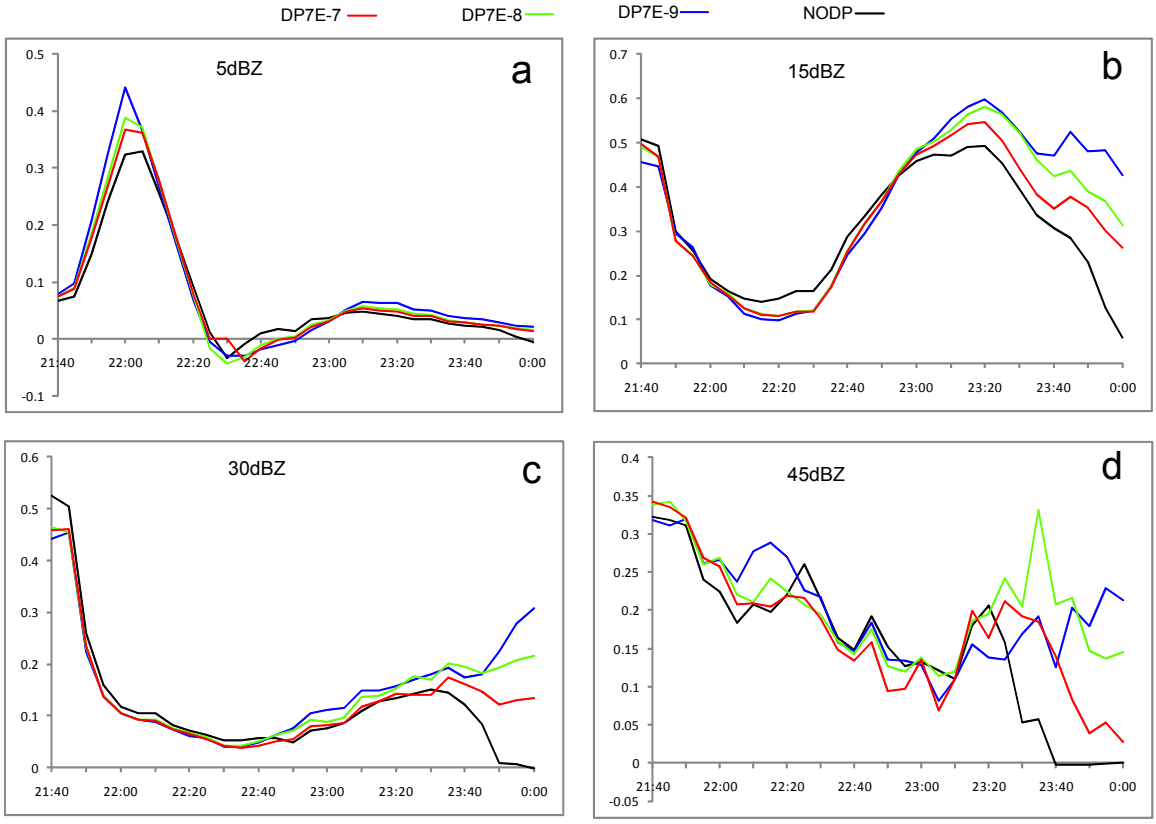

Fig. 4.18. Equitable threat scores of predicted composite reflectivity for the (a) 5-, (b) 15-, (c) 30-, and (d) 45-dBZ thresholds. The black line is from the "NODP" experiment, the red line is form the "DP7E-7" experiment, the green line is from the "DP7E-8" experiment, the blue line is from the "DP7E-9" experiment.

In order to further examine the difference among the experiments, the maximum vertical vorticity in the low two kilometers of the atmosphere is computed from the forecast every one minute. The time series is then plotted in Fig. 4.19. It can be seen that the experiments ("DP7E-7", "DP7E-8", "DP7E-9") using the diagnostic pressure equation constraint generally produce larger low-level vertical vorticity than the "NODP" experiment during the whole forecast period. What does it mean by a larger vertical vorticity? As a demonstration, Fig. 4.20 plots the vertical vorticity field at 2220UTC, which is during the tornado touchdown period. Fig. 4.20a,c,e,g present the vertical vorticity at z=3km MSL while Fig. 4.20b,d,f,h show the vertical vorticity in the cross sections with largest vertical vorticity, i.e. along the lines A-B in their corresponding left panels. It can be seen that the vertical vorticity at z=3km MSL predicted by the "DP7E-7", "DP7E-8" and "DP7E-9" (Fig. 4.20c,e,g) experiments is larger than that predicted by the "NODP" experiment (Fig. 4.20a). The vertical cross sections (Fig. 4.20b,d,f,h) show more difference in the vertical vorticity structure among different experiments. The experiments using the diagnostic pressure equation constraint predict a deeper column of high vertical vorticity ($> 0.008\ s^{-1}$), extending from as low as 1.2 kilometers to as high as 9.5 kilometers. The region of high vertical vorticity ($> 0.008\ s^{-1}$) predicted by the "NODP" experiment is mainly in the mid-upper part of the atmosphere, roughly from 4.5 kilometers to 9.5 kilometers.

Our further examination shows that the deeper rotated vortex column with larger vertical vorticity generally means a better-defined supercell structures. As an example, Fig. 4.21 shows the simulated reflectivity field and the wind vectors at z=3km MSL for all the four experiments at 2200UTC. It can be seen that the "DP7E-8" experiment (Fig.

4.21c) successfully predicts the hook echo signature at this time. The reflectivity hook echo, the clearly visible rotation of the horizontal wind vectors, the area of very large vertical vorticity and the intense updraft (not shown) are very well defined to support each other. This is a very strong signal indicating the location of the mesocyclone associated with the OKC major tornado. The "DP7E-7" and the "DP7E-9" experiments (Fig. 4.21b, d) produce similar forecasts. The hook echo signs are also distinguishable in both experiments. However, this conclusion can barely be drawn from the forecast of the "NODP" experiment (Fig. 4.21a). The hook echo signature is very weak. The rotation in the wind fields near the storm center is much mild than that predicted by the other three experiments that all use the diagnostic pressure equation constraint. Therefore, it can be concluded that the use of the diagnostic pressure equation constraint helps make a better forecast in terms of the mesocyclone rotation and the supercell structures.

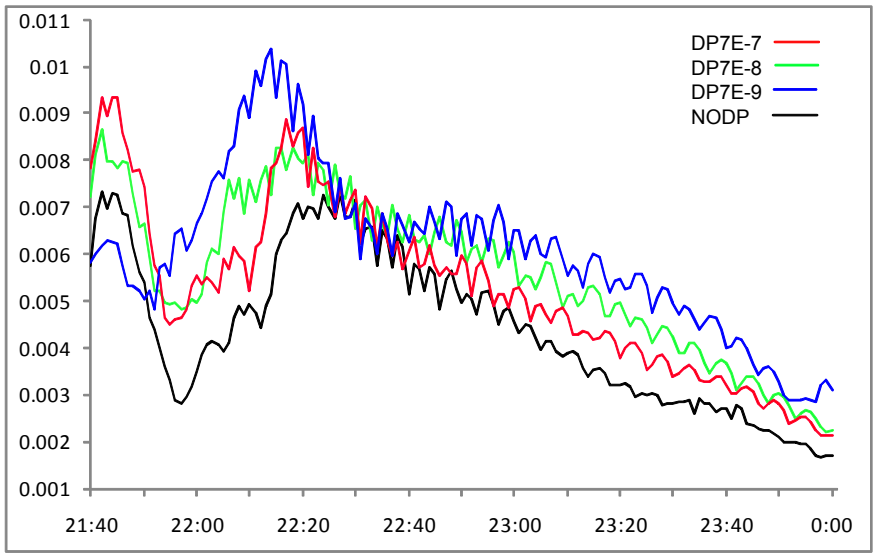

Fig. 4.19. The time series of maximum vertical vorticity below two kilometers from 2140UTC 8 May 2003 to 0000UTC 9 May 2003 every one minute. The horizontal axis shows the time in UTC, the vertical axis shows the vertical vorticity value in unit of s^{-1}. The black line is for the "NODP" experiment, the red line is for the "DP7E-7" experiment, the green line is for the "DP7E-8" experiment, the blue line is for the "DP7E-9" experiment.

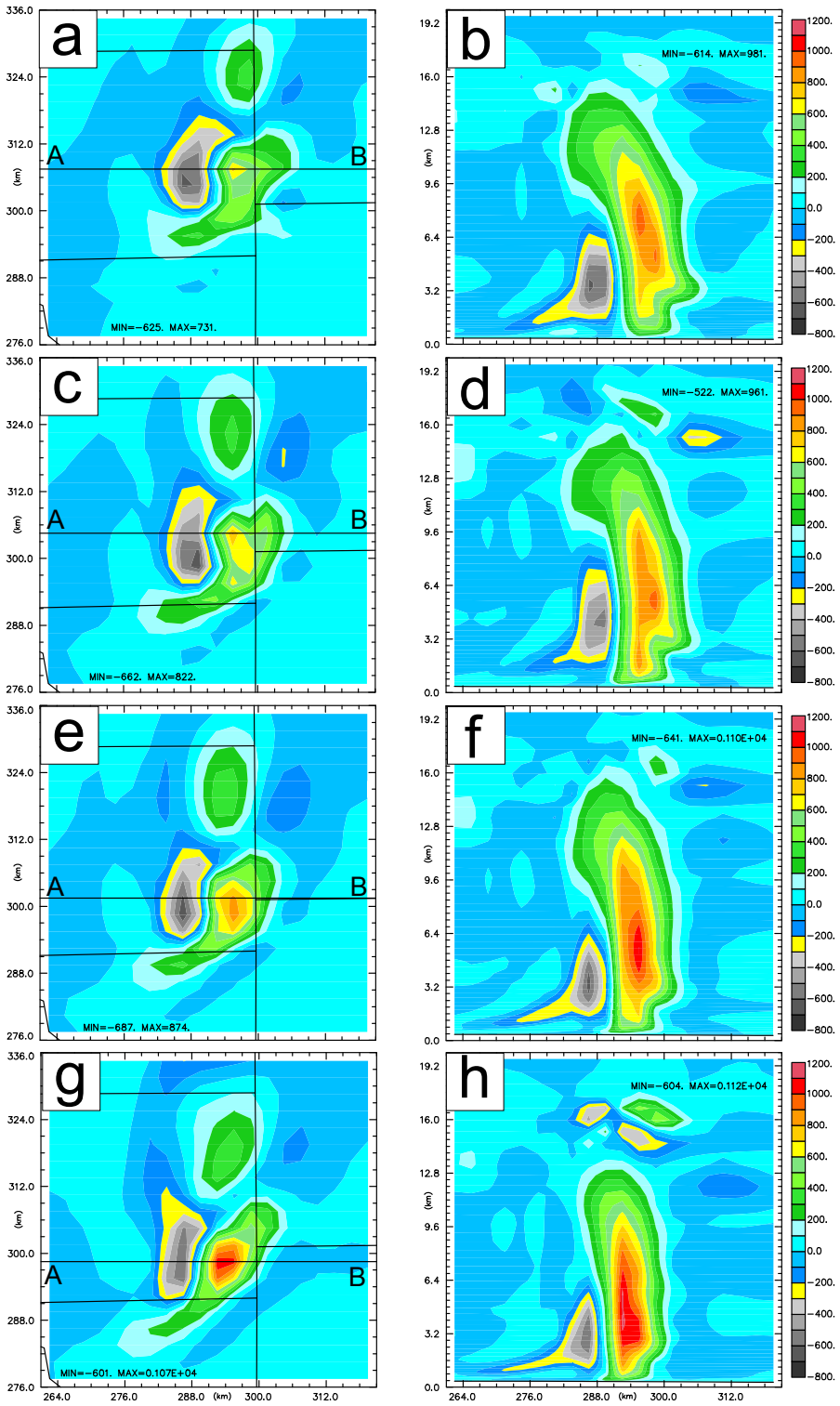

Fig. 4.20. The vertical vorticity (in unit of 10^{-5} s^{-1}) at 2200UTC 8 May 2003 predicted by the four experiments. (a) and (b) are for the "NODP", (c) and (d) are for the "DP7E-7", (e) and (f) are for the "DP7E-8", (g) and (h) are for the "DP7E-9". (a), (c), (e) and (g) are at z=3km MSL; (b), (d), (f) and (h) are the vertical cross sections along the lines A-B in their left corresponding panel respectively.

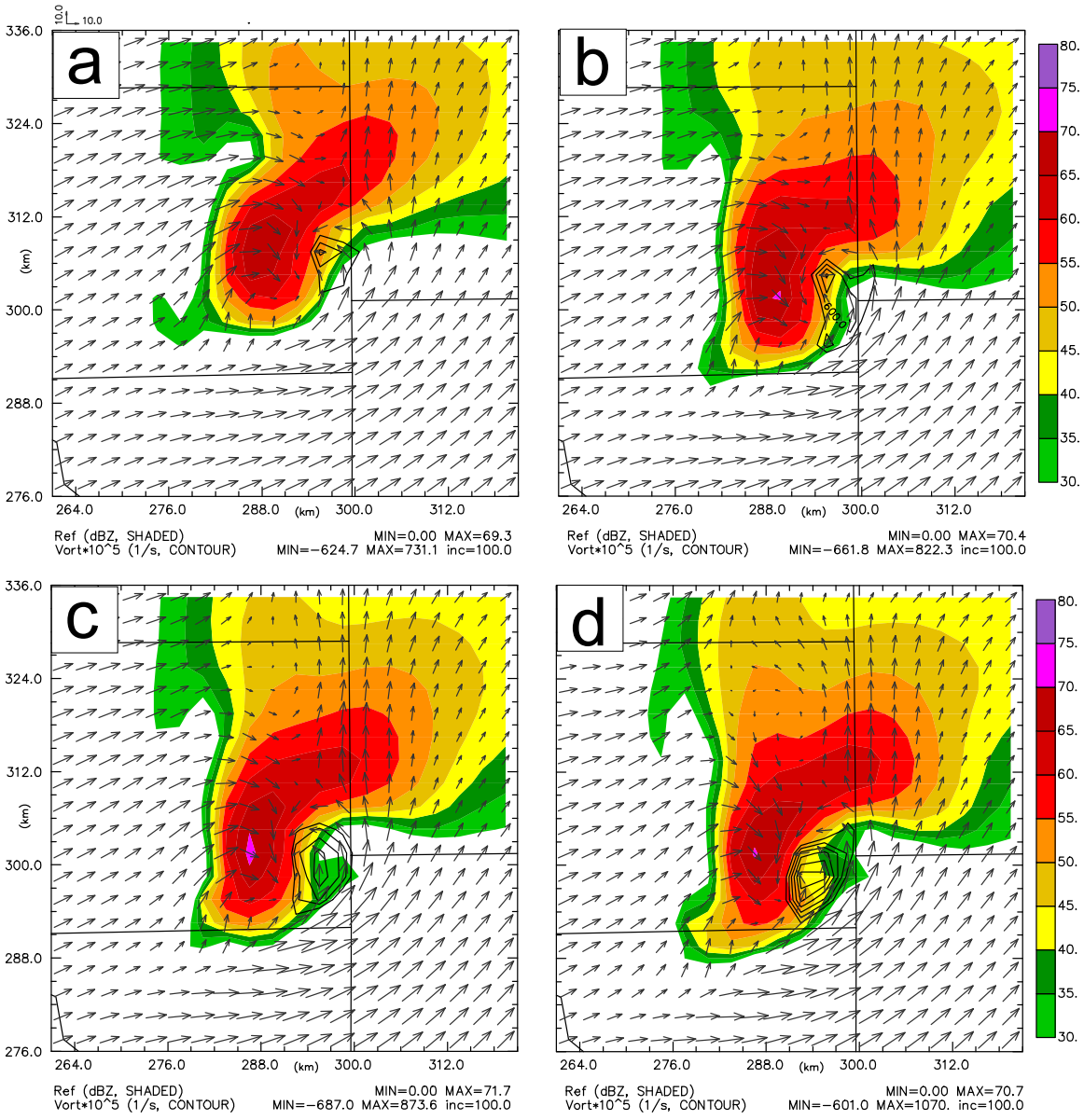

Fig. 4.21. The reflectivity field and wind vectors at z=3km MSL at 2200UTC 8 May 2003 predicted by the experiments (a) NODP, (b) DP7E-7, (c) DP7E-8, (d) DP7E-9. In order to indicate the possible location of mesocyclone, the vertical vorticity larger than $0.005s^{-1}$ is also plotted in (a) and (b).

4.4.4 Conclusions

In this case study, we apply the updated ARPS 3DVAR scheme to the assimilation and forecast of the 8 May 2003 Oklahoma City tornadic supercell thunderstorm in order

to investigate the impact of the diagnostic pressure equation constraint. Four data assimilation experiments are conducted with different weighting coefficients for the diagnostic pressure equation constraint (the "NODP" experiment does not use the constraint). All the experiments assimilate the same amount of observations from the four NEXRAD radars and impose the mass continuity equation constraint.

It is shown that in single analysis step, the diagnostic pressure equation constraint can help control the noise introduced by the assimilation of radar observations and couple different model variables to boost dynamic consistency. The "momentum checking" and "acoustic wave checking" methods are also used to partially justify that the constraint does improve dynamic balance among model variables in single analysis step. After forty-minutes of intermittent data assimilation, the diagnostic pressure equation constraint evidently affects the final analysis. The differences generally lie in the mid-upper levels and larger vertical velocity and potential temperature are produced by the use of the constraint. Because there is no reliable high-resolution analysis of the storm, it is not easy to tell directly which analysis is better. The evaluation of the benefit of this constraint to radar data assimilation in this real case is then further examined mainly based on the ensuing forecast.

It is demonstrated that the experiments using the diagnostic pressure equation constraint predict the intensification and the dissipation process of the storm in a good timing and location manner. On the other hand, the "NODP" experiment, which does not use the constraint, predicts a very fast dissipating stage for the storm, which is not supported by the observations. During the whole forecast period, the experiments using the constraint generally predict high low-level vertical vorticity than the "NODP"

experiment. Further investigation shows that this is the reflection of a better-predicted supercell structure in terms of the hook echo sign, mesocyclone rotation, and updraft intensity. Therefore, it is concluded that the use of the constraint improves the forecast in the general evolution and the supercell characteristics of the major thunderstorm. The experiments adopting different weighting coefficients generate similar results. This suggests that the diagnostic pressure equation constraint is not very sensitive to the weighting coefficients, although very small values should still be avoided as illustrated by the idealized case study in Section 4.3.

4.5 The 5 May 2007 Greensburg tornadic supercell storm case

4.5.1 The case

The general background for this case has been described in detail in the real case part in Chapter 2 (specifically, in Section 2.3.3) and will not be repeated here.

4.5.2 Experimental design

The general rules and data assimilation configurations for the design of experiments is also discussed thoroughly in the real case part in Chapter 2 (specifically, in Section 2.3.4) and will not be repeated here. In order to investigate the impact of the diagnostic pressure equation constraint, more data assimilation experiments are conducted. Table 4.4 lists all experiments performed for the 5 May 2007 Greensburg case. All the experiments assimilate the radial velocity data and impose the mass continuity equation constraint.

Table 4.4. List of experiments.

Experiment Name	**DP weighting coefficient**	**Use of reflectivity**
Vr_NODP		No
Vr_DP_5.0E-8	5.0E-8	No
Vr_DP_1.0E-8	1.0E-8	No
Vr_DP_5.0E-9	5.0E-9	No
Vr_DP_2.5E-9	2.5E-9	No
VrZ_NODP		Yes
VrZ_DP_5.0E-8	5.0E-8	Yes
VrZ_DP_1.0E-8	1.0E-8	Yes
VrZ_DP_5.0E-9	5.0E-9	Yes
VrZ_DP_2.5E-9	2.5E-9	Yes

The first five experiments examine the impact of the diagnostic pressure equation constraint on experiments assimilating radial velocity data alone, a similar situation as the idealized case study in Section 4.3. The last five experiments examine the constraint's impact on experiments assimilating both the radial velocity data and the reflectivity data. It should be noted that the "Vr_NODP" and "VrZ_NODP" experiments here are the same as the "VrOnly" and "Vr&Rf "experiments in Chapter 2 Section 2.3.5.

4.5.3 Results of experiments

4.5.3.1 The impact of the constraint on the experiments assimilating velocity data alone

4.5.3.1.1 The data assimilation results of the experiments

Similar to the idealized case study in Section 4.3, the impact of the diagnostic pressure equation constraint at the first analysis is not easy to demonstrate. The adjustment in model fields is not very obvious. There is barely any pattern or feature difference. However, after 1-h data assimilation, the impact can be seen more clearly. Fig. 4.22 shows the divergence, relative humidity and the cloud water mixing ratio at z=6km MSL valid at 0230UTC. It can be seen that at the end of 1-h data assimilation, the "Vr_NODP" experiment (Fig. 4.22a) produces three moist centers that are saturated or almost saturated. The locations of moist centers correspond well to the observed reflectivity cells as shown in Fig. 2.19a. At the location of the southernmost center, the major storm has already partially developed (Fig. 2.20a and Fig. 4.23a) at this time, but with weaker reflectivity than observation at z=2km MSL. The other two moist centers have not developed into distinguishable storm cells yet (Fig. 4.23a).

For the "Vr_DP_1.0E-8" experiment (Fig. 4.22c), after 1-h data assimilation, it also produces several moist centers. The three major ones correspond to that in Fig. 4.22a but cover broader area and are all saturated. There exists large amount of cloud water co-located with all the three major moist centers. At z=6km MSL, the storm cells have developed rather well in these three locations (Fig. 4.23c). The major storm covers broader area. The "Vr_DP_5.0E-9" and "Vr_DP_2.5E-9" experiments (Fig. 4.22d,e; Fig. 4.23d,e) produces very similar results to that of the "Vr_DP_1.0E-8" experiment except

that the westernmost storm cell develops stronger in the "Vr_DP_5.0E-9" and "Vr_DP_2.5E-9" experiments. For the "Vr_DP_5.0E-8" experiment, although the westernmost and the northernmost storm cell do not develop yet (Fig. 4.22b), it already produces relative broad area of saturated moist air co-located with large amount of cloud water, which speeds up the spin-up problem than the "Vr_NODP" experiment. It will be discussed later.

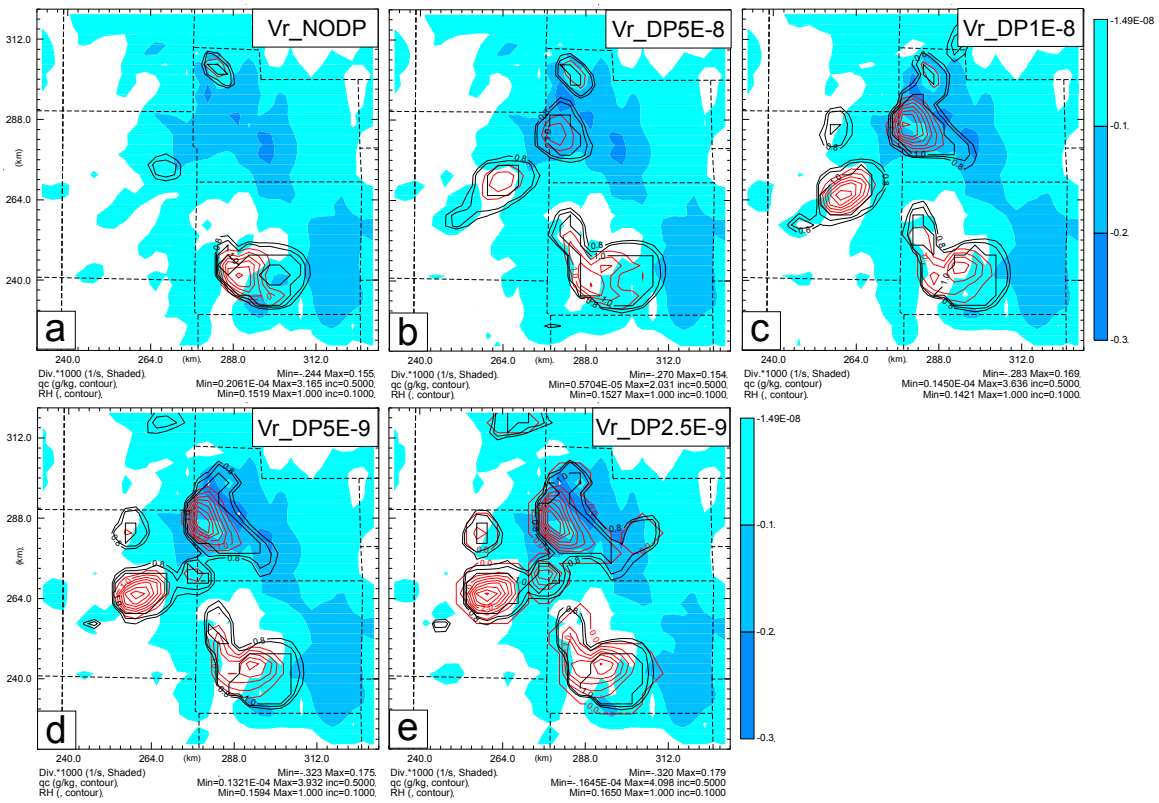

Fig. 4.22. The divergence (shaded), relative humidity (black contours) and the cloud water mixing ration (red contours) at z=6km MSL valid at 0230UTC for the experiments (a) Vr_NODP, (b) Vr_DP_5.0E-8), (c) Vr_DP_1.0E-8, (d) Vr_DP_5.0E-9 and (e) Vr_DP_2.5E-9.

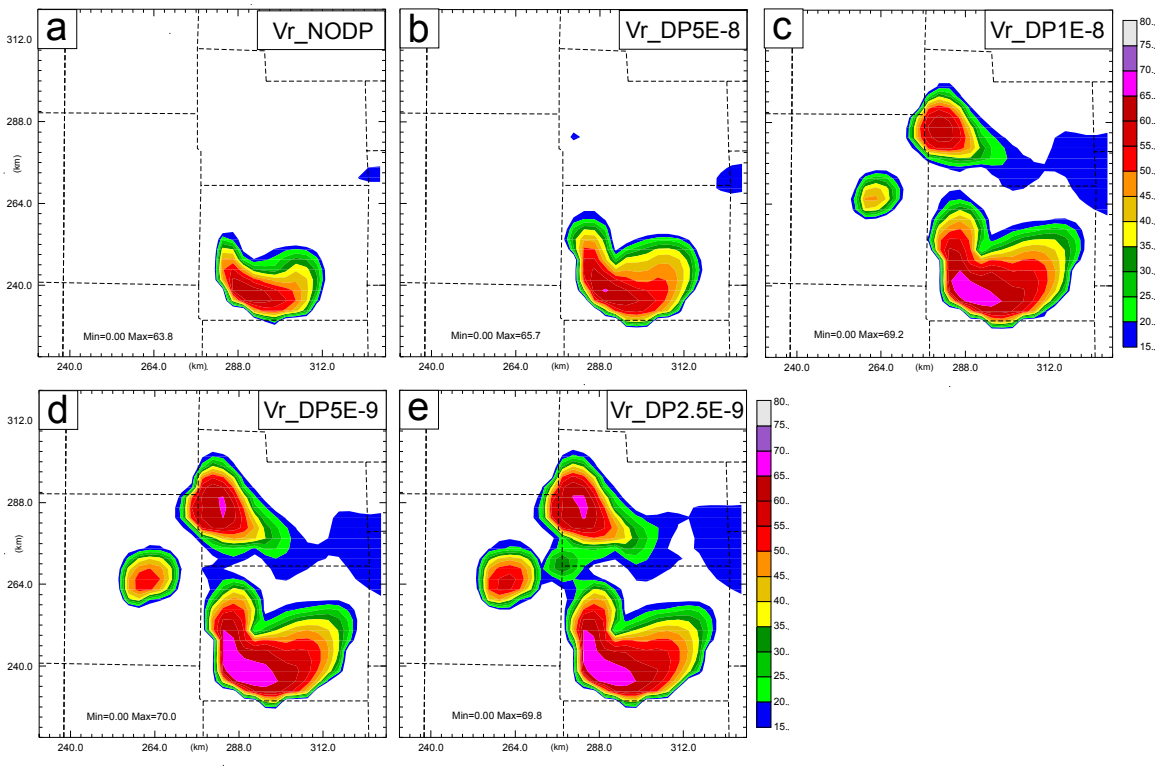

Fig. 4.23. The simulated reflectivity at z=6km MSL valid at 0230UTC for the experiments (a) Vr_NODP, (b) Vr_DP_5.0E-8), (c) Vr_DP_1.0E-8, (d) Vr_DP_5.0E-9 and (e) Vr_DP_2.5E-9.

Overall, it is demonstrated that by imposing the diagnostic pressure equation constraint, the recovered storm rainfall pattern is more close to the observed one at the end of the data assimilation and the spin-up problem is evidently reduced. This impact is also found in the idealized case study in Section 4.3 where it is demonstrated that the use of the diagnostic pressure equation constraint evidently reduce the RMS error of reflectivity after several data assimilation cycles. Therefore, it is illustrated that the use of the diagnostic pressure equation constraint helps build up a better initial condition after 1-h data assimilation of radial velocity data.

4.5.3.1.2 The forecast results of the experiments

It has been shown that the use of diagnostic pressure equation constraint does improve the analysis after 1-h intermittent data assimilation for this case. From the improved analyses, better forecasts should be expected. Fig. 4.24 shows the general evolution of the simulated reflectivity, the horizontal wind vector and the vertical vorticity at z=2km MSL from the forecast of the "Vr_DP_1.0E-8" experiment. It can be seen that at the end of data assimilation, the major storm develops stronger reflectivity than that from the "Vr_NODP" experiment (Fig. 4.24a vs. Fig. 2.20a) and closer to the observation (Fig. 2.19a). North of the major storm, a storm cell also develops in the "Vr_DP_1.0E-8" experiment (Fig. 4.24a). Although it is much weaker than the observation (Fig. 2.19a) at z=2km MSL, it has developed rather well at upper levels and has extended to low levels. On the contrary, the "Vr_NODP" experiment completely miss this cell (Fig. 2.20a) at this time. After thirty minutes at t=0300UTC (Fig. 4.24d), the major storm cells and other observed storm cells develop very well at z=2km MSL in the "Vr_DP_1.0E-8" experiment. On the other hand, the "Vr_NODP" experiment is still during rainfall spin-up period (Fig. 2.20d). In the next half hour, both the "Vr_DP_1.0E-8" and "Vr_NODP" experiments predict the major storm very well in terms of the moving path and the rainfall pattern. The major storm coverage area predicted by the "Vr_DP_1.0E-8" experiment is a little bit broader than that by the "Vr_NODP" experiment and closer to the observation. Both experiments generate some strong spurious cells at the later half hour forecasts. However, this is not our main concern.

To quantitatively evaluate the above two forecasts, the ETS of reflectivity at z=2138m MSL (the model level closest to z=2km) is computed and shown in Fig. 4.25.

The ETS for the experiments “Vr_DP_5.0E-8”, “Vr_DP_5.0E-9” and “Vr_DP_2.5E-9” are also shown in Fig. 4.25 and will be discussed later. It can be seen that for low reflectivity threshold (5-dBZ and 15dBZ, Fig. 4.25a,b), the “Vr_DP_1.0E-8” experiment (the red line) yields much higher scores than the “Vr_NODP” experiment (the black line), which agrees well with our subjective discussions in the above. For the 30-dBZ and 45dBZ thresholds, the scores of “Vr_DP_1.0E-8” experiment (the red line) is higher than that of the “Vr_NODP” experiment (the black line) during the period of 0230UTC~0300UTC. After about 0300UTC, the scores of “Vr_DP_1.0E-8” experiment are lower than that of the “Vr_NODP” experiment. This is because, as mentioned before, during the last half hour of forecast, both experiments generates strong spurious storm cells and the spurious cells are stronger and broader in the forecast of the “Vr_DP_1.0E-8” experiment. Considering the forecast of the major storm, the “Vr_DP_1.0E-8” experiment still outperforms the “Vr_NODP” experiment.

To examine the forecast results of the experiments with different weighting coefficients, Fig. 4.26, Fig. 4.27 and Fig. 4.28 show the general evolution of the simulated reflectivity, the horizontal wind vector, the vertical vorticity at z=2km MSL for the experiments “Vr_DP_5.0E-8”, “Vr_DP_5.0E-9” and “Vr_DP_2.5E-9” respectively. The corresponding ETS for these experiments are shown in Fig. 4.25. It can be seen that changing weighting coefficients does not change the forecast significantly. This indicates that the data assimilation and the forecast are not very sensitive to the weighting coefficient of the diagnostic pressure equation constraint. However, it is also noticeable that the “Vr_DP_2.5E-9” experiment produces much more spurious storm cells than other experiments. It supports our previous conclusion that a small weighting coefficient

for the diagnostic pressure equation constraint should be used very carefully. The comparison of the ETS of these experiments over the "Vr_NODP" experiment leads to similar conclusions as the previous comparison of the ETS of the "Vr_DP_1.0E-8" experiment over the "Vr_NODP" experiment. Therefore, it is demonstrated that the use of the diagnostic pressure equation constraint produces better forecast in terms of the general evolution of the storms.

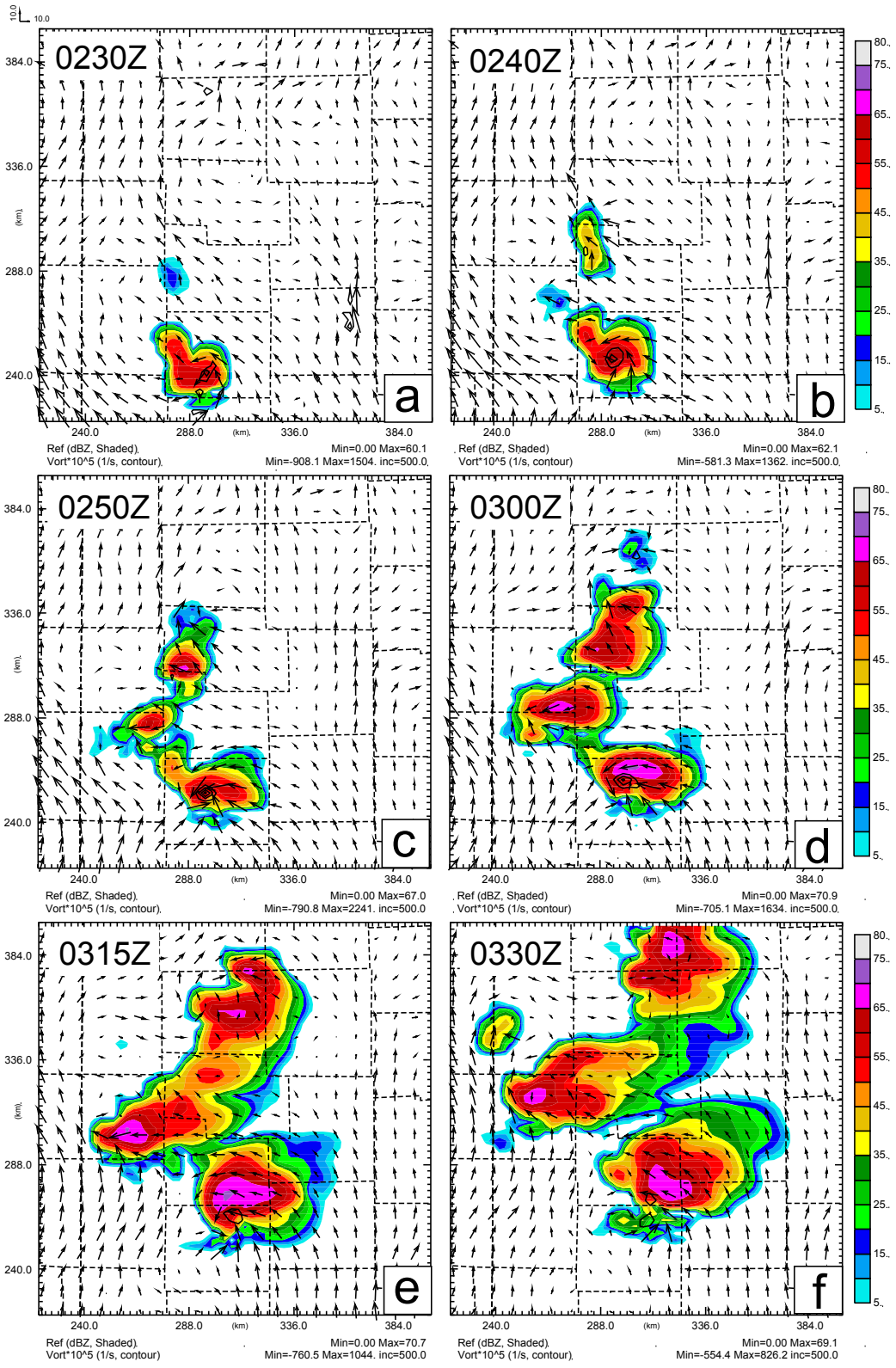

Fig. 4.24. Radar reflectivity (dBZ), horizontal wind vector, and vertical vorticity (contours staring at $0.005s^{-1}$ with an interval of $0.005s^{-1}$) at 2 km MSL from the "Vr_DP_1.0E-8" experiment during 0230~0330UTC 5 May 2007 over western Kansas.

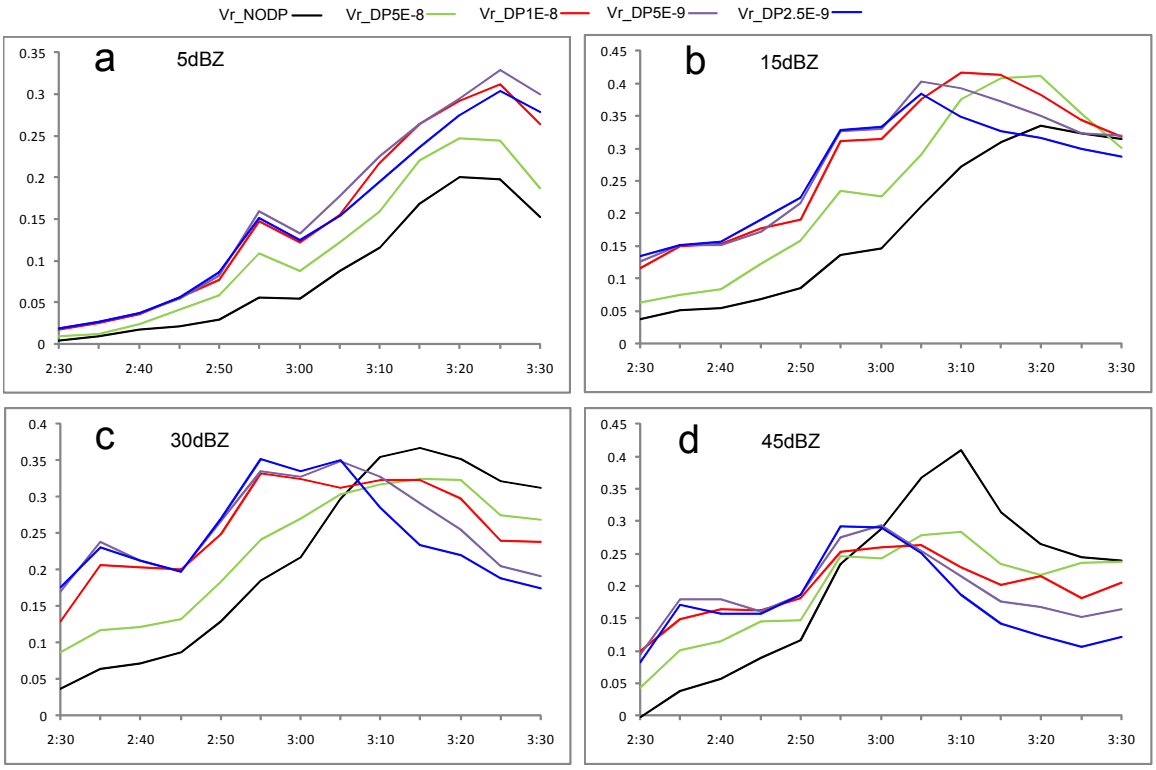

Fig. 4.25. Equitable threat scores of predicted composite reflectivity for the (a) 5-, (b) 15-, (c) 30-, and (d) 45-dBZ thresholds. The black line is from the "Vr_NODP" experiment, the green line is from the "Vr_DP_5.0E-8" experiment, the red line is form the "Vr_DP_1.0E-8" experiment, the purple line is from the "Vr_DP_5.0E-9" experiment, the blue line is from the "Vr_DP_2.5E-9" experiment.

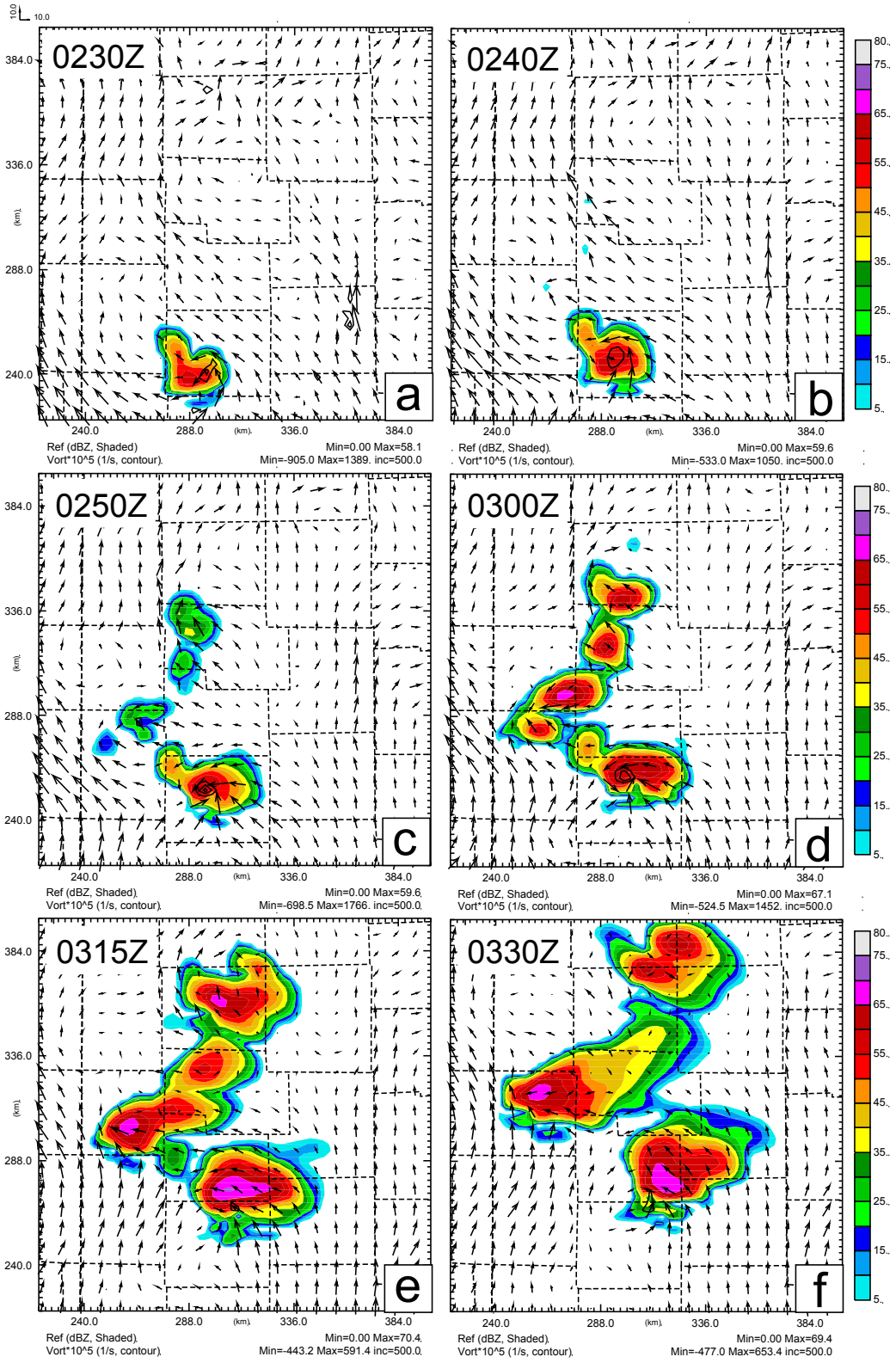

Fig. 4.26. Similar to Fig. 4.24, but for the "Vr_DP_5.0E-8" experiment.

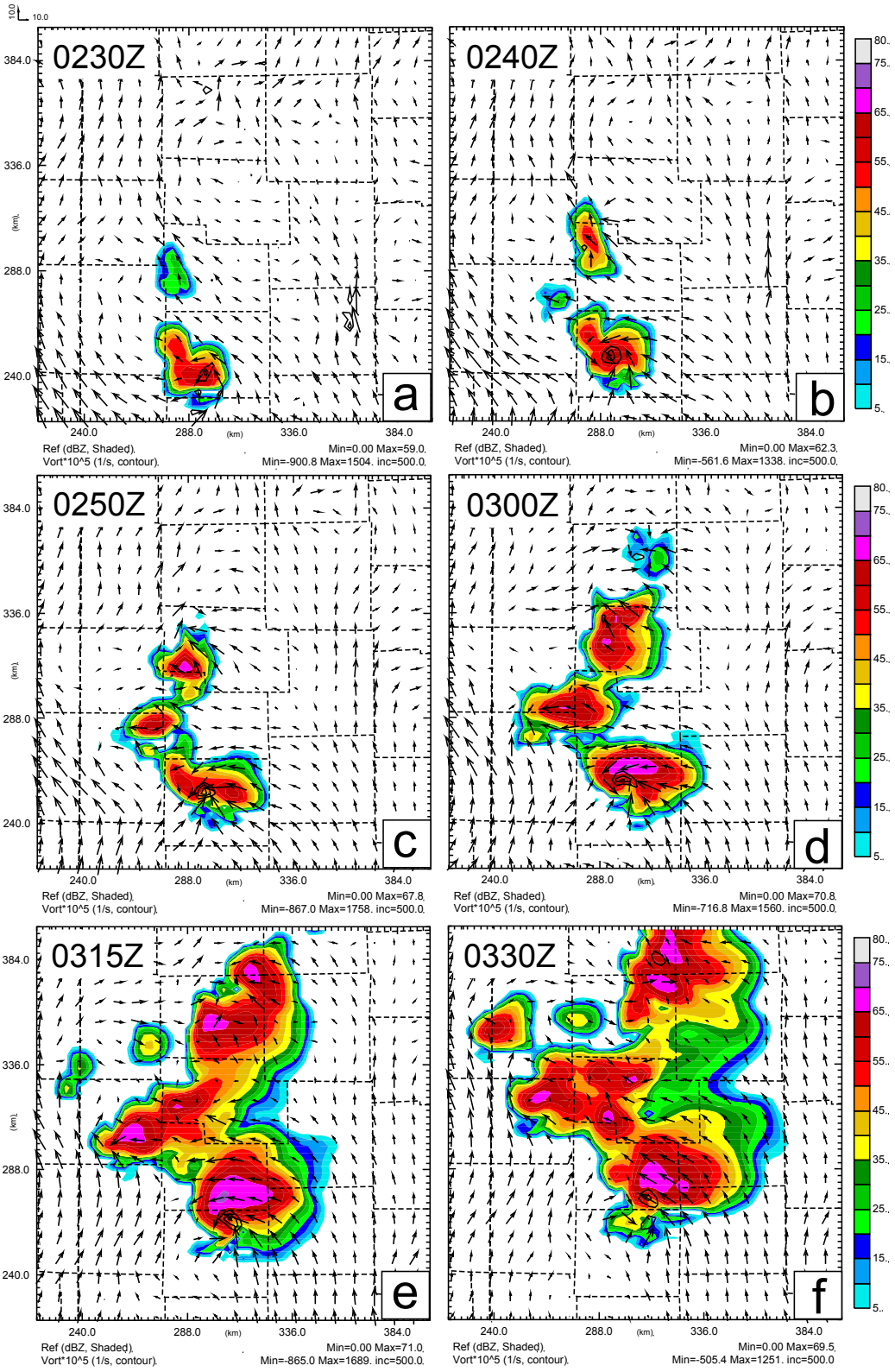

Fig. 4.27. Similar to Fig. 4.24, but for the "Vr_DP_5.0E-9" experiment.

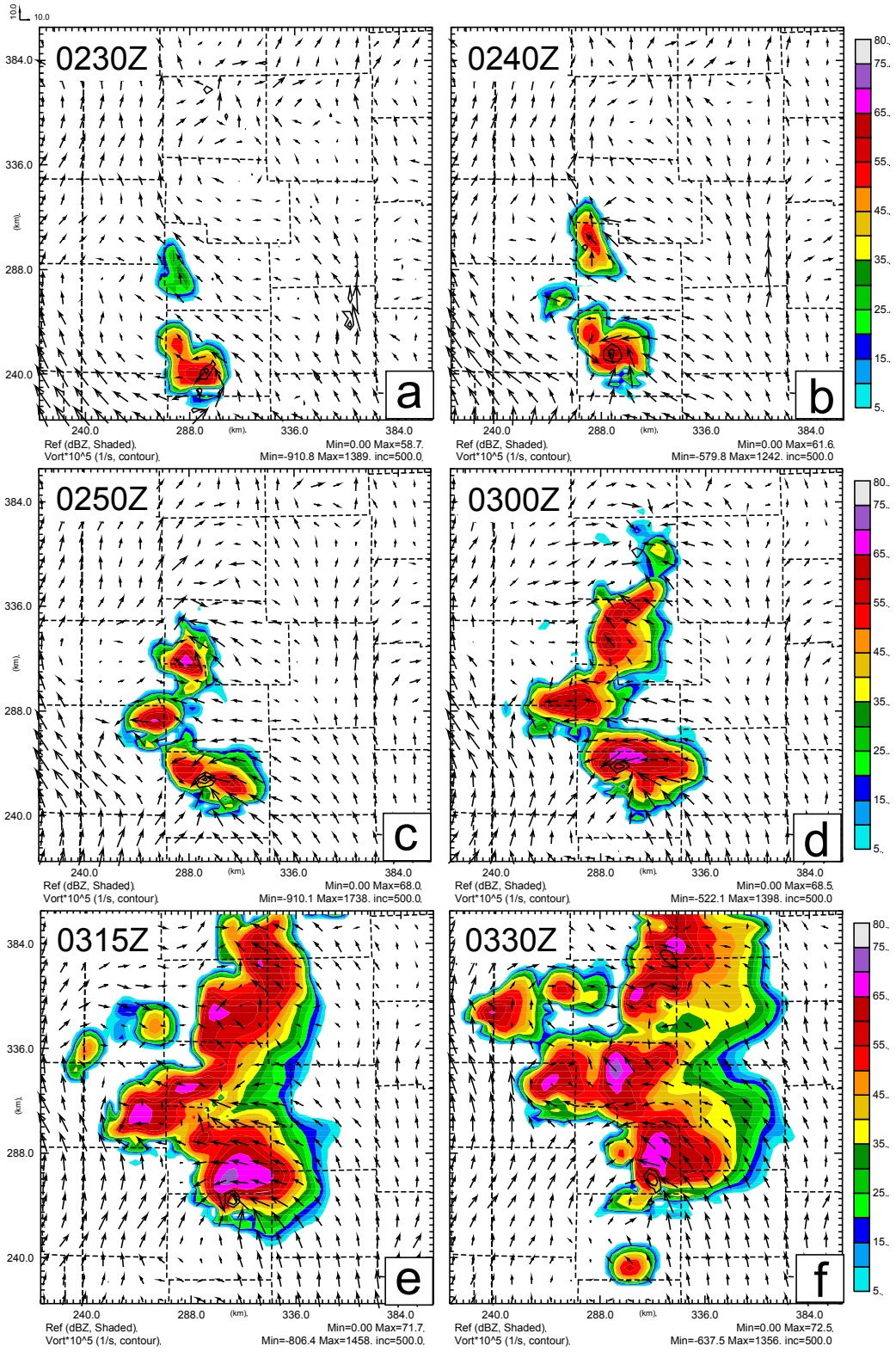

Fig. 4.28. Similar to Fig. 4.24, but for the "Vr_DP_2.5E-9" experiment.

To further examine the difference among the forecasts results of the experiments, the maximum vertical vorticity in the low two kilometers of the atmosphere is computed from the forecast every one minute. The time series is plotted in Fig. 4.29. It is shown that the "Vr_DP_1.0E-8" and "Vr_DP_5.0E-9" experiments predict larger low-level vertical vorticity than that from the "Vr_NODP" experiment during the entire forecast period, especially during the last half hour. The "Vr_DP_5.0E-8" experiment yields similar results as the "Vr_NODP" experiment and the "Vr_DP_2.5E-9" experiment predicts larger low-level vertical vorticity than the "Vr_NODP" experiment during the last half hour of the forecast period. Our examinations show that larger low-level vertical vorticity generally indicates stronger and deeper mesocyclone rotation column, similar to findings in previous 8 May 2003 OKC tornadic supercell thunderstorm case. As an example, Fig. 4.30 shows the vertical vorticity at the vertical cross section through the center of the major storm at y=253.5km at 0250UTC 5 May 2007 for all the five experiments discussed here. It can be seen that the "Vr_DP_1.0E-8" and "Vr_DP_5.0E-9" experiments do predict stronger and deeper rotated column (Fig. 4.30c,d vs. Fig. 4.30a). The "Vr_DP_5.0E-8" experiment also performs a little bit better than the "Vr_NODP" experiment in terms of the prediction of mesocyclone rotation at this time. The "Vr_DP_2.5E-9" experiment outperforms the "Vr_NODP" experiment mainly in the last half hour of forecast period.

Overall, the use of the diagnostic pressure equation constraint improves the data assimilation results. It significantly speeds up the rainfall spin-up. From these improved analyses, the general evolution of the storms and the mesocyclone rotation of the major

storm are predicted in a better fashion. The conclusion is quite similar to that drawn from the previous 8 May 2003 OKC tornadic supercell thunderstorm case.

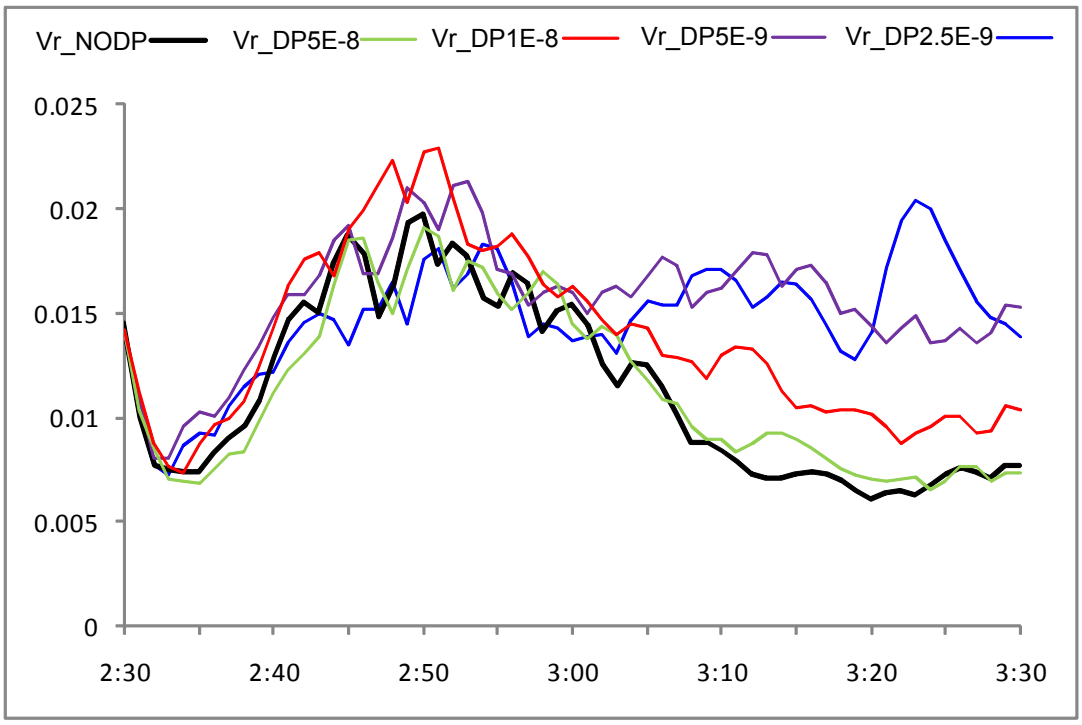

Fig. 4.29. The time series of maximum vertical vorticity below two kilometers from 0230 UTC to 0330UTC 5 May 2007 every one minute. The horizontal axis shows the time in UTC, the vertical axis shows the vertical vorticity value in unit of s^{-1}. The black line is for the "Vr_NODP" experiment, the green line is for the "Vr_DP_5.0E-8" experiment, the red line is for the "Vr_DP_1.0E-8" experiment, the purple line is for the "Vr_DP_5.0E-9" experiment and the blue line is for the "Vr_DP_2.5E-9" experiment.

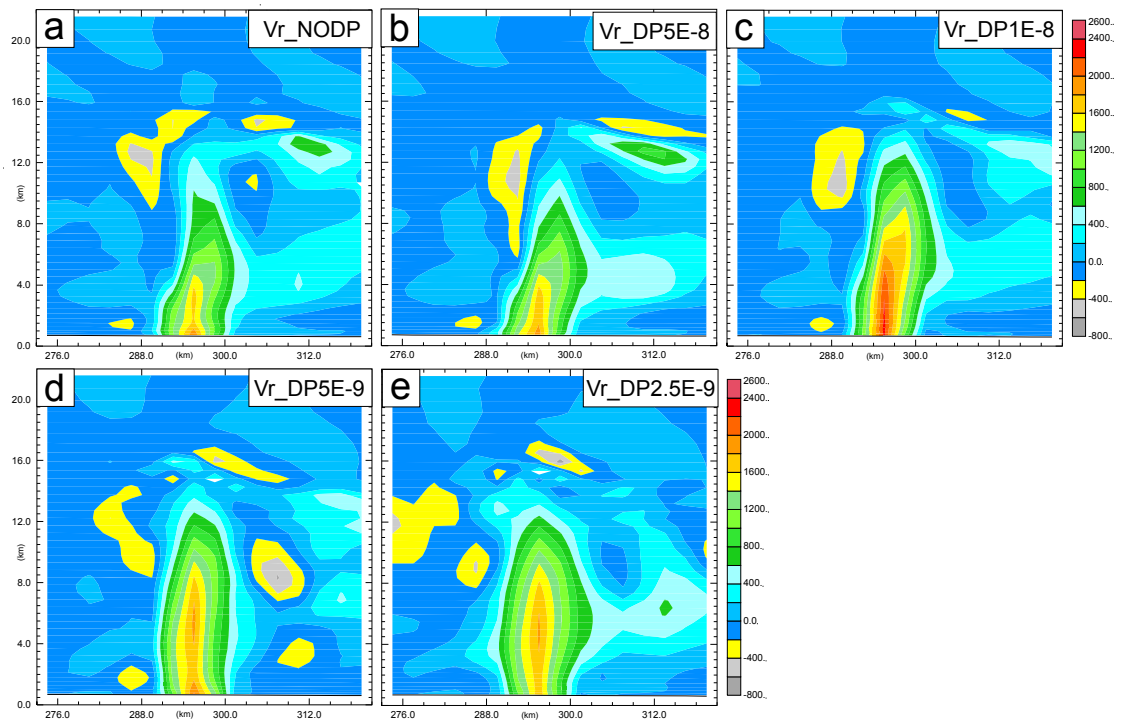

Fig. 4.30. The vertical vorticity (in unit of 10^{-5} s^{-1}) at the vertical cross section through the center of the major storm at y=253.5km at 0250UTC 5 May 2007 for the experiments (a) "Vr_NODP", (b) "Vr_DP_5.0E-8", (c) "Vr_DP_1.0E-8", (d) "Vr_DP_5.0E-9", (e) "Vr_DP_2.5E-9".

4.5.3.2 The impact of the constraint on the experiments assimilating both wind data and reflectivity data

We will now investigate the impact of the constraint on experiments assimilating both the radial velocity data and reflectivity data. Fig. 4.31 shows the reflectivity, horizontal wind vector and vertical vorticity at z=2km MSL from 0230UTC to 0330UTC for the "VrZ_NODP" experiment. It is shown that the assimilation of the reflectivity data in addition to the radial velocity data greatly reduce the spin-up of rainfall (Fig. 4.31a vs. Fig. 2.20a). However, since 0300UTC, the predicted major storm moves faster than that in the forecast of the "Vr_NODP" experiment, which is not very supported by the observation. In spite of this, the "VrZ_NODP" still makes a very reasonable forecast in

terms of the general evolution of the major storm. Based on this experiment, four more experiments are conducted by imposing the diagnostic pressure equation constraints with different weighting coefficients.

The investigation of the analyses leads to similar findings as before. That is, the use of the constraint does introduce some adjustment on the analyses; however, without reliable high-resolution analysis of the storms, it is difficult to tell which analysis is better. Following previous practices, the benefit of this constraint is then examined mainly based on the ensuing forecast. Fig. 4.32 shows the reflectivity, horizontal wind vector and vertical vorticity at z=2km MSL from 0230UTC to 0330UTC for the "VrZ_DP_1.0E-8" experiment. It can be seen that there is no significant difference in the general evolution of the major storm between the "VrZ_DP_1.0E-8" experiment and the "VrZ_NODP" experiment (Fig. 4.31) in terms of reflectivity pattern. The computed forecast scores (ETS, not presented here) also shows little difference, seconding the above finding. However, there is evident difference in the predicted low-level mesocyclone rotation as partly indicated by larger maximum vertical vorticity in Fig. 4.32 c, d, e ,f than that in Fig. 4.31 c, d, e, f. As a further demonstration, Fig. 4.33 shows the time series of the maximum vertical vorticity below two kilometers every one minute from 0230UTC to 0330UTC for both the experiments. The results for other three experiments are also shown in Fig. 4.33 and will be discussed later. It is illustrated in Fig. 4.33 that since 0245UTC and until the end of the forecast, the low level maximum vertical vorticity from the "VrZ_DP_1.0E-8" experiment (the red line) is much higher than that from the "VrZ_NODP" experiment (the black line). Our detailed examinations show that larger low-level vertical vorticity corresponds to a better-defined mesocyclone vortex, which is

stronger and deeper. This kind of behavior is very similar to our previous findings. As an example, Fig. 4.34 presents the vertical vorticity at the vertical cross section through the center of the major storm at y=259.5km at 0250UTC 5 May 2007. It is noticeable that the "VrZ_DP_1.0E-8" experiment predicts a stronger and deeper rotation column than the "VrZ_NODP" experiment.

The experiments using different weighting coefficients from that used by the "VrZ_DP_1.0E-8" experiments produce similar results with that of the "VrZ_DP_1.0E-8" experiments. Therefore, it can be concluded that for the experiments here, although the use of the diagnostic pressure equation constraint does not evidently improve the forecast of the general evolution of the major storm in terms of reflectivity pattern, it does help improve the forecast of the mesocyclone rotation associated with the observed Greensburg tornado.

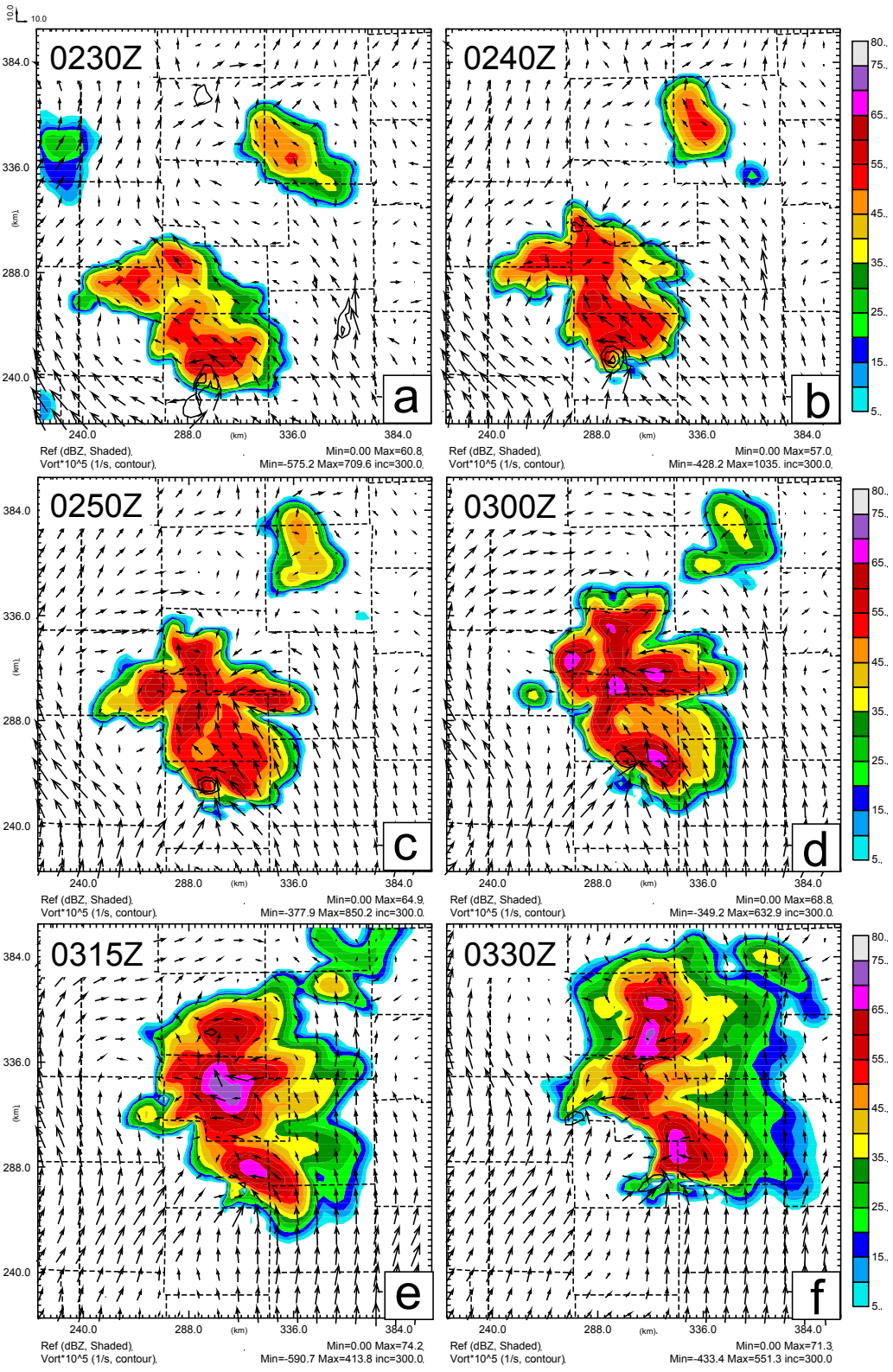

Fig. 4.31. Similar to Fig. 4.24, but for the "VrZ_NODP" experiment.

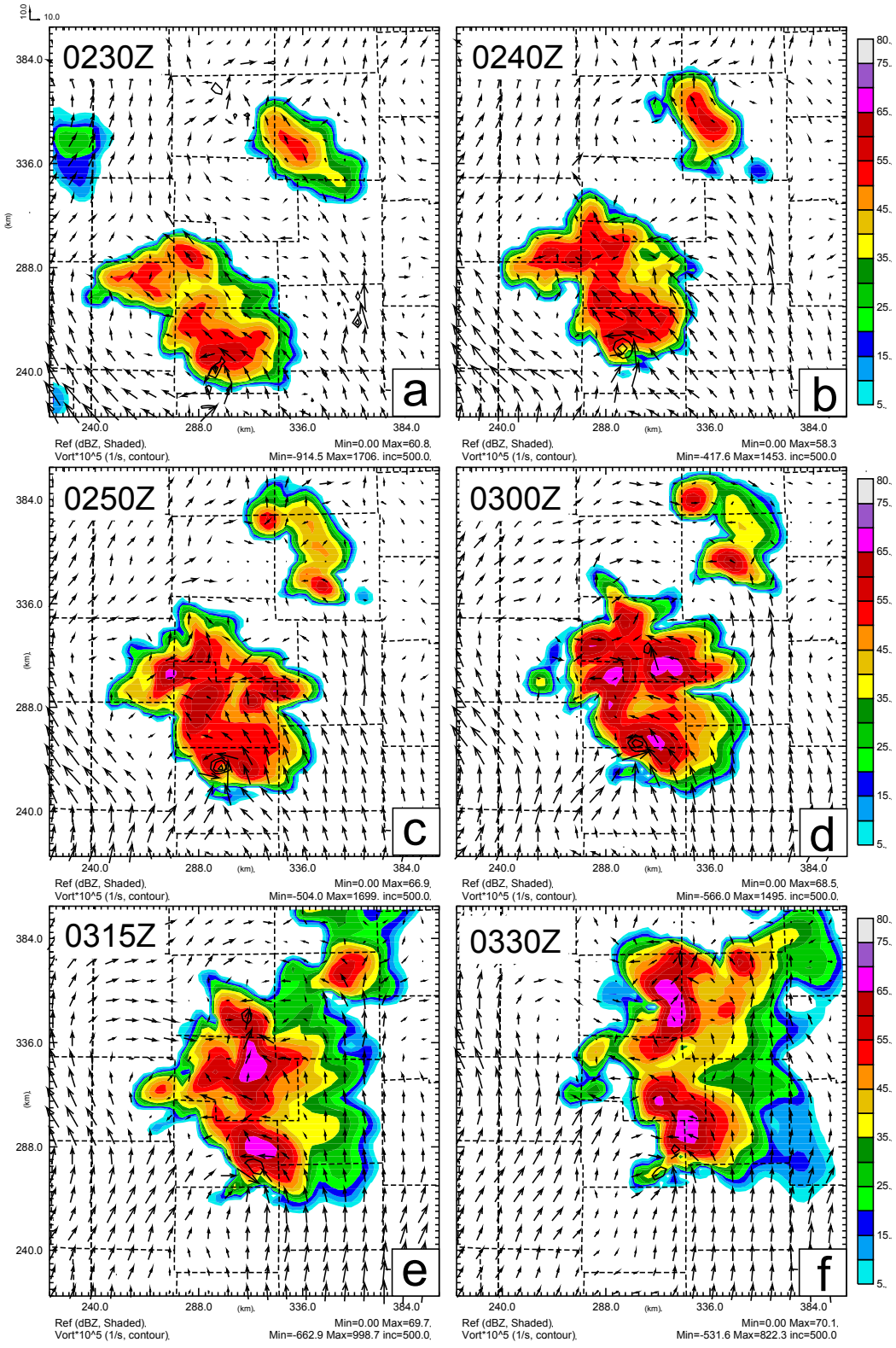

Fig. 4.32. Similar to Fig. 4.24, but for the "VrZ_DP_1.0E-8" experiment.

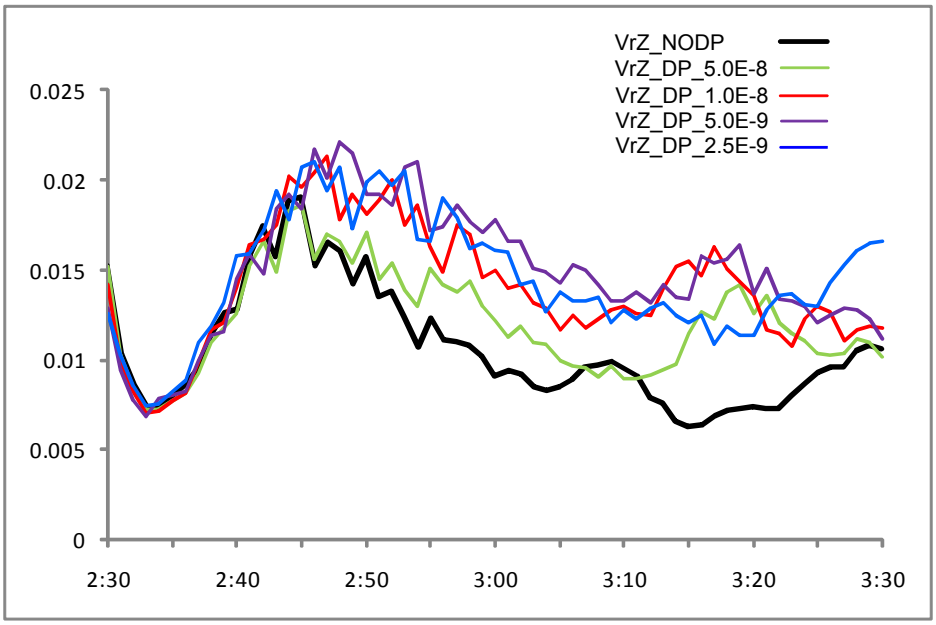

Fig. 4.33. Similar to Fig. 4.29, but the black line is for the "VrZ_NODP" experiment, the green line is for the "VrZ_DP_5.0E-8" experiment, the red line is for the "VrZ_DP_1.0E-8" experiment, the purple line is for the "VrZ_DP_5.0E-9" experiment, the blue line is for the "VrZ_DP_2.5E-9" experiment.

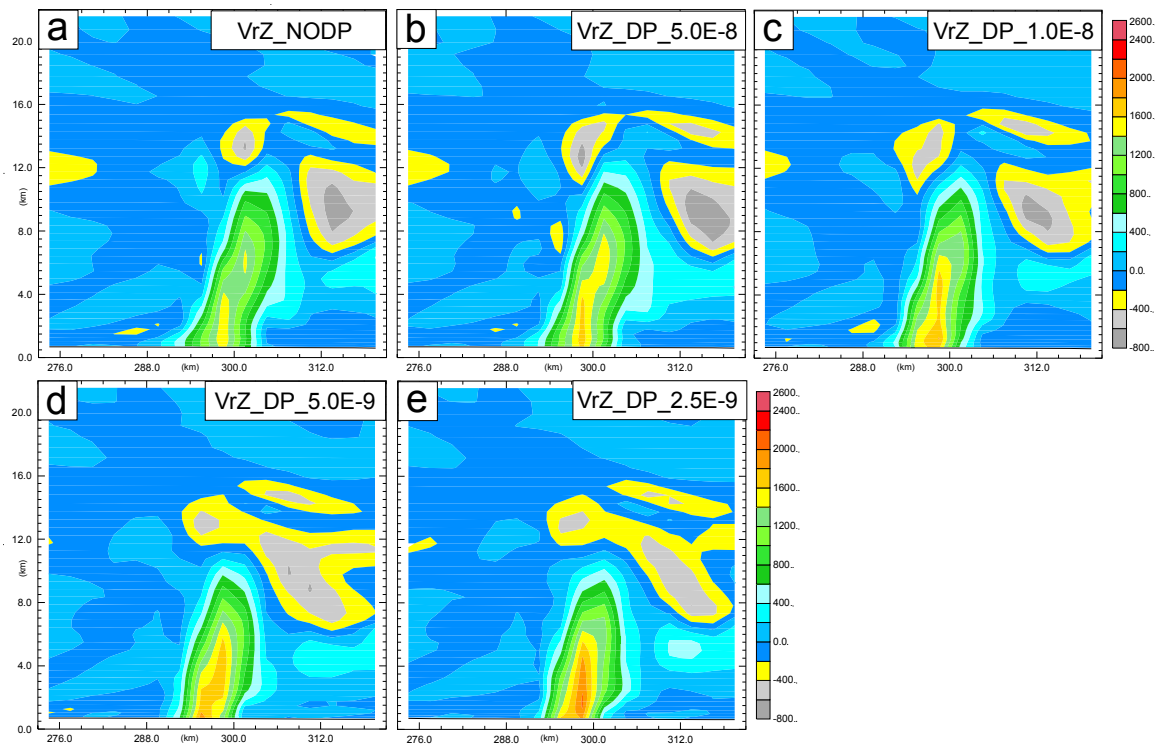

Fig. 4.34. The vertical vorticity (in unit of 10^{-5} s^{-1}) at the vertical cross section through the center of the major storm at y=259.5km at 0250UTC 5 May 2007 for the experiments (a) "VrZ_NODP", (b) "VrZ_DP_5.0E-8", (c) "VrZ_DP_1.0E-8", (d) "VrZ_DP_5.0E-9", (e) "VrZ_DP_2.5E-9".

4.5.4 Conclusion

The updated ARPS 3DVAR system that includes the diagnostic pressure equation as a weak constraint is further applied to the 5 May 2007 Greensburg tornadic supercell storm case under two different data assimilation configurations.

For the experiments assimilating wind data alone, the use of the diagnostic pressure equation constraint can evidently speed up the spin-up of rainfall during the intermittent data assimilation process and hence improve the following forecast in terms of the general evolution of storm cells and the mesocyclone rotation associated with observed tornado. For the experiments assimilating both wind data and reflectivity data, the use of diagnostic pressure equation constraint mainly improves the prediction of the mesocyclone rotation.

4.6 Summary and future work

Storm-scale 3dvar is computationally efficient and operational feasible for utilizing full volume Doppler radar data to predict the thunderstorms. However, it is often challenged by its less optimum theoretically due to its use of static background error covariance and lack of balance between model variables. Some effort has been made to provide flow-dependent background error covariance (Liu and Xue 2006; Liu et al. 2007) using an anisotropic filter or time-evolving forecast error covariance using hybrid ETKF-3DVAR technique (Hamill and Snyder 2000; Wang et al. 2008b, 2008a). Other efforts are made to couple the three components of wind fields using a weak constraint based on the mass continuity equation (Gao et al. 1999; Gao et al. 2001; Gao et al. 2004; Hu et al.

2006a; Hu et al. 2006b; Hu and Xue 2007). However, there is still no suitable link between the dynamic model variables and the thermodynamic model variables.

To mitigate the above problem, this research proposes to incorporate into the 3DVAR cost function a weak constraint based on the diagnostic pressure equation, which is derived from the ARPS full model momentum equations. This method is originated from the thermodynamic retrieval technique pioneered by Gal-Chen (1978) and Hane and Scott (1978). In our research, the main goal of this weak constraint is to help improve dynamic consistency between model variables.

This diagnostic pressure equation constraint is developed within the framework of the ARPS 3DVAR system. The updated ARPS 3DVAR system is then applied to one idealized and two real world tornadic supercell thunderstorms to illustrate the impact of this constraint.

For the idealized case study, it is demonstrated the diagnostic pressure equation constraint helps improve the analysis of wind and pressure fields slightly in single analysis step. After several data assimilation cycles, the impact of the constraint is more prominent. The use of the constraint evidently improves the recovery of all model variables.

For the 8 May 2003 Oklahoma City tornadic supercell thunderstorm case, the intensification and the dissipation process of the major storm are predicted in a better timing and location by the experiments that impose the diagnostic pressure equation constraint. The mesocyclone rotation is also predicted much stronger and deeper by the use of the constraint in the 3DVAR cost function.

For the 5 May 2007 Greensburg tornadic supercell thunderstorm case, the use of the diagnostic pressure equation constraint evidently improves the quality of data assimilation and the following forecast when assimilating wind data alone. When assimilating both the wind data and reflectivity data, the improvement from the use of the constraint lies mainly in the forecast of the mesocyclone rotation.

The sensitivity test of weighting coefficients show that the diagnostic pressure equation constraint is not very sensitive to the choice of weighting coefficient. However, the small values of the weighting coefficient should be used very cautiously.

Overall, it can be concluded that the diagnostic pressure equation constraint has positive impact on storm-scale 3D variational data assimilation and the subsequent forecast. In the future, more case studies on tornadic supercell storms and other storm-scale phenomena will be needed to generalize the conclusions in this research.

Chapter 5 Summary and future plan

5.1 Summary

Unlike relative mature large-scale 3DVAR data assimilation practice, storm-scale 3DVAR still faces many challenges (Droegemeier 1997; Sun 2005b) and calls for more research effort in this area. We are trying to address some fundamental issues in this dissertation: the role of different data fields in storm-scale 3DVAR data assimilation; the impact of imperfect radar radial velocity forward operator; a weak constraint to help improve dynamic balance among model variables.

Chapter 2 reports the work on the impact of different model variables on storm-scale data assimilation and NWP. OSS Experiments are conducted under a simplified 3DVAR framework. The model's first responses at storm scale to the assimilation of different types of observations are thoroughly examined. It is also demonstrated that the horizontal wind fields have the greatest impact on the storm-scale data assimilation. This has a good practical implication since the radar can observe storm-scale wind field (in radial direction) and multiple radar data assimilation can provide relatively accurate horizontal wind information. With the knowledge of accurate horizontal wind fields, extra observations from other model variables will help further improve the quality of data assimilation. Among these "other model variables", the perturbation water vapor field exerts the largest impact. To further examine the effect of wind fields in real world storm-scale NWP, a real case study is also carried out. It is shown that assimilating wind fields observed by six radars, using legacy ARPS 3DVAR system, can successfully predict the

general evolution of a tornadic supercell storm. This confirms the important role of wind fields.

Chapter 3 reports the work on the impact of imperfect radar radial velocity operator, which neglects the factors of beam broadening or earth curvature, on the storm-scale 3DVAR data assimilation. It is shown that the effect of beam broadening can be generally overlooked in storm-scale radar data assimilation without noticeable degradation of assimilation results. However, the effect of earth curvature can only be neglected when the radar is near the storm (within 60 km as demonstrated by this study). The impact of refractive index gradient is also tested and shown to be small.

Chapter 4 reports the development of a diagnostic pressure equation constraint for storm-scale 3DVAR data assimilation and its applications to tornadic supercell thunderstorm cases. This constraint is based on the storm-scale diagnostic pressure equation, which is derived from full ARPS model momentum equations. It serves to help build suitable balance among model variables. The impact of the constraint has been examined by applying it to case studies of one idealized tornadic supercell thunderstorm and two real-world tornadic supercell thunderstorms. It is demonstrated in the idealized case study that at single analysis step, the use of the constraint can help slightly improve the analysis of wind fields and pressure field; after a given period of intermittent data assimilation, the use of the constraint can evidently improve the quality of the data assimilation results. For the 8 May 2003 OKC tornadic supercell thunderstorm case, it is shown that the use of the constraint help improve the forecast in term of the general evolution and the mesocyclone rotation of the major tornadic supercell thunderstorm. For the 5 May 2007 Greensburg tornadic supercell thunderstorm case, two different

assimilation configurations are introduced to examine the impact of the constraint under different situations. It is shown that assimilating the wind data alone produces reasonable forecast and the use of the diagnostic pressure equation constraint evidently improve the forecast. When assimilating both wind data and reflectivity data, the impact of the constraint is also positive and mainly on the improvement of the prediction of the mesocyclone rotation. Overall, it is demonstrated that the diagnostic pressure equation constraint can improve the quality of radar data assimilation and the subsequent forecast.

5.2 Future plan

More case studies will be conducted in the future, beyond this Ph.D. program, to generalize the results in our research, including the dominant role of wind fields in storm-scale NWP, the positive impact of the diagnostic pressure equation constraint. The parallel implement of the updated ARPS 3DVAR, which includes the diagnostic pressure equation constraint, is also planned for future operational testing.

The assimilation of radar reflectivity data is not under a variational framework in this study. A kind of semi-empirical cloud analysis scheme originated from Local Analysis and prediction System (LAPS, Albers et al. 1996) is now commonly used for storm-scale research and operation (Ducrocq et al. 2000; Brewster 2002; Ducrocq et al. 2002; Souto et al. 2003; Hu et al. 2006a). To include the reflectivity data in the 3DVAR cost function using a relative complex radar forward operator other than warm rain version, which is adopted by some previous research (Sun and Crook 1997, 1998; Xiao et al. 2005), is still an unresolved problem. Therefore, further studies on how to best use reflectivity data are still required.

The current WSR-88D radar network is being upgraded with dual-polarization characteristics. This will improve the estimation of hydrometeor fields. Jung et al. (2008) have demonstrated the positive impact of polarimetric radar data in storm analysis through OSS experiments. Li and Mecikalski (2010) demonstrated that the dual-polarization information can improve the short term forecast of moist convections when assimilated by regional-scale WRF-3DVAR with a warm-rain radar forward operator. Inspired by these results, assimilating dual-pol radar data into storm-scale NWP model through a 3DVAR procedure or a cloud analysis system is very worth further investigation in the future.

References

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

Aksoy, A., D. C. Dowell, and C. Snyder, 2010: A Multicase Comparative Assessment of the Ensemble Kalman Filter for Assimilation of Radar Observations. Part II: Short-Range Ensemble Forecasts. *Monthly Weather Review*, 138, 1273-1292.

Albers, S. C., J. A. McGinley, D. A. Birkenheuer, and J. R. Smart, 1996: The local analysis and prediction system (LAPS): Analysis of clouds, precipitation and temperature. *Wea. Forecasting*, 11, 273-287.

Barker, D. M., W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results. *Monthly Weather Review*, 132, 897-914.

Bean, B. R., and E. J. Dutton, 1968: Radio Meteorology. *Dover Publications*, 435.

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

Bodine, D., P. L. Heinselman, B. L. Cheong, R. D. Palmer, and D. Michaud, 2010: A Case Study on the Impact of Moisture Variability on Convection Initiation Using Radar Refractivity Retrievals. *Journal of Applied Meteorology and Climatology*, 49, 1766-1778.

Brandes, E. A., 1984: Vertical vorticity generation and mesocyclone suste-nance in tornadic thunderstorms: The observational evidence. *Mon. Wea. Rev.*, 112, 2253-2269.

Brewster, K., 2002: Recent advances in the diabatic initialization of a non-hydrostatic numerical model. *Preprints, 15th Conf on Numerical Weather Prediction and 21st Conf on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., J6.3.

Brewster, K. A., 2003: Phase-correcting data assimilation and application to storm scale numerical weather prediction. Part II: Application to a severe storm outbreak. *Mon. Wea. Rev.*, 131, 493-507.

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

Caumont, O., and D. Veronique, 2008: What should be considered when simulating Doppler velocities measured by ground-based weather radars? *J. App. Meteor. Climatology*, 47, 2256-2262.

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

Courtier, P., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation. *Quart. J. Roy. Met. Soc.*, 124, 1783-1808.

Crook, N. A., 1994: Numerical simulations initialized with radar-derived winds. Part I: Simulated data experiments. *Mon. Wea. Rev.*, 122, 1189-1203.

Crook, N. A., and J. D. Tuttle, 1994: Numerical simulations initialized with radar-derived winds. Part II: Forecasts of three gust front cases. *Mon. Wea. Rev.*, 122, 1214-1217.

Crook, N. A., and J. Sun, 2002: Assimilating Radar, Surface, and Profiler Data for the Sydney 2000 Forecast Demonstration Project. *Journal of Atmospheric and Oceanic Technology*, 19, 888-898.

Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D operational support facility. *Bull. Amer. Meteor. Soc.*, 74, 1669-1687.

Crum, T. D., R. E. Saffle, and J. W. Wilson, 1998: An update on the NEXRAD program and future WSR-88D support to operations. *Wea. and Forecasting*, 13, 253–262.

Derber, J. C., D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting*, 6, 538-547.

Doviak, R., and D. Zrnic, 1993: *Doppler Radar and Weather Observations.* 2nd ed. Academic Press, 562 pp.

Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman Filter Assimilation of Radar Observations of the 8 May 2003 Oklahoma City Supercell: Influences of Reflectivity Observations on Storm-Scale Analyses. *Monthly Weather Review*, In Press.

Droegemeier, K. K., 1990: Toward a science of storm-scale prediction. *Preprint, 16th conf. on Severe Local Storms*, Kananaskis Park, Alberta, Canada, Amer. Meteor. Soc., 256-262.

Droegemeier, K. K., 1997: The numerical prediction of thunderstorms: Challenges, potential benefits, and results from real time operational tests. *WMO Bulletin*, 46, 324-336.

Ducrocq, V., J.-P. Lapore, J.-L. Redelsperger, and F. Orain, 2000: Initialization of a fine-scale model for convective-system prediction: A case study. *Quarterly Journal of the Royal Meteorological Society*, 126, 3041-3065.

Ducrocq, V., D. Ricard, J.-P. Lafore, and F. Orain, 2002: Storm-Scale Numerical Rainfall Prediction for Five Precipitating Events over France: On the Importance of the Initial Humidity Field. *Weather and Forecasting*, 17, 1236-1256.

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

Fabry, F., 2010: For How Long Should What Data Be Assimilated for the Mesoscale Forecasting of Convection and Why? Part II: On the Observation Signal from Different Sensors. *Monthly Weather Review*, 138, 256-264.

Fabry, F., and J. Sun, 2010: For How Long Should What Data Be Assimilated for the Mesoscale Forecasting of Convection and Why? Part I: On the Propagation of Initial Condition Errors and Their Implications for Data Assimilation. *Monthly Weather Review*, 138, 242-255.

Fabry, F., C. Frush, I. Zawadzki, and A. Kilambi, 1997: On the extraction of near-surface index of refraction using radar phase measurements from ground targets. *J. Atmos. Oceanic Technol.*, 14, 978-987.

Gal-Chen, T., 1978: A method for the initialization of the anelastic equations: Implications for matching models with observations. *Mon. Wea. Rev.*, 106, 587-606.

Gal-chen, T., and C. E. Hane, 1981: Retrieving buoyancy and pressure fluctuations from Doppler radar observations: A status report. *Progress in Radar Meteorology, Atmos. Technol.*, 13, 98-104.

Gal-Chen, T., and R. A. Kropfli, 1984: Buoyancy and pressure perturbation derived from dual-Doppler radar observations of the planetary boundary layer: Applications for matching models with observations. *J. Atmos. Sci.*, 41, 3007-3020.

Gao, J.-D., M. Xue, A. Shapiro, Q. Xu, and K. K. Droegemeier, 2001: Three-dimensional simple adjoint velocity retrievals from single Doppler radar. *J. Atmos. Ocean Tech.*, 18, 26-38.

Gao, J., K. Brewster, and M. Xue, 2006: A comparison of the radar ray path equations and approximations for use in radar data assimilation. *Adv. Atmos. Sci.*, 23, 190-198.

Gao, J., M. Xue, Z. Wang, and K. K. Droegemeier, 1998: The initial condition and explicit prediction of convection using ARPS adjoint and other retrievals methods with WSR-88D data. *12th Conf. Num. Wea. Pred.*, Phoenix AZ, Amer. Meteor. Soc., 176-178.

Gao, J., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the analysis of three-dimensional wind fields from two Doppler radars. *Mon. Wea. Rev.*, 127, 2128-2142.

Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. *Journal of Atmospheric and Oceanic Technology*, 21, 457-469.

Gao, J., M. Xue, K. Brewster, F. Carr, and K. K. Droegemeier, 2002: New development of a 3DVAR system for a nonhydrostatic NWP model. *Preprint, 15th Conf. Num. Wea. Pred. and 19th Conf. Wea. Anal. Forecasting*, San Antonio, TX, Amer. Meteor. Soc., 339-341.

Ge, G., and J. Gao, 2007: Latest development of 3DVAR system for ARPS and its application to a tornadic supercell storm. *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Amer. Meteor. Soc.

Ge, G., J. Gao, K. Brewster, and M. Xue, 2010: Impacts of Beam Broadening and Earth Curvature on Storm-Scale 3D Variational Data Assimilation of Radial Velocity with Two Doppler Radars. *Journal of Atmospheric and Oceanic Technology*, 27, 617-636.

Giering, R., and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM Transactions Math. Software*, 24, 437-474.

Gu, W., H. Gu, and Q. Xu, 2001: Impact of single-Doppler radar observations on numerical predcition of 7 May 1995 Oklahoma squall line. *Fifth Symp. on Integrated Observing Systems*, Albuquerque, NM, Amer. Meteor. Soc., 139-142.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter - 3D variational analysis scheme. *Mon. Wea. Rev.*, 128, 2905-2919.

Hane, C. E., and B. C. Scott, 1978: Temperature and pressure perturbations within convective clouds derived from detailed air motion information: Preliminary testing. *Mon. Wea. Rev.*, 106, 654-661.

Hane, C. E., and P. S. Ray, 1985: Pressure and Buoyancy Fields Derived from Doppler Radar Data in a Tornadic Thunderstorm. *Journal of the Atmospheric Sciences*, 42, 18-35.

Hane, C. E., R. B. Wilhelmson, and T. Gal-Chen, 1981: Retrieval of thermodynamic variables within deep convective clouds: Experiments in three dimensions. *Mon. Wea. Rev.*, 109, 564- 576.

Hodur, R. M., 1997: The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). *Mon. Wea. Rev.*, 125, 1414-1430.

Honda, Y., and K. Koizumi, 2006: The impact of the assimilation of precipitation data and Radar reflectivity with a pre-operational 4DVAR for the JMA nonhydrostatic model. *10th Symp. Integrated Obs. Assim. Sys. Atmos., Oceans, Land Surface (IOAS-AOLS)*, CDROM, P2.1.

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

Hu, M., and M. Xue, 2007: Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case. *Mon. Wea. Rev.*, 135, 507–525.

Hu, M., M. Xue, and K. Brewster, 2006a: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, 134, 675-698.

Hu, M., M. Xue, J. Gao, and K. Brewster, 2006b: 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.

Ide, K., P. Courtier, M. Ghil, and A. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, 75, 181-189.

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

Klemp, J. B., and R. Rotunno, 1983: A study of the tornadic region within a supercell thunderstorm. *J. Atmos. Sci.*, 40, 359-377.

Klemp, J. B., R. B. Wilhelmson, and P. S. Ray, 1981: Observed and numerically simulated structure of a mature supercell thunderstorm. *J. Atmos. Sci.*, 38, 1558-1580.

Lemon, L. R., and M. Umscheid, 2007: THE GREENSBURG, KANSAS TORNADIC STORM: A STORM OF EXTREMES. *24th Conference on Severe Local Storms*, American Meteorological Society.

Li, X., and J. R. Mecikalski, 2010: Assimilation of the dual-polarization Doppler radar data for a convective storm with a warm-rain radar forward operator. *Journal of Geophysical Research*, 115.

Lilly, D. K., 1990: Numerical prediction of thunderstorms - Has its time come? *Quart. J. Roy. Meteor. Soc.*, 116, 779-798.

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

Lin, Y., P. S. Ray, and K. W. Johnson, 1993: Initialization of a modeled convective storm using Doppler radar-derived fields. *Mon. Wea. Rev.*, 121, 2757-2775.

Lin, Y. J., T. C. Wang, and J. H. Lin, 1986: Pressure and Temperature Perturbations within a Squall-Line Thunderstorm Derived from SESAME Dual-Doppler Data. *Journal of the Atmospheric Sciences*, 43, 2302-2327.

Lindskog, M., K. Salonen, H. Järvinen, and D. B. Michelson, 2004: Doppler Radar Wind Data Assimilation with HIRLAM 3DVAR. *Monthly Weather Review*, 132, 1081-1092.

Liou, Y.-C., 2001: The Derivation of Absolute Potential Temperature Perturbations and Pressure Gradients from Wind Measurements in Three-Dimensional Space. *Journal of Atmospheric and Oceanic Technology*, 18, 577-590.

Liou, Y.-C., and Y.-J. Chang, 2009: A Variational Multiple–Doppler Radar Three-Dimensional Wind Synthesis Method and Its Impacts on Thermodynamic Retrieval. *Monthly Weather Review*, 137, 3992-4010.

Liou, Y.-C., T.-C. C. Wang, and K.-S. Chung, 2003: A Three-Dimensional Variational Approach for Deriving the Thermodynamic Structure Using Doppler Wind Observations—An Application to a Subtropical Squall Line. *Journal of Applied Meteorology*, 42, 1443-1454.

Liu, H., and M. Xue, 2006: Retrieval of moisture from slant-path water vapor observations of a hypothetical GPS network using a three-dimensional variational scheme with anisotropic background error. *Mon. Wea. Rev.*, 134, 933–949.

Liu, H., M. Xue, R. J. Purser, and D. F. Parrish, 2007: Retrieval of moisture from simulated GPS slant-path water vapor observations using 3DVAR with anisotropic recursive filters. *Mon. Wea. Rev.*, 135, 1506–1521.

Liu, S., C. Qiu, Q. Xu, P. Zhang, J. Gao, and A. Shao, 2005: An improved method for Doppler wind and thermodynamic retrievals. *Adv. Atmos. Sci.*, 21, 90-102.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, 112, 1177-1194.

McCarthy, D., L. Ruthi, and J. Hutton, 2007: The Greensburg, KS tornado. *22th Conf. on Weather Analysis and Forecasting and 18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc.

Meng, Z., and F. Zhang, 2008a: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part IV: Comparison with 3DVAR in a Month-Long Experiment. *Monthly Weather Review*, 136, 3671-3682.

Meng, Z., and F. Zhang, 2008b: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part III: Comparison with 3DVAR in a Real-Data Case Study. *Monthly Weather Review*, 136, 522-540.

Nascimento, E. L., and K. K. Droegemeier, 2006: Dynamic Adjustment in a Numerically Simulated Mesoscale Convective System: Impact of the Velocity Field. *Journal of the Atmospheric Sciences*, 63, 2246-2268.

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

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, 120, 1747-1763.

Probert-Jones, J. R., 1962: The radar equation in meteorology. *Quart. J. Roy. Met. Soc.*, 88, 485- 495.

Protat, A., and I. Zawadzki, 2000: Optimization of Dynamic Retrievals from a Multiple-Doppler Radar Network. *Journal of Atmospheric and Oceanic Technology*, 17, 753-760.

Protat, A., I. Zawadzki, and A. Caya, 2001: Kinematic and Thermodynamic Study of a Shallow Hailstorm Sampled by the McGill Bistatic Multiple-Doppler Radar Network. *Journal of the Atmospheric Sciences*, 58, 1222-1248.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Monthly Weather Review*, 131, 1524-1535.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Monthly Weather Review*, 131, 1536-1548.

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

Qiu, C.-J., and Q. Xu, 1996: Least-square retrieval of microburst winds from single-Doppler radar data. *Mon. Wea. Rev.*, 124, 1132-1144.

Rabier, F., 2006: Overview of global data assimilation developments in numerical weather-prediction centres. *Quarterly Journal of the Royal Meteorological Society*, 131, 3215-3233.

Rabier, F., and Coauthors, 2006: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Quarterly Journal of the Royal Meteorological Society*, 124, 1809-1829.

Ray, P. S., and Coauthors, 1981: The morphology of severe tornadic storms on 20 May 1977. *J. Atmos. Sci.*, 38, 1643-1663.

Rihan, F. A., C. G. Collier, S. P. Ballard, and S. J. Swarbrick, 2008: Assimilation of Doppler radial winds into a 3D-Var system: Errors and impact of radial velocities on the variational analysis and model forecasts. *Quarterly Journal of the Royal Meteorological Society*, 134, 1701-1716.

Romine, G. S., D. W. Burgess, and R. B. Wilhelmson, 2008: A Dual-Polarization-Radar-Based Assessment of the 8 May 2003 Oklahoma City Area Tornadic Supercell. *Monthly Weather Review*, 136, 2849-2870.

Roux, F., 1985: Retrieval of thermodynamic fields from multiple-Doppler radar data using the equations of motion and the ther-modynamic equation. *Mon. Wea. Rev.*, 113, 2142-2157.

Roux, F., 1988: The West African Squall Line Observed on 23 June 1981 during COPT 81: Kinematics and Thermodynamics of the Convective Region. *Journal of the Atmospheric Sciences*, 45, 406-426.

Roux, F., and J. Sun, 1990: Single-Doppler observations of a West African squall line on 27-28 May 1981 during COPT 81: Kinematics, thermodynamics and water budget. *Mon. Wea. Rev.*, 118, 1826-1854.

Roux, F., V. Marécal, and D. Hauser, 1993: The 12/13 January 1988 Narrow Cold-Frontal Rainband Observed during MFDP/FRONTS 87. Part I: Kinematics and Thermodynamics. *Journal of the Atmospheric Sciences*, 50, 951-974.

Salonen, K., 2002: Observation operator for Doppler radar radial winds in HIRLAM 3D-Var. *ERAD Publication Series*, 1, 405-408.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, 5, 570-575.

Schenkman, A. D., M. Xue, A. Shapiro, K. Brewster, and J. Gao, 2011: The Analysis and Prediction of the 8–9 May 2007 Oklahoma Tornadic Mesoscale Convective

System by Assimilating WSR-88D and CASA Radar Data using 3DVAR. *Monthly Weather Review*, In Press.

Serafin, R. J., and J. W. Wilson, 2000: Operational Weather Radar in the United States: progress and opportunity. *Bull. Amer. Meteor. Soc.*, 81, 501-518.

Shapiro, A., and S. Lazarus, 1993: A Modified Dynamic Recovery Technique for Cloud-scale Numerical Models. *17th Conf. on Severe Local Storms*, St Louis, MO, 455-459.

Shapiro, A., C. K. Potvin, and J. Gao, 2009: Use of a Vertical Vorticity Equation in Variational Dual-Doppler Wind Analysis. *Journal of Atmospheric and Oceanic Technology*, 26, 2089-2106.

Shapiro, A., P. Robinson, J. Wurman, and J. Gao, 2003: Single-Doppler velocity retrieval with rapid-scan radar data. *J. Atmos. Ocean. Tech.*, 20, 1758-1775.

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

Souto, M. J., C. F. Balseiro, V. Pérez-Muñuzuri, M. Xue, and K. Brewster, 2003: Importance of cloud analysis and impact for daily forecast in terms of climatology of Galician region, Spain. *J. App. Meteor.*, 42, 129-140.

Stensrud, D. J., and J. Gao, 2010: Importance of Horizontally Inhomogeneous Environmental Initial Conditions to Ensemble Storm-Scale Radar Data Assimilation and Very Short-Range Forecasts. *Monthly Weather Review*, 138, 1250-1272.

Sun, J., 2005a: Initialization and numerical forecasting of a supercell storm observed during STEPS. *Monthly Weather Review*, 133, 793-813.

Sun, J., 2005b: Convective-scale assimilation of radar data: Progress and challenges. *Q.J.R. Meteorol. Soc.*, 131, 25.

Sun, J., and R. A. Houze, Jr., 1992: Validation of a thermodynamic retrieval technique by application to a simulated squall line with trailing stratiform precipitation. *Mon. Wea. Rev.*, 120, 1003-1018.

Sun, J., and N. A. Crook, 1996: Comparison of thermodynamic retrieval by the adjoint method with the traditional retrieval method. *Mon. Wea. Rev.*, 124, 308–324.

Sun, J., and N. A. Crook, 1997: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: Model development and simulated data experiments. *J. Atmos. Sci.*, 54, 1642-1661.

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

Sun, J., and N. A. Crook, 2001: Real-time low-level wind and temperature analysis using single WSR-88D data. *Wea. Forecasting*, 16, 117-132.

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

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

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

Torn, R. D., 2010: Performance of a Mesoscale Ensemble Kalman Filter (EnKF) during the NOAA High-Resolution Hurricane Test. *Monthly Weather Review*, 138, 4375-4392.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, 136, 5116-5131.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008b: A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiment. *Mon. Wea. Rev.*, 136, 5132-5147.

Wang, Z., 1993: Variational data assimilaition with 2-D shallow water equations and 3-D FSU global spectral models, Department of Mathematics, Florida State University, 140-141 pp.

Weygandt, S., P. Nutter, K. E., P. S., and K. Droegemeier, 1999: The relative importance of different data fields in a numerically simulatd convective storm. *the 8th Conference on Mesoscale Processes*, Boulder, CO., American Meteorological Society, Boston, USA, 310-315.

Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002a: Retrieval of Model Initial Fields from Single-Doppler Observations of a Supercell Thunderstorm. Part II: Thermodynamic Retrieval and Numerical Prediction. *Mon. Wea. Rev.*, 130, 454-476.

Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002b: Retrieval of Model Initial Fields from Single-Doppler Observations of a Supercell Thunderstorm. Part I: Single-Doppler Velocity Retrieval. *Mon. Wea. Rev.*, 130, 433-453.

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

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

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Monthly Weather Review*, 130, 2905-2916.

Xiao, Q., and J. Sun, 2007: Multiple radar data assimilation and short-range quantitative precipitation forecasting of a squall line observed during IHOP_2002. *Mon.Wea.Rev.*, 135, 3381-3404.

Xiao, Q., Ying-Hwa Kuo, J. Sun, W.-C. Lee, E. Lim, Y.-R. Guo, and D. M. Barker, 2005: Assimilation of Doppler radar observations with a regional 3DVAR system: Impact of Doppler velocities on forecasts of a heavy rainfall case. *J. App. Meteor.*, 44, 768-788.

Xu, Q., H. Gu, and S. Yang, 2001a: Simple adjoint method for three-dimensional wind retrievals from single-Doppler radar. *Quart. J. Roy. Meteor. Soc.*, 127, 1053-1068.

Xu, Q., H. Gu, and W. Gu, 2001b: A variational method for Doppler radar data assimilation. *Fifth Symp. on Integrated Observing Systems*, Albuquerque, NM, Amer. Meteor. Soc., 118-121.

Xu, Q., C. Qiu, H. D. Gu, and J. Yu, 1995: Simple adjoint retrievals of microburst winds from single-Doppler radar data. *Mon. Wea. Rev.*, 123, 1822-1833.

Xu, Q., L. Wei, W. Gu, J. Gong, and Q. Zhao, 2010: A 3.5-Dimensional Variational Method for Doppler Radar Data Assimilation and Its Application to Phased-Array Radar Observations. *Advances in Meteorology*, 2010, 1-14.

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

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

Yu, C. K., B. J. D. Jou, and D. P. Jorgensen, 2001: Retrieved Thermodynamic Structure of a Subtropical, Orographically Influenced, Quasi-Stationary Convective Line. *Mon. Wea. Rev.*, 129, 1099-1116.

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

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. *Monthly Weather Review*, 137, 2105-2125.

Zhang, J., F. Carr, and K. Brewster, 1998: ADAS cloud analysis. *Preprints, 12th Conf. on Num. Wea. Pred.*, Phoenix, AZ., Amer. Met. Soc., 185-188.

Zhao, Q., J. Cook, Q. Xu, and P. R. Harasti, 2006: Using Radar Wind Observations to Improve Mesoscale Numerical Weather Prediction. *Weather and Forecasting*, 21, 502-522.

Zhao, Q., J. Cook, Q. Xu, and P. R. Harasti, 2008: Improving Short-Term Storm Predictions by Assimilating both Radar Radial-Wind and Reflectivity Observations. *Weather and Forecasting*, 23, 373-391.