UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ASSIMILATION OF CASA AND WSR-88D RADAR DATA FOR TORNADIC CONVECTIVE STORMS USING AN ENSEMBLE KALMAN FILTER AND APPLICATIONS IN PROBABILISTIC ENSEMBLE FORECASTING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

NATHAN A. SNOOK
Norman, Oklahoma
2011

ASSIMILATION OF CASA AND WSR-88D RADAR DATA FOR TORNADIC CONVECTIVE STORMS USING AN ENSEMBLE KALMAN FILTER AND APPLICATIONS IN PROBABILISTIC ENSEMBLE FORECASTING

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

Dr. Ming Xue, Chair

Dr. Alan Shapiro

Dr. Jerry Brotzge

Dr. Phillip Chilson

Dr. Mark Yeary

Acknowledgements

Financial support for the research contained within this dissertation was provided by the Engineering Research Center (ERC) for Collaborative Adaptive Sensing of the Atmosphere (CASA) through NSF grant EEC-0313747. Supercomputing resources were provided by the Center for Analysis and Prediction of Storms, the OU Supercomputing Center for Research and Education (OSCER), and the National Institute for Computational Sciences (NICS) on their Kraken supercomputer.

I owe a great deal of thanks to my colleagues in CAPS. I wish to thank Dr. Jerry Brotzge, who provided me with CASA radar data and gave me invaluable advice in processing it; and Dr. Youngsun Jung, who offered countless insights into the inner-workings of the ARPS EnKF system, and helped to greatly improve journal publications based on the work contained herein. Finally, I would like to thank my graduate advisor, Dr. Ming Xue, for his continued support and guidance.

Though I benefitted immensely from the wisdom of numerous professors and mentors throughout my college career, there are a few individuals who stand out among the rest to whom I wish to express my gratitude here. To Dr. William Gallus, thank you for introducing me to meteorological research and encouraging me to pursue graduate study. To Drs. Lance Leslie, David Stensrud, and Alan Shapiro, thank you for the knowledge and experience you imparted to me during the courses I took with you which was invaluable in my dissertation research, and for your sound scientific and professional advice.

To conclude, I wish to acknowledge a few of the many family and friends who have supported me. Thank you to Mom, Dad, and Damien, for always being there for me. Thank you to Dan Dawson and Corey Potvin for many, many deep and interesting discussions, both work-related and otherwise. Thank you to Robin Tanamachi for your loyal friendship and insightful advice, both professional and personal. Thank you to Daniel Wilson and Mark McCarthy—the many evenings we spent conversing, joking around, and playing games helped keep me sane and happy through the challenges and stresses of doctoral study. To the many friends I have made as we worked together towards our graduate degrees, thank you for many happy times together; Aaron Botnick, Esther White, Kevin Goebbert, Brad Barrett, Andy Taylor, Chad Shafer, and many, many others. Finally, thank you to Laura Kanofsky, for (among many things) your continual support, understanding, and advice.

Table of Contents

List of Tables

List of Figures

Abstract

Two sets of experiments are performed to examine the ability of a storm-scale numerical weather prediction (NWP) model to predict precipitation and meso-vortices within the tornadic mesoscale convective system (MCS) that occurred over Oklahoma on 8-9 May 2007, when the model is initialized from ensemble Kalman filter (EnKF) analyses assimilating conventional and/or radar observations. Radar data from multiple networks are assimilated, including data from the X-band radar network of the Engineering Research Center for Collaborative and Adaptive Sensing of the Atmosphere (CASA) and the WSR-88D S-band radar network. The first set of experiments uses an independent ensemble for each experiment, while the second set nests each ensemble experiment within an outer-grid ensemble to provide variation in initial and boundary conditions. A variety of ensemble configurations are used during the analysis and forecast to evaluate the impact of assimilation of CASA and WSR-88D radar data, variation within the data assimilation system, and variation of the model, including changes to the model microphysics. In each experiment, radar data and/or conventional weather observations (i.e. surface, wind profiler, and upper-air observations) are assimilated every 5 minutes for 1 hour. Ensemble forecasts are then performed and probabilistic forecast products generated, focusing on prediction of radar reflectivity (a proxy of quantitative precipitation) and meso-vortices (an indication of tornado potential). This work is the first assimilating real CASA data into a NWP model using EnKF.

While the assimilation of WSR-88D data alone is able to produce a reasonably accurate analysis of the convective system, assimilating CASA data in addition to WSR-88D data improves the representation of storm-scale circulations, particularly in the lowest few kilometers of the atmosphere, as evidenced by analyses of gust front position and comparison of modeled and observed radial velocity. Assimilating CASA data decreases RMS innovation of the resulting ensemble mean analyses of radar reflectivity, particularly in early assimilation cycles, suggesting that assimilation of CASA data allows the EnKF system to more quickly achieve a good result. Use of multiple microphysics schemes in the ensemble during data assimilation is found to alleviate under-dispersion by increasing the ensemble spread.

For the single grid experiments, the ensemble assimilating data from both CASA and WSR-88D radars and using a mixed-microphysics ensemble during data assimilation produces the best probabilistic meso-vortex forecast, while the best meso-vortex forecast of the nested grid experiments is produced by the ensemble using an increased rain intercept parameter. Both of these ensembles predict maximum probabilities of greater than 0.65 of a significant near-surface vortex, with the highest probabilities clustered tightly around the location of the tornadic meso-vortex observed at 0400 UTC. Use of multiple microphysics schemes within the ensemble aims to address at least partially the model physics uncertainty and effectively plays a role of flow-dependent inflation (in precipitation regions) during EnKF data assimilation.

Sensitivity to the assumed observation error used during EnKF data assimilation is noted in the nested-grid experiments. Experiments using assumed

observation errors of 2 ms^{-1} for radial velocity and 3 dBZ for radar reflectivity are found to produce substantially better ensemble forecasts (both in qualitative structure and in quantitative skill score analyses) than an experiment using lower values of assumed observation error. The forecast ensemble assimilating only conventional data performs substantially worse than those ensembles assimilating radar data, but still retains useful skill (verified using the relative operating characteristic skill score) throughout the 3-hour forecast period. Though a bias toward stronger precipitation is noted in the ensemble forecasts, all experiments produce skillful probabilistic forecasts of radar reflectivity on a 0-3 hour timescale as evaluated by multiple probabilistic verification metrics. Though only a single case is studied herein, these experiments represent an important step toward an operational EnKF-based ensemble analysis and probabilistic forecast system to support convective-scale warn-on-forecast operations.

Chapter 1: Introduction

1.1 Motivation and Background

Accurate prediction of individual deep, moist convective storms is one of the major challenges of modern numerical weather prediction (NWP) in research and operational settings. Fully resolving all important storm-scale circulations is very expensive from a computational standpoint. In addition, most existing observing networks are quite sparse relative to the spatial scale of the flows being predicted, and offer incomplete observational coverage in both physical and parameter spaces. These challenges are then compounded by uncertainties and errors within NWP models. As computational power continues to increase, and as new high-resolution observing platforms, such as densely networked X-band radars (McLaughlin et al. 2009), are deployed to address the issues described above, a new challenge has arisen: assimilating data from multiple observing systems to best estimate the current state of the atmosphere and initialize storm-scale NWP models.

Because convective-scale errors generally grow very quickly (Lorenz 1969), it is vital to obtain the best possible estimate of the atmospheric state for NWP model initialization, ideally accompanied with an estimate of the uncertainty. While objective analysis can often be used for the purpose of obtaining a gridded analysis when observed and state variables are the same, remote sensing platforms such as radar and satellite do not directly observe most state variables, necessitating advanced data assimilation methods able to 'retrieve' state variables not directly observed. Such methods usually take advantage of physical laws linking various state variables and/or utilize information contained in observations taken at different times. They also try to obtain the “optimal” state estimation by taking into account errors associated with various sources of information. Four-dimensional variational data assimilation (4DVAR) directly uses the numerical model to provide constraints among the estimated state variables, while the ensemble Kalman filter (Evensen 1994) utilizes statistical correlations among variables derived from an ensemble of predictions to achieve a similar goal.

For convective-scale NWP, the only observing platform currently capable of providing spatially and temporally complete coverage of a convective system at a resolution sufficient to capture storm-scale features is Doppler radar. In the United States, the National Weather Service (NWS), together with other collaborating agencies, operates the WSR-88D radar network (Crum et al. 1993), consisting of 158 S-band Doppler radars. While the WSR-88D network is relatively efficient at scanning the precipitating atmosphere, the long-range radars comprising the WSR-88D network cannot reach the lower troposphere beyond a limited distance from the radar site because of the curvature of the earth. Limited low-level radar coverage presents a problem for convective-scale data assimilation and NWP; many aspects of storm- and sub-storm-scale dynamics are sensitive to the near-surface atmospheric state, including the low-level cold pool and its interaction with the surrounding environment (e.g., Rotunno et al. 1988; Markowski et al. 2002; Snook and Xue 2008).

To address the near-surface observation problem described above, the Engineering Research Center (ERC) for Collaborative Adaptive Sensing of the

Atmosphere (CASA) was established to develop short-range, networked X-band radars designed to be deployed in close proximity to one another (McLaughlin et al. 2009). Between 2006 and 2011, the primary CASA testbed was located in southwest Oklahoma and consisted of a network of four dual-polarized X-band radars with a maximum range of 40 km (Brotzge 2010b). This radar network was located roughly halfway between WSR-88D sites KTLX at Oklahoma City and KFDR at Frederick, OK; a location upstream of the Oklahoma City metropolitan area during prevailing westerly and southwesterly flow during the warm season. At the network location, neither KTLX nor KFDR can sample the lowest kilometer of the atmosphere (Xue et al. 2006), maximizing the potential benefit of increased low-level coverage provided by the CASA radars. A comparison of the key specifications of CASA and WSR-88D radars is shown in Table 1. Despite having relatively wide beams compared to WSR-88D, the CASA radars possess such advantages as a radial gate spacing of 100 m, a shorter mean range distance of observations (hence higher mean cross-beam resolutions), and a dynamic adaptive scanning strategy that identifies targets of meteorological interest and chooses an optimal combination of sector scans and full-circle scans at up to eight elevation angles to maximize the spatial and temporal coverage of features of greatest interest (Brotzge et al. 2005).

Commonly used methods for assimilating radar data into storm-scale NWP models include the three dimensional variational (3DVAR, e.g., Xue et al. 2003; Hu et al. 2006b), and four-dimensional variational methods (4DVAR, e.g., Sun et al. 1991; Sun and Crook 1997, 1998), and EnKF (e.g., Snyder and Zhang 2003; Dowell et al. 2004a; Tong and Xue 2005b). Compared to 3DVAR, EnKF has the notable advantage of being able to incorporate multivariate, flow-dependent error covariance, including cross-covariance, as derived from the ensemble (Evensen 2003), allowing effective ‘retrievals’ of state variables from radial velocity and radar reflectivity (Snyder and Zhang 2003; Tong and Xue 2005b). Unlike 4DVAR, EnKF does not require the development of an adjoint model; this development is labor-intensive and the resulting adjoint model often has difficulty handling highly nonlinear processes. In addition, ensemble forecasts are believed to be particularly important for storm-scale NWP (Xue et al. 2007). EnKF naturally provides a set of analyses that in principle best characterizes the analysis uncertainty; such analyses can therefore serve as initial conditions for ensemble forecasts.

One important goal of the CASA project is to evaluate the value and benefit of data collected by its experimental testbed radars. Preliminary results using the ARPS (Xue et al. 2000; Xue et al. 2003) Data Analysis System (ADAS, Brewster 1996), together with its cloud analysis package, for data assimilation (Brewster et al. 2007) show a generally positive impact from the addition of CASA radar data. On 8-9 May 2007, a mesoscale convective system (MCS) with a pronounced line-end vortex (LEV) developed over southwestern Oklahoma and produced several tornadoes shortly after moving out of the CASA domain. Schenkman et al. (2011) studied this case, using ARPS 3DVAR and a cloud analysis to assimilate CASA and WSR-88D reflectivity and radial velocity data, and demonstrated positive impact of CASA data on the prediction of the MCS. In the studies herein, we apply the ARPS EnKF system (Tong and Xue 2005b; Xue et al. 2006) to the 8-9 May 2007 LEV event, further evaluating the impact of CASA radar data, and combining assimilation of CASA data with WSR-88D radar data and other conventional meteorological observations (including surface observations, wind profiler observations, and soundings) to demonstrate the potential of such EnKF analyses for initialization of short-timescale (0-3 hour) deterministic and ensemble convective-scale forecasts.

One of the major goals of the National Weather Service (NWS) in the coming decade is transitioning from the current warn-on-detection strategy to a warn-on-forecast paradigm for convective-scale severe weather warnings (e.g., tornado, severe thunderstorm and flash flood warnings) in order to increase warning lead-time beyond what is possible with nowcasting techniques alone. Such an increase in warning lead-time would allow entities such as hospitals and stadiums sufficient time to respond in the event of a warning (Stensrud et al. 2009). To achieve this goal, reliable short-term (0 – 3 hour) forecasts of meteorological features at the convective scale will be vital.

1.2 Overview

In the studies herein, we seek to accomplish three major goals: (1) assimilating radar data from multiple observation platforms (including WSR-88D and CASA radars, surface observations, and upper-air observations) using an EnKF system to obtain an accurate estimate of the state of the atmosphere for a deep, moist convective system; (2) using an ensemble storm-scale forecast initialized from a set of EnKF analyses to obtain probabilistic predictions for storm-scale and sub-storm-scale processes (including mesovortex prediction); and (3) evaluating the impact of variation in experiment design, including microphysical variation, variation in assumed observation error, and in particular, the impact of CASA radar data on storm-scale NWP. To achieve these goals, we investigate the case of a MCS with an embedded line-end vortex (LEV) that spawned three tornadoes in central Oklahoma on 8-9 May 2007.

Chapter 2 contains a detailed discussion of the data, numerical models, and data assimilation techniques used in this study. We first discuss the characteristics of the radars used, including both the X-band CASA network and the S-band NEXRAD network. We then discuss the Advanced Regional Prediction System (ARPS) NWP model and its EnKF data assimilation system, along with the experimental design applied using ARPS in this study. Finally, ensemble and probabilistic forecasting techniques are introduced and discussed.

In Chapter 3 we begin by introducing the case study examined herein: the tornadic convective system of 8-9 May 2007 over southwestern Oklahoma. An overview of the case is presented, followed by a discussion of severe convective storms observed and reported during the event. Observations, including radar data, taken during the event are interpreted. Next, the experiment design for the 8-9 May 2007 single-grid case study is presented, followed by an in-depth discussion of analysis results. Results of the EnKF data assimilation process are examined, considering innovation and ensemble spread statistics and the accuracy of the resulting reflectivity and wind fields, as well as the impact of CASA data on the assimilation of a near-surface vortex present during the analysis period.

In Chapter 4, we discuss the results of deterministic and ensemble forecasts, including probabilistic forecast products, initialized from the 0200 UTC EnKF analyses obtained in Chapter 3. Forecasts with and without CASA radar data are compared, and the impact of using a mixed-microphysics ensemble is considered. Forecast verification is performed both subjectively, through analysis of radar reflectivity and vertical convective structure; and objectively, using a variety of deterministic and probabilistic skill score metrics.

In Chapter 5 we present the results of a more complex set of data assimilation and forecast experiments performed using the 8-9 May 2007 tornadic MCS case. These experiments differ from those presented in Chapters 3 and 4 in their use of conventional weather observations (such as upper-air data, wind profiler observations, and surface observations) in addition to radar data during EnKF data assimilation. Also, the analysis and forecast experiments are nested within an outer ensemble analysis and forecast at a horizontal resolution of 6 km, included to more accurately represent mesoscale variations and provide perturbed initial conditions and lateral boundary conditions to the 2 km forecast experiments within. These experiments are run on a considerably larger forecast domain identical to that used by Schenkman et al. (2011) in their study of this case using 3DVAR data assimilation, allowing for direct comparison between the EnKF-based forecast results obtained herein and the 3DVAR-based forecast results of Schenkman et al. (2011).

We close with Chapter 6, in which the results of both sets of analysis and forecast experiments are summarized, conclusions are drawn, and thoughts regarding possible future investigations are presented.

Chapter 2: Methodology and Tools

2.1 The CASA project

2.1.1 CASA: An Overview

The Center for Collaborative Adaptive Sensing of the Atmosphere (CASA) is a ten-year National Science Foundation (NSF) Engineering Research Center (ERC), established in 2003. CASA seeks to develop an innovative, low-cost, high-spatial-density, dynamically-adaptive network of Doppler radars with polarimetric capability (Xue et al., 2006) which can be used to detect, track, analyze, and predict tornadoes or processes leading to tornadogenesis (McLaughlin et al. 2009). To these ends, CASA deployed a radar testbed in southwestern Oklahoma which was active between the spring of 2006 and the spring of 2011, consisting of four X-band radars with a maximum range of 40 km, sited in southwestern Oklahoma near Chickasha, Cyril, Lawton, and Rush Springs. The placement of these CASA radars is shown below in Fig. 2.1.

The Weather Surveillance Radar-1998 Doppler (WSR-88D) network is the primary operational network of Doppler weather radars used in the United States. The CASA radars are positioned between the Frederick (KFDR) and Twin Lakes (KTLX) WSR-88D sites, to fill in the gap in radar coverage near the surface caused by the curvature of the earth. CASA seeks to complement the existing WSR-88D network by providing enhanced low-level radar coverage, improving the ability to detect low-level hazardous weather such as tornadoes, and providing a more complete data set for initialization of storm-scale NWP models. The polarimetric data provided by CASA is also an improvement upon the WSR-88D network, which does not have polarimetric capability (though at the time of this writing the WSR-88D radar network is in the process of being upgraded to add this capability). While the radar reflectivity and radial velocity data observed by the WSR-88D network provide a great deal of information, these variables alone do not provide sufficient data to fully describe the microphysical state of a convective storm. Polarimetric data can be helpful in this area, as observations of differential reflectivity and differential phase contain information about the density, shape, orientation, and size distribution of hydrometeors within the storm (Jung et al. 2008).

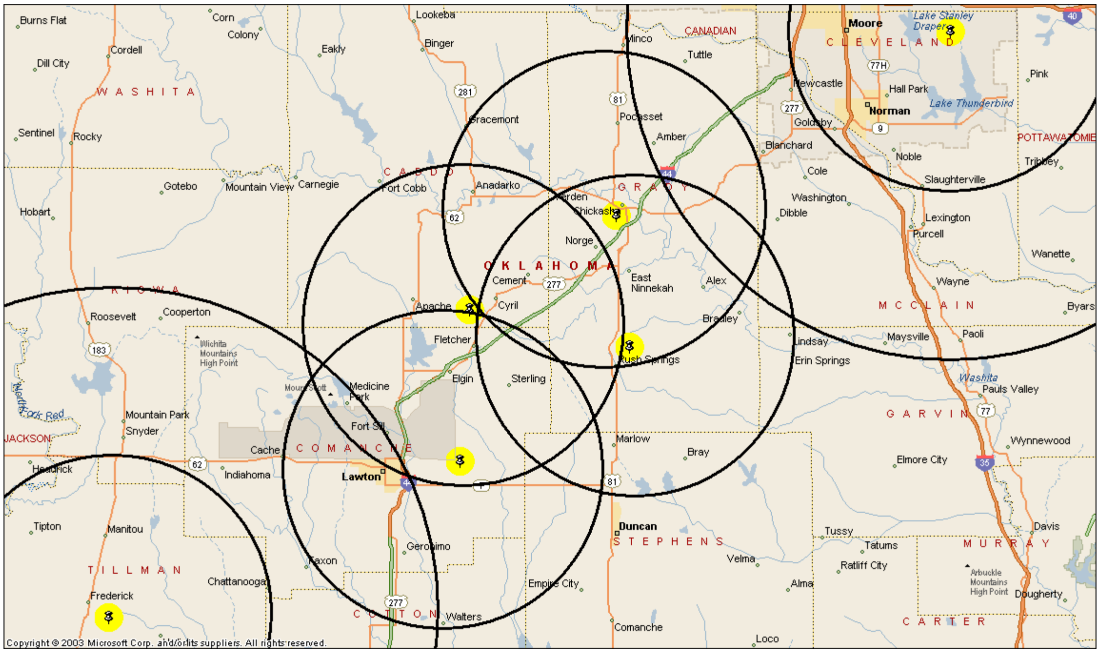

Fig. 2.1: Locations of the 4 radars in the CASA radar testbed. 30 km range rings are shown in black. Also shown are the locations of the nearest two WSR-88D radars; KFDR near Fredrick, OK, and KTLX east of Moore, OK, along with 30 and 60 km range rings. Counties, cities, rivers, and major highways are shown.

2.1.2 Radar Specifications

The CASA network consists of 4 identical X-band radars. While future CASA radars may employ phased-array technology to allow for electronic beam steering (Xue et al., 2006), the existing CASA radars are driven by a single magnetron and steered mechanically. Technical specifications of the CASA radars, and the corresponding specifications of the WSR-88D radar network, are detailed below in Table 2.1.

Table 2.1: Specifications of the radars used in the CASA radar testbed in southwestern Oklahoma, along with corresponding specifications of the operational WSR-88D radar network maintained by the National Weather Service.

	CASA	**WSR-88D**
Wavelength	3.19 cm (X-band)	10.0 cm (S-band)
Maximum Peak Power	25 kW	750 kW
Pulse Repetition Frequency	Variable up to 3.33 kHZ	0.3 – 1.3 kHZ
3 dB Beamwidth	2.0 degrees	0.95 degrees
Polarization	Dual linear (V and H)	Horizontal polarization
Rotation Rate	Variable up to 120 deg./s	36 deg./s
Antenna Gain	38 dB	45 dB
Antenna Diameter	1.5 m	8.5 m
Maximum Range	40 km	459 km

The 4 radars of the CASA network are connected to a central processing server. Every minute, the processing server gathers data from all 4 radars and uses pre-defined rules to generate an optimal scanning strategy for the next one-minute scan cycle. The rules represent the competing interests of various end-users of the CASA network, including research scientists, operational meteorologists, and emergency managers. An example of the output from this process is shown in Fig.

2.2; the rules are used to generate tasks (displayed as polygons), which are then used applied to generate the next scanning strategy. Possible scan modes include a full 360 degree scan at a constant elevation, a sector scan on a constant elevation, an RHI scan, or a combination of the above at varying elevations and azimuths.

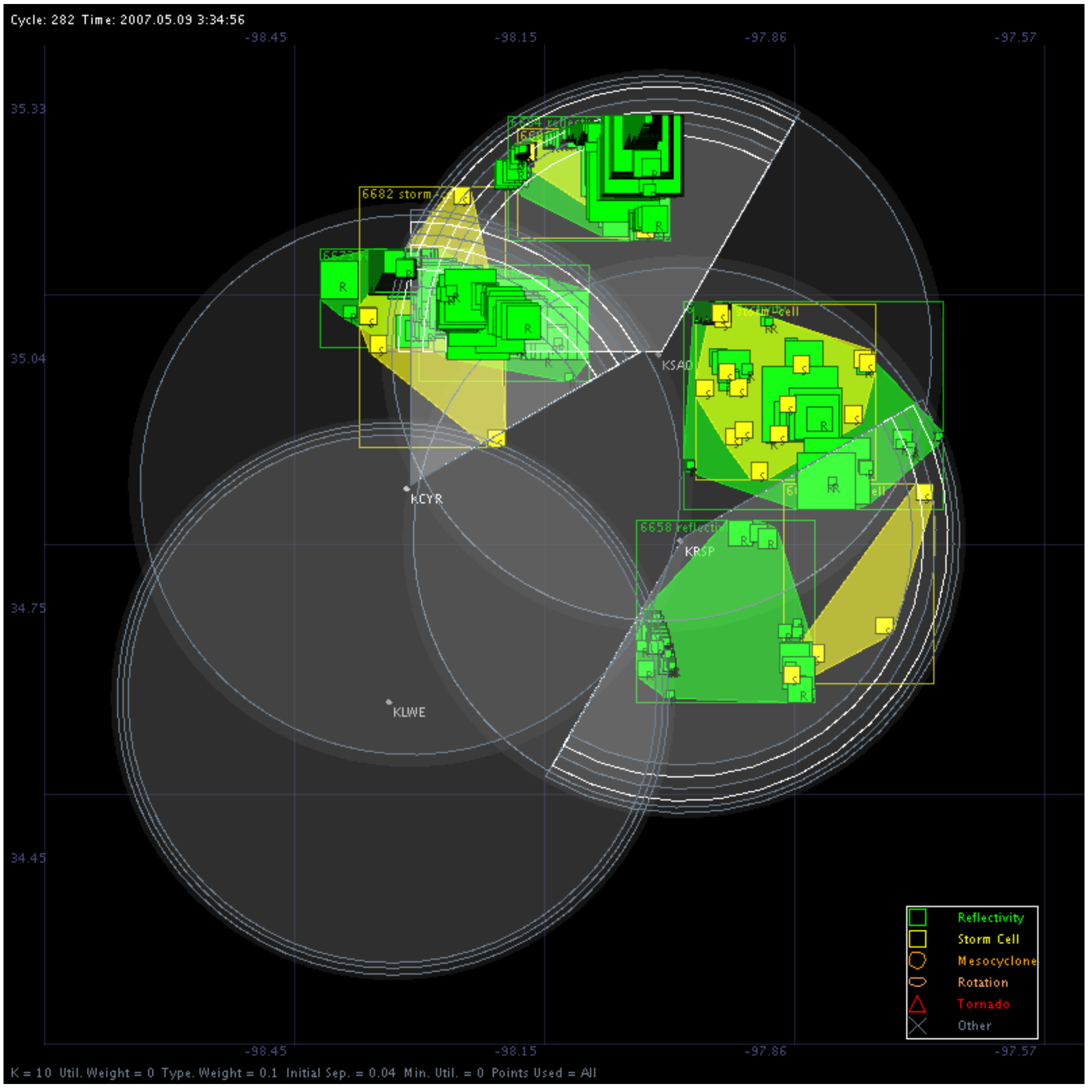

Fig. 2.2: Diagram of one CASA scan cycle. Sector scans are denoted by the shaded sectors for each radar. Scanning strategy was adaptively chosen to best scan tasks generated from the observed meteorological features, shown in color.

Working with X-band radars presents a major challenge in the form of attenuation. In cases of heavy precipitation, significant attenuation of the radar signal can occur. To account for this, CASA uses a built-in attenuation correction algorithm. Also, the relatively wide (2°) beamwidth of the CASA radars limits the azimuthal resolution of the system—this problem is addressed by oversampling in the azimuthal direction.

2.2 The ARPS NWP model

The Advanced Regional Prediction System (ARPS) is used in this study to produce forecasts of convective storms. ARPS is a fully-compressible, non-hydrostatic atmospheric prediction model (Xue et al. 2000). The model state vector contains three velocity components (*u*, *v*, and *w*), potential temperature (θ), air pressure (*p*), turbulence kinetic energy (as used in the model's 1.5-order sub-grid-scale turbulence closure scheme), in addition to microphysical information dependent upon the microphysical scheme used. Both single- and multiple-moment microphysical schemes are available. All microphyicsal schemes used in ARPS assume a three-parameter gamma distribution of the form:

$$n_x(D) = N_{0x} D^{\alpha} e^{-\lambda D} \tag{2.1}$$

where the subscript x denotes the hydrometeor species (rain, snow, or hail/graupel), $n_x(D)$ represents the number of hydrometeors per unit volume with diameter D, N_{0x} is an intercept parameter, and α is a slope parameter. When $\alpha = 0$, the drop size distribution reduces to that of Marshall and Palmer (1948).

The most widely-used single moment schemes available in ARPS are based largely upon on Lin et al. (1983), and track the mixing ratio of water vapor (q_v), as well as mixing ratios for varying combinations of the following hydrometeor species: cloud water (q_c), cloud ice (q_i), rain (q_r), snow (q_s), hail (q_h), and graupel (q_g), while the intercept parameter and slope parameter are set to pre-specified, constant values. ARPS also supports multiple-moment schemes based upon the work of Milbrandt and Yau (2005), which predict the mixing ratios listed above, as well as intercept parameters (N_{0x}), where x denotes hydrometeor species. The slope parameter can be pre-specified, diagnosed, or predicted.

In the studies herein, radar data assimilation and storm-scale forecasts are performed on an ARPS domain with 2 km horizontal spacing; vertical grid stretching is applied, giving a near-surface vertical grid spacing of approximately 100 m. The single-grid experiments presented in Chapters 3 and 4 use a physical domain of 256 × 256 × 40 grid points, while the nested-grid experiments presented in Chapter 5 use a larger 500 × 500 × 40 physical domain for the inner nest. In all experiments, the model top is located at a height of 20 km. Full model physics are used (Xue et al. 2001), including the NASA Goddard Space Flight Center long- and shortwave radiation parameterization, a two-layer soil model, surface fluxes parameterized using predicted surface temperature and water content, and a 1.5-order turbulent kinetic energy (TKE)-based sub-grid-scale turbulence parameterization, along with high-resolution terrain. Initial and boundary conditions are derived from NCEP NAM analyses. The initial ensemble is generated by adding smoothed random perturbations to the model state using the method of Tong and Xue (2008). These smoothed perturbations have a horizontal length scale of 8 km and a vertical length scale of 5 km and are added to the horizontal wind field with a standard deviation of 2 m s^{-1}, to the mixing ratios of hydrometeors, cloud water, and cloud ice with a standard deviation of 0.001 kg kg^{-1}, and to the potential temperature field using positive perturbations only with a standard deviation of 2 K. The initialization of the EnKF system will be discussed in depth later, in section 2.3.2.

2.3 The ensemble Kalman filter for meteorological data assimilation

At the convective scale, the accuracy and rate of divergence of a given NWP forecast depends greatly upon the quality of the initial conditions used. As time passes, the model error increases from that present in the initial state, eventually resulting in a predicted state quite different from the true state of the atmosphere. As a result, one of the top priorities in storm-scale NWP is generating the most accurate set of initial conditions to minimize the initial error in the model state. Of the data assimilation techniques available today, the four-dimensional variational (4DVAR) technique (Le Dimet and Talagrand 1986) and the ensemble Kalman filter (EnKF) technique (Evensen 2003) are considered to hold the most promise.

2.3.1 Data assimilation using EnKF techniques

The EnKF technique was first introduced into the meteorological community around ten years ago, and has rapidly become an attractive technique for many operational and research applications. EnKF and related filtering methods incorporate the computation of flow-dependent error statistics. Instead of solving for the time-dependent probability density function of the model state, EnKF applies the Monte Carlo technique for estimation of the forecast error statistics. By producing a large ensemble of model states and integrating them forward in time, the moments of the probability density function can be calculated for different times (Evensen 2003).

EnKF has proven to be very effective in retrieving accurate and dynamically consistent wind, temperature, and microphysical fields from radar reflectivity and radial velocity observations when using simulated observations (e.g., Snyder and Zhang 2003; Zhang et al. 2004; Tong and Xue 2005a; Xue et al. 2006; Tong and Xue 2008). Snyder and Zhang (2003) demonstrated that simulated radial velocity data could be successfully assimilated using an EnKF method and warm rain microphysics. Tong and Xue (2005a) expanded upon these results, further showing that EnKF could successfully retrieve the microphysical information needed for a simple single-moment ice microphysics scheme from simulated radar reflectivity and radial velocity data, and provide a reduction in root-mean-square error (rmse) despite the highly non-linear nature of the reflectivity observation operator. Obtaining analyses from real data that lead to good short-range forecasts of convective storms remains a challenge (Dowell et al. 2004b; Tong 2006); most storm-scale EnKF studies to date have focused on analysis rather than forecasting; thus relatively few papers showing good forecast results have been published so far, except for Lei et al. (2009) and Dowell et al. (2010).

Real radar observations are almost always contaminated by measurement, sampling, and aliasing errors; compared to simulated radar observations, which use a perfect model and are free from such errors, the challenge of effectively assimilating real radar observations is much greater. Those studies which have used real data have had success in retrieving convective thermodynamic and wind fields, but have often found it necessary to do extensive quality controls, delete erroneous or contaminated data, and perform post-processing such as objective analysis to place data on the model grid (e.g. Dowell et al. 2004b). In the experiments presented here, assimilation of data from multiple radar platforms (WSR-88D and CASA) and the use of assimilated data to update not only warm-rain processes but also ice microphysical processes add additional layers of complexity.

2.3.2 The ARPS EnKF system

The EnKF system of the ARPS model employs an ensemble square-root Kalman filter (EnSRF) similar to that presented in Whitaker and Hamill (2002) and later refined by Tong and Xue (2008), which is a particular variant of the EnKF technique that does not require perturbation of the observations being assimilated. At the beginning of the assimilation period, spatially smoothed perturbations are added to the background state. As in Tong and Xue (2008), the spatially smoothed perturbation at the grid point (x, y, z) is calculated using:

$$\varepsilon(x,y,z) = E\sum_{S} r(i,j,k)W(i,j,k) \tag{2.2}$$

where $r(i, j, k)$ is a random value independently sampled from a normal distribution with a mean of 0 and standard deviation of 1, $W(i, j, k)$ is a 3D distance-dependent weighting function, and E is a scaling parameter for obtaining correct variance in the perturbation field. S represents a domain of grid points used for summation, consisting of all points within a specified radius of *(x, y, z)*. This distance is chosen using the typical decorrelation length of background error scales. For WSR-88D radar data, this was assumed to be approximately 6 km, and for CASA radar data, approximately 2 km.

The smoothed initial perturbations are next rescaled, by evaluating E in equation 2.2 such that the standard deviation of each perturbation field is equal to a desired value. These desired values are 1 m s^{-1} for velocity components u, v, and w; 2 K for potential temperature θ; and 1 g kg^{-1} for microphysical mixing ratios q_v, q_c, q_r, q_i, q_s, and q_h. These values were determined via experimentation and consultation of previous studies (e.g. Tong and Xue 2008). Such perturbations were added to all grid points within a cutoff radius of 6 km of points containing precipitation.

2.4 Ensemble-based Probabilistic Forecasting Tools

To achieve the lead-times necessary for the warn-on-forecast paradigm currently under development within the NWS (Stensrud et al. 2009), reliable short-term (0 – 3 hour) probabilistic forecasts at the convective scale will be vital. While it is possible to obtain probabilistic forecast products from a single deterministic forecast by examining the occurrence of an event at surrounding grid points (Theis et al. 2005; Schwartz et al. 2009b), the capability of such a method is limited by inherent biases and errors in the deterministic forecast used. One such source of error is the microphysical parameterization of the NWP model; Snook and Xue (2008) and Dawson et al. (2009) found that the choice of microphysical scheme (and the parameter settings therein) strongly impact the mode and intensity of convection predicted. Furthermore, such probabilistic forecast products are inadequate in addressing uncertainty in the initial condition and dynamic error growth. Convective-scale errors generally grow very quickly (Lorenz 1969), greatly limiting the utility of a single deterministic forecast.

The use of an ensemble of model predictions allows us to mitigate the impact of errors and biases in initial conditions and model parameterizations, while providing valuable probabilistic data regarding the forecast. EnKF methods have proven effective in generating dynamically consistent wind, temperature, and microphysical fields when assimilating Doppler radar reflectivity and radial velocity data for storm-scale NWP (e.g. Dowell et al. 2004; Houtekamer et al. 2005; Tong 2006). Because EnKF is inherently an ensemble method, it provides a natural starting point for a storm-scale ensemble forecast beginning from the final assimilated state of each of the ensemble members during the data assimilation period.

Because convective cells are highly localized, even small displacement errors in a storm-scale forecast can result in very low objective skill scores when verified on a point-by-point basis, even though the forecast being scored may still be quite valuable to researchers and operational meteorologists (Baldwin et al. 2001; Schwartz et al. 2009b). To alleviate this problem, Schwartz et al. (2009b) applied a probabilistic “neighborhood” approach following Roberts and Lean (2008), where the probability of an event (e.g., radar reflectivity > 40 dBZ; rainfall rate > 2.0 cm/hr; hail mixing ratio > 0) at a grid point is determined by conditions at all grid points within a given radius of influence r from that grid point; this collection of points comprises the neighborhood for the selected grid point. In the studies presented herein, as was done in Schwartz et al. (2010), the neighborhood is extended to include all grid points on the same model level within the radius of influence from every ensemble member. Thus, drawing from Schwartz et al. (2009b; 2010), the forecast probability (P_i) of an event at the i^{th} grid point of the ensemble forecast is defined by:

$$P_i = \frac{1}{N}\sum_{j=1}^{N} B_j \qquad (2.3)$$

where N is the unique collection of all points comprising the neighborhood for point i, and B_j is the binary probability at the j^{th} grid point of the neighborhood, defined to be 1 if the event was observed at that grid point, and 0 if it was not. Schwartz et al. (2010) call P_i neighborhood ensemble probability (NEP); further discussion can be found in that paper.

Meso-vortices and other circulations cannot easily be treated as point variables, and are therefore not well-suited to a NEP method. Thus, for prediction of meso-vortices, we instead apply an object-based ensemble verification approach. Significant low-level circulations are first manually identified for each ensemble member. For a feature to count as a significant circulation, three criteria must be met: (1) vertical vorticity must exceed 0.02 s^{-1} at 2 km above ground level; (2) discernable rotation must be present in the horizontal wind field at this level, and (3) the feature must be located within a convective region with radar reflectivity exceeding 30 dBZ. These criteria were chosen to discriminate true mesovortices from other phenomena, such as shear zones occurring along outflow boundaries.

The binary probability B_{im} for the i^{th} model grid volume of the m^{th} ensemble member is then calculated such that $B_{im} = 1$ if the center of that grid volume lies within a predefined distance (in this study, 25 km) of a circulation center identified in the forecast for that ensemble member, and $B_{im} = 0$ otherwise. A purely ensemble-based probability results,

$$P_i = \frac{1}{N_{ens}} \sum_{m=1}^{N_{ens}} B_{im} \qquad (2.4)$$

where N_{ens} is the number of members in the ensemble. In essence, P_i from equation 2.4 can be viewed as the predicted probability of a strong near-surface vortex being present within 25 km of a given point; this forecast methodology is analogous to that used operationally by the Storm Prediction Center in their Day 1 tornado, hail, and wind outlook products, which forecast the probability of an event occurring within a 25 mile radius of a given point (Edwards et al. 2002).

Chapter 3: 8-9 May 2007—EnKF Radar Data Analysis

3.1 An Overview of the 8-9 May 2007 Tornadic MCS

The tornadic MCS of 8-9 May 2007 occurred over northern Texas and portions of western and central Oklahoma. Three confirmed tornadoes were reported during the event; two of these were of EF-1 intensity, the first of which occurred near Minco, OK at 0354 UTC, 9 May 2007, and the second of which was reported near El Reno, OK at 0443 UTC. The El Reno tornado was by far the most destructive of the three, causing an estimated $3 million of damage. In addition to these two tornadoes, a weaker, EF-0 tornado was confirmed near Union City, OK, reported at 0426 UTC. Of these three tornadoes, only for the El Reno tornado did the NWS issue a tornado warning prior to tornado occurrence, underscoring the challenge of forecasting tornadoes associated with MCVs. The location, time of occurrence, and intensity of the observed tornadoes are summarized in Fig. 3.1.

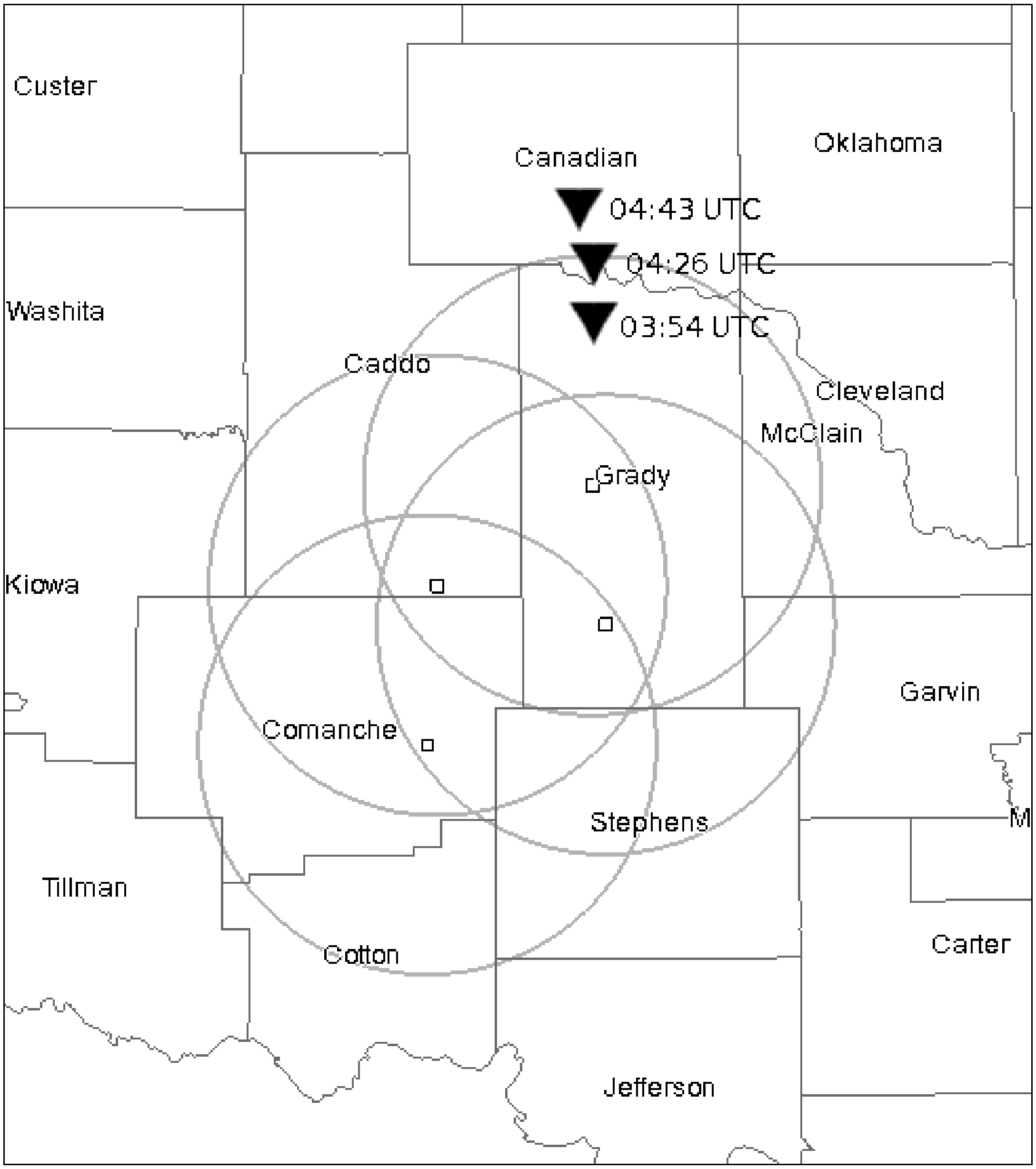

Fig. 3.1: Locations of tornadic activity associated with the 8-9 May 2007 convective system. CASA radars are indicated by black squares; 40 km CASA range rings are indicated in gray. Oklahoma counties are shown in black and labeled. Confirmed tornadoes the forecast period (0200 UTC to 0500 UTC, 9 May 2007) are indicated by black triangles with the time of occurrence noted (all times shown are for 9 May 2007). The tornadoes reported at 0354 and 0443 UTC were of EF-1 intensity on enhanced Fujita scale; the tornado reported at 0426 UTC was of EF-0 intensity.

During the afternoon and evening hours of 8 May 2007, a surface low developed in southwest Oklahoma near the intersection of a slow-moving east-west frontal boundary and an advancing cold front to the west. A mesoscale convective system (MCS) developed from pre-existing multicellular convection along the cold front (Kumjian and Ryzhkov 2008) beginning around 1200 UTC on 8 May 2007, in an area of upper-level divergence associated with a cyclonically curved jet streak. This system grew in geographic extent and by 0000 UTC on 9 May 2007 was located over much of central and north Texas and southwestern Oklahoma and featured a surging bow echo located along its leading edge (Fig. 3.2a). While the portion of the system in Texas began to weaken after 0100 UTC, the northern portion of the system in Oklahoma persisted until approximately 0730 UTC (Fig. 3.2b-f). Though strong low-level rotation was observed in the system as early as 0021 UTC, and a brief tornado was reported by a media-based storm chaser west of Lake Elsworth, OK, at 0115 UTC, subsequent damage survey efforts were unable to confirm this report (Brotzge et al., 2009), and no further tornadic activity was reported until 0354 UTC.

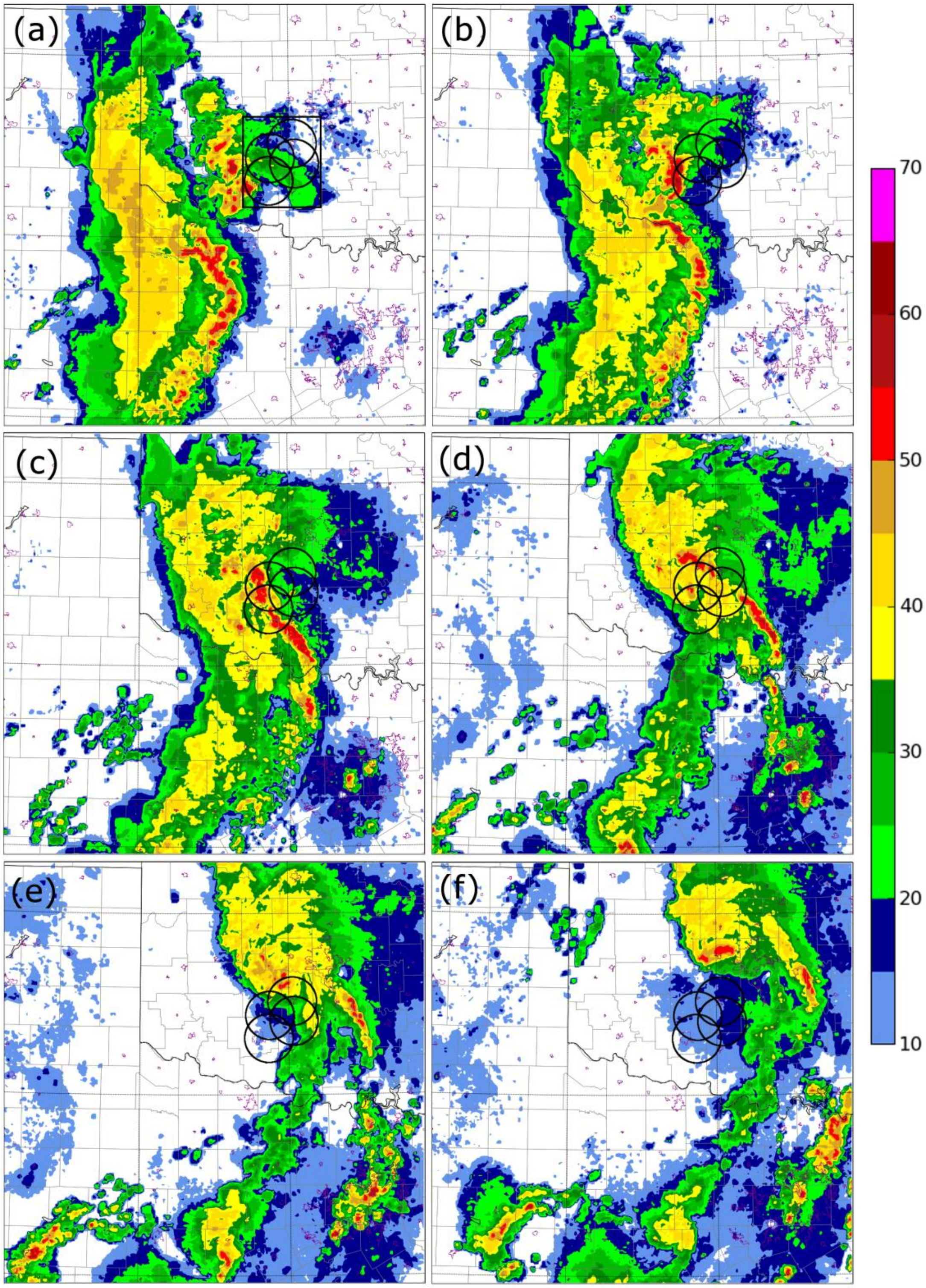

Fig. 3.2: Composite radar reflectivity mosaic (dBZ) as observed by WSR-88D radars KAMA, KDYX, KFWS, KLBB, and KTLX at (a) 0000 UTC, (b) 0100 UTC, (c) 0200 UTC, (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC, 9 May 2007. 30 km CASA radar range rings are included for reference. Urban boundaries are shown in purple.

Beginning approximately 2200 UTC on 8 May, a LEV formed near the northern end of the bow echo, just south of the Red River in northwest Texas. The development of the LEV occurred as the MCS merged with a supercell to its northeast (Schenkman et al. 2010). The LEV moved north-northeast and contracted as it moved into southwestern Oklahoma (Brotzge 2010a). The LEV intensified between 0230 and 0300 UTC as it interacted with and absorbed a supercell in Comanche County OK (Fig. 3.2b, c); evidence of this intensification was present in both WSR-88D radar reflectivity and mesonet observations (Schenkman et al. 2010). The LEV reached its peak intensity between 0330 and 0530 UTC (Fig. 3.2e, f), during which time all three reported tornadoes occurred. Observations from the Oklahoma mesonet indicate that at its peak, the LEV contained a well-defined surface circulation with approximately 25 ms^{-1} of horizontal wind shear (Schenkman et al. 2010). The observed evolution of the MCS and its associated LEV closely fits the conceptual model of an asymmetric convective system presented by Houze et al. (1989).

3.2 Radar Data Assimilation Study: Data, Methods, and Goals

In this radar data assimilation study, we apply the EnKF system of the ARPS NWP model to CASA and WSR-88D radar data gathered during the LEV and mini-supercell tornado case of 8-9 May 2007. During this event, as discussed above, a LEV developed and moved through much of southwest/central Oklahoma, passing directly through CASA's Integrative Project One (IP1) radar network (McLaughlin et al. 2009). This system spawned two confirmed EF-1 tornadoes and one confirmed

EF-0 tornado in central Oklahoma between 0354 UTC and 0443 UTC. This LEV was observed by several WSR-88D radars, as well as all four CASA radars.

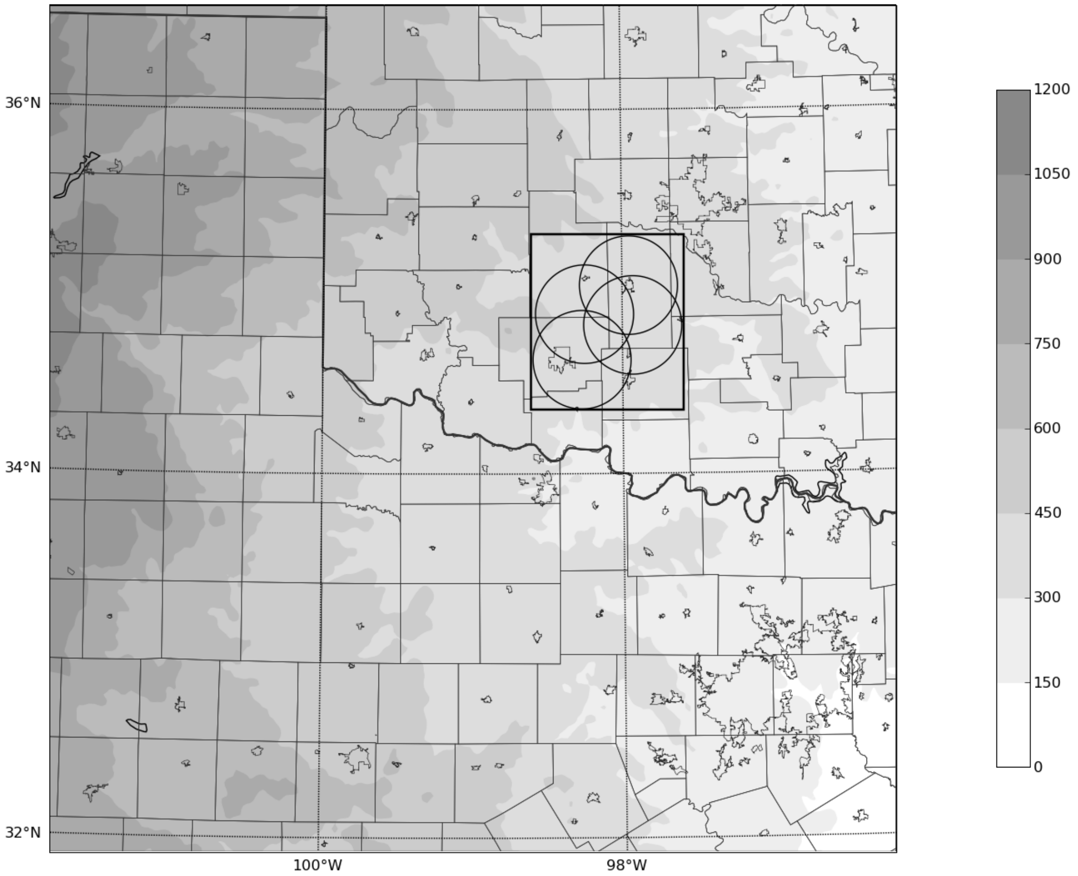

Fig. 3.3: Forecast domain used in the 2km horizontal resolution radar data assimilation experiments for the 9 May 2007 case. Terrain height, in meters above mean sea level, is indicated by the grayscale shading. 40 km CASA radar range rings are included for reference. The black box closely enclosing the CASA IP-1 radar network denotes the extent of the CASA sub-domain used in calculation of updraft mass flux presented in Fig. 3.9.

In this study, a 259 × 259 × 43 ARPS grid with 2 km horizontal spacing is used for analyses and forecasts; vertical grid stretching is applied, giving a near-surface vertical grid spacing of approximately 100 m. The model top is at a height of 20 km. The extent of the domain is shown above in Fig. 3.3. A one-hour-long pre-forecast is performed before EnKF data assimilation cycles begin, initialized from the 9 May 2007 NCEP 0000 UTC NAM analysis. At 0100 UTC, smoothed random perturbations are added to the 1-hour forecast using the method of Tong and Xue (2008) to create a set of initial conditions from which ensemble forecasts are launched. The smoothed perturbations with a horizontal length scale of 8 km and a vertical length scale of 5 km are added to the horizontal wind field with a standard deviation of 2 m s^{-1}, to the mixing ratios of hydrometeors, cloud water, and cloud ice with a standard deviation of 0.001 kg kg^{-1}, and to the potential temperature field using positive perturbations only with a standard deviation of 2 K.

The EnKF algorithm used is the ensemble square-root filter (EnSRF) of Whitaker and Hamill (2002). Radar data are assimilated every 5 minutes from 0100 to 0200 UTC. The observation error standard deviations are assumed to be 1 m s^{-1} for radial velocity and 2 dBZ for radar reflectivity. The observation operator used to map the model state to observation space for radar reflectivity and radial velocity follows that of Jung et al. (2008). To sample radar data on the radar elevation angles, a Gaussian power-gain function following Wood and Brown (1997) is used as in Xue et al. (2006). The covariance localization radius is set to 6 km. Lateral boundary conditions are provided by the NCEP NAM 6-hourly analyses and intervening 3 hour forecasts.

Radar data are assimilated every 5 minutes between 0100 and 0200 UTC in three experiments: a control experiment (hereafter referred to as “CNTL”) using a mixture of several microphysical schemes (as described below) and assimilating both WSR-88D and CASA data, an experiment using a mixed-microphysics ensemble and

WSR-88D data only (hereafter "NoCASA"), and one using both WSR-88D and CASA data, but only a single microphysical scheme (hereafter "NoMMP"). The differences in model setup between these three experiments are summarized in Table 3.1.

Table 3.1: Summary of experiments for the radar data assimilation and ensemble forecast studies presented in chapters 3 and 4.

	CNTL	**NoCASA**	**NoMMP**
WSR-88D radar used?	Yes	Yes	Yes
CASA radar used?	Yes	No	Yes
Ensemble size	40	40	40
Number of Lin microphysics members	16	16	40
Number of WSM-6 microphysics members	16	16	0
Number of NEM microphysics members	8	8	0

In all experiments, level-II volume scans of WSR-88D radial velocity and reflectivity from five WSR-88D radars are assimilated at 5-minute intervals; the WSR-88D radars used are those located at Oklahoma City, OK (KTLX), Vance Air Force Base, OK (KVNX), Amarillo, TX (KAMA), Dyess Air Force Base, TX (KDYX), and Lubbock, TX (KLBB). In experiments using CASA data, aggregate volumes of radial velocity and radar reflectivity are assimilated, also at 5 minute intervals, from each of the four CASA radars: Cyril, OK (KCYR); Lawton, OK (KLWE); Rush Springs, OK (KRSP); and Chickasha, OK (KSAO). Aggregate CASA radar volumes are created by first interpolating raw CASA sector scan data on observed elevations to a uniform radial grid with azimuthal spacing of 1 degree, and then interpolating in time to the center of a five-minute window valid at the assimilation time. For each radial, the nearest data before and after the assimilation time within the five-minute window are linearly interpolated in time to obtain the corresponding radial in the aggregate volume scan. If only one scan is available for a given radial, that scan is used. If no scans are available, that radial is marked as missing. A summary of the radars used and their locations are provided in Table 3.2.

Table 3.2: List of radars used for the single-grid radar data assimilation study.

	Radar Type	**Latitude**	**Longitude**	**Elevation**
KAMA	WSR-88D	35.2333 N	101.7092 W	1113 m
KDYX	WSR-88D	32.5383 N	99.2544 W	357 m
KLBB	WSR-88D	33.6542 N	101.8142 W	1013 m
KTLX	WSR-88D	35.3331 N	97.2778 W	384 m
KVNX	WSR-88D	36.7408 N	98.1278 W	379 m
KCYR	CASA	34.8739 N	98.2522 W	448 m
KLWE	CASA	34.6239 N	98.2708 W	396 m
KRSP	CASA	34.8128 N	97.9306 W	436 m
KSAO	CASA	35.0314 N	97.9562 W	356 m

Because of the strong reflectivity attenuation inherent to X-band radar data in areas of heavy precipitation, attenuation correction using polarimetric differential phase (Chandrasekar et al. 2004) is applied to CASA data before they are used. Accurate attenuation correction is vital; if uncorrected, attenuated radar data were assimilated, the erroneously low values of reflectivity in the attenuated regions would negatively impact the analysis. While the attenuation correction algorithm used has been shown to accurately retrieve un-attenuated reflectivity values (Chandrasekar et al. 2004), it can only do so when the reflected power is above the noise floor of the radar receiver. When total attenuation occurs, the resulting radar data cannot be objectively distinguished from true clear-air data; these areas of total attenuation appear as “shadows” of near-zero reflectivity in areas which may actually contain significant precipitation.

To avoid erroneous assimilation of completely attenuated reflectivity data, CASA reflectivity and radial velocity data are assimilated only in regions where attenuation-corrected reflectivity exceeded 20 dBZ. Unfortunately, this constraint eliminates the ability of CASA reflectivity data to suppress spurious storms that occur in regions free of observed reflectivity; Tong and Xue (2005) showed that the assimilation of reflectivity data in non-precipitation regions is very beneficial in suppressing spurious storms. Furthermore, though CASA data are not assimilated when attenuation-corrected reflectivity is less than 20 dBZ, because no reliable way exists to objectively distinguish fully-attenuated regions from clear air echo regions, attenuated areas were included in the RMS innovation computation at CASA sites, resulting in increased RMS innovation values for the CASA radars. Finding new ways to more effectively use X-band reflectivity data remains an important research topic (Xue et al. 2009).

In order to counteract the inherent tendency for the ensemble to converge on a solution different from the true state of the atmosphere, a method for maintaining ensemble spread is needed (Dowell and Wicker 2009). In order to achieve this, we apply a covariance inflation factor (Anderson and Anderson 1999) of $\gamma = 1.25$ to the prior deviation of each ensemble member from the ensemble mean. This value of γ was chosen to be large enough to broaden the ensemble distribution, but not so large as to result in an unstable or unrealistic model state. Recent work has applied other techniques to increase ensemble spread—for example, Zhang et al. (2004) used a "relaxation" technique that restored a pre-set fraction of the ensemble spread lost during the assimilation period. Also, additive perturbations to various model fields (e.g. horizontal wind, potential temperature) have been employed. Dowell and Wicker (2009), found that applying smoothed additive perturbations to the horizontal wind, potential temperature, and water vapor fields yielded a significant increase in the resulting ensemble spread. While ensemble spread maintenance used in these studies presented here is limited to covariance inflation, investigation of the optimal combination of multiplicative inflation, additive noise, relaxation, and/or adaptive inflation is a promising area for future study.

At the end of the assimilation period, three sets forecasts are performed from 0200 to 0500 UTC for each of the three experiments: a single, deterministic forecast initialized from the 0200 UTC ensemble mean analysis of each experiment, three deterministic forecasts initialized from the 0200 UTC ensemble mean analysis of experiment CNTL, each using a different microphysical scheme, and a 40 member ensemble forecast with members initialized from the final analysis at 0200 UTC of each ensemble member used during the EnKF data assimilation process. The results of these forecasts will be addressed in Chapter 4.

3.3 Radar Data Assimilation Results

By the end of the assimilation period at 0200 UTC 9 May 2007, all three experiments produce a MCS with reflectivity structure very similar to that observed by radar. Composite radar data calculated from model fields for CNTL, NoCASA, and NoMMP (Fig. 3.4a-c) correspond well to composite radar reflectivity measured by WSR-88D (Fig. 3.4d). Analyzed composite reflectivity at 0200 UTC for all three experiments compares closely with observed reflectivity in terms of the intensity and location of the main convective cells and stratiform rain region and the overall shape of the bow echo (Fig. 3.4). In both model simulations and observations, a LEV is present at the northern end of the line of strongest convection, located in the western portion of the CASA radar network (Fig. 3.4). Subtle differences between experiments are present in the composite reflectivity fields near the CASA network (Fig. 3.4a-c). Minor differences between the analyses (Fig. 3.4a-c) and the observations (Fig. 3.4d) of radar reflectivity are notable in the southern portion of the domain, where all three experiments underestimate the coverage and intensity of the strong echo region where the reflectivity is greater than 35 dBZ. Insufficient low-level radar coverage in southwestern portion of the domain is believed to have contributed to the model error there, while underestimation of the intensity of the main convective line is likely due to under-correction to the background forecast by the ensemble filter, which can occur as a result of under-dispersion in the ensemble.

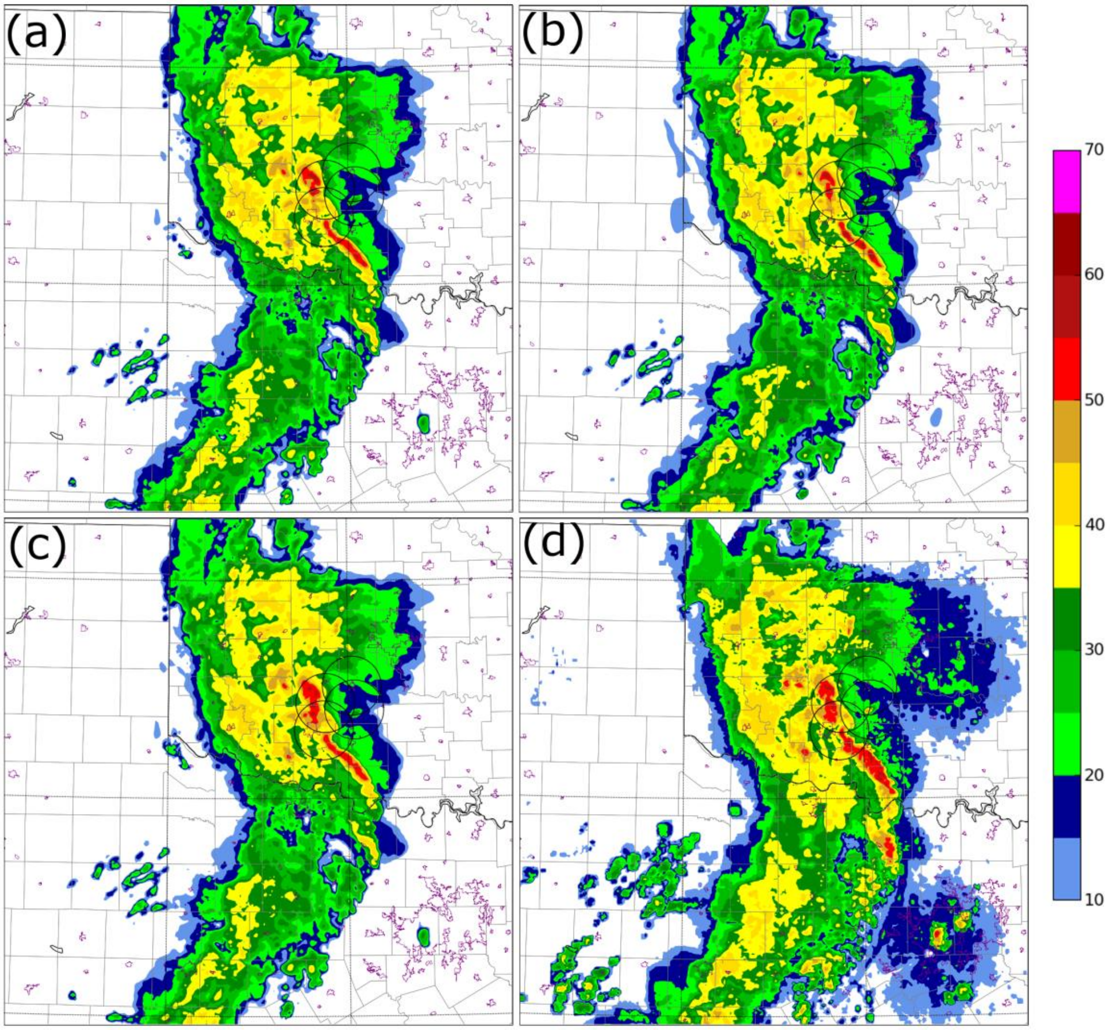

Fig. 3.4: Composite radar reflectivity (dBZ) of the final analysis state at 0200 UTC 9 May 2007 for (a) CNTL, (b) NoMMP, and (c) NoCASA; also (d) composite radar reflectivity mosaic (dBZ) observed by WSR-88D radars KAMA, KDYX, KFWS, KLBB, and KTLX at 0200 UTC 9 May 2007. 30 km CASA radar range rings are included for reference.

Despite the overall qualitative similarity noted in Fig. 3.4, there are important differences between the three experiments that occur throughout the assimilation period. Both the inclusion of CASA data and the use of a multi-microphysics ensemble produce notable changes in the forecast and analysis states. We will now address both of these factors in turn, beginning with the impact of assimilating CASA radar data in addition to WSR-88D radar data.

3.3.1 Impact of assimilating CASA data

CASA seeks to improve storm-scale analyses and forecasts by sampling the near-surface flow at high resolution. Experiment NoCASA is designed to evaluate the impact of withholding CASA data during assimilation. While CNTL and NoCASA produce qualitatively similar reflectivity fields (Fig. 3.5), the impact of assimilating CASA data can be seen in horizontal wind fields of CNTL and NoCASA, particularly in lower levels of the atmosphere (Fig. 3.5); strong southerly and southeasterly flow is present at one kilometer above mean sea level (approximately 700 m above the surface) within the northern portion of the leading convective line in CNTL (located in the region shared by the two western CASA radars, Fig. 3.5a), while the corresponding flow in NoCASA is much weaker (Fig. 3.5c). In addition, in experiments assimilating CASA data (Fig. 3.5a, b), in the southern portion of the CASA domain, southwest winds are present within and just ahead the convective line just to the south of the notch in the line near the southwestern-most CASA radar (KLWE). By contrast, this notch is less noticeable in NoCASA (Fig. 3.5c), and the low-level winds in the region are from the southeast. These differences represent the accumulated effects of assimilating CASA data. This result agrees well with results reported by Schenkman et al. (2011), who found that assimilation of CASA V_r data for this case using a 3DVAR and cloud analysis package had a strong impact on low-level winds and gust front structure.

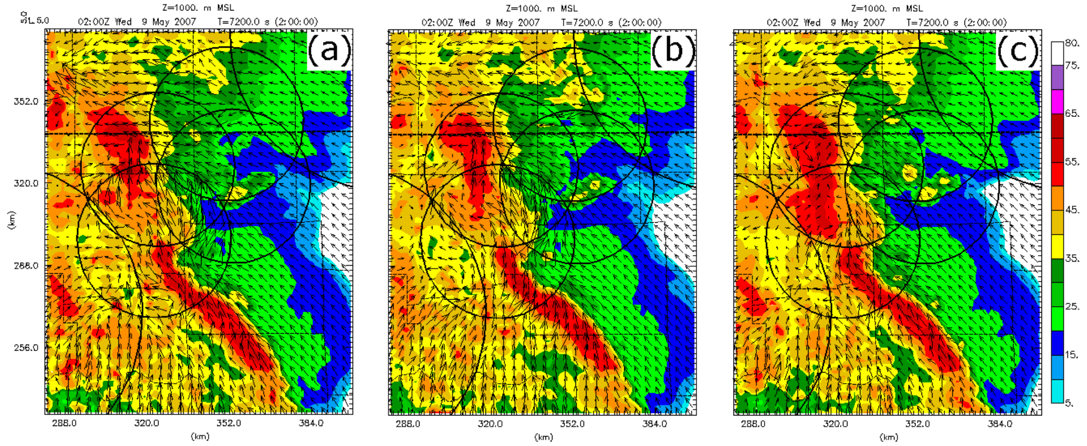

Fig. 3.5: Plots of radar reflectivity (color-shading) and the horizontal wind field (vectors) at 1 km above ground level for the final analysis states of (a) CNTL, (b) NoMMP, and (c) NoCASA at 0200 UTC. 30 km CASA range rings are also shown for reference.

The assimilation of CASA data results in a marked increase in maximum vertical vorticity in the lowest several kilometers of the atmosphere that sets the stage for tornadic processes. Fig. 3.6 shows the time-height cross-section of maximum vertical vorticity within a box tightly surrounding the CASA domain (depicted in Fig. 2a), for experiments CNTL and NoCASA. In CNTL, where CASA data were assimilated, much higher maximum values of low-level vertical vorticity are consistently present within this domain in both the forecast priors and EnKF analyses as compared to NoCASA. In particular, strong vertical vorticity is present in CNTL between 5100 and 6600 s of forecast time (0125 to 0150 UTC) between the surface and the 3 km level; a much weaker maximum is also present in NoCASA, but it is not discernible until 5400 s (0130 UTC).

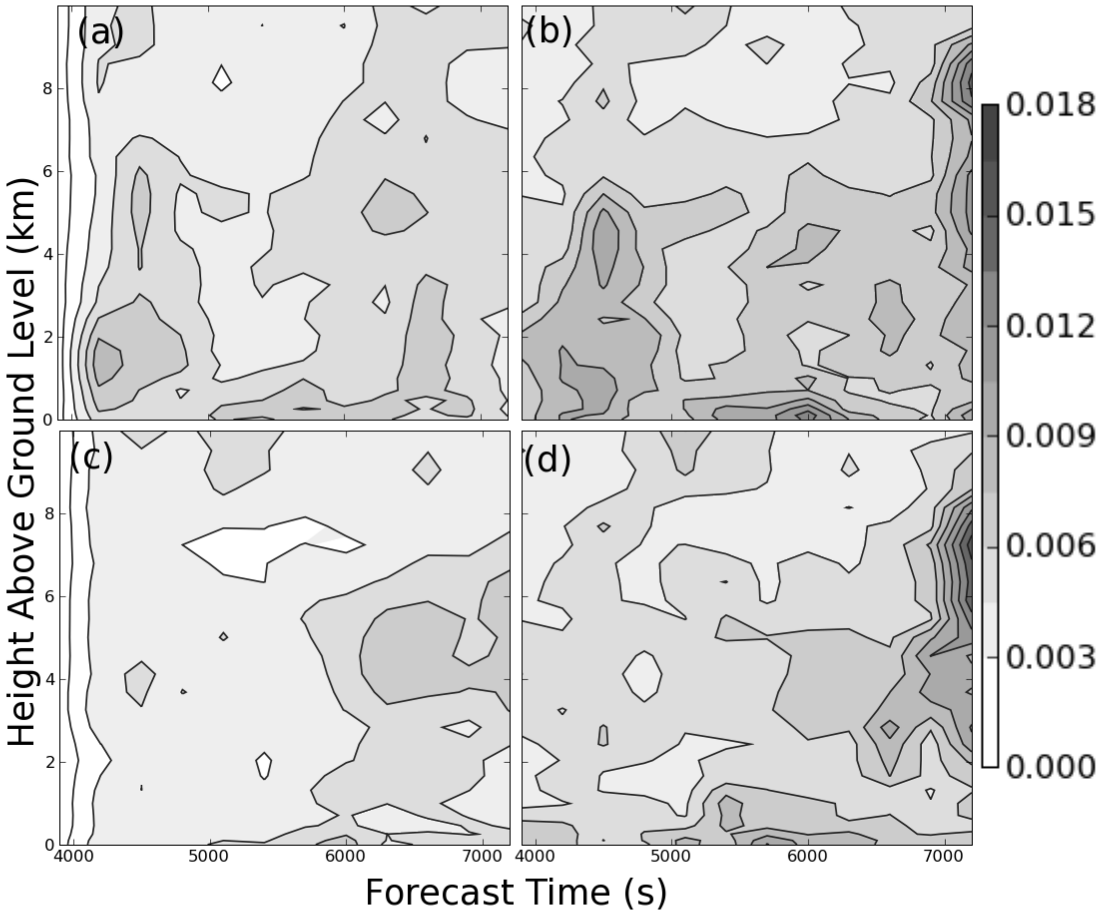

Fig. 3.6: Time-height plot of maximum vertical vorticity (s-1) for the forecast (a) and analysis (b) of CNTL, and the forecast (c) and analysis (d) of NoCASA during the analysis period. Time is denoted in seconds since forecast initialization and ranges from 3900 s (01:05 UTC) to 7200 s (02:00 UTC). Height is shown in kilometers above ground level.

Between 0125 and 0150 UTC, a strong low-level circulation is present west-southwest of KTLX, within the CASA domain. This circulation is visible in CASA and WSR-88D radar observations between 0120 and 0150 UTC, but is much better resolved by CASA radars due in large part to shorter range (not shown). The NWS Norman forecast office issued a tornado warning for the storm cell containing this circulation at 0126 UTC, although a later storm survey found no evidence of an actual tornado at this particular time and location. Inclusion of CASA data resulted in the analysis of a stronger low-level rotation within the CASA domain, matching more closely with the observed evolution of the MCS. In CNTL, where CASA data were assimilated, this circulation is present throughout the time it was observed by radar (Fig. 3.6a); by contrast, NoCASA is slower in developing such a circulation and the resulting low-level vertical vorticity is weaker (Fig. 3.6c).

The tornado-warned meso-vortex is well-observed by CASA radar KCYR. The 2° elevation V_r observations from KCYR at 0140 UTC (Fig. 3.7a) show a strong circulation present between 5 and 20 km to the west of the radar site, with 45 m s^{-1} of horizontal wind shear over a distance of approximately 12 km. Simulated KCYR V_r observations from the 0140 UTC ensemble mean analysis of CNTL (Fig. 3.7b) also indicate the presence of a meso-vortex circulation which closely matches the location and size of that seen in the KCYR observations, but with a slightly weaker maximum horizontal shear of around 40 m s^{-1} across the vortex. Simulated KCYR V_r observations from NoCASA (Fig. 3.7c) show no strong circulation at the 2° elevation; instead, a convergent signature with only very weak rotation is present near the location of the observed meso-vortex. In addition, NoCASA (Fig. 3.7c) greatly underestimates the region of positive V_r observed to the northwest of KCYR (Fig. 3.7a); by contrast, the ensemble mean analysis of CNTL (Fig. 3.7b) indicates a flow that closely matches KCYR observations. These results highlight the importance of assimilated near-surface CASA radar data in accurately capturing the near-surface flow in this convective system.

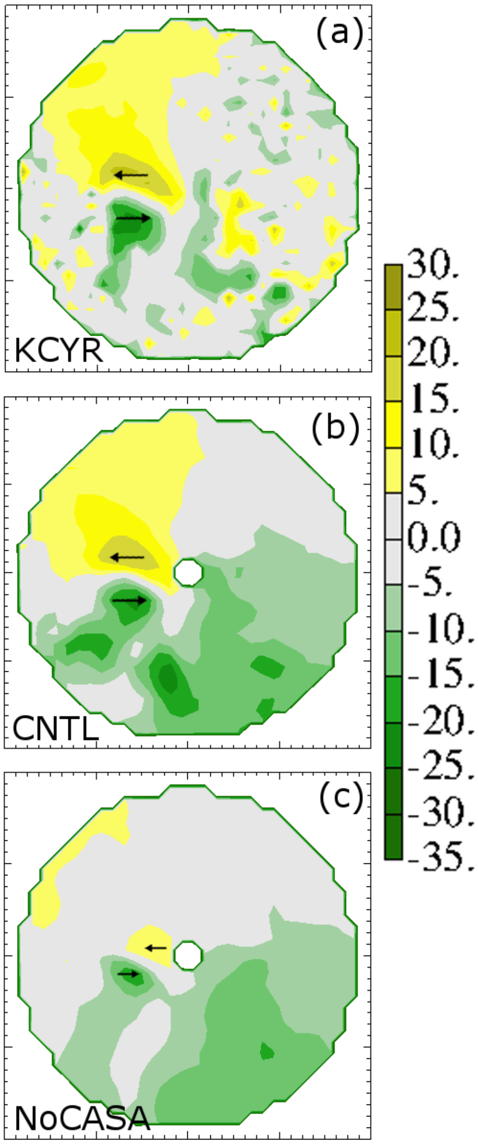

Fig. 3.7: Radial velocity for the 0140 UTC at the 2 degree elevation (a) as observed by CASA radar KCYR, and simulated from the 0140 UTC EnKF analyses of (b) CNTL and (c) NoCASA. The arrows highlight the circulation in the region of the tornado-warned mesovortex.

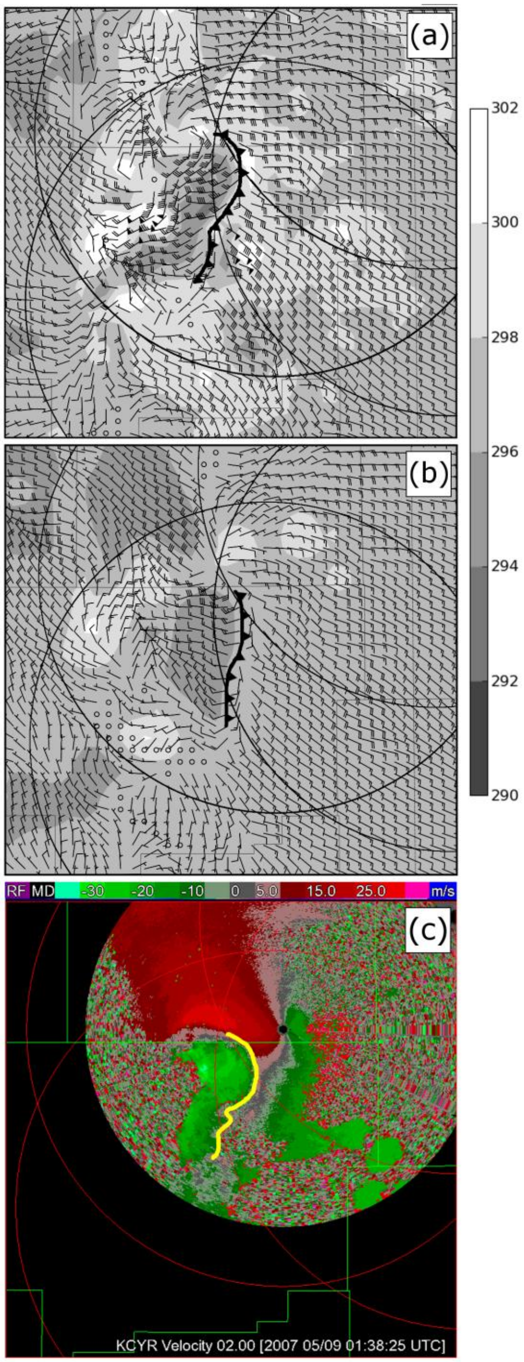

Fig. 3.8: Horizontal winds (barbs) and potential temperature (shaded) at the first model level above the surface near CASA radar KCYR for the 0140 UTC ensemble mean analysis in (a) CNTL and (b) NoCASA . The position of a gust front associated with the embedded mesovortex is indicated. Also shown are (c) full-resolution radial velocity observations from the 2° elevation of CASA radar KCYR shortly before 0140 UTC. The gust front position indicated by the radial velocity observations is indicated by the yellow line.

The impact of assimilated CASA data is also evident in the surface wind field and cold pool structure in the 0140 UTC analysis. In experiment CNTL (Fig. 3.8a), a moderately intense surface circulation is present, horizontally co-located with that indicated by KCYR V_r observations. In the CNTL analysis, a moderately strong gust front is present to the south and southeast of the surface circulation, with strong inflow of between 15 and 25 ms^{-1} ahead of the gust front. The location of the gust front in the 0140 UTC CNTL analysis (Fig. 3.8a) is similar to that indicated by the full-resolution 0139 UTC KCYR 2° V_r observations (Fig. 3.8c); at the location of the meso-vortex, these observations were 500 to 700 m above the surface. The surface wind field in the 0140 UTC ensemble mean analysis of NoCASA shows only weak rotation within a convergent shear zone (Fig. 3.8b), consistent with the simulated V_r observations of Fig. 3.7c. While the gust front present in NoCASA is positioned similarly to that in CNTL, it is much weaker, with a cross-frontal temperature difference of less than 2 K; this is too weak compared to potential temperature decreases of 3 to 4 K as measured by nearby Oklahoma Mesonet stations during passage of the gust front (not shown).

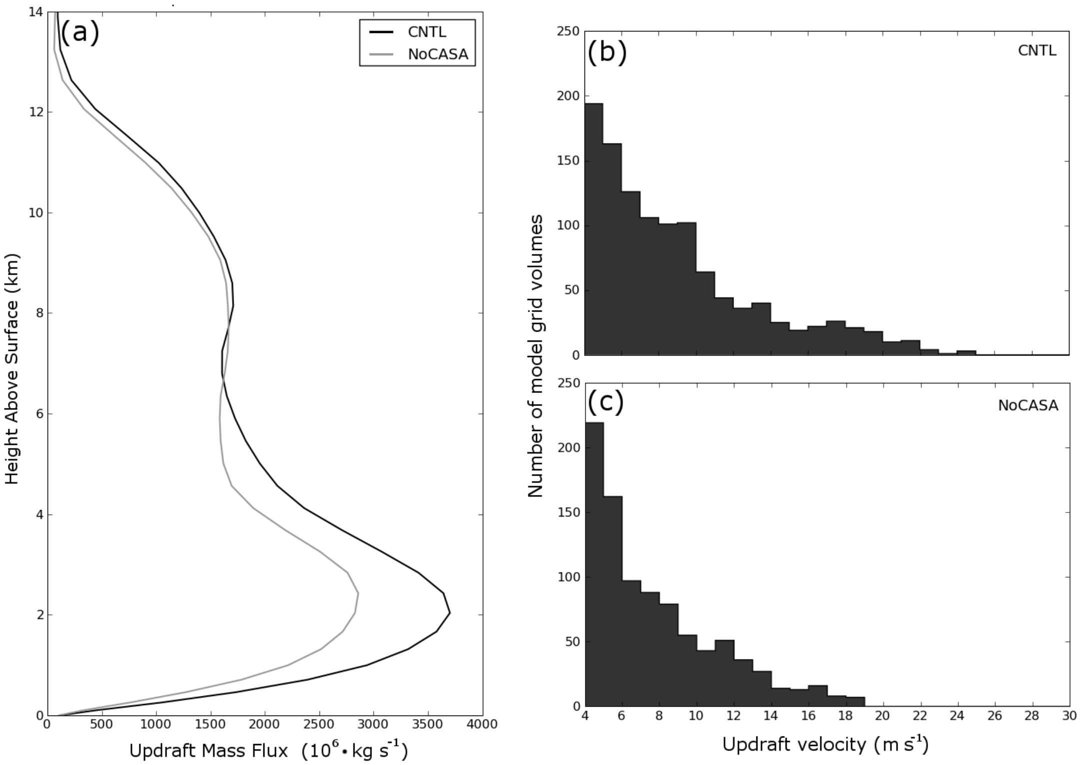

Fig. 3.9: (a) Total updraft mass flux profiles within the CASA sub-domain (as denoted in Fig. 3.3a) at 0140 UTC for experiments CNTL and NoCASA. Histograms of vertical velocity exceeding 4 m s^{-1} within updraft regions in the CASA sub-domain are plotted for (b) CNTL and (c) NoCASA.

The stronger low-level circulation of CNTL is accompanied by more vigorous convective updrafts over the CASA sub-domain. Total updraft flux is calculated at each model level over the CASA sub-domain outlined in Fig. 3.3a; the resulting vertical profiles of updraft flux for the CNTL and NoCASA 0140 UTC ensemble mean analyses are plotted in Fig. 3.9a. Greater updraft flux is present in CNTL than in NoCASA, particularly below the 5 km level. Much of the difference in updraft flux between CNTL and NoCASA can be attributed to greater updraft velocities in CNTL; histograms of updraft velocity for the 0140 UTC analyses of CNTL (Fig. 3.9b) and NoCASA (Fig. 3.9c) in model grid cells where the vertical velocity was greater than or equal to 4 m s^{-1} indicate that more regions of strong updrafts are present in CNTL than in NoCASA. In the 0140 UTC CNTL analysis, updrafts in excess of 16 m s^{-1} are present in more than 100 grid cells; the maximum updraft velocity observed within the CASA sub-domain exceeds 24 m s^{-1}. In NoCASA, only about 30 grid cells have updrafts exceeding 16 m s^{-1}, and the maximum updraft velocity within the CASA domain is less than 19 m s^{-1}. Similar behavior was noted at other analysis times and during the forecast cycles, with stronger updrafts and greater updraft fluxes present in CNTL than in NoCASA (not shown).

To more quantitatively assess the behavior of the EnKF analyses, average root-mean-square (RMS) values of observation innovation (the difference between observations and the model state in the form of observed quantities) and ensemble spread are examined. Observation innovations and ensemble spread are calculated for each of the 4 CASA radars, as well as WSR-88D radars KTLX and KVNX, for radar reflectivity (Z) (Fig. 3.10) and radial velocity (V_r) (Fig. 3.11), in experiments CNTL and NoCASA. Innovations in Fig. 3.10 and Fig. 3.11 are calculated for the ensemble mean fields at locations where either observed or model reflectivity is greater than or equal to 15 dBZ. The calculation is further limited to within the CASA sub-domain (c.f., Fig. 3.3a). In NoCASA, RMS innovations for the CASA radars are calculated against CASA data that were not assimilated; these observations are therefore from independent sources. Nevertheless, EnKF data assimilation in NoCASA was able to decrease the average innovations at all CASA sites for Z during every assimilation cycle, and for V_r during almost every assimilation cycle (Fig. 3.10, Fig. 3.11). Given that different radars measure different components of the velocity field, the reduction in innovation against independent, unassimilated radial velocity measurements indicates good performance of the EnKF.

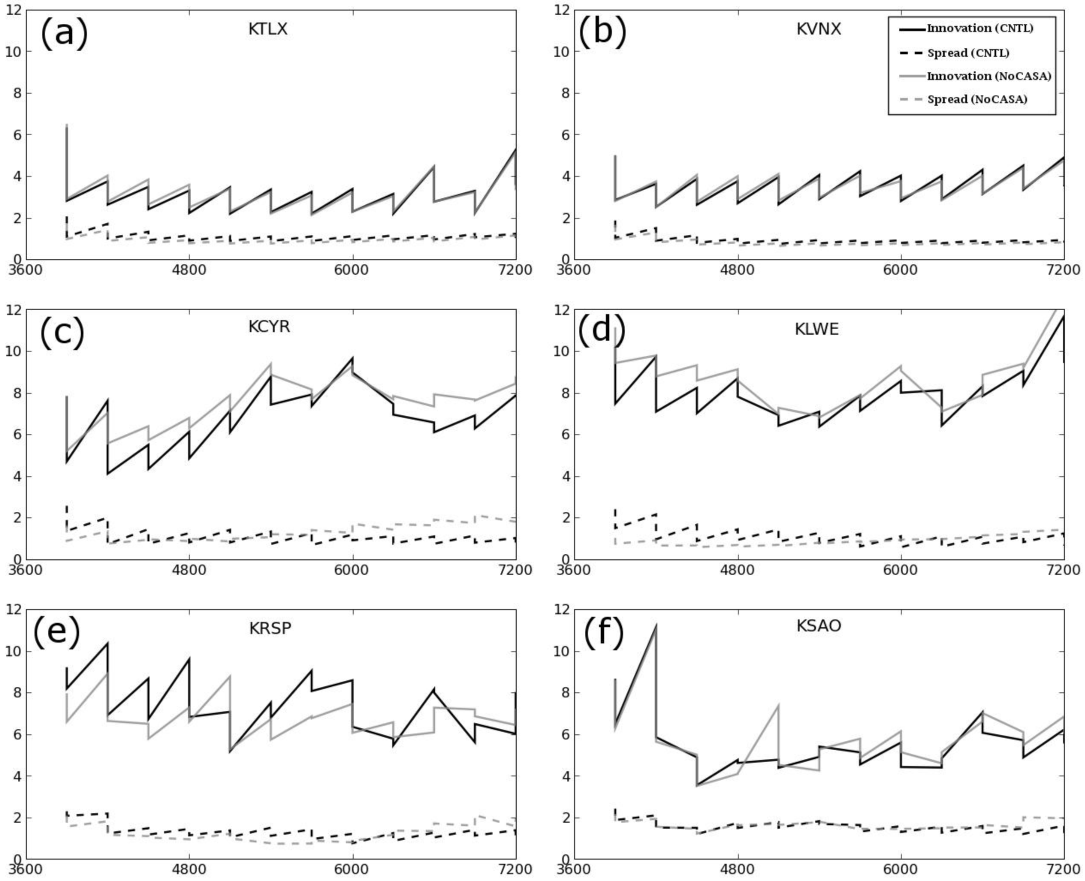

Fig. 3.10: Average root-mean-square (RMS) innovation (solid lines) and spread (dotted lines) of radial velocity (in m/s) for each of the 5 WSR-88D and 4 CASA radars for experiments CNTL (black lines) and NoCASA (gray lines) calculated every 5 minutes during the assimilation period. The assimilation period lasts from 01:00 UTC (3600 seconds of forecast time) to 02:00 UTC (7200 seconds of forecast time).

Assimilation of CASA data resulted in a slight but notable decrease in RMS innovation in the analysis of *Z* in CNTL as compared to NoCASA (Fig. 3.11). The overall decrease was greater for the CASA radars, due to the absence of CASA data in NoCASA. Among the WSR-88D radars, only KTLX and KVNX are included for the comparisons between CNTL and NoCASA in Fig. 10-12 because they are located close to the center of the convective system during the assimilation period and share the greatest overlap with the CASA domain. The RMS innovations of V_r at KTLX and KVNX differ little between CNTL and NoCASA (Fig. 3.10a, b), though the RMS innovation of the CNTL analysis is very slightly lower than that of NoCASA at KTLX during early assimilation cycles (Fig. 3.10a). In contrast, a larger improvement is seen in Z for CNTL during early cycles for KTLX and early to middle cycles for KVNX as compared to NoCASA (Fig. 3.11). These results suggest that the inclusion of CASA data modestly improved the analyzed reflectivity field within the model, particularly during early assimilation cycles, with less improvement to the analysis of radial velocity.

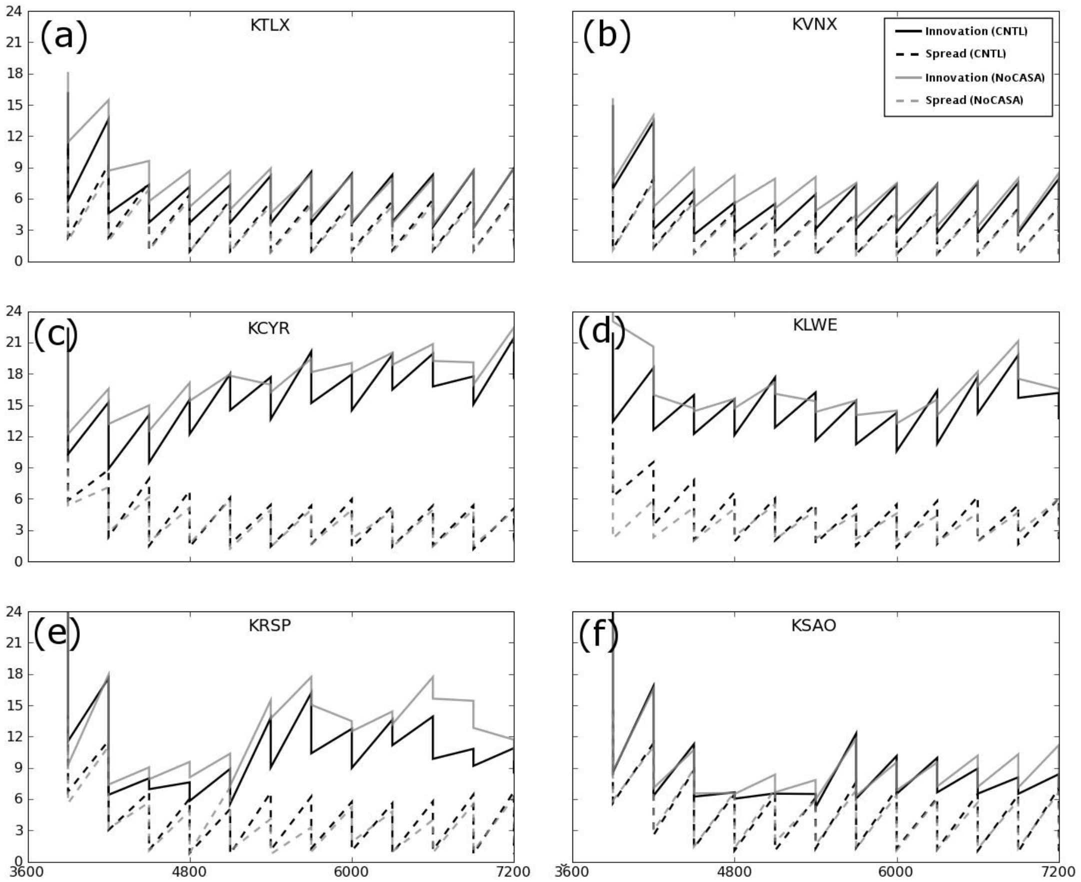

Fig. 3.11: As Fig. 3.10, but for radar reflectivity (in dBZ) instead of for radial velocity.

For the WSR-88D sites (KTLX and KVNX), the greatest differences in RMS innovations of Z and V_r between CNTL and NoCASA occurred in the first six assimilation cycles (Fig. 3.10a, b; Fig. 3.11a, b). In addition, fewer cycles were needed for the analysis to reach its minimum RMS innovation value for Z in CNTL than in NoCASA. While the minimum RMS innovation of the analysis for Z was not reached until around the 8th assimilation cycle at KTLX (Fig. 3.11a) and the 11th assimilation cycle for KVNX (Fig. 3.11b) in NoCASA, the RMS innovation of the analysis reached its minimum value for these radars in CNTL by the 4th and 3rd cycles respectively (Fig. 3.11a, b). Assimilation of CASA data reduces the number of cycles needed for the EnKF analysis to reach a relatively stable and low level of RMS innovation in *Z*.

One important measure of the performance of an EnKF data assimilation system is statistical consistency, as discussed in Snyder and Zhang (2003) and Dowell et al. (2004a). For forecasts and observations with independent error characteristics, the variance of the innovation should be equal to the sum of the observation and forecast error variances:

$$\sigma_d^2 = \sigma_o^2 + \sigma_f^2 \tag{3.1}$$

Following Dowell et al. (2004a), we arrive from equation 3.1 at a consistency relation valid for observations y^o and model forecast state x^f, with angle brackets representing an average over all available observations at a time, and overbars denoting an ensemble mean:

$$\frac{\sigma_o^2 + \left\langle \frac{1}{N-1} \sum_{i=1}^{N} \left[H(x_i^f) - \overline{H(x^f)} \right]^2 \right\rangle}{\left\langle \left[\left(y^o - \overline{H(x^f)} \right) - \left\langle \left(y^o - \overline{H(x^f)} \right) \right\rangle \right]^2 \right\rangle} \approx 1 \tag{3.2}$$

Here, *N* is the ensemble size, *i* is the ensemble index, and *H* is the observation operator. In practice, values of consistency ratio well below 1 are often seen in EnKF studies (Dowell et al. 2004a; Dowell et al. 2004b), indicating a general tendency for under-dispersion in the ensemble.

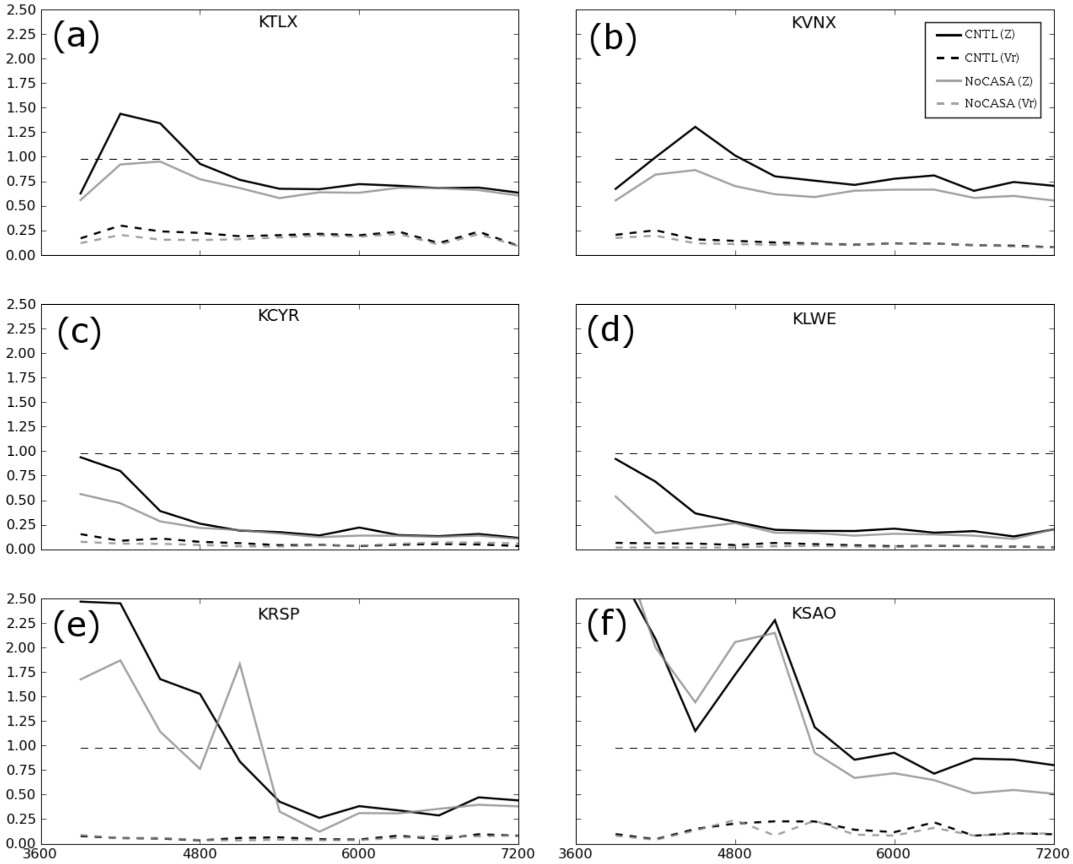

Fig. 3.12: Consistency ratio of Z (solid lines) and Vr (dotted lines) for the 4 CASA radars and the 2 WSR-88D radars nearest to the CASA radar network for CNTL (black lines) and NoCASA (gray lines) calculated every 5 minutes during the assimilation period (0100 to 0200 UTC). The thin black dashed line indicates the theoretically-expected consistency ratio defined in equation (3.2).

Time series of consistency ratio for CNTL and NoCASA, calculated during the assimilation period for four CASA and two WSR-88D (KTLX and KVNX) radars are shown in Fig. 3.12. Values of consistency ratio for V_r and Z in both CNTL and NoCASA fall below the optimal value of approximately 1 throughout much of the period (Fig. 3.12) with the exception being for Z in early cycles at the WSR-88D radar sites (Fig. 3.12a, b) and CASA sites KRSP and KSAO (Fig. 3.12e, f). Consistency ratio for Z was much higher at the WSR-88D radar sites than at CASA sites in both CNTL and NoCASA. Lower values of consistency ratio were observed for V_r than for Z, with V_r consistency ratio values of between 0.1 and 0.3 common for the CASA radars; WSR-88D sites KTLX and KVNX yielded V_r consistency ratios ranging from 0.2 to 0.3. Consistency ratios for Z were higher, ranging between 0.5 and 1.0 for WSR-88D radars, and 0.2 and 1.0 for CASA radars. Very high values (greater than 2.0) of consistency ratio for Z were present during the first few assimilation cycles due to the very high values of RMS ensemble spread for Z at these times (see Fig. 3.12).

Values of consistency ratio in experiments NoCASA and CNTL are slightly lower than those seen in previous real data studies using a similar EnKF setup, such as Dowell et al. (2004b). One can infer from the particularly low values of consistency ratio seen for V_r (Fig. 3.12) that a significant amount of under-dispersion exists in the radial velocity field in both NoCASA and CNTL. In this study, we assumed an observation error standard deviation of 1 m s^{-1} for radial velocity observations. The relatively small assumed observation error may be a contributing factor in the low values of consistency ratio observed. In future studies, we will consider increasing the assumed observation error to 2 m s^{-1} for V_r. Values of consistency ratio for Z are also below 1, suggesting insufficient ensemble spread in the reflectivity field, but this deficiency is not as severe as that in the radial velocity field. Dowell et al. (2009) addressed under-dispersion in radial velocity by using additive perturbations to the horizontal wind field; however, initial tests for this case including additive perturbations to the wind field did not show improvement in RMS innovation for radar reflectivity and radial velocity observations when compared against analyses using multiplicative covariance inflation alone; further tests using perturbations with different perturbation magnitudes and scales will be explored in future work on this case.

The difference in the consistency ratio time series (Fig. 3.12) between CASA and NoCASA for Z (and to a lesser extent V_r) indicates that the under-dispersion is slightly less severe in CNTL than in NoCASA, particularly during early assimilation cycles and at the WSR-88D radar sites (Fig. 3.12a, b). Assimilation of CASA data slightly decreases under-dispersion of radar reflectivity within the ensemble; this is a somewhat counter-intuitive result, as increasing the amount of data assimilated usually results in decreased spread within the ensemble. This is likely to be due to the way that initial perturbations are added. In this study, initial perturbations were only added to grid points within 2 km in the horizontal and 1 km in the vertical of observed radar reflectivity exceeding a threshold of 5 dBZ, following the methodology of Tong and Xue (2005). Because CNTL includes CASA data in addition to WSR-88D radar data, the region containing initial perturbations is slightly larger in CNTL than in NoCASA, particularly at low-levels where only CASA radar data is available. Accordingly, the initial difference in RMS spread is greater for CASA radars and very small for WSR-88D radars (Fig. 3.10). The effect of this slight difference in the initial perturbation region fades as assimilation cycles are performed; by the end of the assimilation window CNTL shows smaller spread in the later cycles due to faster spread reduction, as expected when assimilating more observations.

3.3.2 Impact of using a mixed-microphysics ensemble

In previous studies, using different parameterization schemes among ensemble members (e.g., Meng and Zhang 2007) and including perturbations of microphysical parameters within the ensemble (e.g., Ge et al. 2010) have been shown effective in increasing ensemble spread and reducing under-dispersion within the ensemble. However, the use of multiple microphysics schemes for real-case storm-scale radar data assimilation has, to our knowledge, not been reported in the literature. In this section, different microphysics schemes are used among ensemble members and the effect on the analysis is investigated. Experiment NoMMP was performed to evaluate the effect of using a mixed-microphysics ensemble; NoMMP differed from CNTL only in that it used Lin microphysics for all members in the ensemble forecast (see Table 3.1).

Time-series of RMS innovation and spread during assimilation for experiments CNTL and NoMMP are presented in Fig. 3.13 for V_r and Fig. 3.14 for Z. Since the impact of the mixed-microphysics ensemble is present throughout the model domain, RMS spread and innovation calculations were not limited to the CASA sub-domain (see Fig. 3.3a) for comparisons between CNTL and NoMMP. Thus, unlike in the comparison between CNTL and NoCASA, calculations are presented for all five WSR-88D radars in addition to the four CASA radars; data from all these radar sites were assimilated in both CNTL and NoMMP.

The impact of the mixed-microphysics ensemble on RMS innovation of V_r (Fig. 3.13) is relatively small. The RMS innovation of the V_r analysis of CNTL is slightly lower than that of NoMMP at KTLX during the first five assimilation cycles

(Fig. 3.13d); however farther to the west at radar site KAMA (Fig. 3.13a), NoMMP actually produces a slightly lower RMS innovation for V_r during later cycles of the assimilation period. At most sites no significant difference in RMS innovation of V_r can be seen. Likewise, RMS ensemble spread of V_r is virtually unchanged between CNTL and NoMMP.

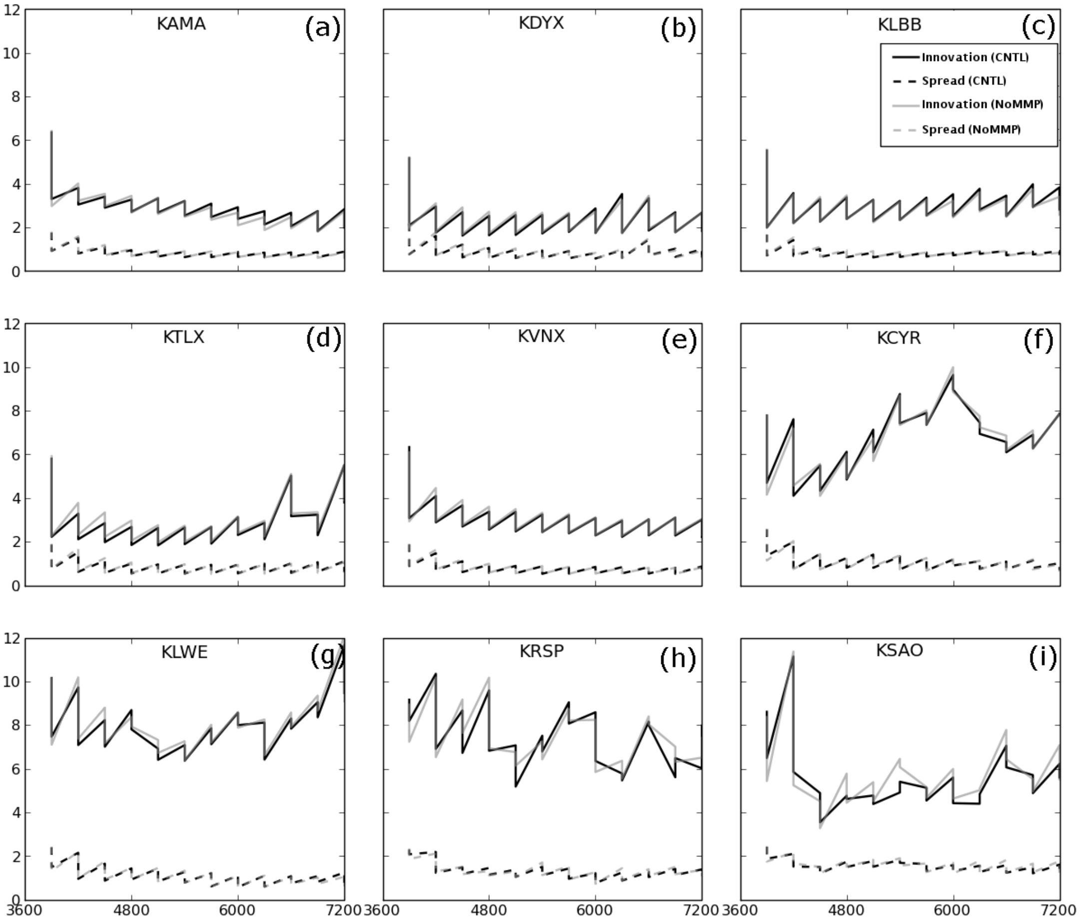

Fig. 3.13: As Fig. 3.10, but for experiments CNTL (black lines) and NoMMP (gray lines).

In contrast to V_r, differences between the RMS innovation and ensemble spread of Z in NoMMP and CNTL (Fig. 3.14) are much more prominent. Compared to NoMMP, ensemble spread of Z in CNTL grows faster during the forecast step and remains higher during the analysis step; greater ensemble spread is consistently present in CNTL during forecasts and analyses than in NoMMP at every radar site. Average RMS ensemble spread of *Z* during the forecast step decreases in the first several cycles and remains largely constant during the remainder of the assimilation period. Average RMS ensemble spread values for *Z* at the end of assimilation period range between about 3 to 5 dBZ in CNTL and between 1 and 4 dBZ in NoMMP.

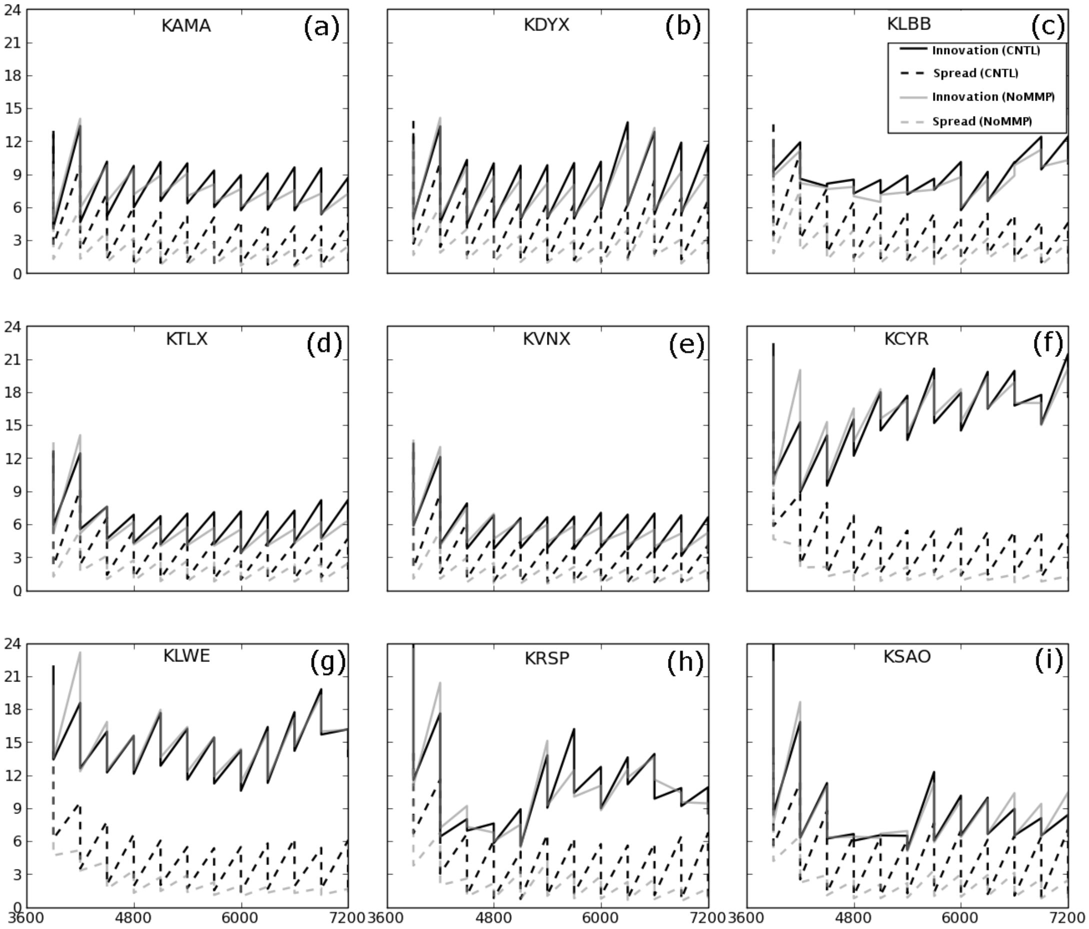

Fig. 3.14: As Fig. 3.13, but for experiments CNTL (black lines) and NoMMP (gray lines).

Differences in the RMS innovation values of Z between CNTL and NoMMP (Fig. 3.14) are also noticeable but not as prominent as differences in RMS ensemble spread. Error in the forecast ensemble grows more quickly in CNTL than in NoMMP, evidenced by a steeper increase between each analysis and the subsequent forecast at every WSR-88D radar site, as members using different microphysics schemes arrive at varying solutions because of differences in treatment of microphysics processes. The faster growth of RMS innovation in CNTL (Fig. 3.14) can be attributed in part to variation in reflectivity formulation between the Lin, WSM, and NEM microphysical schemes—for this case, the NEM microphysics scheme greatly under-predicts the coverage of stratiform rain, thus members using the NEM microphysics scheme within the CNTL ensemble act to increase the RMS innovation during the forecast cycles. When innovation statistics for Z were derived for subsets of CNTL members using individual microphysical schemes, the subset consisting of NEM members within CNTL had the most rapid increase in RMS innovation of Z during forecast steps, while the subset consisting of Lin members within CNTL had the slowest increase (not shown). However, despite the higher RMS innovation values of Z present during the forecast step in CNTL, the RMS innovation of the analysis of Z in CNTL is equal to or lower than that of NoMMP for almost every analysis cycle at all radar sites. The greatest differences can be seen at KAMA and KVNX, where CNTL produces analyses of Z with an average RMS innovation of between 0.3 and 1 dBZ lower than corresponding analyses in NoMMP for most of the assimilation period. At the CASA radar sites differences between

CNTL and NoMMP are more difficult to discern; at these sites the two experiments produced qualitatively similar RMS innovation and ensemble spread time series.

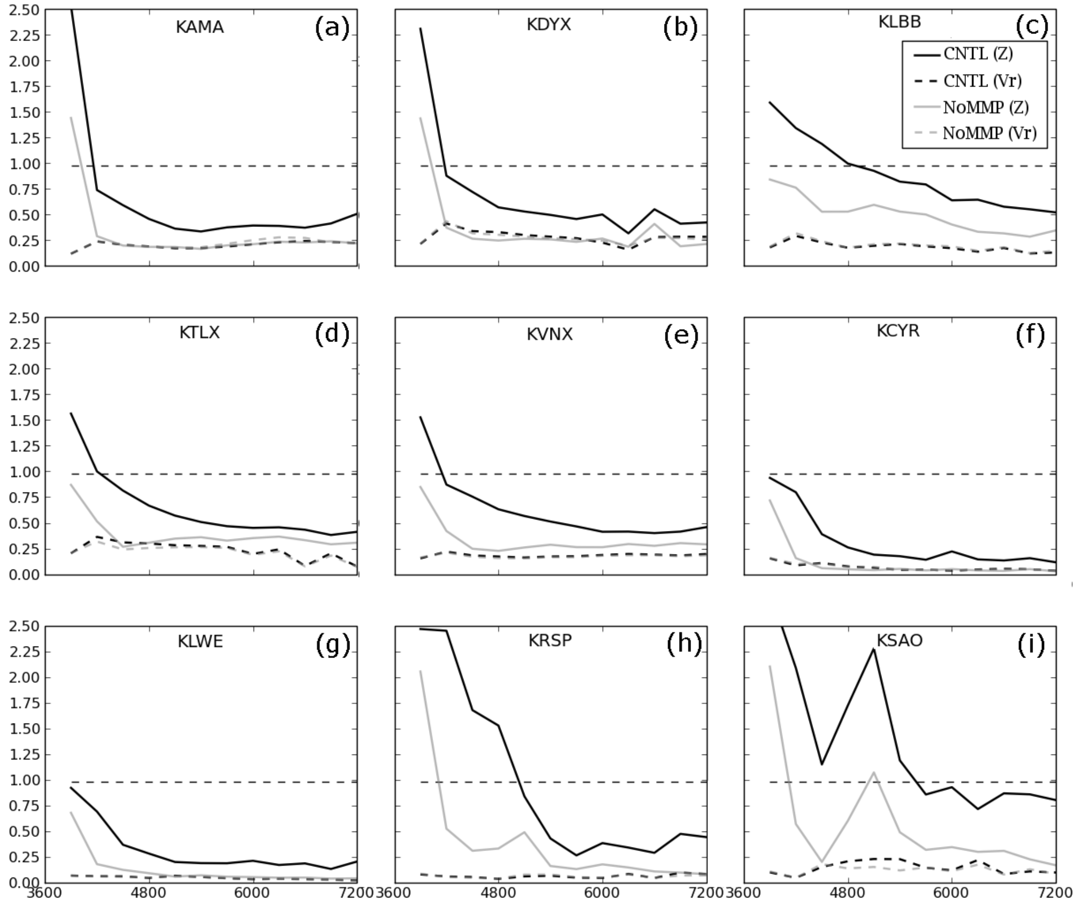

Fig. 3.15: As Fig. 3.12, but for experiments CNTL (black lines) and NoMMP (gray lines).

Comparison of consistency ratio calculated for Vr and Z for experiments CNTL and NoMMP (Fig. 3.15) reveal that use of the mixed-microphysics ensemble results in a higher consistency ratio than the single-microphysics ensemble for Z because of increased ensemble spread of radar reflectivity in the mixed-microphysics case. In both CNTL and NoMMP the consistency ratio of Vr is well below 1.0, ranging between 0.25 and 0.5 for WSR-88D radars and 0.1 and 0.25 for CASA radars. While the consistency ratio of Vr is virtually unchanged between CNTL and NoMMP, the consistency ratio of Z is considerably higher in CNTL than in NoMMP at all radar sites throughout the assimilation period. Though the consistency ratio of Z for CNTL still remains below the optimal value of 1.0 at most radar sites, particularly late in the assimilation period, the higher consistency ratio values for Z in CNTL suggest that CNTL exhibits significantly less under-dispersion than NoMMP (Fig. 3.15).

Chapter 4: 8-9 May 2007—Forecasts from EnKF Radar Analyses

Since the inception of explicit numerical weather prediction (NWP) of severe convective storms (Lilly 1990), assimilation of Doppler weather radar data has been shown to be critical and often effective for initializing such model predictions (e.g., Sun et al. 1991; Sun and Crook 1998; Xue et al. 2003; Hu et al. 2006a). Recent studies have produced promising results assimilating Doppler radar data for convective-scale NWP in real-time and over large domains (e.g., Xue et al. 2008). The ensemble Kalman filter (EnKF) technique, initially developed by Evensen (1994, 2003), has been gaining popularity as an effective method of radar data assimilation for storm-scale NWP (e.g., Snyder and Zhang 2003; Dowell et al. 2004b; Tong and Xue 2005a; Tong and Xue 2008).

Though EnKF is more computationally expensive than the 3-dimensional variational method (3DVAR) widely used operationally, it provides flow-dependent multivariate background error covariances and cross-covariances that 3DVAR cannot. Such cross-covariances are essential for radar data assimilation, because most state variables are not directly observed (Tong and Xue 2005b, 2008). Additional comments on the relative merits of various radar data assimilation methods including 3DVAR, 4DVAR (4-D variational) and EnKF for convective storm analysis can be found in Tong and Xue (2005b). As available computational power increases, it will soon become feasible to run a real-time convective-scale forecast system which assimilates data via EnKF (Zhang et al. 2009) and produces convective-scale ensemble forecasts (e.g., Xue et al. 2008).

Due to the chaotic nature of the atmosphere, and inevitable errors in observations and prediction models, weather forecasts always contain uncertainty. No forecast is therefore complete without a description of its uncertainty (NRC 2006), which is often expressed in terms of forecast probability. Ensemble forecasting offers a practical way to provide a probabilistic forecast (Leith 1974). Global and regional ensemble forecasting has been operational for nearly two decades (e.g., Toth and Kalnay 1993; Houtekamer et al. 1996; Du et al. 2003; Bowler and Mylne 2009); by comparison, convective-scale ensemble forecasting is still in its infancy (Kong et al. 2006; Xue et al. 2011). Convective scale weather poses a greater prediction challenge due to its intermittent nature, smaller spatial and temporal scale, higher nonlinearity, and often due to incomplete observation coverage; these challenges increase the forecast uncertainty, making probabilistic forecasting even more crucial (Stensrud et al. 2009; Xue et al. 2011).

EnKF provides a set of analyses that, in principle, best characterize the analysis uncertainty, making them desirable initial conditions for ensemble forecasts (and ensemble-based probabilistic predictions). At the global scale, ensembles using EnKF analysis initial conditions have shown superior probabilistic forecasting performance compared to those using more traditional perturbation methods (Houtekamer et al. 2005; Hamill et al. 2011). EnKF methods have proven effective in generating dynamically consistent wind, temperature, and microphysical fields for convective storms when assimilating Doppler radar reflectivity and radial velocity data (e.g., Dowell et al. 2004b; Tong 2006; Snook et al. 2011) but probabilistic forecasts at the convective scale using EnKF analyses has so far received limited attention. Zhang et al. (2010) is an example where convection-permitting-resolution ensemble forecasts of a tropical cyclone were initialized from global EnKF analyses.

In this chapter, we detail the results of deterministic and ensemble forecasts produced from the EnKF analyses of the 8-9 May 2007 tornadic mesoscale convective system presented above in Chapter 3, with the goal of evaluating the suitability of EnKF analyses of radar data for initializing an ensemble for the short-term convective-scale probabilistic forecast goals of “warn-on-forecast” (Stensrud et al. 2009). We use the neighborhood ensemble probability (NEP) approach (Schwartz et al. 2010) described in Chapter 2 to obtain probabilistic forecasts of radar reflectivity, and an object-based ensemble approach to obtain probabilistic forecasts of near-surface meso-vortices. The value of assimilating X-band radar data by the Engineering Research Center for Collaborative Adaptive Sensing of the Atmosphere (CASA) (McLaughlin et al. 2009) for improving the forecasts is evaluated, and the impact of microphysical parameterization during analysis and forecast periods is examined.

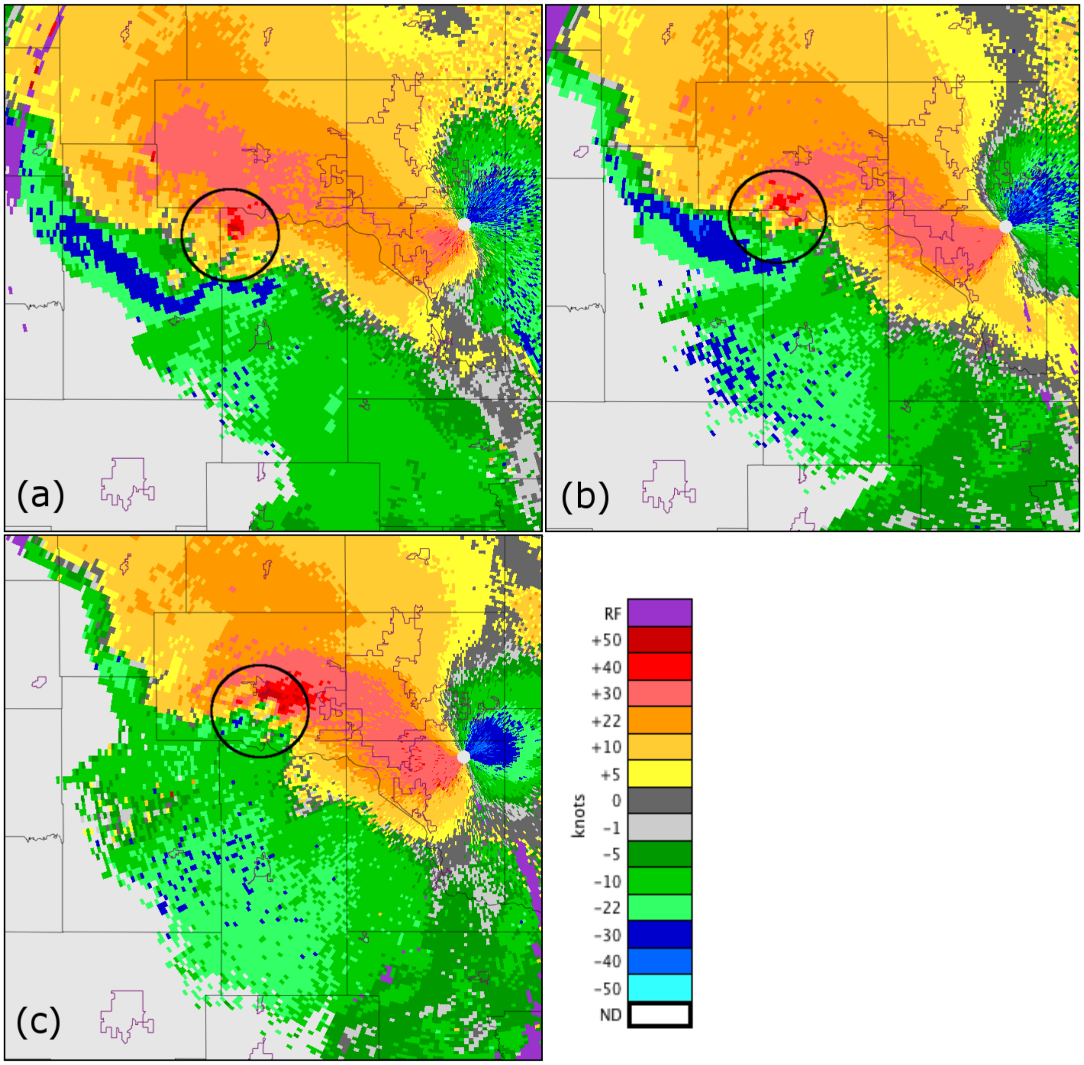

Fig. 4.1: Storm Relative Velocity as observed by the 1.5 degree elevation scan of WSR-88D radar KTLX at 04:00 UTC (upper left), 04:20Z (upper right), and 04:40 UTC (lower left). Urban and county boundaries are shown.

We will begin by analyzing the results of the deterministic, ensemble, and probabilistic forecasts of radar reflectivity (*Z*) for CNTL, NoMMP and NoCASA, and then move on to the results of ensemble and probabilistic forecasts of significant near-surface mesovortices. With regards to the choice of focus, *Z* was chosen because it can be directly verified against WSR-88D radar observations over the entire area of the convective system, while low-level mesovortices were closely co-located with the observed tornadoes in this event, as can be noted from the match between the location of rotational signatures of mesovortices in the KTLX storm-relative velocity observations (Fig. 4.1) and the tornadoes reported in association with them (see Fig. 3.1). We will also address important differences between the deterministic forecast from the ensemble mean and the mean of the ensemble forecast. Finally, we will discuss the effects of variation of the model microphysical scheme, both for the deterministic forecasts from the 0200 UTC CNTL analysis and the single- and mixed-microphysics ensemble forecasts performed for CNTL, NoMMP, and NoCASA.

4.1 Forecast experiment setup and design

In the ensemble forecast experiments presented in this chapter, the EnKF ensemble analyses presented in Chapter 3, valid at 0200 UTC 9 May 2007, are used to initialize 3-hour ensemble forecasts. The model setup and naming convention used for these forecast experiments are identical to those presented in Chapter 3. The computational domain has 256 × 256 × 40 grid points with a 2 km horizontal grid spacing and stretched vertical grid spacing (see Fig. 4.2). Results from three ensemble forecast experiments are presented here. The control experiment (hereafter "CNTL") assimilates both CASA and WSR-88D radar data and contains 40 ARPS ensemble members; 16 of these use the Lin ice microphysical scheme (Lin et al. 1983), 16 use the WRF single-moment 6-class (WSM6) ice microphysics scheme (Hong and Lim 2006), and the remaining 8 members use the NWP explicit microphysics (NEM) scheme developed by Schultz (1995). The second experiment, NoMMP, assimilates the same data as CNTL but uses the Lin microphysics scheme in all 40 of its ensemble members. The third experiment, NoCASA, uses the same ensemble setup as CNTL but does not assimilate CASA data.

To isolate the impact of using a mixed-microphysics ensemble during the forecast period, two more ensemble forecast experiments are run in addition to the three mentioned above. Experiments CNTL_LIN and NoCASA_LIN are initialized from the CNTL and NoCASA initial conditions (respectively), but use a single-microphysics forecast ensemble consisting of 40 ARPS members using the Lin microphysics scheme, as in NoMMP. In all experiments, a reduced rain intercept parameter of 8×10^5 was used, consistent with Snook and Xue (2008), who found that reducing the rain intercept parameter yielded more realistic cold-pool structure. Lateral boundary conditions for all ensemble members are obtained from the NCEP NAM 6-hourly analyses and intervening 3 hour forecasts.

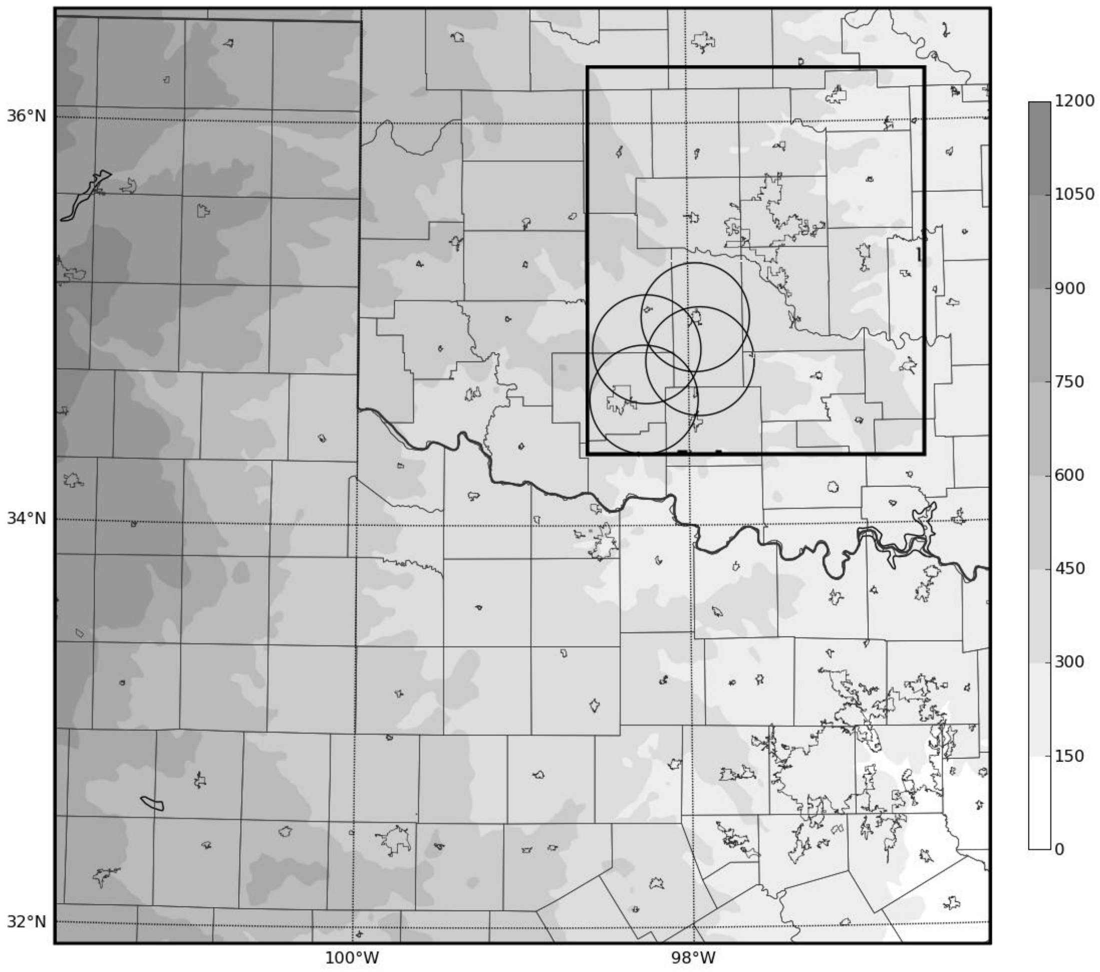

Fig. 4.2: Forecast domain used in forecast experiments for the 9 May 2007 case initialized from analyses obtained in the experiments presented in Chapter 3. Terrain height, in meters above mean sea level, is indicated by the grayscale shading. 40 km CASA radar range rings are included for reference. The black box in the northeast portion of the domain denotes the extent of the forecast verification sub-domain used in calculation of forecast skill scores.

In all ensemble members, a reduced rain intercept parameter of 8×10^5 was used, consistent with Snook and Xue (2008), who found that reducing the rain intercept parameter yielded more realistic cold-pool structure. A timeline for these experiments is presented in Fig. 4.3. Each ensemble forecast experiment uses, as its initial condition, the final ensemble analyses at 0200 UTC of the corresponding EnKF data assimilation experiment (i.e., CNTL, NoMMP or NoCASA).

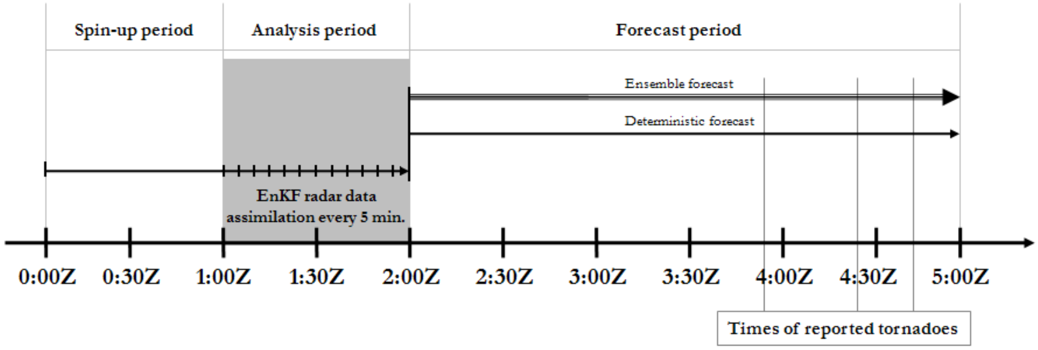

Fig. 4.3: Forecast diagram detailing the analysis and forecast periods for experiments CNTL, NoMMP, and NoCASA.

Forecast verification is performed for radar reflectivity at 0300, 0400, and 0500 UTC, and for low-level meso-vortices at 0400, 0420, and 0440 UTC; the latter times correspond closely to tornado reports received during this event at 0354, 0426, and 0443 UTC (see Fig. 3.1). Given the 2-km horizontal grid spacing used, tornado-scale circulations cannot be resolved, we therefore focus on prediction of resolvable low-level circulations linked to the observed tornadoes, rather than on the tornadoes themselves. With the 2-km horizontal grid spacing used, the mesovortices that were present in this case can be resolved.

4.2 Deterministic forecast experiments

While this study focuses primarily on ensemble forecasting, it is often valuable to perform a deterministic forecast from the ensemble mean of the initial conditions. Such deterministic forecasts are produced for each experiment from the ensemble mean analyses valid at 0200 UTC, using the Lin microphysical scheme in the forecast model. Given that the choice and configuration of the microphysical parameterization scheme is known to greatly affect the forecast solution (Dawson et al. 2009), particularly with regard to convective dynamics (Snook and Xue 2008), two additional deterministic forecasts are launched from the ensemble mean analysis of CNTL, one using the WSM6 microphysics scheme and one using the NEM microphysics schemes; these forecasts are called CNTL-W and CNTL-N, respectively.

4.2.1 Deterministic predictions of radar reflectivity

Deterministic forecasts of radar reflectivity and horizontal winds at approximately 2 km above ground level are displayed in for forecasts valid at 0300, 0400, and 0500 UTC from CNTL (Fig. 4.4a, e, i), NoMMP (Fig. 4.4b, f, j), and NoCASA (Fig. 4.4c, g, k). Also plotted are mosaics of observed radar reflectivity from the WSR-88D network at the corresponding times (rightmost column of Fig. 4.4). All experiments predict a MCS with a line-end vortex (LEV) that moves northeast through the CASA domain by around 0400 UTC (Fig. 4.4). In all experiments, the convective line moves more slowly than in observations, with the greatest position error occurring after 0400 UTC. The general structure and strength of the northern portion of the convective line remains reasonably well-predicted even by 0500 UTC, though the southern portion of the line dissipates more quickly in the forecasts than in the observations (Fig. 4.4i-l).

Though the convective mode is predicted rather well, there is noticeable error in all experiments with regard to the convective lines extending to the south and southeast. WSR-88D observed two convective lines associated with the system (see

Fig. 4.4d, h, l): a stronger leading convective line containing the main bow-echo, initially oriented from northwest to southeast (Fig. 4.4d) then later oriented more north-south (Fig. 4.4h), and a second weaker line of convection and stratiform precipitation oriented from north-northeast to south-southwest, approximately 50-100 km behind the leading convective line. While model reflectivity in the initial conditions of all three experiments generally matches well with WSR-88D composite reflectivity observations at 0200 UTC (Fig. 3.4), the subsequent forecasts (Fig. 4.4a-c, e-g, i-k) have only a single convective line, positioned roughly along the axis of the trailing line in the WSR-88D observations. This line is present in all experiments, and is most intense in NoMMP, particularly later in the forecast (Fig. 4.4i, j, k). The convective line moves eastward more slowly in the forecasts than in the observations, possibly due in part to an underestimation of the intensity of the LEV during the later portion of the forecast period, as will be discussed later. Error in the mesoscale environment surrounding the convective system probably also plays a role; because no data is assimilated outside the radar echo region, environmental errors present in the NAM analysis background cannot be reduced by data assimilation. This issue is addressed in the results that will be presented in Chapter 5 via assimilation of conventional observations.

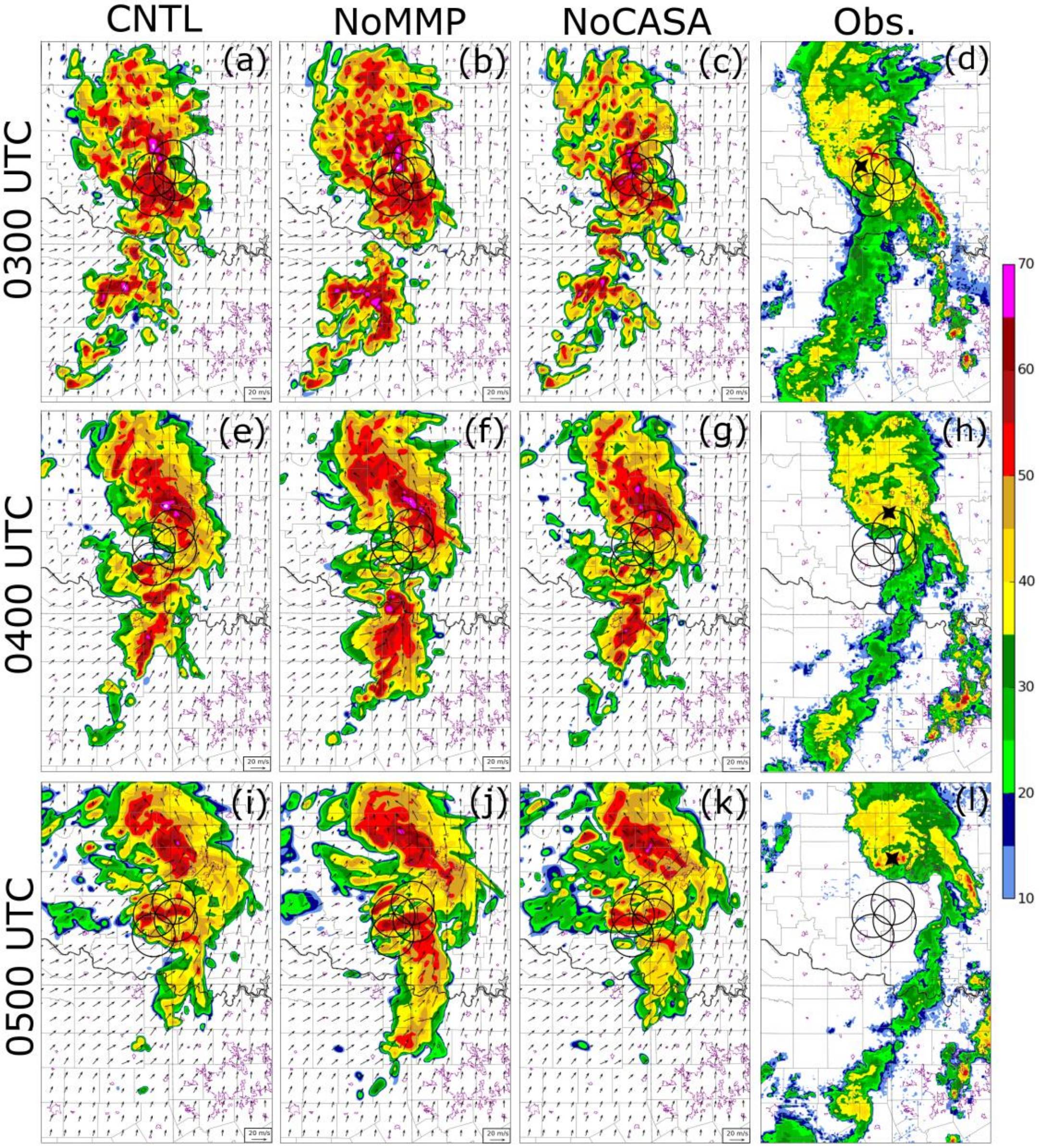

Fig. 4.4: Forecasts of reflectivity (shaded) and winds (vectors) at model grid level 10 (approximately 2 km above ground level) for CNTL at (a) 03:00 UTC, (e) 04:00 UTC, and (i) 05:00 UTC; NoMMP at (b) 03:00 UTC, (f) 04:00 UTC, and (j) 05:00 UTC; and NoCASA at (c) 03:00 UTC, (g) 04:00 UTC, and (k) 05:00 UTC. Also shown is reflectivity as observed by the WSR-88D network, interpolated to the same model grid level at (d) 03:00 UTC, (h) 04:00 UTC, and (l) 05:00 UTC.

To more quantitatively evaluate the deterministic reflectivity forecasts, the Equitable Threat Score (ETS), also known as the Gilbert Skill Score (GSS) (Schaefer 1990), is calculated for radar reflectivity during the forecast period using thresholds of 25 (Fig. 4.5a) and 40 dBZ (Fig. 4.5b). ETS is calculated on model vertical grid level 10 (slightly more than 2 km above the surface) within a verification sub-domain (shown in Fig. 3.3) chosen to focus on the region where the impact of CASA radar data on the forecast is likely to be significant. The LEV and its accompanying tornadoes pass through this sub-domain during the forecast period. Unlike the critical success index, also known as the threat score (Wilks 2006), the ETS is less sensitive to event climatology and thus does not assign unduly poor scores for forecasts of rare events (Schaefer 1990), although issues do exist in the verification of high-resolution precipitation forecasts using ETS (Baldwin et al. 2001; Ebert 2008). Possible values of ETS range from -0.33 to 1.0, with 1.0 representing a perfect forecast and 0.0 indicating a forecast with no skill. ETS is commonly used for accumulated precipitation verification; we apply it here to instantaneous reflectivity fields with the understanding that instantaneous fields are much more sensitive to timing errors than accumulated fields.

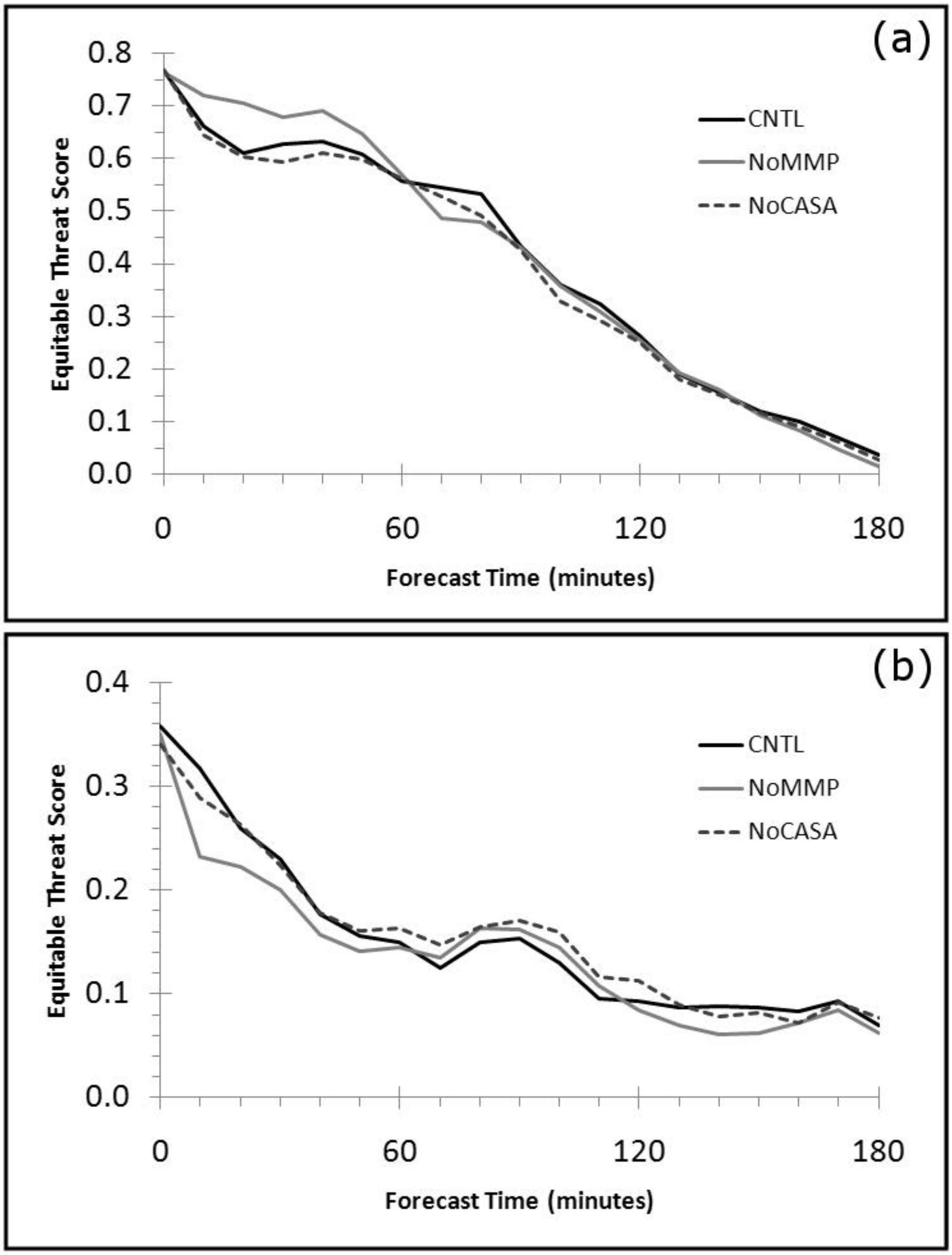

Fig. 4.5: Time-series of equitable threat scores (ETSs) calculated between a mosaic of WSR-88D reflectivity observations and deterministic forecasts from experiments CNTL, NoMMP, and NoCASA at vertical grid level k = 10 (approximately 2 km above ground level) for thresholds of (a) 25 dBZ and (b) 40 dBZ. Calculations were performed over the verification sub-domain indicated by the black outline in Fig. 4.2.

The ETS is considerably higher at the 25 dBZ threshold (Fig. 4.5a) than at the 40 dBZ threshold (Fig. 4.5b) at all times in all experiments. The 25 dBZ threshold includes both stratiform and convective precipitation regions, while the 40 dBZ threshold includes only heavier precipitation associated with convective cores; these small convective cores are much more sensitive to small errors in position and timing than the widespread stratiform precipitation regions also present in this case (see Fig. 4.4). In general, ETS tends to be lower for smaller fractional area coverage within a verification domain. ETS at the 25 dBZ threshold in the initial conditions (0200 UTC; forecast time zero in Fig. 4.5a) is approximately 0.76 in all three experiments—this relatively high value of ETS underscores the overall good quality of the EnKF ensemble mean analysis. The ETSs at 0200 UTC for the 40 dBZ threshold (Fig. 4.5b) are lower but still show considerable skill, with values ranging from 0.34 for NoCASA to 0.36 for CNTL.

During the forecast period, the 25 dBZ threshold ETS scores (Fig. 4.5a) decline modestly during the first 80 minutes the forecast, falling from an initial value of approximately 0.77 in all experiments to between 0.48 (for NoMMP) and 0.53 (for CNTL) at 80 minutes of forecast time. While NoMMP has a similar or slightly lower ETS than CNTL at the 25 dBZ threshold after 0300 UTC, during the first hour of the forecast, NoMMP actually has a slightly higher ETS than both CNTL and NoCASA. The initially lower ETS values of NoCASA and CNTL for the 25 dBZ threshold can be attributed to the influence of ensemble members using the NEM microphysics scheme during the EnKF data assimilation process that generated the ensemble mean analysis from which these forecasts were initialized; ensemble members using the NEM microphysics scheme greatly under-predicted the geographic extent of stratiform precipitation within the MCS, leading to a deficiency in low to moderate reflectivity areas in the CNTL and NoCASA ensemble mean analyses; this deficiency remains present during the first 30 minutes of the CNTL and NoCASA forecasts (not shown). This initial reduction of ETS in CNTL and NoCASA is only present at weak to moderate reflectivity thresholds; at a higher threshold of 40 dBZ (Fig. 4.5b) this trend is actually reversed, with CNTL and NoCASA consistently producing higher ETS values than NoMMP during the first 60 minutes of the forecast.

After 0320 UTC (80 minutes of forecast time), 25 dBZ ETS scores in all experiments remain comparable to one another and decline steadily for the remainder of the forecast period, falling to approximately 0.26 by 0400 UTC, and to below 0.1 by 0500 UTC. Development of spurious convection near and within the CASA domain after 0300 UTC (see Fig. 4.4) contributes significantly to this decline. In all experiments, 40 dBZ ETS scores (Fig. 4.5b) declined most rapidly during the first hour of the forecast period, dropping from their initial values of between 0.34 and 0.36 at 0200 UTC to between 0.14 and 0.17 at 0300 UTC, then declining only modestly afterward, falling only to slightly below 0.10 by 0500 UTC. In addition to the negative impact of spurious convection that developed near the CASA network, the relatively low ETS values later in the forecast period can also be attributed to modest position errors in the location of the heavy convective cores in the forecasts (see Fig. 4.4); as mentioned earlier, even in an operationally-useful forecast, small position errors often result in very low skill scores when verifying isolated features using ETS (Baldwin et al. 2001; Schwartz et al. 2009a).

4.2.2 Evolution of the MCS and LEV in the deterministic forecasts

All experiments underestimate the strength of the LEV, particularly during the last hour of the forecast period, contributing to forecast error in the convective lines in southern and central Oklahoma and northern Texas (see Fig. 4.4i-l). To more closely examine the evolution of the LEV circulation and its associated convection, time-height plots of (horizontal plane) maximum vertical vorticity within the verification sub-domain (the black outline in Fig. 1) are calculated for the deterministic forecasts of CNTL (Fig. 13a), NoMMP (Fig. 13b), and NoCASA (Fig. 13c). Similar time-height plots of maximum updraft velocity are presented in Fig. 14.

During the EnKF assimilation of radar reflectivity and radial velocity data, at 5 minute intervals between 0100 and 0200 UTC, maximum vertical vorticity values decrease throughout much of the troposphere during each forecast step, particularly at the lower levels (Fig. 3.6); the forecast model is unable to adequately sustain the intensity of circulation observed in the LEV. Despite model deficiency (due to resolution, etc.), the inclusion of CASA radar observations in CNTL reduces the underestimation of LEV intensity in the forecast step as compared to NoCASA. During the forecast period (Fig. 4.6), the maximum vertical vorticity values associated with the LEV remain relatively steady, and are similar to those seen during the forecast steps of the analysis period. By contrast, Schenkman et al.

(2011), who performed deterministic forecasts of this case based on 3DVAR analyses, noted that the observed LEV intensified noticeably during the forecast period, reaching its peak intensity between 0400 and 0530 UTC. The LEV was discernable in Oklahoma Mesonet surface observations at 0510 and 0520 UTC (not shown) as a well-defined surface circulation with around 25 m s^{-1} of horizontal shear (Schenkman et al. 2011). The LEV is present in the 0500 UTC deterministic forecasts (Fig. 4.4i-k), but it is weaker the LEV observed by the Oklahoma Mesonet, and located slightly farther north (Fig. 4.4i-l). The maximum vertical vorticity within the verification sub-domain is higher in CNTL and NoMMP than in NoCASA (Fig. 4.6), particularly between 0200 UTC to 0400 UTC in the lowest 6 km of the atmosphere, indicating that CASA data assimilated between 0100 and 0200 UTC contributed to greater sustained intensity of the LEV in the forecasts.

Several factors likely contributed to the underestimation of the intensity of the LEV in the model. Mesonet and surface observations are not included in our EnKF assimilation process; Schenkman et al. (2011) found these observations useful in improving their 3DVAR analyses of this event. In addition, the relatively coarse 2 km grid spacing used in this study limits the ability of the model to resolve smaller convective-scale processes.

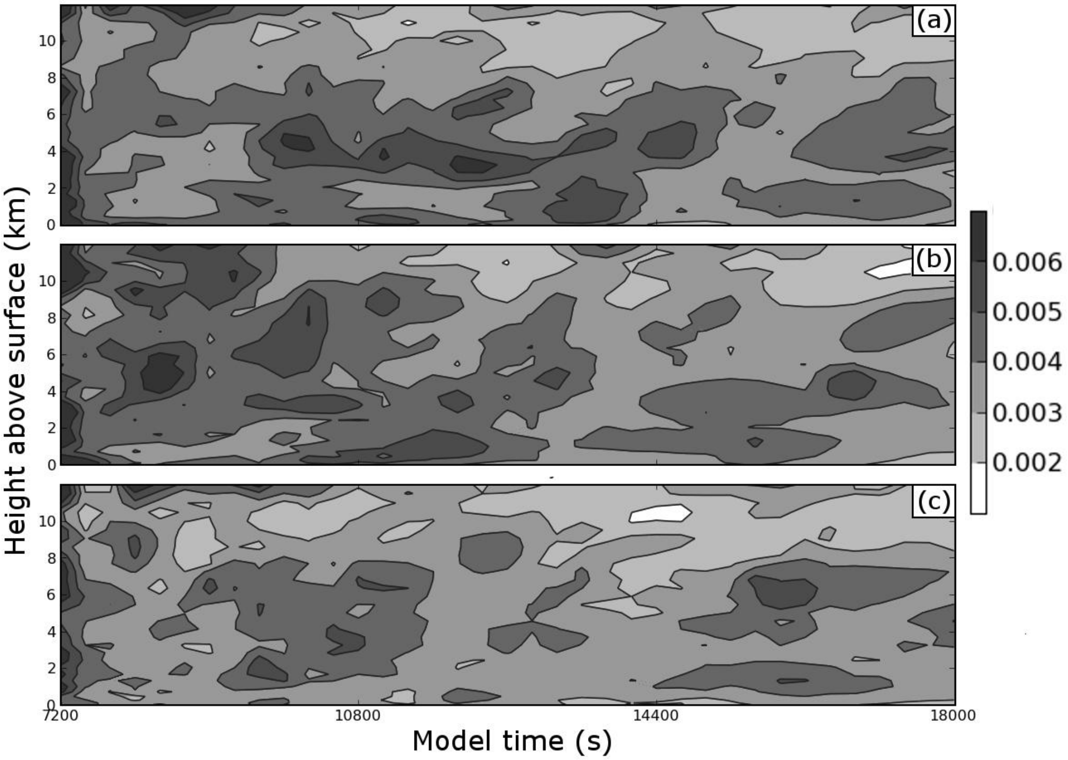

Fig. 4.6: Time-height plot of maximum vertical vorticity (s^{-1}) within the forecast verification sub-domain (depicted by the black outline in Fig. 4.2) during the forecast period for deterministic forecasts of (a) CNTL, (b) NoMMP, and (c) NoCASA. Time is denoted in seconds since forecast initialization and ranges from 7200 s (0200 UTC), which corresponds to the end of the analysis period, to 18000 s (0500 UTC). Height above the surface is shown in kilometers.

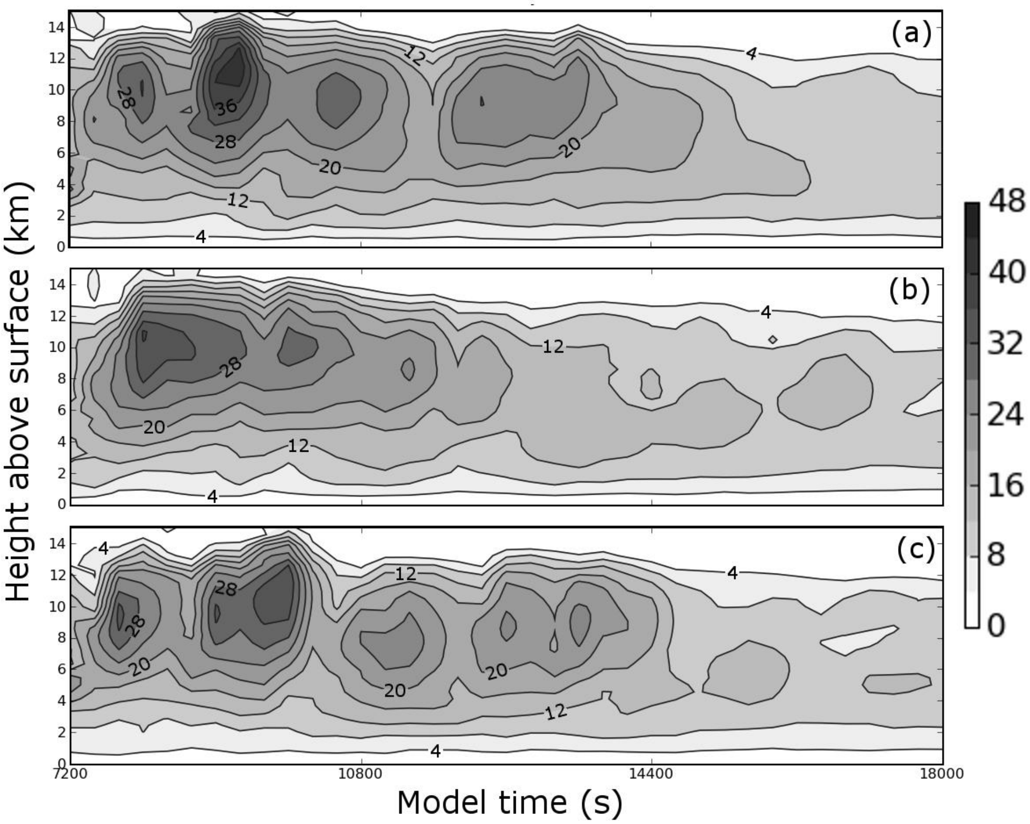

Fig. 4.7: As Fig. 4.6 but for maximum updraft velocity (m s^{-1}).

Despite the underestimation of LEV intensity, strong convective updrafts were maintained in all experiments from the beginning of the forecast at 0200 UTC until approximately 0420 UTC (Fig. 4.7); after this time the peak updraft velocity decreased considerably. Peak updraft speed in CNTL and NoCASA (Fig. 4.7a, c) occurred around 0240 UTC, while peak updraft speed occurred slightly earlier, around 0220 UTC, in NoMMP (Fig. 4.7b). Variation of peak updraft intensity with time is similar in CNTL and NoCASA (Fig. 4.7a, c). Both have similarly-timed peak updraft velocities and show five similar distinct maxima in updraft speed, occurring in CNTL around 0220, 0240, 0300, 0320 and 0345 UTC (Fig. 4.7a), and in NoCASA at around 0215, 0245, 0305, 0325, and 0345 UTC. Updraft speed maxima in

NoMMP (Fig. 4.7b) do not correspond well to those in CNTL and NoCASA, particularly late in the forecast period; in NoMMP there are only three discrete peaks in updraft intensity, occurring at 0220, 0245, and 0310 UTC. The second and third peaks are much weaker than those in CNTL and NoCASA. The maximum updraft speed is also quicker to decrease in NoMMP, as evidenced by the lower peak updraft intensity in NoMMP between 0330 and 0420 UTC (Fig. 4.7). The similarity of CNTL and NoCASA and the relative difference of NoMMP suggest that during the forecast period, evolution of convective updrafts is relatively sensitive to the microphysical makeup of the ensemble during the EnKF data assimilation process, but only weakly sensitive to inclusion of CASA data during the assimilation period. Previous studies by Dawson (2009) and Snook and Xue (2008) found that the choice of microphysical scheme and parameter settings greatly impact forecasts of convective storms. These results suggest reducing error in the forecasting model is at least as important as providing low level data coverage in addition to existing WSR-88D radar network observations, at least for the current case using the given data assimilation and forecasting systems.

4.2.3 Effects of microphysical parameterization in the deterministic forecasts

To further examine the effects of microphysical parameterization on the deterministic forecasts, we performed two additional deterministic forecasts from the 0200 UTC ensemble mean analysis of CNTL using the WSM6 and NEM microphysical schemes. To distinguish these experiments from the deterministic forecasts discussed earlier, we label the forecasts using the WSM6 and NEM schemes CNTL-W and CNTL-N (respectively). Forecast radar reflectivity fields and horizontal wind vectors at vertical model level 10 are plotted at 0500 UTC and compared with the deterministic forecast of CNTL and with WSR-88D observed reflectivity in Fig. 4.8.

While all three deterministic forecasts feature similar positioning of the northern portion of the MCS (Fig. 4.8a-c), the horizontal extent of lighter precipitation (between 10 and 30 dBZ) is considerably greater in CNTL and CNTL-W (Fig. 4.8a, b) than in CNTL-N Fig. 4.8c). All three forecasts underestimate the coverage of the lighter precipitation to some extent (see Fig. 4.8d); this underestimation is least in CNTL (Fig. 4.8a), and greatest in CNTL-N (Fig. 4.8c). The 2 km horizontal wind fields are similar over much of the domain in all three experiments, with southwesterly flow dominating behind the MCS, though notable differences exist in the region of the LEV; the LEV is strongest and best defined at this level in CNTL (Fig. 4.8a), and weakest and most poorly defined in CNTL-N. All experiments show a LEV weaker than that suggested by the Oklahoma mesonet observations cited in Schenkman et al. (2011). Though all three microphysics schemes used are single-moment ice microphysics schemes, the Lin and WSM6 schemes attempt to accurately model various complex conversions between ice and liquid hydrometeor classes, while the NEM microphysical scheme includes simplified conversion processes for better computational efficiency (Schultz 1995). Despite reduced computational cost of the NEM scheme, it does not appear to be well-suited for the case investigated here.

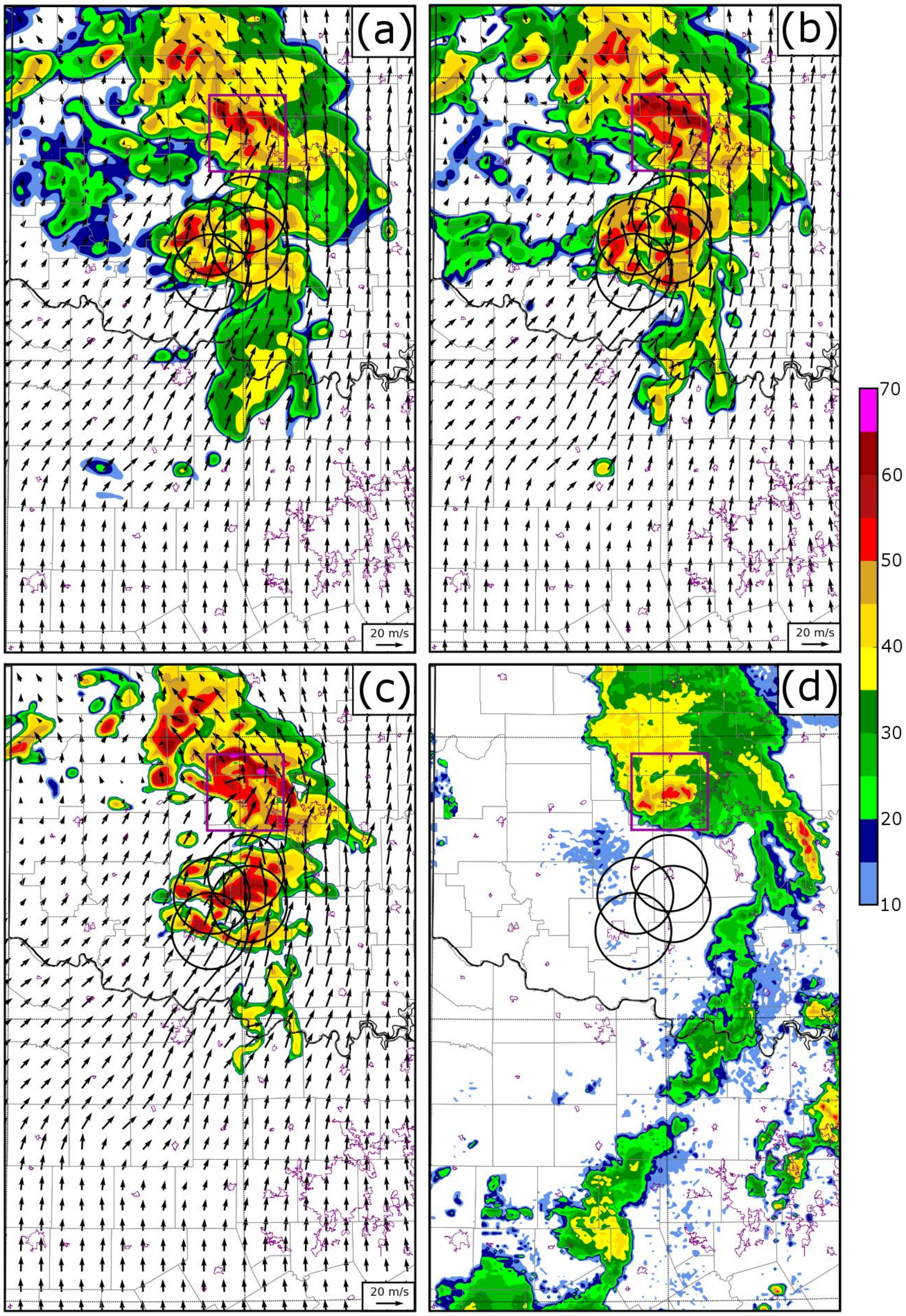

Fig. 4.8: Deterministic forecasts of reflectivity (shaded) and winds (vectors) at model grid level 10 (approximately 2 km above ground level) at 0500 UTC for experiments (a) CNTL-L, (b) CNTL-W, and (c) CNTL-N. Also shown: (d) reflectivity as observed by the WSR-88D network at 0500 UTC interpolated to the same model grid level. The purple box in each panel indicates the approximate location of the observed line-end vortex at 0500 UTC.

4.3 Ensemble and probabilistic forecasts

We begin by presenting the results of ensemble and probabilistic forecasts of radar reflectivity (Z) for CNTL, CNTL_LIN and NoMMP. Radar reflectivity, closely linked to precipitation, is a field of meteorological interest that can be directly verified against WSR-88D radar observations over the entire area of the convective system. When verifying radar reflectivity forecasts, emphasis is placed on the impact of the microphysical makeup of the ensemble during the analysis and forecast period, since the impact of CASA data is limited at later forecast times and at greater distances from the CASA domain. Ensemble and probabilistic forecasts of low-level vortices are then analyzed; near-surface meso-vortices were closely co-located with the observed tornadoes in this event, as demonstrated by the proximity of rotational signatures in the KTLX storm-relative radial velocity observations (Fig. 4.1) to the tornadoes reported in association with them (Fig. 3.1). Both the impact of assimilating CASA data and the use of mixed-microphysics during the analysis and forecast periods are considered in evaluating the performance of the ensemble meso-vortex forecasts.

4.3.1 Probability-matched ensemble mean radar reflectivity

In ensemble forecasting, particularly at the convective scale, averaging the individual ensemble members to produce ensemble mean fields for precipitation-related variables (such as radar reflectivity) often leads to a smoother distribution with increased geographic extent and a low bias in intensity. To counteract the low-bias tendency in the ensemble mean, it is often desirable to reassign the values of precipitation-related fields using values from the component ensemble members used to compute the mean; the result of this process (Ebert 2001) is the probability-matched (PM) mean. The PM mean is a useful tool for forecasters, producing a single "best estimate" deterministic forecast from an ensemble of forecasts; such ensemble-mean products for precipitation fields often outperform most or all of the ensemble members used to produce them (Ebert 2001). PM mean *Z* is calculated at approximately 2 km above the surface at 1 (Fig. 4.9a-c), 2 (Fig. 4.9e-g), and 3 (Fig. 4.9i-k) hours of forecast time for CNTL, CNTL_LIN, and NoMMP, and compared with *Z* as observed by the WSR-88D radar network and interpolated to the model grid at the corresponding times (Fig. 4.9d, h, l).

In all experiments, the dominant convective mode is predicted with reasonable success (Fig. 4.9); the forecast ensembles predict a large mesoscale convective system with a trailing convective line in a similar location to the corresponding features in the WSR-88D observations (Fig. 4.9d, h, l). There are noticeable errors in the PM mean forecasts, however; in particular, the leading convective line, extending southeast from the region of the LEV in the observations (Fig. 4.9d, h, l) is not captured in the forecast ensemble. In addition, the southern portion of the trailing convective line decays too quickly in the ensemble forecast at 0400 UTC and beyond in all forecast experiments (Fig. 4.9b-c, f-g, j-k). Because of the prevailing south and south-southwesterly flow during this case, the decay of the southern portion of the trailing convective line can likely be attributed to domain-boundary interaction in proximity to the southern boundary of the model domain.

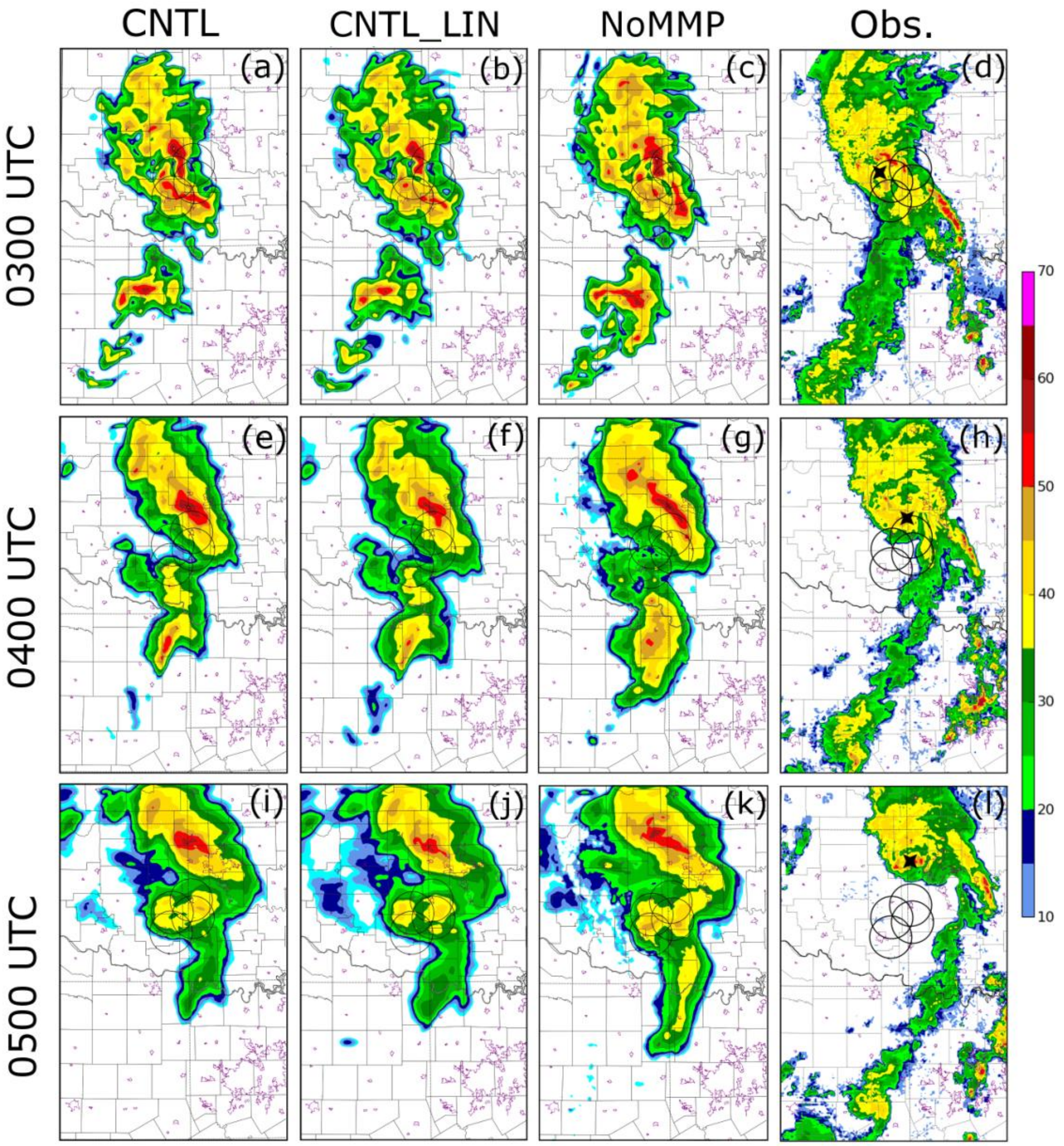

Fig. 4.9: Probability-matched ensemble mean forecasts of reflectivity (shaded) at model grid level 10 (approximately 2 km above ground level) at (a-c) 0300 UTC, (e-g) 0400 UTC, and (i-k) 0500 UTC for CNTL, CNTL_LIN, and NoMMP. Also shown is reflectivity observed by the WSR-88D network, interpolated to the same model grid level at (d) 0300 UTC, (h) 0400 UTC, and (l) 0500 UTC. The center of the line-end vortex in the observations is indicated by the black marker in panels (d), (h), and (l).

The Lin microphysics scheme, used in all 40 members of the NoMMP experiment during both the analysis and forecast periods, produces greater precipitation coverage than NEM and WSM schemes for this case (not shown). In addition, the reduced spread among NoMMP ensemble members as compared to the CNTL ensemble, results in much less variation in position (and thus higher ensemble mean values) in the trailing line, resulting in a more intense trailing line in the PM mean (Fig. 4.9i-k). In contrast, CNTL, used a mixed-microphysics ensemble containing members from the WSM-6 and NEM microphysical schemes, both of which favored solutions with less light precipitation. Combined with exhibited greater variation among ensemble members, the result is a PM mean field with a weaker trailing line and less extensive regions of lighter precipitation in the northern portion of the system in CNTL (Fig. 4.9a-c).

4.3.2 Probabilistic reflectivity forecasts

Fast error growth and high levels of uncertainty from various sources make probabilistic forecast products potentially very valuable at the convective scale. Using a NEP method, as described in chapter 2, P[Z > 40 dBZ] (Fig. 4.10) and P[Z > 25 dBZ] (Fig. 4.11) are calculated at 0300, 0400, and 0500 UTC for CNTL, CNTL_LIN, and NoMMP, and are compared with the corresponding reflectivity contour observed by the WSR-88D network. A neighborhood radius of 5 km was used, resulting in a neighborhood consisting of 21 points in each ensemble member; thus, for the 40 member ensemble, the ensemble-wide neighborhood consists of 840 forecast values.

The observed 25 dBZ threshold (the thick contour in Fig. 4.10) encompasses large areas, including the entire region surrounding the LEV circulation and much of the convective line extending to the south and southwest; areas with a high NEP of reflectivity exceeding 25 dBZ closely match regions where observed reflectivity exceeds 25 dBZ over the northern portion of the system throughout the forecast period (Fig. 4.10), though the forecast precipitation region does not extend quite as far north and east as in observations at 0300 and 0400 UTC. The presence of the trailing convective line to the south and southwest is strongly indicated in the probabilistic forecasts, particularly in the NoMMP ensemble (Fig. 4.10d-f), though the decay of the southern portion of the line discussed in section 3a is evident in all experiments. The greater coverage of light precipitation in members using the Lin scheme and less variation among ensemble members result in a large region of very high probability associated with the trailing line in NoMMP, particularly at 0400 and 0500 UTC (Fig. 4.10h, i).

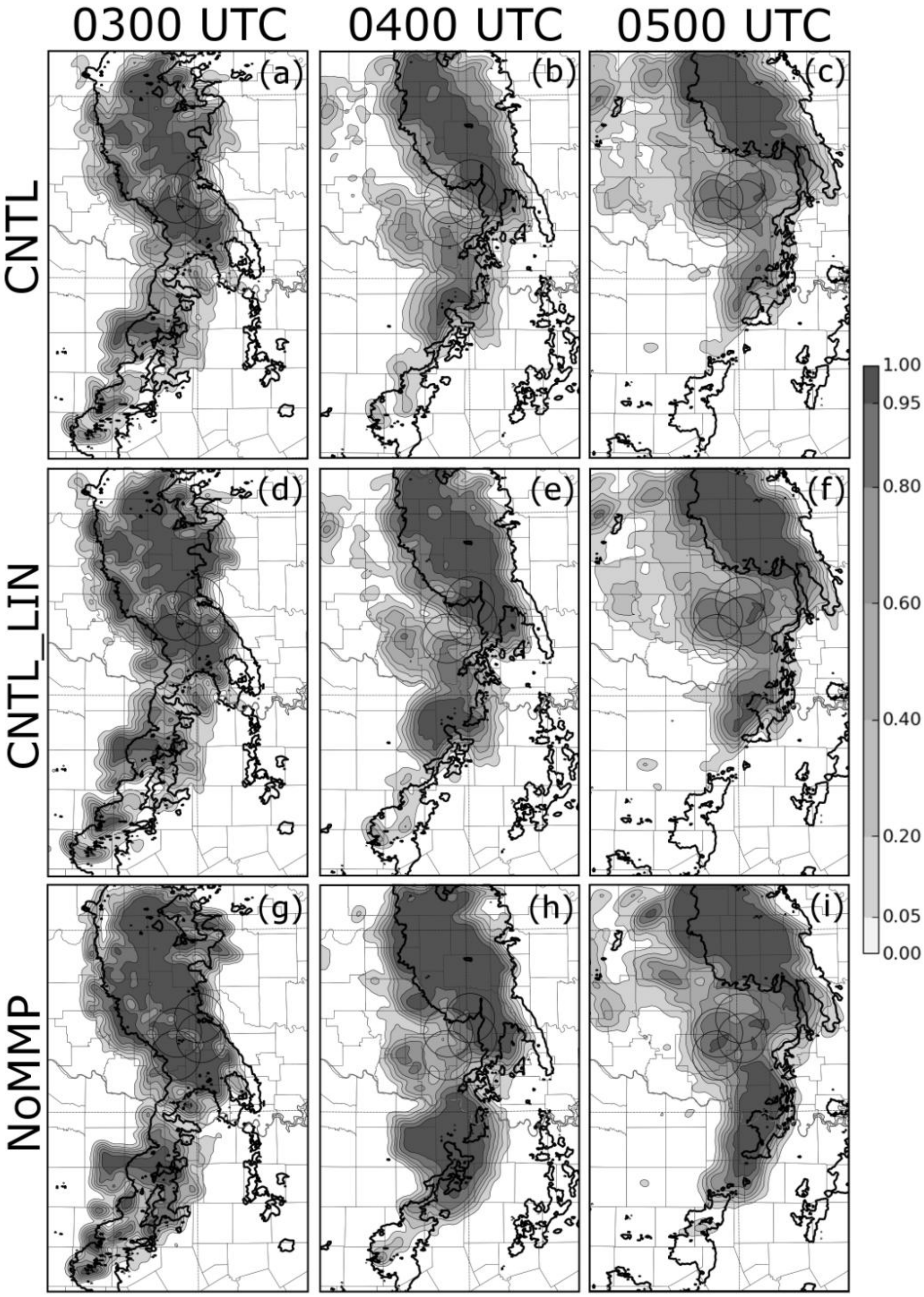

Fig. 4.10: Neighborhood ensemble probabilities (shaded) of reflectivity exceeding 25 dBZ, P[Z > 25 dBZ], at model grid level 10 (approximately 2 km above the surface), in an area surrounding the CASA domain, for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC, CNTL_LIN at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC, and NoMMP at (g) 0300 UTC, (h) 0400 UTC, and (i) 0500 UTC. The 25 dBZ radar reflectivity contours observed by the WSR-88D radars at the same time are in bold black contours.

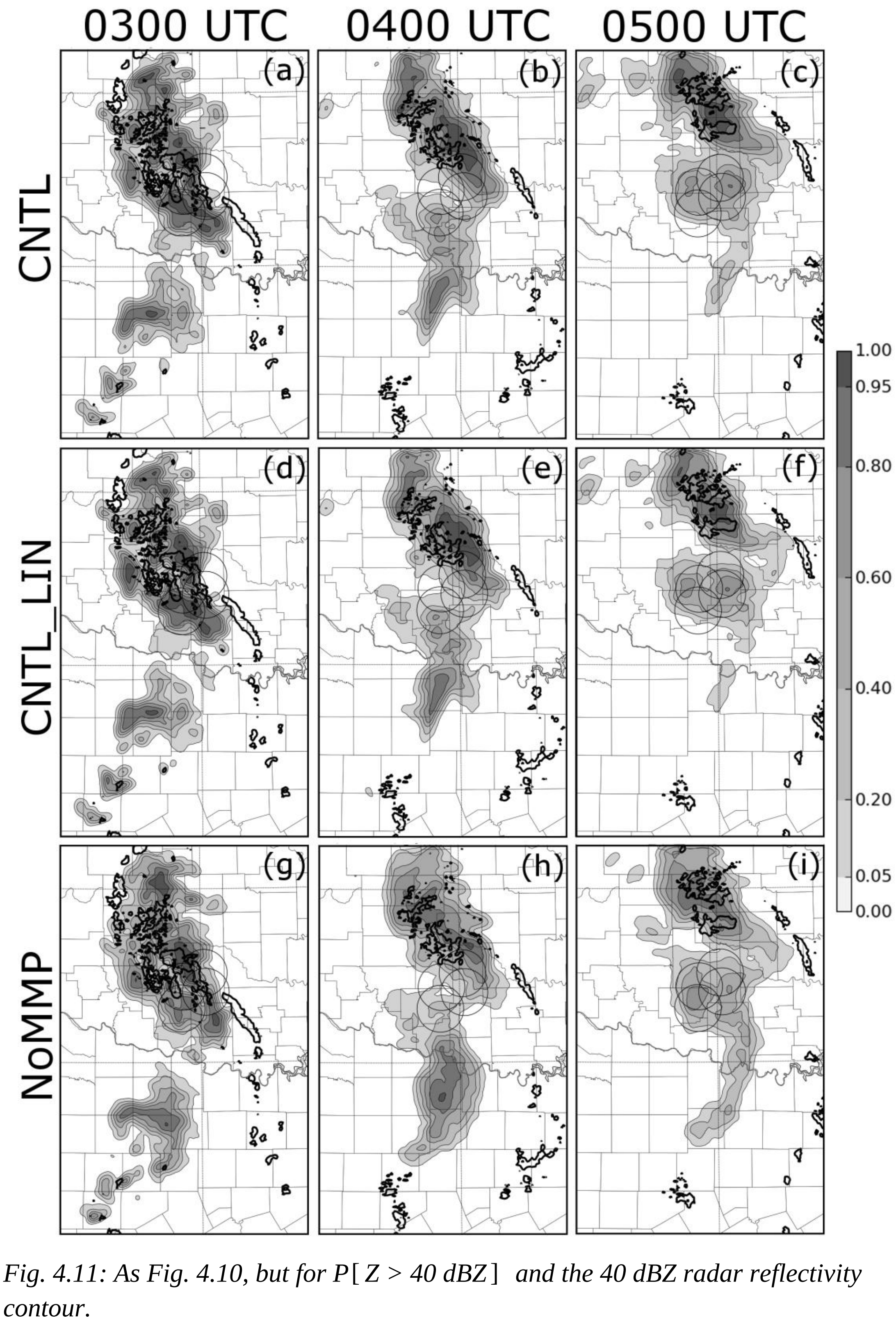

Fig. 4.11: As Fig. 4.10, but for $P[Z > 40\ \mathrm{dBZ}]$ and the 40 dBZ radar reflectivity contour.

The 40 dBZ threshold is exceeded only in small, localized areas in the WSR-88D observations (the single thick contour in Fig. 4.11), in individual convective elements within the core of the LEV and the leading convective line. Regions near the LEV (c.f. Fig. 4.9d, h, l) where observed radar reflectivity exceeds 40 dBZ are found in regions of moderate to high NEP (Fig. 4.11), although observed regions exceeding 40 dBZ in the leading convective line are largely missed by the forecast ensembles. Particularly at later forecast times, all experiments strongly overestimate the geographic coverage of 40 dBZ echoes, due in part to spurious convection that develops near the CASA radar network (Fig. 4.11c, f, i). The causes of this spurious convection are considered below in section 4.3.3. The intensity of the trailing convective line extending southeast and south of the CASA domain into north-central Texas (Fig. 4.11) is also overestimated; moderate to high NEP values are present in areas where no reflectivity exceeding 40 dBZ was observed. The ensemble adequately forecasts many of the large-scale reflectivity features of the system (Fig. 4.10), but a combination of position error, intensity error, and development of spurious convection limits forecast skill for smaller-scale convective structures (Fig. 4.11). This result is consistent with the fact that smaller-scale features have shorter ranges of predictability (Lorenz 1969).

To more quantitatively evaluate the skill of the NEP forecasts of radar reflectivity, we next examine their performance using two skill-score metrics: the relative operating characteristic (ROC) skill score (Mason and Graham 1999), and the Brier score (Brier 1950). The ROC skill score (RSS) is derived from the area under the ROC curve (Mason and Graham 1999), and is given by $\mathrm{RSS} = 2A_{ROC} - 1$, where A_{ROC} is the area under the ROC curve. The RSS has a maximum value of 1.0 for a perfect probabilistic forecast, and a minimum of -1.0, with scores at or below 0.0 indicating forecasts with no skill. The ROC measures hit and false alarm rates at varying probability thresholds (Mason and Graham 1999); thus RSS is a summary statistic providing information on the resolution of the probabilistic forecast system. RSS is calculated for forecasts of reflectivity exceeding 25 dBZ at grid level 10, over the entire model domain (Fig. 4.12a), as well as over the verification sub-domain (Fig. 4.12b) located within and downwind of the CASA radar network (see Fig. 4.2). In all experiments, the RSS is considerably higher over the verification sub-domain (Fig. 4.12b) than over the entire domain (Fig. 4.12a), because of error associated with the trailing convective line south and southwest of the CASA network.

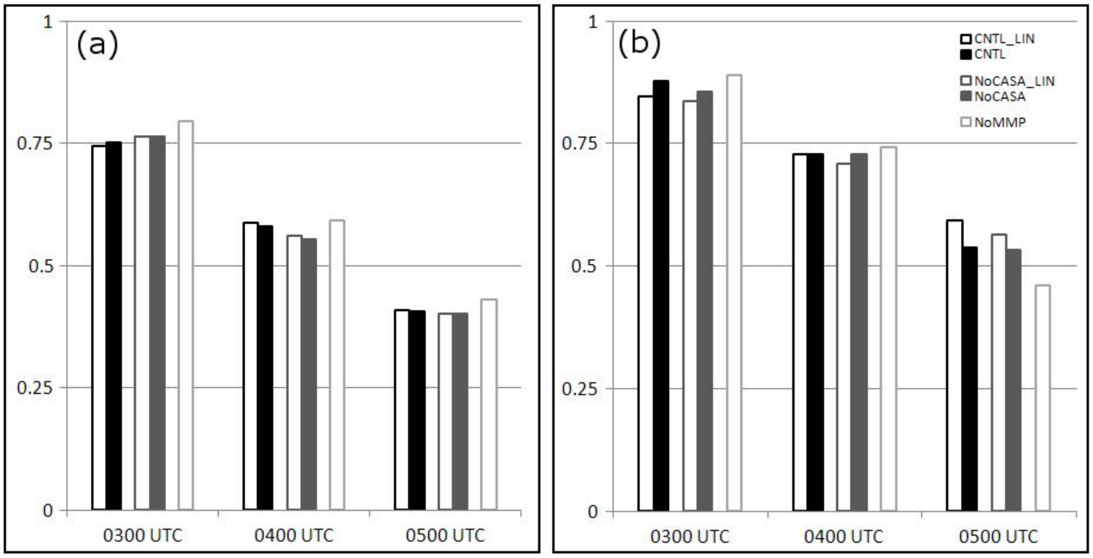

Fig. 4.12: ROC skill score for 1-, 2-, and 3-hour forecasts of radar reflectivity at the 25 dBZ threshold on vertical grid level k = 10 (slightly more than 2 km above mean sea level) calculated over (a) the entire horizontal model domain and (b) the ETS verification sub-domain as depicted by the black outline in Fig. 4.2.

All forecast ensembles show considerable skill in forecasting reflectivity exceeding 25 dBZ. The 1-hour ensemble forecasts (valid at 0300 UTC) show the greatest skill, with verification sub-domain RSSs ranging from 0.84 to 0.89. RSSs steadily decrease between 0300 and 0500 UTC (Fig. 4.12), with domain-wide values (Fig. 4.12a) remaining slightly lower than those over the verification sub-domain (Fig. 4.12b). The minimum RSS of 0.40, obtained for NoCASA_LIN on the full domain at 0500 UTC, corresponds to an area under the ROC curve of 0.70; a ROC area of 0.70 is often considered to be the lower bound for a skillful forecast (Buizza 1997; Kong et al. 2011), thus all experiments produced skillful 1, 2, and 3 hour NEP forecasts for regions of *Z* exceeding 25 dBZ.

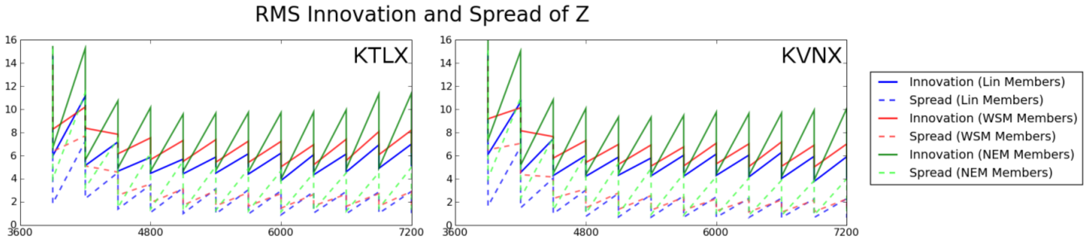

Fig. 4.13: RMS innovation and ensemble spread of radar reflectivity for three sub-ensembles, each containing members using the same microphysical scheme, within the CNTL experiment. The sub-ensembles consist of (1) the 16 Lin members, (2) the 16 WSM-6 members, and (3) the 8 NEM members within CNTL, respectively.

NoMMP has the highest domain-wide RSS (Fig. 4.12a) throughout the forecast period and the highest RSS within the verification sub-domain (Fig. 4.12b) at 0300 and 0400 UTC; at 0500 UTC, however, NoMMP actually has the lowest RSS in the sub-domain, while CNTL_LIN has the highest. For this case, ensemble members using Lin microphysics showed, on average, slightly better agreement with observed *Z* than members using WSM6 or NEM microphysics. When ROC skill scores are calculated for three sub-ensembles within CNTL, grouped by the microphysics scheme used, the Lin sub-ensemble produces the highest RSS, while the NEM sub-ensemble produces the lowest (Fig. 4.13). Similarly, within the verification sub-domain, RSSs decrease more slowly with time for CNTL_LIN and NoCASA_LIN than for their mixed-microphysics counterparts CNTL and NoCASA (Fig. 4.12b). The slightly improved RSS performance in CNTL_LIN and NoCASA_LIN at later times appears to be in part due to slightly faster movement of the trailing convective line, in better agreement with observations. These results suggest a positive impact of using a mixed-microphysics ensemble during the assimilation period, but slightly improved forecast performance when using a Lin-only ensemble during the subsequent forecast, at least for this case.

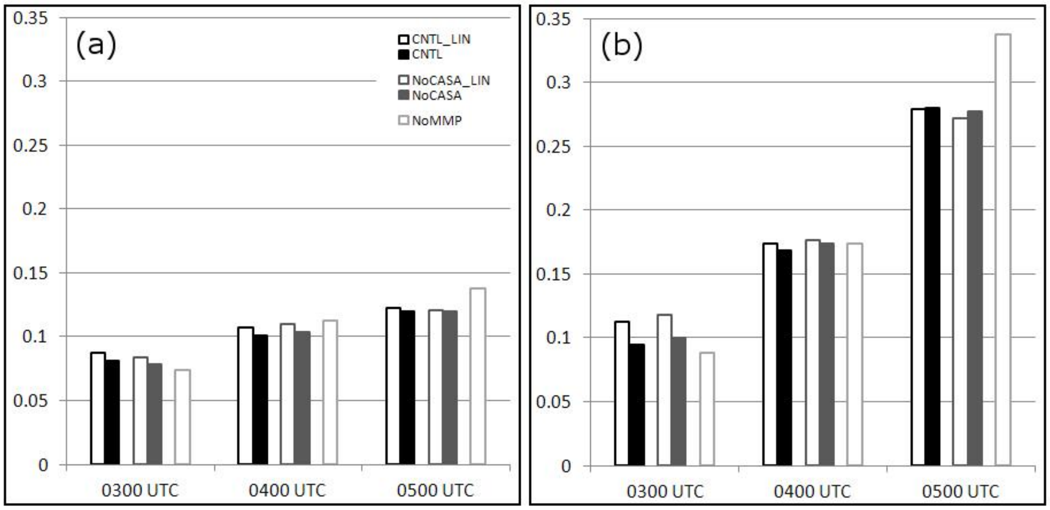

Fig. 4.14: As Fig. 4.12 but for Brier Score.

The Brier score (Brier 1950), another metric commonly used in probabilistic forecast evaluation, measures both forecast bias and uncertainty by verifying the forecast probability at each point against the observed binary realization of the event at that point (1 if the event occurred, 0 if it did not). Possible values of Brier score range from 0 to 1, with lower values indicating better agreement of the forecast probability with the observed binary outcome. As with RSSs, Brier scores are presented for forecasts of radar reflectivity exceeding 25 dBZ over the full model domain (Fig. 4.14a) and over the verification sub-domain (Fig. 4.14b). Because correctly forecasting non-events improves the Brier score, full domain Brier scores (Fig. 4.14a), which include large clear-air regions over western Oklahoma and the panhandle of Texas, are significantly lower (better) than Brier scores calculated over the verification sub-domain (Fig. 4.14b).

Domain-wide Brier scores range from between 0.07 and 0.09 at 0300 UTC, increasing modestly to between 0.11 and 0.14 by 0500 UTC (Fig. 4.14a). Brier scores over the verification sub-domain increase more quickly and are somewhat higher (worse) at later times. Factors negatively impacting the Brier score included development of spurious convection near and southwest of the CASA radar network, particularly at later times and particularly in NoMMP, as well as underestimation of the extent of the stratiform precipitation region in the ensemble (see Fig. 4.9), particularly in CNTL and NoCASA. Brier score deteriorates most quickly in NoMMP, which has the lowest (best) Brier score at 0300 UTC, but the highest (worst) at 0500 UTC both on the full domain and within the verification sub-domain (Fig. 4.14). Compared to their mixed-microphysics counterparts CNTL and

NoCASA, Brier scores for CNTL_LIN and NoCASA_LIN are slightly worse at 0300 and 0400 UTC, but nearly identical at 0500 UTC (Fig. 4.14). The worse performance of NoMMP at later forecast times suggests that increased spread in hydrometeor fields during data assimilation in the mixed-microphysics ensembles of CNTL and NoCASA may be beneficial in improving probabilistic forecasts of fields closely related to hydrometeor distribution, such as radar reflectivity. The faster deterioration of the Brier scores of the mixed-microphysics forecast ensembles (CNTL and NoCASA) as compared to those of their single-microphysics counterparts (CNTL_LIN and NoCASA_LIN), suggest that using a mixed-microphysics ensemble that contains poorer-performing scheme(s) during the forecast period can negatively impact the forecast.

To examine the spread among ensemble members in the reflectivity forecasts, spaghetti plots of the 25 dBZ (Fig. 4.15) and 40 dBZ (Fig. 4.16) contours at vertical model level 10 (approximately 2 km above the surface) are presented at 0300, 0400, and 0500 UTC for the forecast ensembles of CNTL, CNTL_LIN, and NoMMP, and compared to the corresponding observed WSR-88D contour. Though the analyses of the RSS and Brier score demonstrate that the probabilistic reflectivity forecasts possess considerable skill, variations among the ensemble members indicate significant uncertainty and spread within the ensemble near the LEV at the 40 dBZ threshold (Fig. 4.16). Ensemble spread, particularly at the 25 dBZ threshold and at earlier forecast times, is relatively low (Fig. 4.15a, d, g). The position and extent of reflectivity exceeding 25 dBZ in the northern portion of the convective system is well-predicted by most ensemble members in all experiments; most members also predict such regions in the trailing convective line with moderate success at 0300 and 0400 UTC. Spread and error in the ensemble forecast of the 25 dBZ threshold increase throughout the forecast period (Fig. 4.15); greater spread among ensemble members is expected for longer forecasts as initial differences among ensemble members grow with time.

The impact of model microphysics on ensemble spread is evident in the spaghetti plots of the 25 dBZ (Fig. 4.15) and 40 dBZ (Fig. 4.16) contours. Particularly at 25 dBZ, the spread among ensemble members in the single-microphysics forecast ensemble of NoMMP (Fig. 4.15g-i) is considerably less than in the mixed-microphysics ensemble of CNTL (Fig. 4.15a-c), and also somewhat less than in CNTL_LIN (Fig. 4.15d-f) which uses the same microphysical makeup at NoMMP but begins from an initial condition obtained using a mixed-microphysics analysis ensemble. In NoMMP, even after 3 hours of forecast time, close agreement among the ensemble members is evident in the 25 dBZ contours (Fig. 4.15i). Both CNTL (Fig. 4.15c) and CNTL_LIN (Fig. 4.15f) show greater variation among ensemble members; the difference is most prominent in the trailing convective line. Almost all members of NoMMP agree on a wide area of precipitation in excess of 25 dBZ in the trailing convective line at 0400 and 0500 UTC (Fig. 4.15h, i). In CNTL (Fig. 4.15b, c) many members predict the precipitation in the trailing line, but there is much greater variation in the location and extent of that precipitation than in NoMMP. Variation among ensemble members in CNTL_LIN (Fig. 4.15e, f) is greater than NoMMP, but far less than in CNTL. While starting from an initial condition that used mixed-microphysics appears to induce greater variation among members in the single-microphysics CNTL_LIN forecast ensemble, far greater variation results from using mixed-microphysics during both the analysis and forecast as in CNTL (Fig. 4.15a-c, Fig. 4.16a-c). In general, use of a mixed-microphysics ensemble results in increased ensemble spread in a variety of model fields, particularly those closely related to precipitation processes, due to variation in hydrometeor type and distribution resulting from different treatment of microphysical processes. Increased spread alone does not necessarily improve the derived probabilistic forecast products, however, as indicated in the earlier skill-score analyses.

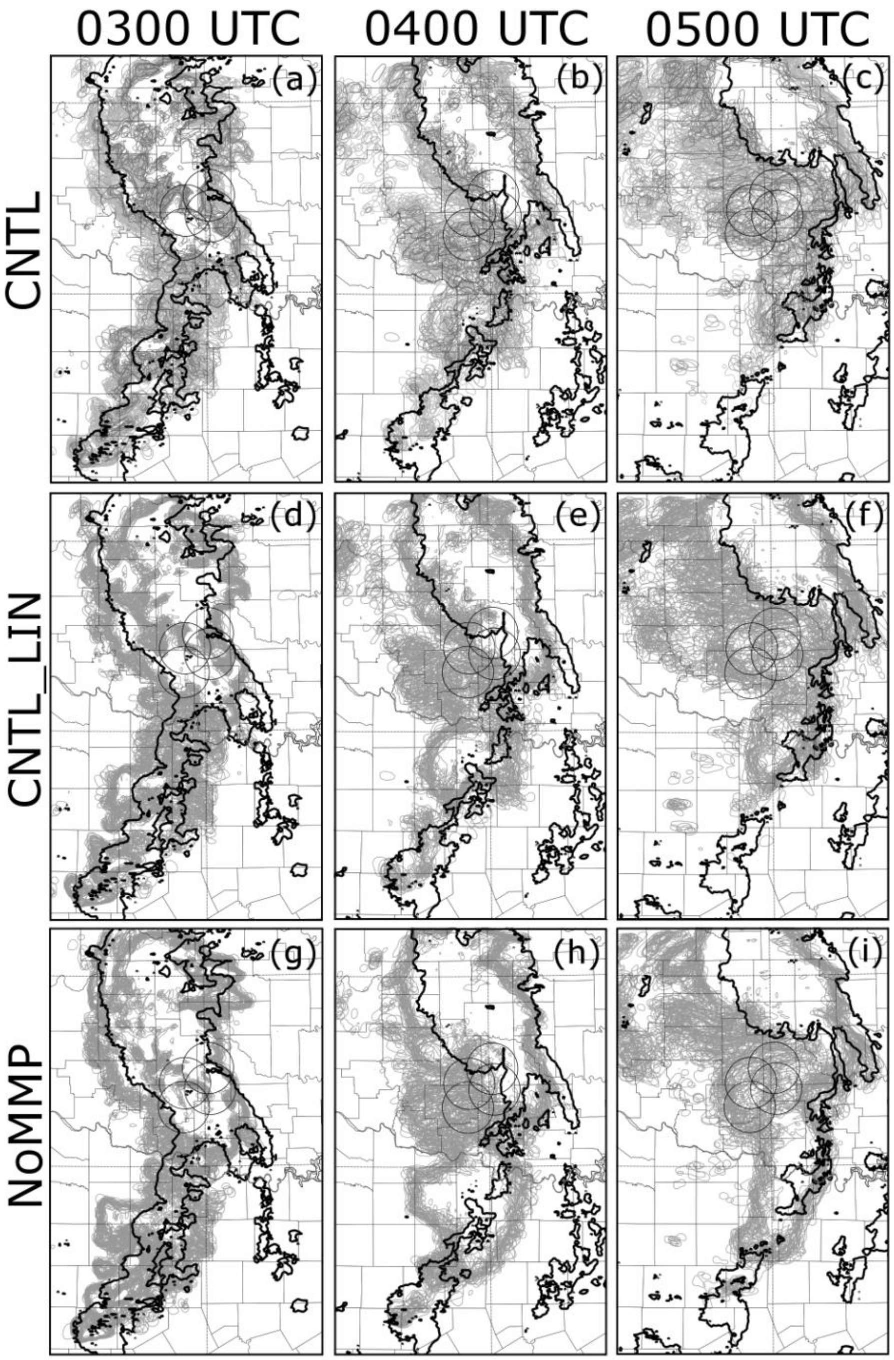

Fig. 4.15: Spaghetti plots of the ensemble forecasts of the 25 dBZ radar reflectivity contours on model grid level 10 (approximately 2 km above the surface) in the area surrounding the CASA domain for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; CNTL_LIN at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC; and NoMMP at (g) 0300 UTC, (h) 0400 UTC, and (i) 0500 UTC. Contours for individual ensemble members are shown as thin gray lines; those for the deterministic forecasts initialized from the ensemble mean at 0200 UTC are in thick gray lines. The 25 dBZ radar reflectivity contour observed by the WSR-88D network at the corresponding time is indicated by the thick, black contour.

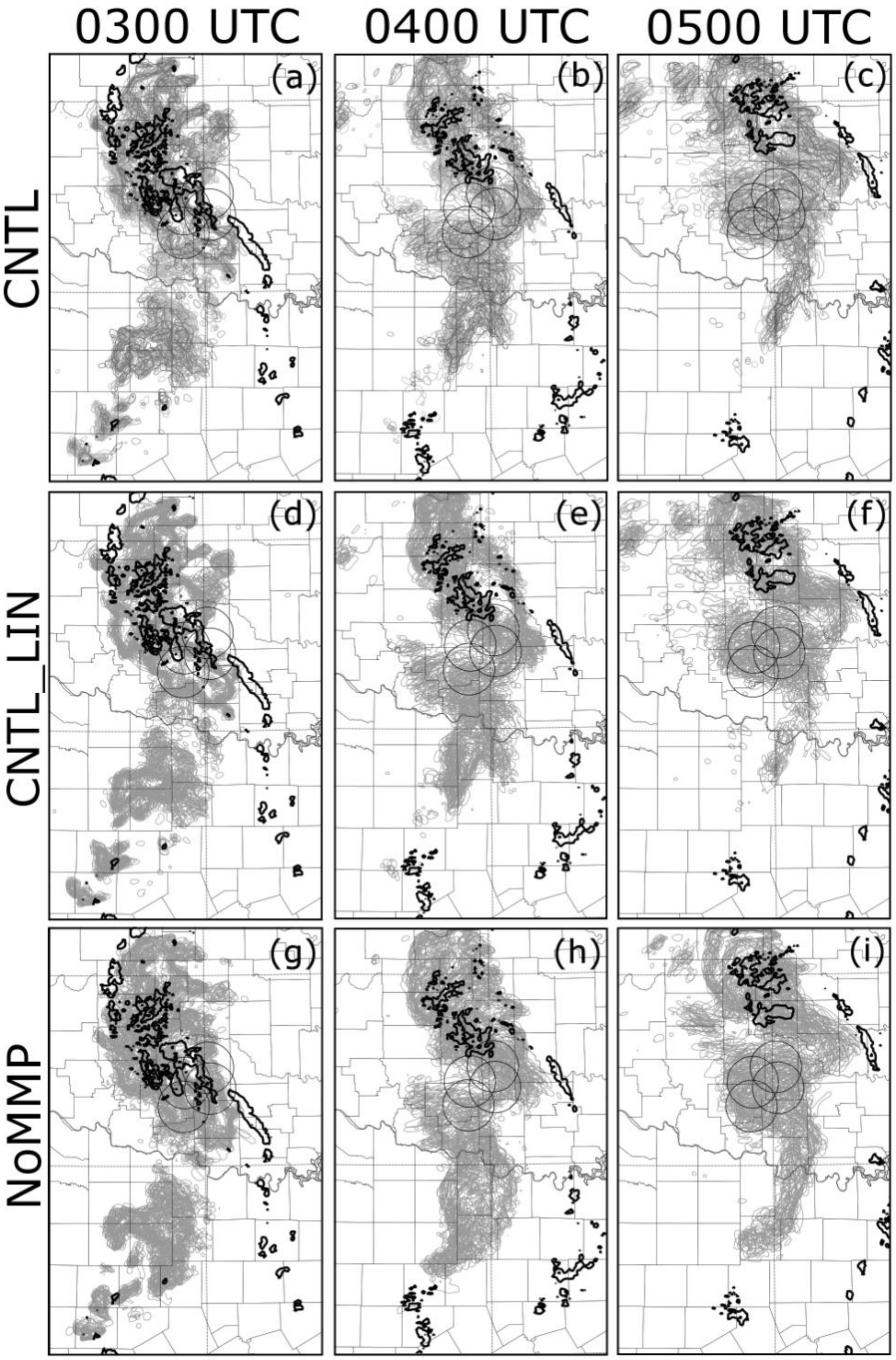

Fig. 4.16: As Fig. 4.15, but for the 40 dBZ radar reflectivity contours.

4.3.3 Bias and error in the ensemble forecasts

During the discussion of the NEP forecasts of radar reflectivity exceeding 40 dBZ (Fig. 4.11), we noted that the ensemble forecasts predict high probabilities of reflectivity exceeding 40 dBZ in regions where no 40 dBZ reflectivity values were observed, particularly at 0400 and 0500 UTC. To more closely examine biases in reflectivity within the ensemble forecasts, we construct domain-wide histograms of radar reflectivity each hour from 0200 UTC to 0500 UTC in the ensemble mean and ensemble member forecasts of CNTL, and compare them to histograms constructed from radar reflectivity observed by the WSR-88D radar network and interpolated to the model grid. Bins are placed at intervals of 1 dBZ, and histograms constructed using the individual ensemble members are normalized by the size of the ensemble. The resulting histograms are shown in Fig. 4.17.

Two prominent biases are evident in the histograms of Z for the both ensemble mean and the ensemble members: an under-prediction (low bias) of weak (<25 dBZ) precipitation regions, and an over-prediction (high bias) of areas of intense (>40 dBZ) precipitation (Fig. 4.17). At the initial time (0200 UTC), only the under-prediction of low Z is present in the ensemble members and the ensemble mean; the histograms match observations well for Z greater than approximately 25 dBZ (Fig. 4.17). By 0300 UTC both biases are present in both the ensemble mean and the ensemble members; the magnitude of these biases remains relatively constant from 0300 to 0500 UTC (Fig. 4.17d-l).

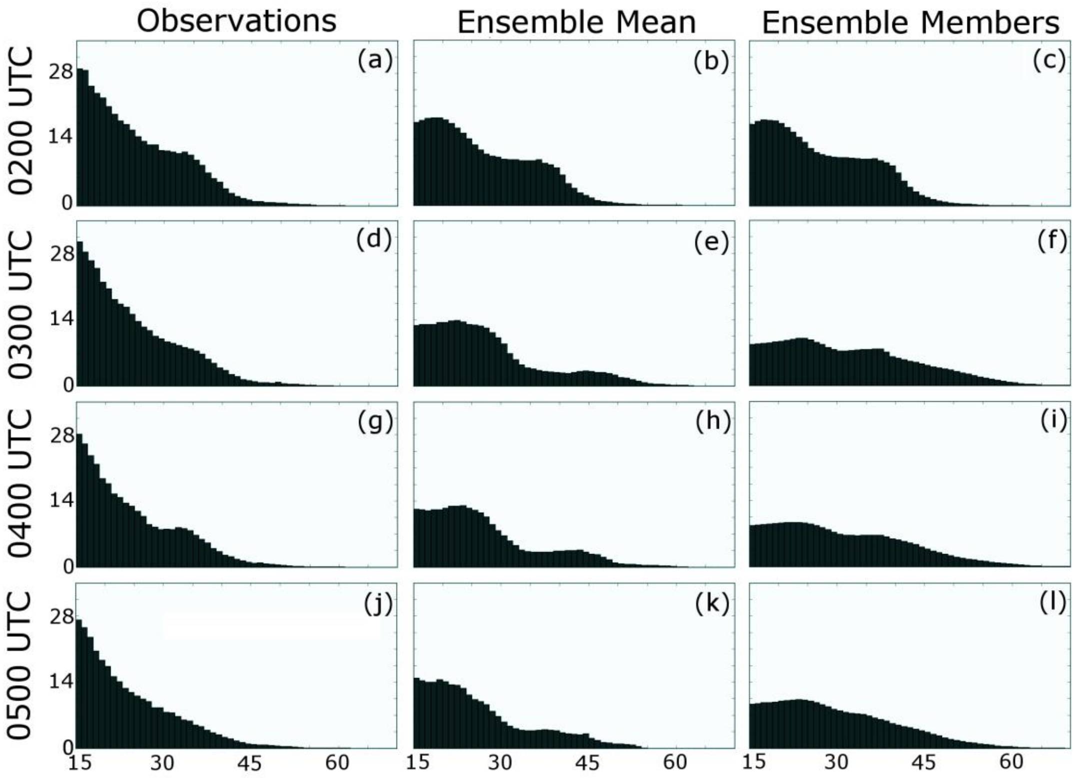

Fig. 4.17: Domain-wide histograms of reflectivity intensity in ensemble forecast experiment CNTL in the initial ensemble state and during the 3-hour forecast period. Shown are WSR-88D observations interpolated to the model grid (left column), the ensemble mean (center column), and individual ensemble members normalized by the size of the ensemble (right column). The vertical axis indicates the number of model grid volumes (in thousands) containing reflectivity of a given intensity.

Both the low-bias in weak precipitation and the high bias in strong precipitation are smaller in magnitude for the ensemble mean than for the individual ensemble members at and after 0300 UTC (Fig. 4.17d-l). The primary source of these differences is smoothing and decreased intensity of Z in the ensemble mean due to differences in the position of small-scale reflectivity features among individual ensemble members. While this smoothing leads to a distribution of Z that agrees more closely with observations in this case, it is not universally desirable; for example, in the absence of the high bias in intense precipitation seen in the ensemble members in this case, the smoothing would degrade, not improve, the distribution of Z in the ensemble mean.

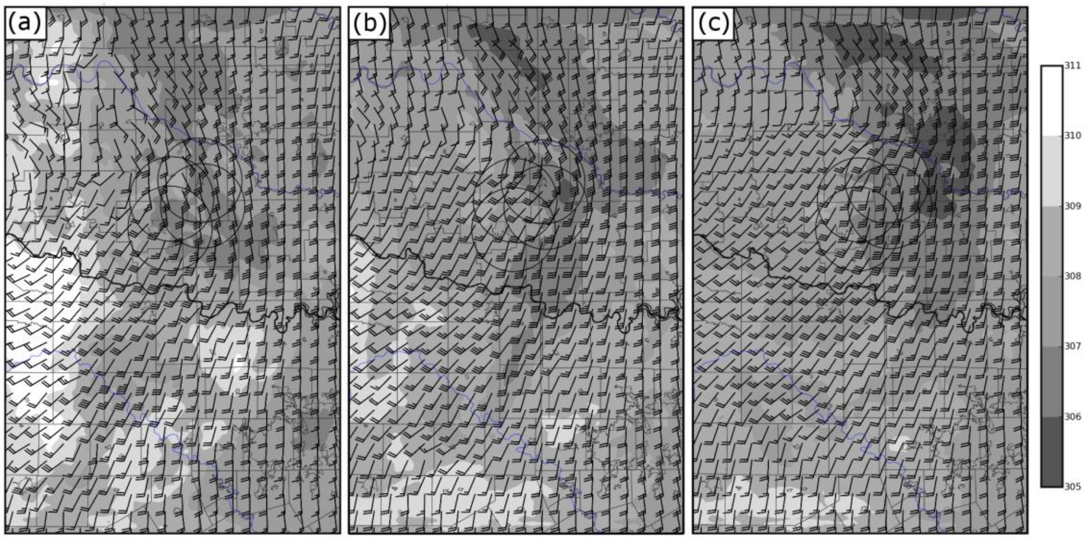

Fig. 4.18: Ensemble mean horizontal winds (barbs) and potential temperature (shaded) in CNTL at model grid level 10 (approximately 2 km above ground level) at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC.

The high bias in intense precipitation observed in the ensemble can be in part attributed to spurious convection that developed in and near the CASA radar network, as noted in section 4.3.1. To provide additional insight into the source of this spurious convection, we calculate ensemble mean potential temperature and horizontal winds approximately 2 km above the surface in CNTL at 0300, 0400, and 0500 UTC (Fig. 4.18). In the ensemble, the cold pool in the wake of the MCS (cf. Fig. 4.9) is relatively weak, with temperatures at and near the surface under the MCS only around 2 to 3 K cooler than the surrounding area (Fig. 4.18). By comparison, many Oklahoma mesonet sites in southwestern Oklahoma recorded temperature drops of around 4K during the passage of the MCS (not shown). In addition, particularly at 0300 UTC (Fig. 4.18a), the strongest convergence at low levels occurred behind the MCS, in and near the CASA domain, where strong south-southwesterly flow just southwest of the CASA domain transitions to weaker southerly flow to the north of the CASA domain (Fig. 4.18). The combination of low-level convergence and relatively warm, buoyant air in the wake of the MCS set the stage for the spurious convection that developed in this region in the ensemble between 0300 and 0500 UTC.

4.3.4 Impact of microphysics on ensemble reflectivity forecasts

To examine the impact of a model microphysics during the ensemble forecasts, we calculate and compare RMS innovation of radar reflectivity (Z) for ensemble forecast experiments CNTL and CNTL_Lin. CNTL and CNTL_Lin ensembles start from the same initial conditions, but CNTL uses a mixed-microphysics ensemble, while CNTL_Lin uses a single-microphysics ensemble consisting only of Lin microphysics members. Innovation is calculated using all grid volumes where the reflectivity observed by the WSR-88D radar network (interpolated to the model grid) exceeds 15 dBZ; this analysis technique is similar to that used in Aksoy et al. (2010). Innovation values for each individual ensemble member, as well as the ensemble mean and probability-matched (PM) ensemble mean, are shown at 1, 2, and 3 hours of forecast time (0300, 0400, and 0500 UTC respectively) in Fig. 4.19.

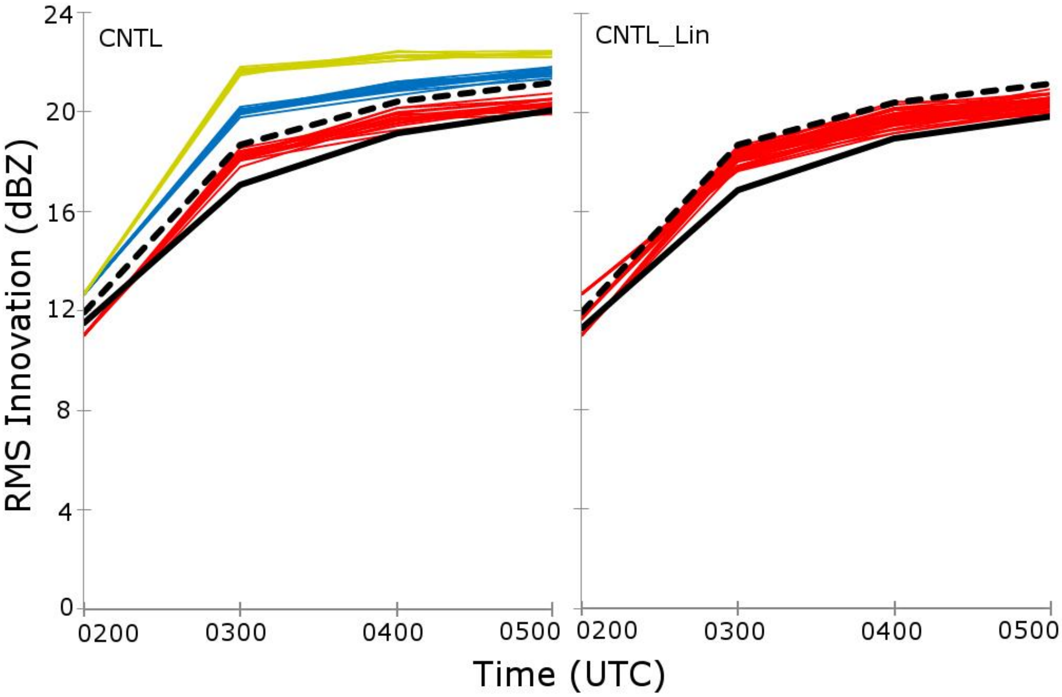

Fig. 4.19: RMS innovation of radar reflectivity during the forecast period for ensemble forecasts CNTL (left) and CNTL_LIN (right). The ensemble mean forecast is indicated by the thick, solid black line, while the probability-matched ensemble mean forecast is indicated by the dashed black line. The thin gray lines indicated innovation within individual ensemble members, with different shades of gray used for members using differing microphysical parameterization schemes.

Both CNTL and CNTL_Lin exhibit similar evolution of RMS innovation of *Z* for the ensemble mean and PM mean forecasts. The ensemble mean consistently outperforms the PM mean, as well as most individual ensemble members (Fig. 4.19). This result is in agreement with Aksoy et al. (2010), whose ensemble mean RMS innovation of *Z* was also lower than most of their ensemble members. The relatively good performance of the ensemble mean can be attributed to the high bias is the reflectivity forecast (Fig. 4.17) discussed earlier, since position differences in the strongest reflectivity cores between individual members give an ensemble mean *Z* field that is smoother and contains lower reflectivity values than the individual members. The PM mean, though it contains the same spatial structure as the ensemble mean, has the same high bias as the individual ensemble members. Because the ensemble mean consistently outperforms the PM mean, we can conclude that it is the absence of the high bias, rather than improved spatial structure of the Z field, that results in lower RMS innovation of Z in the ensemble mean.

In the CNTL ensemble, the individual ensemble members exhibit a trimodal distribution, grouped into three clusters segregated by the microphysical scheme used in the model. The cluster of members using the Lin microphysical scheme has the lowest RMS innovation, followed by the cluster of members using WSM-6 microphysics, and the cluster using NEM microphysics has the highest RMS innovation values. The strong clustering in CNTL indicates that error associated with the microphysical parameterization is a dominant factor in model error in Z during the ensemble forecast; in CNTL_LIN, where all members used the Lin microphysical scheme, the ensemble members have a unimodal distribution. Though the increased ensemble dispersion of CNTL is desirable, ensemble dispersion could also be increased using other methods, such as perturbed boundary conditions, perturbations within a microphysical scheme, and the use of different spread-maintenance techniques during the assimilation period; such methods will be a subject for future work. Because the microphysical parameterization appears to be a dominant source of error in the forecasts, using a more sophisticated 2- or 3-moment microphysical scheme might also be desirable to reduce overall RMS error in the ensemble; this too could be a subject for future studies.

4.3.5 Ensemble and probabilistic forecasts of tornadic mesovortices

In recent years, explicit prediction of tornadic circulations in short-term convective forecasts has become a reality (Xue et al. 2007). The capability to perform such prediction in real-time is a cornerstone of the "warn-on-forecast" paradigm being explored by the NWS for use in future operations (Stensrud et al. 2009). For forecasts of tornadoes, the most likely initial operational implementation would be a probabilistic approach focusing on predicting the probability of tornadoes on a timescale of no more than a few hours (Stensrud et al. 2009). At a horizontal grid spacing of 2 km we cannot hope to resolve torandoes, but we can resolve meso-vortices on the scale of 10 km and greater. Such meso-vortices were present in the 9 May 2007 LEV case, and, as noted above, the location and timing of these vortices correspond well to the location and timing of tornadoes reported in association with the system (see Fig. 4.1; also Fig. 3.1). As all forecasts were initialized at 02:00 UTC, the reported tornadoes (which occurred at 03:54, 04:26, and 04:43 UTC) fall at forecast times of approximately 2 to 3 hours.

Forecast ensemble members of CNTL, NoMMP, and NoCASA were manually examined at 0400, 0420, and 0440 UTC to identify significant low-level vortices, as outlined in section 2. Such low-level vortices are considered objects for the purpose of verification. Equation 2.2 was then applied to generate forecasts of the probability of a significant (vertical vorticity $\zeta > 0.02\ s^{-1}$) low-level vortex being present within 25 km of a point. The probabilities thus calculated are shown in Fig. 4.20 at 0400, 0420, and 0440 UTC for CNTL (Fig. 4.20a-c), NoMMP (Fig. 4.20d-f), and NoCASA (Fig. 4.20g-i), along with the locations of reported tornadoes (which were co-located with their parent mesovortices).

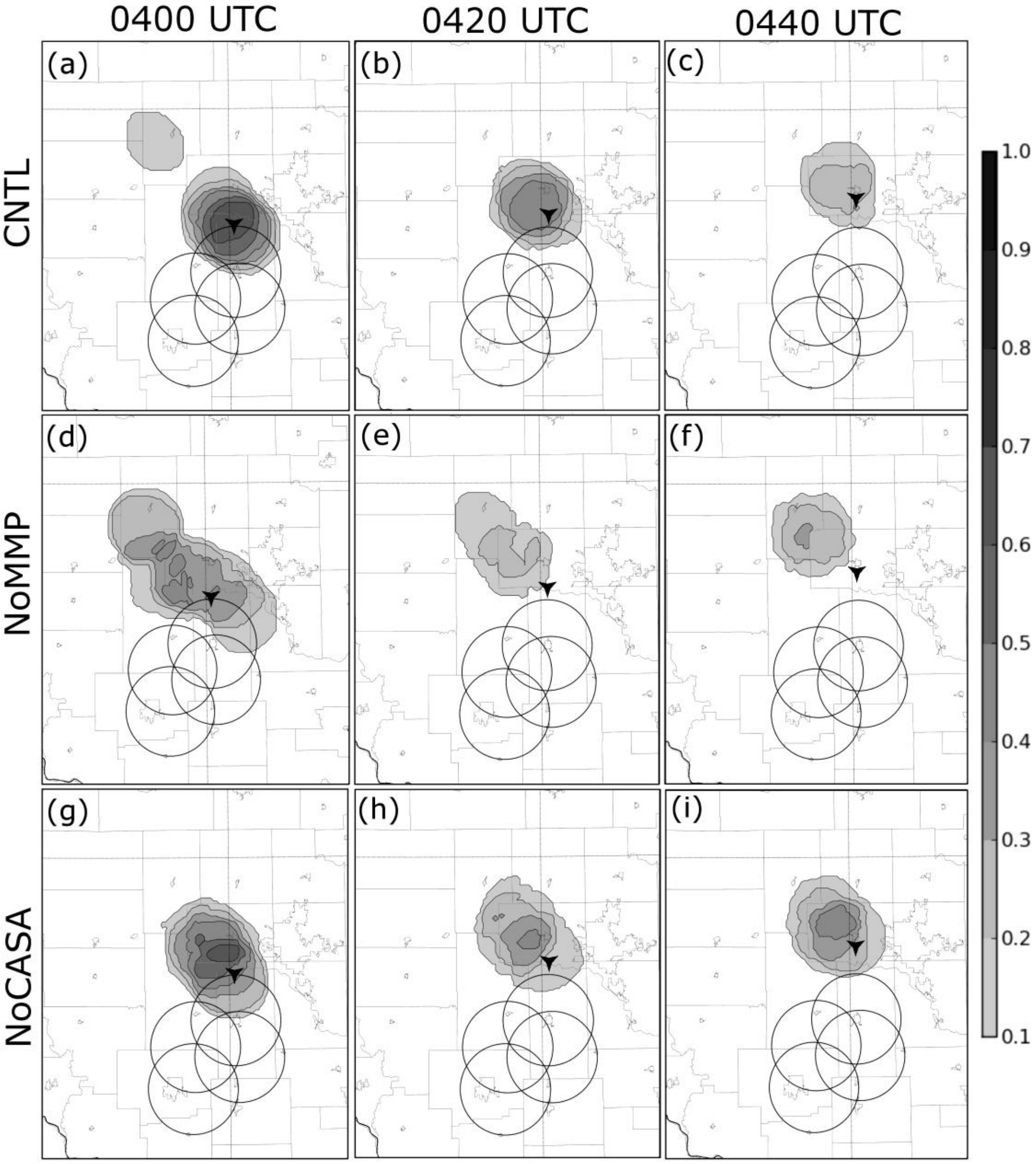

Fig. 4.20: Ensemble-based forecast of P[*low-level circulation with vertical vorticity* $\zeta > 0.02\ s^{-1}$ *within 25 km of a point*] *for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; NoMMP at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC; and NoCASA at (g) 0300 UTC, (h) 0400 UTC, and (i) 0500 UTC. The triangles indicate the reported locations of the tornadoes reported at 0354 UTC (in panels (a), (d), and (g)), at 0426 UTC (in panels (b), (e), and (h)), and at 0443 UTC (in panels (c), (f), and (i)).*

All three forecast ensembles predict near-surface vortices in the vicinity of the reported tornadic activity at 0400, 0420, and 0440 UTC (Fig. 4.20). At 0400 and 0420 UTC, the probability field of CNTL exhibits a tight concentration of the highest probability values close to the reported tornado location (Fig. 4.20a, b). At 0400 UTC, the location of the maximum near-surface vortex probability for CNTL of 0.65 (Fig. 4.20a) is located within 3 km of reported tornado, and its probability field is tightly concentrated around the location of observed tornado. In NoMMP and NoCASA (Fig. 4.20d, g) maximum probabilities are located approximately 10 and 20 km, respectively, from the reported tornado location. NoCASA and NoMMP forecast probabilities of only 0.43 and 0.35 (respectively) at the reported tornado location at 0400 UTC, and, especially in NoMMP (Fig. 4.20d), have relatively diffuse probability distributions with lower maxima. In NoMMP, this difference results from a wider spread of vortex location predictions seen within the single-microphysics ensemble. The only difference between NoCASA and CNTL, however, was the absence or presence of CASA radar data during the assimilation period, suggesting that assimilating CASA data positively impacts the prediction of low-level vortex features, yielding a higher maximum probability prediction at 0400 UTC in the vicinity of the observed tornado. This finding is consistent with that of Schenkman et al (2011), where low-level radial velocity data from CASA radars were found to improve the LEV forecast.

The results at 0420 UTC (Fig. 4.20b, e, and h) are similar to those at 0400 UTC: CNTL (Fig. 4.20b) produces the highest maximum probability (between 0.4 and 0.5) centered very close to the reported tornado location. NoCASA (Fig. 4.20h) and NoMMP (Fig. 4.20e) have position predict lower probabilities (approximately 0.2 and 0.1 respectively) of a significant low-level vortex being present in the vicinity of the observed tornado. The probability in the vicinity of the observed tornado is lower in all three experiments at 0420 UTC than at 0400 UTC. At 0440 UTC (Fig. 4.20g-i), CNTL features a probability field with the highest predicted probabilities in close proximity to the observed tornado, but NoCASA predicts a slightly higher probability near the observed tornado. NoMMP performs worst at 0440 UTC, (Fig. 4.20h), predicting a low probability of vortices, with the highest probabilities located several tens of kilometers away from the reported tornado. Because of the small size of the CASA domain and increasing impact of model factors (such as microphysical parameterization) as the forecast period proceeds, the greatest impact of CASA data from the initial condition of the ensemble would be expected at shorter forecast times, consistent with the results shown in Fig. 4.20.

One issue that arises when considering prediction of tornadic meso-vortices is a question of the limits of predictability for these features. Smaller features within atmospheric flows tend to exhibit error growth on shorter timescales (Lorenz 1969), and indeed, the tornadoes observed during this case formed and decayed with timescales of less than an hour. The tornadic meso-vortices within the 8-9 May 2007 MCS, however, were strongly forced by larger-scale features within the MCS that spawned them. Because MCS’s often exhibit predictability on a significantly longer timescale (Carbone et al. 2002), we believe that is therefore reasonable to consider ensemble predictions of the tornadic meso-vortices within it at timescales of 2-3 hours because of this larger-scale forcing.

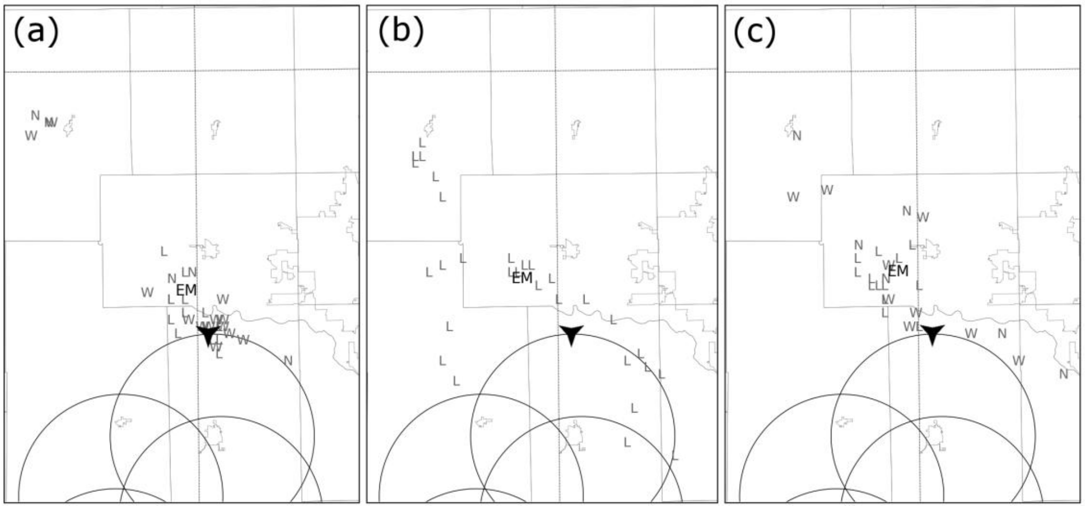

Fig. 4.21: Locations of significant (vertical vorticity $\zeta > 0.02\ s^{-1}$), discernible, low-level (grid level 10, approximately 2 km above the surface) mesovortex centers for all ensemble members at 0400 UTC in experiments (a) CNTL, (b) NoMMP, and (c) NoCASA. Individual members are coded by microphysical type; members using Lin microphysics are marked "L", members using WSM6 microphysics are marked "W", and members using NEM microphysics are marked "N". The ensemble mean vortex location is marked "EM". The location of the EF1 tornado reported at 0354 UTC is marked by the black triangle.

We further examine the role of microphysics in vortex prediction by plotting locations of low-level vortex centers in individual ensemble members at 0400 UTC, along with the ensemble mean vortex location and the location of the EF-1 Minco tornado (reported to have begun at 0354 UTC), for CNTL (Fig. 4.21a), NoMMP (Fig. 4.21b), and NoCASA (Fig. 4.21c). Each low-level vortex center is marked by a letter corresponding to the microphysical scheme used in the member that produced it ("L" for Lin, "W" for WSM6, and "N" for NEM). CNTL (Fig. 4.21a) exhibits the closest clustering of predicted low-level vortices around the location of the observed tornado; 26 of 40 (65%) ensemble members predict a low-level vortex center within 25 km. These 26 members include 12 of 16 Lin members and 11 of 16 WSM6 members, but only 3 of 8 NEM members. In NoCASA (Fig. 4.21c) 17 of 40 (43%)

members predict vortices within 25 km of the observed tornado; these 17 members consist of 10 Lin, 5 WSM6, and 2 NEM members. In CNTL and NoCASA, Lin members have the least error in vortex location, while NEM members have the most; many NEM members are outliers (see Fig. 4.21a, c), or fail to produce significant low-level vortices. In NoMMP (Fig. 4.21b), predicted vortex locations vary substantially among ensemble members; only 14 members (35%) predict low-level vortices within 25 km of the observed tornado location, underscoring the positive impact of a mixed-microphysics forecast ensemble in meso-vortex prediction.

Despite differences in the distribution of individual members, the forecast ensemble mean low-level vortex locations (marked by “EM” in Fig. 4.21a-c) were similar in all three experiments. In each case the ensemble mean location was located slightly to the northwest of the observed tornado location, with a position error ranging from approximately 12 km in CNTL (Fig. 4.21a) to 18 km in NoCASA (Fig. 4.21c). This result underscores the importance of probabilistic information on forecast uncertainty; while CNTL produced a much sharper probabilistic forecast of the low-level vortex than NoMMP, ensemble mean vortex locations differed by only about 3 km, or 1.5 grid intervals (Fig. 4.21a, b) between the two experiments. Also, NoCASA forecasted a higher vortex probability with more individual members clustered around the true vortex location than NoMMP but its ensemble mean vortex location had a larger position error (Fig. 4.21b, c). Probabilistic information is important in assessing the reliability of these forecasts; an ensemble mean or single deterministic forecast cannot provide such information.

Chapter 5: 8-9 May 2007—Nested-Grid EnKF Analysis of Conventional and Radar Data and Comparison with 3DVAR

5.1 Data and Methodology

While reasonable analyses and forecasts are obtained in the single-grid experiments presented in Chapters 3 and 4, those single-grid experiments are subject to a number of limitations. Due to technical restrictions and the limits of available computing resources at the time they were run, the single-grid experiments are limited to a relatively small domain, causing interaction between the trailing convective line of the MCS and the southern domain boundary, as noted in Chapter 4. Furthermore, the single-grid experiments use a single lateral boundary condition for all ensemble members, reducing spread within the ensemble, and make no use of conventional weather observations.

Recent advances in the ARPS EnKF system and an increase in available computing resources have made it possible to move beyond the limitations noted above. To that end, we perform a set of nested-grid EnKF data assimilation and forecast experiments that expand upon the single-grid experiments of Chapters 3 and 4. In addition to addressing the limitations of the single-grid experiments, the nested-grid experiments also investigate additional influences on the EnKF analysis and resulting ensemble forecasts, including the impact of perturbing the rain drop size distribution within a single-microphysics ensemble, and the impact of varying the assumed observation error for radar reflectivity and radial velocity data during EnKF data assimilation.

5.1.1 Model Setup

Like the experiments presented in Chapters 3 and 4, the nested-grid experiments presented here use the ARPS NWP model for assimilation of data and generation of forecasts. Full terrain, soil, surface, and radiation physics are used, with settings matching those of the previous experiments (see section 3.1). Unlike the previous experiments, however, a set of two nested domains is used, with 40-member ensembles used both on an outer domain with a horizontal resolution of 6 km, and an inner domain with a horizontal resolution of 2 km. This setup allows for the introduction of mesoscale perturbations on the outer domain and the inclusion of boundary condition perturbation on the inner domain by using each outer domain member to provide the boundary and initial conditions for the corresponding member on the inner domain. The geographic extent of the two domains is shown in Fig. 5.1. The outer (6 km horizontal grid spacing) domain has a physical extent of 320 × 320 × 40 grid points and covers much of the Great Plains, the Mississippi Valley, and the southern Rocky Mountains, as well as portions of northern Mexico and the northwestern Gulf of Mexico. The inner (2 km horizontal grid spacing) domain has a physical extent of 500 × 500 × 40 grid points and covers all of Oklahoma, much of Texas and Kansas, and portions of surrounding states. The inner domain is identical to that used in Schenkman et al. (2011), facilitating direct comparison between the results of their 3DVAR-based study of the 8-9 May 2007 tornadic MCS and similar analysis and forecast results obtained in this study.

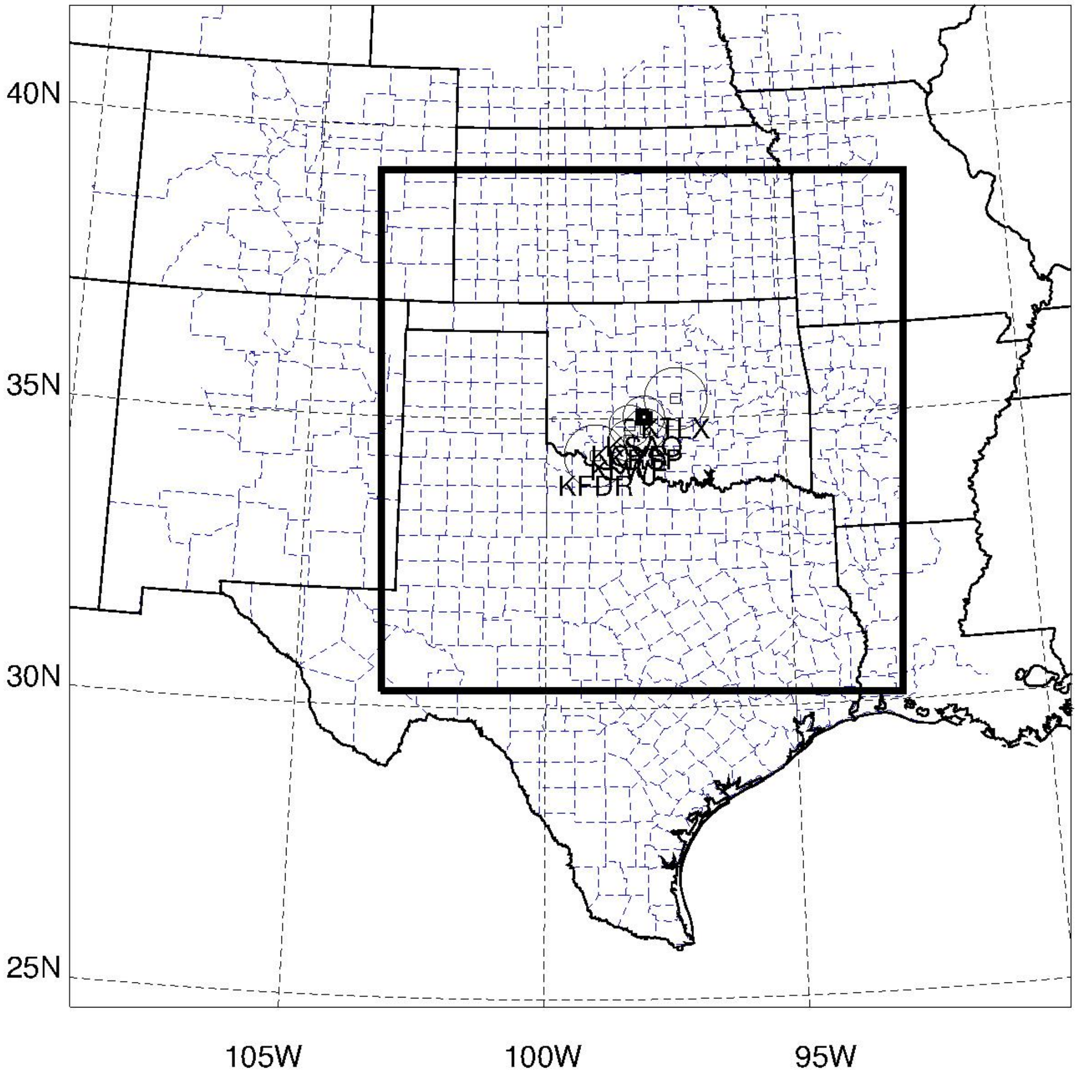

Fig. 5.1: Geographic extent of the outer domain with 6 km horizontal resolution and the nested domain with 2 km horizontal resolution (represented by the thick black box in the center of the figure) used in experiments presented in Chapter 5. CASA IP-1 radar 30 km range rings and 60 km range rings for WSR-88D sites KTLX and KFDR are included for reference.

5.1.2 Experiment Design

In this study, a set of data assimilation and forecast experiments are performed on a 500 × 500 × 40 physical domain with 2 km horizontal grid spacing and stretched vertical grid spacing; the average vertical grid spacing is 500 m, with a minimum vertical spacing of 100 m near the surface. This grid is identical to that used in Schenkman et al. (2011), who performed cycled data assimilation using the ARPS 3DVAR system and a cloud analysis package. As in the experiments of Chapters 3 and 4, full model physics are used (Xue et al. 2001), including the NASA Goddard Space Flight Center long- and shortwave radiation parameterization, a two-layer soil model, surface fluxes parameterized using predicted surface temperature and water content, and a 1.5-order turbulent kinetic energy (TKE)-based sub-grid-scale turbulence parameterization, along with high-resolution terrain. Like the experiments presented in chapters 3 and 4, a 40 member ARPS ensemble is used for each experiment. However, unlike those prior experiments, no mixed-microphysics experiments are performed; instead, all experiments use a variant of the Lin et al. (1983) microphysics scheme with the default rain intercept parameter reduced to $8.0 \cdot 10^5$ in accordance with the findings of Snook et al. (2008); this is a reduction by a factor of 10 from its original setting.

Three experiments are performed with the goal of direct comparison with the results of Schenkman et al. (2011) obtained using 3DVAR. The first of these three experiments is a control experiment, hereafter referred to as CNTL, which will serve as a basis for comparison for all other experiments detailed in this chapter. In CNTL, both CASA and WSR-88D radar data are assimilated, along with conventional observations. The EnKF data assimilation system is configured following the experiments from Chapter 3, with the exception of the assumed observation error for V_r, which is increased from 2 ms^{-1} to 3 ms^{-1} and the assumed observation error for Z, which is increased from 1 dBZ to 2 dBZ. The second experiment, NoCASA, is identical to CNTL except that only WSR-88D radar data and conventional observations are assimilated. The third experiment, NoRAD, assimilates conventional data only. These three experiments correspond closely to three experiments of Schenkman et al. (2011), facilitating direct comparison of analysis and forecast results; the comparison will be presented in section 5.4.

Two more experiments are conducted focusing on the impact of microphysics during the analysis and subsequent ensemble forecast. Experiment R6 uses an identical setup to CNTL, except that the rain intercept parameter of the Lin microphysics scheme used in the ensemble members is increased from 8.0×10^5 to 4.0×10^6, the default value of rain intercept parameter corresponding to the Marshall-Palmer distribution. The cold pool observed in the single-grid CNTL experiment of Chapter 3 was found to be relatively weak. Results from Snook and Xue (2008) indicate that an increased rain intercept parameter corresponds to more intense cold pools due to increased evaporative cooling, thus the increased rain intercept parameter of R6 may produce a cold pool closer in intensity to that observed by Oklahoma Mesonet sites during the 8-9 May 2007 MCS.

During the forecast period, the innovation of Z within the mixed-microphysics single-grid forecast ensembles of Chapter 4 exhibited a trimodal distribution where the modes were segregated by microphysical scheme. Though the enhanced ensemble spread imparted by the mixed-microphysics ensemble is desirable, the trimodal behavior is not. Another possible method of increasing ensemble spread by microphysical variation is perturbation of the intercept parameters within a single-microphysics ensemble; previous work by Snook and Xue

(2008) has shown that variation of the intercept parameters within a single-moment microphysics scheme can have a large impact on the dynamics and structure of the resulting storm. Experiment PERT applies this technique, randomly varying the rain intercept parameter of the Lin microphysical scheme in each member between $8.0 \cdot 10^5$ and $8.0 \cdot 10^6$. The variation in rain intercept parameter is achieved by selecting the intercept parameter for each member randomly from a Gaussian distribution with a mean of $4.4 \cdot 10^6$ and a standard deviation of $1.6 \cdot 10^6$. Random samples of 40 members from this distribution were performed until a set was obtained that met two criteria: (1) the set contained no values more than two standard deviations from the mean, and (2) the absolute value of the sum of deviations of all members was less than one half of the standard deviation of the distribution from which the samples were generated. These two criteria were chosen to ensure that the randomly selected rain intercept parameter values fall within the range specified above, and that their distribution is balanced around the mean. Once selected, the intercept parameter value is kept fixed for each member.

Finally, experiment CNTL_LOW is performed using reduced values of assumed observation error for V_r and Z of 1 ms^{-1} and 2 dBZ, respectively. These values are identical to those used in the control experiment in Chapter 3. When the results of the experiments in Chapter 3 were published (Snook et al. 2011), it was suggested that the relatively low assumed observation errors may have negatively impacted the analysis, a suggestion that will be investigated through comparison of results from CNTL and CNTL_LOW. For reference, a summary of all experiments is provided below in Table 5.1.

Table 5.1: Summary of the inner-domain experiments presented in Chapter 5.

	Use of CASA	Use of WSR-88D	Assumed observation error		Rain intercept parameter
			V_r ($\mathrm{ms^{-1}}$)	Z (dBZ)	
CNTL	Yes	Yes	2.0	3.0	$8.0{\cdot}10^5$
NoCASA	No	Yes	2.0	3.0	$8.0{\cdot}10^5$
NoRAD	No	No	2.0	3.0	$8.0{\cdot}10^5$
R6	Yes	Yes	2.0	3.0	$4.0{\cdot}10^6$
PERT	Yes	Yes	2.0	3.0	Varied
CNTL_LOW	Yes	Yes	1.0	2.0	$8.0{\cdot}10^5$

All experiments (CNTL, NoCASA, NoRAD, R6, PERT, and CNTL_LOW) are initialized from a 40-member ensemble of EnKF analyses performed on the outer grid described in section 5.1.1; the outer domain has a horizontal grid spacing of 6 km, and its geographic extent is shown in Fig. 5.1. The outer-grid forecast is initialized at 1800 UTC on 8 May 2007 from the 8 May 2007 NCEP 1800 UTC NAM analysis, and a subsequent 3-hour pre-forecast is performed from the initialization. At 2100 UTC, an ensemble of 40 members is generated by introducing smoothed, random, Gaussian mesoscale perturbations using the method of Tong and Xue (2008). Perturbations are added to the horizontal wind (u, v) with an average magnitude of 2.0 $\mathrm{ms^{-1}}$, the potential temperature (θ) using positive perturbations only with a mean magnitude of 1K, and the mixing ratio of water vapor (q_v) with a mean magnitude of 10% of the q_v value at the given grid point. The horizontal length scale of these perturbations is 36 km, and the vertical length scale is 7.2 km.

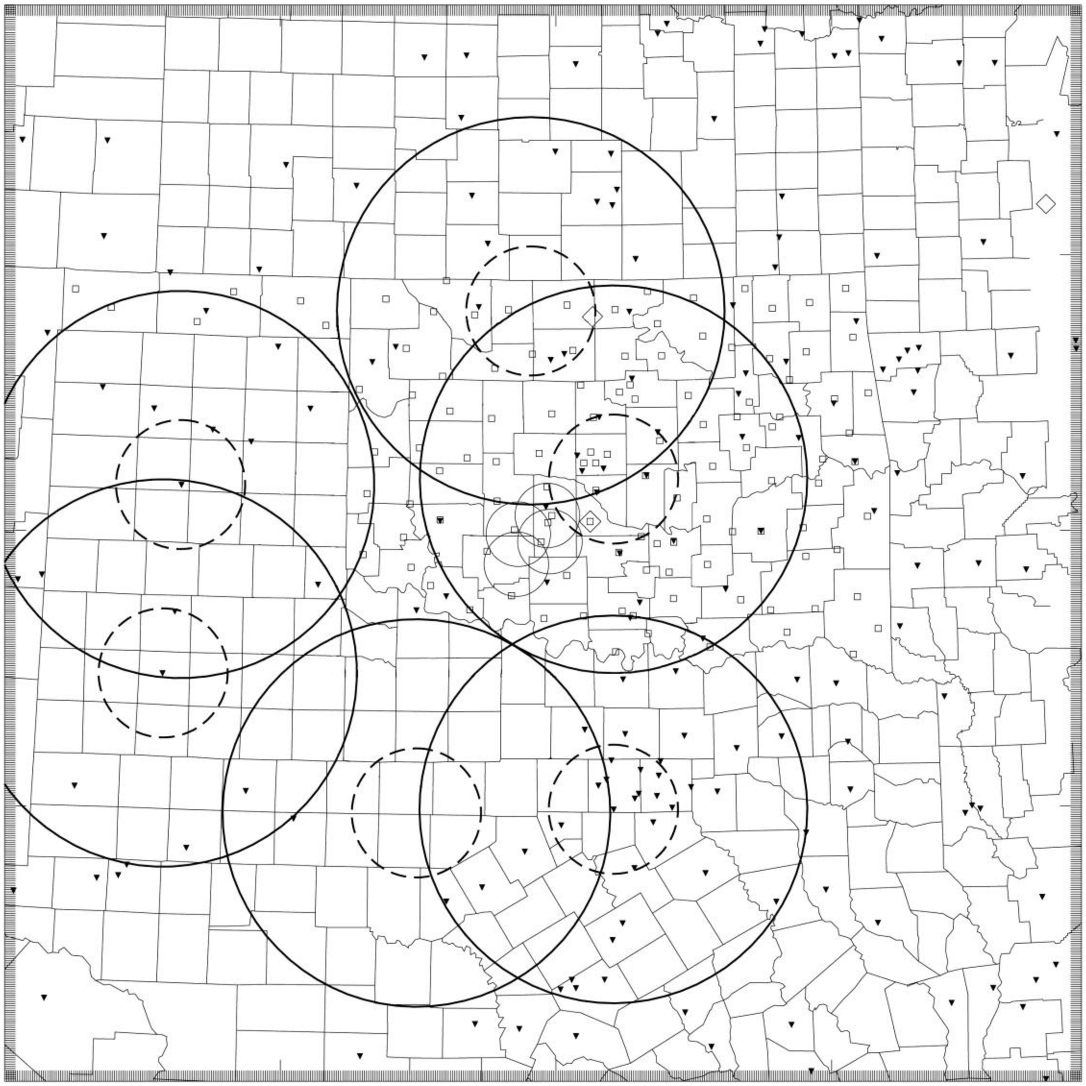

Fig. 5.2: Observations used in the three experiments described in this chapter performed on the 2 km horizontal resolution domain. Locations of Oklahoma Mesonet sites are indicated by open squares, locations of ASOS sites by filled triangles, and locations of wind profilers by open diamonds. The dotted and solid black circles represent 60 km and 180 km range rings of WSR-88D radar sites respectively. The small gray circles indicate 30 km range rings of the CASA radar sites. This figure is adapted, with permission, from Fig. 3a of Schenkman et al. (2011).

Conventional observations, including ASOS, mesonet observations, wind profiler data, and upper air observations (including soundings at 0000 UTC, 9 May 2007) are assimilated hourly on the outer grid from 2200 UTC, 8 May 2007 to 0100

UTC, 9 May 2007 using the ensemble square-root filter (EnSRF) algorithm based on that of Whitaker and Hamill (2002). Assumed observation errors used for conventional data vary by observation type as indicated in Table 5.1.

Table 5.2: Assumed observation error magnitude for conventional observations.

Data Type	u (ms^{-1})	v (ms^{-1})	Temperature (K)	Dewpoint (K)	Pressure (hPa)
Surface	1.5	1.5	1.5	2.0	2.0
Upper-air	2.5	2.5	1.2	2.0	0.6
Profiler	2.5	2.5			

In order to maintain ensemble spread during the assimilation of conventional data on the outer 6 km grid, we apply a multiplicative covariance inflation factor of $\gamma = 1.03$ to the prior deviation of each ensemble member from the ensemble mean (Anderson and Anderson, 1999; Tong and Xue, 2005). Covariance inflation is performed throughout the entire model domain. In addition to covariance inflation, we also apply the relaxation technique of Zhang et al. (2004) with a relaxation coefficient of 0.5.

At 0100 UTC, 9 May 2007, the inner-grid experiments are initialized from the interpolated 6-km ensemble forecasts valid at the same time. Each inner-domain ensemble member uses boundary conditions derived from the corresponding outer-domain member; for example, ensemble member #3 of any inner-grid experiment (e.g. CNTL, NoCASA, NoRAD) derives its initial conditions and boundary conditions from ensemble member #3 of the outer-domain ensemble.

Boundary condition files from the outer-domain members are provided at 15 minute intervals throughout the analysis and forecast period (0100 to 0500 UTC).

In each of the inner-grid experiments, conventional data are assimilated using the same EnSRF algorithm. Data are assimilated at five minute intervals; the first analysis occurs at 0105 UTC, 9 May 2007, and the final analysis occurs at 0200 UTC. Conventional data used in this study include surface observations from Automated Surface Observing System (ASOS) stations and the Oklahoma Mesonet (Brock et al. 1995); as well as upper-air observations from wind profilers. Mesonet data are assimilated every 5 minutes, while ASOS observations and profiler observations, which have a coarser temporal frequency, are assimilated only at 0200 UTC. No sounding data is assimilated because none is available during the data assimilation period. In addition to these conventional data, some of the experiments also assimilate radar observations from the CASA and WSR-88D radar networks.

In experiments assimilating radar data, observations are assimilated along with conventional data at 5 minute intervals from 0105 UTC to 0200 UTC. Level-II radial velocity and radar reflectivity observations are assimilated from six WSR-88D radar sites: Twin Lakes, OK (KTLX); Vance Air Force Base, OK (KVNX), Dyess Air Force Base, TX (KDYX); Amarillo, TX (KAMA); Dallas-Fort Worth, TX (KFWS); and Lubbock, TX (KLBB). In experiments using CASA data, radial velocity and radar reflectivity data are also assimilated from the four X-band radars of the CASA radar network: Cyril, OK (KCYR); Lawton, OK (KLWE); Rush Springs, OK (KRSP); and Chickasha, OK (KSAO). CASA data are assimilated as aggregate 5-minute pseudo-volumes produced using temporal and spatial interpolation, as discussed in section 3.2. The locations of conventional observation sites and radar sites are summarized in Fig. 5.2.

To help maintain ensemble spread, multiplicative covariance inflation is applied to the prior deviation of each ensemble member from the ensemble mean using an inflation factor of $\gamma = 1.03$ in experiment NoRAD (which assimilated only conventional data), and an inflation factor of $\gamma = 1.25$ in all other experiments (which assimilated both conventional and radar data). Unlike on the outer-grid domain, relaxation is not used for the maintenance of ensemble spread on the inner-grid domain because preliminary experiments using relaxation while assimilating radar data resulted in relatively poor ensemble forecast performance compared to experiments using multiplicative covariance inflation only. The larger covariance inflation factor in the experiments assimilating radar data is necessary to maintain ensemble spread while assimilating large quantities of data from multiple radar sites.

At the end of the data assimilation period, a three-hour ensemble forecast is performed for each inner-grid experiment, running from 0200 UTC to 0500 UTC, using the final ensemble member analyses at 0200 UTC as initial conditions. As in Chapter 4, forecast verification is performed for radar reflectivity (as a proxy for quantitative precipitation) and low-level mesovortices (as an indicator of tornado potential) using both qualitative and quantitative methods. Quantitative verification methods include calculation of relative operating characteristic skill score (RSS) and Brier scores for ensemble forecasts. Additional discussion of these skill score metrics can be found in section 4.3.2. Finally, comparisons are performed between the experiments presented herein and the corresponding experiments of Schenkman et al. (2011).

5.2 Analysis Results

For the six inner-grid experiments (see Table 5.1), conventional and/or radar data are assimilated every 5 minutes between 0105 UTC and 0200 UTC. No analysis is performed at 0100 UTC; this is to prevent double-counting of observations used in the outer-grid 0100 UTC analysis ensemble, from which the inner-grid analysis ensembles are initialized. The 0200 UTC analysis ensemble mean for radar reflectivity at model grid level 10 (approximately 2 km above the surface) is shown below in Fig. 5.3 for all experiments. By 0200 UTC, the final analysis of radar reflectivity shows a very similar ensemble mean analysis state for all experiments that assimilated radar data (Fig. 5.3a, b, d-f). The final analysis state of these ensemble members agrees reasonably well the observed radar reflectivity at 0200 UTC (Fig. 5.3g) in the overall placement of convective features, including the leading convective line, the trailing convective line extending south into Texas, and the large stratiform region to the west of the leading convective line.

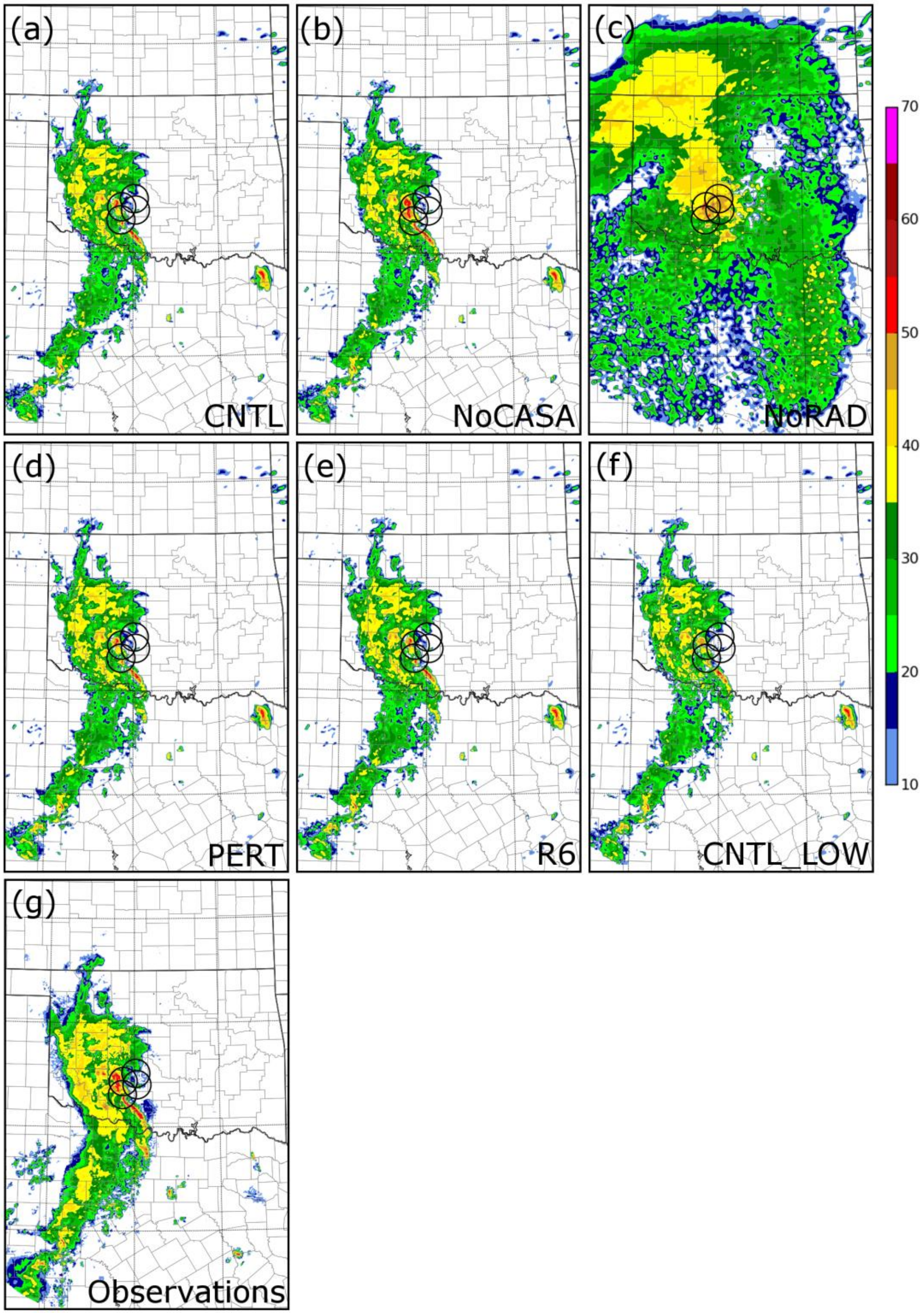

Fig. 5.3: Ensemble mean reflectivity at model level k = 10 (approximately 2 km above ground level) for the 0200 UTC analyses of the inner-grid experiments (panels (a)-(f)), compared to WSR-88D observations from 0200UTC interpolated to the ARPS grid (panel (g)). The experiment name is included in the lower right of each panel.

Despite the relatively good agreement in the location and structure of the convective system, there are some notable errors in the ensemble mean analyses of the experiments assimilating radar data. Compared to the observations (Fig. 5.3g), all of the radar-assimilating experiments (Fig. 5.3a, b, d-f) exhibit slightly weaker reflectivity values in their 0200 UTC ensemble mean analyses throughout much of the stratiform region and trailing convective line. The intensity of the leading convective line is captured well in NoCASA (Fig. 5.3b), but underestimated compared to observations (Fig. 5.3g) within the CASA domain in those experiments assimilating CASA radar data (Fig. 5.3a, d-f). The underestimation in these experiments is most likely the result of attenuated regions within the convective line in the CASA data. In an attempt to limit the impact of completely attenuated regions, CASA reflectivity data are not assimilated when the observed reflectivity is less than 20 dBZ, but some impact still remains.

Each of the ensemble mean analyses from the radar-assimilating experiments contains a convective cell to the east of the Dallas-Fort Worth metro area, just south of the Oklahoma border. This cell occurs near the edge of the coverage areas of the KFWS and KTLX radars; while this cell may be impacted by this data boundary, data from the neighboring WSR-88D radar site to the east at Shreveport (KSHV) indicates that a convective cell was present at 0200 UTC just to the east of the cell indicated in the ensemble mean analyses, just beyond the eastern extent of the KFWS radar observations. The westernmost fringe of this convective cell can be seen at the eastern extent of the KFWS coverage region in Fig. 5.3g, near the location of the convective cell in the ensemble analyses. Also, the analyses underestimate the intensity of two cells ahead of the convective line observed by KFWS, visible in the lower right of Fig. 5.3g. These cells developed late in the assimilation period and likely did not have sufficient time to spin up in the model.

In contrast with the experiments that assimilate radar data, the 0200 UTC ensemble mean analysis of the NoRAD experiment (Fig. 5.3c), which assimilated only conventional data, compares quite poorly to radar reflectivity observations. While most NoRAD ensemble members contained a convective system in the general region of the MCS observed by the WSR-88D radar network, there is a great deal of variation among individual ensemble members in the 0200 UTC analysis regarding the position, intensity, and structure of this convective line (not shown). The result is an ensemble mean analysis with large, smooth, relatively weak areas of reflectivity; most of the convective structure from individual members is lost when computing the ensemble mean (Fig. 5.3c). Because of the large spread among the ensemble members, the NoRAD ensemble analysis can benefit from probability matching when calculating the ensemble mean, as described earlier in section 4.3.1.

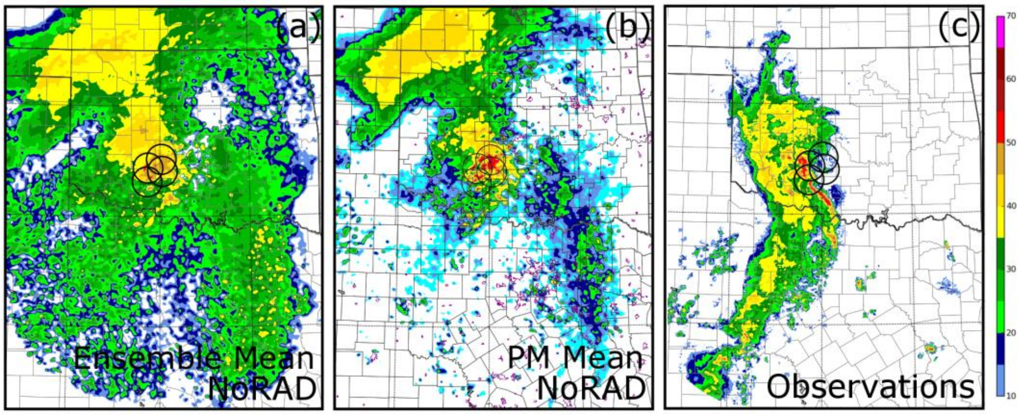

Fig. 5.4: Comparison of (a) ensemble mean radar reflectivity and (b) probability-matched mean radar reflectivity in the 0200 UTC ensemble analysis of NoRAD at model grid level k = 10 (approximately 2 km above the surface). Also shown are (c) WSR-88D radar observations, interpolated to the ARPS grid at model level k = 10.

When probability matching is performed, the overall coverage of rainfall in the probability-matched ensemble mean (Fig. 5.4b) is greatly reduced compared to the pure ensemble mean (Fig. 5.4a), though even then the structure of the convective system remains quite poor compared to the observations (Fig. 5.4c). By contrast, in the experiments that assimilate radar data the ensemble mean and probability-matched ensemble mean are quite similar, due to much closer agreement between ensemble members due to the constraint by radar observations (not shown).

Much of the heavy precipitation in the NoRAD ensemble is located to the north of the observed convective system (Fig. 5.4), while little to no precipitation is indicated in the NoRAD ensemble (Fig. 5.4a, b) in the region of the observed trailing convective line extending south into Texas (Fig. 5.4c). The failure of the NoRAD experiment during the analysis period to capture the structure of the observed MCS suggests that EnKF assimilation of conventional data alone is grossly insufficient to initialize the model for this case. The coverage and frequency of conventional data are greatly limited, while radar can sample regions throughout the storm at comparatively high spatial and temporal resolution. Because convective storms vary on the scale of several minutes and hundreds of meters, assimilation of radar data is vital for convective-scale NWP. This result agrees with the findings of Schenkman et al. (2011), whose similarly-configured experiments using 3DVAR to assimilate conventional and radar data also performed well with comparison to observations only when radar data was assimilated. A more detailed comparison of the results presented in this chapter to those of Schenkman et al. (2011) will be included below in section 5.4.

5.3 Ensemble Forecast Results

As in the single-grid experiments presented in Chapters 3 and 4, a three-hour ensemble forecast is performed from the 0200 UTC ensemble analyses of all six inner-grid experiments. Examination and verification of these ensemble forecasts will follow a similar methodology to that of Chapters 3 and 4. The primary focus will be on ensemble forecasts of radar reflectivity, a variable which serves as a proxy for precipitation and can be easily verified against observations from the WSR-88D radar network. Objective verification will be performed against reflectivity observed by the WSR-88D network (Fig. 5.5) using the ROC skill score (RSS) and Brier score, and both probabilistic and deterministic products will be considered, and prominent biases in the ensemble forecasts of radar reflectivity will be examined.

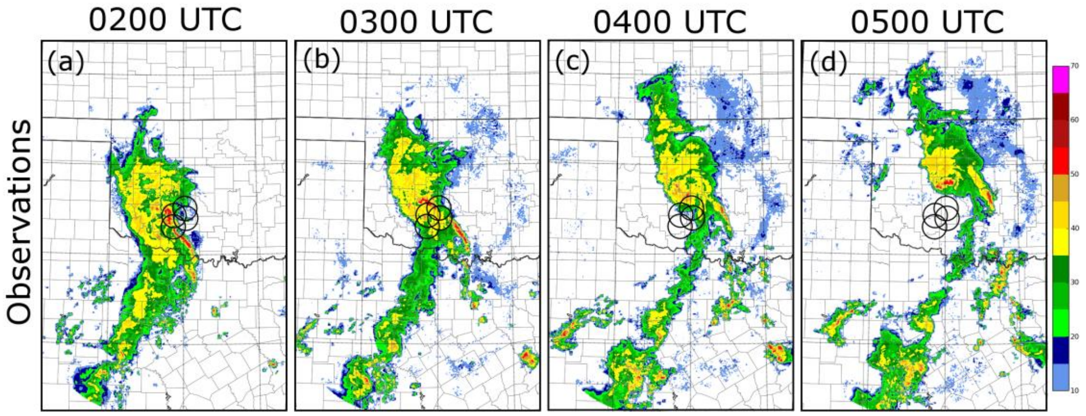

Fig. 5.5: Radar data observed by WSR-88D radar sites KAMA, KDYX, KFWS, KLBB, KTLX, and KVNX at (a) 0200UTC, (b) 0300 UTC, (c) 0400 UTC, and (d) 0500 UTC, interpolated to level k = 10 of the model grid used by the inner-grid experiments.

The inner-grid experiments (see Table 5.1) can be divided into three primary categories: those examining the impact of assimilating various radar data (NoCASA, NoRAD); those examining the impact of varying the model microphysics (PERT, R6); and finally those examining the impact of changing the assumed observation error for radial velocity and radar reflectivity (CNTL_LOW). The CNTL experiment serves as a default against which these three categories of experiments are to be compared. Each of these impacts will be considered in turn, beginning with the impact of assimilating radar data. Finally, we will consider ensemble forecasts of mesovortices, as we did for the single-grid experiments of Chapters 3 and 4, using the object-based ensemble forecasting technique described in equation 2.2.

While we will focus separately on the impacts of the several different variations in ensemble forecast design mentioned in the previous paragraph, it is desirable to include RSS and Brier Score statistics for all inner-grid experiments in a single chart to facilitate inter-comparison of all experiments together. For this reason, graphs of RSS and Brier Score are included below in Fig. 5.6 and Fig. 5.7 for all six inner-grid experiments; these figures will be referred to in each of sections 5.3.1-5.3.4. As in section 4.3.2, RSS and Brier Score are calculated both for the full model domain and for a sub-domain encompassing the CASA radar domain and areas directly to the east and north of the CASA domain (regions downwind of the CASA domain in this case where the greatest impact of assimilated CASA radar data would be expected). The methodology used to calculate RSS and Brier Score is identical to that used in section 4.3.2, and the verification sub-domain used here is identical in geographic extent to the sub-domain indicated in Fig. 4.2.

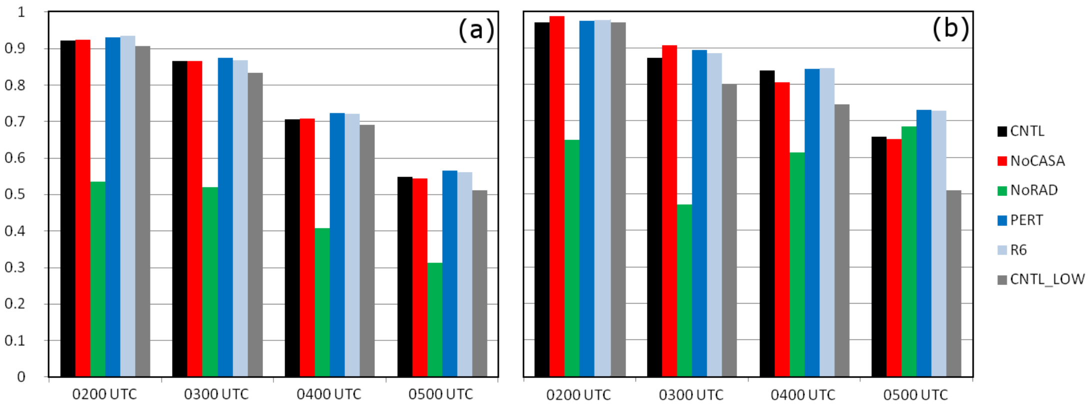

Fig. 5.6: ROC skill score for the ensemble forecast initial condition (at 0200 UTC) and for 1-, 2-, and 3-hour forecasts of radar reflectivity at the 25 dBZ threshold on vertical grid level k = 10 (slightly more than 2 km above mean sea level) for all six inner-grid ensemble forecast experiments. Shown are ROC skill scores calculated over (a) the entire horizontal model domain depicted in Fig. 5.1 and (b) the ETS verification sub-domain as depicted by the black outline in Fig. 4.2.

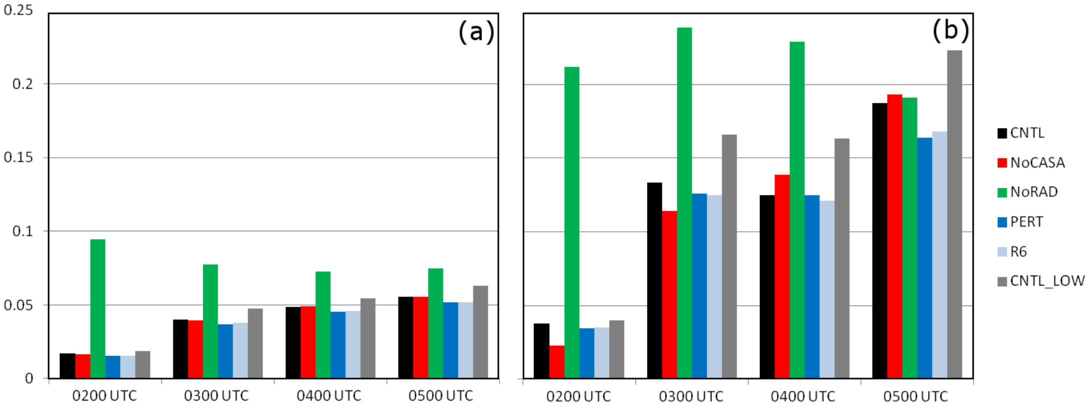

Fig. 5.7: As Fig. 5.6 but for Brier Score.

5.3.1 Impact of Radar Data Assimilation

During the analysis period, the ensemble forecast of NoRAD performed relatively poorly compared to the experiments assimilating radar data. Without assimilating radar data, the structure of the MCS in the 0200 UTC ensemble analysis agreed poorly with observations, and varied considerably among ensemble members (Fig. 5.3, Fig. 5.4). This trend continues throughout the 3-hour forecast period from 0200 to 0500 UTC in both deterministic forecasts from the NoRAD ensemble.

Probability-matched (PM) ensemble mean reflectivity is shown below in Fig. 5.8 for CNTL (Fig. 5.8a-c) and NoRAD (Fig. 5.8d-f) at 0300, 0400, and 0500 UTC. While in the PM mean of NoRAD the heaviest precipitation associated with MCS is poorly defined and displaced to the north (Fig. 5.8d-f), the CNTL ensemble produces a PM mean with a well-defined MCS at 0400 and 0500 UTC (Fig. 5.8b, c) which compares relatively well in location and extent with the observed MCS at those times (Fig. 5.5c, d). At 0300 UTC the PM mean of CNTL exhibits a relatively discrete, cellular structure throughout much of the MCS (Fig. 5.8a) which does not match the more consolidated MCS observed by the WSR-88D network at that time (Fig. 5.5b). The problem of overly cellular organization early during the forecast period is common to many of the experiments assimilating radar data.

The better performance of the ensemble when radar data are assimilated is also evident in probabilistic forecasts produced for CNTL and NoRAD using the neighborhood ensemble probability (NEP) method described in equation (2.1) using a neighborhood radius of 5 km. Regions predicted in the CNTL ensemble to have high probability of Z > 25 dBZ (Fig. 5.9a-c) show much greater agreement with the observed regions of *Z* > 25 dBZ than those of NoRAD (Fig. 5.9d-f). Using a 40 dBZ threshold for NEP calculations, corresponding to heavier convective precipitation, it becomes evident that the CNTL ensemble predicts an overabundance of heavy precipitation regions with *Z* > 40 dBZ (Fig. 5.10a-c), though these regions, at least over the northern part of the convective system, are not too far displaced from areas of observed *Z* > 40 dBZ. The NoRAD ensemble (Fig. 5.10d-f) predicts a more reasonable coverage of *Z* > 40 dBZ over central Oklahoma, but also erroneously predicts significant probability of *Z* > 40 dBZ over southern Kansas.

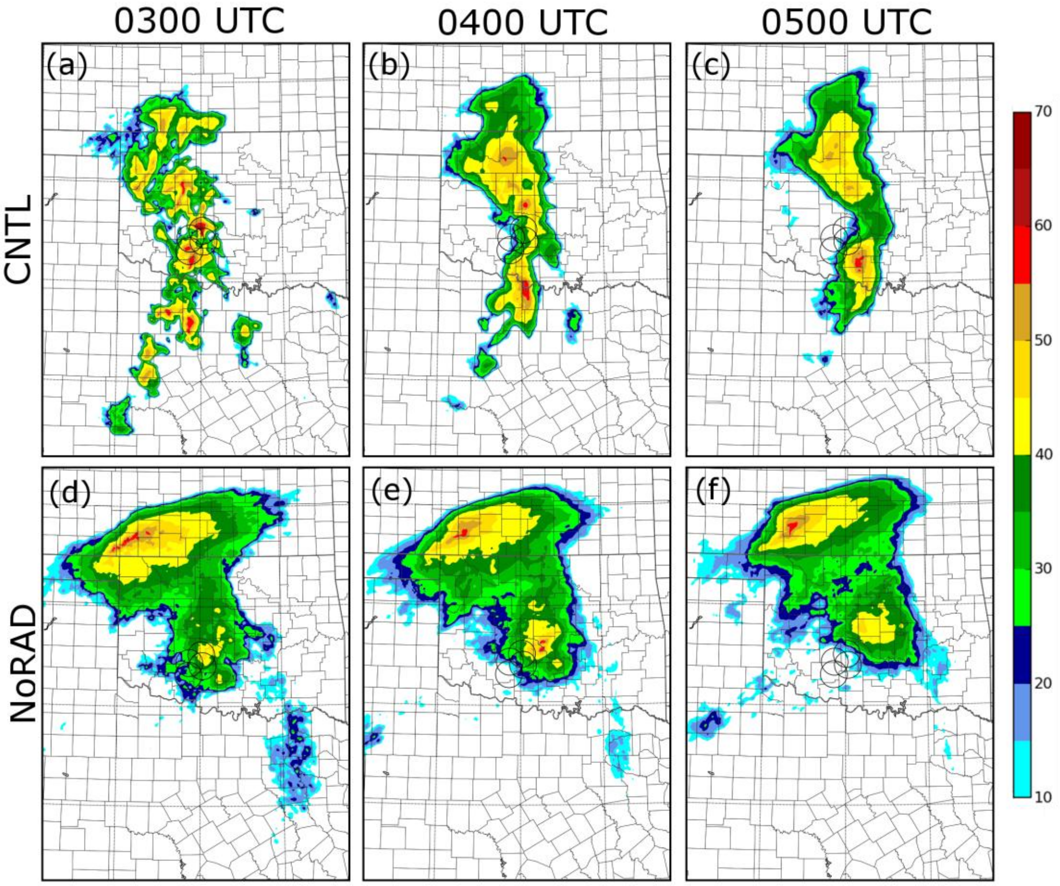

Fig. 5.8: Probability-matched ensemble mean radar reflectivity on model grid level k = 10 (approximately 2 km above the surface) for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and for NoRAD at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC.

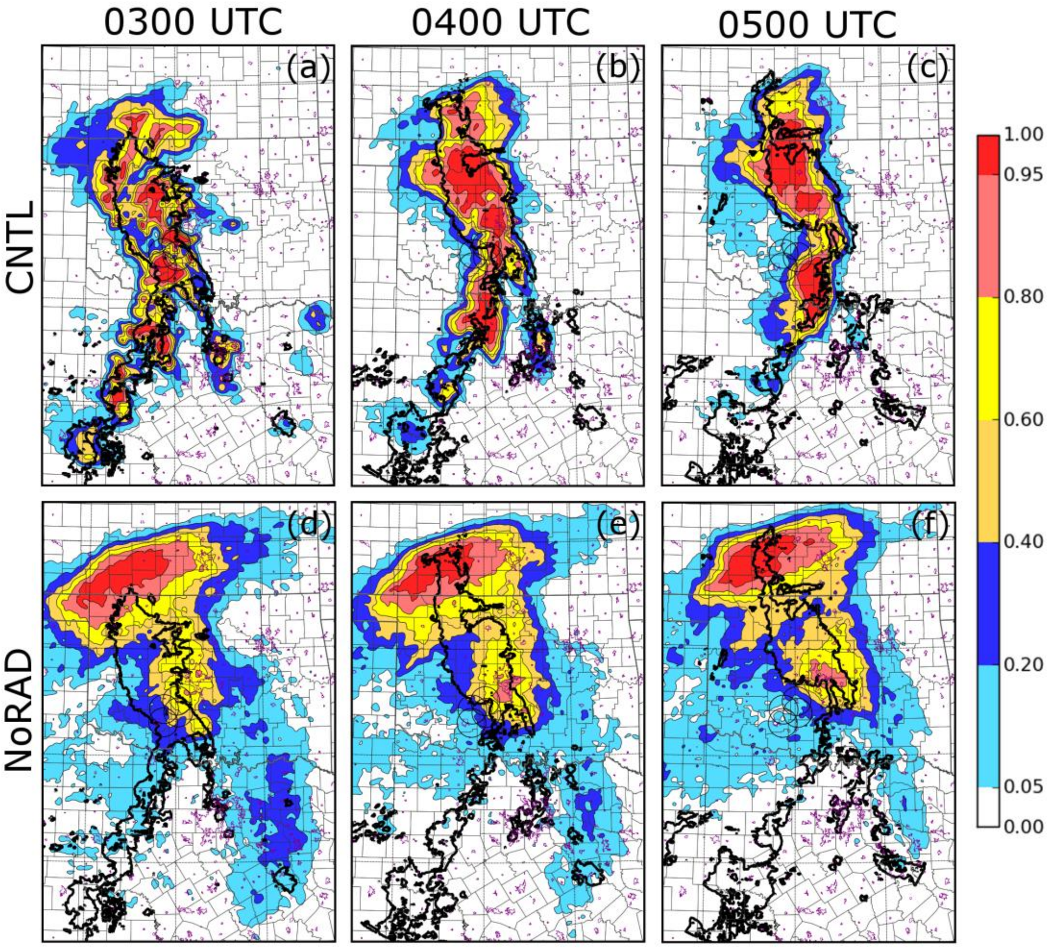

Fig. 5.9: Neighborhood ensemble probabilities (color-shading) of reflectivity exceeding 25 dBZ, $P[Z > 25 \text{ dBZ}]$*, at model grid level 10 (approximately 2 km above the surface) for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC, and for NoRAD at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC. The 25 dBZ radar reflectivity contour observed by the WSR-88D radar network at the corresponding time is indicated by the bold black contours. Urban areas are shown with thin, purple contours.*

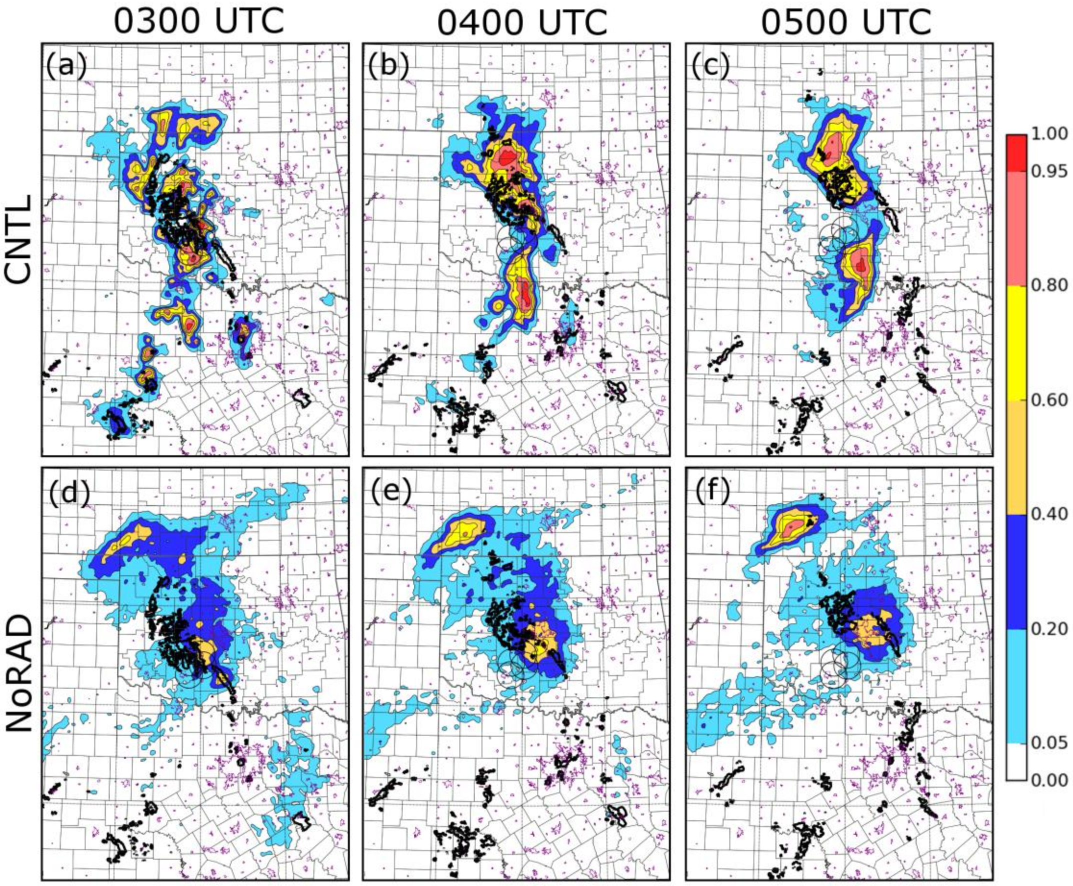

Fig. 5.10: As Fig. 5.9 but for neighborhood ensemble probability calculated using a threshold of 40 dBZ and the observed 40 dBZ contour.

Looking at individual ensemble member forecasts, it can be seen that variation in the position of the MCS is largely responsible for the smooth, widespread, relatively weak precipitation in the NoRAD PM mean. Radar reflectivity is shown below in Fig. 5.11 for three typical NoRAD ensemble members: #3 (Fig. 5.11a-c), #16 (Fig. 5.11d-f), and #32 (Fig. 5.11g-i) at 0300, 0400, and 0500 UTC. These individual NoRAD ensemble members each feature a relatively discrete MCS compared to the PM mean (Fig. 5.8d-f) and agree upon a large region of stratiform rain in northwestern Oklahoma and southwestern Kansas, but disagree among one another in the position and structure of the MCS over central Oklahoma. When these relatively divergent members are averaged to obtain the ensemble mean (and from there the PM ensemble mean), the position differences in the MCS result in the spatial smoothing observed in Fig. 5.8.

While the assimilation of conventional data alone in NoRAD is sufficient to produce organized multicellular convection in all ensemble members (see Fig. 5.11), the predictions of the individual members compare poorly with WSR-88D observations between 0300 and 0500 UTC (Fig. 5.5b-d). Convection extends into southwestern Oklahoma in many of the NoRAD members between 0300 and 0500 UTC, in regions where no convection was observed (see Fig. 5.5), and the stratiform rain region in most members is shifted too far to the north. Some NoRAD members, such as member #32, predict the presence of an organized convective line (Fig. 5.11g, h), but the convective lines predicted vary widely in orientation and do not extend nearly as far south as those seen in observations. Very few of the NoRAD members featured a well-defined line-end vortex.

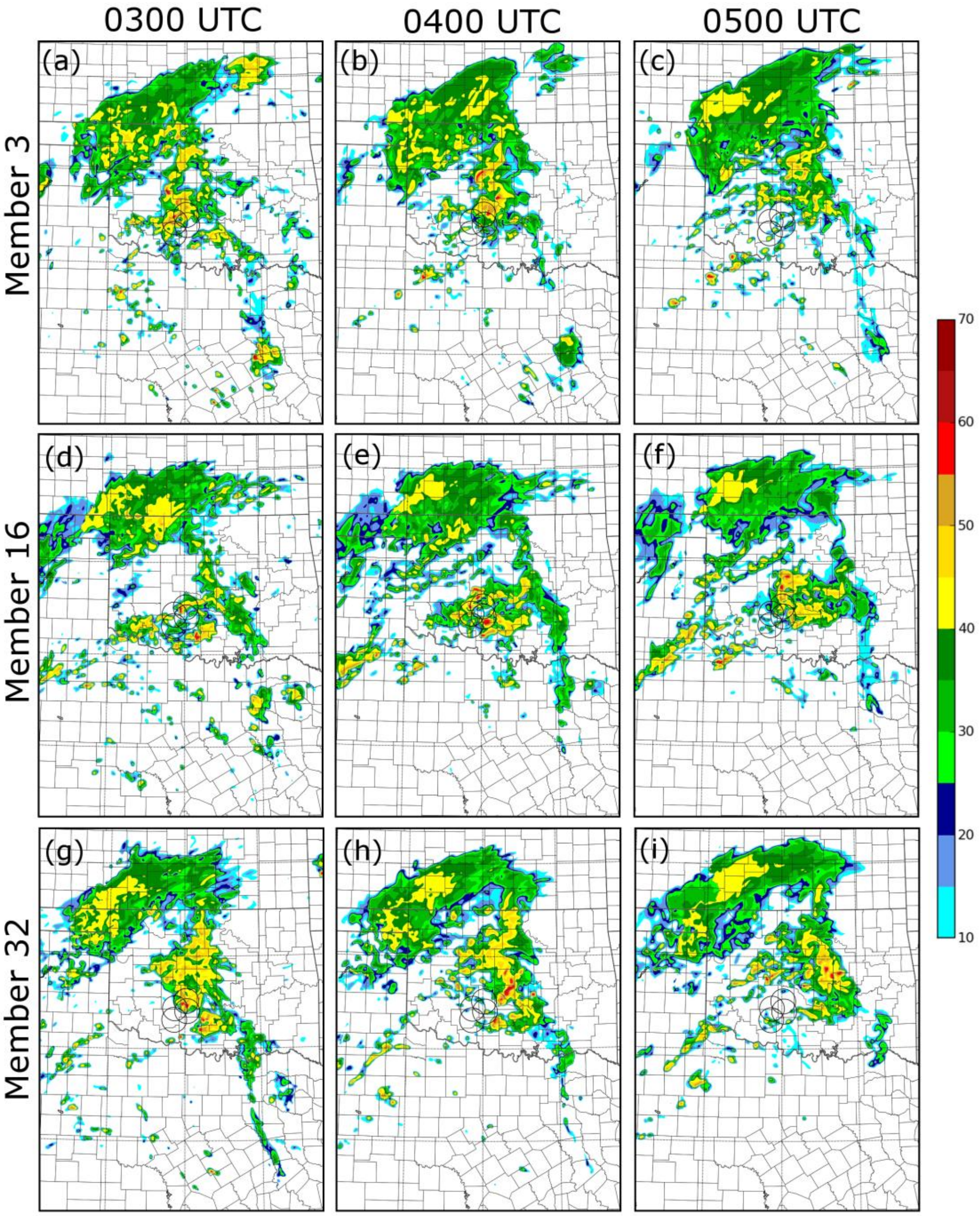

Fig. 5.11: Radar reflectivity at model grid level k = 10 (approximately 2 km above the surface) for three NoRAD ensemble members at 0300, 0400, and 0500 UTC. Shown are (a-c) NoRAD member #3, (d-f) NoRAD member #16, and (g-i) NoRAD member #32.

In CNTL, where radar data is assimilated in addition to conventional data, there is much greater agreement between different forecast ensemble members, as well as between forecast ensemble members and observations. Radar reflectivity forecasts from three typical CNTL ensemble members are shown below in Fig. 5.12 between 0300 and 0500 UTC. Though the members vary moderately in the extent and intensity of the trailing line, particularly at 0500 UTC (Fig. 5.12c, f, i), all members feature an organized MCS whose position agrees relatively well with that observed by the WSR-88D network (Fig. 5.5). Unlike in NoRAD, most CNTL members feature a line-end vortex located at the northern end of the trailing convective line, similar to the observed MCS. Prediction of the line-end vortex will be considered in greater detail below in section 5.3.6.

The relatively poor quality of the NoRAD initial condition is reflected in objective forecast verifications using the RSS (Fig. 5.6) and Brier score (Fig. 5.7). When the entire forecast domain is considered, the initial NoRAD ensemble (valid at 0200 UTC) generates a RSS of just 0.53, compared to 0.92 for CNTL (Fig. 5.6a). Both ensembles perform slightly better on the verification sub-domain (Fig. 5.6b), though the RSS of 0.65 produced by NoRAD still compares poorly to the 0.97 produced by CNTL. The difference is even more pronounced in the 0200 UTC Brier score, where CNTL has a lower (better) value by a factor of 5 on both the full domain and the verification sub-domain (Fig. 5.7). The reason for the greater difference in Brier score between the two experiments is due to the extensive over-prediction of light precipitation in the 0200 UTC NoRAD ensemble. The Brier score rewards correct prediction of clear-air regions, a metric which heavily penalizes the 0200 UTC NoRAD ensemble.

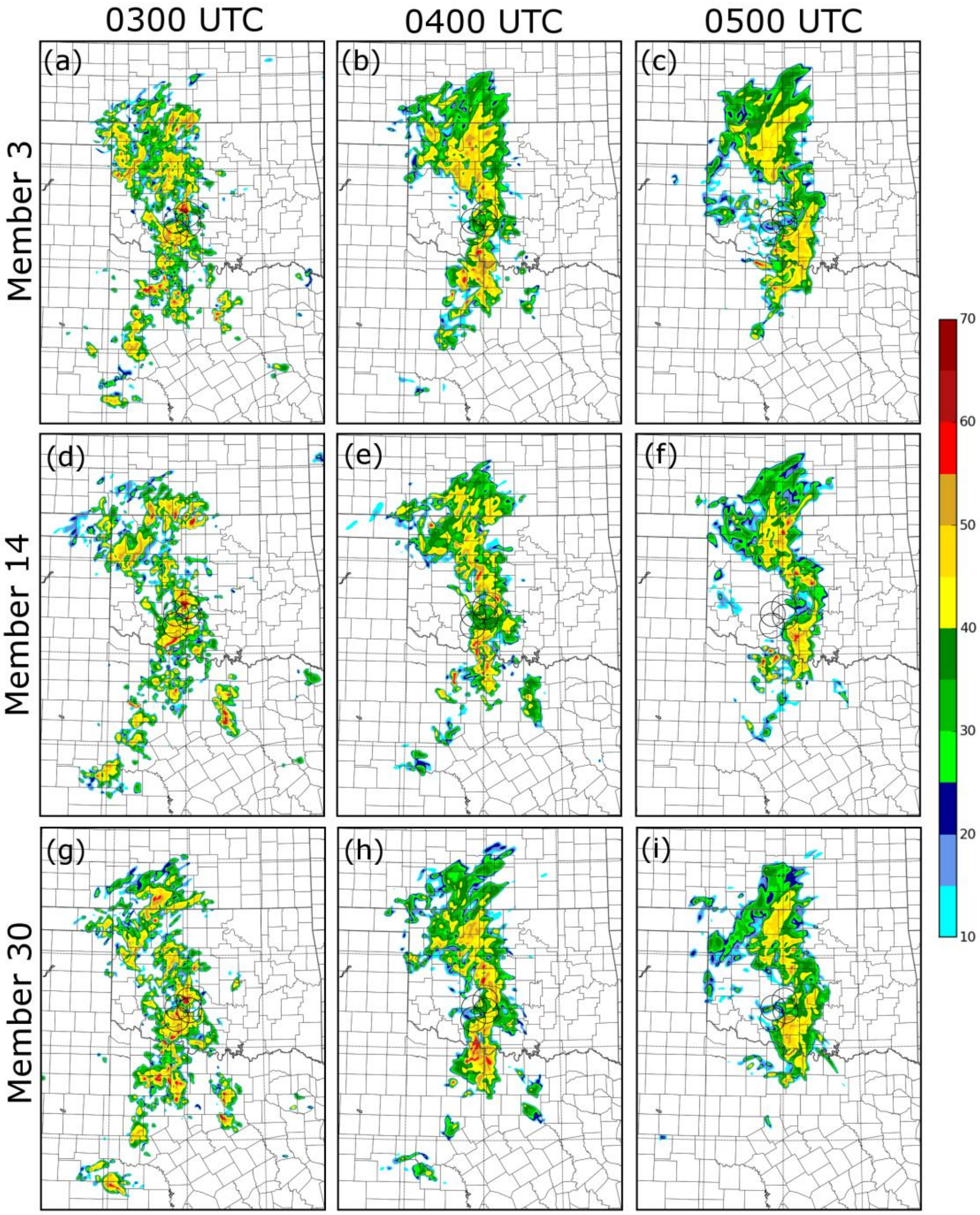

Fig. 5.12: As Fig. 5.11, but for three typical ensemble members from CNTL. Shown are (a-c) CNTL member #3, (d-f) CNTL member #14, and (g-i) CNTL member #30.

During the forecast period, CNTL consistently outperforms NoRAD in both RSS and Brier score calculated on the full domain (Fig. 5.6a, Fig. 5.7a) by a considerable margin, though this margin decreases with increasing forecast time, especially for the Brier score. On the verification sub-domain, CNTL substantially outperforms NoRAD at 0300 and 0400 UTC (Fig. 5.6b, Fig. 5.7b), but by 0500 UTC NoRAD actually slightly outperforms CNTL in both RSS and Brier score. While the quality of the CNTL forecast ensemble declines throughout the forecast period as measured by both skill scores, the skill scores of NoRAD remain relatively steady throughout the forecast period, dropping only slightly below their initial 0200 UTC values by 0500 UTC. In CNTL, the relatively rapid decline in both skill scores on the sub-domain at 0500 UTC is largely due to over-prediction of the intensity of the trailing convective line (see Fig. 5.12c, f, i), a tendency which was not shared by NoRAD (see Fig. 5.11c, f, i). The over-prediction of heavy precipitation in the trailing convective line in CNTL is also clearly visible in NEP forecasts of $Z > 40$ dBZ, where CNTL predicts a wide swath of P[$Z > 40$ dBZ] exceeding 0.8 to the east and southeast of the CASA domain where no precipitation exceeding 40 dBZ was observed by the WSR-88D radar network (Fig. 5.10c).

While NoRAD compares poorly to CNTL overall, this does not imply that the NoRAD ensemble forecast was entirely unskilled. Throughout the forecast period, on the verification sub-domain, NoRAD RSS values range from 0.47 to 0.68 (Fig. 5.6b)—even at the minimum of this range, the NoRAD forecast ensemble exceeds the RSS of 0.4 considered to be the minimum for a useful forecast (Buizza 1997; Kong et al. 2011). Even on the full domain, the NoRAD forecast remains above the

0.40 RSS threshold until after 0400 UTC. The fact that the NoRAD forecast retains useful skill throughout most, if not all, of the forecast period indicates that structures and forcing at the mesoscale (and at larger scales) having control on the precipitation features were present in this case which could be partially captured by the assimilation of conventional observations.

5.3.2 Impact of CASA Data

The primary objective of experiment NoCASA is to evaluate the impact of including CASA data. NoCASA differed from CNTL only in the exclusion of CASA radar during data assimilation on the inner grid between 0105 and 0200 UTC. In this section we will examine the impact on the reflectivity forecast. Impacts on ensemble forecasts of mesovortices will be considered later in section 5.3.6.

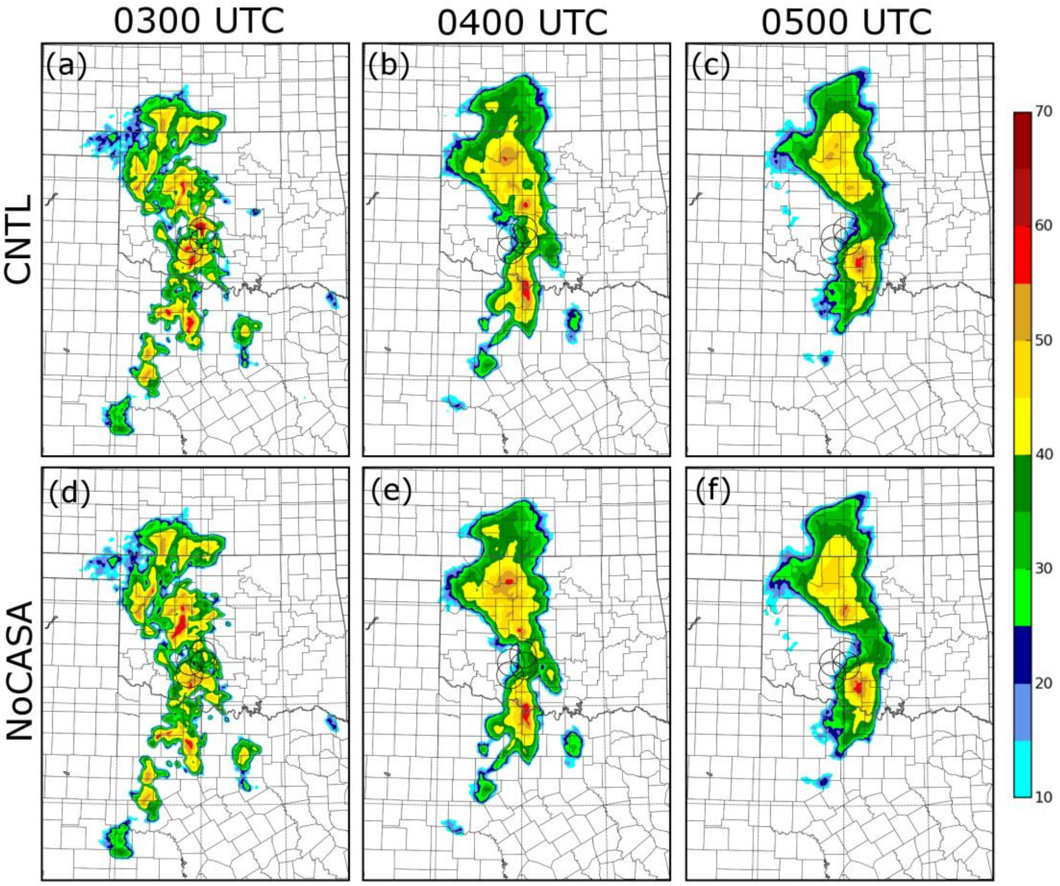

Fig. 5.13: As Fig. 5.8 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and NoCASA at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC.

On a system-wide scale, there is relatively little difference in the overall structure and evolution of the MCS predicted by the ensemble mean of NoCASA as compared to that predicted by the ensemble mean of CNTL. The PM ensemble means of radar reflectivity in CNTL and NoCASA at 0300, 0400, and 0500 UTC (1, 2, and 3 hours of forecast time, respectively) are shown in Fig. 5.13, and the similarity of the two ensembles is apparent. Both ensembles predict a MCS over central Oklahoma, in good agreement with observations (Fig. 5.5), with a trailing line extending south and south-southwestward into northern Texas.

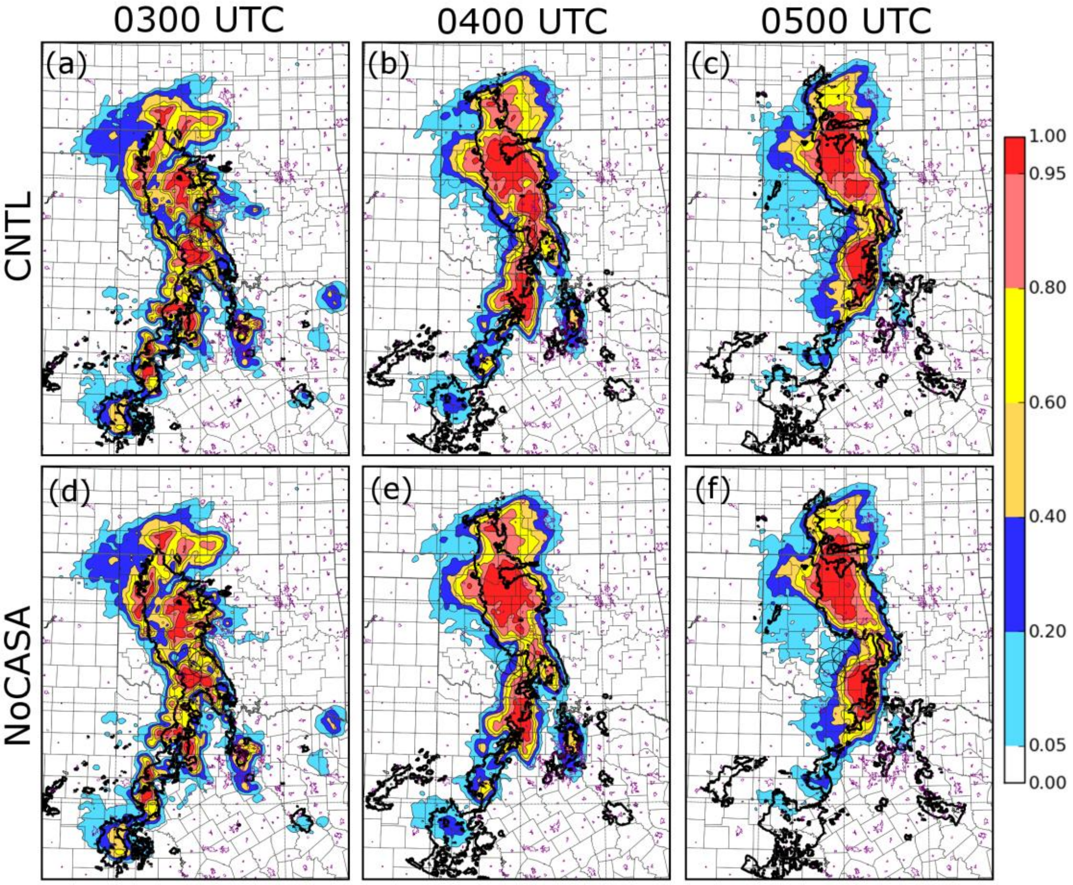

Fig. 5.14: As Fig. 5.9 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and NoCASA at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC.

Many members of both the CNTL and NoCASA ensembles also predict some convective activity associated with the leading convective line located in the observations to the east and southeast of the CASA radar network (Fig. 5.5b, c) at 0300 and 0400 UTC, as indicated by the areas of moderately high NEP of Z > 25 dBZ in these regions in both CNTL and NoCASA (Fig. 5.14). NoCASA also shares the tendency of the CNTL ensemble to predict a more cellular organization than seen in observations (Fig. 5.5) at 0300 UTC (Fig. 5.13a, d), and to over-predict the intensity of the trailing convective line, particularly at later times, as is evident in the NEP of $Z > 40$ dBZ in both ensembles south and southeast of the CASA radar network (Fig. 5.15).

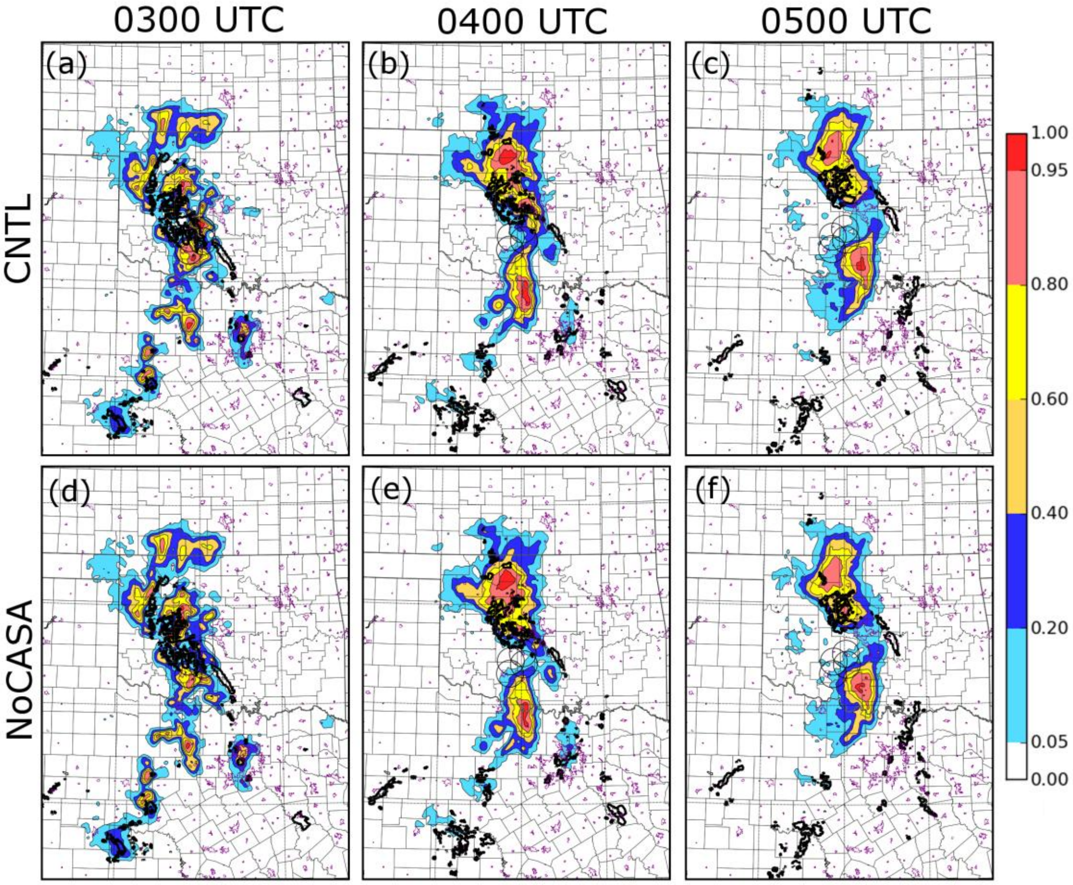

Fig. 5.15: As Fig. 5.10 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and NoCASA at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC.

When the ensemble forecasts are verified using the RSS and Brier score, almost no difference can be seen between the two ensembles for scores calculated on the full domain (Fig. 5.6a, Fig. 5.7a). The two forecasts also perform similarly on the verification sub-domain (Fig. 5.6b, Fig. 5.7b), though NoCASA slightly outperforms CNTL in both RSS and Brier score in the initial condition at 0200 UTC and in the 1-hour forecast at 0300 UTC; this trend is reversed in the 2- and 3-hour forecasts at 0400 and 0500 UTC, at which time CNTL slightly outperforms NoCASA in both RSS and Brier score. The initially slightly better performance of NoCASA on the verification sub-domain can likely be attributed to the impact of attenuation in observations from CASA radars, as was discussed in section 5.2. As indicated by the better performance of CNTL at 0400 and 0500 UTC, this impact appears to be temporary, and outweighed at later times by the positive impact of assimilating additional near-surface radar data from the CASA network.

The impact of assimilating CASA data becomes more apparent when we consider the spread of the ensemble forecast of radar reflectivity and focus on the region near and downwind of the CASA radar network. Ensemble spread of radar reflectivity at 0300, 0400, and 0500 UTC is shown for the CNTL and NoCASA ensembles in Fig. 5.16. In both ensembles, higher spread occurs near the fringes of the predicted MCS. Because of differences in the position and movement speed of the MCS between individual ensemble members (see Fig. 5.12), points near the fringe of the MCS contained moderate to heavy reflectivity in some ensemble members, but clear air in others, leading to large RMS spread values. By contrast, in the core of the MCS and the trailing line, all ensemble members indicate that moderate to heavy precipitation is present (see Fig. 5.12), thus the relative difference in reflectivity between ensemble members, and hence the RMS ensemble spread, is lower.

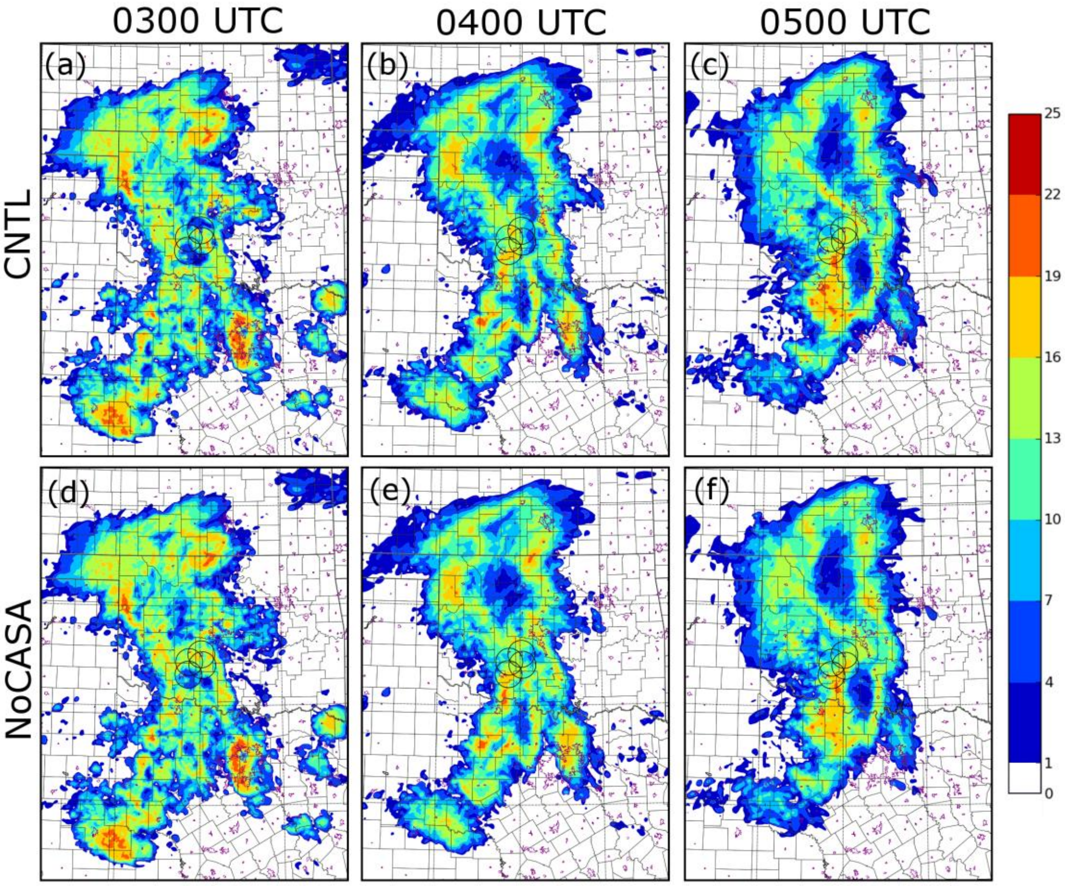

Fig. 5.16: RMS Ensemble spread of radar reflectivity on model grid level k = 10 (approximately 2 km above the surface) for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and for NoCASA at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC. 30 km CASA range rings are shown in black, and urban boundaries are indicated in purple.

When the difference between the RMS ensemble spread of *Z* in CNTL and that in NoCASA is calculated for 1-, 2, and 3-hour forecasts at 0300, 0400, and 0500 UTC (Fig. 5.17), the largest differences in RMS ensemble spread occur near and downwind of the CASA radar network. In Fig. 5.17, positive values indicate regions in which CNTL has a greater RMS spread of *Z*, while negative values indicate regions where NoCASA has a greater RMS ensemble spread of *Z*. At 0300 UTC (Fig. 5.17a), a there is a considerable region in the northeastern portion of the

CASA radar network and just to its north where NoCASA has a much higher ensemble spread of *Z* than CNTL, with difference in RMS ensemble spread of more than 10 dBZ in a few regions. This region of higher ensemble spread in NoCASA moves to the north-northeast over time, and is located just to the east and northeast of the CASA radar network by 0400 UTC (Fig. 5.17b). By 0500 UTC the region of higher ensemble spread in NoCASA has decreased in magnitude, with peak differences of only slightly less than 10 dBZ (compared to more than 13 dBZ at 0300 UTC), but the region is still well-defined to the northeast of the CASA domain (Fig. 5.17c).

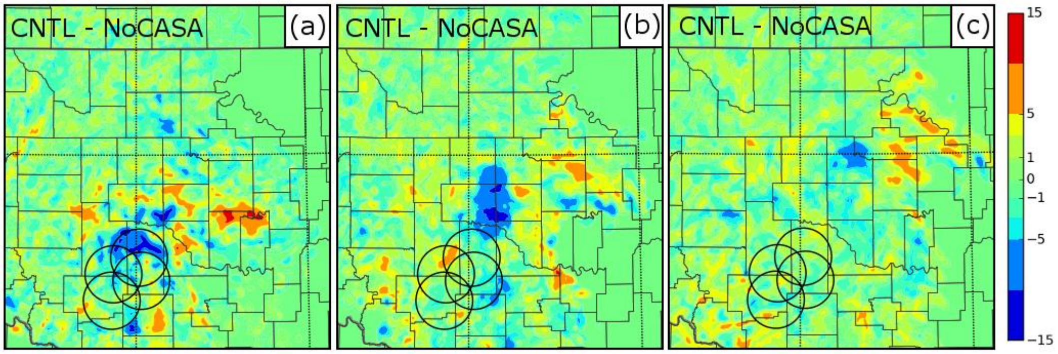

Fig. 5.17: Difference in RMS Ensemble spread of radar reflectivity on model grid level k = 10 (approximately 2 km above the surface) for (CNTL – NoCASA) at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC in a region near and downwind of the CASA radar network. 30 km CASA range rings are shown in black.

This region of higher RMS spread in NoCASA than in CNTL represents a plume of reduced ensemble spread in CNTL extending downwind from the CASA radar network with time. The reason for this plume is the impact of the CASA radar data assimilated in CNTL (which was not assimilated in NoCASA); the wealth of densely-spaced CASA data assimilated reduced the ensemble spread of *Z* in the region. This reduced spread was advected to the north-northeast by the prevailing south-southwesterly wind in the region, moving in roughly the same direction and at the same speed as the observed MCS (see Fig. 5.5). The persistence of this plume at 0400 and even 0500 UTC suggests that, at least on a regional scale, the impact of CASA radar data assimilated between 0105 and 0200 UTC remains quite evident in the forecast two hours later at 0400 UTC, and has still not fully disappeared by the end of the forecast period at 0500 UTC. The regional impact of this data at 0400 and will be further considered in section 5.3.6, where ensemble forecasts of the tornadic mesovortex occurring near the southern extent of the plume at 0400 UTC will be examined. In the field of targeted observation research, expected reduction in ensemble forecast variance is often used as a measure of the impact of added observations (Majumdar et al. 2001).

5.3.3 Impact of Microphysical Variations

Two inner-grid experiments were performed to investigate the impact of varying the model microphysics of the forecast ensemble. In R6, the rain intercept parameter of the Lin et al. (1983) microphysics scheme used in all ensemble members is increased from $8.0 \cdot 10^5$ to $4.0 \cdot 10^6$. While Snook and Xue (2008) found that a rain intercept parameter of $8.0 \cdot 10^5$ produced the best results for a supercell simulation, it is possible that a larger value of the rain intercept parameter may be optimal for a squall-line simulation such as the 8-9 May 2007 MCS case studied here. In PERT, the rain intercept parameter is varied within the ensemble in an effort to produce the same type of benefit observed when using mixed-microphysics in the single-grid experiments presented in Chapters 3 and 4 without the drawback of a multi-modal distribution resulting from the use of different microphysical schemes within the ensemble (see Fig. 4.19).

Unlike the impact of CASA data, which was mostly localized in a region near and downwind of the CASA domain, the impact of changes to the microphysical parameterization appears on a system-wide scale. While subtle differences can be observed throughout the MCS, the greatest differences between CNTL and R6 in the PM mean reflectivity field (Fig. 5.18) can be seen in the trailing convective line extending southward into northern Texas, especially at 0400 and 0500 UTC. Compared to the trailing line in the PM mean of CNTL (Fig. 5.18b, c), the trailing line in R6 (Fig. 5.18e, f) is not as intense, particularly at 0400 UTC. The less intense line agrees better with the observed intensity of precipitation at this time (Fig. 5.5c). The trailing line in the PM mean of PERT is also weaker than that in the PM mean of CNTL, but the difference is even more pronounced (Fig. 5.18h, i).

CNTL, R6, and PERT all indicate the presence of convective activity in the leading convective line located just to the east and southeast of the CASA radar network at 0300 and 0400 UTC (Fig. 5.5b, c; Fig. 5.18). When the NEP of $Z > 25$ dBZ is calculated, all three experiments predict a moderate probability of $Z > 25$ dBZ in the vicinity of the observed leading convective line at 0300 and 0400 UTC (Fig. 5.19a-b, d-e, g-h). Both PERT and R6 predict higher probabilities than CNTL in the region of the leading convective line. At the 40 dBZ threshold (Fig. 5.20), all three experiments predict significant probabilities of $Z > 40$ dBZ in the trailing convective line where no precipitation of that intensity was observed, but the greatest over-prediction is in CNTL, followed by R6; the over-prediction in PERT is the least.

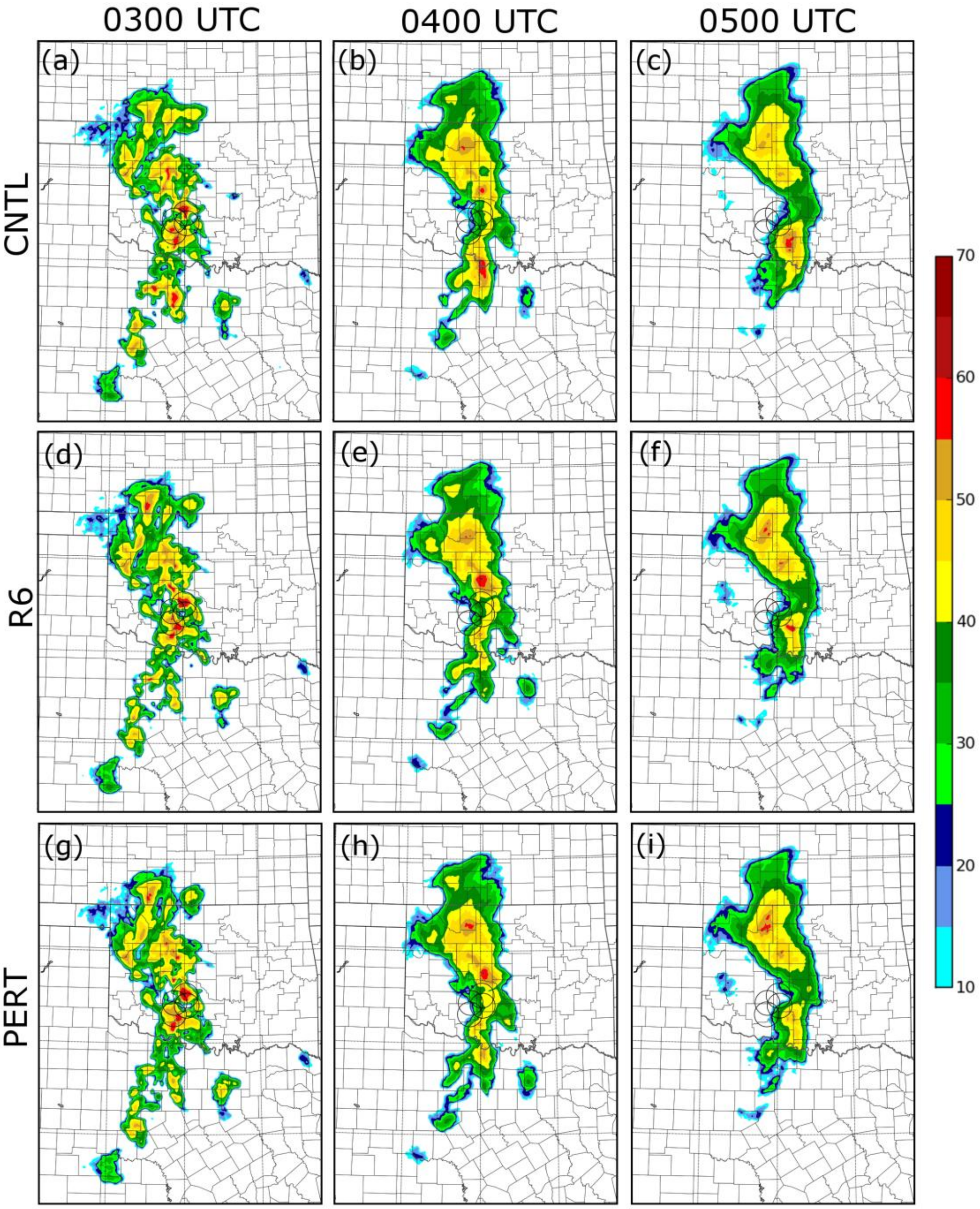

Fig. 5.18: As Fig. 5.8 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; R6 at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC; and PERT at (g) 0300 UTC, (h) 0400 UTC, and (i) 0500 UTC.

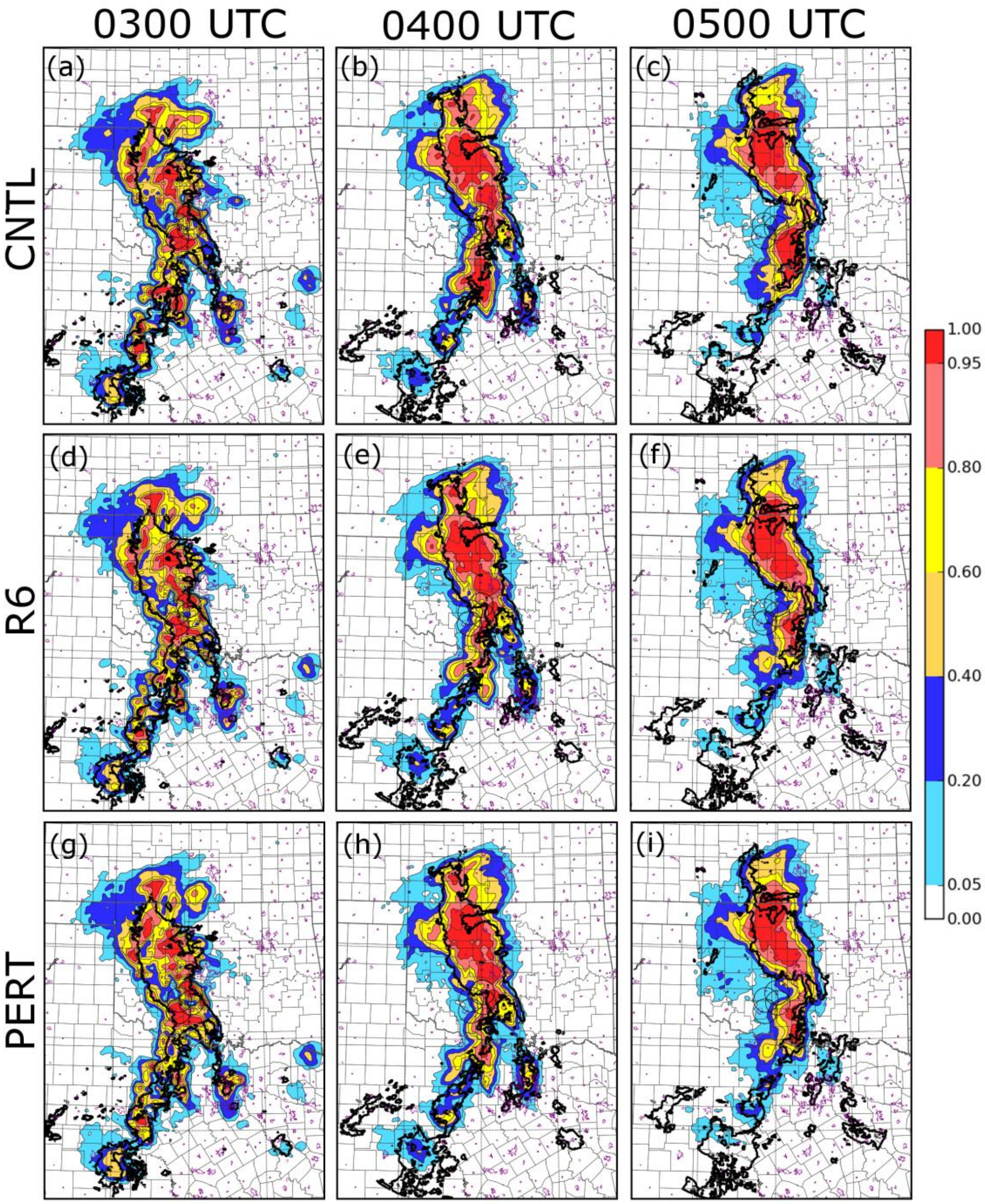

Fig. 5.19: As Fig. 5.9, but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; R6 at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC; and PERT at (g) 0300 UTC, (h) 0400 UTC, and (i) 0500 UTC.

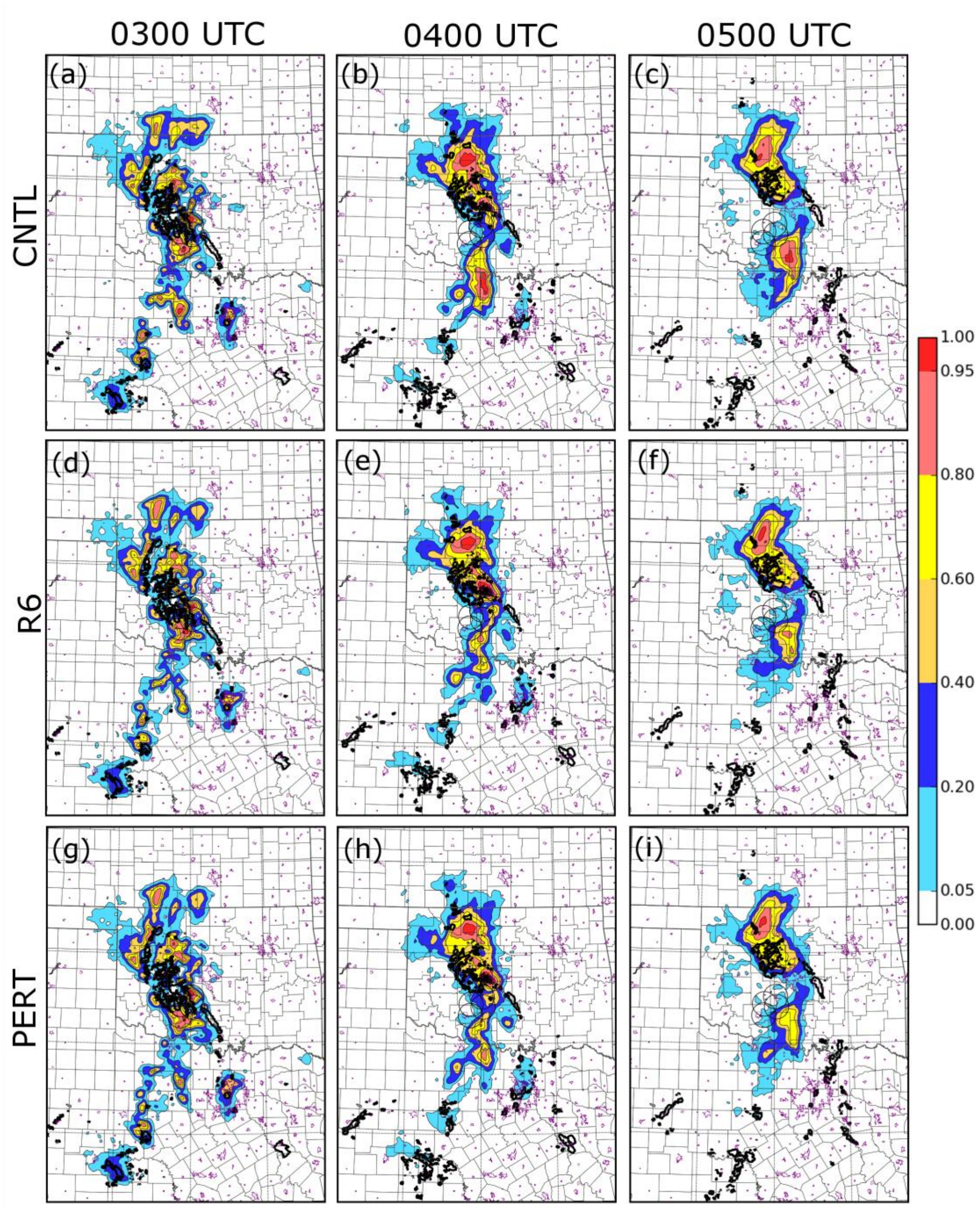

Fig. 5.20: As Fig. 5.10 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; R6 at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC; and PERT at (g) 0300 UTC, (h) 0400 UTC, and (i) 0500 UTC.

The better agreement with observations seen in the PM mean radar reflectivity and NEP reflectivity forecasts for PERT and R6 compared to CNTL is reflected in the ROC skill score and Brier score statistics for the three experiments. On the full domain, both R6 and PERT slightly outperform CNTL at all hours between 0200 and 0500 UTC in both ROC skill score (Fig. 5.6a) and Brier score (Fig. 5.7a). On the verification sub-domain, PERT outperforms CNTL in both RSS and Brier score at all times (Fig. 5.6b; Fig. 5.7b), with the greatest improvement being at 0500 UTC. Similarly R6 outperforms CNTL at all times on the verification sub-domain in RSS (Fig. 5.6b), and outperforms CNTL in Brier score (Fig. 5.7b) at all times except 0400 UTC, at which time the two perform equally well.

In terms of the skill score metrics considered in Fig. 5.6 and Fig. 5.7, R6 and PERT perform similarly for most of the forecast period, both on the full domain and on the verification sub-domain. The relative similarity in the skill scores of R6 and PERT suggests that much of the improvement of these experiments over CNTL is likely derived from the increased rain intercept parameter they employ (in PERT, the mean rain intercept parameter is $4.4 \cdot 10^6$, which is very close to that of R6), rather than the variation in rain intercept parameter of PERT. It is also possible that forecast improvement may be imparted by variation of the rain intercept parameter, but that this improvement is not captured by bulk skill statistics such as the RSS and Brier score.

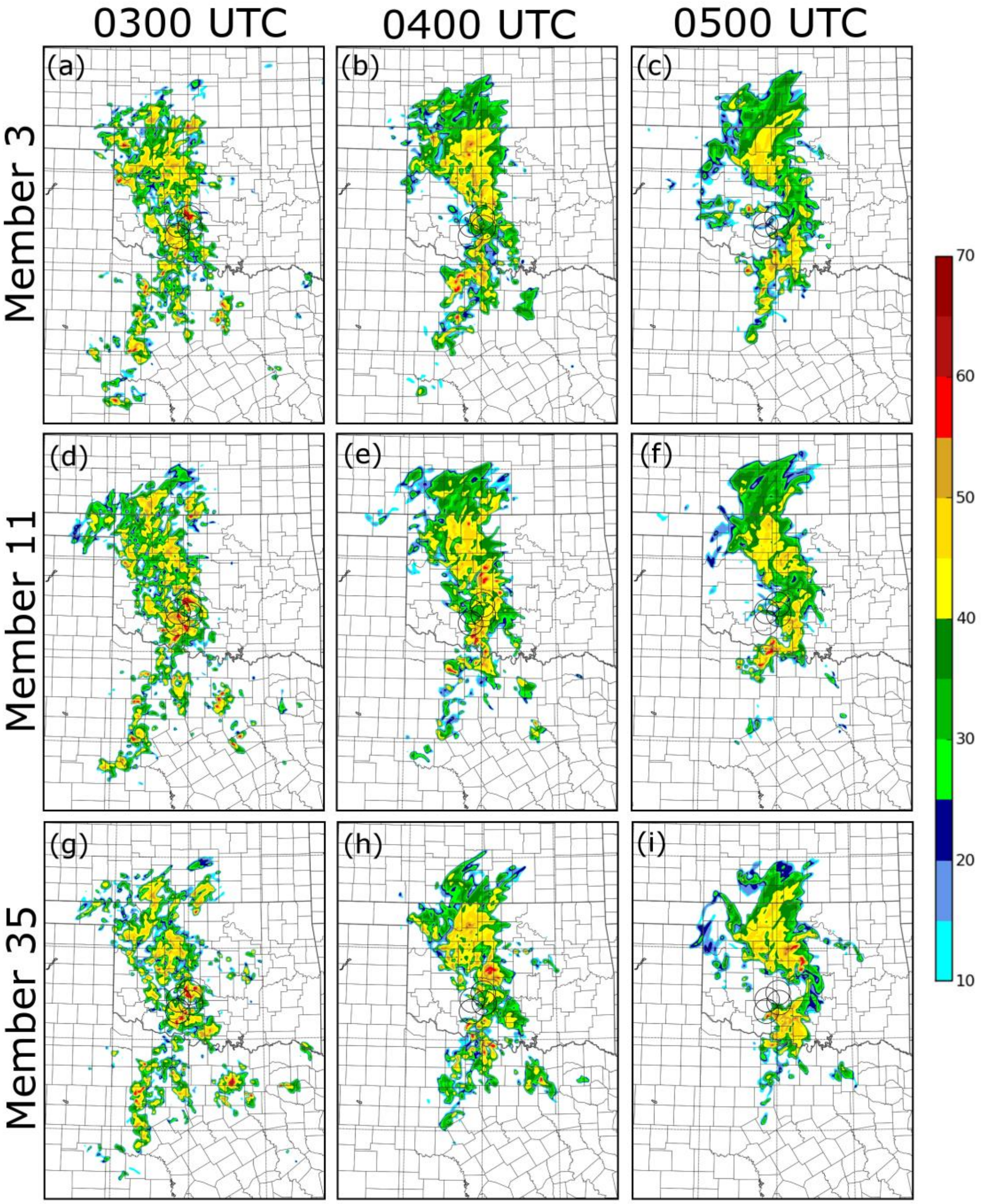

Fig. 5.21: As Fig. 5.11, but for three typical ensemble members from PERT. Shown are (a-c) PERT member #3, (d-f) PERT member #11, and (g-i) PERT member #35.

Though there is little difference in skill scores between R6 and PERT, variation of the rain intercept parameter in PERT does appear to noticeably impact the behavior of the ensemble members. The weaker trailing line in PERT as compared to CNTL can be traced to greater variation among PERT members in the position, orientation, and intensity of the trailing line, particularly at 0400 and 0500 UTC. These differences can be seen in Fig. 5.21, which shows predicted reflectivity at vertical model level k = 10 (approximately 2 km above the surface) for three members of the PERT ensemble. All three of these members predict a relatively robust trailing convective line that remains present through 0500 UTC, but by 0500 UTC the line has evolved quite differently between the three. In PERT member #3 (Fig. 5.21c), a thin, moderately intense line is predicted. In PERT member #11 (Fig. 5.21f), the trailing line does not extend as far south, contains slightly more intense convection, and exhibits a bowing pattern. PERT member #35 (Fig. 5.21i) predicts a thicker intense line in southern Oklahoma, with only weak precipitation extending south into Texas. While differences exist between the three ensemble members in treatment of other portions of the MCS, the differences in the trailing line are the most pronounced.

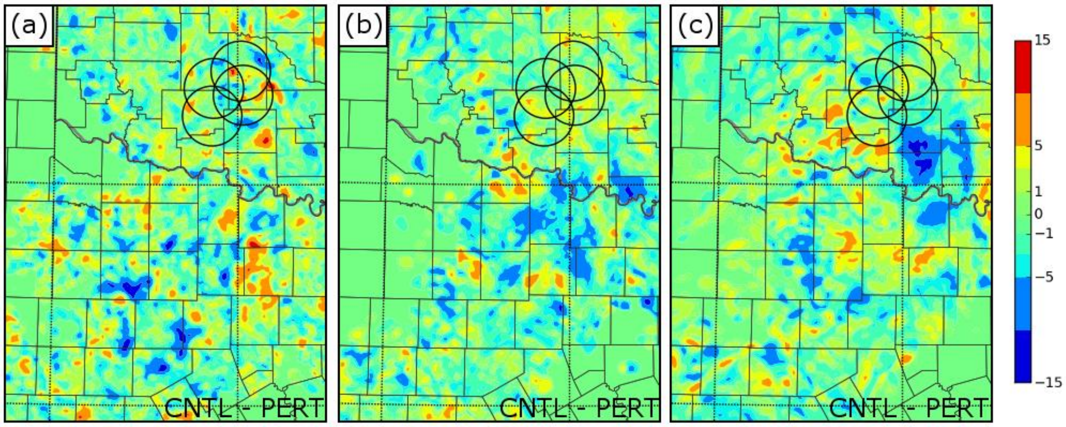

Fig. 5.22: As Fig. 5.17, but for (CNTL – PERT).

The greater variation in the trailing convective line in PERT can be seen in the difference between RMS ensemble spread of radar reflectivity between CNTL and PERT (Fig. 5.22). Early in the forecast period at 0300 UTC, there is little pattern in the difference in RMS error between CNTL and PERT (Fig. 5.22a). At 0400 and 0500 UTC, however, considerably higher RMS ensemble spread of *Z* is present over much of the region in and near the trailing convective line, as evidenced by the large area of negative values to the south and southeast of the CASA radar network in Fig. 5.22b-c. Though some positive values can be seen in other regions of the convective system (Fig. 5.22b-c), there are more negative regions than positive, indicating the presence of slightly greater spread in PERT than in CNTL.

5.3.4 Impact of Assumed Observation Error

One notable difference between the nested-grid experiments presented here and the single-grid experiments presented in Chapters 3 and 4 is a difference in the assumed observation error for radar reflectivity and radial velocity observations assimilated between 0105 and 0200 UTC. In the single-grid experiments, relatively low observation errors of 2 dBZ for radar reflectivity and 1 ms^{-1} for radial velocity were assumed during EnKF data assimilation. While these values have been used successfully in observation system simulation experiments (OSSE's) such as Jung et al. (2010), higher values may be optimal in experiments assimilating real data. Thus, in five of the inner-grid experiments, including CNTL and those discussed above in sections 5.3.1-5.3.3, increased observation errors of 3 dBZ for radar reflectivity and 2 ms^{-1} for radial velocity are assumed (see Table 5.1). Inner-grid experiment CNTL_LOW, however, uses the same reduced values of assumed observation error (2 dBZ for radar reflectivity and 1 ms^{-1} for radial velocity), as the single-grid experiments during EnKF radar data assimilation. During the forecast period, CNTL and CNTL_LOW share an identical ensemble design and model setup.

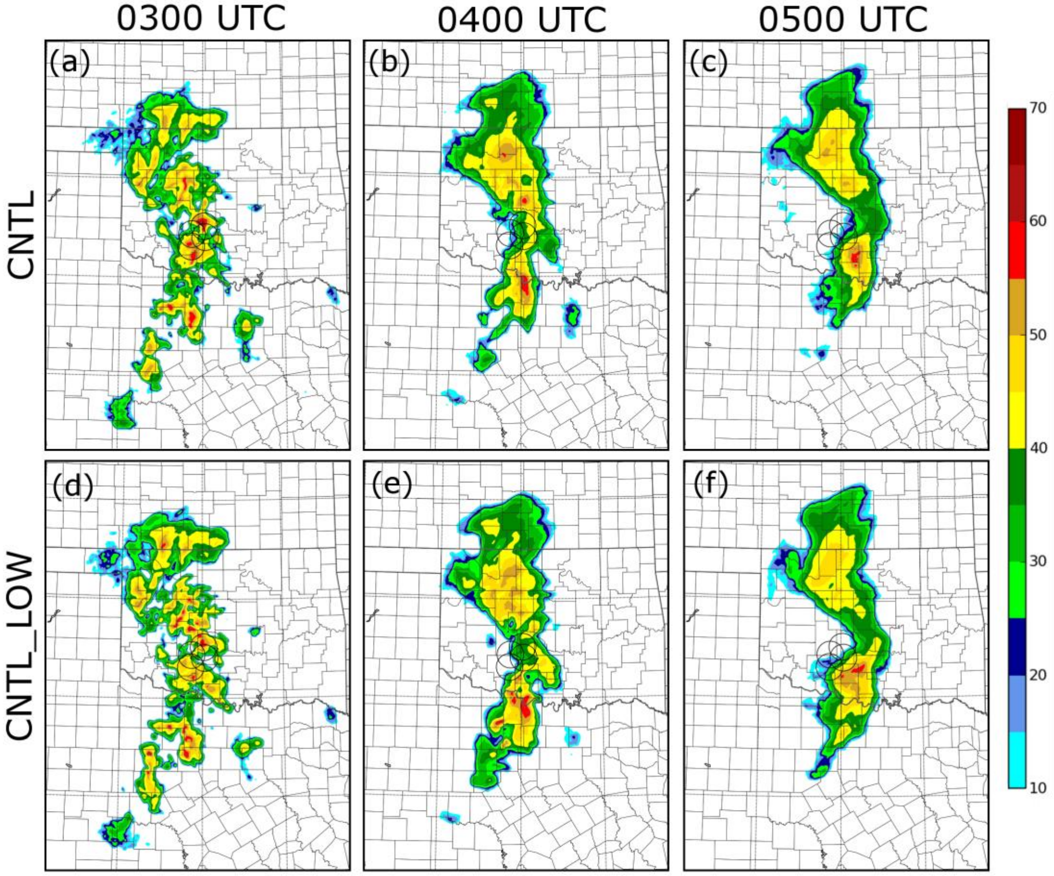

Fig. 5.23: As Fig. 5.8 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and CNTL_LOW at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC.

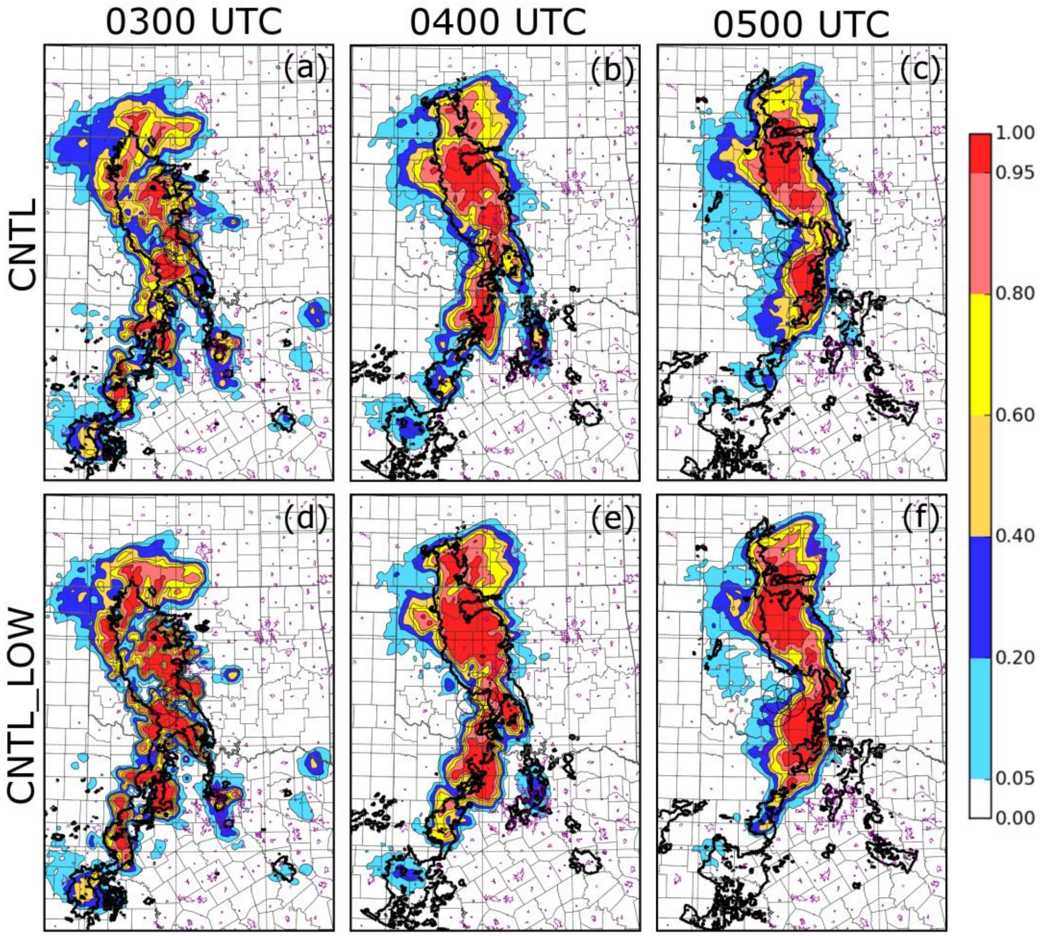

Fig. 5.24: As Fig. 5.9 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and CNTL_LOW at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC.

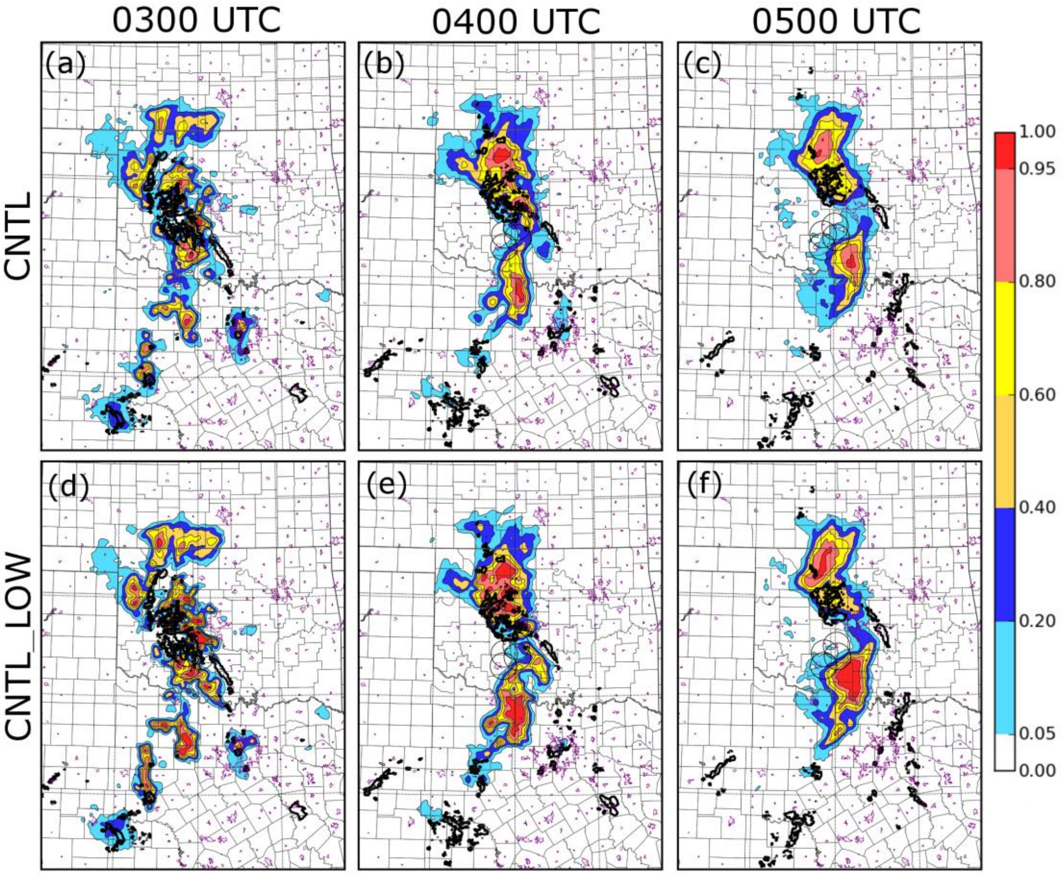

Fig. 5.25: As Fig. 5.10 but for CNTL at (a) 0300 UTC, (b) 0400 UTC, and (c) 0500 UTC; and CNTL_LOW at (d) 0300 UTC, (e) 0400 UTC, and (f) 0500 UTC.

In the PM ensemble mean of *Z*, differences between the two experiments are relatively minor. Compared to CNTL, CNTL_LOW predicts slightly stronger and more widespread convection in the region of the trailing convective line (Fig. 5.23b-c, e-f; Fig. 5.5c-d); elsewhere, the two experiments produce similar PM ensemble mean values. When the ensemble is used to generate a neighborhood ensemble probability forecast, however, the differences between the two experiments become more pronounced. For both NEP of $Z > 25$ dBZ (Fig. 5.24) and $Z > 40$ dBZ (Fig. 5.25), areas with NEP values exceeding 0.5 are similar between the two experiments, but NEP values are considerably greater in CNTL_LOW than in CNTL in many of these areas. Also, while both experiments overestimate the coverage of precipitation exceeding 40 dBZ (Fig. 5.25), especially late in the forecast period, the error is greater in CNTL_LOW, which predicts higher probabilities of Z > 40 dBZ over larger regions where no precipitation exceeding 40 dBZ was observed (Fig. 5.25f).

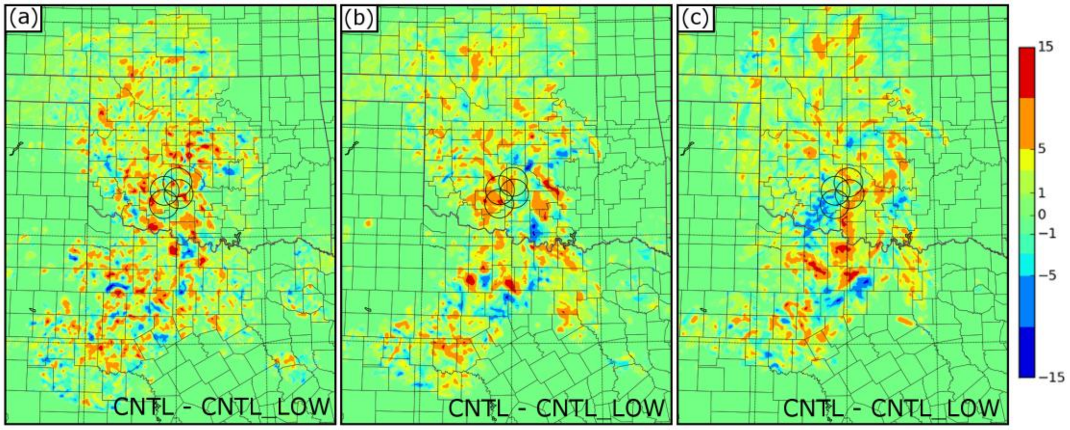

Fig. 5.26: As Fig. 5.17, but for (CNTL – CNTL_LOW).

The higher NEP values in many regions of the forecast in CNTL_LOW can be largely attributed to reduced ensemble spread in CNTL_LOW compared to CNTL. Fig. 5.26 shows the difference in RMS ensemble spread in Z between CNTL and CNTL_LOW on vertical model level 10 (at approximately 2 km above the surface) over the full extent of the MCS. Especially at 0300 UTC (Fig. 5.26a), the difference field is dominated by positive values, indicating that ensemble spread of Z is considerably higher in CNTL than in CNTL_LOW throughout the MCS. Higher ensemble spread of Z in CNTL is less pronounced later in the forecast period, but still persists to a lesser extent at 0400 UTC (Fig. 5.26b) and 0500 UTC (Fig.

5.26c), particularly near and to the south of the CASA domain in the region of the trailing convective line (Fig. 5.5c, d). This decreasing difference with time is not unexpected, because CNTL and CNTL_LOW share the same model and ensemble configuration during the forecast period; the only differences between the two experiments are the assumed observation errors used for Z and V_r during EnKF assimilation of radar data. During the forecast, as model error increases, the impact of the initial state (at 0200 UTC) is likely to be diminished.

In terms of objective forecast verification using the RSS and Brier score, the lower assumed observation errors used in CNTL_LOW appear to have a negative impact on the forecast quality. At 0200 UTC, CNTL and CNTL_LOW begin with similar RSS and Brier scores (Fig. 5.6, Fig. 5.7), both on the full domain and the verification sub-domain. During the forecast period, however, CNTL consistently performs slightly better than CNTL_LOW on the full domain in terms of both RSS (Fig. 5.6a) and Brier score (Fig. 5.7a). On the verification sub-domain (Fig. 5.6b, Fig. 5.7b) the difference is more pronounced; CNTL substantially outperforms CNTL_LOW at all hours during the forecast, and the difference increases with time. By 0500 UTC, at 3 hours of forecast time, CNTL retains a relatively skillful RSS of 0.66 on the verification sub-domain, while the RSS of CNTL_LOW has decreased to 0.51. At 0500 UTC, CNTL_LOW was the poorest performing of all six forecast experiments on the verification sub-domain in terms of both the RSS (Fig. 5.6b) and the Brier score (Fig. 5.7b). These results suggest that the higher values of assumed observation error for Z and V_r used in CNTL do indeed produce better forecast results for this real-data case.

5.3.5 Biases in the Ensemble Forecasts

In the single-grid forecast experiments presented earlier, two primary biases are identified in the reflectivity forecasts: a high bias in intense (> 35 dBZ) reflectivity, and a low bias in weak (< 20 dBZ) reflectivity (Fig. 4.17). These biases result, respectively, from over-prediction of strong convective regions (especially in the trailing line), and under-prediction of stratiform precipitation in the trailing line and trailing stratiform region. Biases in radar reflectivity are examined in a similar fashion for the inner-grid experiments by constructing reflectivity histograms using the method outlined at the beginning of section 4.3.3. The resulting histograms for inner-grid experiment CNTL are shown below in Fig. 5.27.

Compared to the similar reflectivity histograms produced for the single-grid experiments discussed in Chapter 4 (Fig. 4.17), several differences are apparent. In the initial forecast state at 0200 UTC, there is very little difference between the histograms of the ensemble mean and the ensemble members for the single-grid experiment (Fig. 4.17b, c). By contrast, for inner-grid experiment CNTL, while similar numbers of grid volumes contain reflectivity values greater than approximately 30 dBZ in the ensemble mean and ensemble members, the individual members contain significantly more low-reflectivity values (< 30 dBZ) than the ensemble mean (Fig. 5.27). This difference indicates greater spread in the inner-grid CNTL ensemble presented here, because spread in the ensemble leads to smoothing of the ensemble mean and difference in the reflectivity distribution between the ensemble mean and ensemble members. This difference is likely the result of the mesoscale perturbations added to the nested-grid ensemble on the outer grid, as discussed in section 5.1.

The two biases identified in Chapter 4 appear to be present in the inner-grid CNTL ensemble. During the forecast period (Fig. 5.27d-l), both the ensemble mean and ensemble members of CNTL exhibit a low bias weak (< 30 dBZ) precipitation and a high bias in intense (> 45 dBZ) precipitation at 0300, 0400, and 0500 UTC. Unlike the single grid experiments, however, the high bias in intense precipitation appears to be present primarily in individual ensemble members, and does not appear in the ensemble mean. This can also be attributed to greater ensemble spread in the inner-grid experiments than in the single grid experiments, both from mesoscale perturbations added on the outer grid, and from the use of higher assumed observation errors for Z and V_r, as discussed in section 5.3.4. The low bias in weak precipitation is present in both the ensemble members and ensemble mean, resulting from model error—specifically, the tendency of the model to predict too much strong convection compared to the observations.

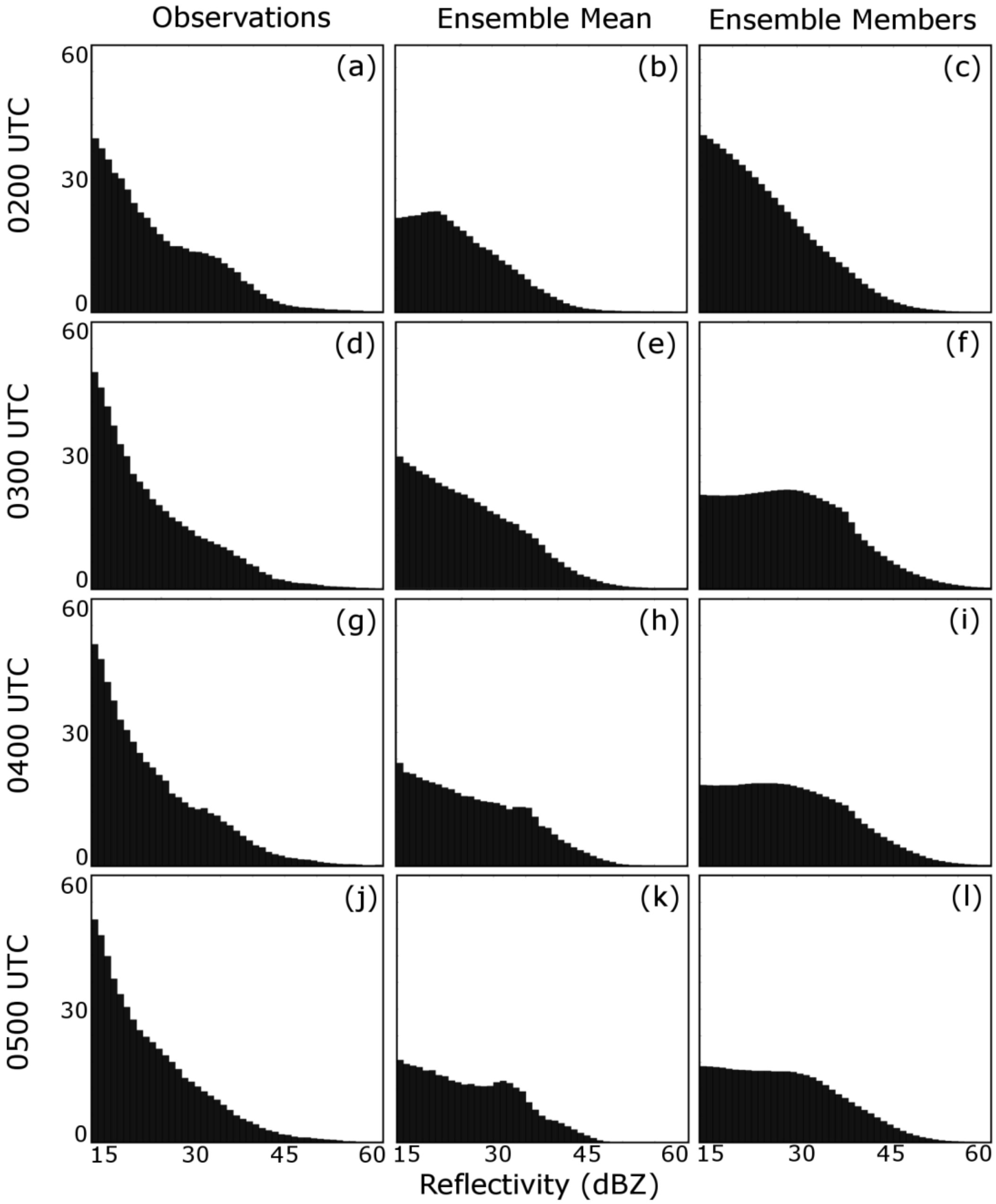

Fig. 5.27: Domain-wide histograms of reflectivity intensity in CNTL at 0200, 0300, 0400, and 0500 UTC. Shown are WSR-88D observations interpolated to the model grid (left column), the ensemble mean (center column), and individual ensemble members normalized by the size of the ensemble (right column). The vertical axis indicates the number of model grid volumes (in thousands) containing reflectivity of a given intensity.

While similar biases are observed in the five inner-grid experiments assimilating radar data (see Table 5.1), the biases in experiment NoRAD, which assimilates only conventional data, differ significantly from those of the other five inner-grid experiments. Domain-wide reflectivity histograms are shown below in Fig. 5.28 for NoRAD. The initial ensemble mean in NoRAD, obtained from the 0200 UTC ensemble member analyses, exhibits biases opposite to those seen in CNTL—specifically, the NoRAD ensemble mean has a low bias in intense precipitation (> 40 dBZ) and a strong high bias in light precipitation (< 30 dBZ). The 0200 UTC ensemble members in NoRAD show a high bias for light and moderate precipitation (< 40 dBZ).

The high bias of weak precipitation in NoRAD at 0200 UTC is particularly pronounced and is present in both the ensemble mean and in individual ensemble members. For precipitation between 15 and 20 dBZ, the model in NoRAD has 2 to 3 times as many grid volumes with precipitation of a given intensity as are present in observations. This bias quickly disappears, however; at later times, very little bias can be seen in light precipitation in either the NoRAD ensemble mean (Fig. 5.28e, h, k) or ensemble members (Fig. 5.28f, i, l). At and after 0300 UTC the ensemble mean of NoRAD consistently contains substantially less moderate and heavy precipitation than the ensemble members, a quality indicative of significant smoothing in the mean due to variation within the ensemble (see Fig. 5.11).

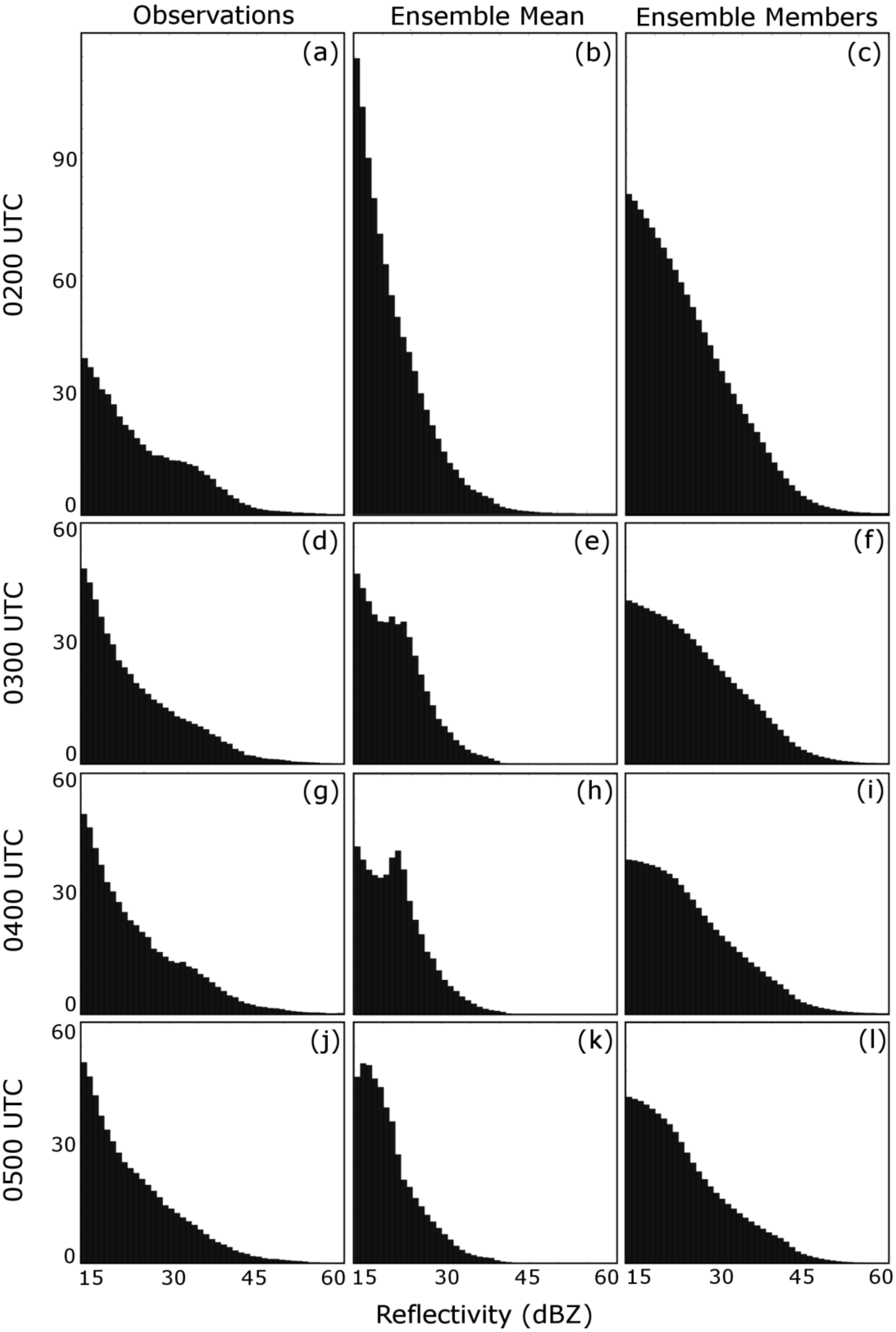

Fig. 5.28: As Fig. 5.27 but for NoRAD.

Even at 0500 UTC, significant differences can be seen in the reflectivity histograms between CNTL and NoRAD. In the ensemble members (Fig. 5.27l, Fig. 5.28l), CNTL shows a notable low bias of light precipitation that is absent in NoRAD. This suggests that assimilation of radar data, while it substantially improves the representation of the storm in the ensemble (as discussed in earlier sections), is also responsible for altering the model state in such a way that a low bias in weak (< 30 dBZ) reflectivity is introduced. In this case, this bias appears to be the result of an underestimation of stratiform regions of precipitation in the model and an overabundance of convection, particularly in the trailing line (see Fig. 5.13, Fig. 5.18, Fig. 5.23). Identifying the cause of this bias and considering ways to mitigate it can be topics for future investigations.

5.3.6 Ensemble Forecasts of Mesovortices

As in the single-grid forecast experiments presented in Chapter 4, each of the inner-grid forecast experiments (CNTL, NoCASA, NoRAD, PERT, R6, and CNTL_LOW) are manually examined at 0400 to identify significant low-level vortices, the method outlined in section 2.4. The vortices identified are considered objects for the purpose of verification, and their locations are used along with equation 2.2 to generate forecasts of the probability of one or more significant (vertical vorticity $\zeta > 0.02\ \mathrm{s}^{-1}$) low-level vortices being present within 25 km of a point. The probabilities thus calculated are shown for each of the inner-grid ensemble forecast experiments below in Fig. 5.29. When comparing these results to those presented for the single-grid experiments in Fig. 4.20, it should be noted that the inner-grid CNTL experiment (Fig. 5.29a) uses a single-microphysics ensemble, and thus, of the three single-grid experiments, most closely resembles the single-grid NoMMP experiment (Fig. 4.20d), and not the single-grid CNTL experiment (Fig. 4.20a) which uses a mixed-microphysics ensemble. Also, it should be remembered that the single-grid experiments presented in Fig. 4.20 use lower values of assumed observation error for V_r and Z during EnKF data assimilation, identical to those used in CNTL_LOW.

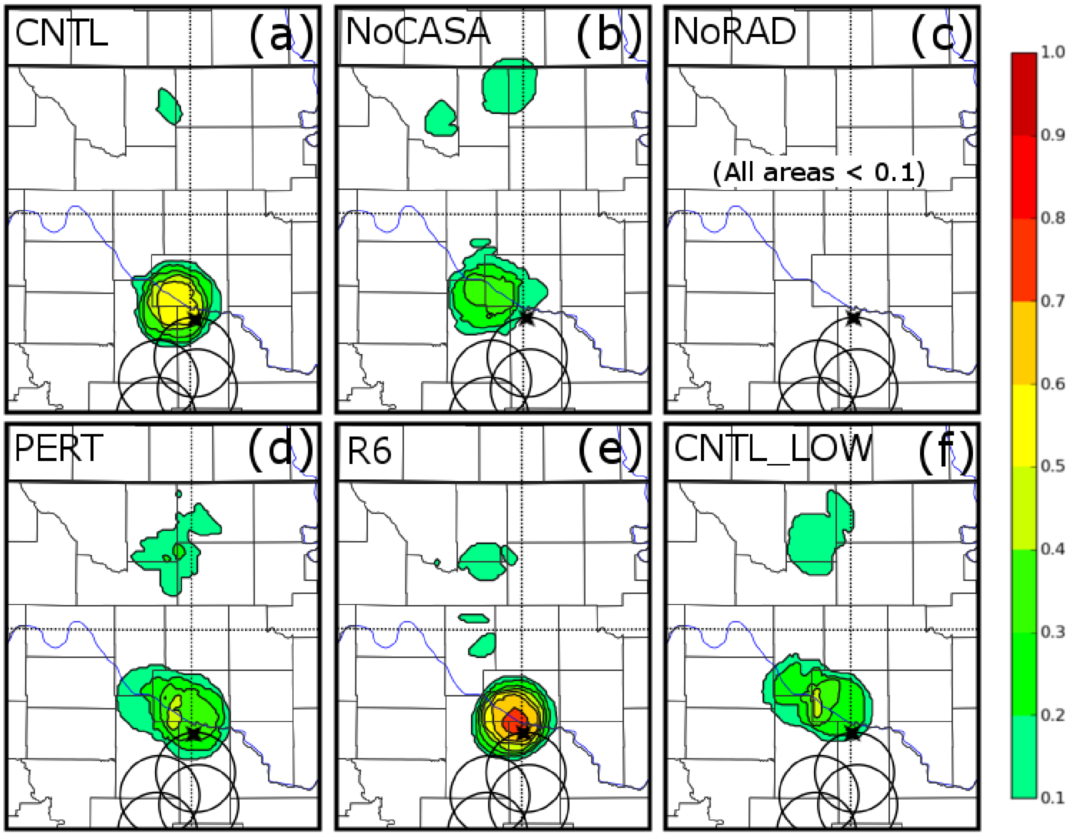

Fig. 5.29: Probability of significant low-level vortices within 25 km of a point at 0400 UTC for (a) CNTL, (b) NoCASA, (c) NoRAD, (d) PERT, (e) R6, and (f) CNTL_LOW. 30 km CASA radar range rings are shown. The position of the EF-1 tornado reported beginning at 0354 UTC is indicated by the black star.

Of the six inner-grid ensemble forecast experiments, five indicate probabilities of greater than 0.3 of significant low-level vortices being present in the vicinity of the tornadic mesovortex observed at 0400 UTC. Only NoRAD (Fig. 5.29c), which did not assimilate radar data, fails to indicate the presence of this significant low-level circulations in the vicinity of the mesovortex. The greatest probability is indicated by experiment R6 (Fig. 5.29e), which predicts a maximum probability of greater than 0.7 and a probability field tightly concentrated around the location of the observed mesovortex. CNTL (Fig. 5.29a) also exhibits a probability field concentrated near the observed mesovortex, though the maximum probability value in CNTL is lower, between 0.5 and 0.6. The remaining three experiments (PERT, NoCASA, and CNTL_LOW) exhibit somewhat less-concentrated probability fields with maxima ranging between 0.3 and 0.5 (Fig. 5.29b, e, f).

The higher probability predicted in the vicinity of the tornadic mesovortex in R6 (Fig. 5.29e) as compared to CNTL (Fig. 5.29a) suggests that increasing the rain intercept parameter improves the ensemble forecast of the tornadic mesovortex observed in this case. CNTL uses a fixed rain intercept parameter of $8.0 \cdot 10^5$, while R6 increases this by a factor of 5, using a fixed rain intercept parameter of $4.0 \cdot 10^6$. In the single-grid experiments presented in chapter 4, the ensemble forecasts are found to produce relatively weak cold pools of approximately 2 K (Fig. 4.18), compared to observed temperature drops of approximately 4 K during the passage of the MCS at nearby Oklahoma mesonet surface observation sites (not shown). The increased rain intercept parameter used in R6 results in a drop size distribution that favors smaller raindrops. Because these smaller raindrops have a greater total surface area, evaporation of raindrops is increased in R6, thus increasing the amount of cooling occurring as a result of the evaporation of raindrops. While Snook and Xue (2008) found that a rain intercept parameter of $8.0 \cdot 10^5$ produced more realistic cold-pool structure for a supercell case, this does not appear to hold true for the 8-9 May 2007 MCS case studied here.

The probabilistic forecast of mesovortices did not appear to benefit from a perturbed rain intercept parameter, as evidenced by the lower maximum and more diffuse probability distribution in PERT (Fig. 5.29d) as compared to that in CNTL (Fig. 5.29a). CNTL already exhibits a relatively good forecast of the low-level mesovortex at 0400 UTC (Fig. 5.29a). For this reason, the increased ensemble spread of PERT induced by the perturbation of the rain intercept parameter among its ensemble members has a negative impact on the mesovortex forecast.

The inclusion of CASA radar data significantly improves the 0400 UTC mesovortex forecast; this finding agrees with results from the single-grid forecast experiments presented in Chapter 4 (Fig. 4.20). Compared to CNTL (Fig. 5.29a), NoCASA (Fig. 5.29b) predicts a lower maximum probability in the vicinity of the observed mesovortex, and slightly higher probabilities well to the north of the CASA domain. These false alarms, indicated by the regions of probability between 0.1 and 0.25 near the Oklahoma-Kansas border in several of the experiments, are the result of small circulations in the northern stratiform region of the MCS within a few ensemble members. Inclusion of CASA radar data provides valuable information about the near-surface flow, which is important for prediction of features, such as mesovortices, that are strongly sensitive to low-level dynamics.

Decreasing the assumed observation error for V_r and Z, as was done in CNTL_LOW (Fig. 5.29f) resulted in a poorer ensemble forecast of the near-surface mesovortex observed at 0400 UTC. The probability field predicted by CNTL_LOW (Fig. 5.29f) somewhat resembles that of the NoMMP single-grid forecast (Fig. 4.20f)—this is not surprising considering that the single-grid NoMMP ensemble also used the smaller observation errors (2 dBZ for Z and 1 m s^{-1} for V_r) that CNTL_LOW uses. This result supports the speculation of Snook et al. (2011) that the smaller observation errors used in that study and in CNTL_LOW are likely lower than optimal.

5.4 Comparison to 3DVAR-based Forecast Results

One goal of the current set of nested grid experiments is to provide a platform for directly comparing EnKF and 3DVAR results. To facilitate this, three of the inner-grid experiments are configured to correspond closely to 3DVAR data assimilation and forecast experiments performed by Schenkman et al. (2011). Schenkman et al. (2011) use the ARPS 3DVAR system with a cloud analysis to assimilate conventional and radar data for the same case studied here. The cloud analysis uses reflectivity data to estimate hydrometeor fields for rainwater, hail, and snow and replaces the background with these estimates anywhere that observations are available. The cloud analysis provides a way of updating temperature, moisture and cloud microphysical fields within a 3DVAR framework. With EnKF, flow-dependent cross-covariances are used to directly assimilate reflectivity data for the update of all state variables including temperature and moisture.

Schenkman et al. (2011) report results from five experiments assimilating different combinations of V_r and Z observations from CASA and WSR-88D radars along with conventional observations. The inner-grid experiments CNTL, NoCASA, and NoRAD presented in this chapter use similar data and model settings to experiments CASAVrZ, 88DONLY, and NORAD from Schenkman et al. (2011). The relationship between experiments presented in this work and those of Schenkman et al. (2011) is detailed below in Table 5.3.

Table 5.3: Comparison of similar experiments from this study and in Schenkman et al. (2011).

Experiment from this work	**Experiment from Schenkman et al. (2011)**	**Data Assimilated**		
		Conventional Data?	**WSR-88D Data?**	**CASA Data?**
CNTL	*CASAVrZ*	Yes	Yes	Yes
NoCASA	*88DONLY*	Yes	Yes	No
NoRAD	*NORAD*	Yes	No	No

The experiment design used in CNTL, NoRAD, and NoCASA is chosen to closely follow that of Schenkman et al. (2011). Both investigations use the same ARPS model domain of 503 × 503 × 43 grid points with horizontal grid spacing of 2 km in both the x- and y-directions, and the same stretched vertical coordinate with minimum vertical grid spacing of 100 m and average vertical grid spacing of 500 m. In both investigations, conventional and radar data are assimilated from 0100 to 0200 UTC at 5 minute intervals, and the same types of radar data are assimilated in corresponding experiments in both investigations.

Though CNTL, NoCASA, and NoRAD are designed to correspond closely to experiments from Schenkman et al. (2011) as described above, some significant differences between the EnKF and 3DVAR experiments still exist. Most notably, the EnKF experiments are nested within an outer 6-km domain initialized at 1800 UTC, 8 May 2007 from the 8 May 2007 NCEP 1800 UTC NAM analysis; conventional data are assimilated hourly on this outer-grid ensemble between 2100 UTC, 8 May 2007 and 0100 UTC, 9 May 2007. By contrast, the experiments of Schenkman et al. (2011) are directly initialized from a one-hour spinup forecast starting from a 3DVAR analysis using the 9 May 2007 NCEP 0000 UTC NAM analysis as its background. The outer grid in the EnKF experiments is necessary to properly introduce mesoscale perturbations into the ensemble and provide needed variation in boundary conditions to the inner-grid ensemble.

Some differences also exist in the data assimilated between the two investigations. Two of the experiments of Schenkman et al. (2011) assimilate level-III radar reflectivity observations from WSR-88D radar site KFDR in lieu of level-II data from that site (the level-II data are unavailable for KFDR during the data assimilation period). Because level-III data store radar reflectivity on a rather coarse interval (5 dBZ), and thus can be expected to contain significantly larger observation errors than level-II radar data, no data from KFDR were assimilated in the EnKF experiments of this work. Also, the format of radar data assimilated differs between the experiments presented here and those of Schenkman et al. (2011). In Schenkman et al. (2011), WSR-88D and CASA radar data are interpolated directly to the Cartesian ARPS model grid for use in 3DVAR data assimilation. For the

EnKF experiments of this work, WSR-88D data are interpolated to Cartesian model coordinates in the x- and y-directions, but allowed to remain on radar elevation surfaces in the vertical direction. CASA radar data are interpolated both radial by radial in time before being interpolated to the ARPS grid in the x- and y-directions, as described in section 5.1. This choice of interpolation methodology is made to reduce interpolation error during EnKF data assimilation, though the impact of this difference in radar data format between the experiments presented here and those of Schenkman et al. (2011) is not expected to be large. A third difference of note pertains to the use of Oklahoma mesonet data; in the EnKF experiments of this work, Oklahoma mesonet data are assimilated every 5 minutes, while in Schenkman et al. (2011) these data are assimilated hourly. In future work with Alex Schenkman, 3DVAR experiments are planned which will eliminate the differences in the use of Oklahoma Mesonet and KFDR level-III radar reflectivity data.

Another challenge in comparing EnKF and 3DVAR results of the two investigations is the use of ensemble forecasts in this work compared to deterministic forecasts in Schenkman et al. (2011). Because EnKF is an ensemble data assimilation method and produces an ensemble of analyses as its final state, it is natural to perform an ensemble forecast from an EnKF ensemble analysis. By contrast, 3DVAR is a variational method that produces a single, deterministic best-guess analysis. While it is possible to perform an ensemble forecast based on a 3DVAR analysis, the process is not as straightforward as it is when using EnKF. Because ensemble mean fields are significantly smoother than ensemble member fields at the same time and may differ significantly in structure (see Fig. 5.12, Fig.

5.13), it is not feasible to directly compare ensemble mean forecasts from EnKF experiments to the deterministic forecasts of Schenkman et al. (2011). For this reason, direct comparisons with deterministic forecasts will be limited to individual ensemble members. Ensemble member #35, a member which performed reasonably well in CNTL, NoCASA, and NoRAD, is chosen for this purpose.

At 0350 UTC, just prior to the start of the reported tornadic activity at 0354 UTC, the three experiments of Schenkman et al. (2011) each predict a MCS in the approximate location of the observed MCS as shown below in Fig. 5.30. The 3DVAR experiment assimilating only conventional data (Fig. 5.30b) predicts a very small line of convection near the center of the LEV, with little to no stratiform precipitation. By contrast, at 0400 UTC, ensemble member #35 of EnKF experiment NoRAD predicts a much larger MCS (Fig. 5.31b), with the core of the convection slightly located several tens of kilometers south of the observed LEV. While the structure of the MCS predicted is in NoRAD is less linear than the MCS predicted in the 3DVAR experiment, the area of coverage of precipitation agrees much better with observations (Fig. 5.30a, Fig. 5.31a). Furthermore, in NoRAD, the PM ensemble mean at this time (Fig. 5.8e) predicts the extent of the MCS better than the individual ensemble members, suggesting a benefit from the ensemble forecast that cannot be realized by the single deterministic forecast of Schenkman et al. (2011). These results suggest that, on a system-wide scale, the EnKF forecast benefits from cross-covariances that are computed when using EnKF (and not computed when using 3DVAR) for assimilation of conventional data.

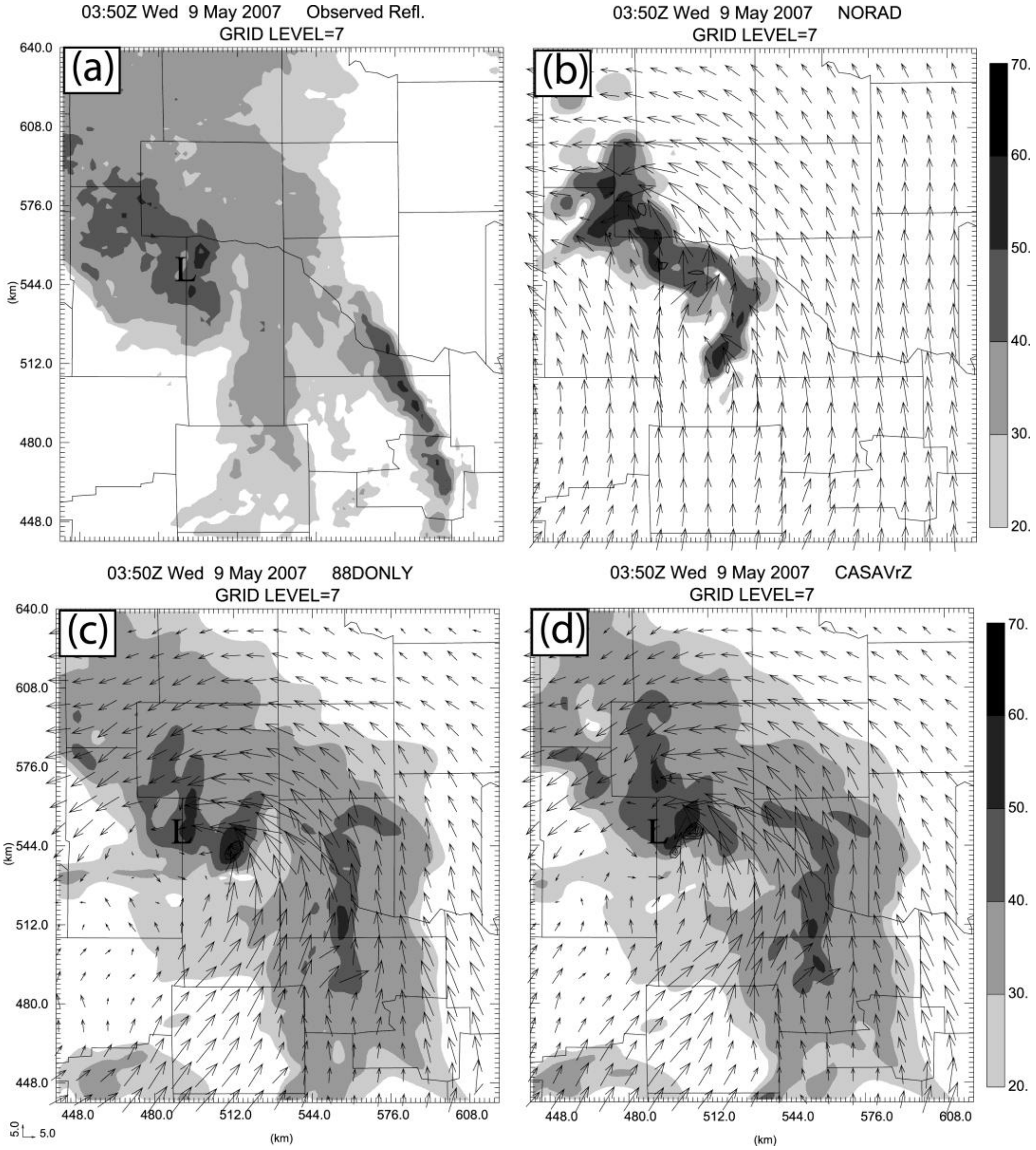

Fig. 5.30: 0350 UTC 9 May 2007 (a) observed reflectivity (shaded, dBZ) and forecast reflectivity (shaded, dBZ), and horizontal wind vectors (ms-1) from experiments (b) NORAD, (c) 88DONLY, and (d) CASAVrZ of Schenkman et al. (2011). Reflectivity is plotted on vertical model grid level 7 (approximately 1100 m above the surface). The "L" in panels (a), (c), and (d) marks the approximate observed location of the line-end vortex at 0350 UTC. This figure is reprinted with permission from Fig. 11 of Schenkman et al. (2011).

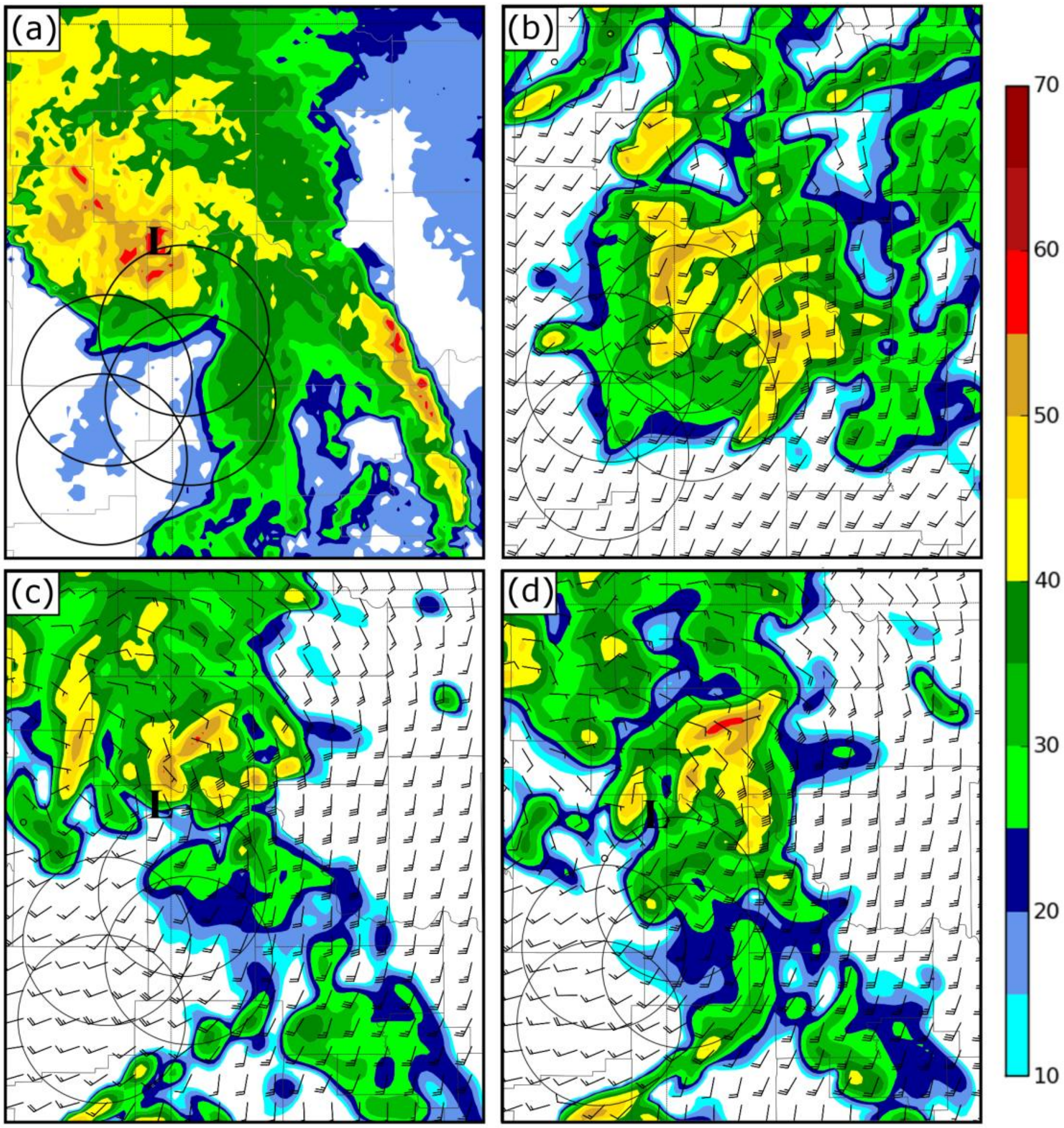

Fig. 5.31: 0400 UTC 9 May 2007 (a) observed reflectivity (shaded, dBZ) and forecast reflectivity (shaded, dBZ), and horizontal wind (barbs) from ensemble member #35 of experiments (b) NoRAD, (c) NoCASA, and (d) CNTL. Reflectivity is plotted on vertical model grid level 7 (approximately 1100 m above the surface). The "L" in panels (a), (c), and (d) marks the approximate observed location of the line-end vortex at 0400 UTC.

In the 3DVAR experiment assimilating WSR-88D radar data and conventional data (but not CASA radar data), the predicted MCS (Fig. 5.30c) has a relatively well-defined line-end vortex (LEV) circulation, located just to the southeast of the observed LEV (Fig. 5.30a). The MCS in the forecast from the 3DVAR experiment has a relatively large region of stratiform rain to the north of the LEV, in good agreement with observations, as well as a single line of trailing convection located between the observed trailing and leading convective lines (Fig. 5.30a). The NoCASA ensemble member forecast (Fig. 5.31c) also predicts a MCS with a well-defined LEV, though the LEV in the NoCASA member is located approximately 15 km to the west of the observed LEV. NoCASA also predicts a large region of stratiform rain to the north of the LEV, and unlike the forecast from the 3DVAR experiment of Schenkman et al. (2011), NoCASA predicts elements of both the trailing and leading convective lines, though the trailing convective line is stronger in NoCASA (Fig. 5.31c) while the leading line is stronger in observations (Fig. 5.31a). At least near the CASA domain, however, neither of the convective lines predicted by NoCASA is as intense or well-organized as the single line in the 3DVAR forecast.

Both the CNTL ensemble member (Fig. 5.31d) and the forecast from the 3DVAR experiment assimilating CASA and WSR-88D radar data (Fig. 5.30d) predict a well-organized MCS with a LEV located very near the observed LEV. In both investigations, the experiment assimilating both CASA and WSR-88D data produces the most accurate LEV forecast. As for the experiments that assimilated WSR-88D radar data along with conventional observations, the 3DVAR experiment predicts a single, strong trailing convective line, while the CNTL ensemble member predicts both the leading and trailing convective lines.

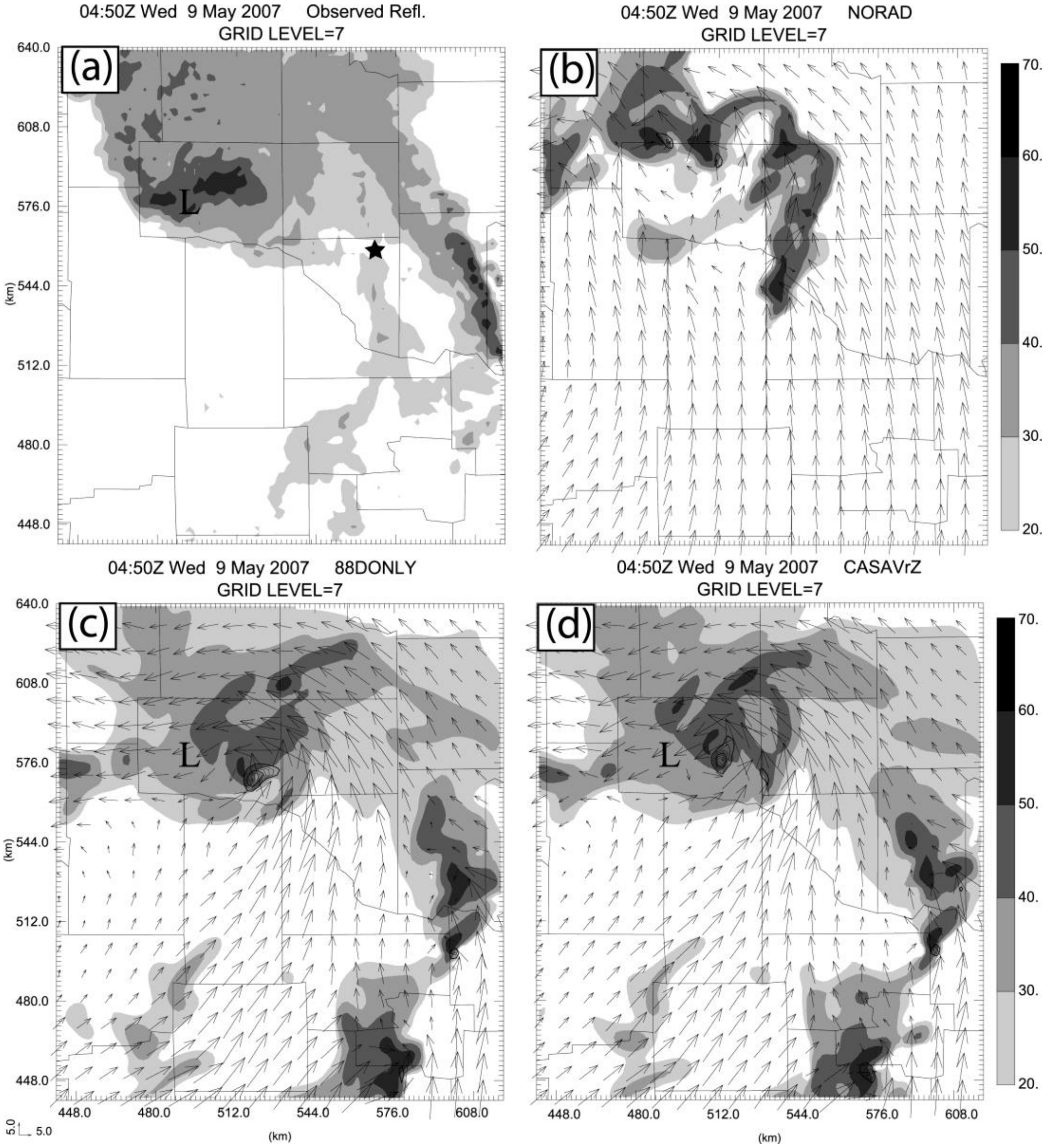

Fig. 5.32: As Fig. 5.30, but at 0450 UTC 9 May 2007. This figure is reprinted with permission from Fig. 13 of Schenkman et al. (2011).

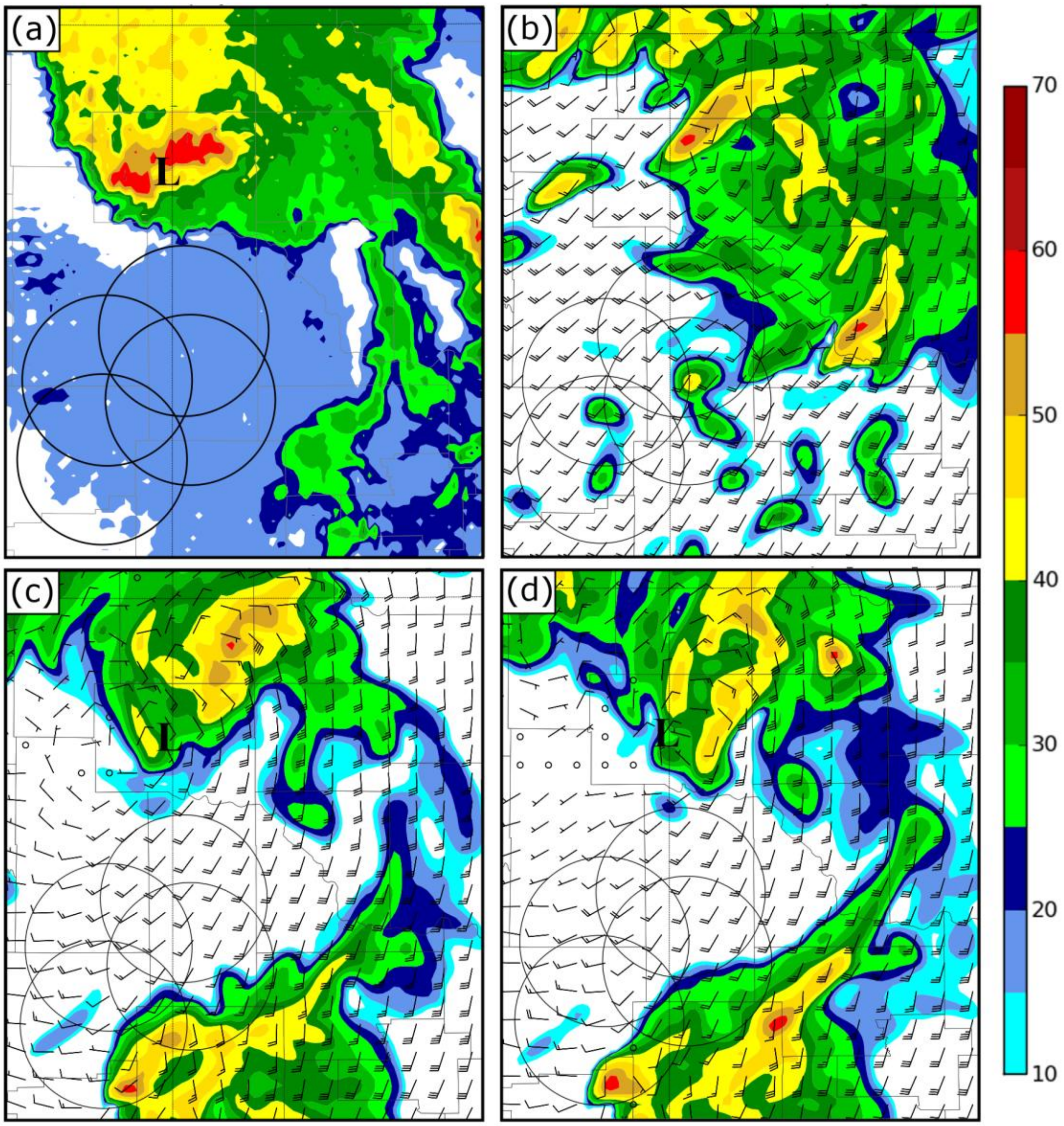

Fig. 5.33: As Fig. 5.31, but at 0500 UTC 9 May 2007.

Near the end of the forecast period, the observed LEV has moved approximately 30 km to the north from its position at 0400 UTC (Fig. 5.32a). The 0450 UTC forecast from the 3DVAR experiment assimilating only conventional data (Fig. 5.32b) still predicts a much smaller MCS than observed, with a reflectivity pattern that compares poorly to observations (Fig. 5.32a). The NoRAD ensemble member at 0500 UTC (Fig. 5.33b) predicts a much larger MCS, with a large region of stratiform rain and, compared to 0400 UTC, improved linear structure. While the NoRAD ensemble member compares only moderately well to observations (Fig. 5.32a), the forecast of precipitation coverage in the NoRAD appears qualitatively superior to that of the forecast from the 3DVAR experiment assimilating only conventional data.

For both experiments assimilating radar data (Fig. 5.32c, d), the 3DVAR-based forecasts predict an MCS that agrees well with observations, containing a strong LEV located 15-20 km to the east of the observed LEV location (Fig. 5.32a). The LEV is slightly stronger, and the position error slightly less, in the forecast from the 3DVAR experiment assimilating both CASA and WSR-88D radar data (Fig. 5.32d). Likewise, the NoCASA (Fig. 5.33c) and CNTL (Fig. 5.33d) ensemble members both predict an MCS with a well-defined LEV. The LEV in the NoCASA ensemble member is located approximately 5-10 km to the southwest of the observed LEV, while in the CNTL ensemble member, it is located approximately 5-10 km to the east of the observed LEV; in both cases, the position error in the LEV forecast is less than in the corresponding 3DVAR-based forecast. At 0500 UTC, the NoCASA and CNTL ensemble member forecasts predict only a single trailing line of convection, much like the 3DVAR-based forecasts, and both NoCASA and CNTL strongly over-predict the intensity of the trailing convective line (Fig. 5.33). While the 3DVAR-based forecasts (Fig. 5.32) also over-predict the intensity of the trailing convective line, the amount and areal extent of the over-prediction is somewhat less.

Chapter 6: Summary and Discussions

6.1 Summary of Results

Two sets of experiments are conducted using an ensemble Kalman filter to assimilate conventional and/or radar data into the ARPS numerical weather prediction model for the case of a tornadic MCS/LEV that occurred over Texas and Oklahoma on 8-9 May 2007. Three-hour forecasts are produced from the resulting ensemble analyses, with the goal of obtaining accurate short-term forecasts of radar reflectivity and meso-vortices. The first set of experiments, conducted on a single model domain with 2-km grid spacing, examines the impact of assimilating CASA radar data and of using a mixed-microphysics ensemble. The second set of experiments, conducted on a pair of nested ensemble domains using 6- and 2-km grid spacing, examines the impact of assimilating various conventional and radar data sources, of variation in the rain intercept parameter of the single-moment ice microphysics scheme (Lin et al. 1983) used in the model, and of variation in the assumed observation error used during EnKF data assimilation for radar data. The results of these experiments are summarized below, beginning with the results of the single-grid experiments.

6.1.1 Single-grid EnKF Analysis of Radar Data

In the single-grid EnKF data assimilation study, the impact of CASA radar data and the use of a mixed-microphysics ensemble are examined for the 8-9 May 2007 tornadic MCS/LEV that occurred over Texas and Oklahoma. Inclusion of

CASA data imparted a noticeable improvement on the mean ensemble analysis, particularly within the CASA domain. Using a multi-microphysics ensemble helps to improve the spread of the ensemble. The best analysis is obtained in CNTL, which used both CASA data and a multi-microphysics ensemble.

As evidenced by a time-height analysis of vertical vorticity and updraft intensity, as well as the low-level wind field, assimilation of radar data from the network of 4 CASA radars improves the resulting analysis, with the most significant improvement in the ensemble mean analysis seen in the lowest few kilometers of the atmosphere where WSR-88D coverage is poorest. The assimilation of CASA data does not have a very large impact on average RMS innovation and ensemble spread statistics, however the effect of assimilating CASA data is a slight reduction in RMS innovation at the WSR-88D radar sites whose coverage areas overlap the CASA network domain (KTLX and KVNX). This reduction is present in data for both V_r and Z, though the reduction is larger for Z, and is observed primarily during the first six assimilation cycles. Assimilation of CASA data thus allows the EnSRF system to more quickly achieve its best estimate of the atmospheric state. Assimilation of CASA data does not appear to have any significant effect on RMS ensemble spread of V_r or Z.

Using a mixed-microphysics ensemble during data assimilation greatly increases the average RMS ensemble spread of Z at all radar sites; in many cases the RMS ensemble spread of Z when using a mixed-microphysics ensemble is more than twice that observed when a single-microphysics ensemble is used. This increase constitutes a significant improvement in the analysis, helping to alleviate the marked under-dispersiveness of the analysis ensemble; the consistency ratio for reflectivity in the mixed-microphysics experiment is much closer to the theoretically-expected value of approximately 1.0 than that of the single-microphysics experiment. Even in the mixed-microphysics experiment, however, under-dispersiveness is still present in Z, and all three experiments show marked under-dispersiveness in $V_{r,}$ despite covariance inflation of 25% being applied. The use of a mixed-microphysics ensemble, while beneficial, is not alone sufficient to counteract the low ensemble spread often observed in EnKF studies, at least for this case. To address this issue, use of more robust multi-moment microphysical parameterizations within the ensemble, such as those used in Putnam et al. (2010), may be effective, as well as use of additive perturbations to horizontal wind variables, and application the relaxation technique of Zhang et al. (2004); such techniques can be topics for future study.

6.1.2 Single-grid Deterministic and Ensemble Forecasts

Following the EnKF analysis of CASA and WSR-88D radar data, we perform 3-hour deterministic and ensemble forecasts for the 8-9 May 2007 tornadic MCS, starting from ensemble Kalman filter analyses assimilating CASA and WSR-88D radar data. We examine the impact of assimilating CASA (in addition to WSR-88D) radar data on the forecast, and the impact of variation in the choice of model microphysics during both the assimilation and forecast periods. The choice of microphysical scheme makes a significant impact on forecast evolution for predictions of radar reflectivity and mesovortices. Deterministic forecast experiments initialized from the ensemble mean of analysis experiments using single- and mixed-microphysics ensembles differ noticeably in their low-level wind and reflectivity fields, and in particular the evolution of the vertical updraft structure within the system. When model microphysics is varied for deterministic forecasts initialized from a single 0200 UTC EnKF ensemble mean analysis, the forecast using the NEM microphysical scheme is found to greatly under-forecast stratiform precipitation regions. Similar results were seen in the ensemble forecast experiments, where members using the Lin or WSM6 scheme performed considerably better in prediction of low-level mesovortex locations than members using the NEM microphysical scheme.

From the ensemble forecast experiments, probabilistic predictions for radar reflectivity and low-level circulations are obtained. A neighborhood ensemble probability approach (Schwartz et al. 2010) is applied to generate 60-180 minute probabilistic forecasts for radar reflectivity, and an object-oriented ensemble forecast approach is used to generate 120-160 minute probabilistic forecasts of low-level mesovortices. Both the assimilation of CASA data and the use of a mixed-microphysics ensemble during the EnKF data assimilation cycles have significant impacts on the resulting forecasts in terms of convective structure, and both have a positive impact on the forecast of low-level vortices.

All ensemble forecast experiments successfully predict the dominant convective mode during the forecast period, indicating a MCS with an embedded LEV, and a convective line extending to the south, though the southern end of the convective line decays too quickly in the forecasts as a result of interaction with the southern domain boundary. The ensemble forecasts show varying degrees of success in predicting the location of tornadic mesovortices between 120 and 160 minutes of forecast time; the control forecast, which uses a mixed-microphysics ensemble and assimilates CASA radar data, performs particularly well, predicting a strong, localized probability maximum of 0.65 less than 5 km from the observed mesovortex that spawned the 0354 UTC EF-1 Minco tornado. The experiment that does not assimilate CASA data and the experiment using a single-microphysics ensemble predict comparatively diffuse probability fields, with lower probabilities of 0.43 and 0.35 respectively at the observed location of the tornadic mesovortex.

6.1.3 Nested-grid Experiments Assimilating Conventional and Radar Data

Expanding upon the single-grid EnKF data assimilation and ensemble forecast experiments presented in chapters 3 and 4, we perform a set of nested-grid EnKF experiments, also using the 8-9 May 2007 tornadic MCS as a case study. In these experiments, conventional data are assimilated in addition to radar data, and mesoscale perturbations are introduced on the outer-nest ensemble in addition to the assimilation of conventional data. A 40-member outer-nest ensemble with 6 km horizontal grid spacing is used, within which are embedded six inner-nest experiments. In addition to the control experiment, the experiments performed can be broken down into three categories: (1) two experiments to investigate the impact of assimilating radar data from the WSR-88D and CASA radar networks, (2) two experiments to investigate the impact of varying the rain intercept parameter of the model microphysical scheme, and (3) one experiment investigating the impact of changing the assumed observation error for radar reflectivity and radial velocity observations. Ensemble forecasts of radar reflectivity were produced, analyzed, and verified against WSR-88D observations, and biases in these forecasts were considered. Forecasts of the tornadic mesovortex observed at 0400 UTC were also performed and compared among experiments and against the observed mesovortex.

As expected, the control experiment, which assimilated CASA and WSR-88D radar data in addition to conventional data, substantially outperforms the experiment that assimilated only conventional data (i.e. no radar data) in terms of the structure and location of the MCS predicted. In particular, many members of the control experiment predict a well-defined line-end vortex that follows a similar path to the observed line-end circulation, while almost no members do in the experiment assimilating only conventional data. Even when no radar data are assimilated, however, the resulting ensemble forecast is found to retain useful skill throughout the 3-hour forecast period. When WSR-88D radar data and conventional data are assimilated, but CASA data are not, the ensemble forecast performs similarly to the control experiment in qualitative prediction of the MCS evolution and in objective verification using the ROC skill score and Brier score. With regard to probabilistic forecasts of the tornadic mesovortex, however, the control experiment outperforms the experiment that did not assimilate CASA data.

Experiments are performed to test two different variations to the model microphysical scheme: (1) increasing the rain intercept parameter by a factor of five, and (2) randomly varying the rain intercept parameter between the default value and 10 times the default value. Both variations are found to improve the resulting ensemble forecasts of radar reflectivity. The predicted MCS looks qualitatively similar in the probability-matched ensemble mean forecast of the control experiment and the experiments with variation in the model microphysics, but the trailing convective line of the MCS is better predicted in the two experiments using an altered rain intercept parameter. The control experiment performs slightly better than the experiment with the randomly varying rain intercept parameter in prediction of the tornadic mesovortex occurring at 0400 UTC, but the experiment with a uniformly increased rain intercept parameter performs the best of all six experiments in its probabilistic prediction of the tornadic mesovortex.

Decreasing the assumed observation error of radar reflectivity from 3 dBZ to 2 dBZ, and of radial velocity from 2 ms^{-1} to 1 ms^{-1}, results in negative impacts to both reflectivity forecasts and probabilistic forecasts of the tornadic mesovortex. The experiment using the reduced observation errors during assimilation produces an ensemble forecast that has larger errors in the trailing convective line than the control experiment. The control experiment has a substantially higher ensemble spread for radar reflectivity compared to the experiment with reduced observation error, and produces a substantially better probabilistic forecast of the tornadic mesovortex.

6.2 Concluding Remarks

Particularly for a tornadic system of this nature, the horizontal grid spacing used in this study (2 km) is insufficient to fully capture all important sub-storm-scale processes occurring within the domain. While 2 km horizontal grid spacing is sufficient to capture the line-end vortex and near-surface mesovortex circulations observed in the 8-9 May 2007 case, a significantly smaller horizontal grid spacing (on the order of 100 m) would be necessary to capture tornado-scale circulations.

In addition, we note the challenges of working with X-band Doppler radar data, perhaps the greatest of which is attenuation. While S-band radars, such as the WSR-88D network suffer very little attenuation, even through heavy precipitation, attenuation in X-band radar data is much more significant. An X-band radar beam passing through more than approximately 10 km of heavy precipitation (> 50 dBZ) is attenuated completely, leaving a “shadow” beyond the range where complete attenuation occurred. Additionally, even when attenuation is incomplete, error within the attenuation correction algorithm leads to discrepancy between X-band and S-band observations of the same volume. While CASA X-band data are assimilated with reasonable success when a minimum threshold of 20 dBZ is applied to the CASA data, applying this threshold eliminates the ability of the CASA radar data to suppress spurious convection via assimilation of clear-air data. Finding effective ways to better correct for attenuation, objectively identify areas where complete attenuation has occurred, and to mitigate the effects of attenuation upon data assimilation will be vital in improving future data assimilation endeavors using X-band radar data.

We also point out the value of polarimetric radar data for purposes of data assimilation. Polarimetric variables such as differential reflectivity and specific differential phase have been shown to be useful in hydrometeor classification, an area which would help to address uncertainty within the microphysical parameterizations of the model. While challenges remain in developing well-suited forward operators

(Jung et al. 2008), such dual-polarimetric radar data hold great promise for improving the analysis of the atmospheric state. While this study was limited by the single-polarization data available from the WSR-88D network for the case study chosen, and by the limited nature of existing polarimetric forward operators, the upcoming dual-polarization WSR-88D upgrade holds great promise for future EnKF studies using polarimetric radar data.

Future improvement in the representation of microphysical processes within NWP models also holds promise for improved storm-scale EnKF data assimilation. While skillful forecasts are obtained in this study using single-moment model microphysical schemes, such schemes are suboptimal for convective-scale EnKF. Single-moment microphysical schemes assume a constant drop size distribution for rain, hail, and snow; for example, in the Lin et al. (1983) scheme, an exponential distribution is assumed, tunable by changing the intercept parameter for each hydrometeor species. In nature, however, the drop size distribution does not necessarily conform to an exponential distribution, and the intercept parameter can vary significantly within a convective system, particularly for large, linear systems such as the 8-9 May 2007 tornadic MCS studied here. Multiple-moment microphysical schemes hold promise for improving the representation of hydrometeors within the model; preliminary work by Putnam et al. (2010) suggests that using a two-moment microphysical scheme during EnKF assimilation of radar data and during the subsequent forecast can improve the quality of forecasts for this case.

The results of this study are encouraging with regard to development of a future warn-on-forecast severe weather warning system (Stensrud et al. 2009), suggesting that, at least for the case studied, it is possible to provide useful ensemble predictions of convective hazards, including areas of heavy precipitation and tornadic mesovortices, with lead times of 1 to 3 hours. We do note, however, that the robustness of such predictions still needs to be tested and calibrated over a large number of cases including a wide range of convective tornadic and non-tornadic storms, and that objective identification and verification of tornado-scale vortices is non-trivial (Potvin et al. 2009). At present, the high computational cost of performing such high-resolution predictions in real-time remains a barrier to operational implementation which will need to be addressed with algorithm and code optimizations and access to petascale computing systems. Further improvement in the prediction model and the inclusion of other available observations are also important; such work is ongoing with this and other cases. Though only a single case is used in the studies presented herein, careful case studies of this nature represent an essential first step towards full systematic testing and operational implementation of a warn-on-forecast system.

References

Aksoy, A., D. C. Dowell, and C. Snyder, 2010: A multi-case comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Short-range ensemble forecasts. *Monthly Weather Review***,** In press.

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

Baldwin, M. E., S. Lakshmivarahan, and J. S. Kain, 2001: Verification of mesoscale features in NWP models. *Preprint, 9th Conf. Mesoscale Proc.*, Fort Lauderdale, FL, Amer. Meteor. Soc., 255-257.

Bowler, N. E. and K. R. Mylne, 2009: Ensemble transform Kalman filter perturbations for a regional ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **135,** 757-766.

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

Brewster, K. A., K. W. Thomas, J. Brotzge, Y. Wang, D. Weber, and M. Xue, 2007: High resolution data assimilation of CASA X-band radar data for thunderstorm forecasting. *22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Pred.*, Salt Lake City, Utah, Amer. Meteor. Soc., CDROM 1B.1.

Brier, G. W., 1950: Verification of weather forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78,** 1-3.

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M.D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Tech.*, **12,** 5-19.

Brotzge, J., D. Westbrook, K. Brewster, K. Hondl, and M. Zink, 2005: The Meteorological Command and Control Structure of a Dynamic, Collaborative, Automated Radar Network. *21st International Conf. on Interactive Information Processing Systems (IIPS) for Meteor., Ocean., and Hydrology*, American Meteorological Society.

Brotzge, J., K. Hondle, B. Philips, L. Lemon, E. Bass, D. Rude, and D. Andra, Jr., 2010a: Evaluation of Distributed Collaborative Adaptive Sensing for detection of low-level circulations and implications for severe weather operations. *Wea. Forecasting*, **25,** 173-189.

Brotzge, J., L.R. Lemon, 2010b: An anticyclonic tornado observed 10 May 2010 with CASA radar. *25th Conf. Severe Local Storms*, Amer. Meteor. Soc., Paper P10.11.

Buizza, R., 1997: Potential Forecast Skill of Ensemble Prediction and Spread and Skill Distributions of the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **125,** 99-119.

Carbone, R. E., J. D. Tuttle, D. A. Ahijevych, and S. B. Trier, 2002: Inferences of Predictability Associated with Warm Season Precipitation Episodes. *J. Atmos. Sci.*, **59,** 2033-2056.

Chandrasekar, V., S. Lim, N. Bharadwaj, W. Li, D. McLaughlin, V. N. Bringi, and E. Gorgucci, 2004: Principles of networked weather radar operation at attenuating frequencies. *Proc. Third European Conf. on Radar Meteorology and Hydrology*, 109-114.

Crum, T. D., R. L. Albert, and D. W. Burgess, 1993: Recording, archiving, and using WSR-88D data. *Bull. Amer. Meteor. Soc.*, **74,** 645-653.

Dawson, D., 2009: Impacts of Single- and Multi-moment Microphysics on Numerical Simulations of Supercells and Tornadoes of the 3 May 1999 Oklahoma Tornado Outbreak, School of Meteorology, University of Oklahoma, 173.

Dawson, D. T., II, M. Xue, and J. A. Milbrandt, 2009: High resolution real-data simulations of the 3 May 1999 tornadic storms with multi-moment microphysics**,** Under preparation.

Dowell, D., G. Romine, and C. Snyder, 2010: Ensemble Storm Scale Data Assimilation and Prediction for Severe Convective Storms. *25th Conference on Severe Local Storms*.

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

Dowell, D. C. and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *Journal of Atmospheric and Oceanic Technology***,** 911-927. DOI: 10.1175/2008JTECHA1156.1.

Dowell, D. C., L. J. Wicker, and D. J. Stensrud, 2004b: High-resolution analyses of the 8 May 2003 Oklahoma City storm. Part II: EnKF data assimilation and forecast experiments. *Preprints, 22nd Conf. on Severe Local Storms*, Hyannis, MA,, Amer. Meteor. Soc., CDROM, 12.5.

Du, J., G. DiMego, S. Tracton, and B. Zhou, 2003: NCEP short-range ensemble forecasting (SREF) sys-tem: Multi-IC, multi-model, and multi-physics approach. Research Activities in Atmospheric and Oceanic Modelling, J. Cote, Ed., CAS/JSC Working Group Numerical Experimentation (WGNE), Rep. 33, WMO Tech. Doc. 1161, 5.09-5.10.

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

Ebert, E. E., 2008: Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework. *Meteorological Applications*, **15,** 51-64.

Edwards, R., S. F. Corfidi, R. L. Thompson, J. S. Evans, J. P. Craven, J. P. Racy, D. W. McCarthy, and M. D. Vescio, 2002: Storm Prediction Center Forecasting Issues Related to the 3 May 1999 Tornado Outbreak. *Weather and Forecasting,* **17,** 544-558.

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

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

Ge, G., J. Gao, K. Brewster, and M. Xue, 2010: Impacts of beam broadening and earth curvature on 3D variational radar data assimilation of radial velocity with two Doppler radars. *J. Atmos. Ocean Tech.*, **27,** 617-636.

Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Monthly Weather Review***,** In press.

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

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

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

Houze, R. A., Jr., S. A. Rutledge, M. I. Biggerstaff, and B. F. Smull, 1989: Interpretation of Doppler weather radar displays of midlatitude mesoscale convective systems. *Bull. Amer. Meteor. Soc.*, **70,** 607-619.

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

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

Jung, Y., G. Zhang, and M. Xue, 2008: Assimilation of simulated polarimetric radar data for a convective storm using ensemble Kalman filter. Part I: Observation operators for reflectivity and polarimetric variables. *Mon. Wea. Rev.*, **136,** 2228-2245.

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

Kong, F., K. K. Droegemeier, and N. L. Hickmon, 2006: Multi-resolution ensemble forecasts of an observed tornadic thunderstorm system, Part I: Comparison of coarse and fine-grid experiments. *Mon. Wea. Rev.*, **134,** 807-833.

Kong, F., M. Xue, K. W. Thomas, Y. Wang, K. A. Brewster, X. Wang, J. Gao, S. J. Weiss, A. J. Clark, J. S. Kain, M. C. Coniglio, and J. Du, 2011: CAPS multi-model storm-scale ensemble forecast for the NOAA HWT 2010 Spring Experiment. *24th Conf. Wea. Forecasting/20th Conf. Num. Wea. Pred.*, Amer. Meteor. Soc., Paper 457.

Kumjian, M. R. and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47,** 1940-1961.

Le Dimet, F. X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A,** 97-110.

Lei, T., M. Xue, and T. Yu, 2009: Multi-scale analysis and prediction of the 8 May 2003 Oklahoma City tornadic supercell storm assimilating radar and surface network data using EnKF. *Extended abstract, 13th Conf. of IOAS-AOLS, AMS Meetings 2008*, Paper 6.4.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102,** 409-418.

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

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

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21,** 289-307.

Majumdar, S. J., C. H. Bishop, B. J. Etherton, I. Szunyogh, and Z. Toth, 2001: Can an ensemble transform Kalman filter predict the reduction in forecast-error variance produced by targeted observations? *Quart. J. Roy. Meteor. Soc.*, **127,** 2803-2820.

Markowski, P. M., J. M. Straka, and E. N. Rasmussen, 2002: Direct surface thermodynamic observations within the rear-flank downdrafts of nontornadic and tornadic supercells. *Mon. Wea. Rev.*, **130,** 1692-1721.

Marshall, J. S. and W. M. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5,** 165-166.

Mason, S. J. and N. E. Graham, 1999: Conditional probabilities, relative operating characteristics, and relative operating levels. *Wea. Forcasting*, **14,** 713-725.

McLaughlin, D., D. Pepyne, V. Chandrasekar, B. Philips, J. Kurose, M. Zink, K. Droegemeier, S. Cruz-Pol, F. Junyent, J. Brotzge, D. Westbrook, N. Bharadwaj, Y. Wang, E. Lyons, K. Hondl, Y. Liu, E. Knapp, M. Xue, A. Hopf, K. Kloesel, A. DeFonzo, P. Kollias, K. Brewster, R. Contreras, B. Dolan, T. Djaferis, E. Insanic, S. Frasier, and F. Carr, 2009: Short-wavelength technology and the potential for distributed networks of small radar systems. *Bull. Amer. Meteor. Soc.*, **90,** 1797-1817.

Meng, Z. and F. Zhang, 2007: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part II: Imperfect Model Experiments *Mon. Wea. Rev.*, **135,** 1403–1423

Milbrandt, J. A. and M. K. Yau, 2005: A multi-moment bulk microphysics parameterization. Part I: Aanlysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62,** 3051-3064.

NRC, 2006: *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts*. The National Academies Press, 112 pp.

Potvin, C. K., A. Shapiro, T.-Y. Yu, J. Gao, and M. Xue, 2009: Using a low-order model to detect and characterize tornadoes in multiple-Doppler radar data. *Mon. Wea. Rev.*, **137,** 1230-1249.

Putnam, B. J., M. Xue, G. Zhang, Y. Jung, N. Snook, and A. D. Schenkman, 2010: Comparison of the structural evolution and polarimetric variable fields of a forecasted MCV using a single and two-moment microphysics scheme. *25th Conf. Severe Local Storms*, Amer. Metero. Soc., Paper P7.6.

Roberts, N. M. and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev*, **136,** 78-97.

Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A theory for strong long-lived squall lines. *J. Atmos. Sci.*, **45,** 463-485.

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

Schenkman, A., M. Xue, A. Shapiro, K. Brewster, and J. Gao, 2010: The analysis and prediction of the 8-9 May 2007 Oklahoma tornadic mesoscale convective system by assimilating WSR-88D and CASA radar data using 3DVAR. *Mon. Wea. Rev.*, **139,** 224-246.

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

Schwartz, C., J. Kain, S. Weiss, M. Xue, D. Bright, F. Kong, K. Thomas, J. Levit, and M. Coniglio, 2009a: Next-day convection-allowing WRF model guidance: A second look at 2 vs. 4 km grid spacing. *Mon. Wea. Rev.*, **137,** 3351-3372.

Schwartz, C. S., J. S. Kain, D. R. Bright, S. J. Weiss, M. Xue, F. Kong, J. J. Levit, M. C. Coniglio, and M. S. Wandishin, 2009b: Optimizing probabilistic high resolution ensemble guidance for hydrologic prediction. *Preprint, 23rd Conf. Hydrology*, Phoenix, AZ., Amer. Meteor. Soc., Paper 9.4.

Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, M. C. Coniglio, and M. S. Wandishin, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25,** 263-280.

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

Snook, N., M. Xue, and J. Jung, 2011: Analysis of a tornadic meoscale convective vortex assimilating CASA X-band and WSR-88D radar data using an ensemble Kalman filter. *Mon. Wea. Rev.***,** Conditionally accepted.

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

Stensrud, D. J., M. Xue, L. J. Wicker, K. E. Kelleher, M. P. Foster, J. T. Schaefer, R. S. Schneider, S. G. Benjamin, S. S. Weygandt, J. T. Ferree, and J. P. Tuell, 2009: Convective-scale Warn on Forecast System: A Vision for 2020. *Bull. Am. Meteor. Soc.*, **90,** 1487-1499.

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

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

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

Theis, S. E., A. Hense, and U. Damrath, 2005: Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteor. Appl.*, **12,** 257-268.

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

Tong, M. and M. Xue, 2005a: Simultaneous retrieval of microphysical parameters and atmospheric state variables with radar data and ensemble Kalman filter method. *Preprint, 17th Conf. Num. Wea. Pred.*, Washington DC, Amer. Meteor. Soc., CDROM P1.30.

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

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

Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74,** 2317-2330.

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

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. Second ed. Academic Press, Elsevier, Amsterdam.

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

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

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

Xue, M., M. Tong, and G. Zhang, 2009: Simultaneous state estimation and attenuation correction for thunderstorms with radar data using an ensemble Kalman filter: Tests with simulated data. *Q. J. Roy. Meteor. Soc.*, **135,** 1409-1423.

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

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

Xue, M., F. Kong, D. Weber, K. W. Thomas, Y. Wang, K. Brewster, K. K. Droegemeier, J. S. K. S. J. Weiss, D. R. Bright, M. S. Wandishin, M. C. Coniglio, and J. Du, 2007: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2007 spring experiment. *22nd Conf. Wea. Anal. Forecasting/18th Conf. Num. Wea. Pred.,* Amer. Meteor. Soc., CDROM 3B.1.

Xue, M., F. Kong, K. W. Thomas, J. Gao, Y. Wang, K. Brewster, K. K. Droegemeier, J. Kain, S. Weiss, D. Bright, M. Coniglio, and J. Du, 2008: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment. *24th Conf. Several Local Storms*, Savannah, GA, Ameri. Meteor. Soc., Paper 12.2.

Xue, M., F. Kong, K. W. Thomas, Y. Wang, K. Brewster, J. Gao, X. Wang, S. J. Weiss, A. J. Clark, J. S. Kain, M. C. Coniglio, J. Du, T. L. Jensen, and Y. H. Kuo, 2011: CAPS Realtime Storm Scale Ensemble and High Resolution Forecasts for the NOAA Hazardous Weather Testbed 2010 Spring Experiment. *24th Conf. Wea. Forecasting/20th Conf. Num. Wea. Pred.*, Amer. Meteor. Soc., Paper 9A.2.

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

Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-Resolving Hurricane Initialization and Prediction through Assimilation of Doppler

Radar Observations with an Ensemble Kalman Filter. *Monthly Weather Review*, **137,** 2105-2125.

Zhang, F., Y. Weng, Y.-H. Kuo, J. S. Whitaker, and B. Xie, 2010: Predicting typhoon Morakot's catastrophic rainfall with a convection-permitting mesoscale ensemble system. *Weather and Forecasting*, **25,** 1816-1825.