

STORM-SCALE RADAR DATA ASSIMILATION AND HIGH RESOLUTION NWP

GUEST EDITORS: JIDONG CAO, DAVID J. STENSRUD, LOUIS WICKER, MING XUE, AND KUN ZHAO





Storm-Scale Radar Data Assimilation and High Resolution NWP

Advances in Meteorology

Storm-Scale Radar Data Assimilation and High Resolution NWP

Guest Editors: Jidong Gao, David J. Stensrud, Louis Wicker,
Ming Xue, and Kun Zhao



Copyright © 2014 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "Advances in Meteorology." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Paulo Artaxo, Brazil
Guy Brasseur, USA
Mladjen Ćurić, Serbia
Raymond Desjardins, Canada
Klaus Dethloff, Germany
P.C.S. Devara, India
Julio Diaz, Spain
Shouting Gao, China
Luis Gimeno, Spain
Sven-Erik Gryning, Denmark

Ismail Gultepe, Canada
Hiroyuki Hashiguchi, Japan
Didier Hauglustaine, France
I. S A Isaksen, Norway
Yasunobu Iwasaka, Japan
Hann-Ming Henry Juang, USA
George Kallos, Greece
Harry D. Kambezidis, Greece
Markku Kulmala, Finland
Richard Leaitch, Canada

Monique Leclerc, USA
Zhanqing Li, USA
Gwo-Fong Lin, Taiwan
Edward Llewellyn, Canada
Kyaw Tha Paw, USA
Sara C. Pryor, USA
Eugene Rozanov, Switzerland
Zhihua Zhang, China

Contents

Storm-Scale Radar Data Assimilation and High Resolution NWP, Jidong Gao, David J. Stensrud, Louis Wicker, Ming Xue, and Kun Zhao
Volume 2014, Article ID 213579, 3 pages

Analysis and Forecast of a Tornadoic Thunderstorm Using Multiple Doppler Radar Data, 3DVAR, and ARPS Model, Edward Natenberg, Jidong Gao, Ming Xue, and Frederick H. Carr
Volume 2013, Article ID 281695, 18 pages

Prediction of Convective Storms at Convection-Resolving 1 km Resolution over Continental United States with Radar Data Assimilation: An Example Case of 26 May 2008 and Precipitation Forecasts from Spring 2009, Ming Xue, Fanyou Kong, Kevin W. Thomas, Jidong Gao, Yunheng Wang, Keith Brewster, and Kelvin K. Droegemeier
Volume 2013, Article ID 259052, 9 pages

The Development of a Hybrid EnKF-3DVAR Algorithm for Storm-Scale Data Assimilation, Jidong Gao, Ming Xue, and David J. Stensrud
Volume 2013, Article ID 512656, 12 pages

Impact of a Diagnostic Pressure Equation Constraint on Tornadoic Supercell Thunderstorm Forecasts Initialized Using 3DVAR Radar Data Assimilation, Guoqing Ge, Jidong Gao, and Ming Xue
Volume 2013, Article ID 947874, 12 pages

Retrieving 3D Wind Field from Phased Array Radar Rapid Scans, Xiaobin Qiu, Qin Xu, Chongjian Qiu, Kang Nai, and Pengfei Zhang
Volume 2013, Article ID 792631, 16 pages

Low-Level Polarimetric Radar Signatures in EnKF Analyses and Forecasts of the May 8, 2003 Oklahoma City Tornadoic Supercell: Impact of Multimoment Microphysics and Comparisons with Observation, Daniel T. Dawson II, Louis J. Wicker, Edward R. Mansell, Youngsun Jung, and Ming Xue
Volume 2013, Article ID 818394, 13 pages

The Impact of Mesoscale Environmental Uncertainty on the Prediction of a Tornadoic Supercell Storm Using Ensemble Data Assimilation Approach, Nusrat Yussouf, Jidong Gao, David J. Stensrud, and Guoqing Ge
Volume 2013, Article ID 731647, 15 pages

Correcting Fast-Mode Pressure Errors in Storm-Scale Ensemble Kalman Filter Analyses, Corey K. Potvin and Louis J. Wicker
Volume 2013, Article ID 624931, 14 pages

Improved Doppler Velocity Dealiasing for Radar Data Assimilation and Storm-Scale Vortex Detection, Qin Xu, Kang Nai, Shun Liu, Chris Karstens, Travis Smith, and Qingyun Zhao
Volume 2013, Article ID 562386, 10 pages

Impact of 3DVAR Data Assimilation on the Prediction of Heavy Rainfall over Southern China, Tuanjie Hou, Fanyou Kong, Xunlai Chen, and Hengchi Lei
Volume 2013, Article ID 129642, 17 pages



WRF-ARW Variational Storm-Scale Data Assimilation: Current Capabilities and Future Developments,

Juanzhen Sun and Hongli Wang

Volume 2013, Article ID 815910, 13 pages

Using WSR-88D Polarimetric Data to Identify Bird-Contaminated Doppler Velocities, Yuan Jiang,

Qin Xu, Pengfei Zhang, Kang Nai, and Liping Liu

Volume 2013, Article ID 769275, 13 pages

Editorial

Storm-Scale Radar Data Assimilation and High Resolution NWP

Jidong Gao,¹ David J. Stensrud,¹ Louis Wicker,¹ Ming Xue,² and Kun Zhao³

¹ NOAA/National Severe Storm Laboratory, 120 David L. Boren Boulevard, Norman, OK 73072, USA

² Center for Analysis and Prediction of Storms and School of Meteorology, University of Oklahoma, 120 David L. Boren Boulevard, Norman, OK 73072, USA

³ Key Laboratory of Mesoscale Severe Weather/MOE, School of Atmospheric Sciences, Nanjing University, 22 Hankou Road, Nanjing 210093, China

Correspondence should be addressed to Jidong Gao; jidong.gao@noaa.gov

Received 1 December 2013; Accepted 1 December 2013; Published 30 January 2014

Copyright © 2014 Jidong Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the steady increase in computing power, operational centers throughout the world are planning to run their numerical models with resolutions that begin to resolve the convective scale. This situation opens the opportunity and increases the demand for using radar data in convective-scale data assimilation in order to insert convective structures into model initial conditions. The potential for successfully assimilating radar data into convective-scale numerical weather prediction (NWP) models is predicated on the solution of several key challenges, including data quality control, proper estimation of the background error statistics, and, most notably, the estimation of atmospheric state variables that are not directly observed by radar. This special issue of 12 papers focuses on progresses in some of these important areas.

Radar data quality control is critical for radar data applications, including the use of radar data for severe weather warning, assimilation of radar data into storm-scale NWP models, and diagnostic research on thunderstorm mechanisms and structures. Velocity dealiasing is an especially important and yet often very difficult part of the radar data quality control problem. In this special issue, Q. Xu et al. report on a Doppler velocity dealiasing technique based on alias-robust Velocity-Azimuth-Display (VAD) and variational analyses at the NOAA/National Severe Storms Laboratory (NSSL). This technique has been tested extensively with observations from the operational WSR-88D radars under various weather conditions and data from a phased array radar (PAR) operated by NSSL. In another study related to radar data quality control, an automated

technique is developed by Y. Jiang et al. to identify and remove contaminated velocities by birds, especially migrating birds in addition to those contaminated by ground clutter.

The effective assimilation of radar data into a NWP model requires advanced data assimilation (DA) techniques, such as variational and ensemble Kalman filter (EnKF) methods. A three-dimensional variational (3DVAR) system, which includes a mass continuity equation and the Advanced Regional Prediction System (ARPS) model equations as weak constraints, has been developed over the years at the Center for Analysis and Prediction of Storms (CAPS). The 3DVAR system was designed with special considerations for assimilating radar data into a convective-scale nonhydrostatic model, the ARPS, and has been used to provide initial conditions for numerous real time convective-scale NWP and research studies, including several papers in this special issue (G. Ge et al.; T. Hou et al.; E. Natenberg et al.; N. Yussouf et al.; M. Xue et al.). A major advantage of a 3DVAR system is its computational efficiency and the ease by which equation constraints can be included. However, one of the major shortcomings of a 3DVAR system is that it usually employs background error covariances that are not flow-dependent, limiting their ability to “retrieve” unobserved state variables.

Compared to 3DVAR, the more advanced 4DVAR technique incorporates a full prediction model into the assimilation system and implicitly includes the effects of flow-dependent error covariances through the use of both the forward prediction and backward adjoint models. In J. Sun

and H. Wang, a variational storm-scale data assimilation system for the Advanced Research Weather Research and Forecasting (WRF-ARW) model, including both 3DVAR and 4DVAR components, is described. The paper provides a review of the basics of the system and various studies that have been conducted to evaluate and improve its performance. Future activities that are required to further improve the system and to make it operational are also discussed. In another 4DVAR related work, X. Qiu et al. describe an application of the so-called simple adjoint method for retrieving three-dimensional winds from PAR data. They demonstrate that the method is computationally efficient and can be used for real-time applications with PAR rapid scans.

The EnKF is another advanced data assimilation method that shares many of the advantages of 4DVAR. For convective storms, some encouraging results are reported using EnKF data assimilation in this special issue. The impact of model microphysics on storm-scale DA is a very important topic. In D. T. Dawson et al., the impact of increasing the number of predicted moments in a multimoment bulk microphysics scheme is investigated using EnKF analyses. They find that a triple-moment version of a microphysics scheme produces the best results in comparison to the single- and double-moment versions. In another study, C. K. Potvin and L. J. Wicker apply procedures that have traditionally been used to retrieve the pressure and buoyancy fields from dual-Doppler wind analyses to rebalance the pressure fields produced by EnKF. The method is shown to provide an effective and low-cost way to remove pressure oscillations in the storm-scale EnKF analyses.

As pointed out earlier, 3DVAR is computationally efficient and can easily incorporate equation constraints that can be very beneficial. However, its use of static background error covariances limits its ability to retrieve unobserved variables. The EnKF directly uses flow-dependent error covariances derived from forecast ensemble but has a rank-deficiency problem due to the typically small ensemble size. To blend the advanced features of both variational and EnKF methods and to overcome their respective shortcomings, hybrid ensemble-variational (EnVar) algorithms had been proposed. Such an approach has been shown to be effective by a number of studies for large-scale data assimilation, but studies have been very limited for the storm scale. J. Gao et al. demonstrate the potential usefulness of a hybrid En3DVar method for convective-scale data assimilation. The algorithm uses the extended control variable approach to combine the static and ensemble-derived flow-dependent forecast error covariances, and the ensemble perturbations are provided by an EnKF that is coupled with the hybrid En3DVar. The method is applied to the assimilation of simulated radar data for a supercell storm. The results indicate that the hybrid method provides the best results among the 3DVAR, EnKF, and hybrid methods for hydrometeor-related state variables in terms of root-mean squared errors. For other state variables, the performance of the hybrid is very close to that of EnKF but much better than that of 3DVAR. More studies are needed in this direction.

The purpose of storm-scale data assimilation is to provide initial conditions for high resolution explicit convective storm prediction and quantitative precipitation forecasting.

In this aspect, several studies of high impact severe weather events are included in this special issue. To explore the impact of model physics on the mesoscale environmental variability and its uncertainty, combined mesoscale convective scale ensemble DA and forecast experiments are conducted for the 8 May, 2003, Oklahoma City tornadic supercell storm by N. Yussouf et al. Two sets of 36-member WRF-ARW model mesoscale ensemble adjustment Kalman filter (EAKF) DA experiments with continuous cycling on a continental United States domain are conducted to provide background environmental conditions, using either fixed physics or multiple physics parameterization schemes across the ensemble members, respectively. Results show that the ensemble with background fields from the multiple physics mesoscale ensemble provides more realistic forecasts of significant tornado parameter, dryline structure, and near surface variables than the ensemble from fixed physics mesoscale background fields. This suggests that incorporating physics diversity across the ensemble can be important to successful probabilistic convective-scale forecast of supercell thunderstorms. With the same 8 May, 2003, Oklahoma tornadic case, E. Natenberg et al. find that the use of data from all available radars provides the best forecast in which a strong low-level mesocyclone develops and tracks in close proximity to the actual tornado damage path. T. Hou et al. examine the impact of 3DVAR on the prediction of two heavy rainfall events over southern China by using the storm-scale WRF-ARW model. Observations from Doppler radars, surface Automatic Weather Station (AWS) network, and radiosondes are used. The assimilation of all radar, surface, and radiosonde data together produces more positive impacts on the forecast skill than the assimilation of only one type of data for the two rainfall events.

M. Xue et al. report on the results of real-time convection-resolving forecasts at 1km horizontal spacing over the continental United States in spring 2009 and a test case from spring 2008, produced by CAPS at the University of Oklahoma as part of the NOAA Hazardous Weather Testbed (HWT) Spring Experiment. The forecasts started from initial conditions which included the assimilation of radial velocity and reflectivity data from all operational WSR-88D radars within a large model domain. The assimilation used the ARPS 3DVAR with a cloud analysis package. Significant positive impact of radar data assimilation is found to last for at least 12 hours. Clear advantages of running the model at the convection-resolving 1km grid over a coarser 4km convection permitting grid spacing are also demonstrated. Convection-allowing storm-scale data assimilation and forecasting, including ensemble forecasting that provides probabilistic forecast information and radar data that provides valuable information on convective storms and precipitation, clearly represent the future of operational NWP.

Acknowledgments

As the guest editors for this special issue, we appreciate the contributions of all of the authors, reviewers, and journal staff. We sincerely hope that the publication of this special

issue will encourage and invite more advanced research on storm-scale radar data assimilation and high resolution NWP.

Jidong Gao
David J. Stensrud
Louis Wicker
Ming Xue
Kun Zhao

Research Article

Analysis and Forecast of a Tornadic Thunderstorm Using Multiple Doppler Radar Data, 3DVAR, and ARPS Model

Edward Natenberg,^{1,2} Jidong Gao,³ Ming Xue,^{1,2} and Frederick H. Carr^{1,2}

¹ Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK 73072, USA

² School of Meteorology, University of Oklahoma, Norman, OK 73072, USA

³ NOAA/National Severe Storms Laboratory, Norman, OK 73072, USA

Correspondence should be addressed to Edward Natenberg; enatenberg@meso.com

Received 30 May 2013; Accepted 25 September 2013

Academic Editor: Kun Zhao

Copyright © 2013 Edward Natenberg et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A three-dimensional variational (3DVAR) assimilation technique developed for a convective-scale NWP model—advanced regional prediction system (ARPS)—is used to analyze the 8 May 2003, Moore/Midwest City, Oklahoma tornadic supercell thunderstorm. Previous studies on this case used only one or two radars that are very close to this storm. However, three other radars observed the upper-level part of the storm. Because these three radars are located far away from the targeted storm, they were overlooked by previous studies. High-frequency intermittent 3DVAR analyses are performed using the data from five radars that together provide a more complete picture of this storm. The analyses capture a well-defined mesocyclone in the midlevels and the wind circulation associated with a hook-shaped echo. The analyses produced through this technique are used as initial conditions for a 40-minute storm-scale forecast. The impact of multiple radars on a short-term NWP forecast is most evident when compared to forecasts using data from only one and two radars. The use of all radars provides the best forecast in which a strong low-level mesocyclone develops and tracks in close proximity to the actual tornado damage path.

1. Introduction

Doppler radar observations became more widely used as an analysis tool since the 1970s, by mapping polar coordinate radial velocity data to a Cartesian grid space. Results in Brandes [1] showed the evolution of a supercell including an intensifying mesocyclone and development of the rear-flank downdraft (RFD), as observed by a dual-Doppler analysis. Additional study was done using this technique in Brandes [2] where the airflow within a tornadic thunderstorm was compared with the observed tornado damage path. Brandes [2] also found that, at the beginning of the tornadic phase of a supercell thunderstorm, an increase in radial flow near the tornado was observed in the radar data analysis. This helped determine that there was a strongly convergent flow beneath the mesocyclone during this time. Another major development in the use of Doppler radar was the discovery of a unique Tornado Vortex Signature (TVS), that coincided with

the development of low-level rotation and tornadogenesis in the 1973 Union City tornado [3]. The development of this signature was closely linked to the mesocyclone structure model found in Lemon and Doswell [4].

With the frequent use of dual-Doppler analysis and increased resolution of storm scale models, data assimilation has become an important area of research that integrates observations into a numerical weather prediction (NWP) model. The development of variational data assimilation techniques for Doppler analysis has shown better results over other methods of analysis [5]. Several methods to obtain an initial state by thermodynamic retrieval using observations were originally proposed by Gal-Chen [6] and tested by Hane et al. [7] for simulated data and Gal-Chen and Kropfli [8] for radar observations of the planetary boundary layer.

Crook [9] tested a variational technique that would allow the retrieval of thermodynamic variables using radar data and a background sounding. The technique used a time

tendency term to estimate the thermodynamic variables and then estimated the deviation of those variables from the background values. This technique was then tested using observational data from three gust front cases [10]. Several data interpolation methods were tested to create an initial analysis, followed by a forecast to see if the analysis could help accurately predict the propagation of several outflow boundaries. The findings from these experiments showed that the analysis and forecast were greatly improved when surface data were used in addition to radar data and upper air soundings.

Weygandt et al. [11] tested a single velocity Doppler retrieval technique on a supercell thunderstorm with the use of the Taylor frozen turbulence hypothesis as a weak constraint. The results showed that there was reasonable agreement in the analysis with the exception of the magnitude of the vertical velocities. In an additional experiment performed by Weygandt et al. [12], thermodynamic retrieval was included using three separate retrieval times for a supercell thunderstorm. The initial analysis from this study showed a supercell which exhibited pressure perturbations consistent with Rotunno and Klemp [13] linear theory. When a non-hydrostatic model was initialized using a set of retrieved fields as initial conditions, the evolution of the storm seemed reasonably well predicted [12]. Results, obtained from another experiment using a simplified version of the analysis technique which only included radial velocities, estimated mean horizontal wind, and perturbation radial divergence, were poor. One problem noted by this study was the speed of the cold pool propagation, which was faster than the observed speed. This could have been due to the lack of ice microphysics in the model [12].

A four-dimensional variational (4DVAR) analysis technique was developed by Sun and Crook [14, 15], whose goal was to accurately retrieve wind and thermodynamic fields by minimizing the difference between the background and the observations defined by a cost function. By applying this technique, all the model variables are determined at the same time. Initially, the method was tested with simulated observations which showed good results for the thermodynamic structure, and the addition of a penalty term further increased the accuracy of the model state [14]. The results from the latter studies using 4DVAR analysis of several Florida thunderstorms showed good agreement with observational data [15]. Sun and Crook [16] also used the 4DVAR technique to assimilate data from a line of strong thunderstorms. Their findings showed that the penalty term helps to decrease noise in the analysis and the background term increases the forecast accuracy (after the first 10 minutes) in areas of sparse observational data. It was also found that using a previous analysis as a background for the analysis produced better results than those obtained by using a short-term forecast as a background [16]. However, the 4DVAR analysis technique is computationally expensive.

Gao et al. [17] tested a less costly computational technique known as three-dimensional variational (3DVAR). In this method, the cost function is composed of several terms: an observational term, a background term, and several penalty terms. Simulated observations showed that this analysis

technique was less sensitive to the boundary conditions than other methods that used mass continuity as a strong constraint, thereby mitigating the effects of error accumulation through explicit integration. The use of a recursive filter has been tested with the 3DVAR technique to improve the quality and efficiency of the analysis [18]. Each pass of the recursive filter employs the use of a filter in two directions, a left moving filter and a right moving filter [19]. When applied using dual-Doppler data of a supercell thunderstorm, it was shown that the use of a recursive filter produced analysis results similar to that in Gao et al. [17] but greatly improved the efficiency of the 3DVAR method.

The method developed in Gao et al. [18] was also used to run an analysis and numerical prediction cycle of a tornadic thunderstorm outbreak in north Texas [20, 21]. Two cloud analysis schemes along with the 3DVAR technique were tested to determine the sensitivity of these schemes to the analysis and subsequent forecast. The use of radial velocity in the analysis was found to play an important role in the development of vorticity in the forecasts. Also, the role of reflectivity in the analysis was shown to have a major impact on the development and maintenance of intense convection.

Recently the ensemble Kalman filter (EnKF) technique, originally developed by Evensen [22], has been widely used as a new method to assimilate observations into the model state. The EnKF technique seeks to accurately determine the flow dependent background error covariance through the use of an ensemble of nonlinear forecasts. The application of the EnKF method to storm-scale data assimilation has been examined by Snyder and Zhang [23], Zhang et al. [24], Tong and Xue [25], and Xue et al. [26] using simulated data, and by Dowell et al. [27], Snook et al. [28, 29], and other authors using observed radar data. Although EnKF avoids the linearization of the background error covariance, the use of a large ensemble is computationally expensive. Also, when decreasing the amount of members, the covariance can become underestimated.

Most of the above research used observations of severe thunderstorms from only one or two radars located close to the convection being studied, thereby limiting the information available to provide a complete picture of the storm, and which does not take full advantage of the WSR-88D radar network. By only using data from one or two nearby radars in an analysis, a large part of atmosphere located above the radar cannot be effectively observed due to the limits in the elevation of a radar volume scan, known as the "cone of silence".

The 8 May 2003 tornado in central Oklahoma has been widely studied because of the high impact on society and the high density of weather observations. The storm spawned a F4 tornado that tracked through southern Oklahoma City causing \$370 million in damage and 134 injuries. Located in the region are four WSR-88D radars: KVNIX, located in northwest Oklahoma at Vance Air Force Base; KINX, located just northeast of Tulsa; KFDR, located in southwest Oklahoma near the town of Frederick; and KTLX, located southeast of Oklahoma City (Figure 1). Two additional radars are also located in close proximity to Oklahoma City: KOUN,

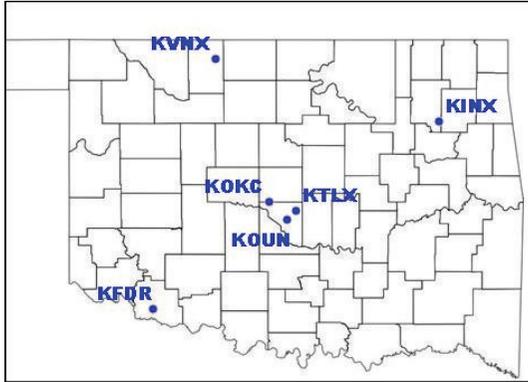


FIGURE 1: Map showing weather radars over Oklahoma. Shown are four WSR-88D Doppler radars KVNK, KTLX, KFDR, and KINX as well as two other radars: KOUN, an experimental dual-polarization WSR-88D radar, and KOKC, a TDWR radar operated by Will Rogers Airport in Oklahoma City.

a research dual-polarization WSR-88D radar located in Norman, and KOKC, a terminal Doppler weather radar (TDWR) used by the FAA near Will Rogers Airport in Oklahoma City. Because these two radars are too close to each other, we chose to use data from one of them, the KOKC TDWR radar, in this study. The Oklahoma Mesonet also provides surface observations every five minutes for 119 sites in all 77 of Oklahoma counties [30]. This, along with four upper-air wind profilers in the state, allows for high temporal resolution of atmospheric phenomena.

There have been several studies examining the evolution of the 8 May 2003 Oklahoma City supercell through use of radar data, assimilation techniques, and numerical models [27, 31–36]. In these studies, observations from only one or two radars close to the storm were assimilated, thus omitting the upper-level structures of this supercell.

In this study, multiple radars both close to and far from the storm are used to perform an analysis using the 3DVAR technique. The use of the 3DVAR technique to study this case was chosen over the EnKF or 4DVAR techniques because the 3DVAR technique is not computationally costly and has shown good results in other related studies [18, 20, 21]. In Section 2, the advanced regional prediction system (ARPS) and its 3DVAR technique are briefly discussed. Section 2 also describes the 8 May tornadic storm case, the processing of radar data, and experiment configurations. Results about the analysis and forecast experiments are presented in Section 3. Finally, conclusions are discussed in Section 4.

2. Models, Methods, and Data

2.1. ARPS Model, ARPS 3DVAR, and Cloud Analysis Algorithm. The advanced regional prediction system (ARPS) has been developed at the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma over the past 15 years [37–39]. The ARPS model was designed as a system suitable for explicit prediction of convective storms. It is a three-dimensional, nonhydrostatic, compressible model

formulated in generalized terrain-following coordinates. The model employs advanced numerical techniques, including monotonic advection schemes for scalar transport and variance-conserving fourth-order advection for other variables. The model also includes state-of-the-art physics parameterization schemes that are important for explicit prediction of convective storms. The system has been used in real-time high-resolution prediction experiments for convective scales in the past several years over the continental United States [40, 41].

A detailed description of the ARPS 3DVAR system can be found in Gao et al. [18]. In the current version of the ARPS 3DVAR system, the spatial covariances of the background error are modeled by a recursive filter [19], and the square root of the matrix is used for preconditioning [42]. The corresponding covariance matrix is diagonal, and its diagonal elements are specified according to the estimated observation errors.

Radar data can easily be incorporated into the cost function in the observation term; the observed values represent the radial component of wind even though the data have been mapped to a Cartesian grid during preprocessing. The ARPS 3DVAR data assimilation system can ingest data from a number of different sources including surface air observations, upper air soundings, profiler data, and aircraft data. Quality control in ARPS 3DVAR is performed by determining if an innovation vector is less than a threshold value. An observation whose innovation exceeds the threshold is rejected. The quality control threshold is a function of the specified background error, in which a larger background error value corresponds to a larger threshold. With the 3DVAR analysis, the background error also affects the relative weight of the observations and background. Observations can also be rejected based on climatological error statistics of the observation stations.

One unique feature of the ARPS 3DVAR system is that multiple analysis passes can be used to analyze different data types with different spatial correlation scales to account for the variations in the observation density among different data sources. Upper-air rawinsonde data and radar observations are examples of two observing systems with very different spatial density. An initial analysis pass can be performed using only upper air observations, using a large correlation scale. A second pass is then performed with radar data with a smaller correlation scale. Such a procedure allows for the retention of multi-scale information contained in observations of vastly different spatial density. A mass divergence constraint is also employed within the 3DVAR cost function to help couple the three wind components. This constraint permits the retrieval of vertical velocity from the mostly horizontal wind components as measured by radar and also helps spread observational information to nearby grid points in the analysis [21].

After completion of the 3DVAR analysis, a cloud analysis is performed. The procedure is based on that from the local analysis and prediction system [43] and includes modifications from Zhang et al. [44] and Brewster [45]. The cloud analysis package estimates mixing ratios of water vapor, rain-water, and cloud water based on reflectivity measurements.

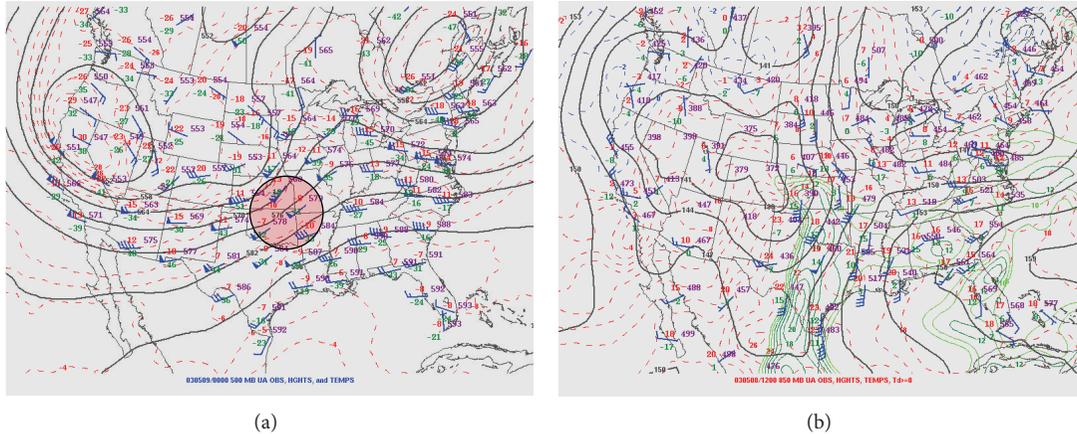


FIGURE 2: (a) 500 mb analysis and upper air observations, (b) 850 mb analysis and upper air observations valid at 0000 UTC, 9 May 2003. The red-shaded circle denotes the position of a short-wave trough. Heights are shown as black contours (in decameters), temperature is shown by dashed red contours, and dewpoint temperatures greater than 8 degrees celsius are contoured in green. Note also the 850 mb 50 knot wind observation at Dallas/Fort Worth (image from <http://www.spc.noaa.gov/>).

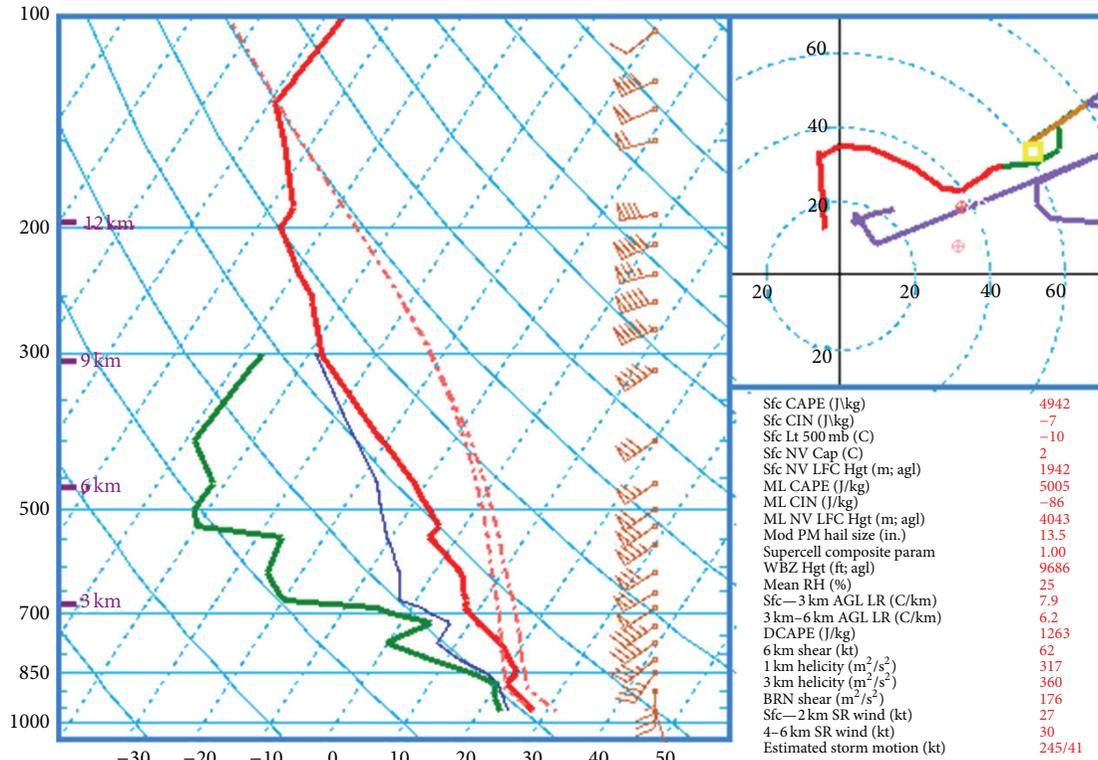
The cloud analysis can also adjust in-cloud temperature based on reflectivity [20].

2.2. The 8 May Storm Environment. In the first ten days of May 2003, there were 434 tornadoes, more than any month in the US on record (although the record was broken in 2008). On three of the ten days, central and eastern Oklahoma were hit by significant tornado outbreaks. The 1200 UTC upper-air observations on the morning of 8 May 2003 showed a significant positively tilted trough in the Pacific Northwest characterized by 500 mb heights near 540 decameters (not shown). This trough exhibited a double-barrel structure with two regions of minimum heights, one in central Alberta and a second approaching the coast of Washington State. Several short-waves rotated around the large trough and propagated downstream. The most prominent of these short waves (shaded in red on Figure 2(a)) provided enhanced vorticity advection over Oklahoma in the hours previous to 0000 UTC 9 May which induced midtropospheric rising motion and pressure falls at the surface. These pressure falls enhanced those associated with lee troughing downstream of the Rocky Mountains, over which a strong upper-level jet existed. A poleward meridional flow then developed on the eastern side of the developing low that advected large quantities of unstable warm moist air from the Gulf of Mexico in the lower levels of the atmosphere (Figure 2(b)). An important aspect of this dynamical setting was the continued presence of the trough in the northwestern United States during the week prior to the 8 May event and the subsequent reinforcement of the lee trough and southerly winds. This continuous flow of moist air from the Gulf of Mexico helped reinforce thermodynamic instability over the central plains throughout the week.

At the surface, several outflow boundaries existed in the warm sector caused by a complex of nocturnal tornadic thunderstorms. These outflow boundaries served as a focal point for convergence and subsequent convective initiation later

that afternoon. As the convection from the previous night moved east along the Red River and dissipated, skies over Oklahoma cleared, contributing to strong surface heating during the day. The low-level environment through the late morning of 8 May was characterized by temperatures near 28°C and dewpoints around 22°C. A sounding from Norman, Oklahoma, (OUN) at 0000 UTC 9 May 2003 showed an extreme level of instability within the warm sector (Figure 3). Surface convective available potential energy (CAPE) reached values near 5000 J kg⁻¹, but convection was inhibited initially by a capping inversion. The OUN sounding also shows a well-defined “quarter circle” hodograph representative of strong directional and speed shear with associated storm relative environmental helicity (SREH) values of 317 m² s⁻² in the lowest 1 km. This strongly sheared environment with large values of CAPE was conducive to the formation of strong tornadoes.

At 3 pm local daylight time (LDT) or 2000 UTC, satellite data showed the upper-level short-wave moving eastward into western Oklahoma. Observations from the Oklahoma Mesonet indicated a dryline located just to the west of OUN, in central Oklahoma. Along this boundary dry air from the western plains which converged with moist air from the Gulf (not shown). By 4 pm LDT, towering cumulus had penetrated the capping inversion and initiated strong surface-based convective thunderstorms along a bulge in the dryline. After 4 pm the storms that formed along the dryline became severe in Grady County, just southwest of Oklahoma City and by 4:30 pm LDT the most prominent of the convective cells split and deviated in storm motion by about 9 degrees to the right of the mean motion. At this time both cells showed moderate rotation at midlevels. By 5 pm LDT, the right-split storm was entering Oklahoma County and radar detected strong low-level convergence and rotation. It was at this time that sporadic wind damage was caused by the advancing RFD. Minutes later, at 5:11 pm LDT (2211 UTC) a tornado was confirmed on the south side of Moore, Oklahoma.



OUN-030509/0000
Observed sounding

NWS/NCEP/SPC

FIGURE 3: 0000 UTC 9 May 2003 sounding from Norman, Oklahoma (image from <http://www.spc.noaa.gov/>).

The tornado increased in intensity as it tracked to the east-northeast just south of Interstate 240 and into Oklahoma City. At 5:38 pm LDT, after traveling 30 km, the tornado dissipated 4 km south-southeast of Choctaw, Oklahoma.

2.3. Radar Data Ingest and Preprocessing. Radar data must undergo several procedures before being used in the analysis. For this study radar data in several formats were ingested into the ARPS 3DVAR system. Level II data were obtained from the National Climate Data Center (NCDC) in binary format which included full volume scans every 5 to 6 minutes. Data from TDWR were ingested in network common data form (NetCDF) format with full volume scans being reconstructed using each tilt during the preprocessing procedure. TDWR data included 9 elevations, while WSR-88D radars had 13 elevation levels. After the volume scans are read into the program, the program checks the quality of the velocity and reflectivity data and makes corrections to the data if needed. For raw velocity data, the absolute values of velocities greater than the Nyquist velocity of the radar are folded (Figure 4(a)). This means that inbound velocities greater than the Nyquist velocity in the raw data will be displayed as if they were outbound velocities (of the opposite sign) with a magnitude

less than the Nyquist velocity. To correct the velocity information, a dealiasing algorithm is used to unfold the raw velocity data by comparing the data to a background sounding or by comparing the data to velocity measurements surrounding the region being dealiased. If the velocity is determined to be folded, the deviation from the Nyquist velocity will be found and then added to the Nyquist velocity of the opposite sign. Figure 4(b) shows a sweep of velocity data after dealiasing which clearly depicts the location of the circulation. The initial radar reflectivity includes ground clutter (displayed in Figure 4(c)), anomalous propagation, and transient clear air echoes. These artifacts are removed by additional radar reflectivity quality control procedures by detecting both large changes in mean radial velocity and reflectivity over neighboring range gates at the lowest elevations.

After being cleaned, the data are projected onto a Cartesian grid with the same resolution as the analysis being performed. Radar data are then interpolated to a grid using a least square method. During the interpolation, a linear fit is used in the vertical and a quadratic polynomial fit is used in the horizontal. The domain of this method is bounded by the range of available radar data, precluding extrapolation of the data.

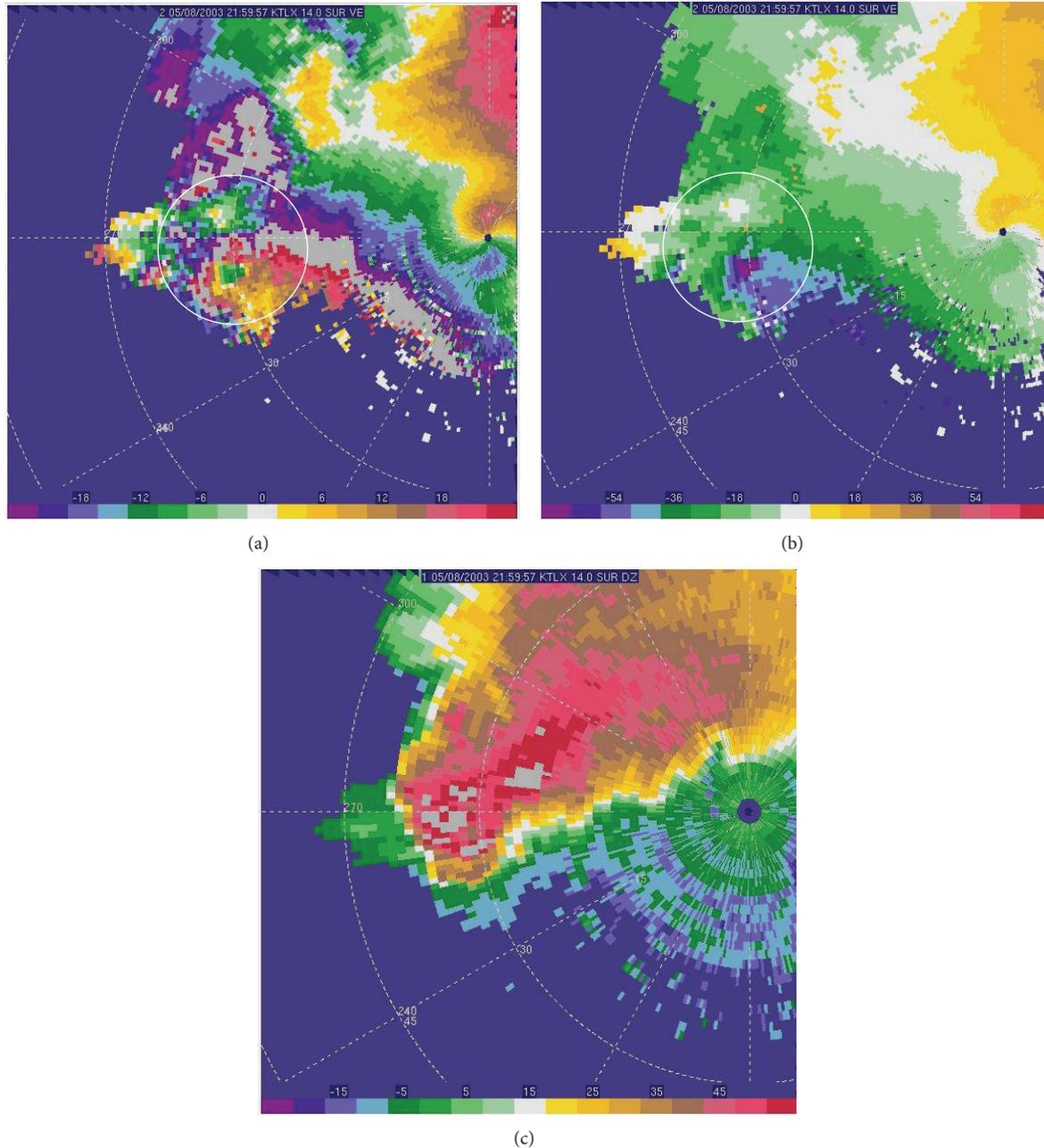


FIGURE 4: Raw (a) and dealiased (b) radial velocity and (c) reflectivity data on the 14 degree elevation surface from KTLX valid at 2159 UTC on 8 May 2003. The region of twice-folded velocities is shown in the circled region. Note that the color scale for (a) has increments of 3 m/s while (b) has increments of 9 m/s.

2.4. Experiment Design. A number of analyses are performed on the 8 May 2003 Moore/Oklahoma City supercell at 5-minute intervals starting from 2145 UTC to 2240 UTC. For all experiments, the grid resolution in the horizontal direction is 1 km, with a lower bound vertical grid resolution of 100 meters which is stretched vertically to 23 km using a cubic function with an average grid spacing of 500 meters. The domain size is $157 \times 143 \times 47$ grid points with the center of the domain located just east-southeast of Oklahoma City at the location where spotters reported a large tornado.

The background state is obtained from a 9 km ARPS forecast using a background from the operational NCEP ETA model and upper air observations. A forecast is then performed using intermittent analyses that assimilate only sounding and profiler data starting from 18 UTC to produce the background for the 1 km analyses. The observational data that are used in the 1 km analysis consist of Oklahoma Climate Survey Mesonet observations, profiler data, and radar data. Upper air sounding data were only included in the 9 km forecast and not the 1 km analyses due to the sparse temporal resolution

of the data. The radar data used in this study consist of four WSR-88D radars and one TDWR, as discussed in the introduction. Since radar observations are obtained at different times, scans closest to the time of the analysis were used.

In the control analysis experiment, all available observations are ingested into the assimilation system using four analysis passes. During the first pass the radius of influence is set at 50 km in the horizontal, 3 grid points in the vertical, and only profiler data are used. On the second analysis pass, the radius of influence is changed to 25 km in the horizontal and 2 grid points in the vertical, and Mesonet and profiler data are used. In the third pass, the radius of influence is decreased to 10 km in the horizontal and 2 grid points in the vertical, and Mesonet and radar data are used. On the fourth and final pass, the radius of influence is changed to 5 km in the horizontal and 1 grid point in the vertical, and only radar data are used.

Two other analysis experiments are performed to determine the impact of the number of radars used in the analysis at one time, 2200 UTC (Table 1). These experiments show the role of each radar on the control analysis and their impact on the divergence and vertical velocity fields. The second experiment (IRAD) is the same as the control analysis, but with the use of radar data only from KOKC, the closest radar to the storm during a large duration of the analysis period. The third experiment (2RAD) includes the two radars that only observed the low levels of the thunderstorm, KOKC and KTLX.

Finally, several forecast experiments are performed to examine the sensitivity of a forecast to the number of radars using the previous 3DVAR analyses at 2155 utc (Table 1). These analyses are then used as initial conditions for integration of the ARPS model. To demonstrate that use of data from radars that are close to the storm may not always benefit the forecast, we perform two more forecast experiments; one with only the KOKC radar, and a second with both the KOKC and KTLX radars.

The model was run using the Lin 3-ice microphysics scheme [46]. A large time step of 2 seconds and a small time step of 5 seconds are used to perform forward integration of slow and acoustic wave modes, respectively. The drop size distribution's intercept parameter for rain was adjusted from the default value of 8×10^6 to $8 \times 10^5 \text{ m}^{-4}$ to allow for expected larger raindrops associated with the supercell storm that will result in less evaporative cooling and a weaker cold pool [47].

3. Results and Discussion

3.1. The Control Analysis with all Radars. In the control analysis, we use data from all available radars except for KOUN (which is very close to KTLX and therefore has similar viewing angles), so the amount of data used in the analysis is quite large. This analysis best resolves the storm-scale features throughout the analysis period and therefore serves as a benchmark for other experiments. The evolution of the storm from the stage before tornado production until the tornado lifted is shown in Figure 5. The development of a hook feature around 2200 UTC about 10 minutes before the first

TABLE 1: List of analysis and forecast experiments showing the number of radars used.

Name	No. of radars	Radars used
Control	5	KOKC KINX
		KFDR KTLX
		KVNX
1RAD	1	KOKC
2RAD	2	KOKC KTLX

touchdown and dissipation just after 2230 UTC about 10 minutes before the tornado lifts during the final stages of mesocyclone occlusion are very clear. Halfway through the life of the tornado, the storm weakens and the reflectivity decreases quite rapidly after 2230 UTC. This is due to the full occlusion of the RFD around the tornadic circulation [4] and decreased updraft intensity associated with a downward oriented pressure gradient [48, 49]. Wind vectors shown in Figure 5 depict intense inflow during the mature period of the storm and weak inflow during the decaying period.

The evolution of the low-level mesocyclone and occluding RFD is visible in the divergence field (Figure 6). The storm evolves from an elongated diffuse area of weak convergence on the south flank of the reflectivity contour at 2150 UTC (shown in Figure 6(a)) to a more concentrated area of convergence beneath the area of the mesocyclone by 2220 UTC (Figure 6(d)). Just before this at 2200 utc (Figure 6(b)) and then at 2210 UTC (Figure 6(c)) divergence is strongest wrapping around (to the west and south of) the area of convergence associated with the RFD. The advancement of the RFD around the southwestern region of the updraft at 2200 UTC coincides with the development of the tornado. The mesocyclone tracks across an area of enhanced low-level convergence on the front flank of the storm similar to that in the Lemon and Doswell model [4]. Strong upper-level divergence is visible at the 12 km level in the analysis (Figure 7), which can be attributed to the assimilation of data from radars that are located far away from the storm. As the storm progressed to the east-northeast, the divergent nature of the outflow weakened.

The dipole vorticity structure simulated by Klemp and Wilhelmson [49] and discussed by Rotunno [50] is clearly depicted in this analysis (Figure 8). The right split storm exhibits strong positive vertical vorticity at the beginning of the period with a weaker anticyclonic counterpart on the northwest side of the mesocyclone. The left split storm also exhibits signs of the dipole vorticity structure opposite to that of the right split but with weaker magnitudes of vorticity and vertical velocities. Since the hodograph in Figure 3 favors a right split storm, the left split counterpart never evolves into a mature storm [13]. It is shown that the vertical vorticity is strongest at 2210 UTC when the tornado touched down (Figure 8(c)).

3.2. Sensitivity of the Analysis to Different Number of Radars. To investigate the contribution of individual radars to the control analysis, two more 3DVAR analyses are performed only at 2200 UTC. One uses data from KOKC only (IRAD),

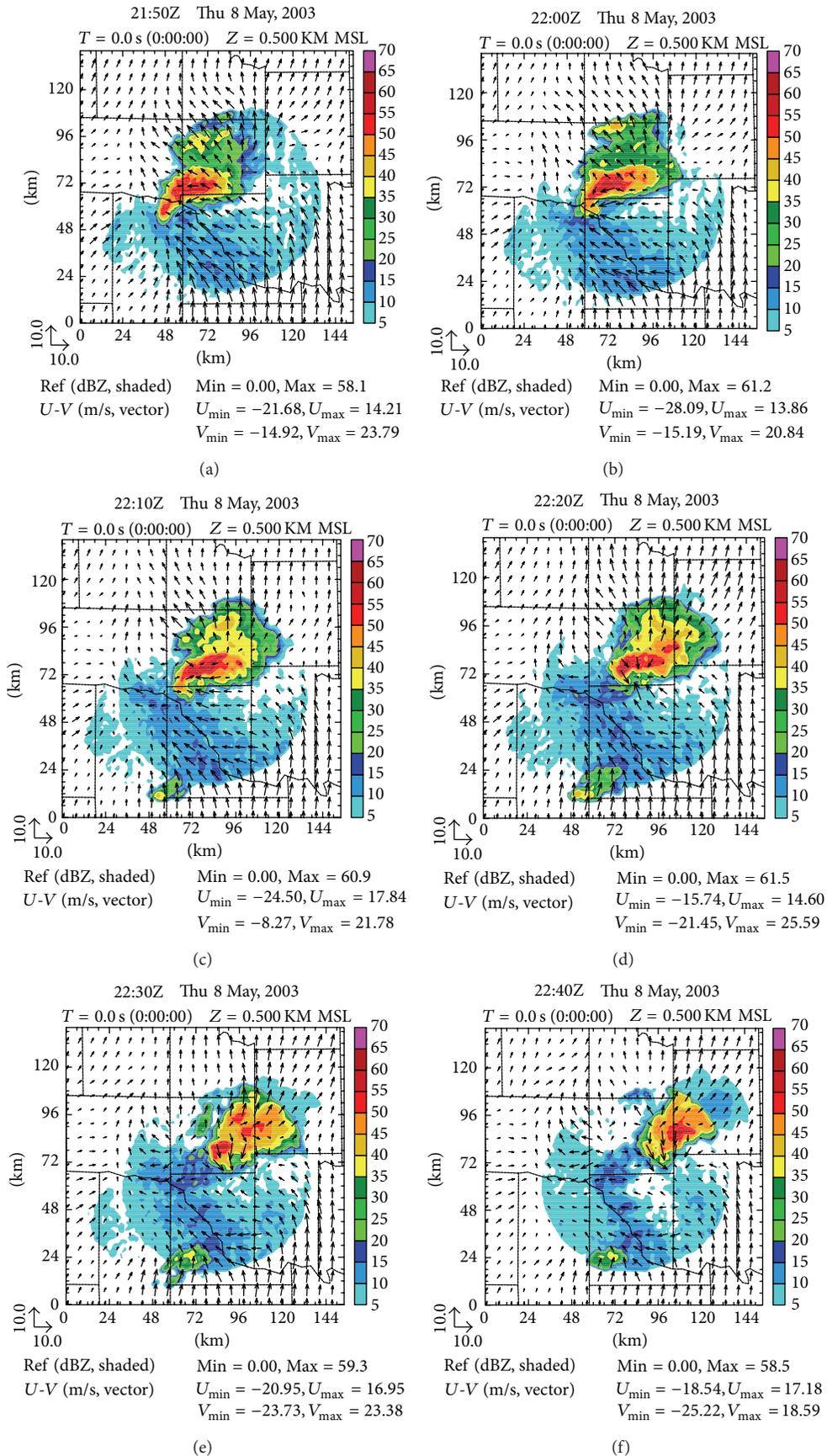


FIGURE 5: 0.5 km radar reflectivity and horizontal wind vectors from the control analysis from (a) 2150 UTC to (f) 2240 UTC at 10-minute intervals.

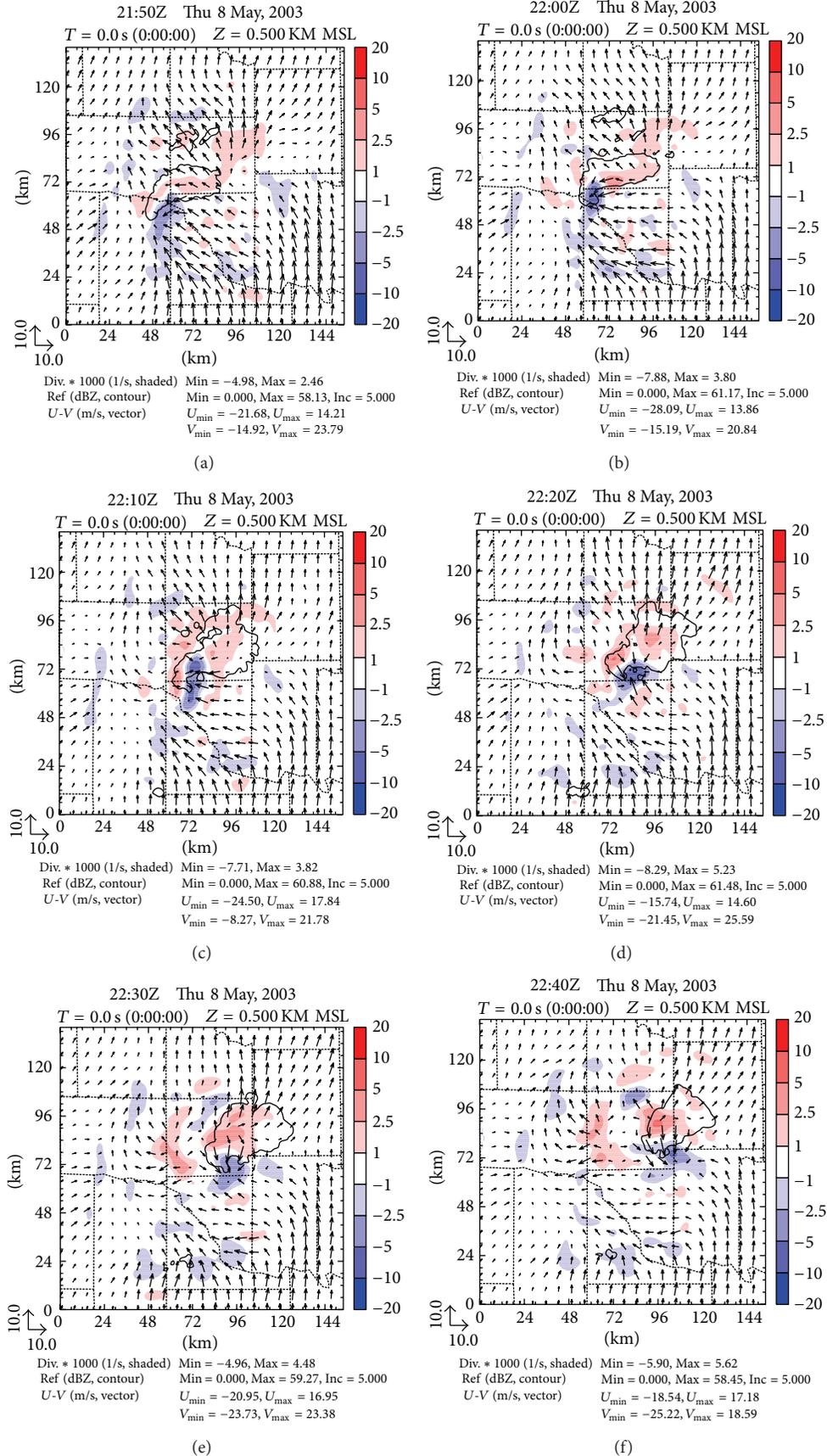


FIGURE 6: 0.5 km horizontal divergence (10^{-3} s^{-1}) and horizontal wind vectors with the 35 dBZ reflectivity contour outlined from the control analysis from (a) 2150 UTC to (f) 2240 UTC at 10-minute intervals.

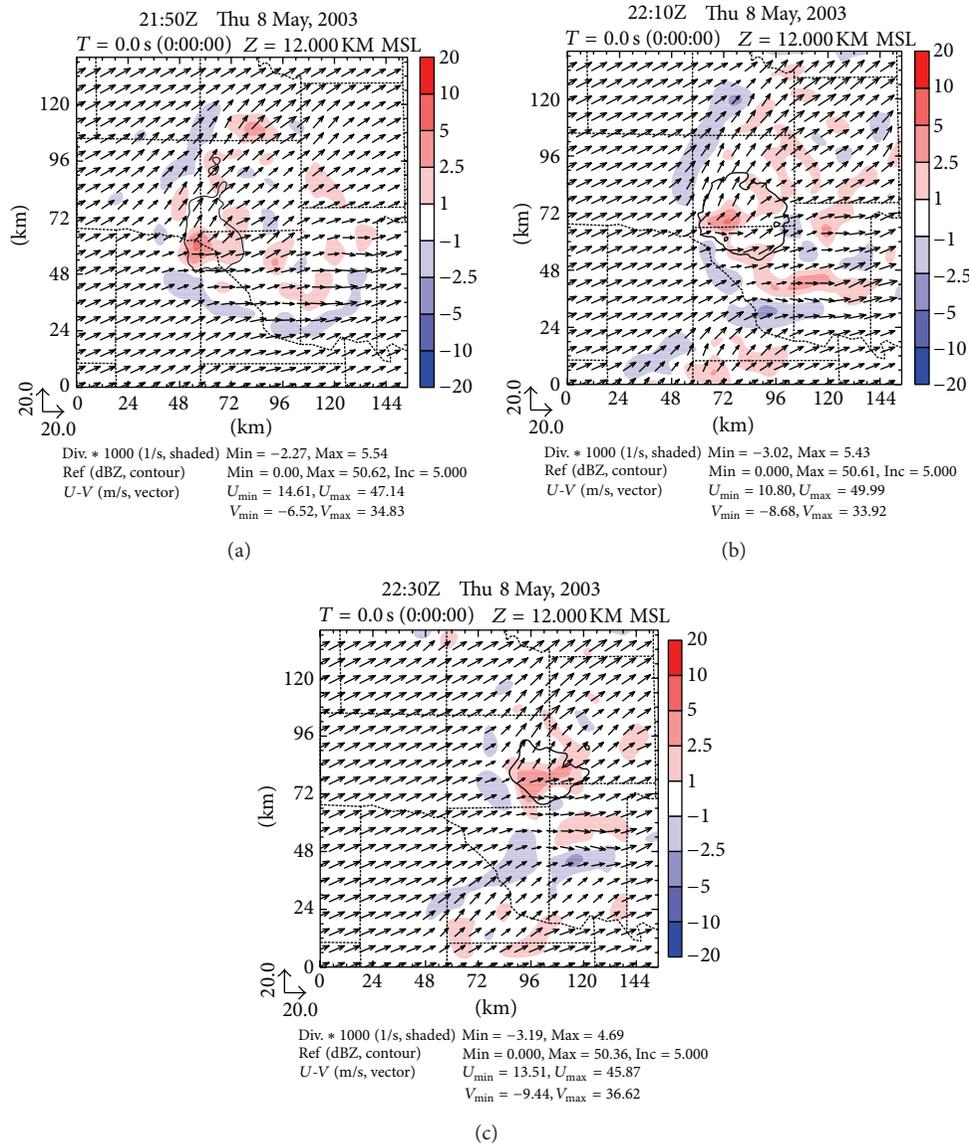


FIGURE 7: Same as Figure 6 (horizontal divergence, 10^{-3} s^{-1}) but for 12 km height and 20-minute intervals.

and another uses data from KOKC and KTLX (2RAD). When assimilating radar data over a period of time, the effects of unavailable data above the highest elevation level of the volume scan, referred to as the cone of silence, may be minimized by using a storm scale NWP model to estimate quantities in regions that are unobservable [34–36]. For this experiment, however, the analysis does not make use of data assimilation cycles from a storm-scale NWP model to fill in the gaps between observations. Therefore, in regions of no radar coverage, the analysis relies heavily on the background, the spatial spreading of observation information through background error correlations, and the mass continuity constraint to fill in the gaps (conventional upper air data are usually very limited). The impact of decreased radar coverage is most clearly depicted in Figure 9.

When radar data from KOKC only (1RAD), located closest to the storm, are used, the ability to obtain cross-beam winds is very limited. Even at the 3 km level (Figure 9(e)), a significant part of the circulation in the storm is missed. At the top level (12 km), both the reflectivity and storm outflow are almost completely missed (Figure 9(h)). When data from both KOKC and KTLX (2RAD), which intersect the hook of the storm at about 45 degrees, are used, the accuracy of the analyzed wind is increased dramatically. This is most visible along the front flank of the thunderstorm at the lowest levels just east of the mesocyclone where winds are parallel to the cross-beam component of the KOKC radar (Figures 9(a), 9(b), and 9(c)). The most notable impact of the radars is in the reflectivity structure. Both experiments with one and two radars that are close to the storm miss significant parts of

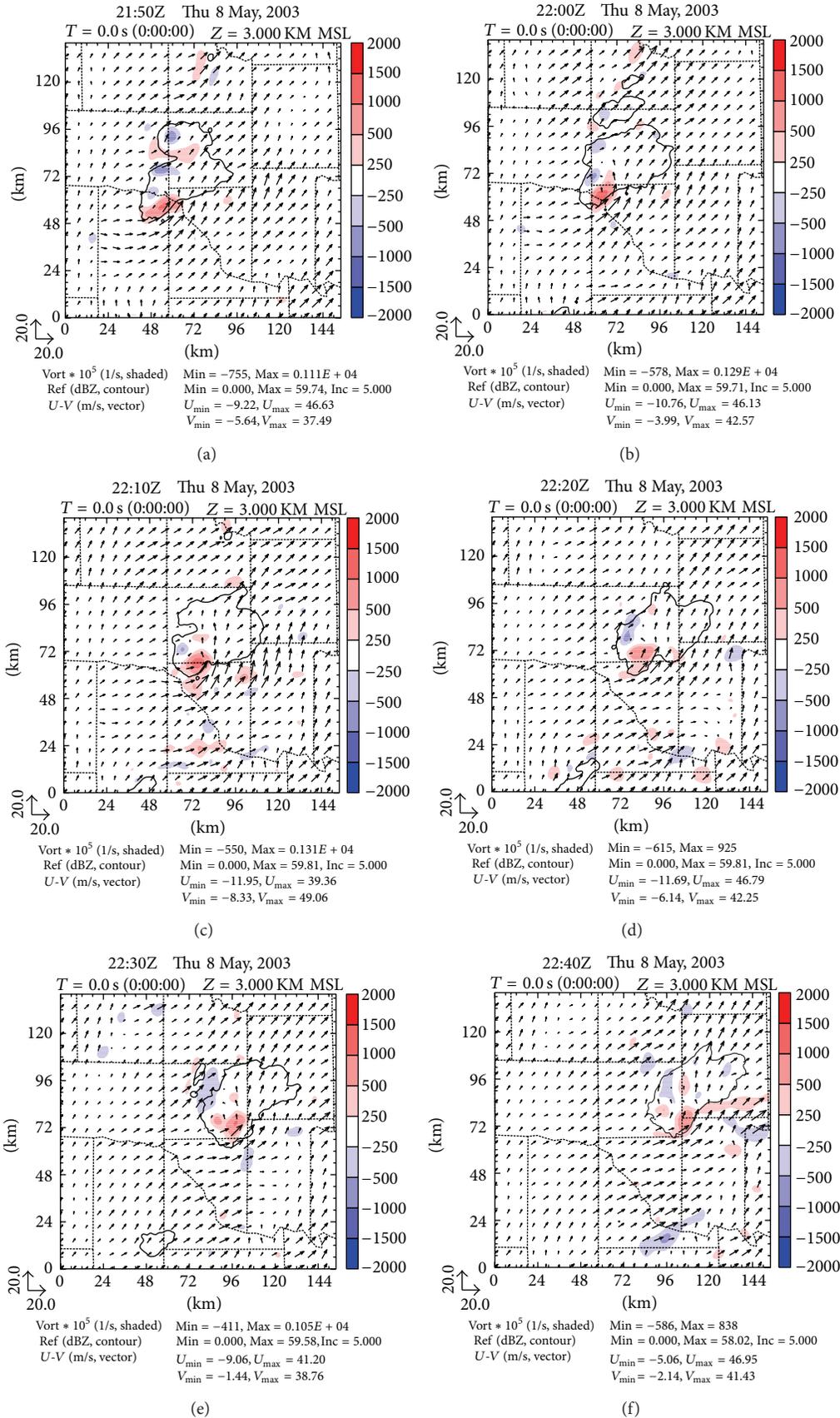


FIGURE 8: 3 km vertical vorticity (10^{-5} s^{-1}) and horizontal wind vectors with the 35 dBZ reflectivity contour outlined from the control analysis from (a) 2150 UTC to (f) 2240 UTC at 10-minute intervals.

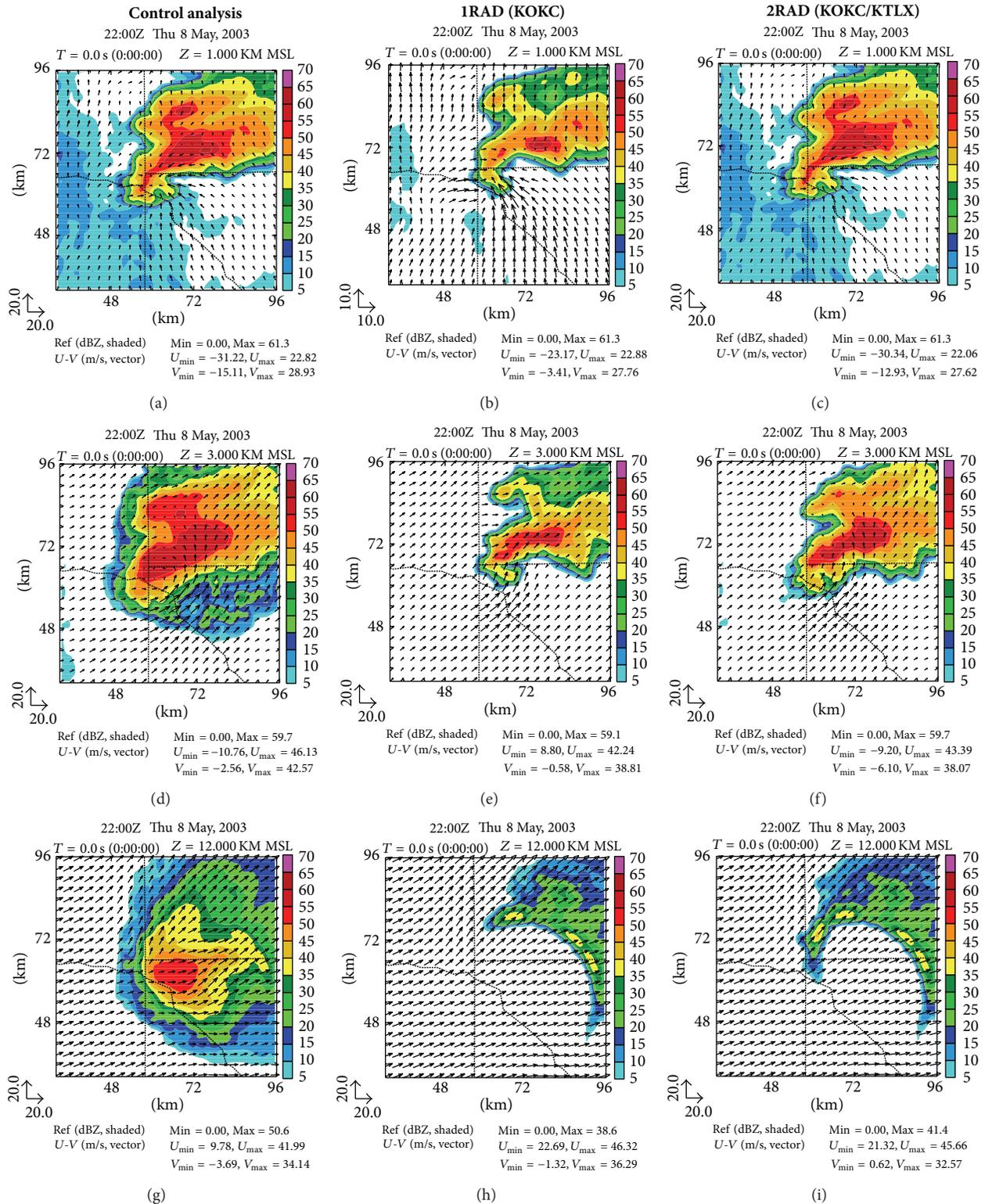


FIGURE 9: Reflectivity and horizontal wind vectors at 2200 UTC for the control analysis at (a) 1 km, (d) 3 km, and (g) 12 km level and for the IRAD experiment ((b), (e), (h)) and for the 2RAD experiment ((c), (f), (i)).

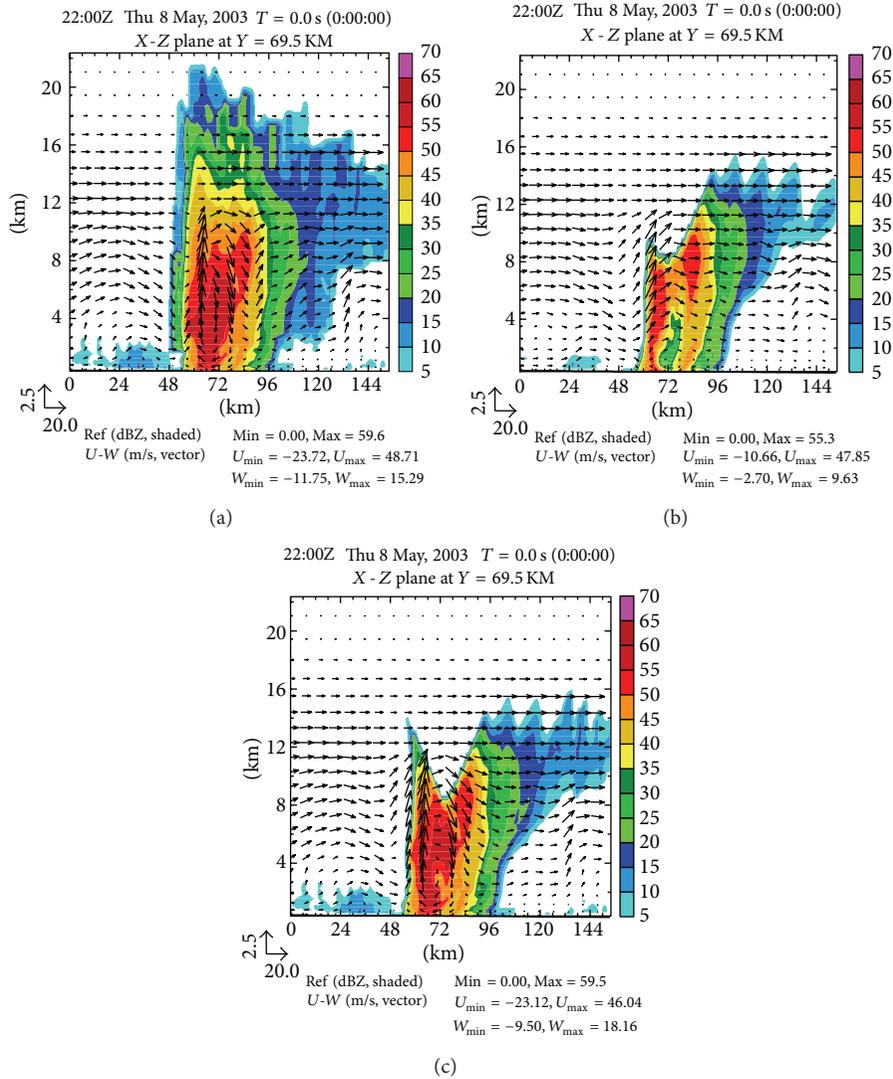


FIGURE 10: East-west cross section of reflectivity and wind vectors through the center of the storm at 2200 UTC for the (a) control analysis, (b) the 1RAD analysis, and (c) the 2RAD analysis.

the reflectivity structure owing to the cone of silence (Figures 9(h), and 9(i)). This artifact is eliminated when data from three other radars (KVNK, KINX, and KFDR) far away from the storm are used, as depicted in the control analysis (Figures 9(a), 9(d), and 9(g)). The ability of the analysis scheme to resolve the structure of the updrafts and downdrafts within the storm is also improved when using all available radars. This is most evident when looking at a cross section as shown in Figure 10. The higher vertical velocities in the updraft are resolved only when at least two radars are used. In the control analysis, the additional radar data from radars located far from the storm help resolve the anvil structure above the 10 km level (Figure 10(a)) up to 20 km.

When assimilating data from only KOKC, low-level convergence is weak, and upper-level divergence is almost missing (Figures 11(b) and 11(e)). When two close radars KTLX and KOKC are used, the low-level convergence is more pronounced when compared to using a single radar (Figure 11(b)

versus Figure 11(c)), but the upper-level divergence is still missing. Only when assimilating data from radars located further away from the storm are used, do areas of upper-level divergence, which are consistent with the Lemon and Doswell [4] model, appear in the analysis (Figure 11(d)).

When observations of the storm structure are incomplete (as with the analyses performed by using data from only one radar), the variational analysis technique acts to fill in the gaps with values obtained from other sources (surface and upper air), as well as information from the background state. While providing a reasonable estimate to the state of the atmosphere in synoptic and mesoscale data assimilation, sounding and profiler data, when available, are rather crude estimates of the true state of the atmosphere when deep convection is present.

In summary, the control analysis captures the structure and evolution of the supercell using information from multiple radars. Features at both the low and upper levels are

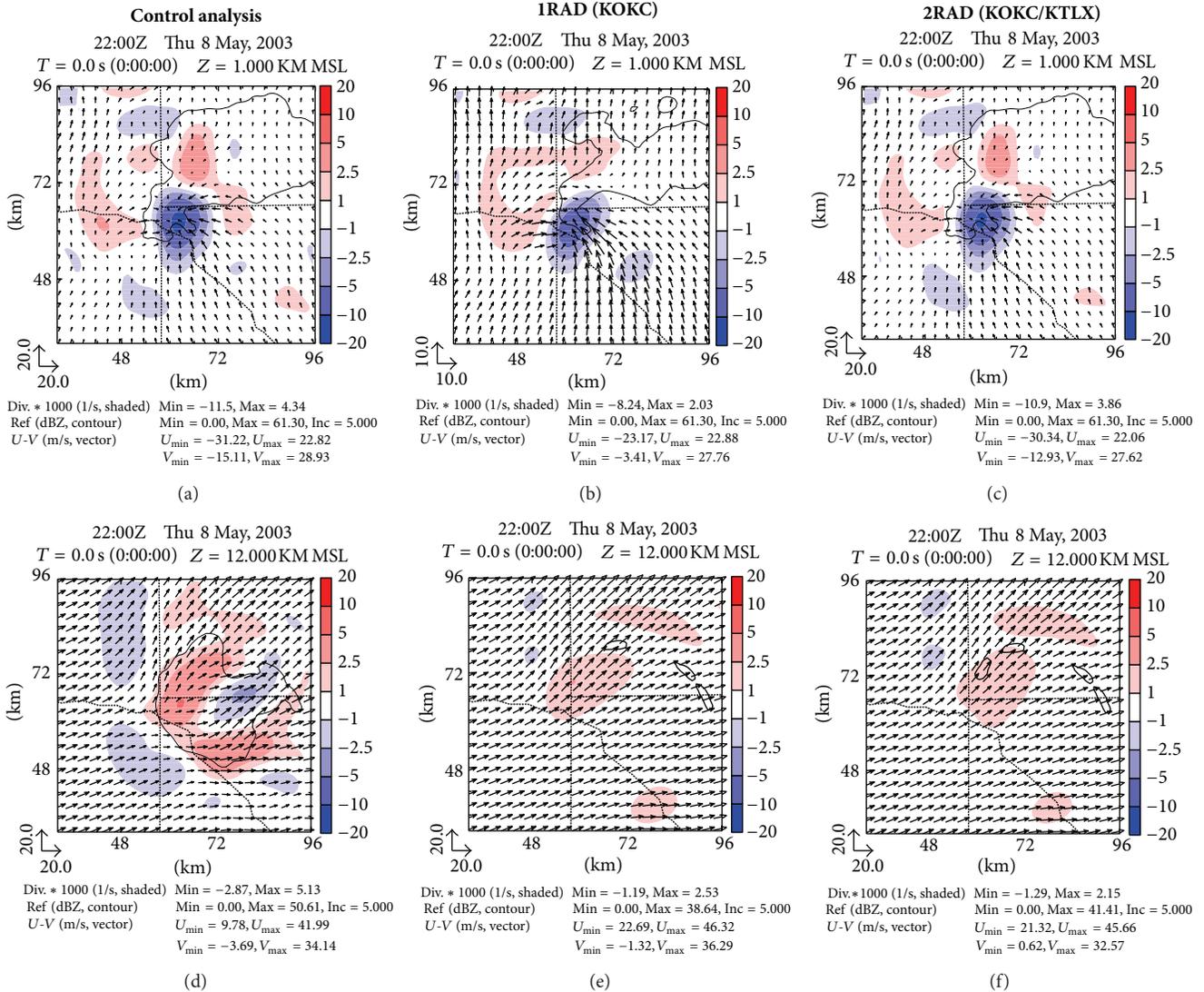


FIGURE 11: Divergence (10^{-3} s^{-1}) and horizontal wind vectors at 2200 UTC with 35 dBZ reflectivity contour outlined in black for the control analysis at (a) 1 km, (d) 12 km, and level and for the 1RAD experiment ((b), (e)) and for the 2RAD experiment ((c), (f)).

well resolved. These storm scale features are most apparent when examining convergence and vorticity within the analysis that depict a developing and strengthening low-level mesocyclone. Thus, a complete set of analyses of the storm is obtained when using data from radars located both close and far away from the storm, a strategy often overlooked in previous investigations.

3.3. Sensitivity of the Forecast to Different Number of Radars. Three forecast experiments are performed using initial conditions similar to those produced with data from one, two, and five radars to examine the sensitivity of a forecast to the number of radars used. To limit the amount of noise from radar data, only radar reflectivity and radial velocity in regions of high reflectivity (greater than 50 dBZ) and close to the convection were assimilated. All three forecasts are initialized using the 3DVAR analysis at 2155 UTC.

The forecast 0.5 km vertical vorticity and horizontal wind vectors between 2200 UTC and 2220 UTC are shown in Figure 12 at ten-minute intervals with the observed tornado damage path. The KOKC forecast, assimilating data from only one radar, shows the storm moving in a north-northeasterly direction, away from the actual tornado damage path. In addition, the simulation produces only low levels of positive and negative vorticity. The 2RAD forecast (Figures 12(d)–12(f)) used data from only the KTLX and KOKC radars, and although it captured more of the low-level wind structure than the KOKC experiment, their volume scans still did not observe the top portion of the storm. This forecast predicts storm movement closer to the observed damage path than the KOKC experiment, but still to the north of the observed tornado damage path. In general, both these forecasts without the observations in the upper levels of the thunderstorm deviated to the north of the tornado damage path. The control

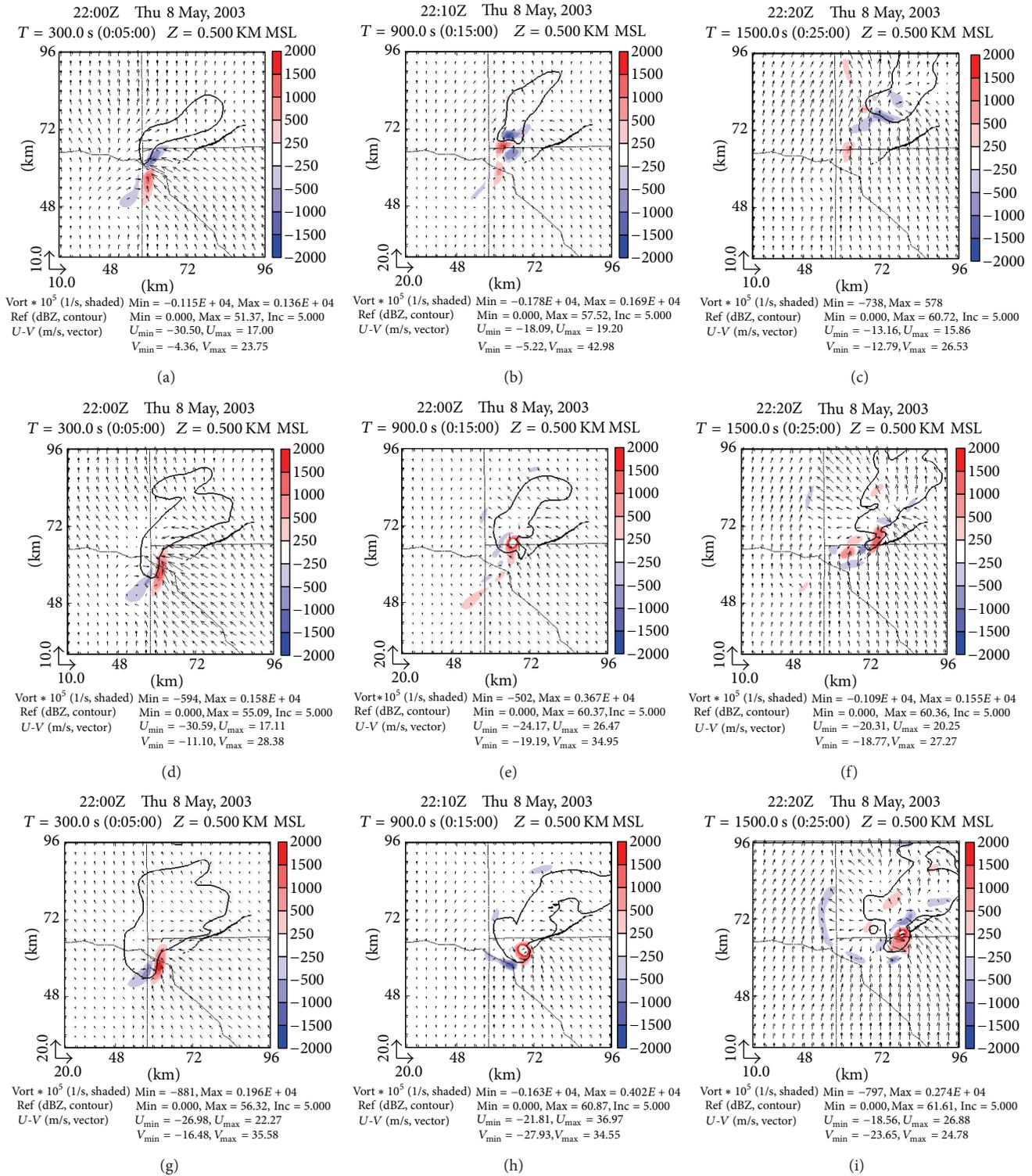


FIGURE 12: 0.5 km vertical vorticity (10^{-5} s^{-1}) and horizontal wind vectors for the IRAD forecast ((a)–(c)), 2RAD forecast ((d)–(f)), and control forecast ((g)–(i)) initialized at 2155 UTC and shown at 10-minute intervals starting at 2200 UTC until 2220 UTC. Vorticity values greater than $2 \times 10^{-2} \text{ s}^{-1}$ are shown as transparent. The thicker black line denotes the location of the observed damage path of the tornado.

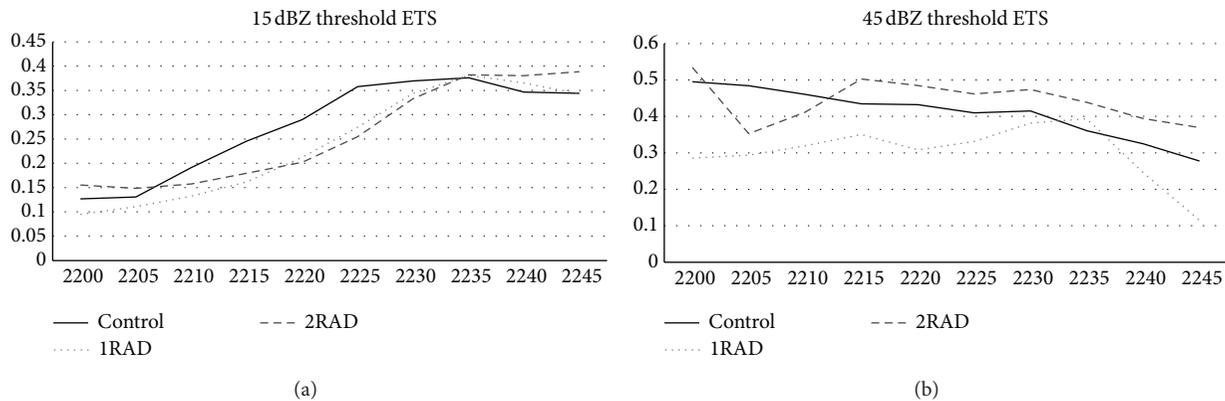


FIGURE 13: Equitable threat scores (ETS) of predicted radar reflectivity for (a) 15 and (b) 45 dBZ validated using KTLX radar composite reflectivity.

forecast (using 5 radars) shows a strong maximum in vorticity (Figures 12(g)–12(i)) close to the actual storm track. The storm propagates in an east-northeasterly direction closer to the observed tornado damage path than the previous two experiments. The inclusion of radar data in the upper levels of the convection seems to greatly improve the accuracy of prediction of the simulated storms mesocyclone in relation to the actual tornado damage path.

To quantify the improvement of the predicted storm for the three experiments the equitable threat scores (ETS; [51]) of the predicted reflectivity were calculated using KTLX radar composite reflectivity. The ETS scores using a 15 dBZ threshold are shown in Figure 13(a). There is a pronounced improvement in the ETS scores from 2210 through 2235 UTC for the control experiment when compared to the experiments which only assimilated data from one or two radars. Between 2205 and 2215 UTC the control run does produce higher scores for the 45 dBZ threshold than the other two experiments but the improvement is degraded after 2215 UTC.

These results are similar to that of Xue et al. [36] for a 1 km forecast of the same thunderstorm but with the use of a cycling 3DVAR analysis. In their study, only one radar was used, KTLX, and their storm also propagated to the north of the observed tornado damage path. The cycling of the 3DVAR analysis in Xue et al. [36] seemed to have a positive impact when compared to the single-time analysis used in this study. In contrast, our study shows that the track of the storm is closer to the actual tornado damage path when assimilating all available radar data.

4. Conclusion

In this study, the ARPS 3DVAR system with a cloud analysis package is used to assimilate radar observations for the 8 May 2003 tornadic supercell thunderstorm in central Oklahoma. The analysis includes observations from five radars, four operational WSR-88D radars (KTLX, KVNK, KINX, and KFDR) and one terminal Doppler weather radar (KOKC). Results show that the whole storm can be properly analyzed

by including radar observations further away from the storm in addition to those located close to the storm. This aspect has not received enough attention in previous studies of convective storm data assimilation. The radars located further away from the storm are able to observe the upper levels of the storm that cannot be captured by the radars located very close to the storm owing to the “cone of silence”, an unobservable volume centered directly above the radar due to the upper limit of the scanning elevation angles. In analyses performed using data from one or two radars located close to the storm, the vertical vorticity and upper-level divergence structure are not analyzed as well when compared to the control analyses using all five radars.

The resulting analyses are then used as initial conditions for short-term storm-scale forecasts. Forecasts initialized from analyses using data from fewer radars proved to be less accurate. These forecasts had noticeably lower values of maximum low-level vertical vorticity and the track of maximum vorticity deviated from the observations more than that from a forecast using data from five radars. The computational efficiency of the 3DVAR technique makes it very suitable for short-term operational storm-scale forecasts, as has also been demonstrated by real-time radar data assimilation experiments over the continental United States [40].

References

- [1] E. A. Brandes, “Flow in severe thunderstorms observed by dual-Doppler radar,” *Monthly Weather Review*, vol. 105, pp. 113–120, 1977.
- [2] E. A. Brandes, “Mesocyclone evolution and tornadogenesis: some observations,” *Monthly Weather Review*, vol. 106, pp. 995–1011, 1978.
- [3] D. W. Burgess, L. R. Lemon, and R. A. Brown, “Tornado characteristics revealed by Doppler radar,” *Geophysical Research Letters*, vol. 2, pp. 183–184, 1975.
- [4] L. R. Lemon and C. A. Doswell III, “Severe thunderstorm evolution and mesocyclone structure as related to tornadogenesis,” *Monthly Weather Review*, vol. 107, no. 9, pp. 1184–1197, 1979.
- [5] J. S. Juazhen Sun, D. W. Flicker, and D. K. Lilly, “Recovery of three-dimensional wind and temperature fields from simulated

- single-Doppler radar data,” *Journal of the Atmospheric Sciences*, vol. 48, no. 6, pp. 876–890, 1991.
- [6] T. Gal-Chen, “A method for the initialization of the anelastic equations: implications for matching models with observations,” *Monthly Weather Review*, vol. 106, pp. 587–606, 1978.
- [7] C. E. Hane, R. B. Wilhelmson, and T. Gal-Chen, “Retrieval of thermodynamic variables within deep convective clouds: experiments in three dimensions,” *Monthly Weather Review*, vol. 109, no. 3, pp. 564–576, 1981.
- [8] T. Gal-Chen and R. A. Kropfli, “Buoyancy and pressure perturbations derived from dual-Doppler radar observations of the planetary boundary layer: applications for matching models with observations,” *Journal of the Atmospheric Sciences*, vol. 41, no. 20, pp. 3007–3020, 1984.
- [9] A. Crook, “Numerical simulations initialized with radar-derived winds. Part I: simulated data experiments,” *Monthly Weather Review*, vol. 122, no. 6, pp. 1189–1203, 1994.
- [10] A. Crook and J. D. Tuttle, “Numerical simulations initialized with radar-derived winds. Part II: forecasts of three gust-front cases,” *Monthly Weather Review*, vol. 122, no. 6, pp. 1204–1217, 1994.
- [11] S. S. Weygandt, A. Shapiro, and K. K. Droegemeier, “Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part I: single-Doppler velocity retrieval,” *Monthly Weather Review*, vol. 130, no. 3, pp. 433–453, 2002.
- [12] S. S. Weygandt, A. Shapiro, and K. K. Droegemeier, “Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part II: thermodynamic retrieval and numerical prediction,” *Monthly Weather Review*, vol. 130, no. 3, pp. 454–476, 2002.
- [13] R. Rotunno and J. B. Klemp, “The influence of shear-induced pressure gradient on thunderstorm motion,” *Monthly Weather Review*, vol. 110, no. 2, pp. 136–151, 1982.
- [14] J. Sun and N. A. Crook, “Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: model development and simulated data experiments,” *Journal of the Atmospheric Sciences*, vol. 54, no. 12, pp. 1642–1661, 1997.
- [15] J. Sun and N. A. Crook, “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,” *Journal of the Atmospheric Sciences*, vol. 55, no. 5, pp. 835–852, 1998.
- [16] J. Sun and N. A. Crook, “Real-time low-level wind and temperature analysis using single WSR-88D data,” *Weather and Forecasting*, vol. 16, no. 1, pp. 117–132, 2001.
- [17] J. Gao, M. Xue, A. Shapiro, and K. K. Droegemeier, “A variational method for the analysis of three-dimensional wind fields from two Doppler radars,” *Monthly Weather Review*, vol. 127, no. 9, pp. 2128–2142, 1999.
- [18] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, “A three-dimensional variational data analysis method with recursive filter for Doppler radars,” *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [19] R. J. Purser, W.-S. Wu, D. F. Parrish, and N. M. Roberts, “Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: spatially homogeneous and isotropic Gaussian covariances,” *Monthly Weather Review*, vol. 131, pp. 1524–1535, 2003.
- [20] M. Hu, M. Xue, and K. Brewster, “3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: cloud analysis and its impact,” *Monthly Weather Review*, vol. 134, no. 2, pp. 675–698, 2006.
- [21] M. Hu, M. Xue, J. Gao, and K. Brewster, “3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: impact of radial velocity analysis via 3DVAR,” *Monthly Weather Review*, vol. 134, no. 2, pp. 699–721, 2006.
- [22] G. Evensen, “Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics,” *Journal of Geophysical Research*, vol. 99, no. 5, pp. 10143–10162, 1994.
- [23] C. Snyder and F. Zhang, “Assimilation of simulated Doppler radar observations with an ensemble Kalman filter,” *Monthly Weather Review*, vol. 131, no. 8, pp. 1663–1677, 2003.
- [24] F. Zhang, C. Snyder, and J. Sun, “Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter,” *Monthly Weather Review*, vol. 132, pp. 1238–1253, 2004.
- [25] M. Tong and M. Xue, “Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments,” *Monthly Weather Review*, vol. 133, no. 7, pp. 1789–1807, 2005.
- [26] M. Xue, M. Tong, and K. K. Droegemeier, “An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting,” *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 1, pp. 46–66, 2006.
- [27] D. C. Dowell, L. J. Wicker, and D. J. Stensrud, “High-resolution analyses of the 8 May 2003 Oklahoma City storm. Part II: EnKF data assimilation and forecast experiments,” in *Proceedings of the 22nd Conference on Severe Local Storms*, The American Meteorological Society, Hyannis, Mass, USA, 2004, Extended Abstract 12.5.
- [28] N. Snook, M. Xue, and Y. Jung, “Analysis of a tornadic mesoscale convective vortex based on ensemble kalman filter assimilation of CASA X-band and WSR-88D radar data,” *Monthly Weather Review*, vol. 139, no. 11, pp. 3446–3468, 2011.
- [29] N. Snook, M. Xue, and Y. Jung, “Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data,” *Monthly Weather Review*, vol. 140, pp. 2126–2146, 2012.
- [30] R. A. McPherson, C. A. Fiebrich, K. C. Crawford et al., “State-wide monitoring of the mesoscale environment: a technical update on the Oklahoma Mesonet,” *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 3, pp. 301–321, 2007.
- [31] L. J. Wicker and D. Dowell, “High-resolution analyses of the 8 May 2003 Oklahoma City storm. Part III: an ultra-high resolution forecast experiment,” in *Proceedings of the 22nd Conference on Severe Local Storms*, The American Meteorological Society, Hyannis, Mass, USA, 2004, Extended Abstract 12.6.
- [32] D. W. Burgess, “High-resolution analyses of the 8 May 2003 Oklahoma City storm. Part I: storm structure and evolution from radar data,” in *Proceedings of the 22nd Conference on Severe Local Storms*, The American Meteorological Society, Hyannis, Mass, USA, 2004, Extended Abstract 12.4.
- [33] A. Witt, R. A. Brown, and V. Lakshmanan, “Real-time calculation of horizontal winds using multiple Doppler radars: a new WDSS-II module,” in *Proceedings of the 11th Conference on Mesoscale Processes and the 32nd Conference on Radar Meteorology*, pp. 1171–1181, Albuquerque, NM, USA, October 2005.

- [34] M. Hu and M. Xue, "Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case," *Monthly Weather Review*, vol. 135, no. 2, pp. 507–525, 2007.
- [35] M. Hu and M. Xue, "Analysis and prediction of the 8 May 2003 Oklahoma City tornadic thunderstorm and embedded tornado using ARPS with assimilation of WSR-88D radar data," in *Proceedings of the 18th Conference on Numerical Weather Prediction*, The American Meteorological Society, Park City, Utah, USA, 2007, Extended Abstract 1B.4.
- [36] M. Xue, M. Hu, and A. Schenkman, "Numerical prediction of 8 May 2003 Oklahoma City tornadic supercell and embedded tornado using ARPS with assimilation of WSR-88D radar data," *Weather and Forecasting*. In press.
- [37] M. Xue, K. K. Droegemeier, and V. Wong, "The advanced regional prediction system (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: model dynamics and verification," *Meteorology and Atmospheric Physics*, vol. 75, no. 3-4, pp. 161–193, 2000.
- [38] M. Xue, K. K. Droegemeier, V. Wong et al., "The advanced regional prediction system (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: model physics and applications," *Meteorology and Atmospheric Physics*, vol. 76, no. 3-4, pp. 134–165, 2001.
- [39] M. Xue, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, "The advanced regional prediction system (ARPS), storm-scale numerical weather prediction and data assimilation," *Meteorology and Atmospheric Physics*, vol. 82, no. 1–4, pp. 139–170, 2003.
- [40] M. Xue, O. K. Norman, F. Kong et al., "CAPS realtime storm scale ensemble and high resolution forecasts for the NOAA Hazardous Weather Testbed 2010 Spring Experiment," in *Proceedings of the 24th Conference on Weather and Forecasting and 20th Conference on Numerical Weather Prediction*, The American Meteorological Society, 2011, Paper 9A.2.
- [41] A. J. Clark, S. J. Weiss, J. S. Kain et al., "An overview of the 2010 Hazardous Weather Testbed experimental forecast program spring experiment," *Bulletin of the American Meteorological Society*, vol. 93, pp. 55–74, 2012.
- [42] P. Courtier, J.-N. Thépaut, and A. Hollingsworth, "A strategy for operational implementation of 4D-Var, using an incremental approach," *Quarterly Journal*, vol. 120, no. 519, pp. 1367–1387, 1994.
- [43] S. C. Albers, J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, "The local analysis and prediction system (LAPS): analyses of clouds, precipitation, and temperature," *Weather and Forecasting*, vol. 11, no. 3, pp. 273–287, 1996.
- [44] J. Zhang, F. H. Carr, and K. Brewster, "ADAS cloud analysis," in *Proceedings of the 12th Conference on Numerical Weather Prediction*, pp. 185–188, The American Meteorological Society, Phoenix, Ariz, USA, 1998.
- [45] K. Brewster, "Recent advances in the diabatic initialization of a non-hydrostatic numerical model," in *Proceedings of the 15th Conference on Numerical Weather Prediction and 21st Conference on Severe Local Storms and 19th Conference on Weather Analysis and Forecasting*, The American Meteorological Society, San Antonio, Tex, USA, 2002.
- [46] Y.-L. Lin, R. D. Farley, and H. D. Orville, "Bulk parameterization of the snow field in a cloud model," *Journal of Climate & Applied Meteorology*, vol. 22, no. 6, pp. 1065–1092, 1983.
- [47] N. Snook and M. Xue, "Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms," *Geophysical Research Letters*, vol. 35, no. 24, Article ID L24803, 2008.
- [48] L. R. Lemon, D. W. Burgess, and R. A. Brown, "Tornado production and storm sustenance," in *Proceedings of the 9th Conference on Severe Local Storms*, pp. 100–104, The American Meteorological Society, Boston, Mass, USA, 1975.
- [49] J. B. Klemp and R. B. Wilhelmson, "Simulations of right- and left-moving storms produced through storm splitting," *Journal of the Atmospheric Sciences*, vol. 35, pp. 1097–1110, 1978.
- [50] R. Rotunno, "On the evolution of thunderstorm rotation," *Monthly Weather Review*, vol. 109, no. 3, pp. 577–586, 1981.
- [51] J. T. Schaefer, "The critical success index as an indicator of warning skill," *Weather and Forecasting*, vol. 5, pp. 570–575, 1990.

Research Article

Prediction of Convective Storms at Convection-Resolving 1km Resolution over Continental United States with Radar Data Assimilation: An Example Case of 26 May 2008 and Precipitation Forecasts from Spring 2009

Ming Xue, Fanyou Kong, Kevin W. Thomas, Jidong Gao, Yunheng Wang, Keith Brewster, and Kelvin K. Droegemeier

Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David Boren Boulevard, Norman, OK 73072, USA

Correspondence should be addressed to Ming Xue; mxue@ou.edu

Received 31 May 2013; Revised 2 November 2013; Accepted 3 November 2013

Academic Editor: Kun Zhao

Copyright © 2013 Ming Xue et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For the first time ever, convection-resolving forecasts at 1 km grid spacing were produced in realtime in spring 2009 by the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma. The forecasts assimilated both radial velocity and reflectivity data from all operational WSR-88D radars within a domain covering most of the continental United States. In preparation for the realtime forecasts, 1 km forecast tests were carried out using a case from spring 2008 and the forecasts with and without assimilating radar data are compared with corresponding 4 km forecasts produced in realtime. Significant positive impact of radar data assimilation is found to last at least 24 hours. The 1 km grid produced a more accurate forecast of organized convection, especially in structure and intensity details. It successfully predicted an isolated severe-weather-producing storm nearly 24 hours into the forecast, which all ten members of the 4 km real time ensemble forecasts failed to predict. This case, together with all available forecasts from 2009 CAPS realtime forecasts, provides evidence of the value of both convection-resolving 1 km grid and radar data assimilation for severe weather prediction for up to 24 hours.

1. Introduction

Accurate prediction of convective-scale hazardous weather continues to be a major challenge. Efforts to explicitly predict convective storms using numerical models dated back to Lilly [1] and began with the establishment in 1989 of an NSF Science and Technology Center, the Center for Analysis and Prediction of Storms at the University of Oklahoma (CAPS). Over the past two decades, steady progress has been made, aided by steady increases in available computing power. Still, the resolutions of the current-generation operational numerical weather prediction (NWP) models remain too low to explicitly resolve convection, limiting the accuracy of quantitative precipitation forecasts.

For over a decade, the research community has been producing experimental real time forecasts at 3-4 km convection-allowing resolutions (e.g., [2-4]). Roberts and

Lean [5] documented that convection forecasts of up to 6 hours are more skillful when run on a 1 km grid than on a 12 km grid, and more so than on a 4 km grid. On the other hand, Kain et al. [2] found no appreciable improvement with 2 km forecasts compared to 4 km forecasts beyond 12 hours.

In the spring seasons of 2007 and 2008, CAPS conducted more systematic real-time experiments. Daily forecasts of 30 h or more were produced for 10-member 4 km ensembles and 2 km deterministic forecasts ([6, 7], X07 and X08 hereafter). In 2008, radial velocity (V_r) and reflectivity (Z) data from all operational radars in a domain covering most of the CONUS (continental US) were assimilated [7] using a combined 3DVAR-cloud analysis method [8, 9]. Standard precipitation verification scores show that significant positive impact of radar data lasts up to 9 hours but the difference in scores between the 4 and 2 km forecasts is relatively small [7, 10].

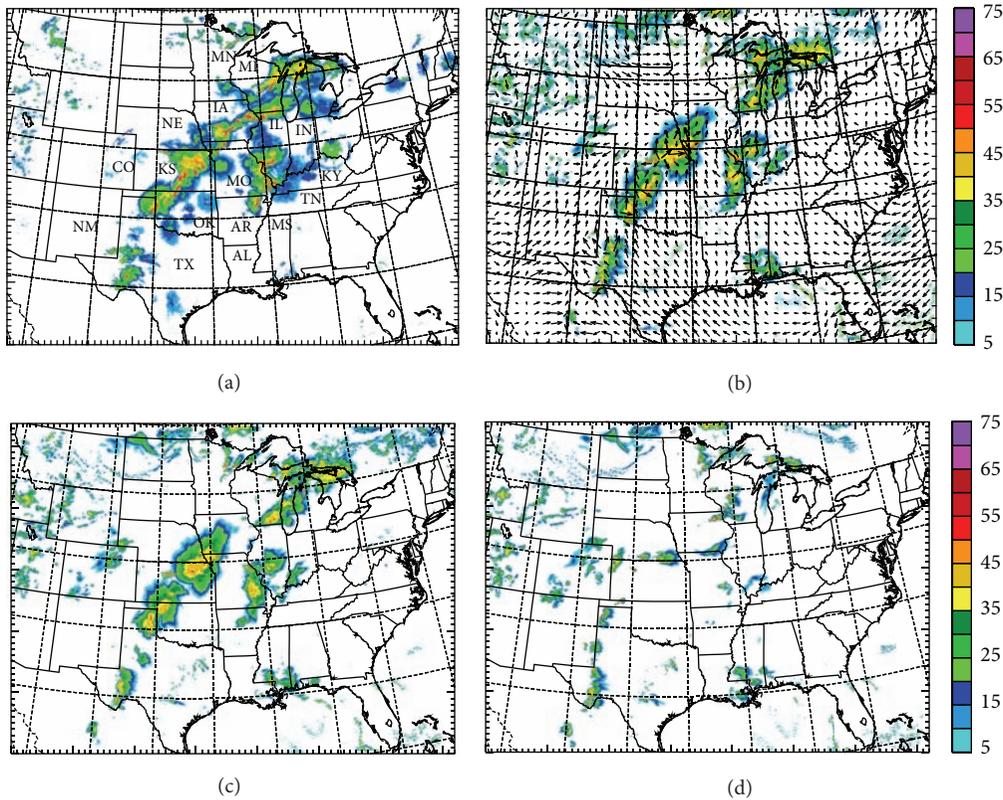


FIGURE 1: Observed composite radar reflectivity at 0300 UTC, May 26, 2008 (a) and 3-hour forecasts of the same field valid at the same time from (b) the 1 km forecast with radar data assimilation, (c) 4 km control forecast with radar data assimilation, and (d) 4 km forecast without radar data. Panel (b) includes surface wind vectors at 10 m AGL plotted at every 80th grid point.

Recognizing that producing better convective forecasts requires accurately resolving the internal structures of convective storms, the CAPS team carried out real-time 1 km resolution forecasts assimilating radar data from mid-April through early June, 2009 [11]. Daily 30-hour forecasts used 9600 processor cores of a Cray XT5 supercomputer at the National Institute of Computational Science, University of Tennessee. Each forecast took about 5.5 hours to complete. In preparation for such forecasts, tests were made using cases from the spring of 2008 and they represented the first time ever that forecasts at a 1 km resolution were produced for a large domain covering the entire CONUS, assimilating all available data from the operational weather radars in the domain (see Figure 1). In this paper, we document the results of one of the 1 km tests as they were produced in early 2009 in preparation of the 2009 CAPS spring forecast experiment and compare the forecasts produced at 4 km grid spacing with and without radar data assimilation that were produced in realtime in 2008. We also present briefly the mean precipitation skill scores from the spring 2009 forecasts, produced at 1 km and 4 km grid spacing with radar data assimilation and 4 km forecasts without radar data, together with their comparisons with the reference NAM forecasts.

The rest of this paper is organized as follows. Section 2 describes the forecast model configurations, and Sections 3 and 4 present and discuss the results. A summary is given in Section 5.

2. Forecast Configurations

The 26 May 2008 test case is a more weakly forced case highlighted in X08. The 4 km realtime forecasts correspond to the control members of the 4 km storm-scale ensemble forecasts (SSEF, X08, [12]), with and without radar data (named CN4 and C04, resp.). In 2008, the CAPS forecasts used version 2.2 of the Advanced Research Weather Research and Forecast (WRF-ARW) [13] model while in 2009 version 3.0 of WRF-ARW was used. For this reason, the 4 km and 1 km forecasts presented in this paper used versions 2.2 and 3.0 of WRF, respectively, but with the same set of physics parameterization options that correspond to the control member of the CAPS SSEFs of the two years [12, 14]. (The physics options used by the control forecasts of the two years were the same. Furthermore, version 3.0 differs from 2.2 mainly in the addition of new physics parameterization schemes while the dynamic core remains the same. For the configurations used, version 3.0 produced essentially the same forecasts results as version 2.2 for the 4 km forecasts based on later tests.) They are, specifically, the RRTM short-wave and NASA GSFC long-wave radiation; the NOAA land surface model, the Thompson microphysics, and the Mellor-Yamada-Jancic (MYJ) PBL schemes were used (see X08 for references), together with monotonic advection for water variables [15]. Cumulus parameterization scheme was not used, since 4 km and 1 km grid spacings are generally

considered convection permitting and convection resolving, respectively, while cumulus parameterization schemes are usually designed for grid spacings larger than 10 km [16].

All forecasts were initialized at 0000 UTC of 26 May 2008 for the test case. Forecasts C04 and C01 are, respectively, 4 and 1 km forecasts without radar data assimilation and were initialized by interpolation from the operational National Centers for Environmental Prediction (NCEP) North America Mesoscale (NAM) model 0000 UTC analysis on a 12 km grid. The 4 and 1 km forecasts with radar data assimilation, that is, CN4 and CN1, started from the analyses produced on the native model grid by the Advanced Regional Prediction System (ARPS) [8] three-dimensional variational (3DVAR) system [17] and its complex cloud analysis package [9, 18], using the same NAM analysis as the background. Full-volume level-2 V_r data from 57 WSR-88D radars running in precipitation mode (63 additional radars ran in clear mode) were analyzed by the 3DVAR. The Z data entered the system through the ARPS complex cloud analysis package, which analyzes cloud and hydrometeor fields and then adjusts in-cloud temperature and moisture based on a 1D parcel model with entrainment in areas of diagnosed cloud and rising motion [18]. The radar data were first automatically quality controlled, including velocity dealiasing, then “remapped” to the model grid through a least-squares fitting procedure [19] before being analyzed. Hence, the data were essentially super-obbed to the model grid first. Additionally, wind profiler and standard surface observations including the Oklahoma (OK) Mesonet data were also analyzed. The lateral boundary conditions came from the NAM forecasts. Both grids had 50 vertical layers with a near-surface vertical resolution of 20 m.

3. Forecast Results and Subjective Evaluation

3.1. The 26 May 2008 Case. At 0000 UTC, 26 May 2008 (not shown), a low was centered over Minnesota (MN), and a weak, quasistationary cold front extended from the low center southwestward to the western Kansas (KS) border, where it intersected a dryline that extended southward along eastern New Mexico (NM) border into northern Mexico (the point where a dryline intercept a front is often referred to as the front-dryline triple point, e.g., [20]). Fully developed quasi-linear convection existed through central KS about 100 km ahead of the cold front. Another SSW-NNE-oriented quasi-linear convective line existed in the Texas (TX) panhandle area, about 150 km east of the dryline at the TX-NM border. Over the next three hours, these lines evolved into a long connected line that was further linked with the convection in the Great Lakes (GL) region (Figure 1(a)). This squall line propagated eastward and maintained its identity until 0000 UTC, May 27 (not shown), when it was found over eastern Mississippi (MS), northern Alabama (AL), and eastern Tennessee (TN). During the entire period, the cold front was nearly stationary; the squall line was therefore mostly self-propagating, driven by the progression of its own cold pool. The initial convection-initiating forcing along the front and dryline was lost during this stage. This line quickly dissipated after 0000 UTC, May 27.

During this 24 hour period, there were other regions of convection that interacted with each other. As documented by X08, the evolution of convection during this period was rather complex and the morphology of many of the convective storms was modulated by their own cold pools and gust fronts and interactions with those of other storms. Such a situation is more difficult to predict than cases where strong propagating synoptic-scale features, such as a strong cold front, play more controlling roles. We demonstrate here that in the absence of strong large-scale control, the impact of radar data can be long-lasting.

3.2. Prediction Results. At the initial time (not shown), the composite (vertical column maximum) Z fields in CN4 and CN1 look very similar to the observed, which is due to the direct assimilation of Z data. C04 and C01, however, had no reflectivity in the initial condition (not shown). In addition to the quasi-linear convection ahead of the dryline and cold front, there was a large bow-shaped echo extending from central Missouri (MO) to central Arkansas (AR) at this time. There was also a line of cells in far southwestern TX, also east of the dryline.

Being properly initialized in CN1 and CN4, these groups of convection were accurately predicted over the first three hours (Figures 1(b) and 1(c)). The characteristics and pattern of convection predicted by CN1 (Figure 1(b)) in the TX panhandle, northwest OK, and KS regions at 0300 UTC compare very well with those of observation (Figure 1(a)). The associated narrow-line structures in CN1 agree particularly well with the observations. The forecast did miss the development of a new line segment in eastern Iowa (IA) at this time, which developed in the model later, at 0500 UTC. The model also predicted the bow-echo in the MO-AR region well, moving it from the initial central-MO-central-AR location to the Mississippi River at 0300 UTC (Figure 1(b)). The broad pattern of CN4-predicted convection is similar, but many fine-scale details are missing. The line segments in the TX, OK, and KS regions are not as well organized. This shows the noticeable advantage of the 1 km grid in resolving storm-scale structures.

The 4 km forecast without assimilating radar or additional surface Mesonet data (C04) is clearly inferior at 3 hours (Figure 1(d)). Essentially all of the line segments in TX, OK, and KS are missing. Instead, the model was trying to initiate new convection along the dryline at the TX-NM and KS-Colorado (CO) borders and along the cold front now located at the KS-Nebraska (NE) border and intersecting the dryline at the northwest corner of KS. In C04, the bow in MO-AR region is mostly missing, and the convection in the GL region is too weak. In this case, the convection that developed in the first few hours of forecast near the cold front and dryline was at wrong locations; as we will see later, this has long-term consequences.

At 9 hours, a time when the direct impact of radar data measured by standard skill scores for the season average starts to diminish (X08), the positive impact of radar data is still very clear in this case in both CN1 and CN4 (Figure 2). Figure 2(b) shows that CN1 predicted the strong, narrow

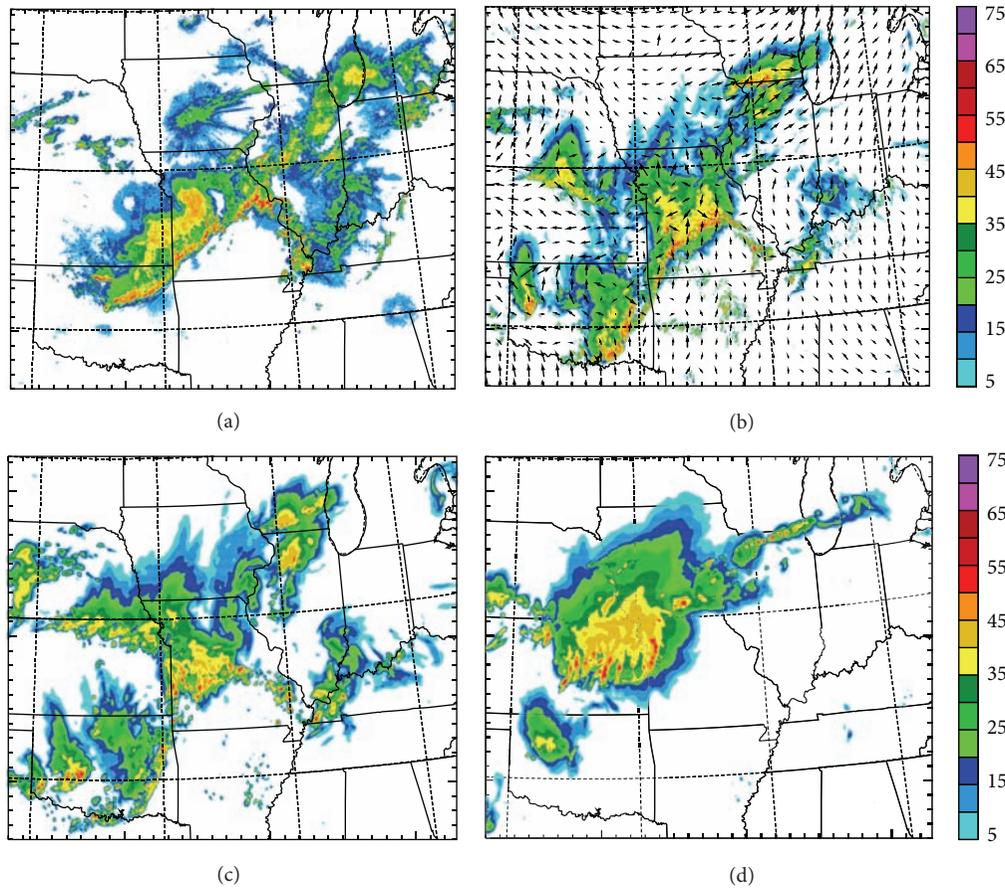


FIGURE 2: As Figure 1 but valid at 0900 UTC, 26 May 2008, corresponding to 9-hour forecast time and for a selected Midwest subdomain.

squall line extending from central OK through eastern-central MO very well, including the structure of embedded intense convection. Its southern end advanced too fast though, placing it about 150 km ahead of the one observed in southeast Oklahoma. One possible reason for the too fast advancement of the line is the cold pool which may be too strong. Cold pool intensity has been found to be rather sensitive to the microphysics, especially the drop/particle size distributions of rain and graupel, which through evaporative and melting affects cold pool intensity [21, 22].

Along the Mississippi River is another narrow line of cells that was observed and also predicted accurately in CN1. An examination of radar data and satellite imagery indicates that these cells developed along the back edge of the cold pool left behind by the northeastward propagating bow-shaped convection, which is at this time barely identifiable in northwestern Kentucky (KY, Figure 2(a)). This line intersected with the main squall line northwest of St. Louis, MO, forming a Γ -shaped echo. The CN1 forecast of this pattern matches the observation very well. In addition, there is indication that the 1 km forecast is producing stratiform precipitation trailing the leading convective line at the northern portion (Figure 2(b)), while the observation shows a clear secondary precipitation maximum behind the convective line somewhat near the southern end of the line. Such secondary precipitation maxima are prominent features within mature squall lines (e.g.,

[23]) but are notoriously difficult to predict in numerical models, and lack of model resolution and deficiency in the microphysics had been suspected to be the cause [24–26]. The fact that the 1 km forecast shows a somewhat better ability in producing the trailing stratiform precipitation is encouraging. The evolution of convection in other parts of the domain not shown, including those in southwest TX, the northern US Rockies, and near the GL, generally agrees with observations also.

The general pattern of predicted convection in CN4 (Figure 2(c)) is similar to that in CN1 (Figure 2(b)), although significant differences exist in detail. CN4 also captured the general Γ -shaped echo, but the embedded cells are clearly weaker. The southern portion of the main line also propagated too fast. In general, the 1 km forecast is noticeably superior to the 4 km forecast; it provides a much clearer indication of the intensity of the strongest embedded convective cells.

The forecast of C04 at this time is much poorer (Figure 2(d)). This forecast never managed to “spin up” the prefront and predryline convection. It simply evolved the convection that was incorrectly initiated along the front and dryline during the first few hours of the forecast, missing the most significant areas of convection. As discussed in X08, this failure continued to affect the subsequent evolution of a complex sequence of convective activities, for the remainder of the forecast.

By noon of 26 May (1800 UTC), all of the convective systems from the previous evening and night have moved out of the central Plains. The quasistationary front remained running across central KS, intersecting the dryline that extended north from the TX panhandle near the CO border (not shown). In the afternoon, convection was initiated along the dryline and, to a lesser extent, along the front. These processes were captured well in both CN1 and CN4 (Figure 3).

In the late afternoon hours, many hail events associated with the above convective storms were reported. Two brief tornadoes were reported near Dodge City, KS, between 2300 UTC, 26 May, and 0000 UTC, 27 May, emerging from storms that developed near the dryline-cold front triple point. At 2300 UTC, the observed composite reflectivity map of the OK-KS region shows three groups of convective cells (labeled A, B, and C in Figure 3(a)), one near the western OK border (A), one in southwestern KS near Dodge City (B), and one in the form of more isolated cells at the central OK-KS border (C). Groups A and B were initiated along the dryline and B near the front-dryline triple point (the east-west frontal location can be inferred from the surface wind field in, e.g., Figure 3(b), while the north-south dryline is located near the east edge of the plotting domain), and they were captured in both CN4 and CN1 (Figures 3(b) and 3(c)) but not in C04 (Figure 3(d)). In C04, the convection that was incorrectly initiated along the front over 20 hours earlier was organized into an east-west oriented line and moved to northern OK by this time (Figure 3(d)); it dissipated over the next couple of hours. This line obviously interfered with the conditions producing the actual dryline convective initiation in the afternoon of the second day. In fact, in C04 no initiation occurred at all along the dryline, except for an isolated cell near the triple point (Figure 3(d)).

Group C, consisting of more isolated cells, formed in the warm sector south of the front and east of the dryline near KS-OK border (Figure 3(a)). It is interesting that the main cell with this group is successfully predicted in CN1 (Figure 3(b)), but not in CN4, C04, nor in any other member of the 4 km ensemble produced in real time (X08). The observed cell became fully developed at 1900 UTC, while in CN1 it reached maturity at 2100 UTC. The observed storm propagated slowly south-southeastward and maintained its identity until 0300 UTC, 27 May. It generated many hail reports and a high-wind report of over 40 m s^{-1} at 2340 UTC.

The corresponding storm in the CN1 prediction maintained its full intensity until after 0100 UTC. It gained some supercell characteristics in terms of the shape of the reflectivity by 2300 UTC (Figure 3(b)), consistent with severe weather reports. Despite some difference in the exact timing and longevity between the observed and prediction storms, the ability of a 1 km model to predict, about 20 hours into the forecast, an isolated severe storm that developed in the absence of obvious mesoscale forcing is very remarkable. None of the ten 4 km ensemble forecasts that included initial and boundary condition perturbations as well as variations in physics schemes captured this storm. In fact the 4 km member without radar data assimilation completely missed

the initiation along the dryline on the second day. Finally, the 1 km forecast without radar data assimilation, C01, is similarly poor as C04, and this can be seen from the precipitation forecast scores presented in the next section.

4. Precipitation Verifications

To complement the earlier subjective evaluation of the forecasts for May 26, 2008 test case, we calculate the equitable threat scores (ETSs) verified against hourly radar-estimated precipitation produced on a 1 km grid by the National Severe Storms Laboratory in real time [27]. Such data were first interpolated to the forecast model grid before the ETS scores are calculated. Figure 4 shows the ETSs for hourly accumulated precipitation, at the 0.1 and 0.5 inch per hour thresholds, for the entire model domain. Clear evident is that the radar-assimilating CN1 and CN4 start with much higher ETSs initially, while the scores of C01 and C04 are around zero before 12 hours. For the 0.1 inch per hour threshold (Figure 4(a)), the ETS score for the first hour is about 0.45 for CN1 and 0.3 for CN4, indicating large difference in the short-range precipitation forecasts of 1 and 4 km grids. For the higher 0.5 inch per hour threshold (Figure 4(b)), the scores for the first hour are 0.29 versus 0.14, respectively. In general, the ETS scores decrease quickly during the first 5 hours, and the decrease is the fastest during the first two hours, especially for the higher thresholds. Such behaviors are actually expected and are consistent with the shorter range of predictability for more intense, smaller-scale convection, since errors associated with smaller scale, unstable motion grow the fastest (e.g., [28]). As errors associated with very short spatial scales present in the radar-assimilated initial condition grow quickly, predictability associated with such scales is quickly lost, causing initially rapid decrease of the precipitation forecast skill scores. Another possible cause for the initially rapid decrease in the skill score is insufficient dynamic and thermodynamic consistency among the model state variables within clouds when initialized by the single-time 3DVAR/cloud analysis. More advanced, four-dimensional, data assimilation methods that are closely coupled with the prediction model are expected to slow down the initial error growth to some degree. The forecast model error is another source although such an error tends to have larger impacts on longer forecasts.

The scores of C04 and C01 remain very low throughout the 30-hour-long forecasts and never exceed 0.03 (0.02 for the higher threshold). Between 2 and 19 hours, the scores of CN1 are up to 0.05 higher than those of CN4 for the lower threshold (Figure 4(a)). After 19 hours, the scores are comparable. For the higher threshold (Figure 4(b)), the differences between CN1 and CN4 become small after three hours. For grid point-based skill scores such as the ETS, position errors in small scale features can significantly impact the skill scores. In general, beyond the life cycle of the initial convective storms present in the initial condition, it is difficult for an NWP model to predict accurately the timing and location of new storm cells, especially when they are not forced by fixed features such as local terrain. Therefore skill

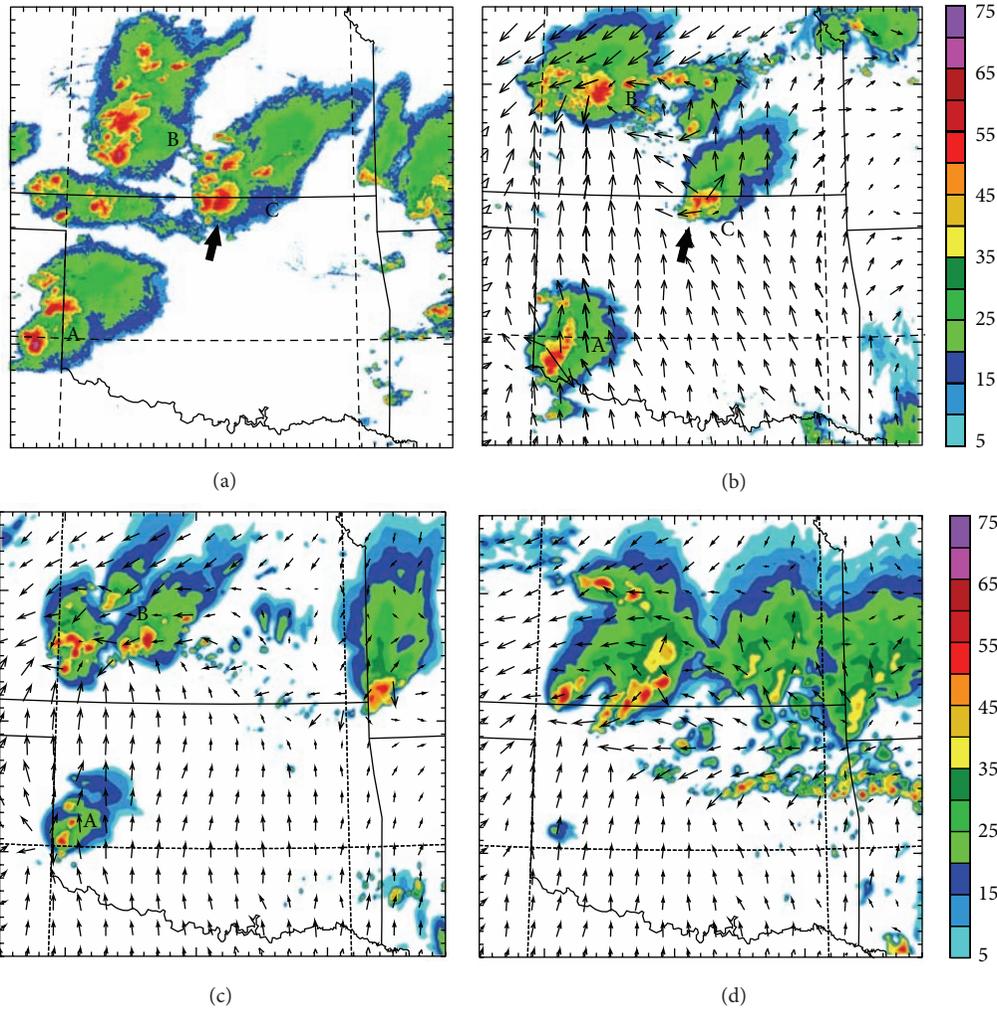


FIGURE 3: As Figure 1 but valid at 2300 UTC, 26 May 2008, corresponding to 23-hour forecast time and for a zoomed-in subdomain. The bold arrow near “C” points to the supercell discussed in the text.

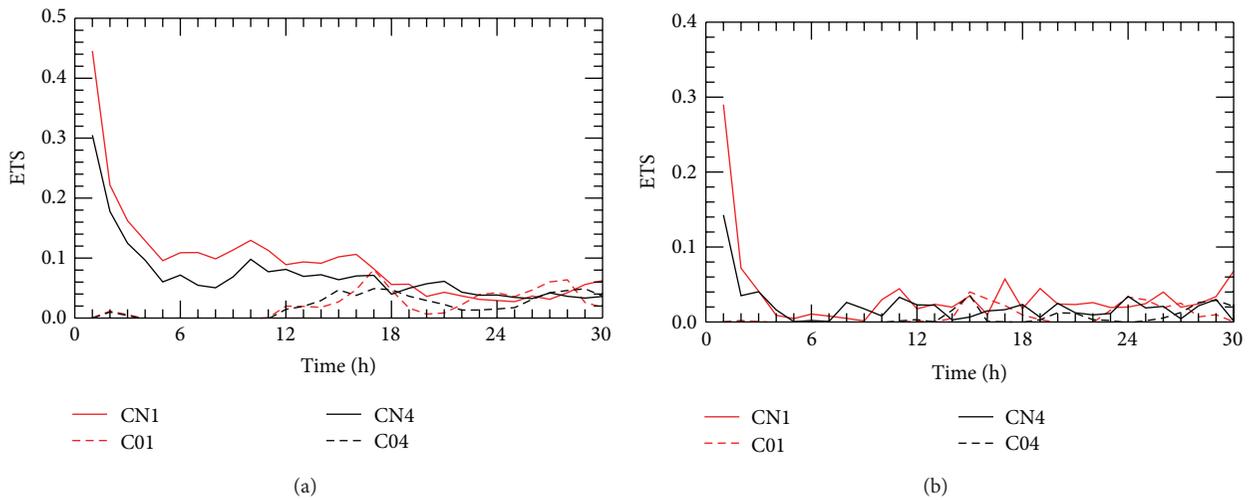


FIGURE 4: Equitable threat scores (ETSs) of hourly accumulated precipitation at 0.1 inch (a) and 0.5 inch (b) thresholds, for the 1 km forecast with radar data (CN1, solid red), 1 km forecast without radar data (C01, dashed red), 4 km control (CN4, solid black), and 4 km run without radar data (C04, dashed black), for the 26 May 2008 case, for hour 1 through 30.

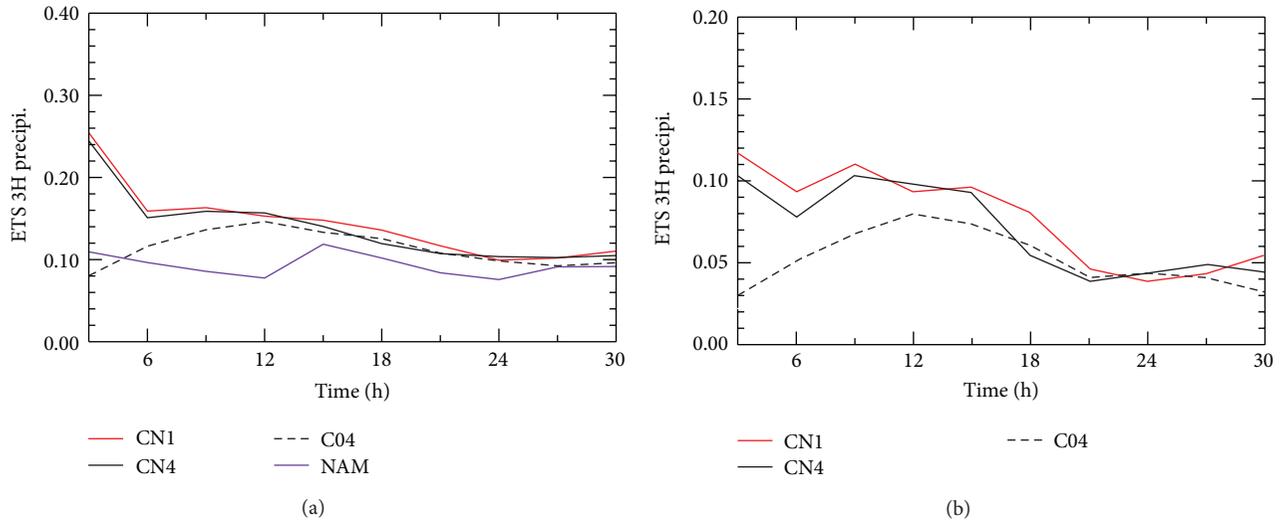


FIGURE 5: Mean equitable threat scores (ETs) of 3-hour accumulated precipitation at 0.1 inch (a) and 0.5 inch (b) thresholds, for the 1 km forecast with radar data (CN1, solid red), 4 km control (CN4, solid black), and 4 km run without radar data (C04, dashed black), for forecasts of 23 days from the CAPS 2009 spring forecast experiment.

scores that would allow for a certain degree of position error are often more useful (e.g., [5]).

To examine the precipitation forecast skill scores for the 4 and 1 km grids and the impact of radar data on the 4 km grid beyond the single test case present above, we discuss briefly here ETS scores for forecasts from 23 days of the 2009 CAPS spring forecast experiment on which all three forecasts are available; they are presented in Figure 5 for three-hour accumulated precipitation and for the 0.1 and 0.5 inch thresholds. For the ETS calculations, the 1 km precipitation fields were averaged to the 4 km grid.

Figure 5(a) shows that for the lower threshold, the mean ETS scores for CN1 are slightly higher than those of CN4 before 21 hours except for hour 12 when the score of CN1 dips slightly below that of CN4. For later hours, the scores are similar. The same comparison holds for the higher threshold (Figure 5(b)) although the relative difference is larger. This suggests that more intensive convection typically associated with smaller, more localized storms benefits more from the increased spatial resolution on average. For the May 26, 2008 test case, the difference between CN1 and CN4 is larger for the lower threshold, but it should be pointed out that the threshold for Figure 4(a) is 0.1 inch per hour rather than the 0.1 inch per three hours; therefore it actually corresponds to a higher precipitation intensity. In general, the ETS scores for all forecasts of spring 2009 are consistent with those of May 26, 2008 test case.

The ETS scores for the operational 12 km NAM forecasts are consistently lower than all high resolution forecasts for the lower threshold shown (Figure 5(a)), except for the first three hours when compared to the no-radar 4 km run (C04). For the initial hours, the NAM might have benefited from the consistency of its own analysis with its prediction model. Still, with the assimilation of radar data on either 4 or 1 km grid, the precipitation scores are much higher even during the initial hours (Figure 5(a)).

5. Summary

In this paper, we report on the results of the first ever test forecasts performed for a case from May 2008, at 1 km grid spacing in a domain covering almost the entire continental U.S., and the comparison of such forecasts with similarly configured forecasts produced at 4 km grid spacing in real time. These forecasts were 30 hours long, and a pair of forecasts assimilated both radial velocity and reflectivity data from all operational U.S. WSR-88D radars within the model domain, while another pair did not assimilate radar data. These 1 and 4 km forecasts with and without radar data assimilation are compared. Based on subjective evaluations, significant positive impact of radar data assimilation is found to last at least 24 hours for the test case. The 1 km forecast with radar data assimilation more accurately reproduced the observed convection than the corresponding 4 km forecast, especially in structure and intensity. It successfully predicted an isolated severe storm nearly 24 hours into the forecast, while the corresponding 4 km forecast, as well as all other 4 km members from the CAPS realtime storm-scale ensemble forecasts, failed to do so. The positive impact of radar assimilation on the precipitation forecast is even larger on both 4 and 1 km grids. Similar conclusions hold for precipitation forecasts based on mean equitable threat scores for 23 forecast days from spring 2009. This study provides evidence of the value of both convection-resolving resolution and radar data assimilation for severe weather prediction for up to 24 hours. We do want to point out that the equitable threat score examined in this paper has many limitations when applied to high-resolution precipitation forecasts due to large penalty associated with position errors. Object-based verification methods (e.g., [29]) and methods that account for position errors (e.g., [5]) will be explored in the future. In fact, an initial effort has been made to compare the number and size characteristics of storm cells predicted on the 4 and 1 km grids during the CAPS realtime forecasts [30].

Acknowledgments

This research was supported by a NOAA Collaborative Science and Technology Applied Research (CSTAR) Grant NA17RJ1227 and by the National Science Foundation Grants AGS-0738370, AGS-0802888, and EEC-0313747. The forecasts were produced at the National Institute of Computational Sciences, University of Tennessee, as part of the national Teragrid (currently Xsede) supercomputing allocation.

References

- [1] D. K. Lilly, "Numerical prediction of thunderstorms: has its time come?" *Quarterly Journal of the Royal Meteorological Society*, vol. 116, no. 494, pp. 779–798, 1990.
- [2] J. S. Kain, S. J. Weiss, D. R. Bright et al., "Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP," *Weather and Forecasting*, vol. 23, no. 5, pp. 931–952, 2008.
- [3] M. Xue, K. Brewster, K. K. Droegemeier et al., "The 1996 CAPS spring operational forecasting period: realtime storm-scale NWP, part II: operational summary and examples," in *Proceedings of the 11th Conference on Numerical Weather Prediction*, pp. 297–300, Norfolk, Va, USA, 1996.
- [4] M. L. Weisman, C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, "Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model," *Weather and Forecasting*, vol. 23, no. 3, pp. 407–437, 2008.
- [5] N. M. Roberts and H. W. Lean, "Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events," *Monthly Weather Review*, vol. 136, no. 1, pp. 78–97, 2008.
- [6] M. Xue, F. Kong, D. Weber et al., "CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA hazardous weather testbed 2007 spring experiment," in *Proceedings of the 22nd Conference on Weather Analysis and Forecasting/18th Conference Numerical Weather Prediction*, Salt Lake City, Utah, USA, 2007.
- [7] M. Xue, F. Kong, K. W. Thomas et al., "CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA hazardous weather testbed 2008 spring experiment," in *Proceedings of the 24th Conference on Severe Local Storms*, Savannah, Ga, USA, 2008.
- [8] M. Xue, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, "The advanced regional prediction system (ARPS), storm-scale numerical weather prediction and data assimilation," *Meteorology and Atmospheric Physics*, vol. 82, no. 1–4, pp. 139–170, 2003.
- [9] M. Hu, M. Xue, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: cloud analysis and its impact," *Monthly Weather Review*, vol. 134, no. 2, pp. 675–698, 2006.
- [10] C. S. Schwartz, J. S. Kain, S. J. Weiss et al., "Next-day convection-allowing WRF model guidance: a second look at 2-km versus 4-km grid spacing," *Monthly Weather Review*, vol. 137, no. 10, pp. 3351–3372, 2009.
- [11] M. Xue, F. Kong, K. W. Thomas et al., "CAPS realtime multi-model convection-allowing ensemble and 1-km convection-resolving forecasts for the NOAA hazardous weather testbed 2009 spring experiment," in *Proceedings of the 23rd Conference on Weather Analysis and Forecasting/19th Conference on Numerical Weather Prediction*, Omaha, Neb, USA, 2009.
- [12] F. Kong, M. Xue, M. Xue et al., "Real-time storm-scale ensemble forecast experiment-analysis of 2008 spring experiment data," in *Proceedings of the 24th Conference on Severe Local Storms*, Savannah, Ga, USA, 2008.
- [13] W. C. Skamarock, J. B. Klemp, J. Dudhia et al., "A description of the advanced research WRF version 2," NCAR2005.
- [14] F. Kong, M. Xue, K. Thomas et al., "A real-time storm-scale ensemble forecast system: 2009 spring experiment," in *Proceedings of the 23rd Conference on Weather Analysis and Forecasting/19th Conference Numerical Weather Prediction*, Omaha, Neb, USA, 2009.
- [15] W. C. Skamarock and M. L. Weisman, "The impact of positive-definite moisture transport on NWP precipitation forecasts," *Monthly Weather Review*, vol. 137, no. 1, pp. 488–494, 2009.
- [16] A. J. Clark, W. A. Gallus Jr., M. Xue, and F. Kong, "A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles," *Weather and Forecasting*, vol. 24, no. 4, pp. 1121–1140, 2009.
- [17] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [18] K. Brewster, "Recent advances in the diabatic initialization of a non-hydrostatic numerical model," in *Proceedings of the 15th Conference on Numerical Weather Prediction/21st Conference Severe Local Storms*, San Antonio, Tex, USA, 2002.
- [19] K. Brewster, M. Hu, M. Xue, and J. Gao, "Efficient assimilation of radar data at high resolution for short-range numerical weather prediction," in *Proceedings of the WWRP Symposium on Nowcasting and Very Short Range Forecasting (WSN '05)*, CDROM 3.06, Toulouse, France, 2005.
- [20] R. M. Wakimoto, H. V. Murphey, E. V. Browell, and S. Ismail, "The "triple point" on 24 May 2002 during IHOP. Part I: airborne Doppler and LASE analyses of the frontal boundaries and convection initiation," *Monthly Weather Review*, vol. 134, no. 1, pp. 231–250, 2006.
- [21] N. Snook and M. Xue, "Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms," *Geophysical Research Letters*, vol. 35, no. 24, 2008.
- [22] D. T. Dawson II, M. Xue, J. A. Milbrandt, and M. K. Yau, "Comparison of evaporation and cold pool development between single-moment and multimoment bulk microphysics schemes in idealized simulations of tornadic thunderstorms," *Monthly Weather Review*, vol. 138, no. 4, pp. 1152–1171, 2010.
- [23] R. H. Johnson and P. J. Hamilton, "The relationship of surface pressure features to the precipitation and airflow structure of an intense midlatitude squall line," *Monthly Weather Review*, vol. 116, no. 7, pp. 1444–1472, 1988.
- [24] G. H. Bryan, J. C. Wyngaard, and J. M. Fritsch, "Resolution requirements for the simulation of deep moist convection," *Monthly Weather Review*, vol. 131, no. 10, pp. 2394–2416, 2003.
- [25] G. H. Bryan and H. Morrison, "Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics," *Monthly Weather Review*, vol. 140, no. 1, pp. 202–225, 2012.
- [26] B. J. Putnam, M. Xue, Y. Jung, N. A. Snook, and G. Zhang, "The analysis and prediction of microphysical states and polarimetric variables in a mesoscale convective system using double-moment microphysics, multi-network radar data, and the ensemble Kalman filter," *Monthly Weather Review*, 2013.
- [27] J. Zhang, K. Howard, and J. J. Gourley, "Constructing three-dimensional multiple-radar reflectivity mosaics: examples of

- convective storms and stratiform rain echoes,” *Journal of Atmospheric and Oceanic Technology*, vol. 22, no. 1, pp. 30–42, 2005.
- [28] F. Zhang, N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, “Mesoscale predictability of moist baroclinic waves: convection-permitting experiments and multistage error growth dynamics,” *Journal of the Atmospheric Sciences*, vol. 64, no. 10, pp. 3579–3594, 2007.
- [29] C. A. Davis, B. Brown, and R. Bullock, “Object-based verification of precipitation forecasts. Part I: methodology and application to mesoscale rain areas,” *Monthly Weather Review*, vol. 134, no. 7, pp. 1772–1784, 2006.
- [30] A. Johnson, X. Wang, F. Kong, and M. Xue, “Object based evaluation of the impact of horizontal grid spacing on storm-scale forecasts,” *Monthly Weather Review*, vol. 141, no. 10, pp. 3413–3425, 2013.

Research Article

The Development of a Hybrid EnKF-3DVAR Algorithm for Storm-Scale Data Assimilation

Jidong Gao,^{1,2} Ming Xue,^{1,3} and David J. Stensrud²

¹ Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK 73072, USA

² NOAA/National Severe Storm Laboratory, National Weather Center 120 David L. Boren Boulevard, Norman, OK 73072, USA

³ School of Meteorology, University of Oklahoma, Norman, OK 73072, USA

Correspondence should be addressed to Jidong Gao; jidong.gao@noaa.gov

Received 17 May 2013; Revised 15 September 2013; Accepted 25 September 2013

Academic Editor: Kun Zhao

Copyright © 2013 Jidong Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A hybrid 3DVAR-EnKF data assimilation algorithm is developed based on 3DVAR and ensemble Kalman filter (EnKF) programs within the Advanced Regional Prediction System (ARPS). The hybrid algorithm uses the extended alpha control variable approach to combine the static and ensemble-derived flow-dependent forecast error covariances. The hybrid variational analysis is performed using an equal weighting of static and flow-dependent error covariance as derived from ensemble forecasts. The method is first applied to the assimilation of simulated radar data for a supercell storm. Results obtained using 3DVAR (with static covariance entirely), hybrid 3DVAR-EnKF, and the EnKF are compared. When data from a single radar are used, the EnKF method provides the best results for the model dynamic variables, while the hybrid method provides the best results for hydrometeor related variables in term of rms errors. Although storm structures can be established reasonably well using 3DVAR, the rms errors are generally worse than seen from the other two methods. With two radars, the results from 3DVAR are closer to those from EnKF. Our tests indicate that the hybrid scheme can reduce the storm spin-up time because it fits the observations, especially the reflectivity observations, better than the EnKF and the 3DVAR at the beginning of the assimilation cycles.

1. Introduction

The effective assimilation of radar data into a numerical weather prediction (NWP) model requires advanced data assimilation (DA) techniques, such as variational and ensemble Kalman filter methods. A three-dimensional variational (3DVAR) system, which includes a mass continuity equation and other appropriate model equations as weak constraints, has been developed in recent years [1–5]. This system was designed with special considerations for assimilating radar data into a convective-scale nonhydrostatic model—the Advanced Regional Prediction System (ARPS)—and has been used to provide initial conditions for numerous real-time convective-scale data forecasts. These forecasts have been produced since 2008 using grid spacing that varied from 4 to 1 km for domains covering the entire continental United States as part of the NOAA Hazardous Weather Testbed (HWT) Spring Experiments [6, 7]. For the HWT Spring Experiments, Level-II radial velocity and reflectivity data

from over 120 operational Weather Surveillance Radar-1988 Doppler (WSR-88D) radars were analyzed using the 3DVAR system, and ensemble forecasts were produced by adding additional initial condition perturbations to this 3DVAR analysis. The ARPS 3DVAR system has also been used in a large number of real case studies with encouraging results [2, 3, 8, 9]. Barker et al. [10] and Xiao et al. [11] also applied the 3DVAR method to assimilate Doppler radar observations into the Weather Research and Forecasting (WRF) model [12]. The major advantage of the 3DVAR method is its computational efficiency and the ease by which weak constraints can be included. However, the truly flow-dependent background error covariances were not included in either ARPS 3DVAR or WRF 3DVAR systems at that time.

Compared to 3DVAR, the more advanced 4DVAR technique incorporates the full prediction model into the assimilation system and implicitly includes the effects of flow-dependent error covariances through the use of both the

forward and backward models. In recent years, the 4DVAR technique has helped improve global forecasts at several operational NWP centers, including the European Centre for Medium-Range Weather Forecasts, Meteo-France, Meteorological Service of Canada, and Japan Meteorological Agency (JMA) [13]. Research has also focused on storm-scale radar data assimilation using the 4DVAR method by Sun and Crook [14–16]. In these studies, both radial velocity and reflectivity data were assimilated into a convective cloud-resolving model. Despite some encouraging results, 4DVAR for convective-scale applications has been limited to the use of simple microphysics in almost all cases because the strong nonlinearity within sophisticated microphysics schemes makes the minimization process difficult. Honda and Koizumi [17] report difficulties, including slow convergence, when including complex ice microphysics within the inner loop of the 4DVAR system when using a nonhydrostatic model at JMA.

The ensemble Kalman filter (EnKF) is an advanced data assimilation method that shares many of the advantages of 4DVAR. It has gained considerable popularity in recent years in meteorology and oceanography since first proposed by Evensen [18]. For convective storms, very encouraging results have been obtained in recent studies using the ensemble Kalman filter method in analyzing wind, temperature, moisture fields, and even microphysics variables from radar observations of convective storms [19–26]. One of the advantages of the EnKF method over variational methods is that it can explicitly evolve and carry the background error covariances through the assimilation cycles. However, one of the major sources of error with ensemble-based DA is covariance matrix rank deficiency or sampling error as a result of a relatively small ensemble size [27, 28]. This problem can be more severe with storm-scale data assimilation because the degrees of freedom of the system are typically even larger relative to the practical ensemble size. The commonly utilized remedy to the rank deficiency problem is to apply covariance localization by a Schur product as introduced by Houtekamer and Mitchell [27]. This solution, however, prevents the use of distant correlations that are physically meaningful. Further, the modification to the spatial covariances within a cut-off radius by a Schur product also introduces imbalances, and the effect is more substantial when the localization is more restrictive [28]. This problem may be remedied or reduced when using a hybrid 3DVAR and EnKF method.

As discussed above, the 3DVAR method is attractive for convective scale assimilation because of its computational efficiency and the ease by which weak constraints can be added. However, the major shortcoming is that the background error covariances are stationary and isotropic and error covariances related to the model equations cannot be simply defined. In addition, for convective-scale radar data assimilation, only observations of radial velocity and reflectivity are typically measured, while all other state variables have to be “retrieved”; in this case, the flow-dependent background error covariances, such as that derived from a forecast ensemble, are especially important. One way to blend the advanced features of both variational and EnKF methods and to overcome their respective shortcomings is to employ

a hybrid ensemble 3DVAR framework. In such a framework, a combination of the static background error covariance and the flow-dependent error covariance derived from an ensemble is used within the variational analysis. For large-scale data assimilation, such an approach was initially demonstrated for a quasigeostrophic system by Hamill and Snyder [29] and further developed by Lorenc [30], Buehner [31], and Zupanski [32] with different formulations. Another relatively new approach estimates the four-dimensional background-error covariances from the ensemble members to produce a 4D analysis with the variational data assimilation approach. In this method, the tangent-linear or adjoint versions of the forecast model are no longer needed. This approach was called the En-4DVar approach [33–35] but was recently renamed as 4DVar [36].

Wang et al. [37] showed that the formulations proposed by Hamill and Snyder [29], Lorenc [30], and Buehner [31], though different in implementation and computational cost, are mathematically equivalent. Barker et al. [38], Li et al. [39], and Zhang et al. [40] recently reported the capability of the WRF hybrid system for mesoscale applications. Further studies have demonstrated the potential advantages of the hybrid method over both the pure variational and pure ensemble methods for mesoscale and global applications, especially for small ensemble size [41–44]. However, the application of hybrid methods to convective scale data assimilation has so far been limited. The purpose of this paper is to demonstrate the potential usefulness of the hybrid EnKF-3DVAR method for convective scale data assimilation, especially when assimilating radar data.

The rest of this paper is organized as follows. In Section 2, we introduce the hybrid EnKF-3DVAR system developed in this study. Section 3 describes the DA experiment design. Experiment results and quantitative performance are assessed in Section 4. We conclude in Section 5 with a summary and outlook for future work.

2. The Hybrid EnKF-3DVAR Scheme

In the implementation of the hybrid method for convective scale, the ensemble covariance is incorporated in the variational framework using the extended control variable method [30, 31, 37]. A convenient approach, initially suggested by Buehner [31], is to combine the ensemble-derived and static covariance matrices through the augmentation of state vector, from \mathbf{v} to (\mathbf{v}, \mathbf{w}) within the 3DVAR cost function, which can be written as

$$J = \frac{1}{2} \mathbf{v}^T \mathbf{v} + \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} \left[H(\bar{\mathbf{x}}^b + \Delta \mathbf{x}) - \mathbf{y}^o \right]^T \times \mathbf{R}^{-1} \left[H(\bar{\mathbf{x}}^b + \Delta \mathbf{x}) - \mathbf{y}^o \right] + J_c, \quad (1)$$

where

$$\Delta \mathbf{x} = \Delta \mathbf{x}_1 + \Delta \mathbf{x}_2 = \beta_1 \mathbf{B}^{1/2} \mathbf{v} + \beta_2 \mathbf{P}^{1/2} \mathbf{w} \quad (2)$$

is the analysis increment of state vector \mathbf{x} , \mathbf{B} is the static 3DVAR background error covariance matrix, and \mathbf{P} is the covariance matrix derived from an ensemble of forecasts. The

control variable \mathbf{v} is defined in association with \mathbf{B} , and \mathbf{w} is the augmented control vector associated with \mathbf{P} . The size of \mathbf{v} is the number of analysis variables multiplied by their dimension, and the size of \mathbf{w} is the ensemble size multiplied by the dimension of variables. By using control variables \mathbf{v} and \mathbf{w} , instead of $\Delta\mathbf{x}_1$ and $\Delta\mathbf{x}_2$ in (2), the minimization procedure is preconditioned by $\mathbf{B}^{1/2}$ and $\mathbf{P}^{1/2}$, respectively. This technique was first proposed in the context of data assimilation by Derber and Rosati [45]. The definition of $(\mathbf{B})^{1/2}$ is the same as Gao et al. [1]. If no localization is applied to the ensemble covariance, $\mathbf{P}^{1/2}$ is simply a rectangular matrix whose columns are the ensemble perturbation vectors divided by $\sqrt{N-1}$, where N is the ensemble size. The localization of the ensemble covariance in a variational system with preconditioning is discussed in Lorenc [30], Buehner [31], and Wang et al. [37]. The procedure and cost of doing so were also discussed in these papers. For computational efficiency, we also use the recursive filter for covariance localization, as suggested in Wang et al. [41].

In (2), there are two factors β_1 and β_2 that define the weights placed on the static background error covariance and the ensemble covariance. To conserve the total background-error variance, β_1 and β_2 are constrained by

$$\beta_1^2 + \beta_2^2 = 1. \quad (3)$$

A similar constraint was applied in Hamill and Snyder [29]. This approach for combining two covariance matrices to form a hybrid covariance provides flexibility since it allows for different relative contributions from two covariance matrices. When $\beta_1 = 1$, the analysis is back to a 3DVAR analysis scheme, when $\beta_2 = 1$, the analysis is mathematically equivalent to an EnKF scheme, and in between, we have a hybrid scheme that incorporates a mixture of both static and flow-dependent error covariances. When $\beta_2 = 1$, the scheme is essentially a variational formulation of an ensemble-based analysis scheme, and it can be called 3DENVAR. Though the dimension of the control variables is increased, the form of the background term of the cost function remains unchanged from that of 3DVAR, so that codes from an existing 3DVAR system can readily be utilized [30].

In the current study, the hybrid system will assimilate both radar reflectivity and radial velocity data. Within this system, flow-dependent background-error covariances, in particular cross covariances between microphysical and dynamic variables, will be derived and utilized. The single-resolution version of the EnKF system of Gao and Xue [46] is used for updating the ensemble perturbations in the data assimilation cycles. In Gao and Xue [46], an efficient dual-resolution (DR) data assimilation algorithm was developed based on the ensemble square root Kalman filter method and tested using simulated radar radial velocity data for a supercell storm. Within the algorithm, radar observations were assimilated on both high-resolution and lower-resolution grids using ensemble Kalman filter algorithms and the flow-dependent background error covariance estimated from the lower resolution ensemble. In that paper, the DR method was compared to a standard full-resolution ensemble square root Kalman filter method which is used in this study.

Different from other hybrid systems [40, 41], for this hybrid method, an extra model integration for the length of the analysis cycle is needed to produce a control forecast and analysis cycle. The EnKF analyses are performed to update analysis perturbations for each ensemble member. Then, the cost function (1) is minimized to obtain optimal analyses of control vectors \mathbf{v} and \mathbf{w} , and the optimal analysis increment, $\Delta\mathbf{x}$, is derived from (2). The ensemble mean analysis is replaced with the hybrid EnKF-3DVAR analysis. Finally, the initial conditions for the ensemble and one control forecast are obtained. The above steps are repeated for each data assimilation cycle (Figure 1).

3. Model and Experimental Design

3.1. Prediction Model and Truth Simulation for OSSEs. We test our hybrid EnKF-3DVAR algorithm and compare its results with those of 3DVAR and EnKF schemes, using simulated data from a classic supercell storm of May 20, 1977, near Del City, Oklahoma [47]. The ARPS prediction model is used in a 3D cloud model mode, and the prognostic variables include three velocity components u , v , and w , perturbation potential temperature θ' , pressure p , and six categories of water substances, that is, water vapor specific humidity q_v , and mixing ratios for cloud water q_c , rainwater q_r , cloud ice q_i , snow q_s , and hail q_h . The microphysical processes are parameterized using the single-moment, three-category ice scheme of Ying Lin et al. [48]. More details on the model can be found in Xue et al. [49, 50].

For our experiments, the model domain is $57 \times 57 \times 16 \text{ km}^3$. The horizontal grid spacing is 1 km, and the mean vertical grid spacing is 500 m. The truth simulation run is initialized from a modified real sounding plus a 4 K ellipsoidal thermal bubble centered at $x = 48$, $y = 16$, and $z = 1.5 \text{ km}$, with radii of 10 km in x and y and 1.5 km in the z direction. Open conditions are used at the lateral boundaries. The length of simulation is 2 hours. A constant wind of $u = 3 \text{ ms}^{-1}$ and $v = 14 \text{ ms}^{-1}$ 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. [50]. During the truth simulation, the initial convective cell strengthens over the first 30 min. The strength of the cell then decreases over the next 30 min or so, which is associated with the splitting of the cell at around 55 min. The right moving (relative to the storm motion vector which is towards north-northeast) cell tends to dominate the system, and its updraft reaches a peak value of over 40 ms^{-1} at 90 min. The initial cloud starts to form at about 10 min, and rainwater forms at about 15 min. Ice phase fields appear at about 20 min. A similar truth simulation was also used in Gao et al. [51], Tong and Xue [21], and Gao and Xue [46].

3.2. Simulation of Radar Observations. The simulated radial velocity observations are assumed to be available on the grid points. The simulated radial velocity, v_r , is calculated from

$$v_r = u \sin \phi \cos \mu + v \cos \phi \cos \mu + w \sin \mu, \quad (4)$$

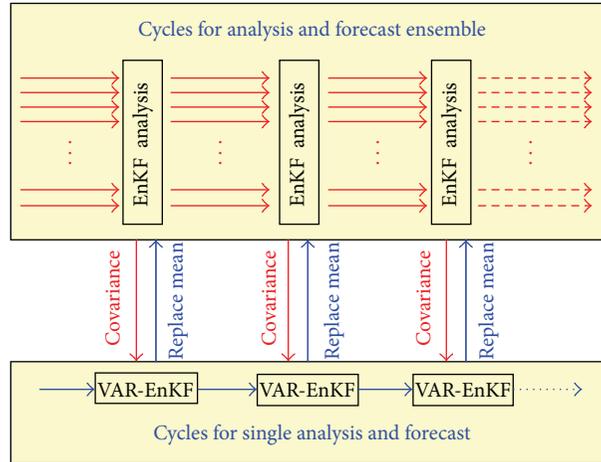


FIGURE 1: Illustration of cycle used in a hybrid EnKF-3DVAR analysis scheme.

where μ is the elevation angle ϕ is the azimuth angle of radar beams, and u , v , and w are the model-simulated velocities interpolated to the scalar points of the staggered model grid. Random errors drawn from a normal distribution with zero mean and a standard deviation of 1 ms^{-1} are added to the simulated data. Since v_r is sampled directly from the model velocity fields, hydrometeor sedimentation is not involved. The ground-based radar is located at the southwest corner of the computational domain, that is, at the origin of the x - y coordinates. The simulated reflectivity observations are calculated based on Smith et al. [52] and Ferrier [53]. For reflectivity, random errors drawn from a normal distribution with zero mean and a standard deviation of 3 dBZ are added to the simulated data. The radial velocity data are assimilated and are only available where the truth reflectivity is greater than zero in the analysis domain. We also use only the data at every other grid point from the 1 km truth simulation grid in horizontal, so that the total data used are one-fourth of total model grid points.

3.3. Design of Assimilation Experiments. We start the initial ensemble forecast at 20 min of the model integration time when the storm cell is well developed. To initialize the ensemble members, random noise is first added to the initially horizontally homogeneous first guess defined using the environmental sounding. A 2D five-point smoother is applied to the resultant fields, similar to a method used by Zupanski et al. [54]. The random noise is sampled from Gaussian distributions with zero mean and standard deviations of 5 ms^{-1} for u , v , and w and 3 K for potential temperature. These perturbation variances are somewhat larger than those used in Tong and Xue [21], but the standard deviation of the final perturbations is not necessarily larger because of the smoothing. Other variables, including the microphysical variables, are not perturbed at the initial time. The radial and reflectivity observations are calculated and assimilated using a 5 min cycle in all three data assimilation schemes. The first analysis is performed at 20 min, and 20 ensemble members are used. A cut-off radius of 8 km is used in most of our experiments.

We perform two set of experiments. The first group of experiments is performed to compare the performance of three different schemes when observations from a single Doppler radar are used. The second group of experiments will be performed when observations from two Doppler radars are used. For comparison purposes, all three methods (3DVAR, EnKF, and Hybrid EnKF-3DVAR) are performed with 16 data assimilation cycles where each cycle has a 5 min analysis-prediction interval. The total assimilation period is 75 min.

4. Results

4.1. Single Observation Experiment. Figure 2 provides analysis results of a single observation with three model variables, showing that ensemble information can provide flow-dependent estimates of the background-error covariance and that both the EnKF and hybrid 3DVAR-EnKF methods can utilize such information to provide flow-dependent analysis increments. Because mass continuity equation is used as a weak constraint in 3DVAR [1], the 3DVAR method can also provide a kind of flow-dependent anisotropic non-Gaussian type covariance structure for both u component and w component (Figures 2(a) and 2(b)). However, the 3DVAR cannot provide increments for potential temperature (Figure 2(d)), though updated potential temperature can be obtained through a cycled 3DVAR analysis (built up by integration of a convective NWP model, ARPS in this study). The EnKF provides a flow-dependent covariance structure (Figures 2(b), 2(e), and 2(h)), and the hybrid 3DVAR-EnKF provides a covariance structure in between the other two structures. In addition, both EnKF and hybrid 3DVAR-EnKF can provide increments for unobserved variables, such as potential temperature which is not directly related to radial velocity (Figures 2(h) and 2(i)). Because the mass continuity equation is used as a weak constraint in 3DVAR, this actually provides a physical constraint for three components of wind field. Similar to Buehner [31] and to take advantage of both 3DVAR and EnKF methods, 50/50 weightings are chosen in the cost function.

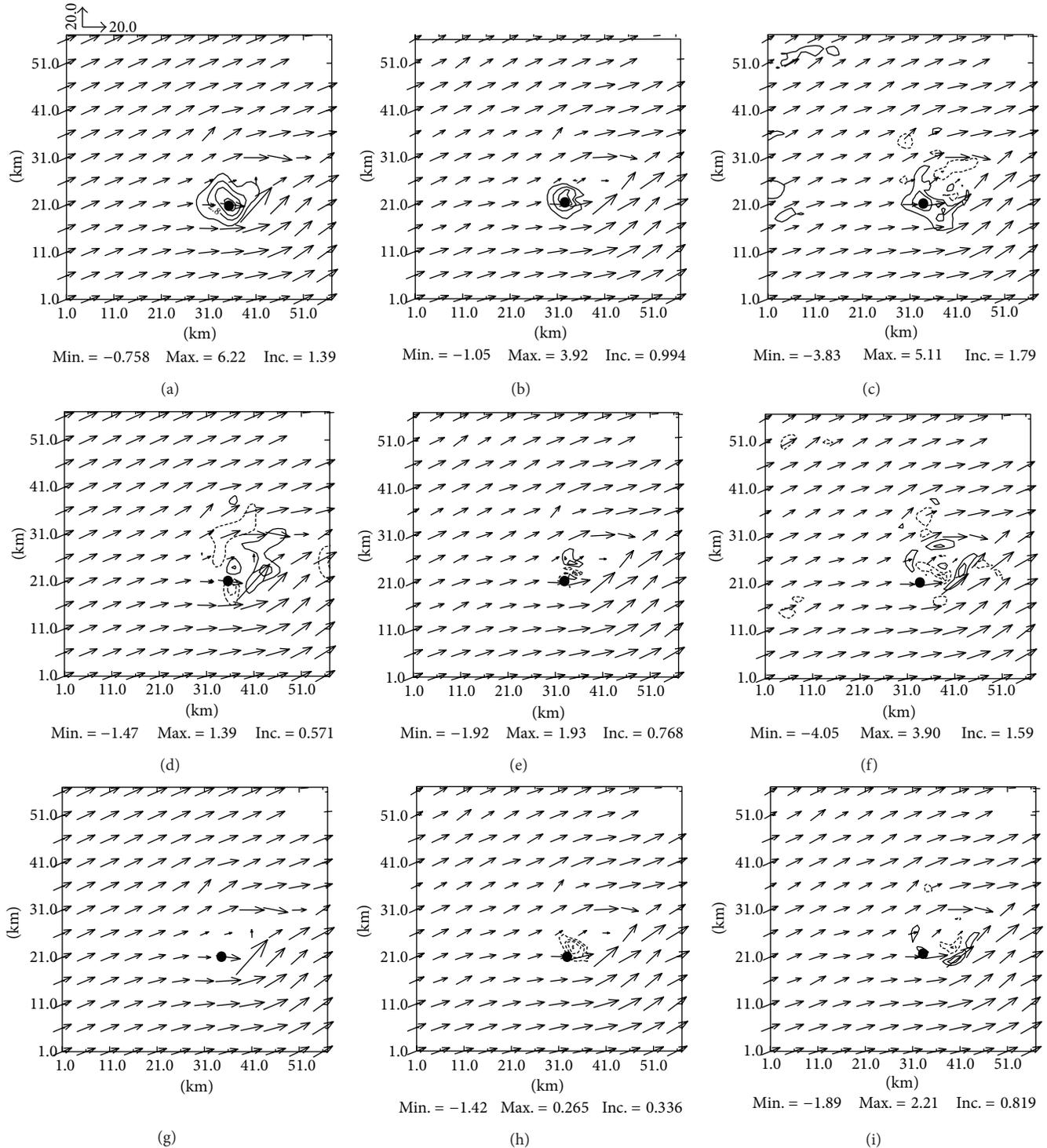


FIGURE 2: Wind vectors: u -component increment by using (a) 3DVAR, (b) EnKF, and (c) hybrid 3DVAR-EnKF, w -component increment by using (d) 3DVAR, (e) EnKF, and (f) hybrid 3DVAR-EnKF, and potential temperature increment by using (g) 3DVAR, (h) EnKF, and (i) hybrid 3DVAR-EnKF by assimilating a single radial velocity at the black dot.

4.2. Experiments with Single Radar. As stated above, the first group of experiments is performed with radial velocity and reflectivity data from a single radar. Figure 3 shows the final assimilation results after 16 assimilation cycles with 5 min prediction-analysis intervals. The low-level flow, reflectivity

patterns, and the strength of the cold pool from both EnKF and hybrid EnKF-3DVAR agree very well with the simulated truth (Figure 3(a)) and are better than the result using 3DVAR (Figure 3(b)), although this 3DVAR can also establish the storm structures reasonably well. The most obvious difference

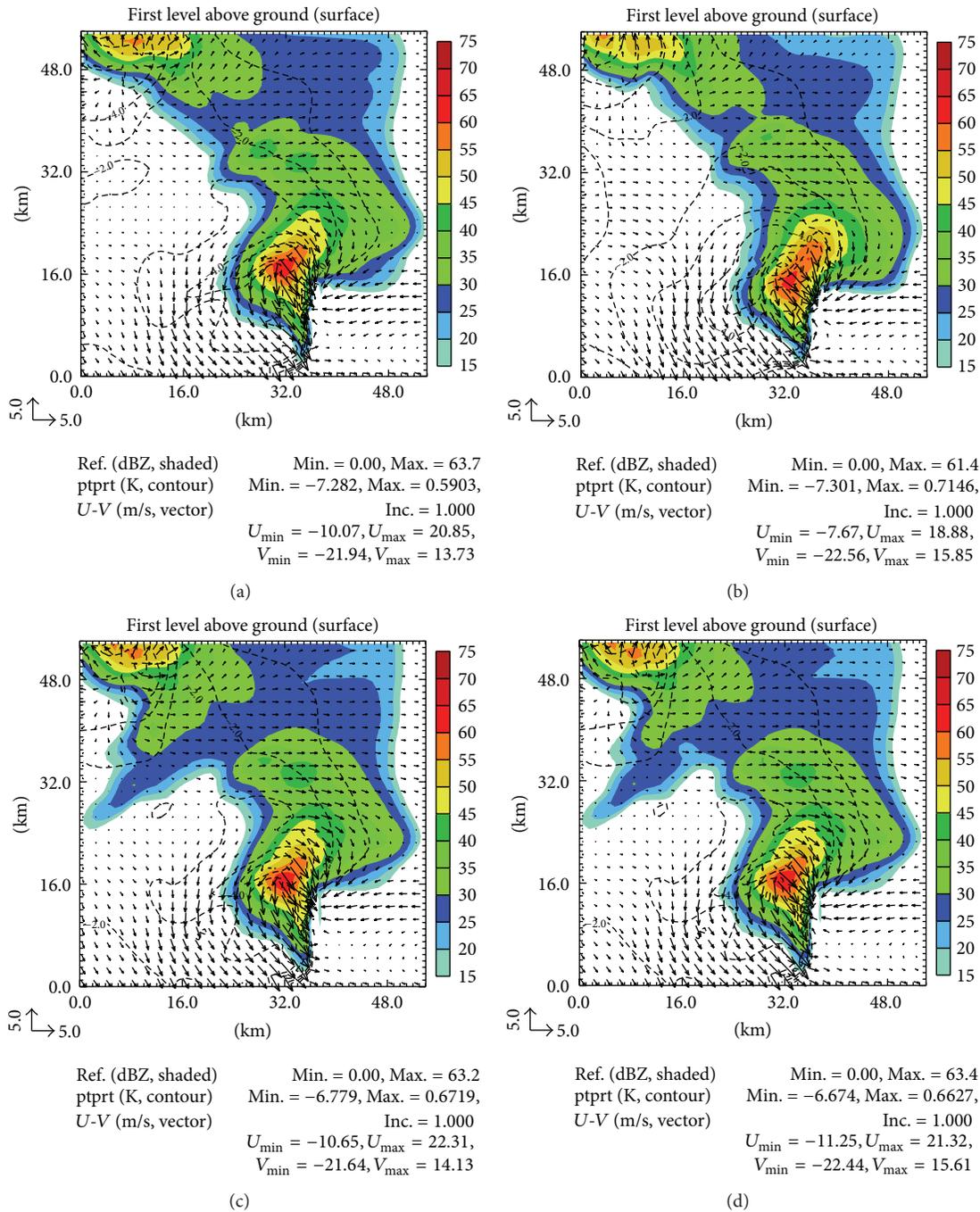


FIGURE 3: Horizontal winds (vectors; ms^{-1}), perturbation potential temperature (contours at 1-K intervals), and simulated reflectivity (shaded contours; dBZ) at 250 m AGL for (a) the truth simulation; (b) the 3DVAR analysis; (c) the EnKF analysis; and (d) the hybrid EnKF-3DVAR analysis for the single radar experiment. The time shown is at 100 min (the end of data assimilation cycles). Wind vectors are shown every 2 km.

is the reflectivity field in the center of model domain. The area of reflectivity values greater than 55 dBZ is over extended in a peanut-shaped region for 3DVAR. The spread of potential temperature is little bit far to the south-southwest direction in the southwest corner (Figure 3(b)). But the strength of the cold pool in 3DVAR, as indicated by minimum perturbation potential of -7.30° , is closer to the truth simulation (-7.28°) than seen in either EnKF or the hybrid EnKF-3DVAR.

The rms errors of the analyzed fields with data from a single radar are shown in Figure 4. The rms error calculation is limited to the regions where the truth reflectivity exceeds 10 dBZ. Figure 4 shows that the rms errors for model variables u , v , w , θ , and q_v and reflectivity Z (derived from the hydrometeor mixing ratios) generally decrease with the cycles in all three experiments. The errors for 3DVAR decrease more slowly and remain at a higher level at the end of assimilation

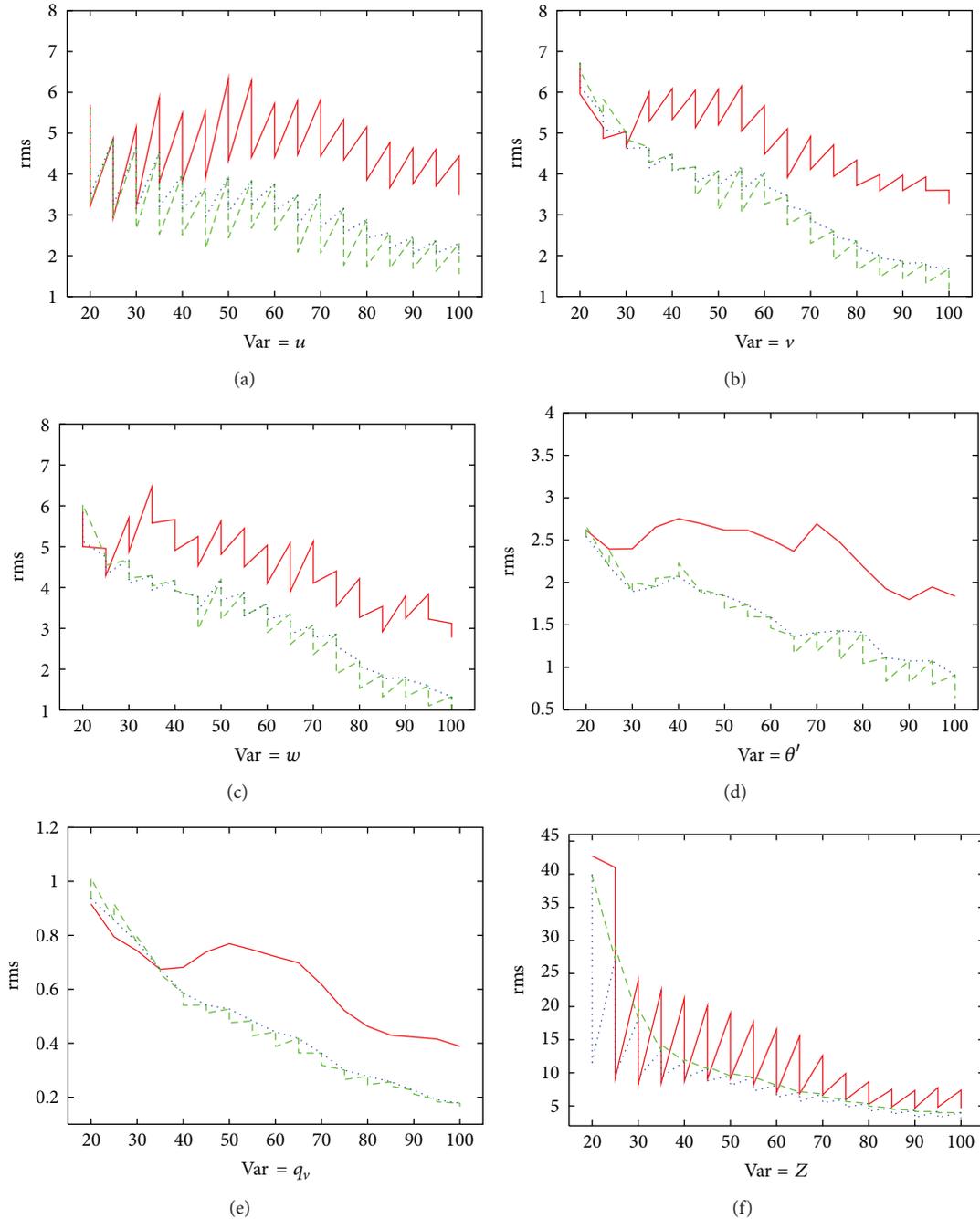


FIGURE 4: The rms errors of the analysis and forecast for the 3DVAR, (red) EnKF, (green) hybrid EnKF-3DVAR, (blue) methods averaged over points at which the reflectivity is greater than 10 dBZ for (a) u -wind component, (b) v -wind component, (c) vertical wind speed, (d) potential temperature, (e) water vapor mixing ratio, and (f) reflectivity.

cycles than those for the ensemble based methods for most of model variables. For example, the rms error of w is close to 3 ms^{-1} at 100 min for 3DVAR method, while that in EnKF and hybrid EnKF-3DVAR is close to 1.3 ms^{-1} . The rms errors of q_v for 3DVAR is 0.4 g/kg , and that in EnKF and hybrid EnKF-3DVAR is below 0.2 g/kg . While these differences are significant, the error levels late in the assimilation period for EnKF and hybrid EnKF-3DVAR are unrealistically low due to the perfect model assumption. For real data cases where model error exists, the analysis errors are likely to be much

larger (see, for example, Dowell et al. [22, 23]). For systems containing discrete intense updrafts, the rms error tends to exaggerate errors because of small spatial displacement and/or structure discrepancies, such as those seen in Figure 4. So the results for 3DVAR may still be reasonable. It should be noted that for most of model variables, the performance of EnKF and hybrid methods is very close to each other, with EnKF a little bit better. Interestingly, the differences among the rms errors for Z in different experiments are smallest (Figure 4(f)). The rms error of Z is decreased to about

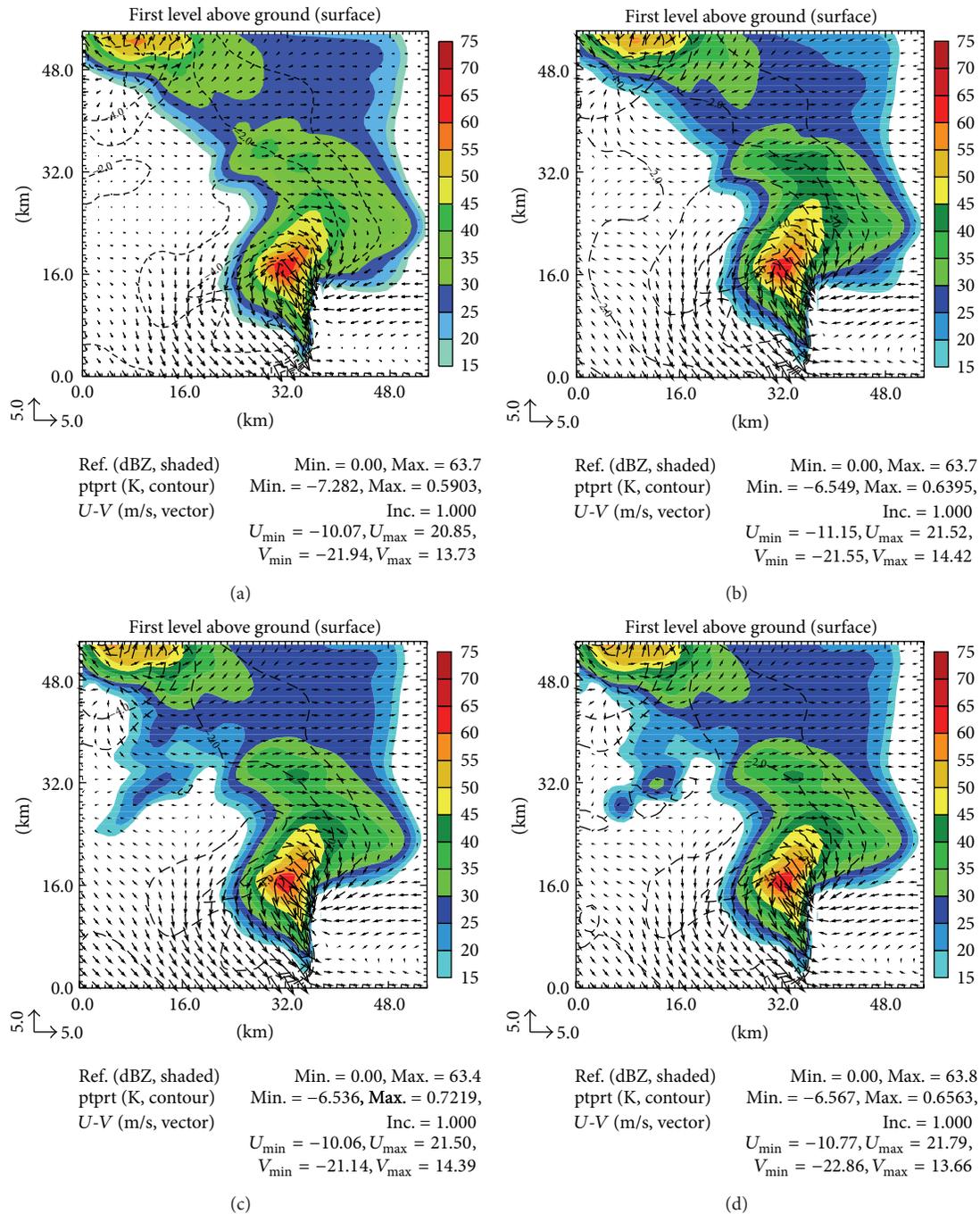


FIGURE 5: The same as Figure 3 but for the experiment with two radars.

5 dBZ in all three experiments. The variation of rms errors is volatile for 3DVAR, especially near the very beginning of the assimilation. The method can decrease the errors from about 40 dBZ to 10 dBZ in two data assimilation cycles, but the errors quickly increase to above 20 dBZ after the 5 min model integration step. The rms errors for the EnKF method decrease more smoothly throughout the data assimilation cycles because of its statistical nature. Perhaps the advantage of hybrid method is most obvious for reflectivity, as it fits the observed reflectivity field more closely than the other two

methods. Though the evolution of rms errors is also volatile for the first 10 minutes, it quickly settles down, and its rms errors are the lowest among all three methods.

4.3. Experiments with Two Radars. The second group of experiments is performed with radar data from two simulated Doppler radars. Figure 5 shows the final assimilation results after 16 assimilation cycles. As expected, the low-level flow, reflectivity patterns, and the strength of the cold pool look

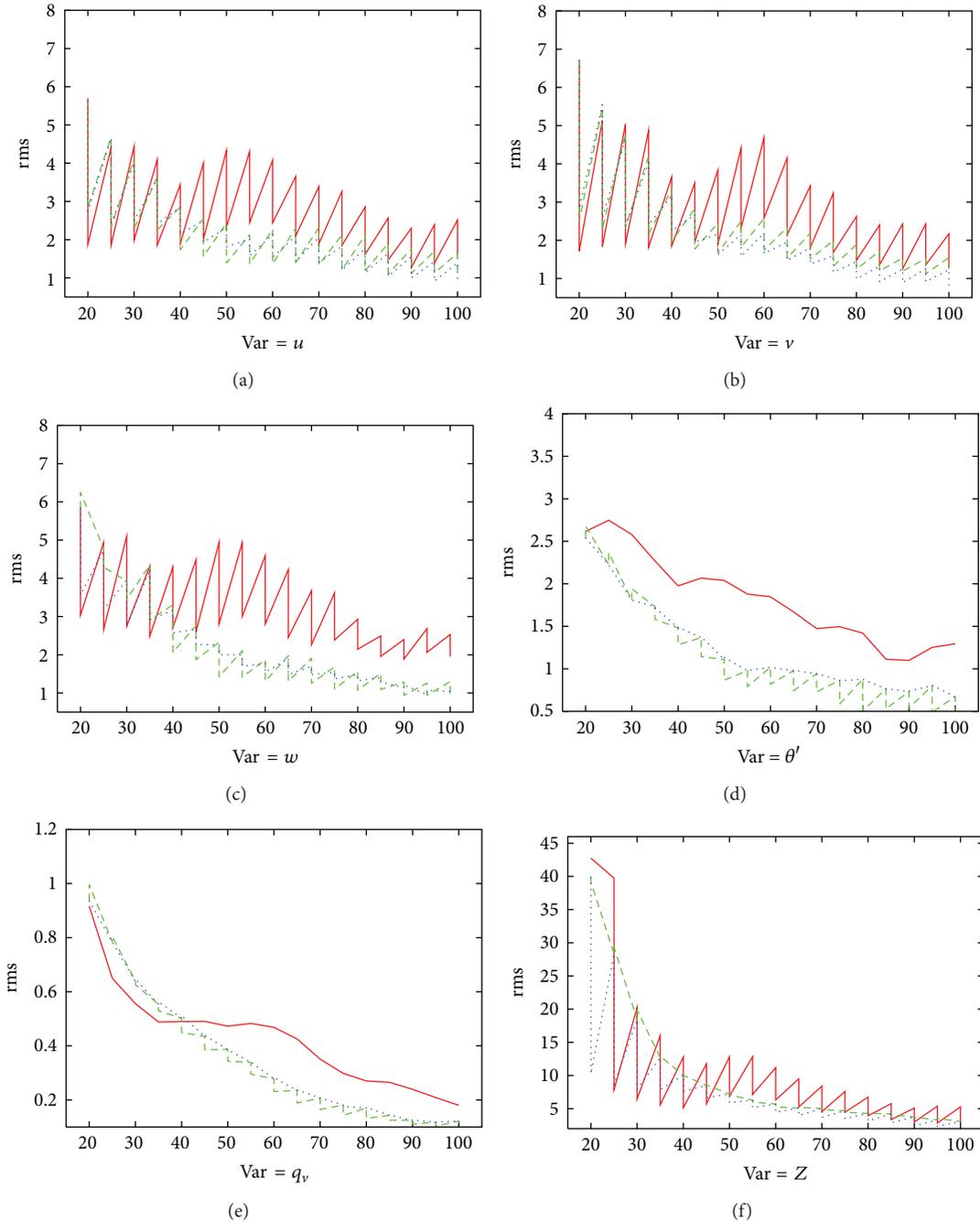


FIGURE 6: The same as Figure 4 but for the experiment with two radars.

much better (Figure 5(b)) for 3DVAR (Figure 5(b)), especially for the reflectivity field. The pattern for potential temperature is improved when compared with the single radar experiment (Figure 3(b)) but is still not as good as the truth simulation (Figure 5(a)) and that for EnKF (Figure 5(c)) and the hybrid EnKF-3DVAR (Figure 5(d)). So with more data used, the results for 3DVAR are improved. Again, the most obvious improvement is for the reflectivity field in the center of model domain. The area with reflectivity values larger than 55 dBZ is more similar to the shape of truth simulation. The storm structure for all three methods is well established by

the end of data assimilation at 100 min of reference model assimilation time. The variation of rms errors for the analyzed fields using data from two radars is shown in Figure 6. It is not surprising that the rms errors for model u and v are much improved for 3DVAR. For the first several data assimilation cycles, the errors for 3DVAR are the lowest. With more cycles, the errors for the hybrid method become the lowest among three methods. For most of variables (except potential temperature), the errors for 3DVAR decrease more quickly than seen in the other two methods for the first several data assimilation cycles but then remain at higher levels for later

DA cycles. The variation of rms errors is less volatile when data from two radars are used compared to when data from a single radar is used for 3DVAR. The other features are quite similar to the cases when data from a single radar are used.

5. Summary and Future Work

A hybrid EnKF-3DVAR data assimilation system has been developed based on existing 3DVAR and ensemble Kalman filter (EnKF) programs within the ARPS model. The algorithm uses the extended control variable approach to combine the static and ensemble-derived flow-dependent forecast error covariances [30, 31, 37].

The method is applied to the assimilation of radar data from a simulated supercell storm. Two groups of experiments are performed using different amounts of radar data. Results obtained using 3DVAR (with static covariances entirely), hybrid EnKF-3DVAR, and EnKF are compared. When data from a single radar are used, results show that after 16 cycles of data assimilation, the EnKF and hybrid schemes provide similar results. When evaluated in term of rms errors, the EnKF provides slightly better results for the model dynamic variables, while the hybrid provides slightly better results for the hydrometeor related variables. Though the storm structures can be established reasonably well using 3DVAR, its rms errors are generally worse than those from the other two methods. When data from two radars are used, the rms errors for the hybrid method are smallest for most of the model variables. With two radars, the results from 3DVAR are close to those from EnKF. These tests also indicate that the hybrid scheme can reduce the storm spin-up time because it fits the observations, especially the reflectivity observations, better than the EnKF and the 3DVAR at the beginning of the assimilation cycles. Thus, precipitation exists from the beginning of the model integration.

Our future studies will try to answer a number of key questions within the hybrid EnKF-3DVAR framework just described. They include the following. (1) What is the optimal choice for the relative weight of the static and flow-dependent covariances for storm scale radar data assimilation? (2) What is the optimal combination of ensemble size and grid spacing for a specific computational cost? (3) How does the overall performance of the proposed method compare with 3DVAR and EnKF methods when model error is present? More sensitivity experiments will be performed to answer these questions in the near future, and results will likely help us to solve the challenges of applying this method to real-world scenarios. Even if these questions are successfully answered, the high computational cost of this method is still likely to be a big hurdle. For this, we will apply the dual-resolution strategy as developed for the EnKF scheme in Gao and Xue [46]. A new strategy for hybrid data assimilation proposed by Penny [55] also will be tested within a storm scale data assimilation framework in the near future.

Acknowledgments

This research was funded by the NOAA Warn-on-Forecast project. The first two authors were partially supported by

NSF Grants EEC-0313747, ATM-0738370, ATM-0331594, and AGS-0802888.

References

- [1] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [2] M. Hu, M. Xue, J. Gao, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: impact of radial velocity analysis via 3DVAR," *Monthly Weather Review*, vol. 134, no. 2, pp. 699–721, 2006.
- [3] M. Hu and M. Xue, "Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma city tornadic thunderstorm case," *Monthly Weather Review*, vol. 135, no. 2, pp. 507–525, 2007.
- [4] G. Ge and J. Gao, "Latest development of 3DVAR system for ARPS and its application to a tornadic supercell storm," in *Proceedings of the 22nd Conference on Weather Analysis and Forecasting/18th Conference on Numerical Weather Prediction*, 2007.
- [5] G. Ge, J. Gao, and M. Xue, "Diagnostic pressure equation as a weak constraint in a storm-scale three dimensional variational radar data assimilation system," *Journal of Atmospheric and Oceanic Technology*, vol. 29, pp. 1075–1092, 2012.
- [6] M. Xue, F. Kong, K. Thomas et al., "CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA hazardous weather testbed spring experiment," in *Proceedings of the 24th Conference on Severe Local Storms*, AMS, Savannah, Ga, USA, 2008.
- [7] F. Kong, M. Xue, K. W. Thomas et al., "A realtime storm-scale ensemble forecast system: 2009 spring experiment," in *Proceedings of the 23rd Conference on Weather Analysis and Forecasting/19th Conference on Numerical Weather Prediction*, American Meteorological Society, Omaha, Neb, USA, 2009.
- [8] A. D. Schenkman, M. Xue, A. Shapiro, K. Brewster, and J. Gao, "Impact of CASA radar and Oklahoma Mesonet data assimilation on the analysis and prediction of tornadic mesovortices in an MCS," *Monthly Weather Review*, vol. 139, no. 11, pp. 3422–3445, 2011.
- [9] A. D. Schenkman, M. Xue, A. Shapiro, K. Brewster, and J. Gao, "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*, vol. 139, no. 1, pp. 224–246, 2011.
- [10] D. M. Barker, W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. Xiao, "A three-dimensional variational data assimilation system for MM5: implementation and initial results," *Monthly Weather Review*, vol. 132, pp. 897–914, 2004.
- [11] Q. Xiao, Y.-H. Kuo, J. Sun et al., "Assimilation of Doppler radar observations with a regional 3DVAR system: impact of Doppler velocities on forecasts of a heavy rainfall case," *Journal of Applied Meteorology*, vol. 44, no. 6, pp. 768–788, 2005.
- [12] W. C. Skamarock, J. B. Klemp, J. Dudhia et al., "A description of the advanced research WRF," Version 2, 2005.
- [13] F. Rabier, "Overview of data assimilation developments in numerical weather prediction centers," in *Proceedings of the 4th WMO International Symposium: Assimilation of Observations in Meteorology and Oceanography*, Prague, Czech Republic, 2005.

- [14] J. Sun and N. A. Crook, "Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: model development and simulated data experiments," *Journal of the Atmospheric Sciences*, vol. 54, no. 12, pp. 1642–1661, 1997.
- [15] J. Sun and N. A. Crook, "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," *Journal of the Atmospheric Sciences*, vol. 55, no. 5, pp. 835–852, 1998.
- [16] J. Sun, "Initialization and numerical forecasting of a supercell storm observed during STEPS," *Monthly Weather Review*, vol. 133, no. 4, pp. 793–813, 2005.
- [17] Y. Honda and K. Koizumi, "The impact of the assimilation of precipitation data and Radar reflectivity with a pre-operational 4DVAR for the JMA nonhydrostatic model," in *Proceedings of the 10th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, Land Surface (IOAS-AOLS '06)*, 2006.
- [18] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, no. 5, pp. 10143–10162, 1994.
- [19] C. Snyder and F. Zhang, "Assimilation of simulated Doppler radar observations with an ensemble Kalman filter," *Monthly Weather Review*, vol. 131, no. 8, pp. 1663–1677, 2003.
- [20] F. Zhang, C. Snyder, and J. Sun, "Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filter," *Monthly Weather Review*, vol. 132, pp. 1238–1253, 2004.
- [21] M. Tong and M. Xue, "Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments," *Monthly Weather Review*, vol. 133, no. 7, pp. 1789–1807, 2005.
- [22] D. Dowell, F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, "Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: ensemble Kalman filter experiments," *Monthly Weather Review*, vol. 132, pp. 1982–2005, 2004.
- [23] D. C. Dowell, L. J. Wicker, and C. Snyder, "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*, vol. 139, no. 1, pp. 272–294, 2011.
- [24] N. Yussouf and D. J. Stensrud, "Impact of phased-array radar observations over a short assimilation period: observing system simulation experiments using an ensemble Kalman filter," *Monthly Weather Review*, vol. 138, no. 2, pp. 517–538, 2010.
- [25] N. Snook, M. Xue, and Y. Jung, "Analysis of a tornadic mesoscale convective vortex based on ensemble kalman filter assimilation of CASA X-band and WSR-88D radar data," *Monthly Weather Review*, vol. 139, no. 11, pp. 3446–3468, 2011.
- [26] N. Snook, M. Xue, and Y. Jung, "Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data," *Monthly Weather Review*, vol. 140, pp. 2126–2146, 2012.
- [27] P. L. Houtekamer and H. L. Mitchell, "A sequential ensemble Kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 129, no. 1, pp. 123–137, 2001.
- [28] H. L. Mitchell, P. L. Houtekamer, and G. Pellerin, "Ensemble size, balance, and model-error representation in an ensemble Kalman filter," *Monthly Weather Review*, vol. 130, no. 11, pp. 2791–2808, 2002.
- [29] T. M. Hamill and C. Snyder, "A hybrid ensemble Kalman filter-3D variational analysis scheme," *Monthly Weather Review*, vol. 128, no. 8, pp. 2905–2919, 2000.
- [30] A. C. Lorenc, "The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var," *Quarterly Journal of the Royal Meteorological Society*, vol. 129, no. 595, pp. 3183–3203, 2003.
- [31] M. Buehner, "Ensemble-derived stationary and flow-dependent background-error covariances: evaluation in a quasi-operational NWP setting," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 607, pp. 1013–1043, 2005.
- [32] M. Zupanski, "Maximum likelihood ensemble filter: theoretical aspects," *Monthly Weather Review*, vol. 133, no. 6, pp. 1710–1726, 2005.
- [33] X. Tian, A. Dai, D. Yang, and Z. Xie, "Effects of precipitation-bias corrections on surface hydrology over northern latitudes," *Journal of Geophysical Research*, vol. 112, no. 14, pp. 1984–2012, 2007.
- [34] C. Liu, Q. Xiao, and B. Wang, "An ensemble-based four-dimensional variational data assimilation scheme. Part I: technical formulation and preliminary test," *Monthly Weather Review*, vol. 136, no. 9, pp. 3363–3373, 2008.
- [35] C. Liu, Q. Xiao, and B. Wang, "An ensemble-based four-dimensional variational data assimilation scheme. Part II: observing system simulation experiments with advanced research WRF (ARW)," *Monthly Weather Review*, vol. 137, no. 5, pp. 1687–1704, 2009.
- [36] N. E. Bowler, J. Flowerdew, and S. R. Pring, "Tests of different flavours of EnKF on a simple model," *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 675, pp. 1505–1519, 2013.
- [37] X. Wang, C. Snyder, and T. M. Hamill, "On the theoretical equivalence of differently proposed ensemble—3DVAR hybrid analysis schemes," *Monthly Weather Review*, vol. 135, no. 1, pp. 222–227, 2007.
- [38] D. M. Barker and Coauthors, "The Weather Research and Forecasting model's community variational/ensemble data assimilation system: WRFDA," *Bulletin of the American Meteorological Society*, vol. 93, pp. 831–843, 2012.
- [39] Y. Li, X. Wang, and M. Xue, "Assimilation of radar radial velocity data with the WRF hybrid ensemble-3DVAR system for the prediction of hurricane Ike (2008)," *Monthly Weather Review*, vol. 140, pp. 3507–3524, 2012.
- [40] F. Zhang, M. Zhang, and J. Poterjoy, "E4DVar: coupling an ensemble kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar," *Monthly Weather Review*, vol. 141, pp. 900–917, 2013.
- [41] X. Wang, D. M. Barker, C. Snyder, and T. M. Hamill, "A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: observation system simulation experiment," *Monthly Weather Review*, vol. 136, pp. 5116–5131, 2008.
- [42] X. Wang, D. M. Barker, C. Snyder, and T. M. Hamill, "A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II: real observation experiments," *Monthly Weather Review*, vol. 136, pp. 5116–5131, 2008.
- [43] M. Buehner, P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, "Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: description and single-observation experiments," *Monthly Weather Review*, vol. 138, no. 5, pp. 1550–1566, 2010.

- [44] M. Buehner, P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, "Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: one-month experiments with real observations," *Monthly Weather Review*, vol. 138, no. 5, pp. 1567–1586, 2010.
- [45] J. C. Derber and A. Rosati, "A global oceanic data assimilation system," *Journal of Physical Oceanography*, vol. 19, pp. 1333–1347, 1989.
- [46] J. Gao and M. Xue, "An efficient dual-resolution approach for ensemble data assimilation and tests with simulated Doppler radar data," *Monthly Weather Review*, vol. 136, no. 3, pp. 945–963, 2008.
- [47] P. S. Ray, "The morphology of several tornadic storms on 20 May 1977," *Journal of the Atmospheric Sciences*, vol. 38, no. 8, pp. 1643–1663, 1981.
- [48] Y. L. Ying Lin, P. S. Ray, and K. W. Johnson, "Initialization of a modeled convective storm using Doppler radar-derived fields," *Monthly Weather Review*, vol. 121, no. 10, pp. 2757–2775, 1993.
- [49] M. Xue, K. K. Droegemeier, and V. Wong, "The advanced regional prediction system (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: model dynamics and verification," *Meteorology and Atmospheric Physics*, vol. 75, no. 3-4, pp. 161–193, 2000.
- [50] M. Xue, K. K. Droegemeier, V. Wong et al., "The advanced regional prediction system (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: model physics and applications," *Meteorology and Atmospheric Physics*, vol. 76, no. 3-4, pp. 143–165, 2001.
- [51] J. Gao, M. Xue, A. Shapiro, Q. Xu, and K. K. Droegemeier, "Three-dimensional simple adjoint velocity retrievals from single-Doppler radar," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 1, pp. 26–38, 2001.
- [52] P. L. Jr. Smith, C. G. Myers, and H. D. Orville, "Radar reflectivity factor calculations in numerical cloud models using bulk parameterization of precipitation processes," *Journal of Applied Meteorology*, vol. 14, pp. 1156–1165, 1975.
- [53] B. S. Ferrier, "A double-moment multiple-phase four-class bulk ice scheme. Part I: description," *Journal of the Atmospheric Sciences*, vol. 51, no. 2, pp. 249–280, 1994.
- [54] M. Zupanski, S. J. Fletcher, I. M. Navon et al., "Initiation of ensemble data assimilation," *Tellus A*, vol. 58, no. 2, pp. 159–170, 2006.
- [55] S. Penny, "The hybrid local ensemble transform kalman filter," *Monthly Weather Review*. Accepted.

Research Article

Impact of a Diagnostic Pressure Equation Constraint on Tornadoic Supercell Thunderstorm Forecasts Initialized Using 3DVAR Radar Data Assimilation

Guoqing Ge,¹ Jidong Gao,² and Ming Xue^{1,3}

¹ Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David L. Boren Boulevard, Norman, OK 73072, USA

² National Severe Storms Laboratory, Norman, OK 73072, USA

³ School of Meteorology, University of Oklahoma, Norman, OK 73072, USA

Correspondence should be addressed to Guoqing Ge; gegaoqing@ou.edu

Received 9 April 2013; Revised 25 July 2013; Accepted 13 September 2013

Academic Editor: Louis Wicker

Copyright © 2013 Guoqing Ge et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A diagnostic pressure equation constraint has been incorporated into a storm-scale three-dimensional variational (3DVAR) data assimilation system. This diagnostic pressure equation constraint (DPEC) is aimed to improve dynamic consistency among different model variables so as to produce better data assimilation results and improve the subsequent forecasts. Ge et al. (2012) described the development of DPEC and testing of it with idealized experiments. DPEC was also applied to a real supercell case, but only radial velocity was assimilated. In this paper, DPEC is further applied to two real tornadoic supercell thunderstorm cases, where both radial velocity and radar reflectivity data are assimilated. The impact of DPEC on radar data assimilation is examined mainly based on the storm forecasts. It is found that the experiments using DPEC generally predict higher low-level vertical vorticity than the experiments not using DPEC near the time of observed tornadoes. Therefore, it is concluded that the use of DPEC improves the forecast of mesocyclone rotation within supercell thunderstorms. The experiments using different weighting coefficients generate similar results. This suggests that DPEC is not very sensitive to the weighting coefficients.

1. Introduction

A dynamic consistent initial condition is very important for making a quality storm-scale numerical weather prediction (NWP) forecast. For this purpose, a large number of studies have been focused on utilizing high-resolution radar data to provide better storm-scale initial conditions (e.g., [1–7]). Since radars primarily observe the radial velocity and reflectivity, most state variables have to be “retrieved” in the data assimilation (DA) process. This makes the assimilation of radar data a very challenging problem.

Three-dimensional variational (3DVAR), four-dimensional variational (4DVAR), and ensemble Kalman filter (EnKF) methods have been applied to the previously mentioned radar DA problem. The 4DVAR method uses a NWP model as a strong constraint and hence naturally produces a dynamically consistent analysis. Sun and Crook [8, 9] and Sun [10] have shown encouraging results using a 4DVAR

cloud model. However, it is very difficult to develop and maintain complex adjoint codes for NWP models. Complex ice microphysics, which are important for storm-scale applications but contain discontinuities and strong nonlinearities, introduce more difficulties in this situation. All of these difficulties limit the adoption of the 4DVAR method in storm-scale NWP operations. The EnKF technique is expected to generate similar analysis quality as 4DVAR but avoids the coding of an adjoint model. Many EnKF-based radar DA studies have been carried out in recent years (e.g., [11–24]). These studies have shown great potential for the EnKF method. However, EnKF is not as mature as the variational methods and so far successful applications to real data assimilation problems are still limited. Computationally it has similar cost as the 4DVAR approach.

The 3DVAR method is more efficient than 4DVAR and EnKF in terms of computation cost. For this reason, 3DVAR has been applied in many real-time applications. Past studies

(e.g., [4–6, 25]) have used 3DVAR to assimilate radar data for the prediction of tornadic supercell storms. The advanced regional prediction system (ARPS, [26–28]) 3DVAR system, and its cloud analysis package have been used to make continental-US-scale real-time weather predictions at up to 1 km resolution [29, 30]. However, the 3DVAR method is still challenged by its theoretical suboptimality due to the use of a static diagonal background error covariance matrix and the lack of suitable balances among model variables. Research has been done to address this problem. For example, hybrid approaches (e.g., [31–35]) have been proposed to provide flow-dependent background covariances obtained from a forecast ensemble, for the variational system. Because of the use of an ensemble, the method is still computationally expensive and requires further research.

Suitable weak constraints can also be built into the 3DVAR cost function to improve the balance among model variables and help spread observational information to state variables which are not directly observed. Gao et al. [36–38], Hu et al. [4, 5], and Hu and Xue [39] incorporated an anelastic mass continuity equation into the ARPS 3DVAR system in the form of a weak constraint to couple three wind components. Ge et al. [7] further incorporated the diagnostic pressure equation into the 3DVAR cost function in addition to the previously mentioned mass continuity equation constraint (MCEC). The main goal of this diagnostic pressure equation constraint (DPEC) is to improve the consistency between dynamic and thermodynamic fields. Ge et al. [7] demonstrated, using observational system simulation experiments (OSSEs), that DPEC can improve the data assimilation results slightly after a given period of data assimilation. The application of DPEC to a real tornadic supercell thunderstorm case showed that DPEC can improve the forecast in terms of the general evolution of storm cells and mesocyclone rotation near the time of the observed tornado.

For the real case study by Ge et al. [7], only radial velocity data was assimilated. In this paper, we will further examine the impact of DPEC by applying it to two real tornadic supercell thunderstorm cases by assimilating the radial velocity and radar reflectivity data together. The rest of the paper is organized as follows. Section 2 will brief the updated ARPS 3DVAR system with the implementation of DPEC. Section 3 applies the updated system to the 5 May 2007 Greensburg, Kansas tornadic supercell thunderstorm case, while Section 4 applies the system to the 8 May 2003 Oklahoma City tornadic supercell thunderstorm case. The summary and future work will be presented in Section 5.

2. The Scheme for the 3DVAR System

A 3DVAR system within the ARPS model framework [26–28] has been developed and applied to the assimilation of weather radar and other data [4–6, 29, 36, 38, 39]. The system consists of two components: one is the 3DVAR subsystem, whose purpose is to assimilate radar radial velocity data as well as other conventional observations in a three-dimensional variational framework; the other is the cloud analysis subsystem whose purpose is to assimilate the radar

reflectivity data based mainly on semiempirical rules. The cloud analysis system not only updates the hydrometeor fields, but also can adjust the in-cloud temperature and water vapor fields.

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

$$J(x) = J_b + J_o + J_c,$$

$$J_b = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b), \quad (1)$$

$$J_o = \frac{1}{2}[H(\mathbf{x}) - \mathbf{y}^o]^T \mathbf{R}^{-1} [H(\mathbf{x}) - \mathbf{y}^o].$$

Following the standard notion of Ide et al. [40], \mathbf{x} and \mathbf{x}^b are the analysis and background state vectors, and \mathbf{y}^o is the observation vector. \mathbf{B} and \mathbf{R} are the background and observation error covariance matrices, respectively. $H(\mathbf{x})$ is the nonlinear observation operator. To improve the conditioning of the J minimization problem and avoid the need for the inverse of \mathbf{B} , a new control variable \mathbf{v} is introduced, which is related to the analysis increment $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b$ according to

$$\delta\mathbf{x} = \mathbf{B}^{1/2}\mathbf{v}, \quad (2)$$

where $\mathbf{B}^{1/2}$ is the square root of \mathbf{B} ; that is, $\mathbf{B} = (\mathbf{B}^{1/2})^T (\mathbf{B}^{1/2})$.

In terms of \mathbf{v} , the background term becomes

$$J_b = \left(\frac{1}{2}\right) \mathbf{v}^T \mathbf{v}. \quad (3)$$

Consequently, the minimization is performed in the space of \mathbf{v} . The recursive filter proposed by Purser et al. [41, 42] is used to model the effect of the background error covariance, or more precisely, the square root of \mathbf{B} . Currently, in our 3DVAR system, the background state vector, \mathbf{x}^b , can be provided by a sounding profile, a previous ARPS model forecast, or a forecast from another model. The analysis vector \mathbf{x} contains the three wind components (u , v , and w), potential temperature (θ), pressure (p), and water vapor mixing ratio (q_w). The observations include Doppler radar radial velocity, single-level (such as surface observations), and multiple-level conventional observations (such as those of rawinsondes and wind profilers). For the study in this paper, only Doppler radial velocity data is used by the 3DVAR subsystem and a 2 m s^{-1} observation error variance is assumed for the data.

Term J_c in (1) includes any penalty or equation constraint terms. Currently, it includes two terms as defined in the following:

$$J_c = J_{MC} + J_{DP}. \quad (4)$$

The first term J_{MC} is intended to minimize the 3D anelastic mass divergence so as to provide the key coupling among the three wind components. The definition and impact of this constraint have been investigated by Gao et al. [36, 38] and Hu et al. [5].

The second term J_{DP} is the DPEC term defined as follows:

$$J_{DP} = P(\mathbf{x})^T \mathbf{A}_P^{-1} P(\mathbf{x}), \quad (5)$$

$$\begin{aligned} P \equiv \nabla \cdot \vec{E} \equiv & -\nabla^2 p' - \nabla \cdot (\bar{\rho} \vec{V} \cdot \nabla \vec{V}) \\ & + g \frac{\partial}{\partial z} \left(\bar{\rho} \left[\frac{\theta'}{\bar{\theta}} - \frac{p'}{\bar{\rho} c_s^2} + \frac{q'_v}{\varepsilon + \bar{q}_v} - \frac{q'_v + q_{\text{liquid+ice}}}{1 + \bar{q}_v} \right] \right) \\ & + \nabla \cdot \vec{C} + \nabla \cdot \vec{D}, \end{aligned} \quad (6)$$

where

$$\begin{aligned} \vec{E} &= \frac{\partial(\bar{\rho} \vec{V})}{\partial t} = \hat{i} \frac{\partial(\bar{\rho} u)}{\partial t} + \hat{j} \frac{\partial(\bar{\rho} v)}{\partial t} + \hat{k} \frac{\partial(\bar{\rho} w)}{\partial t}, \\ \vec{V} &= \hat{i} u + \hat{j} v + \hat{k} w, \\ \vec{C} &= \hat{i}(\bar{\rho} f v - \bar{\rho} \tilde{f} w) - \hat{j}(\bar{\rho} f u) + \hat{k}(\bar{\rho} \tilde{f} u), \\ \vec{D} &= \hat{i} D_u + \hat{j} D_v + \hat{k} D_w. \end{aligned} \quad (7)$$

The vector \vec{E} is the forcing term of the vector Euclidian momentum equation. The $q_{\text{liquid+ice}}$ includes hydrometeor mixing ratios. The \hat{i} , \hat{j} , and \hat{k} are unit vectors in the x , y , and z directions. The overbar represents base state and 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 are $f = 2\Omega \sin \phi$ and $\tilde{f} = 2\Omega \cos \phi$, 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 in the x , y , and z directions, respectively. When the mass continuity equation is applied, (6) becomes $P = 0$ where P represents the right-hand side (R.H.S) of (6).

Equation (6) is derived by applying the divergence operator to the three momentum equations of the ARPS model [26]:

$$\begin{aligned} \bar{\rho} \frac{\partial u}{\partial t} &= -\bar{\rho} \vec{V} \cdot \nabla u - \frac{\partial p'}{\partial x} + (\bar{\rho} f v - \bar{\rho} \tilde{f} w) + D_u, \\ \bar{\rho} \frac{\partial v}{\partial t} &= -\bar{\rho} \vec{V} \cdot \nabla v - \frac{\partial p'}{\partial y} - \bar{\rho} f u + D_v, \\ \bar{\rho} \frac{\partial w}{\partial t} &= -\bar{\rho} \vec{V} \cdot \nabla w - \frac{\partial p'}{\partial z} \\ &+ \bar{\rho} g \left[\frac{\theta'}{\bar{\theta}} - \frac{p'}{\bar{\rho} c_s^2} + \frac{q'_v}{\varepsilon + \bar{q}_v} - \frac{q'_v + q_{\text{liquid+ice}}}{1 + \bar{q}_v} \right] \\ &+ \bar{\rho} \tilde{f} u + D_w. \end{aligned} \quad (8)$$

The \mathbf{A}_P in (5) is the error covariance matrix associated with the DPEC constraint, which is assumed to be diagonal with empirically defined constant diagonal elements as the variances. The inverse of \mathbf{A}_P is called the weighting coefficient and determines the relative importance of the DPEC

constraint and its optimal value can be determined through numerical experiments, similar to the way to determine certain weights in cloud-scale variational data assimilation systems (e.g., [3]). Usually, the constraint terms with their weights should be of similar orders of magnitude as other terms in J for them to be effective.

2.2. The Cloud Analysis Subsystem. The cloud analysis subsystem is based on the local analysis and prediction system (LAPS, [43]) with significant modifications by Zhang et al. [44], Brewster [45], and Hu et al. [4]. It is used to assimilate radar reflectivity data into the model. It updates the following model fields: rain water mixing ratio, snow mixing ratio, hail mixing ratio, cloud ice mixing ratio, cloud water mixing ratio, water vapor mixing ratio, and temperature. The general procedure is as follows. For each model grid point inside an area with observed reflectivity, a precipitation type (rain, snow, or hail) is first determined according to the reflectivity observation and the background state. After the precipitation type is determined, its mixing ratio is computed using reflectivity equations that link the precipitation species with reflectivity data. The cloud water and cloud ice mixing ratios are estimated by assuming that air parcels ascend moist-adiabatically from cloud base to cloud top. The water vapor mixing ratio is adjusted so that the air is saturated in precipitation area. The temperature field is also changed in order to make the in-cloud temperature consistent with the changed cloud fields. Readers can refer to Hu et al. [4] for more details.

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 q_v , rain water mixing ratio q_r , snow water mixing ratio q_s , hail mixing ratio q_h , cloud water mixing ratio q_c , and ice mixing ratio q_i . Currently, the cloud analysis subsystem is a follow-up step after the 3DVAR subsystem finished running. These two subsystems are separate from each other and there is no suitable coupling between the wind fields and the thermodynamic fields. Therefore, there may be inconsistencies between the different model variables in the data analysis. These inconsistencies may harm the quality of subsequent data assimilation cycles and the ensuing forecast.

To alleviate these kinds of inconsistencies, 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 pseudoobservations and be ingested, as well as the radar radial velocity data, by the 3DVAR subsystem. DPEC will then act to couple all model variables so as 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 the following data assimilation cycles and the subsequent forecast will be improved. In practice, the hydrometeors are updated directly by the cloud analysis subsystem. The temperature and water vapor mixing ratio

fields are updated in the 3DVAR subsystem by treating the corresponding results obtained from the cloud analysis subsystem as pseudoobservations. The assumed error variances for the pseudoobservations θ and q_v are 0.5K and $5 \times 10^{-4} \text{g}(\text{kg})^{-1}$, respectively.

3. The 5 May 2007 Greensburg Tornadoic Supercell Storm Case

The 5 May 2007 Greensburg, Kansas (KS) tornadic thunderstorm complex produced 18 tornadoes in the Dodge City area and additional 47 tornado reports in Kansas, Nebraska, and Missouri. One tornado was the strongest observed in recent years. This tornado started moving through Greensburg at 0245 UTC 5 May 2007 (2145 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 [46]. A detailed description of the supercell that spawned this tornado and its environmental setting can be found in Bluestein [47] and Stensrud and Gao [6].

For this real data case, we used a 3 km grid spacing with 200×200 grid points in the horizontal. The ARPS model domain is shown in Figure 1. The domain was selected with sufficient coverage to contain the principal features of interest while maintaining some distance between the primary storms and the lateral boundaries. The model uses 47 terrain following vertical layers with nonlinear vertical 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 Lin three-ice microphysical scheme [48] was used together with a 1.5-order turbulent kinetic energy subgrid parameterization. A wave radiation condition was applied at the top boundary and rigid-wall conditions were applied to the bottom boundary.

The impact of the DPEC will be discussed in terms of the quality of ensuing forecasts instead of the analysis because no truth or high-resolution observation is available for verification of the analysis. Four experiments were conducted for this case (Table 1). The first experiment did not include DPEC in J and will be referred to as experiment NoDP1. The second experiment used DPEC with the DP weighting coefficient of $1.0E8$ and is referred to as experiment DP1. The third and fourth experiments were similar to DP1 except that the DP weighting coefficients were multiplied and divided by 5, respectively. They are referred to as experiments DP1m5 and DP1d5, respectively. In all of the previously four experiments the mass continuity equation constraint was used with the MC weighting coefficient of $1.0E8$.

For all the previously mentioned four experiments, data from five radars at Dodge City, Kansas (KDDC), Vance Air Force Base, Oklahoma (KVNXX), Wichita, Kansas (KICT), Oklahoma City, Oklahoma (KTLX), and Amarillo, Texas (KAMA), was used (Figure 1). A quality control procedure was applied before the use of the radar data, which included clutter removal and velocity dealiasing using SOLOII software from the National Center for Atmospheric Research (NCAR). The initial analysis background and the boundary conditions came from the mean of a mesoscale ensemble

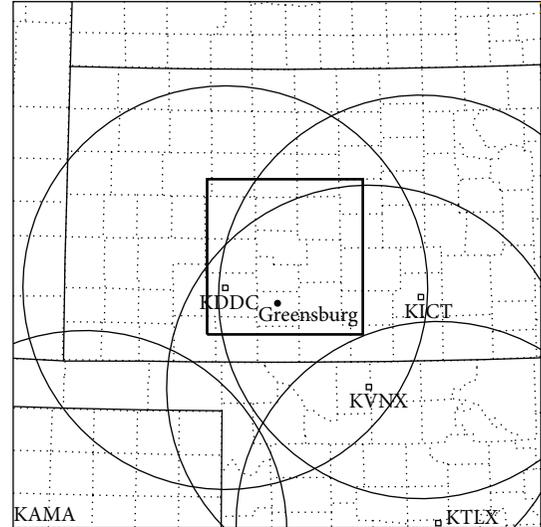


FIGURE 1: The model domain with county boundaries for the 5 May 2007 Greensburg tornadoic supercell thunderstorm case study. The five radars as well as their 230 km range circles are also shown. The black dot shows the location of the town of Greensburg. The black bold inner box illustrates the domain coverage in Figure 2.

TABLE 1: List of data assimilation experiments (DP stands for “diagnostic pressure equation”).

Case name	Experiment name	DP weighting coefficient
5 May 2007 Greensburg case	NODP1	0
	DP1	$1E8$
	DP1d5	$2E7$
	DP1m5	$5E8$
8 May 2003 OKC case	NODP2	0
	DP2	$1E8$
	DP2d5	$2E7$
	DP2m5	$5E8$

assimilation system run at 30 km grid spacing [6]. While Stensrud and Gao [6] performed a 3DVAR analysis at only one time before the launch of the forecast, the present study performed cycled 3DVAR analyses with a 1 h long assimilation period before the forecast. A five-minute ARPS forecast followed each analysis, and this process was repeated until the end of the 1 h assimilation period. From the final analysis, a 1 h forecast was launched. In this way, each experiment consisted of a 1 h assimilation period (from 0130 and 0230 UTC) and a 1 h forecast period (0230–0330 UTC).

We now focus the discussion on the dominant supercell thunderstorm at the southernmost portion of the storm complex, which produced the EF-5 tornado that hit the Greensburg area between 0245 UTC–0305 UTC. A hook echo signature was evident at 0230 UTC. As the storm reached Greensburg, the hook echo signature became less prominent due to the tornado moving in a more northerly direction and toward the storm’s main core. During this period, the radar velocity observations indicated strong cyclonic rotation

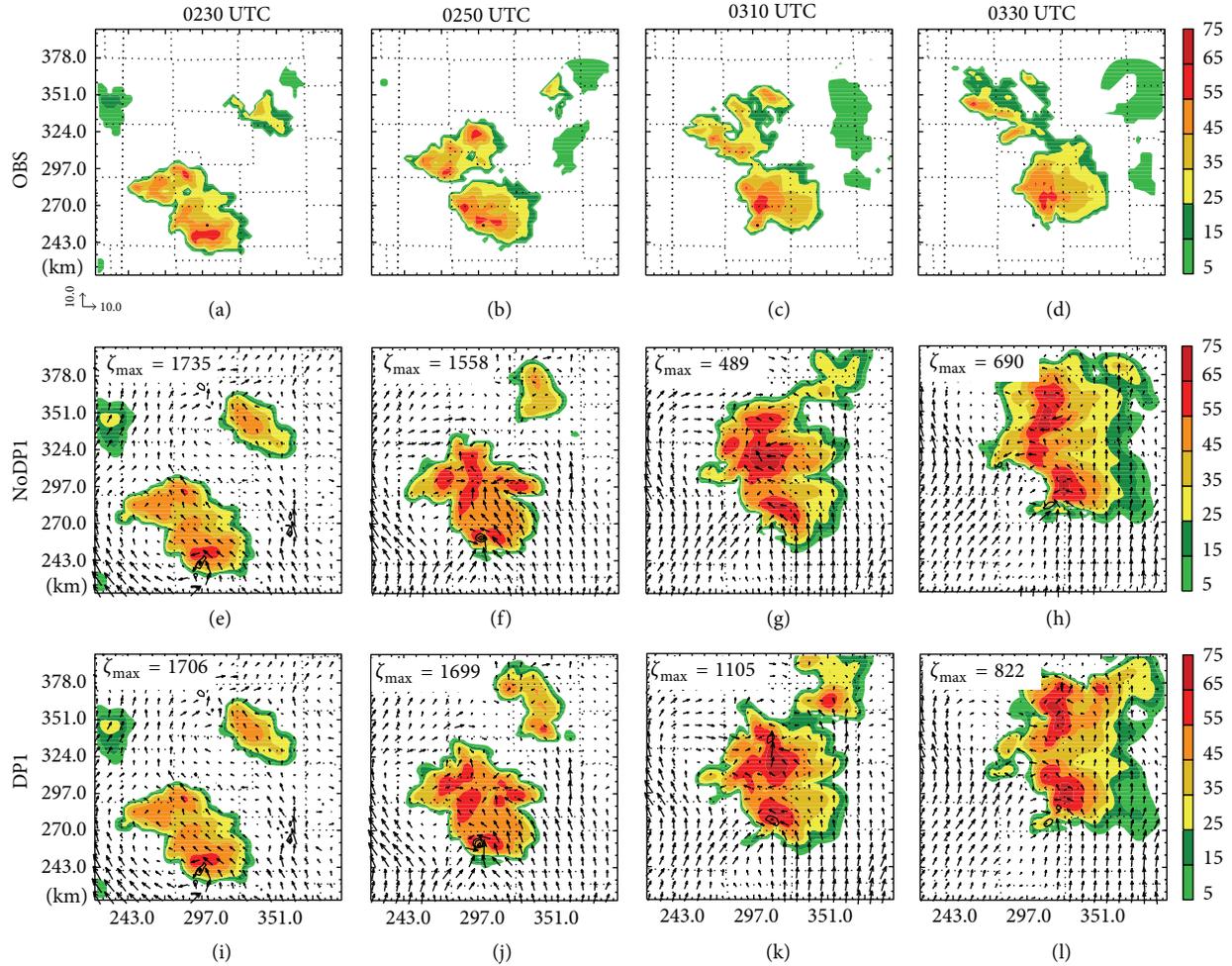


FIGURE 2: Observed radar reflectivity mosaic (dBZ) at 2 km MSL from KDDC, KICT, KVNK, KAMA, and KTLX Doppler radars valid at (a) 0230, (b) 0250, (c) 0310, and (d) 0330 UTC; simulated radar reflectivity (dBZ), horizontal winds, and vertical vorticity (contours starting at 0.005 s^{-1} with an interval of 0.005 s^{-1}) at 2 km MSL from NoDP1 valid at (e) 0230, (f) 0250, (g) 0310, and (h) 0330 UTC, and from DP1 valid at (i) 0230, (j) 0250, (k) 0310, and (l) 0330 UTC. The duration 0230–0330 UTC covers the 1 h forecast period. The black dots in (a)–(d) indicate the location of the town of Greensburg. The maximum vertical vorticity is shown for NoDP1 and DP1 experiments with the unit of 10^{-5} s^{-1} .

associated with the violent tornado. The entire storm complex moved gradually toward the northeast. After passing the town of Greensburg, a second tornado, rated EF-3, developed at the end of the Greensburg tornado’s path just northeast of the town [46]. A radar reflectivity mosaic was created from the aforementioned five WSR-88D radars by interpolating reflectivity data from all radars onto model grid points and keeping the largest reflectivity value for each grid point. The reflectivity mosaic was then used for forecast verification. The evolution of the storm as indicated by the radar reflectivity mosaic at 2 km MSL is shown in Figure 2 from 0230 to 0330 UTC every twenty minutes. Note that the hook echo is not evident in these figures due to the use of 3 km resolution and a smoothing procedure applied in the mosaic generating process.

To demonstrate the impact of DPEC, we investigated these data assimilation experiments ingesting both the radial velocity data and reflectivity data. Figures 2(e)–2(l) show the reflectivity, horizontal wind vector, and vertical vorticity at

$z = 2 \text{ km MSL}$ from 0230 UTC to 0330 UTC every 20 minutes for the NoDP1 and DP1 experiments. It is shown that after 1 hour of data assimilation (Figures 2(e) and 2(i)), the storm had already spun up in terms of the reflectivity pattern. The reflectivity pattern, strength, and location agree well with the observed values (Figure 2(a)). A rotating circulation and a strong vertical vorticity column are collocated at the observed hook-echo region. The storm then moves gradually toward the northeast, which also agrees with the observations. After 0300 UTC, the predicted storm moves faster than what was observed. In spite of this, both NODP1 and DP1 still made reasonable forecasts in terms of the general evolution of the major storm. DP1d5 and DP1m5 produced very similar forecasts as DP1 and are therefore not shown in Figure 2.

In Figure 2, in terms of the reflectivity pattern, there is no significant difference in the general evolution of the storm between the NODP1 and DP1 experiments. The computed forecast scores (not presented here) also show little difference, consistent with the previous result. However, there is some

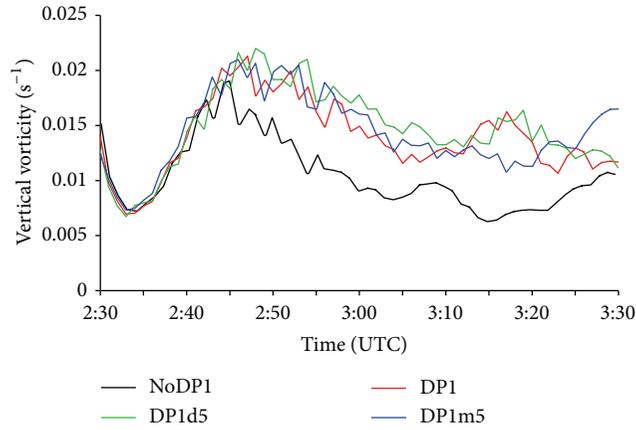


FIGURE 3: The time series of maximum vertical vorticity below two kilometers from 0230 UTC to 0330 UTC 5 May 2007 every one minute. The horizontal axis shows the time in UTC; the vertical axis shows the vertical vorticity values in units of s^{-1} . The black line is for experiment NoDP1, the red line for experiment DP1, the blue line for DP1m5, and the green line for DP1d5.

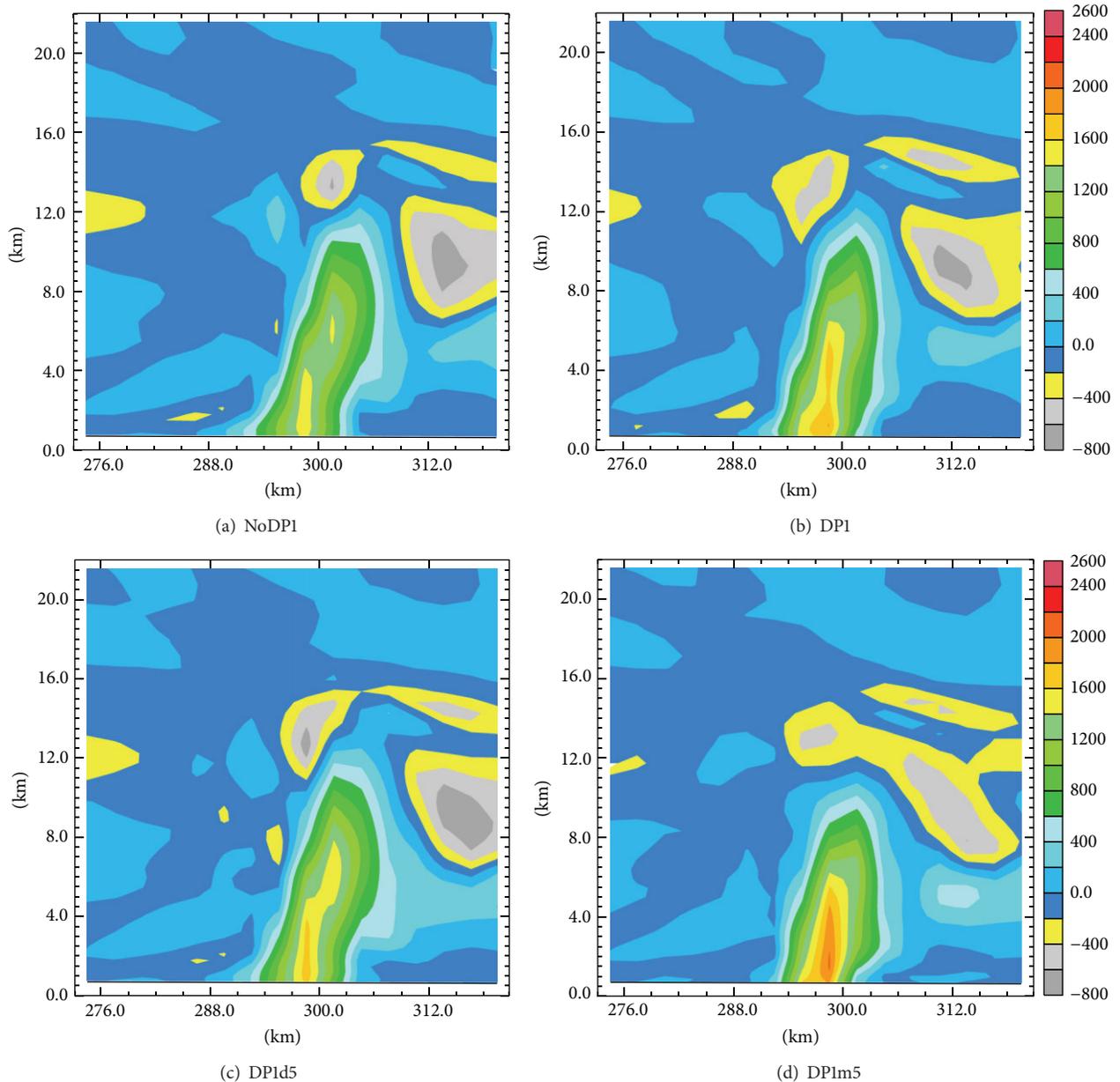


FIGURE 4: The vertical vorticity (in units of $10^{-5} s^{-1}$) at the vertical cross-section through the center of the major storm at $y = 253.5$ km at 0250 UTC 5 May 2007 for the (a) NoDP1, (b) DP1, (c) DP1d5, and (d) DP1m5.

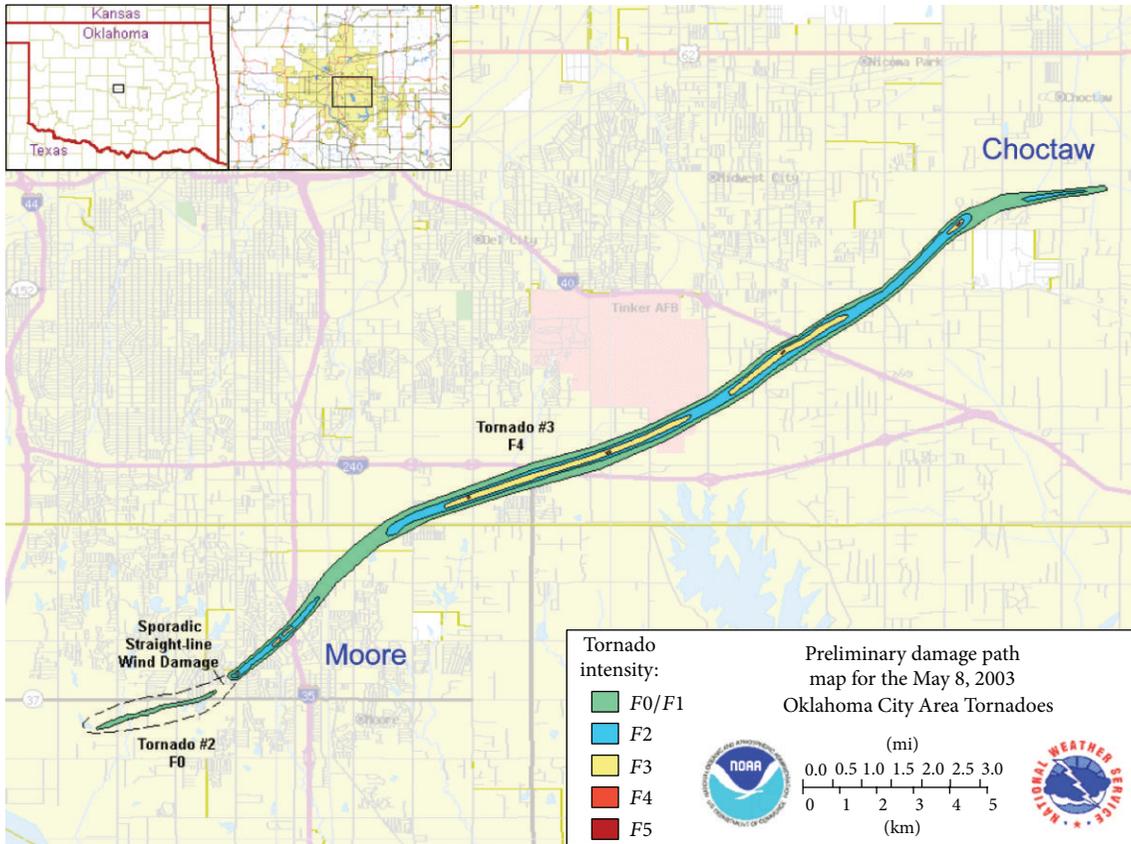


FIGURE 5: The damage path map for the 8 May 2003 Oklahoma City Area Tornadoes (National Weather Service, Norman).

difference in the predicted low-level mesocyclone rotation as indicated by larger maximum vertical vorticity in Figures 2(j), 2(k), and 2(l) than that in Figures 2(f), 2(g), and 2(h). As a further demonstration, Figure 3 shows the time series of the maximum vertical vorticity below two kilometers every minute from 0230 UTC to 0330 UTC for all four experiments. It is illustrated in Figure 3 that beginning at 0245 UTC and through the end of the forecast, the low-level maximum vertical vorticity from the experiments applying DPEC (the red, blue and green lines) is much larger than that from the NODP1 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 than those with smaller values of vertical vorticity. This kind of behavior is very similar to findings in Ge et al. [7]. As an example, Figure 4 presents the vertical vorticity at the vertical cross-section through the center of the major storm at $y = 259.5$ km at 0250 UTC 5 May 2007. It is noticeable that the experiments using DPEC (Figures 4(b), 4(c), and 4(d)) predicted stronger and deeper rotation than the “NODP1” experiment (Figure 4(a)). Therefore, it can be concluded that, for the experiments here, although the use of DPEC does not evidently improve the forecast of the general evolution of the major storm in terms of the reflectivity pattern, it does help improve the forecast of the mesocyclone rotation associated with the observed Greensburg tornado.

4. The 8 May 2003 Oklahoma City Tornadoic Supercell Storm Case

During the late afternoon on 8 May 2003, a major tornado hit the southern Oklahoma City metropolitan area (Figure 5). 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 was about 28 minutes from 2210 UTC to 2238 UTC. It caused up to F4 (Fujita scale) damages but no deaths. The tornado is hereafter referred to as the OKC tornado and the parent storm as the OKC tornadoic thunderstorm.

The synoptic environment on 8 May 2003 over Oklahoma was very favorable for the development of supercell storms and even tornadoes, as discussed by Hu and Xue [39] and Romine et al. [49]. The low-level flow was southerly over Oklahoma all day. A meridionally oriented dryline moved eastward approaching Moore, Oklahoma. A large amount of potential instability with 4004 J kg^{-1} convective available potential energy (CAPE), 1 J kg^{-1} convective inhibition (CIN) and about 25 m s^{-1} vertical shear over the lowest 6 km was present in the 1800 UTC 8 May Norman, Oklahoma (OUN) sounding. All of these conditions indicated that there was a high possibility for tornadoic supercell thunderstorms to develop.

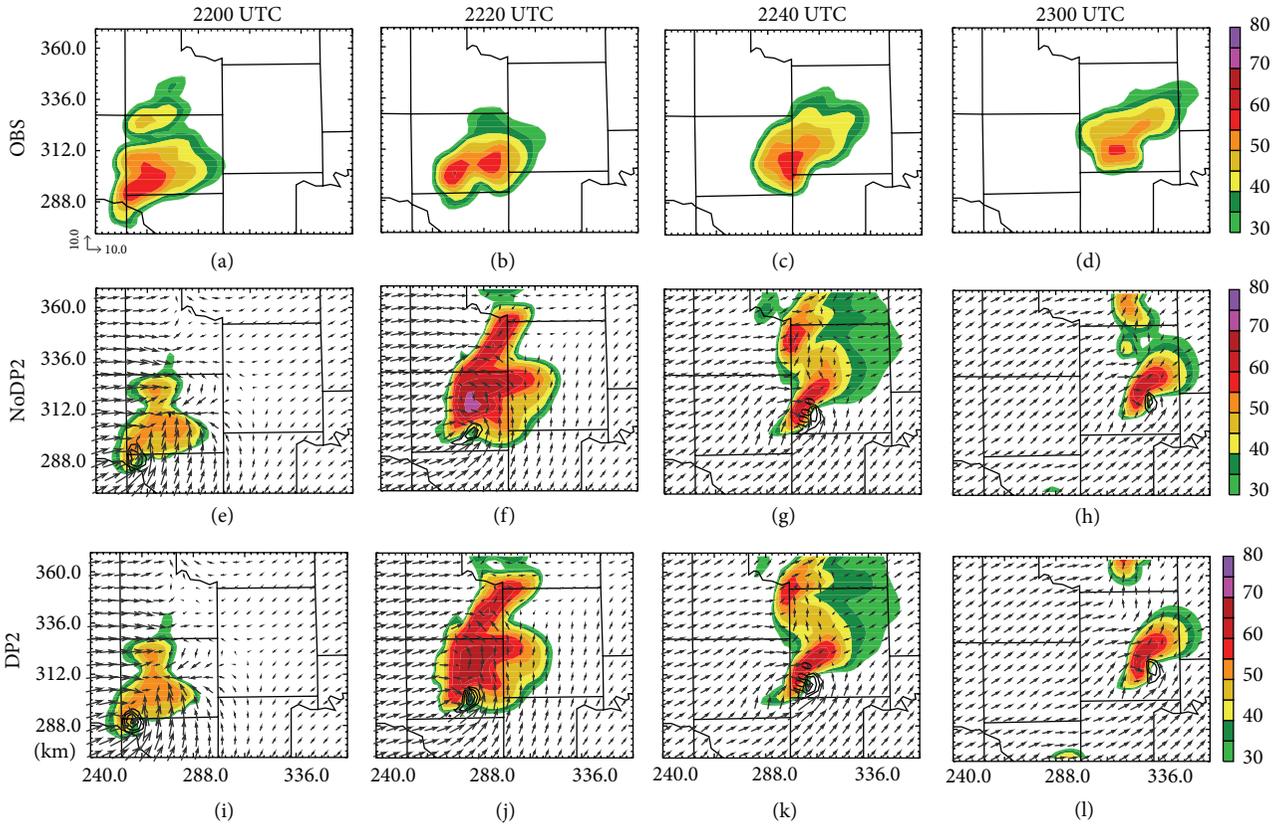


FIGURE 6: Observed radar reflectivity mosaic (dBZ) at 2 km MSL from KTLX, KVNx, KFDR, and KINX Doppler radars valid at (a) 2200, (b) 2220, (c) 2240, and (d) 2300 UTC; simulated radar reflectivity (dBZ), horizontal winds, and vertical vorticity (contours starting at 0.004 s^{-1} with an interval of 0.002 s^{-1}) at 2 km MSL from NoDP2 valid at (e) 2200, (f) 2220, (g) 2240, and (h) 2400 UTC, and from DP2 valid at (i) 2200, (j) 2220, (k) 2240, and (l) 2300 UTC. The duration 2200–2300 UTC covers the 1 h forecast period.

At about 2030 UTC, the first sign of the OKC tornadic storm showed up as a weak echo in the KTLX radar reflectivity field. By 2101 UTC, the storm had developed into a strong cell. In the following hour, the storm grew rapidly and moved northeastward. By 2201 UTC, the storm displayed an obvious hook echo signature at its southwestern end. The hook echo at this time was located just northwest of Moore, only several miles away from the center of the city. The pronounced hook echo signature was present until at least 2235 UTC while the parent supercell storm propagated east northeastward. The storm began weakening at 2240 UTC and dissipated by 0020 UTC 9 May. In addition to the OKC tornadic thunderstorm, there were three other short-lived storms (not shown). Here, we will just focus on the dominant thunderstorm. Figures 6(a), 6(b), 6(c), and 6(d) show the general evolution of the major thunderstorm every twenty minutes from 2200 UTC to 2300 UTC represented by the radar reflectivity mosaic at 2 km MSL.

All experiments were conducted with a horizontal resolution of 3 km. There were 195 grid points in both x and y directions. In the vertical direction, a stretched grid scheme was used. It contained 53 layers with an average grid spacing of 400 m, stretching from about 20 m at the surface to 770 m at the model top. The model domain is shown in Figure 7.

It covers nearly the entire state of Oklahoma. The evolution of the 8 May 2003 Oklahoma City tornadic supercell thunderstorm was roughly at the center of the domain. The four WSR-88D radars KTLX, KVNx, KINX, and KFDR and their associated coverage region are also shown in Figure 7. The outline near the KTLX radar is the damage path of the 8 May 2003 OKC tornado.

The ARPS system was used as the prediction model. The parameterization schemes and vertical boundary conditions used in Section 3 were adopted for the present experiments. The initial first guess and the lateral boundaries were provided by a 9 km data assimilation experiment. This 9 km experiment was done in the same way as in Hu and Xue [39]. It assimilates rawinsonde data and wind profiler data every hour for a total of six hours. The Eta model analysis and forecast provide the background and lateral boundaries for the 9 km experiment.

Similar to before, four experiments (Table 1), that is, NODP2, DP2, DP2d5, and DP2m5, were conducted in order to examine the impact of DPEC and the sensitivity of DPEC to different weighting coefficients. The assimilation experiments start at 2100 UTC and assimilate radial velocity data and radar reflectivity data every 5 minutes in a cycled manner similar to the procedure described in Section 3.

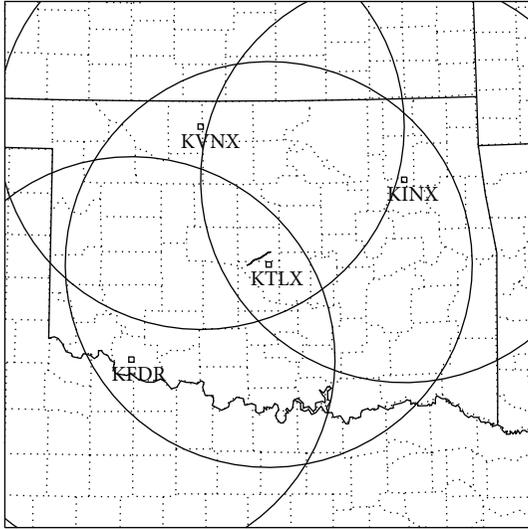


FIGURE 7: The model domain with county boundaries for the 8 May 2003 OKC tornadic supercell thunderstorm case. The five radars as well as their 230 km range circles are also shown. The outline near KTLX radar illustrates the damage path of the 8 May 2003 OKC tornado.

The assimilation window was 1 h long and the final analysis was at 2200 UTC. From the final analysis, a 1 h free forecast (2200 UTC–2300 UTC) was made.

Figures 6(e)–6(l) show the reflectivity, horizontal wind vector, and vertical vorticity at $z = 2$ km MSL from 2200 UTC to 2300 UTC every 20 minutes for the NODP2 and DP2 experiments. After 1 h of data assimilation, the storm had been successfully spun up (Figures 6(e) and 6(i) versus Figure 6(a)). The area of strong vertical vorticity is located where a hook echo is observed. The storm then moves east-northeastward. The direction and speed of the predicted storm are very close to what was observed (Figures 6(f), 6(g), 6(h) and 6(j), 6(k), and 6(l) versus 6(b), 6(c), and 6(d)). Therefore, both NODP2 and DP2 performed well in predicting the general evolution of the storm. DP2d5 and DP2m5 made similar forecasts to DP2 (not shown).

Comparing Figures 6(j), 6(k), and 6(l) (for DP2) with Figures 6(f), 6(g), and 6(h) (for NODP2), we can see that DP2 predicted larger low-level vertical vorticity than NODP2. This is further confirmed by examining the evolution of low-level vertical vorticity. Figure 8 shows the time series of the maximum vertical vorticity below two kilometers every one minute from 2200 UTC to 2300 UTC for all four experiments. It is illustrated in Figure 8 that after 22:17 UTC and until the end of the forecast, the low-level maximum vertical vorticity from the experiments applying DPEC (the red, blue, and green lines) was generally larger than that from the NODP2 experiment (the black line). As mentioned before, larger low-level vertical vorticity corresponds to a better-defined mesocyclone vortex, which is stronger and deeper. As an example, Figure 9 presents the vertical vorticity at the vertical cross-section through the center of the storm at $y = 301.5$ km at 2220 UTC, which is during the tornado touchdown period. The experiments using DPEC

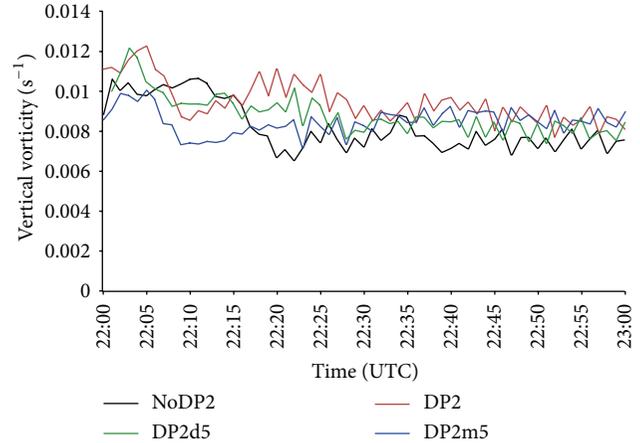


FIGURE 8: The time series of maximum vertical vorticity below two kilometers from 2200 UTC to 2300 UTC 8 May 2003 every one minute. The horizontal axis shows the time in UTC; the vertical axis shows the vertical vorticity values in units of s^{-1} . The black line is for experiment NoDP2, the red line for experiment DP2, the blue line for DP2m5, and the green line for DP2d5.

(Figures 9(b), 9(c), and 9(d)) predict a deeper column of high vertical vorticity ($>0.008 s^{-1}$), extending from as low as 1.0 kilometers to as high as 8.5 kilometers. The region of high vertical vorticity ($>0.008 s^{-1}$) predicted by the NODP2 experiment is mainly in the middle part of the atmosphere, roughly from 2.4 kilometers to 6.0 kilometers. Therefore, it can be concluded that the use of DPEC helps make a better forecast of low-level mesocyclone rotation.

5. Summary and Conclusions

A diagnostic pressure equation was added into the ARPS 3DVAR system as a weak constraint with the goal of coupling the dynamic and thermodynamic variables so as to improve the analysis of convective storms and their subsequent forecast. The updated ARPS 3DVAR system was tested using OSSEs in Ge et al.'s. [7] and applied to a real tornadic supercell case where only radial velocity has been assimilated. This study further applied the newly updated ARPS 3DVAR system to tornadic supercell thunderstorm studies by assimilating both the radial velocity and radar reflectivity data.

For both the 5 May 2007 Greensburg tornadic supercell storm case and the 8 May 2003 Oklahoma City tornadic supercell storm case, four data assimilation experiments were conducted with three of them using different DP weighting coefficients and the other one without DPEC. The four experiments assimilated the same amount of observations from multiple Doppler radars and imposed the mass continuity equation constraint.

After 1 h of intermittent data assimilation, it was found that DPEC did affect the final analysis. However, since there is no reliable high-resolution analysis of the storm, it is not easy to tell directly which analysis is better. The evaluation of

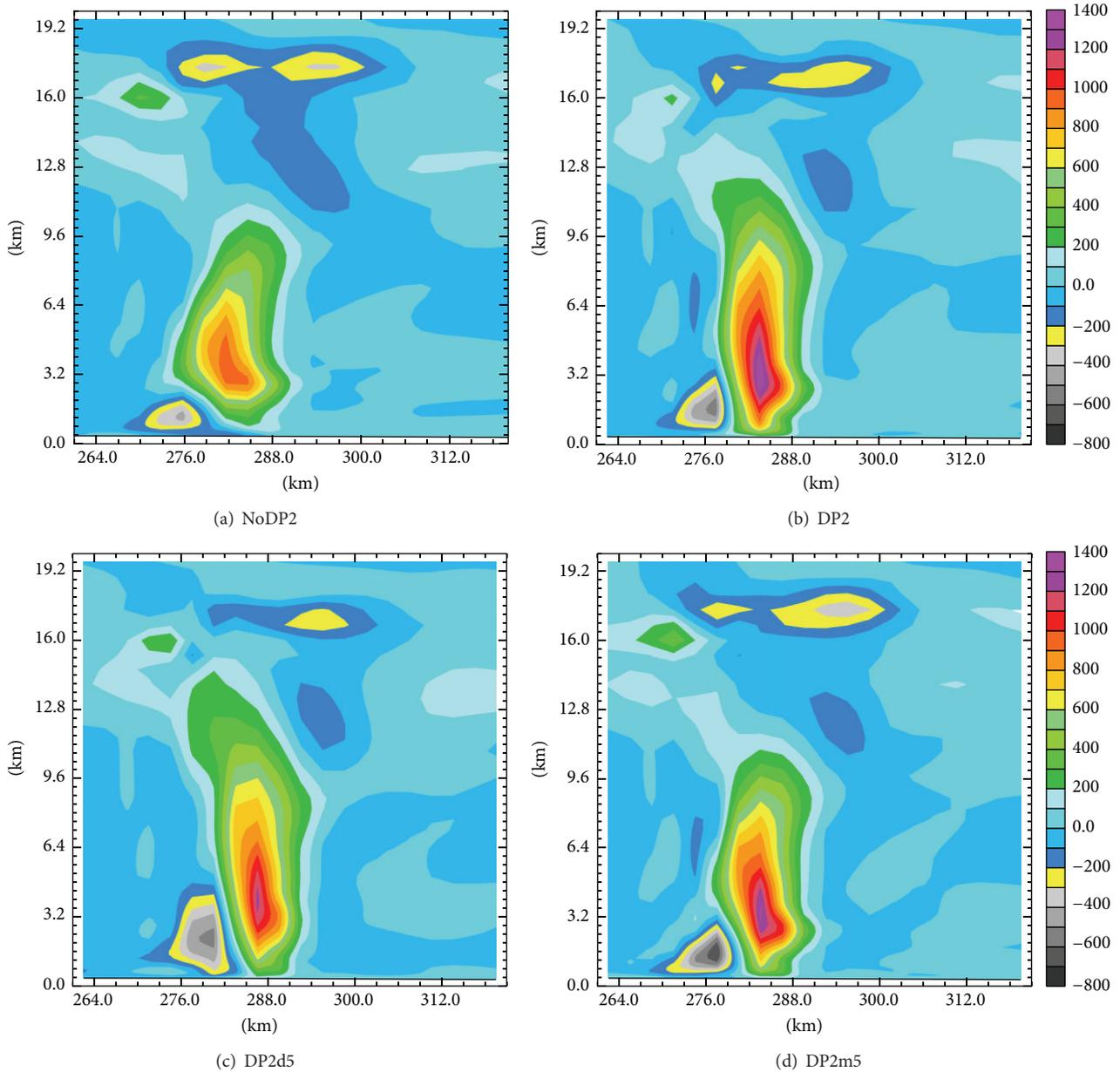


FIGURE 9: The vertical vorticity (in units of 10^{-5} s^{-1}) at the vertical cross section through the center of the major storm at $y = 301.5 \text{ km}$ at 2220 UTC 8 May 2003 for the (a) NoDP2, (b) DP2, (c) DP2d5, and (d) DP2m5.

the benefit of DPEC to radar data assimilation in these real cases is examined mainly based on the ensuing forecasts.

It was demonstrated that the experiments using DPEC generally predict larger low-level vertical vorticity than the experiments not using DPEC. Therefore, it is concluded that the use of DPEC improves the forecast of supercell mesocyclone rotation of the major thunderstorm. The experiments using different weighting coefficients generated similar results. This suggests that DPEC is not very sensitive to the weighting coefficients, although very small values should still be avoided as found in Ge et al's. [7].

Overall, the addition of DPEC in the ARPS 3DVAR system had a positive impact on storm-scale 3DVAR data

assimilation of Doppler radar data and on the subsequent forecast. In the future, the system needs to be tested with more real data cases, including tornadic supercell thunderstorms and other storm-scale phenomena, to further demonstrate the robustness of these conclusions.

Acknowledgments

This work was supported by NSF Grants ATM-0331756, ATM-0738370, NSF ATM-0530814, NSF AGS-0802888, and NOAA's Warn-on-Forecast Project. The third author was also supported by NSF OCI-0905040, AGS-0941491, and AGS-1046171. Robin Tanamachi kindly provided manually

dealised radial velocity dataset of the Dodge City, Kansas (KDDC) radar. Computations were performed at the Pittsburgh Supercomputing Center (PSC) and Oklahoma Supercomputing Center for Research and Education (OSCER). The authors appreciate two anonymous reviewers and the Editor Louis Wicker for their very thoughtful comments which helped significantly to improve the quality of this paper from its original form. The authors also thank Jacob Carlin and Edward Natenberg for proofreading the paper.

References

- [1] A. Crook and J. D. Tuttle, "Numerical simulations initialized with radar-derived winds—part II: forecasts of three gust-front cases," *Monthly Weather Review*, vol. 122, no. 6, pp. 1204–1217, 1994.
- [2] J. Sun and N. A. Crook, "Comparison of thermodynamic retrieval by the adjoint method with the traditional retrieval method," *Monthly Weather Review*, vol. 124, no. 2, pp. 308–324, 1996.
- [3] J. Sun and N. A. Crook, "Real-time low-level wind and temperature analysis using single WSR-88D data," *Weather and Forecasting*, vol. 16, no. 1, pp. 117–132, 2001.
- [4] M. Hu, M. Xue, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms—part I: cloud analysis and its impact," *Monthly Weather Review*, vol. 134, no. 2, pp. 675–698, 2006.
- [5] M. Hu, M. Xue, J. Gao, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms—part II: impact of radial velocity analysis via 3DVAR," *Monthly Weather Review*, vol. 134, no. 2, pp. 699–721, 2006.
- [6] D. J. Stensrud and J. Gao, "Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short-range forecasts," *Monthly Weather Review*, vol. 138, no. 4, pp. 1250–1272, 2010.
- [7] G. Ge, J. Gao, and M. Xue, "Diagnostic pressure equation as a weak constraint in a storm-scale three-dimensional variational radar data assimilation system," *Journal of Atmospheric and Oceanic Technology*, vol. 29, pp. 1075–1092, 2012.
- [8] J. Sun and N. A. Crook, "Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint—part I: model development and simulated data experiments," *Journal of the Atmospheric Sciences*, vol. 54, no. 12, pp. 1642–1661, 1997.
- [9] J. Sun and N. A. Crook, "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," *Journal of the Atmospheric Sciences*, vol. 55, no. 5, pp. 835–852, 1998.
- [10] J. Sun, "Initialization and numerical forecasting of a supercell storm observed during STEPS," *Monthly Weather Review*, vol. 133, no. 4, pp. 793–813, 2005.
- [11] C. Snyder and F. Zhang, "Assimilation of simulated Doppler radar observations with an ensemble Kalman filter," *Monthly Weather Review*, vol. 131, no. 8, pp. 1663–1677, 2003.
- [12] F. Zhang, C. Snyder, and J. Sun, "Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filter," *Monthly Weather Review*, vol. 132, pp. 1238–1253, 2004.
- [13] A. Caya, J. Sun, and C. Snyder, "A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation," *Monthly Weather Review*, vol. 133, no. 11, pp. 3081–3094, 2005.
- [14] M. Tong and M. Xue, "Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments," *Monthly Weather Review*, vol. 133, no. 7, pp. 1789–1807, 2005.
- [15] M. Xue, M. Tong, and K. K. Droegemeier, "An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 1, pp. 46–66, 2006.
- [16] A. Aksoy, D. C. Dowell, and C. Snyder, "A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations—part I: storm-scale analyses," *Monthly Weather Review*, vol. 137, no. 6, pp. 1805–1824, 2009.
- [17] F. Zhang, Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, "Cloud-resolving hurricane initialization and prediction through assimilation of doppler radar observations with an ensemble Kalman filter," *Monthly Weather Review*, vol. 137, no. 7, pp. 2105–2125, 2009.
- [18] A. Aksoy, D. C. Dowell, and C. Snyder, "A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations—part II: short-range ensemble forecasts," *Monthly Weather Review*, vol. 138, no. 4, pp. 1273–1292, 2010.
- [19] D. C. Dowell, L. J. Wicker, and C. Snyder, "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*, vol. 139, no. 1, pp. 272–294, 2011.
- [20] N. Snook, M. Xue, and Y. Jung, "Analysis of a tornadic mesoscale convective vortex based on ensemble kalman filter assimilation of CASA X-band and WSR-88D radar data," *Monthly Weather Review*, vol. 139, no. 11, pp. 3446–3468, 2011.
- [21] D. T. Dawson, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, "Impact of the environmental low-level wind profile on ensemble forecasts of the 4 may 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones," *Monthly Weather Review*, vol. 140, no. 2, pp. 696–716, 2012.
- [22] J. Marquis, Y. Richardson, P. Markowski, D. Dowell, and J. Wurman, "Tornado maintenance investigated with high-resolution dual-doppler and EnKF analysis," *Monthly Weather Review*, vol. 140, no. 1, pp. 3–27, 2012.
- [23] R. L. Tanamachi, L. J. Wicker, D. C. Dowell, H. B. Bluestein, D. T. Dawson, and M. Xue, "EnKF assimilation of high-resolution, mobile Doppler radar data of the 4 May 2007 Greensburg, Kansas, supercell into a numerical cloud model," *Monthly Weather Review*, vol. 141, pp. 625–648, 2012.
- [24] C. K. Potvin and L. J. Wicker, "Assessing ensemble forecasts of low-level supercell rotation within an OSSE framework," *Weather and Forecasting*, 2013.
- [25] M. Hu and M. Xue, "Analysis and prediction of 8 May 2003 Oklahoma City tornadic thunderstorm and embedded tornado using ARPS with assimilation of WSR-88D radar data," in *Proceedings of the 22nd Conference on Weather Analysis and Forecasting/18th Conference on Numerical Weather Prediction*, American Meteorological Society, Salt Lake City, Utah, USA, 2007.

- [26] M. Xue, K. K. Droegemeier, and V. Wong, "The Advanced Regional Prediction System (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction model—part I: model dynamics and verification," *Meteorology and Atmospheric Physics*, vol. 75, no. 3-4, pp. 161–193, 2000.
- [27] M. Xue, K. K. Droegemeier, V. Wong et al., "The Advanced Regional Prediction System (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction tool—part II: model physics and applications," *Meteorology and Atmospheric Physics*, vol. 76, no. 3-4, pp. 143–165, 2001.
- [28] M. Xue, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, "The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation," *Meteorology and Atmospheric Physics*, vol. 82, no. 1–4, pp. 139–170, 2003.
- [29] M. Xue, "CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment," in *Proceedings of the 24th Conference on Severe Local Storms*, American Meteorological Society, Savannah, Ga, USA, 2008.
- [30] M. Xue, "CAPS realtime storm scale ensemble and high resolution forecasts for the NOAA hazardous weather testbed," in *Proceedings of the 25th Conference on Severe Local Storms*, American Meteorological Society, Seattle, Wash, USA, 2011.
- [31] T. M. Hamill and C. Snyder, "A hybrid ensemble Kalman filter-3D variational analysis scheme," *Monthly Weather Review*, vol. 128, no. 8, pp. 2905–2919, 2000.
- [32] A. C. Lorenc, "The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var," *Quarterly Journal of the Royal Meteorological Society*, vol. 129, no. 595, pp. 3183–3203, 2003.
- [33] M. Buehner, "Ensemble-derived stationary and flow-dependent background-error covariances: evaluation in a quasi-operational NWP setting," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 607, pp. 1013–1043, 2005.
- [34] X. Wang, D. M. Barker, C. Snyder, and T. M. Hamill, "A hybrid ETKF-3DVAR data assimilation scheme for the WRF model—part II: real observing experiments," *Monthly Weather Review*, vol. 136, no. 12, pp. 5132–5147, 2008.
- [35] X. Wang, D. M. Barker, C. Snyder, and T. M. Hamill, "A hybrid ETKF-3DVAR data assimilation scheme for the WRF model—part I: observation system simulation experiment," *Monthly Weather Review*, vol. 136, no. 12, pp. 5116–5131, 2008.
- [36] J. Gao, M. Xue, A. Shapiro, and K. K. Droegemeier, "A variational method for the analysis of three-dimensional wind fields from two Doppler radars," *Monthly Weather Review*, vol. 127, no. 9, pp. 2128–2142, 1999.
- [37] J. Gao, M. Xue, A. Shapiro, Q. Xu, and K. K. Droegemeier, "Three-dimensional simple adjoint velocity retrievals from single-Doppler radar," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 1, pp. 26–38, 2001.
- [38] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [39] M. Hu and M. Xue, "Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case," *Monthly Weather Review*, vol. 135, no. 2, pp. 507–525, 2007.
- [40] K. Ide, P. Courtier, M. Ghil, and A. C. Lorenc, "Unified notation for data assimilation: operational, sequential and variational," *Journal of the Meteorological Society of Japan*, vol. 75, no. 1 B, pp. 181–189, 1997.
- [41] R. J. Purser, W.-S. Wu, D. F. Parrish, and N. M. Roberts, "Numerical aspects of the application of recursive filters to variational statistical analysis—part I: spatially homogeneous and isotropic Gaussian covariances," *Monthly Weather Review*, vol. 131, pp. 1524–1535, 2003.
- [42] R. J. Purser, W.-S. Wu, D. F. Parrish, and N. M. Roberts, "Numerical aspects of the application of recursive filters to variational statistical analysis—part II: spatially inhomogeneous and anisotropic general covariances," *Monthly Weather Review*, vol. 131, no. 8, pp. 1536–1548, 2003.
- [43] S. C. Albers, J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, "The local analysis and prediction system (LAPS): analyses of clouds, precipitation, and temperature," *Weather and Forecasting*, vol. 11, no. 3, pp. 273–287, 1996.
- [44] J. Zhang, F. Carr, and K. Brewster, "ADAS cloud analysis," in *Proceedings of the Conference on Numerical Weather Prediction*, pp. 185–188, American Meteorological Society, Phoenix, Ariz, USA, 1998.
- [45] K. Brewster, "Recent advances in the diabatic initialization of a non-hydrostatic numerical model," in *Proceedings of the 15th Conference on Numerical Weather Prediction and 21st Conference on Severe Local Storms*, American Meteorological Society, San Antonio, Tex, USA, 2002.
- [46] D. McCarthy, L. Ruthi, and J. Hutton, "The greensburg, KS tornado," in *Proceedings of the 22nd Conference on Weather Analysis and Forecasting/18th Conference on Numerical Weather Prediction*, American Meteorological Society, Park City, Utah, USA, 2007.
- [47] H. B. Bluestein, "The formation and early evolution of the Greensburg, Kansas, tornadic supercell on 4 May 2007," *Weather and Forecasting*, vol. 24, no. 4, pp. 899–920, 2009.
- [48] Y.-L. Lin, R. D. Farley, and H. D. Orville, "Bulk parameterization of the snow field in a cloud model," *Journal of Climate & Applied Meteorology*, vol. 22, no. 6, pp. 1065–1092, 1983.
- [49] G. S. Romine, D. W. Burgess, and R. B. Wilhelmson, "A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma City Area tornadic supercell," *Monthly Weather Review*, vol. 136, no. 8, pp. 2849–2870, 2008.

Research Article

Retrieving 3D Wind Field from Phased Array Radar Rapid Scans

Xiaobin Qiu,^{1,2} Qin Xu,³ Chongjian Qiu,² Kang Nai,⁴ and Pengfei Zhang^{3,4}

¹ Tianjin Institute of Meteorological Science, Tianjin 300074, China

² College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China

³ NOAA/OAR National Severe Storms Laboratory, David L. Boren Boulevard, Norman, OK 73072, USA

⁴ Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, OK 73072, USA

Correspondence should be addressed to Qin Xu; qin.xu@noaa.gov

Received 11 May 2013; Revised 1 July 2013; Accepted 19 July 2013

Academic Editor: Jidong Gao

Copyright © 2013 Xiaobin Qiu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The previous two-dimensional simple adjoint method for retrieving horizontal wind field from a time sequence of single-Doppler scans of reflectivity and/or radial velocity is further developed into a new method to retrieve both horizontal and vertical winds at high temporal and spatial resolutions. This new method performs two steps. First, the horizontal wind field is retrieved on the conical surface at each tilt (elevation angle) of radar scan. Second, the vertical velocity field is retrieved in a vertical cross-section along the radar beam with the horizontal velocity given from the first step. The method is applied to phased array radar (PAR) rapid scans of the storm winds and reflectivity in a strong microburst event and is shown to be able to retrieve the three-dimensional wind field around a targeted downdraft within the storm that subsequently produced a damaging microburst. The method is computationally very efficient and can be used for real-time applications with PAR rapid scans.

1. Introduction

Updrafts and downdrafts are the essential components of storms. Their strengths often determine the type and evolution stage of storms. Quickly detecting updrafts and downdrafts and estimating their strengths in storm wind fields will make timely and accurate assessments of hazardous weather conditions. It is thus desirable to develop an efficient method to retrieve both the horizontal and vertical winds, including updrafts and downdrafts, in real time from phased array radar (PAR) rapid scans of storms. A key advantage of PAR over Weather Surveillance Radar 1988-Doppler (WSR-88D) is the capability to rapidly and adaptively scan storms. With its agile electronic beam steering, the PAR scan strategy can be optimized on particular weather phenomena with the volume scan time reduced from minutes to seconds (Zrnich et al. [1], Torres et al. [2]). High spatial and temporal resolution volumetric radar data are often necessary to resolve very fine echo structures, their transient developments, and movements inside storms (Heinselman et al. [3]). Previous research also indicates that the retrieval errors can be reduced

if the reflectivity and radial-velocity fields are sampled more frequently (Qiu and Xu [4], Shapiro et al. [5]).

Since Doppler radar observations are limited mainly to reflectivity and radial-component velocity (along the radar beam) and there is no direct measurement of the remaining two wind components perpendicular to the radar beam, a two-dimensional simple adjoint (2D-SA) method was developed by Qiu and Xu [6] to retrieve the horizontal wind field from radar scans at low-elevation angles. In this 2D-SA method, a simplified reflectivity advection equation which is used to predict the reflectivity and the time-mean velocity that advects the reflectivity field in this equation is estimated by minimizing the difference between the radar observed and predicted reflectivity fields. The method was then refined and successfully tested with many real radar observations (Xu et al. [7, 8]). Built on the above success, a three-dimensional simple adjoint (3D-SA) method was developed by Xu et al. [9]. By using the full momentum equations and the mass continuity equation as weak constraints, this 3D-SA method can retrieve the three-dimensional wind field and the perturbation potential temperature field, similarly to

the four-dimensional adjoint (4DA) method with a complete system of dynamic and thermodynamic equations (Sun and Crook [10]). Computationally, this 3D-SA method is more efficient than the 4DA method but still too expensive to apply to real-time observations from PAR rapid scans. As the control variable dimension in the 3D-SA method is much larger than that in the 2D-SA method, the 3D-SA method is not only computationally more expensive but also less flexible to adapt to incomplete data coverage than the 2D-SA method. In view of the above limitations, the 3D-SA method has not been applied to PAR data. Instead, the 2D-SA method can be further developed into a two-step SA method in which the vertical velocity is retrieved in a selected vertical cross-section along the radar beam in the second step after the horizontal velocity is retrieved in the first step on the conical surface at each tilt of PAR scans in a targeted domain. The basic idea and formulations of this two-step SA method are described in the next section. The method is applied to PAR observations for a selected case in Section 3. The benefits of rapid scans and the usefulness of mesoscale background wind field are examined in Section 4. Conclusions follow in Section 5.

2. Description of the Method

2.1. Basic Idea. To reduce the computational cost, the horizontal and vertical winds will be retrieved separately in two steps. In the first step, the 2D-SA method is used to retrieve the horizontal winds on each conical surface of the radar scans in a targeted domain of convective scale, while the mesoscale background horizontal wind field is provided by the existing radar wind analysis system that was developed based on the statistic interpolation for real-time applications with the operational WSR-88D radars (Xu et al. [11]). The vertical velocity is then retrieved in the second step in the along-beam vertical cross-section that cuts through the concerned feature at the center of the targeted domain. Since the horizontal winds are retrieved in the first step and their related terms are known in the forecast equation in the second step, the control variable dimension is reduced in the second step.

2.2. Horizontal Wind Retrieval in the First Step. The 2D-SA method is used in the first step to estimate the incremental time-mean quasi-horizontal velocity $(\Delta u_m, \Delta v_m)$ with respect to the background time-mean quasi-horizontal velocity (u_b, v_b) , the time-mean source term S_m (that includes the effect of vertical advection), and the horizontal turbulent diffusivity coefficient k_h in the equation of quasi-horizontal advection of reflectivity over the time period of N consecutive volume scans. Here, the quasi-horizontal velocity is defined as the nearly horizontal component of the three-dimensional vector velocity projected onto the conical surface of radar scan, and the quasi-horizontal advection is the advection produced by the quasi-horizontal velocity on the conical surface of radar scan. In this first step, $(\Delta u_m, \Delta v_m, S_m, k_h)$ are estimated by minimizing the following cost function:

$$J = J_\eta + J_{vr} + J_d + J_v + J_b. \quad (1)$$

The first term in (1) measures the difference between the predicted reflectivity η and the observed reflectivity η_{ob} , and this term is given by

$$J_\eta = \int_\tau \iint_{D_h} W_\eta (\eta - \eta_{ob})^2 dx dy dt, \quad (2)$$

where D_h is the horizontal area on the conical surface of radar scan in the targeted retrieval domain, W_η is the weight, $\tau = (N - 1)\Delta\tau$ is the time period covering the N sequential volume scans, and $\Delta\tau$ is the time elapsed for each volume scan. Here, η is predicted by the following quasi-horizontal advection equation:

$$\partial_t \eta + u_m \partial_x \eta + v_m \partial_y \eta - k_h (\partial_x^2 \eta + \partial_y^2 \eta) = S_m, \quad (3a)$$

$$\text{with } \eta(t, x, y) = \eta_{ob}(t, x, y) \text{ at the boundary of } D_h, \quad (3b)$$

$$\eta(0, x, y) = \eta_{ob}(0, x, y) \text{ at the initial time,} \quad (3c)$$

where $(u_m, v_m) = (u_b + \Delta u_m, v_b + \Delta v_m)$ is the estimated time-mean quasi-horizontal velocity (including the projection of hydrometeors' terminal velocity onto the radar beam).

The second term in (1) measures the difference between the estimated time-mean radial velocity v_{rm} and observed radial velocity v_{rob} . This term is given by

$$J_{vr} = \int_\tau \iint_{D_h} W_{vr} (v_{rm} - v_{rob})^2 dx dy dt, \quad (4)$$

where W_{vr} is the weight, v_{rob} is the radar observed radial velocity, $v_{rm} = u_m \sin \phi + v_m \cos \phi$ is the along-beam radial component of (u_m, v_m) , and ϕ is the azimuthal angle of the radar beam (positive for clockwise rotation from the y -coordinate pointing to the north) at the observation point in D_h (on the conical surface).

The third and fourth terms in (1) are given by

$$J_d = \iint_{D_h} W_d (\text{Div})^2 dx dy, \quad (5)$$

$$J_v = \iint_{D_h} W_v (\text{Vor})^2 dx dy, \quad (6)$$

respectively, where $\text{Div} = \partial_x \Delta u_m + \partial_y \Delta v_m$ and $\text{Vor} = \partial_x \Delta v_m - \partial_y \Delta u_m$. These terms impose weak constraints on the divergence and vorticity of the incremental time-mean velocity $(\Delta u_m, \Delta v_m)$ to suppress spurious divergence and vorticity caused by data noises in the same way as in Xu et al. [7]. Their associated differential operators enhance the background error correlations, according to Xu [12], in addition to the Gaussian correlations used for (ψ_m, χ_m) in (7).

The last term in (1) is the background term for $(\Delta u_m, \Delta v_m)$ and S_m . This term is given by

$$J_b = \iint_{D_h} (W_\psi \psi_c^2 + W_\chi \chi_c^2 + W_S S_c^2) dx dy, \quad (7)$$

where W_ψ , W_χ , and W_S are the weights (given by the inverses of the background error variances associated with ψ_m , χ_m , and S_m , resp.), (ψ_c, χ_c, S_c) are the control variables related to (ψ_m, χ_m, S_m) by $(G_\psi * \psi_c, G_\chi * \chi_c, G_S * S_c) = (\psi_m, \chi_m, S_m)$, ψ_m and χ_m are the stream function and velocity potential defined

by $\Delta u_m = -\partial\psi_m/\partial y + \partial\chi_m/\partial x$ and $\Delta v_m = \partial\psi_m/\partial x + \partial\chi_m/\partial y$, ($G_\psi * G_\psi, G_\chi * G_\chi, G_S * G_S$) are the background error correlation functions with given decorrelation lengths (L_ψ, L_χ, L_S) for (ψ_m, χ_m, S_m) , and $*$ denotes the spatial convolution between the two functions (on the two sides of $*$) over D_h . Numerically, these convolutions are computed by a recursive filter (Purser et al. [13, 14]). Their discrete formulations (in matrix forms) are similar to those in (8a), (8b), (8c), and (8d) of Xu et al. [15]. The time integrations for J_η in (2) and J_{vr} in (4) are computed by summing their respective integrands at each time step Δt over the time period τ , where Δt is the time step used for the numerical integration of (3a) with (3b)-(3c). The spatial integrations for the cost-function terms in (2) and (4)-(7) are computed by summing their respective integrands at each grid point over the horizontal area D_h .

The gradients of the first cost-function term J_η in (2) with respect to the control variables $(\psi_c, \chi_c, S_c, k_h)$ are computed from their gradients with respect to $(\Delta u_m, \Delta v_m, S_m, k_h)$ by using the recursive filter and related transformations, while the latter gradients are given by

$$\frac{\partial J_\eta}{\partial \Delta u_m} = \int_\tau \eta^* \partial_x \eta dt, \quad (8a)$$

$$\frac{\partial J_\eta}{\partial \Delta v_m} = \int_\tau \eta^* \partial_y \eta dt, \quad (8b)$$

$$\frac{\partial J_\eta}{\partial S_m} = - \int_\tau \eta^* dt, \quad (8c)$$

$$\frac{\partial J_\eta}{\partial k_h} = - \int_\tau \iint_{D_h} \eta^* (\partial_x^2 \eta + \partial_y^2 \eta) dx dy dt, \quad (8d)$$

where η^* is the adjoint variable obtained by integrating (backward in time) the following adjoint equation:

$$-\partial_t \eta^* - \partial_x (u_m \eta^*) - \partial_y (v_m \eta^*) - k_h (\partial_x^2 \eta^* + \partial_y^2 \eta^*) = -2W_\eta (\eta - \eta_{ob}), \quad (9a)$$

$$\text{with } \eta^*(t, x, y) = 0 \text{ at the boundary of } D_h, \quad (9b)$$

$$\text{and } \eta^*(\tau, x, y) = 0 \text{ at the final time.} \quad (9c)$$

The derivations of (8a)-(9c) follow (2.5)-(2.13) of Qiu and Xu [6].

The gradients of the subsequent three cost-function terms in (4)-(6) with respect to the control variables (S_c, k_h) are obviously zero. Their gradients with respect to the control variables (ψ_c, χ_c) are computed from their gradients with respect to $(\Delta u_m, \Delta v_m)$ by using the recursive filter and related transformations, while the latter gradients are given by

$$\frac{\partial J_{vr}}{\partial \Delta u_m} = \int_\tau 2 \sin \phi W_{vr} (v_{rm} - v_{rob}) dt, \quad (10a)$$

$$\frac{\partial J_{vr}}{\partial \Delta v_m} = \int_\tau 2 \cos \phi W_{vr} (v_{rm} - v_{rob}) dt, \quad (10b)$$

$$\frac{\partial (J_d + J_v)}{\partial \Delta u_m} = \frac{-2W_d \partial \text{Div}}{\partial x} + \frac{2W_v \partial \text{Vor}}{\partial y}, \quad (10c)$$

$$\frac{\partial (J_d + J_v)}{\partial \Delta v_m} = \frac{-2W_d \partial \text{Div}}{\partial y} - \frac{2W_v \partial \text{Vor}}{\partial x}. \quad (10d)$$

The gradient of the background term in (7) with respect to k_h is zero. Its gradients with respect to the control variables (ψ_c, χ_c, S_c) are derived directly from (7) in the following forms:

$$\frac{\partial J_b}{\partial \psi_c} = 2W_\psi \psi_c, \quad (11a)$$

$$\frac{\partial J_b}{\partial \chi_c} = 2W_\chi \chi_c, \quad (11b)$$

$$\frac{\partial J_b}{\partial S_c} = 2W_S S_c. \quad (11c)$$

The standard conjugate-gradient descending algorithm is used with the above computed gradients to minimize the cost-function in (1).

2.3. Vertical Wind Retrieval in the Second Step. The second step retrieves the time-mean vertical crossbeam velocity component w_m (that includes the projection of hydrometeors' terminal velocity but with zero background vertical velocity) in the selected vertical cross-section (x, z) along the radar beam in the targeted domain. In addition to w_m , the control variables also include a time-mean source term S_{m2} and the vertical turbulent diffusivity coefficient k_z in the reflectivity advection equation. The cost function is formulated similarly to that in (1), but there is no J_{vr} term since the observed radial velocity v_{rob} has zero projection on the crossbeam velocity component w_m . Thus, the cost function consists of only four terms, that is,

$$J = J_\eta + J_d + J_v + J_b. \quad (12)$$

These four terms have similar forms as those in (2) and (4)-(7), but the spatial integrations in these terms are over the vertical cross-section D_z , instead of D_h in the retrieval domain. In addition to this difference, there are several other differences as described below.

The first cost-function term J_η in (12) has the same form as that in (2) except that the spatial integration is over D_z and η is predicted by the following advection equation:

$$\partial_t \eta + u_{m2} \partial_x \eta + v_m \partial_y \eta + w_{m2} \partial_z \eta - k_h (\partial_x^2 \eta + \partial_y^2 \eta) - k_z \partial_z^2 \eta = S_{m2}, \quad (13a)$$

$$\text{with } \eta(t, x, z) = \eta_{ob}(t, x, z) \text{ at the boundary of } D_z, \quad (13b)$$

$$\text{and } \eta(0, x, z) = \eta_{ob}(0, x, z) \text{ at the initial time,} \quad (13c)$$

where $u_{m2} = u_m \cos \theta - w_m \sin \theta$ is the horizontal time-mean velocity component along the azimuth of radar beam in the selected (x, z) cross-section, $w_{m2} = u_m \sin \theta + w_m \cos \theta$ is the vertical time-mean velocity (including the hydrometeors' terminal velocity), and θ is the slope angle of the radar beam (see Figure 1). Here, (u_m, v_m) is the previously defined time-mean quasi-horizontal velocity vector (projected onto the conical surface of radar scan), while w_m is the time-mean quasi-vertical velocity component that is perpendicular to (u_m, v_m) in the three-dimensional space. Since the vertical

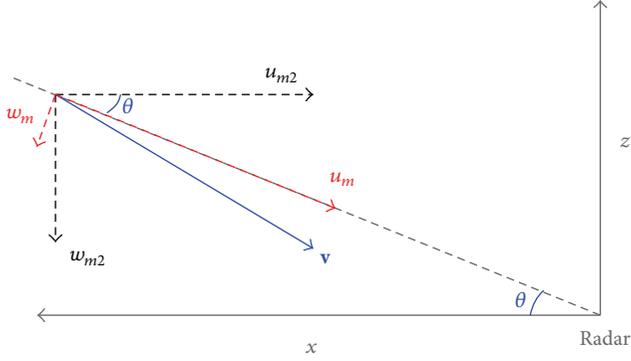


FIGURE 1: Sketch of geometric relationship between the components of (u_{m2}, w_{m2}) and (u_m, v_m) , plotted by the dashed black and red arrows, respectively, for the same vector \mathbf{v} plotted by the solid blue arrow. The radar beam is plotted by the gray dashed line, and θ is the slope angle of the radar beam.

cross-section (x, z) is along the radar beam, u_m is the along-beam radial component, and hence $u_m = v_{rm}$, v_m is the cross-beam component perpendicular to the vertical cross-section and hence is exactly horizontal. The w_m -component is also perpendicular to the radar beam but is confined within (rather than perpendicular to) the vertical cross-section. Note that (u_m, v_m) and k_h are already retrieved in the first step while $\partial_x \eta$, $\partial_y \eta$, $\partial_x^2 \eta$, and $\partial_y^2 \eta$ can be computed directly from the observed reflectivity field, so all their related terms, that is, $u_m \cos \theta \partial_x \eta + v_m \partial_y \eta + u_m \sin \theta \partial_z \eta - k_h (\partial_x^2 \eta + \partial_y^2 \eta)$, can be treated as known and thus moved to the right-hand side of (13a). This treatment reduces the control variable dimension and increases the computational efficiency significantly.

The J_d term has the same form as that in (5) except that the spatial integration is over D_z , and Div is replaced by the mass continuity constraint:

$$\text{Div}_3 = \partial_x u_{m2} + \partial_y v_m + \partial_z w_m^a, \quad (14)$$

where $w_m^a = w_{m2} - w_T$ is the time-mean vertical velocity of the air motion and w_T (≤ 0) is the hydrometeors' terminal velocity estimated from the observed reflectivity η_{ob} by using the empirical formula of Kessler [16] as that in (5) of Xu et al. [15]. The J_v term has the same form as that in (6) except that that the spatial integration is over D_z , and Vor is replaced by

$$\text{Vor}_z = \partial_z u_{m2} - \partial_x w_m^a. \quad (15)$$

The background term J_b is now changed into the following form:

$$J_b = \iint_{D_z} (W_w w_c^2 + W_{S_2} S_{c2}^2) dx dz, \quad (16)$$

where W_w and W_{S_2} are the weights (given by the inverses of the background error variances associated with w_m and S_{m2} , resp.), (w_c, S_{c2}) are the final control variables related to (w_m, S_{m2}) by $(G_w * w_c, G_{S_2} * S_{c2}) = (w_m, S_{m2})$, (G_w, G_{S_2}) are the Gaussian correlation functions associated with the square roots of the background error covariance functions

with given decorrelation lengths (L_w, L_{S_2}) for (w_m, S_{m2}) , and $*$ denotes the spatial convolution between the two functions (on the two sides of $*$) over D_z . Again, like the constraints in (5)-(6), the constraints in (14)-(15) can also suppress the spurious divergence and vorticity caused by data noises, and their associated differential operators also enhance the background error correlation in addition to the Gaussian correlation used above for w_m .

The gradients of the first cost-function term J_η with respect to the control variables (w_c, S_{c2}, k_z) in this second step are computed from their gradients with respect to (w_m, S_{m2}, k_z) by using the recursive filter as described in the first step, while the latter gradients are given by

$$\frac{\partial J_\eta}{\partial w_m} = \int_\tau \eta^* (-\sin \theta \partial_x \eta + \cos \theta \partial_z \eta) dt, \quad (17a)$$

$$\frac{\partial J_\eta}{\partial S_m} = - \int_\tau \eta^* dt, \quad (17b)$$

$$\frac{\partial J_\eta}{\partial k_z} = - \int_\tau \iint_{D_z} \eta^* \partial_z^2 \eta dx dz dt, \quad (17c)$$

where η^* is the adjoint variable obtained by integrating the following adjoint equation:

$$\begin{aligned} -\partial_t \eta^* + \partial_x (w_m \sin \theta \eta^*) - \partial_z (w_m \cos \theta \eta^*) + k_z \partial_z^2 \eta^* \\ = -2W_\eta (\eta - \eta_{ob}), \end{aligned} \quad (18a)$$

$$\text{with } \eta^*(t, x, z) = 0 \text{ at the boundary of } D_z, \quad (18b)$$

$$\eta^*(\tau, x, z) = 0 \text{ at the final time.} \quad (18c)$$

The derivations of (17a)–(18c) are similar to those that lead to (8a)–(9c).

The gradients of $J_d + J_v$ with respect to the control variables (S_{c2}, k_z) are zero. Their gradient with respect to the control variable w_c is computed from their gradient with respect to w_m by using the recursive filter as described in the first step, while the latter is given by

$$\begin{aligned} \frac{\partial (J_d + J_v)}{\partial w_m} = -2W_d \left(\frac{\cos \theta \partial \text{Div}_3}{\partial z} - \frac{\sin \theta \partial \text{Div}_3}{\partial x} \right) \\ + 2W_v \left(\frac{\cos \theta \partial \text{Vor}_z}{\partial x} - \frac{\sin \theta \partial \text{Vor}_z}{\partial z} \right). \end{aligned} \quad (19)$$

The gradient of the background term J_b with respect to k_z is zero. Its gradients with respect to the control variables (w_c, S_{c2}) are derived directly from (16) in the following forms:

$$\frac{\partial J_b}{\partial w_c} = 2W_w w_c, \quad (20)$$

$$\frac{\partial J_b}{\partial S_{c2}} = 2W_{S_2} S_{c2}.$$

The standard conjugate-gradient descending algorithm is used again to minimize the cost-function in (12) in this second step. The initial guesses of (w_c, S_{c2}) are set to zero, and the algorithm converges in less than 100 iterations.

3. Application to PAR Data

3.1. PAR Observations and Mesoscale Background Wind Field. The multimission PAR, located at the National Weather Radar Testbed (NVRT), Norman, Oklahoma, is a research radar using a converted U.S. Navy SPY-1A phased array antenna. This PAR has essentially the same wavelength (9.4 cm in S band) and range resolution (250 m) as the WSR-88D radars, and it can mimic WSR-88D volume coverage patterns (VCPs) and collect data at similar pulse repetition intervals. The most significant difference between the PAR and the WSR-88D is that the phased array antenna forms each beam electronically by controlling the phases of transmit-receive elements and thus can scan storms rapidly and adaptively. With its rapid scan capability, the PAR captured the rapid evolution of a severe storm produced microburst at the 20 km radial range to the south-southwest during the early evening of July 10, 2006. The PAR applied a beam multiplexing scanning strategy to volume coverage pattern 12 (VCP12) for a 90° sector scan, so each volume scan contained up to 53 tilts (with the elevation angles from 0.51° to 41°) and is completed in just 34 seconds. Since the retrieval domain is small and not very close to the PAR, only 14 tilts (from 0.51° to 19.5°) are intercepted by the retrieval domain and will be used for the retrievals in this paper. The storm and its produced microburst were fully sampled in time and space by the PAR. In particular, the PAR detected a reflectivity core aloft between 19:40:21 and 19:42:20 UTC at about 6 km above ground level. This reflectivity core produced a strong downdraft maximized around 19:49:07 UTC as shown in Figure 2, and this downdraft generated a strong outflow near the surface sampled by the PAR from 19:50:15 to 19:58:07 UTC (Heinselman et al. [3]).

As mentioned in Section 2.1, the radar wind analysis system (Xu et al. [11]) is used to produce the mesoscale background horizontal wind field. For the storm case shown in Figure 2, this system is applied to radial-velocity data from the operational KTLX radar at the Oklahoma City Twin Lakes in combination with surface wind data from the Oklahoma Mesonet. The analysis domain is centered at the KTLX radar site. The domain size is $160 \times 160 \times 8 \text{ km}^3$ covered by a $81 \times 81 \times 32$ grid, the horizontal resolution is $\Delta x = \Delta y = 2 \text{ km}$, and the vertical resolution is $\Delta z = 0.25 \text{ km}$. The background wind fields produced at 19:40:24 UTC are shown by white arrows in Figure 3(a) at $z = 0.25 \text{ km}$ and Figure 3(b) at $z = 5 \text{ km}$ (above the KTLX radar site) superimposed on the KTLX reflectivity images at 0.5° and 3.3° tilts, respectively. The small red square in panel (b) of Figure 3 marks the nested domain used by the two-step SA method. As shown in Figure 3(a), the low-level background flow is divergent (or convergent) on the mesoscale to the southwest (or northeast) on the upstream (or downstream) side of the nested domain. This mesoscale divergence (or convergence) is consistent with the weakened (or enhanced) upper-level reflectivity in Figure 3(b), relative to the lower-level reflectivity in Figure 3(a), in the area to the southwest (or northeast) on the upstream (or downstream) side of the nested domain, because the weakened (or enhanced) upper-level reflectivity indicates that its associated convective-cell cluster had started to decay (or grow) and

thus produced lower-level divergence (or convergence). In the nested domain, the background wind field is smooth and nearly constant at each vertical level. Clearly, the mesoscale background wind field cannot resolve the small-scale flow structures associated with the convective cell in the nested domain. As we can see from Figure 3, the lower-level and upper-level reflectivity fields have about the same intensity in the nested domain, and this suggests that the convective cell in the nested domain was fully developed around 19:40:00 UTC. The small-scale flow structure associated the downdraft produced by the convective cell in the nested domain can be retrieved by the two-step SA method as shown in the next two subsections.

3.2. Experiment Design. The nested horizontal domain on the conical surface of each tilt of PAR scan is centered at the radial range of $r = 20 \text{ km}$ from the PAR (marked by the blue + sign in Figure 3) with the x -axis along the PAR beam at the 210.7° azimuth (as shown by the yellow dashed line in Figure 3). The domain size is $16 \times 16 \text{ km}^2$ (as shown by the white dashed boundary lines in the right panel of Figure 2(a)) covered by a 65×65 grid with a horizontal resolution of $\Delta x = \Delta y = 250 \text{ m}$. The x -axis of the vertical cross section is thus also along the 210.7° azimuth and centered at $r = 20 \text{ km}$ from the PAR, while the vertical domain size is $15 \times 7.5 \text{ km}^2$ (as shown by the white dashed boundary lines in the left panel of Figure 2(a)) covered by a 61×31 grid with $\Delta x = \Delta z = 250 \text{ m}$. The total of 26 volume scans (one volume every $\Delta\tau = 34\text{--}64 \text{ s}$) for the time period from 19:40:00 to 20:00:00 UTC are used by the two-step SA method to retrieve the wind fields over 24 time windows. Each time window contains $N = 3$ consecutive volume scans, so $\tau = (N - 1)\Delta\tau \approx 1 \text{ min}$. The time step is $\Delta t = 2.5 \text{ s}$ for the forward integrations of (3a), (3b), and (3c) and (13a), (13b), and (13c) and the backward integrations of (9a), (9b), and (9c) and (18a), (18b), and (18c).

The weights and decorrelation length scales used by the cost function in (1) are specified as follows:

$$\begin{aligned} W_\eta &= P(\sigma_\eta^2 \tau D_h)^{-1}, \\ W_{vr} &= (10^2 \sigma_v^2 \tau D_h)^{-1}, \\ W_\psi &= W_\chi = 10^2 (L^2 \sigma_v^2 D_h)^{-1} \quad \text{with } L = L_\psi = L_\chi = 8\Delta x, \\ W_d &= W_v = \frac{(10s^2)}{D_h}, \\ W_S &= (\sigma_S^2 D_h)^{-1} \quad \text{with } L_S = 4\Delta x, \end{aligned} \tag{22a}$$

where $P = 0.5[\tau/(t + \Delta t)]^{1/2}$ is the time-dependent factor for the weight W_η and is specified similarly to that in (3.2) of Xu and Qiu [17], σ_η is the reflectivity observation error standard deviation and is set to 1 dBZ as in Xu et al. [15], σ_v is the error standard deviation for the background wind field, and σ_S is the error standard deviation for the (zero) background reflectivity source field.

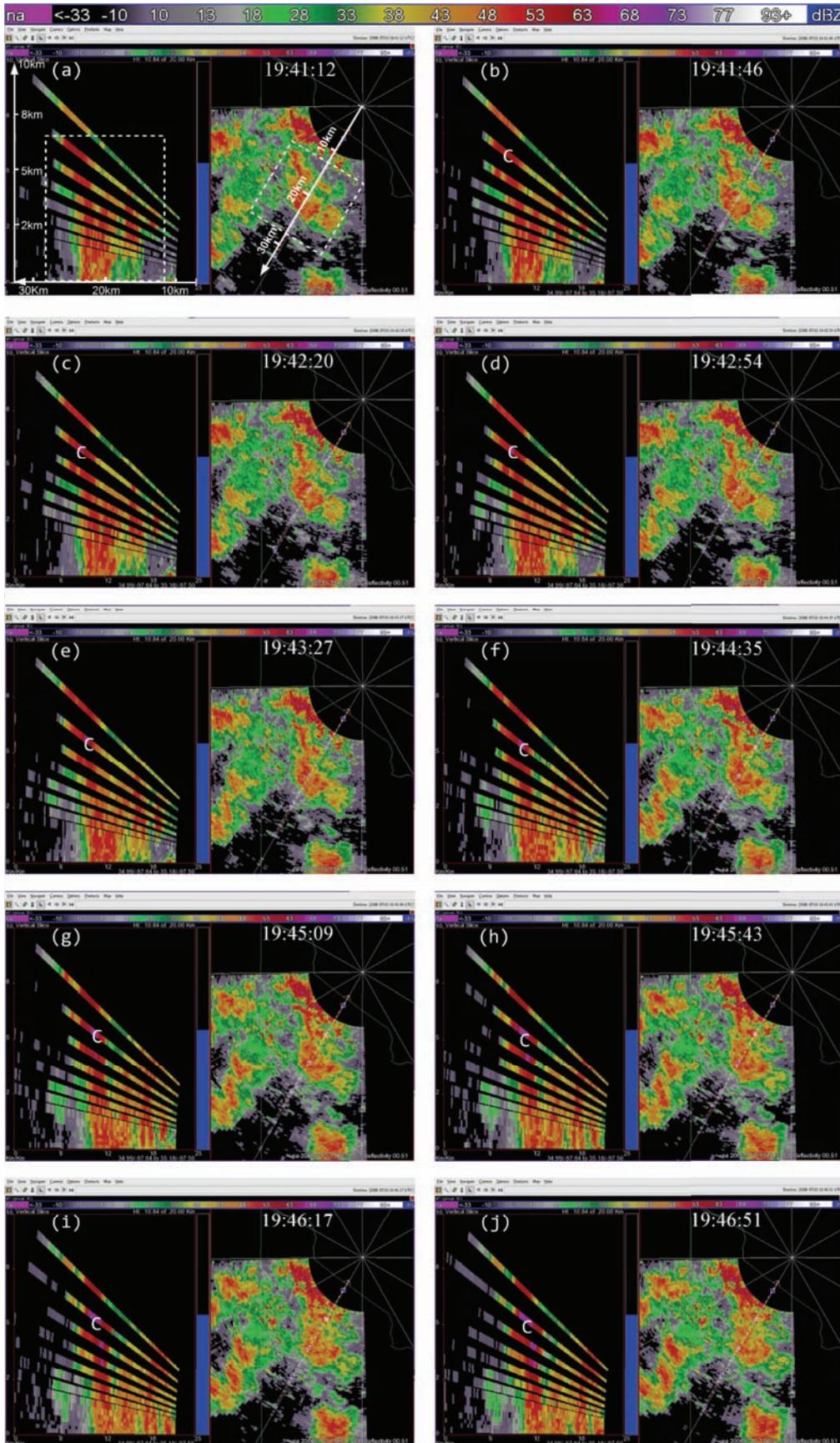


FIGURE 2: Continued.

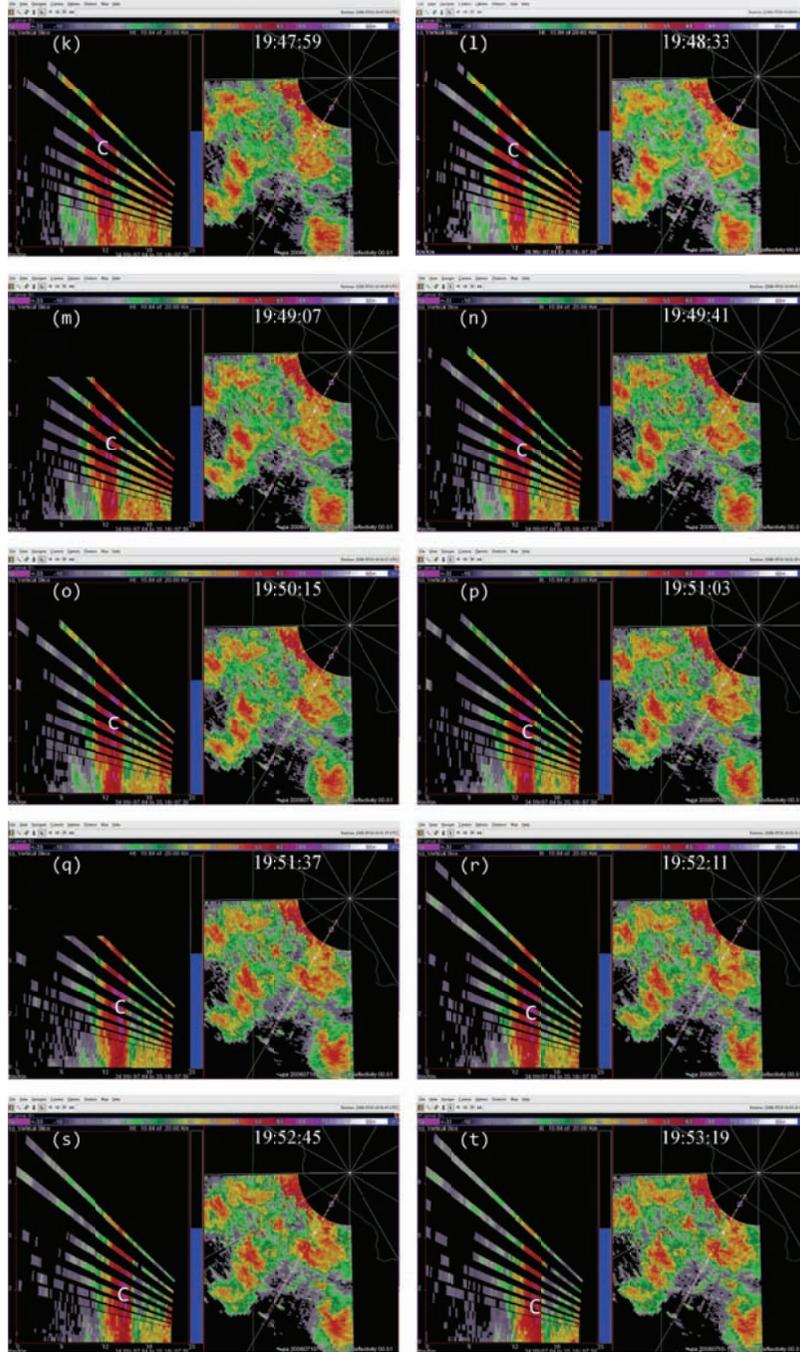


FIGURE 2: Reflectivity images from 20 consecutive PAR 90° sector scans of a severe microburst event from 19 : 41 : 12 to 19 : 53 : 19 UTC on July 10, 2006. Each panel consists of two subpanels: (i) the range height indicator (RHI) display along the 210.7° azimuth on the left side and (ii) the plan position indicator (PPI) display at 0.51° elevation on the right side. For the RHI display, the radial range is leftward from $r = 10$ to 35 km and the height range is upward from $z = 0$ to 10 km. The horizontal (or vertical) domain used in the first (or second) step of retrieval is shown by the white dashed boundary lines in the right (or left) subpanel of panel (a). The white letter C in each left subpanel marks the reflectivity core. As shown from panel (a) to (t), the reflectivity core falls from $z \approx 6$ km to 1.8 km in 667 s, so the estimated falling speed of the reflectivity core is $w_{m2} = (w_m^t + w_T) \approx (4.2 \text{ km}) / (667 \text{ s}) = 6.3 \text{ m s}^{-1}$, where $w_T (\leq 0)$ is the hydrometeors' terminal velocity as defined in (14).

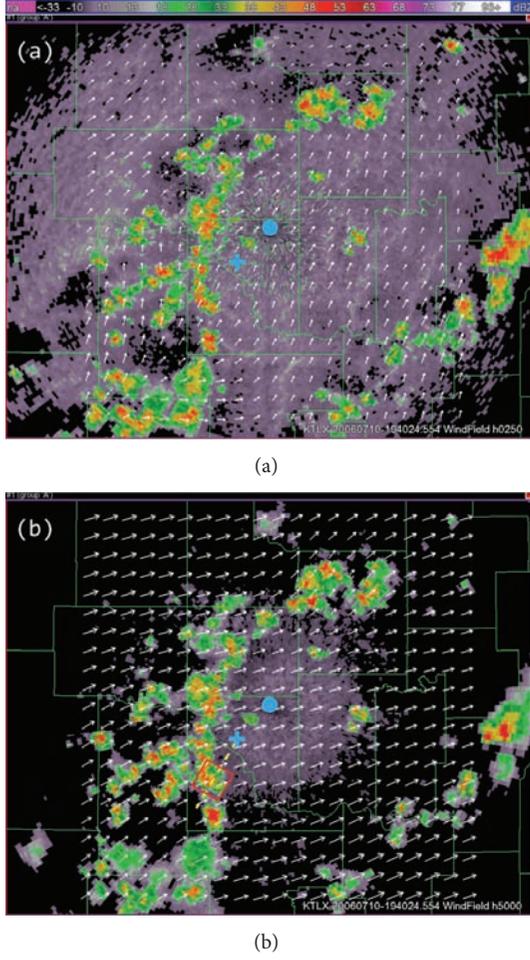


FIGURE 3: Background wind fields (shown by white arrows) produced by the radar wind analysis system at 19:40:42 UTC at (a) $z = 0.25$ km and (b) $z = 5$ km superimposed on the KTLX reflectivity images at 0.5° and 3.3° tilts, respectively. The domain size is $160 \times 160 \times 8$ km³ covered by an $81 \times 81 \times 32$ grid centered at the KTLX radar site (marked by the blue dot). The blue + sign marks the PAR site. In panel (b), the red square box marks the nested domain used by the two-step SA method, and the yellow dashed line shows the radar beam direction from PAR along the 210.7° azimuth.

The weights and decorrelation length scales used by the cost function in (12) are specified as follows:

$$\begin{aligned}
 W_\eta &= P(\sigma_\eta^2 \tau D_z)^{-1}, \\
 W_d &= W_v = \frac{(10s^2)}{D_z}, \\
 W_w &= 10^2(\sigma_v^2 D_z)^{-1} \quad \text{with } L_w = 8\Delta x, \\
 W_{S_2} &= (\sigma_S^2 D_z)^{-1} \quad \text{with } L_{S_2} = 4\Delta x,
 \end{aligned} \tag{22b}$$

where P , σ_η , σ_v , and σ_S are the same as defined in (22a). Note that the variance of the radial velocity innovation (observation minus background in the observation space),

denoted by σ_{vri}^2 , should be the sum of the radial-velocity background error variance and radial-velocity observation error variance, denoted by σ_{vro}^2 . This implies that σ_v can be estimated by $\sigma_v^2 \approx \sigma_{vri}^2 - \sigma_{vro}^2 \approx \sigma_{vri}^2$ if $\sigma_{vri}^2 \gg \sigma_{vro}^2$ ($\approx 1 \text{ m}^2 \text{ s}^{-2}$), while σ_{vri}^2 can be estimated by the spatially averaged RMS amplitude of the radial-velocity innovation. Similarly and loosely, σ_S^2 is estimated by the spatially averaged RMS amplitude of the time derivative of η_{ob} . With these estimates, the weights and decorrelation length scales are properly tuned to the above-specified values.

3.3. Results Obtained in the First Step. Figures 4(a)–4(f) show the retrieved (u_m, v_m) fields in the nested domain on the conical surfaces of six tilts (selected from the total 14 tilts from 0.51° to 19.5°) from the PAR 90° sector scans over the time period of 19:45:43–19:46:13 UTC. Note that the x -coordinate is along the 210.7° azimuth (as shown by the dashed yellow line in Figure 3(b)). Thus, as shown in Figure 4, the prevailing winds in the nested domain were southwesterly on the conical surface of $\theta = 0.51^\circ$ (around $z \approx 0.2$ km) but veered with height and gradually changed to westerly on the conical surface of $\theta = 15.6^\circ$ (around $z \approx 8$ km). This feature is consistent with the vertical variation of the background winds in Figure 3. In addition to this feature, the retrieved quasi-horizontal wind field exhibits strong storm-scale variations inside and around the main reflectivity core area (with reflectivity >40 dBZ) on each conical surface. The detailed storm-scale flow structures and associated horizontal divergence/convergence patterns are examined below.

From Figures 4(a) and 4(b) we can see that before the lower-level inflow (from the left boundary of the domain) reached the reflectivity core area, the flow was not only deflected rightward but also became strongly convergent (shown by the dashed negative contours of horizontal divergence) in a narrow elongated area to the left of the reflectivity core. Inside the reflectivity core area, the lower-level flow was strongly divergent, as shown by the solid positive contours of the horizontal divergence in the reflectivity core area in Figure 4(a). This strong lower-level divergence was caused by and tied up with a strong downdraft aloft, while this strong downdraft is not only revealed by the downward movement of the reflectivity core as exhibited by the time series of RHI display in Figures 2(a)–2(t) but also retrieved in the second step (as shown later in Figure 6(c)).

On the other hand, as we can see from Figures 4(e) and 4(f), the upper flow was divergent in a narrow elongated area to the left of the reflectivity core and became convergent inside the reflectivity core. From Figures 4(c) and 4(d), we can see that the middle-level flow was weakly convergent or divergent, and the convergence-divergence pattern for the middle-level flow is intermediate between the two nearly opposite patterns at the lower and upper levels. Thus, at least qualitatively, the retrieved quasi-horizontal wind fields in the first step captured the storm scale convergence-divergence patterns associated with the strong downdraft generated by the storm.

Figures 5(a) and 5(b) show the reflectivity source fields (see the S_m term on the right-hand side of the reflectivity

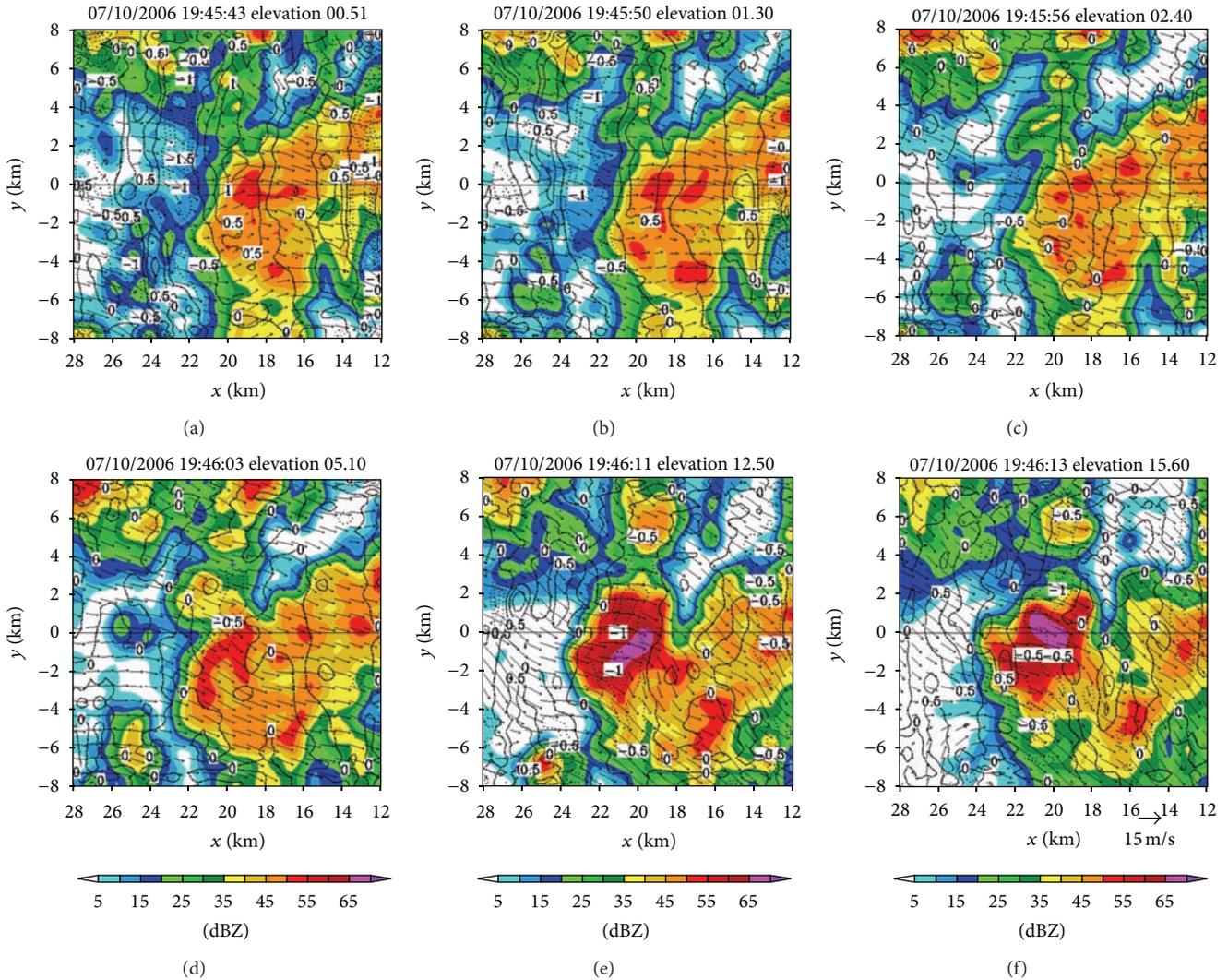


FIGURE 4: Quasi-horizontal (u_m, v_m) velocity fields (plotted by black arrows) retrieved in the nested domain on the conical surfaces of (a) $\theta = 0.51^\circ$, (b) 1.3° , (c) 2.4° , (d) 5.1° , (e) 12.5° , and (f) 15.6° from PAR 90° sector scans over the time period from 19:44:43 to 19:46:13 UTC on July 10, 2006. In each panel, the colored field shows the PAR-observed reflectivity (gridded and smoothed by the spatial interpolation on each tilt), and the black contours (with solid for positive and dashed for zero and negative) plot the field of horizontal divergence. The reflectivity of color scale is shown at the bottom of the figure, and the velocity vector scale is shown at the low-right corner of the figure. The nested domain size is $16 \times 16 \text{ km}^2$ (as marked by the red square box in Figure 3(b)). The x-coordinate is originated from the PAR site and directed along the 210.7° azimuth (as shown by the dashed yellow line in Figure 3(b)).

advection equation in (3a) obtained as by-products of the retrievals in Figures 4(a) and 4(e) on the conical surfaces of $\theta = 0.51^\circ$ and 12.5° , respectively. As shown in Figure 5(a), the retrieved reflectivity source field on the 0.51° tilt is characterized by small-scale fluctuations around zero, and the RMS amplitude of these small-scale fluctuations is smaller by several times than the RMS amplitudes of the local time derivative and advection terms (not shown) in the reflectivity advection equation. Note that the vertical advection term is implicitly absorbed into the source term S_m on the right-hand side of (3a), and this vertical advection term includes the effect of hydrometeors' downward terminal velocity, whereas the vertical gradient of reflectivity was small and more or less positive in the lower level (around $z \approx 0.2 \text{ km}$ on

the 0.51° tilt as shown by the RHI display in Figure 2(d)). Thus, the contribution of the vertical advection to S_m is weakly positive in the lower level. This positive contribution could offset the negative contribution to S_m produced by precipitation evaporation (since the evaporation could be also weak as implied by the small vertical gradient of reflectivity around $z \approx 0.2 \text{ km}$). This may explain the smallness of the retrieved S_m in Figure 5(a), although the source term S_m also absorbs residual errors of the reflectivity advection equation in (3a) especially those caused by the imperfect reflectivity observations and imperfect velocity retrievals.

In the upper level (around $z \approx 8 \text{ km}$ on the 12.5° tilt), as shown in Figure 5(b), the retrieved source term S_m is maximized in the narrow curved area along and inside

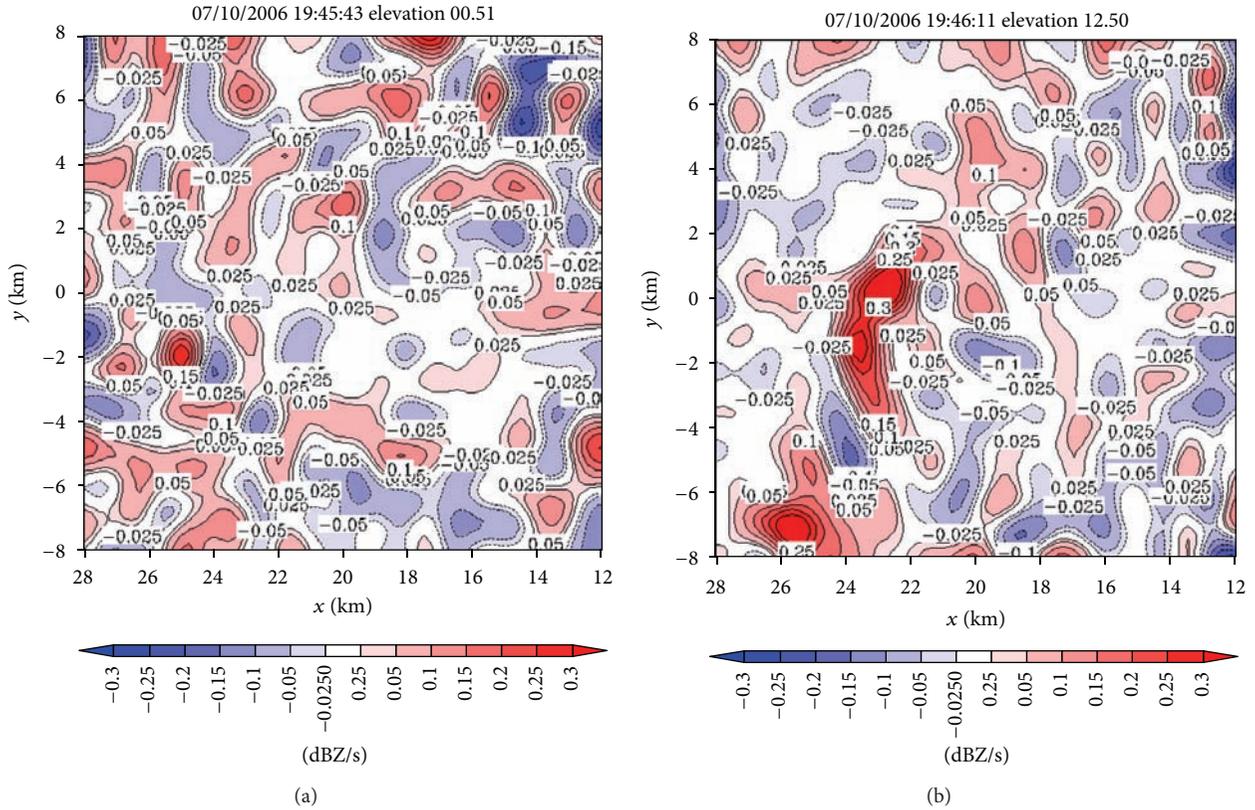


FIGURE 5: Reflectivity source fields (for the S_m term on the right hand-side of (3a)) produced as by-products of the retrievals in Figures 4(a) and 4(e) on the conical surfaces of (a) $\theta = 0.51^\circ$ and (b) $\theta = 12.5^\circ$. The color scale is shown on the bottom of the figure, and the unit is dBZ s^{-1} for the labeled contour values.

the left boundary of the reflectivity core (as seen by overlapping Figure 5(b) onto Figure 4(e)). In this curved upper-level area, the upward vertical velocity of air largely cancelled the downward terminal velocity for hydrometeors (as seen later from the retrieved vertical velocities in Figure 6(c)), and the vertical gradient of reflectivity was very small (on the 12.5° tilt as shown by the RHI display in Figure 2(d)). The vertical advection of reflectivity was thus very small, so only the condensation caused by the upward air motion may explain why the source term S_m is maximized in the curved upper-level area as we have seen from Figure 5(b).

As another by-product, the horizontal turbulent diffusivity coefficient k_h is also retrieved, and the retrieved value varies slightly with the scan elevation and time window. The mean value of the retrieved k_h is $274 \text{ m}^2 \text{ s}^{-1}$ with a standard deviation of $30 \text{ m}^2 \text{ s}^{-1}$. These retrieved values of k_h are in the same range as those obtained or used in the previous studies of the 2D-SA method (see the second paragraph in Section 2 of Xu et al. [8] and reference cited there).

3.4. Results Obtained in the Second Step. Figures 6(a)–6(h) show the time series of the retrieved (u_{m2}, w_m^a) fields for air motions (plotted by black arrows) in the vertical cross-section along the x -coordinate in the nested domain superimposed on the PAR-observed reflectivity fields from 19:42:03 to 19:55:43 UTC on July 10, 2006. As we can see from

Figures 6(a)–6(c), the air motions were upward within the reflectivity core, especially in the upper levels, and their produced condensation may largely explain the rapid intensification of the reflectivity core. However, as the reflectivity core moved down to the middle levels and finally reached the ground, the upward air motions inside the reflectivity core became weak and even changed to slightly downward motions (as shown in Figures 6(d)–6(h) and thus ceased to produce condensation during the later times. This may largely explain why the reflectivity core ceased to grow and even became slightly weak as it fell into and below the middle levels during the later time period from 19:48:50 to 19:55:43 UTC, as shown in Figures 6(d)–6(h). Furthermore, as shown by the white arrows in Figures 6(a)–6(h), the vector velocities of hydrometeors were overwhelmingly downward especially inside the reflectivity core, and their vertical-component velocities are around the value ($\approx 6.3 \text{ m s}^{-1}$) as estimated by the downward movements of the reflectivity core from the time series of the RHI displays in Figure 2.

Note that the vertical advection term is explicitly considered in (13a), so the source term S_{m2} on the right-hand side of (13a) should be related primarily to the production/reduction of hydrometeors caused by condensation/evaporation, although the source term S_{m2} also absorbs the residual errors caused in (13a) by the imperfect reflectivity observations and imperfect velocity retrievals. Figure 7 shows the S_{m2} field

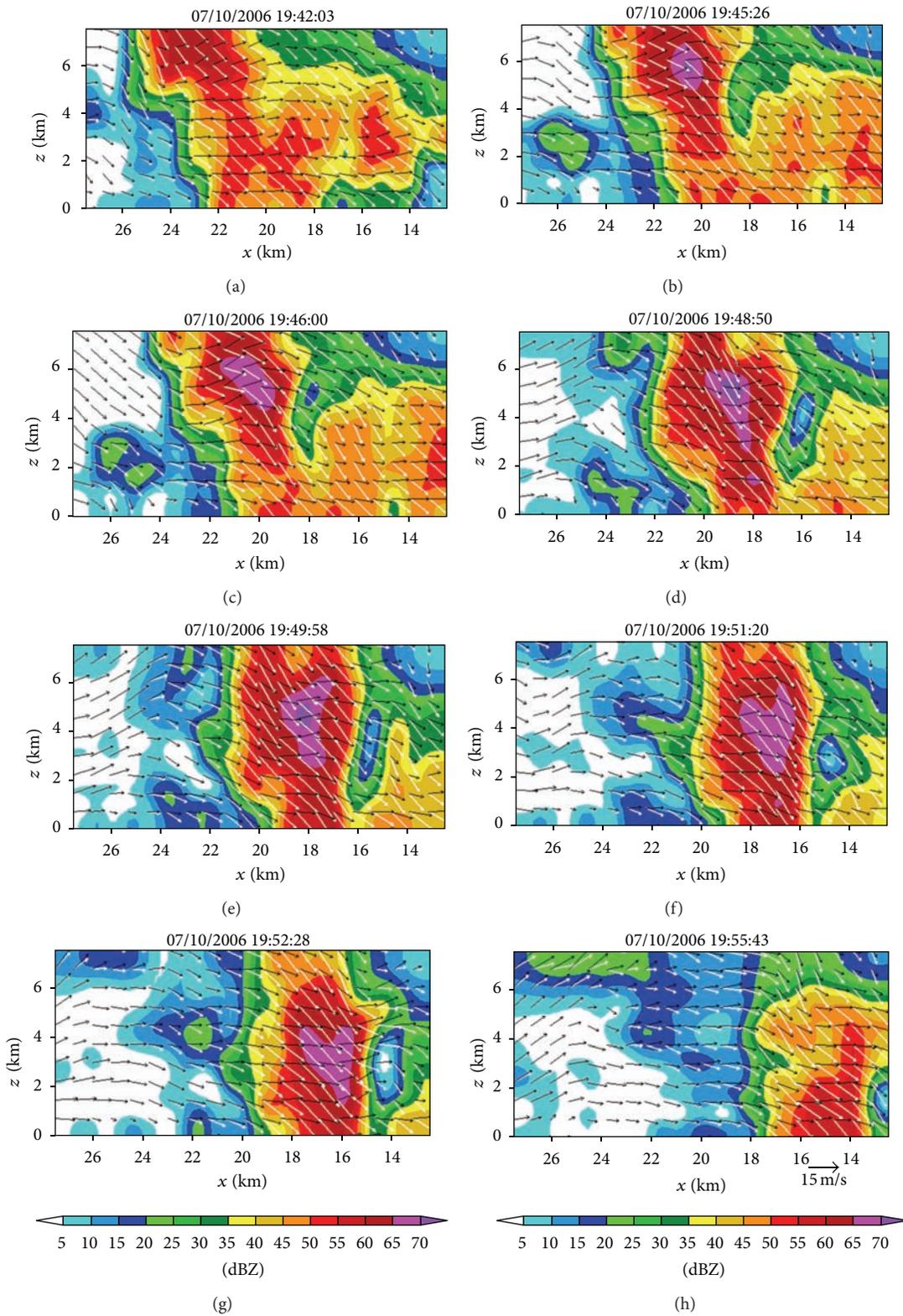


FIGURE 6: Time series of retrieved (u_{m2}, w_m^a) fields for air motions (plotted by black arrows) in the vertical cross-section (along the x -coordinate in Figure 4) superimposed on the PAR-observed reflectivity fields (gridded and smoothed by the spatial interpolation in the vertical cross-section) from 19:42:03 to 19:55:43 UTC on July 10, 2006. The white arrows plot the vector velocities of hydrometeors (i.e., air velocity plus the hydrometeors' downward terminal velocity).

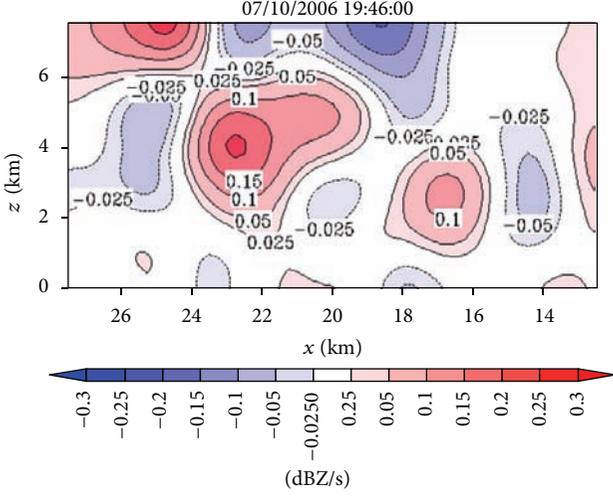


FIGURE 7: Reflectivity source field (for the S_{m2} term on the right-hand side of (13a)) produced in the vertical cross-section as a by-product of the retrieval in Figure 6(c). The color scale is shown on the bottom, and the unit is dBZ s^{-1} for the labeled contour values.

produced in the vertical cross-section as a by-product of the retrieval in Figure 6(c). By overlapping Figure 7 onto Figure 6(c), it is easy to see that the red-colored positive core of $S_{m2} > 0.2 \text{ dBZ s}^{-1}$ in the middle levels is inside the main updraft (where the upward air motions are maximized) within the reflectivity core, so this middle-level positive core of S_{m2} can be largely explained by the condensation and associated production of hydrometeors in the main updraft. Similarly, the upper-level negative core of $S_{m2} < -0.2 \text{ dBZ s}^{-1}$ between $20 \text{ km} > x > 18 \text{ km}$ along the top boundary of Figure 7 can be largely explained by the evaporation and associated reduction of hydrometeors due to the downward advection of hydrometeors into a relatively dry area (with reduced reflectivity) as shown in Figure 6(c). However, all other relatively weak positive and negative cores in Figure 7 do not seem to be physically meaningful and they may merely represent the residual errors caused in (13a) by the imperfect reflectivity observations and imperfect velocity retrievals.

As another by-product, the vertical turbulent diffusivity coefficient k_z is also retrieved. The retrieved values for the different time windows are almost the same. The mean value of the retrieved k_z is $201 \text{ m}^2 \text{ s}^{-1}$, and the standard deviation is merely $0.25 \text{ m}^2 \text{ s}^{-1}$. Note that all values of k_z are retrieved for the same along-beam vertical domain over a short time period, and the bulk effect of the vertical turbulent diffusion could remain nearly the same over this short time period. This may partially explain why the standard deviation is small. The retrieved values of k_z are roughly within the same range as those of k_h retrieved in the first step.

4. Benefits of PAR Rapid Scans and Usefulness of Mesoscale Background Wind Field

4.1. Benefits of PAR Rapid Scans. As mentioned in the introduction, the retrieval errors can be reduced if the reflectivity

and radial velocity fields are scanned more rapidly than the operational WSR-88D radar scans (Qiu and Xu [4], Shapiro et al. [5]). A similar benefit is expected for the two-step SA method developed in this paper and applied to PAR rapid scans. To demonstrate this, an additional experiment is performed in this subsection with the input PAR sector scan data used every 5 min instead of every 30 s, so the temporal resolution of the input observations is reduced by 10 times and becomes about the same as that of the operational WSR-88D radar scans. The parameter settings in this experiment are the same as those described for the experiment in Section 3.2 except that $\Delta\tau$ is increased to 5 min and thus $\tau = (N - 1)\Delta\tau$ is increased to 10 min (since $N = 3$ is unchanged). This experiment will be called Expt-5 min, while the experiment performed in Section 3 will be used as the benchmark.

Figures 8(a) and 8(b) show the (u_m, v_m) fields retrieved from the first step of Expt-5 min on the conical surfaces of $\theta = 0.51^\circ$ and 12.50° , respectively, and they are the 10 min time-mean wind fields centered at the same times (19:45:43 and 19:46:11 UTC) as the 1 min time-mean wind fields in Figures 4(a) and 4(e), respectively. As we can see, the retrieved wind field in Figure 8(a) (or Figure 8(b)) is weaker and slightly smoother than that in Figure 4(a) (or Figure 4(e)) although their gross patterns are similar. The reduced intensity and spatial variability for the retrievals in Figures 8(a) and 8(b) can be attributed largely to the reduced temporal resolution of the input observations and its caused increase of the time window (from $\tau = 1$ to 10 min) for the time-mean wind retrieval. Figures 9(a) and 9(b) show the difference fields obtained by subtracting the (u_m, v_m) fields in Figures 4(a) and 4(e) from those in Figures 8(a) and 8(b), respectively. The spatially averaged RMS values of the two difference fields are 3.08 and 4.25 m s^{-1} , respectively. These RMS differences can represent roughly the RMS errors caused by coarsening the temporal resolution of the input observations from $\Delta\tau = 0.5$ to 5 min, because the retrievals from Expt-5 min are significantly less accurate than those from the benchmark experiment, as evaluated indirectly below.

The 1 min time-mean wind fields retrieved in the benchmark experiment are less smeared in time and therefore expected to be more accurate than the 10 min time-mean wind fields retrieved in Expt-5 min. However, since the true fields are not known, it is difficult to directly evaluate whether and how much the retrieval accuracy is improved by rapid scans. To overcome this difficulty, we resort to the square root of the first (or second) cost-function term J_η (or J_{v_r}) in (1) that measures the RMS difference of the predicted η (or retrieved v_{rm}) to the observed reflectivity η_{ob} (or radial-velocity v_{rob}) averaged over the retrieval time window τ . Note that J is nondimensional and the weight W_η in J_η (or W_{v_r} in J_{v_r}) is inversely proportional to τ as shown in (22a), so the aforementioned RMS difference is measured by $R_\eta \equiv J_\eta^{1/2}$ (or $R_{v_r} \equiv J_{v_r}^{1/2}$) in the same way for different τ . Thus, the relative accuracies of retrievals obtained with different τ can be evaluated indirectly by comparing the initial values of R_η (or R_{v_r}) and subsequent reductions achieved through the iterative descending of J for different τ . To facilitate the comparison, the terms $R \equiv J^{1/2}$, R_η , and R_{v_r} computed from

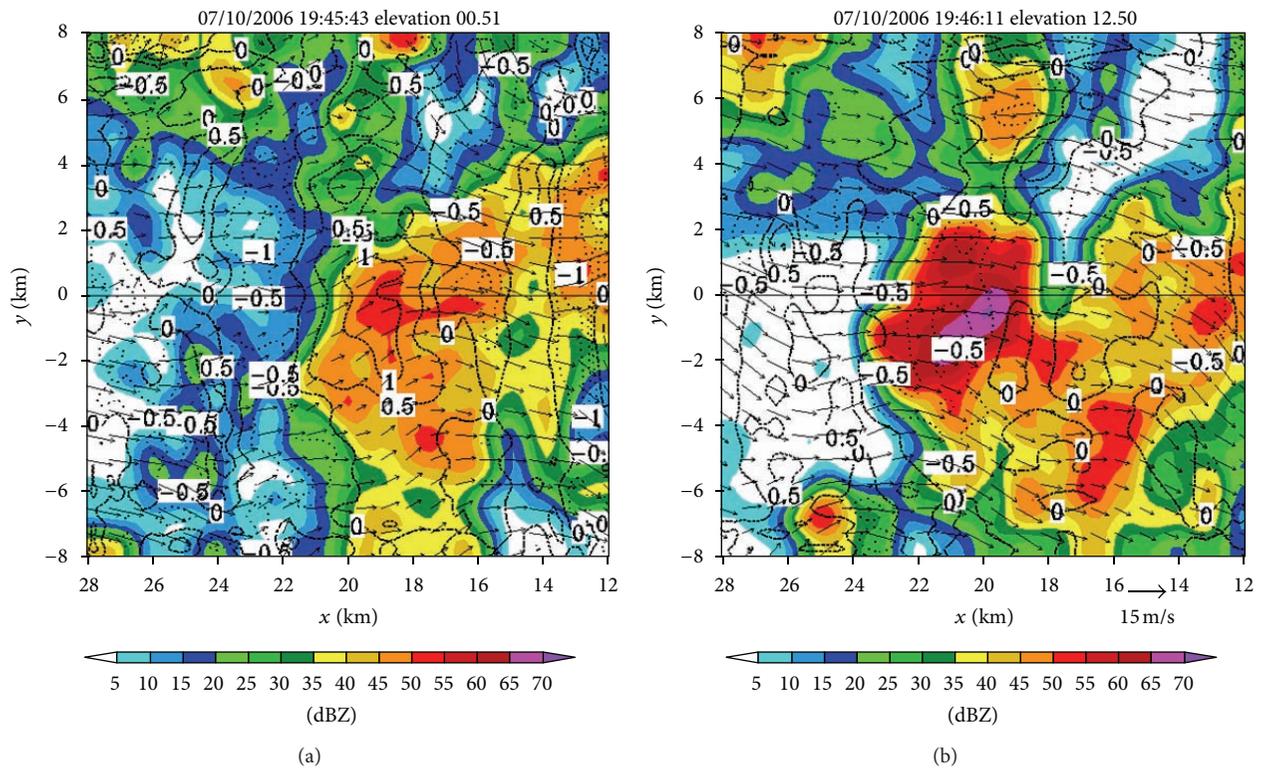


FIGURE 8: As in Figure 4 but the retrieved (u_m, v_m) fields are from the first step of Expt-5 min on the conical surfaces of (a) $\theta = 0.51^\circ$ and (b) 12.50° at the same times (19:45:43 and 19:46:11 UTC) as those in Figures 4(a) and 4(e), respectively.

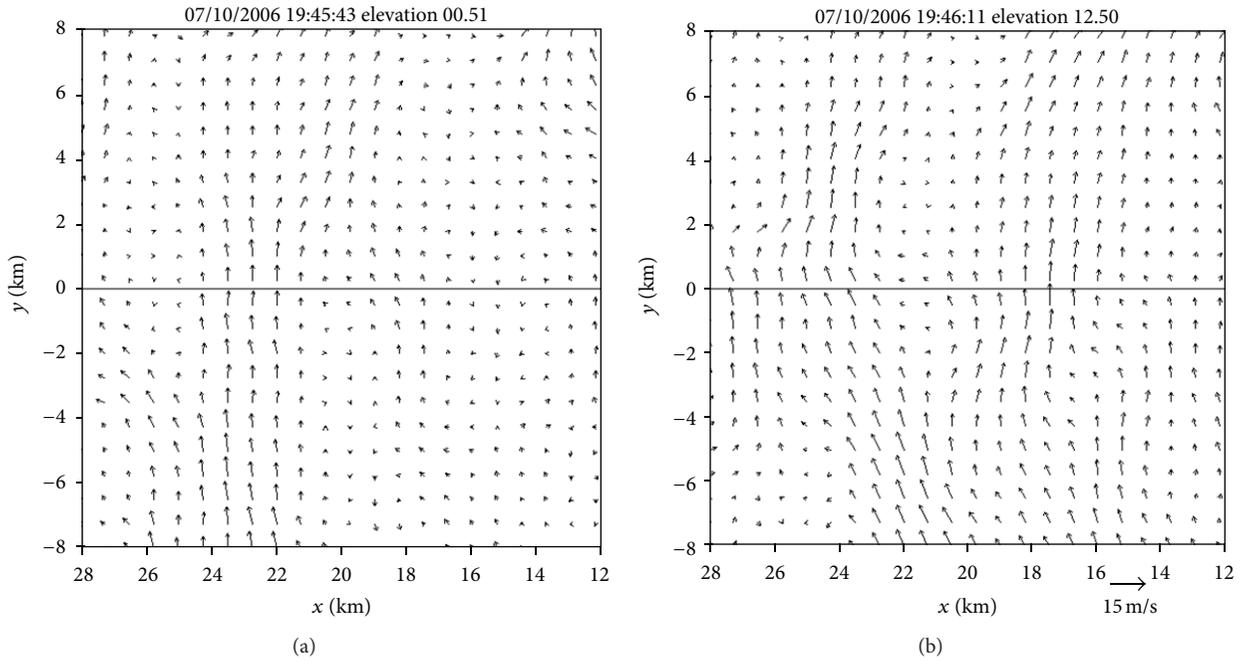


FIGURE 9: The difference fields obtained by subtracting the (u_m, v_m) fields in Figures 4(a) and 4(e) from those in Figures 8(a) and 8(b), respectively.

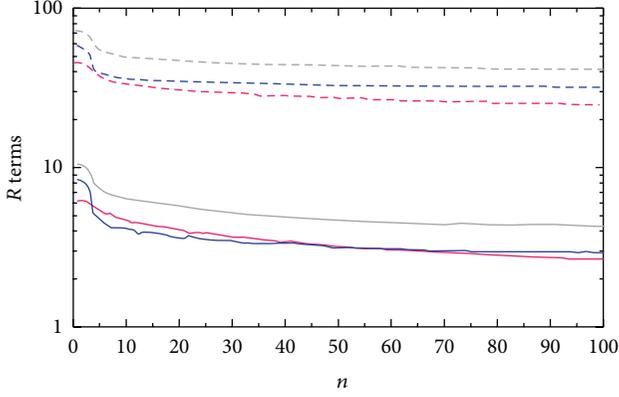
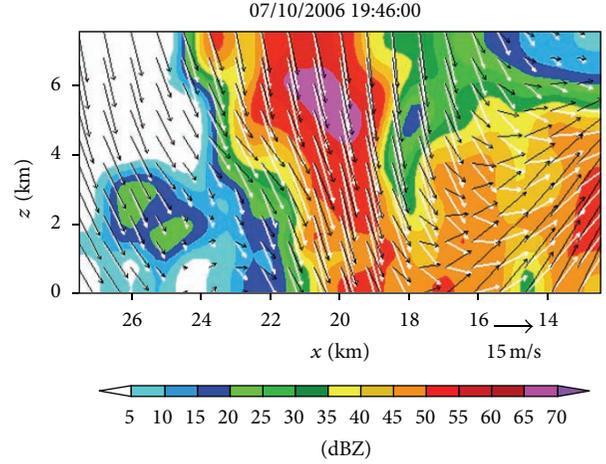


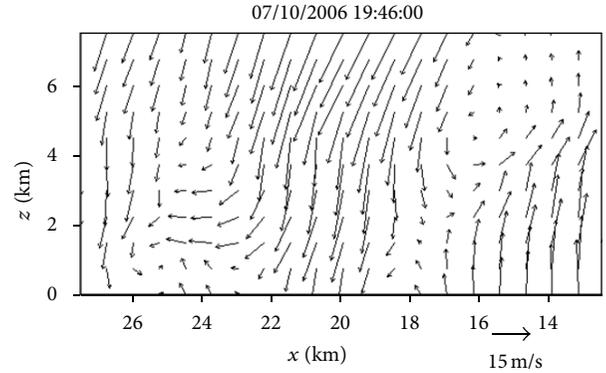
FIGURE 10: R -30 s, R_η -30 s, and R_{vr} -30 s for the retrieval in Figure 4(a) plotted by the gray, red, and blue solid curves, respectively, as functions of iteration number n . R -5 min, R_η -5 min, and R_{vr} -5 min for the retrieval in Figure 8(a) plotted by the gray, red, and blue dashed curves, respectively, as functions of n .

the first step of the benchmark experiment (or Expt-5 min) are denoted by R -30 s, R_η -30 s, and R_{vr} -30 s (or R -5 min, R_η -5 min and R_{vr} -5 min), respectively. These terms are plotted in Figure 10 as functions of iteration number n by the dashed curves for the retrieval in Figure 8(a) versus those plotted by the solid curves for the retrieval in Figure 4(a). By comparing the red (or blue) solid curve with the red (or blue) dashed curve in Figure 10, we can see that the ratio of R_η -5 min (or R_{vr} -5 min) to R_η -30 s (or R_{vr} -30 s) is as large as 7.3 (or 6.9) at $n = 0$ and increases to 9.3 (or 10.9) at $n = 100$. Similar large ratios are seen (not shown) from the retrieval in Figure 8(b) versus that in Figure 4(e). These large ratios imply that the first-step retrievals from Expt-5 m are significantly less accurate than those from the benchmark experiment.

Figure 11(a) shows the (u_{m2}, w_m^a) field for air motions (plotted by black arrows) retrieved from the second step of Expt-5 min around the same time as that (19:46:00 UTC) in Figure 6(c). As shown, the retrieved vertical-velocity field in Figure 11(a) has much stronger variations than that in Figure 6(c). In particular, the hydrometeors' downward vertical velocities (shown by the white arrows) inside the reflectivity core in Figure 11(a) are much larger than the value of 6.3 m s^{-1} , estimated from the downward movements of the reflectivity core in Figure 2 (see figure's caption for the estimated value), whereas the downward vertical velocities of hydrometeors inside the reflectivity core in Figure 6(c) are very close to and around the estimated value of 6.3 m s^{-1} . Thus, the retrieval in Figure 11(a) is significantly less accurate than that from the benchmark experiment in Figure 6(c). Figure 11(b) shows the difference field obtained by subtracting the (u_{m2}, w_m^a) field in Figure 11(a) from that in Figure 6(c). The spatially averaged RMS value of this difference field is 13.31 m s^{-1} . This large RMS difference is caused mostly by the RMS error of the retrieval in Figure 11(a), and the latter is caused both directly by the reduced temporal resolution of the input observations in the second step and indirectly by the reduced accuracies in the horizontal velocities retrieved in the first step and used in the second step. This large increase



(a)



(b)

FIGURE 11: (a) As in Figure 6(c) but the retrieved (u_{m2}, w_m^a) field is from the second step of Expt-5 min (around 19:46:00 UTC). (b) The difference field obtained by subtracting the (u_{m2}, w_m^a) field in Figure 6(c) from that in (a).

of retrieval error is also reflected by the large ratio of R_η -5 min/ R_η -30 s (≈ 15), where R_η -5 min (or R_η -30 s) is the square root of the first cost-function term in (12) computed from Expt-5 min (or benchmark experiment). Thus, the PAR rapid scans can improve not only the horizontal-wind retrieval in the first step but also the vertical-velocity retrieval in the second step, and the improvement is more significant in the second step.

4.2. Usefulness of Mesoscale Background Wind Field. For the particular case considered in this paper, the targeted retrieval domain is small and the mesoscale background velocity (u_b, v_b) is smooth and nearly constant on each tilt within this small domain. Because of this, the retrieved (u_m, v_m) fields in Figure 4 change insignificantly when the background wind field is not used, while the retrieved (u_{m2}, w_m^a) fields in Figure 4 change even less due to the fact that the background vertical velocity is zero. These properties are verified by another additional experiment performed with no background wind, that is, $(u_b, v_b) = 0$ and thus $(\Delta u_m, \Delta v_m) = (u_m, v_m)$. This experiment will be called Expt-NB.

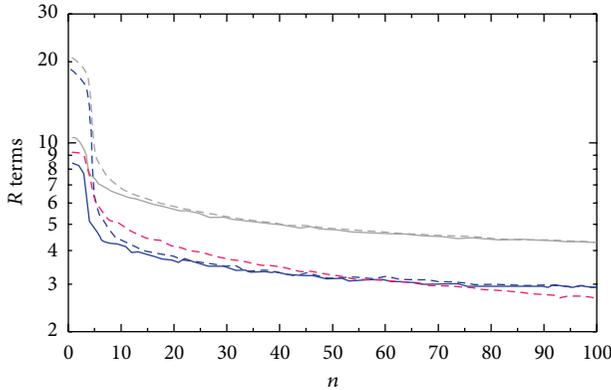


FIGURE 12: As in Figure 10 but the gray, red, and blue solid curves plot R -30 s NB, R_η -30 s NB, and R_{v_r} -30 s NB, respectively, for the retrieval from the first step of Expt-NB (instead of Expt-5 min).

The terms R , R_η , and R_{v_r} computed from the first step of Expt-NB are denoted by R -30 s NB, R_η -30 s NB, and R_{v_r} -30 s NB, respectively. These terms are plotted in Figure 12 as functions of the iteration number n by the dashed curves versus those plotted by the solid curves for the retrieval in Figure 4(a). By comparing the red (or blue) solid curve with the red (or blue) dashed curve in Figure 12, we can see that the ratio of R_η -30 s NB (or R_{v_r} -30 s NB) to R_η -30 s (or R_{v_r} -30 s) is 1.5 (or 2.2) at $n = 0$, decreases rapidly to 1.09 (or 1.03) at $n = 10$, and then gradually approaches 1.000 (or 1.000) as n further increases toward 100. These ratios are large initially, and this is because the unknown control variables $(\Delta u_m, \Delta v_m) = (u_m, v_m)$ are zero initially in Expt-NB and thus have larger errors than those in the benchmark experiment where $(\Delta u_m, \Delta v_m) = (u_m, v_m) - (u_b, v_b)$ are zero initially. Thus, using the mesoscale background velocity field can reduce the initial errors of the control variables $(\Delta u_m, \Delta v_m)$. The mesoscale background velocity should be more useful than just reducing the initial errors of the control variables for the two-step SA method if the retrieval domain is not too small (as that in this paper) to resolve some mesoscale variations from the background field. This speculation needs to be verified in future applications of the two-step SA method.

5. Conclusions

In this paper, the simple adjoint method (Qiu and Xu [6], Xu et al. [7, 8, 17]) is further developed into a two-step method to retrieve high-resolution horizontal and vertical wind fields from the PAR rapid scans of convective storm cells in a targeted domain. The first step retrieves the horizontal vector wind field on the conical surface of radar scan at each elevation angle, and the second step retrieves the vertical velocity in an along-beam vertical cross-section. As the horizontal winds can be retrieved in parallel on different elevations in the first step and only the vertical velocity needs to be retrieved in the second step, this two-step method is computationally efficient. In particular, on a workstation with Intel(R) Xeon(R) CPU X7550 (2.00 GHz), the computer time for the first step is about 20 s (with the retrievals

performed in parallel for different tilts) and the computer time for the second step is merely 10 s. Thus, the method can be applied very efficiently to real-time PAR rapid scans. The performance and expected capability of the method are demonstrated by a real-data example in Section 3 where the method is applied to PAR rapid 90° sector scans of a severe storm that produced a strong downdraft and subsequent damaging microburst during the early evening of July 10, 2006.

The above severe storm was scanned not only by the PAR but also by the operational KTLX radar (about every 5 minutes per volume) and the terminal Doppler weather radar (also about every 5 minutes per volume). These two operational radars, however, are all located in the same northeast quadrant as the PAR relative to the storm, so real dual-Doppler observations are not available for quantitative verifications of the single-Doppler retrievals obtained in this paper. Nevertheless, the method used in the first step is a refinement of the previous 2D-SA method, and the previous method has been tested for many real cases with the retrieved wind fields well verified by dual-Doppler analyses (Xu et al. [7, 8]). Thus, the method in the first step should remain at least as reliable as the previous 2D-SA method, and this speculation is supported by the detailed analysis of the retrievals obtained in the first step (see Section 3.3). The method developed for the second step is new. This new method performs well with the PAR rapid scans as suggested by the detailed analysis of the retrievals obtained in the second step (see Section 3.4), but the performance deteriorates significantly when the temporal resolution of the input observations is coarsened from 0.5 to 5 min to mimic the operational WSR-88D radar scans (see Section 4.1).

Note that the 2D-SA can be also formulated in the polar coordinates centered at the radar (instead of the local Cartesian coordinates as shown in Figure 2(a)) with the retrieval performed in a targeted sector area on each tilt of radar scan essentially in the same way as that in Xu et al. [18] and Xu and Qiu [17]. After these sector-area retrievals are performed in parallel for different tilts in the first step, the second step can be performed efficiently in parallel for different along-beam vertical cross-sections. In this way, the method can retrieve high-resolution three-dimensional wind fields in real time from PAR rapid scans over various adaptively-nested sector areas threatened by damaging winds generated by severe storms. This real-time capability will be developed and tested in future studies.

Acknowledgments

The authors are thankful to Dr. Alan Shapiro and anonymous reviewers for their comments and suggestions that improved the representation of the results and to GuangYu Wang for his help in editing Figure 4. The research was supported by ONR Grant N000141010778 to the University of Oklahoma (OU), NOAA-OU Cooperative Agreement NA17RJ1227, and Tianjin Municipal Meteorological Bureau Grant 201316.

References

- [1] D. S. Zrnic, J. F. Kimpel, D. E. Forsyth et al., "Agile-beam phased array radar for weather observations," *Bulletin of the American Meteorological Society*, vol. 88, no. 11, pp. 1753–1766, 2007.
- [2] S. Torres, P. Heinselman, R. Adams et al., "ADAPTS implementation: can we exploit phased-array radar's electronic beam steering capabilities to reduce update times?" in *Proceedings of the 28th Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, American Meteorological Society, New Orleans, La, USA, 2012.
- [3] P. L. Heinselman, D. L. Priegnitz, K. L. Manross, T. M. Smith, and R. W. Adams, "Rapid sampling of severe storms by the national weather radar testbed phased array radar," *Weather and Forecasting*, vol. 23, no. 5, pp. 808–824, 2008.
- [4] C. Qiu and Q. Xu, "Least squares retrieval of microburst winds from single-Doppler radar data," *Monthly Weather Review*, vol. 124, no. 6, pp. 1132–1144, 1996.
- [5] A. Shapiro, P. Robinson, J. Wurman, and J. Gao, "Single-Doppler velocity retrieval with rapid-scan radar," *Journal of Atmospheric and Oceanic Technology*, vol. 20, no. 12, pp. 1758–1775, 2003.
- [6] C. Qiu and Q. Xu, "A simple adjoint method of wind analysis for single-Doppler data," *Journal of Atmospheric and Oceanic Technology*, vol. 9, no. 5, pp. 588–598, 1992.
- [7] Q. Xu, C. Qiu, H. Gu, and J. Yu, "Simple adjoint retrievals of microburst winds from single-Doppler radar data," *Monthly Weather Review*, vol. 123, no. 6, pp. 1822–1833, 1995.
- [8] Q. Xu, H. Gu, and C. Qiu, "Simple adjoint retrievals of wet-microburst winds and gust-front winds from single-Doppler radar data," *Journal of Applied Meteorology*, vol. 40, no. 8, pp. 1485–1499, 2001.
- [9] Q. Xu, H. Gu, and S. Yang, "Simple adjoint method for three-dimensional wind retrievals from single-Doppler radar," *Quarterly Journal of the Royal Meteorological Society*, vol. 127, no. 573, pp. 1053–1067, 2001.
- [10] J. Sun and A. Crook, "Wind and thermodynamic retrieval from single-Doppler measurements of a gust front observed during Phoenix II," *Monthly Weather Review*, vol. 122, no. 6, pp. 1075–1091, 1994.
- [11] Q. Xu, K. Nai, L. Wei, P. Zhang, Q. Zhao, and P. R. Harasti, "A real-time radar wind data quality control and analysis system for nowcast application," in *Proceedings of the International Symposium on Nowcasting and Very Short Range Forecasting (WSN '09)*, WMO, Whistler, Canada, 2009, http://www.nowcasting2009.ca/images/stories/Presentations/WSN09_Presentations.htm.
- [12] Q. Xu, "Representations of inverse covariances by differential operators," *Advances in Atmospheric Sciences*, vol. 22, no. 2, pp. 181–198, 2005.
- [13] R. J. Purser, W. S. Wu, D. F. Parrish, and N. M. Roberts, "Numerical aspects of the application of recursive filters to variational statistical analysis—part I: spatially homogeneous and isotropic Gaussian covariances," *Monthly Weather Review*, vol. 131, no. 8, pp. 1524–1535, 2003.
- [14] R. J. Purser, W. Wu, D. F. Parrish, and N. M. Roberts, "Numerical aspects of the application of recursive filters to variational statistical analysis—part II: spatially inhomogeneous and anisotropic general covariances," *Monthly Weather Review*, vol. 131, no. 8, pp. 1536–1548, 2003.
- [15] Q. Xu, L. Wei, W. Gu, J. Gong, and Q. Zhao, "A 3.5-dimensional variational method for Doppler radar data assimilation and its application to phased-array radar observations," *Advances in Meteorology*, vol. 2010, Article ID 797265, 14 pages, 2010.
- [16] E. Kessler, "On the distribution and continuity of water substance in atmospheric circulation," *Meteorological Monographs*, vol. 10, no. 32, p. 84, 1969.
- [17] Q. Xu and C. Qiu, "Adjoint-method retrievals of low-altitude wind fields from single-Doppler reflectivity and radial-wind data," *Journal of Atmospheric and Oceanic Technology*, vol. 12, no. 5, pp. 1111–1119, 1995.
- [18] Q. Xu, C. Qiu, and J. Yu, "Adjoint-method retrievals of low-altitude wind fields from single-Doppler reflectivity measured during Phoenix II," *Journal of Atmospheric and Oceanic Technology*, vol. 11, no. 2, pp. 275–288, 1994.

Research Article

Low-Level Polarimetric Radar Signatures in EnKF Analyses and Forecasts of the May 8, 2003 Oklahoma City Tornado Supercell: Impact of Multimoment Microphysics and Comparisons with Observation

Daniel T. Dawson II,^{1,2,3} Louis J. Wicker,² Edward R. Mansell,²
Youngsun Jung,¹ and Ming Xue^{1,4}

¹ Center for Analysis and Prediction of Storms, University of Oklahoma, 120 David L. Boren Boulevard, Norman, OK 73072, USA

² NOAA/National Severe Storms Laboratory, National Weather Center, 120 David L. Boren Boulevard, Norman, OK 73072, USA

³ Cooperative Institute for Mesoscale Meteorological Studies, National Weather Center, 120 David L. Boren Boulevard, Norman, OK 73072, USA

⁴ School of Meteorology, University of Oklahoma, Norman, OK 73072, USA

Correspondence should be addressed to Daniel T. Dawson II; ddawson@ou.edu

Received 31 May 2013; Accepted 26 August 2013

Academic Editor: David J. Stensrud

Copyright © 2013 Daniel T. Dawson II et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The impact of increasing the number of predicted moments in a multimoment bulk microphysics scheme is investigated using ensemble Kalman filter analyses and forecasts of the May 8, 2003 Oklahoma City tornadic supercell storm and the analyses are validated using dual-polarization radar observations. The triple-moment version of the microphysics scheme exhibits the best performance, relative to the single- and double-moment versions, in reproducing the low- Z_{DR} hail core and high- Z_{DR} arc, as well as an improved probabilistic track forecast of the mesocyclone. A comparison of the impact of the improved microphysical scheme on probabilistic forecasts of the mesocyclone track with the observed tornado track is also discussed.

1. Introduction

The assimilation of radar data into storm scale models using the Ensemble Kalman Filter (EnKF) [1] approach has proven to be an extremely useful tool for the analysis and prediction of convective storms in recent years. There have been many recent successful uses of this approach for both analyses [2–7] and short-range forecasts [8–10] based on these analyses. In general, these studies have focused on improving techniques for assimilation of radar data, on the design of the overall data assimilation system, or on the impact of initial and boundary conditions. High-resolution numerical weather prediction has progressed during the past decade such that prediction of the dynamics of individual convective storms is now routinely attempted. One substantial challenge is the improvement and validation of the microphysics parameterization and the

associated impacts on storm structure and behavior (e.g., through the development of the cold pool). Errors from the model's microphysical parameterization can significantly impact forecasts of these storms. Polarimetric radar observations offer a rich source of data to validate the output of such schemes within this context.

Several storm-scale simulation studies have shown that the microphysics parameterization has a profound impact on simulated storm structure and behavior [11–17] and even on tornadic potential [18]. Here, we restrict our discussion to bulk microphysics schemes, which assume *a priori* a certain functional form for the underlying drop or particle size distribution (DSD/PSD) for several hydrometeor categories. Typically, one or more moments of the PSD for a given category are explicitly predicted within a scheme, with single-moment schemes that predict the mass mixing ratio

(proportional to the third moment) being the most common. Double-moment schemes that typically predict total number concentration (0th moment) in addition to mixing ratio are becoming increasingly common [19–22], and triple-moment schemes that additionally predict the sixth moment (proportional to radar reflectivity factor) are also available [23]. Milbrandt and Yau [11] and Dawson et al. [12] found that a multimoment bulk microphysics parameterization, as opposed to the more traditional single-moment schemes, better simulated the reflectivity structure and cold pool intensity as compared with observations, in their respective supercell storm simulation experiments. Dawson et al. [8] briefly investigated the impact of decreasing the complexity of the microphysics parameterization (from double- to single-moment) on their EnKF-based ensemble forecasts of the May 4, 2007 Greensburg, KS, USA tornadic supercell and found that doing so substantially degraded the forecast track of the simulated mesocyclone across the ensemble. Similarly, Yussouf et al. [10] showed that the double-moment schemes they considered in their EnKF-based forecast experiments generally performed better than the single-moment schemes in predicting the forecast track of the May 8, 2003 Oklahoma City tornadic supercell.

Polarimetric radar offers an unprecedented amount of information about microphysical characteristics within convective storms [24–33]. Jung et al. [34] utilized a polarimetric emulator to successfully evaluate the performance of a single- and double-moment microphysics scheme. They found that the double-moment scheme performed substantially better than its single-moment counterpart in simulating commonly observed polarimetric radar signatures in supercell storms, such as the Z_{DR} arc and the low- Z_{DR} hail signature [35], due at least in part to its ability to parameterize the important process of gravitational size sorting of hydrometeors.

In the current study, we investigate the impact of systematically increasing the microphysics complexity on EnKF analyses and ensemble forecasts of the May 8, 2003 Oklahoma City tornadic supercell, the subject of several other recent studies [3, 10, 36–38]. We emphasize in particular the benefits of systematically increasing the number of PSD moments predicted from one to three, pertaining to the improved simulation of size sorting. Milbrandt and Yau [39] showed that, for idealized hydrometeor sedimentation, a single-moment scheme is unable to simulate size sorting at all, a double-moment scheme can vastly overestimate size sorting, and a triple-moment scheme closely approximates an analytical bin solution. In this study, we show that accurately representing size sorting is very important for accurate simulation of low-level polarimetric signatures, focusing on the Z_{DR} arc and low- Z_{DR} hail core. Additionally, we demonstrate an improved probabilistic mesocyclone track prediction for the double- and triple-moment microphysics schemes.

This paper is organized as follows. In Section 2, the methodology of the EnKF assimilation and forecast experiments and an improved version of the Jung et al. [34] polarimetric emulator are briefly described. Section 3 describes the results of the EnKF analysis and forecast experiments. Summary and conclusions are given in Section 4.

2. Methodology

2.1. EnKF Analysis and Forecast Experiments. All experiments use the NSSL Collaborative Model for Multiscale Atmospheric Simulation (COMMAS) [20, 40, 41] and its associated EnKF radar data assimilation system [3, 38]. The model is run using a horizontally homogeneous background environment representative of the inflow conditions during the mature phase of the May 8, 2003 tornadic supercell storm (not shown). No surface or radiation physics are included in the simulations, and the bottom and top boundaries are free slip. We use a horizontal domain size of 100 km in each direction. The grid spacings are 1 km in the horizontal directions and utilize a stretched vertical grid with 60 total levels with the lowest 10 levels set to 150-m grid spacing, stretched thereafter to 600 m at the model top at 26 km AGL. This storm produced a long-track F4 tornado in the city of Moore, OK USA and the reader is referred to the work of Romine et al. [37] for a thorough overview of the storm. While some previous studies of this storm assimilate both radar reflectivity (Z) and radial velocity (V_r) data from nearby S-band radars [3, 10], we choose to assimilate only V_r data from the KOUN polarimetric S-band radar. However, reflectivity data *were* used to define the regions for additive noise and thermal bubbles, in order to spin up the analysis and increase ensemble spread [38]. These data were first quality controlled, dealiased, and objectively analyzed to a regular 2 km horizontal grid, but left on the original conical radar sweep surfaces [42]. The V_r observations are then aggregated into 2 min bins and assimilated every two minutes from 2040 to 2208 UTC on May 8, 2003. This period covers the time from the beginning of the first echoes associated with the storm to just before tornadogenesis [38]. Furthermore, we restrict the covariance updates of the model state variables by the assimilated V_r to only the three wind components (u, v, w), potential temperature θ , and water vapor mixing ratio q_v . The goal of this “minimalist” or “quasi-kinematic” EnKF update strategy is to allow the microphysics state variables as much freedom as possible to evolve according to the microphysics scheme itself, without any *direct* updating from the EnKF portion of the data assimilation system. In this regard, our study differs from the recent study of Jung et al. [6] who also verified their analyzed storm against polarimetric observations but directly assimilated reflectivity throughout the analysis period.

A total of four assimilation experiments with 30 members each are performed, each utilizing an increasing number of predicted moments using the NSSL multi-moment microphysics scheme [20, 43], from one (1M) to two (2M and 2MSC) to three (3M). In 1M, all fixed intercept parameter values for rain, snow, and hail are set as in Dawson et al. [12] for their “LINB” experiment, with the intercept parameter for rain n_{0r} reduced from its default value of $8.0 \times 10^6 \text{ m}^{-4}$ to $4.0 \times 10^5 \text{ m}^{-4}$, to reduce the overall strength of the cold pool. A reduction in n_{0r} by a factor of 10 reduces the evaporation rate by roughly half. The single-moment graupel category has an intercept of $4.0 \times 10^5 \text{ m}^{-4}$ and particle density of 500 kg m^{-3} . For the double-moment configuration, two separate experiments are performed, one without correction

to the sedimentation (2M) and one using the sedimentation correction method (I+II, 2MSC) of Mansell [44], which compensates for the excessive size sorting inherent in double-moment sedimentation [39, 44–46]. Finally, from the 2200 UTC analyses of each experiment, 1 hr forecasts are launched for each ensemble member, from which vorticity probabilities are computed (Section 3.2).

Excessive size sorting is characterized by increased reflectivity (unreasonably large particles) when the mass moment (q) falls too much faster than the number concentration (N), particularly at the downward leading edge of a precipitation shaft, and is more pronounced for small fixed values of the shape parameter (e.g., [38]). Briefly, the correction scheme of Mansell [44] creates a temporary field of reflectivity (Z) moments, and then all three moments sediment using the corresponding moment-weighted fall speeds. The sedimented mass and reflectivity moments (q' and Z') are then used to generate a concentration number (N''), which is compared to the sedimented number concentration (N'). The final number concentration is the maximum of N' and N'' at each point. The result is to artificially increase N to prevent Z from increasing [43]. In a 3-moment scheme, however, the size distribution shape parameter increases automatically in response to size sorting, which decreases the differences between the moment-weighted fall speeds and thus decelerates further size sorting.

2.2. Microphysics and Polarimetric Radar Emulator. We use an updated version of the multi-moment microphysics scheme described in Mansell et al. [20] and originally based on Ziegler [43]. The updated scheme predicts up to three moments (mixing ratio, total number concentration, and radar reflectivity factor-proportional to the 3rd, 0th, and 6th moments) of the assumed three-parameter gamma size distribution [47] for the hydrometeor categories of rain, graupel, and hail and up to two moments for cloud, cloud ice, and snow. The 6th moment closure mainly follows Milbrandt and Yau [23] and is described in detail in Dawson et al. [48]. Additionally, as described in Mansell et al. [20] and Mansell and Ziegler [49], the bulk densities of graupel and hail are predicted for the double- and triple-moment versions of the scheme but are held fixed at 500 kg m^{-3} and 900 kg m^{-3} , respectively, for the single-moment scheme.

To derive polarimetric variables from the predicted model microphysics state variables, we utilize an updated version of the polarimetric emulator described in Jung et al. [34]. The emulator makes use of the T-matrix method [50–52] to compute scattering amplitudes across size bins for each precipitating hydrometeor category, where the size bins have been discretized from the model predicted PSDs at each grid point. From these scattering amplitudes, the standard polarimetric variables of reflectivity at horizontal and vertical polarization (Z_H , Z_V), differential reflectivity Z_{DR} , specific differential phase K_{DP} , and cross-correlation coefficient ρ_{HV} are computed.

The updates primarily apply to how the water fraction on wet graupel and hail is diagnosed. Like Jung et al. [34], we diagnose a water fraction at each grid point where both ice

and rain are present, where the ice can be any or all of the snow, graupel, and hail categories, by borrowing a portion of the available water from the rain field. However, whereas Jung et al. [34] assumed that the computed water fraction was applied evenly across the sizes in a given graupel or hail distribution, we make use of the empirical relationship of critical water fraction derived by Rasmussen et al. [53], which allows for a varying water fraction across the size distribution (possibly leading to completely melted hail or graupel that is smaller than 8 mm). This updated method provides more realistic polarimetric signatures in regions where partially melted hail and/or graupel are present. A complete description of the diagnostic water fraction technique is provided in Dawson et al. [48].

3. Results

3.1. Low-Level Polarimetric Signatures. From each of the four assimilation experiments, we choose the final analysis time (2208 UTC) and compute the polarimetric variables for the prior analysis (i.e., the 2 min mean forecast from the previous EnKF analysis update) using the emulator. The variables are then vertically interpolated to the 2.5° elevation sweep surface corresponding to the KOUN radar, to facilitate direct comparison with the objectively analyzed observations (Figure 1). For brevity, we focus only on the Z , Z_{DR} , and ρ_{HV} fields in this paper. The observed Z_{DR} signatures (Figure 1(b)) feature a well-defined Z_{DR} “arc” near the southeast edge of the forward flank and a low- Z_{DR} (and high Z) hail core between the hook echo region and the Z_{DR} arc [35, 37]. The Z_{DR} arc has been hypothesized to be a result of size sorting of large rain drops toward the upwind (in a storm-relative sense) side of the forward flank in the presence of low-level environmental wind shear by Kumjian and Ryzkhov [54]. In ρ_{HV} , the region of hail near the core is highlighted by relatively low values (~ 0.90 – 0.96 , Figure 2(a)), commensurate with the presence of mixed rain and tumbling hail.

Substantial differences in the simulated polarimetric fields are seen across each of the four experiments (Figures 1(c)–1(j) and 2(b)–2(e)). In general, the triple-moment experiment 3M performs the best of all in qualitatively reproducing the Z_{DR} arc signature and performs reasonably well for the ρ_{HV} pattern (Figures 1(j) and 2(e)), though it tends to over predict both the magnitude and coverage of high reflectivity in the core (Figure 1(i)). In contrast, the single-moment experiment 1M performs poorly in Z_{DR} and ρ_{HV} (Figures 1(d) and 2(b)), as does the double-moment without sedimentation correction (2M, Figures 1(f) and 2(c)), but on the other hand has better Z in the hook region of the storm. In 1M, the forward flank region is oriented differently than the observations (from southwest to northeast, Figure 1(c)) and the Z_{DR} arc and low- Z_{DR} and ρ_{HV} hail core are entirely absent (Figures 1(d) and 2(b)). The pattern of mean volume diameter D_m of rain and hail for 1M (Figures 3(a) and 3(b)) mirrors the Z_{DR} pattern; D_{mr} and D_{mh} are both relatively small across the storm (1 mm or less). (The mean volume diameter D_m is the diameter of the (spherical) particle with a volume equal to the mean volume of the distribution). (As stated

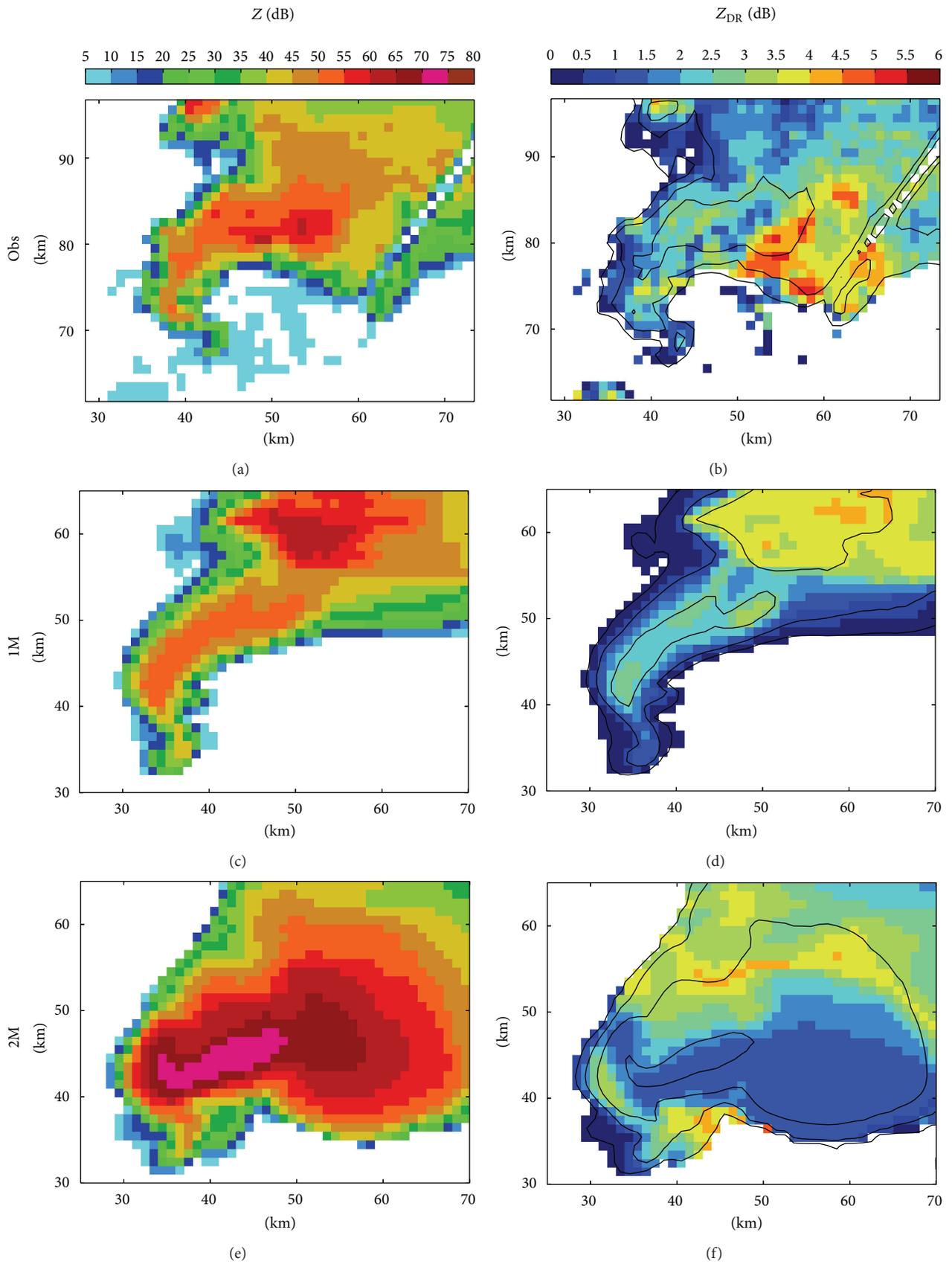


FIGURE 1: Continued.

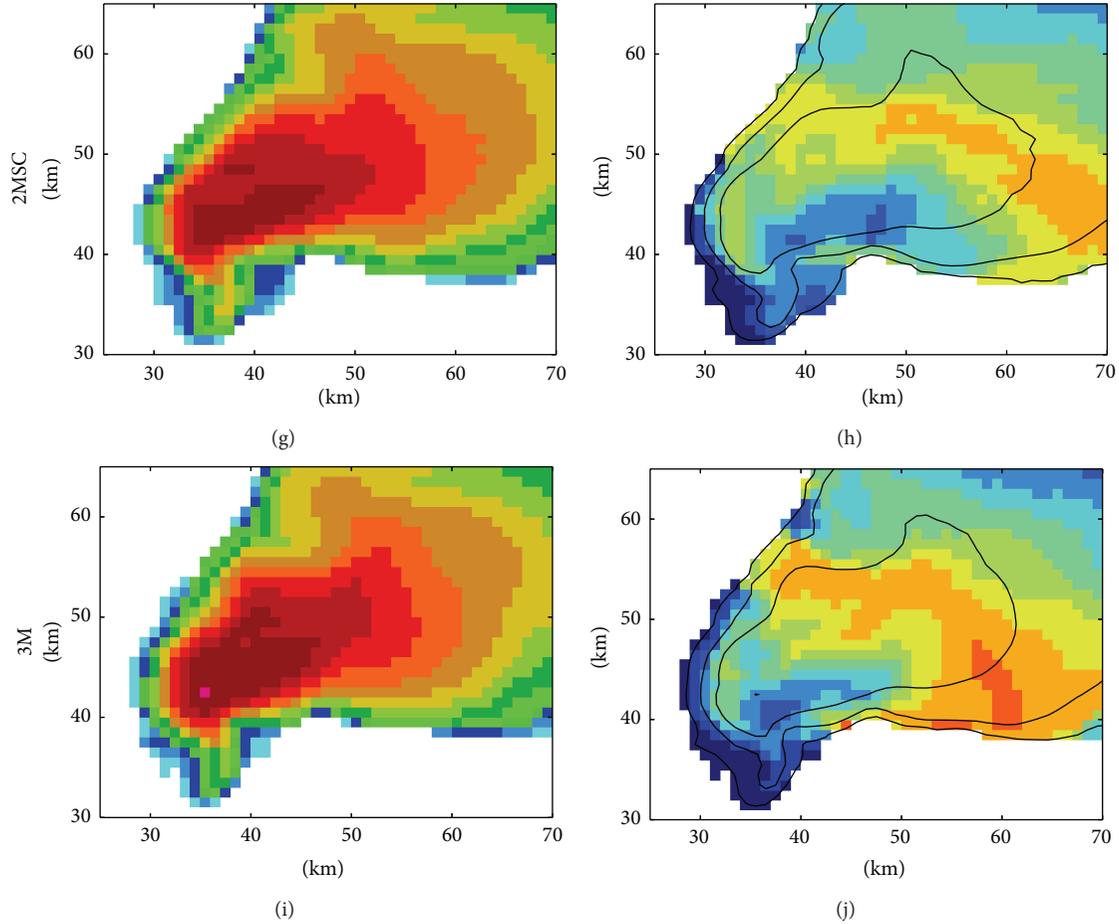


FIGURE 1: Radar reflectivity Z (left column, color fill, dBZ) and differential reflectivity Z_{DR} (right column, color fill, dB) for (a) and (b) objective analysis of the 2.5 degree elevation scan from KOUN at approximately 2208 UTC May 8, 2003, (c) and (d) EnKF prior analysis mean for experiment 1 M, (e) and (f) 2 M, (g) and (h) 2 MSC, and (i) and (j) 3 M. For reference in the right column, Z is overlaid in black contours in 20 dBZ increments, starting at 10 dBZ.

previously, separate graupel and hail categories are predicted in the double- and triple-moment versions of the scheme. The graupel in the low levels, however, has hail-like characteristics due to a high predicted bulk density (not shown), and so is referred to as “hail” here. The separate hail category shows similar behavior (not shown), and so is omitted for the sake of brevity). In 2 M, the Z_{DR} arc is almost completely absent due to very large predicted mean hail diameters (~ 20 mm) in the forward flank region where the Z_{DR} arc would otherwise be (Figure 3(d)), despite the presence of large ($D_{mr} \sim 6$ mm) rain here (Figure 3(c)). Additionally, the reflectivity values in the core are much too high (>70 dBZ, Figure 1(e)), and the ρ_{HV} magnitudes are too much low (~ 0.90 – 0.96) over a broad region in the forward flank (Figure 2(c)). The double-moment experiment with sedimentation correction (2 MSC, Figures 1(g) and 1(h)) is intermediate between 2 M (Figures 1, 3(e), 3(f), and 2(d)), and 3 M (Figures 1, 3(g), 3(h), and 2(e)), and in fact 2 MSC appears to perform slightly better than 3 M in the magnitude and coverage of the low- ρ_{HV} region (compare Figures 2(d) and 2(e) with Figure 2(a)), although this may come at the expense of a muted Z_{DR} arc.

The patterns of mean volume diameter of both rain and hail in the double- and triple-moment experiments (Figures 3(c)–3(h)) reflect the action of size sorting in these experiments, with the largest hail and rain falling out closer to the updraft and south edge of the forward flank in each case [54, 55], while the smaller rain and hail are advected further downstream in a storm-relative sense (i.e., to the north and northeast). The somewhat smaller D_{mr} in the forward flank in 2 MSC relative to 3 M (compare Figures 3(e) and 3(g)) may be a result of the correction scheme increasing total number concentration N_{tr} to prevent spurious growth in Z via size sorting, whereas in 3 M, the shape parameter of the gamma distribution is also allowed to increase, allowing larger D_{mr} for the same Z [44]. For rain in 3 M (Figure 3(g)), this leads to larger Z_{DR} values in the Z_{DR} arc region than elsewhere in the storm, due to the well-known increase in Z_{DR} with increasing drop oblateness and the fact that the hail sizes in this region are relatively small ($D_{mh} < 10$ mm) and relatively wet (Figure 4). The region of larger hail sizes in 3 M (Figure 3(h)) corresponds well with the low- Z_{DR} hail core, as expected, due to the assumed tumbling characteristics for large hail

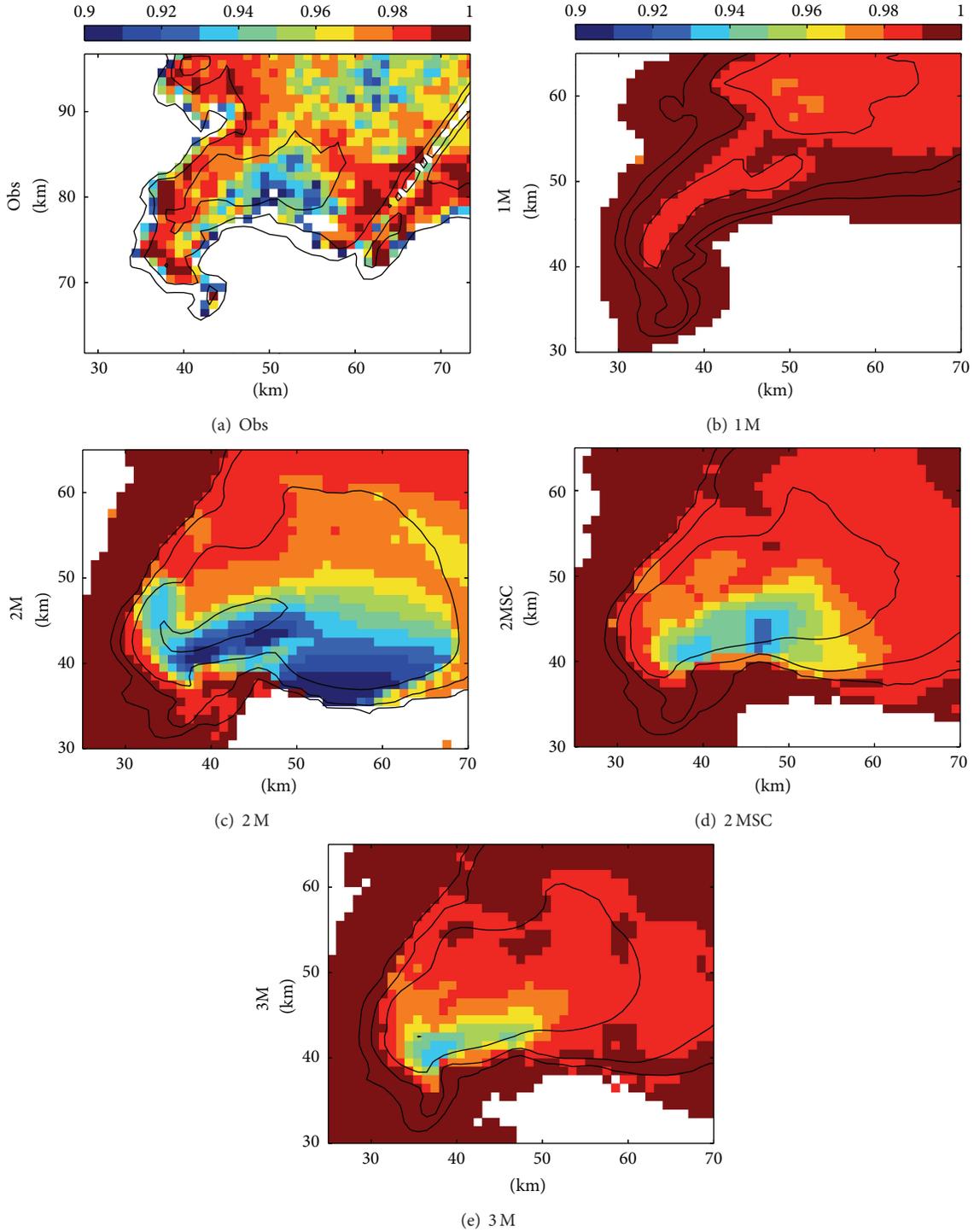


FIGURE 2: As in Figure 1 but for cross-correlation coefficient ρ_{HV} .

in the emulator, which, combined with less diagnosed water fraction (Figure 4), tends to drive Z_{DR} towards lower values [34].

These results demonstrate that a triple-moment bulk microphysics scheme, in which hydrometeor size sorting is “fully” parameterized (i.e., all three parameters of the assumed gamma size distribution can vary independently),

performs well in reproducing the salient polarimetric signatures in an analyzed supercell thunderstorm. In contrast, microphysics schemes that do not allow size sorting (single-moment schemes) or parameterize it only partially (double-moment schemes without some sort of correction mechanism to take into account the effects of the narrowing of the distribution during size sorting) exhibit degraded

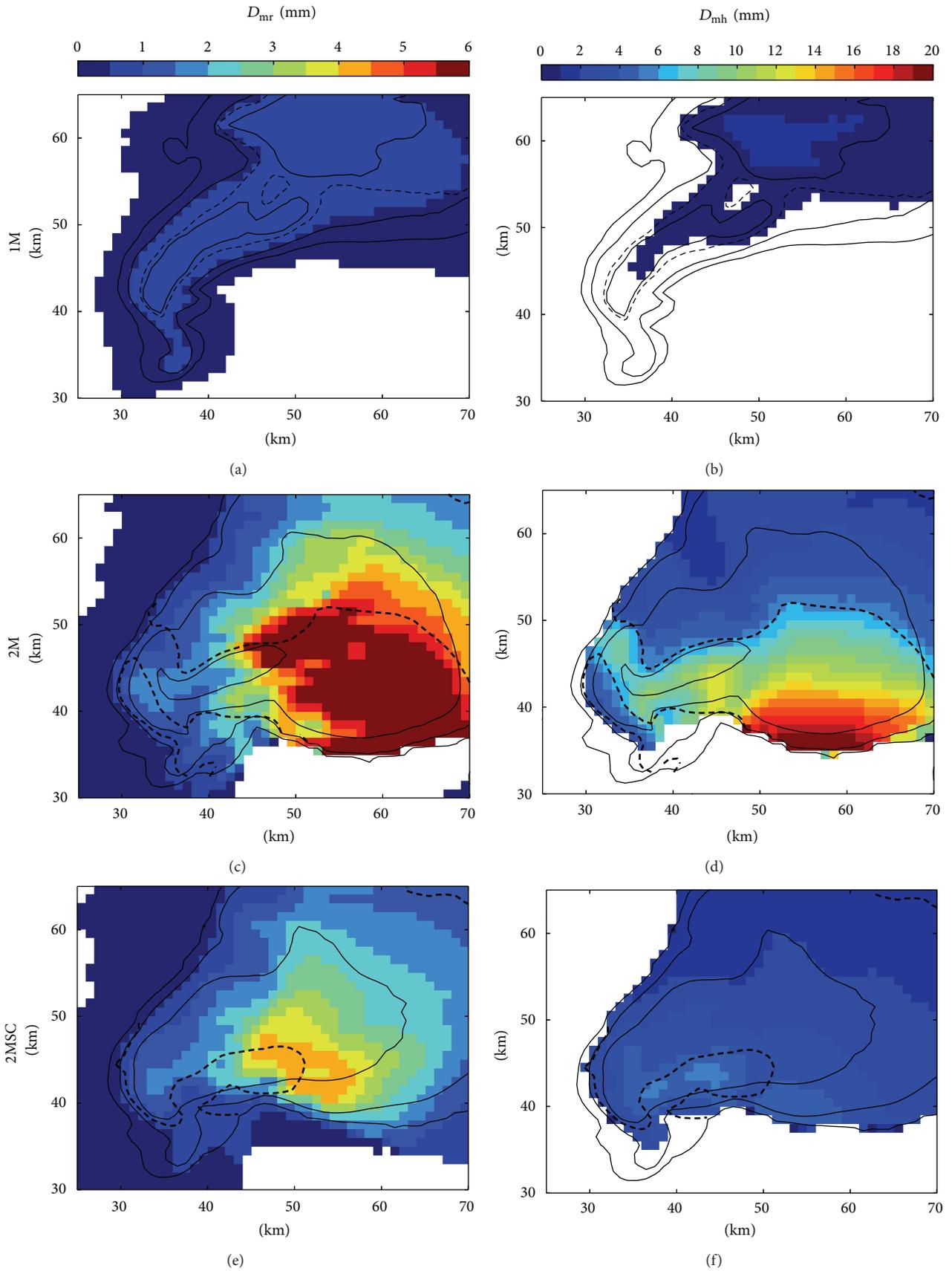


FIGURE 3: Continued.

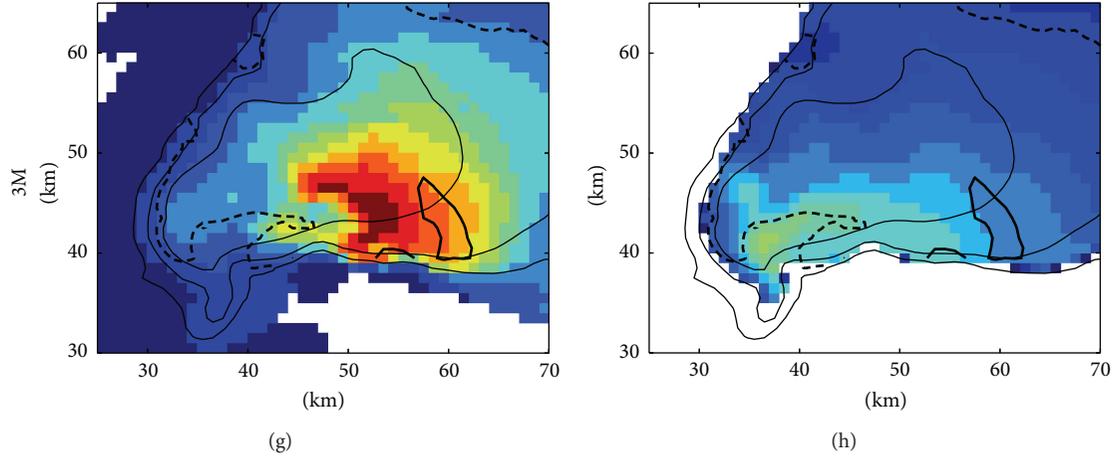


FIGURE 3: As in Figure 1, but for (left column) rain mean volume diameter D_{mr} (color fill, mm) and (right column) hail mean volume diameter D_{mh} (color fill, mm) for (a) and (b) 1 M, (c) and (d) 2 M, (e) and (f) 2 MSC, and (g) and (h) 3 M. In each panel, reflectivity Z is overlaid in 20 dBZ increments, starting at 10 dBZ (thin black contours). The 1.5 dB (bold black dashed) and 4.5 dB (bold black solid) Z_{DR} contours are also shown in each panel for reference denoting the low- Z_{DR} hail core and Z_{DR} arc, respectively.

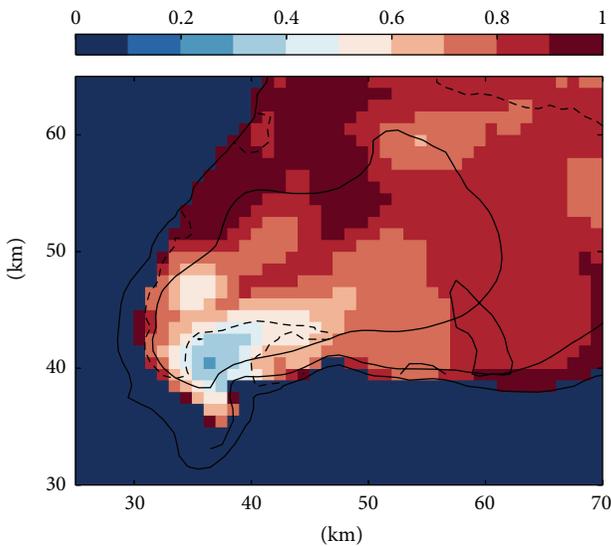


FIGURE 4: As in Figure 3, but for the diagnosed water fraction on hail (color fill) for 3 M.

polarimetric signatures as compared with observations [39, 44, 45]. The seemingly better Z in the hook region in the 1M case is not typical of pure simulation tests, and the 15 min forecast (below) shows a broadened reflectivity region more similar to the other cases. The 1M ensemble members produce much less hail and have graupel that tends to be smaller and lower in density than the multimoment cases, both of which contribute to the lower reflectivity in the hook region. These results are consistent with those of Dawson et al. [48] who similarly investigated the impact of size sorting on the low-level polarimetric signatures in idealized simulations of a different supercell thunderstorm; they also found that the triple-moment scheme, by virtue of its complete parameterization of size sorting, performed

better than the single-moment scheme as compared with observations.

3.2. Forecast Probabilistic Vorticity Swaths. To further evaluate the impact of the microphysics scheme complexity, we investigate four ensemble forecast experiments, each of which is launched from the 2200 UTC prior ensemble state of the above four EnKF experiments and run out to 2300 UTC. Stensrud and Gao [56], Dawson et al. [8], and Yussouf et al. [10] made use of probabilistic vorticity swaths to evaluate the ability of their ensemble supercell forecasts to predict the overall track of the mesocyclone. We follow the approach of these studies and compute the ensemble probability of vertical vorticity ζ at the surface and 1 km AGL exceeding a given threshold (0.01 s^{-1}) for each of the four experiments, at any time over the 1 hr forecast period. The procedure is on a grid point-by-grid point basis for each level. For example, at a given grid point, if 3 out of the 30 ensemble members exceed the given threshold of ζ , the ensemble probability at that point is 10%. In this manner, a probabilistic vorticity “swath” is derived. The 1 km level is used to indicate low-level mesocyclones, whereas the surface (lowest model level) yields an indication of how well the model can maintain the enhanced vorticity imposed by the data assimilation. The results (Figure 5) show that while each experiment shows an overall southward bias as compared with the observed tornado track (also seen in Yussouf et al. [10]), especially at the surface, nevertheless the double- and triple-moment experiments outperform the single-moment experiment, with the largest improvement seen for the surface mesocyclone (The presence of high probabilities of significant rotation at the surface is used as a proxy for the tornado track, since grid spacings of 1 km are far too coarse to explicitly resolve the tornadic circulations themselves [8, 10].) Investigation of the surface cold pool at the 16 min forecast time for a select member for each experiment (Figure 6) reveals one possible

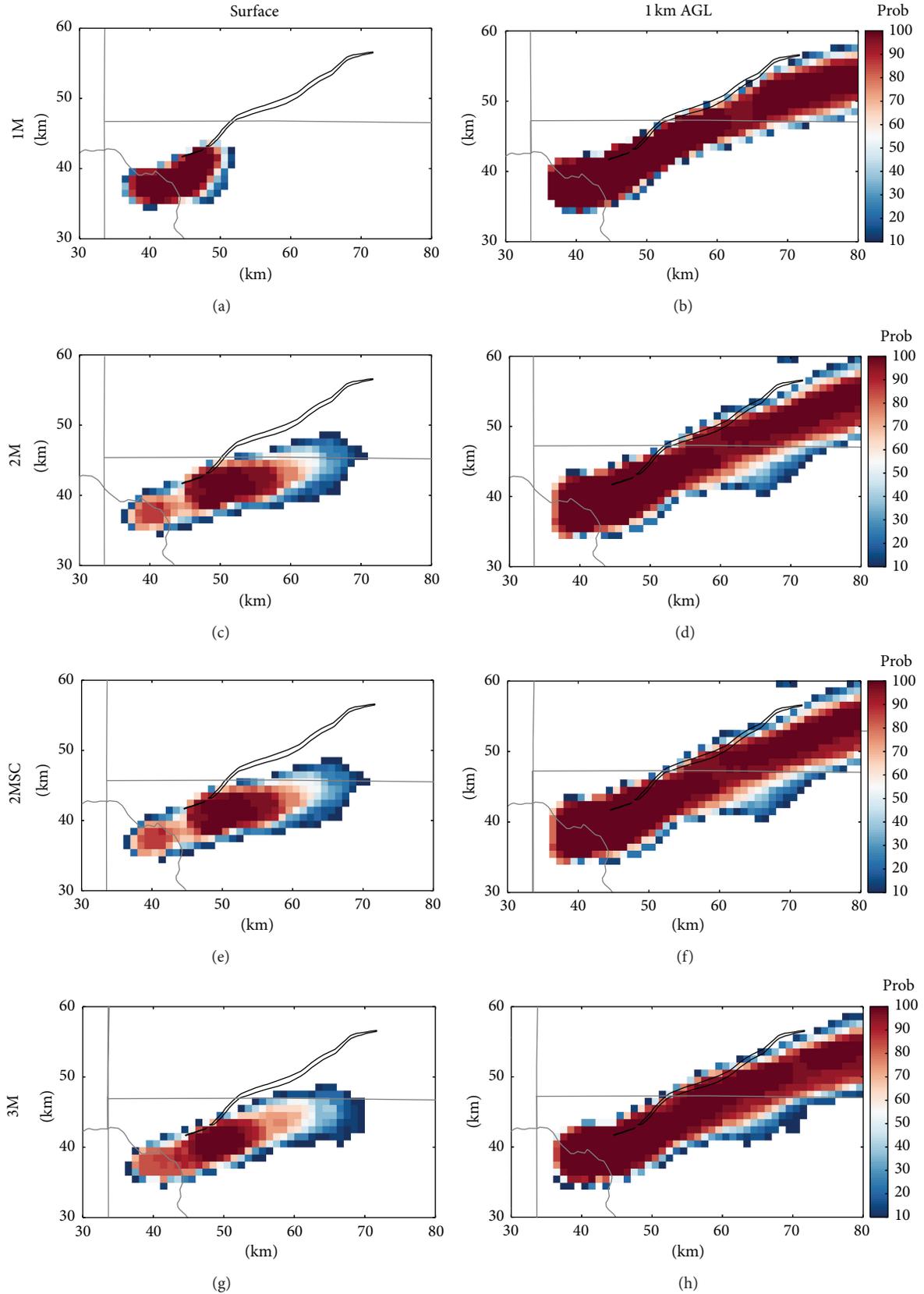


FIGURE 5: Ensemble probability of vertical vorticity ζ exceeding 0.01 s^{-1} (color fill) for the forecast period 2200–2300 UTC at (left column) the surface and (right column) 1 km AGL for each of the four ensemble forecast experiments: (a) and (b) 1M, (c) and (d) 2M, (e) and (f) 2MSC, and (g) and (h) 3M. Also overlaid in each panel is the observed May 8, 2003 F4 tornado track (bold line) and Oklahoma county boundaries (thin gray lines).

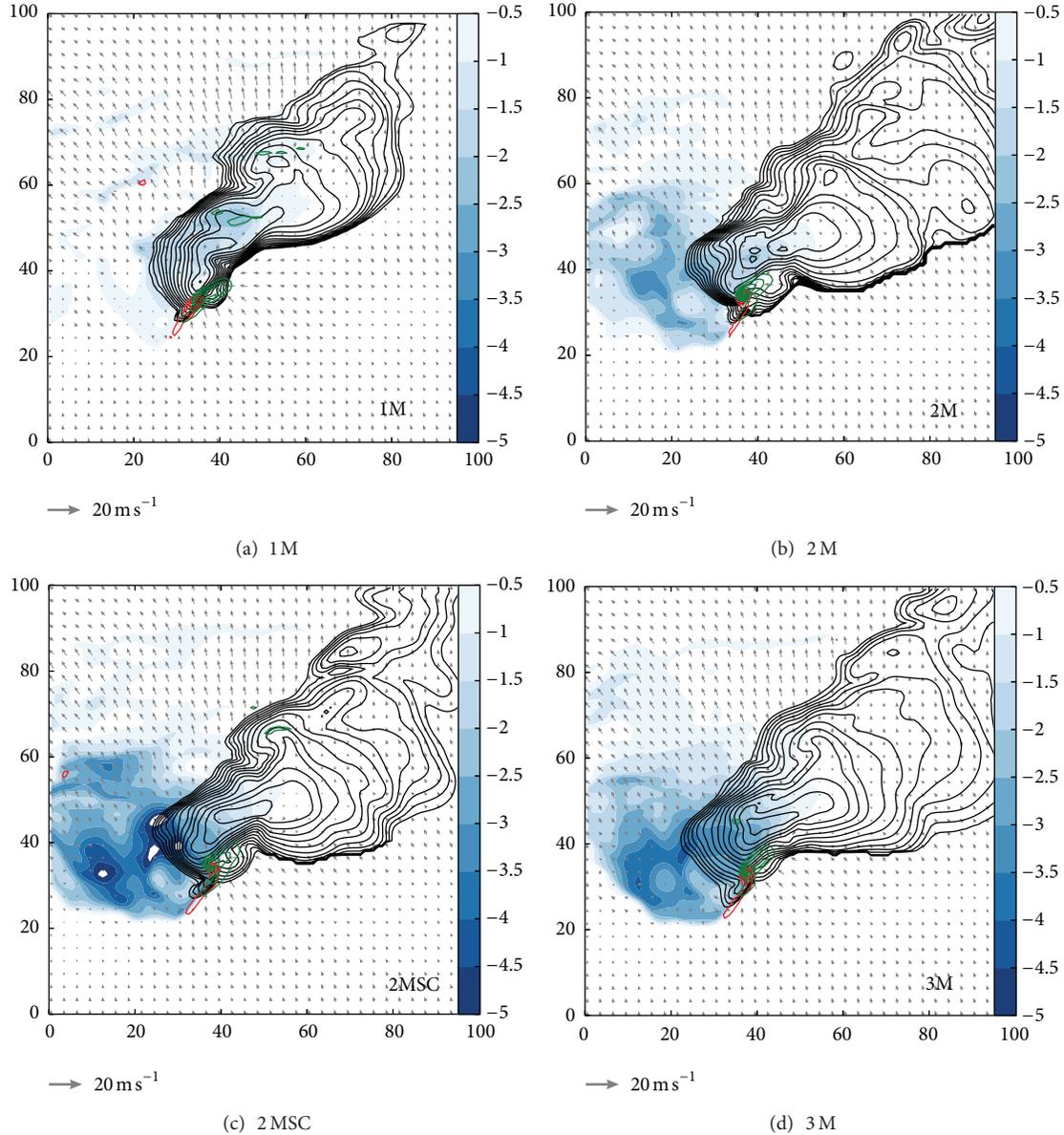


FIGURE 6: Surface potential temperature perturbation (defined relative to the horizontally homogeneous base state) θ' (color fill, K), radar reflectivity (thick black contours, 5 dBZ increment), surface and 1 km AGL vertical vorticity ζ (red and green contours, resp., 0.00375 s^{-1} increment, starting at 0.00375 s^{-1}), and storm-relative wind vectors plotted every 3 km (key in lower right) for ensemble member 13 at 2216 UTC (960 s forecast) for (a) 1M, (b) 2M, (c) 2MSC, and (d) 3M.

reason for the differences in the mesocyclone track forecast at the surface: the cold pool strength in the double- and triple-moment schemes is somewhat larger than that in the corresponding single-moment scheme. This allows stronger rotation at the surface to persist, while it is suppressed in the single-moment scheme due possibly to lack of surface convergence. These results are overall consistent with the results of Dawson et al. [8] who also found an improved track forecast for the double-moment scheme they investigated. We hasten to point out, however, that whether the single-moment scheme will exhibit stronger or weaker cold pools than the multi-moment schemes depends strongly on the choice of

tunable parameters in the scheme, particularly the intercept parameter for the various hydrometeor categories [8, 14, 18]; it may be difficult a priori to choose reasonable values, whereas the multi-moment schemes offer substantially more flexibility in allowing these parameters to vary in time and space. In any case, it appears that in the current study, the cause of the stronger cold pool in the multi-moment schemes relative to that of the single-moment scheme is tied to the presence of more hydrometeor mass in the low levels (not shown) which in turn results in more evaporation and melting. Further investigation of these differences is left for future work.

4. Conclusions

The goals of this study were to investigate the impact of varying the number of moments predicted in a multi-moment bulk microphysics scheme on EnKF analyses and ensemble forecasts of a tornadic supercell thunderstorm. We focused on two main aspects: (1) the ability of the microphysics to qualitatively reproduce the low-level polarimetric (specifically Z , Z_{DR} , and ρ_{HV}) signatures seen in the observed storm and (2) the impact on probabilistic vorticity swath forecasts (as a proxy for the tornado track). For the former aspect, the triple-moment scheme, at least in part by virtue of its complete parameterization of size sorting [48], was able to reproduce well the overall placement and magnitudes of the Z_{DR} arc and low- Z_{DR} and ρ_{HV} hail core. For the latter aspect, both the double- and triple-moment experiments performed much better than the single-moment scheme in predicting the overall track of the surface mesocyclone out to 1 hour, as compared with the observed tornado track. Though all experiments had a south-of-track bias, the single-moment scheme failed to maintain significant rotation near the surface during the duration of the forecast, while both the double- and triple-moment schemes maintained rotation. The weaker cold pool in the single-moment experiment likely led to the suppression of significant near-ground rotation due to weaker surface convergence. We stress, however, that this reported sensitivity is not general to single-moment schemes, but rather depends substantially on the *a priori* choices of fixed parameters in such schemes (typically the intercept parameters of the assumed exponential size distributions).

Our overall results indicate that multi-moment schemes, in broad agreement with several recent studies, generally outperform their single-moment counterparts when various methods of analysis and forecast performance are assessed. In the current study, the simulated polarimetric signatures and ensemble vorticity forecasts were investigated. For future work, a more robust examination of the impact of microphysics might be undertaken across multiple cases. Additionally, sensitivity to various microphysical processes that become important in a multi-moment context such as the rain drop breakup parameterization [17] is needed. The results shown here and in related work demonstrate the large sensitivity of forecast output to the choice of microphysical parameters and/or parameterization that will need to be addressed prior to any operational implementation to the warn-on-forecast paradigm [57].

Acknowledgments

This work was primarily supported by the National Research Council and National Science Foundation Postdoctoral Fellowships (AGS-1137702) awarded to the first author at the NOAA National Severe Storms Laboratory and was partially supported by NOAA/OAR under NOAA-University of Oklahoma Cooperative Agreement NA11OAR4320072, NSF Grant AGS-1046171, and by a research grant from the Weather Radar Center, Korea Meteorological Administration. Graphics were generated using Matplotlib [58].

References

- [1] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, no. 5, pp. 10143–10162, 1994.
- [2] A. Aksoy, D. C. Dowell, and C. Snyder, "A multicase comparative assessment of the Ensemble Kalman filter for assimilation of radar observations. Part I: storm-scale analyses," *Monthly Weather Review*, vol. 137, no. 6, pp. 1805–1824, 2009.
- [3] D. C. Dowell, L. J. Wicker, and C. Snyder, "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*, vol. 139, no. 1, pp. 272–294, 2011.
- [4] R. L. Tanamachi, L. J. Wicker, D. C. Dowell, H. B. Bluestein, D. T. Dawson, and M. Xue, "EnKF assimilation of high-resolution, mobile doppler radar data of the 4 May 2007 Greensburg, Kansas, supercell into a numerical cloud model," *Monthly Weather Review*, vol. 141, no. 2, pp. 625–648, 2012.
- [5] J. Marquis, Y. Richardson, P. Markowski, D. Dowell, and J. Wurman, "Tornado maintenance investigated with high-resolution dual-doppler and EnKF analysis," *Monthly Weather Review*, vol. 140, no. 1, pp. 3–27, 2012.
- [6] Y. Jung, M. Xue, and M. Tong, "Ensemble Kalman filter analyses of the 29-30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric radar data," *Monthly Weather Review*, vol. 140, no. 5, pp. 1457–1475, 2012.
- [7] D. Dowell, F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, "Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: ensemble Kalman filter experiments," *Monthly Weather Review*, vol. 132, pp. 1982–2005, 2004.
- [8] D. T. Dawson, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, "Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones," *Monthly Weather Review*, vol. 140, no. 2, pp. 696–716, 2012.
- [9] A. Aksoy, D. C. Dowell, and C. Snyder, "A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: short-range ensemble forecasts," *Monthly Weather Review*, vol. 138, no. 4, pp. 1273–1292, 2010.
- [10] N. Yussouf, E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, "The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma city tornadic supercell storm using single and double moment microphysics schemes," *Monthly Weather Review*, 2013.
- [11] J. A. Milbrandt and M. K. Yau, "A multimoment bulk microphysics parameterization. Part IV: sensitivity experiments," *Journal of the Atmospheric Sciences*, vol. 63, no. 12, pp. 3137–3159, 2006.
- [12] D. T. Dawson II, M. Xue, J. A. Milbrandt, and M. K. Yau, "Comparison of evaporation and cold pool development between single-moment and multimoment bulk microphysics schemes in idealized simulations of tornadic thunderstorms," *Monthly Weather Review*, vol. 138, no. 4, pp. 1152–1171, 2010.
- [13] H. Morrison and J. Milbrandt, "Comparison of two-moment bulk microphysics schemes in idealized supercell thunderstorm simulations," *Monthly Weather Review*, vol. 139, no. 4, pp. 1103–1130, 2011.

- [14] M. S. Gilmore, J. M. Straka, and E. N. Rasmussen, "Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme," *Monthly Weather Review*, vol. 132, no. 11, pp. 2610–2627, 2004.
- [15] S. C. van den Heever and W. R. Cotton, "The impact of hail size on simulated supercell storms," *Journal of the Atmospheric Sciences*, vol. 61, pp. 1596–1609, 2004.
- [16] H. Morrison, G. Thompson, and V. Tatarskii, "Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one- and two-moment schemes," *Monthly Weather Review*, vol. 137, no. 3, pp. 991–1007, 2009.
- [17] H. Morrison, S. A. Tessoroff, K. Ikeda, and G. Thompson, "Sensitivity of a simulated midlatitude squall line to parameterization of raindrop breakup," *Monthly Weather Review*, vol. 140, no. 8, pp. 2437–2460, 2012.
- [18] N. Snook and M. Xue, "Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms," *Geophysical Research Letters*, vol. 35, no. 24, Article ID L24803, 2008.
- [19] H. Morrison, J. A. Curry, and V. I. Khvorostyanov, "A new double-moment microphysics parameterization for application in cloud and climate models. Part I: description," *Journal of the Atmospheric Sciences*, vol. 62, no. 6, pp. 1665–1677, 2005.
- [20] E. R. Mansell, C. L. Ziegler, and E. C. Bruning, "Simulated electrification of a small thunderstorm with two-moment bulk microphysics," *Journal of the Atmospheric Sciences*, vol. 67, no. 1, pp. 171–194, 2010.
- [21] G. Thompson, P. R. Field, R. M. Rasmussen, and W. D. Hall, "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization," *Monthly Weather Review*, vol. 136, no. 12, pp. 5095–5115, 2008.
- [22] K.-S. S. Lim and S.-Y. Hong, "Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models," *Monthly Weather Review*, vol. 138, no. 5, pp. 1587–1612, 2010.
- [23] J. A. Milbrandt and M. K. Yau, "A multimoment bulk microphysics parameterization. Part II: a proposed three-moment closure and scheme description," *Journal of the Atmospheric Sciences*, vol. 62, no. 9, pp. 3065–3081, 2005.
- [24] N. Balakrishnan and D. S. Zrnica, "Use of polarization to characterize precipitation and discriminate large hail," *Journal of the Atmospheric Sciences*, vol. 47, no. 13, pp. 1525–1540, 1990.
- [25] P. H. Herzegh and A. R. Jameson, "Observing precipitation through dual-polarization radar measurements," *Bulletin of the American Meteorological Society*, vol. 73, no. 9, pp. 1365–1374, 1992.
- [26] A. V. Ryzhkov and D. S. Zrnica, "Discrimination between rain and snow with a polarimetric radar," *Journal of Applied Meteorology*, vol. 37, no. 10, pp. 1228–1240, 1998.
- [27] D. S. Zrnica and A. V. Ryzhkov, "Polarimetry for weather surveillance radars," *Bulletin of the American Meteorological Society*, vol. 80, no. 3, pp. 389–406, 1999.
- [28] J. M. Straka, D. S. Zrnica, and A. V. Ryzhkov, "Bulk hydrometeor classification and quantification using polarimetric radar data: synthesis of relations," *Journal of Applied Meteorology*, vol. 39, no. 8, pp. 1341–1372, 2000.
- [29] V. N. Bringi and V. Chandrasekar, *Polarimetric Doppler Weather Radar: Principles and Applications*, Cambridge University Press, Cambridge, UK, 2001.
- [30] D. S. Zrnica, A. Ryzhkov, J. Straka, Y. Liu, and J. Vivekanandan, "Testing a procedure for automatic classification of hydrometeor types," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 6, pp. 892–913, 2001.
- [31] S. A. Tessoroff, L. J. Miller, K. C. Wiens, and S. A. Rutledge, "The 29 June 2000 supercell observed during STEPS. Part I: kinematics and microphysics," *Journal of the Atmospheric Sciences*, vol. 62, no. 12, pp. 4127–4150, 2005.
- [32] P. L. Heinselman and A. V. Ryzhkov, "Validation of polarimetric hail detection," *Weather and Forecasting*, vol. 21, no. 5, pp. 839–850, 2006.
- [33] H. S. Park, A. V. Ryzhkov, D. S. Zrnica, and K.-E. Kim, "The hydrometeor classification algorithm for the polarimetric WSR-88D: description and application to an MCS," *Weather and Forecasting*, vol. 24, no. 3, pp. 730–748, 2009.
- [34] Y. Jung, M. Xue, and G. Zhang, "Simulations of polarimetric radar signatures of a supercell storm using a two-moment bulk microphysics scheme," *Journal of Applied Meteorology and Climatology*, vol. 49, no. 1, pp. 146–163, 2010.
- [35] M. R. Kumjian and A. V. Ryzhkov, "Polarimetric signatures in supercell thunderstorms," *Journal of Applied Meteorology and Climatology*, vol. 47, no. 7, pp. 1940–1961, 2008.
- [36] M. Hu and M. Xue, "Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case," *Monthly Weather Review*, vol. 135, no. 2, pp. 507–525, 2007.
- [37] G. S. Romine, D. W. Burgess, and R. B. Wilhelmson, "A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma city area tornadic supercell," *Monthly Weather Review*, vol. 136, no. 8, pp. 2849–2870, 2008.
- [38] D. C. Dowell and L. J. Wicker, "Additive noise for storm-scale ensemble data assimilation," *Journal of Atmospheric and Oceanic Technology*, vol. 26, no. 5, pp. 911–927, 2009.
- [39] J. A. Milbrandt and M. K. Yau, "A multimoment bulk microphysics parameterization. Part I: analysis of the role of the spectral shape parameter," *Journal of the Atmospheric Sciences*, vol. 62, no. 9, pp. 3051–3064, 2005.
- [40] L. J. Wicker and R. B. Wilhelmson, "Simulation and analysis of tornado development and decay within a three-dimensional supercell thunderstorm," *Journal of the Atmospheric Sciences*, vol. 52, no. 15, pp. 2675–2703, 1995.
- [41] M. C. Coniglio, D. J. Stensrud, and L. J. Wicker, "Effects of upper-level shear on the structure and maintenance of strong quasi-linear mesoscale convective systems," *Journal of the Atmospheric Sciences*, vol. 63, no. 4, pp. 1231–1252, 2006.
- [42] J. Sun and N. A. Crook, "Real-time low-level wind and temperature analysis using single WSR-88D data," *Weather and Forecasting*, vol. 16, no. 1, pp. 117–132, 2001.
- [43] C. L. Ziegler, "Retrieval of thermal and microphysical variables in observed convective storms. Part I: model development and preliminary testing," *Journal of the Atmospheric Sciences*, vol. 42, no. 14, pp. 1487–1509, 1985.
- [44] E. R. Mansell, "On sedimentation and advection in multimoment bulk microphysics," *Journal of the Atmospheric Sciences*, vol. 67, no. 9, pp. 3084–3094, 2010.
- [45] J. A. Milbrandt and R. Mctaggart-Cowan, "Sedimentation-induced errors in bulk microphysics schemes," *Journal of the Atmospheric Sciences*, vol. 67, no. 12, pp. 3931–3948, 2010.
- [46] U. Wacker and A. Seifert, "Evolution of rain water profiles resulting from pure sedimentation: spectral vs. parameterized description," *Atmospheric Research*, vol. 58, no. 1, pp. 19–39, 2001.

- [47] C. W. Ulbrich, "Natural variations in the analytical form of the raindrop size distribution," *Journal of Climate & Applied Meteorology*, vol. 22, no. 10, pp. 1204–1215, 1983.
- [48] D. T. Dawson II, E. R. Mansell, Y. Jung, L. J. Wicker, M. R. Kumjian, and M. Xue, "Low-level ZDR signatures in supercell forward flanks: the role of size sorting and melting of hail," *Journal of the Atmospheric Sciences*, 2013.
- [49] E. R. Mansell and C. L. Ziegler, "Aerosol effects on simulated storm electrification and precipitation in a two-moment bulk microphysics model," *Journal of the Atmospheric Sciences*, vol. 70, no. 7, pp. 2032–2050, 2013.
- [50] P. C. Waterman, "Scattering by dielectric obstacles," *Alta Frequenza*, vol. 38, pp. 348–352, 1969.
- [51] J. Vivekanandan, W. M. Adams, and V. N. Bringi, "Rigorous approach to polarimetric radar modeling of hydrometeor orientation distributions," *Journal of Applied Meteorology*, vol. 30, no. 8, pp. 1053–1063, 1991.
- [52] M. I. Mishchenko, "Calculation of the amplitude matrix for a nonspherical particle in a fixed orientation," *Applied Optics*, vol. 39, no. 6, pp. 1026–1031, 2000.
- [53] R. M. Rasmussen, V. Levizzani, and H. R. Pruppacher, "A wind tunnel and theoretical study of the melting behavior of atmospheric ice particles: III: experiment and theory for spherical ice particles of radius $>500\ \mu\text{m}$," *Journal of the Atmospheric Sciences*, vol. 41, no. 3, pp. 381–388, 1984.
- [54] M. R. Kumjian and A. V. Ryzhkov, "Storm-relative helicity revealed from polarimetric radar measurements," *Journal of the Atmospheric Sciences*, vol. 66, no. 3, pp. 667–685, 2009.
- [55] M. R. Kumjian and A. V. Ryzhkov, "The impact of size sorting on the polarimetric radar variables," *Journal of the Atmospheric Sciences*, vol. 69, no. 6, pp. 2042–2060, 2012.
- [56] D. J. Stensrud and J. Gao, "Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short-range forecasts," *Monthly Weather Review*, vol. 138, no. 4, pp. 1250–1272, 2010.
- [57] D. J. Stensrud, X. Ming, L. J. Wicker et al., "Convective-scale warn-on-forecast system: a vision for 2020," *Bulletin of the American Meteorological Society*, vol. 90, no. 10, pp. 1487–1499, 2009.
- [58] J. D. Hunter, "Matplotlib: a 2D graphics environment," *Computing in Science and Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

Research Article

The Impact of Mesoscale Environmental Uncertainty on the Prediction of a Tornadic Supercell Storm Using Ensemble Data Assimilation Approach

Nusrat Yussouf,^{1,2} Jidong Gao,² David J. Stensrud,² and Guoqing Ge³

¹ Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, OK 73072, USA

² NOAA/OAR/National Severe Storms Laboratory, National Weather Center, 120 David L. Boren Boulevard, Norman, OK 73072, USA

³ Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK 73072, USA

Correspondence should be addressed to Nusrat Yussouf; nusrat.yussouf@noaa.gov

Received 10 May 2013; Revised 1 August 2013; Accepted 22 August 2013

Academic Editor: Ming Xue

Copyright © 2013 Nusrat Yussouf et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Numerical experiments over the past years indicate that incorporating environmental variability is crucial for successful very short-range convective-scale forecasts. To explore the impact of model physics on the creation of environmental variability and its uncertainty, combined mesoscale-convective scale data assimilation experiments are conducted for a tornadic supercell storm. Two 36-member WRF-ARW model-based mesoscale EAKF experiments are conducted to provide background environments using either fixed or multiple physics schemes across the ensemble members. Two 36-member convective-scale ensembles are initialized using background fields from either fixed physics or multiple physics mesoscale ensemble analyses. Radar observations from four operational WSR-88Ds are assimilated into convective-scale ensembles using ARPS model-based 3DVAR system and ensemble forecasts are launched. Results show that the ensemble with background fields from multiple physics ensemble provides more realistic forecasts of significant tornado parameter, dryline structure, and near surface variables than ensemble from fixed physics background fields. The probabilities of strong low-level updraft helicity from multiple physics ensemble correlate better with observed tornado and rotation tracks than probabilities from fixed physics ensemble. This suggests that incorporating physics diversity across the ensemble can be important to successful probabilistic convective-scale forecast of supercell thunderstorms, which is the main goal of NOAA's Warn-on-Forecast initiative.

1. Introduction

The development and evolution of severe thunderstorm events are strongly tied to the environment, and therefore incorporating mesoscale environmental variability and its uncertainty is crucial for successful convective-scale data assimilation and forecasts [1–3]. Several studies illustrate the importance of incorporating the influence of environmental variability and mesoscale forcing on the storm scale flows for accurate prediction of tornadic supercell thunderstorms ([4, 5]). In particular, when Stensrud and Gao [4] use a more realistic inhomogeneous mesoscale environment as initial and boundary conditions for their convective-scale three-dimensional variational (3DVAR) data assimilation

and forecast system, substantial improvement in forecast accuracy is obtained over a similar convective-scale system using a homogeneous, single-sounding environment, which is typical of idealized storm modeling studies. Yussouf et al. [6] investigate the benefits of using a combined mesoscale-convective scale cycled ensemble data assimilation and prediction system to investigate the accuracy of a very short-range (0-1 h) ensemble forecast of a tornadic supercell storm. The same suite of physics parameterization schemes is applied to the members of the mesoscale ensemble, which is used to provide environmental initial and boundary conditions for a convective-scale ensemble system, and the results are very encouraging. The convective-scale ensemble captures the structure and propagation of the main supercell storm

and predicts the probability of a strong low-level vorticity track for the tornadic supercell that correlates well with the observed rotation track.

However, while providing mesoscale environmental variability is critical to severe weather forecasts, model bias errors due to the uncertainties associated with the physical parameterization schemes are unavoidable and are a known problem in convective-scale forecasting [7, 8]. Romine et al. [9] show that using the same set of physical parameterization suites across mesoscale ensemble members leads to unique bias errors, and when these mesoscale ensembles are used as a background field for convective-scale model, the forecast skill and accuracy degrade. Due to our limited understanding of atmospheric processes, it is likely that the model physics parameterizations schemes will face challenges in some convective environments. Removing model biases in a data assimilation system is very difficult and is an active area of research [10]. One approach to account for the model biases due to its uncertainties associated with physics parameterizations schemes is to allow for the inclusion of multiple physical parameterization schemes amongst the ensemble members [11]. Fujita et al. [12] find that an ensemble with both physics and initial condition uncertainties shows considerable improvement in forecasts of storm environment with improved location and intensity of drylines, frontal boundaries, and planetary boundary layer height and structure. Since the quality of convective-scale analyses and forecasts is so sensitive to background environmental variability, using an ensemble that contains uncertainties in both initial and model physics parameterization schemes is important and can positively impact the forecasts of convective events.

To study the impact of environmental variability and its uncertainty in the forecasts of severe thunderstorm events, an ensemble-based mesoscale and convective-scale data-assimilation and prediction system is developed for May 8, 2003, Oklahoma City (OKC), Oklahoma (OK) tornadic supercell storm. The OKC tornado is one of the most destructive events that occurred during a multiday tornado outbreak across the central and eastern United States in early May 2003 [13] and several data assimilation and forecast studies have focused on this particular storm [6, 14–16]. Two 36-member ensemble data assimilation experiments are conducted at mesoscale resolution to provide background environments for convective-scale ensembles: a FixedPhysics ensemble with the same set of physics parameterization schemes amongst the members [6, 9] and a MultiPhysics ensemble with members having different physical parameterization schemes to account for model physics uncertainty [12, 17–19]. In addition, each member from the two ensemble systems has slightly perturbed initial conditions to account for uncertainties in the atmospheric state. These two mesoscale ensembles are used to provide the initial and boundary conditions for the convective-scale ensemble data assimilation system centered in OKC and covering parts of the surrounding states of Kansas, Missouri, Arkansas, and Texas.

The main objective of this study is to investigate the accuracy of a very short-range (0–1 h) ensemble forecast of the OKC tornadic storm due to two different ensemble depictions of storm environmental conditions. A brief overview of

the OKC tornadic supercell thunderstorm event followed by the experiment design for both the mesoscale and convective-scale data assimilation systems is discussed in Section 2. Section 3 assesses the quantitative and qualitative results of the forecasts from the ensembles. A discussion of key results is found in Section 4.

2. Experiment Design

2.1. Overview of the Event. On May 8, 2003, a violent tornado passed through portions of Moore, a suburban city south of OKC, as well as the southeast OKC metropolitan area with F4 damage reported along its path. Prior to tornado formation in the mid- to late afternoon, the synoptic scale environment became increasingly favorable for severe tornadic thunderstorms [14, 20]. At around 2050 UTC, several small cells initiated along the dryline in west central Oklahoma with one of the cells maturing into an isolated supercell storm by 2130 UTC. Over the next hour, this supercell moved northeastward and intensified. A violent tornado developed around 2210 UTC and tracked east-northeastward for about 30 km until it dissipated at around 2238 UTC, leaving a damage path stretching from Moore to Choctaw, Oklahoma (Figure 1(b)). The National Weather Service (NWS) Office in Norman OK issued a tornado warning for the path of the storm, including Cleveland, McClain, and south Oklahoma counties at 2149 UTC, with approximately 21-minute lead time for Moore in Cleveland County and approximately 30-minute lead time for citizens in Oklahoma County.

2.2. Mesoscale Ensembles and Cycled EnKF Data Assimilation System. The Advanced Research Weather Research and Forecasting (WRF-ARW core version 3.3.1; [21]) model is used to create the mesoscale ensemble data assimilation system. The model domain covers the continental United States (Figure 1(a)) with a horizontal grid spacing of 12 km and 51 vertical grid levels with vertically stretched grids from the surface to 50 hPa aloft. Two sets of 36-member ensembles are initialized at 1200 UTC May 8, 2003, using The National Centers for Environmental Prediction's (NCEP) Eta model for the ensemble mean initial and boundary conditions. Random samples of the horizontal components of wind, water vapor mixing ratio, and temperature are drawn from a default background error covariance file estimated by the NMC method [22] using the WRF data assimilation software. These samples are then added to each ensemble member to account for uncertainties in the initial and boundary conditions [23]. One experiment uses the same sets of physics parameterization schemes (FixedPhysics) across all 36 ensemble members. The physics options used are Thompson [24] for microphysics, Tiedtke [25, 26] for cumulus parameterization, YSU [27] for planetary boundary layer parameterization, RRTMG for both longwave and shortwave radiation, and Noah [28] for the land surface parameterization scheme. A second 36-member ensemble experiment uses different combinations of physics schemes (MultiPhysics) amongst the ensemble members to address the uncertainties in model physics parameterization schemes (e.g., [11, 12, 18, 19]). The diversity in physics options

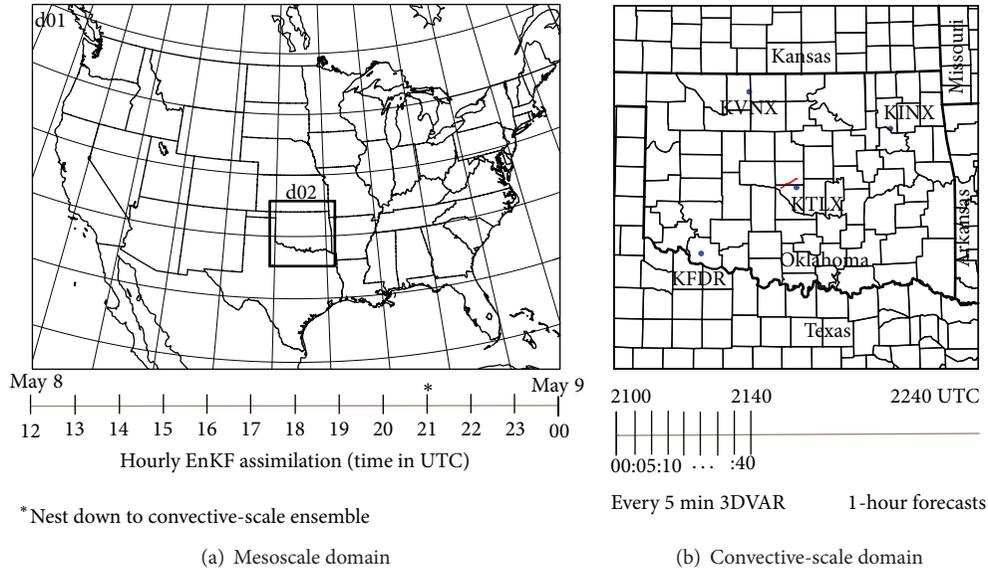


FIGURE 1: (a) The mesoscale domain (d01) covering the CONUS, the nested convective-scale domain (d02), and the time line of the hourly mesoscale data assimilation experiments; (b) the convective-scale domain with county borders (d02, enlarged), location of the four WSR-88D radars (blue dots), the NWS surveyed storm damage path (in red), and the time line of the convective-scale data assimilation and forecast experiments.

includes land surface, planetary boundary layer, radiation, convection, and microphysical parameterizations schemes and is shown in Table 1.

Both ensemble systems assimilate routinely available observations from NOAA’s Meteorological Assimilation Data Ingest System (MADIS) every hour starting at 1300 UTC May 8, 2003, and extending to 0000 UTC May 9, 2003 (Figure 1(a)), using the ensemble adjustment Kalman filter (EAKF; [29]) within the Data Assimilation Research Testbed (DART) software [30, 31].

A half radius of 230 km in the horizontal and a half radius of 4 km in the vertical are used for the covariance localization function (the fifth order correlation function from [32]). The observations assimilated in the ensembles are the surface altimeter setting, pressure, temperature, dewpoint, and horizontal wind components from land and marine surface stations, rawinsondes, and aircraft. The predicted variables updated by the data assimilation scheme include the three wind components, perturbation temperature, perturbation geopotential, perturbation surface pressure of dry air, potential temperature tendency due to microphysics, water vapor, and hydrometeors. Also updated are the 10 m wind fields, 2 m temperature and water vapor fields, and total surface pressure variables, which are diagnosed by the surface and boundary layer schemes using state variables on the model grid. The FixedPhysics and MultiPhysics mesoscale ensemble analyses are then used to create the initial background and boundary conditions for their associated convective-scale ensembles.

2.3. Convective-Scale Ensembles and Cycled 3DVAR Data Assimilation System. The model used for the two convective-scale ensemble data assimilation and forecasts experiments is the Advanced Regional Prediction System (ARPS; [33, 34])

and its 3DVAR [4, 14, 35–37] and cloud analyses scheme [14, 38]. The ARPS 3DVAR system has been successfully used in NOAA’s Hazardous Weather Testbed (HWT) Spring Forecast experiments [39–41] for the past several years to analyze and detect convective-scale severe weather events [42]. Two 36-member convective-scale ensembles are initialized from the FixedPhysics and MultiPhysics mesoscale ensemble analyses at 2100 UTC. Thus, the mesoscale ensembles provide environmental background fields and boundary conditions for their associated convective-scale 3DVAR data assimilation system. The convective-scale domain is centered in OKC using 3 km horizontal grid spacing with $192 \times 192 \times 50$ grid points and is selected such that sufficient distance is maintained between the supercell storm and lateral boundaries (Figure 1(b)). Radar observations are assimilated into each of the individual convective-scale ensemble members using the 3DVAR system. The convective-scale ensembles are referred to as FixedPhysics and MultiPhysics in reference to the mesoscale ensemble system that provides the initial and boundary conditions. The physics options used for both FixedPhysics and MultiPhysics convective-scale ensembles are identical and include Lin et al. [43] for microphysics, Noah [28] for land surface, Mellor-Yamada-Janjic (MYJ; [44, 45]) for planetary boundary layer, Dudhia [46] for shortwave, and RRTM [47] for longwave radiation parameterization schemes. Cumulus parameterization is turned off for the convective-scale ensemble. The only differences in the two convective-scale ensembles result from the use of different mesoscale environmental conditions provided by either the FixedPhysics or MultiPhysics 12 km mesoscale ensembles.

Reflectivity and radial velocity observations from four operational Weather Surveillance Radar-1988 Doppler (WSR-88D) radars located at Vance Air Force Base (KVNK),

TABLE 1: Physics options for the MultiPhysics and FixedPhysics WRF mesoscale ensemble system.

Member	Cumulus	MicroPhysics	PBL	Land surface	LW/SW Rad.
MultiPhysics ensemble					
1			YSU		
2			MYJ		RRTM/Dudhia
3			MYNN		
4			ACM2		
5			YSU		
6	BMJ	Thompson	MYJ	Noah	RRTMG/RRTMG
7			MYNN		
8			ACM2		
9			YSU		
10			MYJ		New Goddard/New Goddard
11			MYNN		
12			ACM2		
13			YSU		
14			MYJ		RRTM/Dudhia
15			MYNN		
16			ACM2		
17			YSU		
18	GD	Thompson	MYJ	Noah	RRTMG/RRTMG
19			MYNN		
20			ACM2		
21			YSU		
22			MYJ		New Goddard/New Goddard
23			MYNN		
24			ACM2		
25			YSU		
26			MYJ		RRTM/Dudhia
27			MYNN		
28			ACM2		
29			YSU		
30	Tiedtke	Thompson	MYJ	Noah	RRTMG/RRTMG
31			MYNN		
32			ACM2		
33			YSU		
34			MYJ		New Goddard/New Goddard
35			MYNN		
36			ACM2		
FixedPhysics ensemble					
1-36	Tiedtke	Thompson	YSU	Noah	RRTMG/RRTMG

Twin Lakes (KTLX), Tulsa (KINX), and Frederick (KFDR) are assimilated into the two convective-scale ensembles (Figure 1(b)). The radar observations are processed using the 88D2ARPS software with the necessary quality control steps, including velocity dealiasing and ground clutter removal [48]. The quality controlled radar observations are then projected into the model grid space in the form of a series of column observations. In order to mitigate the negative impact of small spurious cells, the noisy data in the radar observations are discarded if the reflectivity is smaller than 25 dBZ. The latent heat (LH) release based method from

the ARPS cloud analysis package is used for in-cloud temperature adjustment and all hydrometeor variables are updated during every analysis in the assimilating window. The ARPS 3DVAR uses the radar radial velocity and Oklahoma Mesonet [49] surface observations of temperature, pressure, wind speed and direction, and dewpoint temperature to update the three wind components (u , v , and w), potential temperature (θ), pressure (p), and water vapor mixing ratio (q_v), while the cloud analysis procedure uses the reflectivity observations to update the hydrometeor variables and adjust the in-cloud temperature and moisture fields. Additional quality control

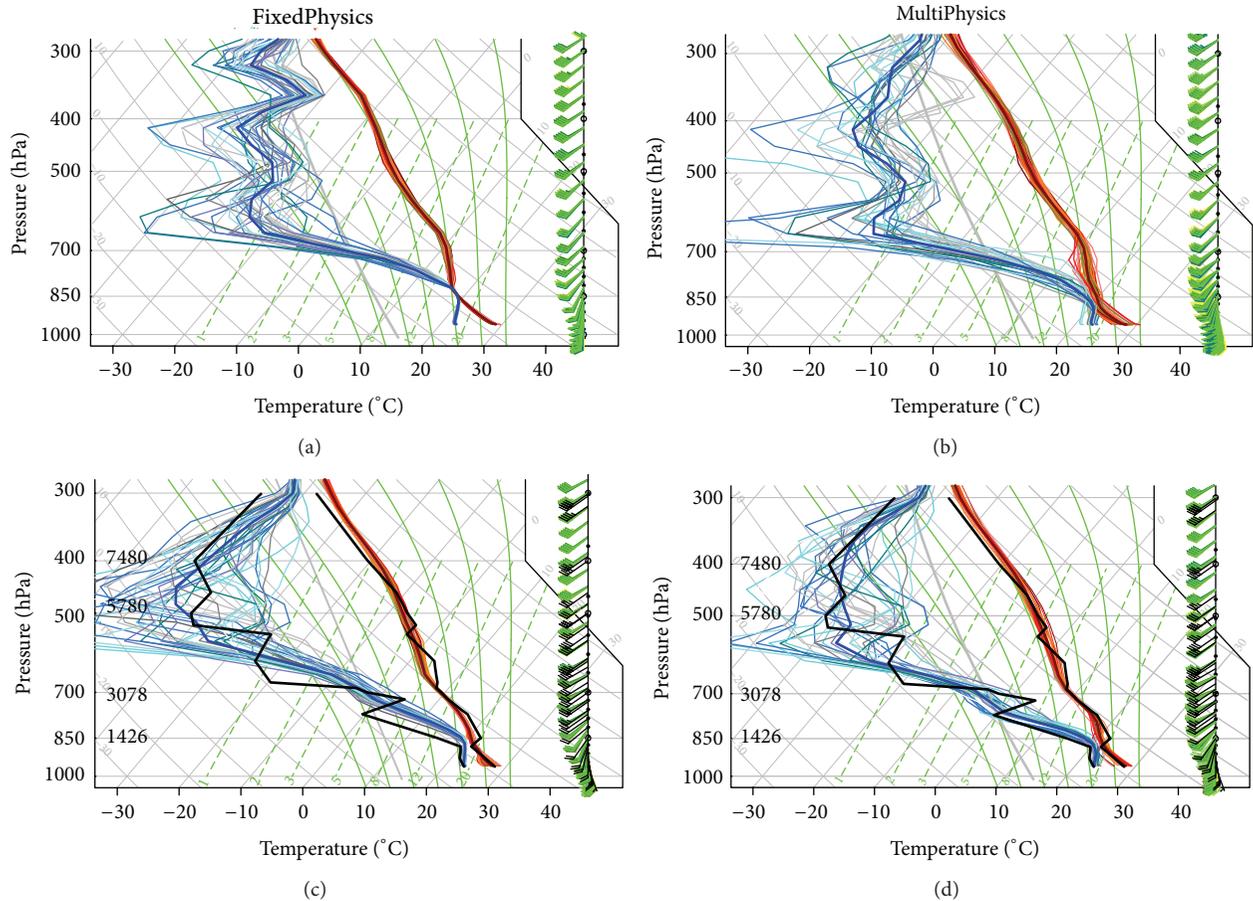


FIGURE 2: Environmental soundings from the 36-member ensembles with temperature (thin red lines) and dewpoint temperature (thin blue lines) at 2100 UTC May 8 ((a) and (b)) and 0000 UTC May 9, 2003 ((c) and (d)) from the Oklahoma City (KOKC) station from the FixedPhysics and MultiPhysics mesoscale ensemble data assimilation systems. Overlaid are the ensemble mean (thick lines) and radiosonde observations (black) at 00 UTC May 9, 2003 ((c) and (d)).

of the radial velocity observations is conducted during ARPS 3DVAR assimilation, such that if the absolute difference between a gridded radial velocity and the background is too high (greater than 20 m s^{-1}), that observation is rejected. Radar observations are assimilated into each member of the FixedPhysics and MultiPhysics ensemble members via a 5 min cycling procedure that lasts for a 40 min period starting at 2100 UTC and ending at 2140 UTC with a total of nine assimilation cycles. Each cycle begins with an application of the 3DVAR and cloud analysis, followed by a 5 min ARPS forecast, which is then used as the background for the next 3DVAR and cloud analysis. One-hour ensemble forecasts are launched from each of the 36 convective-scale ensemble analyses valid at the end of the cycling period at 2140 UTC. This time is 30 min prior to the time the OKC tornado first developed in the city of Moore, Oklahoma.

3. Results

The accuracy of the forecasts from both mesoscale and convective scale experiments using either FixedPhysics or MultiPhysics ensemble is evaluated using both quantitative

and qualitative perspectives. Statistical measures include root-mean-square (RMSE) error, bias (forecast observations), and equitable threat scores (ETS) [50]. The environmental soundings, dryline structures, significant tornado parameter (STP), and forecast probability of low-level updraft helicity track from the two ensemble systems are also compared to quantify the accuracy of the storm forecasts using two different inhomogeneous mesoscale storm environments.

3.1. Environmental Soundings from the Mesoscale Ensembles.

Soundings from the mesoscale ensembles at 2100 UTC May 8, and 0000 UTC May 9, 2003, from Oklahoma City (KOKC) indicate that the two ensembles produce different storm environments (Figure 2). The environmental soundings at 2100 UTC, the time when convective-scale ensembles are initialized from the two mesoscale ensembles, show noticeable differences between the two ensemble systems (Figures 2(a) and 2(b)). The soundings from the MultiPhysics ensemble that incorporates physics parameterizations diversity across the members show larger variability amongst the members than those from the FixedPhysics ensemble with the same single suite of parameterization schemes among the members. All

36 ensemble members from the FixedPhysics ensemble show saturated air around 850 hPa while the MultiPhysics ensemble members show greater variability in temperature and humidity from the surface to 700 hPa. The winds in the lowest 3 km are also more variable in MultiPhysics, with stronger backing of the surface winds in MultiPhysics. Due to the lack of radiosonde observations at 2100 UTC, it is not known which soundings are more realistic. Soundings from the two ensemble systems later in the evening at 0000 UTC show that both ensemble systems fail to capture the observed capping inversion (Figures 2(c) and 2(d)). Accurately capturing the capping inversion is a common forecasting problem faced by the modeling community. However, the larger variability within MultiPhysics captures the observed temperature and moisture profiles within the member envelope for most vertical levels, an improvement over that seen from FixedPhysics. The observations more often lie on the edge or outside the ensemble envelope for the FixedPhysics experiment.

3.2. Location of Drylines in the Convective-Scale Ensembles.

The forecast locations of the dryline—the feature that helped initiate the OKC supercell storm—and their associated dryline bulges also are important to compare between the two ensembles. Isolines of 10°C 2 m dewpoint temperature forecasts (a reasonable proxy for dryline location) from each member of FixedPhysics and MultiPhysics along with the analyzed isoline from Oklahoma Mesonet observations are shown in Figure 3. The MultiPhysics ensemble has dryline bulges (areas where dry air is advancing eastward more rapidly yielding an eastward bulge in the isodrosotherm) in Oklahoma as early as 10 min into the forecasts at 2150 UTC, in reasonable agreement with observations, while the FixedPhysics ensemble has no dryline bulges at this time (Figures 3(a)–3(c)). The MultiPhysics ensemble captures the dryline location within the ensemble envelope better than the FixedPhysics ensemble throughout the 1 h forecast (Figures 3(d)–3(l)). Most importantly, the MultiPhysics ensemble also produces two distinct dryline bulges in Oklahoma that compare well with the two observed dryline bulges. Dryline bulges are an indication of the development of deep moist convection and they develop due to enhanced low-level convergence, helping parcels reach their level of free convection [51].

3.3. Forecast Error Statistics of Near Surface Variables.

Bias and RMSE of 2 m temperature, 2 m dewpoint temperature, and 10 m wind speed are calculated from the two convective-scale ensembles and corresponding Oklahoma Mesonet observations at 5 min intervals using the 112 available Mesonet observations stations within the model domain (Figure 4). The RMSE from the MultiPhysics ensemble is smaller compared to the FixedPhysics ensemble throughout the entire forecast period for both 2 m temperature and 2 m dewpoint temperature (Figures 4(a) and 4(b)). The differences in the magnitude of the RMSE errors are as high as 0.35°C and 0.60°C for 2 m temperature and 2 m dewpoint temperature, respectively, at the beginning of the forecast period, with the differences reducing to 0.08 and 0.10,

respectively, at the end of forecast period. The differences in the RMSE values between the two ensembles for the 10 m wind speed are very small with slightly smaller values for the MultiPhysics ensemble (Figure 4(c)). These results are consistent with the findings from Fujita et al. [12] and Zhiyong and Zhang [17] in which the benefits of a MultiPhysics ensemble over a single-scheme ensemble are found to be more pronounced in the thermodynamic variables than in the wind fields. For 2 m temperature, the FixedPhysics ensemble has a larger warm bias while the MultiPhysics ensemble has a smaller cold bias at all forecast times. For 2 m dewpoint temperature both FixedPhysics and MultiPhysics ensembles have a moist bias with the bias from MultiPhysics being larger. However, for the 10 m wind speed, the bias in MixedPhysics is consistently larger than that in FixedPhysics. These statistics indicate that using physics diversity across the ensemble can have a positive impact on the forecast of near surface thermodynamic variables but a mixed impact on the forecast of near surface wind field.

3.4. Ensemble Mean Forecasts of Significant Tornado Parameter (STP).

One of the severe weather parameters used to evaluate tornadic supercell environments by the NOAA/NWS/Storm Prediction Center is the significant tornado parameter (STP; [52]). The STP helps discriminate between significantly tornadic (F2 or greater damage) and nontornadic supercell environments, with proximity soundings yielding STP values greater than 1 in association with a majority of F2 or greater tornadic supercell storms. The STP equation is defined as

$$\text{STP} = \frac{\text{CAPE}}{1000 \text{ J} \cdot \text{kg}^{-1}} \times \frac{\text{SHR}}{20 \text{ m} \cdot \text{s}^{-1}} \times \frac{\text{SREH}}{100 \text{ m}^2 \cdot \text{s}^{-2}} \times \frac{(2000 \text{ m} - \text{LCL})}{1500 \text{ m}} \times \frac{(150 \text{ J} \cdot \text{kg}^{-1} + \text{CIN})}{125 \text{ J} \cdot \text{kg}^{-1}}, \quad (1)$$

where CAPE is the convective available potential energy, SHR is 0–6 km vector vertical shear magnitude, SREH is 0–1 km storm-relative helicity, CIN is convective inhibition, and LCL is the lifting condensation level. The ensemble-mean forecast of STP derived from MultiPhysics ensemble at 2150 (20 minutes prior to tornadogenesis) is very large around the OKC area, with values approaching 50, suggestive of a severe storm environment with significant tornado threat (Figure 5(b)). Thompson et al. [52] show that the largest values of STP are below 10 when using proximity soundings from the hourly 40 km Rapid Update Cycle-2 (RUC-2), suggesting that the high temporal frequency 3 km convective-scale model forecasts over an area that includes both the supercell storm and its surrounding environments may be providing new and useful information. The maximum value of STP continues to increase over the next 20 minutes out to 2200 UTC (Figure 5(d)) with values higher than 75. By the time the observed tornado forms at ~2210 UTC, the values of STP start to decrease (Figures 5(f) and 5(h)). In contrast, the FixedPhysics ensemble generates smaller STP values around OKC at 2150 UTC indicating a less favorable storm environment (Figure 5(a)). By 2200 UTC, the FixedPhysics

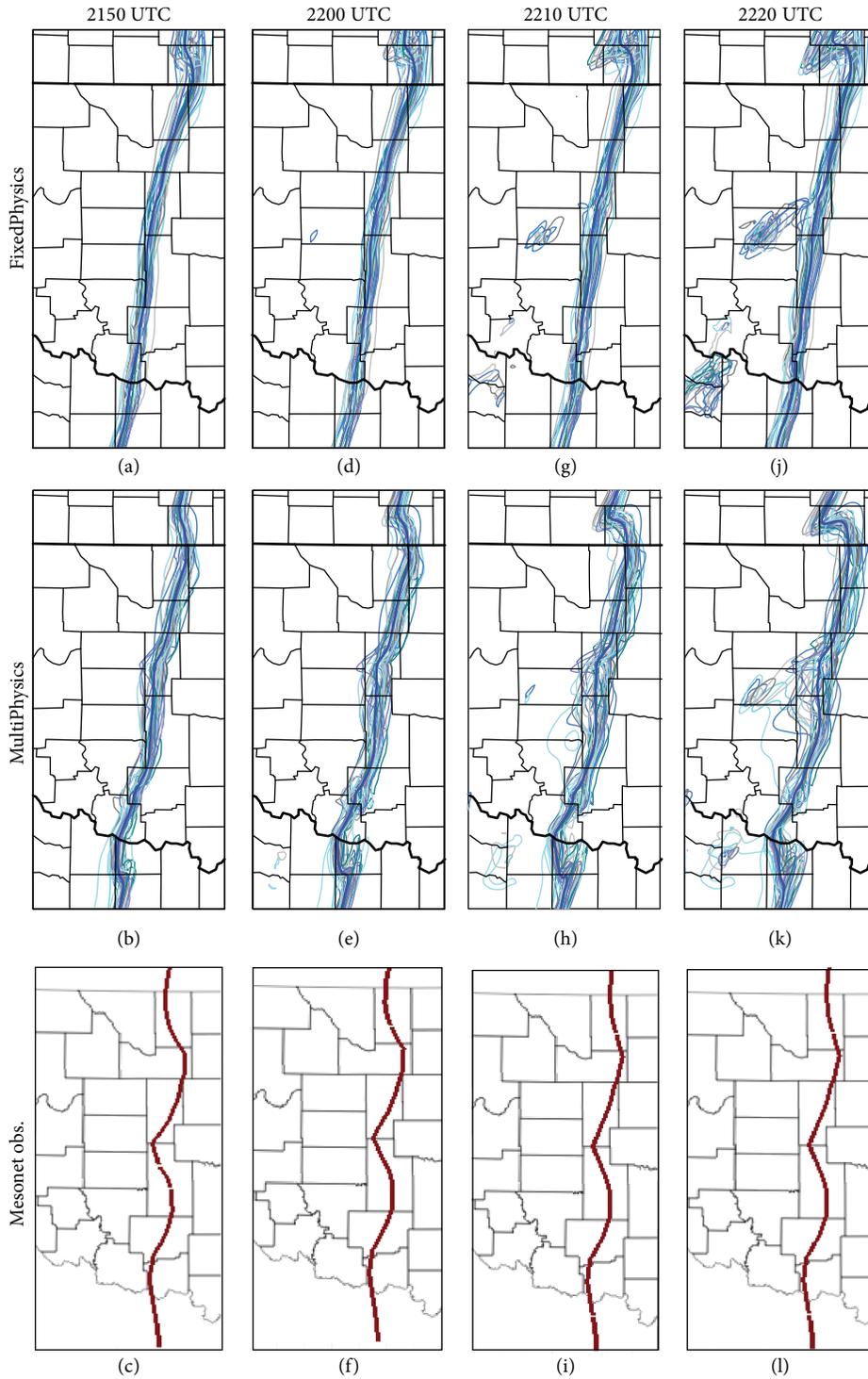


FIGURE 3: Isolines of 10°C 2 m dewpoint temperature forecasts from FixedPhysics and MultiPhysics convective-scale ensemble members (thin blue lines), ensemble mean (thick blue lines), and Oklahoma mesonet observations (red line) at ((a), (b), (c)) 2150, ((d), (e), (f)) 2200, ((g), (h), (i)) 2210, and ((j), (k), (l)) 2220 UTC May 8, 2003. The portion of the domain shown here is 201×435 km wide.

ensemble produces high values of STP in south central Kansas (Figure 5(c)) indicating severe tornadic environment in that area and small values of STP around OKC area. Thus, the STP values from the FixedPhysics ensemble could have

diverted forecasters attention to the north of Kansas where no significant tornadoes were observed until over 30 min after the end of the forecast period. The behavior of the large magnitude STP fields to the south of regions of forecast

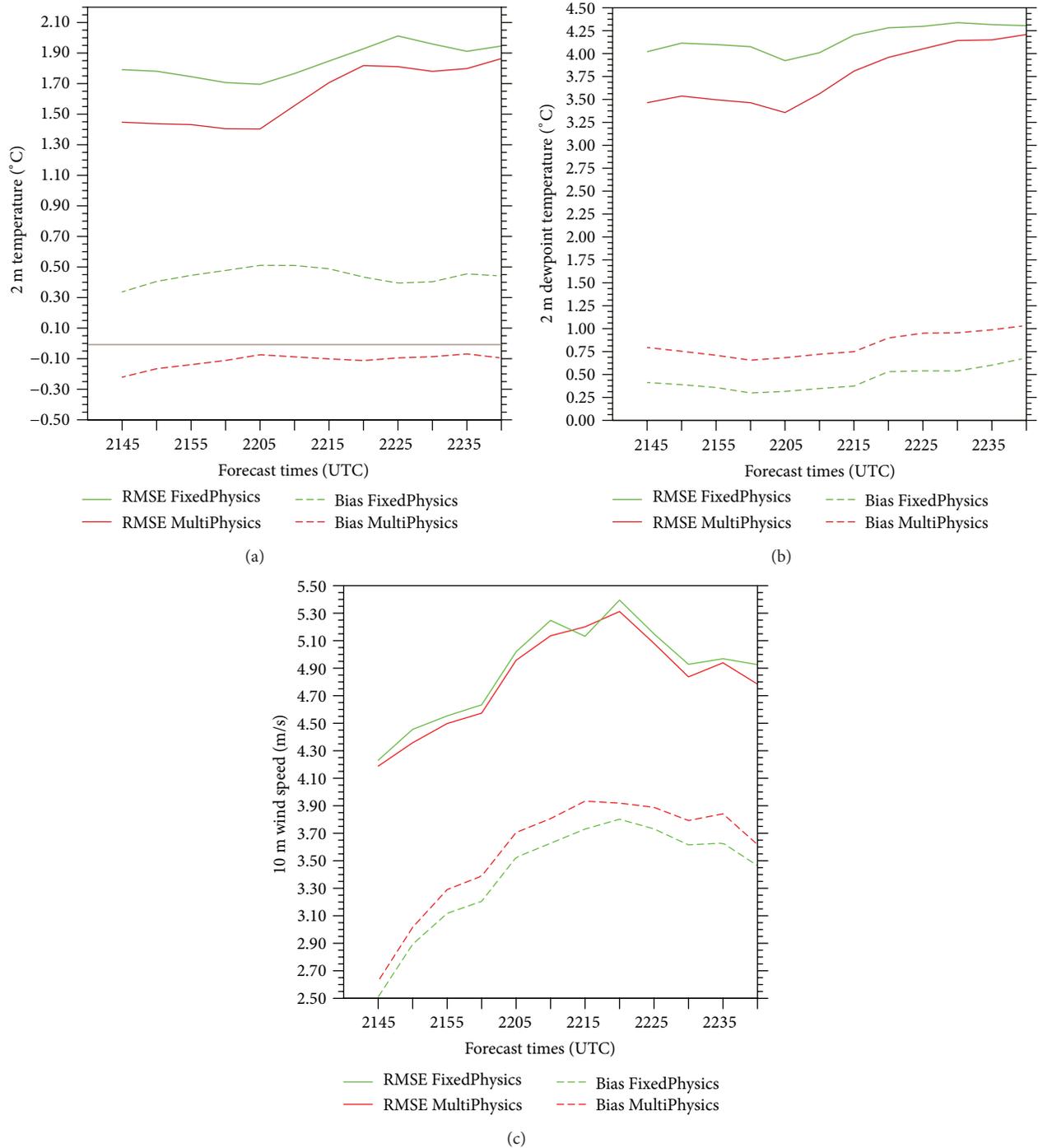


FIGURE 4: The time series of RMSE and bias (forecast observations) during 1 h forecast period for (a) 2 m temperatures ($^{\circ}\text{C}$), (b) 2 m dewpoint temperature ($^{\circ}\text{C}$), and (c) 10 m wind speed (m s^{-1}) for the FixedPhysics (green lines) and MultiPhysics (red lines) convective-scale ensemble system.

convection suggests that they are produced by the model supercells modifying the surrounding environment as also seen in Brooks et al. [53]. The correlation coefficient between the maximum values of STP surrounding the supercell region and the maximum values of 0–3 km updraft helicity within the storm (a measure of low-level storm intensity) during the

forecast period is 0.86 for FixedPhysics ensemble and 0.95 for MultiPhysics ensemble. These high correlations suggest that the intensity of the environmental modification is related to the intensity of the low-level mesocyclone. This relationship deserves further study to evaluate whether or not it could be used to evaluate the likelihood of tornado formation.

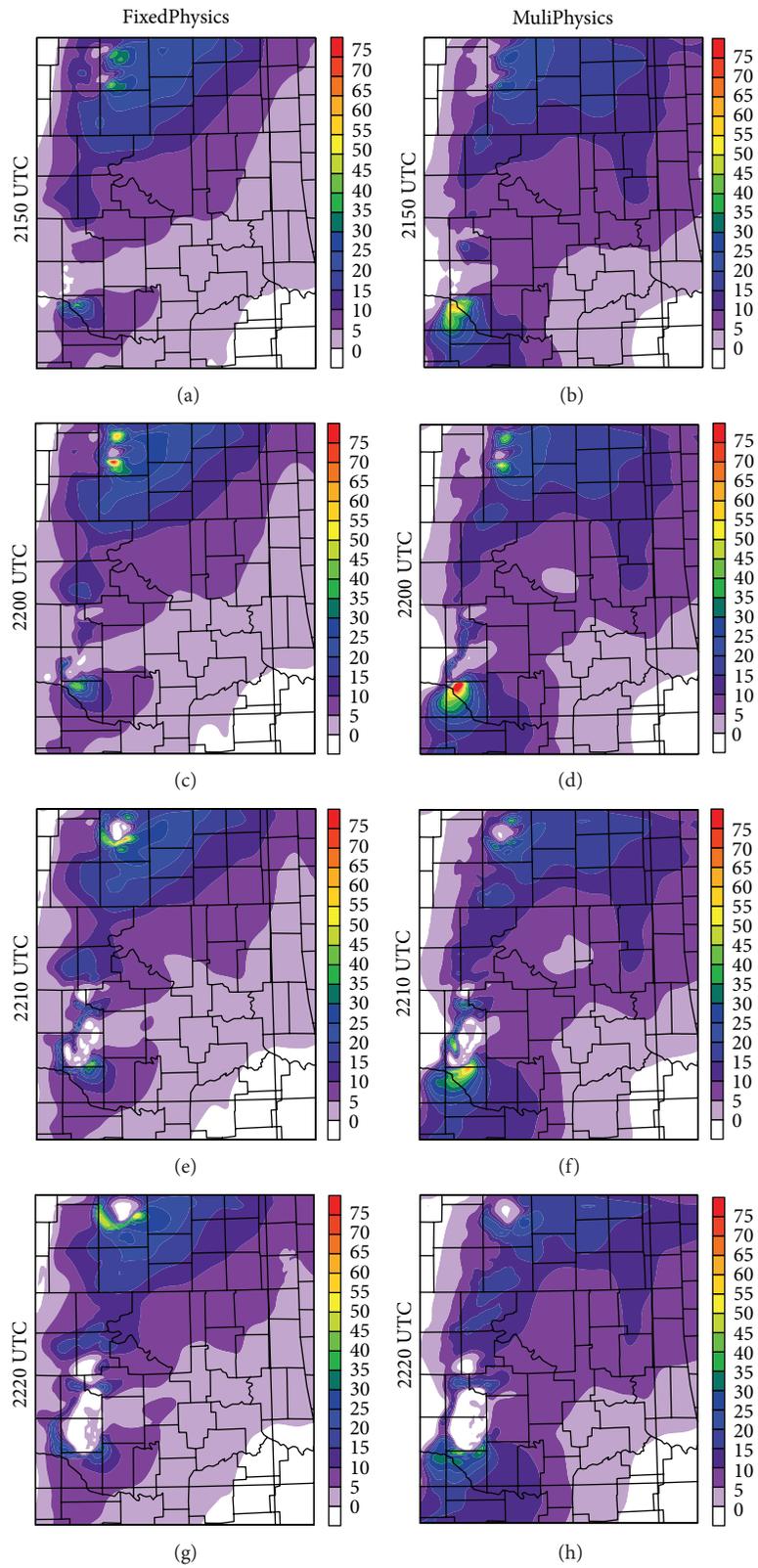


FIGURE 5: Ensemble-mean forecasts of STP parameter (colorfill, 5 increments) from FixedPhysics and MultiPhysics convective-scale experiments. The portion of the domain shown here is 306×363 km wide.

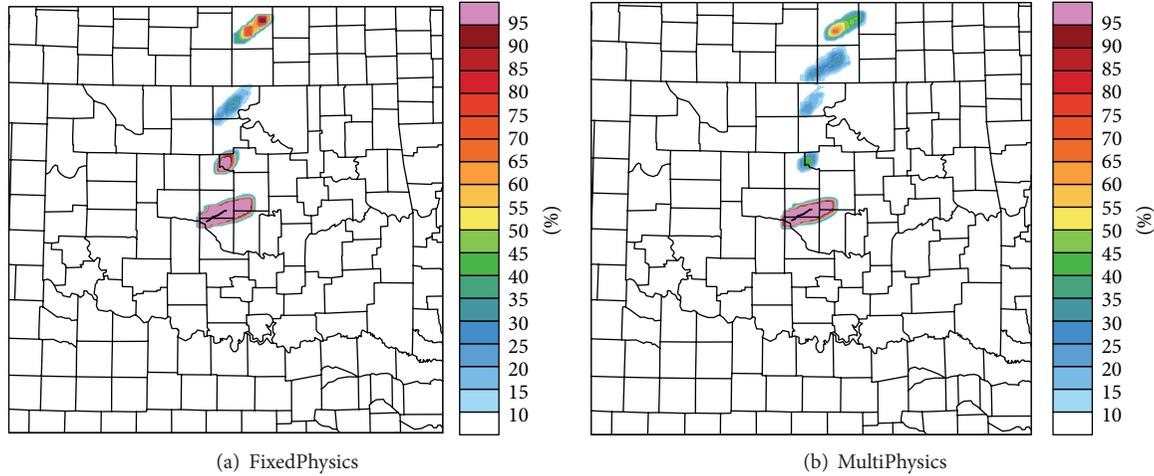


FIGURE 6: Neighborhood ensemble probability forecasts of 0–3 km updraft helicity from (a) FixedPhysics and (b) MultiPhysics convective-scale ensembles exceeding thresholds of $50 \text{ m}^2 \text{ s}^{-2}$ starting at 2200 UTC and ending at 2240 UTC over the entire convective-scale domain. Overlaid in each panel is the NWS observed tornado damage track (black outline) that starts at 2210 UTC and ends at 2238 UTC.

3.5. Ensemble Probabilistic Forecast of Updraft Helicity of the Supercell. The 3 km model horizontal grid spacing used in this study is far too coarse to explicitly resolve a tornado circulation. However, one good measure of the amount of rotation within the supercell storm is the updraft helicity (UH; [39–41, 54]), as it tends to highlight the main rotating storm updraft within a specified layer. A 0–3 km UH is selected to evaluate forecasts of low-level mesocyclones associated with tornadic supercell storms. Neighborhood ensemble probabilities of UH exceeding predetermined thresholds are calculated during the 1 h forecast period from both experiments (Figures 6 and 7), with a 9 km radius used to calculate the probabilistic forecasts of UH around each horizontal grid point to account for the small displacement errors across the ensemble members [6]. Results using a threshold UH of $50 \text{ m}^2 \text{ s}^{-2}$, a reasonable value for identifying mesocyclonic features in a convective-scale model [55, 56], show several regions of interest (Figure 6). Both experiments show maximum probabilities (100%) of significant rotation over OKC that covers the NWS surveyed OKC tornado observed damage track (black line) and extends farther northeastward. Close examination reveals that the 100% probabilities of a low-level mesocyclone from the FixedPhysics ensemble (Figure 6(a)) encompass a broader area than those from the MultiPhysics ensemble (Figure 6(b)) for the OKC area. In addition, both experiments show several additional rotation tracks north of OKC in north central Oklahoma, near the Oklahoma-Kansas border and in south-central Kansas. The FixedPhysics ensemble experiments show two high probability rotation tracks, one just north of OKC with 100% probabilities at several points and another in south-central Kansas with probabilities as high as 95%. In contrast, the MultiPhysics ensemble generates low probabilities of rotation with values below 45% on the storm north of OKC and values below 70% for the longer mesocyclone track in south Kansas. The high probabilities of UH in Kansas from the FixedPhysics ensembles correlate with the high STP values in

that area as shown in Figures 5(a), 5(c), 5(e), and 5(g). These results suggest that compared to the FixedPhysics ensemble, the MultiPhysics ensemble is able to better discriminate the region of tornadic supercell threat during this 1-hour forecast period.

Using UH track as a proxy for tornado path length forecasts, Clark et al. [41] show that the UH forecast path length from convective-scale models is strongly related to the track length of observed tornadoes. Therefore, to evaluate the forecasts of low-level tornadic rotation for the OKC supercell storm, 0–3 km neighborhood UH probabilities with higher threshold values of $150 \text{ m}^2 \text{ s}^{-2}$, $200 \text{ m}^2 \text{ s}^{-2}$, and $250 \text{ m}^2 \text{ s}^{-2}$ from both FixedPhysics and MultiPhysics convective-scale ensemble are evaluated (Figures 7(a)–7(f)) and are compared against the 0–3 km mesocyclone circulations [57] from KTLX radar observations (Figure 7(g)) generated using the Warning Decision Support System-Integrated Information software (WDSS-II; [58]). Results indicate that the low-level mesocyclone persists during the 0-1h forecast for all threshold values, with higher probabilities of UH qualitatively correlating well with the observed rotation track (Figure 7(g)). Maximum probabilities (100%) are seen at all grid points covering the NWS damage path and correlating well with the radar observed rotation path for $150 \text{ m}^2 \text{ s}^{-2}$ threshold value for both ensemble experiments (Figures 7(a), 7(b), and 7(g)). However, the UH track from FixedPhysics (Figure 7(a)) extends well beyond the observed mesocyclone track with 100% probabilities stretching northeastward, while the MultiPhysics (Figure 7(b)) ensemble correctly forecasts the length of observed rotation with 100% probabilities and has lower UH probabilities beyond the observed rotation track. The probabilities remain above 90% for a UH threshold of $200 \text{ m}^2 \text{ s}^{-2}$ and above 50% for a threshold of $250 \text{ m}^2 \text{ s}^{-2}$ for the entire path length of the observed damage track (Figures 7(d) and 7(f)) in MultiPhysics. In contrast, the FixedPhysics ensemble indicates lower UH probabilities with values below 65% for a $200 \text{ m}^2 \text{ s}^{-2}$ threshold and below 30%

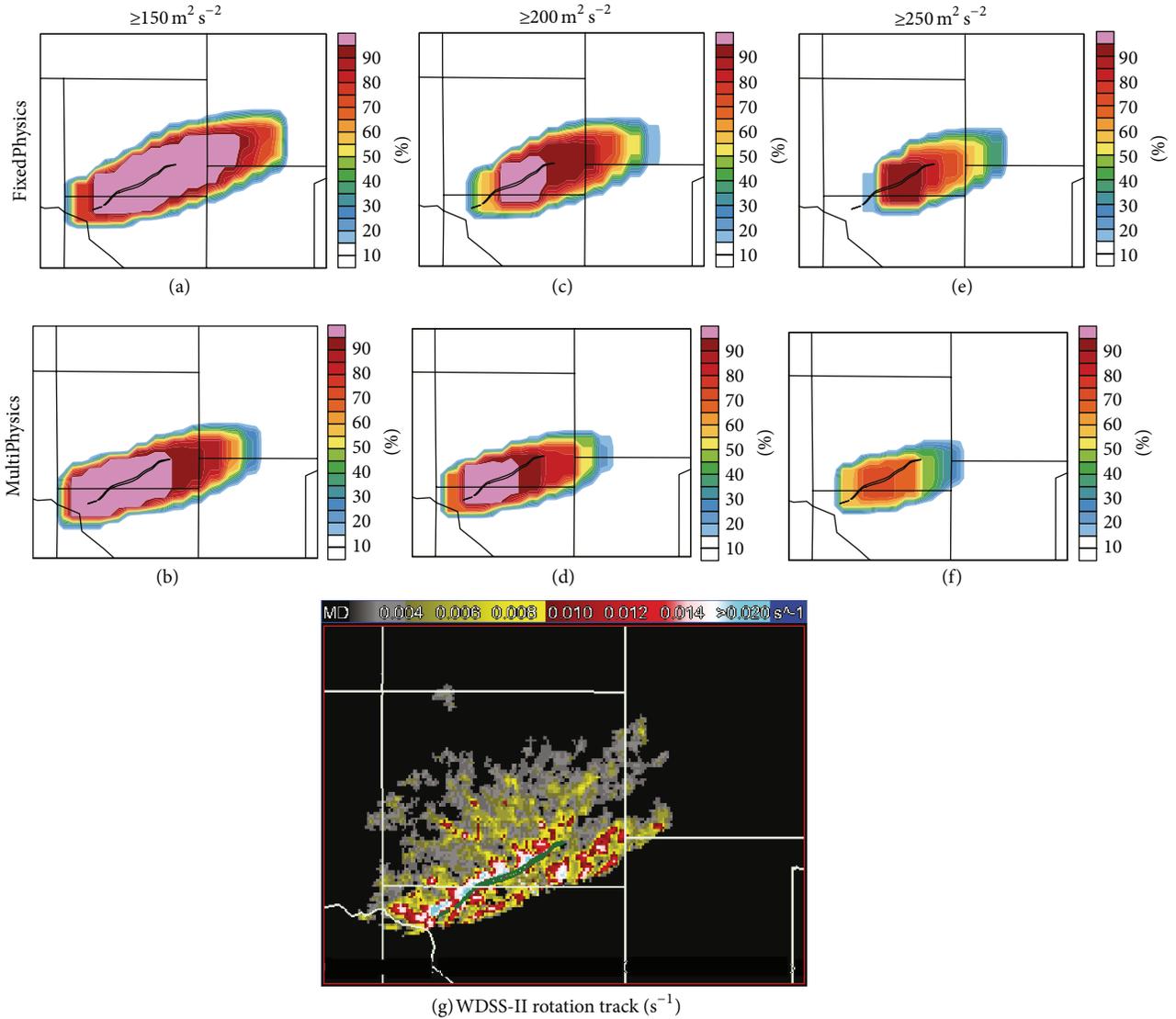


FIGURE 7: Neighborhood ensemble probability forecasts of 0–3 km updraft helicity from FixedPhysics and MultiPhysics convective-scale ensembles exceeding thresholds of ((a), (b)) $150 \text{ m}^2 \text{ s}^{-2}$, ((c), (d)) $200 \text{ m}^2 \text{ s}^{-2}$, and ((e), (f)) $250 \text{ m}^2 \text{ s}^{-2}$ starting at 2200 UTC and ending at 2240 UTC. The bottom panel (g) is the WDS-II generated KTLX radar observed low level (0–3 km AGL) mesocyclone track during 2200–2240 UTC (MD is missing data). Overlaid in each panel is the NWS observed tornado damage track (black outline in (a)–(f) and green outline in (g)) that starts at 2210 UTC and ends at 2238 UTC. The portion of the domain shown here is $120 \times 90 \text{ km}$ wide.

for a $250 \text{ m}^2 \text{ s}^{-2}$ threshold near the beginning of the observed tornado. Thus, the UH probability track from the MultiPhysics ensemble better captures the observed tornado and rotation track extent than from the FixedPhysics ensemble. These results highlight the potential benefit of background environmental variability in predicted 0–3 km UH forecast probabilities violent tornadoes, one of the goals of NOAA’s Warn-on-Forecast initiative [18].

3.6. Forecast Time Series of Equitable Threat Scores (ETS). To quantify the accuracy of precipitation forecasts from the ensembles, the ETS is calculated from both FixedPhysics and MultiPhysics convective-scale ensembles for radar reflectivity exceeding threshold values of 35 and 45 dBZ (Figure 8).

The ETS is calculated using continuously cycled 3DVAR analyses produced throughout the 1h forecast period as observations. An ETS score of 1 indicates a perfect forecast, with the ETS value decreasing to 0 as forecast accuracy declines. Results indicate that both ensemble systems start with ETS values of ~ 0.70 for 35 dBZ threshold (Figures 8(a) and 8(b)) and ~ 0.55 for 45 dBZ threshold (Figures 8(c) and 8(d)) at the beginning of the forecast. The ETS accuracy decreases with forecast lead times as expected. However, the variability in the ETS score amongst the members is larger and increases with forecast lead times for the MultiPhysics ensembles compared to that for the FixedPhysics ensemble. At the end of the forecast period at 2240 UTC, the mean ETS values for MultiPhysics ensemble are ~ 0.15 and ~ 0.20

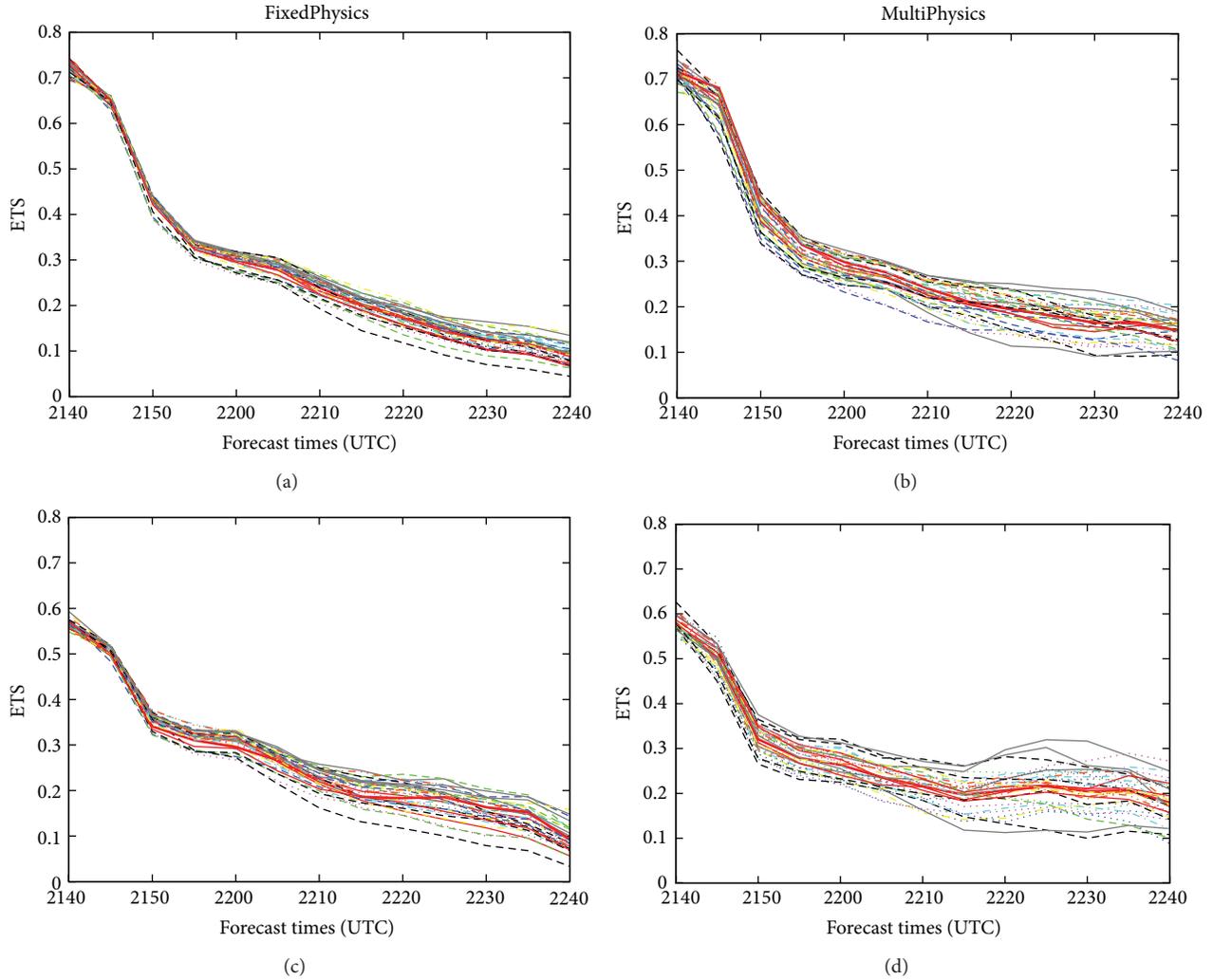


FIGURE 8: Values of equitable threat score (ETS) for reflectivity thresholds of ((a), (b)) 35 dBZ and ((c), (d)) 45 dBZ as a function of forecast times (UTC) from the convective-scale FixedPhysics and MultiPhysics 36-member ensembles (thin lines) and ensemble mean (thick lines). The independent 3DVAR analyses of reflectivity are used as observations.

(Figures 8(b) and 8(d)), while the mean ETS values for FixedPhysics ensembles are 0.09 and 0.10 (Figures 8(a) and 8(c)) for 35 and 45 dBZ thresholds, respectively. Thus the MultiPhysics ensemble maintains higher ETS accuracy than the FixedPhysics ensemble at the end of 1-hour-long forecasts. This is more pronounced for 45 dBZ threshold (Figure 8(d)), in which the MultiPhysics ensemble maintains the 0.20 ETS values during the last 25 minutes of the forecasts.

4. Discussion

In this study, experiments are conducted to assimilate radar observations within a convective-scale ensemble using background storm environments from two different mesoscale ensembles for May 8, 2003, Oklahoma City tornadic supercell storm event. The two sets of 36-member 12 km mesoscale ensembles using either single (FixedPhysics) or multiple physical parameterization (MultiPhysics) schemes are produced. The FixedPhysics ensemble uses the same land

surface, planetary boundary layer, radiation, convection, and microphysical parameterizations amongst all the ensemble members, whereas the MultiPhysics ensemble uses a variation of those combinations across the members. Traditional atmospheric observations are assimilated into the ensembles at every hour cycle starting at 1200 UTC on the day of the event and out to 12 h or 0000 UTC, May 9, 2003. The convective-scale 3 km ensembles are created using the mesoscale ensembles as background and assimilating Doppler radial velocity and reflectivity observations from four operational WSR-88D radars every 5 minutes over a 40 min cycling period starting at 2100 UTC and ending at 2140 UTC. Finally, 1h forecasts are launched from the convective-scale ensemble analyses starting at 2140 UTC and extending out to 2240 UTC, thereby covering the entire lifetime of the observed OKC tornado.

Results indicate that the forecast RMSE values for the near surface temperature, dewpoint temperature, and wind variables from the convective-scale MultiPhysics ensemble

are smaller than those from the FixedPhysics ensemble, highlighting the positive impact of the MultiPhysics approach. However, a more qualitative evaluation of specific forecast features, such as the presence of dryline bulges, environmental sounding structures, values of ensemble mean STP, and 0–3 km UH probabilities shows that the MultiPhysics ensemble better captures the important features on this day than the FixedPhysics ensemble. In particular, the convective-scale Multiphysics ensemble forecasts high values of STP around the OKC area before tornadogenesis, suggesting an environment that is very favorable for tornadic supercell storms, while the FixedPhysics experiment forecasts much lower STP values in the same area. The 0–3 km UH values for both FixedPhysics and MultiPhysics ensembles show high probabilities that correlate well with the observed tornado and low-level rotation tracks. However, the UH track in the MultiPhysics ensemble better captures the beginning and ending points of the observed tornado track than seen in the FixedPhysics ensemble. Therefore, convective-scale ensembles with greater diversity in the mesoscale environmental conditions as produced through using multiple physics schemes can provide forecasters with more accurate situational awareness and greater confidence of the tornado threats from very short-range ensemble forecasts.

Although not computationally feasible for this study, convective-scale data assimilation and forecast experiments with horizontal grid spacing less than 1 km are needed to resolve tornadic-scale circulations. Past studies show noticeable differences in storm structures when simulated with a horizontal grid spacing varying between 250 m and 1 km [59, 60]. While the computational demands associated with such small grid spacing are significant at this time, with continued rapid increases in computing power, future work will focus on convective-scale data assimilation and forecast experiments at 1 km or less. The use of more sophisticated double or triple moment microphysics schemes in the convective-scale ensemble with perturbed microphysical parameters within the scheme [61] and applying physics diversity across the ensemble [62, 63] can provide improved short-range forecasts for a wide range of storm systems and will be included in future convective-scale data assimilation studies.

Due to our limited understanding of atmospheric processes, it is likely that the use of even more sophisticated physical parameterization schemes will face challenges when used in some storm environments. However, the results obtained from this study suggest that by using reasonable diversity in physics schemes, an ensemble system is more likely to span the observations and provide improved storm environments for a wide range of storm systems. An ensemble system that accounts for uncertainties both in initial condition and model physical parameterization schemes is important to the successful very short-range probabilistic convective-scale forecast of tornadic supercell thunderstorms, which is the main goal NOAA's Warn-on-Forecast initiative.

Acknowledgments

The constructive comments of two anonymous reviewers greatly improved the paper. Local computer assistance

provided by Brett Morrow, Steven Fletcher, Brad Swagowitz, and Karen Cooper is greatly appreciated. Partial funding for this research was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce.

References

- [1] A. Aksoy, D. C. Dowell, and C. Snyder, "A multicaser comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: storm-scale analyses," *Monthly Weather Review*, vol. 137, no. 6, pp. 1805–1824, 2009.
- [2] A. Aksoy, D. C. Dowell, and C. Snyder, "A multicaser comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: short-range ensemble forecasts," *Monthly Weather Review*, vol. 138, no. 4, pp. 1273–1292, 2010.
- [3] R. M. Cintineo and D. J. Stensrud, "On the predictability of supercell thunderstorm evolution," *Journal of Atmospheric Science*, vol. 70, pp. 1993–2011, 2013.
- [4] D. J. Stensrud and J. Gao, "Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short-range forecasts," *Monthly Weather Review*, vol. 138, no. 4, pp. 1250–1272, 2010.
- [5] C. L. Ziegler, E. R. Mansell, J. M. Straka, D. R. Macgorman, and D. W. Burgess, "The impact of spatial variations of low-level stability on the life cycle of a simulated supercell storm," *Monthly Weather Review*, vol. 138, no. 5, pp. 1738–1766, 2010.
- [6] N. Yussouf, E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, "The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City Tornadic Supercell Storm using single and double moment microphysics schemes," *Monthly Weather Review*, 2013.
- [7] C. S. Schwartz, J. S. Kain, S. J. Weiss et al., "Next-day convection-allowing WRF model guidance: a second look at 2-km versus 4-km grid spacing," *Monthly Weather Review*, vol. 137, no. 10, pp. 3351–3372, 2009.
- [8] C. S. Schwartz, J. S. Kain, S. J. Weiss et al., "Toward improved convection-allowing ensembles: model physics sensitivities and optimizing probabilistic guidance with small ensemble membership," *Weather and Forecasting*, vol. 25, no. 1, pp. 263–280, 2010.
- [9] G. S. Romine, C. S. Schwartz, C. Snyder, J. L. Anderson, and M. L. Weisman, "Model bias in a continuously cycled assimilation system and its influence on convection-permitting forecasts," *Monthly Weather Review*, vol. 141, pp. 1263–1284, 2013.
- [10] B. C. Ancell, "Examination of analysis and forecast errors of high-resolution assimilation, Bias removal, and digital filter initialization with an ensemble Kalman filter," *Monthly Weather Review*, vol. 140, pp. 3992–4004, 2012.
- [11] D. J. Stensrud, J.-W. Bao, and T. T. Warner, "Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems," *Monthly Weather Review*, vol. 128, no. 7 I, pp. 2077–2107, 2000.
- [12] T. Fujita, D. J. Stensrud, and D. C. Dowell, "Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties," *Monthly Weather Review*, vol. 135, no. 5, pp. 1846–1868, 2007.
- [13] T. M. Hamill, R. S. Schneider, H. E. Brooks et al., "The May 2003 extended tornado outbreak," *Bulletin of the American Meteorological Society*, vol. 86, no. 4, pp. 531–542, 2005.

- [14] M. Hu and M. Xue, "Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case," *Monthly Weather Review*, vol. 135, no. 2, pp. 507–525, 2007.
- [15] T. M. Lei, M. Xue, and T. Yu, "Multi-scale analysis and prediction of the 8 May 2003 Oklahoma City tornadic supercell storm assimilating radar and surface network data using EnKF," in *Proceedings of the 13th Conference on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface (IOAS-AOLS '09)*, The American Meteorological Society, 2009.
- [16] D. C. Dowell, L. J. Wicker, and C. Snyder, "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*, vol. 139, no. 1, pp. 272–294, 2011.
- [17] Z. Meng and F. Zhang, "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*, vol. 136, no. 10, pp. 3671–3682, 2008.
- [18] D. J. Stensrud, N. Yussouf, D. C. Dowell, and M. C. Coniglio, "Assimilating surface data into a mesoscale model ensemble: cold pool analyses from spring 2007," *Atmospheric Research*, vol. 93, no. 1–3, pp. 207–220, 2009.
- [19] D. M. Wheatley, D. J. Stensrud, D. C. Dowell, and N. Yussouf, "Application of a WRF mesoscale data assimilation system to springtime severe weather events 2007–09," *Monthly Weather Review*, vol. 140, pp. 1539–1557, 2012.
- [20] G. S. Romine, D. W. Burgess, and R. B. Wilhelmson, "A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma City Area tornadic supercell," *Monthly Weather Review*, vol. 136, no. 8, pp. 2849–2870, 2008.
- [21] W. C. Skamarock, J. B. Klemp, J. Dudhia et al., "A description of the Advanced Research WRF version 2," Tech. Rep., NCAR, TN-468+STR, 2005.
- [22] D. F. Parrish and J. C. Derber, "The National Meteorological Center's spectral statistical- interpolation analysis system," *Monthly Weather Review*, vol. 120, no. 8, pp. 1747–1763, 1992.
- [23] R. D. Torn, G. J. Hakim, and C. Snyder, "Boundary conditions for limited-area ensemble Kalman filters," *Monthly Weather Review*, vol. 134, no. 9, pp. 2490–2502, 2006.
- [24] G. Thompson, P. R. Field, R. M. Rasmussen, and W. D. Hall, "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization," *Monthly Weather Review*, vol. 136, no. 12, pp. 5095–5115, 2008.
- [25] M. Tiedtke, "A comprehensive mass flux scheme for cumulus parameterization in large-scale models," *Monthly Weather Review*, vol. 117, no. 8, pp. 1779–1800, 1989.
- [26] C. Zhang, Y. Wang, and K. Hamilton, "Improved representation of boundary layer clouds over the southeast pacific in ARW-WRF using a modified tiedtke cumulus parameterization scheme," *Monthly Weather Review*, vol. 139, no. 11, pp. 3489–3513, 2011.
- [27] S.-Y. Hong, Y. Noh, and J. Dudhia, "A new vertical diffusion package with an explicit treatment of entrainment processes," *Monthly Weather Review*, vol. 134, no. 9, pp. 2318–2341, 2006.
- [28] F. Chen and J. Dudhia, "Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity," *Monthly Weather Review*, vol. 129, no. 4, pp. 569–585, 2001.
- [29] J. L. Anderson, "An ensemble adjustment Kalman filter for data assimilation," *Monthly Weather Review*, vol. 129, no. 12, pp. 2884–2903, 2001.
- [30] J. L. Anderson and N. Collins, "Scalable implementations of ensemble filter algorithms for data assimilation," *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 8, pp. 1452–1463, 2007.
- [31] J. Anderson, T. Hoar, K. Raeder et al., "The data assimilation research testbed a community facility," *Bulletin of the American Meteorological Society*, vol. 90, no. 9, pp. 1283–1296, 2009.
- [32] G. Gaspari and S. E. Cohn, "Construction of correlation functions in two and three dimensions," *Quarterly Journal of the Royal Meteorological Society*, vol. 125, no. 554, pp. 723–757, 1999.
- [33] M. Xue, K. K. Droegemeier, and V. Wong, "The Advanced Regional Prediction System (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: model dynamics and verification," *Meteorology and Atmospheric Physics*, vol. 75, no. 3–4, pp. 161–193, 2000.
- [34] M. Xue, K. K. Droegemeier, V. Wong et al., "The Advanced Regional Prediction System (ARPS)—a multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: model physics and applications," *Meteorology and Atmospheric Physics*, vol. 76, no. 3–4, pp. 143–165, 2001.
- [35] J. Gao, M. Xue, A. Shapiro, and K. K. Droegemeier, "A variational method for the analysis of three-dimensional wind fields from two Doppler radars," *Monthly Weather Review*, vol. 127, no. 9, pp. 2128–2142, 1999.
- [36] J. Gao, M. Xue, K. Brewster, F. Carr, and K. K. Droegemeier, "New development of a 3DVAR system for a nonhydrostatic NWP model," in *Proceedings of the 15th Conference of Numerical Weather Prediction and 19th Conference of Weather Analyses and Forecasting*, pp. 339–341, The American Meteorological Society, San Antonio, Tex, USA, 2002.
- [37] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [38] K. A. Brewster, "Recent advances in diabatic initialization of a non-hydrostatic numerical model," in *Proceedings of the 15th Conference on Numerical Weather Prediction and 21st Conference on Severe Local Storms*, American Meteorological Society, San Antonio, Tex, USA, 2001.
- [39] A. J. Clark, S. J. Weiss, J. S. Kain et al., "An overview of the 2010 Hazardous Weather Testbed experimental forecast program spring experiment," *Bulletin of the American Meteorological Society*, vol. 93, no. 1, pp. 55–74, 2012.
- [40] A. J. Clark, J. S. Kain, P. T. Marsh, J. Correia, M. Xue, and F. Kong, "Forecasting tornado pathlengths using a three-dimensional object identification algorithm applied to convection-allowing forecasts," *Weather and Forecasting*, vol. 27, pp. 1090–1113, 2012.
- [41] A. J. Clark, J. Gao, P. T. Marsh et al., "Tornado path length forecasts from 2010–2011 using ensemble updraft helicity," *Weather and Forecasting*, vol. 28, pp. 387–407, 2013.
- [42] J. Gao, T. T. Smith, and D. J. Stensrud, "A realtime weather-adaptive 3DVAR analysis system for severe weather detections and warnings," *Weather and Forecasting*, vol. 28, pp. 727–745, 2013.
- [43] Y.-L. Lin, R. D. Farley, and H. D. Orville, "Bulk parameterization of the snow field in a cloud model," *Journal of Climate & Applied Meteorology*, vol. 22, no. 6, pp. 1065–1092, 1983.

- [44] G. L. Mellor and T. Yamada, "Development of a turbulence closure model for geophysical fluid problems," *Reviews of Geophysics & Space Physics*, vol. 20, no. 4, pp. 851–875, 1982.
- [45] Z. I. Janjic, "Nonsingular implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model," NCEP Office Note 437, 61, 2002.
- [46] J. Dudhia, "Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model," *Journal of the Atmospheric Sciences*, vol. 46, no. 20, pp. 3077–3107, 1989.
- [47] E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, "Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave," *Journal of Geophysical Research D*, vol. 102, no. 14, pp. 16663–16682, 1997.
- [48] K. Brewster, M. Hu, M. Xue, and J. Gao, "Efficient assimilation of radar data at high resolution for short-range numerical weather prediction," in *Proceedings of the WWRP International Symposium on Nowcasting and Very Short Range Forecasting*, Toulouse, France, 2005.
- [49] R. A. McPherson, C. A. Fiebrich, K. C. Crawford et al., "Statewide monitoring of the mesoscale environment: a technical update on the Oklahoma Mesonet," *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 3, pp. 301–321, 2007.
- [50] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*, Academic Press, Boston, Mass, USA, 2nd edition, 2006.
- [51] C. E. Hane, R. M. Rabin, T. M. Crawford, H. B. Bluestein, and M. E. Baldwin, "A case study of severe storm development along a dryline within a synoptically active environment. Part II: multiple boundaries and convective initiation," *Monthly Weather Review*, vol. 130, no. 4, pp. 900–920, 2002.
- [52] R. L. Thompson, R. Edwards, and J. A. Hart, "Evaluation and interpretation of the supercell composite and significant tornado parameters at the Storm Prediction Center," in *Proceedings of the 21st Conference on Severe Local Storms*, American Meteorological Society, San Antonio, Tex, USA, 2002.
- [53] H. E. Brooks, C. A. Doswell III, and J. Cooper, "On the environments of tornadic and nontornadic mesocyclones," *Weather & Forecasting*, vol. 9, no. 4, pp. 606–618, 1994.
- [54] J. S. Kain, S. J. Weiss, D. R. Bright et al., "Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP," *Weather and Forecasting*, vol. 23, no. 5, pp. 931–952, 2008.
- [55] J. R. Carley, B. R. J. Schwedler, M. E. Baldwin et al., "A proposed model-based methodology for feature-specific prediction for high-impact weather," *Weather and Forecasting*, vol. 26, no. 2, pp. 243–249, 2011.
- [56] R. J. Trapp, E. D. Robinson, M. E. Baldwin, N. S. Diffenbaugh, and B. R. J. Schwedler, "Regional climate of hazardous convective weather through high-resolution dynamical downscaling," *Climate Dynamics*, vol. 37, no. 3, pp. 677–688, 2011.
- [57] M. L. Miller, V. Lakshmanan, and T. Smith, "An automated method for depicting mesocyclone paths and intensities," *Weather and Forecasting*, vol. 28, pp. 570–585, 2013.
- [58] V. Lakshmanan, T. Smith, G. Stumpf, and K. Hondl, "The warning decision support system-integrated information," *Weather and Forecasting*, vol. 22, no. 3, pp. 596–612, 2007.
- [59] G. H. Bryan, J. C. Wyngaard, and J. M. Fritsch, "Resolution requirements for the simulation of deep moist convection," *Monthly Weather Review*, vol. 131, pp. 2394–2416, 2003.
- [60] G. H. Bryan and H. Morrison, "Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics," *Monthly Weather Review*, vol. 140, no. 1, pp. 202–225, 2012.
- [61] N. Yussouf and D. J. Stensrud, "Comparison of single-parameter and multiparameter ensembles for assimilation of radar observations using the ensemble kalman filter," *Monthly Weather Review*, vol. 140, no. 2, pp. 562–586, 2012.
- [62] N. Snook, M. Xue, and Y. Jung, "Analysis of a tornadic mesoscale convective vortex based on ensemble kalman filter assimilation of CASA X-band and WSR-88D radar data," *Monthly Weather Review*, vol. 139, no. 11, pp. 3446–3468, 2011.
- [63] N. Snook, M. Xue, and J. Jung, "Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data," *Monthly Weather Review*, vol. 140, pp. 2126–2146, 2012.

Research Article

Correcting Fast-Mode Pressure Errors in Storm-Scale Ensemble Kalman Filter Analyses

Corey K. Potvin^{1,2} and Louis J. Wicker²

¹Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, USA

²NOAA/OAR/National Severe Storms Laboratory, National Weather Center,
120 David L. Boren Boulevard, Norman, OK 73072, USA

Correspondence should be addressed to Corey K. Potvin; corey.potvin@noaa.gov

Received 29 May 2013; Revised 26 July 2013; Accepted 19 August 2013

Academic Editor: Ming Xue

Copyright © 2013 C. K. Potvin and L. J. Wicker. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A typical storm-scale ensemble Kalman filter (EnKF) analysis/forecast system is shown to introduce imbalances into the ensemble posteriors that generate acoustic waves in subsequent integrations. When the EnKF is used to research storm-scale dynamics, the resulting spurious pressure oscillations are large enough to impact investigation of processes driven by nonhydrostatic pressure gradient forces. Fortunately, thermodynamic retrieval techniques traditionally applied to dual-Doppler wind analyses can be adapted to diagnose the balanced portion of an EnKF pressure analysis, thereby eliminating the fast-mode pressure oscillations. The efficacy of this approach is demonstrated using a high-resolution supercell thunderstorm simulation as well as EnKF analyses of a simulated and a real supercell.

1. Introduction

The EnKF [1] has become a popular and valuable tool for storm-scale research [2–12]. Particularly when dual-Doppler radar data are available, EnKF data assimilation can provide reliable analyses of wind and, to a lesser degree, temperature and microphysical variables in convective storms. EnKF analyses of pressure, on the other hand, are subject to severe errors, at least with some compressible model configurations (the first tests of the EnKF with a compressible model were performed by Tong and Xue [5]). This problem is illustrated in Figure 1 using output from the National Severe Storms Laboratory Collaborative Model for Multiscale Atmospheric Simulation (NCOMMAS; [13, 14]) ensemble square root filter [15]. Similar behavior occurs using the Data Assimilation Research Testbed [16] EnKF with the Advanced Research Weather Research and Forecasting (WRF-ARW; [17]) model (James Marquis and Thomas Jones, personal communication 2013). The pressure analysis errors severely impede investigation of critical storm processes that are, in part, driven by dynamic pressure gradient forces, including supercell occlusion downdrafts [18], lifting of negatively buoyant air

[19], supercell propagation [20], the descending rear inflow and ascending front-to-rear flow in mesoscale convective systems [21], and possibly descending reflectivity cores [22].

The pressure oscillations shown in Figure 1 are associated with acoustic waves generated within the data assimilation region. The waves are presumably excited as each ensemble member adjusts to an updated initial condition that is dynamically inconsistent with the model (i.e., unbalanced). This hypothesis is supported by two observations. First, the acoustic waves occur whether or not pressure is updated during the data assimilation and therefore cannot be attributed to erroneous ensemble covariances between the pressure and other variables (though the latter could conceivably exacerbate the problem in cases where pressure is updated). Second, spurious waves are not evident in perfect-model EnKF observing system simulation experiments (OSSEs) with the NCOMMAS (not shown), indicating that the waves arise only when analysis increments are substantially unbalanced. The generation of high-amplitude fast modes due to unbalanced initial conditions is a long-recognized problem in numerical weather prediction, and many approaches have been used to improve dynamical balance during the data assimilation

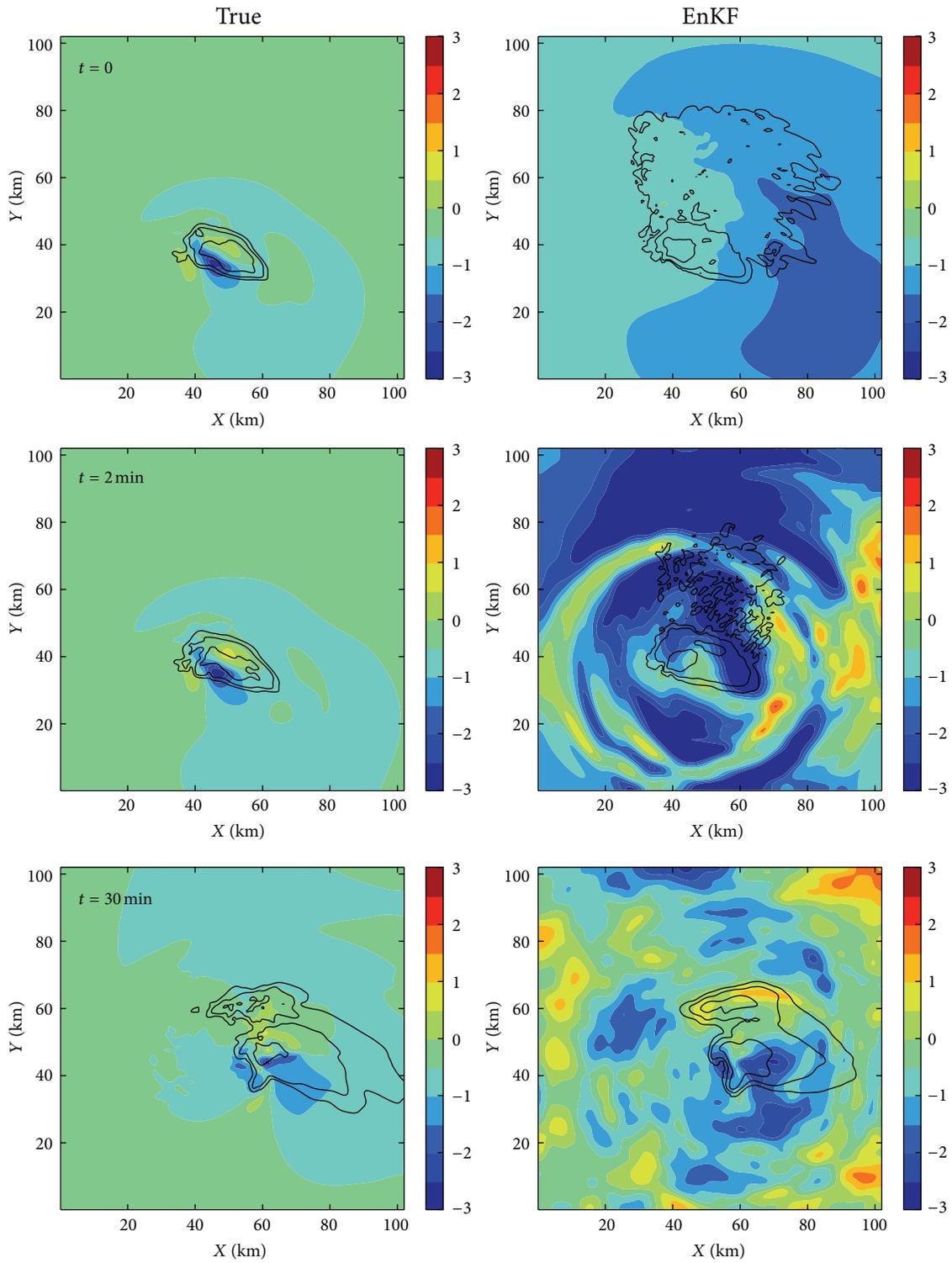


FIGURE 1: True (left column) and EnKF mean posterior (right column) p' (shading; hPa) and radar reflectivity factor (contoured at 10, 30, and 50 dBZ) at $z = 0.9 \text{ km}$. Fields are valid after (top row) zero, (middle row) one, and (bottom row) fifteen 2 min forecast cycles. The true p' were filtered and averaged as in Potvin et al. [23] to permit more direct comparisons with the (coarser) EnKF p' .

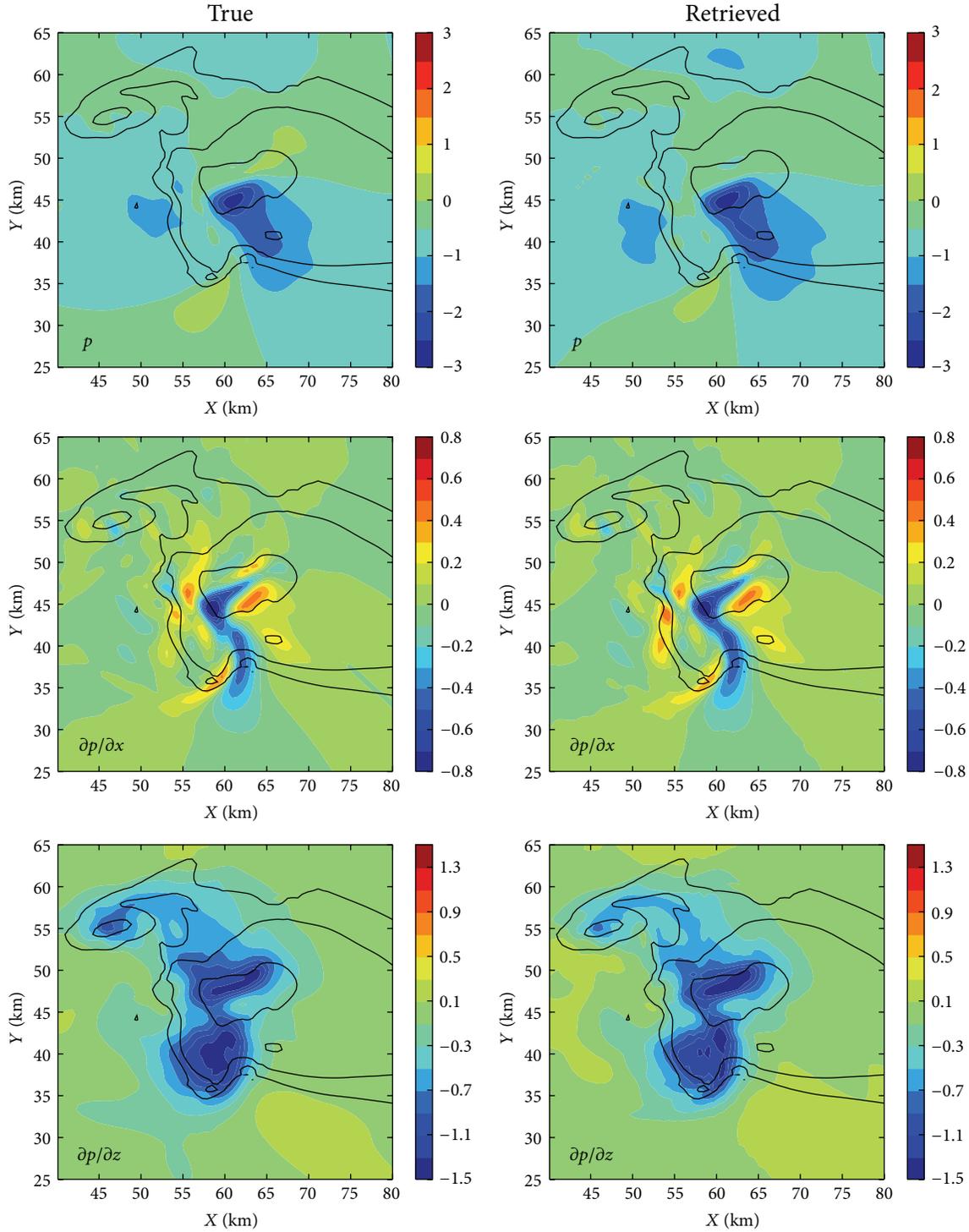


FIGURE 2: True (left) and retrieved (right) p (top; hPa), $\partial p/\partial x$ (middle; hPa km⁻¹), and $\partial p/\partial z$ (bottom; hPa km⁻¹) at $z = 0.9$ km. Radar reflectivity factor is contoured at 10, 30, and 50 dBZ.

procedure [24, 25]. Given the small influence of the pressure field on the remaining state variables in certain compressible cloud models on the relevant spatiotemporal scales, errors due to the acoustic waves are largely confined to the pressure field in storm-scale EnKF analyses. This makes it possible to

retrieve the portion of the pressure field that is in balance with the remaining model fields (hereafter, the “balanced” pressure). This obviates the need to mitigate the acoustic waves during the data assimilation procedure, at least when the analysis is not being used to initialize a numerical forecast.

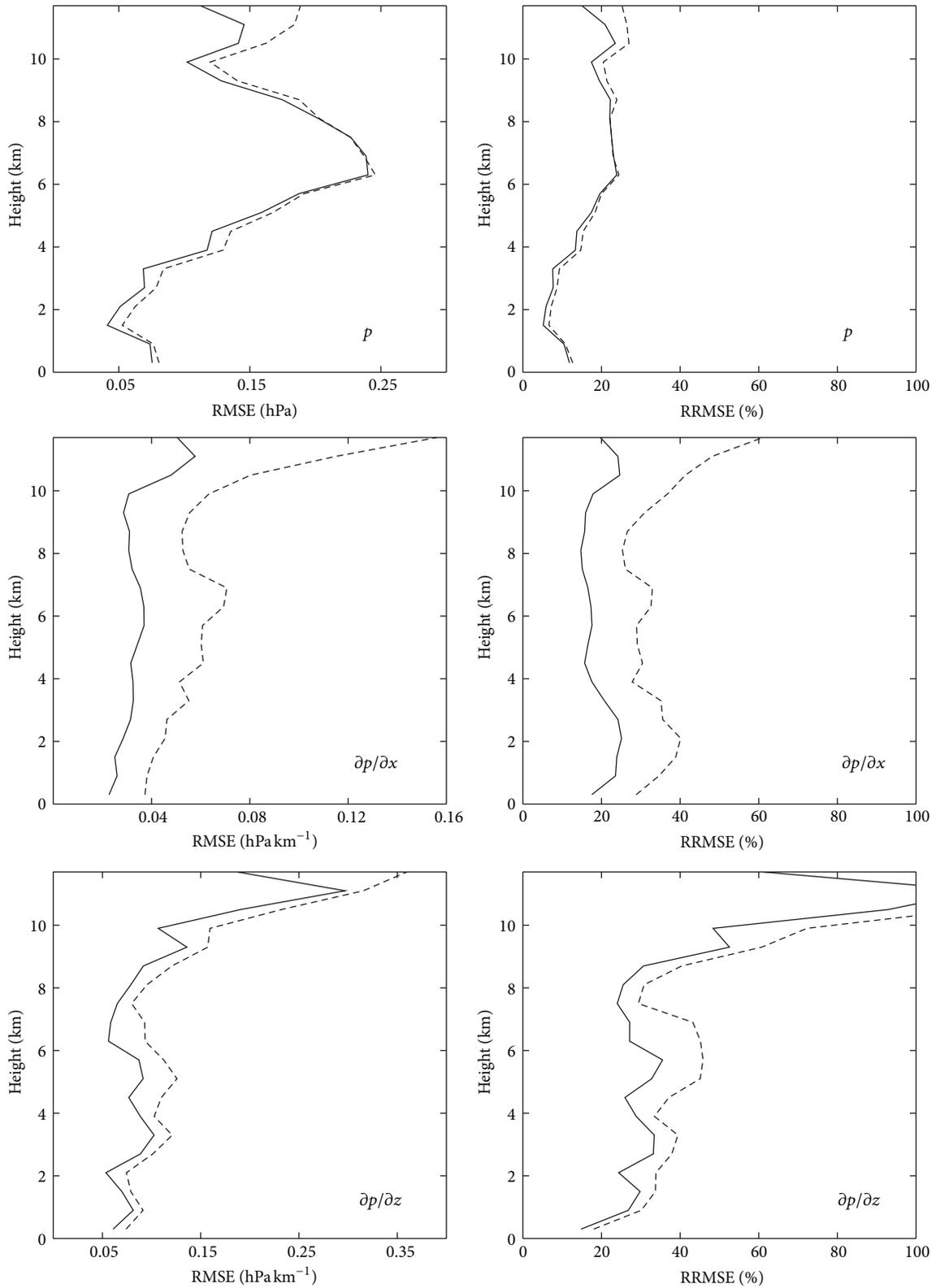


FIGURE 3: Vertical profiles of RMSE (left column) and RRMSE (right column) p_{final} (top row), $\partial p_{\text{final}}/\partial x$ (middle row), and $\partial p_{\text{final}}/\partial z$ (bottom row) for default retrieval ($\Delta t = 30$ s; solid) and retrieval with $\Delta t = 120$ s (dashed). All quantities valid where the true dBZ > 0.

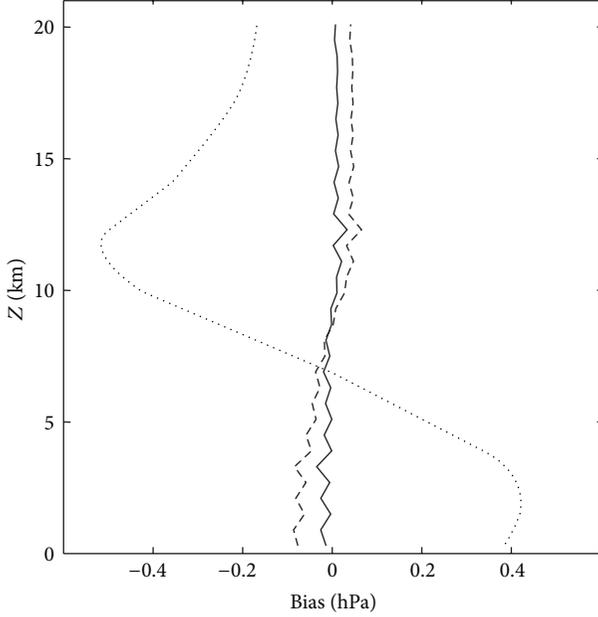


FIGURE 4: Vertical profiles of horizontal domain-wide bias in p_{2D} (dotted), p_{final} (solid), and p_{final} retrieved with the hydrometeor drag term omitted in the cost function (dashed).

Instead, the balanced pressure can be retrieved after the data assimilation is complete.

Several methods are available for diagnosing perturbation (from the hydrostatic base state) pressure, typically in its nondimensional Exner function form π (the prime symbol is omitted herein to simplify the notation), from the equations of motion. Early methods satisfied the horizontal equations of motion on individual horizontal planes [26, 27]. A major limitation of that approach is that the analyzed π is offset from the true π by a vertically varying constant, precluding unique solution of $\partial\pi/\partial z$ unless independent pressure measurements are available at each analysis level. This problem is avoided when all three equations of motion are satisfied [28–30], in which case the retrieved π is offset by a volume-wide (rather than vertically varying) constant, thereby permitting the impact of vertical pressure gradients to be quantitatively considered. The tradeoff is that errors in the analyzed local derivative and buoyancy terms in the vertical momentum equation typically result in $\partial\pi/\partial x$ and $\partial\pi/\partial y$ analyses that are inferior to those obtained using two-dimensional retrieval. As will be shown, however, there is a very simple procedure for obtaining the advantages of both the 2D and 3D approaches. This combined method permits useful pressure retrievals to be obtained from acoustic wave-contaminated EnKF analyses.

2. Pressure Retrieval Method

We adopt a variational framework for our pressure retrieval scheme. The momentum equation constraints use the Klemp and Wilhelmson [31] formulation of the equations of motion except with the Coriolis term (which is negligible) omitted.

A horizontal smoothness constraint is imposed to filter noise. The cost function we seek to minimize can therefore be expressed as

$$J = \frac{1}{2} \sum_i \sum_j \sum_k C_1 J_1^2 + C_2 J_2^2 + C_3 J_3^2 + C_4 J_4^2, \quad (1)$$

where

$$J_1 = \left(\frac{\partial\pi}{\partial x} - F \right), \quad J_2 = \left(\frac{\partial\pi}{\partial y} - G \right), \quad (2)$$

$$J_3 = \left(\frac{\partial\pi}{\partial z} - g \frac{\theta'}{\theta_0} - H \right), \quad J_4 = \left(\frac{\partial^2\pi}{\partial x^2} + \frac{\partial^2\pi}{\partial y^2} \right); \quad (3)$$

$i, j,$ and k are the model grid indices; $F, G,$ and H are the sums of the local time derivative, advection, and turbulent mixing terms for the u -, v -, and w -equations, respectively; g is the gravitational acceleration; and θ_0 and θ' are the horizontally homogeneous base state and perturbation potential temperature, respectively. The subscripted C 's represent the weighting coefficients for each respective constraint, computed similarly to the coefficients in [30]:

$$C_1 = \left[\frac{\sum_i \sum_j \sum_k (F^2) \Delta x \Delta y \Delta z}{\sum_i \sum_j \sum_k \Delta x \Delta y \Delta z} \right]^{-1},$$

$$C_2 = \left[\frac{\sum_i \sum_j \sum_k (G^2) \Delta x \Delta y \Delta z}{\sum_i \sum_j \sum_k \Delta x \Delta y \Delta z} \right]^{-1}, \quad (4)$$

$$C_3 = \left[\frac{\sum_i \sum_j \sum_k (H^2) \Delta x \Delta y \Delta z}{\sum_i \sum_j \sum_k \Delta x \Delta y \Delta z} \right]^{-1},$$

$$C_4 = k_0^{-4},$$

where $\Delta x, \Delta y,$ and Δz are the analysis grid spacings and k_0 is the cutoff wavenumber, for which the theoretical filter response is 0.5. The present study uses a $3\Delta x$ cutoff wavelength.

In the first step of the pressure retrieval scheme, a 2D retrieval is performed (i.e., C_3 is set to zero) over each analysis level. The resulting analysis, π_{2D} , is stored, and then a 3D retrieval is performed, yielding π_{3D} . The vertically varying constant $A(k)$ by which π_{2D} is offset from the balanced π , π_{bal} , is then estimated by computing the mean difference between π_{3D} and π_{2D} for each k . To see why $A(k)$ can be estimated in this way, consider that the 2D and 3D analyses can be written as

$$\pi_{2D} = \pi_{\text{bal}} + A(k) + \pi_{2D}^{\text{error}}, \quad (5)$$

$$\pi_{3D} = \pi_{\text{bal}} + \pi_{3D}^{\text{error}},$$

where π_{2D}^{error} and π_{3D}^{error} are the retrieval errors, apart from the vertically varying constant. Assuming that the 2D and 3D retrievals are approximately unbiased relative to each other (apart from the vertically varying constant), averaging the differences between the retrievals over each horizontal analysis plane (symbolized by brackets) yields

$$\langle \pi_{2D} - \pi_{3D} \rangle = \langle \pi_{2D}^{\text{error}} - \pi_{3D}^{\text{error}} \rangle + A(k) \approx A(k). \quad (6)$$

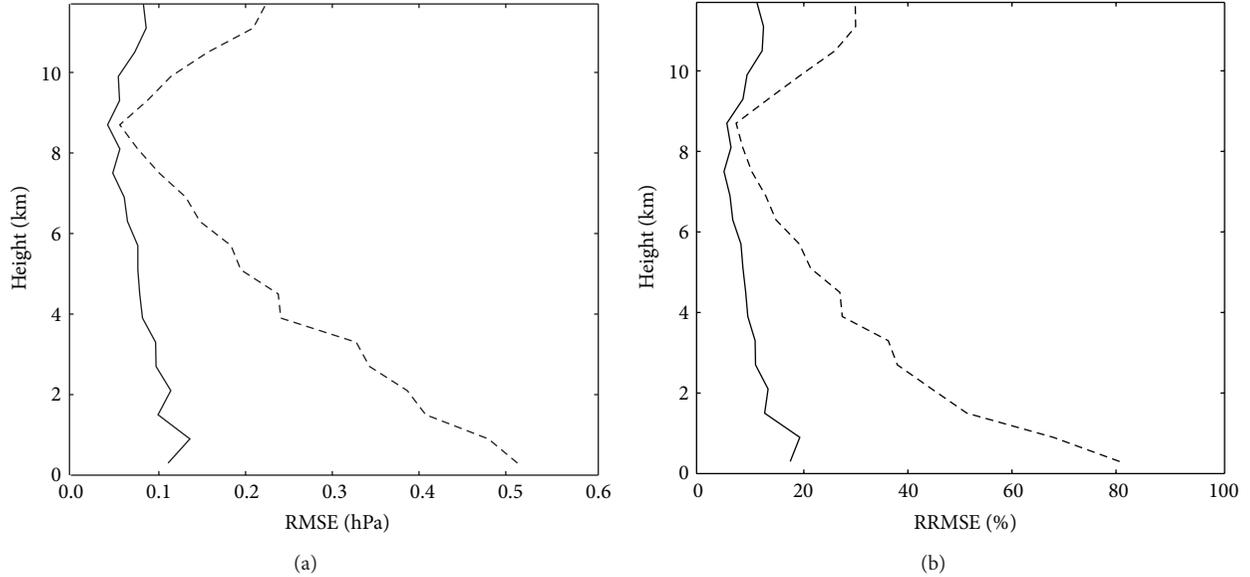


FIGURE 5: As in Figure 3 (top row) but for p_{3D} (solid) and p_{3D} retrieved with hydrometeor drag term omitted (dashed).

The final analysis π_{final} is then obtained by subtracting the $A(k)$ estimate from π_{2D} :

$$\pi_{\text{final}} = \pi_{2D} - \langle \pi_{2D} - \pi_{3D} \rangle. \quad (7)$$

The π_{final} is still subject to a volume-wide constant, which can be estimated as the volume-mean difference between π_{final} and the π from the simulation or EnKF analysis to which the retrieval method is applied. This estimate is reasonable to the extent that the errors in the input π are unbiased. Errors in the estimate of this constant will not impact most applications since it is typically only the spatial derivatives of π that are sought. The estimated volume-wide constant has been subtracted from the π_{final} in the experiments below. To further facilitate interpretation, π has been converted to dimensional perturbation pressure p .

3. Experiments

3.1. Verification of Retrieval Procedure. An NCOMMAS supercell simulation was used to test the robustness of the retrieval procedure and the integrity of the computer code. The simulation was performed on a stationary $102.6 \text{ km} \times 102.6 \text{ km} \times 20.4 \text{ km}$ domain with 600 m grid spacing in all three dimensions and a large (small) time step of 4 (2/3) s. A fully dual-moment version of the Ziegler [32] microphysical parameterization (Ziegler Variable Density or ZVD) scheme [33] was used. The remaining model settings were identical to those of Potvin et al. [23]. The supercell in the present simulation appears reasonably realistic, and its evolution qualitatively resembles that of Potvin et al. [23].

Pressure retrievals were performed over the entire simulation domain. The local derivatives of u , v , and w were computed using centered finite differences with default

$\Delta t = 30 \text{ s}$ (e.g., $\partial u / \partial t = (u(t+30) - u(t-30)) / 60$). In general, spatial derivatives were computed using centered differencing on the Arakawa A (unstaggered) grid, necessitating averaging of the model u , v , and w from the Arakawa C (staggered) grid. At the model boundaries, the π derivatives in $J_1 - J_3$ were computed using one-sided differences. In preliminary retrievals, setting the turbulence terms in the equations of motion to zero generally had negligible or mildly positive impact, presumably due to discretization errors being of similar magnitude to the turbulence terms themselves, which are typically much smaller than the remaining momentum equation terms. The turbulence terms were consequently omitted in the experiments shown.

Visual comparisons of the model and retrieved p_{final} , $\partial p_{\text{final}} / \partial x$, and $\partial p_{\text{final}} / \partial z$ reveal high fidelity in the retrieval technique (a representative example is shown in Figure 2), as do vertical profiles of root mean square error (RMSE) and relative (i.e., as a percentage of RMS p_{true}) RMSE (RRMSE) computed within the storm (Figure 3). (Both the model and retrieved pressure gradients were computed using centered differences on the Arakawa A grid.) Calculations of the horizontal domain-wide p bias at each level indicate that the 3D retrieval largely eliminates the vertically varying constant present in the 2D retrieval (Figure 4). To assess the sensitivity of the technique to hydrometeor mixing ratio errors (which are often large in EnKF analyses due to gross imperfections in current microphysical parameterization schemes), the retrieval was repeated with the hydrometeor drag term (in the vertical equation of motion) omitted. While ignoring water loading substantially degraded p_{3D} within the storm (Figure 5), p_{3D} was barely impacted in precipitation-free regions (not shown), resulting in relatively small domain-wide p_{3D} biases at each level and, thus, only minor error increases in the $A(k)$ estimates (Figure 4). The spurious vertical trend introduced to p_{final} (generally <

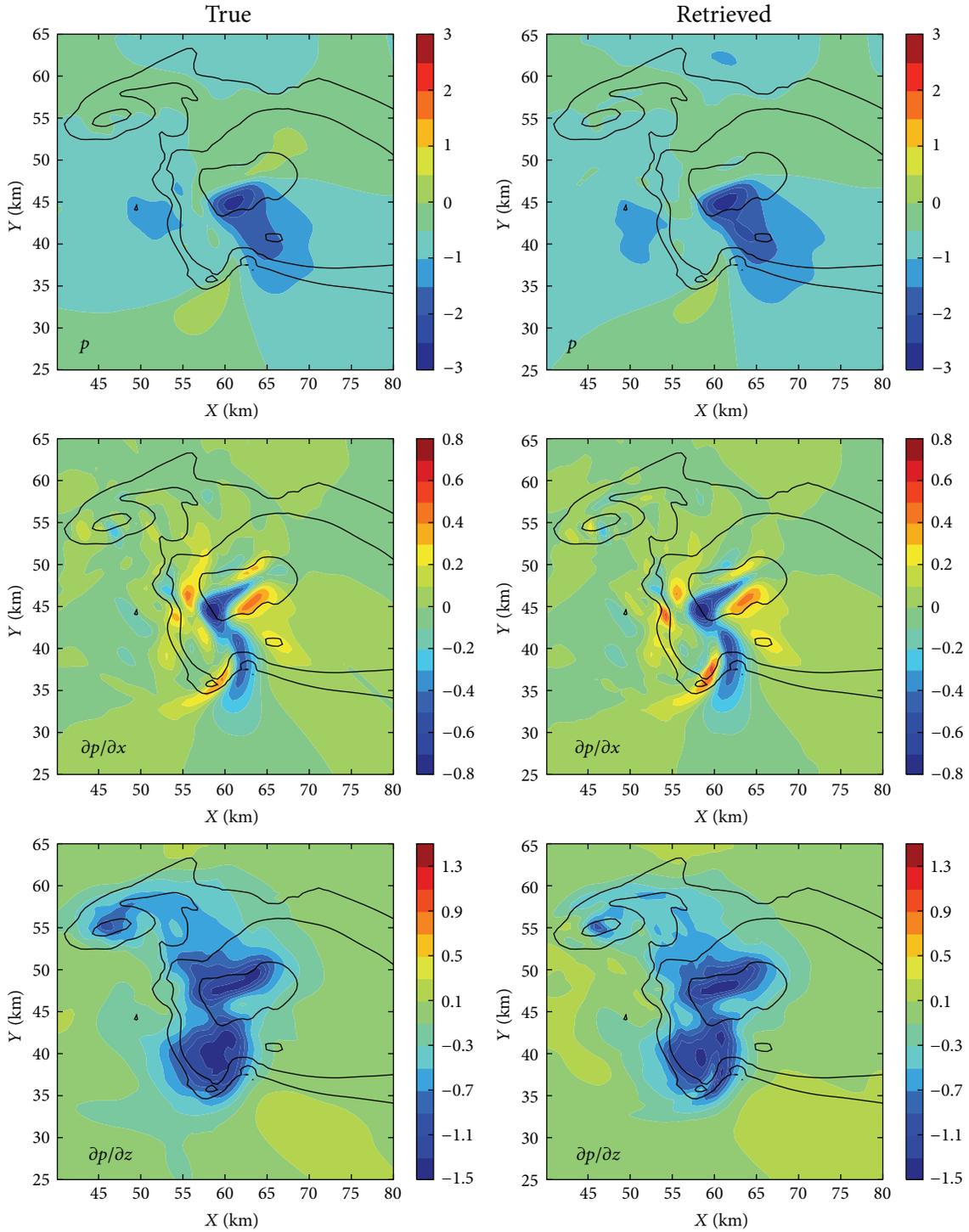


FIGURE 6: As in Figure 2 except for $\Delta t = 120$ s in the pressure retrieval.

0.05 hPa km^{-1}) by the degraded $A(k)$ was practically negligible. The relative insensitivity of the proposed retrieval method to hydrometeor errors is a major advantage over existing 3D approaches.

The retrievals were also relatively insensitive to temporal discretization errors in the local wind derivatives. Computing the latter using $\Delta t = 120$ s (rather than 30 s) to match

typical storm-scale EnKF analysis cycle periods substantially increased the RMSE $\partial p_{\text{final}}/\partial x$ and $\partial p_{\text{final}}/\partial z$ (Figure 3) but did not seriously qualitatively degrade the retrieval (Figure 6). While temporal discretization errors will be much larger in cases of rapidly moving storms (the present storm travels eastward at $\sim 10 \text{ m s}^{-1}$), such errors may be substantially reduced using advection correction methods [34–36].

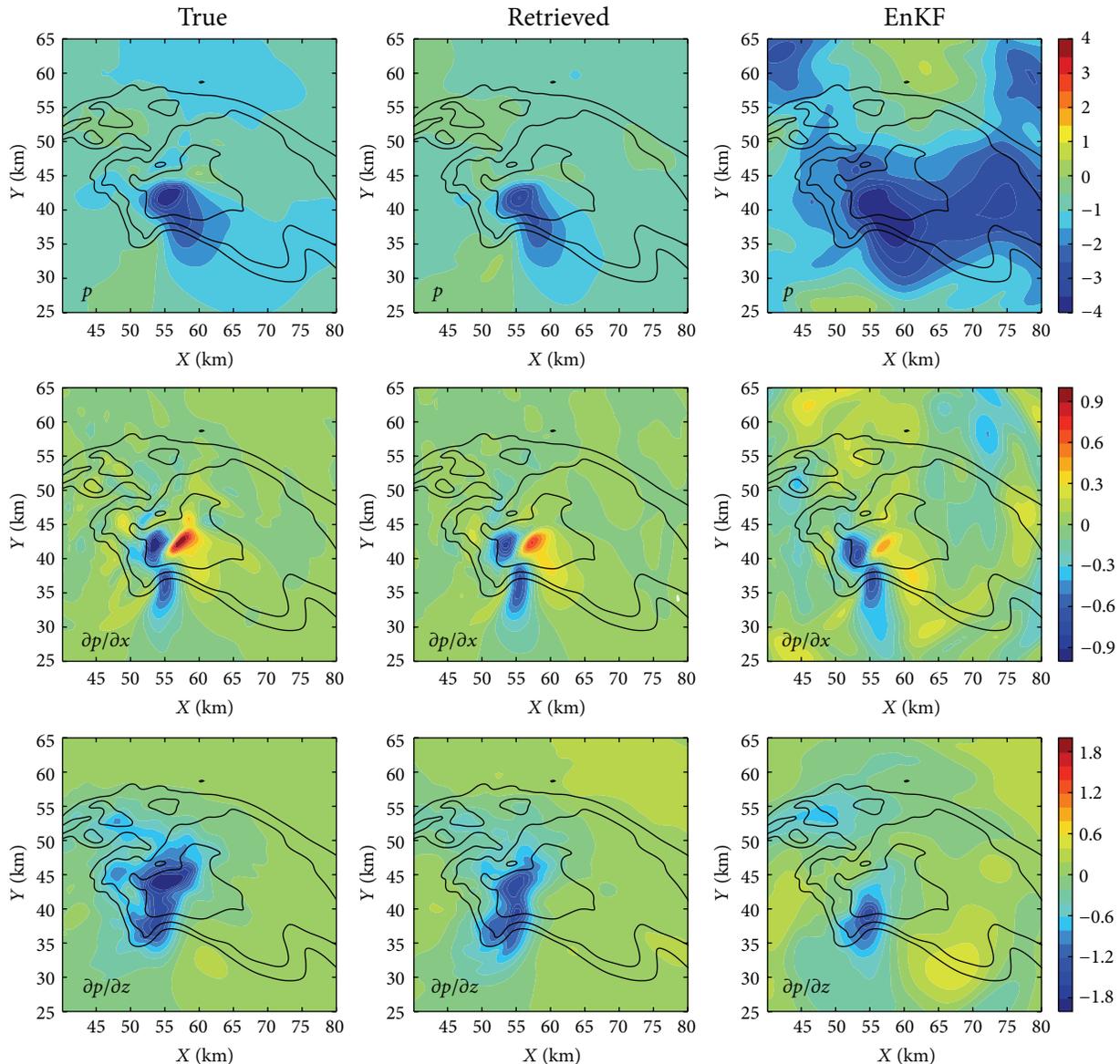
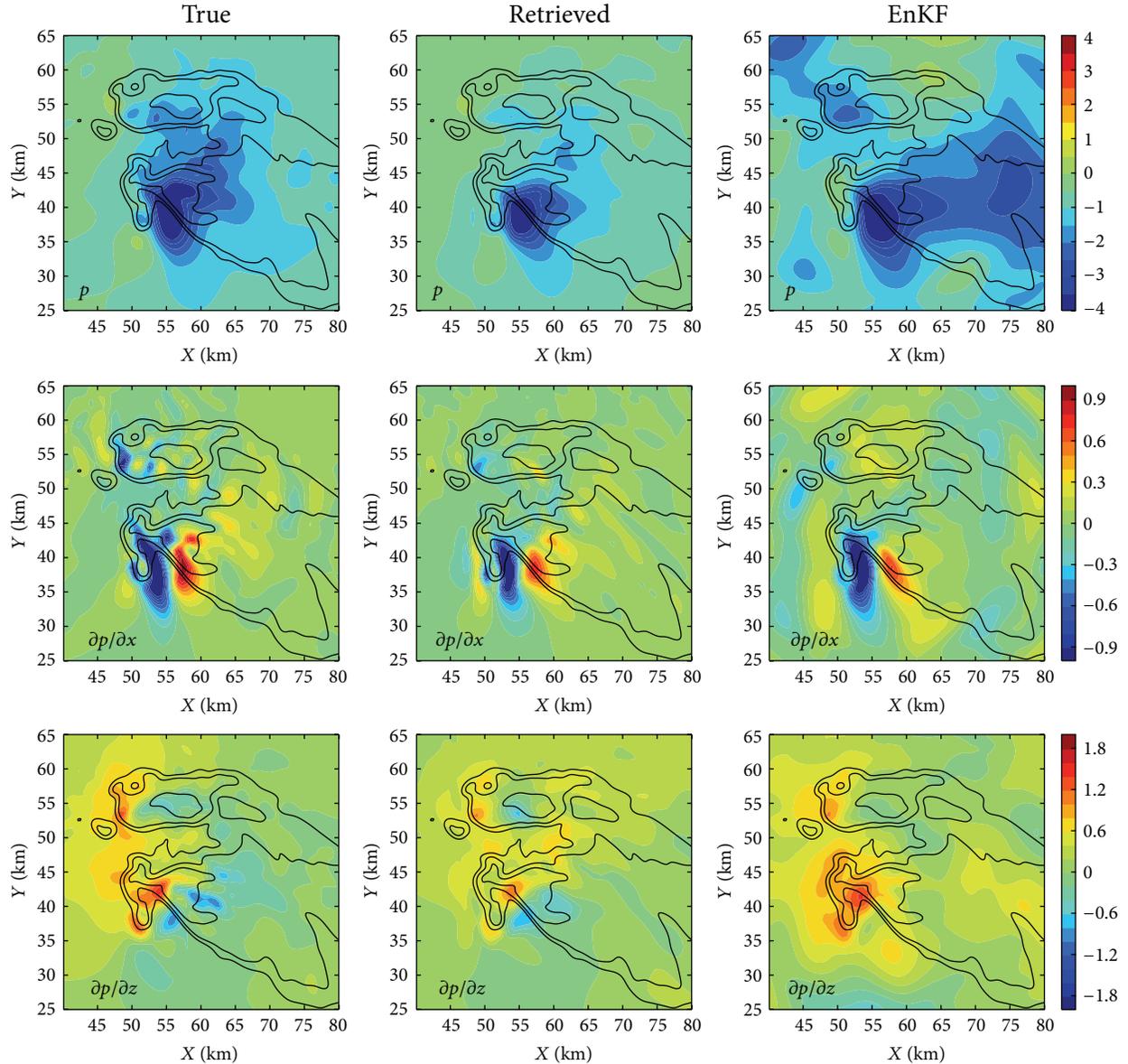


FIGURE 7: True (left column), retrieved (middle column), and EnKF mean posterior (right column) p (top row; hPa km^{-1}), $\partial p/\partial x$ (middle row; hPa km^{-1}), and $\partial p/\partial z$ (bottom row; hPa km^{-1}) at $z = 0.9 \text{ km}$, $t = 50 \text{ min}$. The true p were filtered and averaged as in Potvin et al. [23] to permit more direct comparisons with the (coarser) EnKF p . Radar reflectivity factor is contoured at 10, 30, and 50 dBZ.

3.2. Retrievals from a Simulated EnKF Mean Analysis. Having verified the robustness of the pressure retrieval scheme, we then applied the technique to mean posteriors from an EnKF OSSE conducted by Potvin and Wicker [11] (pictured in Figure 1). In that experiment (labeled “2-LFO” in [11]), mobile dual-Doppler velocity pseudoobservations were generated from the “truth” simulation of Potvin et al. [23], perturbed with random errors, and assimilated at two-minute intervals. Model error was introduced by using coarser ensemble grid spacing (600 m versus 200 m) and different microphysical parameterization (the Gilmore et al. [37] version of the Lin et al. [38] scheme versus the ZVD scheme) from those in the “truth” simulation. The use of an imperfect model in the EnKF resulted in dynamically unbalanced analysis

increments that instigated acoustic waves during the data assimilation. The ensemble was integrated for 30 min prior to the first analysis update to develop physically realistic covariance structures.

Local wind derivatives for the pressure retrieval scheme were computed using mean posteriors separated by 4 min (i.e., $\Delta t = 2 \text{ min}$). This interval is similar to the volume scan periods for Weather Surveillance Radar—1988 Doppler (WSR-88D) convection sampling patterns. All fields required for the momentum equation constraints were obtained from the EnKF posteriors (i.e., treated as known quantities). The pressure retrievals were performed offline from the EnKF analysis; that is, the retrieved posterior pressure field at each time was *not* used to initialize the subsequent forecast cycle

FIGURE 8: As in Figure 7 except for $z = 3.3$ km.

in the data assimilation experiment. Due to the very weak impact of the pressure field on the remaining state variables in compressible models (the property exploited by the proposed method to remove the fast-mode pressure oscillations), rebalancing the pressure field at each analysis cycle should have generally negligible impact on the remaining model fields throughout the data assimilation. We have verified this with separate, real data experiments (not shown), in which the differences due to the pressure rebalancing produced only tiny differences in the final ensemble analysis and in subsequent 30 min forecasts.

Given the insensitivity of the pressure retrievals in Section 3.1 to discretization and hydrometeor mixing ratio errors, it was expected that p_{final} obtained from the EnKF analyses would well capture the balanced portion of the

EnKF p (p_{bal}) and thereby improve upon the EnKF p and generally resemble p_{true} (keeping in mind that errors in the other EnKF fields will create some discrepancies between the EnKF p_{bal} and p_{true}). This was indeed the case throughout the assimilation period, except at higher altitudes. After 20 min of data assimilation ($t = 50$ min), the EnKF wind analyses are sufficiently accurate that p_{final} reasonably resembles p_{true} at lower and middle levels, much more so than does the EnKF p (Figures 7, 8, and 9). The improvement of p_{final} over the EnKF p , however, decreases with height (Figure 9), presumably due in large part to the shallower fast-mode pressure oscillations aloft (not shown). Within the top half of the storm, p_{final} is generally mildly inferior to the EnKF p , indicating that retrieval errors dominate improvements from removing the (small) fast-mode pressure errors. This

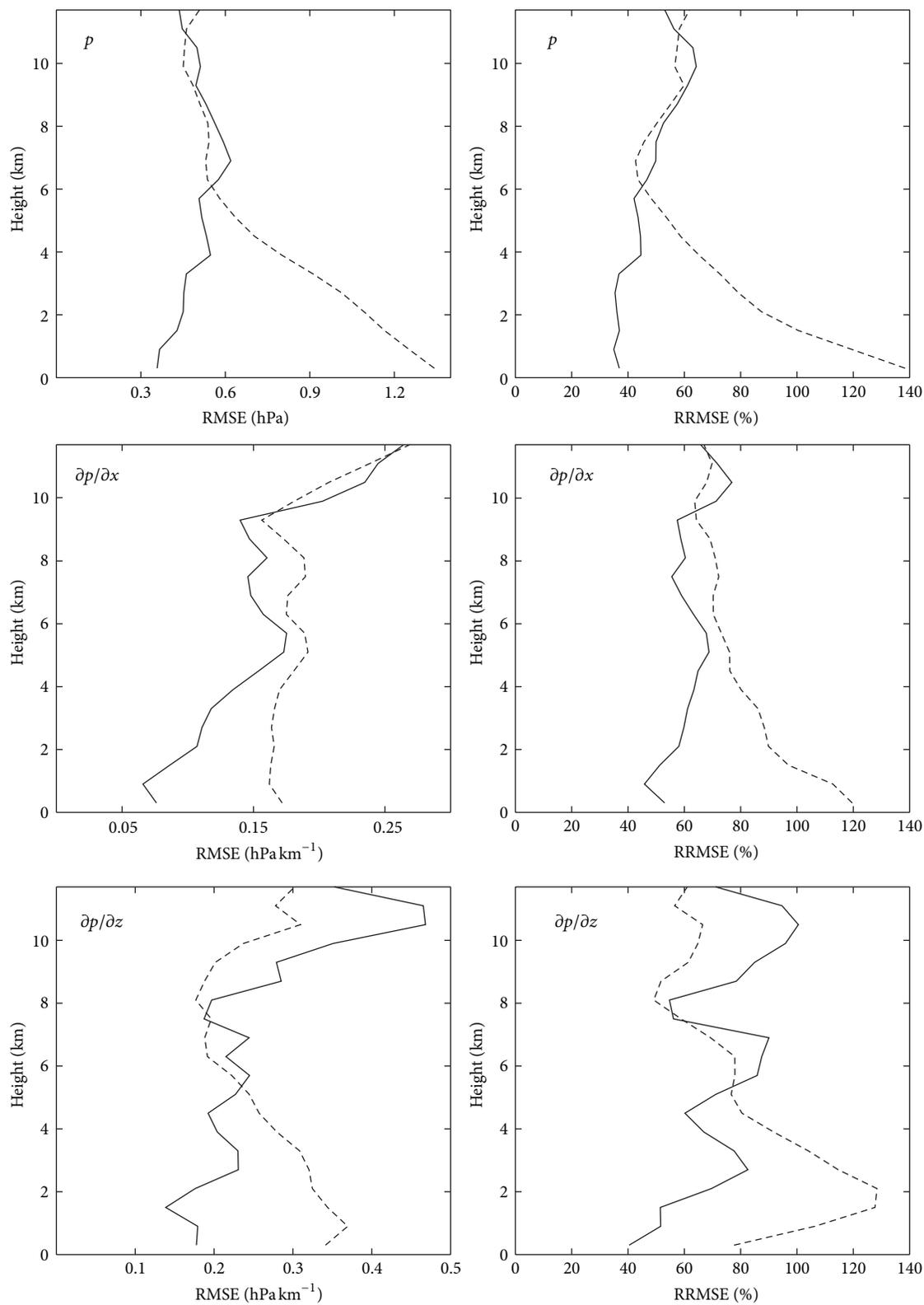


FIGURE 9: As in Figure 3 but for p_{final} retrieved from mean EnKF posterior (solid) and mean EnKF posterior p (dashed).

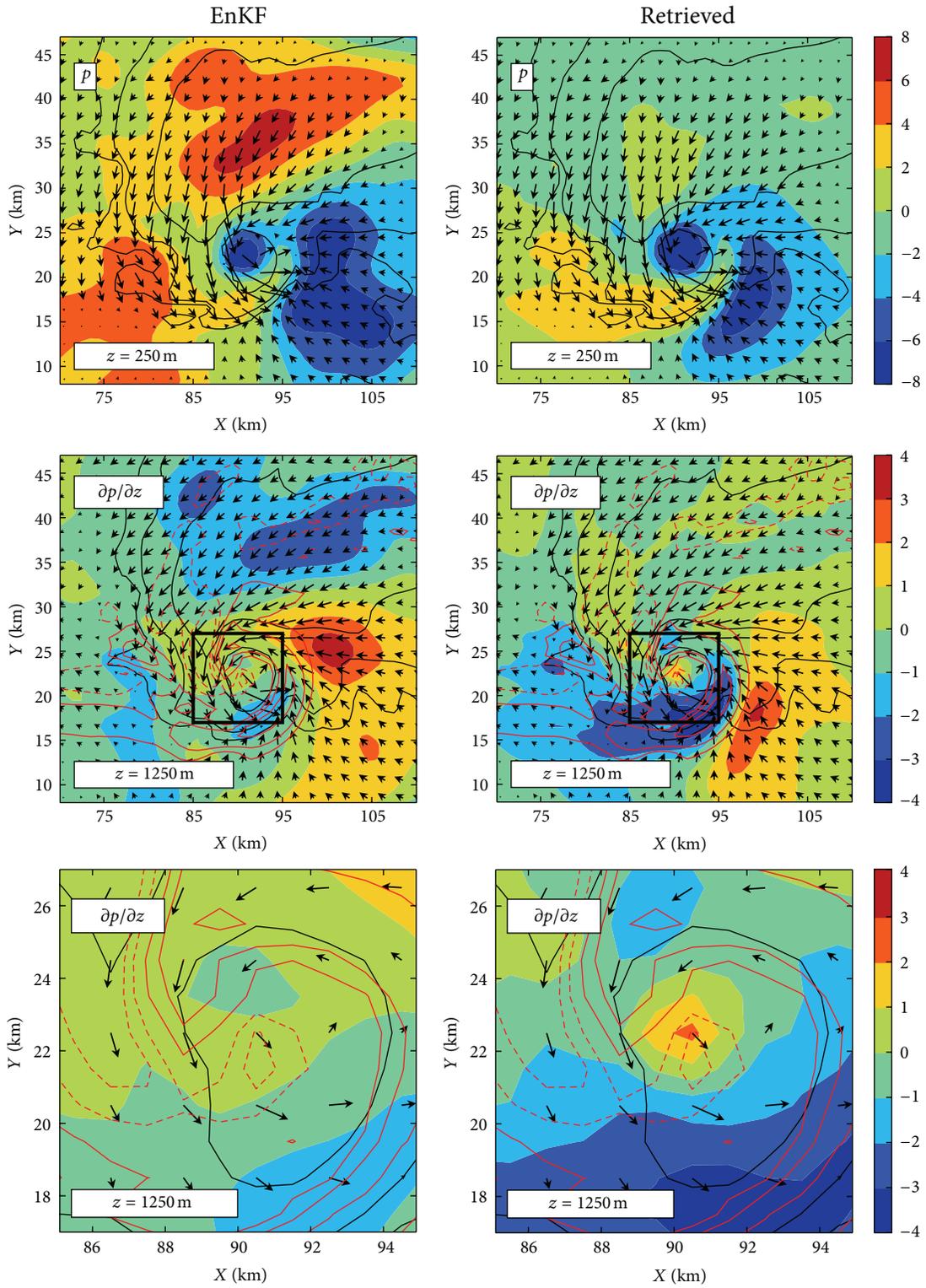


FIGURE 10: EnKF mean posterior (left column) and retrieved (right column) p at $z = 250$ m (hPa; top row) and $\partial p / \partial z$ at $z = 1250$ m (hPa km^{-1} ; middle and bottom rows). The black squares in the middle row outline the domain (centered near the occlusion updraft) in the bottom row. EnKF mean posterior vertical velocity (red; dashed = negative) is contoured at $-5, -2, 2,$ and 5 m s^{-1} and radar reflectivity factor (black) at 10, 30, and 50 dBZ. Arrows represent EnKF mean posterior horizontal winds.

suggests that application of the proposed pressure retrieval method be restricted to altitudes where spurious pressure oscillations are large.

While the retrieval results from this OSSE are likely somewhat optimistic despite our efforts to mitigate the “identical twin” problem, larger EnKF analysis errors in real cases should not substantially impede the technique’s ability to recover the EnKF p_{bal} and, thus, to remove the fast-mode pressure errors (which is the objective of the proposed method). Furthermore, the previous results, combined with the resilience of dual-mobile-Doppler EnKF wind analyses to microphysical and background wind errors [11, 39], suggest that pressure retrievals obtained from high-quality radar datasets such as those collected during the second Verification of the Origins of Rotation in Tornadoes Experiment (VORTEX2; [40]) may be useful for investigating pressure-gradient-driven storm processes. Rigorous testing of this hypothesis, unfortunately, is currently precluded by the lack of dense 3D pressure observations within storms.

3.3. Retrievals from a Real Data EnKF Mean Analysis. While our ability to verify the real-world performance of the proposed technique is quite limited, qualitative evaluation of pressure retrievals from EnKF analyses of the May 29–30, 2004, Geary, OK, USA, tornadic supercell supports cautious optimism. The EnKF analyses were obtained by assimilating quality-controlled Doppler velocity data from two Shared Mobile Atmospheric Research and Teaching (SMART; [41]) radars using the NCOMMAS EnKF with the ZVD microphysics scheme (see Potvin et al. [39] for further details). Pressure retrievals were performed in the same manner as with the simulated EnKF analyses in Section 3.2.

A representative comparison of the p_{final} and EnKF p , valid after 44 min of data assimilation (0037 UTC 30 May), is shown in Figure 10. As in the EnKF OSSE (Section 3.2), high-amplitude pressure oscillations generated during the data assimilation are not evident in the retrieved pressure fields. Furthermore, the p_{final} comports with present understanding of the pressure distribution within supercells, which suggests retrieval errors are not unduly large. For example, both the p_{final} and EnKF p exhibit an inflow low and a deep pressure minimum associated with the strong mesocyclonic rotation (Figures 10(a) and 10(b)), as well as upward-directed perturbation pressure gradient force along the gust front (Figures 10(c) and 10(d)). The p_{final} , but not the EnKF p , persistently indicates a region of substantial downward-directed perturbation pressure gradient force near the analyzed occlusion downdraft (Figures 10(c)–10(f)). This is consistent with current understanding of occlusion downdraft formation [42], suggesting that the EnKF pressure analysis near the occlusion downdraft is contaminated with fast-mode errors that the retrieval successfully corrects.

4. Conclusion

As in other data assimilation frameworks, storm-scale EnKF radar data assimilation can introduce dynamical imbalances

that induce severe pressure errors in analyses, inhibiting investigation of important storm processes. Retrieval procedures that have traditionally been used to retrieve the pressure and buoyancy fields from dual-Doppler wind analyses can be used to rebalance the pressure field in such cases. Using a combination of the 2D and 3D pressure retrieval approaches that mitigates the deficiencies of both, we have demonstrated that useful analyses of the perturbation pressure field and its derivatives can be obtained from EnKF analyses despite model and discretization errors. The described procedure could be modified to separately retrieve, for example, the linear and nonlinear components of the dynamic perturbation pressure [43]. The method could also be extended to other nonhydrostatic equation sets including that used by the WRF-ARW model.

It should be noted that some of the existing methods for suppressing acoustic and gravity waves in model initial conditions could be adapted to removing pressure oscillations in EnKF analyses. For example, during data assimilation, digital filter initialization (DFI; [44]) could be applied to the analysis increments at each cycle prior to integrating the ensemble forward to the next time. All such approaches, however, would increase computational cost, and it is not clear whether storm-scale analyses would improve given the model approximations used in these methods (e.g., the adiabatic backward model integration in DFI). Alternatively, a diagnostic pressure equation (e.g., [45]) could be solved (either exactly, or weakly as in the proposed method), thereby avoiding temporal discretization errors, but at the cost of increased spatial discretization errors. While we recommend evaluation of the relative efficacy of alternative methods for recovering the balanced pressure component, we also emphasize that the proposed method provides an existing, effective, simple, and low-cost way to remove pressure oscillations from storm-scale EnKF analyses and does not require data assimilation experiments to be rerun.

Acknowledgments

Funding was provided by the NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement no. NA11OAR4320072, US Department of Commerce. This work benefitted from discussions with Patrick Skinner and Alan Shapiro. The authors thank the three anonymous reviewers for their helpful comments that substantially improved the paper and Michael French for his useful comments on an earlier version of the paper. The SMART radar data were collected and provided by Michael Biggerstaff and edited by Kristin Kuhlman and Daniel Betten under NSF Grants ATM-0619715 and ATM-0802717.

References

- [1] G. Evensen, “Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics,” *Journal of Geophysical Research*, vol. 99, no. 5, pp. 10–162, 1994.

- [2] C. Snyder and F. Q. Zhang, "Assimilation of simulated Doppler radar observations with an ensemble Kalman filter," *Monthly Weather Review*, vol. 131, no. 8, pp. 1663–1677, 2003.
- [3] F. Q. Zhang, C. Snyder, and J. Z. Sun, "Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter," *Monthly Weather Review*, vol. 132, no. 5, pp. 1238–1253, 2004.
- [4] D. C. Dowell, F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, "Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: ensemble Kalman filter experiments," *Monthly Weather Review*, vol. 132, no. 8, pp. 1982–2005, 2004.
- [5] M. J. Tong and M. Xue, "Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments," *Monthly Weather Review*, vol. 133, no. 7, pp. 1789–1807, 2005.
- [6] M. Xue, M. J. Tong, and K. K. Droegemeier, "An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 1, pp. 46–66, 2006.
- [7] A. Aksoy, D. C. Dowell, and C. Snyder, "A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: storm-scale analyses," *Monthly Weather Review*, vol. 137, no. 6, pp. 1805–1824, 2009.
- [8] D. T. Dawson, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, "Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones," *Monthly Weather Review*, vol. 140, no. 2, pp. 696–716, 2012.
- [9] D. C. Dowell, L. J. Wicker, and C. Snyder, "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*, vol. 139, no. 1, pp. 272–294, 2011.
- [10] J. Marquis, Y. Richardson, P. Markowski, D. Dowell, and J. Wurman, "Tornado maintenance investigated with high-resolution dual-doppler and EnKF analysis," *Monthly Weather Review*, vol. 140, no. 1, pp. 3–27, 2012.
- [11] C. K. Potvin and L. J. Wicker, "Comparison between Dual-Doppler and EnKF Storm-Scale Wind Analyses: observing system simulation experiments with a supercell thunderstorm," *Monthly Weather Review*, vol. 140, no. 12, pp. 3972–3991, 2012.
- [12] R. L. Tanamachi, L. J. Wicker, D. C. Dowell, H. B. Bluestein, D. T. Dawson, M. Xue et al., "EnKF assimilation of high-resolution, mobile Doppler radar data of the 4 May 2007 Greensburg, Kansas, Supercell into a numerical cloud model," *Monthly Weather Review*, vol. 141, no. 2, pp. 625–648, 2013.
- [13] L. J. Wicker and W. C. Skamarock, "Time-splitting methods for elastic models using forward time schemes," *Monthly Weather Review*, vol. 130, no. 8, pp. 2088–2097, 2002.
- [14] M. C. Coniglio, D. J. Stensrud, and L. J. Wicker, "Effects of upper-level shear on the structure and maintenance of strong quasi-linear mesoscale convective systems," *Journal of the Atmospheric Sciences*, vol. 63, no. 4, pp. 1231–1252, 2006.
- [15] J. S. Whitaker and T. M. Hamill, "Ensemble data assimilation without perturbed observations," *Monthly Weather Review*, vol. 130, no. 7, pp. 1913–1924, 2002.
- [16] J. Anderson, T. Hoar, K. Raeder et al., "The data assimilation research testbed a community facility," *Bulletin of the American Meteorological Society*, vol. 90, no. 9, pp. 1283–1296, 2009.
- [17] W. C. Skamarock, J. B. Klemp, J. Dudhia et al., "A description of the Advanced Research WRF version 2," Tech. Rep., NCAR, 2005.
- [18] J. B. Klemp and R. Rotunno, "A study of the tornadic region within a supercell thunderstorm," *Journal of the Atmospheric Sciences*, vol. 40, no. 2, pp. 359–377, 1983.
- [19] J. D. Marwitz, "Trajectories within the weak echo region of hailstorms," *Journal of Applied Meteorology*, vol. 12, no. 7, pp. 1174–1182, 1973.
- [20] R. Rotunno and J. B. Klemp, "The influence of shear-induced pressure gradient on thunderstorm motion," *Monthly Weather Review*, vol. 110, no. 2, pp. 136–151, 1982.
- [21] R. A. Houze Jr., S. A. Rutledge, M. I. Biggerstaff, and B. F. Smull, "Interpretation of Doppler weather radar displays of midlatitude mesoscale convective systems," *Bulletin of the American Meteorological Society*, vol. 70, no. 6, pp. 608–619, 1989.
- [22] Z. Byko, P. Markowski, Y. Richardson, J. Wurman, and E. Adlerman, "Descending reflectivity cores in supercell thunderstorms observed by mobile radars and in a high-resolution numerical simulation," *Weather and Forecasting*, vol. 24, no. 1, pp. 155–186, 2009.
- [23] C. K. Potvin, L. J. Wicker, and A. Shapiro, "Assessing errors in variational dual-doppler wind syntheses of supercell thunderstorms observed by storm-scale mobile radars," *Journal of Atmospheric and Oceanic Technology*, vol. 29, no. 8, pp. 1009–1025, 2012.
- [24] R. Daley, *Atmospheric Data Analysis*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, Cambridge, UK, 1991.
- [25] D. T. Kleist, D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, "Improving incremental balance in the GSI 3DVAR analysis system," *Monthly Weather Review*, vol. 137, no. 3, pp. 1046–1060, 2009.
- [26] T. Galchen, "Method for initialization of anelastic equations—implications for matching models with observations," *Monthly Weather Review*, vol. 106, no. 5, pp. 587–606, 1978.
- [27] C. E. Hane, R. B. Wilhelmson, and T. Gal-Chen, "Retrieval of thermodynamic variables within deep convective clouds: experiments in three dimensions," *Monthly Weather Review*, vol. 109, no. 3, pp. 564–576, 1981.
- [28] E. A. Brandes, "Relationships between radar-derived thermodynamic variables and tornadogenesis," *Monthly Weather Review*, vol. 112, no. 5, pp. 1033–1052, 1984.
- [29] F. Roux, "Retrieval of thermodynamic fields from multiple-Doppler radar data using the equations of motion and the thermodynamic equation," *Monthly Weather Review*, vol. 113, no. 12, pp. 2142–2157, 1985.
- [30] Y. C. Liou, T. C. C. Wang, and K. S. Chung, "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*, vol. 42, no. 10, pp. 1443–1454, 2003.
- [31] J. B. Klemp and R. B. Wilhelmson, "Simulation of 3-dimensional convective storm dynamics," *Journal of the Atmospheric Sciences*, vol. 35, no. 6, pp. 1070–1096, 1978.
- [32] C. L. Ziegler, "Retrieval of thermal and microphysical variables in observed convective storms. Part I: model development and preliminary testing," *Journal of the Atmospheric Sciences*, vol. 42, no. 14, pp. 1487–1509, 1985.

- [33] E. R. Mansell, C. L. Ziegler, and E. C. Bruning, "Simulated electrification of a small thunderstorm with two-moment bulk microphysics," *Journal of the Atmospheric Sciences*, vol. 67, no. 1, pp. 171–194, 2010.
- [34] T. G.-C. Tzvi Gal-Chen, "Errors in fixed and moving frame of references: applications for conventional and Doppler radar analysis," *Journal of the Atmospheric Sciences*, vol. 39, no. 10, pp. 2279–2300, 1982.
- [35] A. Shapiro, K. M. Willingham, and C. K. Potvin, "Spatially variable advection correction of radar data. Part I: theoretical considerations," *Journal of the Atmospheric Sciences*, vol. 67, no. 11, pp. 3445–3456, 2010.
- [36] A. Shapiro, K. M. Willingham, and C. K. Potvin, "Spatially variable advection correction of radar data. Part II: test results," *Journal of the Atmospheric Sciences*, vol. 67, no. 11, pp. 3457–3470, 2010.
- [37] M. S. Gilmore, J. M. Straka, and E. N. Rasmussen, "Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme," *Monthly Weather Review*, vol. 132, no. 11, pp. 2610–2627, 2004.
- [38] Y.-L. Lin, R. D. Farley, and H. D. Orville, "Bulk parameterization of the snow field in a cloud model," *Journal of Climate & Applied Meteorology*, vol. 22, no. 6, pp. 1065–1092, 1983.
- [39] C. K. Potvin, J. W. Louis, M. I. Biggerstaff, D. Betten, and A. Shapiro, "Comparison between Dual-Doppler and EnKF storm-scale wind analyses: the 29–30 May 2004 Geary, Oklahoma, Supercell thunderstorm," *Monthly Weather Review*, vol. 141, no. 5, pp. 1612–1628, 2013.
- [40] J. Wurman, D. Dowell, Y. Richardson et al., "The second verification of the origins of rotation in tornadoes experiment Vortex2," *Bulletin of the American Meteorological Society*, vol. 93, no. 8, pp. 1147–1170, 2012.
- [41] M. I. Biggerstaff, L. J. Wicker, J. Guynes et al., "The shared mobile atmospheric research and teaching radar: a collaboration to enhance research and teaching," *Bulletin of the American Meteorological Society*, vol. 86, no. 9, pp. 1263–1274, 2005.
- [42] P. M. Markowski, "Hook echoes and rear-flank downdrafts: a review," *Monthly Weather Review*, vol. 130, no. 4, pp. 852–876, 2002.
- [43] H. Cai and R. M. Wakimoto, "Retrieved pressure field and its influence on the propagation of a supercell thunderstorm," *Monthly Weather Review*, vol. 129, no. 11, pp. 2695–2713, 2001.
- [44] P. Lynch and X.-Y. H. Xiang-Yu Huang, "Initialization of the HIRLAM model using a digital filter," *Monthly Weather Review*, vol. 120, no. 6, pp. 1019–1034, 1992.
- [45] G. Q. Ge, J. D. Gao, and M. Xue, "Diagnostic pressure equation as a weak constraint in a storm-scale three-dimensional variational radar data assimilation system," *Journal of Atmospheric and Oceanic Technology*, vol. 29, no. 8, pp. 1075–1092, 2012.

Research Article

Improved Doppler Velocity Dealiasing for Radar Data Assimilation and Storm-Scale Vortex Detection

Qin Xu,¹ Kang Nai,² Shun Liu,³ Chris Karstens,^{1,2} Travis Smith,^{1,2} and Qingyun Zhao⁴

¹ National Severe Storms Laboratory, Norman, OK 73072, USA

² Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, OK 73072, USA

³ National Centers of Environmental Prediction and I. M. Systems Group Inc., College Park, MD 20740, USA

⁴ Marine Meteorology Division, Naval Research Laboratory, Monterey, CA 93943, USA

Correspondence should be addressed to Qin Xu; qin.xu@noaa.gov

Received 28 May 2013; Revised 6 August 2013; Accepted 9 August 2013

Academic Editor: David J. Stensrud

Copyright © 2013 Qin Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Doppler velocity dealiasing technique based on alias-robust VAD and variational (AR-Var) analyses developed at the National Severe Storms Laboratory for radar data quality control and assimilation is further improved in its two-step procedures: the reference check in the first step and the continuity check in the second step. In the first step, the alias-robust variational analysis is modified adaptively and used in place of the alias-robust velocity-azimuth display (VAD) analysis for all scan modes (rather than solely the WSR-88D volume coverage pattern 31 with the Nyquist velocity v_N reduced below 12 m s^{-1} and the TDWR Mod80 with v_N reduced below 15 m s^{-1}), so more raw data can pass the stringent threshold conditions used by the reference check in the first step. This improves the dealiasing data coverage without false dealiasing to better satisfy the high data quality standard required by radar data assimilation. In the second step, new procedures are designed and added to the continuity check to increase the dealiasing data coverage over storm-scale areas threatened by intense mesocyclones and their generated tornados. The performances of the improved dealiasing technique versus the existing techniques are exemplified by the results obtained for tornadic storms scanned by the operational KTLX radar in Oklahoma.

1. Introduction

Radar data quality control is critical for radar data assimilation, and dealiasing is an important and yet often very difficult part of radar data quality control. Currently, the local environment dealiasing algorithm (Eilts and Smith [1]) is used on WSR-88D radars (NEXRAD Network) for processing radar data for Doppler velocity data in real-time. When this algorithm was originally developed at the National Severe Storms Laboratory (NSSL), the primary goal was for visual and certain quantitative applications, such as automated mesocyclone detections (Stumpf et al. [2] and Smith and Elmore [3]), with considerable tolerance for bad or poor quality data to retain as much as possible the original data coverage. Therefore, the processed data often do not satisfy the high-quality standard required by data assimilation at the National Centers of Environmental Prediction (NCEP). This problem is common for most early developed dealiasing techniques, especially for those designed primarily for

applications not involving data assimilation including the currently developed and tested techniques (such as the two-dimensional dealiasing of Jing and Wiener [4]) for broad applications at the NWS Radar Operations Center (Burgess and Crum [5] and Witt et al. [6]). Striving to meet the need of radar data assimilation, continued efforts have been made at NSSL in collaboration with other research institutes to develop robust dealiasing techniques (Gong et al. [7], Gao et al. [8], Zhang and Wang [9], Xu et al. [10], and Xu and Nai [11]). In particular, the dealiasing technique of Xu et al. [10] has been tested extensively with real-time observations from operational WSR-88D radars under various weather conditions. From these tests, the technique was found to be capable of correcting alias errors without false dealiasing in most cases, but it occasionally failed to correct or flag severely aliased radial velocities around strongly sheared and often tilted inversion layers in severe winter ice storms scanned by the operational WSR-88D radars using volume coverage

pattern 31 (VCP31) with $v_N < 12 \text{ m s}^{-1}$, where v_N denotes the Nyquist velocity. The occasional failures were caused by the lack of the required accuracy (within $\pm 7v_N/4$ of the true value at every observation point) of the reference radial-velocity field produced by the alias-robust velocity-azimuth display (AR-VAD) analysis (Xu et al. [12]). To solve this problem (caused mainly by reduced v_N), an alias-robust variational (AR-Var) analysis was developed (Xu and Nai [13]) and used in place of the AR-VAD analysis for the reference check to upgrade the VAD-based dealiasing (Xu et al. [10]) adaptively for VCP31 into an AR-Var-based dealiasing (Xu and Nai [11]). As the above dealiasing techniques were developed for data assimilation applications at NCEP (Liu et al. [14]), various stringent threshold conditions must be used to ensure the dealiased data to be completely free of false dealiasing. As these stringent conditions tend to discard many data that cannot pass the thresholds even after dealiasing, the dealiased data often have less or even much less coverage than the raw data, especially in isolated data areas far away from the radar or in localized areas of strong and complex winds such as those around mesocyclones. It is thus necessary to further improve these dealiasing techniques to increase their dealiased data coverage but remain to be completely free of false dealiasing, especially over storm-scale areas threatened by intense mesocyclones and their generated tornados. The above existing dealiasing techniques for data assimilation are reviewed in Section 2. Improvements made to the existing dealiasing techniques are presented in Section 3. Conclusions follow in Section 4.

2. Review of VAD-Based and AR-Var-Based Dealiasing Techniques for Data Assimilation

The VAD-based dealiasing technique (Xu et al. [10]) developed at NSSL for radar wind data quality control was delivered to NCEP for radar data assimilation tests and applications (Liu et al. [14]). This dealiasing performs two steps: (i) the reference check with the reference radial-velocity field produced by the AR-VAD analysis (Xu et al. [12]) to detect and correct alias errors in the all-flagged raw data and (ii) the continuity check with an enhanced use of available seed data (i.e., dealiased or deflagged data that have passed the reference check and thus have been accepted as good data) in a properly enlarged block area near each remaining flagged data point being checked. The technique has been tested with raw radial velocities scanned from operational WSR-88D radars under various weather conditions. In most cases, the algorithm can correct aliased velocities without false dealiasing. However, like the classic VAD analysis, the AR-VAD analysis is inherently limited by its uniform-wind assumption, so the resultant reference radial velocities often do not have the desired flexibility and variability to allow most raw data to pass the stringent threshold conditions, especially when the Nyquist velocity v_N is reduced, and thus, the threshold ranges are further narrowed. Because of this, the VAD-based dealiasing occasionally fails to detect and correct severely aliased velocities around strongly sheared and

tilted inversion layers in severe winter ice storms observed by operational WSR-88D radars using VCP31 with $v_N < 12 \text{ m s}^{-1}$. To solve this problem, the AR-Var analysis (Xu et al. [15]) was refined into a two-step analysis (Xu and Nai [13]) and used in place of the AR-VAD analysis for the reference check. In the first step of the refined AR-Var analysis, the original AR-VAD analysis was modified to fit the raw aliased radial-velocity observations in the $\pm 90^\circ$ vicinity of each of the two paired zero radial-velocity points (about 180° apart) on a selected range circle. This modification resembles the sectorized uniform-wind technique applied with two 180° sectors. The two analyzed radial-velocity fields were then combined into a single radial-velocity field not rigidly constrained by the VAD uniform-wind assumption. In the second step, the combined radial-velocity field was used as the first guess for the refined AR-Var analysis. This upgraded the AR-VAD-based dealiasing into the AR-Var-based dealiasing (Xu and Nai [11]) adaptively for winter ice storms scanned by WSR-88D radars using VCP31 (at 0.5° , 1.5° , 2.5° , 3.5° , and 4.5° tilts) as well as for storms scanned by the FAA Terminal Doppler Weather Radar (TDWR) radars using Mod80 (with $v_N < 15 \text{ m s}^{-1}$ at 0.5° , 1.0° , 2.5° , 5.1° , 7.7° , 11.3° , 15.3° , 20.7° , and 28.2° tilts).

The AR-Var-based dealiasing has been tested successfully with many winter ice storms scanned by operational WSR-88D radars using VCP31 with $v_N < 12 \text{ m s}^{-1}$ and many storms scanned by the FAA OKC airport TDWR radar using Mod80 with $v_N < 15 \text{ m s}^{-1}$. The effectiveness of this dealiasing technique is shown by the example in Figure 1 (in addition to the example in Figure 2 of Xu and Nai [11]). As shown in Figure 1(a)(or Figure 1(b)), the raw radial velocities were severely aliased in most areas outside the 25 (or 10) km radial range, and the aliased velocities were even folded twice in areas to the southwest and northeast of the radar outside the 150 (or 60) km radial range on the 1.5° (or 4.5°) tilt. As shown by the dealiased radial-velocity field in Figure 1(c) (or Figure 1(d)), the wind field is dominated by a strong northeasterly flow below the vertical level of $z \approx 0.7 \text{ km}$ (that is, within the radial range of $r \approx 25$ (or 15) km on 1.5° (or 4.5°) tilt) and then veers sharply to a strong southwesterly flow above $z \approx 0.8 \text{ km}$. This sharp change in wind direction is a typical feature often observed during winter ice storms, and this feature is well captured by the two-step AR-Var analysis but not retrieved by the AR-VAD analysis due to the reduced v_N ($=11.51 \text{ m s}^{-1}$) and the increased discrepancy between the VAD-analyzed uniform wind and the true wind, especially around the vertical-shear layer that is not perfectly flat.

The AR-Var-based dealiasing has been incorporated into the radar data quality control package (Liu et al. [14]) at the National Centers for Environmental Prediction (NCEP) for operational tests. For radar data assimilation applications at NCEP, the method is required to be absolutely free of false dealiasing. This has inevitably sacrificed the data coverage to certain but different degrees for AR-Var-based dealiasing and AR-VAD-based dealiasing. The sacrificed (flagged) data coverage is usually minor and marginal for the AR-Var-based dealiasing applied adaptively to winter ice storms since the raw velocities scanned from ice storms using VCP31 usually

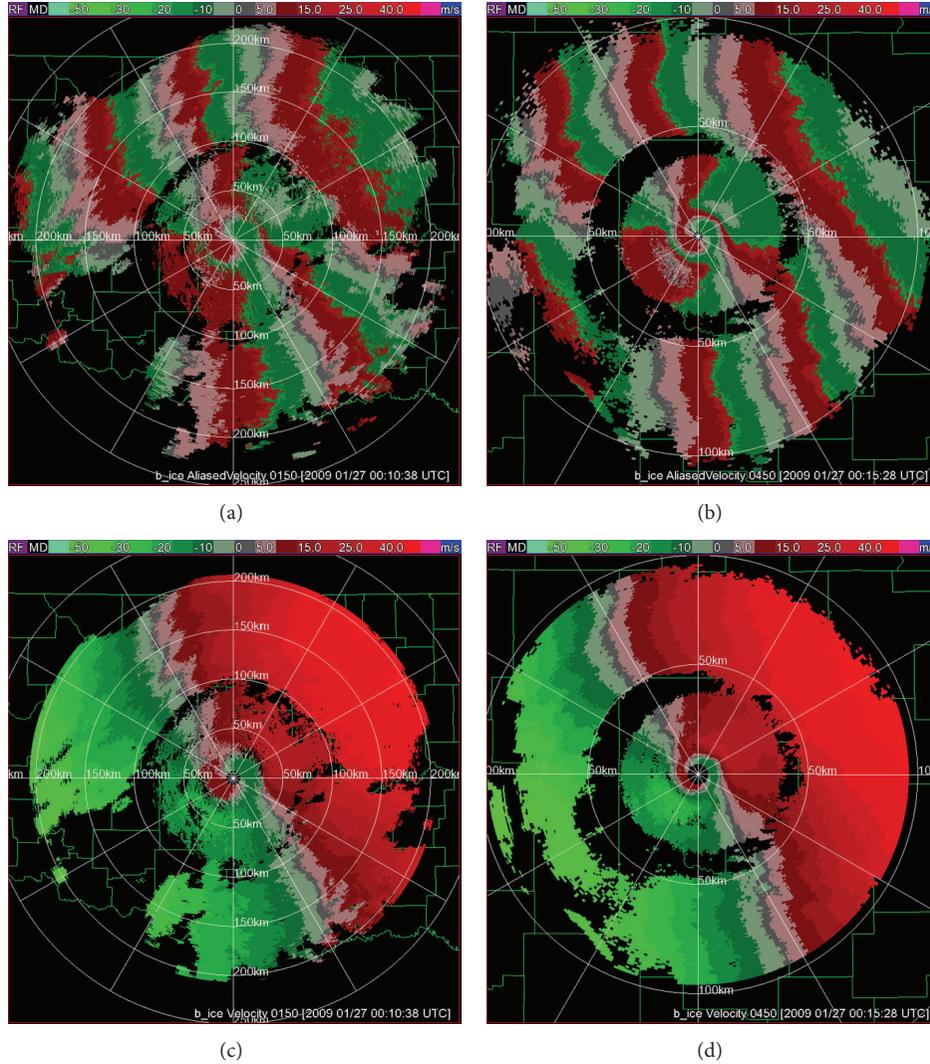


FIGURE 1: Raw velocity image scanned by the KTLX radar with $v_N = 11.51 \text{ m s}^{-1}$ (a) on 1.5° tilt at 00:10:38 UTC and (b) on 4.5° tilt at 00:15:28 UTC for the ice storm on January 27, 2009. Dealiased velocity image on (c) 1.5° and (d) 4.5° tilts by the AR-Var-based dealiasing.

have rather continuous and complete coverage on each tilt, and so do the dealiased data (as shown in Figure 1). However, for the VAD-based dealiasing still used for other VCPs, the reference check in the first main step has to be limited within the cut-off radial range (30 km for 0.5° tilt and 80 km for all other higher tilts; see Section 2b of Xu et al. [10]) to avoid false dealiasing, so isolated data areas away (by 40 gates and 5 beams or more) from the seed data areas produced by the reference check will remain flagged as they cannot be reached by the continuity check in the second step of the VAD-based dealiasing (see Section 2c of Xu et al. [10]). This limitation can become serious for a spring or summer convective storm when the raw velocities scanned from the storm using VCP11, VCP211, VCP12, or VCP121 (with $v_N > 20 \text{ m s}^{-1}$ at 14 tilts from 0.5° to 19.5°) are scattered sparsely in large isolated areas far away from the radar. To reduce this limitation, further improvements are made to both the reference check in the first step and the continuity check in the second step. These new improvements are described in Section 3.

3. New Improvements and Results

To increase the coverage of seed data produced by the reference check, the AR-Var-based dealiasing is extended and used in place of the VAD-based dealiasing for all the remaining VCPs (other than VCP31 for WSR-88D and Mod80 for TDWR) with two new adaptations (for $v_N > 18 \text{ m s}^{-1}$ only): (i) The first-guess background radial-velocity field (used to start the descent algorithm in the second step of the refined AR-Var analysis) is taken directly from the AR-VAD analysis without modification in the first step of the refined AR-Var analysis, because the modified AR-VAD analysis is necessary only for $v_N < 18 \text{ m s}^{-1}$ but it requires sufficient raw-data coverage on each selected range circle as explained in Section 2. (ii) The cut-off radial range limitation (see Section 2b of Xu et al. [10]) is removed for the reference check, because the refined AR-Var analysis is not rigidly constrained by the VAD uniform-wind assumption and thus can fit the data much better than the AR-VAD analysis at

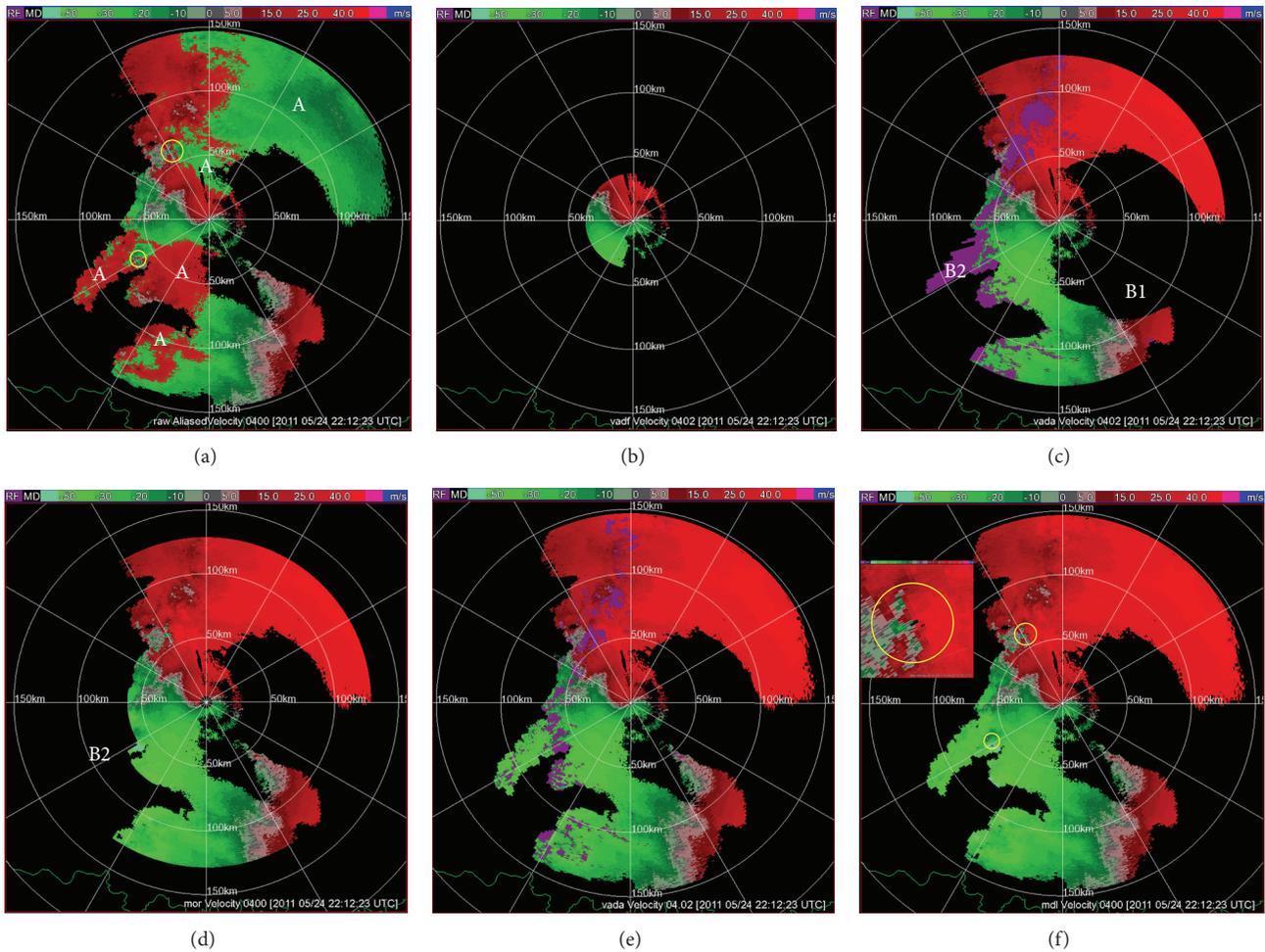


FIGURE 2: (a) Raw radial-velocity image scanned by the KTLX radar using VCPI2 with $v_N = 26.1 \text{ m s}^{-1}$ at 4.0° tilt for the Oklahoma tornadic storm system at 22:12:23 UTC on May 24, 2011. (b) Dealiased radial-velocity (seed data) image produced by the improved reference check alone. (c) Dealiased radial-velocity image (with the special data points plotted in purple) produced by the first one-way forward procedure in the continuity check after the improved reference check. (d) Final dealiased radial-velocity image produced by the new extended AR-Var-based method. (e) As in panel (c) but the first guess is from the RAP-predicted wind field for the refined AR-Var analysis. (f) As in panel (d) but the first guess is from the RAP-predicted wind field for the refined AR-Var analysis. The white letters “A” in panel (a) mark the main aliased-velocity areas. The white letters B1 and B2 in panel (c) mark the two flagged (in black) peninsula areas. The white letter B2 in panel (d) marks the remained flagged peninsula area. The large (or small) yellow circle marks a mature (or emerging) tornadic mesocyclone in panels (a) and (f), and a magnified frame is inserted to show the mature tornadic mesocyclone in panel (f). The green contour on the bottom of each panel shows the boundary between Oklahoma and Texas states. The color scale for the radial velocity is plotted on the top of each panel.

far radial ranges (with an increased analysis accuracy as exemplified in Figures 1 and 2 of Xu and Nai [13]). The above two adaptations can improve the seed data coverage, but the improved seed data coverage is still confined by the often limited radial range of the first-guess background radial-velocity field produced by the AR-VAD analysis. This is shown by the example in Figure 2(b), where the seed data produced from the raw data in Figure 2(a) by the reference check with the above two adaptations are still confined within the 36 km radial range, because the AR-VAD analysis fails to produce the first-guess background radial-velocity field for the AR-Var analysis for $r > 36 \text{ km}$ ($z > 2.6 \text{ km}$ on 4.0° tilt). As listed in column 3 of Table 1, these seed data cover 28.04% of the raw data, which is better than the seed data,

coverage (26.36%) produced by the reference check in the AR-VAD-based dealiasing (not shown). Nevertheless, after the continuity check is performed (one-way forward going away from the radar) in the second main step, the new extended AR-Var-based dealiasing can cover most connected raw-data areas without any false dealiasing as shown in Figure 2(c), and this coverage (76.07% as listed in column 5 of Table 1) is higher than that (74.56%) of the AR-VAD-based dealiasing. As we can see in Figure 2(c), there are still flagged (in black or purple) data holes (especially around the two tornadic mesocyclones marked by the large and small yellow circles in Figure 2(a)) and large amounts of flagged data in the two peninsula areas (marked by the white letters B1 and B2 in Figure 2(b)) and beyond the 130 km radial range.

To recover the flagged data holes (especially around tornadic mesocyclones), the original one-way procedure (going forward away from the radar) of the continuity check (see Section 2c of Xu et al. [10]) is upgraded into a two-way procedure (going forward and backward twice on each tilt). In addition, the block-to-point continuity check in Section 2c(1) of Xu et al. [10] and the three-directional continuity check in Section 2c(2) of Xu et al. [10] are modified and augmented with new step as described below.

(i) The threshold value for the modified block-to-point continuity check is tightened from $v_N/2$ to $v_N/3$ in (6) of Xu et al. [10], and flagged data that fail to pass this tightened threshold condition are marked as special data points (to be used and treated in later steps). The Nyquist folding number correction and related check in (7) of Xu et al. [10] are not performed here but delayed to the new fifth step described below. This can avoid possible false dealiasing in the vicinity of a mesocyclone.

(ii) If at least 40 seed data cannot be found in the $40\Delta r \times 11^\circ$ box, and thus, the modified block-to-point continuity check in (i) cannot perform, then the three-directional continuity check is modified and performed with v_r^{o2} in (8)–(10) of Xu et al. [10] replaced by

$$v_r^{oN} = v_r^o + 2Nv_N, \quad (1)$$

where v_r^o is the raw radial-velocity observation, $N = \text{Int}[(v_r^{\text{seed}} - v_r^o)/(2v_N)]$ is the Nyquist folding number estimated for v_r^o from v_r^{seed} , $\text{Int}[\cdot]$ represents the nearest integer of \cdot , and v_r^{seed} is given by the seed datum (i.e., $v_r^{\text{seed}1}$ in (8), $v_r^{\text{seed}2}$ in (9), or $v_r^{\text{seed}3}$ in (10) of Xu et al. [10]). The flowchart of the above steps (i)–(ii) is shown in Figure 3(a).

(iii) A new third step is added to recheck each seed datum. The rechecking procedure goes forward (away from the radar) along each beam through the entire tilt (clockwise). The seed datum being checked will degrade to a special datum if the following two conditions are both met. (a) There is one or more special data points within 10 gates (along the same beam from the seed datum) in the direction opposite to the searching direction. (b) There is one or more special data point within 15 gates in the searching direction or there are more than 5 flagged data points within 15 gates and more than 20 flagged data points within 40 gates in the searching direction. This third step ensures each accepted seed datum to be not two-sided by special data along the current beam (where the special datum is located) and not neighbored by too many flagged data in the searching direction.

(iv) A new fourth step is added to perform a line-to-point continuity check at each special datum point. The procedure goes forward along each beam through the entire tilt (clockwise). The special datum being checked or its correction v_r^{oN} will be upgraded to a seed datum if the following two conditions are both satisfied. (a) v_r^{oN} is within $\pm v_N/2$ of the averaged value of the seed data (if any) available from the three data points that are on the current beam and its two neighboring beams along either of the two adjacent range circles (to the current range circle where the special datum is located). (b) v_r^{oN} is within $\pm v_N/2$ of the averaged value of

the seed data (if any) available from the three data points that are on the current range circle and its two neighboring range circles along either of the two adjacent beams (to the current beam). Here, v_r^{oN} is computed by (1) with N estimated for v_r^o from v_r^{seed} , v_r^o is the special datum being checked, and v_r^{seed} is given by the averaged value of the seed data.

(v) A new fifth step is added to perform a point-to-point continuity check at each special datum point. The procedure goes in the same way as in the above step (iv). The special datum being checked or its correction v_r^{oN} will upgrade to a seed datum if the following two conditions are both satisfied. (a) v_r^{oN} is within $\pm v_N/2$ of the nearest seed datum that is within ± 5 range gates from the special datum point along the same beam. (b) v_r^{oN} is within $\pm v_N/2$ of the nearest seed datum that is within ± 2 beam positions from the special datum point along the same range circle. Here, v_r^{oN} is computed by (1) with N estimated for v_r^o from v_r^{seed} , v_r^o is the special datum being checked, and v_r^{seed} is given by the seed datum. The above steps (iv)–(v) are performed repetitively through the entire tilt for three times. This completes the first one-way forward procedure, as shown by the flowchart in Figure 3(b).

(vi) After the above first one-way forward procedure, the procedure is restarted with the five steps (i)–(vi) performed one way backward (toward the radar). After this, the above two-way forward-backward procedure is repeated in the same way as described above except that the line-to-point continuity check and point-to-point continuity check in (iv)–(v) are now applied only to those special data that satisfy the four cyclonic-rotation conditions (listed in the appendix) in an 11×41 window (11 beams and 41 range gates) centered at the special datum point being checked.

The effectiveness of the above modifications (i)–(vi) is exemplified by the result in Figure 2(d), where the data holes that are previously identified by the first one-way forward procedure as special data points (plotted in purple in Figure 2(c)) are all recovered around the mature tornadic mesocyclone and the previously flagged data in the first peninsula area (marked by the white letters B1 in Figure 2(c)) are also recovered, although the flagged data and special data in the second peninsula area (marked by the white letters B2 in Figure 2(d)) and the flagged data beyond the 130 km radial range are not recovered yet. The final dealiasied data (including the deflagged nonaliasied data) cover 83.66% of the raw data as listed in column 6 of Table 1. The CPU time required by this new extended AR-Var-based dealiasing is 99 s for processing the entire volume of radial-velocity data for the case in Figure 2, which is nearly 2.5 times of that (40 s) required by the AR-VAD-based dealiasing but is still fast enough for real-time applications. The increased CPU time is caused by the AR-Var analysis that can be but is not yet parallelized in the current code and thus is applied sequentially to each selected range circle on each tilt. The increased part of the CPU time can be reduced by about 10 (or even 10^2) times if the AR-Var analysis is parallelized for all different tilts (or all selected range circles).

In addition to the above improvements, a model-predicted wind field can be interpolated onto each selected

TABLE 1: Numbers of raw data (column 2), seed data (column 3), one-way special data (column 4), one-way dealiased data (column 5), final dealiased data (column 6), and falsely dealiased data (column 7) are produced by the three methods (column 1) for the case in Figure 2. The percentage with respect to the total number of raw data is also listed below the number in each type of data in the same box. “Seed data” are dealiased data produced by the reference check alone. “One-way special data” are special data identified by the first one-way forward procedure (see the flowchart in Figure 3(b)) in the continuity check. “One-way dealiased” data are dealiased data produced by the first one-way forward procedure in the continuity check. All “dealiased” data include de-flagged nonaliased data.

Method	Raw data	Seed data	One-way special data	One-way dealiased	Final dealiased	Falsely dealiased
AR-VAD-based	120971	31882 26.36%	N/A	90191 74.56%	N/A	0
New AR-Var-based	120971	33922 28.04%	10593 8.75%	92028 76.07%	101200 83.66%	0
Model + AR-Var-based	120971	104436 86.33%	6579 5.44%	113712 94.00%	119964 99.17%	0

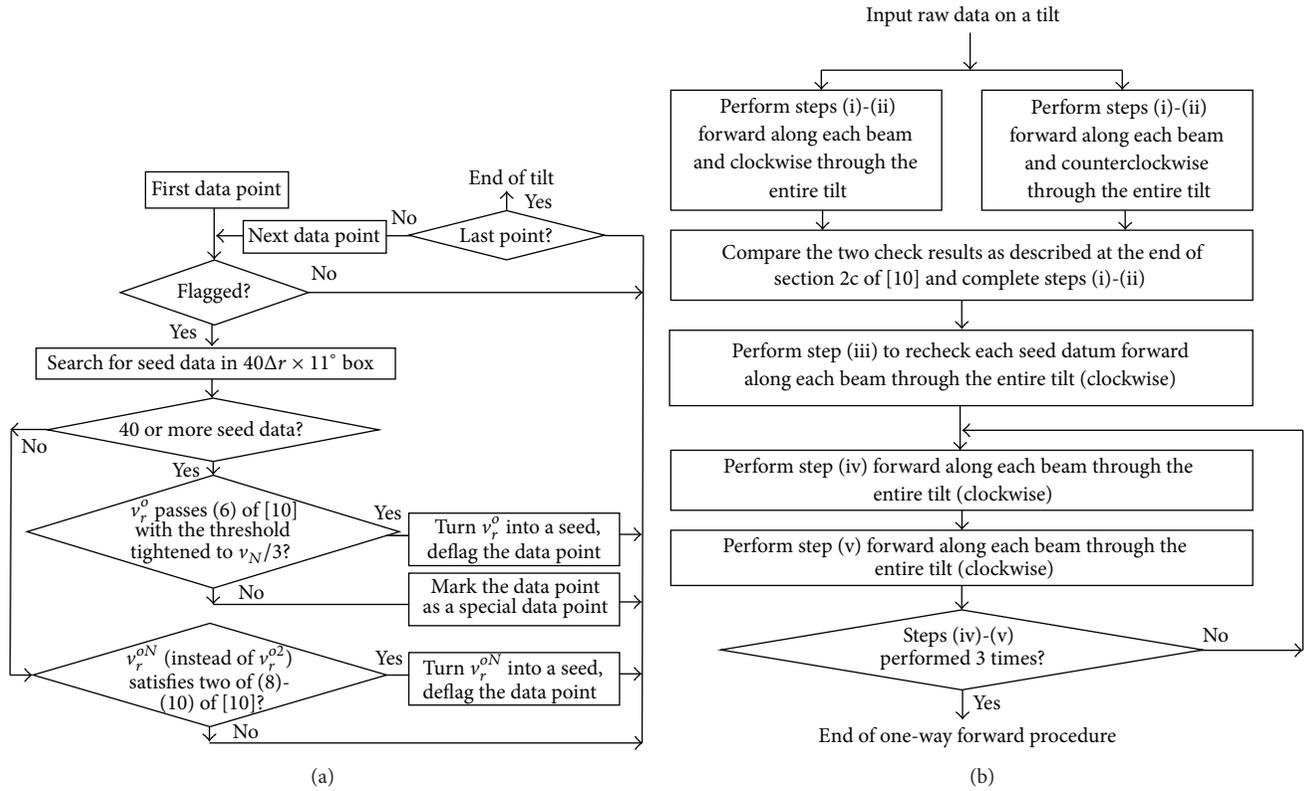


FIGURE 3: (a) Flowchart for the modified block-to-point continuity check and modified three-directional continuity check in steps (i)-(ii). (b) Flowchart for the first one-way forward procedure from step (i) to step (v) in the continuity check.

range circle and used as the first guess for the refined AR-Var analysis to cover the isolated data areas far away from the radar if the predicted field can fit the data closely and allow most of the data to pass the threshold in (14) of Xu and Nai [13] for the refined AR-Var analysis. This approach was recently developed and tested with the wind fields predicted hourly by the NCEP operational Rapid Refresh (RAP) model. The preliminary results are very encouraging. The effectiveness of this approach, called model+AR-Var-based dealiasing, is exemplified by the improved results in Figures 2(e) and 2(f). In this case, the seed data coverage is increased from 28.04% to 86.33% (see column 3 of Table 1), the special data coverage is reduced from 8.75% to 5.44%

(see column 4 of Table 1), the dealiased data coverage is increased from 76.07% to 94.00% (see column 5 of Table 1) after the first one-way forward procedure, and the final dealiased data coverage is increased from 83.66% to 99.17% (see column 6 of Table 1). The CPU time required by this model+AR-Var-based dealiasing is 271 s for processing the entire volume of radial-velocity data for the case in Figure 2, which is 6.8 times of that (40 s) required by the AR-VAD-based dealiasing and is marginal for real-time applications. Again, the increased part of the CPU time can be reduced by about 10 (or even 10^2) times if the AR-Var analysis used in this model+AR-Var-based dealiasing is parallelized for all different tilts (or all selected range circles).

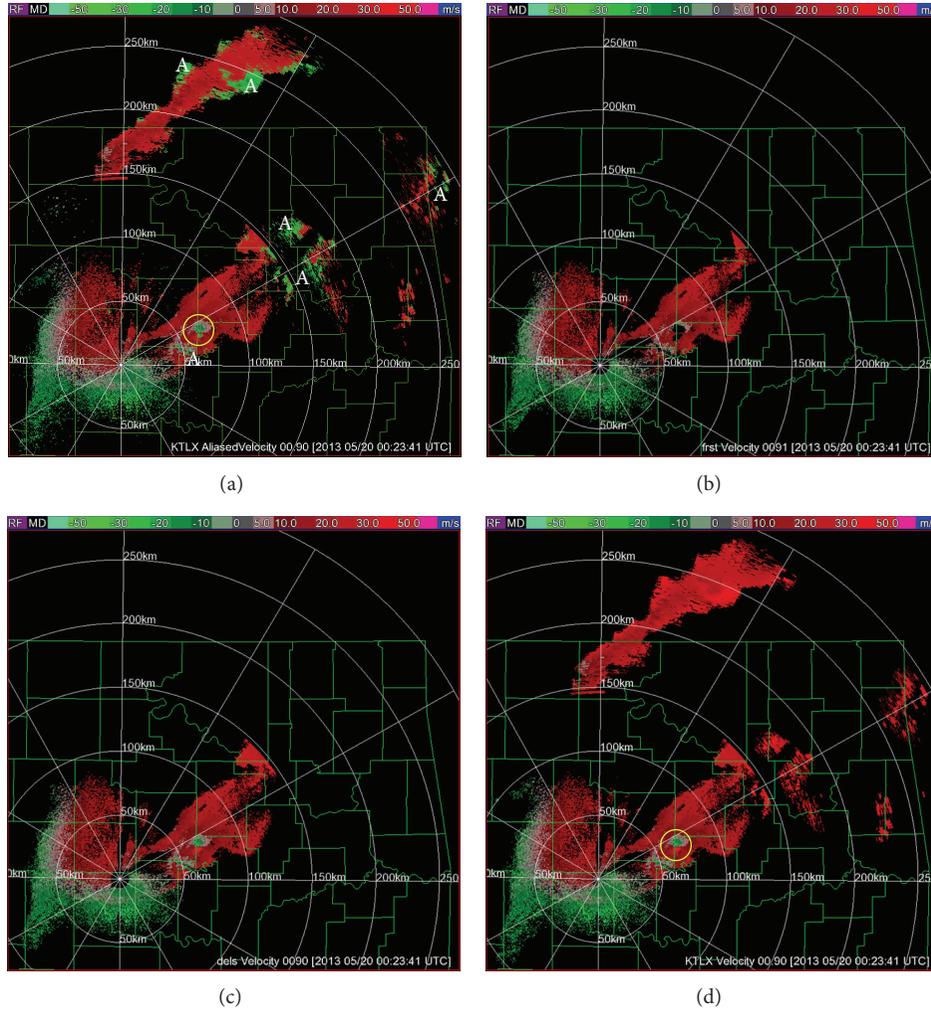


FIGURE 4: Raw radial-velocity image scanned by the KTLX radar using VCP12 with $v_N = 26.1 \text{ m s}^{-1}$ at 0.9° tilt for the Shawnee, Oklahoma tornadic storm at 00:23:41 UTC on May 20, 2013 (local time 7:23:41 pm, on May 19, 2013). (b) Dealiasied radial-velocity image produced by the first one-way forward procedure in the continuity check. (c) Final dealiasied radial-velocity image produced by the new extended AR-Var-based method. (d) as in panel (c) but the first guess is from the RAP-predicted wind field for the refined AR-Var analysis. The white letters “A” in panel (a) mark the main aliased-velocity areas. The yellow circle in panels (a) and (d) encircles the mesocyclone and its produced EF4 tornado that struck Shawnee, Oklahoma, in the evening (local time around 6:45 pm) on May 19, 2013.

The improved dealiasing techniques presented in this paper have been running continuously with real-time data from five operational WSR-88D radars (KTLX, KSRX, KINX, KVNK, and KFDR) and one TDWR radar (TOKC). Their improved performances have been monitored and verified under various weather conditions, especially during severe storms. Examples of their real-time results are shown in Figures 4 and 5 for two recent tornadic storms scanned by the KTLX radar. The example in Figure 4 is the tornadic storm that produced a strong mesocyclone (marked by the yellow circle in Figure 4(d)) and an EF4 tornado that tracked from Norman to Shawnee, Oklahoma, in the evening (local time between 6:00 pm and 7:00 pm) on May 19, 2013. The example in Figure 5 is the tornadic storm that produced a strong mesocyclone (marked by the yellow circle in Figure 5(d)) and an EF5 tornado that struck the cities of Newcastle and Moore,

Oklahoma, in the afternoon (local time between 2:45 pm and 3:35 pm) on May 20, 2013.

4. Conclusions

In this paper, further improvements are made in radar velocity dealiasing atop of the recently published VAD-based dealiasing (Xu et al. [10]) and AR-Var-based dealiasing (Xu and Nai [11]). The improved dealiasing techniques have been running continuously with real-time data from six operational radars (KTLX, KSRX, KINX, KVNK, KFDR, and TOKC). According to the real-time results so far monitored (as sampled in Figures 2, 4–6), the techniques are quite effective in improving the dealiasied data coverage without false dealiasing. The improvement achieved by using the predicted wind field from the operational RAP model as the first

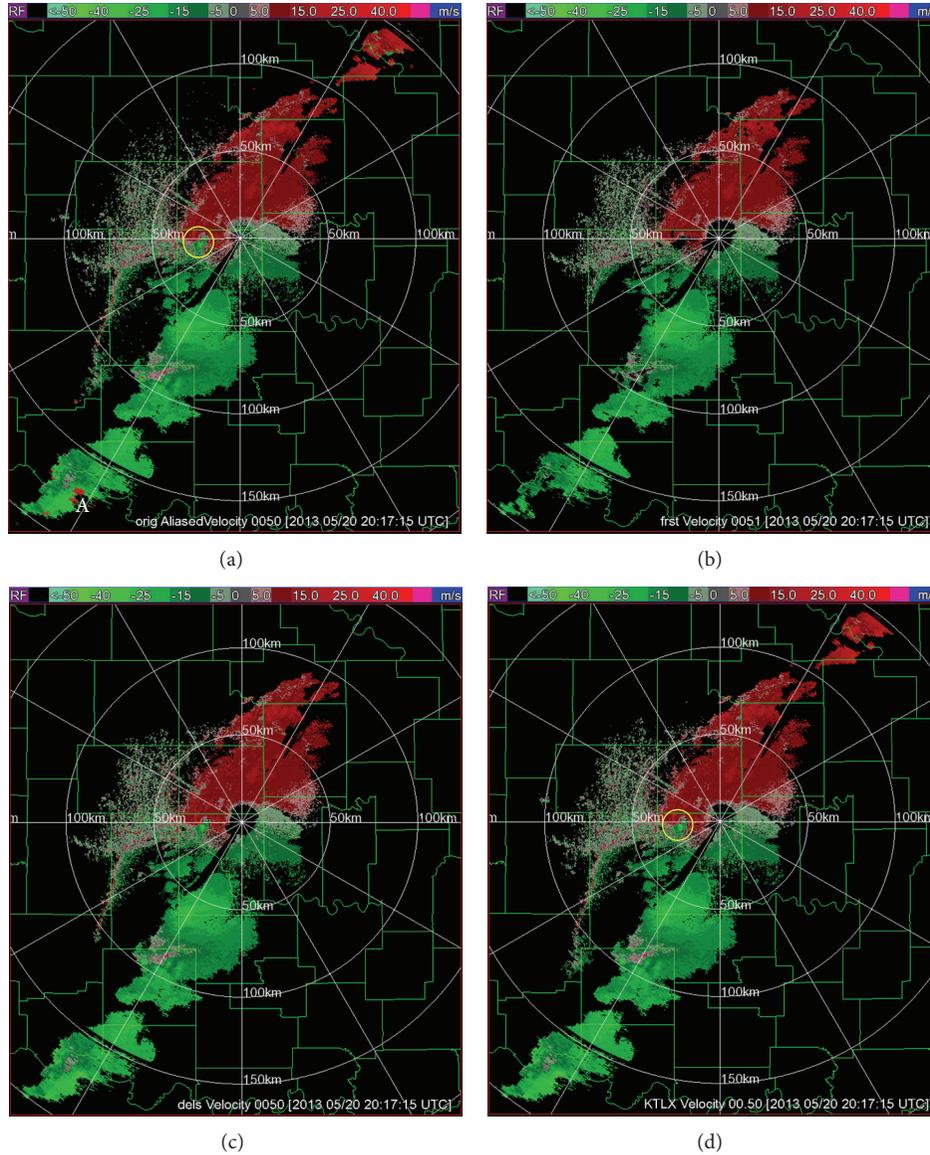


FIGURE 5: As in Figure 4 but for the Moore, Oklahoma, tornadic storm scanned by the KTLX radar using VCP12 with $v_N = 26.1 \text{ m s}^{-1}$ on 0.5° tilt at 20:17:15 UTC on May 20, 2013 (local time 03:17:15 pm on May 20, 2013). The yellow circle in panels (a) and (d) encircles the mesocyclone and its produced EF5 tornado that struck Moore, Oklahoma, in the afternoon (local time around 3:30 pm) on May 20, 2013.

guess for the refined AR-Var analysis (Xu and Nai [13]) to produce a radial-velocity reference field over broad radial ranges for the reference check in the first step is especially encouraging. However, according to our additional tests (not shown in this paper), directly using the RAP-predicted wind field for the reference check without the refined AR-Var analysis can occasionally cause false dealiasing. In particular, directly using the RAP-predicted wind field in the reference check can work well (with no false dealiasing) for almost all the 222 volumes of raw radial-velocity data collected by the KTLX radar for the three cases considered in this paper, but it causes false dealiasing on 10 tilts in two volumes at 23:14:34 UTC on May 19, 2013. Although the false dealiasing rate is extremely low (0.9% in number of volumes and 0.32% in number of tilts) for these three cases due to the stringent

threshold condition used for the reference check (see (4)-(5) of Xu et al. [10]), the false dealiasing area (less than 40 seed data points or 70 final dealiased data points on any of the 10 tilts) is within the mesocyclone (not shown). Thus, the refined AR-Var analysis is still necessary for the reference check to be free of false dealiasing when a model-predicted wind field is used. The current refined AR-Var analysis uses stringent threshold to filter radial-velocity data (see (14) of Xu and Nai [13]) and requires at least 20 valid radial-velocity data on each selected range circle, so it tends to reject scattered data in small isolated areas far away from the radar. This is a minor limitation for the improved reference check in the first step. It is possible to reduce this limitation by performing the AR-Var analysis not over the entire range circle but locally and adaptively around each isolated data area. The improved

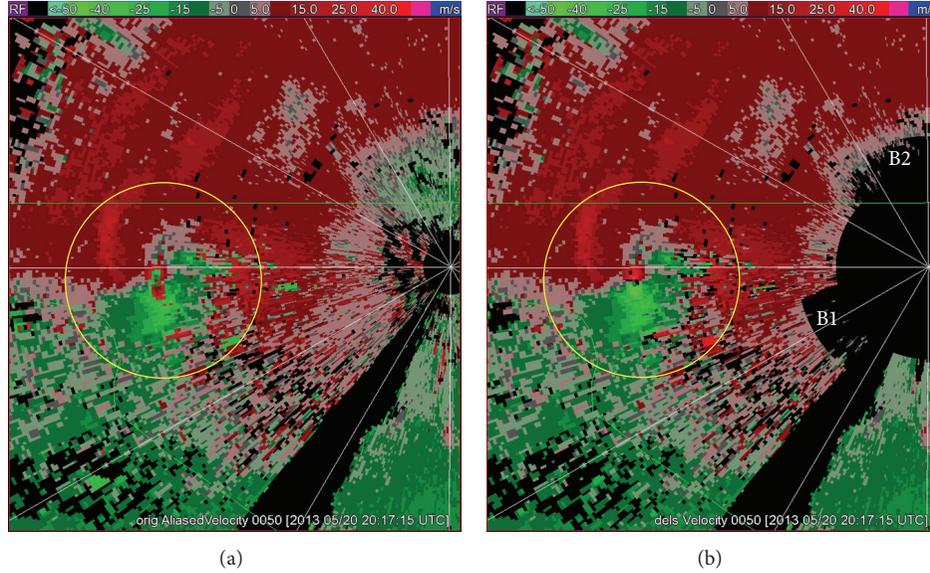


FIGURE 6: Enlarged displays around the mesocyclone of (a) the raw radial-velocity image in Figure 5(a) and (b) the dealiased radial-velocity image in Figure 5(c). The yellow circle in each panel encircles the mesocyclone. The white letters B1 and B2 in panel (b) mark the two flagged (in black) areas immediately outside the 5 km range, while the radial velocities within the 5 km radial range on 0.5° tilt are not used and thus flagged (in black) to avoid possible ground clutter contaminations.

continuity check in the second step also has a limitation in its dealiased data coverage. In particular, it does not always cover all the data points in every small but critical area of tornadic mesocyclone. This limitation is seen clearly when the raw and dealiased radial-velocity images in Figures 5(a) and 5(c) are enlarged around the mesocyclone in Figures 6(a) and 6(b), respectively. Continued research efforts are being made to reduce the above limitations.

Appendix

Cyclonic-Rotation Conditions

- $v_{r\max} - v_{r\min} > v_N$ and $\varphi_{\max} > \varphi_{\min}$, where $v_{r\max}$ (or $v_{r\min}$) is the maximum (or minimum) radial velocity among all the seed data in the 11×41 window and φ_{\max} (or φ_{\min}) is the azimuthal angle of the $v_{r\max}$ (or $v_{r\min}$) data point.
- $v_{r+} - v_{r-} > 0.3v_N$, where v_{r+} (or v_{r-}) is the averaged value of those seed data that are larger (or smaller) than $v_{rm} \equiv (v_{r\max} + v_{r\min})/2$ in the 11×41 window.
- $\varphi_+ > \varphi_-$, where φ_+ (or φ_-) is the averaged azimuthal angle for those seed data that are larger (or smaller) than v_{rm} in the 11×41 window.
- $N_{l-} > 2N_{l+}$ and $N_{r+} > 2N_{r-}$, where N_{l+} (or N_{l-}) is the number of seed data that are larger (or smaller) than v_{rm} in the left window, N_{r+} (or N_{r-}) is the number of seed data that are larger (or smaller) than v_{rm} in the right window, and the 11×41 window is divided into two parts, the left window and the right window, by the radial line at $\varphi = \varphi_m \equiv (\varphi_{\max} + \varphi_{\min})/2$.

If the above four cyclonic-rotation conditions are satisfied for the special datum point being checked, then the 11×41 window centered at the special datum point may very likely cover the vortex center of a mesocyclone. In this case, as by-products, the vortex center location can be estimated by $(r, \varphi) = (r_m, \varphi_m)$ on the current tilt in the radar local coordinate system, and the radius of maximum tangential velocity for the vortex flow can be estimated by $\rho_M \approx (|r_{\max} - r_{\min}|^2 + r_m^2 |\varphi_{\max} - \varphi_{\min}|^2)^{1/2} / 2$, where $r_m \equiv (r_{\max} + r_{\min})/2$ and r_{\max} (or r_{\min}) is the radial range of the $v_{r\max}$ (or $v_{r\min}$) data point.

Acknowledgments

The authors are thankful to Dr. David Stensrud and anonymous reviewers for their comments and suggestions that improved the representation of the results. This research was supported by the NCEP-NSSL radar data QC project and the ONR Grant N000141010778 to the University of Oklahoma (OU). Funding was also provided by DOC/NOAA/OAR under NOAA-OU Cooperative Agreement NAI7RJ1227.

References

- M. D. Eilts and S. D. Smith, "Efficient dealiasing of Doppler velocities using local environment constraints," *Journal of Atmospheric and Oceanic Technology*, vol. 7, no. 1, pp. 118–128, 1990.
- G. J. Stumpf, A. Witt, E. D. Mitchell et al., "The national severe storms laboratory mesocyclone detection algorithm for the WSR-88D," *Weather and Forecasting*, vol. 13, no. 2, pp. 304–326, 1998.

- [3] T. M. Smith and K. L. Elmore, "The use of radial velocity derivative to diagnose rotation and divergence," in *Proceedings of the 11th Conference on Aviation, Range, and Aerospace Meteorology*, pp. 855–861, American Meteorological Society, Hyannis, Mass, USA, October 2004.
- [4] Z. Jing and G. Wiener, "Two-dimensional dealiasing of Doppler velocities," *Journal of Atmospheric and Oceanic Technology*, vol. 10, pp. 798–808, 1993.
- [5] D. W. Burgess and T. D. Crum, "Observed failure modes of the WSR-88d velocity dealiasing algorithm during severe weather outbreaks," in *Proceedings of the 34rd Conference on Radar Meteorology*, American Meteorological Society, Williamsburg, Va, USA, 2009.
- [6] A. Witt, R. A. Brown, and Z. Jing, "Performance of a new velocity dealiasing algorithm for the WSR-88D," in *Proceedings of the 34rd Conference on Radar Meteorology*, American Meteorological Society, Williamsburg, Va, USA, 2009.
- [7] J. Gong, L. Wang, and Q. Xu, "A three-step dealiasing method for Doppler velocity data quality control," *Journal of Atmospheric and Oceanic Technology*, vol. 20, pp. 1738–1748, 2003.
- [8] J. Gao, K. K. Droegemeier, J. Gong, and Q. Xu, "A method for retrieving mean horizontal wind profiles from single-Doppler radar observations contaminated by aliasing," *Monthly Weather Review*, vol. 132, pp. 1399–1409, 2004.
- [9] J. Zhang and S. Wang, "An automated 2D multipass Doppler radar velocity dealiasing scheme," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 9, pp. 1239–1248, 2006.
- [10] Q. Xu, K. Nai, L. Wei, P. Zhang, S. Liu, and D. Parrish, "A VAD-based dealiasing method for radar velocity data quality control," *Journal of Atmospheric and Oceanic Technology*, vol. 28, no. 1, pp. 50–62, 2011.
- [11] Q. Xu and K. Nai, "An adaptive dealiasing method based on variational analysis for radar radial velocities scanned with small Nyquist velocities," *Journal of Atmospheric and Oceanic Technology*, vol. 29, pp. 1723–1729, 2012.
- [12] Q. Xu, K. Nai, and L. Wei, "Fitting VAD winds to aliased doppler radial-velocity observations: a global minimization problem in the presence of multiple local minima," *Quarterly Journal of the Royal Meteorological Society*, vol. 136, no. 647, pp. 451–461, 2010.
- [13] Q. Xu and K. Nai, "A two-step variational method for analyzing severely aliased radar velocity observations with small Nyquist velocities," *Quarterly Journal of the Royal Meteorological Society*, 2013.
- [14] S. Liu, G. DiMego, K. V. Kumar et al., "WSR-88D radar data processing at NCEP," in *Proceedings of the 34rd Conference on Radar Meteorology*, American Meteorological Society, Williamsburg, Va, USA, 2009.
- [15] Q. Xu, K. Nai, L. Wei, and Q. Zhao, "An unconventional approach for assimilating aliased radar radial velocities," *Tellus A*, vol. 61, no. 5, pp. 621–630, 2009.

Research Article

Impact of 3DVAR Data Assimilation on the Prediction of Heavy Rainfall over Southern China

Tuanjie Hou,^{1,2} Fanyou Kong,² Xunlai Chen,^{2,3} and Hengchi Lei¹

¹ *Laboratory of Cloud-Precipitation Physics and Severe Storms, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China*

² *Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK 73072, USA*

³ *Shenzhen Meteorological Bureau, Shenzhen 518040, China*

Correspondence should be addressed to Tuanjie Hou; houtuanjie@gmail.com

Received 26 March 2013; Revised 16 June 2013; Accepted 12 July 2013

Academic Editor: Jidong Gao

Copyright © 2013 Tuanjie Hou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study examines the impact of three-dimensional variational data assimilation (3DVAR) on the prediction of two heavy rainfall events over Southern China by using a real-time storm-scale forecasting system. Initialized from the European Centre for Medium-Range Weather Forecasts (ECMWF) high-resolution data, the forecasting system is characterized by combining the Advanced Research Weather Research and Forecasting (WRF-ARW) model and the Advanced Regional Prediction System (ARPS) 3DVAR package. Observations from Doppler radars, surface Automatic Weather Station (AWS) network, and radiosondes are used in the experiments to evaluate the impact of data assimilation on short-term quantitative precipitation forecast (QPF) skill. Results suggest that extrasurface AWS data assimilation has slight but general positive impact on rainfall location forecasts. Surface AWS data also improve model results of near-surface variables. Radiosonde data assimilation improves the QPF skill by improving rainfall position accuracy and reducing rainfall overprediction. Compared with radar data, the overall impact of additional surface and radiosonde data is smaller and is reflected primarily in reducing rainfall overestimation. The assimilation of all radar, surface, and radiosonde data has a more positive impact on the forecast skill than the assimilation of either type of data only for the two rainfall events.

1. Introduction

Convective storms accompanied with heavy precipitation, hail, and damaging wind occur frequently in summer season in Southern China. To reduce damage from such severe weather, more accurate short-term forecast from convective-scale numerical weather prediction (NWP) models incorporated with robust data assimilation systems have been paid more attention [1–3]. In recent years, several studies have demonstrated that the Advanced Regional Prediction System (ARPS) three-dimensional variational (3DVAR) system is capable of analyzing different data types, by using multiple analysis passes [4–7].

Based on the Advanced Research Weather Research and Forecasting (WRF-ARW) model and the ARPS 3DVAR/Cloud Analysis module, a real-time hourly updated storm-scale forecasting system has been developed collaboratively by the Center for Analysis and Prediction of Storms (CAPS)

in the University of Oklahoma, Shenzhen Meteorological Bureau (SZMB) of China and the Shenzhen Institute of Advanced Technology (SIAT), Chinese Academy of Sciences. The forecasting system, called Hourly Assimilation and Prediction System, or HAPS, has been in daily real-time forecast runs since March 2010. The system was initialized from Global Forecast System (GFS) data and characterized by assimilating reflectivity and radial wind from local Weather Surveillance Radar-1998 Doppler (WSR-98D) radars every hour in real-time forecast in the first phase. During the second phase of upgrading HAPS, not only fine-resolution European Centre for Medium-Range Weather Forecasts (ECMWF) analysis but also forecast data have been used for initial condition (IC) and lateral boundary condition (LBC), and assimilation of surface Automatic Weather Station (AWS) and radiosonde observations were also examined [8].

One way of improving precipitation forecast is thought to use as much available data as possible. Better mesoscale analyses can be provided from frequent intermittent assimilation of surface data [9] or combination of surface and radiosonde data within and above the model planetary boundary layer [10]. Kain et al. [11] demonstrated that assimilation of radar data and other conventional observations had a positive impact on convection-scale model forecasts. Ha et al. [12] found that the assimilation of both Doppler radar and surface data had a more positive impact on quantitative precipitation forecast (QPF) skill than the assimilation of radar data or surface data only. Considering the consistent positive impact of radar data assimilation shown in recent studies [13–16], emphasis is placed on evaluating the effect of surface AWS and radiosonde data assimilation. AWS data have the advantage of better temporal and spatial resolutions, while radiosonde data can provide multilayer meteorological information, both of which should be used effectively in NWP models. As the purpose of deploying HAPS forecast system is to improve convective-scale precipitation forecasts, experiments in the study are focused on the data assimilation results from the convective-scale domain only.

A description of the model and the observation network is given in Section 2. In Section 3, two heavy rainfall events and corresponding experimental design are introduced. In Section 4, impacts of data assimilation on precipitation distribution, near-surface variables, and the QPF skill are evaluated. Finally, summary and conclusions are given in Section 5.

2. Model Description and Observation Network

2.1. Model Description. The HAPS forecast system, using the WRF-ARW (V3.3.1) model as the forecast model and the ARPS 3DVAR package to provide data assimilation, consists of a mesoscale domain with 12 km horizontal grid spacing and a one-way nested convective-scale domain at 4 km grid spacing (Figure 1). The model is initiated from the ECMWF high-resolution data which have a horizontal resolution of 0.25 deg and 20 vertical pressure levels with model top at 10 hPa. The mesoscale domain that covers Southern China region produces 48 h forecasts every 12 h (initiated at 0000 and 1200 UTC) to serve as background and LBC for the convective-scale forecasts. The convective-scale domain produces 12 h forecasts every hour (e.g., initiated at 0000, 0100, 0200 UTC, etc.).

Both model domains have 51 vertical levels. The RRTM longwave radiation [17], Goddard shortwave radiation [18], MYJ planetary boundary layer [19, 20], and Noah land surface model [21] are used in both domains in the real-time system. The Eta microphysics and Kain-Fritsch cumulus parameterization [22] are used in the 12 km domain, while the more complicated New Thompson scheme [23] is used in the 4 km domain.

The 3DVAR method minimizes a cost function that includes the background, observation, and mass conservation constraint terms. The cost function [4] can be

written as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}[H(\mathbf{x}) - \mathbf{y}^o]^T \times \mathbf{R}^{-1} [H(\mathbf{x}) - \mathbf{y}^o] + J_c(\mathbf{x}), \quad (1)$$

where the first term on the right-hand side measures the departure of the analysis vector \mathbf{x} from the background \mathbf{x}^b , weighed by the inverse of the background error covariance matrix \mathbf{B} . The analysis variables include three wind components (u , v , and w), potential temperature, pressure, and water vapor mixing ratio. The second term measures the departure of \mathbf{x} , projected into observation space by H , from the observation vector \mathbf{y}^o . The analyzed radial velocity V_r can be approximated as

$$V_r = \frac{(X - X_o)u + (Y - Y_o)v + (Z - Z_o)w}{r}, \quad (2)$$

where u , v , and w are wind components in Cartesian coordinates (X, Y, Z) , (X_o, Y_o, Z_o) is the radar location, and r is the distance between radar site and data point less than 100 km. The third term $J_c(\mathbf{x})$ represents dynamic or equation constraints. The detailed description of ARPS 3DVAR can be found in [4, 24].

Observation errors can be generated from different sources, such as instrument noise, preprocessing, and representative error. Considering the random noise in radial velocity and the calibration error in reflectivity, Xu et al. [25] discussed the effect of such observational errors. They found that the four-dimensional variational data assimilation technique was not very sensitive to the observational errors, although larger errors in some retrieved fields did occur after increasing errors in radial velocity and reflectivity. Gong et al. [26] discussed partition of observation and background error covariances with the Bessel fitting function and the data from three dense radiosonde observation regions including East Asia, North America, and Europe. For surface AWS data assimilation, Xu et al. [27] conducted experiments to test the impact of the actual elevation difference between observation sites and numerical model surface. Their studies all demonstrated the importance of estimating observation error statistics but can only be classified as sensitivity experiments. To represent the observations accurately from instruments in China, more systematic experiments are needed. Therefore, due to lack of reliable statistics on error information for China instruments, observation errors are assumed to be uncorrelated and specified according to estimated errors for the various observational platforms [7].

A multiscale analysis procedure is included in the ARPS 3DVAR package. It means that multiple analysis passes with each one including different data types and using different filter scales can be used for data assimilation. The filter scale, defined as radius in the package, is guided by the density of observation sites. Considering the distances of those observation sites shown in Figure 1, the horizontal influence radiuses for radar, AWS, and radiosonde data are selected as 20, 50, and 400 km, respectively. The vertical influence radius is four in grid points.

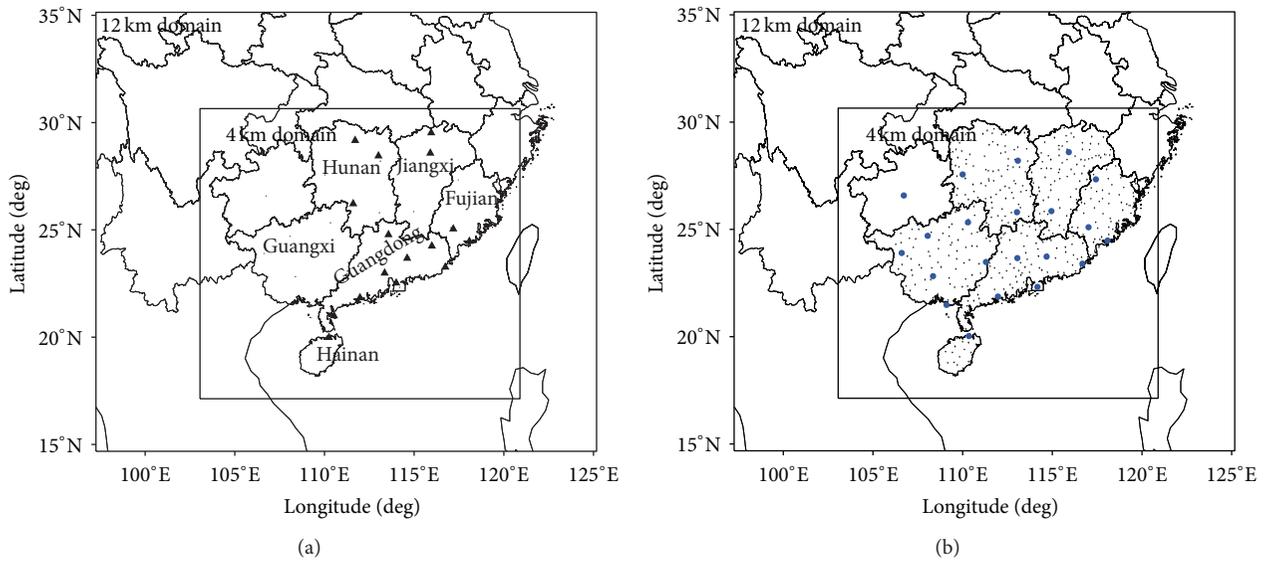


FIGURE 1: Model domains and locations of (a) radar (black triangles) sites, and (b) AWS (small black dots), and radiosonde (big blue dots) sites.

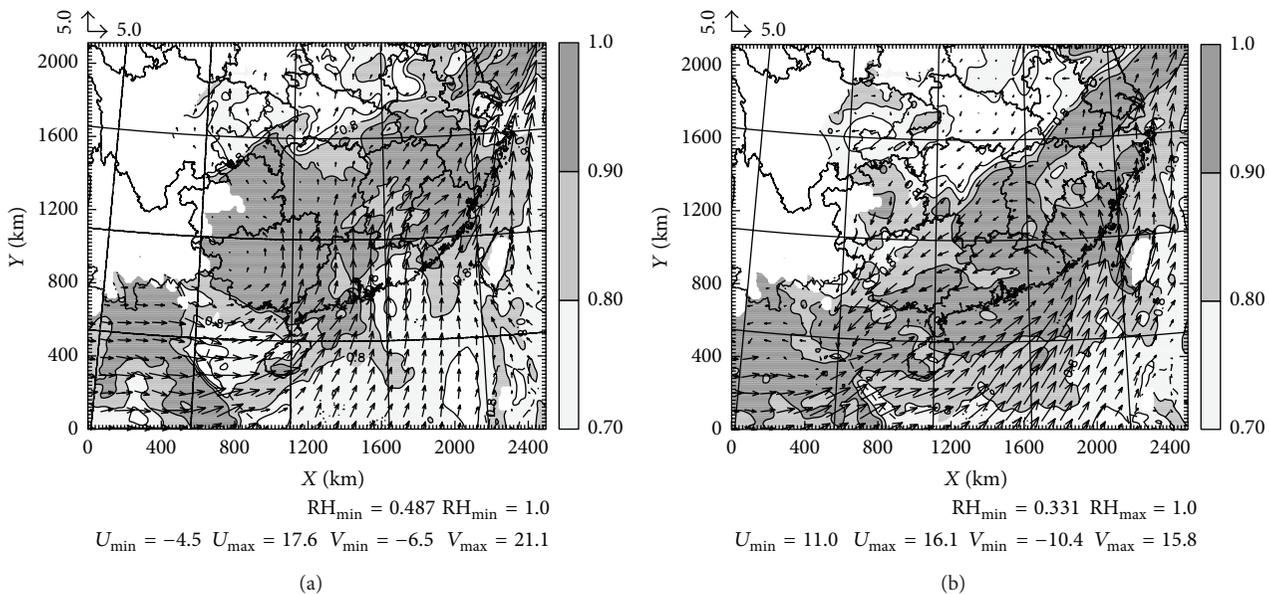


FIGURE 2: 850 hPa wind (arrows, $m s^{-1}$) and relative humidity (shading) at (a) 0000 UTC June 29, 2011 and (b) 0000 UTC July 15, 2011.

2.2. Observation Network. The available observations for this study are from WSR-98D radars, surface AWS, radiosonde, and rain gauge measurements. There are 16 radars in the 4 km domain (triangles in Figure 1(a)), including eight in Guangdong Province and eight in other four provinces. The WSR-98D radars are S-band radars with the similar characteristics to those of the Weather Surveillance Radar-1998 Doppler (WSR-88D) radars in the US operational Doppler radar network [28]. Volumetric radar data from these radars are available every 6 min for 9 elevation angles, with resolutions of radar reflectivity and velocity on the radials of being 1.0 km and 250 m [29]. Raw radar data must be processed through

automated quality control and averaged into data “columns” by using the program *88D2arps*.

The AWS network in China has been used primarily in observing weather conditions but not yet been applied to NWP models [27]. The surface AWS data include near-surface temperature, dewpoint, and wind components with the time interval of every 5 min. Data from 437 surface AWS sites (Figure 1(b)) with a horizontal distance of 3–10 km are used in the study. In addition, data from 21 radiosonde sites (blue dots in Figure 1(b)) in the 4 km domain is also used in data assimilation. These sites are part of the radiosonde network stations distributed throughout China,

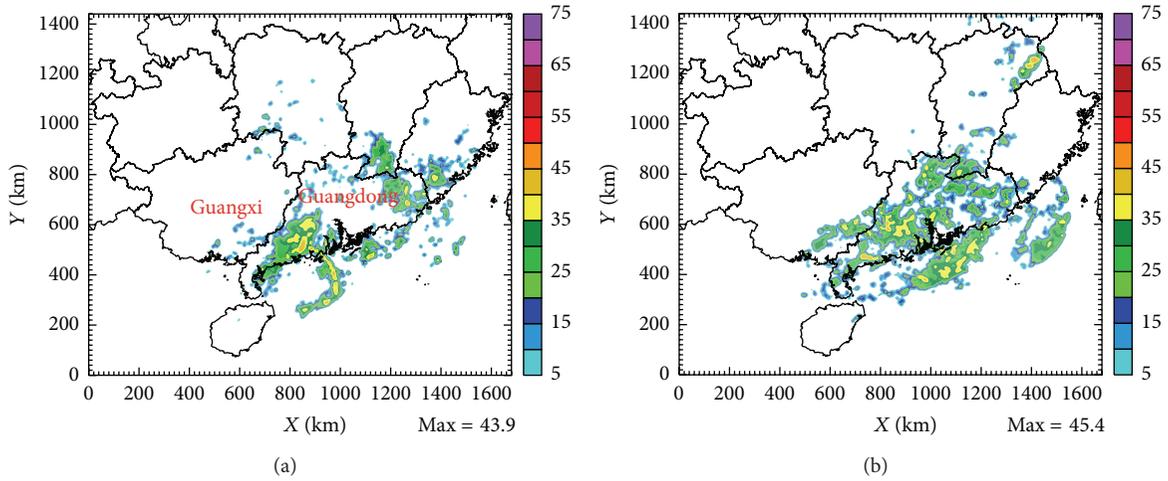


FIGURE 3: Composite reflectivity from observations at (a) 1200 UTC June 29, 2011 and (b) 1200 UTC July 15, 2011.

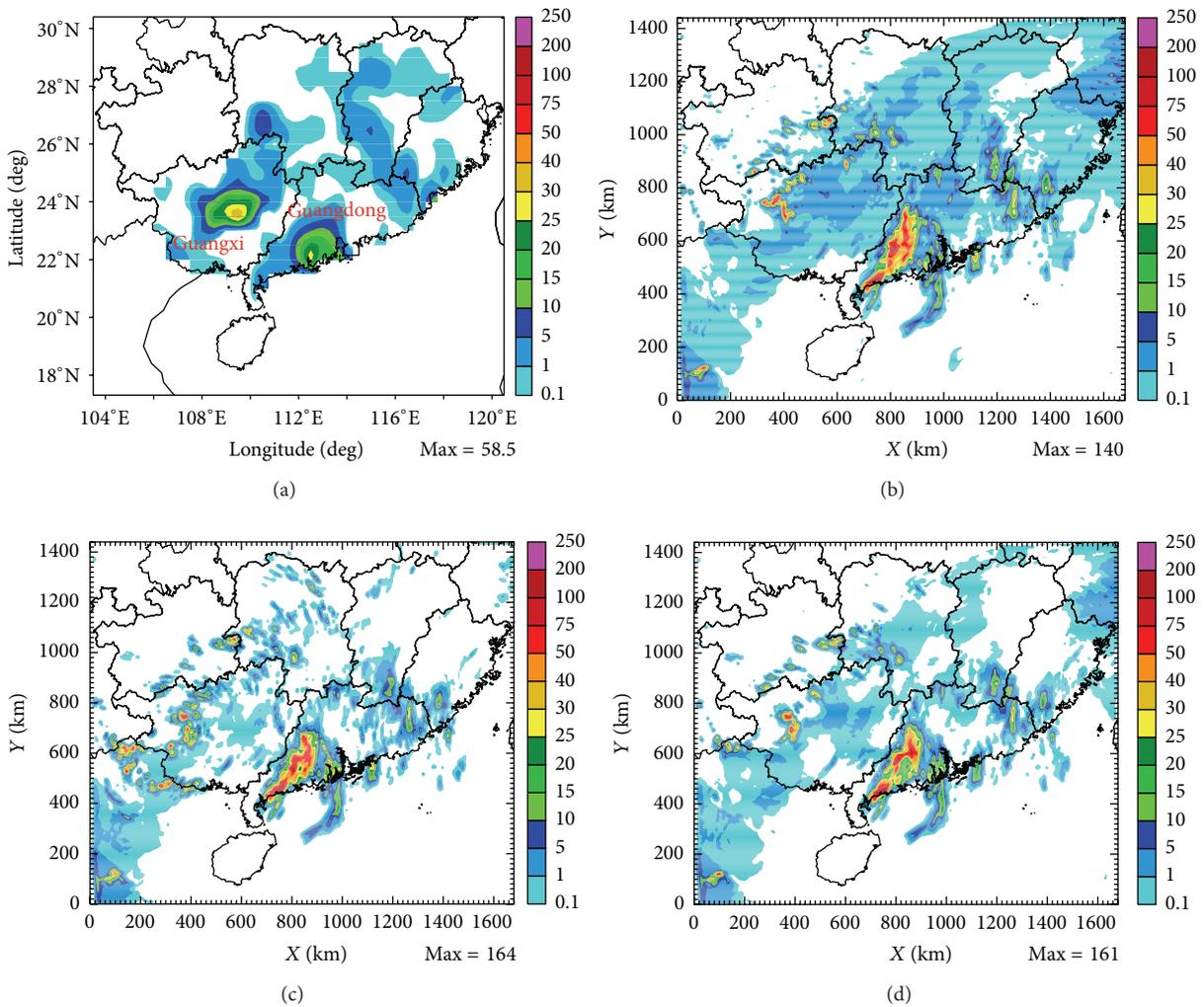


FIGURE 4: 3 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1500 UTC June 29, 2011 for case 20110629.

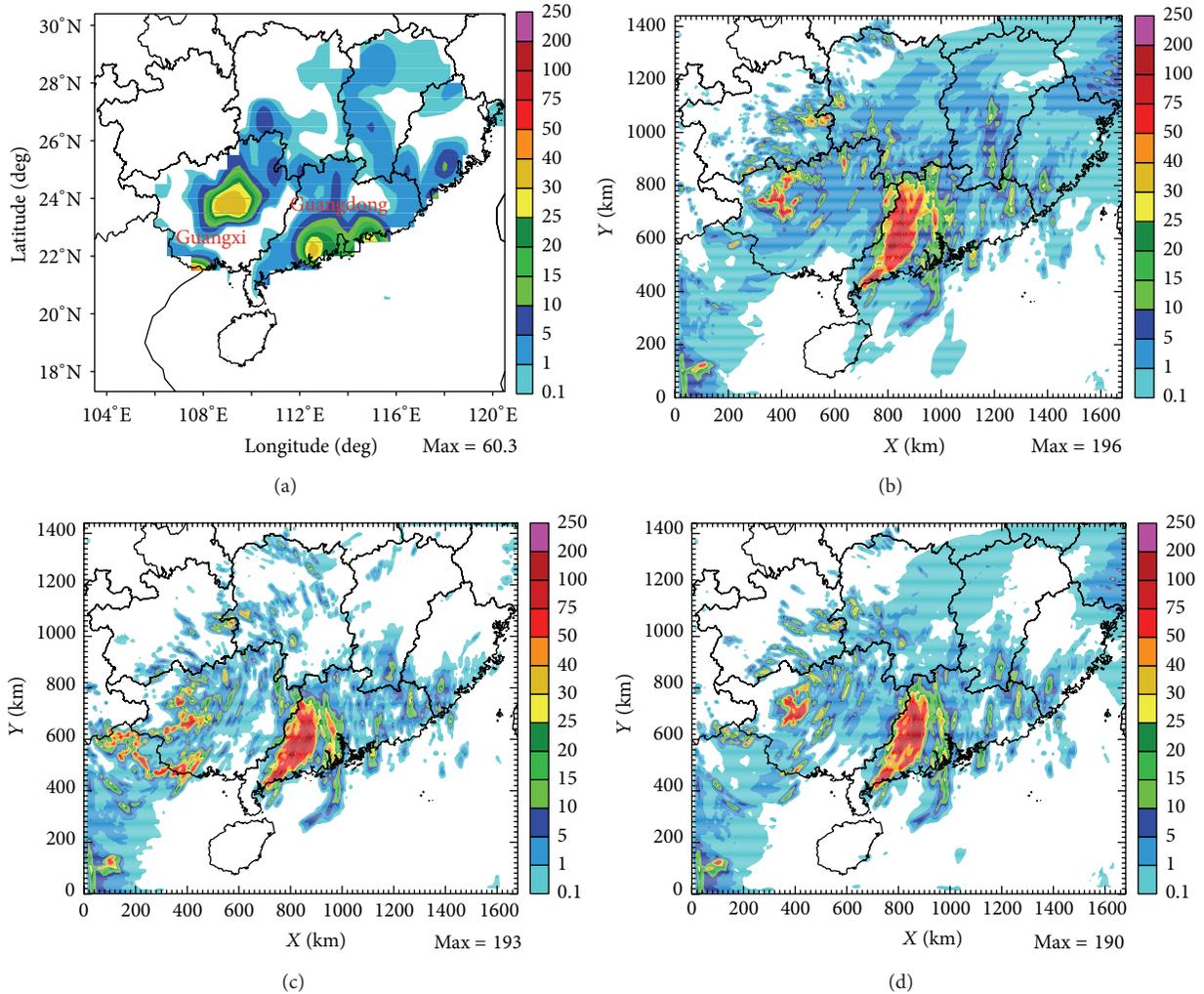


FIGURE 5: 6 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1800 UTC June 29, 2011 for case 20110629.

TABLE 1: Experiment description.

Experiments	Description
CNTL	Radar data assimilation at start time in 4 km domain
RADAWS	CNTL + AWS data assimilation at start time in 4 km domain
RADSND	CNTL + radiosonde data assimilation at start time in 4 km domain
ALLDATA	CNTL + AWS and radiosonde data assimilation at start time in 4 km domain

with improvement in instrumentation model and correction method during the last 50 years [30, 31]. The conventional radiosonde data are available at 0000 and 1200 UTC, including temperature, dewpoint, and wind profiles from 850 to 100 hPa. The total number of rain gauges in the 4 km domain is 1353 (figure not shown). They are used for quantitative precipitation verification in later sections.

3. Overview of Rainfall Events and Experimental Design

3.1. Overview of Rainfall Events. Two heavy rainfall events associated with southwest monsoons are selected in this study, as the southwestern summer monsoons are typical synoptic settings producing rainfall over Southern China in June and July. The first rainfall event (case 20110629) was from June 28, 2011 to June 30, 2011, affecting several provinces in Southern China with heavy rain and severe flooding. The second rainfall event (case 20110715), occurred from July 15, 2011 to July 18, 2011, characterized by nonuniformity and extremely high rainfall rates in localized areas.

Figure 2 shows the 850 hPa wind field and relative humidity at 0000 UTC June 29 and 0000 UTC July 15, 2011 from the data of ECMWF analysis for the two events. At 0000 UTC June 29, 2011 (Figure 2(a)), warm and moist air is transported from coastal areas near Southern China to inland regions through the strong southwest winds, forming a southwest-northeast water vapor band. The high relative humidity of

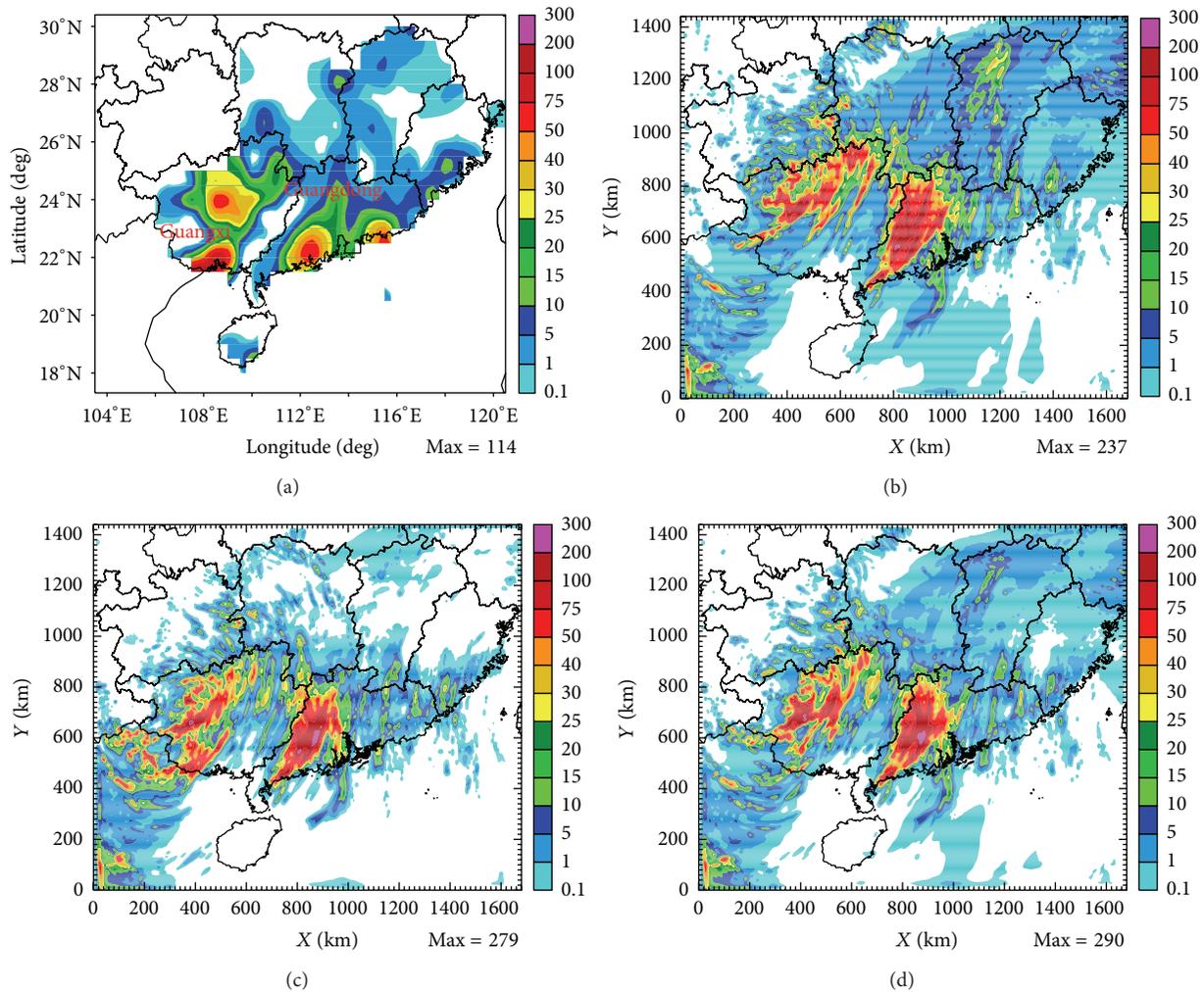


FIGURE 6: 12 h accumulated precipitation from (a) observations, (b) CNTRL, (c) RADSND, and (d) ALLDATA between 1200 UTC 29 June and 0000 UTC June 30, 2011 for case 20110629.

over 80% in several provinces, combined with the gradual deeper trough at higher levels, contributes to the formation of heavy rainfall. Similar to the first case, southwest winds and a water vapor band at 0000 UTC July 15, 2011 appear at coastal areas of Southern China (Figure 2(b)). However, the moist air band in this event is somewhat narrower, due to the wind shear from southwest to southeast over eastern Guangdong and Fujian. Therefore, the low pressure and vapor transport mainly affect the coastal provinces over Southern China.

To have a clear picture of echo structures, Figure 3 shows the observed composite reflectivity at 1200 UTC June 29 and 1200 UTC July 15, 2011, respectively. Composite reflectivity is defined as the maximum reflectivity in the vertical column. The observed reflectivity is derived from the ARPS 3DVAR analysis by using available radar data (with radar station locations shown in Figure 1(a)). At 1200 UTC June 29, 2011 (Figure 3(a)), the main echo regions are at southwest Guangdong and nearby coastal areas, with the maximum value of 43.9 dBZ. In addition, scattered weaker cells can also be seen over surrounding areas of Guangdong.

In comparison, larger reflectivity regions with the similar magnitude cover almost all the Guangdong Province at 1200 UTC July 15, 2011 (Figure 3(b)). Consequently, both events caused severe waterlogging in multiple cities of Guangdong Province.

3.2. Experimental Design. One 12 km and 4 km forecast pair for each rainfall event is conducted. For case 20110629, 12 km forecast is initiated at 0000 UTC 29 June 2011 and runs for 48 h to 0000 UTC 1 July 2011. The 12 h forecast in the 4 km domain is from 1200 UTC 29 June to 0000 UTC 30 June 2011. For the other rainfall event, the 12 km forecast is from 0000 UTC 15 July to 0000 UTC 17 July 2011, and the 4 km run is from 1200 UTC 15 July to 0000 UTC 16 July 2011.

Data assimilation experiments are carried out in the convective-scale domain without cycling. As previous studies have demonstrated that radar data assimilation is useful in improving precipitation pattern and short-term QPF skill, the impact of radar data assimilation will not be examined again in the study. The focus of the present study is to

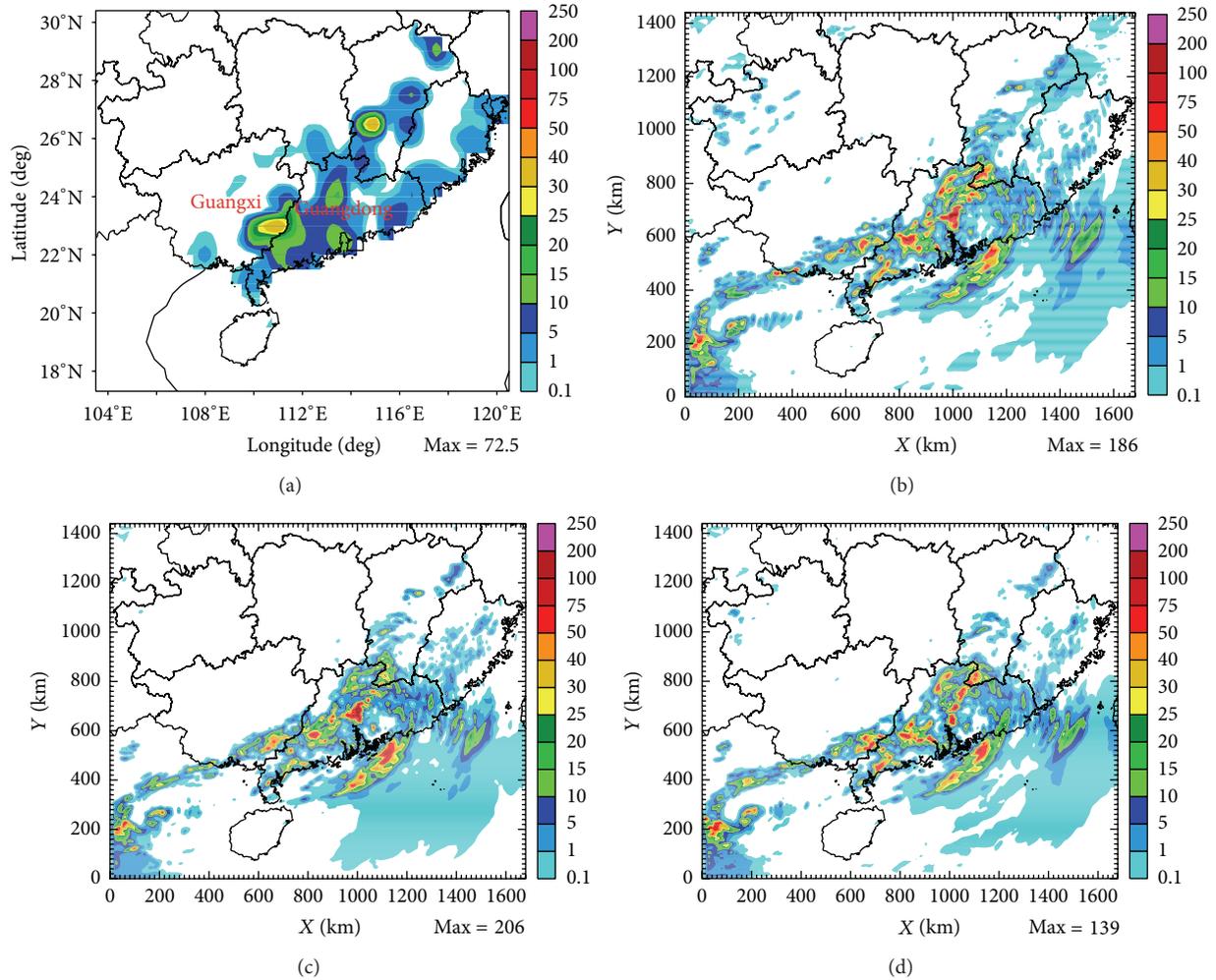


FIGURE 7: 3 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1500 UTC July 15, 2011 for case 20110715.

evaluate the impact of surface AWS and radiosonde data (experiments shown in Table 1). Hence, the basic experiment (or control experiment, CNTL) is a WRF-ARW forecast with radar data assimilation conducted at the initiation time in the 4 km domain. Additional experiments assimilate radar data plus surface AWS and radiosonde data, which are named as RADAWS and RADSND, respectively. To examine the impact of all the combined data, experiments with radar, surface AWS, and radiosonde data assimilated together (ALLDATA) at the initiation time of the 4 km domain are also performed for the two cases.

4. Results

4.1. Impact of Data Assimilation on Precipitation Distribution.

To show the impact of data assimilation on precipitation distribution, 3, 6, and 12 h accumulated precipitations from both observations and three experiments (CNTL, RADSND, and ALLDATA) for the two rainfall cases are provided in this section. Model results are from forecasts in the convective-scale domain. The precipitation distribution from RADAWS

is not shown as little difference is detected compared with that from CNTL. However, the positive impact of AWS data assimilation still can be seen from near-surface variables, which will be presented in Section 4.2.

Figure 4 presents the first 3 h accumulated precipitation between 1200 and 1500 UTC June 29, 2011 from both observations and experiments (CNTL, RADSND, and ALLDATA). According to observations (Figure 4(a)), there are primarily two precipitation regions, one in the central Guangxi Province and the other in the southern Guangdong Province. The observed maximum 3 h rainfall is 58.5 mm. Compared with observations, the main precipitation region in Guangdong generated from experiment CNTL (Figure 4(b)) is a little north and overestimated with the maximum value of up to 140 mm. And the rainfall region in Guangdong becomes much smaller in size and is displaced northwestward, as no radar data are available in Guangxi. Further radiosonde (Figure 4(c)) as well as radiosonde and AWS (Figure 4(d)) data assimilation reduces the size of rainfall overestimation region in Guangdong but still not captures the rainfall region in Guangxi.

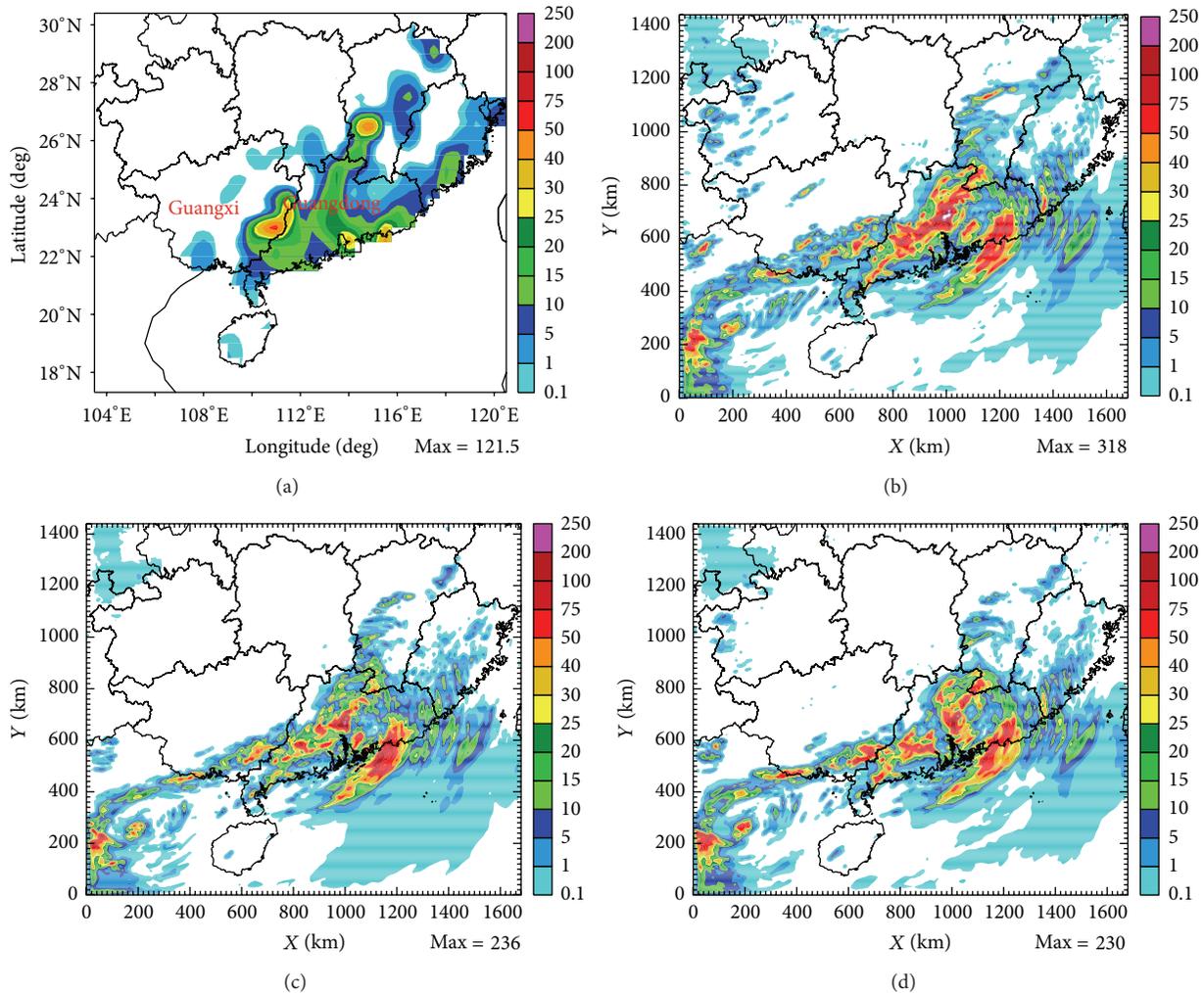


FIGURE 8: 6 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 and 1800 UTC July 15, 2011 for case 20110715.

At 1800 UTC (the time of 6 h forecast), the observed precipitation pattern (Figure 5(a)) is similar to that of 3 h before, with two main precipitation regions. The model forecasts from CNTL, RADSND, and ALL DATA (Figures 5(b)–5(d)) also show one main precipitation region in Guangdong and are different from each other in the small scattered precipitation regions in Guangxi.

Rainfall continues increase during the latter 6 h forecast. The 12 h rainfall records from rain gauge observations (Figure 6(a)) suggest that the two precipitation regions become larger both in size and the central value, characterized by one region with the maximum value of 114 mm in Guangxi and the other rainband with the southwest-northeast direction in Guangdong. Results from radar data assimilation (Figure 6(b)) show two precipitation regions, although the direction of the precipitation region in Guangdong is not closes enough to the coastline. The experiment RADSND (Figure 6(c)) improves the precipitation pattern in Guangxi by increasing the precipitation coverage. Compared with observations and experiment CNTL, experiment

ALLDATA generally improves the forecast by moving the precipitation region in Guangxi southward although still have the problem of rainfall overestimation.

Figures 7–9 present the first 3, 6, and 12 h accumulated precipitations from both observations and three experiments for case 20110715. For the first 3 h between 1200 and 1500 UTC July 15, 2011, Figure 7(a) shows a southwest-northeast rainband over southern coastal provinces, with the maximum value of 72.5 mm. The precipitation patterns from CNTL, RADSND, and ALLDATA (Figures 7(b)–7(d)) are quite close to each other; all are successful in generating the basic rainband. Overall, the experiment ALLDATA is better in terms of the maximum precipitation amount, even though it is higher than that from observations. At 1800 UTC July 15, 2011, the observed 6 h accumulated precipitation areas (Figure 8(a)) are still over Guangdong and its surrounding regions, but the rainfall amount increases with the maximum value of 121.5 mm. At that time, differences among experiments can be seen, as more detailed structures of the precipitation patterns are generated from RADSND

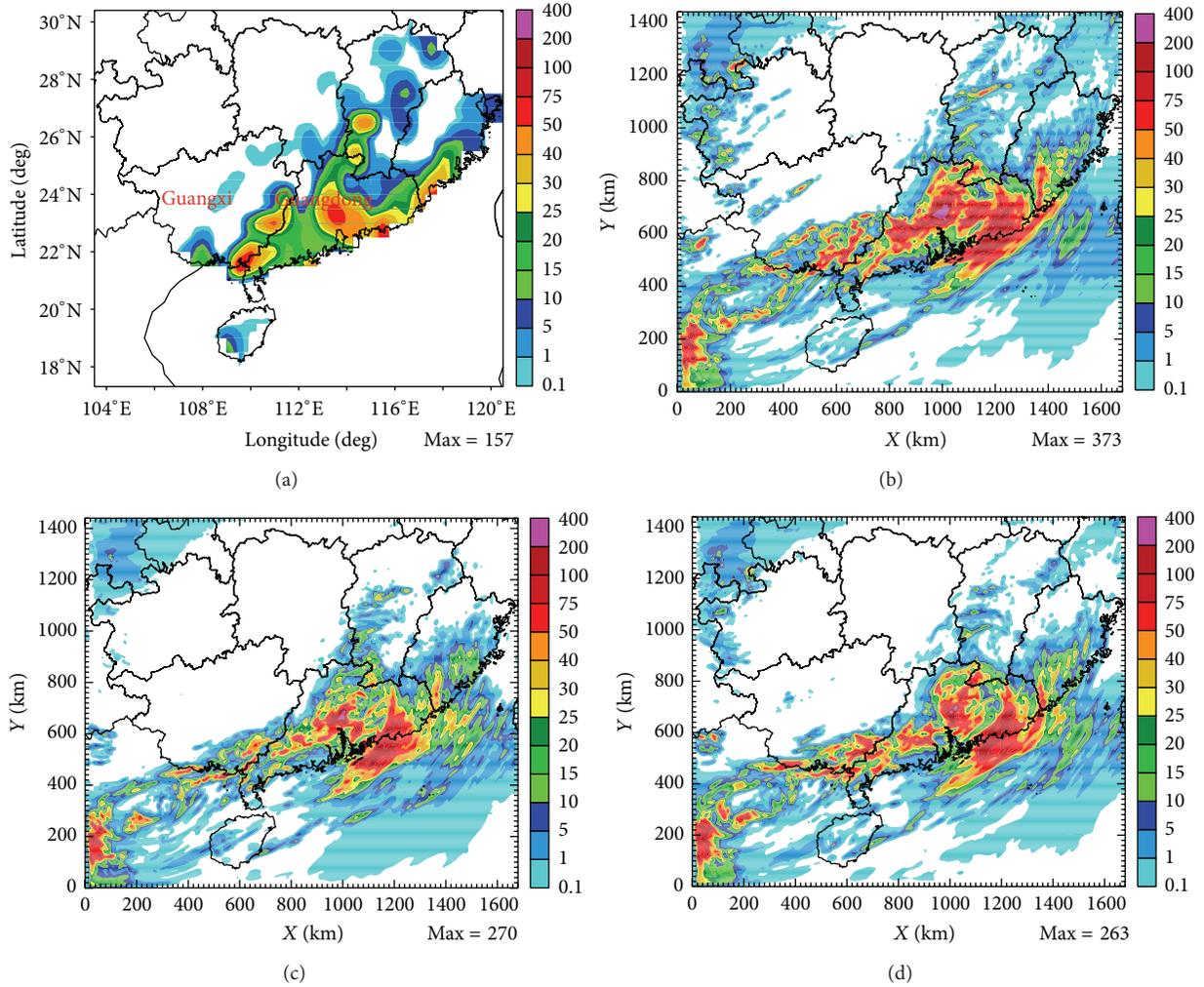


FIGURE 9: 12 h accumulated precipitation from (a) observations, (b) CNTL, (c) RADSND, and (d) ALLDATA between 1200 UTC July 15 and 0000 UTC July 16, 2011 for case 20110715.

and ALLDATA (Figures 8(c) and 8(d)). The discrepancies in precipitation distribution become larger by the time of 12 h forecasts. The 12 h accumulated precipitation distribution (Figure 9(a)) is characterized by separate localized precipitation centers, mainly with one at the southwest edge of Guangdong and the other one in central Guangdong. In comparison with experiment CNTL (Figure 9(b)), RASND (Figure 9(c)) shows positive impact by reducing the size of rainfall overestimation regions and the value of the maximum rainfall amount. Experiment ALLDATA (Figure 9(d)) further improves forecasts by splitting the one rainfall region over central Guangdong into smaller separate precipitation centers, which are closer to observations.

On the whole, forecasts with radar data assimilation are successful in producing the general precipitation patterns but not satisfying in forecasting accurate locations of precipitation centers and rainfall amounts. Addition of radiosonde data assimilation can either reduce the rainfall areal coverage overestimation or the precipitation maxima. Combination of

all the three data types is helpful in producing more detailed precipitation spatial distribution.

4.2. AWS Data Assimilation Impact on Near-Surface Variables.

Surface observations, although they provide data of only one level, have the advantage of better temporal and spatial resolutions with both thermal and wind information. Root mean square errors (RMSEs) for 2 m temperature and relative humidity as well as 10 m wind components are used to evaluate the impact of AWS data assimilation in this section. Interpolation from model grid data to observation locations is performed over the AWS sites.

The RMSE is defined as

$$RMSE = \left(\frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2 \right)^{1/2}, \quad (3)$$

where F_i represents variables, including temperature, relative humidity, and wind components from model results, and O_i is

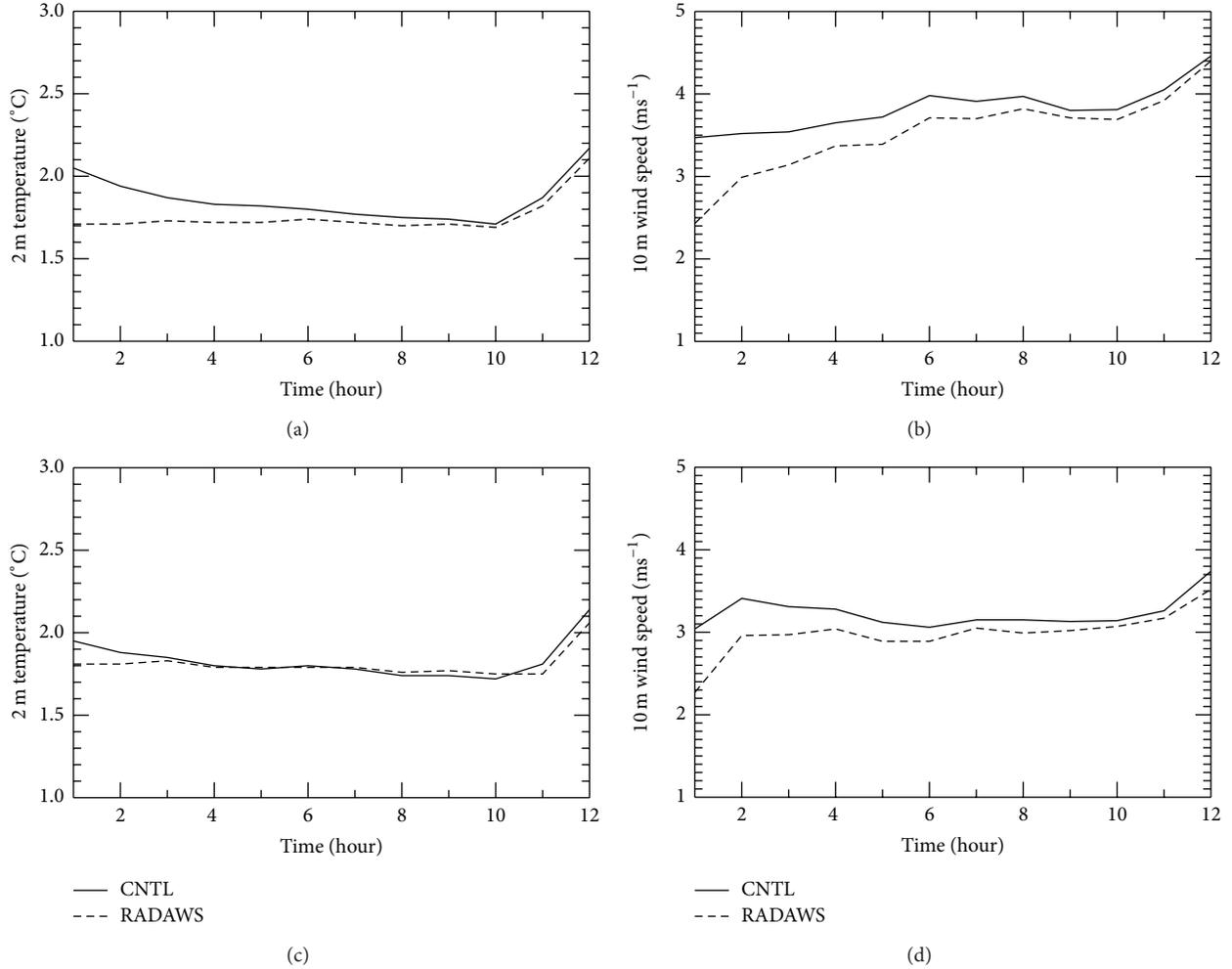


FIGURE 10: Time series of RMSEs for (a) 2 m temperature and (b) 10 m wind speed during the 12 h forecast for case 20110629 valid from 1300 UTC June 29 to 0000 UTC June 30 in the 4 km domain and RMSEs for (c) 2 m temperature and (d) 10 m wind speed during the 12 h forecast for case 20110715 valid from 1300 UTC July 15 to 0000 UTC July 16 in the 4 km domain.

for the same variables from AWS data. N is the total number of observations and is equal to or less than 437, depending on the availability of data.

Figure 10 presents the time series of RMSEs for 2 m temperature and 10 m wind speed from experiments CNTL and RADAWS for both cases. For case 20110629, compared with results in experiment CNTL, the RMSE in experiment RADAWS for temperature (Figure 10(a)) is reduced from 2.1 to 1.7°C at 1 h forecast, and for 10 m wind speed (Figure 10(b)) is reduced from 3.5 to 2.4 ms^{-1} at 1 h forecast. For case 20110715 (Figures 10(c) and 10(d)), both RMSEs for 2 m temperature and 10 m wind speed are reduced, although the positive impact on temperature only lasts for 3 h.

Overall, AWS data assimilation has positive impact on near-surface variables. Ha et al. [12] pointed out that surface data contributed to the occurrence of intensified low-level winds. Stensrud et al. [32] found that the assimilation of only surface observations was beneficial to the creation of more realistic mesoscale features associated with mesoscale convective systems. High-resolution surface data do have the

potential of improving storm-scale convective features [33]. Hence, although our focus is on improving QPF skill, it is still encouraging to add AWS data assimilation in the real-time forecasting system.

4.3. Data Assimilation Impact on QPF Skill. To evaluate the QPF skills of above data assimilation experiments, the bias score [34] and Equitable Threat Score (ETS, [35]) for 1, 3, and 6 h accumulated precipitations with thresholds of 1, 5 and 10 mm from the convective-scale domain are calculated. Similar to the calculation of the RMSE, interpolation from model grid data to station locations is first performed before calculating the biases and ETSs. For each case, the number of available sites is approximately 1000.

The bias is defined as

$$\text{bias} = \frac{F}{O}, \quad (4)$$

where F is the number of events that are forecast and O is the number of the events that occurred. For a specified

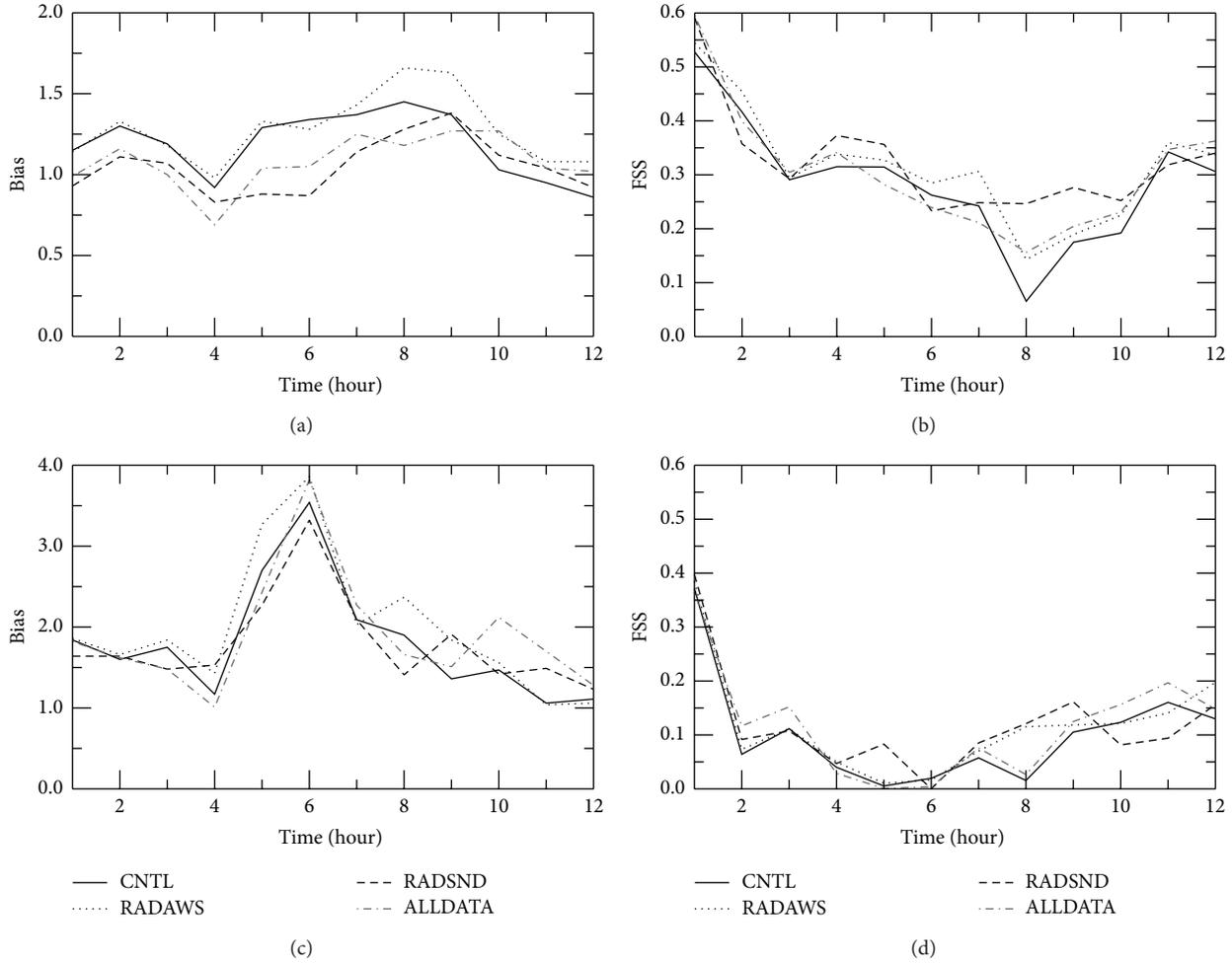


FIGURE 11: Bias score and FSS for hourly accumulated precipitation from experiments CNTL, RADAWS, RADSND, and ALLDATA for case 20110629. (a) Bias and (b) FSS with the threshold of 1 mm and (c) bias and (d) FSS with the threshold of 5 mm.

threshold, a perfect forecast would have a bias of 1, while values of bias less than and greater than one represent rainfall underprediction and overprediction, respectively.

The ETS is defined as

$$\text{ETS} = \frac{C - H}{F + O - C - H}, \quad (5)$$

where $H = (F \cdot O)/N$, C is the number of the events that are correctly forecast, and F and O are defined as above. For a perfect forecast, the ETS is 1. A larger ETS indicates a more skillful forecast.

The neighborhood-based Fractions Skill Score (FSS) is also used for the verification. It is defined as [36]

$$\text{FSS} = 1 - \frac{(1/N) \sum_{i=1}^N [P_{F(i)} - P_{O(i)}]^2}{(1/N) (\sum_{i=1}^N P_{F(i)}^2 - \sum_{i=1}^N P_{O(i)}^2)}, \quad (6)$$

where N is the number of grid points within the neighborhood of grid point i , $P_{F(i)}$ and $P_{O(i)}$ are the neighborhood probabilities at the i th grid box in the model forecast and observed fraction fields, respectively. A perfect forecast

would have a FSS of 1, while a score of 0 means no skill. The radius of influence in the study is specified as 25 km.

Figure 11 shows the bias score and FSS for hourly accumulated precipitation with thresholds of 1 and 5 mm from all the experiments for case 20110629. For the verifications with the threshold of 1 mm, the biases (Figure 11(a)) in RADSND and ALLDATA are generally smaller than those in CNTL and RADAWS, and the experiment RADAWS has the highest bias errors during the latter 5 h. The FSS (Figure 11(b)) are rather mixed, as additional AWS or radiosonde data assimilation produce better forecasts than experiment CNTL during certain hours, but no experiment is consistently better than the others throughout the 12 h forecast. The results from the verifications with the threshold of 5 mm (Figures 11(c) and 11(d)) are also mixed, including the higher degree of overprediction from RADAWS, smaller bias errors in the first few hours but worse results later from ALLDATA and fluctuations of FSS from all experiments.

Comparison of bias score and ETS for 6 h accumulated precipitation from all the experiments is shown in Figure 12. RADAWS produces larger biases than CNTL both in the

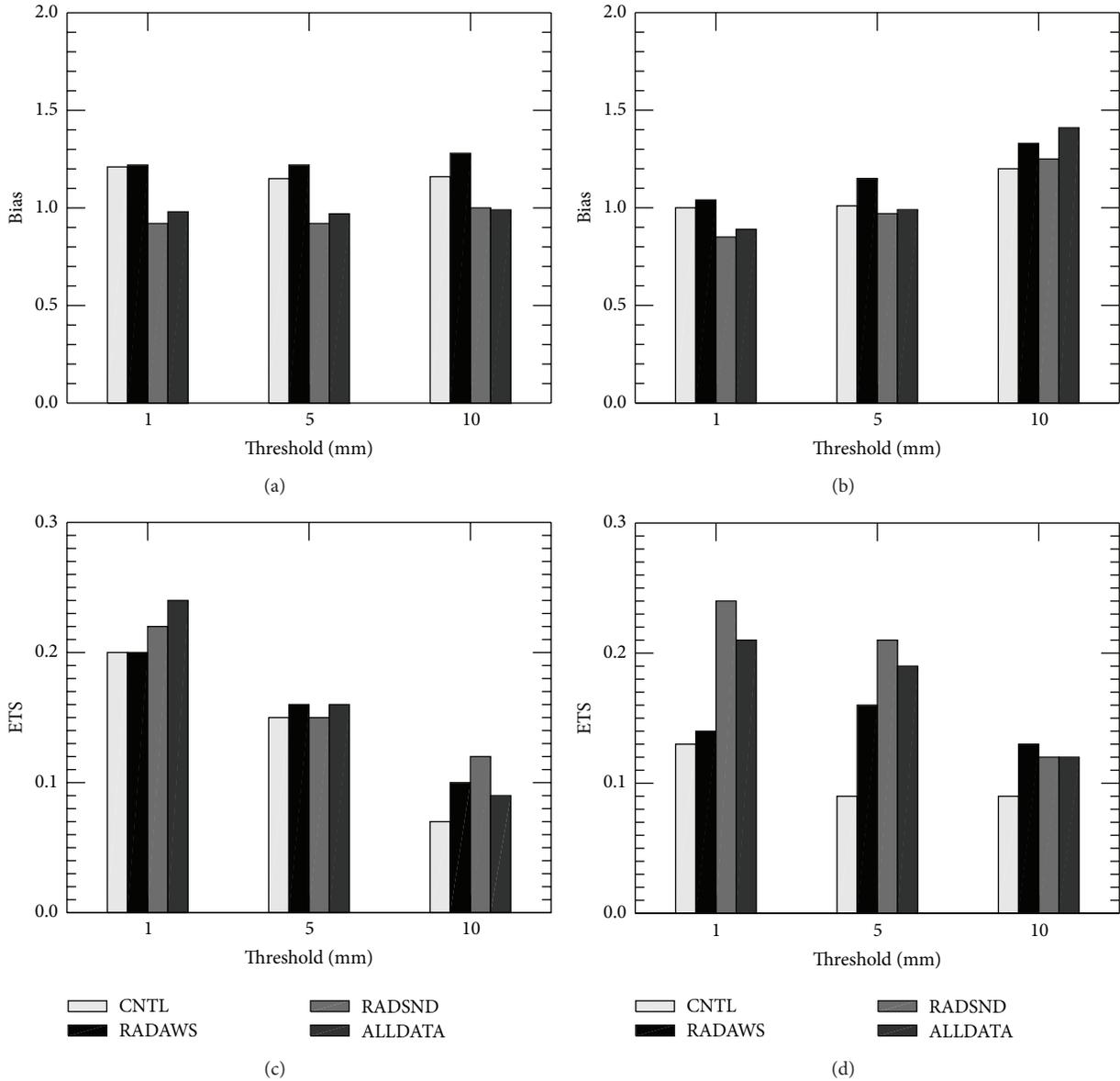


FIGURE 12: Bias score and ETS for 6 h accumulated precipitation with the thresholds of 1, 5, and 10 mm from the four experiments for case 20110629. Bias for (a) the first 6 h and (b) the latter 6 h precipitations and ETS for (c) the first 6 h and (d) the latter 6 h precipitations.

first and latter 6 h forecasts, indicating that addition of AWS data increases the degree of rainfall overprediction for this case. The biases of RADSND and ALLDATA are very similar, showing lower errors with all thresholds in the first 6 h and with the 1 and 5 mm thresholds in the latter 6 h. ETS in RADAWS are higher than those in CNTL, indicating that additional AWS data assimilation improves forecasts in terms of position accuracy. In comparison, RADSND and ALLDATA generally exhibit higher ETS than CNTL. In addition, RADSND and ALL DATA also show more skill than RADAWS with the 1 mm threshold, but not consistent with the other two thresholds.

Figure 13 shows the bias score and FSS for hourly accumulated precipitation with thresholds of 1, 5, and 10 mm from all the experiments for case 20110715. After an overview of

the biases in the left column, we can see that biases are clustered into two groups, including CNTL and RADAWS with larger biases, as well as RADSND and ALLDATA with lower bias errors. That means addition of radiosonde data reduces the degree of rainfall overprediction. FSS shown in Figure 13 also suggest that CNTL and RADAWS have similar skills, while RADSND and ALLDATA exhibit generally but not very consistent higher scores.

Figure 14 shows the bias score and ETS for the first and latter 6 h accumulated precipitations with the three thresholds from all the experiments for case 20110715. It can be seen that RADSND and ALLDATA exhibit lower biases than CNTL and RADAWS (Figures 14(a) and 14(b)), which is consistent with precipitation distribution as shown before. For ETS in the first 6 h (Figure 14(c)), RADAWS produces

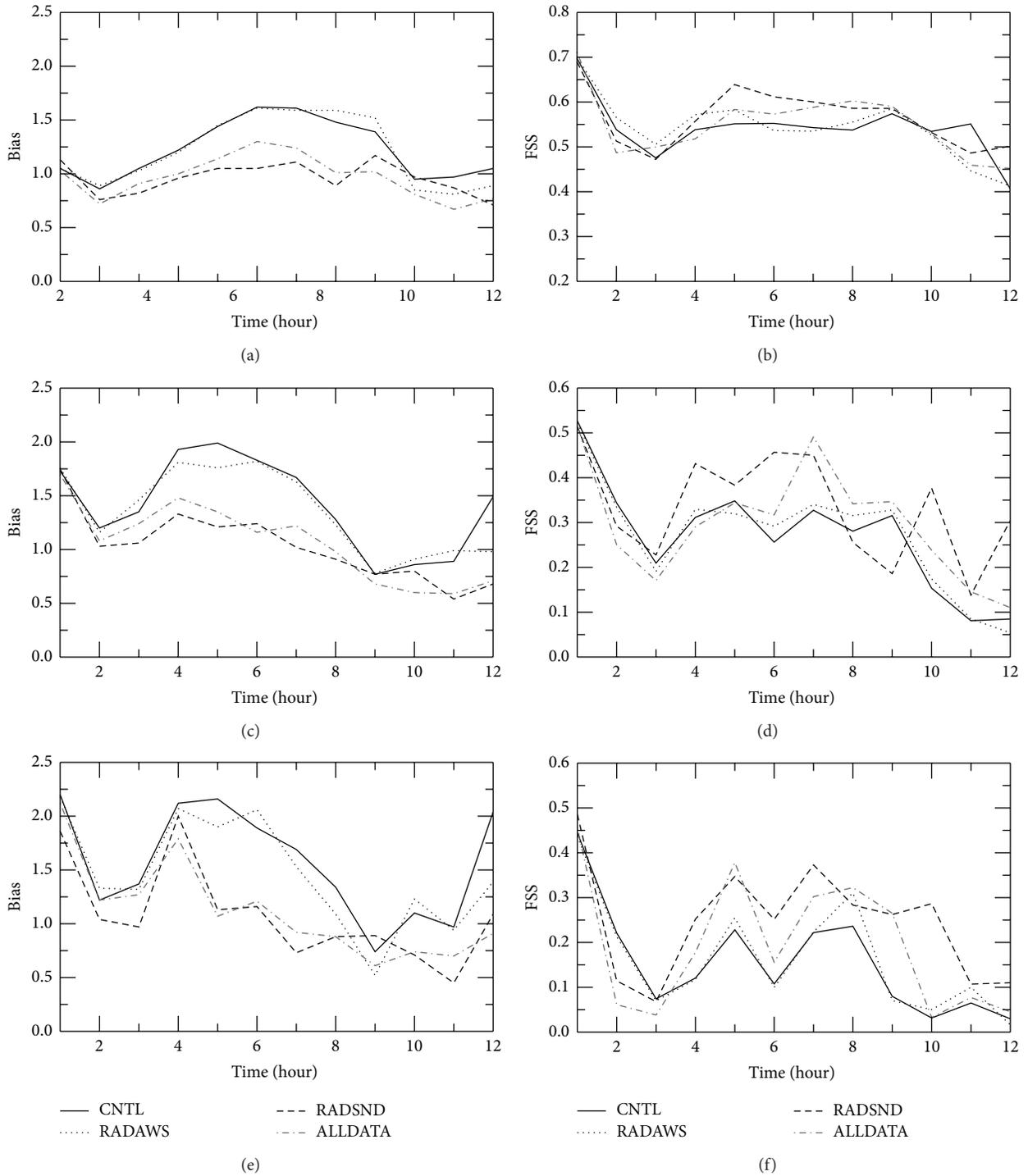


FIGURE 13: Bias score and FSS for hourly accumulated precipitation for case 20110715 from experiments CNTL, RADAWS, RADSND, and ALLDATA. (a) Bias and (b) FSS with the threshold of 1 mm, (c) bias and (d) FSS with the threshold of 5 mm, and (e) bias and (f) FSS with the threshold of 10 mm.

slightly better forecast than CNTL. When compared with RADAWS, RADSND and ALLDATA improve forecast skills slightly only for the thresholds of 5 and 10 mm. In the latter 6 h (Figure 14(d)), ALLDATA exhibit the highest ETS among all the experiments for both light and heavy rainfall.

In summary, additional surface AWS data assimilation improves the location of rainfall forecasting slightly; however, it may increase the degree of rainfall overestimation sometimes. In comparison, radiosonde data assimilation improves the QPF skill mainly by reducing rainfall overprediction.

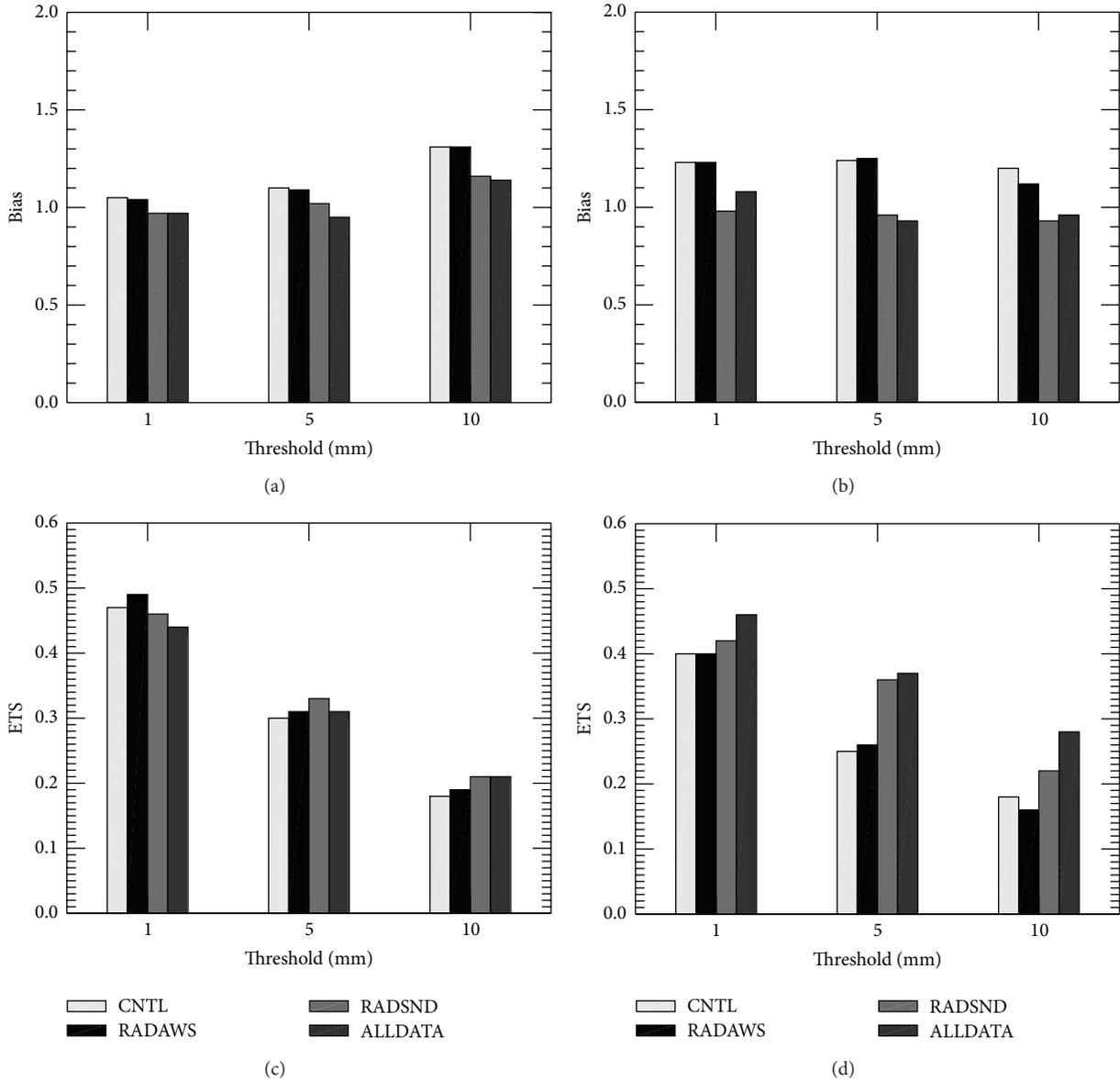


FIGURE 14: Bias score and ETS for 6 h accumulated precipitation with the thresholds of 1, 5, and 10 mm for case 20110715 from the four experiments. Bias for (a) the first 6 h and (b) the latter 6 h precipitations and ETS for (c) the first 6 h and (d) the latter 6 h precipitations.

Therefore, combination of the surface AWS and radiosonde data assimilation has the overall positive impact of improving rainfall position accuracy and reducing the degree of over-prediction. Dawson II and Xue [33] found that additional surface data in the initial conditions produced positive but not significant impact on position error forecast of the mesoscale convective system. The study of Sheng et al. [28] demonstrated that surface and upper-air data assimilation on the 6 km domain forecast was beneficial but rather small. On the whole, our results are in agreement with these previous studies. To compare the impact of surface AWS and radiosonde data with that of radar data more clearly, Figure 15 shows the bias score and ETS for 12 h accumulated precipitation from experiments CNTL and ALLDATA, as

well as forecasts without any data assimilation (short for NOASSIM) for both cases.

The biases scores of NOASSIM for both cases (Figures 15(a) and 15(c)) are less than 0.7 for the 1 mm threshold and even less than 0.5 for the 5 and 10 mm thresholds. In contrast, the bias scores of CNTL are too large, all over the perfect value of 1. Further addition of surface AWS and radiosonde data from experiment ALLDATA suggests the best forecasts in terms of bias errors, as they are closer to 1. The ETS from all experiments (Figures 15(b) and 15(d)) suggest that radar data assimilation improves the forecasts evidently and surface AWS and radiosonde data have the general positive but much smaller impact. The lower ETS in ALLDATA than that in CNTL with the 1 mm threshold (Figure 15(d)) is speculated

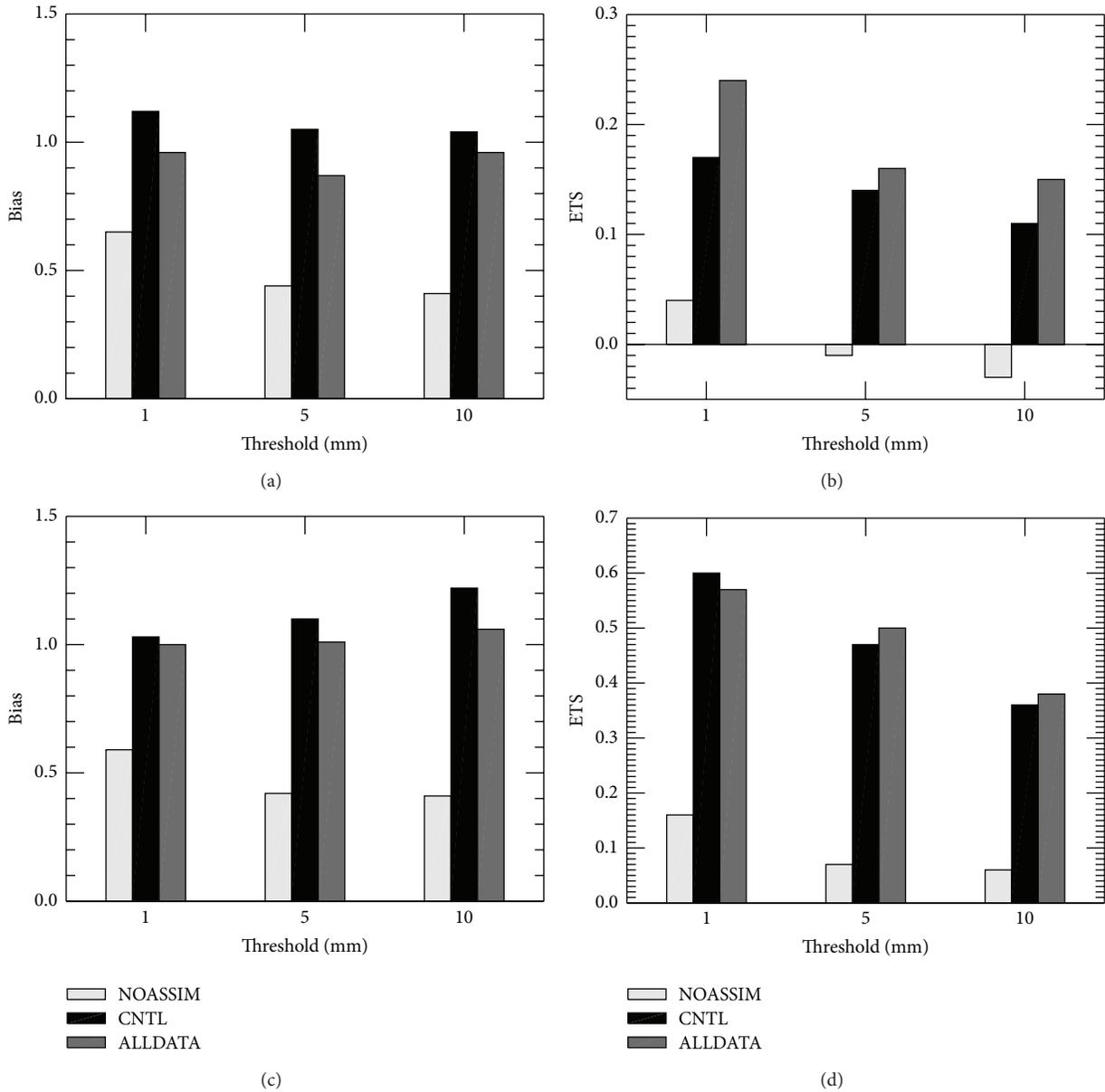


FIGURE 15: Bias score and ETS for 12 h accumulated precipitation with the thresholds of 1, 5, and 10 mm from NOASSIM, CNTL, and ALLDATA. (a) Bias and (b) ETS for case 20110629 and (c) bias and (d) ETS for case 20110715.

from the slight negative impact of radiosonde data, as shown before in Figure 14(c). Overall, additional surface AWS and radiosonde data assimilation reduce rainfall overprediction from radar data assimilation. Even though they have smaller impact than radar data in terms of ETS, they still should be considered in convective-scale forecasts.

5. Summary and Conclusions

This study is part of effort to improve the forecast skill of the HAPS system for the Shenzhen Meteorological Bureau, with the main focus on evaluating the impact of 3DVAR data assimilation on short-term QPF by considering different observation data sources from Doppler weather radars,

surface AWS, and conventional radiosondes. Data assimilation experiments for the June 29 and July 15, 2011 rainfall events over Southern China are carried out. Since the positive impact of radar data assimilation has been demonstrated in previous studies, this study is emphasized on evaluating additional surface AWS and radiosonde data assimilation impacts on short-term QPF skill.

Extrasturface AWS data assimilation has slight but general positive impact on rainfall location forecasts. The positive impact of AWS data can also be shown from the forecast of near-surface variables. At the same time, it cannot be denied that AWS data assimilation may increase the degree of rainfall overprediction sometimes. AWS data are available at the start time of every hour and can be provided timely,

being recommended in the daily real-time forecasting system.

Radiosonde data assimilation improves the QPF skill in terms of improving rainfall position accuracy and reducing rainfall overprediction. Although radiosonde data have relatively coarser resolution and are available every 12 h, they are also beneficial to convective-scale precipitation forecasts. Combination of the surface AWS and radiosonde data assimilation has the overall positive impact of improving rainfall position accuracy and reducing the degree of overprediction.

Radar data assimilation has the obvious positive impact on precipitation forecasts in terms of ETS but also produces higher bias errors for overprediction. Compared with radar data, the overall impact of additional surface and radiosonde data is smaller and is reflected primarily in reducing rainfall overestimation. Therefore, the assimilation of all radar, surface and radiosonde data has a more positive impact on the QPF skill than the assimilation of either type of data only.

Only two cases are investigated in this study. To improve the precipitation forecast more effectively, more research efforts are needed, including more case studies and further quality control of observation data. It should be noted that cycling experiments are also performed with a 6 h AWS data cycle window at 1 h intervals in the 12 km domain, with mixed or even negative impacts. AWS data quality control and improvement of the cycling process might need to be further looked into in the future work. In addition, higher model resolution of 1 km or even sub-100-meter should also be examined in future studies.

Acknowledgments

The research was supported by a grant to CAPS from Shenzhen Meteorological Bureau (SZMB). All numerical experiments and analyses were performed on the SZMB super-computer facility. The authors benefited from discussions with Dr. Jidong Gao for using the ARPS 3DVAR package. The work was also supported by the National Basic Research Program of China (973 Program, Grant no. 2013CB430105), the Knowledge Innovation Program of the Chinese Academy of Sciences (Grant no. KZCX2-EW-203), and the National Natural Science Foundation of China (Grant no. 41105095).

References

- [1] M. Xue, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, "The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation," *Meteorology and Atmospheric Physics*, vol. 82, no. 1-4, pp. 139–170, 2003.
- [2] F. Kong, K. K. Droegemeier, and N. L. Hickmon, "Multiresolution ensemble forecasts of an observed tornadic thunderstorm system. Part II: storm-scale experiments," *Monthly Weather Review*, vol. 135, no. 3, pp. 759–782, 2007.
- [3] A. J. Clark, S. J. Weiss, J. S. Kain et al., "An overview of the 2010 Hazardous Weather Testbed experimental forecast program spring experiment," *Bulletin of the American Meteorological Society*, vol. 93, no. 1, pp. 55–74, 2012.
- [4] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [5] M. Hu, M. Xue, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: cloud analysis and its impact," *Monthly Weather Review*, vol. 134, no. 2, pp. 675–698, 2006.
- [6] K. Zhao and M. Xue, "Assimilation of coastal Doppler radar data with the ARPS 3DVAR and cloud analysis for the prediction of Hurricane Ike (2008)," *Geophysical Research Letters*, vol. 36, no. 12, Article ID L12803, 2009.
- [7] A. D. Schenkman, M. Xue, A. Shapiro, K. Brewster, and J. Gao, "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*, vol. 139, no. 1, pp. 224–246, 2011.
- [8] T. Hou, F. Kong, and X. Chen, "Evaluation of 3DVAR data assimilation and planetary boundary layer scheme sensitivities for two heavy rainfall cases in Southern China," in *Proceedings of the 13th Weather Research and Forecasting Model Workshop*, p. 76, Boulder, Colo, USA, 2012.
- [9] F. H. Ruggiero, K. D. Sashegyi, R. V. Madala, and S. Raman, "The use of surface observations in four-dimensional data assimilation using a mesoscale model," *Monthly Weather Review*, vol. 124, no. 5, pp. 1018–1033, 1996.
- [10] D. R. Stauffer, N. L. Seaman, and F. S. Binkowski, "Use of four-dimensional data assimilation in a limited-area mesoscale model Part II: effects of data assimilation within the planetary boundary layer," *Monthly Weather Review*, vol. 119, no. 3, pp. 734–754, 1991.
- [11] J. S. Kain, M. Xue, M. C. Coniglio et al., "Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting-research environment," *Weather and Forecasting*, vol. 25, no. 5, pp. 1510–1521, 2010.
- [12] J.-H. Ha, H.-W. Kim, and D.-K. Lee, "Observation and numerical simulations with radar and surface data assimilation for heavy rainfall over central Korea," *Advances in Atmospheric Sciences*, vol. 28, no. 3, pp. 573–590, 2011.
- [13] Q. Xiao, Y.-H. Kuo, J. Sun et al., "Assimilation of Doppler radar observations with a regional 3DVAR system: impact of Doppler velocities on forecasts of a heavy rainfall case," *Journal of Applied Meteorology*, vol. 44, no. 6, pp. 768–788, 2005.
- [14] Q. Xiao and J. Sun, "Multiple-radar data assimilation and short-range quantitative precipitation forecasting of a squall line observed during IHOP_2002," *Monthly Weather Review*, vol. 135, no. 10, pp. 3381–3404, 2007.
- [15] J. Gao and D. J. Stensrud, "Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification," *Journal of the Atmospheric Sciences*, vol. 69, no. 3, pp. 1054–1065, 2012.
- [16] S. J. Rennie, S. L. Dance, A. J. Illingworth, S. P. Ballard, and D. Simonin, "3D-Var assimilation of insect-derived Doppler radar radial winds in convective cases using a high-resolution model," *Monthly Weather Review*, vol. 139, no. 4, pp. 1148–1163, 2011.
- [17] E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, "Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave," *Journal of Geophysical Research D*, vol. 102, no. 14, pp. 16663–16682, 1997.

- [18] J. Dudhia, "Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model," *Journal of the Atmospheric Sciences*, vol. 46, no. 20, pp. 3077–3107, 1989.
- [19] Z. I. Janjic, "The step-mountain coordinate: physical package," *Monthly Weather Review*, vol. 118, no. 7, pp. 1429–1443, 1990.
- [20] Z. I. Janjic, "Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso Model," 2001, <http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf>.
- [21] F. Chen and J. Dudhia, "Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity," *Monthly Weather Review*, vol. 129, no. 4, pp. 569–585, 2001.
- [22] J. S. Kain and J. Kain, "The Kain—Fritsch convective parameterization: an update," *Journal of Applied Meteorology*, vol. 43, no. 1, pp. 170–181, 2004.
- [23] G. Thompson, P. R. Field, R. M. Rasmussen, and W. D. Hall, "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization," *Monthly Weather Review*, vol. 136, no. 12, pp. 5095–5115, 2008.
- [24] M. Hu, M. Xue, J. Gao, and K. Brewster, "3DVAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: impact of radial velocity analysis via 3DVAR," *Monthly Weather Review*, vol. 134, no. 2, pp. 699–721, 2006.
- [25] X. Xu, L. Liu, and G. Zheng, "Dynamical and microphysical retrieval from simulated Doppler radar observations using the 4DVAR assimilation technique," *Acta Meteorologica Sinica*, vol. 19, no. 2, pp. 160–173, 2005.
- [26] J. Gong, L. Wei, S. Tao et al., "Accurate estimation and application of 3D error covariance structures in global data assimilation. Part I: accurate estimation of error covariance in observation space," *Acta Meteorologica Sinica*, vol. 64, no. 6, pp. 669–683, 2006 (Chinese).
- [27] Z. Xu, J. Gong, J. Wang et al., "A study of assimilation of surface observational data in complex terrain. Part I: influence of the elevation difference between model surface and observation site," *Chinese Journal of Atmospheric Sciences*, vol. 31, no. 2, pp. 222–232, 2007 (Chinese).
- [28] C. Sheng, S. Gao, and M. Xue, "Short-range prediction of a heavy precipitation event by assimilating Chinese CINRAD-SA radar reflectivity data using complex cloud analysis," *Meteorology and Atmospheric Physics*, vol. 94, no. 1-4, pp. 167–183, 2006.
- [29] H. Liu, J. Xue, J. Gu, and H. Xu, "Radar data assimilation of the GRAPES model and experimental results in a typhoon case," *Advances in Atmospheric Sciences*, vol. 29, no. 2, pp. 344–358, 2012.
- [30] Y. Guo and Y. Ding, "Long-term free-atmosphere temperature trends in China derived from homogenized in situ radiosonde temperature series," *Journal of Climate*, vol. 22, no. 4, pp. 1037–1051, 2009.
- [31] Y. Guo and Y. Ding, "Impacts of reference time series on the homogenization of radiosonde temperature," *Advances in Atmospheric Sciences*, vol. 28, no. 5, pp. 1011–1022, 2011.
- [32] D. J. Stensrud, N. Yussouf, D. C. Dowell, and M. C. Coniglio, "Assimilating surface data into a mesoscale model ensemble: cold pool analyses from spring 2007," *Atmospheric Research*, vol. 93, no. 1–3, pp. 207–220, 2009.
- [33] D. T. Dawson II and M. Xue, "Numerical forecasts of the 15–16 June 2002 southern plains mesoscale convective system: impact of mesoscale data and cloud analysis," *Monthly Weather Review*, vol. 134, no. 6, pp. 1607–1629, 2006.
- [34] R. A. Anthes, "Regional models of the atmosphere in middle latitudes," *Monthly Weather Review*, vol. 111, no. 6, pp. 1306–1335, 1983.
- [35] J. T. Schaefer, "The critical success index as an indicator of warning skill," *Weather and Forecasting*, vol. 5, no. 4, pp. 570–575, 1990.
- [36] N. M. Roberts and H. W. Lean, "Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events," *Monthly Weather Review*, vol. 136, no. 1, pp. 78–97, 2008.

Review Article

WRF-ARW Variational Storm-Scale Data Assimilation: Current Capabilities and Future Developments

Juanzhen Sun and Hongli Wang

National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, USA

Correspondence should be addressed to Juanzhen Sun; sunj@ucar.edu

Received 6 May 2013; Accepted 18 June 2013

Academic Editor: Jidong Gao

Copyright © 2013 J. Sun and H. Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The variational radar data assimilation system has been developed and tested for the Advanced Research Weather Research and Forecasting (WRF-ARW) model since 2005. Initial efforts focused on the assimilation of the radar observations in the 3-dimensional variational framework, and recently the efforts have been extended to the 4-dimensional system. This article provides a review of the basics of the system and various studies that have been conducted to evaluate and improve the performance of the system. Future activities that are required to further improve the system and to make it operational are also discussed.

1. Introduction

In the past two decades active research was conducted on the development of techniques to initialize storm-scale numerical prediction models. It has been recognized that the success will critically depend on the optimal use of the national operational WSR-88D radar network that covers the United States with single Doppler coverage in most areas. Although the network provides observations at a resolution that is able to resolve atmospheric convection, they are only limited to radial wind and reflectivity. Therefore several early studies focused on the feasibility of retrieving meteorological fields from these single Doppler observations. Techniques with different complexities have been developed which aim at obtaining the unobserved meteorological variables such as 3-dimensional (3D) wind, temperature, and microphysical fields from the radar observations of radial velocity and reflectivity (e.g., [1–5]).

The techniques that make use of a numerical model in a data assimilation (DA) context received particular attention because they combine the retrieval, initialization, and forecast in one system. The first radar DA system for the storm-scale was developed based on the 4-dimensional variational data assimilation (4D-Var) technique and a boundary layer fluid dynamics model for the retrieval of the 3D wind and temperature [1]. This system, known as VDRAS (Variational

Doppler Radar Analysis System), was later expanded to include microphysical retrieval, as well as short-term forecasts initialized by these retrieved fields [6–9]. Another variational-based radar DA system was developed by Gao et al. [4] using a 3-dimensional variational data assimilation (3D-Var) technique in the framework of the ARPS (Advanced Research and Prediction System [10]) model. A so-called 3.5-dimensional variational radar data assimilation based on Navy's COAMPS (The Coupled Ocean/Atmosphere Mesoscale Prediction System) was developed and demonstrated through a number of studies [11, 12]. These variational systems showed great potentials in the use of radar observations for initializing high-resolution numerical models through several case studies and real-time demonstrations [13, 14].

The demonstrated potential motivated the development of a radar DA scheme in the variational DA system WRFVAR of the community model ARW-WRF (Advanced Research Weather Research and Forecasting; hereafter refer to WRF). WRFVAR includes both 3D-Var [15, 16] and 4D-Var [17] components. The radar DA scheme was first developed for WRF 3D-Var [18, 19] and recently expanded to 4D-Var [20, 21]. Unlike the ARPS 3D-Var system and VDRAS whose developments were motivated by the assimilation of radar observations for convective-scale analysis, the WRFVAR system was started with conventional data assimilation without

the convective-scale analysis as a priority. It thus emphasized the synoptic-scale balance through the use of model forecast background and its error covariance. It also used the stream function and velocity potential as the momentum control variables instead of the direct velocities as in VDRAS and ARPS 3D-Var. In addition, the incremental formulation [22] that is commonly used in large-scale data assimilation systems was adopted by WRFVAR, which might not be suitable for the storm-scale applications because of the linearization required in the formulation. Therefore, the addition of a radar data assimilation capability to WRF, which is a forecast system aiming at broad research and operational applications, presents new challenges that do not exist in the systems that focus on the storm-scale analysis and forecasting.

In recent years several operational centers throughout the world have also been developing and testing capabilities of assimilating radar observations to initialize high-resolution convection-permitting NWP models. One of the notable efforts was taken by the Met Office of the United Kingdom using the unified model and its variational data assimilation system [23]. The lessons learnt in understanding and finding solutions to the aforementioned challenges in WRFVAR have a wide application to the operational community because some operational data assimilation systems use the similar framework as WRFVAR and will hence face the same challenges.

The purpose of this paper is to provide a review of the progress of the WRFVAR system and its planned future developments. Since the initial development of the system, a good number of studies have been conducted which either aim at the further development of the system or at answering some challenging scientific questions. A review of these studies will benefit the future users of the community system. In Section 2, we provide a review on the overall system design, including the fundamental framework of WRFVAR, radar observation operators, control variables and their background error (BE) estimates, and automated radar data quality control. The performance of the 3D-Var system is reviewed in Section 3 and of the 4D-Var system in Section 4. At the last section, we discuss the planned future efforts for the DA system.

2. System Design of WRFVAR Radar DA

2.1. Fundamental Framework of WRFVAR. The WRFVAR follows the incremental variational formulation [22] that is commonly used in operational systems. In the standard variational formulation, the optimal analysis is obtained by minimizing a cost function measuring the misfit between observation and prediction by a nonlinear model. In the incremental variational formulation, however, a linearized model is used to generate the prediction needed in the cost function. The advantage of the incremental approach is that it not only reduces the computational cost of 4D-Var but also improves the mathematical conditioning of the cost function (i.e., a quadratic cost function) because of the linearization of the forward operator. It facilitates the use of a coarser resolution tangent linear model (TLM) and adjoint model

(ADM) to reduce the computational cost and a simplified representation of physical processes in TLM and ADM in the data assimilation cycle.

The cost function of the incremental formulation can be derived from the standard cost function with some assumptions that will be described later. First we express the standard cost function as follows:

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2} \sum_{k=0}^K \{ \mathbf{y}_k^o - H_k [M_k(\mathbf{x}_0)] \}^T \times \mathbf{R}^{-1} \{ \mathbf{y}_k^o - H_k [M_k(\mathbf{x}_0)] \}. \quad (1)$$

This cost function assumes that the assimilation window covers K observation windows with each represented by the subscript k . The variables of \mathbf{x}_0 , \mathbf{x}_0^b , and \mathbf{y}_k^o represent the initial atmospheric state, the background state which can be previous model forecast, and the observed state, respectively. M_k is the nonlinear prediction model to propagate the initial atmospheric state to that at the k th observation time in the case of 4D-Var. It should be noted that 3D-Var is regarded as a special case in which K can simply set to 0 and M_k set to the identity matrix. H_k is the nonlinear observation operator. \mathbf{B} and \mathbf{R} are the background and observation error covariance matrices, respectively.

To derive the cost function of the incremental formulation, we first introduce the innovation variable \mathbf{d} at the k th observation time as

$$\mathbf{d}_k = \mathbf{y}_k^o - H_k [M_k(\mathbf{x}_0)] \quad (2)$$

and then linearize the operators H_k and M_k in (1) as follows:

$$H_k [M_k(\mathbf{x}_0^n)] \approx H_k [M_k(\mathbf{x}_0^{n-1})] + \mathbf{H}_k \mathbf{M}_k [\mathbf{x}_0^{n-1}] \delta \mathbf{x}_0^n, \quad (3)$$

where the superscripts n and $n-1$ represent the current and previous outer loop iterations, respectively. \mathbf{H}_k and \mathbf{M}_k are the tangent linear operator of H_k and M_k . Using (2) and (3), the cost function (1) can be changed to

$$J^n [\delta \mathbf{x}_0^n] = \frac{1}{2} \{ \delta \mathbf{x}_0^n - [\mathbf{x}_0^b - \mathbf{x}_0^{n-1}] \}^T \mathbf{B}^{-1} \{ \delta \mathbf{x}_0^n - [\mathbf{x}_0^b - \mathbf{x}_0^{n-1}] \} + \frac{1}{2} \sum_{k=0}^K \{ \mathbf{H}_k \mathbf{M}_k [\mathbf{x}_0^{n-1}] \delta \mathbf{x}_0^n - \mathbf{d}_k^{n-1} \}^T \times \mathbf{R}^{-1} \{ \mathbf{H}_k \mathbf{M}_k [\mathbf{x}_0^{n-1}] \delta \mathbf{x}_0^n - \mathbf{d}_k^{n-1} \}. \quad (4)$$

When the increment $\delta \mathbf{x}_0^n$ at the n th outer loop iteration is obtained, the estimate of the atmospheric state is updated by $\mathbf{x}_0^n = \mathbf{x}_0^{n-1} + \delta \mathbf{x}_0^n$ and used to produce the first guess trajectory $M_k[\mathbf{x}_0^n]$ for the next outer loop $n+1$. The basic assumption of the incremental approach is that the solution of the cost function (4) gradually approaches that of the nonlinear cost function (1) given enough number of outer loop iterations.

Assuming the background error covariance matrix is given by $\mathbf{B} = \mathbf{U}\mathbf{U}^T$ and a control variable transform by $\delta\mathbf{x}_0 = \mathbf{U}\mathbf{v}$, the background term of the cost function (4) is simplified to $(1/2)(\sum_{i=1}^N \mathbf{v}^i)^T (\sum_{i=1}^N \mathbf{v}^i)$, where N is the total number of outer loops, whereby avoiding the difficulty in computing the inverse of \mathbf{B} . The control variable transform $\delta\mathbf{x}_0 = \mathbf{U}\mathbf{v}$ is implemented through a series of operations $\mathbf{U} = \mathbf{U}_p \mathbf{U}_v \mathbf{U}_h$ [15]. The horizontal transform \mathbf{U}_h is performed using a recursive filter [24]. The vertical transform \mathbf{U}_v applies an EOF (empirical orthogonal function) decomposition on the vertical component of the background error covariance. The physical transform \mathbf{U}_p converts the increment in control variable space to analysis variable space. The control variables of the WRFVAR system will be described later in this section.

The optimal analysis is obtained by minimizing the cost function (4) by iterating in both the outer and inner loops. The number of inner loop iterations is determined either by a convergence criterion or a prespecified fixed number. The number of the outer loop iterations is typically 2 to 6. The gradient of the cost function required in the minimization process is computed by the adjoint method. For the 4D-Var system with the incremental formulation, a forward model and its TLM and ADM are needed. The forward model is the same as the WRF forecast model but currently only with the physical schemes of the Kessler microphysics and the diffusion. More schemes will be added to the forward model in the future as the adjoint of these schemes has been developed.

2.2. Radar Observation Operators. To assimilate radar observations, the following observation terms are added to the existing cost function:

$$J = J_{\text{old}} + \frac{1}{2} \sum_{k=0}^k \left[(\mathbf{V}_{rk} - \mathbf{V}_{rk}^{\text{ob}})^T \mathbf{R}_v^{-1} (\mathbf{V}_{rk} - \mathbf{V}_{rk}^{\text{ob}}) + \frac{1}{2} (\mathbf{Z}_k - \mathbf{Z}_k^{\text{ob}})^T \mathbf{R}_z^{-1} (\mathbf{Z}_k - \mathbf{Z}_k^{\text{ob}}) \right], \quad (5)$$

where J_{old} is used to represent the existing cost function before radar data assimilation is developed. The variables \mathbf{V}_r stand for the radial velocity and \mathbf{Z} for the reflectivity factor. The superscript “ob” indicates the observations. The symbols \mathbf{R}_v^{-1} and \mathbf{R}_z^{-1} stand for observation error covariance matrices for radial velocity and reflectivity, respectively. Note that the summation over the observation time levels k is not needed in the case of 3D-Var. The observation operator H_k in the cost function (1) links the model variables in a model coordinate to the observation variables in an observation space. For the radar radial velocity, this linkage is formulated with the 3D wind field (u , v , and w), the hydrometeor fall speed V_t , and the distance D between the location of a data point and the radar antenna:

$$V_r = \frac{1}{D} [(x_d - x_r)u + (y_d - y_r)v + (z_d - z_r)(w - V_t)], \quad (6)$$

$$D = \sqrt{(x_d - x_r)^2 + (y_d - y_r)^2 + (z_d - z_r)^2}, \quad (7)$$

where (x_d, y_d, z_d) represents the location of the observation point and (x_r, y_r, z_r) represents the location of the radar station. V_t is calculated from the rainwater mixing ratio with a height correction following Sun and Crook [6]. Relation (6) is linear except that V_t is nonlinearly dependent on q_r , which needs to be linearized to obtain the linear observation operator for V_r . Note that the earth curvature effect [25] must be considered when mapping the radar observations to the model grid, which is done in the radar data preprocessing package that is described in the next subsection.

The formulation of the reflectivity operator is not as straightforward because it depends on the assumption of drop size distribution in a microphysical parameterization scheme and the classification of hydrometeors. Following Sun and Crook [6], Xiao et al. [19] used the following relation for WRF 3D-Var by assuming the Marshall-Palmer drop size distribution for rain:

$$Z = 43.1 + 17.5 \log_{10}(\rho q_r), \quad (8)$$

where ρ is the air density (kg m^{-3}) and q_r is the rainwater mixing ratio (g kg^{-1}). Zhang et al. [26] examined a different relation by assuming a constrained gamma drop size distribution and found some improvement in the analysis of low precipitation.

Equation (8) was used in the cost function (5) to assimilate reflectivity in the WRF 3D-Var developed by Xiao et al. [19]. However, Wang et al. [27] showed that the linearization of (8) as required by the incremental formulation of the cost function can result in a dry bias in rainwater analysis. Thus they proposed to indirectly assimilate the derived rainwater mixing ratio q_r from (8) by replacing Z with q_r in (5).

2.3. Control Variables and Background Error Estimate. The standard control variables (CV) [15] for WRF 3D-Var and 4D-Var are the stream function ψ , the unbalanced component of velocity potential χ_u , the unbalanced component of temperature T_u , the unbalanced component of surface pressure P_{su} , and “pseudo” relative humidity Rhs. The unbalanced component for each of the three variables is the residual after subtracting the respective correlated component with ψ which is obtained by a statistical fitting. The “pseudo” relative humidity is the water vapor mixing ratio divided by its saturated value in the background state. To assimilate radar radial velocity and reflectivity associated with warm rain, vertical velocity and the microphysical parameters of cloud water and rainwater must be added as control variables. In the current WRFVAR system, no cross-correlation with other variables is considered for relative humidity, cloud water mixing ratio, rain water mixing ratio, and vertical velocity. The 4D-Var technique through the use of the WRF model can implicitly produce multivariate correlations for the analysis.

One issue that raises concerns for the convective-scale DA is the selection of momentum control variables. The use of stream function and velocity potential (ψ and χ) was widely accepted by global operational models (e.g., National Center of Environmental Prediction’s Gridpoint Statistical Interpolation and United Kingdom Met Office Unified Model 3D-Var system). However, a recent study by Xie and MacDonals [28]

suggested that the use of stream function and velocity potential as momentum control variables might not be suitable for analysis over a regional domain for small-scale problems. Other mesoscale DA systems that involve radar observations, such as VDRAS and ARPS 3D-Var, used u -wind along x -grid direction and v -wind along y -grid direction as momentum control variables. An effort is being undertaken to compare the impact of the choice of momentum control variables on high-resolution analysis. Some preliminary results will be shown in Section 5.

For control variables with option 5 (CV5) in WRFVAR, a recursive filter and EOFs (empirical orthogonal functions) are used to model background error covariance. The horizontal autocovariance, which is assumed to be spatially homogeneous and isotropic, is modeled by the recursive filter. The vertical error correlations are modeled using the diagnosed EOFs (empirical orthogonal functions) from estimated background errors. Since convective-scale data assimilation using radar observations is commonly done on limited area domains, it is preferable to derive background error statistics over the domain of interest as well as the season of interest. There are two approaches to estimate the background error statistics of the control variables in WRFVAR: the ensemble method [20, 29] and the NMC method [30]. Both methods are included in the utility GEN_BE [15] of the standard released WRF data assimilation system. Both algorithms were used in previous WRFVAR studies.

An example of the estimated background error statistics is given in Figure 1, which shows the horizontal length scales of the control variables with respect to the vertical EOF modes. The error statistics is computed on a domain covering a large part of the USA (N24°–N49° and W116°–W74°) using the NMC method and the June 2012 data from the real-time WRF 3 km runs. The plot shows that the length scales are a few hundred kilometers for the dynamical variables and temperature in the first few modes and reduced to below 100 kilometers in high frequency modes. The length scale of relative humidity is below 50 kilometers for all EOF modes.

2.4. Automated Radar Data Quality Control. An automated radar data quality control and error estimate system that can be executed in real-time is an essential component of an operational radar DA system. The automated radar data preprocessing system embedded in NCAR’s real-time 4D-Var system VDRAS that is based on a cloud-scale model (different from WRF) [6] produces quality controlled radar data for use by WRFVAR. The system along with its radar data quality control package has been used in real-time since 2001. The VDRAS data preprocessing system includes several quality control algorithms as well as filtering, superobbing, and error computation. In Sun [31], some of these preprocessing methods were described. One of the challenging issues in the use of radar radial velocity for various applications is the data quality control of aliased velocities, known as velocity dealiasing. Radial velocity is aliased when the true radial velocity is larger than the unambiguous velocity of a radar or Nyquist velocity [25]. A unique feature of the VDRAS data preprocessing system is a velocity dealiasing algorithm

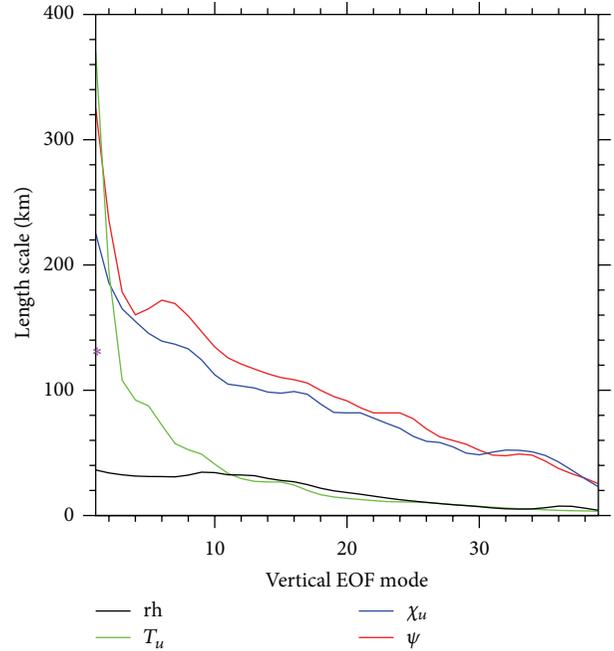


FIGURE 1: Length scales of forecast errors of the WRFVAR control variables of stream function (ψ), unbalanced velocity potential (χ_u), unbalanced temperature (T_u), and relative humidity (rh), with respect to the vertical EOF mode, computed using 3 km WRF forecasts over a domain that covers a region of N24°–N49° and W116°–W74° in the USA. Note that the surface pressure has only one mode (indicated by the purple *) because it is a 2D field.

coupled with the 4D-Var radar data assimilation in which the frequently updated wind analysis through a continuous cycling is used as the reference wind for the dealiasing. Lim and Sun [32] demonstrated that the coupling of the velocity dealiasing algorithm with the radar DA system improved the accuracy of the reference wind and hence resulted in an improved dealiased velocity field by significantly reducing the percentage of the improperly dealiased data from using a reference wind field that had no radar data assimilation.

The velocity dealiasing can be particularly challenging during hurricane events because of the strong wind and the large error in the reference wind, especially near the center where a slight offset of center location can cause a large difference in wind direction. Recent applications of the coupled dealiasing scheme in VDRAS suggest that the scheme is effective to automatically unfold severely aliased velocity observations. Figure 2 shows an example from a large number of radar volumes that were successfully dealiased for Typhoon Morakot that affected Taiwan during 7–9 August 2008.

3. Review of the 3D-Var Performances

The performance and impact of radar DA using WRF 3D-Var were examined through studies of individual cases as well as multiple cases/days under operational settings. The first case

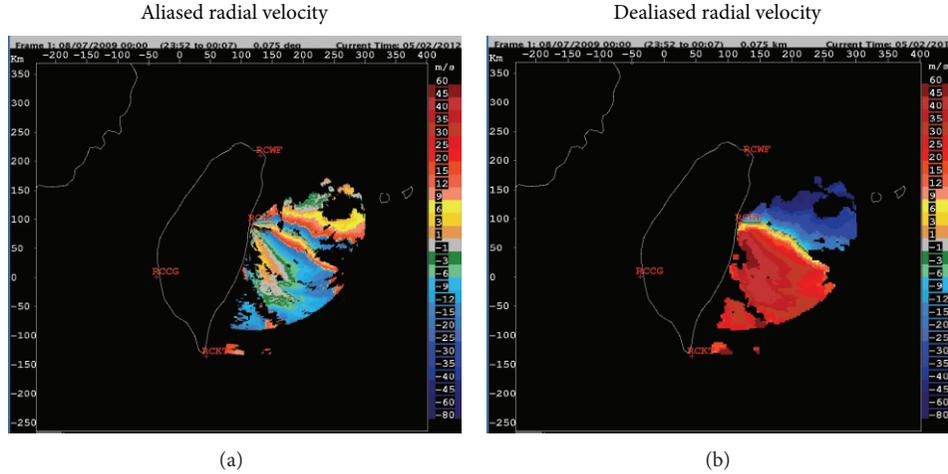


FIGURE 2: Radial velocity observations at 0.5° elevation from RCHL radar of the Central Weather Bureau (CWB) of Taiwan before (a) and after (b) the dealiasing scheme in VDRAS was applied. The thin white contours are coastal lines.

study with radar radial velocity data assimilation in WRF 3D-Var was reported by Xiao et al. [18] using a heavy rainfall case that occurred over the Korean Peninsula. They found that the assimilation of the radial velocity data from Jindo radar operated by the Korean Meteorological Administration improved the wind analysis (especially the northward component) resulting in substantial improvement for 3 h rainfall forecasts and noticeable improvement for 6 h forecasts. In a follow-up paper, Xiao et al. [19] described a methodology to assimilate radar reflectivity data into the WRF 3D-Var. Although the reflectivity is linked to the rain water mixing ratio, they chose the total liquid water q_t as the control variable instead of q_r , not only because the distribution of total water is more Gaussian-like but also the multivariate analyses can be achieved by use of a partition scheme. In this scheme, the analysis increment of q_t is partitioned into the water vapor mixing ratio q_w , the cloud water q_c , and the rain water q_r using a warm rain microphysical scheme. The method was first tested with the landfalling typhoon Rusa (2002) over South Korea. Their results indicated that the positive impact of radar reflectivity lasted only for 3 hours. However, when both radial velocity and reflectivity radar observations were assimilated, noticeable improvement lasted up to 12 hours.

The above 3D-Var radar data assimilation system was further evaluated by Pu et al. [33] in a case study of the Hurricane Dennis near landfall that occurred during 2005 in southeastern Cuba. In this study, the eastward and northward velocities synthesized from an airborne Doppler radar as well as reflectivity were assimilated into the WRF 3D-Var. They demonstrated that the radar data significantly improved the short-term (18 hours) forecasts of the intensity, track, and precipitation. However, similarly as in Xiao et al. [19], the impact was mainly attributed to the assimilation of the Doppler wind observations.

The relatively small impact of the reflectivity data in the scheme developed by Xiao et al. [19] motivated a further examination of the reflectivity assimilation. Wang et al. [27] pointed out some issues with this scheme, namely,

the use of a linearized form of the $Z-q_r$ equation (8) as the observation operator and the warm-rain partition scheme. As aforementioned, the WRF 3D-Var applies the incremental approach that requires the linearization of the forward model. The logarithm $Z-q_r$ equation (8) has a high degree of nonlinearity especially when q_r is small. They found a dry bias was introduced because of the errors in the linearization. In addition, the performance of the partition scheme largely depends on quality of relative humidity in the first guess state because the switches that initiate the warm rain processes may never be turned on during the minimization of the cost function. Moreover, as noted by Xiao et al. [19], the warm-rain partition scheme is not suitable for winter storms in which cold rain process may play an important role. Wang et al. [27] described a new scheme for the assimilation of reflectivity in which the derived rain water mixing ratio from (8) was assimilated instead of the direct assimilation of the reflectivity observations. In addition, an extra observation term was added to the cost function (5) that measures the misfit between the water vapor mixing ratios from the model and an estimated field from radar reflectivity observations. The estimated water vapor was obtained by assuming saturation where radar reflectivity is higher than a specified threshold above the cloud base. In this new scheme, the control variables for the microphysics are pseudorelative humidity and rain water mixing ratios. The partition scheme used in Xiao et al. [19] is not needed.

The performance of the new reflectivity assimilation scheme was demonstrated using an operational WRF 3D-Var DA and forecasting system with rapid update cycles by Wang et al. [27]. Four heavy-rain-producing convective cases that occurred during summer 2009 in Beijing, China, were studied using the new system. Results showed that the indirect assimilation of reflectivity significantly improved the short-term precipitation forecast skill up to 6 hours, extended from the 3 hours shown in Xiao et al. [19]. Figure 3 compares the 4-hour forecasts from two experiments that assimilate radial velocity alone and both radial velocity and reflectivity,

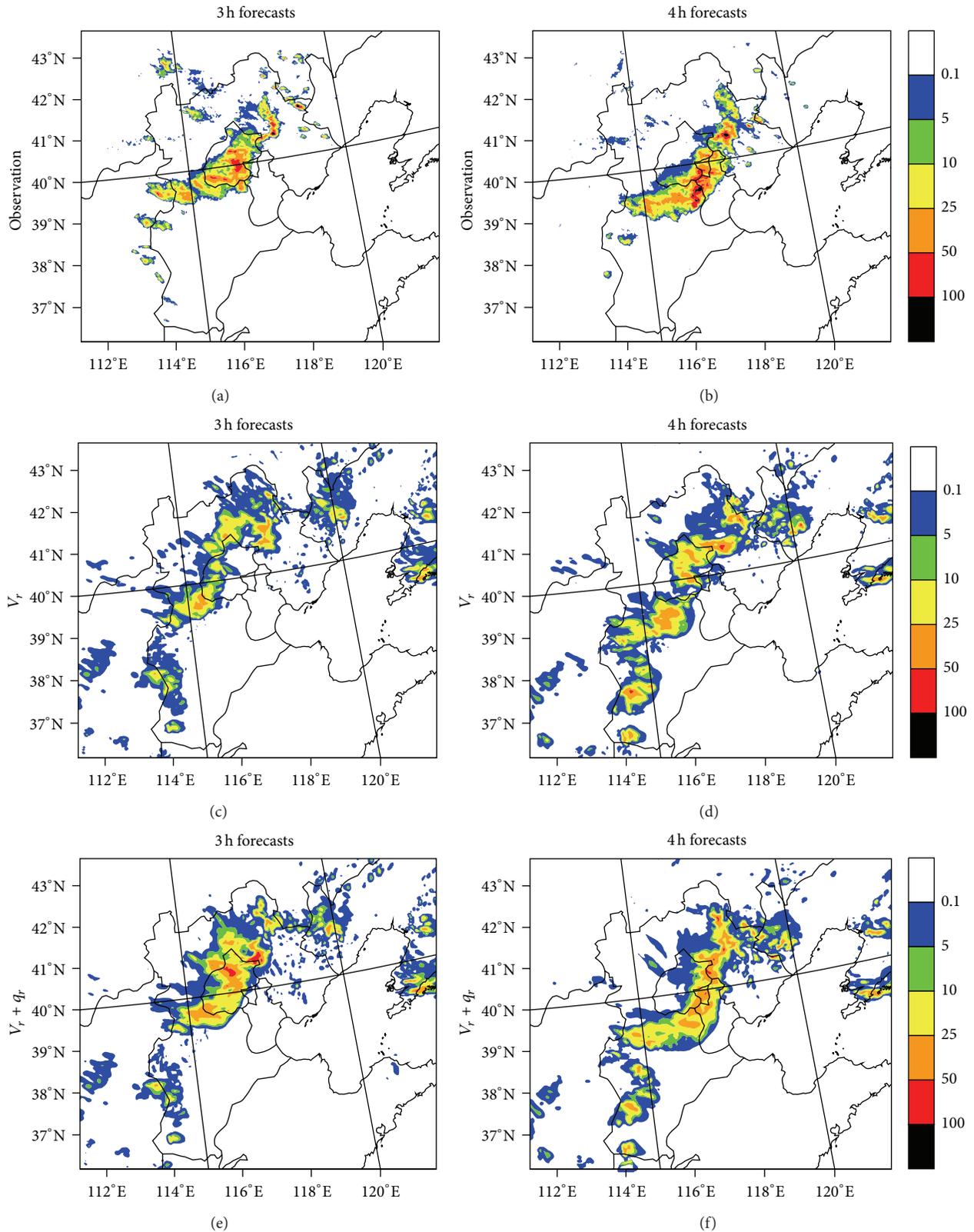


FIGURE 3: Hourly accumulated precipitation (unit: mm) for the convective case that occurred on July 23, 2009, in Beijing, China. The 3-hour (left column) and 4-hour (right column) forecasts are initialized at 0600 UTC by a 3 hourly cycled WRF 3D-Var. The radar QPE ((a) and (b)) is used to verify the experiments that assimilate radar radial velocity ((c) and (d)) only and that assimilate both radial velocity and rain water mixing ratio derived from reflectivity ((e) and (f)) (from [27]).

respectively, for one of the four cases studied in Wang et al. [27]. It is evident from this comparison that the addition of the reflectivity data assimilation improves the forecast by making the area of the precipitation more compact. The improved positive impact of the new reflectivity assimilation scheme was also confirmed by Sun et al. [34] by running the 3D-Var over a consecutive one-week period over the US Great Plains region.

All of the studies reviewed previously assimilated radar observations with one cycle or multiple 3 hourly cycles. Sun et al. [34] demonstrated that the improvement of the short-term precipitation forecast could be achieved by running the WRF 3D-Var with 3 hourly rapid update cycles even without radar observations, although adding radar data further improved the forecasts. It is desirable to run 3D-Var with more frequent cycles (less than one hour) for nowcasting applications, given the fact that radars provide frequent observations. However, initial experiments yielded mixed results with hourly cycles. Digital filter initialization [35] was used in an attempt to reduce noises that may have been the reason for the difficulty with the hourly cycles, but no significant impact was found. Some technical issues need to be resolved and will be discussed in the last section.

4. Some Early Results of the 4D-Var System

The ultimate goal of the WRFVAR radar data assimilation system is the development of a 4D-Var scheme that is capable of using the frequent radar observations at multiple time levels. In their OSSE (Observation System Simulation Experiments) study, Sugimoto et al. [36] showed that the 3D-Var had only limited ability in retrieving the tangential wind component that was not observed by radar. Using the 4D-Var of a cloud-scale model, Sun and Crook [37], Sun and Crook [38], and Crook and Sun [39] all showed that the 4D-Var technique was able to retrieve the tangential wind with reasonable accuracy. We therefore anticipate that the WRF 4D-Var can improve the 3D wind retrieval and hence the convective forecasting over its 3D-Var.

The 4D-Var technique is used in the global operational systems by several major operational weather service centers throughout the world. The European Center for Medium-Range Weather Forecasts (ECMWF) is the first center that implemented a 4D-Var operational system [40–42]. Following the success of ECMWF, other 4D-Var systems based on operational NWP models were developed and implemented at several operational centers throughout the world as their data assimilation systems for global or regional NWP without the convective-scale radar data assimilation. Recently the Met Office has run a real-time demonstration with an hourly cycling 4D-Var system and a 1.5 km resolution over a domain covering Southern England and Wales. This system currently assimilates radar radial velocity in the 4D-Var, but the reflectivity is assimilated with a diabatic initialization following Jones and Macpherson [43].

The basic framework of the WRF 4D-Var was described by Huang et al. [17] and tested using a Typhoon case. Recently, the TLM and its ADM of the system were upgraded to match

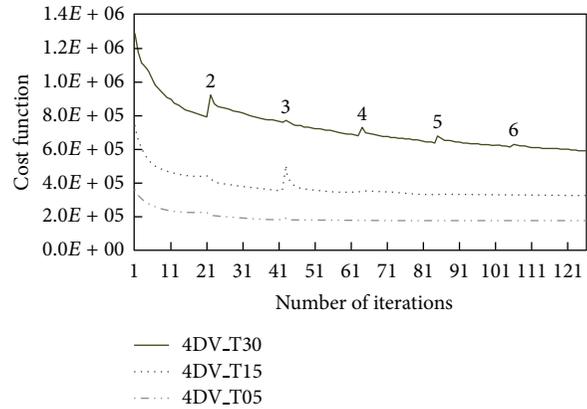


FIGURE 4: Cost functions with respect to the number of total iterations accumulated by the 6 outer loops indicated on top of the solid curve. The three cost functions are from three experiments that use the assimilation window length of 5 minutes (4DV_T05), 15 minutes (4DV_T15), and 30 minutes (4DV_T30) (from [20]).

the updated WRF version 3.3. These upgrades along with some other software enhancements were described in Zhang et al. [44]. Wang et al. [20] described the latest development of the system in regard to radar data assimilation and showed the results of a case study from a squall line precipitation event. Using the same case, Sun and Wang [21] compared the performance of the 4D-Var with that of 3D-Var and found that the precipitation forecast initialized by the 4D-Var noticeably improved over that of 3D-Var.

One question sought to answer by Wang et al. [20] was whether the incremental formulation would be feasible for the application at the convective-scale. They concluded that the incremental formulation worked well in the context of radar data assimilation with a high-resolution configuration when the typical assimilation window lengths of less than 30 minutes were used. Figure 4 shows the reduction of the cost function (5) when the window lengths of 5, 10, and 30 minutes are used. Note that the frequency of the data is the same (5 min) in all the three experiments. In these experiments, six outer loop iterations for the nonlinear base state update were used, and 20 inner loop iterations for the minimization of the cost function were performed. It is seen that the cost function is smoothly reduced when the 5-minute window was used. Some discontinuities are shown as the base state is updated in the case of 15- and 30-minute window lengths, but the cost function is steadily reduced at the end of each outer loop from the previous loop. Figure 4 also suggests that the 30-minute assimilation window results in a largest reduction of the cost function despite of the discontinuity caused by the mismatch between the linearized forward model and the nonlinear base state. Wang et al. [20] showed that the forecast initialized by the 4D-Var with the 30-minute window produced precipitation with smallest bias.

The performance of WRF 4D-Var was compared with 3D-Var by Sun and Wang [21] in a case study. They showed that the 4D-Var substantially improved the precipitation skill over the 3D-Var for the 0–6 h forecasts of the squall line

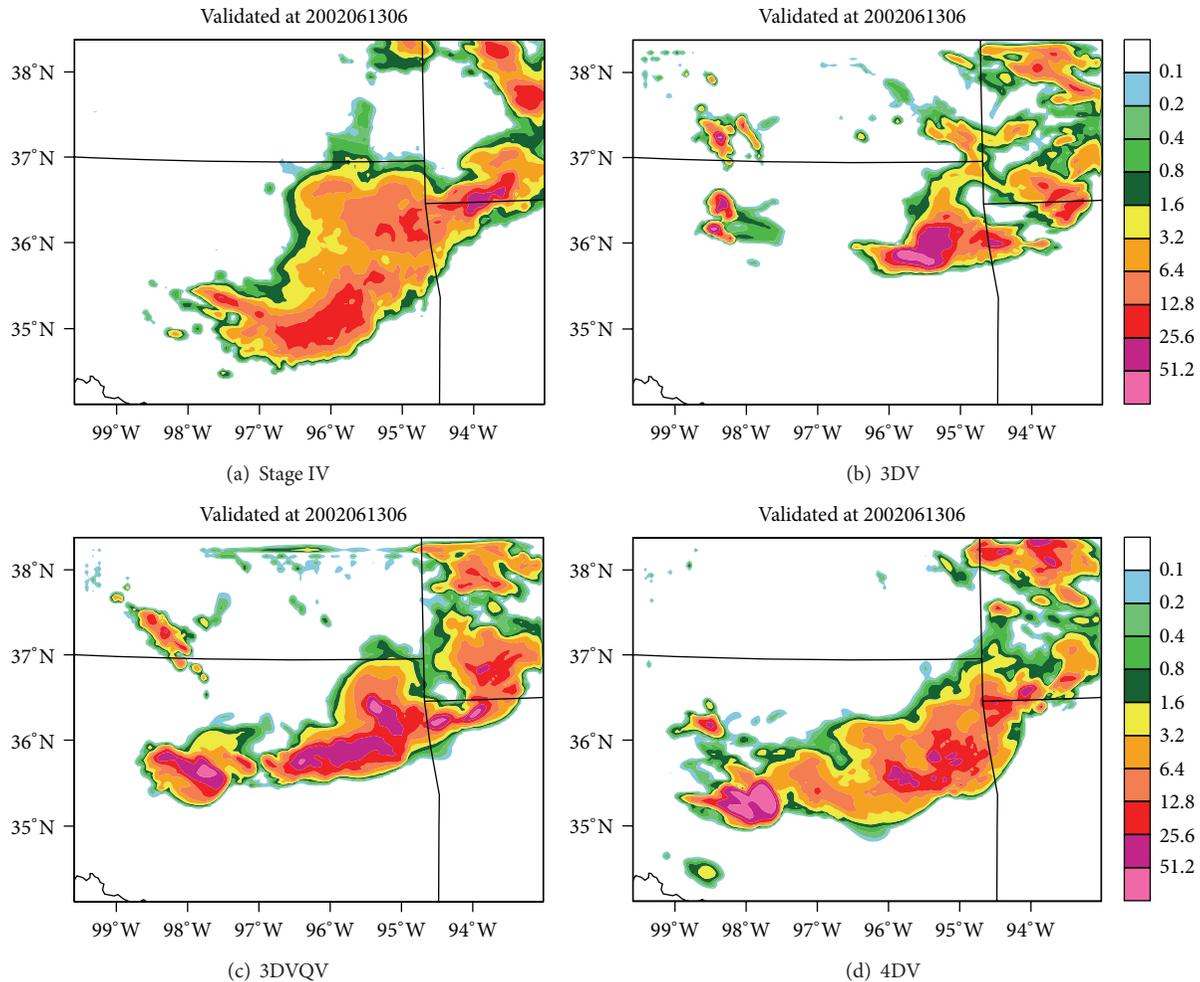


FIGURE 5: 6-hour forecasts of hourly accumulated precipitation (mm) from (b) the basic WRF 3D-Var (3DV), (c) the enhanced WRF 3D-Var (3DVQV), and (d) the 4D-Var (4DV). The Stage IV precipitation analysis is shown by (a) for verification.

studied. Figure 5 compares the 6-hour precipitation forecasts (hourly accumulation) of a 4D-Var experiment with two 3D-Var experiments from their study. The difference between the two 3D-Var experiments 3DV and 3DVQV is that the latter accounts for the effect of latent heat by assimilating saturated in-cloud humidity as described in Wang et al. [20, 27]. The improvement of the precipitation forecast from the 4D-Var over the basic 3D-Var (Figure 5(b)) is clearly shown when compared with the Stage IV data (Figure 5(a)). Although the line structure is forecasted in the experiment 3DVQV (Figure 5(c)), the squall line has slower propagation speed and the precipitation is overforecasted. It is clearly seen that the 4D-Var experiment (Figure 5(d)) has the closest agreement with the observations. The qualitative verifications of the forecasts are compared in Figure 6 using the Fractions Skill Score (FSS) [45] with the radii of influence of 8 km, 24 km, and 48 km. The superior performance of the 4D-Var is shown by the improved FSS and the elimination of the initial skill drop that is present in the 3D-Var techniques. Sun and Wang [21] also found that the radial velocity observations had greater impact on the 4D-Var analyses and

forecasts compared to the 3D-Var, suggesting that the 4D-Var has a better ability in retrieving the 3D wind as expected. Diagnostic analyses in their study suggested that the 4D-Var produced more realistic low-level cold pool, its leading edge convergence, and mid-level latent heating when compared with the radar analyses produced by VDRAS.

5. Future Developments for WRFVAR Radar DA

The encouraging results produced in the past studies warrant further development and improvement of the WRFVAR radar data assimilation system. An immediate next step is to investigate the impact of the analysis and forecast with regard to the selection of momentum control variables. One of the motivations for investigating the impact of the momentum control variables is that we have found that WRF 3D-Var tended to produce degraded results when an hourly cycle frequency was used. A new option that uses the u -wind and v -wind is being added to the WRFVAR system and compared

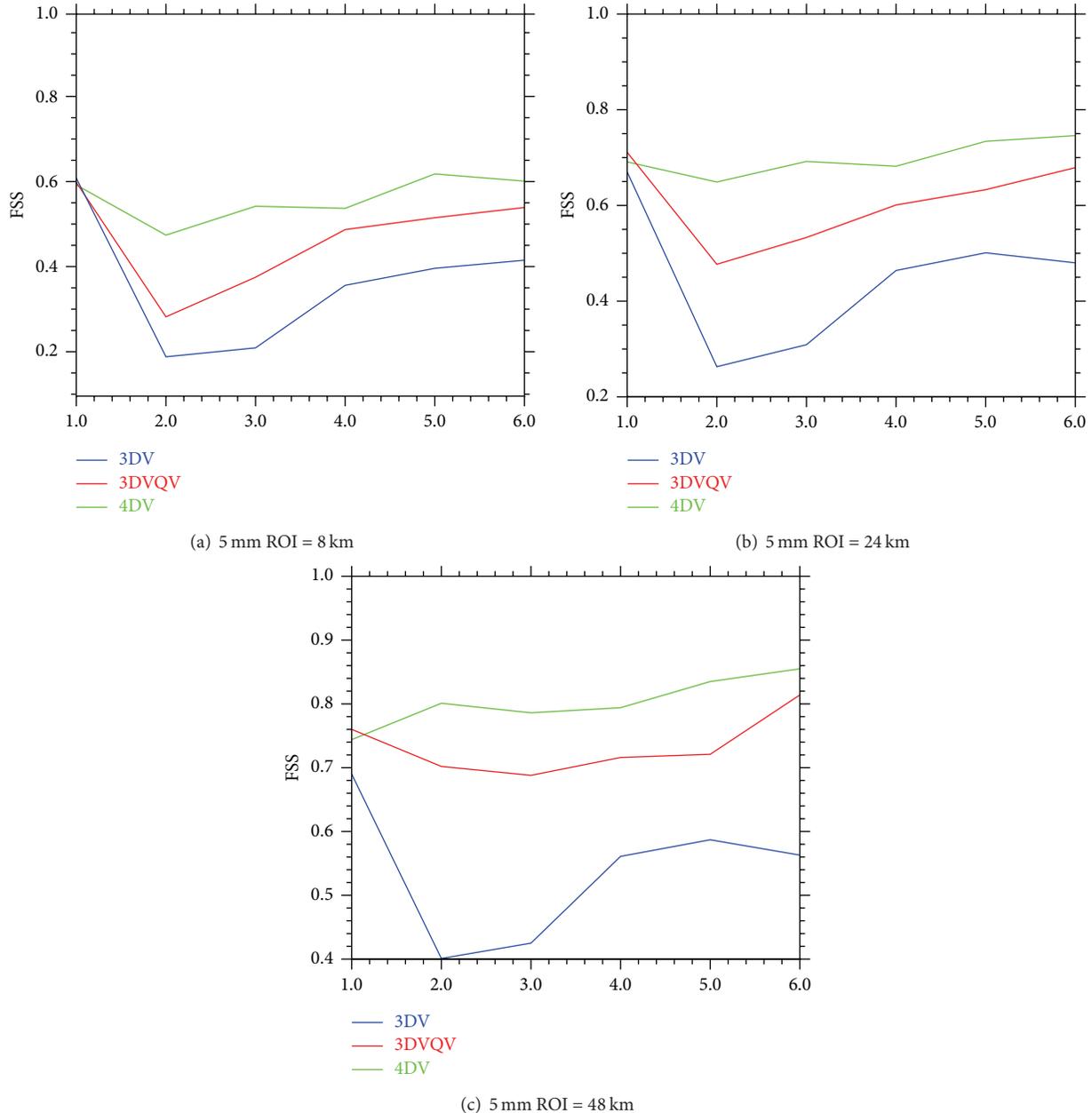


FIGURE 6: Fractions scale score (FSS) versus forecast hour of hourly accumulated precipitation (mm h^{-1}) for the thresholds 5 mm with the radius of influence (ROI) of 8 km (a), 24 km (b), and (c) 48 km, respectively.

with the existing scheme using ψ and χ momentum variables. Note that the unbalance component of the velocity potential is used in WRF 3D-Var although, for simplicity, we use the notation χ in the following description. Preliminary results indicate that the two options can yield significant differences in the characteristics of background error covariance and hence affect the final analysis. Figure 7 compares the analysis increments of u -wind and v -wind from single observation tests using the two momentum control variable options. These single observation tests were conducted by assuming an observation of u (left column) and v (right column) at the point marked by “+” in Figure 7 with an innovation and

observation error of 1 ms^{-1} . The background error covariance values for the two tests were calculated using an ensemble of short-term WRF forecasts. Apparently, the test using the ψ - χ control variables (Figures 7(a) and 7(d)) gives larger horizontal spread than the test using the u - v control variables (Figures 7(b) and 7(e)). By plotting the increments along the x -direction on the observation level (Figures 7(c) and 7(f)) it is clearly shown that the magnitudes of the wind increments are substantially different and much larger values have resulted from the use of the u - v option than the ψ - χ option. In addition, the ψ - χ control variable scheme produces negative increments in the distance far away from the observation

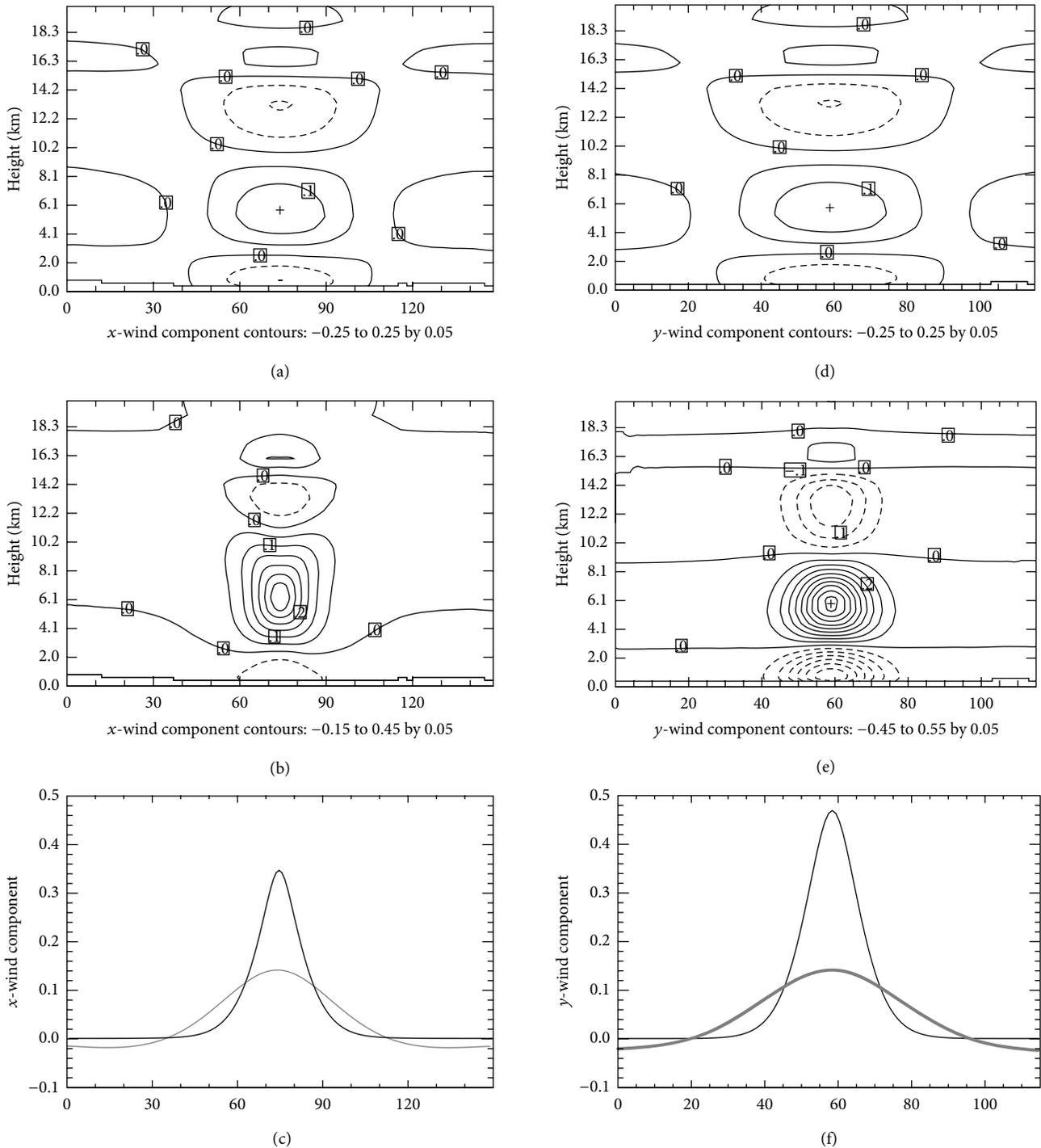


FIGURE 7: Vertical sections of analysis increments from experiments that assume a single observation of u -wind (left columns) and v -wind (right columns) with 1 ms^{-1} innovation located at the point marked by “+” in (a) and (d). The u and v increments (unit: ms^{-1}) resulting from the analyses that use ψ and χ as momentum control variables are shown in (a) and (d), respectively, and from the direct use of u and v as control variables that are shown in (b) and (e), respectively. The magnitudes of u and v increments on the observation level (6 km) are plotted as a function of x - and y -grid points. The horizontal grid spacing is 4 km.

point, which is questionable for the convective-scale. The impact of the different momentum control variables in the context of radar data assimilation for convective forecasting is being currently studied in more depth to investigate what

are the causes of the different error characteristics and is also being studied using real observations.

Although the initial results from Sun and Wang [21] are encouraging, it is obvious that the WRF 4D-Var should be

examined with more cases to confirm its superiority over the 3D-Var. One important issue that was not yet examined is the continuous cycle in WRF 4D-Var. In both Wang et al. [20] and Sun and Wang [21], only one assimilation cycle was performed. Continuous cycles should be designed and tested before the system can be implemented operationally. One issue that requires attention is how to maintain the continuity between the assimilation cycle and the short-forecast cycle in between two assimilation cycles. Difficulty may arise from the use of different microphysics schemes during the assimilation period and the short-forecast period. Another obstacle for operational implementation of 4D-Var is its high computational cost. The incremental formulation of WRF 4D-Var can be implemented with a coarser resolution forward model for the data assimilation while the high-resolution nonlinear model is used for the outer loop background run as well as the forecast run. This strategy will be tested in the near future, and its performance both in terms of computational efficiency and analysis/forecast quality will be evaluated.

Another planned future development is the use of dual-polarization observations to improve the initialization of microphysics variables. Li and Mecicalski [46] examined the impact of assimilating an improved estimate of the rain water mixing ratio through the use of differential reflectivity and specific differential phase. Some impacts were found both in the analyses and subsequent forecasts. A microphysics scheme that includes the ice phase will be developed to facilitate the assimilation and initialization of ice microphysics. Although the preliminary study by Li and Mecicalski [46] was encouraging, they raised the question how much additional information can be retrieved from the dual-polarization observations and to what extent it can improve the NWP model forecast. It is certainly an area requiring more studies.

In recent years, a great deal of research has been devoted to the development and testing of methodologies of hybrid data assimilation that make use of both variational and ensemble data assimilation techniques. Wang et al. [47] developed a hybrid ensemble 3D-Var system for the WRF model. This system was applied to a hurricane prediction, and promising results were obtained [48]. Although the hybrid concept is applicable to WRF 4D-Var, its effectiveness is yet to be tested with radar data assimilation for convective forecasting.

Acknowledgments

This work was supported by the National Science Foundation's US Weather Research Program (USWRP). NCAR is sponsored by the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- [1] J. Sun, D. W. Flicker, and D. K. Lilly, "Recovery of three-dimensional wind and temperature fields from simulated

- single-Doppler radar data," *Journal of the Atmospheric Sciences*, vol. 48, no. 6, pp. 876–890, 1991.
- [2] C. J. Qiu and Q. Xu, "A simple adjoint method of wind analysis for single-Doppler data," *Journal of Atmospheric & Oceanic Technology*, vol. 9, no. 5, pp. 588–598, 1992.
- [3] S. Laroche and I. Zawadzki, "A variational analysis method for the retrieval of three-dimensional wind field from single-Doppler data," *Journal of the Atmospheric Sciences*, vol. 51, pp. 2664–2682, 1995.
- [4] J. Gao, M. Xue, K. Brewster, and K. K. Droegemeier, "A three-dimensional variational data analysis method with recursive filter for Doppler radars," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 3, pp. 457–469, 2004.
- [5] A. Shapiro, S. Ellis, and J. Shaw, "Single-Doppler velocity retrievals with Phoenix II data: clear air and microburst wind retrievals in the planetary boundary layer," *Journal of the Atmospheric Sciences*, vol. 52, no. 9, pp. 1265–1287, 1995.
- [6] J. Sun and N. A. Crook, "Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part I: model development and simulated data experiments," *Journal of the Atmospheric Sciences*, vol. 54, no. 12, pp. 1642–1661, 1997.
- [7] J. Sun and N. A. Crook, "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," *Journal of the Atmospheric Sciences*, vol. 55, no. 5, pp. 835–852, 1998.
- [8] J. Sun, "Initialization and numerical forecasting of a supercell storm observed during STEPS," *Monthly Weather Review*, vol. 133, no. 4, pp. 793–813, 2005.
- [9] J. Sun and Y. Zhang, "Analysis and prediction of a squall line observed during IHOP using multiple WSR-88D observations," *Monthly Weather Review*, vol. 136, no. 7, pp. 2364–2388, 2008.
- [10] M. Xue, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, "The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation," *Meteorology and Atmospheric Physics*, vol. 82, no. 1–4, pp. 139–170, 2003.
- [11] Q. Zhao, J. Cook, Q. Xu, and P. R. Harasti, "Using radar wind observations to improve mesoscale numerical weather prediction," *Weather and Forecasting*, vol. 21, no. 4, pp. 502–522, 2006.
- [12] Q. Xu, L. Wei, W. Gu, J. Gong, and Q. Zhao, "A 3.5-dimensional variational method for Doppler radar data assimilation and its application to phased array radar observations," *Advances in Meteorology*, vol. 2010, Article ID 797265, 14 pages, 2010.
- [13] J. Sun, M. Chen, and Y. Wang, "A frequent-updating analysis system based on radar, surface, and mesoscale model data for the Beijing 2008 forecast demonstration project," *Weather and Forecasting*, vol. 25, no. 6, pp. 1715–1735, 2010.
- [14] M. Xue, F. Kong, K. W. Thomas et al., "CAPS realtime storm scale ensemble and high resolution forecast for the NOAA Hazardous Weather Testbed 2010 Spring Experiment," in *Proceedings of the 25th Conference on Severe Local Storms*, American Meteorological Society, 2010, Paper 7B.3.
- [15] D. M. Barker, W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. N. Xiao, "A three-dimensional variational data assimilation system for MM5: implementation and initial results," *Monthly Weather Review*, vol. 132, pp. 897–914, 2004.
- [16] D. M. Barker, X. Y. Huang, Z. Liu et al., "The Weather Research and Forecasting (WRF) model's community variational/ensemble data assimilation system: WRFDA," *Bulletin of*

- the American Meteorological Society*, vol. 93, no. 6, pp. 831–843, 2012.
- [17] X. Y. Huang, Q. Xiao, D. M. Barker et al., “Four-dimensional variational data assimilation for WRF: formulation and preliminary results,” *Monthly Weather Review*, vol. 137, no. 1, pp. 299–314, 2009.
- [18] Q. Xiao, Y. H. Kuo, J. Sun et al., “Assimilation of Doppler radar observations with a regional 3DVAR system: impact of Doppler velocities on forecasts of a heavy rainfall case,” *Journal of Applied Meteorology*, vol. 44, no. 6, pp. 768–788, 2005.
- [19] Q. Xiao, Y. H. Kuo, J. Sun, W. C. Lee, D. M. Barker, and E. Lim, “An approach of radar reflectivity data assimilation and its assessment with the Inland QPE of Typhoon Rusa (2002) at landfall,” *Journal of Applied Meteorology and Climatology*, vol. 46, no. 1, pp. 14–22, 2007.
- [20] H. Wang, J. Sun, X. Zhang, X. Huang, and T. Auligne, “Radar data assimilation with WRF 4D-Var. Part I: system development and preliminary testing,” *Monthly Weather Review*, 2013.
- [21] J. Sun and H. Wang, “Radar data assimilation with WRF 4D-Var. Part II: comparison with 3D-Var for a squall line over the U.S. Great Plains,” *Monthly Weather Review*, 2012.
- [22] P. Courtier, J. N. Thepaut, and A. Hollingsworth, “A strategy for operational implementation of 4D-Var, using an incremental approach,” *Quarterly Journal*, vol. 120, no. 519, pp. 1367–1387, 1994.
- [23] S. P. Ballard, Z. Li, D. Simonin et al., “Use of radar data in NWP-based nowcasting in the Met Office,” in *Proceedings of the International Symposium on Weather Radar and Hydrology*, R. J. Moore, S. J. Cole, and A. J. Illingworth, Eds., vol. 351, pp. 336–341, IAHS Publications, 2012.
- [24] C. M. Hayden and R. J. Purser, “Recursive filter objective analysis of meteorological fields: applications to NESDIS operational processing,” *Journal of Applied Meteorology*, vol. 34, no. 1, pp. 3–15, 1995.
- [25] R. J. Doviak and D. S. Zrnica, *Doppler Radar and Weather Observations*, Academic Press, 2nd edition, 1993.
- [26] G. Zhang, J. Sun, and E. A. Brandes, “Improving parameterization of rain microphysics with disdrometer and radar observations,” *Journal of the Atmospheric Sciences*, vol. 63, no. 4, pp. 1273–1290, 2006.
- [27] H. Wang, J. Sun, S. Fan, and X. Y. Huang, “Indirect assimilation of radar reflectivity with WRF 3D-Var and its impact on prediction of four summertime convective events,” *Journal of Applied Meteorology and Climatology*, vol. 52, pp. 889–902, 2013.
- [28] Y. Xie and A. E. MacDonald, “Selection of momentum variables for a three-dimensional variational analysis,” *Pure and Applied Geophysics*, vol. 169, pp. 335–351, 2011.
- [29] M. Fisher, “Background error covariance modelling,” in *Proceedings of the ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, pp. 45–63, ECMWF, Reading, UK, 2003.
- [30] D. F. Parrish and J. C. Derber, “The National Meteorological Center’s spectral statistical-interpolation analysis system,” *Monthly Weather Review*, vol. 120, no. 8, pp. 1747–1763, 1992.
- [31] J. Sun, “Convective-scale assimilation of radar data: progress and challenges,” *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3439–3463, 2006.
- [32] E. Lim and J. Sun, “A Velocity dealiasing technique using rapidly updated analysis from a four-dimensional variational doppler radar data assimilation system,” *Journal of Atmospheric and Oceanic Technology*, vol. 27, no. 7, pp. 1140–1152, 2010.
- [33] Z. Pu, X. Li, and J. Sun, “Impact of airborne doppler radar data assimilation on the numerical simulation of intensity changes of hurricane dennis near a landfall,” *Journal of the Atmospheric Sciences*, vol. 66, no. 11, pp. 3351–3365, 2009.
- [34] J. Sun, S. B. Trier, Q. Xiao et al., “Sensitivity of 0-12-h warm-season precipitation forecasts over the central united states to model initialization,” *Weather and Forecasting*, vol. 27, pp. 832–855, 2012.
- [35] P. Lynch and X. Y. Huang, “Initialization of the HIRLAM model using a digital filter,” *Monthly Weather Review*, vol. 120, no. 6, pp. 1019–1034, 1992.
- [36] S. Sugimoto, N. A. Crook, J. Sun, Q. Xiao, and D. M. Barker, “An examination of WRF 3DVAR radar data assimilation on its capability in retrieving unobserved variables and forecasting precipitation through observing system simulation experiments,” *Monthly Weather Review*, vol. 137, no. 11, pp. 4011–4029, 2009.
- [37] J. Sun and N. A. Crook, “Wind and thermodynamic retrieval from single-Doppler measurements of a gust front observed during Phoenix II,” *Monthly Weather Review*, vol. 122, no. 6, pp. 1075–1091, 1994.
- [38] J. Sun and N. A. Crook, “Real-time low-level wind and temperature analysis using single WSR-88D data,” *Weather and Forecasting*, vol. 16, no. 1, pp. 117–132, 2001.
- [39] N. A. Crook and J. Sun, “Assimilating radar, surface, and profiler data for the Sydney 2000 forecast demonstration project,” *Journal of Atmospheric and Oceanic Technology*, vol. 19, no. 6, pp. 888–898, 2002.
- [40] F. Rabier, H. Järvinen, E. Klinker, J. F. Mahfouf, and A. Simmons, “The ECMWF operational implementation of four-dimensional variational assimilation. I: experimental results with simplified physics,” *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 564, pp. 1143–1170, 2000.
- [41] J. F. Mahfouf and F. Rabier, “The ECMWF operational implementation of four-dimensional variational assimilation. II: experimental results with improved physics,” *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 564, pp. 1171–1190, 2000.
- [42] E. Klinker, F. Rabier, G. Kelly, and J. F. Mahfouf, “The ECMWF operational implementation of four-dimensional variational assimilation. III: experimental results and diagnostics with operational configuration,” *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 564, pp. 1191–1215, 2000.
- [43] C. D. Jones and B. Macpherson, “A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model,” *Meteorological Applications*, vol. 4, no. 3, pp. 269–277, 1997.
- [44] X. Zhang, X. Y. Huang, and N. Pan, “Development of the upgraded tangent linear and adjoint of the Weather Research and Forecasting (WRF) model,” *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 6, pp. 1180–1188, 2013.
- [45] N. M. Roberts and H. W. Lean, “Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events,” *Monthly Weather Review*, vol. 136, no. 1, pp. 78–97, 2008.
- [46] X. Li and J. R. Mecikalski, “Impact of the dual-polarization radar data on two convective storms with a warm-rain radar forward operator,” *Monthly Weather Review*, vol. 140, pp. 2147–2167, 2012.

- [47] X. Wang, D. M. Barker, C. Snyder, and T. M. Hamill, "A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: observation system simulation experiment," *Monthly Weather Review*, vol. 136, no. 12, pp. 5116–5131, 2008.
- [48] Y. Li, X. Wang, and M. Xue, "Assimilation of radar radial velocity data with the WRF ensemble-3DVAR hybrid system for the prediction of hurricane Ike," *Monthly Weather Review*, vol. 140, pp. 3507–3524, 2008.

Research Article

Using WSR-88D Polarimetric Data to Identify Bird-Contaminated Doppler Velocities

Yuan Jiang,^{1,2,3} Qin Xu,⁴ Pengfei Zhang,^{3,4} Kang Nai,³ and Liping Liu¹

¹ State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China

² Nanjing University of Information Science and Technology, Jiangsu Province 210044, China

³ Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, OK 73072, USA

⁴ National Severe Storms Laboratory, David L. Boren Boulevard, Norman, OK 73072, USA

Correspondence should be addressed to Qin Xu; qin.xu@noaa.gov

Received 6 May 2013; Revised 21 May 2013; Accepted 21 May 2013

Academic Editor: Jidong Gao

Copyright © 2013 Yuan Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As an important part of Doppler velocity data quality control for radar data assimilation and other quantitative applications, an automated technique is developed to identify and remove contaminated velocities by birds, especially migrating birds. This technique builds upon the existing hydrometeor classification algorithm (HCA) for dual-polarimetric WSR-88D radars developed at the National Severe Storms Laboratory, and it performs two steps. In the first step, the fuzzy-logic method in the HCA is simplified and used to identify biological echoes (mainly from birds and insects). In the second step, another simple fuzzy logic method is developed to detect bird echoes among the biological echoes identified in the first step and thus remove bird-contaminated velocities. The membership functions used by the fuzzy logic method in the second step are extracted from normalized histograms of differential reflectivity and differential phase for birds and insects, respectively, while the normalized histograms are constructed by polarimetric data collected during the 2012 fall migrating season and sorted for bird and insects, respectively. The performance and effectiveness of the technique are demonstrated by real-data examples.

1. Introduction

Radar echoes from migrating birds can severely contaminate Doppler velocity measurements (Jungbluth et al. [1]; Gauthreaux et al. [2]). For meteorological applications, especially quantitative applications in radar data assimilation, it is necessary to remove bird-contaminated velocities using an automated identification technique. Such a technique was developed previously before the operational WSR-88D radars in the US were upgraded with dual-polarization capability (Zhang et al. [3]; Liu et al. [4]), but the technique could only crudely detect and remove bird-contaminated velocities volume-wise or tiltwise (for each sweep) because the usable input data from operational WSR-88D radars were limited to reflectivity and velocity measurements. A major drawback of this previous technique is that it rejects the

entire volume (or tilt) of velocity observations even if only a fraction of the volume (or tilt) is contaminated by migrating birds. Now, almost all the operational WSR-88D radars in the US are upgraded with dual-polarization capability, which is essential for discriminating between meteorological and nonmeteorological scatterers and for distinguishing different hydrometeor types. In the hydrometeor classification algorithm (HCA) developed for polarimetric radars at the National Severe Storms Laboratory (NSSL) (Straka and Zrnić [5]; Zrnić and Ryzhkov [6]; Zrnić et al. [7]; Schuur et al. [8]; Ryzhkov et al. [9]; Park et al. [10]), birds and insects are not distinguished and labeled as biological scatterers entirely. This HCA is designed to remove biological scatterer contamination from meteorological scatterers for the purpose of quantitative precipitation estimation (QPE). It is thus timely and desirable to develop an improved technique that

takes advantage of polarimetric measurements to provide a pixel-wise identification and removal of bird-contaminated velocities. As a continuation of the previous efforts of Zhang et al. [11, 12], this paper strives to develop such a desired technique by leveraging the NSSL existing HCA.

The most recent version of the HCA (Park et al. [10]) utilizes all available polarimetric variables to discern ten different classes of radar echoes with eight classes for various hydrometeors, one class for ground clutter (GC, including that due to anomalous propagation) and one class for biological scatterers (BS). This classification was designed mainly for using radar observations to improve the quantitative precipitation estimation. The GC class is useful for both reflectivity and velocity data quality controls, but the BS class is directly useful only for reflectivity data quality control. For radar velocity data quality control, it is necessary to differentiate insects from birds, because insects can be treated as passive tracers of air motions and thus are useful for wind measurements in most cases, while birds and especially migrating birds (mostly flying at night during the migrating seasons) can cause significant biases (to the order of 10 m s^{-1}) in radar measured Doppler velocities (Gauthreaux et al. [2]; Collins [13]; Bi et al. [14]). It has been known that the polarimetric signatures of insects and birds are different due to the shape and size differences between insects and birds (Riley [15]; Vaughn [16]; Zrnić and Ryzhkov [17]). In general, birds have higher differential phase and lower differential reflectivity than insects (Zhang et al. [11, 12]). These properties will be used in this paper to develop an automated technique for pixel-wise identification and removal of bird-contaminated velocities in two steps. In the first step, the fuzzy logic in the HCA is simplified to identify 3 classes: BS, GC, and MS, where MS represents meteorological scatterers and is essentially a combined category containing all meteorological hydrometeor categories from the original algorithm (see Section 2.1 of Schuur et al. [8]). In the second step, a new simple fuzzy logic method is developed and used to divide the BS class into two subclasses: (i) BS due to birds (including bats) especially migrating birds (BSb) and (ii) BS due to insects (BSi). The next section describes the simplified HCA developed for the first step. Section 3 presents the fuzzy logic method developed for the second step. The performance and effectiveness of the two-step method are demonstrated by real-data examples in Section 4. Conclusions follow in Section 5.

2. Simplified HCA for Identifying BS in the First Step

As explained in the introduction, the ten classes of radar echoes in the original HCA are consolidated into three classes using a simple fuzzy logic method that focuses on identifying BS, GC, and MS. Among the six polarimetric variables used by the original HCA, the simplified fuzzy logic method uses the following five variables: (i) the radar reflectivity Z (at horizontal polarization), (ii) the differential reflectivity Z_{DR} , (iii) the cross-correlation coefficient ρ_{hv} between horizontally and vertically polarized radar returns, (iv) the texture parameter

$\text{SD}(Z)$ of Z , and (v) the texture parameter $\text{SD}(\Phi_{\text{DP}})$ of the differential phase Φ_{DP} . The specific differential phase K_{DP} is used in the original HCA for identifying different types of hydrometeors, so this variable is not used here.

As described in Park et al. [10], the parameter $\text{SD}(Z)$ [or $\text{SD}(\Phi_{\text{DP}})$] characterizes the magnitude of the small-scale fluctuations of Z (or Φ_{DP}) along the radial. It is estimated by smoothing Z (or Φ_{DP}) data along the radial using a 1 km (or 2 km) running-average window, subtracting the smoothed Z (or Φ_{DP}) from the original values and calculating the root-mean-square (RMS) value of the residuals. Before applying the classification procedure, Z , Z_{DR} , and ρ_{hv} are smoothed along each radial using a 1 km averaging window for Z and a 2 km window for Z_{DR} and ρ_{hv} . In addition, the Doppler velocity v_r is used to flag possible GC prior to the fuzzy logic classification procedure. In particular, if $v_r < 1 \text{ m s}^{-1}$ (or $\geq 1 \text{ m s}^{-1}$) for a pixel, then this pixel will be (or not be) checked for GC in the fuzzy logic classification procedure. To account for attenuation in precipitation, the biases of Z and Z_{DR} are estimated using heavily filtered Φ_{DP} for WSR-88D radars (Bringi et al. [18]):

$$\begin{aligned}\Delta Z^{\text{att}} (\text{dBZ}) &= 0.04\Phi_{\text{DP}} (^{\circ}), \\ \Delta Z_{\text{DR}}^{\text{att}} (\text{dB}) &= 0.004\Phi_{\text{DP}} (^{\circ}).\end{aligned}\quad (1)$$

From the above parameter values, the simplified fuzzy logic method computes the aggregation value A_i for the i th class of radar echoes at each pixel by

$$A_i = \frac{\sum_j W_{ij} P^{(i)}(V_j)}{\sum_j W_{ij}}, \quad (2)$$

where $P^{(i)}(V_j)$ is the membership function that characterizes the distribution of the j th variable for the i th class and W_{ij} is a weight between 0 and 1 assigned to the i th class and the j th variable. The classification decision is based on the maximal aggregation value. The formulation in (2) is similar to that in (3) of Park et al. [10] except that the confidence vector assigned to each variable is now simply set to 1.

Each membership function $P^{(i)}(V_j)$ in (2) has a trapezoidal shape described by four parameters: x_1 , x_2 , x_3 , and x_4 . This trapezoidal function increases linearly from 0 to 1 as V_j increases from x_1 to x_2 , remains to be 1 until V_j increases x_3 , and then decreases linearly from 1 to 0 as V_j further increases from x_3 to x_4 (see Figure 1 of Schuur et al. [8] or Park et al. [10]). The parameter values are listed in Table 1 for each of the 5×3 membership functions, where f_1 and f_2 are functions of Z (in dBZ) given by (4) of Park et al. [10] or originally by (7) and (8) of Schuur et al. [8] in the following forms:

$$f_1 = -0.50 + 2.50 \times 10^{-3} Z + 7.50 \times 10^{-4} Z^2, \quad (3)$$

$$f_2 = 0.08 + 3.64 \times 10^{-2} Z + 3.57 \times 10^{-4} Z^2. \quad (4)$$

Here, the five membership functions for GC are largely the same as those in Table 1 of Park et al. [10] or originally Table 1 of Schuur et al. [8] except for the following modifications according to our observations in the presence of GC: (i) the

TABLE 1: Parameter values of the membership functions used by the simplified HCA for the three classes. Reflectivity-dependent parameters f_1 and f_2 are defined in (3) and (4), respectively.

	GC	BS	RA
		$P(Z)$ (dBZ)	
x_1	5	5	5
x_2	20	10	10
x_3	70	20	65
x_4	80	30	75
		$P(Z_{DR})$ (dB)	
x_1	-3	0	$f_1 - 0.3$
x_2	-2	2	f_1
x_3	1	10	f_2
x_4	2	12	$f_2 + 0.3$
		$P(\rho_{hv})$	
x_1	0.5	0.3	0.85
x_2	0.8	0.5	0.97
x_3	0.9	0.8	1
x_4	0.95	1.01	1.05
		$P[SD(Z)]$ (dBZ)	
x_1	2	1	0
x_2	4	2	0.5
x_3	10	4	3
x_4	15	7	6
		$P[SD(\Phi_{DP})]$ (°)	
x_1	30	8	0
x_2	40	10	1
x_3	50	40	15
x_4	60	60	30

lower-limit value x_1 for Z is lowered from 15 to 5 dBZ, (ii) the lower-limit value x_1 for Z_{DR} is increased from -4 to -3 dB, and (iii) the lower-shoulder value x_2 for ρ_{hv} is increased from 0.6 to 0.8 to reduce the possibility of misclassifying BS with ρ_{hv} between 0.5 to 0.8 as GC. The five membership functions for BS are mostly the same as those in Table 1 of Park et al. [10] or Table 1 of Schuur et al. [8], except that the upper-limit value x_4 for ρ_{hv} is increased from 0.83 to 1.01, because it has been found that ρ_{hv} occasionally reaches to 1.01 for BS in the operational polarimetric radar observations. The five membership functions for MS are mostly the same as those in Table 1 of Schuur et al. [8] except that the upper-limit value x_4 for ρ_{hv} is increased from 1.01 to 1.05 because incorrect values of $\rho_{hv} > 1$ and even almost up to 1.05 have been seen from the polarimetric data so far collected in the presence of MS, and these incorrect values of $\rho_{hv} (>1)$ are caused by excessive noise corrections (see (5) of Schuur et al. [8]).

The weights of the 5×3 membership functions are listed in Table 2. The weights of the five membership functions for GC are largely the same as those in Table 2 Park et al. [10] except for the following modifications: (i) the weight assigned to Z is increased from 0.2 to 0.4 to enhance and ensure the GC detection, (ii) the weight assigned to ρ_{hv} is decreased from 1.0 to 0.4, and (iii) the weight assigned to $SD(Z)$ is fine tuned slightly from 0.6 to 0.5. The weights of the five membership

TABLE 2: Specified weights for the 5×3 membership functions in the simplified HCA.

Class	Variables				
	Z	Z_{DR}	ρ_{hv}	$SD(Z)$	$SD(\Phi_{DP})$
GC/AP	0.4	0.4	0.4	0.5	0.8
BS	0.4	0.6	1	0.8	0.8
RA	1.0	1.0	0.6	0.2	0.2

TABLE 3: Typical ranges of Z_{DR} , Φ_{DP} , and ρ_{hv} and height for BSb, BSi, and MS.

	Z_{DR}	Φ_{DP}	ρ_{hv}	Height
BSb	<3 dB	70–100°	<0.97	Mostly below melting layer
BSi	≥2 dB	40°	<0.97	Below melting layer and mainly in boundary layer
MS	-2~6 dB		>0.97	

functions for BS are the same as those in Table 2 of Park et al. [10]. The weights of the five membership functions for MS are the same as the maximum weights among the eight classes of hydrometeors in Table 2 of Park et al. [10].

3. Fuzzy-Logic Method for Identifying BSb in the Second Step

3.1. “Ground Truth” Data for Birds versus Insects. In the second step, the BS class in the HCA is divided into two subclasses: (i) BS due to birds especially migrating birds (BSb) and (ii) BS due mainly to insects (BSi). The goal is to develop a fuzzy logic method for further differentiating BSb from BSi for each BS pixel identified by the simplified HCA (described in Section 2). As reviewed in the introduction, insects and birds have different polarimetric signatures due to their differences in shape and size. In particular, birds have higher Φ_{DP} and lower Z_{DR} than insects. These signatures are further quantified in Table 3 in comparison with those for MS. As an example, Figure 1 shows the color images of Z_{DR} , Φ_{DP} , and ρ_{hv} from the KVNx 0.5° scan at 0102 UTC (local time 8 pm) on 27 October 2011. In this case, the following two types of pixels are readily identified by the simplified HCA: (i) BS pixels within the 120 km radial range around the radar and (ii) MS pixels far beyond 120 km radial range to the northwest of the radar. The identified BS pixels have relatively low Z_{DR} , high Φ_{DP} , and low ρ_{hv} , so they should be further classified as BSb pixels. The identified MS pixels reveal a precipitation system to the northwest of the radar.

To develop a fuzzy-logic method to differentiate BSb from BSi, it is necessary to find proper membership functions that characterize the distributions of Φ_{DP} and Z_{DR} for BSb versus BSi within the BS class based on the probability density function estimated from normalized histograms of Φ_{DP} and Z_{DR} for the two subclasses. To construct these normalized histograms, it is necessary to collect large numbers of “ground truth” data for BSb versus BSi in the presence of BS, and this is done by the following selection procedure. First, it has long been recognized and well observed that BS echoes are caused

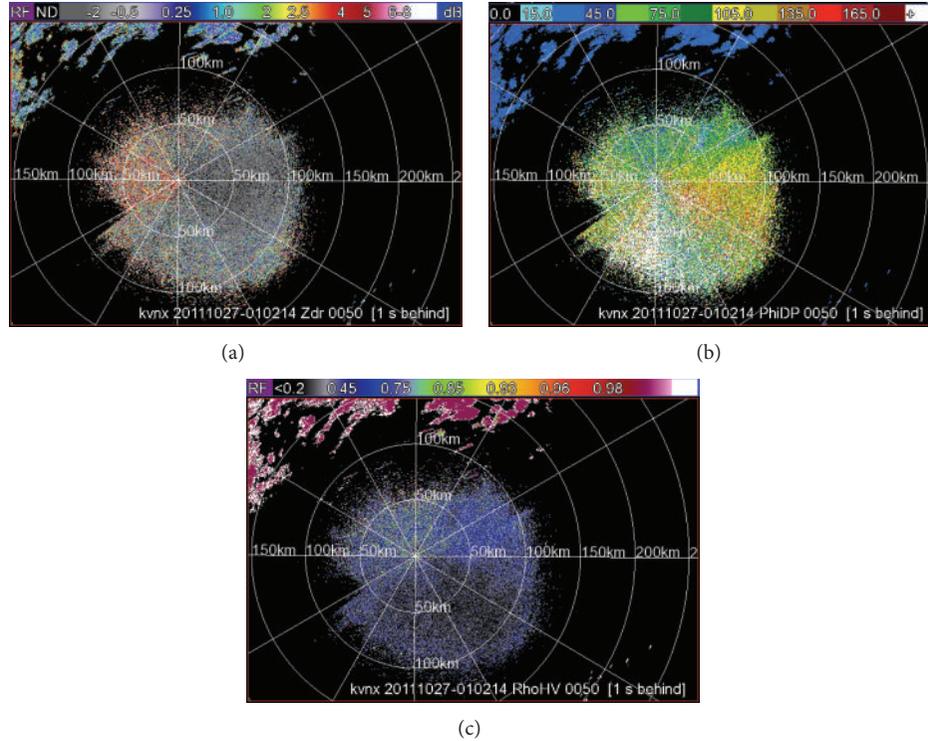


FIGURE 1: Imagers of (a) Z_{DR} , (b) Φ_{DP} , and (c) ρ_{HV} from the KVNx 0.5° scan at 0102 UTC (local time 8 pm) on 27 October 2011. The spatial resolutions of the original data are 250 m in the radial direction and 0.5° in the azimuthal direction.

dominantly by migrating birds during the nighttime but by insects during the daytime in the spring and fall migrating seasons, and the transition from birds to insects (or from insects back to birds) occurs around the sunrise time (or sunset time) almost every nonstormy day during a migrating season, especially over the southern great plain in the central United States (Zhang et al. [11]). Based on this well-observed fact (Dingle [19]; Drake and Gatehouse [20]), most BS pixels identified from nighttime (or daytime) radar scans during migrating seasons by the simplified HCA can be attributed to BSb (or BSi). Thus, we can sample midnight (or midday) BS data tilt-by-tilt and treat each sampled tilt of data as proxy “ground-truth” data for BSb (or BSi). Then, by visually examining color images of the sampled polarimetric data on each tilt, we can manually pick specific areas for BSb and BSi where the polarimetric signatures are distinct in Φ_{DP} and Z_{DR} (with $\rho_{HV} < 0.97$) for BSb and BSi, respectively, according to Table 3. This completes the selection procedure.

As an example, Figures 2(a) and 2(b) show the Z_{DR} - Φ_{DP} scattergrams for BS pixels identified by the simplified HCA from the KVNx 0.5° scans at 0706 UTC (local midnight) and 1901 UTC (local midday), respectively, on 17 October 2011. As shown, the midnight Z_{DR} - Φ_{DP} scattergram in Figure 2(a) has relatively low Z_{DR} and relatively high Φ_{DP} , and these are the expected characters for BSb. On the other hand, the midday Z_{DR} - Φ_{DP} scattergram in Figure 2(b) has relatively high Z_{DR} and relatively low Φ_{DP} , and these are the expected characters for BSi caused by insects. Thus, by plotting and examining

their Z_{DR} - Φ_{DP} scattergrams, we can further check and ensure the quality of the “ground truth” data.

3.2. Scattergrams and Normalized Histograms for Migrating Birds and Insects. To quantify and differentiate the polarimetric signatures between insects and birds, especially migrating birds, we started to monitor and sample real-time operational polarimetric observations from the operational KVNx radar since September 2011 (shortly after this radar was upgraded with dual-polarization capability, the first among the 159 operational WSR-88D radars in the US) and from the operational KICT radar since September 2012 (shortly after this radar was upgraded with dual-polarization capability). After polarimetric data are sampled from these radars during the fall and spring migrating seasons and processed by using the simplified HCA to identify BS pixels, “ground-truth” data are gathered for BSb (or BSi) through the two selection steps described in Section 3.1. The “ground-truth” data gathered from KICT and KVNx radars for the 2012 fall migrating season have better qualities than those gathered early from KVNx radar for the 2011 fall migrating season (before KVNx radar was fully calibrated), so the scattergrams constructed from the “ground-truth” data gathered from KICT radar for the 2012 fall migrating season will be used to extract the membership functions for BSb and BSi in this subsection.

The red (or blue) dots in Figure 3 plot the Z_{DR} - Φ_{DP} scattergram for BSb (or BSi) constructed by 1692359 (or 1385090)

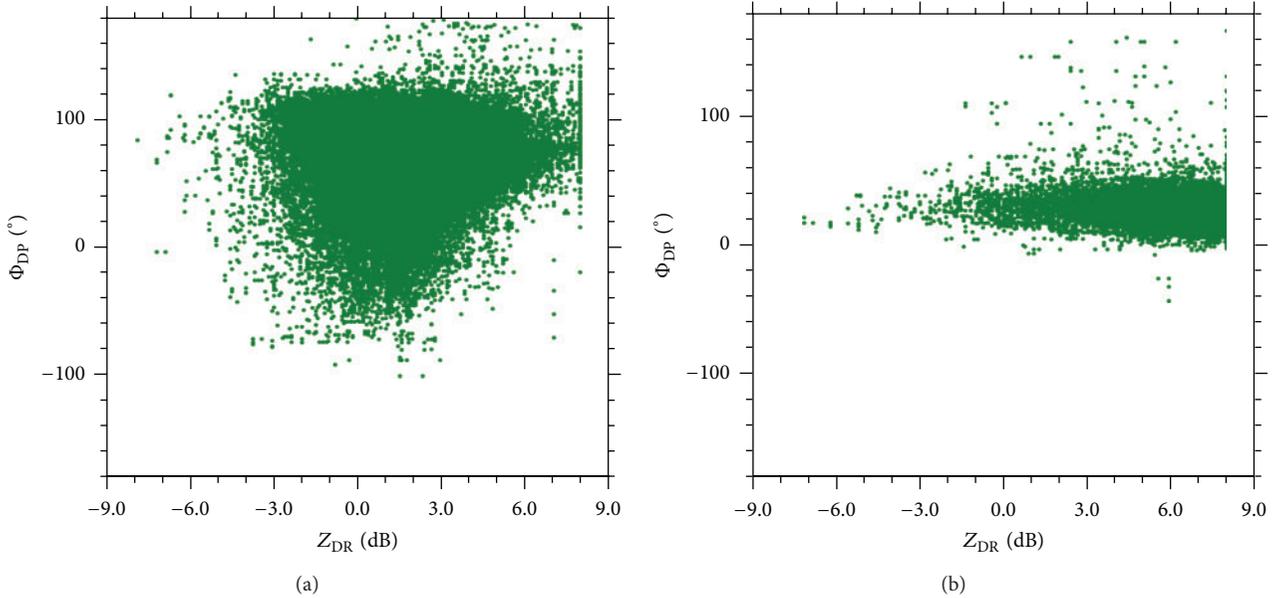


FIGURE 2: Z_{DR} - Φ_{DP} scattergrams for BS pixels identified by the simplified HCA from the KVNx 0.5° scans at (a) 0706 UTC (local midnight) and (b) 1901 UCT (local midday), respectively, on 17 October 2011.

pixels of “ground-truth” data gathered from the KICT observations during the nighttime (or daytime) on 2 October 2012. Here, again as expected, the scattergram for BSb (or BSi) has relatively low (or high) Z_{DR} and relatively high (or low) Φ_{DP} , and the main differences between the two scattergrams for BSb and BSi are similar to those showed in Figure 3 of Zrnić and Ryzhkov [17], although the core area of their scattergram for BSb is relatively narrow (roughly within $-2 \text{ dB} \leq Z_{DR} \leq 3 \text{ dB}$) and the core area of their scattergram for BSi is not bounded by $Z_{DR} = 8 \text{ dB}$. Note that the input Φ_{DP} is obtained by subtracting the radar system Φ_{DP} from the radar measured Φ_{DP} , but the radar system Φ_{DP} can be estimated and recorded only through a rainy event. The true system Φ_{DP} often drifts slowly away from the recorded value, but the latter will be still used as the true system Φ_{DP} until it is estimated and recorded through the next rainy event. The slow drift of the true system Φ_{DP} can cause uncertainties in the input Φ_{DP} (obtained by subtracting the recorded system Φ_{DP} from the radar measured Φ_{DP}). For the scattergrams in Figure 3, the subtracted system Φ_{DP} is 14.81° . Scattergrams constructed from “ground-truth” data gathered from the KICT observations for other days during the 2012 fall migrating season are similar to those in Figure 3 for BSb and BSi, respectively, but the core areas often shift to slightly different Φ_{DP} ranges due to the above explained uncertainties in the input Φ_{DP} .

As we can see from Figure 3, the two scattergrams for BSb and BSi are not completely separated. Nevertheless, their overlapped area is small which may imply a small degree of true inseparability between BSb and BSi. The normalized histograms of Z_{DR} (or Φ_{DP}) that estimate the probability density functions of Z_{DR} (or Φ_{DP}) are extracted from the scattergrams in Figure 3 and plotted for BSb by the solid curve and for BSi by the dashed curve in Figures 4(a) (or 4(b)).

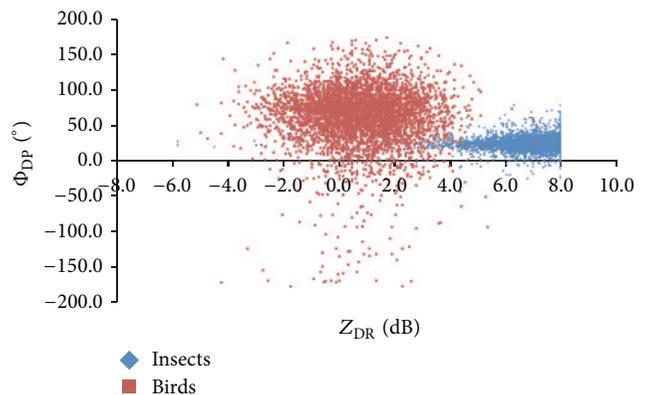


FIGURE 3: Z_{DR} - Φ_{DP} scattergrams for BSb (plotted by red dots) and BSi (plotted by blue dots) constructed by 1692359 and 1385090 pixels of “ground-truth” data gathered from the KICT observations during the nighttime and daytime, respectively, on 2 October 2012.

As shown in Figure 4(a), BSb is dominant and the probability for BSi is almost zero when $-4 \text{ dB} \leq Z_{DR} < 4 \text{ dB}$, whereas BSi is dominant and the probability for BSb is almost zero when $Z_{DR} > 4 \text{ dB}$. As shown in Figure 4(b), BSb is dominant and the probability for BSi is nearly zero when $\Phi_{DP} > 50^\circ$, whereas BSi is dominant and the probability for BSb is much smaller than that for BSi when $\Phi_{DP} < 35^\circ$, so the situation becomes ambiguous (with no clear dominance in either BSb or BSi) only when $30^\circ < \Phi_{DP} < 50^\circ$. Normalized histograms of ρ_{hv} are also plotted and examined in pairs for BSb and BSi (not shown), and the dominant range of ρ_{hv} for BSb is found to be slightly lower than that for BSi. Since ρ_{hv} has a significantly overlapped range between BSb and BSi, it will not be used as an input variable for the fuzzy logic

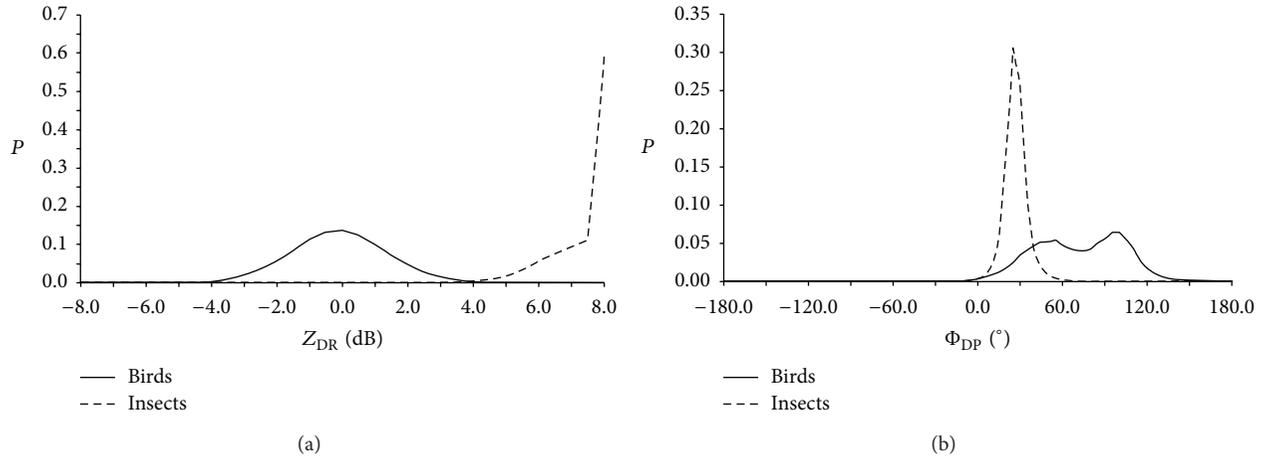


FIGURE 4: Normalized histograms of (a) Z_{DR} and (b) Φ_{DP} exacted from the scattergrams in Figure 3 and plotted by the solid curve for BSb and by the dashed curve for BSi.

method in the next subsection. Normalized histograms of other variables, such as Z , v_r , $SD(Z)$, $SD(\Phi_{DP})$, and $SD(v_r)$, exhibit largely overlapped ranges between BSb and BSi (not shown), so they are not very useful here. Thus, only Z_{DR} and Φ_{DP} will be used as input variables for the fuzzy logic method in the next subsection.

3.3. Fuzzy-Logic Method for Identifying Migrating Birds. The fuzzy logic formulation for identifying migrating birds is similar to that in (2) but is further simplified by using only 2 membership functions to characterize the distributions of the two variables Z_{DR} and Φ_{DP} for the subclass BSb. Based on the normalized histogram of Z_{DR} for BSb in Figure 4(a), the four parameters x_1 , x_2 , x_3 , and x_4 are set to -5 , -3 , 2 , and 4 dB, respectively, for the membership functions of Z_{DR} for BSb. Based on the normalized histogram of Φ_{DP} for BSb in Figure 4(b), the four parameters x_1 , x_2 , x_3 , and x_4 are set to 0° , 40° , 120° and 150° , respectively, for the membership function of Φ_{DP} for BSb. The weights are set to 1 and 0.8 for the membership functions of Z_{DR} and Φ_{DP} , respectively. Because we only need to identify BSb (versus BSi) among BS, it is convenient and sufficient to compute a single aggregation value A_i in (2) for $i = \text{BSb}$ only and then check the computed A_{BSb} with a properly tuned threshold value (which is 0.3 in this case). Thus, if A_{BSb} is larger than this threshold value for a BS pixel, then this pixel is identified as a BSb pixel.

Since the method is developed as a part of the radar velocity data quality control for operational data assimilation applications (Liu et al. 2009 [21]), the goal is to detect and remove all pixels contaminated by migrating birds. To achieve this goal, the threshold is purposely tightened to a relatively low value (0.3) to enhance the detection of BSb. In this case, some BSi pixels may be also removed as they are incorrectly identified as “BSb” pixels due to the tightened threshold, and this is a price paid for the enhanced detection and removal of BSb to ensure that the processed and accepted velocity data are free of bird contaminations. Thus, the method is designed and tasked to detect and remove all true BSb pixels rather

than to identify and retain all true BSi pixels, and this task is implemented effectively by considering only the subclass BSb in the fuzzy logic formulation.

4. Effectiveness of the Method

The two-step method developed in this paper has been tested with polarimetric data collected from the operational KVNx and KICT radars during the 2011 fall and 2012 spring and fall migrating seasons. The results indicate that the method is as effective as, but more efficient than, the original HCA for identifying GC and BS in the first step. The method is also demonstrated to be effective and efficient for identifying BSb in the second step. Examples are presented below to show the effectiveness of this two-step method.

4.1. Effectiveness of the Simplified HCA. As mentioned in Section 2, the simplified HCA is designed to identify and discriminate BS from GC and MS. An example is given in this subsection to illustrate that this simplified HCA is effective for this designed purpose in comparison with the original HCA. Figures 5(a) and 5(b) show the images of reflectivity Z and radial velocity v_r , respectively, at the lowest tilt (0.5°) scanned by the operational KVNx radar at 013037 UTC on 28 October 2011. The local time was 8:30 pm. As shown, the radar echoes were weak over most of the area within 100 km radial range. Z_{DR} and Φ_{DP} images (not show) indicate that these weak echoes are BS from migrating birds. They were also GC identified empirically (with $|v_r| < 1 \text{ m s}^{-1}$ as shown in Figure 5(b)) around $r = 50 \text{ km}$ to the south-southeast and west-northwest of the radar. Around $r = 120 \text{ km}$ to the north of radar, there was an isolated area of weak echo that lasted only for a few volume scans (not shown) around sunrise and sunset, and these weak echoes were BS of local birds, as identified by human expertise. The scattered areas with relatively high reflectivity (around 20 dBZ) outside $r = 150 \text{ km}$ to the southeast of radar are clearly MS, again as identified with human expertise. Figure 5(c) shows the classification

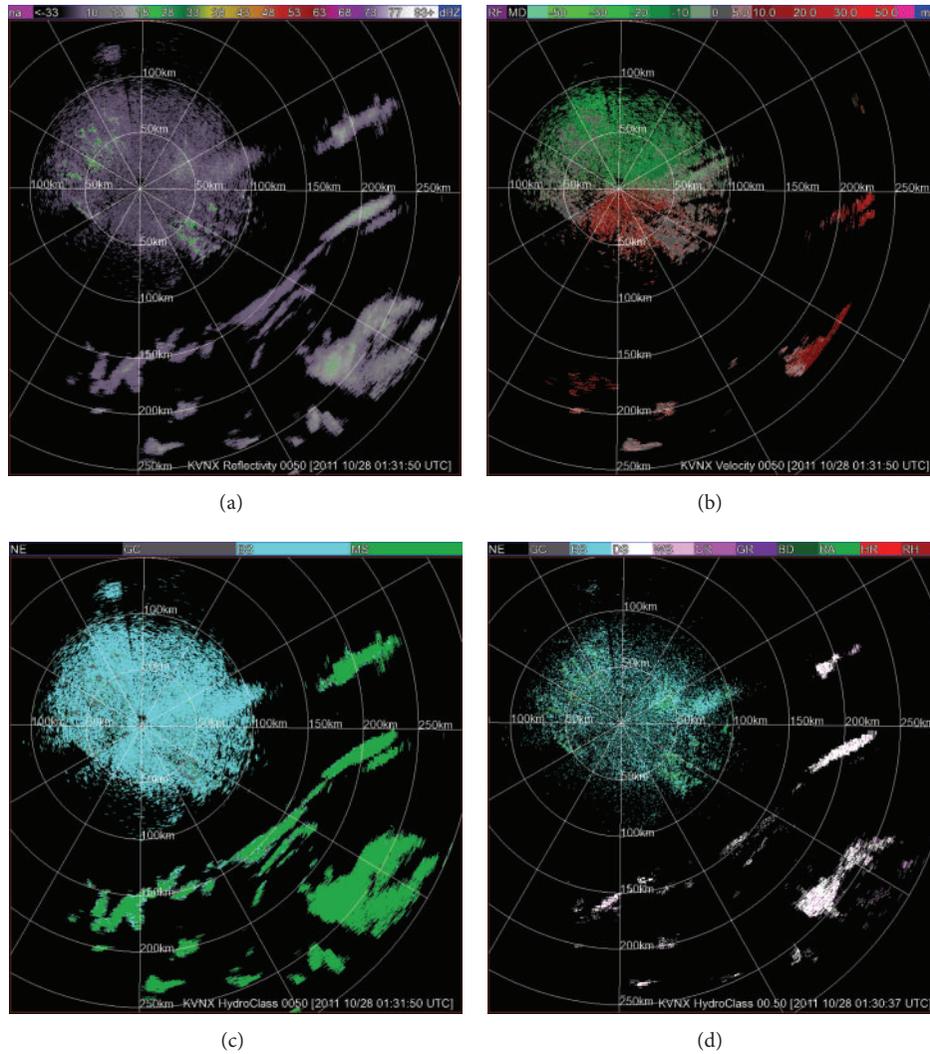


FIGURE 5: Images of (a) Z , (b) v_r , (c) classification results produced by the simplified HCA, and (d) classification results produced by the original HCA from the KVNIX 0.5° scan at 013150 UTC on 28 October 2011. The 10 classes in the original HCA are: (1) ground clutter including that due to anomalous propagation (GC/AP); (2) biological scatterers (BS); (3) dry aggregated snow (DS); (4) wet snow (WS); (5) crystals of various orientations (CR); (6) graupel (GR); (7) “big drops” (BD); (8) light and moderate rain (RA); (9) heavy rain (HR); and (10) a mixture of rain and hail (RH). The spatial resolutions of the original data are the same as in Figure 1.

results from the simplified HCA. Here, the BS, GC, and MS are colored in cyan, gray, and green, respectively. The results are consistent with those identified above based on human expertise. The classification results from the original HCA are plotted in Figure 5(d), where the BS and GC are again in cyan and gray, respectively, and the eight different classes of MS are in other different colors. By comparing Figure 5(c) with Figure 5(d), we can see that the BS and GC areas identified by the simplified HCA cover the BS and GC areas identified by the original HCA, respectively, and the MS areas identified by the simplified HCA cover the areas of hydrometeors identified by the original HCA. Thus, the BS pixels identified by the simplified HCA in the first step can be used to further identify BSb pixels in the second step.

4.2. Effectiveness of the Fuzz-Logic Method in the Second Step

4.2.1. Effectiveness for Identifying BSb. After BS pixels are identified by the simplified HCA in the first step, the fuzz-logic method described in Section 3.3 is applied to discriminate BSb from BSi among the BS pixels. An example is given in this subsection to illustrate the effectiveness of the method for identifying and removing BSb contaminations. Figures 6(a)–6(e) show the images of reflectivity Z , radial velocity v_r , cross-correlation coefficient ρ_{hv} , differential reflectivity Z_{DR} , and differential phase Φ_{DP} , respectively, at the lowest tilt (0.5°) scanned by the operational KICT radar at 090142 UTC on 30 September 2012. The local time was around 4 am during the fall migrating season, and the sky over the 150 km radial range from the KICT radar was

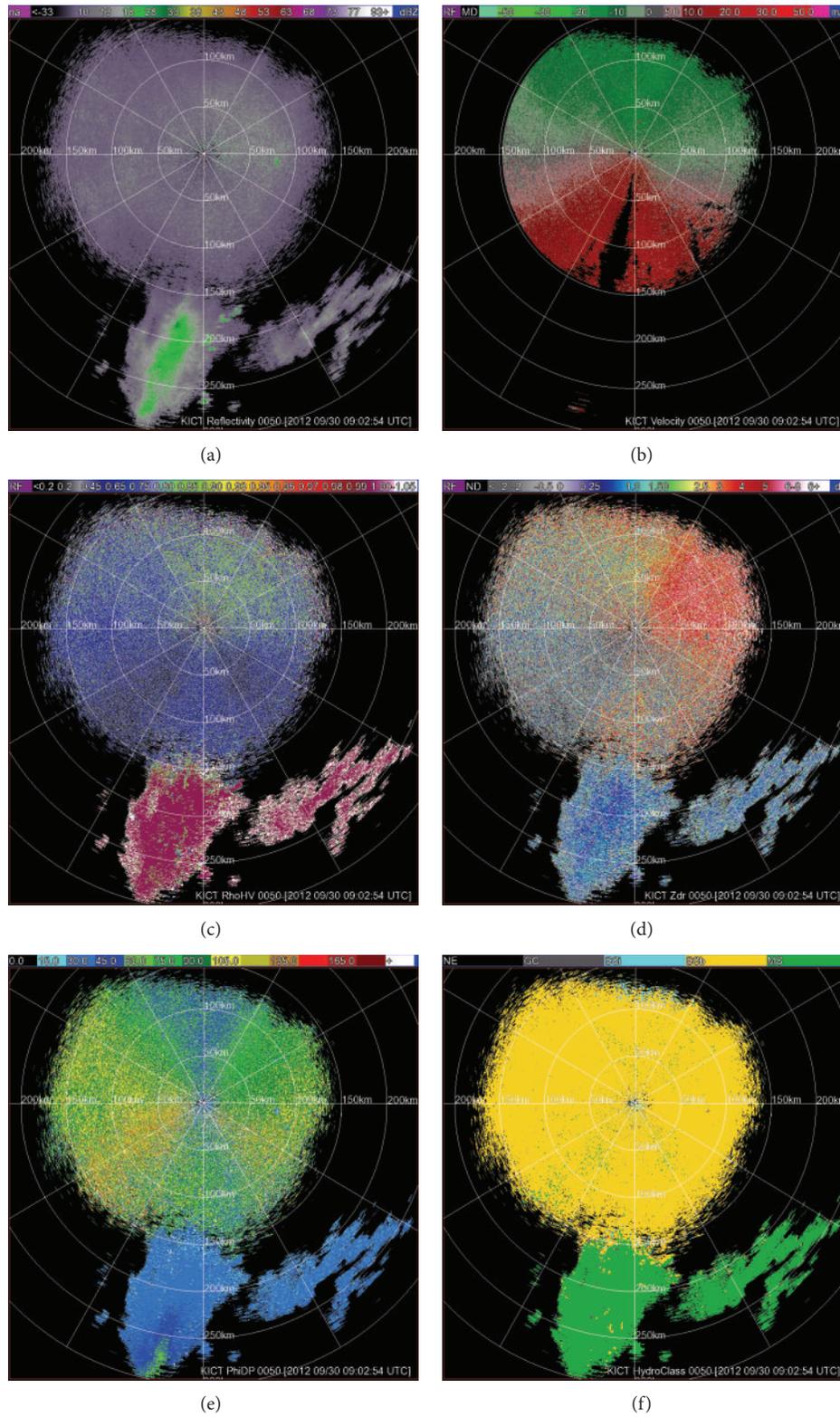


FIGURE 6: Images of (a) Z , (b) v_r , (c) ρ_{HV} , (d) Z_{DR} , (e) Φ_{DP} , and (f) final classification results produced in the second step from the KICT 0.5° scan at 090142 UTC (local time around 4 am) on 30 September 2012. The spatial resolutions of the original data are the same as in Figure 1.

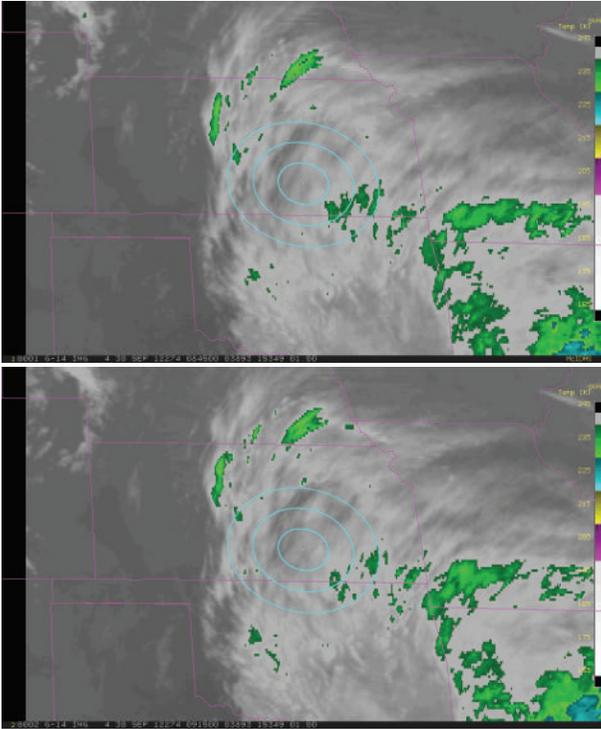


FIGURE 7: Infrared images from GOES-14 at 084500 UTC (upper panel) and 091500 UTC (lower panel) on 30 September 2012 over Kansas State and its neighboring states. The state boundaries are plotted in purple. The three cyan circles plot the 50, 100, and 200 km range rings cocentered at the KICT radar site (shown by the cyan dot). The dark areas show relatively warm brightness temperature remotely sensed from the ground, indicating clear sky over each dark area. The green areas highlight cloud tops colder than 240°K ($\approx -30^{\circ}\text{C}$ around 7 km height).

covered by elongated rapidly-moving patches of thin clouds aloft with no precipitation or, at least, no significant precipitation based on GOES-14 infrared observations as shown in Figure 7. The presence of migrating birds is apparent and prevailing as judged by human expertise from the radar reflectivity image in Figure 6(a) and the GOES-14 infrared image in Figure 7. In particular, as shown in Figure 6(a), the gray colored reflectivity (<18 dBZ) over the broad area within 150 km radial range, which is identified as BS by the simplified HCA in the first step, was caused by migrating birds, whereas the green colored reflectivity (>18 dBZ) beyond 150 km to the south of radar, which is identified as MS in the first step, was caused by precipitation. Figure 6(c) shows that the ρ_{hv} values (mostly around 1) in the MS area are larger than the ρ_{hv} values (mostly below 0.7) in the broad BSb area. The Z_{DR} field in Figure 6(d) and Φ_{DP} field in Figure 6(e) are smoother and more uniform in the MS area than in the BSb area. Moreover, the Z_{DR} values are mostly below 3 dB and the Φ_{DP} values are larger than 60° in the broad BSb area. These polarimetric features further confirm that the BS pixels identified in the first step are indeed BSb pixels rather than BSi pixels. After MS and BS pixels are identified by the simplified HCA in the first step, BSb pixels are further

identified from the BS pixels by the fuzz-logic method in the second step. The final results are shown in Figure 6(f), where the MS pixels are in green and the BSb pixels are in yellow. These results are consistent with those identified above by human expertise. This exemplifies the effectiveness of the method for identifying BSb and removing BSb-contaminated velocities.

From the classification results in Figure 6(f), we can also see that the method is not perfect and its performance is affected by noises in polarimetric measurements. For example, ideally, all the pixels should be identified as MS in the green-dominated areas beyond the 175 km radial range to the south and southeast of the radar in Figure 6(f), but the method identifies 37562 MS pixels among the total of 39452 pixels in these green-dominated area, so the hit rate is 95.21% slightly below the perfect 100%. The remaining pixels are 4.79% of the total, and they are identified as GC, BSi, and BSb in 0.68%, 0.25%, and 3.86%, respectively. Thus, the correctly retained (MS and BSi) pixels are 95.46% of the total, and the incorrectly rejected (GC and BSb) pixels are merely 4.54% of the total. Similarly, all the pixels should be identified as BSb in the yellow-dominated area within the 150 km radial range in Figure 6(f), but the method identifies 372021 BSb pixels among the total of 392649 pixels in this yellow-dominated area, so the hit rate is 94.75%. The remaining pixels are 5.25% of the total, and they are identified as GC, BSi, and MS in 0.56%, 1.95%, and 2.74%, respectively. Thus, the correctly retained (MS and BSi) pixels are merely 4.69% of the total. These retained (MS and BSi) pixels are scattered and mostly isolated from each other, so they can be easily removed by a simple continuity check after the classification.

Figure 8 shows the vertical profiles of absolute value and azimuthal angle for the following two types of horizontal velocities: (i) the horizontal velocity of air plus birds (plotted by the solid profiles) produced by the velocity azimuth display (VAD) analysis (which is a by-product of the VAD-based dealiasing of Xu et al. [22]) from the same volume of velocity data as that displayed (on 0.5° tilt only) in Figure 6(b) but within the 150 km radial range to cover BSb pixels only and (ii) the horizontal velocity of air (plotted by the dashed profiles) above the KICT radar site at 090142 UTC interpolated from the hourly analyzed wind fields produced by the operational Rapid Refresh (RAP) model (see <http://rapidrefresh.noaa.gov/>) on 30 September 2012. Subtracting the second velocity from the first gives the differential velocity caused by migrating birds in the BSb area. The absolute value and direction of the differential velocity are plotted by the dotted vertical profiles in Figure 8. As shown, when the height increases from 250 to 800 m, the absolute value of the differential velocity increases from 4.5 to 6.3 m s^{-1} and the azimuthal angle changes slightly from 190° to 176° (that is, within $\pm 10^{\circ}$ of the southward direction). This result is consistent with the well-recognized fact that migrating birds fly southward at speeds of $5\text{--}10\text{ m s}^{-1}$ nighttime during the fall migrating season over the southern great plain in central United States. The differential velocities diagnosed in Figure 8 are also in the range of the previously

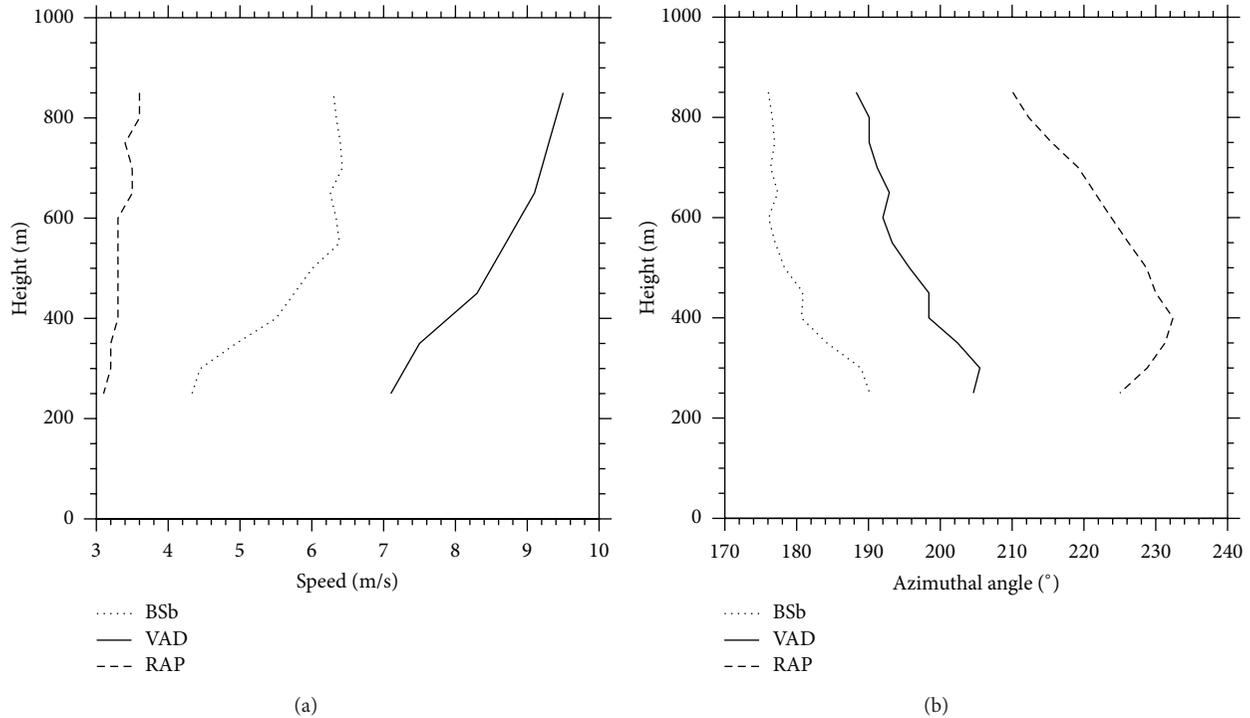


FIGURE 8: Vertical profiles of (a) absolute value (m s^{-1}) and (b) azimuthal angle (clockwise from the north) for two types of horizontal velocities: (i) the horizontal velocity of air plus birds (plotted by the solid profiles) produced by the VAD analysis from the KICT-scanned volumetric velocity data at 090142 UTC but within the 150 km radial range to cover BSb pixels only and (ii) the horizontal velocity of air (plotted by the dashed profiles) interpolated from RAP hourly wind analyses to the KICT radar location at 090142 UTC on 30 September 2012. The absolute value and azimuthal angle of the vector difference of the above two velocities (the first minus the second velocity) are plotted by the dotted vertical profiles in panels (a) and (b), respectively.

documented VAD wind biases caused by migrating birds (Jungbluth et al. [1]; Gauthreaux et al. [2]; Collins [13]). Clearly, the differential velocities caused by migrating birds are not small. It is thus necessary to identify BSb pixels and remove them from velocity data, and this task can be performed effectively by the method in the second step.

4.2.2. Effectiveness for Identifying BSi. The method is also effective for identifying and retaining BSi pixels in the second step, and this is illustrated by the example in Figure 9 where the images were from the daytime KICT 0.5° scan at 215544 UTC (4 pm local time) on 30 September 2012. As shown in Figure 9(a), most pixels within the 150 km radial range are gray colored with $Z < 15$ dBZ. As these gray-colored low-reflectivity pixels are already classified as BS by the simplified HCA in the first step, they can be further identified as BSi pixels judged by human expertise based on the time of the day (4 pm local time). Furthermore, as we can see from Figures 9(c)–9(e), on these low-reflectivity BS pixels, the ρ_{hv} values are mostly below 0.97 (but occasionally become unrealistically larger than 1 and even reach 1.05 for the reason explained in Section 2), the Z_{DR} values are mostly above 6 dB, and the Φ_{DP} values are mostly below 60° . These polarimetric features further confirm that the low-reflectivity BS pixels can be identified as BSi pixels according to Table 3. From Figure 9(a), we can also see scattered areas of green-colored

pixels (with $Z > 15$ dBZ) near the 100 km range to the north of radar, around the 70 km range to the northeast of radar, and inside the 50 km range to the northwest of the radar. These scattered areas ($Z > 15$ dBZ) are also already classified as BS by the simplified HCA in the first step. In these scattered areas (with $Z > 15$ dBZ), again as shown in Figures 9(b)–9(e), the observed radial velocities are highly nonuniform with large-amplitude irregular fluctuations, the ρ_{hv} values are mostly below 0.7, the Z_{DR} values are mostly below 3 dB, and the Φ_{DP} values are mostly above 60° . These features indicate that the BS pixels in the scattered areas of $Z > 15$ dBZ can be further identified as BSb pixels according to Table 3, but these BSb pixels indicate the presence of local nesting birds rather than migrating birds as identified by human expertise based on the local time (around 4 pm) and the highly fluctuated radial velocities in these scattered areas ($Z > 15$ dBZ). The BSb and BSi areas identified above by human expertise are also well identified automatically by the method in the second step as shown in Figure 9(f) by the yellow and cyan arrears, respectively.

Figure 10 shows the vertical profiles of absolute value and azimuthal angle for (i) the VAD velocity (plotted by the solid profiles) computed from the same volume of velocity data as that displayed (on 0.5° tilt only) in Figure 9(b) but with all the BSb pixels removed, (ii) the interpolated velocity from RAP hourly wind analyses (plotted by the dashed profiles) at

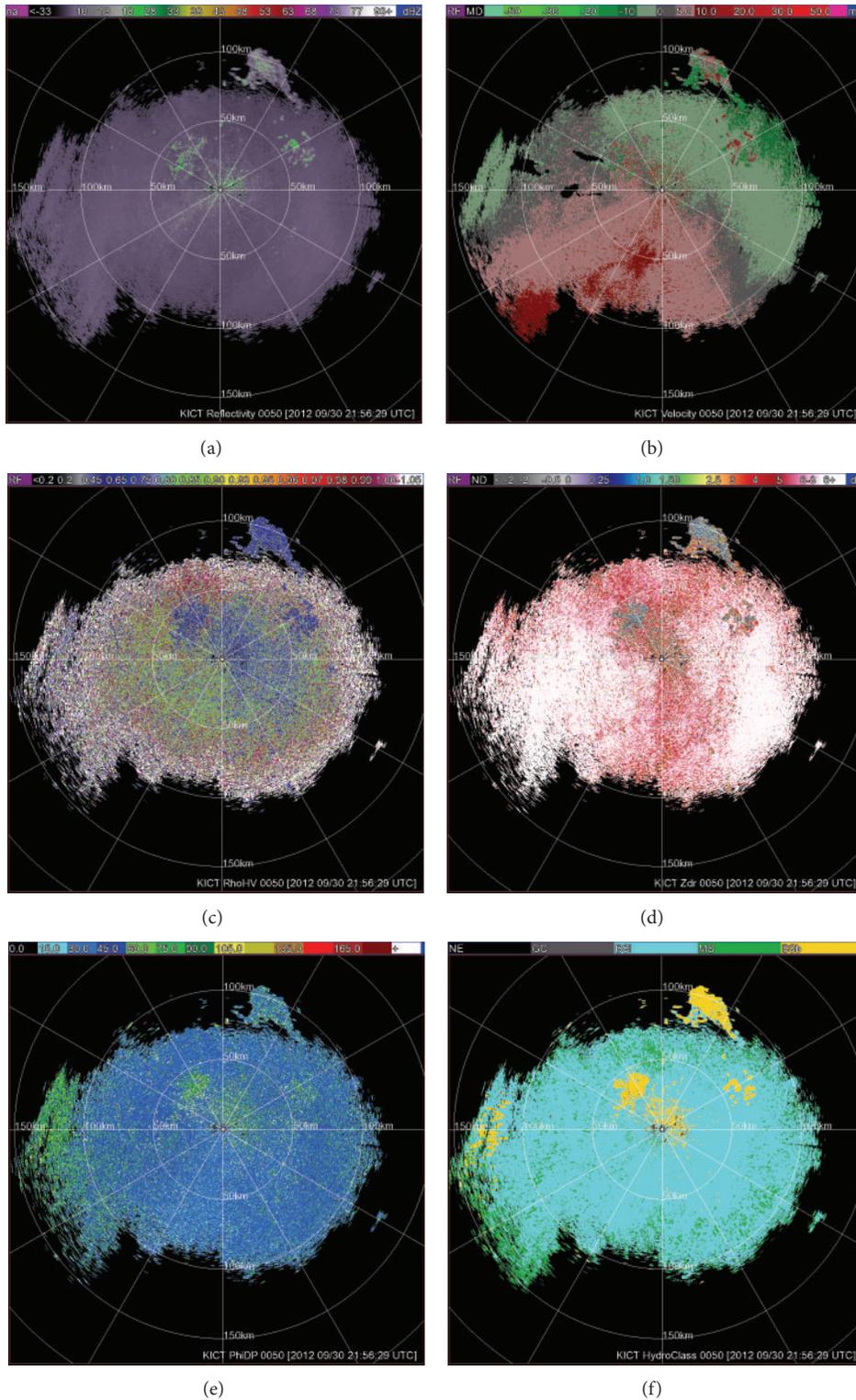


FIGURE 9: As in Figure 6 but from the KICT 0.5° scan at 215542 UTC (local time around 4 pm) on 30 September 2012.

215542 UTC on 30 September 2012, and (iii) the differential velocity (plotted by the dotted vertical profiles) obtained by subtracting the second velocity from the first. As shown, the differential velocity is small, and its absolute value

decreases from 3.4 to 1.6 m s⁻¹ and its azimuthal angle is nearly constant around 150~155° when the height increases from 250 to 800 m. These diagnosed differential velocities could be due to the RAP hourly wind analysis error plus

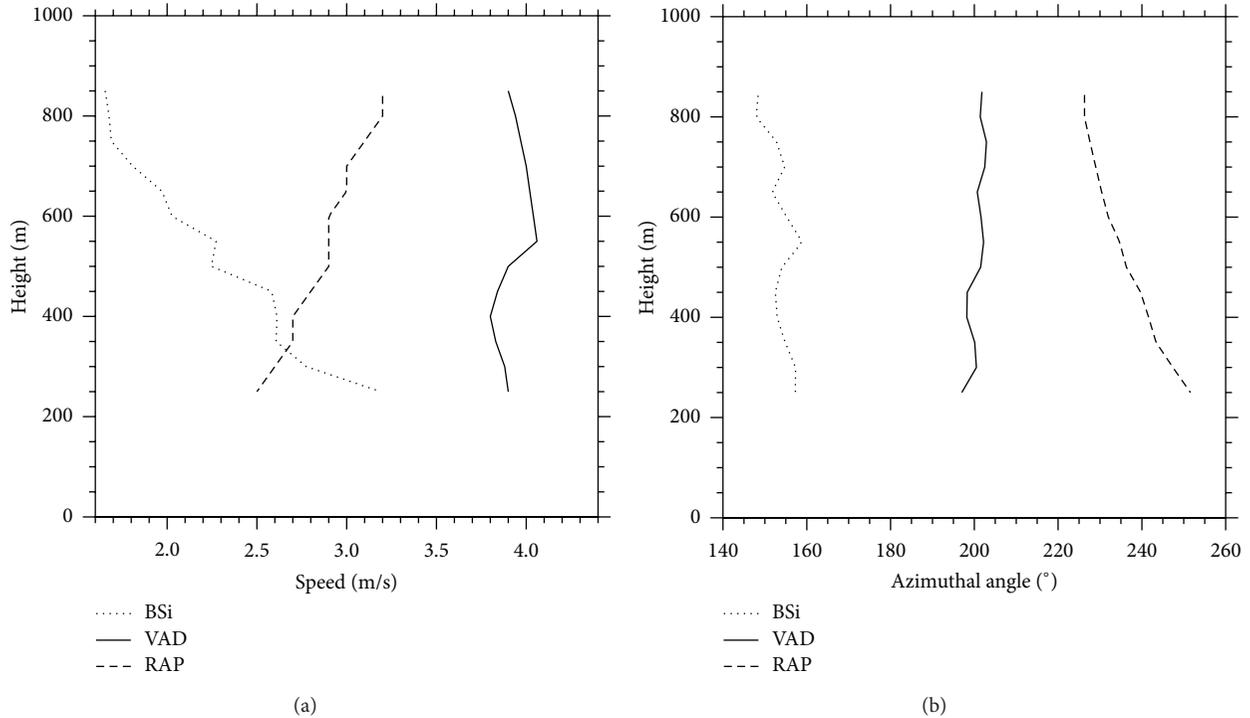


FIGURE 10: As in Figure 8 except that the VAD velocity (plotted by the solid profiles) is produced from the KICT-scanned volumetric velocity data at 215542 UTC with all the BSi pixels removed and the interpolated velocity from RAP hourly wind analyses (plotted by the dashed profiles) is also at 215542 UTC on 28 October 2011.

insects' flying speeds (Drake and Gatehouse [20]). This further confirms that BSi pixels are correctly identified by the method in the second step. As the differential velocities caused by insects are small, the observed radial velocities on BSi pixels are retained by the method for data assimilation applications.

5. Conclusions

In this paper, a two-step method is developed to identify and remove contaminated velocities by birds, especially migrating birds, in addition to those contaminated by ground clutter (GC, including that due to anomalous propagation). In the first step, the existing hydrometeor classification algorithm (HCA) developed for polarimetric radars at the NSSL (Zrnić et al. [7]; Schuur et al. [8]; Ryzhkov et al. [9]; Park et al. [10]) is simplified to identify three classes of radar echoes: (i) GC, (ii) biological scatterers (BS), and (iii) meteorological scatterers (MS). In the second step, a fuzzy-logic method is developed and used to further identify scatterers of birds (including bats) especially migrating birds (BSb) versus scatterers of insects (BSi) among BS. The simplified HCA in the first step uses five input polarimetric variables to identify BS, GC, and MS, and the associated membership functions (listed in Table 1) are condensed from those used in the original HCA. Each BS pixel identified in the first step is used as a starting point to further identify BSb versus BSi on that BS pixel in the second step. The fuzzy-logic method in the second

step uses two input polarimetric variables, that is, Z_{DR} and Φ_{DP} , and the associated membership functions are extracted from normalized histograms of Z_{DR} and Φ_{DP} that estimate the probability density functions of Z_{DR} and Φ_{DP} , respectively, while the normalized histograms are constructed from "ground truth" data selected empirically by human expertise (as described and exemplified in Section 3.1).

The method has been tested with polarimetric data collected from the operational KVNx and KICT radars during the 2011 fall and 2012 spring and fall migrating seasons. The simplified HCA in the first step is found to be as effective as the original HCA, and the effectiveness is shown by the example in Section 4.1. This simplified HCA has been incorporated into the operational radar reflectivity data quality control package (Liu et al. [21]) for radar data assimilation application at the national centers for environmental prediction (NCEP). The fuzzy-logic method in the second step is also found to be effective in further identifying BSb versus BSi for each BS pixel identified in the first step, and the effectiveness is shown by the examples in Section 4.2. The two-step method has been continuously applied to real-time polarimetric data from the operational KVNx and KICT radars with the classification results verified indirectly by comparing the interpolated RAP wind analyses with the two types of VAD velocities produced from volumetric velocity data (i) on BSb pixels only and (ii) on non-BSb pixels only (as illustrated by the examples in Figures 8 and 10). Through this real-time run, more and better "ground truth" data will be accumulated and used to further improve the method in the second step. The improved

method will be incorporated into the operational radar velocity data quality control package (Liu et al. [21]) for radar data assimilation application at NCEP.

Acknowledgments

The authors are thankful to Bob Rabin at NSSL for kindly providing the GOES-14 infrared images in Figure 7, Terry Schuur at NSSL, and anonymous reviewers for their comments and suggestions that improved the representation of the results. The research was supported by the ONR Grant N000141010778 to the University of Oklahoma (OU). Funding was also provided by NOAA/OAR under NOAA-OU Cooperative Agreement NA17RJ1227 U.S. Department of Commerce, by the China's National Natural Science Foundation Grant 41175038, by the National Key Program for Developing Basic Sciences Grant 2012CB417202, and by the Chinese Academy of Meteorological Sciences Basic Scientific and Operational Project "Observation and retrieval methods of micro-physics and dynamic parameters of cloud and precipitation with multi-wavelength Remote Sensing."

References

- [1] K. Jungbluth, J. Belles, and M. Schumacher, "Velocity contamination of WSR-88D and wind profiler data due to migrating birds," in *Proceedings of the 27th Conference on Radar Meteorology*, pp. 666–668, American Meteorological Society, Vail, Colo, USA, October 1995.
- [2] S. A. Gauthreaux Jr., D. S. Mizrahi, and C. G. Belser, "Bird migration and bias of WSR-88D wind estimates," *Weather and Forecasting*, vol. 13, no. 2, pp. 465–481, 1998.
- [3] P. Zhang, S. Liu, and Q. Xu, "Identifying Doppler velocity contamination caused by migrating birds—part I: feature extraction and quantification," *Journal of Atmospheric and Oceanic Technology*, vol. 22, no. 8, pp. 1105–1113, 2005.
- [4] S. Liu, Q. Xu, and P. Zhang, "Identifying Doppler velocity contamination caused by migrating birds—part II: bayes identification and probability tests," *Journal of Atmospheric and Oceanic Technology*, vol. 22, no. 8, pp. 1114–1121, 2005.
- [5] J. M. Straka and D. S. Zrnic, "Algorithm to deduce hydrometeor types and contents from multi-parameter radar data," in *Proceedings of the 26th International Conference on Radar Meteorology*, pp. 513–515, American Meteorological Society, Norman, Okla, USA, May 1993.
- [6] D. Zrnic and A. Ryzhkov, "Polarimetry for weather surveillance radars," *Bulletin of the American Meteorological Society*, vol. 80, no. 3, pp. 389–406, 1999.
- [7] D. S. Zrnic, A. Ryzhkov, J. Straka, Y. Liu, and J. Vivekanandan, "Testing a procedure for automatic classification of hydrometeor types," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 6, pp. 892–913, 2001.
- [8] T. J. Schuur, A. V. Ryzhkov, and P. L. Heinselman, "Observations and classification of echoes with the polarimetric WSR-88D radar," Publications related to algorithms and upgrades for the WSR-88D radars, NOAA/National Severe Storms Laboratory, 2003, http://www.nssl.noaa.gov/publications/wsr88d_reports/.
- [9] A. V. Ryzhkov, T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrnic, "The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification," *Bulletin of the American Meteorological Society*, vol. 86, no. 6, pp. 809–824, 2005.
- [10] H. S. Park, A. V. Ryzhkov, D. S. Zrnic, and K. E. Kim, "The hydrometeor classification algorithm for the polarimetric WSR-88D: description and application to an MCS," *Weather and Forecasting*, vol. 24, no. 3, pp. 730–748, 2009.
- [11] P. Zhang, A. Ryzhkov, and D. Zrnic, "Detection of birds and insects using polarimetric radar observation," in *Proceedings of the 11th Conference on Aviation, Range, and Aerospace Meteorology*, pp. 897–901, October 2004.
- [12] P. Zhang, A. V. Ryzhkov, and D. S. Zrnic, "Observations of insects and birds with a polarimetric prototype of the WSR-88D radar," in *Proceedings of the 32nd Conference on Radar Meteorology*, CD-ROM, 9R6, pp. 365–372, American Meteorological Society, Albuquerque, NM, USA, October 2005.
- [13] W. G. Collins, "The quality control of velocity azimuth display (VAD) winds at the national centers for environmental prediction," in *Proceedings of the 11th Symposium Meteorological Observations and Instruments*, pp. 317–320, American Meteorological Society, Albuquerque, NM, USA, 2001.
- [14] L. Bi, A. Shapiro, P. Zhang, and Q. Xu, "Quality control problems for VAD winds and NEXRAD Level-II winds in the presence of migrating birds," in *Proceedings of the 19th Conference on Weather Analysis and Forecasting and 15th Conference on Numerical Weather Prediction*, pp. 143–145, American Meteorological Society, San Antonio, Tex, USA, August, 2002.
- [15] J. R. Riley, "Radar cross sections of insects," *Proceedings of the IEEE*, vol. 73, no. 2, pp. 228–232, 1985.
- [16] C. R. Vaughn, "Birds and insects as radar targets: a review," *Proceedings of the IEEE*, vol. 73, no. 2, pp. 205–227, 1985.
- [17] D. S. Zrnic and A. V. Ryzhkov, "Observations of insects and birds with a polarimetric radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 2, pp. 661–668, 1998.
- [18] V. N. Bringi, V. Chandrasekar, and D. S. Zrnic, "An examination of propagation effects in rainfall on polarimetric variables at microwave frequencies," *Journal of Atmospheric and Oceanic Technology*, vol. 7, pp. 829–840, 1990.
- [19] H. Dingle, *Migration: The Biology of Life on the Move*, Oxford University Press, New York, NY, USA, 1996.
- [20] V. A. Drake and A. G. Gatehouse, *Insect Migration*, Cambridge University Press, New York, NY, USA, 1995.
- [21] S. Liu, G. DiMego, K. V. Kumar et al., "WSR-88D radar data processing at NCEP," in *Proceedings of the 34th Conference on Radar Meteorology*, CD-ROM, 14. 2, American Meteorological Society, Williamsburg, Va, USA, 2009.
- [22] Q. Xu, K. Nai, L. Wei, P. Zhang, S. Liu, and D. Parrish, "A VAD-based dealiasing method for radar velocity data quality control," *Journal of Atmospheric and Oceanic Technology*, vol. 28, no. 1, pp. 50–62, 2011.