UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

FOUR-DIMENSIONAL ENSEMBLE KALMAN FILTER ASSIMILATION OF HIGH TEMPORAL RESOLUTION PHASED ARRAY RADAR DATA WITH REAL CONVECTIVE STORM CASES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

TIMOTHY A SUPINIE
Norman, Oklahoma
2020

FOUR-DIMENSIONAL ENSEMBLE KALMAN FILTER ASSIMILATION OF HIGH TEMPORAL RESOLUTION PHASED ARRAY RADAR DATA WITH REAL CONVECTIVE STORM CASES

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Ming Xue, Chair

Dr. Youngsun Jung

Dr. Alan Shapiro

Dr. Guifu Zhang

Dr. Yan Zhang

Acknowledgements

Many people have made this dissertation possible. I need to thank first my committee members, both past and present, for their guidance: my advisor, Ming Xue, and members Youngsun Jung, Alan Shapiro, Guifu Zhang, Yan Zhang, Dave Stensrud, Brian Fielder, and Sridhar Lakshmivarahan. I also need to thank CAPS scientists and collaborators Nate Snook, Derek Stratman, Bryan Putnam, Nusrat Yussouf for their support and ideas during my Ph.D. Patrick Skinner did (mind-numbing, I'm sure!) manual quality control on the PAR data used for the second part of the dissertation. My good friends Greg Blumberg, David John Gagne, and Cathy Bolene have provided moral support during the more difficult parts of the journey. My parents, Andy and Lori, provided financial support during my undergraduate years and also made me the person I am.

Finally, but most certainly not least, I need to thank my wife, Heather. She has been very patient as I've wrapped up this dissertation, which has involved me coming home late many nights of many weeks. She has also been very attentive as I've ranted about research problems and has given me encouragement to overcome these problems. Thank you, Heather, for being so supportive!

Table of Contents

List of Figures

List of Tables

Abstract

With the US National Weather Service (NWS) Weather Surveillance Radar-1988 Doppler (WSR-88D) radar network aging, a candidate for replacing it is phased array radar (PAR). The defining feature of PAR is an electronically steered beam, which allows for much faster volumetric updates than the mechanically steered beam from the WSR-88Ds. The higher temporal resolution afforded by PAR could be useful for initializing convective-scale numerical forecasts for purposes of issuing warnings based on forecasts of severe convective hazards, rather than on detection. The purpose of this work is to evaluate the performance of PAR versus the current operational WSR-88Ds in this context. Data collected during two supercell severe weather events in central Oklahoma are used in two sets of ensemble data assimilation (DA) and forecast experiments.

The first case is the 22 May 2011 Ada, Oklahoma supercell. These experiments use 40 ensemble members at 2 km grid spacing and a 4-dimensional ensemble square root filter (4D EnSRF). The Advanced Regional Prediction System (ARPS) is used for the forecast model. The WSR-88D elevations are grouped into 1-minute batches to compare to the 1-minute PAR volumes. For this case, after 30 minutes of DA, the experiment using PAR data outperformed the experiment with WSR-88D data in the placement and rough intensity of the mesocyclone track and using several objective performance metrics, such as equitable threat score (ETS) and area under the relative operator characteristic (ROC) curve (AUC). After 45 minutes of DA, the advantage of PAR over the WSR-88D is reduced, suggesting that PAR data are most beneficial when the DA period must be short, perhaps to quickly spin up a storm in the ensemble. This would afford a longer lead time for a severe weather warning. A supplemental experiment using

PAR data thinned in elevation angles to imitate the WSR-88D data batches confirms that the advantage is due to the extra data volumes provided by the PAR.

The second case is the 31 May 2013 El Reno, Oklahoma supercell. The data processing for the WSR-88D and 1-minute PAR data is nearly identical to the 21 May 2011 case. For these experiments, the 4D EnSRF is used with the Weather Resesarch and Forecasting (WRF) forecast model. 36 members at 1 km grid spacing are used. In this case, the PAR experiment results in stronger mesocyclones because of a stronger mid-level temperature perturbation. Examination of the analysis increments by PAR volume reveals that in the default experiment, only the first 3 PAR volumes provide significant changes to the ensemble state. This is because the reflectivity innovation is reduced by having updated previous volumes, and an experiment with additional covariance inflation between PAR volumes is run that confirms this. In order to alleviate this, experiments are conducted assimilating the radar volumes in different orders, such as reversed, starting from the analysis time and working towards the edges of the DA window ("inside out") and the reverse of that ("outside in"). The results suggest that the different orders can improve track forecasts due to warmer temperatures in the analyzed cold pool near the mesocyclone. Additionally, colder forecast cold pools were found to be due to drier analyses in the near-storm inflow. The colder forecast cold pools also contributed to track displacement errors.

Chapter 1: Introduction

Numerical Weather Prediction (NWP) is an initial-condition problem: the outcome of a forecast is heavily dependent on the initial conditions. As such, the creation of an accurate forecast depends greatly on the accuracy of the initial conditions. Numerous methods have been used to create initial conditions for forecasts. Early numerical forecasts used several methods, including interpolation from raw observations (Richardson 1922), interpolation from manual analyses (Charney et al. 1950), and Cressman analysis (Cressman 1959), which uses the observations to make corrections to a first guess or "background" field. Variational methods (Sasaki 1970) attempt to minimize a cost function that includes a term for the difference between the analysis and the observations and a term for the estimated error in the background field, but may have more terms as needed.

The major disadvantage of the above methods is that the notion of background error does not depend on meteorological context, meaning the increments provided by the data assimilation method do not depend on meteorological context. Additionally, it is difficult to prescribe climatological static background error covariance between different fields, such as temperature and vertical velocity, when such error covariance typically varies strongly in space and time. One method to address this is the Kalman filter (Kalman 1960), which in general is a method for estimating the value of a random process with only noisy observations of that process. This has many applications in engineering and geosciences, and it can be applied to an ensemble of NWP forecasts, which is used to estimate the background error covariance.

One common ensemble data assimilation method is the ensemble Kalman filter (EnKF; Evensen 2003) and its variants. The basic formulation common to most variants is

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}[\mathbf{y}_\mathrm{o} - H(\mathbf{x}_b)], \tag{1}$$

where $\mathbf{x}_a$ is the analysis ensemble state vector, $\mathbf{x}_b$ is the background ensemble state vector, $\mathbf{y}_o$ is the observation vector, $\mathbf{K}$ is the Kalman gain matrix, and $H(\cdot)$ denotes the forward operator responsible for projecting the background state into observation space. An alternate formulation is

$$\bar{\mathbf{x}}_a = \bar{\mathbf{x}}_b + \mathbf{K}[\mathbf{y}_o - \mathbf{H}\bar{\mathbf{x}}_b], \tag{2}$$

$$\mathbf{x'}_a = \mathbf{x'}_b + \widetilde{\mathbf{K}}[\mathbf{y'}_o - \mathbf{H}\mathbf{x'}_b], \tag{3}$$

where $\bar{\mathbf{x}}_a$ and $\mathbf{x'}_a$ are the analysis ensemble mean and perturbation (i.e. the difference from the ensemble mean; $\mathbf{x'} = \mathbf{x} - \bar{\mathbf{x}}$), $\bar{\mathbf{x}}_b$ and $\mathbf{x'}_b$ are the background ensemble mean and perturbation, $\mathbf{H}$ is the linearized version of the forward operator, $\widetilde{\mathbf{K}}$ is the Kalman gain matrix used for updating the perturbations (in the regular EnKF, $\widetilde{\mathbf{K}} = \mathbf{K}$), and $\mathbf{y'}_o$ represents perturbations to observations (in the regular EnKF, $\mathbf{y'}_o = \mathbf{0}$). This formulation is useful because it gives a direct update equation for the ensemble mean, and ensemble DA methods are formulated such that the ensemble mean represents the most likely state of the atmosphere.

The Kalman gain matrix is calculated by

$$\mathbf{K} = \mathbf{P}_b\mathbf{H}^\mathrm{T}(\mathbf{H}\mathbf{P}_b\mathbf{H}^\mathrm{T} + \mathbf{R})^{-1}, \tag{4}$$

where $\mathbf{P}_b$ is the background error covariance matrix and $\mathbf{R}$ is the observation error covariance matrix. In this formulation, $\mathbf{H}\mathbf{P}_b\mathbf{H}^\mathrm{T}$ is the background error projected into observation space and is often implemented as the covariance matrix of observation priors. Additionally, $\mathbf{P}_b\mathbf{H}^\mathrm{T}$ is related to the covariance between the observation priors at the observation locations and the model

grid points. Thus, at its core, the EnKF estimates the covariance between the observed quantity and the model state variables and uses that information to update the ensemble. The covariance estimates are dependent on the state of the ensemble ("flow-dependent").

Because the number of ensemble members is finite, there is some error in sampling the background error distribution and therefore in the estimates of covariance (this is known as "sampling error"; Gasperoni and Wang 2015). This sampling error can produce spurious covariances between a given observation and grid points that are far away, which can degrade the quality of the analysis (Houtekamer and Mitchell 1998, 2001). In the limit of infinite ensemble members, this error would go to zero, but having large numbers of ensemble members is impractical. A more practical workaround is to localize the covariance such that an observation can only affect grid points nearby. Houtekamer and Mitchell (2001) introduced formal notation for covariance localization:

$$\mathbf{K} = \rho \circ (\mathbf{P}_b \mathbf{H}^{\mathrm{T}})[\rho \circ (\mathbf{H}\mathbf{P}_b \mathbf{H}^{\mathrm{T}}) + \mathbf{R}]^{-1}, \quad (5)$$

where ρ denotes a localization function whose output is between 0 and 1, and the $\circ$ operator denotes the Schur product, or elementwise multiplication of a matrix. Often, the function of Gaspari and Cohn (1999), which is a piecewise 5th order polynomial with local support, is used for the localization function. It has a single tunable parameter: the localization radius and can be applied to produce spherical or oblate spheroidal localization in three-dimensional models. This method is widely used because of its simplicity, however prescribing a fixed region that each observation can update slightly defeats the advantage of using flow-dependent background error covariance. For this reason, many adaptive localization methods have been proposed which

attempt to use the ensemble to define what the localization should be [see the review by Gasperoni and Wang (2015)].

Another problem with the EnKF is that by itself, it tends to result in a posterior distribution that is too narrow. This can result in a condition known as "filter divergence," where subsequent applications of the DA algorithm ignore the observations, and the DA becomes ineffective. A few techniques have been used to increase the posterior spread. One potential technique is to perturb the observations by drawing from a distribution of observation errors (i.e. making $\mathbf{y'}_o \neq \mathbf{0}$ in Eqn. $\mathbf{x'}_a = \mathbf{x'}_b + \widetilde{\mathbf{K}}[\mathbf{y'}_o - \mathbf{H}\mathbf{x'}_b]$, (3); Burgers et al. 1998). This is often assumed to be a Gaussian distribution with mean 0 and standard deviation equal to the assumed observation error (Houtekamer and Mitchell 1998). However, this method exacerbates the error in sampling error described above and introduces noise into the DA process. Another potential technique is to reformulate the Kalman gain (Eqn. $\mathbf{K} = \mathbf{P}_b\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_b\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$, (4) to better match the expected spread in the posterior distribution (Whitaker and Hamill 2002). Whitaker and Hamill (2002) set $\mathbf{y'}_o = \mathbf{0}$ and define $\widetilde{\mathbf{K}} = \alpha\mathbf{K}$, in Eqn. $\mathbf{x'}_a = \mathbf{x'}_b + \widetilde{\mathbf{K}}[\mathbf{y'}_o - \mathbf{H}\mathbf{x'}_b]$, (3, where

$$\alpha = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R}}}\right)^{-1}. \quad (6)$$

Because the Kalman gain matrix contains a square root, their formulation is known as the "ensemble square root filter" (EnSRF). Additionally, assimilating multiple observations at the same time, which is possible in the original EnKF formulation, would require finding matrix square roots; for this reason, the EnSRF assimilates observations sequentially, which is much less computationally expensive.

Another way to combat the narrow posterior distribution is to artificially inflate its variance. This process is known as "covariance inflation," and there are several methods to do it. One is a relaxation to the prior state of the ensemble, known as "relaxation to prior perturbation" [RTPP; introduced by Zhang et al. (2004) and named by Whitaker and Hamill (2012)], which works because the prior state generally has greater spread than the posterior state. Other methods are based on multiplying the ensemble perturbations by a value greater than 1 to increase the variance. One method (known simply as "multiplicative inflation"; Anderson 2001) multiplies the ensemble perturbations by a constant value everywhere in the domain. This method can be applied to either the prior ensemble or the posterior ensemble. Another method multiplies the posterior ensemble perturbations by a value that is proportional to the amount by which spread was decreased during the DA process. This known as "relaxation to prior spread" (RTPS; Whitaker and Hamill 2012). Yet another method is the adaptive inflation method (Anderson 2007, 2009), which multiplies the prior ensemble perturbations by a spatially-varying factor. This spatially varying factor is updated by the observations during the DA process and optionally evolved forward in time by the model during the forecast step.

The above methods implicitly assume that all the observations considered in the DA are valid at the analysis time. In some problems, this may not be the case, or one may wish to analyze observations valid at multiple times. For this, four-dimensional (4D) EnKF-based methods were developed (Sakov et al. 2010; Hunt et al. 2007; Wang et al. 2013a). While 3D EnKF methods consider covariances only in space, 4D EnKF methods consider covariances in time as well as in space. The time period over which observations are used for a single analysis is usually referred to as the "DA window." Using temporal covariances means the ensemble mean is the most likely

state of the atmosphere assuming the atmosphere evolves approximately linearly in time over the DA window. To illustrate, Wang et al. (2013a) developed a 4D EnSRF method and ran observing system simulation experiments (OSSEs) to test it. Their results indicated that the 4D EnSRF outperformed the 3D EnSRF, particularly for long DA windows. They also found that a 5- to 10-minute DA window and cycle length is optimal for their test.

As described above, these methods readily provide initial conditions for ensemble forecasts. Ensemble forecasts are essential for convective scale prediction, such as that in the Warn-on-Forecast project (Stensrud et al. 2009). The goal of the Warn-on-Forecast project is to provide warnings for severe convective storms and their associated hazards based on forecasts of those hazards, rather than detection. To produce direct forecasts of convective scale weather, one must provide observation information on the convective scale. Radar is a common source of convective-scale observations.

A complication with using radar observations is that the forward operator for radar reflectivity is non-trivial to formulate. In the language of data assimilation, the role of the forward operator is to project the model state into the observation space to obtain a background estimate of the observed quantity. Practically, this means the forward operator must compute reflectivity from the model state, which is a non-linear operation, and the exact formulation depends on how the forecast model represents hydrometeors (commonly called the "microphysics"). However, some methods have been developed. Tong and Xue (2005) provided forward operators for reflectivity for single-moment ice microphysics (for example, Lin et al. 1983), and Jung et al. (2008) introduced forward operators for reflectivity and polarimetric variables suitable for multi-moment microphysics (Putnam et al. 2014).

Despite the difficulties assimilating radar data, many radar DA and forecast experiments have been performed with 3D EnKF-based methods. Early work on storm scale EnKF DA includes Snyder and Zhang (2003), who demonstrated assimilation of simulated Doppler velocity observations in an OSSE framework. Dowell et al. (2004) ran DA experiments using real reflectivity and radial velocity observations, though their forecast model used a warm-rain microphysics scheme, which simplified the reflectivity forward operator at the cost of being physically unrealistic. Because in a real-data experiment truth is unknown, Dowell et al. (2004) and Dowell and Wicker (2009) also introduced diagnostics such as the consistency ratio for evaluating the performance of the EnKF in real-data experiments. Dowell and Wicker (2009) also tested an additive noise method for maintaining spread in a storm-scale ensemble. Additionally, one study which focused on physical processes in supercells is Marquis et al. (2012), who created EnKF analyses of mobile Doppler radar data to investigate vorticity generation in supercell thunderstorms. Also, Dawson et al. (2012) used EnKF analyses of radar data to investigate the impact of the low-level wind field on forecasts of a supercell thunderstorm. Labriola et al. (2017, 2019) used EnSRF analyses to investigate forecasts of hail from various microphysics schemes. Snook et al. (2011, 2015) used EnSRF to show that the use of localized gap-filler X-band radars improved analyses and forecasts of a tornado-producing mesoscale convective system. The NSSL Experimental Warn-on-Forecast System for ensembles (NEWS-e; Wheatley et al. 2015; Jones et al. 2016), an on-demand 3 km ensemble forecast system used a flavor of EnKF called the ensemble adjustment Kalman filter (EAKF) to initialize forecasts, and its successor Warn-on-Forecast System (WoFS; Yussouf and Knopfmeier 2019; Stratman et al. 2020) uses the EnSRF for the same purpose.

Most of the above work uses radar volume scans that are ~5 minutes apart, which is comparable to the operational Weather Surveillance Radar-1988 Doppler (WSR-88D) network. However, the WSR-88D network is aging, and one candidate to replace the network is phased-array radar (PAR). Instead of a mechanically steered beam in the WSR-88D radars, the PAR uses an electronically steered beam, which allows for volumetric update times of about 1 minute, compared to the ~5-minute update times for WSR-88D. This is very useful for forecasters to maintain awareness of rapidly evolving storms and to provide data for warning on short-lived tornadoes that might occur entirely between volume scans of the WSR-88Ds. Additionally, the shorter time between volumetric updates makes it an ideal candidate for use of 4D methods, as frequent starting and stopping of the model when using 3D EnKF methods has been shown to cause imbalances in the model and degrade the quality of the forecast (Wang et al. 2013a). In contrast, with 4D EnKF methods, multiple radar scans can be assimilated in a single step.

Comparatively few radar DA and forecast experiments have been performed with 4D EnKF-based methods in real convective-storm cases. Stratman et al. (2020) included one experiment using the 4D EnSRF in a real-data DA and forecast experiment using the 31 May 2013 El Reno tornadic supercell case. They found that, in contrast to Wang et al. (2013a), assimilating PAR volumes every 1 minute using a 3D EnKF produces a better forecast than assimilating 1-minute volumes and producing analyses every 5 minutes using a 4D EnSRF. They attributed this to suboptimal covariance inflation in the 4D EnSRF.

Because few studies have examined the performance of PAR, it is unknown whether and how much the PAR data provide a benefit over the current operational WSR-88D data when used for convective-scale DA. The goal of this work is to evaluate this benefit. This will be done using

DA and forecast experiments from two cases: The 21 May 2011 supercell near Ada, Oklahoma (OK), and the 31 May 2013 supercell near El Reno, OK. Chapter 2 will provide a meteorological overview of the two cases. Chapter 3 will discuss experiments with the 22 May 2011 case, and chapter 4 will discuss the 31 May 2013 case. Chapter 5 will provide a discussion and conclusions from both cases.

Chapter 2: Cases Studied

Two cases are examined using the PAR data. The first is the 22 May 2011 Ada, Oklahoma, supercell. This was a weakly tornadic event with a tornado rating of EF0 of the Enhanced Fujita (EF) scale. The tornadoes started at ~0119 UTC and ended at ~0141 UTC (based on tornado reports, Table 1). Several operational WSR-88D radars and the NWRT PAR documented the life cycle of this supercell storm. During the late afternoon and evening of 21 May 2011, a dryline extended from the Big Bend area of Texas (southwest Texas) through central Oklahoma and into south-central Kansas where it intersected the Pacific front. A midlevel jet rotating through the base of upper low increased the effective shear. The bulk wind difference through 0-6 km AGL in the 0000 UTC 22 May 2011 Norman, Oklahoma, sounding was 43 kts (22.1 m s^{-1}); values greater than 40 kts (20.6 m s^{-1}) favor supercell storms.

Table 1: Tornado reports associated with storm B on 22 May 2011 obtained from Storm Prediction Center (http://www.spc.noaa.gov/climo/reports/110521_rpts.html)

Time (UTC)	Location	County in Oklahoma	Lat	Lon
0119	3.22 km NW Vanoss	Pontotoc	34.79°	-96.89°
0130	12.88 km W Ada	Pontotoc	34.78°	-96.80°
0134	8.05 km W Ada	Pontotoc	34.78°	-96.75°
0141	6.44 km WNW Ada	Pontotoc	34.80°	-96.73°

Convection initiated along the dryline in south-central Oklahoma and by 0000 UTC on 22 May, a storm (storm A) evolved into a supercell in south-central Oklahoma (Figure 1a). Several smaller cells developed north of storm A (Figure 1a-c), and the one (storm B) closest to storm A matured into a supercell with a hook echo feature by 0045 UTC (Figure 1d). There was one tornado report associated with storm A at 00:40 UTC at Murray county, Oklahoma (SPC storm report database). During the next 30 min, storm B grew stronger and storm A quickly dissipated

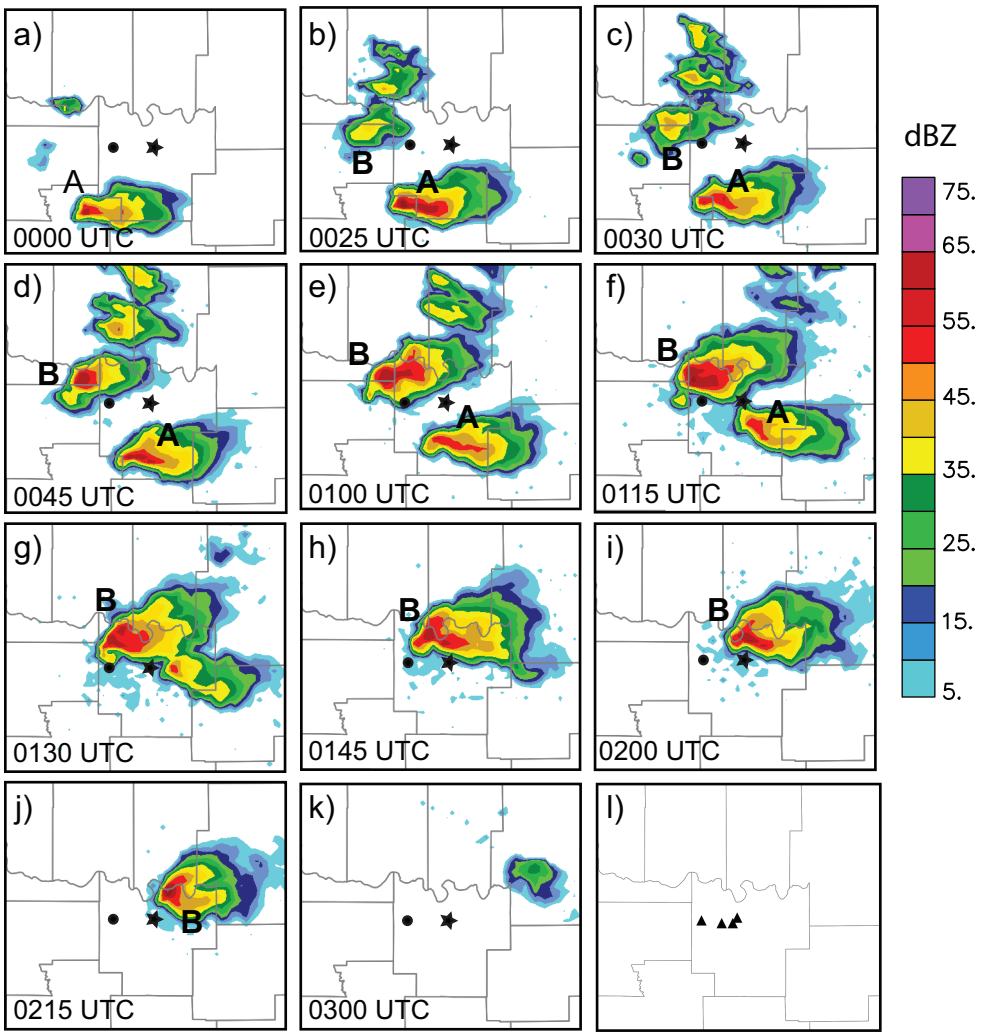

Figure 1: Observed reflectivity (dBZ, from the NSSL National Mosaic and Multi-Sensor system) at 2 km MSL from (a) 0000 UTC, (b) 0025 UTC, (c) 0030 UTC, (d) 0045 UTC, (e) 0100 UTC, (f) 0115 UTC, (g) 0130 UTC, (h) 0145 UTC, (i) 0200 UTC, (j) 0215 UTC, and (k) 0215 UTC 22 May 2011 over the area of interest. The black circle indicates the location of Vanoss, Oklahoma and black star indicates the location of Ada, Oklahoma. The black triangles in (l) are the four tornado reports from storm B (Table 1).

when it encountered storm B (Figure 1e, f, g, h). There were four tornado reports (Table 1) associated with storm B, all of which were rated 0 on the EF scale. The first tornado report was at 0119 UTC located 3.2 km northwest of Vanoss, Oklahoma. The supercell continued east towards Ada, Oklahoma. The last tornado report from storm B was at 0141 UTC located 6.44 km west-northwest of Ada, Oklahoma. There is no damage track available for this event. The storm starts to dissipate after 0200 UTC (Figure 1i, j, k). The storms were sampled both by WRD-88D

radars and by the NWRT PAR. The observations from both radar platforms are assimilated separately into a storm-scale ensemble forecast system in order to compare the ability of the radars in forecasting the storm and its associated low-level mesocyclone.

The second case is the 31 May 2013 El Reno, Oklahoma supercell. This event included several tornadoes, the strongest of which was rated EF3 and tracked near El Reno, Oklahoma. The synoptic situation included a broad upper-level low centered over the northern Great Plains. A stalled front was draped across northern and western Oklahoma, and a dryline extended across southwestern Oklahoma, intersecting the front ~75 km west of El Reno. The warm side of these boundaries was characterized by very moist and unstable air with 100-mb mixed-layer convective available potential energy (MLCAPE) values over 4000 J kg^{-1} and 2-m dewpoint temperatures at or above 72 °F (22.2 °C). The 0-6 km AGL bulk wind difference was 56 kts (28.8 m s^{-1}) on the 00Z 1 Jun sounding from Norman, Oklahoma, well above the threshold for supercells.

The aforementioned boundaries provided a lifting mechanism to initiate convection between 2130 and 2230 UTC, with the cell on the southern end of the line developing supercell characteristics west of El Reno (Figure 2a-d). At 2255 UTC, the storm produced a brief EF0 tornado, and at 2302 UTC, the large EF3 tornado developed west of El Reno (Bluestein et al. 2019). This tornado ended at 2344 UTC, and this supercell went on to produce several more tornadoes over the next hour (Bluestein et al. 2015). One of the more notable tornadoes was an EF1 that occurred from 2351-0009 UTC in western Oklahoma City. During this time period, convection continued to develop to the west of the initial supercell, some of which briefly had supercell characteristics (Figure 2h-k). Eventually, after 0000 UTC 1 June, the convection grew

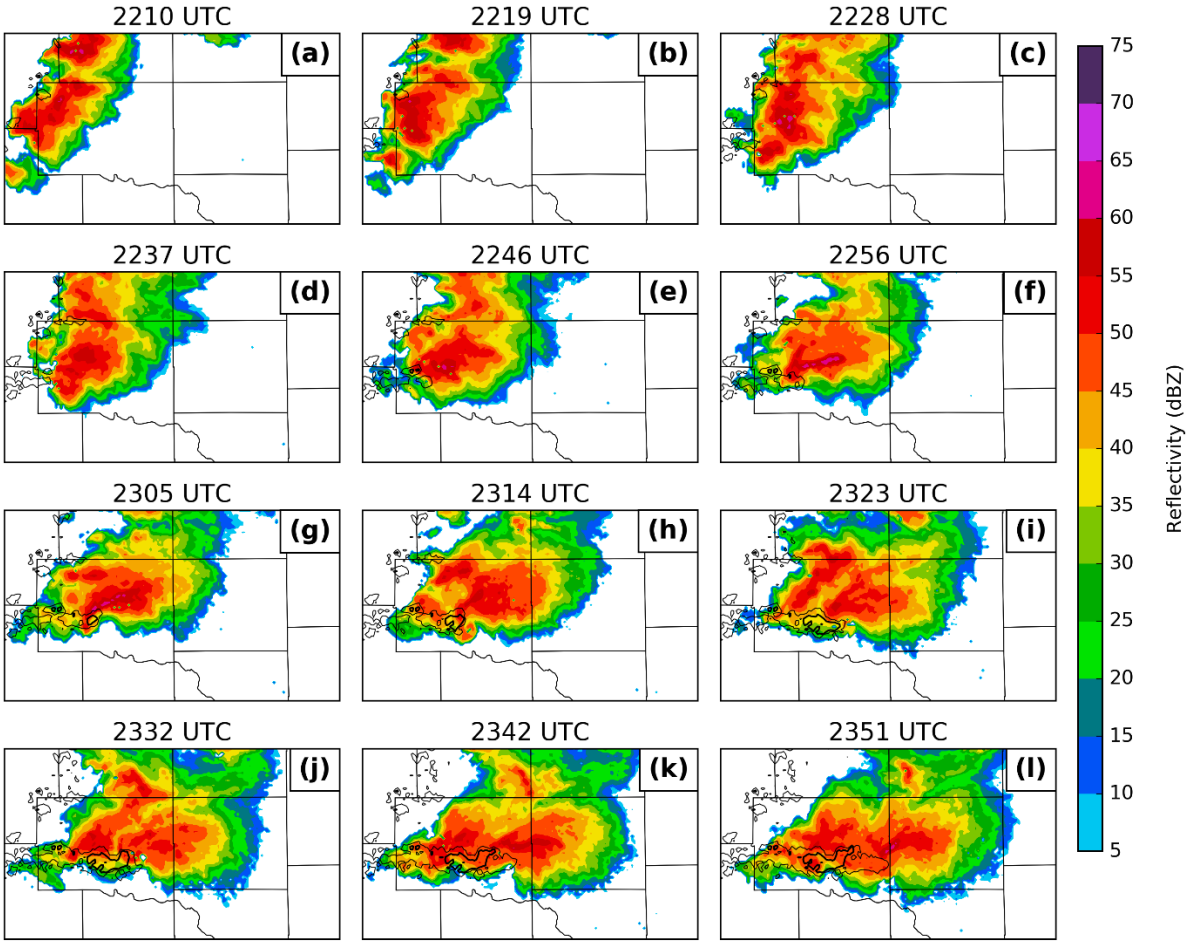

Figure 2: Observed base reflectivity from the KTLX radar, valid at 2210 (a), 2219 (b), 2228 (c), 2237 (d), 2246 (e), 2256 (f), 2305 (g), 2314 (h), 2323 (i), 2332 (j), 2342 (k), and 2351 (l) UTC. The 0.006 and 0.012 s^{-1} contours from MRMS azimuthal shear track from 2200 UTC to the valid time in each panel are given in black.

into a backbuilding mesoscale convective system, resulting in a major flash flooding event in the

Oklahoma City metropolitan area.

Chapter 3: Analysis of the 22 May 2011 Case[1]

3.1. Data and Methods

3.1.1. Data Used

The WSR-88D radar used is the operational radar at Twin Lakes, Oklahoma (KTLX). The storm associated with the Ada tornado was ~80 km away from this radar. The WSR-88D radar sampled the storm at ~4.2-min intervals with a volume coverage pattern that consisted of 14 tilts (elevation angles) ranging from 0.46^o to 19.45^o (Table 2). The level II reflectivity and radial velocity observations from the WSR-88D KTLX are quality controlled using an automatic procedure within the ARPS package to eliminate noise and ground clutter from the reflectivity observations and to unfold the aliased radial velocity observations (Brewster et al. 2005). The quality controlled reflectivity and radial velocity observations from both radars are interpolated to the model grid horizontally, but remain on the radar elevation levels vertically; this procedure is used in many real data studies with the ARPS EnKF DA system (e.g., Jung et al. 2012; Snook et al. 2012).

The NWRT PAR sampled the evolution of the Ada supercell with a 90^o sector at a rapid ~50 sec volume updates. The volume coverage pattern spanned tilts from 0.5^o to 60^o with higher number of tilts at lower levels of the atmosphere. In this study, the lowest 14 elevation angles from 0.5° to 19.5° are assimilated (Table 2) into the model during the 45-min DA period. However, due to mechanical failures, several volume scans of NWRT PAR observations are missing between 0045 and 0108 UTC (listed in Table 2) and therefore not assimilated in the experiments. The

[1] This chapter was published as sections 3, 4, and 5 of Supinie et al. (2017)

NWRT PAR observations are quality controlled using the same ARPS quality control package and interpolated in the same manner as the WSR-88D observations.

Table 2: List of experiments conducted and the number of volume scans of radar observations assimilated during the 45-min data assimilation period from 0030 through 0115 UTC 22 May 2011.

Experiment Names	**WSR-88D**	**PAR**	**PAR-reducedtilts**
Tilts (elevation angles in degrees) in a complete volume scan	0.46, 0.88, 11.25, 1.75, 2.35, 3.08, 3.98, 5.11, 6.35, 7.97, 10.00, 12.46, 15.57, 19.45	0.50, 0.90, 1.30, 1.80, 2.40, 3.10, 4.00, 5.10, 6.40, 8.00, 10.00, 12.50, 15.60, 19.50	0.50, 0.90, 1.30, 1.80, 2.40, 3.10, 4.00, 5.10, 6.40, 8.00, 10.00, 12.50, 15.60, 19.50
Total number of volume scans in 15-min assimilation period (valid 0030—0045 UTC)	5	17 (0045—0047 UTC scans missing)	4 (0045—0047 UTC scans missing)
Total number of volume scans in 30-min assimilation period (valid 0030—0100 UTC)	8	26 (0045—0051, 0053—0055 UTC scans missing)	6 (0045—0051, 0053—0055 UTC scans missing)
Total number of volume scans in 45-min assimilation period (valid 0030—0115 UTC)	12	38 (0045—0051, 0053—0055, 0103, 0107, and 0108 UTC scans missing)	10 (0045—0051, 0053—0055, 0103, 0107, and 0108 UTC scans missing)

There are a few caveats to using the PAR data. The first is that the scan sector is free to change between volume scans; this occurs twice for this case. The first and most substantial change occurs between the 0045 and 0050 UTC analyses, when the sector shifts approximately 21 degrees counterclockwise. The second sector change is between the 0050 and 0055 UTC analyses; the scan sector shifts a further 2 degrees counterclockwise. The effect of a sector change is that storms that were outside of the radar coverage area are now observed by the radar, which affects the observation space statistics. This will be discussed in more detail in later

sections. In addition, because the PAR scans over a limited sector in azimuth, there is no way to suppress spurious convection outside the scan sector from the PAR. This can result in spurious convection interacting with the assimilated storms and degrading the quality of the forecast. To mitigate this, the zero-reflectivity observations from the most recent KTLX volume are used outside the PAR scan sector in the following manner: if the composite reflectivity from KTLX is less than or equal to 0 dBZ, the corresponding column of PAR observations is set to 0 dBZ. Otherwise, the entire column of PAR observations is set to "missing."

3.1.2. Experiment Design

The storm-scale radar DA experiments are performed on a grid with 2 km grid spacing, nested within a mesoscale grid of 18 km grid spacing that covers the continental United States (CONUS, Figure 3a). The Advanced Regional Prediction System (ARPS; Xue et al. 2000, 2001, 2003) is used as the prediction model for the DA and forecast experiments while the EnSRF system developed for the ARPS system (Xue et al. 2006; Wang et al. 2013b; Jung et al. 2012; Snook et al. 2012, 2015; Putnam et al. 2014) and extended to the 4DEnSRF algorithm (Wang et al. 2013a) is used for the ensemble DA.

The mesoscale domain contains 243×163×51 grid points over the continental United States (Figure 3a). The vertical grids are stretched with a vertical spacing of 50 m near the surface and 900 m at the model top. The model top is at 25 km with a Rayleigh damping layer above 12 km. A 40-member mesoscale ensemble is initialized from the 20-member National Centers for Environmental Prediction (NCEP) Short-Range Ensemble Forecasting (SREF; Du et al. 2006) analyses (at 40 km horizontal grid spacing) at 2100 UTC 21 May 2011, but using two sets of physics parameterizations (Table 3). The schemes used in the first set of 20 ensemble members are the

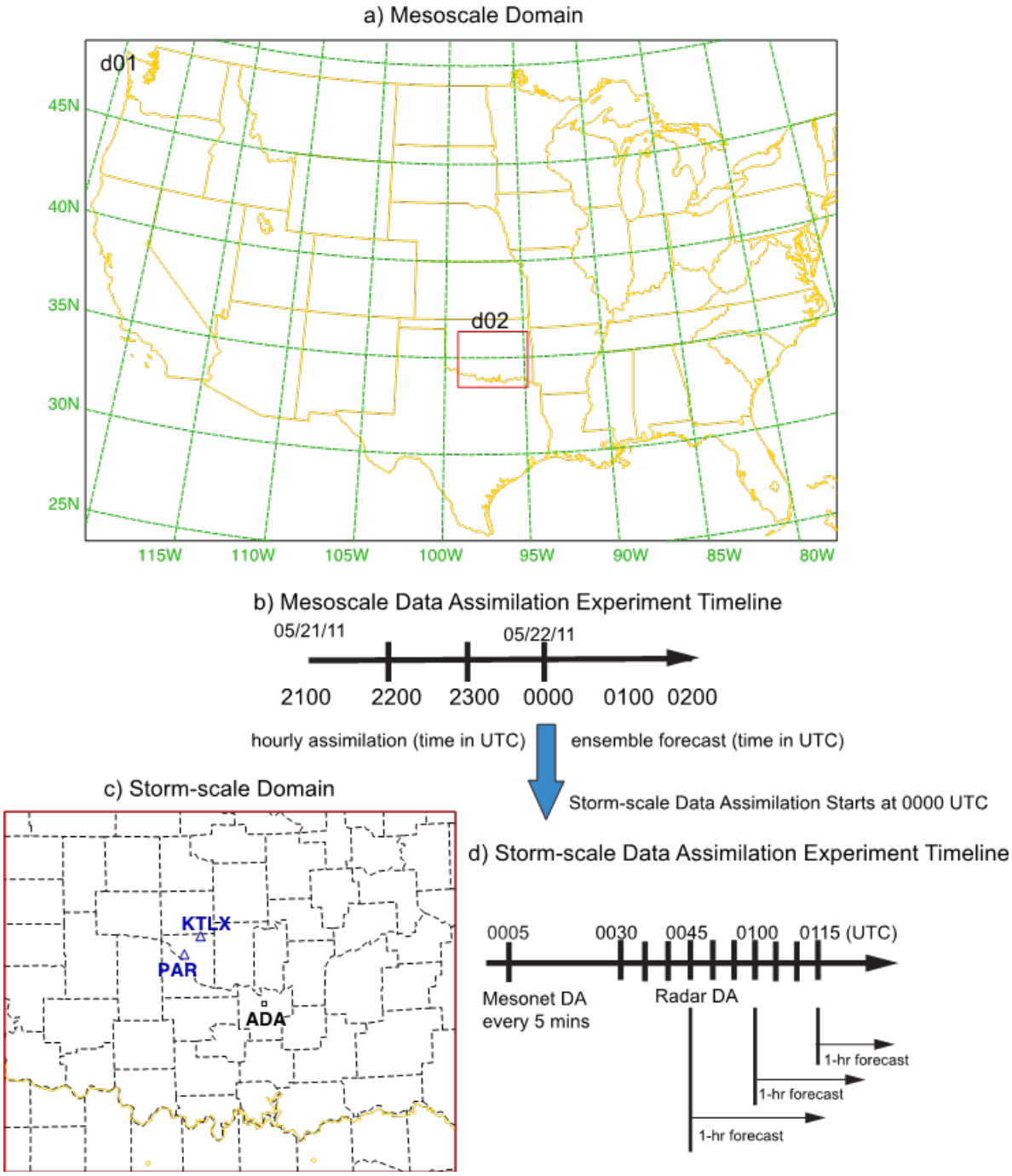

Figure 3: (a) The mesoscale domain (d01) covering the CONUS, the nested storm-scale domain (d02), (b) the timeline of the mesoscale data-assimilation experiments, (c) The storm-scale domain with county borders (d02, enlarged), location of WSR-88D (KTLX, blue triangle) and PAR (blue triangle), location of Ada, Oklahoma (black square) and the (d) timeline of the storm-scale data-assimilation and forecast experiments.

Betts-Miller-Janjic cumulus parameterization scheme (Betts 1973) and the Sun and Chang TKE-based planetary boundary layer (PBL; Sun and Chang 1986; Xue et al. 1996) scheme. The

remaining 20 ensemble members use the Kain-Fritsch cumulus parameterization scheme for (Kain and Fritsch 1990, 1993) and the Yonsei University (YSU) PBL scheme (Hong and Pan 1996). In addition, the single-moment Lin (Lin et al. 1983) microphysics scheme, stability-dependent surface-layer physics, a two-layer soil model initialized from NCEP's Eta Model analysis, and the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) (Chou 1990, 1992; Chou and Suarez 1994) longwave and short wave radiation schemes are used in all 40 ensemble members except for differences in the transmission functions used. The transmission function in the radiation scheme is computed using k-distribution method for the first 20 members and using a look-up table for the remaining 20 ensemble members. The physics diversity in the mesoscale ensemble is used to represent model error in the ensemble system and to help account for ensemble under dispersion in the assimilation system (Snook et al. 2012; Yussouf and Stensrud 2012). This mesoscale ensemble is used to provide initial and boundary conditions for the one-way nested storm-scale ensemble system.

Routinely available surface (i.e., horizontal u and v wind components, temperature T and dewpoint T_d), wind profiler (u and v) and radiosondes (u, v, potential temperature ϑ and T_d; only available at 0000 UTC 22 May 2011) observations are assimilated in the mesoscale domain every 1 h starting at 2200 UTC 21 May 2011 out to 0000 UTC 22 May 2011 (Fig. 2b) using the ARPS EnSRF (Xue et al. 2006; Wang et al. 2013b) system. Radar data are not assimilated on the mesoscale grid. The covariance localization function is based on the 5th order correlation function from Gaspari and Cohn (1999) with a horizontal cutoff radius of 300 km and 800 km used for the surface and upper air (both wind profiler and radiosonde) observations, respectively. A 6 km cutoff radius in the vertical is used for all observations (Snook et al. 2015). To help

maintain the ensemble spread, multiplicative inflation (Anderson 2001) with a factor of 5% is applied everywhere in the model domain to all model state variables. The mesoscale ensemble analyses at 0000 UTC are used to create the initial conditions for the storm-scale ensemble. Meanwhile, 2-h ensemble forecasts are generated from the mesoscale ensemble analyses at 0000 UTC 22 May 2011 out to 0200 UTC 22 May 2011 to provide ensemble boundary conditions for the storm-scale ensemble.

Table 3: Physics options for the 40-member ARPS mesoscale and storm-scale ensemble system.

Parameterization schemes	**Mesoscale Ensemble Members 1-20**	**Mesoscale Ensemble Members 21-40**	**Storm-scale Ensemble Members 1-40**
Cumulus	Betts-Miller-Janjic	Kain-Fritsch	N/A
Planetary Boundary Layer	Sun and Chang	Yonsei University	Yonsei University
Longwave and shortwave radiation	NASA GSFC: transmission functions are computed using the k-distribution method	NASA GSFC: transmission functions are computed using look-up table	NASA GSFC: transmission functions are computed using the k-distribution method
Cloud microphysical scheme	Lin scheme	Lin scheme	Milbrandt-Yau 2-moment scheme

The physics packages used by the ARPS on the 2-km nested grid are the same as the first 20 members of the mesoscale ensemble, except that the double-moment Milbrandt and Yau (2006a,b; MY hereafter) microphysics scheme and the YSU PBL scheme are used, while the cumulus parameterization scheme is turned off (Table 3).

A 40-member 2-km storm-scale ensemble is downscaled from the mesoscale ensemble analyses at 0000 UTC 22 May 2011 in a one-way nested configuration. The domain is centered on Ada, Oklahoma and covers most of Oklahoma and northern part of Texas with 203 × 163 × 51

grid points (Figure 3c). To introduce smaller scale perturbations into the convective-scale ensemble, samples from a Gaussian distribution with zero mean are drawn at each grid point in the domain. Then, a two-dimensional recursive filter (Jung et al. 2012) with a decorrelation length scale of 6 km is applied to these samples horizontally, and a homogeneous Gaussian filter with a decorrelation length scale of 4 km is applied vertically. This generates smoothed, spatially correlated perturbations, and these are added to the interpolated mesoscale ensemble at 0000 UTC. The smoothed perturbations have zero mean and standard deviations of 2 m s^{-1} for the horizontal velocity components (u and v) and 1 K for potential temperature (ϑ) over the entire domain. The perturbation standard deviations of the mixing ratios for water vapor (q_v), cloud water (q_c), rain water (q_r), ice (q_i), snow (q_s), hail (q_h), and graupel (q_g) are set to 10% of their values at each grid point where the observed reflectivity is larger than 10 dBZ. Only positive perturbations are retained for ϑ, q_v, q_c, q_r, q_i, q_s, q_h and q_g. The effect of the recursive filter is similar to but computationally more efficient than that used in (Tong and Xue 2008).

Oklahoma Mesonet (Brock et al. 1995; McPherson et al. 2007) observations (i.e., 10-m u, v, 2-m temperature T and relative humidity r_h) are assimilated every 5-min into the storm-scale ensemble starting from 0005 UTC and ending at 0020 UTC using 3D EnSRF (Figure 3d) DA. A cutoff radius of 50 km is used in the horizontal and 6 km in the vertical for covariance localization [as in Snook et al. (2015), similar to Sobash and Stensrud (2015)]. A 15% multiplicative inflation and relaxation-to-prior spread (RTPS; Whitaker and Hamill 2012) with a relaxation factor of 0.85 are applied over the whole domain to all model variables to help maintain ensemble spread during the period of Mesonet data assimilation. These values are selected based on several sensitivity studies (not shown).

The radar reflectivity and radial velocity observations from the WSR-88D or PAR are assimilated into the storm-scale ensemble starting at 0030 UTC using the 4DEnSRF (Figure 3d), and the assimilation continues until 0115 UTC with 5-min assimilation window. In order to show the effect of the radar data assimilation without the effect of other observation types, Oklahoma Mesonet data are not assimilated during this period.

Both the PAR and WSR-88D radar observation times are reassigned into 1-min batches, as in a traditional EnSRF approach. For example, a ~50 sec complete volume scan of PAR observations are reassigned to the nearest minute and the ~4.2 min volume scan WSR-88D observations, which consist of 14 elevation angles, are separated into single elevation angles and re-grouped at 1 minute intervals based on the times of the tilts. Therefore these 1-min data batches contains ~2-3 tilts of for WSR-88D observations rather than full volume scans that are often treated as being collected simultaneously in 3D EnSRF (Wang et al. 2013a; Yussouf and Stensrud 2010). The 4DEnSRF scheme pre-calculates the observation prior, $H(\mathbf{x})$ for all these observations every 1-min to perform an analysis every 5-minutes (Wang et al. 2013a). Specifically, to produce an analysis valid at time t, the 4DEnSRF uses 5 batches of data and their priors at t - 2 min, t - 1 min, t, t + 1 min and t + 2 min, with each batch consisting of observations within 30 seconds of the batch time.

The standard deviations of the WSR-88D reflectivity and radial velocity observation errors are assumed to be 5 dBZ and 3 m s^{-1}, respectively, following Putnam et al. (2014). After several sensitivity experiments, the same observation error standard deviations as those for the WSR-88D data are used for the PAR data as well. The cutoff radius used in covariance localization for the storm-scale ensemble is 6 km in both the horizontal and vertical, the same as those used in

Jung et al. (2012), while the time localization is assumed to be 5 minutes, following Wang et al. (2013a). To help maintain ensemble spread, RTPS with a relaxation factor of 0.9 is applied over the entire region, and multiplicative inflation of 20% is applied to the prior ensemble in regions directly influenced by the radar data during the period of radar data assimilation. These values are selected based on several sensitivity studies and guidance of earlier studies.

Table 4: An example of tilts of PAR observations that are assimilated in the PAR-reducedtilts experiment. The table lists the tilts and corresponding times of observations used in WSR-88D, PAR, and PAR-reducedtilts experiments from 0028 UTC to 0031 UTC.

UTC Forecast Time	**WSR-88D Tilts in Degrees (Valid Times)**	**PAR Tilts in Degrees (Valid Times)**	**PAR-reducedtilts Tilts in Degrees (Valid Times)**
0028	0.5-0.9 (0027:34-0028:07)	0.51-19.5 (0027:37-0028:15)	0.51-0.9 (0027:37-0027:44)
0029	1.4-3.0 (0028:39-0029:23)	0.51-19.5 (0029:02-0029:40)	1.3-3.1 (0029:15-0029:29)
0030	4.0-8.0 (0029:38-0030:19)	0.51-19.5 (0029:45-0030:24)	4.0-8.0 (0030:15-0030:20)
0031	10.2-19.7 (0030:33-0031:14)	0.51-19.5 (0030:30-0031:21)	10.0-19.5 (0031:17-0031:21)

Three DA and forecast experiments are conducted (Table 2) on the 2-km grid, forced by the same mesoscale ensemble at the lateral boundaries. The first experiment assimilates the KTLX WSR-88D radar observations only (referred to as WSR-88D experiment hereafter), while the second experiment assimilates the full-volume PAR observations (referred to as PAR experiment hereafter). The third experiment assimilates PAR observations but uses only those tilts that closely match the times of the tilts of WSR-88D observations (referred to as PAR-reducedtilts experiment hereafter). Four to five volume scans of PAR observations are needed to reproduce one volume scan of PAR-reducedtilts observations that has the same temporal frequency as in WSR-88D. An example radar observation dataset that is assimilated in the WSR-88D, PAR, and

PAR-reducedtilts experiments during a 5-min DA window is shown in Table 4. As mentioned earlier, while a volume scan of PAR observations has higher temporal resolution than the WSR-88D volume scan, the PAR observations have a slightly coarser spatial resolution than the WSR-88D observations. Therefore, while experiments WSR-88D and PAR-reducedtilts have similar temporal resolutions, the latter has coarser spatial resolutions. The differences between PAR and PAR-reducedtilts will reveal the impact of the higher temporal frequency of PAR observations. Additional information regarding the experiments is found in Table 3. Finally, 1-h deterministic forecasts are initialized from the ensemble mean analyses and three sets of 1-h ensemble forecasts are initialized from the ensemble analyses for all three experiments after 15-, 30- and 45-min of ensemble radar DA (Figure 3d). The analyses and forecasts are used to investigate how quickly and accurately the rapid scan PAR observations initialize and forecast the storm.

3.2. Results

3.2.1. Innovation statistics during the DA period

The three DA and forecast experiments assimilate radar observations from two different radar platforms, i.e., the WSR-88D and PAR. To evaluate the performance of the ensemble filter during the 45-min assimilation period, the root mean square innovation (RMSI), mean innovation, total ensemble spread (TES), and consistency ratio statistics are calculated in observation space against the observations that each experiment assimilated and are shown in Figure 4. The statistics are calculated at the time of the observations and only at locations where the observed or model reflectivity is greater than 15 dBZ. The RMSI is calculated using the following equation (Dowell and Wicker 2009),

$$\mathrm{RMSI} = \sqrt{\langle (d - \langle d \rangle)^2 \rangle}, \tag{7}$$

where

$$d = y^o - \overline{H(x^f)} \text{ or } d = y^o - \overline{H(x^a)}, \tag{8}$$

and the brackets indicate an average over all observations in a radar volume. y^o is the observation, H is the observation operator, which maps the model state to the observation type and location, x represents the model state vector, superscript f indicates a prior estimate (i.e., before the observation is assimilated), superscript a indicates a posterior estimate (i.e., after the observation is assimilated). The TES is defined in terms of the observation prior as

$$\text{TES} = \sqrt{\sigma^2 + \langle \frac{1}{N-1} \sum_{n=1}^{N} \left[H(x_n) - \overline{H(x)} \right]^2 \rangle}, \tag{9}$$

where *N* is the number of ensemble members (40 in our experiment), *n* is an index that identifies a particular ensemble member, and the observation error standard deviation σ is assumed 5.0 dB for both WSR-88D and PAR reflectivity observations, as mentioned earlier. Finally, the consistency ratio is defined as in Dowell and Wicker (2009):

$$\text{consistency ratio} = \left(\frac{\text{TES}}{\text{RMSI}} \right)^2 = \frac{\sigma^2 + \langle \frac{1}{N-1} \sum_{n=1}^{N} \left[H\left(x_n^f\right) - \overline{H(x^f)} \right]^2 \rangle}{\langle (d - \langle d \rangle)^2 \rangle}. \tag{10}$$

All three experiments start with an analysis RMSI of ~15 dBZ (Figure 4); by the end of the assimilation period the analysis RMSI is reduced to ~5-7 dBZ. The forecast and analysis RMSI from the PAR experiment is generally less than the other experiments. Additionally, the effect of the missing PAR volumes between 0045-0055 UTC (see Table 2) is apparent in the increased RMSI near 0055 UTC in both the analyses and forecasts in the PAR-reducedtilts experiment. However, by the 0115 UTC cycle, the RMSI is stabilized (Figure 4a), which suggests that the missing PAR observations have minimal effect by the end of the assimilation period.

Furthermore, the TES (Figure 4c and d) is similar for all experiments in either the forecasts or the analyses, suggesting the ensembles are fairly configured. The consistency ratios are also comparable in magnitude between experiments, with values between 0.1 and 0.4 (Figure 4g). These values are well below 1, suggesting underdisperson in the ensemble, which is commonly seen in storm-scale real data ensemble DA studies (e.g., Jung et al. 2012; Yussouf et al. 2013 and references therein). Despite the spread deficiency, the filter shows no sign of divergence during the 1-hr assimilation period, indicating reasonable robustness of the DA system.

The mean innovation (Figure 4e and f) is largely positive indicating that model reflectivity is lower than the observed reflectivity. The underprediction of reflectivity in the model during the assimilation period is also seen in other studies (Dawson et. al 2012; Yussouf et al. 2015). As expected, the forecast and analysis mean innovation generally decrease with time as more data are assimilated. There are some apparent oscillations in the forecast mean innovation in the early part of the assimilation period that are the result of overprediction of spurious echoes and underprediction of the geographical extent of storms before the forecast is sufficiently improved. The intensity of the storms is underpredicted at 0035 UTC, while the spatial coverage of reflectivity is overpredicted, leading to a small mean innovation. Radar DA at this time is too aggressive at removing reflectivity, resulting in a large positive mean innovation and large RMSI at 0040 UTC. Both spurious echoes and the underprediction of storm intensity are improved at 0045 UTC as more data are assimilated at 0040 UTC, leading to a smaller mean innovation. The main reason for the large positive mean innovation at 0050 UTC is different. As mentioned earlier, the PAR scan sector shifts approximately 21 degrees counterclockwise at 0050 UTC,

meaning the PAR is observing storms that were previously not in the PAR coverage area. Those storms have yet to spin up in the model, leading to underprediction and a large positive mean innovation (Figure 4e and f) at the 0050 UTC assimilation. After 0055 UTC, both mean innovation and RMSI continue to decrease, indicating that the filter is stable.

3.2.2. Analyzed near surface cold pool and vertical vorticity

Studies such as Dawson et al. (2010) and Putnam et al. (2014) have shown the importance of the analyzed and predicted cold pool in convective storms on their dynamics and evolution. To evaluate the analyzed cold pools and associated low-level mesocyclones, the ensemble mean analyses of equivalent potential temperature (ϑ_e), horizontal winds and vertical vorticity at the

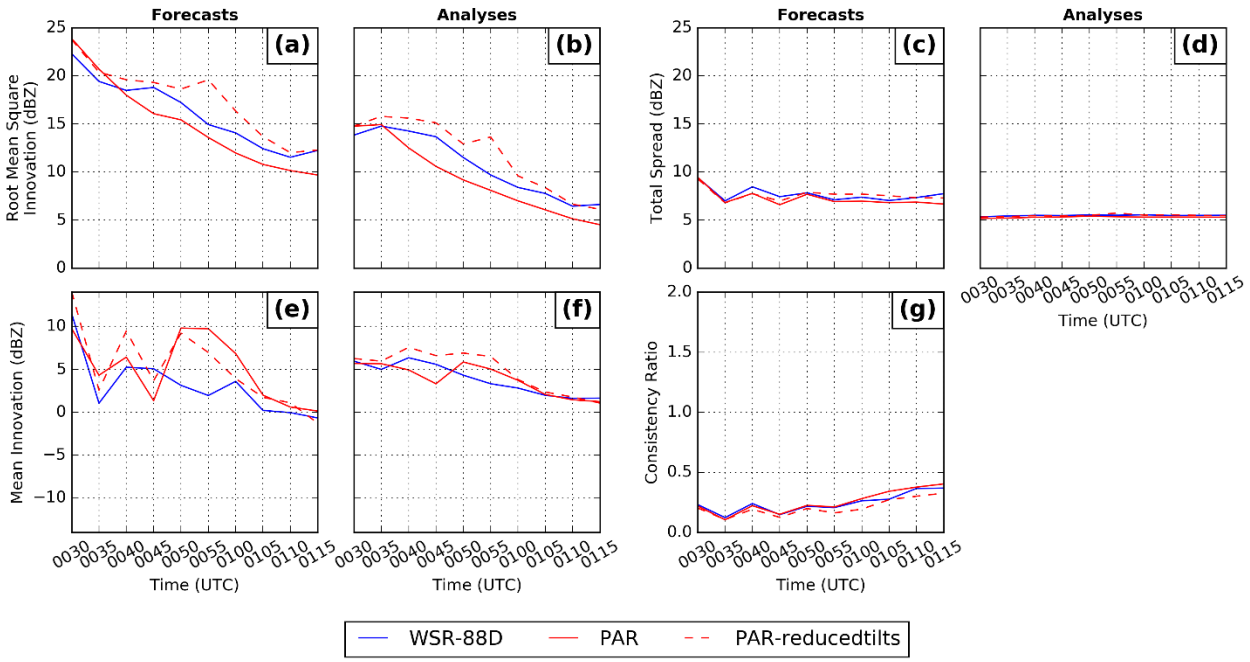

Figure 4: Observation-space diagnostic statistics: forecast and analysis root mean square innovation (RMSI; a and b, respectively) forecast and analysis total ensemble spread (TES; c and d, respectively), forecast and analysis mean innovation (e and f, respectively; positive values denote that the observations are larger than the ensemble mean), and forecast consistency ratio (g), for reflectivity (dBZ) from the three data assimilation and forecast experiments. Calculations are limited to locations where observed and/or ensemble mean reflectivity exceeds 15 dBZ.

first model level or about 25 m above ground from the three experiments after 15 (valid 0045 UTC), 30 (valid 0100 UTC), 45 min (valid 0115 UTC) of DA cycles are shown in Figure 5.

After 15 minutes of DA, the cold pool in storm B in the PAR experiment is the coldest out of all three experiments (minimum ϑ_e of 319 K in Figure 5d). Assimilating more reflectivity observations in the PAR experiment leads to a stronger cold pool, likely due to the negative correlation between reflectivity and temperature in the cold pool (Dowell et al. 2011). The low-level mesocyclones are also relatively weak in all experiments, though the WSR-88D experiment has the strongest (1.9×10^{-3} s^{-1}), followed by the PAR experiment (1.7×10^{-3} s^{-1}) and then by the PAR-reducedtilts experiment (1.4×10^{-3} s^{-1}). As 4DEnSRF uses 5-min assimilation window, the missing PAR observations at the last assimilation cycle from the initial 15 minutes of DA period (Table 2) directly impact the analysis and may be partly responsible for weaker low-level mesocyclone PAR-reducedtilts.

After 30 minutes of DA, the PAR (Figure 5e) experiment has stronger cold pools than the other two experiments in general in storm A and storm B. Additionally, the PAR experiment has the strongest low-level mesocyclone for both storm A and storm B. The peak of 5.4×10^{-3} s^{-1} is in storm A. By this time, 8, 26, and 6 volume scans of observations are assimilated in WSR-88D, PAR, and PAR-reducedtilts experiments, respectively (Table 2), leading to the weakest mesocyclone in the PAR-reducedtilts experiment.

After an additional 15-min of DA (valid 0115 UTC), the PAR experiment continues to have the strongest vorticity values associated with storm B, with the maximum vertical vorticity reaching 6.6×10^{-3} s^{-1} (Figure 5f) while the other two experiments produce much smaller maximum values (Figure 5c and i). In addition, the minimum ϑ_e for both the WSR-88D and PAR experiments

are in the cold pool for storm A; the minimum ϑ_e is similar between these two experiments. For storm B, the cold pool has approximately the same minimum ϑ_e in all three experiments. However, approximating the outline of the cold pool as the 336 K contour, the PAR experiment has a slightly more widespread cold pool.

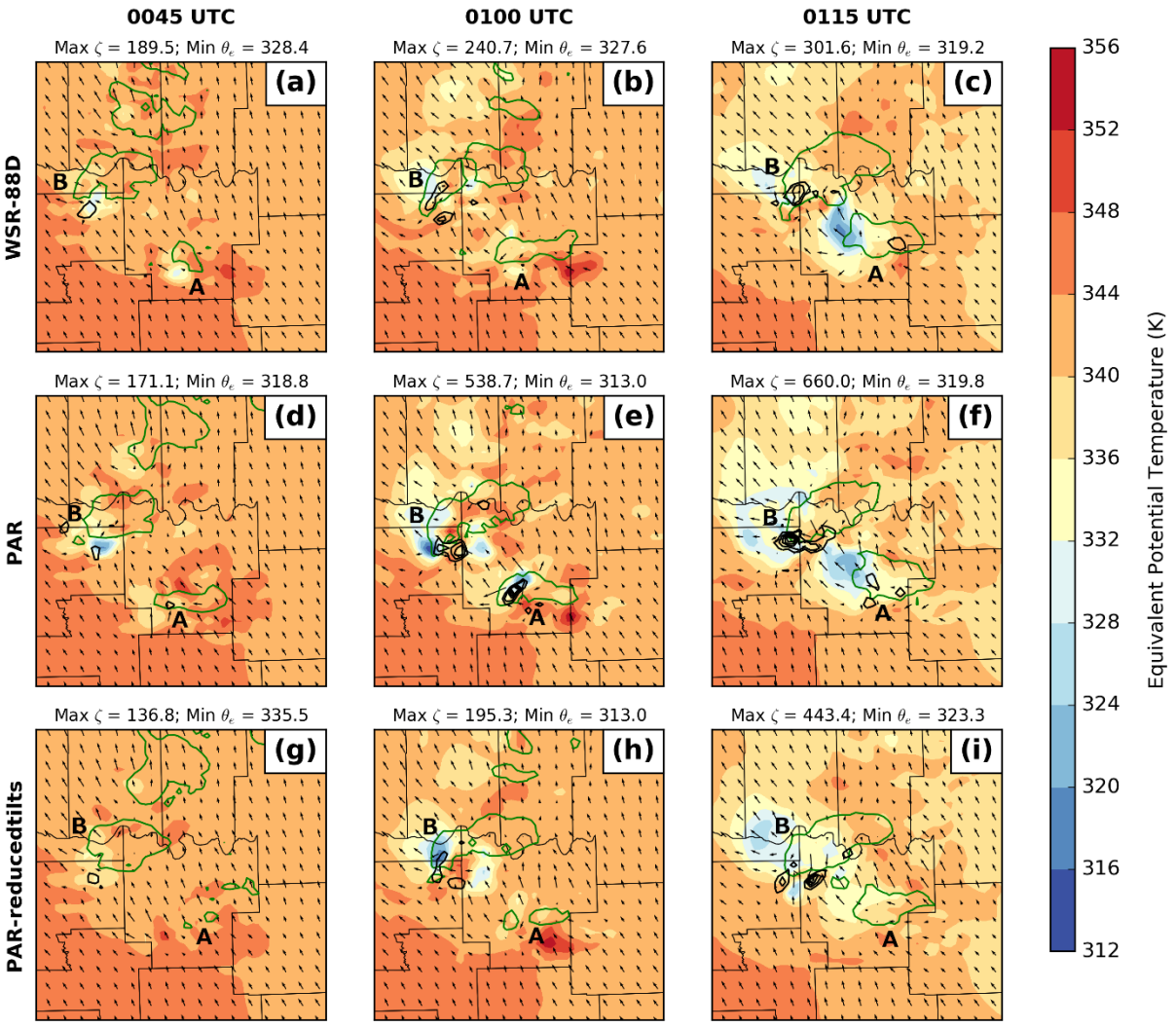

Figure 5: The ensemble mean analyses of vertical vorticity (black contours at intervals of 100×10^{-5} s^{-1}, max vorticity is shown on each panel with unit $\times10^{-5}$ s^{-1}), horizontal wind vectors and equivalent potential temperature (colored; at 4 K increments; min is shown on each panel) after 15 (valid 0045 UTC), 30 (valid 0100 UTC) and 45 (valid 0115 UTC) minutes of data assimilation from (a, b and c) WSR-88D, (d, e and f) PAR and (g, h, and i) PAR-reducedtilts experiments at the first model grid level (~25 m) above ground. Green contours represent 35 dB*Z* reflectivity analyses.

3.2.3. Analyses and forecasts of reflectivity

To evaluate the impacts of DA on the forecasts, the observed National Mosaic and Multi-Sensor QPE (*NMQ*) 3D radar reflectivity mosaic product (Zhang et al. 2011) reflectivity is compared to the analyses and 15-minute forecasts from the ensemble member closest to the mean (Figure 6 and Figure 7). This member is chosen in a similar way to Yussouf et al. (2013). It is defined as the member with the smallest normalized root mean square difference from the mean in potential temperature and *u* and *v* wind components. The reflectivity analyses at 2 km MSL valid at 0045 UTC (after 15 minutes of DA, or three assimilation cycles) are able to capture both storms A and B in all three experiments, though storm A in particular has lower reflectivity in the analyses than the observations (Figure 6a, d, g, and j). Because the PAR does not observe the cells to the north of storm B until 0050 UTC, they are not well-represented at this time in the PAR and PAR-reducedtilts experiments (Figure 6g and j). Particularly, both experiments have one intense storm instead of two weaker storms.

As PAR observations of the two cells to the north of storm B start when the scan sector changes at 0050 UTC, the reduction in intensity of these cells induced by the PAR observations is apparent in the PAR experiment 0100 UTC (Figure 6h). While 7 full volumes are assimilated in the PAR experiment, only two partial volumes are assimilated in the PAR-reducedtilts experiment due to the missing data discussed earlier. As a result, those storms are still poorly resolved at 0100 UTC (Figure 6k). Otherwise, only minor structural differences are evident between experiments. By the 0115 UTC analyses, all three experiments show similar storm structure, suggesting sufficient observations are assimilated by this time (Figure 6f, i and l).

The 15-minute forecast from the member closest to the ensemble mean is examined in Figure 7. In the forecasts initialized at 0045 UTC, the intensity and structure of storm A in the PAR

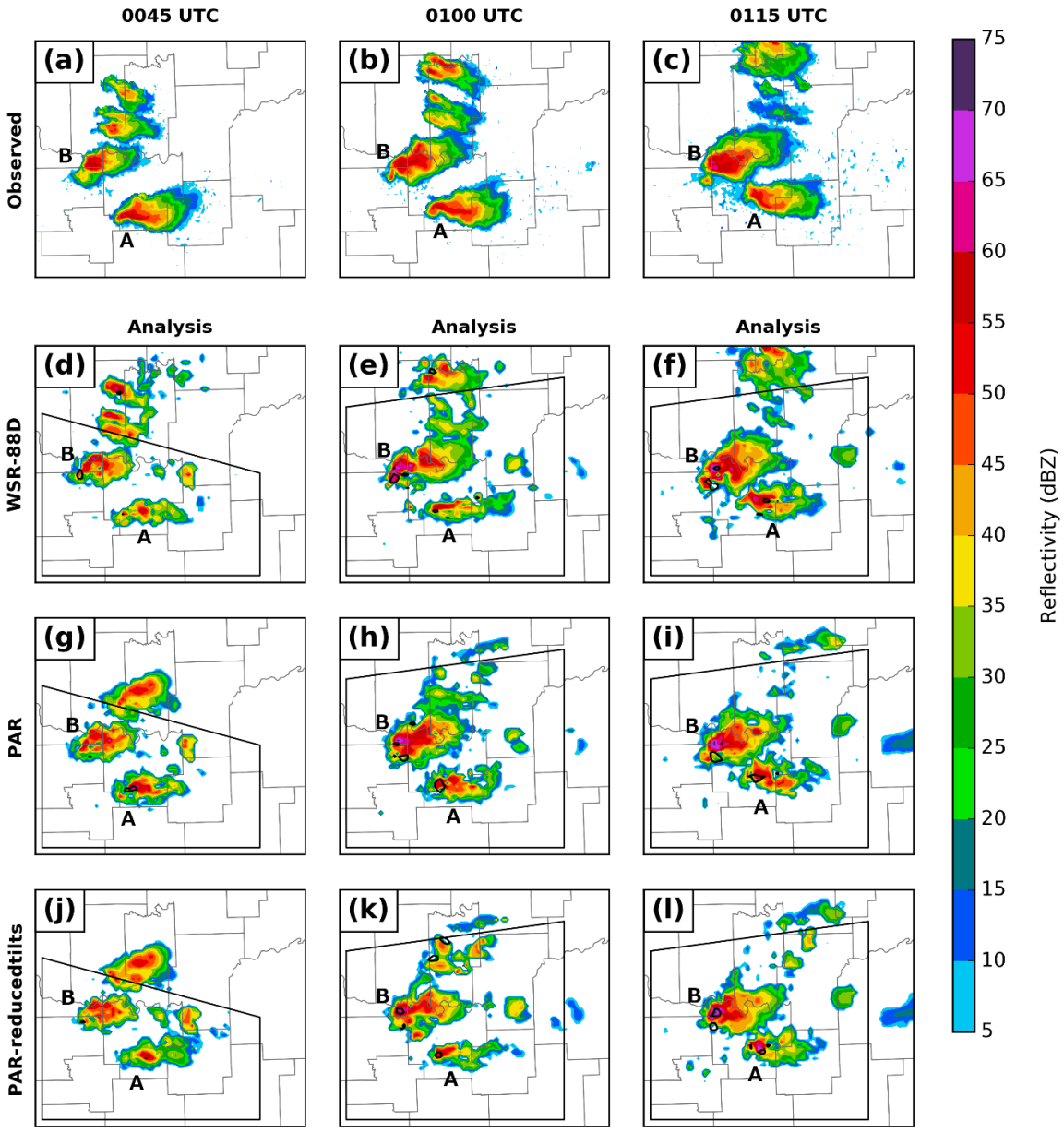

Figure 6: The NMQ reflectivity observations (first row, colors at 5-dBZ increment) and analyses from the member closest to the mean at 0045 UTC (first column), 0100 UTC (second column), and 0115 UTC (third column) from the WSR-88D (second row), PAR (third row) and PAR-reducedtilts (fourth row) experiments. Reflectivity at 2 km AGL is color filled. The 0.004 s^{-1} vertical vorticity contours are given in solid black. The verification domain for the skill scores in Figs. 8 and 9 are shown by the black box. The portion of the domain shown here is over southeast Oklahoma.

experiment matches better with observations (Figure 7a, d, g, and j). However, storm B, which is located near the boundary of the PAR coverage area, is merging with a spurious supercell outside of the PAR coverage area. In the forecasts initialized 0100 UTC, the overall structure of storm B in the WSR-88D and PAR-reduced tilts experiments is more realistic than the corresponding initialization (see Figure 6e and k). Additionally, the ensemble member closest to the mean happens to develop much less spurious cells to the east of storms A and B in the PAR and PAR-reducedtilts experiments compared to the WSR-88D experiment. All other ensemble members (not shown) show similar results, and this is consistent with our observation that northern storms outside of the PAR coverage area are underpredicted in the PAR and PAR-reducedtilts experiments, producing less convection.

Additionally, the reflectivity in storm A is much more intense in all three experiments than in the observations, and the simulated storms continue to display hook echoes, whereas the observed storm does not. In fact, storm A does not dissipate through the end of the 1-hr forecast in any of the experiments (not shown). Thus, the reflectivity forecasts from 15 minutes through the end of the run are considered poor. Furthermore, the 15-minute forecasts initialized at 0115 UTC (Figure 7f, i, and l) are largely similar, except that the PAR experiment and, to a lesser extent, the PAR-reducedtilts experiment develop a spurious vorticity maximum in the northeastern quadrant of storm B. Also, storm A in the 0115 UTC forecasts is still much stronger than the observed, though it does begin to dissipate near the end of the run.

One problem seen in all three experiments is the faster propagation of supercells to the north and east compared to the observations. Similar problems have been seen in many prior storm-scale radar DA studies (e.g. Dawson et al. 2012, 2015; Xue et al. 2014; Yussouf et al. 2015,

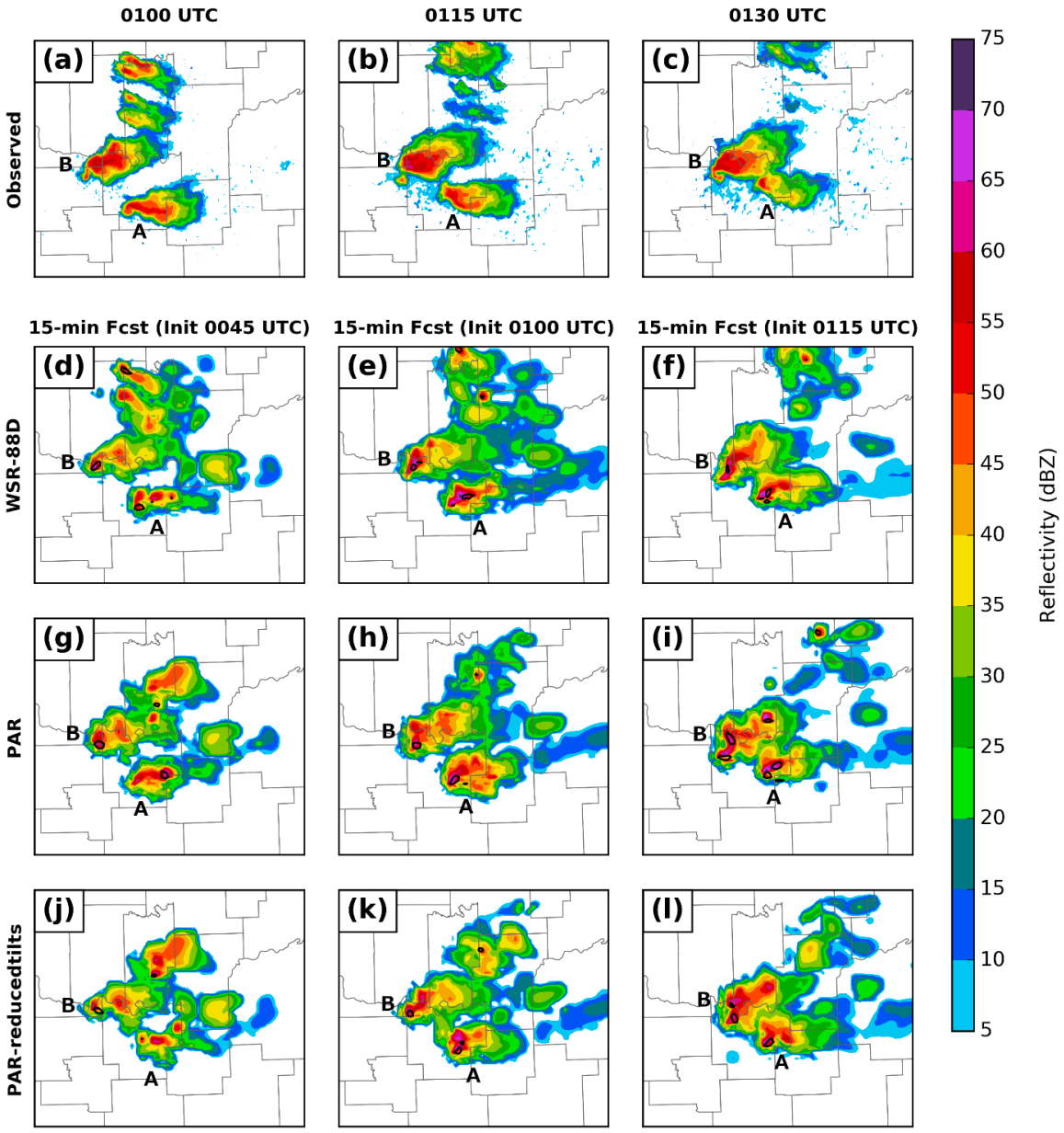

Figure 7: As in Figure 6, but for 15-minute deterministic forecasts initialized from the ensemble member closest to the mean. Note that the verification domain for the skill scores is omitted from this figure.

2016; Stensrud and Gao 2010, to name a few). Stensrud and Gao (2010) suggest that this may be because the model is unable to develop mid-level pressure perturbations that lead to deviant motion in supercells. Xue et al. (2014) speculate that this may be due to errors in analyzing the storm environment, which leads to incorrect steering winds in the model. Fiori et al. (2011) found

that the storm motion depends on the turbulence closure model and grid spacing. Reducing errors associated with supercell motion in storm-scale NWP is worthy of further investigation, however addressing it is beyond the scope of this paper.

3.2.4. Forecast probability of low-level vorticity

The 2-km model horizontal grid spacing used in this study is far too coarse to explicitly resolve any tornado circulation and is barely capable of resolving mesocyclone scale features of the tornadic storm. To infer the amount of low-level rotation from the three experiments, the ensemble forecast probability of low level vorticity is compared with the Warning Decision Support System-Integrated Information (WDSS-II; Lakshmanan et al. 2007) generated surface to 2 km MSL mesocyclone circulations (Miller et al. 2013) from KTLX radar observations (Dawson et al. 2012; Stensrud and Gao 2010; Yussouf et al. 2013, 2015; Xu et al. 2015). The vorticity probabilities are calculated at first model level above ground (~25 m AGL) during the 0-1 h forecast period. The forecast output (which is written out at 5-min intervals) from each ensemble member initialized from a certain analyses time is checked to see whether the vorticity exceeds a threshold value within a radius of 4 km in the horizontal direction around each horizontal grid point at any output time during the 1-h forecast period, and the vorticity probabilities are calculated from the number of members exceeding the threshold values. A threshold of 0.0015 s^{-1} is used for vertical vorticity. The 1-h forecast probability of vorticity are examined every 15 min from after 15, 30 and 45 min (valid 0045, 0100 and 0115 UTC initialization times) of radar DA and is compared with the WDSS-II rotation track (Figure 8).

The 1-hr forecast probability of low-level vorticity exceeding 0.0015 s^{-1} initialized at 0045 UTC shows largely low (<40%) probabilities of vorticity (Figure 8a, d and g). The PAR experiment

(Figure 8d) has the largest overlap between the forecast vorticity swath and the WDSS-II rotation track. The WSR-88D experiment (Figure 8a) shows the highest probability values (~50%) displaced well to the north and west of the WDSS-II track, and the PAR-reducedtilts experiment shows only weak probabilities overlapping the WDSS-II track.

The forecast probabilities of strong low-level rotation increase with an additional 15 minutes of DA (Figure 8b, e, and h), and the forecast vorticity swaths are aligned better with the observed WDSS-II track compared to the 0045 UTC forecasts. The WSR-88D experiment (Figure 8b) has very low probability over the beginning part of the observed strong rotation track but increases to high probabilities over the later part of the track. The PAR-reducedtilts experiment (Figure 8h) has coverage of mid-range probabilities (40-60%) over the first part of the track, but the probabilities never increase above ~60%. On the other hand, the PAR experiment (Figure 8e) has coverage of high probabilities over the first part of the observed rotation track and at least mid-range probabilities over the rest of the track. Thus, the assimilation of the higher temporal frequency PAR observations clearly produces more reasonable probabilistic forecasts of vorticity with only 30 minutes of DA compared to the assimilation of WSR-88D data.

The forecast of low-level vorticity initialized after 45 minutes of DA generates relatively large areas with probability >90% in all three experiments (Figure 8c, f, and i). These are displaced slightly south of the WDSS-II track in all experiments. As in the forecasts initialized at 0115 UTC, the PAR experiment (Figure 8f) has better coverage in the early part of the track as compared to the WSR-88D and PAR-reducedtilts experiments (Figure 8c and i). Additionally, the spurious vorticity maximum seen in the 2 km MSL vorticity from the PAR and PAR-reduced tilts

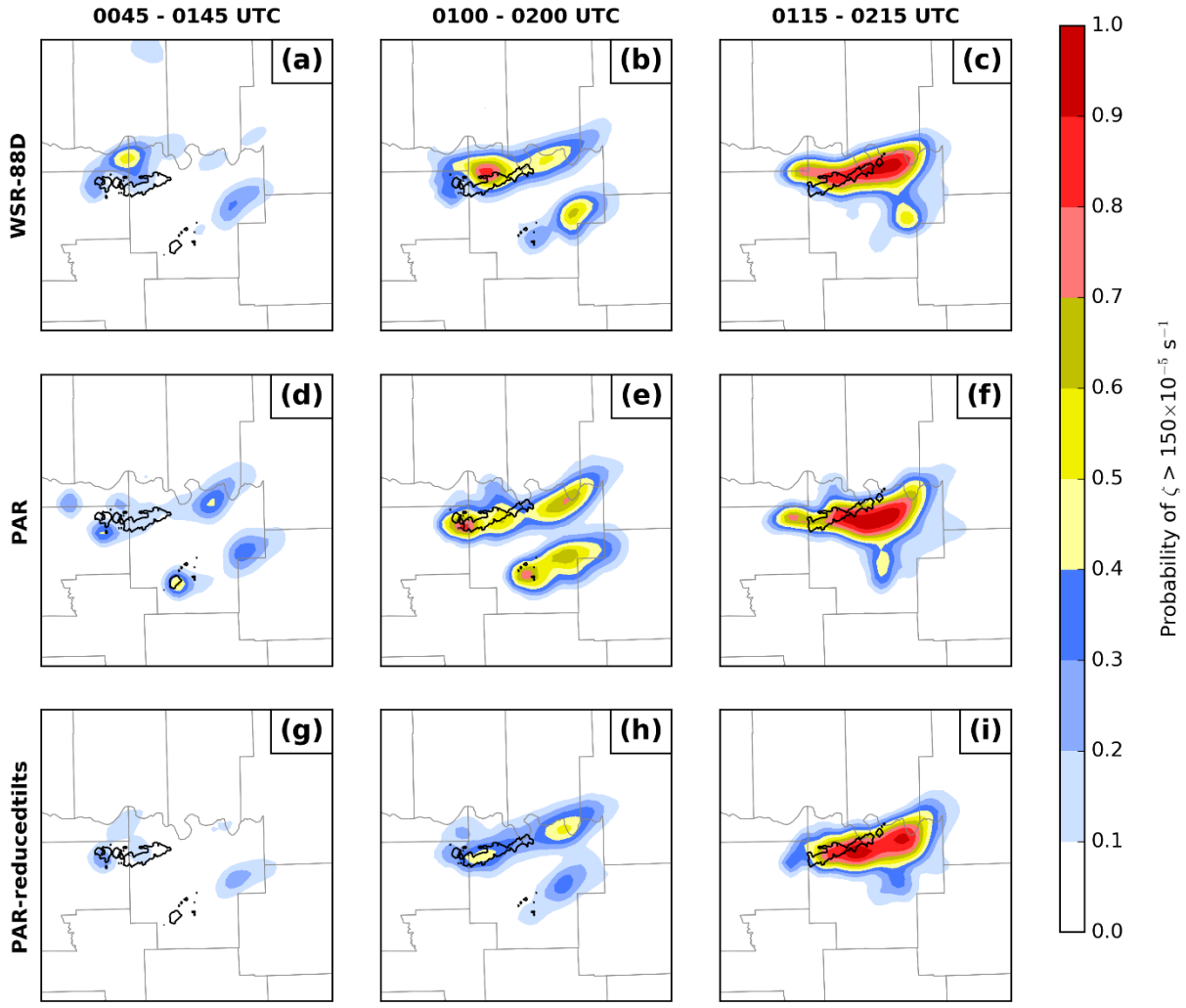

Figure 8: 1-hr forecast neighborhood ensemble probability of vertical vorticity exceeding a threshold of 0.0015 s^{-1} at the first model level above ground from WSR-88D (top row), PAR (middle row) and PAR-reducedtilts (bottom row) experiments initialized after 15 min (valid 0045 UTC, left column), 30 min (valid 0100 UTC, middle column) and 45 min (valid 0115 UTC, right column) of data assimilation. The black contours overlaid are the WDSS-II generated KTLX low-level mesocyclone rotation exceeding 0.01 s^{-1} vorticity during the 0-1 hr forecast periods.

experiments (see Figure 7i and l) does not extend down to the surface. However, aside from the coverage in the early part of the rotation track, few differences are apparent between experiments in the forecasts initialized at 0115 UTC.

Though the above analysis has primarily focused on storm B, the forecasts for storm A show similar results. Particularly, in the forecasts initialized at 0100 UTC, the PAR experiment shows higher probabilities of vorticity for storm A than the WSR-88D and PAR-reducedtilts in the early part of the swath, where there is observed rotation from the WDSS-II track. Additionally,

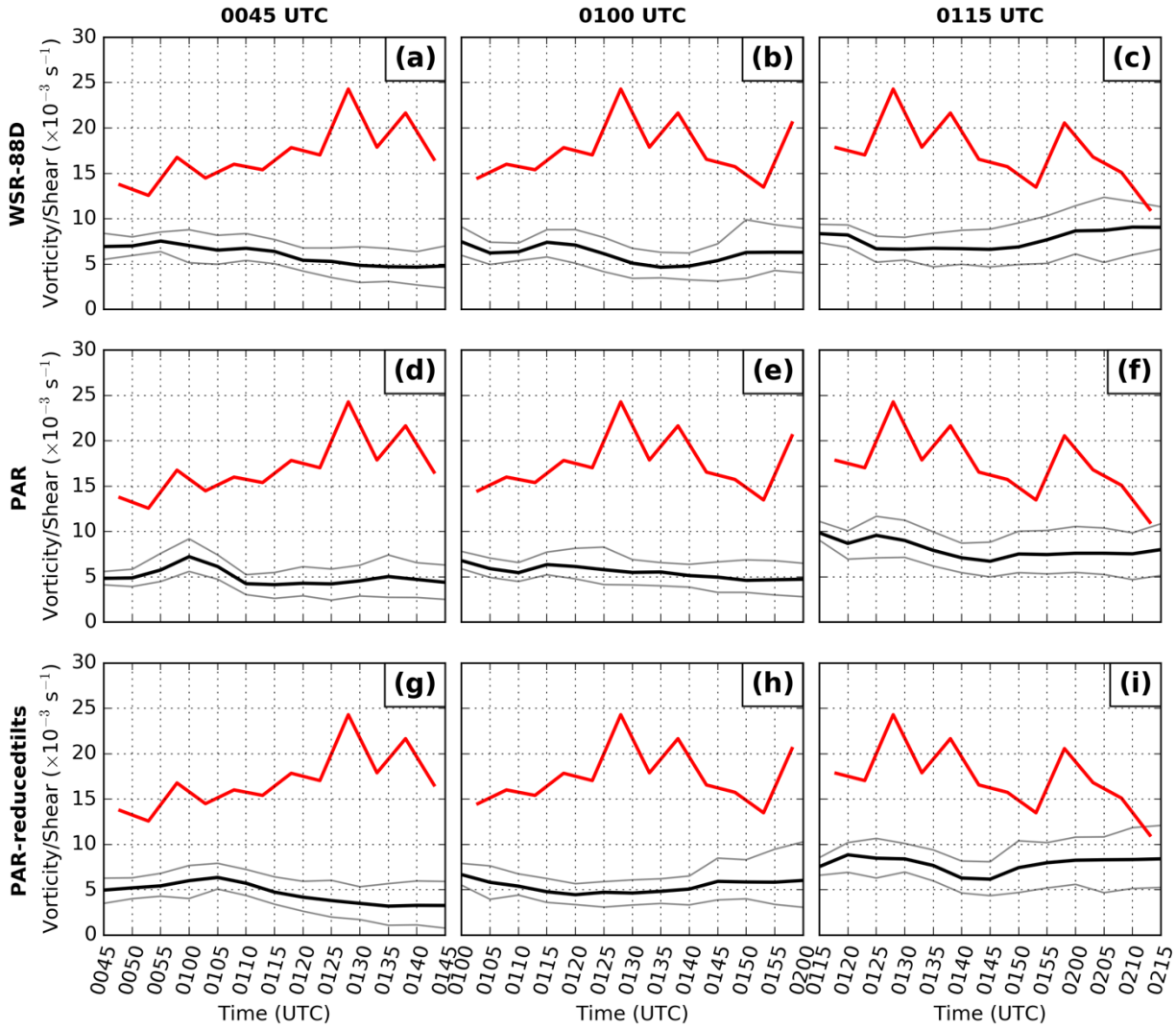

Figure 9: The maximum azimuthal shear for storm B (red line). The maximum vorticity at 2 km AGL within 15 km of the observed location of the mesocyclone of storm B is given in the black and gray lines. The ensemble mean is given in the black line, and the ensemble 10th and 90th percentiles are given in the gray lines. The forecasts for the WSR-88D, PAR, and PAR-reducedtilts experiments are given in the top, middle, and bottom rows, respectively. The 0045, 0100, and 0015 UTC forecasts are given in the left, middle, and right columns, respectively.

the probabilities of vorticity for storm A increase with eastward extent in the WSR-88D and PAR-reducedtilts experiments, which is not reflected in the WDSS-II rotation track. In contrast, the vorticity probabilities in the PAR experiment remain approximately the same with a slight reduction in value with eastward extent. This is consistent with storm A being underpredicted during early DA cycles in the WSR-88D and PAR-reducedtilts experiments and dissipating much later than observations. However, the PAR experiment was able to retrieve the intensity and the spatial coverage much better at early cycles although dissipation was delayed as in other two experiments.

The probabilities are indirectly related to mesocyclone intensity, but to get a direct measure of intensity, the magnitudes of the vorticity in the forecasts and the azimuthal shear in the rotation track are compared for storm B (Figure 9). While it is difficult to directly compare these values, the trends can be compared to perhaps extract useful information. For the 0045 UTC forecast, the trend in the ensemble in the WSR-88D and PAR-reduced tilts experiments are slightly downward, which contrasts with the observed strengthening of the mesocyclone throughout this set of forecasts. The intensity in the PAR experiment is approximately steady, with no net change over the forecast. For the 0100 UTC forecast, all experiments depict an approximately steady mesocyclone intensity, which is the same trend as observed. The 0115 UTC forecast shows the mesocyclone maintaining approximately steady in the PAR experiment with maybe a slight decrease and approximately steady with maybe a slight increase in the WSR-88D and PAR-reducedtilts experiment. None of the ensemble intensity changes are particularly strong, so it is difficult to draw definitive conclusions from this figure about which experiment is

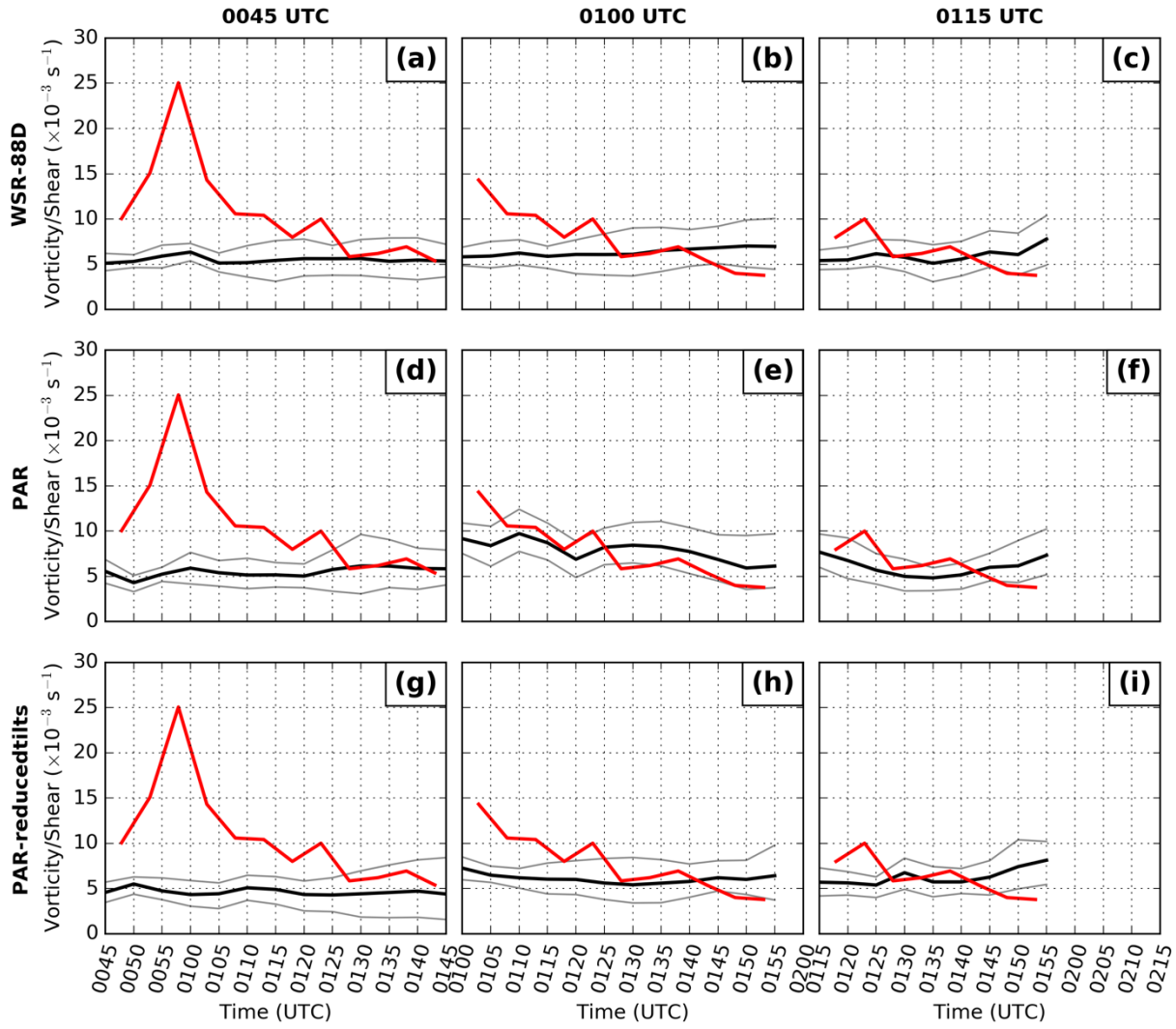

Figure 10: As Figure 9, but for storm A.

better. However, the WSR-88D and PAR-reducedtilts experiments again perform most similar to each other.

The same analysis can be performed for storm A (Figure 10). All experiments miss the spike in intensity in the 0045 UTC forecast, which may be because both storms are poorly initialized in the ensemble after only 15 minutes of DA. The WSR-88D and PAR-reducedtilts experiments keep the intensity steady or slightly increase during the 0100 UTC experiment,

coincident with a weakening of the observed mesocyclone. The 0100 UTC forecast for the PAR experiment does decrease the intensity of the mesocyclone. All three forecasts are approximately steady or slightly increasing during the 0115 UTC forecast, during which the mesocyclone intensity decreased until it could no longer be distinguished at 0200 UTC. The correctness of the PAR experiment at the 0100 UTC forecast may illustrate some skill with capturing the intensity of the mesocyclone, but it is again very difficult to draw definitive conclusions from this analysis.

3.2.5. Skill scores of ensemble reflectivity forecasts

To quantify the accuracy of reflectivity forecasts from the 40-member ensembles during the 1-h forecast period, the equitable threat scores (ETS; Wilks 2006) are calculated from the three experiments for reflectivity exceeding the thresholds of 25 and 40 dBZ (Figure 11) using the NMQ reflectivity observations. These thresholds are chosen to focus on light precipitation areas and heavy convective cores, respectively. Ensemble members are aggregated by summing the number of hits, misses, and false alarms from each ensemble member in the ETS calculation. That is,

$$\mathrm{ETS} = \frac{\bar{h} - h_r}{\bar{h} + \bar{m} + \bar{f} - h_r}; \quad h_r = \frac{(\bar{h} + \bar{m})(\bar{h} + \bar{f})}{\bar{n}}, \tag{11}$$

where $\bar{h} = \sum_i h_i$, and h_i is the number of hits in ensemble member *i*. The quantities $\bar{m}$, $\bar{f}$, and $\bar{n}$ are misses, false alarms, and the total number of grid points, respectively, and are defined in a similar manner to the hits. This is similar, though not equivalent, to taking the mean of the ETS from the members. An ETS of 0 indicates no skill, while an ETS of 1 indicates a perfect forecast by all ensemble members. In order to eliminate areas that were not observed by the PAR, which would unfairly penalize that experiment, verification regions were created at the 0045, 0100, and

0115 UTC analysis times (see Figure 6). Each region is entirely contained within the PAR scan sector at that time. The verification regions were translated during the respective forecasts along with the storms to keep the entirety of the observed and forecast storms in the region. Additionally, ETS is computed over the entire depth of the storm.

The ETSs are found largely to increase with decreasing forecast lead time as expected in all three experiments, reducing to near or below 0 by the end of the forecast periods. In general, the PAR experiment has a higher ETS for both the 25 and 40 dBZ thresholds than the WSR-88D experiment for the forecasts (Fig. 8) despite the missing observations and lower spatial

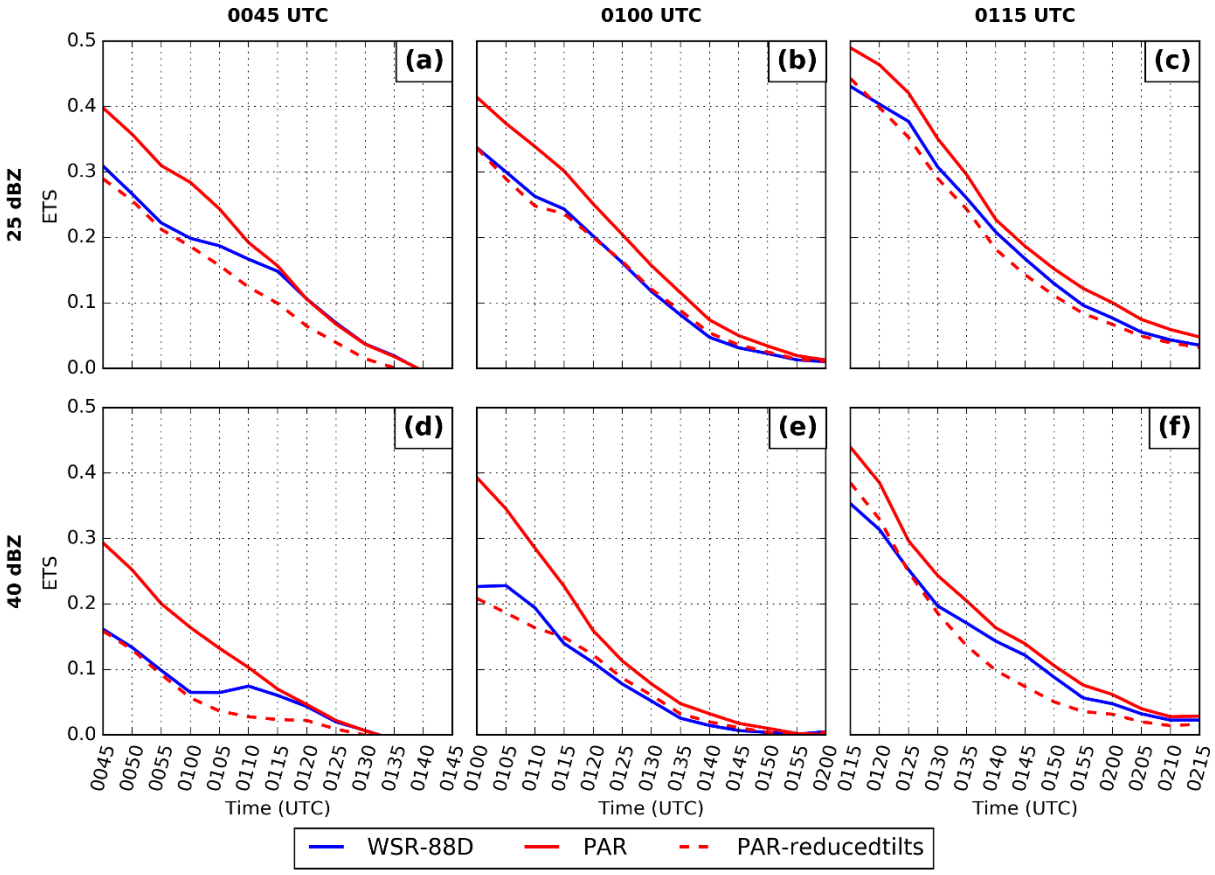

Figure 11: The equitable threat scores (ETS) versus forecast times (UTC) for reflectivity threshold of 25 (first row) and 40 (second row) dBZ from the three experiments. Forecasts are initialized after 15 (valid 0045 UTC; a, d), 30 (valid 0100 UTC; b, e) and 45 min (valid 0115 UTC; c, f) of data assimilation. Calculations are performed over a small subdomain (see text for details).

resolution. We suspect that those limitations are responsible for a similar or slightly higher ETS in the WSR-88D experiment in general compared to the PAR-reducedtilts experiment. The differences between the WSR-88D and PAR experiments are evident only for the first 30 minutes of the forecasts initialized at 0045 UTC (Figure 11a and b), while differences are noticeable almost to the end of the 1-hr forecast period for forecasts initialized at 0100 and 0115 UTC (Figure 11b,

c, e, and f). The ETS of the PAR-reducedtilts experiment remains lower than that of the PAR experiment for the duration of all 1-hr forecasts, indicating the benefit of high temporal frequency observations.

The area under the relative operating characteristic (ROC; Mason 1982) curve (AUC) is calculated for a range of reflectivity thresholds (10 to 50 dBZ) are shown in Figure 12 for the

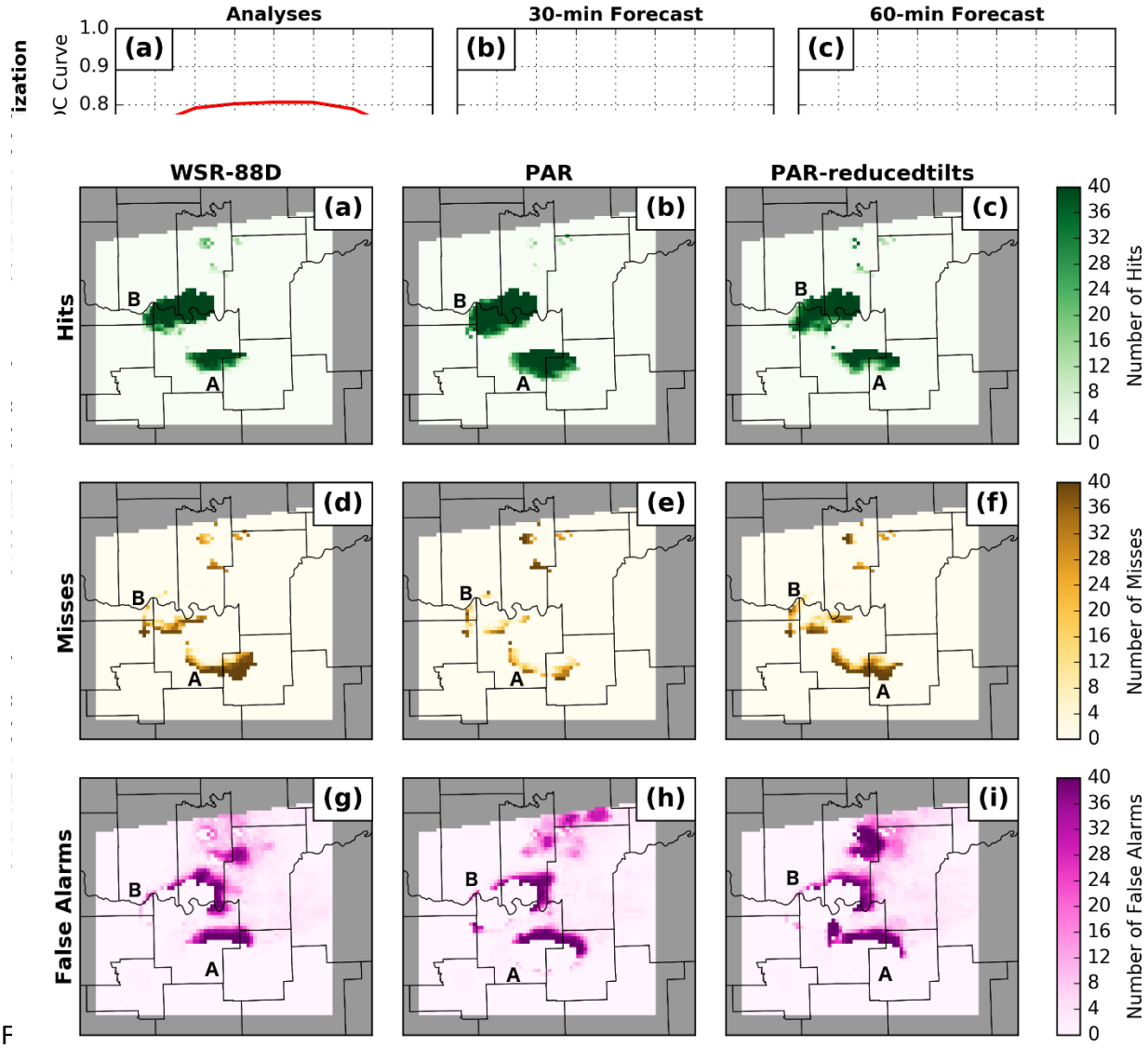

Figure 13: The 10-minute forecast from the 0100 UTC initialization of the number of hits (first row), misses (second row), and false alarms (third row) at the 25 dBZ threshold at each grid point, summed over all ensemble members, at 2 km MSL.

analyses and forecasts. As discussed in Snook et al. (2012), ROC measures hit and false alarm rates at varying probability thresholds (Mason and Graham 1999); it therefore provides information on the ability of a probabilistic forecast system to correctly discriminate between events and non-events (no resolution). A forecast with perfect skill has an AUC of 1.0 and 0.5 means zero skill.

All three experiments have high AUCs in the analyses (Figure 12a, d and g) and the value decreases as the forecast lead times increases (Figure 12b, e, h, c, f, and i). The 30-minute forecasts after 15 minutes of DA are just below the skillful range (AUC > 0.7), and the 60-minute forecasts have very small AUCs, indicating no ability to discriminate (AUC < 0.5). The 30-minute forecasts after 30 and 45 min of DA increases the AUCs to above 0.7 for most thresholds for the PAR experiment, suggesting that at least 8 radar volume scans (~30-min of DA) are needed to produce operationally useful forecasts.

The reason for the improved ETS scores and AUCs can be seen in Figure 13, which shows the 10-minute forecast from the 0100 UTC initialization of the number of hits, misses, and false alarms at the 25 dBZ threshold at each grid point, summed over all members of the ensemble. All three experiments have large numbers of hits where the observed storms overlap with the storms in the ensemble. The majority of the misses in all three experiments occur to the south and west of storms A and B (Figure 13d, e, f), and the majority of false alarms occur to the northeast of storms A and B (Figure 13g, h, i), likely a result of the faster propagation of storms in the model, noted earlier. The three experiments are largely similar, except the PAR experiment better captures the southern extent of both storms A and B, resulting in fewer misses and more

hits in those regions. This directly increases the ETS score and results in larger AUCs because of a higher probability of detection, which is used in computing the ROC curve.

Overall, the PAR experiment generates the highest AUCs and the PAR-reducedtilts experiment the lowest AUCs during the 1-h forecast period. In the analyses, these differences are largest after 15 minutes of DA and decrease as more data are assimilated. However, the largest differences in the 30-minute forecasts are after 30 minutes of DA, and in the 60-minute forecasts, the AUCs are largely the same between experiments. These results suggest that the PAR data are most useful in DA windows that are ~30 minutes long.

3.3. Summary

To evaluate the impact of high-temporal-resolution PAR observations compared to the current operational WSR-88D observations on convective storm prediction, three DA and forecast experiments are conducted for the 22 May 2011 Ada, Oklahoma tornadic supercell event using radial velocity and reflectivity observations from both rapid-scan NWRT PAR and WSR-88D KTLX. The experiments are conducted with the ARPS model and its 4DEnSRF DA system using a heterogeneous mesoscale environment. The first experiment assimilates traditional WSR-88D radar observations, the second experiment assimilates observations from the NWRT PAR and the third experiment assimilates PAR observations with WSR-88D-like temporal frequency (PAR-reducedtilts experiment). The WSR-88D and PAR experiments are designed to compare the accuracy of the forecasts due to two different temporal resolutions from the different radar platforms. The PAR-reducedtilts experiment picks a subset of elevations from the rapid-scan PAR observations from the observation data set to mimic the WSR-88D-like coarser temporal resolution observations (even though spatial resolution differences exist between PAR and WSR-

88D observations) and is designed to examine the differences in the analyses and forecasts due to high-temporal-resolution PAR observations. All three experiments assimilate radar observations for a continuous 45 min DA period at every 5-min assimilation frequency. There are some missing PAR volume scans during this period. Finally, 1-hr forecasts are launched at the end of the 15-, 30- and 45-min assimilation periods, respectively. The goal is to evaluate how quickly and accurately the rapid-scan PAR observations can initialize and forecast the tornadic supercell storm.

The observation-space diagnostic statistics compared against the assimilated observations reveal that the filter shows no sign of forecast divergence during the 45-minutes assimilation period. All three experiments are able to analyze reflectivity structures that are similar to the observations even after a short 15-min DA period. The 15-minute forecasts from the member closest to the ensemble mean largely captured the storm evolution, though storm A was too strong in the 0100 and 0115 UTC initializations. The 1-h forecast probability of low-level vorticity after 15, 30 and 45 min of DA from all three experiments indicate that the placement and alignment of the swath of high probabilities from the PAR experiment more closely match the WDSS-II derived rotation track, particularly from the forecast initialized after only 30 min of DA. For this forecast, the low-level rotation from the WSR-88D does not cover the initial part of the WDSS-II rotation track, and the probability values from the PAR-reducedtilts experiment are comparatively low. The ETS score and the area under the ROC curve (AUC) also indicate that PAR experiment forecasts improved skill compared to the WSR-88D and PAR-reducedtilts experiments, particularly over 15- and 30-min long assimilation periods. The

improved skill in the PAR experiment is the result of better capturing the southern extent of reflectivity coverage in both storms A and B.

The overall results indicate that the WSR-88D and PAR-reducedtilts experiments perform most similarly to each other, and the PAR experiment performs better than those two experiments. From this, we conclude that the assimilation of high-frequency PAR observations results in improved accuracy over traditional WSR-88D observations. Furthermore, we conclude that this improvement in accuracy is the result of the increased data volume from the PAR. The improved accuracy is more pronounced with a short or moderately long (15-30 min) assimilation period. The positive impact of higher temporal resolution of the PAR observations decreases for a longer assimilation period (45 min in this study). This is similar to the findings of the OSSE studies of Xue et al. (2006) and Yussouf and Stensrud (2010).

Chapter 4: Analysis of the 31 May 2013 Case

4.1. Experiment Design and Observations

4.1.1. Observations assimilated

The observations assimilated in these experiments fall into two categories: surface data and radar data. For surface data, we use u, v, potential temperature (ϑ), and water vapor mixing ratio (q_v) observations from the Oklahoma Mesonet (Brock et al. 1995; McPherson et al. 2007). Radar data come from two sources: the NWRT PAR and the operational WSR-88D (KTLX).

The NWRT PAR contains a single panel creating an electronically steered beam, which provides continuous ~70 s volumetric updates over a 90° sector for the duration of the assimilation period. The limited sector poses a potential problem: the lack of observations outside the scan sector means that the EnSRF does not have observations to suppress convection outside the scan sector. This could lead to growth of that convection, which could interfere with the assimilated storms during forecast period. Supinie et al. (2017) mitigated this by transplanting zero-reflectivity observations from the KTLX scan wherever the composite reflectivity from KTLX is zero. We do the same here, but add the additional step that, in columns where the KTLX composite reflectivity is non-zero, the following algorithm is applied:

1) Linearly interpolate the column of KTLX observations to the heights of the PAR observations.
2) Set PAR observations to 0 where the interpolated column of KTLX observations is 0, set the PAR observations to "missing" where the interpolated column of KTLX observations is greater than 0.

3) If the lowest level of PAR observations is below the lowest level of KTLX observations and the lowest level of KTLX observations is 0, set the lowest level of PAR observations to 0. Otherwise, set the lowest level of PAR observations to "missing"

This additional step is designed to provide zero-reflectivity observations below the expansive anvil coverage of the storms.

In addition to the difference in scanning coverage, the NWRT PAR beam width varies with azimuth from 1.6° for beams normal to the panel to 2.3° at ±45° away from normal. Conversely, the WSR-88D has a constant 0.95° beam width over the full 360° scan. Thus, at worst case, the PAR has half the spatial resolution of the WSR-88D.

Quality control procedures are applied to both sets of data. Automated processing, consisting of automated velocity unfolding and clutter removal is applied to the KTLX data. Details on the velocity unfolding procedure can be found in Brewster et al. (2005). The automated clutter removal procedure is to remove gates classified as ground clutter, biological scatterers, or three-body scatter spikes by a polarimetric hydrometeor classification algorithm (Park et al. 2009; Mahale et al. 2014). Manual clutter removal and velocity unfolding are performed on the PAR data. Data from both radars are bilinearly interpolated to the model grid points in the horizontal direction and left on the radar tilts in the vertical.

4.1.2. Model Configuration and Experiment Design

4.1.2.1. Mesoscale Domain

The mesoscale DA system used for these experiments is identical to that used by Stratman et al. (2020). Briefly, the mesoscale DA system consists two domains, one with a grid spacing of 15 km covering CONUS and one with a grid spacing of 3 km covering the south-central US. The

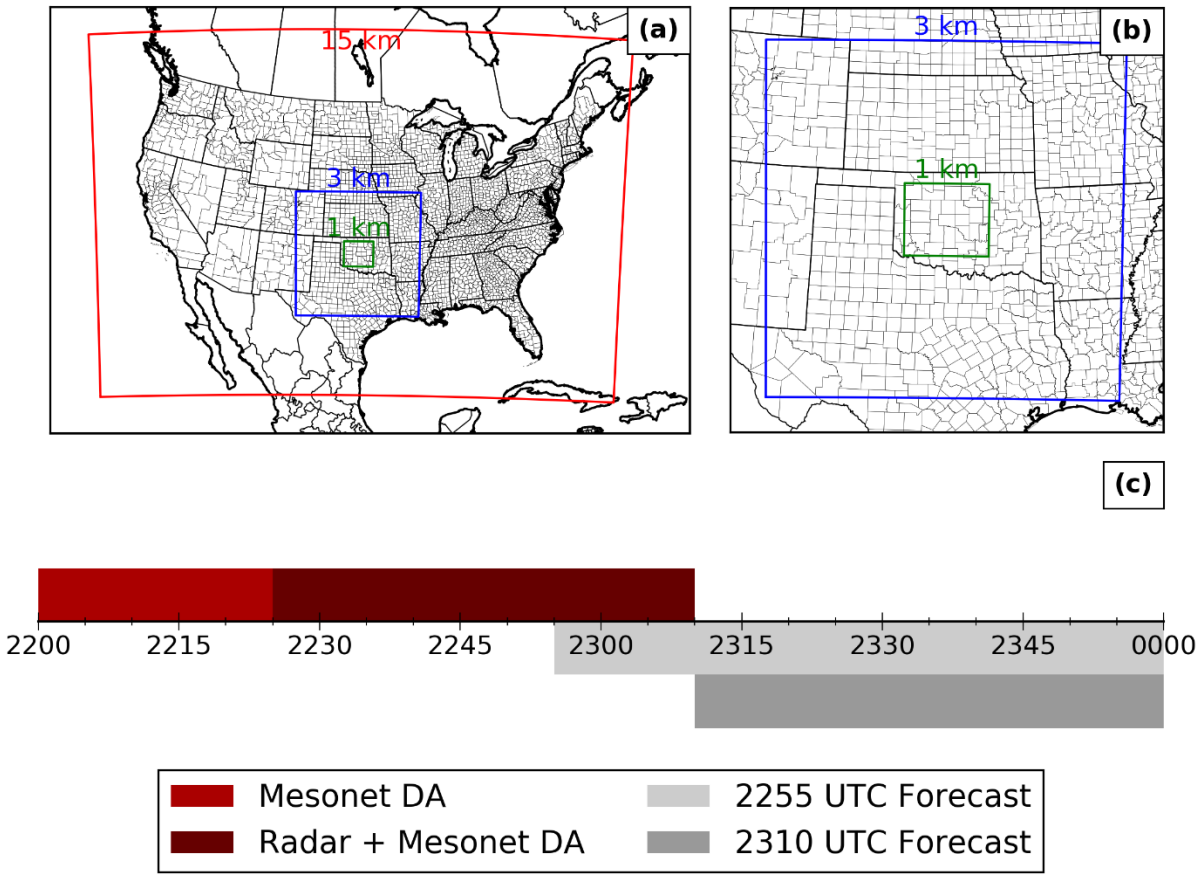

Figure 14: Experiment domains and timeline. The red and blue lines in (a) and (b) give the boundaries of the 15 km and 3 km mesoscale domains. The green line gives the boundary storm-scale domain at 1 km grid spacing. The times in (c) are in UTC.

first 18 members of the National Centers for Environmental Prediction's Global Ensemble Forecast System (GEFS) provides the atmospheric initial and boundary conditions for the 15 km domain, while the North American Mesoscale Forecast System (NAM) provides the soil state. Physics diversity is used to provide spread in the mesoscale ensembles. Both ensembles are initialized at 0000 UTC 31 May 2013 and cycled hourly until 0000 UTC 1 June 2013. Surface, radiosonde, and aircraft observations are assimilated on both domains. More detail can be found in Stratman et al. (2020).

4.1.2.2. Storm Scale Domain

All experiments are performed using 36 ensemble members on a 282 × 240 × 51 grid with 1 km grid spacing. The storm-scale domain is centered on El Reno, OK (Figure 14). Ensemble members from the mesoscale domain are downscaled at 2200 UTC to initialize the storm-scale domain. The storm-scale ensemble is then integrated until 2220 UTC, assimilating Oklahoma Mesonet data every 5 minutes. Next, radar and Oklahoma Mesonet data are assimilated until 2310 UTC, with analyses produced every 5 minutes. This gives a total of 4 analyses with only Oklahoma Mesonet data and 10 analyses with both Oklahoma Mesonet data and radar data.

The Weather Research and Forecasting (WRF) model version 3.8.1 is used as the prediction model for each experiment. Thompson microphysics (Thompson et al. 2004, 2008) is used for all members, as are the Rapid Radiative Transfer Model (RRTM) longwave (Mlawer et al. 1997), Dudhia shortwave (Dudhia 1989) radiation parameterizations and Noah land surface model (Tewari et al. 2004). Of the 36 members, 12 use the Mellor-Yamada-Janjić (MYJ; Janjić 1994), 12 use the Yonsei University (YSU; Hong and Pan 1996), and 12 use the Mellor-Yamada-Nakanishi-Niino (MYNN; Nakanishi and Niino 2009) planetary boundary layer parameterizations. No cumulus parameterization is used.

The Advanced Regional Prediction System (ARPS) version 5.4.2 4DEnSRF (Wang et al. 2013a) is used as the data assimilation algorithm. Spread in the ensemble is maintained using a combination of multiplicative covariance inflation applied to the prior ensemble and relaxation to prior spread (RTPS; Whitaker and Hamill 2012). A multiplicative inflation factor of 20% is used, and multiplicative inflation is only applied to regions directly influenced by radar data. A relaxation factor of 0.98 is used for the RTPS inflation. Assumed observation errors for radar data are 7 dBZ for reflectivity, 3 m s^{-1} for radial velocity; for Mesonet data, they are 2 m s^{-1} for *u* and

v, 2 K for ϑ, and 1 g kg^{-1} for q_v. Observation localization radii are 6 km in the horizontal and vertical for radar data, and 50 km in the horizontal and 6 km in the vertical for Mesonet observations. The temporal localization half-width for all observations is 2.5 min, corresponding to a 5-min DA window. The 5th-order polynomial function of Gaspari and Cohn (1999) is used for both spatial and temporal covariance localization.

4.1.3. Experiments Run

Several ensemble DA and forecast experiments are conducted. One experiment uses only data from KTLX, and another uses only data from the PAR. Both experiments apply RTPS inflation once at the end of the assimilation window. To determine the optimal use of PAR observations in the EnSRF, four additional experiments are run. The first is identical to the PAR experiment, except for the order that the radar observations are assimilated. In the PAR experiment, radar observations are assimilated in order by volume, then by elevation angle. The first additional experiment assimilates these in reverse order, meaning later volumes and higher elevation angles are assimilated first. Thus, this first additional experiment is called "PAR_reversed." Another experiment assimilates volumes starting at the analysis time and working outwards, prioritizing the latest volume. For example, at the 2310 UTC analysis, the PAR volume valid at 2310 UTC is assimilated first, followed by 2311, 2309, 2312, and then 2308 UTC. This experiment is called "PAR_inout." A third additional experiment assimilates volumes starting at the edges of the DA window and working inward towards the analysis time, again prioritizing the later volumes. For example, at the 2310 UTC analysis, the 2312 UTC volume is assimilated first, followed by the 2308, 2311, 2309, and then the 2310 UTC volume.

The final additional experiment is also based on the PAR experiment but deals with the covariance inflation. The PAR experiment applies RTPS inflation once at the end of the assimilation procedure. In order to determine the effect of the large number of PAR observations on the ensemble spread, we created this second additional experiment which applies RTPS inflation after assimilating each volume of PAR data. We call this technique "Inter-Volume Covariance Inflation" (IVCI), and thus this experiment is named "PAR+IVCI."

While IVCI is straightforward conceptually, in practice, the IVCI technique is complicated by the parallelization algorithm used in the ARPS EnSRF implementation (Wang et al. 2013b). As part of this algorithm, each parallel subdomain is subdivided into four patches, and assimilation occurs on each patch in sequence. However, the complicating factor is that the algorithm assimilates data for all volumes on patch 1, then for all volumes on patch 2, and so on. Data on the boundaries of the parallel subdomains are exchanged after assimilating all volumes on each patch. To handle this, we apply covariance inflation after every volume on every patch. This means that instead of applying RTPS inflation 5 times for 5 volumes of PAR data, it is applied 20 (= 5 volumes × 4 patches per parallel subdomain) times. For the patches on which radar data are not assimilated, the formulation of RTPS means that the covariance inflation has no effect.

4.2. Results

4.2.1. Observation Space Statistics

First, we examine various statistics computed in observation space (Figure 15) as a sanity check on the performance of the EnSRF. The statistics computed here are root mean square innovation (RMSI), mean innovation (also known as bias), total ensemble spread (TES), and consistency ratio. Statistics are computed for observations at all times considered in the analysis

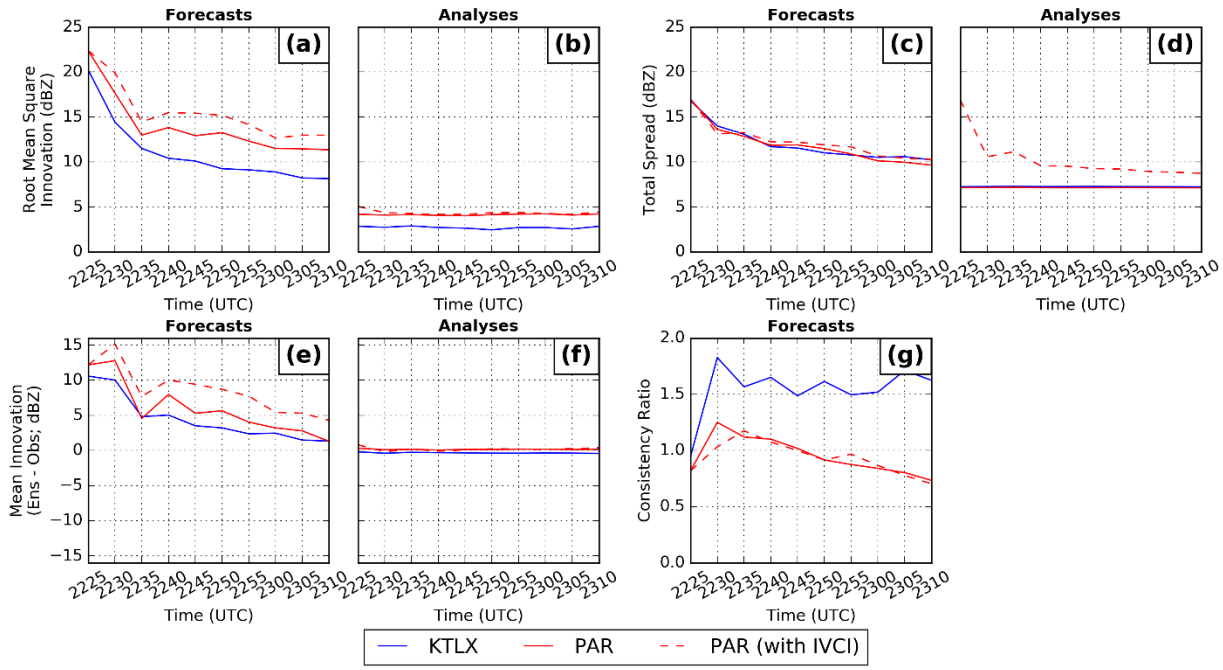

Figure 15: The forecast and analysis root-mean-square innovation in observation space (a) and (b), the forecast and analysis mean innovation in observation space (e) and (f), the forecast and analysis total spread in observation space (c) and (d), and the consistency ratio (g).

and wherever reflectivity ≥ 15 dBZ in either the ensemble mean or the observations, except where the observations are missing.

The forecast RMSI (Figure 15a) for each analysis decreases from ~21 dBZ to ~10 dBZ over the course of the DA period. The RMSI in the KTLX experiment is always 2-4 dB less than in the PAR-based experiments. This is partially a result of the zero-reflectivity observations added to the PAR volumes outside the PAR scan area and where KTLX composite reflectivity is non-zero. In this case, the PAR analyses are removing hydrometeors from the low levels of the atmosphere, but not the anvil region aloft, where no zero-reflectivity observations have been added. This results in the remaining hydrometeors falling into the area where the zero-reflectivity observations have been added, which appears in the RMSI statistics. This can also be seen in the forecast bias statistics (Figure 15e), in which the KTLX experiment always has a lower bias than

the PAR and PAR+IVCI experiments. The PAR_reversed, PAR_inout, and PAR_outin experiments display nearly the same behavior as the PAR experiment and are left out of Figure 15 for clarity.

The effect of the additional covariance inflation is apparent in the analysis TES (Figure 15d); the PAR+IVCI experiment generally has higher analysis TES than the other three experiments. However, the enhanced TES for reflectivity does not feed back into the ensemble, as the forecast TES (Figure 15c) does not differ greatly between the PAR and PAR+IVCI experiments. In the ARPS implementation of the 4DEnSRF, the reflectivity priors are updated during the DA period independently of the model state, and the analysis statistics are computed on these updated priors. In general, if the observed variable is a linear function of the model state, the updated priors and the priors recomputed from the updated model state should be the same. For radar reflectivity, this is not true, so we do not expect that the updated priors will be reflected in the model state.

The consistency ratio (Figure 15g) is designed to show how large model errors are in relation to the model spread. Values near 1 indicate that the ensemble spread is about the same magnitude as the forecast error, and this is considered ideal. Values over (under) 1 mean that the ensemble has too much (not enough) spread compared to the magnitude of the forecast errors, and the ensemble is considered over- (under-) dispersive. Here, the KTLX experiment is generally around 1.5 (indicating over-dispersion) and the three PAR experiments are generally around 1, decreasing with more DA cycles (indicating correct dispersion, becoming under-dispersed with more DA cycles). As in Supinie et al. (2017), the assumed observation error standard deviation for reflectivity (7 dBZ) used here is larger than in other storm-scale DA and forecast experiments (usually 3-5 dBZ, as in Snook et al. 2015; Wheatley et al. 2015; Yussouf et

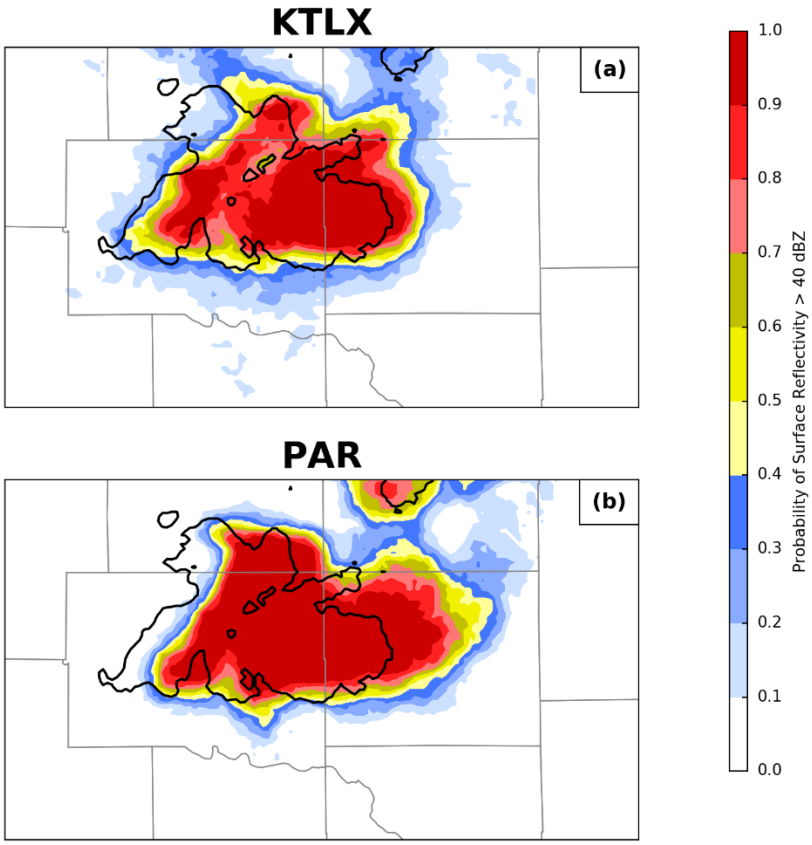

Figure 16: Forecast probability of reflectivity > 40 dBZ at the lowest model level at 2330 UTC for the KTLX (a) and PAR (b) experiments. The forecast was initialized at 2310 UTC. The observed 40 dBZ contour of 0.5° reflectivity from KTLX is given in black.

al. 2016). Therefore, it follows that our consistency ratios are larger than other DA and forecast experiments, which generally report under-dispersion. Overall, the observation-space statistics do not reveal any systemic problems with any of the experiments, so we feel comfortable continuing with the analysis.

4.2.2. Forecast Behavior

To examine the behavior of the forecasts, we will first examine the reflectivity presentation in the 2310 UTC forecast (Figure 16). Both the KTLX and PAR experiments maintain large areas of reflectivity > 40 dBZ which are overall coincident with the observed areas of 40 dBZ. In both experiments, the storms appear to move slightly too fast, which is a well-known

feature of storm-scale forecasts, as mentioned in Chapter 3:. The hook reflectivity shape appears to be at roughly the same longitude in both experiments (collocated with the inflow notch in the observed reflectivity), and implying the updrafts are at roughly the same longitude at this forecast time. However, there are a couple differences between the two experiments. The PAR experiment has much higher reflectivity extending downstream from the storm, perhaps indicating stronger upper-level westerly winds. In addition, the western flank of the storm has a much sharper probability gradient in the PAR experiment than the KTLX experiment. This same behavior is present in the 15-minute forecast from the 2255 UTC initialization. However, in the 2310 UTC forecast, the PAR experiment does capture the development to the northeast of the El Reno supercell better than the KTLX experiment.

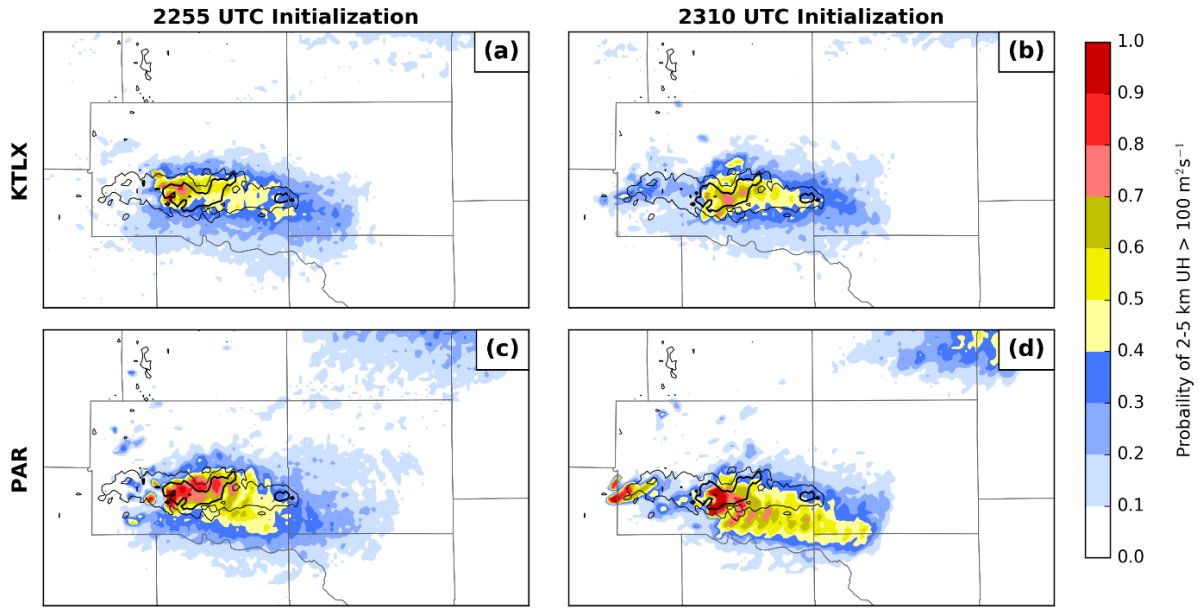

Figure 17: Probability swaths of 2-5 km updraft helicity > 100 m^2 s^{-2} for forecasts initialized at 2255 UTC (left column) and 2310 UTC (right column) for the KTLX (a,b) and PAR (c,d) experiments. Swaths begin at forecast initialization and end at 0000 UTC. The MRMS mid-level azimuthal shear track is contoured in black. The thin and thick lines are the 0.006 and 0.012 s^{-1} contours, respectively.

In order to diagnose the behavior of the mesocyclone, we use the probability of 2-5 km updraft helicity (UH) exceeding 100 m^2 s^{-2} (Figure 17). UH is defined as

$$UH = \int_{z_l}^{z_u} w\zeta dz, \tag{12}$$

where w is vertical velocity, ζ is vertical vorticity, and z_u and z_l are the heights of the bottom and top of the layer over which UH is considered (2 and 5 km here, respectively). This is a popular metric in storm-scale forecasting intended to show the presence of a rotating updraft. The UH swath from the 2255 UTC initialization shows UH probability coinciding with the MRMS mid-level rotation tracks for the main supercell in both the KTLX and PAR experiments (Figure 17a and c). The PAR experiment shows a higher UH probability than the KTLX experiment over the early portion of the track of the El Reno supercell. The westernmost portion of the MRMS rotation track is associated with a secondary supercell that develops to the west of the El Reno supercell, and none of the 2255 UTC forecasts capture this development. Furthermore, the 2310 UTC experiment is similar, with the PAR experiment having a higher probability of UH from the main supercell than the KTLX experiment. However, one difference from the 2255 UTC initializations is the mesocyclone from secondary supercell that develops west of the El Reno supercell is better captured with higher probabilities in the PAR experiment than in the KTLX experiment.

4.2.3. Mesocyclone Intensity Differences

To diagnose the reasons for differences in mesocyclone intensity, we look at 2-5 km mean vertical vorticity and vertical velocity (Figure 18) to determine the contributions from each variable to the UH. Both PAR and KTLX experiments display a single vertical velocity swath (Figure 18a and c) that covers the eastern portion of the observed rotation track. The probabilities in the PAR experiment are stronger than the KTLX experiment, implying larger vertical velocities in the

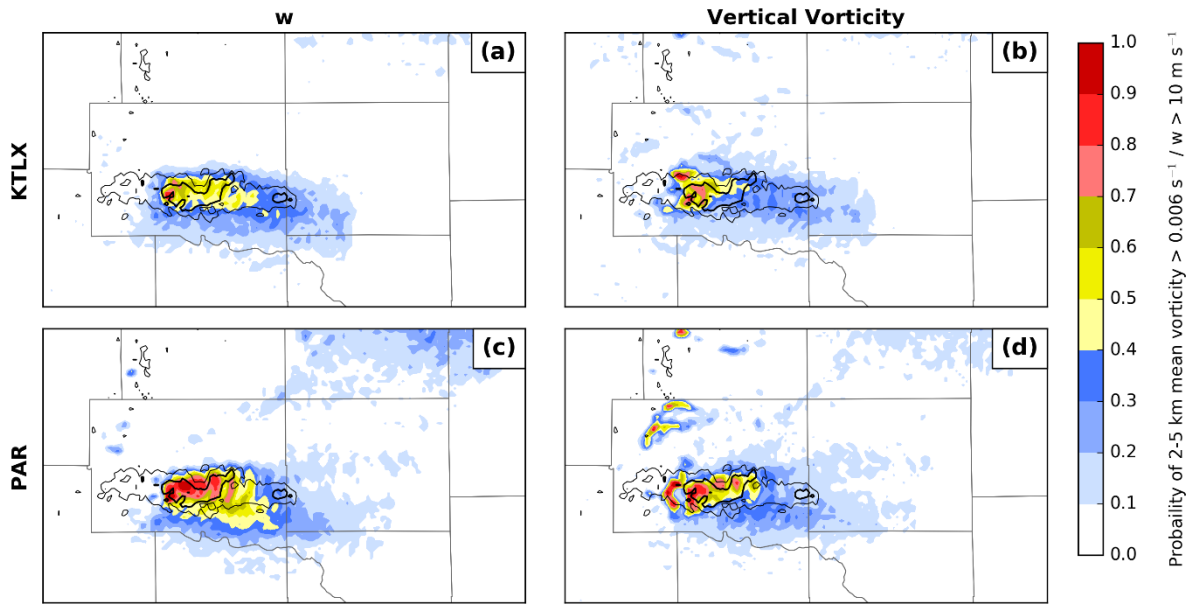

Figure 18: As Figure 17, but with probability of 2-5 km mean vertical velocity > 10 m s-1 (left column) and probability of 2-5 km mean vertical vorticity > 0.006 s-1 (right column) for the 2255 UTC initialization only.

updraft in the PAR experiment. The same is largely true for the vertical vorticity swaths (Figure 18b and d). The probabilities in the PAR experiment is higher than the KTLX experiment, implying stronger vertical vorticity in the PAR experiment. For both variables, there are differences at the initial time (the far western end of the forecast swath), but the largest differences appear to be during the forecast. This implies that not only does the PAR DA directly produce stronger mesocyclones, but it also produces conditions that keep the mesocyclone stronger into the forecast.

One mechanism by which this might happen is by strengthening the mid-level temperature perturbation associated with the updraft (Figure 19). At 2310 UTC, the 2-5 km mean potential temperature increment in all experiments shows a positive-negative dipole on the southern flank of the storm, implying a relocation of the updraft to the southwest. The increments to all fields are generally stronger in the PAR-based experiments than in the KTLX

experiment, which implies that the PAR experiment does indeed create a stronger mid-level temperature perturbation associated with the updraft. Also, all experiments display a cyclonic pattern to the 2-5 km mean wind increments, implying a strengthening of the mid-level mesocyclone. However, the cyclonic pattern in the KTLX experiment is much weaker, and there is a weak anti-cyclonic pattern, implying a weakening of vorticity, just to the north of the cyclonic increments. Furthermore, the positive increments in vertical velocity are generally better co-located with the positive increments in potential temperature and vertical vorticity in the PAR-

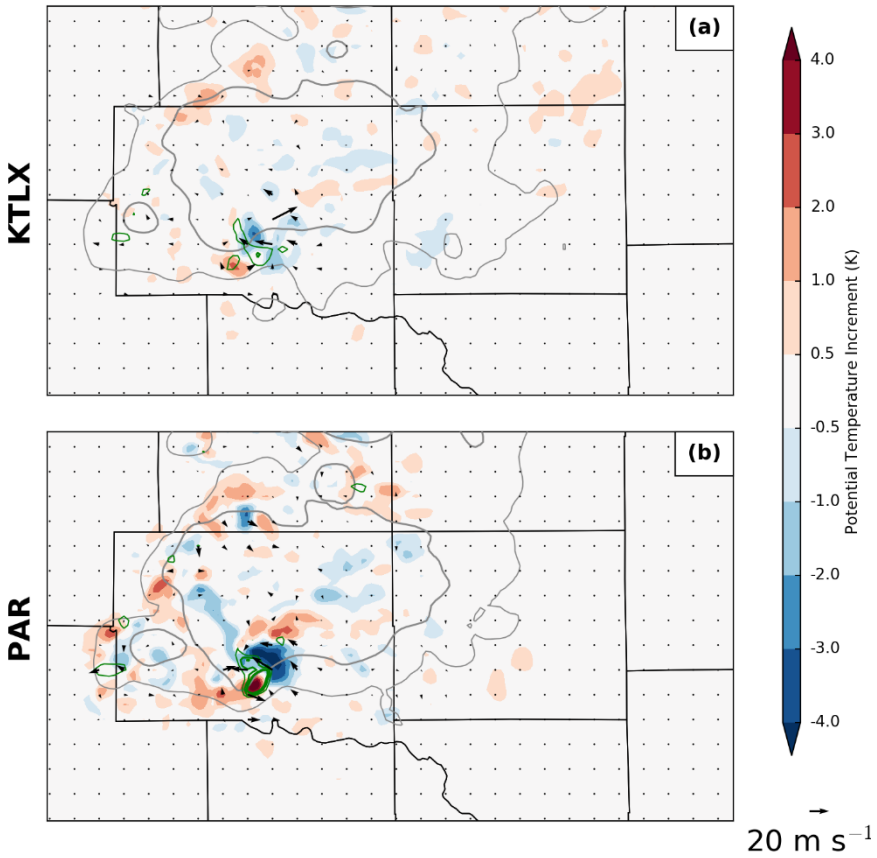

Figure 19: Ensemble mean increment of 2-5 km mean potential temperature at the 2310 UTC analysis from the KTLX (a) and PAR (b) experiments. The ensemble mean 2-5 km horizontal wind increment is given in vectors, the vertical wind increment is given in green contours, and ensemble mean forecast reflectivity is contoured in gray.

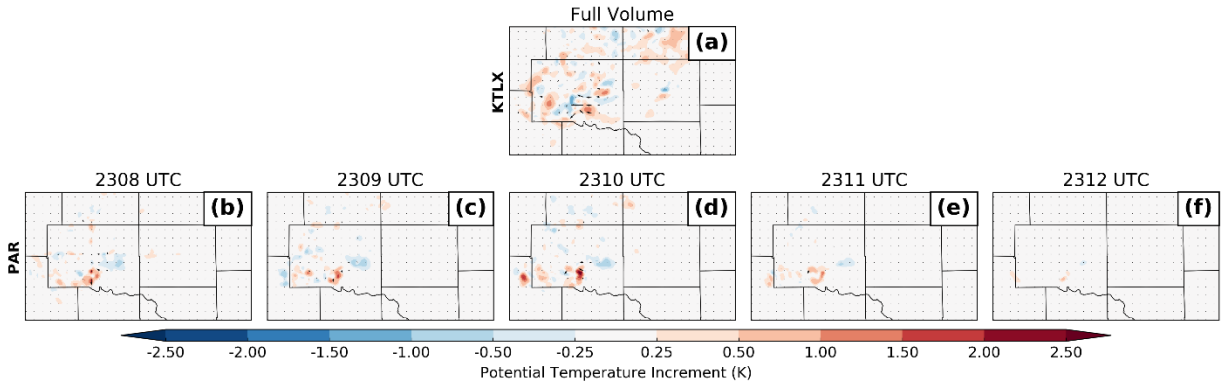

Figure 20: Lowest-model-level potential temperature increments at the 2310 UTC analysis. The full increments from the KTLX experiment are given in (a), and the increments for the PAR experiment by radar volume are given in (b-f). Note that none of these include increments from the Oklahoma Mesonet DA.

based experiments, implying that the PAR experiments develop more physically intuitive covariance structures than the KTLX experiment.

Additionally, we can look at the contributions to the increments from each PAR volume to determine which volumes affect the model state the most (Figure 20). For the PAR experiment, each of the first three volumes assimilated (the 2308, 2309, and 2310 UTC volumes) contributes approximately the same amount to the total increment, whereas the last two (the 2311 and 2312 UTC volumes) contribute very little. Some of this can be explained by the temporal localization weight, which increases from 0.38 at the first volume to 1 at the third volume and then decreases back to 0.38 at the fifth volume. However, this cannot explain why the 2308 UTC volume has much more of an impact on the analysis state than the 2312 UTC volume (compare Figure 20b to Figure 20f), even though they both receive the same temporal weight in the 2310 UTC analysis.

Ideally, each analysis would have roughly equal contributions from all radar volumes, as this means the DA is using all the information available to it. Thus, it is helpful to look at the reasons that some volumes contribute very little to the analysis. To this end, we can look at how the prior reflectivity changes as more volumes of radar observations are assimilated (Figure 21).

For the 2310 UTC analysis for the PAR experiment, the prior mean reflectivity is too high outside of the observed storm and too low inside the observed storm (Figure 21a-e). After the first volume (the one valid at 2308 UTC) is assimilated (Figure 21f-j), the magnitude of the mean reflectivity innovation is much smaller, indicating that the priors have been updated as expected. Because of the time localization, the weight applied to the observations decreases towards the 2310 UTC volume and goes to 0 at the 2311 UTC volume. However, the weight on the 2309 UTC volume from assimilating the 2308 UTC volume is quite high, so the magnitude of the innovations from the 2309 UTC volume are decreased a large amount just by assimilating the 2308 UTC

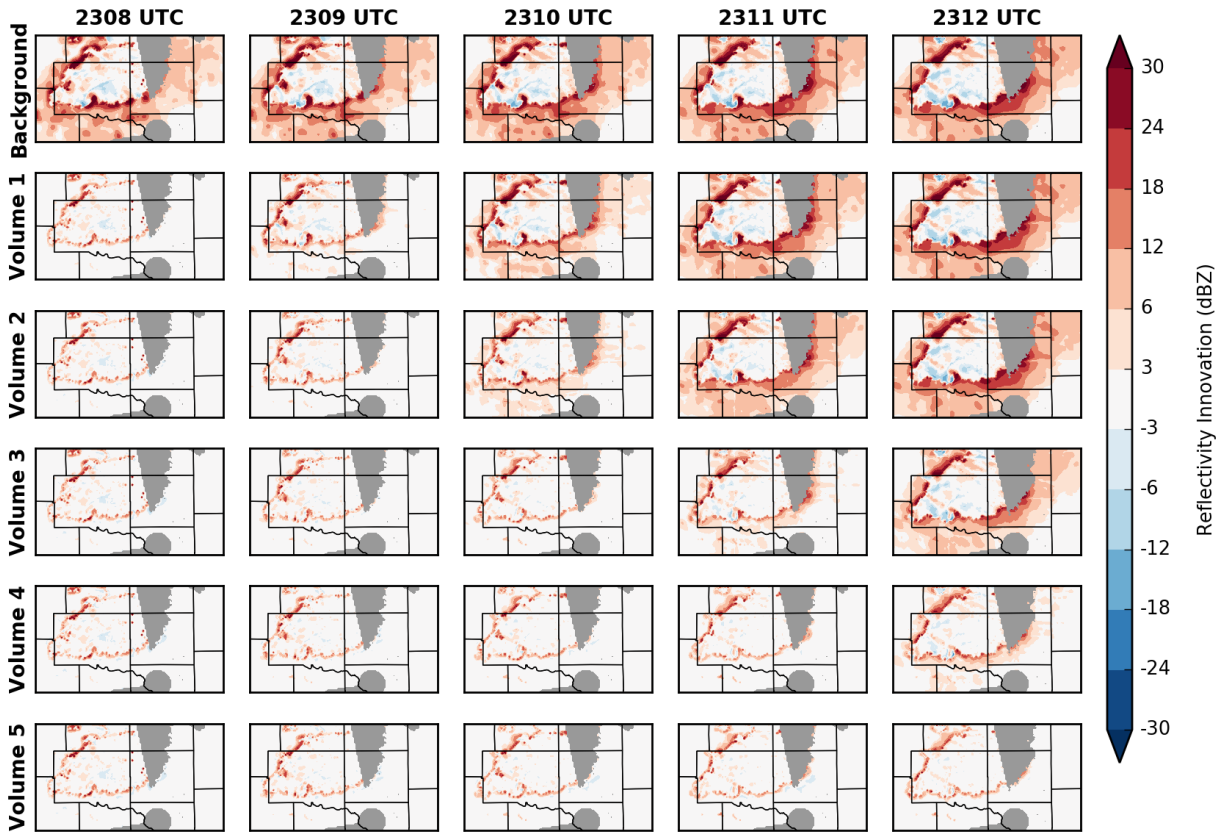

Figure 21: Progression of innovation in the reflectivity priors at the lowest elevation angle during the 2310 UTC analysis for the PAR experiment. The first row is the background, the second row is after assimilating the first volume of radar data, the third row is after assimilating the second volume, etc.

volume. This means that when assimilating the second volume, valid at 2309 UTC (Figure 21k-o), the innovation magnitude has already been decreased by the assimilation of the first volume, which decreases the analysis increments from the second volume. This process of the assimilation of one volume decreasing the innovation, and therefore the increment for the next volume, continues for the rest of the assimilation procedure. The same is essentially true for the prior reflectivity spread (not shown).

This motivates further experiments in order to determine the contribution of various factors to the distribution of increments among radar volumes (Figure 22). The fourth and fifth volumes in the PAR+IVCI experiment (Figure 22d,e) do not contribute significantly more to the potential temperature increments than the same volumes from the PAR experiment. The PAR+IVCI experiment applies RTPS inflation after every volume of radar assimilated in order to

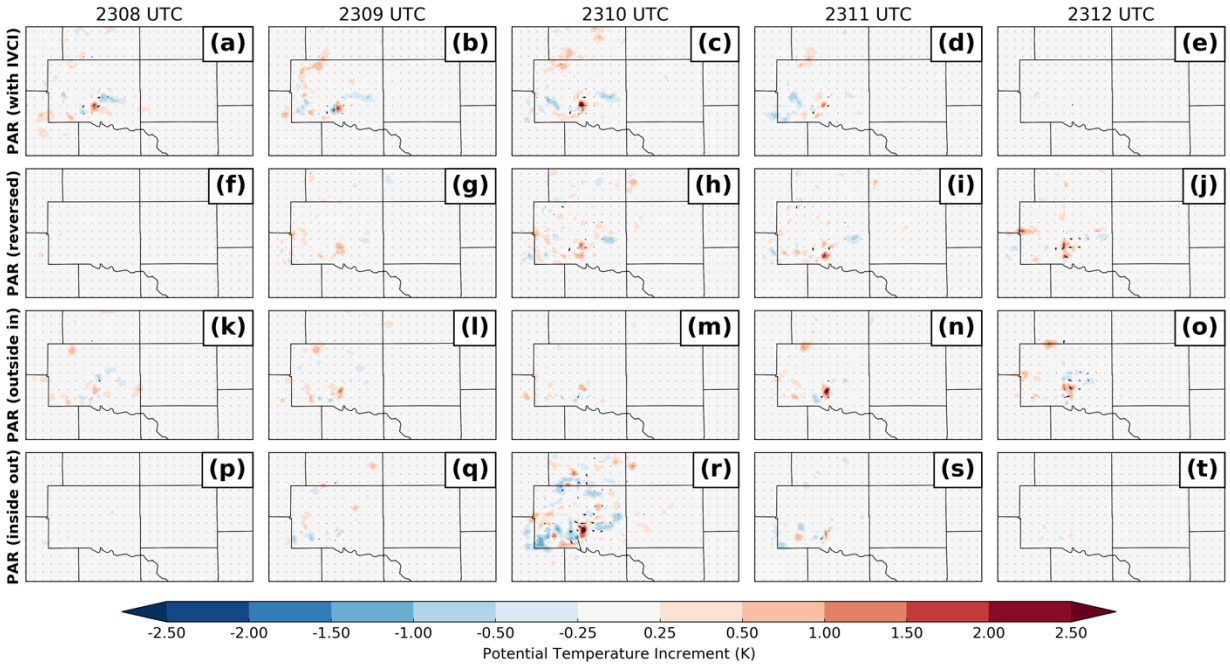

Figure 22: As Figure 20, but for the PAR+IVCI (a-e), PAR_reversed (f-j), PAR_inout (k-o), and PAR_outin (p-t) experiments.

bring back the spread inherently lost during the DA process. Thus, the spread reduction by itself is not a significant cause of the uneven distribution of increments.

Another option is to assimilate the radar volumes in non-chronological order. Assimilating the volumes in reverse order (PAR_reversed; Figure 22f-j) still retains the behavior where assimilating one volume reduces the prior innovation in temporally adjacent volumes. However, assimilating more recent observations first might be desirable, as it puts a higher weight on the more recent observations. Additionally, the initial spread in the prior reflectivity is greater at the end of the DA window, which might allow for larger ensemble mean increments. Furthermore, assimilating the volumes in inside-out order (PAR_inout; Figure 22p-t) results in much of the increments coming from the volume valid at the analysis time; closer to a 3D EnKF implementation. Finally, assimilating volumes in outside-in order (PAR_outin; Figure 22k-o) spreads the increments out the most evenly out of any experiment among the five volumes used in the analysis. The disadvantage of the PAR_outin experiment is that the volume valid at the analysis time is assimilated last, so it contributes the least to the analysis.

Next, we can examine how the alternate assimilation orders affect the forecasts of updraft helicity (Figure 23). As in the PAR experiment, (cf. Figure 17c and d) all experiments display a large probability of UH > 100 $m^2 s^{-2}$ early in the forecast, tapering off as the forecast progresses. However, all experiments using the alternate assimilation orders appear to have improved track forecasts over the original PAR experiment. In particular, the overprediction on the southern periphery of the UH swath is reduced in the 2310 UTC forecasts from the experiments that prioritize later volumes over earlier volumes (PAR_reversed, PAR_inout, and PAR_outin). This suggests there is some utility to assimilating later volumes first.

One other noteworthy aspect to these experiments is that the sign of the increments is generally consistent across the volumes (note, for example, the consistent placement of the positive and negative increments in Figure 22a-d). This implies that further observations do not typically reverse the underlying correlations in the ensemble, meaning that additional

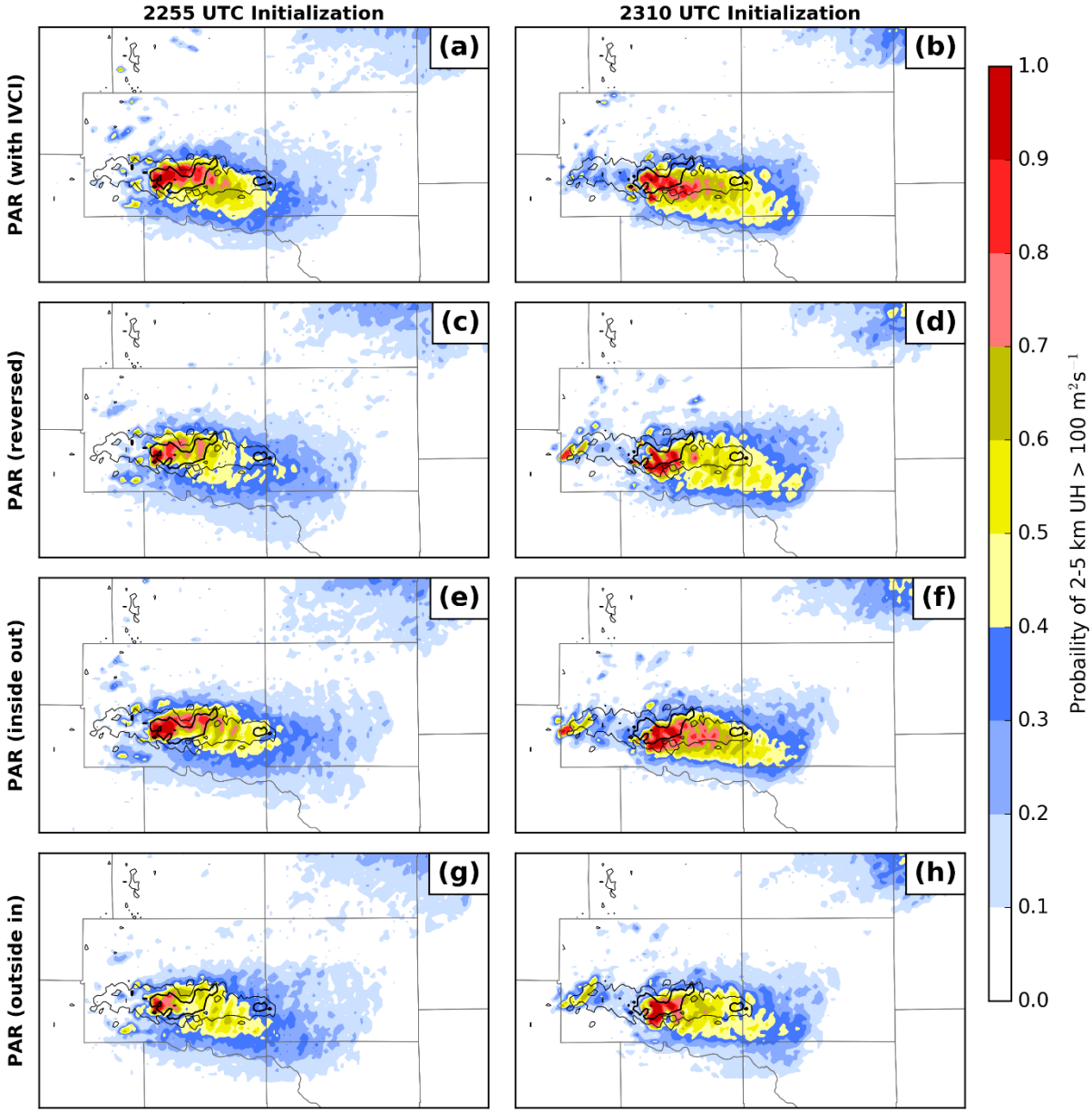

Figure 23: As Figure 17, but for the PAR+IVCI (a,b), PAR_reversed (c,d), PAR_inout (e,f), and PAR_outin (g,h) experiments.

observations simply make the total increments larger. Thus, to some extent, the larger magnitude increments in the PAR experiment are simply because there are more observations to assimilate.

4.2.4. Cold Pool Differences

In order to explain some of the differences noted above, we examine the cold pool analyses (Figure 24). As in Supinie et al. (2017), the coldest analyzed cold pools are generally in the PAR-assimilating experiments, though this is not universally the case. The PAR+IVCI

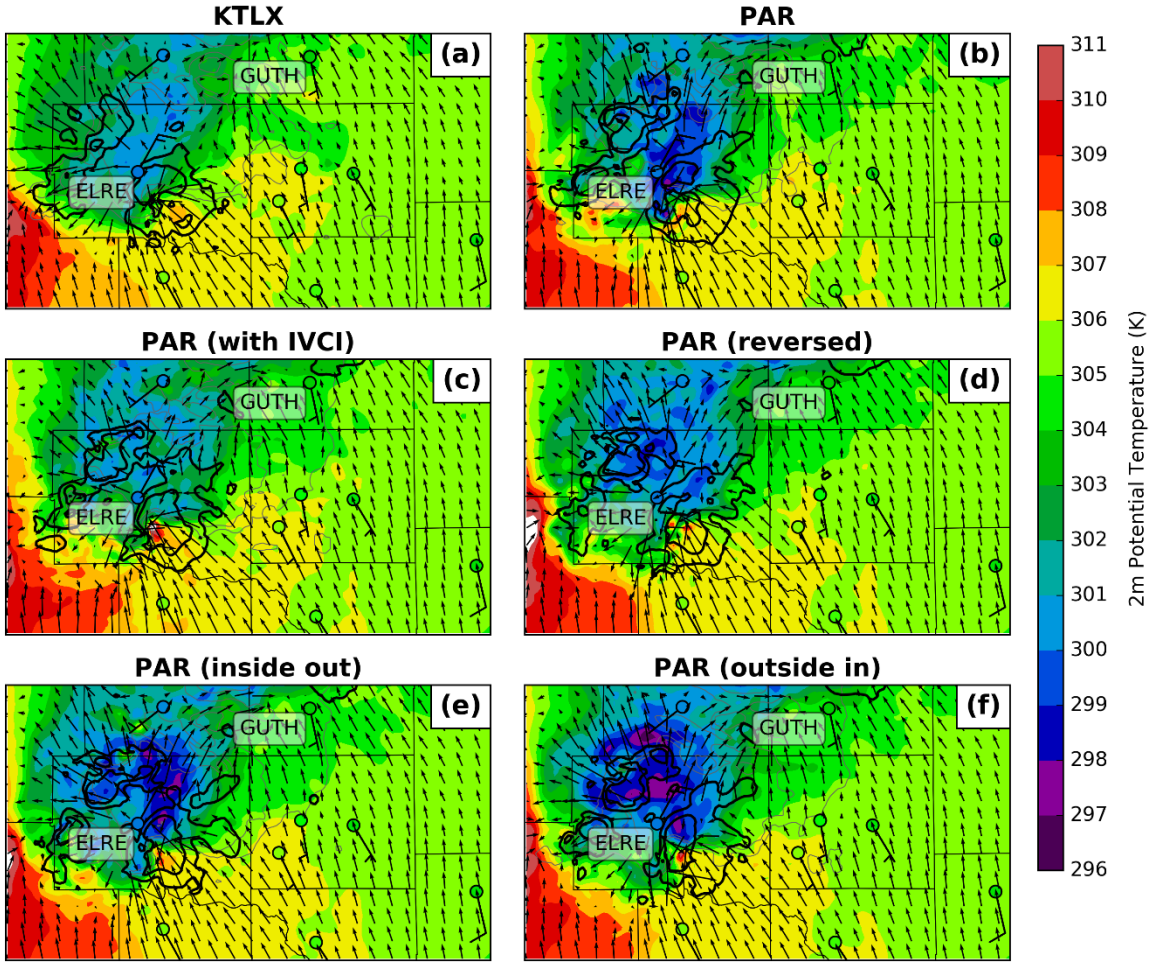

Figure 24: Analysis ensemble mean potential temperature (color fills) and horizontal wind (vectors) at the lowest model level for the KTLX (a), PAR (b), PAR+IVCI (c), PAR_reversed (d), PAR_inout (e), and PAR_outin (f) experiments, valid 2310 UTC. The 10 and 20 m s^{-1} contours of column-max vertical velocity are given in black, and the 20 and 40 dBZ contours of ensemble mean reflectivity are given in gray. Additionally, Oklahoma Mesonet observations are plotted in colored circles for potential temperature and barbs for winds.

experiment (Figure 24c) as a cold pool strength comparable to the KTLX experiment (Figure 24a), while the other experiments are much colder. All experiments are close to the observed potential temperatures at the Oklahoma Mesonet sites; most of the differences between experiments lie between the sites. One major difference that impacts the forecasts is the temperature in the rear-flank cold pool in the PAR experiment (Figure 24b) is much colder than any of the other experiments. This difference in the PAR experiment persists into the early part of the forecast (not shown), and the enhanced forcing for ascent from the strong cold pool may be responsible for the larger updraft helicity probably in the PAR experiment south of the observed MRMS

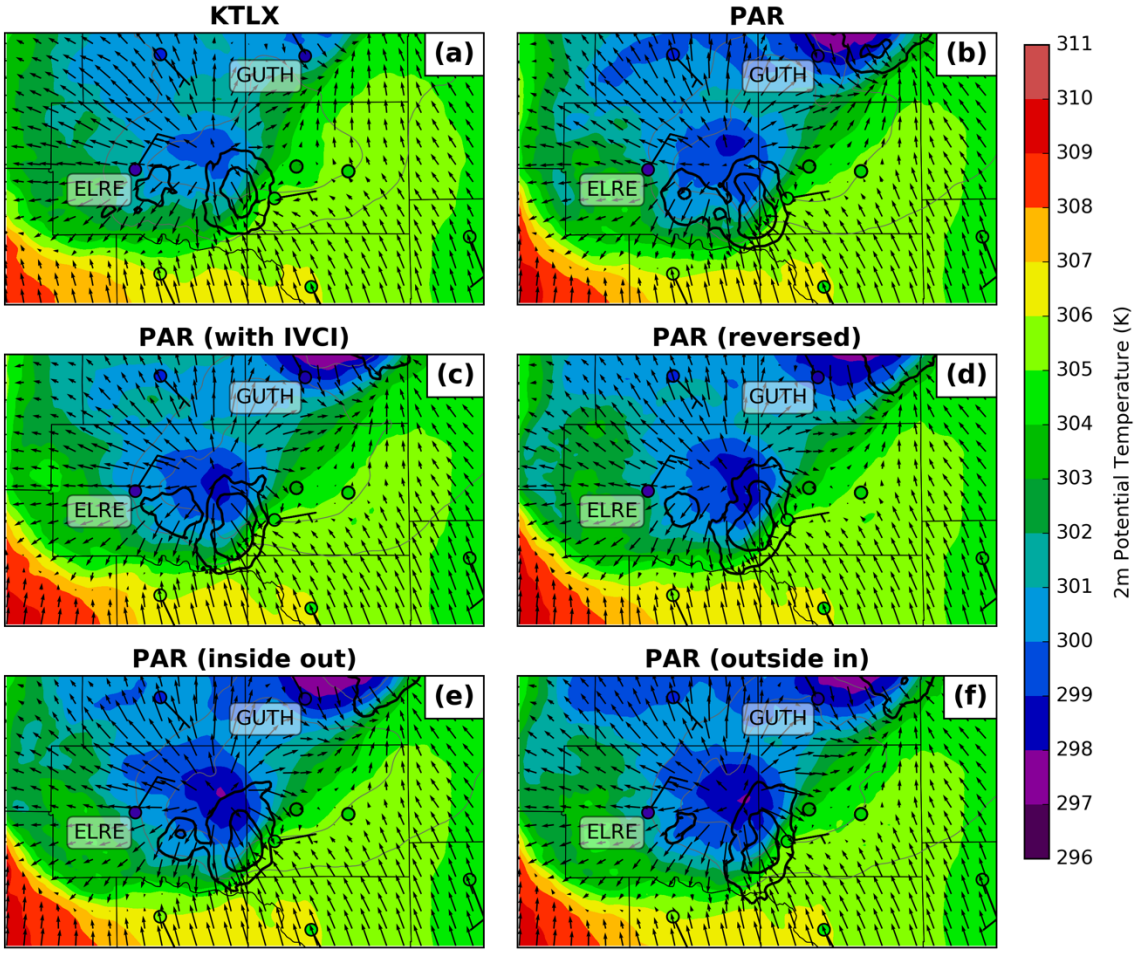

Figure 25: As Figure 24, but for a 30-minute forecast, valid 2340 UTC.

rotation track (see Figure 17d). Thus, the reason the alternate assimilation orders improve the track forecast for the mesocyclone is because of differences in analyzing the rear-flank cold pool.

Next, we examine the surface cold pool strength in the 30-minute forecast from the 2310 UTC analysis (Figure 25). Again, the PAR-based experiments have a colder cold pool for the El Reno supercell. It is difficult to tell which is closer to the Oklahoma Mesonet observations, as none of the experiments capture the training convection that occurred over the region. This convection reinforced the real-world surface cold pool, leading to much colder temperatures observed at the El Reno Mesonet site (labeled "ELRE") than captured in any of the forecasts. For the cell to the north of the main supercell (the northern edge of Figure 25), the PAR DA also results in a colder cold pool, more in line with the Guthrie Mesonet site (labeled "GUTH"). This is also likely related to the PAR experiment better capturing the reflectivity northeast of the El Reno supercell as in Figure 16b.

The explanation that Supinie et al. (2017) gave for the colder cold pool in their PAR experiment is that in the low levels, one would expect a negative correlation between temperature and reflectivity. Thus, if the data assimilation systematically increases reflectivity, as one would expect when the forecast mean reflectivity is smoother and lower in magnitude than the observed storm, it should also systematically decrease low-level temperature. Thus, more PAR observations than KTLX observations results in more decreases to low-level temperature in the PAR experiment. This explanation was originally proposed by Dowell et al. (2011), who found this to be occurring in their real-data DA and forecast experiments. This explanation is also consistent with the finding here that additional observations tend to increase the magnitude of increments previously applied to the ensemble state.

This DA-based mechanism for colder cold pools is occurring in these experiments, though to a much lesser degree (Figure 26). Most potential temperature increments at the surface in both experiments are positive, particularly in the cold pool and in the final location of the rear-flank downdraft. The PAR experiment (Figure 26b) does have some areas of overall negative potential temperature increments south of the rear flank of the storm and near the rear-flank gust front which do not appear in the KTLX experiment (Figure 26a). However, overall magnitudes of these negative increments are low. While Figure 26 shows the sum of the increments applied

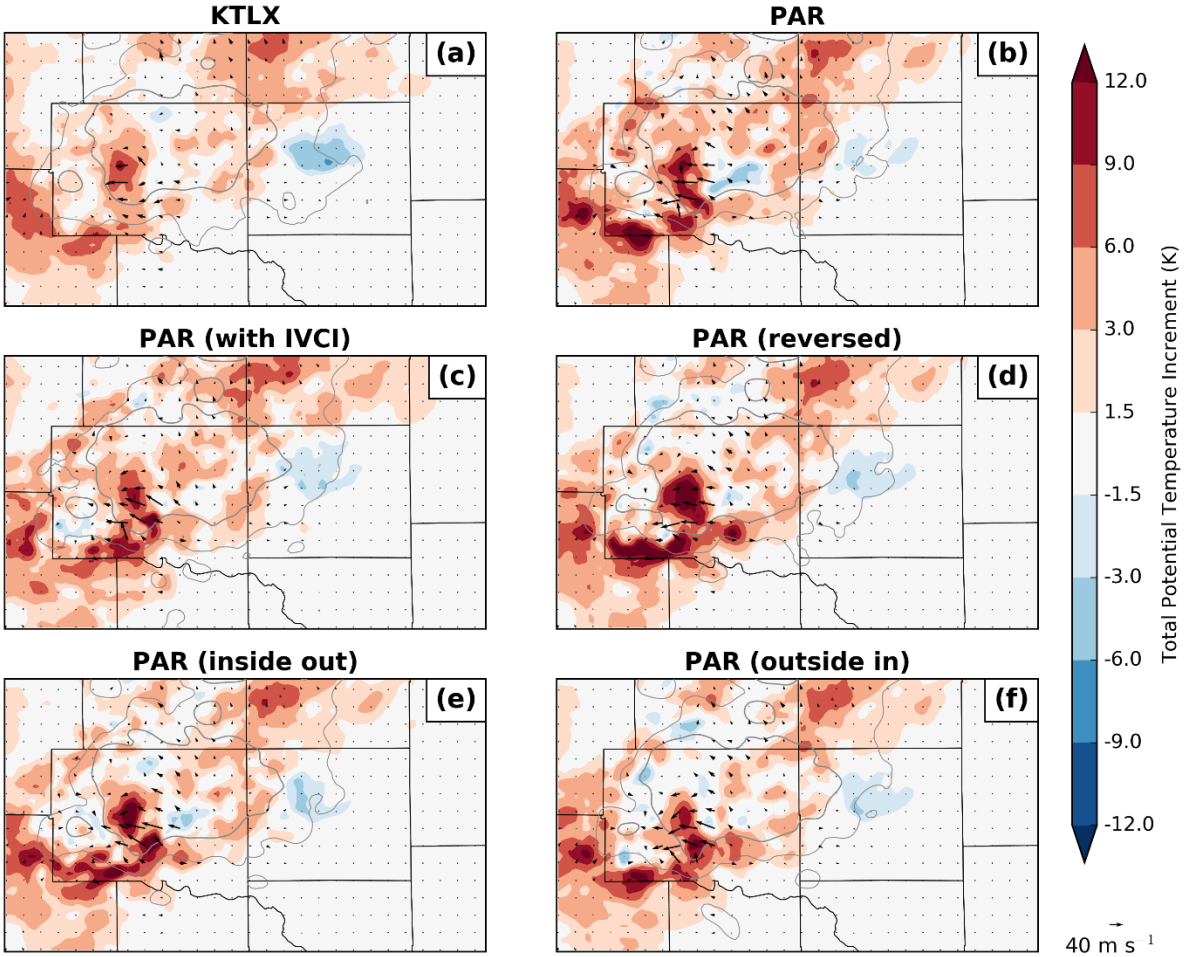

Figure 26: Increments to ensemble mean potential temperature (color fills) and wind (vectors) at the lowest model level, summed over the entire DA period, for the KTLX (a), PAR (b), PAR+IVCI (c), and PAR_reversed (d) experiments. The 20 and 40 dBZ contours of ensemble mean reflectivity at 2310 UTC are given in light gray.

throughout the DA period, the increments applied in individual analyses (not shown) are generally consistent with the sum in showing large areas of positive increments in the cold pool and very few negative increments. Thus, the DA by itself is not consistent with the PAR experiment's much colder cold pool.

If not the DA, another reason for the PAR experiment cold pool being colder than the KTLX experiment is if the PAR DA changes the model state in such a way as to promote greater latent cooling via different microphysical processes than in the KTLX experiment (Figure 27). The mean thermal energy change due to evaporative cooling by rain in the downdraft for the PAR experiment is larger in magnitude than the KTLX experiment. Here, the downdraft is defined as areas below 4 km AGL where vertical velocity is less than -0.5 m s^{-1} (Dawson et al. 2010). At many

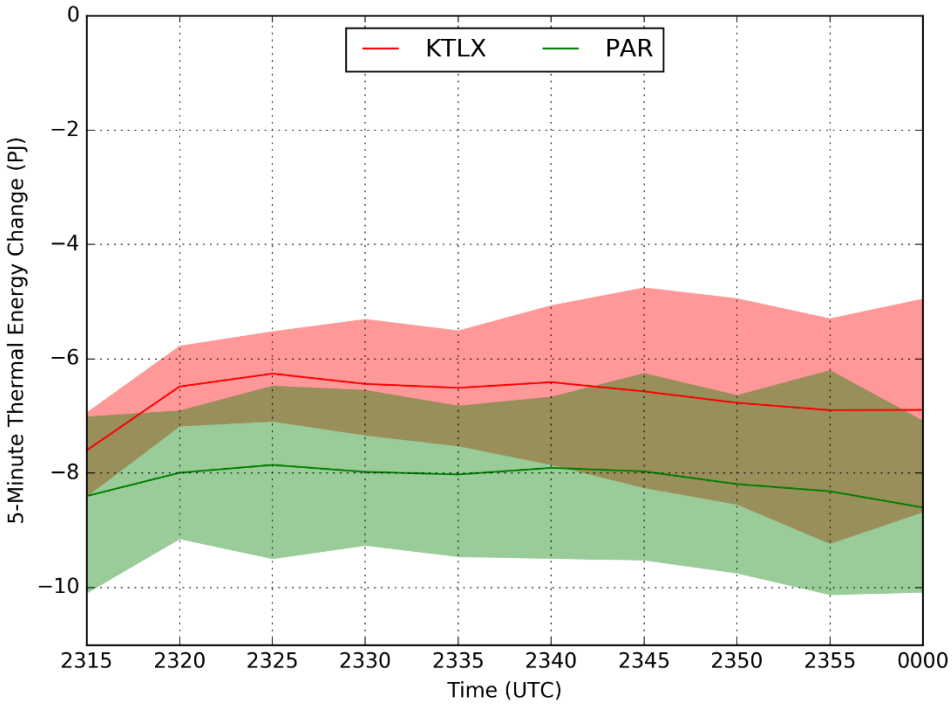

Figure 27: 5-minute thermal energy change (in petajoules) by evaporation of rain in the low-level downdraft during the 2310 UTC forecast. The low-level downdraft is defined as regions below 4 km AGL where vertical velocity is less than -0.5 m s^{-1}. The solid line is the ensemble mean and the shading is the 5^{th}-95^{th} percentiles.

times, particularly early in the forecast, the mean for the KTLX experiment is outside the 5th-95th percentile range for the ensemble members from the PAR experiment and vice versa. This indicates that the ensembles have significantly different evaporative cooling rates. The other PAR+IVCI, PAR_reversed, PAR_inout, and PAR_outin experiments do not differ significantly from the PAR experiment and are therefore not shown for clarity. Other microphysical processes than evaporative cooling by rain were examined, but none had nearly as big an effect. Given that evaporative cooling is the dominant process, there are two ways larger evaporative cooling in the PAR experiments could occur: 1) the rain drop-size distribution (DSD) in the PAR experiment

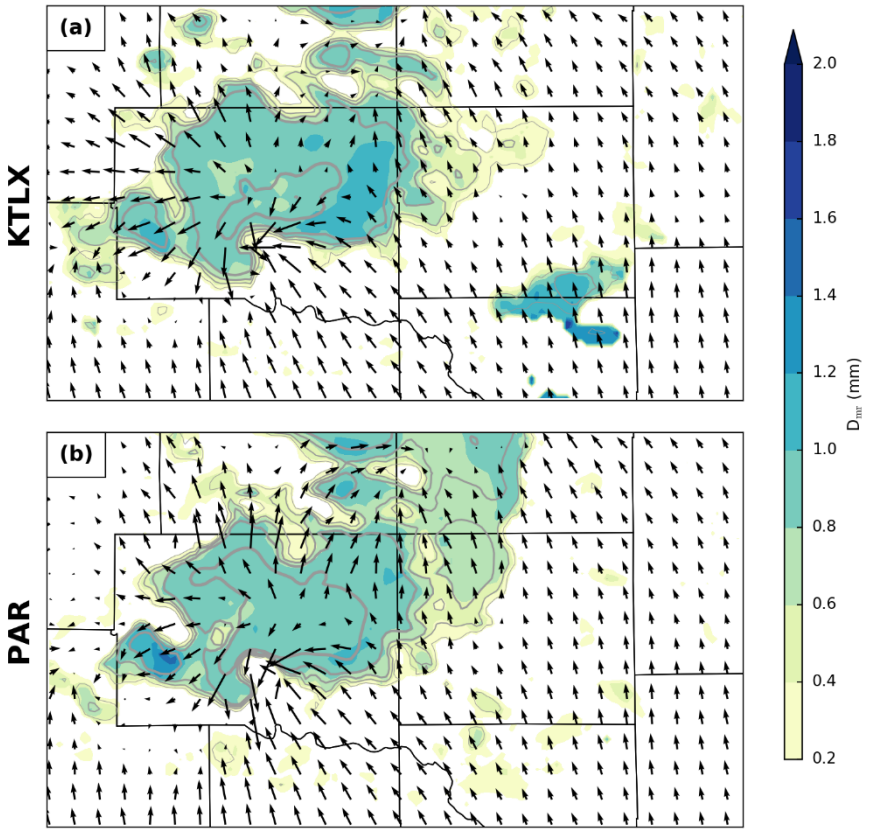

Figure 28: Ensemble mean mean-mass diameter for rain at the first model level above the surface at the 2310 UTC analysis for the KTLX (a), PAR (b) experiments. Analyzed reflectivity is given in gray contours, and horizontal wind is given in vectors.

may have more small drops, or 2) the ambient environment may be drier. We will examine both hypotheses here.

To evaluate differences in the rain DSD, we examine the mass-weighted mean diameter of the DSD (Figure 28). Both KTLX and PAR experiments show a region of slightly larger mean-mass diameter along the forward flank of the main supercell, indicative of larger drops. This is a common feature of storms in strongly sheared environments and is an indicator of size sorting and commonly manifests in differential reflectivity (ZDR) as a ZDR arc (Kumjian and Ryzhkov 2008). This signature is weaker than one might expect; the reason for this is that low-level ZDR signatures are much more sensitive to graupel size sorting than rain size sorting (Dawson et al. 2014; Putnam et al. 2017). Thompson microphysics is single-moment in graupel and therefore cannot reproduce size-sorting behavior with graupel. In some places, the PAR experiment (Figure 28b) has a slightly lower mean-mass diameter than the KTLX experiment (Figure 28a). However, these differences are not that large, and many areas of the storm have the same or slightly higher mean-mass diameter, indicating more large drops in the PAR experiment. The supplemental PAR experiments do not show significant differences from either the PAR or the KTLX experiment. Overall, this does not provide strong evidence in support of hypothesis (1) above.

With respect to the ambient environment around the supercell, there are major differences in the ensemble mean relative humidity increments over the lowest ~1 km AGL (Figure 29). The DA in the PAR experiment results in much larger decreases in water vapor mixing ratio in the inflow region (Figure 29b), whereas increments in the same location in the KTLX experiment (Figure 29a) are much closer to neutral. This difference is likely the result of more aggressive clear-air suppression in the PAR-assimilating experiments. Additionally, drier inflow

has previously been associated with colder cold pools via evaporative cooling (Parker 2014), which is consistent with Figure 27. This supports hypothesis (2) above.

The choice of microphysics scheme has a large influence on the cold pool strength, primarily driven by cooling from evaporating rain and cloud water (Dawson et al. 2010). The experiments run by Supinie et al. (2017) used Milbrandt and Yau 2-moment microphysics, which had the smallest evaporative cooling out of the experiments by Dawson et al. (2010), and therefore had the weakest cold pools. Thus, it is possible that this evaporative cooling mechanism was much weaker in experiments by Supinie et al. (2017) because of their choice of microphysics scheme for their experiments. Additionally, the different process might explain why the cold pool differences affected the mesocyclone tracks in these experiments and did not affect the

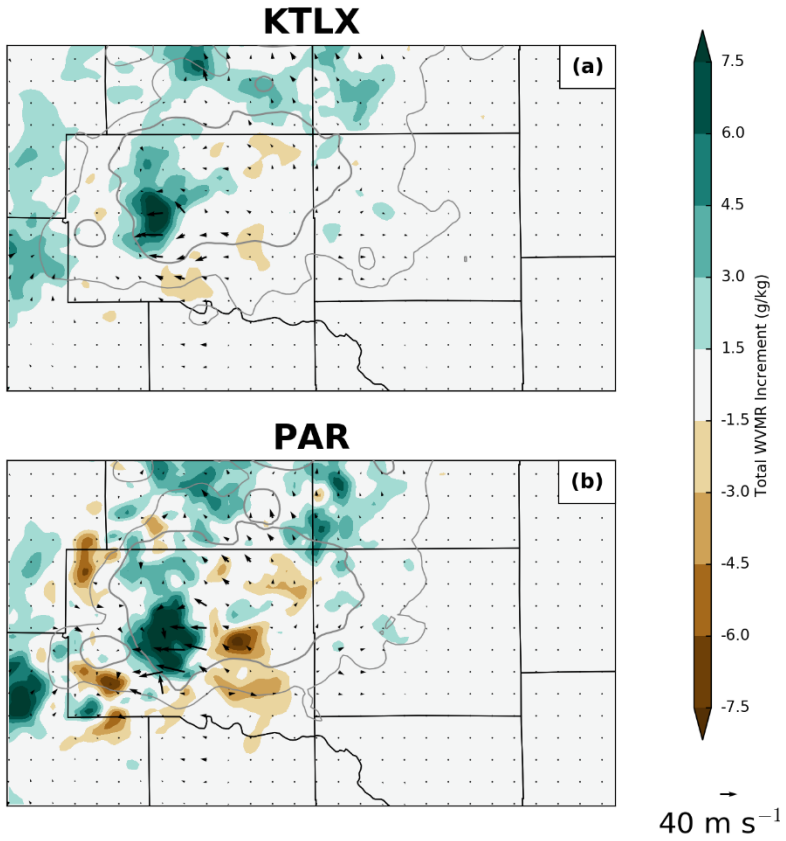

Figure 29: As Figure 26, but for 0-1 km mean water vapor mixing ratio.

mesocyclones in experiments by Supinie et al. (2017). The DA covariance process produces a change to the initial conditions which can be quickly removed by the microphysics scheme. However, the evaporative cooling process can modify the forecast throughout the forecast, producing a greater effect.

4.3. Summary

This chapter presents results from storm-scale DA and forecast experiments, one using 5-min volumes from the operational WSR-88D (KTLX), and another using 70-sec volumes from the NWRT PAR. Data were taken from the 31 May 2013 central Oklahoma tornado and flash flood event. The 4DEnSRF algorithm is used to assimilate reflectivity and radial velocity observations from both radars into a heterogeneous environment initialized from an outer mesoscale ensemble that ultimately gets its diversity from the NCEP GEFS ensemble. All experiments assimilate data for 50 min, producing analyses every 5 min starting at 2225 UTC. DA windows are also 5 min in length and centered on the analysis time. Forecasts are launched at 2255 and 2310 UTC after 35 and 50 minutes of radar DA. The objective is to examine the benefit of assimilating PAR observations over the current WSR-88D observations on a different case to Supinie et al. (2017) and to provide insight how differences between the two experiments arise in the 4DEnSRF. To do this, four experiments have been completed: one assimilating KTLX data, one assimilating PAR data, one assimilating PAR data, but with the observation order reversed, and one assimilating PAR data with RTPS covariance inflation applied between each volume of PAR data.

The probability of updraft helicity is generally greater in experiments that assimilate PAR data than the KTLX experiment, which is primarily driven by a stronger updraft in the PAR-based

experiments. This in turn is driven by a combination of the DA directly creating a stronger updraft and the DA also creating a stronger mid-level temperature perturbation associated with the updraft. The stronger mid-level temperature perturbation implies more physically realistic covariances between temperature and reflectivity. The stronger temperature perturbation allows the updraft strength to be maintained throughout the forecast.

Additionally, the way in which the 4DEnSRF uses the large number of PAR observations is examined. In the PAR and PAR_reversed experiments, most of the changes to the ensemble state come from the first three volumes of data that are assimilated. In the PAR_inout experiment, most of the increments come from the first volume assimilated. This is because of the changes in the prior reflectivity field as more observations are assimilated. Assimilation of one volume decreases the prior innovation and spread for temporally adjacent volumes, so assimilating the next volume has proportionally less impact. However, in the PAR_outin experiment, increments are distributed approximately evenly over the volumes considered in the analysis. This is because the assimilation order minimizes the updates to one volume because of assimilating previous volumes. The volumes contributing to the increments in the PAR+IVCI experiment are more similar to the PAR and PAR_reversed experiments than the others, indicating that spread reduction in the analysis is not the primary reason for the uneven distribution of increments among volumes.

Some effects of this are apparent in the analyses and forecasts of the cold pool. The PAR experiment has a much stronger rear-flank cold pool which results in a UH swath that is too far south compared to the observed. The alternate assimilation orders result in a warmer analysis in the rear-flank cold pool, which results in a better forecast of the mesocyclone track. While the

rear-flank cold pool specifically is colder in the PAR experiment, all the PAR-assimilating experiments develop a stronger cold pool in the free forecasts. The stronger cold pool in the forecast is driven mostly by additional evaporative cooling of rain in the PAR experiments. This in turn is due primarily to the drier environment created by the PAR DA to the south of the storm, which drives the evaporative cooling. Changes in rain drop-size distributions are probably not responsible for the differences in the behavior between the KTLX and PAR experiments. Broadly speaking, these results are consistent with previous studies, such as Supinie et al. (2017) and Yussouf and Stensrud (2010), which show more rapid spin-up of convection using 1-min PAR data as compared to ~5-min WSR-88D data. However, some key differences are apparent. In these experiments, the track forecast for the mesocyclone was not as good, which was traced to the microphysics handling of evaporative cooling. The different microphysics scheme used here (Thompson) vs. Supinie et al. (2017) (Milbrandt-Yau 2-moment) could explain some of the differences in the track forecasts created by experiments in the respective papers.

Chapter 5: Conclusions[2]

5.1. Summary

This work looked at data assimilation and forecast experiments for two cases using a 4D ensemble square root filter (EnSRF) and high-temporal resolution data from an experimental phased array radar. The control experiment for both cases uses an operational WSR-88D radar. The primary goal is to examine whether the PAR data improve storm-scale forecasts for severe convective storms and examine the reasons for those improvements. The findings are that generally, PAR data improve forecasts of mesocyclone tracks and, as far as can be inferred from the ensemble probabilities, generally increases the intensity of the mesocyclones. Additionally, the PAR data improve the reflectivity forecasts, as measured by several skill scores. The higher intensity of the mesocyclones in the PAR forecasts is tied to better capturing the mid-level temperature perturbation associated with the updraft. The track forecasts related to the cold-pool behavior, with warmer cold pools generally improving the track forecasts. The cold pool forecasts are then related to the mixing ratio in the inflow. Experiments with lower mixing ratio in the inflow tended to have colder cold pools because of enhanced evaporative cooling by rain, which is consistent with past work.

Other experiments revealed some previously unseen behavior with the assimilation of several volumes of data during one analysis. Chiefly, the default behavior in the ARPS 4D EnSRF is to assimilate volumes in chronological order, which results in the earliest volumes assimilated in each DA cycle having the most effect on the analysis, while the latest volumes assimilated in

[2] Parts of this chapter were published in section 5 of Supinie et al. (2017)

each DA cycle having little effect on the analysis. The primary reason for this is that updating the priors for one time updates the priors for adjacent times, as well. This lessens the innovation for subsequent volumes, which therefore lessens the increments. Assimilating volumes in an alternate order can change this. Assimilating volumes in reverse order results in the most weight being placed on the most recent volumes. Assimilating volumes from the outside of the DA window inward to the valid time results in the most even usage of information from all radar volumes in the analysis. Assimilating volumes starting with the one valid at the analysis time and working outwards towards the edge of the DA window results in most of the weight being placed on the volume valid at the analysis time. All three of these methods produce better results than naively assimilating volumes in chronological order.

The overall results from these two cases shows promise for the PAR data in initializing storm-scale forecasts of severe convective storms. The ultimate goal of these forecasts is to extend severe weather warning lead times (Stensrud et al. 2009). A hypothetical real-time prediction system would require at least 12 volume scans, or about 45 minutes of operational WSR-88D data to produce a reasonably good forecast. In contrast, the length of the NWRT PAR data assimilation period could be reduced to as short as 15 minutes, potentially increasing severe weather forecast lead times. These rapid-scan observations would be particularly beneficial in a rapidly evolving severe weather situation. Additionally, the impact of the unique flexible adaptive scanning capability from NWRT PAR [e.g., dense vertical scanning and elevation prioritized scanning strategies (Heinselman and Torres 2011)], particularly when the storm is close to or far away from the radar, on the model analyses and forecasts also needs to be investigated. Moreover, errors in storm motion in storm-scale modeling warrant further investigation.

Accurate forecasts of storm position are necessary for skillful prediction of storm-associated hazards, a primary research goal of WoF.

5.2. Future Research

The results from Chapter 3: are consistent with the findings from the OSSE study in Yussouf and Stensrud (2010). However, the case we used for this study involves only two supercell storms. To lay a foundation for the value of the next generation PAR technology beyond the current WSR-88D network in storm-scale modeling, more rigorous testing on how to best assimilate PAR observations in a variety of severe weather systems (e.g., quasi-linear convective systems, microbursts, hailstorms, nontornadic as well as tornadic supercell storms, etc.) is still needed.

Furthermore, the results of Chapter 4: motivate the question of the most effective way to use very large numbers of observations in an ensemble DA system, which would be a good topic for future research. Other potential topics include examination of the flash flooding aspect the 31 May 2013 case, which was not a priority here. These questions will be useful to address under Warn-on-Forecast to make the best possible use of the observations. Additionally, with the latest series of GOES satellites (GOES-16 and 17) having the capability to send images every 30 seconds over a mesoscale window, the assimilation of rapid-scan satellite imagery would be another good topic for future research.

Another potential avenue for future research is to examine alternate ensemble DA methods that remove some of the assumptions built into the EnKF and related methods. One such method is the particle filter (Poterjoy 2016; Poterjoy et al. 2017, 2019). The particle filter updates the ensemble using weighted sums of the prior members, where the weights are derived

from Bayes' theorem. The advantage of using a particle filter is that it makes no assumptions about the prior distribution of ensemble members. This contrasts with Kalman-filter-based methods, which assume a Gaussian form for the prior distribution. Radar reflectivity from a storm-scale model is frequently non-Gaussian (Poterjoy et al. 2017), so particle filters in theory have an advantage in this application. However, the formulation of particle filters results in the ensemble members collapsing onto a single value if the ensemble size is too small. For many NWP applications, the size of the ensemble considered "too small" is so large as to be computationally infeasible without some means of reducing the dimensionality. The local particle filter (Poterjoy 2016; Poterjoy et al. 2017, 2019) is one means to overcome this high dimensionality. As the name implies, it accomplishes this by restricting the region over which it computes the update weights to the region surrounding a given observation. This method is promising, and future work may focus on comparing performance of EnKF-based methods to particle-filter-based methods.

References

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Weather Rev.*, **129**, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **59**, 210–224, https://doi.org/10.1111/j.1600-0870.2006.00216.x.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **61 A**, 72–83, https://doi.org/10.1111/j.1600-0870.2008.00361.x.

Betts, A. K., 1973: A Composite Mesoscale Cumulonimbus Budget. *J. Atmos. Sci.*, **30**, 597–610.

Bluestein, H. B., J. C. Snyder, and J. B. Houser, 2015: A multiscale overview of the El Reno, Oklahoma, tornadic supercell of 31 May 2013. *Weather Forecast.*, **30**, 525–552, https://doi.org/10.1175/WAF-D-14-00152.1.

——, K. J. Thiem, J. C. Snyder, and J. B. Houser, 2019: Tornadogenesis and early Tornado evolution in the El Reno, Oklahoma, supercell on 31 May 2013. *Mon. Weather Rev.*, **147**, 2045–2066, https://doi.org/10.1175/MWR-D-18-0338.1.

Brewster, K. A., M. Hu, M. Xue, and J. Gao, 2005: Efficient Assimilation of Radar Data at High Resolution for Short Range Numerical Weather Prediction. *World Weather Research Program Symposium and Nowcasting and Very Short-Range Forecasting WSN05*, 1–14 http://www.meteo.fr/cic/wsn05/DVD/resumes/longs/3.06-249.pdf.

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. Ocean. Technol.*, **12**, 5–19, https://doi.org/10.1175/1520-0426(1995)012<0005:tomato>2.0.co;2.

Burgers, G., P. Jan van Leeuwen, and G. Evensen, 1998: Analysis Scheme in the Ensemble Kalman Filter. *Mon. Weather Rev.*, **126**, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Charney, J. G., R. Fjörtoft, and J. Von Neumann, 1950: Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, **2**, 237–254, https://doi.org/10.3402/tellusa.v2i4.8607.

Chou, M.-D., 1990: Parameterizations for the Absorption of Solar Radiation by O2 and CO2 with Application to Climate Studies. *J. Clim.*, **3**, 209–217, https://doi.org/10.1175/1520-0442(1990)003<0209:pftaos>2.0.co;2.

——, 1992: A Solar Radiation Model for Use in Climate Studies. *J. Atmos. Sci.*, **49**, 762–772.

——, and M. J. Suarez, 1994: An Efficient Thermal Infrared Radiation Parameterization for Use in General Circulation Models. 85.

Cressman, G. P., 1959: An Operational Objective Analysis System. *Mon. Weather Rev.*, **87**, 367–374, https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Dawson, D. T., M. Xue, J. A. Milbrandt, and M. K. Yau, 2010: Comparison of Evaporation and Cold Pool Development between Single-Moment and Multimoment Bulk Microphysics Schemes in Idealized Simulations of Tornadic Thunderstorms. *Mon. Weather Rev.*, **138**, 1152–1171, https://doi.org/10.1175/2009MWR2956.1.

——, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the Environmental Low-Level Wind Profile on Ensemble Forecasts of the 4 May 2007 Greensburg, Kansas, Tornadic Storm and Associated Mesocyclones. *Mon. Weather Rev.*, **140**, 696–716,

https://doi.org/10.1175/MWR-D-11-00008.1.

——, E. R. Mansell, Y. Jung, L. J. Wicker, M. R. Kumjian, and M. Xue, 2014: Low-level ZDR signatures in supercell forward flanks: The role of size sorting and melting of hail. *J. Atmos. Sci.*, **71**, 276–299, https://doi.org/http://dx.doi.org/10.1175/JAS-D-13-0118.1.

——, M. Xue, J. A. Milbrandt, and A. Shapiro, 2015: Sensitivity of Real-Data Simulations of the 3 May 1999 Oklahoma City Tornadic Supercell and Associated Tornadoes to Multimoment Microphysics. Part I: Storm- and Tornado-Scale Numerical Forecasts. *Mon. Weather Rev.*, **143**, 2241–2265, https://doi.org/10.1175/MWR-D-14-00279.1.

Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Ocean. Technol.*, **26**, 911–927, https://doi.org/10.1175/2008JTECHA1156.1.

——, F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and Temperature Retrievals in the 17 May 1981 Arcadia, Oklahoma, Supercell: Ensemble Kalman Filter Experiments. *Mon. Weather Rev.*, **132**, 1982–2005, https://doi.org/10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2.

——, L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman Filter Assimilation of Radar Observations of the 8 May 2003 Oklahoma City Supercell: Influences of Reflectivity Observations on Storm-Scale Analyses. *Mon. Weather Rev.*, **139**, 272–294, https://doi.org/10.1175/2010MWR3438.1.

Du, J., J. McQueen, G. DiMego, Z. Toth, D. Jovic, B. Zhou, and H.-Y. Chuang, 2006: *New Dimension of NCEP Short-Range Ensemble Forecasting (SREF) System: Inclusion of WRF Members*. 5 pp. https://www.wmo.int/pages/prog/www/DPFS/Meetings/ET-EPS_Exeter2006/Doc6(5).pdf.

Dudhia, J., 1989: Numerical Study of Convection Observed During the Winter Monsoon

Experiment Using a Mesoscale Two-Dimensional Model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Evensen, G., 2003: The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, https://doi.org/10.1007/s10236-003-0036-9.

Fiori, E., A. Parodi, and F. Siccardi, 2011: Uncertainty in prediction of deep moist convective processes: Turbulence parameterizations, microphysics and grid-scale effects. *Atmos. Res.*, **100**, 447–456, https://doi.org/10.1016/j.atmosres.2010.10.003.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125**, 723–757, https://doi.org/10.1256/smsqj.55416.

Gasperoni, N. A., and X. Wang, 2015: Adaptive localization for the ensemble-based observation impact estimate using regression confidence factors. *Mon. Weather Rev.*, **143**, 1981–2000, https://doi.org/10.1175/MWR-D-14-00272.1.

Heinselman, P. L., and S. M. Torres, 2011: High-temporal-resolution capabilities of the National Weather Radar Testbed Phased-Array Radar. *J. Appl. Meteorol. Climatol.*, **50**, 579–593, https://doi.org/10.1175/2010JAMC2588.1.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal Boundary Layer Vertical Diffusion in a Medium-Range Forecast Model. *Mon. Weather Reivew*, **124**, 2322–2339, https://doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.*, **126**, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

——, and H. L. Mitchell, 2001: A Sequential Ensemble Kalman Filter for Atmospheric Data

Assimilation. *Mon. Weather Rev.*, **129**, 123–137, https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Phys. D Nonlinear Phenom.*, **230**, 112–126, https://doi.org/10.1016/j.physd.2006.11.008.

Janjić, Z. I., 1994: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Mon. Weather Rev.*, **122**, 927–945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Jones, T. A., K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-Scale Data Assimilation and Ensemble Forecasting with the NSSL Experimental Warn-on-Forecast System. Part II: Combined Radar and Satellite Data Experiments. *Weather Forecast.*, **31**, 297–327, https://doi.org/10.1175/WAF-D-15-0107.1.

Jung, Y., G. Zhang, and M. Xue, 2008: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble kalman filter. Part I: Observation operations for reflectivity and polarimetric variables. *Mon. Weather Rev.*, **136**, 2228–2245, https://doi.org/10.1175/2007MWR2083.1.

——, M. Xue, and M. Tong, 2012: Ensemble Kalman Filter Analyses of the 29–30 May 2004 Oklahoma Tornadic Thunderstorm Using One- and Two-Moment Bulk Microphysics Schemes, with Verification against Polarimetric Radar Data. *Mon. Weather Rev.*, **140**, 1457–1475, https://doi.org/10.1175/MWR-D-11-00032.1.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802,

https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2.

Kain, J. S., and J. M. Fritsch, 1993: Convective Parameterization for Mesoscale Models: The Kain-Fritcsh Scheme. *The Representation of Cumulus Convection in Numerical Models, Meter. Monogr.*, 165–170.

Kalman, R. E., 1960: A New Approach to Linear Filtering and Prediction Problems. *J. Fluids Eng.*, **82**, 35–45, https://doi.org/10.1115/1.3662552.

Kumjian, M. R., and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteorol. Climatol.*, **47**, 1940–1961, https://doi.org/10.1175/2007JAMC1874.1.

Labriola, J., N. Snook, Y. Jung, B. Putnam, and M. Xue, 2017: Ensemble Hail prediction for the storms of 10 May 2010 in South-Central Oklahoma using single- and double-moment microphysical schemes. *Mon. Weather Rev.*, **145**, 4911–4936, https://doi.org/10.1175/MWR-D-17-0039.1.

——, ——, ——, and M. Xue, 2019: Explicit ensemble prediction of hail in 19 May 2013 Oklahoma city thunderstorms and analysis of hail growth processes with several multimoment microphysics schemes. *Mon. Weather Rev.*, **147**, 1193–1213, https://doi.org/10.1175/MWR-D-18-0266.1.

Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The warning decision support system-integrated information. *Weather Forecast.*, **22**, 596–612, https://doi.org/10.1175/WAF1009.1.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model. *J. Clim. Appl. Meteorol.*, **22**, 1065–1092, https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Mahale, V. N., G. Zhang, and M. Xue, 2014: Fuzzy logic classification of S-band polarimetric radar echoes to identify three-body scattering and improve data quality. *J. Appl. Meteorol. Climatol.*, **53**, 2017–2033, https://doi.org/10.1175/JAMC-D-13-0358.1.

Marquis, J., Y. Richardson, P. Markowski, D. Dowell, and J. Wurman, 2012: Tornado Maintenance Investigated with High-Resolution Dual-Doppler and EnKF Analysis. *Mon. Weather Rev.*, **140**, 3–27, https://doi.org/10.1175/MWR-D-11-00025.1.

Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag*, **30**, 291–303.

Mason, S. J., and N. E. Graham, 1999: Conditional probabilities, relative operating characteristics, and relative operating levels. *Weather Forecast.*, **14**, 713–725, https://doi.org/10.1175/1520-0434(1999)014<0713:CPROCA>2.0.CO;2.

McPherson, R. A., and Coauthors, 2007: Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *J. Atmos. Ocean. Technol.*, **24**, 301–321, https://doi.org/10.1175/JTECH1976.1.

Milbrandt, J. A., and M. K. Yau, 2006a: A multimoment bulk microphysics parameterization. Part III: Control simulation of a hailstorm. *J. Atmos. Sci.*, **63**, 3114–3136, https://doi.org/10.1175/JAS3816.1.

——, and ——, 2006b: A multimoment bulk microphysics parameterization. Part IV: Sensitivity experiments. *J. Atmos. Sci.*, **63**, 3137–3159, https://doi.org/10.1175/JAS3817.1.

Miller, M. L., V. Lakshmanan, and T. M. Smith, 2013: An Automated Method for Depicting Mesocyclone Paths and Intensities. *Weather Forecast.*, **28**, 570–585, https://doi.org/10.1175/WAF-D-12-00065.1.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer

for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.*, **102**, 16663–16682, https://doi.org/10.1029/97JD00237.

Nakanishi, M., and H. Niino, 2009: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *J. Meteorol. Soc. Japan*, **87**, 895–912, https://doi.org/10.2151/jmsj.87.895.

Park, H. S., A. V. Ryzhkov, D. S. Zrnić, and K. E. Kim, 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Weather Forecast.*, **24**, 730–748, https://doi.org/10.1175/2008WAF2222205.1.

Parker, M. D., 2014: Composite VORTEX2 Supercell Environments from Near-Storm Soundings. *Mon. Weather Rev.*, **142**, 508–529, https://doi.org/10.1175/MWR-D-13-00167.1.

Poterjoy, J., 2016: A Localized Particle Filter for High-Dimensional Nonlinear Systems. *Mon. Weather Rev.*, **144**, 59–76, https://doi.org/10.1175/MWR-D-15-0163.1.

——, R. A. Sobash, and J. L. Anderson, 2017: Convective-scale data assimilation for the weather research and forecasting model using the local particle filter. *Mon. Weather Rev.*, **145**, 1897–1918, https://doi.org/10.1175/MWR-D-16-0298.1.

——, L. Wicker, and M. Buehner, 2019: Progress toward the application of a localized particle filter for numerical weather prediction. *Mon. Weather Rev.*, **147**, 1107–1126, https://doi.org/10.1175/MWR-D-17-0344.1.

Putnam, B. J., M. Xue, Y. Jung, N. Snook, and G. Zhang, 2014: The Analysis and Prediction of Microphysical States and Polarimetric Radar Variables in a Mesoscale Convective System Using Double-Moment Microphysics, Multinetwork Radar Data, and the Ensemble Kalman Filter. *Mon. Weather Rev.*, **142**, 141–162, https://doi.org/10.1175/MWR-D-13-00042.1.

——, ——, ——, G. Zhang, and F. Kong, 2017: Simulation of polarimetric radar variables from 2013 CAPS spring experiment storm-scale ensemble forecasts and evaluation of microphysics schemes. *Mon. Weather Rev.*, **145**, 49–73, https://doi.org/10.1175/MWR-D-15-0415.1.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, 236 pp.

Sakov, P., G. Evensen, and L. Bertino, 2010: Asynchronous data assimilation with the EnKF. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, **62**, 24–29, https://doi.org/10.1111/j.1600-0870.2009.00417.x.

Sasaki, Y., 1970: Some Basic Formalisms in Numerical Variational Analysis. *Mon. Weather Rev.*, **98**, 875–883.

Snook, N., M. Xue, and Y. Jung, 2011: Analysis of a Tornadic Mesoscale Convective Vortex Based on Ensemble Kalman Filter Assimilation of CASA X-Band and WSR-88D Radar Data. *Mon. Weather Rev.*, **139**, 3446–3468, https://doi.org/10.1175/MWR-D-10-05053.1.

——, ——, and ——, 2012: Ensemble Probabilistic Forecasts of a Tornadic Mesoscale Convective System from Ensemble Kalman Filter Analyses Using WSR-88D and CASA Radar Data. *Mon. Weather Rev.*, **140**, 2126–2146, https://doi.org/10.1175/MWR-D-11-00117.1.

——, M. Xue, and Y. Jung, 2015: Multiscale EnKF Assimilation of Radar and Conventional Observations and Ensemble Forecasting for a Tornadic Mesoscale Convective System. *Mon. Weather Reivew*, **143**, 1035–1057, https://doi.org/10.1175/MWR-D-13-00262.1.

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

https://doi.org/10.1175//2555.1.

Sobash, R. A., and D. J. Stensrud, 2015: Assimilating surface mesonet observations with the EnKF to improve ensemble forecasts of convection initiation on 29 May 2012. *Mon. Weather Rev.*, **143**, 3700–3725, https://doi.org/10.1175/MWR-D-14-00126.1.

Stensrud, D. J., and J. Gao, 2010: Importance of Horizontally Inhomogeneous Environmental Initial Conditions to Ensemble Storm-Scale Radar Data Assimilation and Very Short-Range Forecasts. *Mon. Weather Rev.*, **138**, 1250–1272, https://doi.org/10.1175/2009MWR3027.1.

——, and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Am. Meteorol. Soc.*, **90**, 1487–1499, https://doi.org/10.1175/2009BAMS2795.1.

Stratman, D. R., N. Yussouf, Y. Jung, T. A. Supinie, M. Xue, P. S. Skinner, and B. J. Putnam, 2020: Optimal Temporal Frequency of NSSL Phased Array Radar Observations for an Experimental Warn-on-Forecast System. *Weather Forecast.*, **35**, 193–214, https://doi.org/10.1175/waf-d-19-0165.1.

Sun, W.-Y., and C.-Z. Chang, 1986: Diffusion Model for a Convective Layer. Part I: Numerical Simulation of Convective Boundary Layer. *J. Clim. Appl. Meteorol.*, **25**, 1445–1453.

Supinie, T. A., N. Yussouf, Y. Jung, M. Xue, J. Cheng, and S. Wang, 2017: Comparison of the Analyses and Forecasts of a Tornadic Supercell Storm from Assimilating Phased-Array Radar and WSR-88D Observations. *Weather Forecast.*, **32**, 1379–1401, https://doi.org/10.1175/WAF-D-16-0159.1.

Tewari, M., and Coauthors, 2004: Implementation and Verification of the Unified Noah Land Surface Model in the WRF Model. *20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part I: Description and Sensitivity Analysis. *Mon. Weather Rev.*, **132**, 519–542, https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

——, P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Weather Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Tong, M., and M. Xue, 2005: Ensemble Kalman Filter Assimilation of Doppler Radar data with a Compressible Nonhydrostatic Model: OSS Experiments. *Mon. Weather Reivew*, **133**, 1789–1807, https://doi.org/10.1175/MWR2898.1.

——, and ——, 2008: Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Weather Rev.*, **136**, 1630–1648, https://doi.org/10.1175/2007MWR2070.1.

Wang, S., M. Xue, and J. Min, 2013a: A four-dimensional asynchronous ensemble square-root filter (4DEnSRF) algorithm and tests with simulated radar data. *Q. J. R. Meteorol. Soc.*, **139**, 805–819, https://doi.org/10.1002/qj.1987.

Wang, Y., Y. Jung, T. A. Supinie, and M. Xue, 2013b: A Hybrid MPI-OpenMP Parallel Algorithm and Performance Analysis for an Ensemble Square Root Filter Designed for Multiscale Observations. *J. Atmos. Ocean. Technol.*, **30**, https://doi.org/10.1175/JTECH-D-12-00165.1.

Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-Scale Data

Assimilation and Ensemble Forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar Data Experiments. *Weather Forecast.*, **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0107.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Mon. Weather Rev.*, **130**, 1913–1924, https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

——, and ——, 2012: Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. *Mon. Weather Rev.*, **140**, 3078–3089, https://doi.org/10.1175/MWR-D-11-00276.1.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. Second ed. Atmospheric Press, Elsevier, 627 pp.

Xu, X., M. Xue, and Y. Wang, 2015: Mesovortices within the 8 may 2009 bow echo over the Central United States: Analyses of the characteristics and evolution based on doppler radar observations and a high-resolution model simulation. *Mon. Weather Rev.*, **143**, 2266–2290, https://doi.org/10.1175/MWR-D-14-00234.1.

Xue, M., J. Zong, and K. K. Droegemeier, 1996: Parameterization of PBL Turbulence in a Multi-Scale Non-Hydrostatic Model. *Preprints, 11th Conf. on Numerical Weather Prediction*, Norfolk, VA, 363–365.

——, K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteorol. Atmos. Phys.*, **75**, 161–193, https://doi.org/10.1007/s007030170027.

——, and Coauthors, 2001: The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteorol. Atmos. Phys.*, **76**, 143–165, https://doi.org/10.1007/s007030170027.

——, D. Wang, J. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteorol. Atmos. Phys.*, **82**, 139–170, https://doi.org/10.1007/s00703-001-0595-6.

——, M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Ocean. Technol.*, **23**, 46–66, https://doi.org/10.1175/JTECH1835.1.

——, M. Hu, and A. D. Schenkman, 2014: Numerical prediction of the 8 May 2003 Oklahoma City tornadic supercell and embedded tornado using ARPS with the assimilation of WSR-88D data. *Weather Forecast.*, **29**, 39–62, https://doi.org/10.1175/WAF-D-13-00029.1.

Yussouf, N., and D. J. Stensrud, 2010: Impact of phased-array radar observations over a short assimilation period: Observing system simulation experiments using an ensemble Kalman filter. *Mon. Weather Rev.*, **138**, 517–538, https://doi.org/10.1175/2009MWR2925.1.

——, and D. D. J. Stensrud, 2012: Comparison of single-parameter and multiparameter ensembles for assimilation of radar observations using the ensemble kalman filter. *Mon. Weather Rev.*, **140**, 562–586, https://doi.org/10.1175/MWR-D-10-05074.1.

——, and K. H. Knopfmeier, 2019: Application of the Warn-on-Forecast system for flash-flood-producing heavy convective rainfall events. *Q. J. R. Meteorol. Soc.*, **145**, 2385–2403,

https://doi.org/10.1002/qj.3568.

——, E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The Ensemble Kalman Filter Analyses and Forecasts of the 8 May 2003 Oklahoma City Tornadic Supercell Storm Using Single- and Double-Moment Microphysics Schemes. *Mon. Weather Rev.*, **141**, 3388–3412, https://doi.org/10.1175/MWR-D-12-00237.1.

——, D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Weather Rev.*, **143**, 3044–3066, https://doi.org/10.1175/MWR-D-14-00268.1.

——, J. S. Kain, and A. J. Clark, 2016: Short-Term Probabilistic Forecasts of the 31 May 2013 Oklahoma Tornado and Flash Flood Event Using a Continuous-Update-Cycle Storm-Scale Ensemble System. *Weather Forecast.*, **31**, 957–983, https://doi.org/10.1175/WAF-D-15-0160.1.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Weather Rev.*, **132**, 1238–1253, https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

Zhang, J., and Coauthors, 2011: National mosaic and multi-sensor QPE (NMQ) system description, results, and future plans. *Bull. Am. Meteorol. Soc.*, **92**, 1321–1338, https://doi.org/10.1175/2011BAMS-D-11-00047.1.