JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2022MS003307

Key Points:

- A multivariate additive-inflation approach is introduced to address the deficient ensemble for ensemble-based radar data assimilation
- The multivariate approach improves the analysis and forecast of a tornadic supercell relative to a previous univariate approach
- The multivariate approach by additionally perturbing hydrometeors and vertical velocity is beneficial in analysis and forecast

Correspondence to:
X. Wang,
xuguang.wang@ou.edu

Citation:
Wang, Y., & Wang, X. (2023). A multivariate additive inflation approach to improve storm-scale ensemble-based data assimilation and forecasts: Methodology and experiment with a tornadic supercell. *Journal of Advances in Modeling Earth Systems*, *15*, e2022MS003307. https://doi.org/10.1029/2022MS003307

Received 14 JUL 2022
Accepted 2 JAN 2023

A Multivariate Additive Inflation Approach to Improve Storm-Scale Ensemble-Based Data Assimilation and Forecasts: Methodology and Experiment With a Tornadic Supercell

Yongming Wang[1] and Xuguang Wang[1]

[1]School of Meteorology, University of Oklahoma, Norman, OK, USA

Abstract Ensemble-based convective-scale radar data assimilation commonly suffers from an underdispersive background ensemble. This study introduces a multivariate additive-inflation method to address such deficiency. The multivariate additive inflation (AI) approach generates coherent random perturbations drawn from a newly constructed convective-scale static background error covariance matrix for all state variables including hydrometeors and vertical velocity. This method is compared with a previously proposed univariate AI approach, which perturbs each variable individually without cross-variable coherency. Comparisons are performed on the analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell. Within assimilation cycles, the multivariate approach is more efficient in increasing reflectivity spread and thus has a reduced spinup time than the univariate approach; the additional inclusion of hydrometeors and vertical velocity results in more background spread for both reflectivity and radial velocity. Significant differences among AI experiments also exist in the subsequent forecasts and are more pronounced for the forecasts initialized from the earlier assimilation cycles. The multivariate approach yields better forecasts of low-level rotation, reflectivity distributions, and storm maintenance for most lead times. The additional inclusion of hydrometeor and vertical velocity in the multivariate method is beneficial in forecasts. Conversely, the additional inclusion of hydrometeor and vertical velocity in the univariate method poses negative impacts for the majority of forecast lead times.

Plain Language Summary Data assimilation (DA) requires an accurate estimation of the background uncertainty. In ensemble-based DA, such uncertainty is represented by the statistics of a background ensemble. However, it is difficult to construct ideal background ensembles that truly represent forecast errors. The deficient background ensemble becomes more problematic for convective-scale radar DA when all ensemble members miss the observed storms. This study proposes a multivariate additive-inflation approach to address such deficiency. The mitigation of ensemble deficiency is achieved by drawing spatially and physically coherent random perturbations from a recently constructed convective-scale static background error covariance matrix and adding them to each ensemble member. This study assesses the new approach by comparing it with a previously proposed univariate approach, which perturbs each variable individually without cross-variable coherency. Results and diagnostics from a tornadic supercell study show that the multivariate approach improves the analysis and forecast compared to the univariate approach. The multivariate approach by additionally perturbing hydrometeors and vertical velocity further improves the forecast. In contrast, the univariate approach including the hydrometeors and vertical velocity perturbations degrades the forecast.

1. Introduction

Recently, multiple ensemble-based data assimilation (DA) methods, such as the ensemble Kalman filter (e.g., Aksoy et al., 2009; Dowell et al., 2004, 2011; Johnson et al., 2015; Jung et al., 2008; Lei et al., 2009; Tong & Xue, 2005; Yokota et al., 2018; Yussouf et al., 2013; Zeng et al., 2020) and the hybrid ensemble-variational (EnVar, e.g., Duda et al., 2019; Gasperoni et al., 2022; Y. Wang & Wang, 2017, 2020, 2021) methods, have been widely adopted to initialize convection-allowing numerical weather prediction models through assimilating radar observations. Many results have demonstrated that the flow-dependent background error covariances (BECs) derived from ensembles are useful in convective scale data assimilation (e.g., Dowell et al., 2011; Johnson et al., 2015; Snyder & Zhang, 2003). However, the effectiveness of the flow-dependent BECs can be heavily degraded by the ensemble misrepresentation of model error, including imperfect parameterizations of subgrid-scale physical processes. Such an inadequate representation of uncertainties associated with the model error implies ensemble

sampling deficiency, which leads to degraded performance of storm-scale ensemble-based DA and the quality of the subsequent forecasts.

Many earlier DA studies have developed a series of techniques to account for the ensemble deficiency, including multiplicative inflation (Anderson & Anderson, 1999; Houtekamer et al., 2005; Whitaker & Hamill, 2012), additive inflation (AI) (Dowell & Wicker, 2009; Houtekamer et al., 2005; X. Wang et al., 2009; Whitaker et al., 2008; Zhang et al., 2004), multi-physics–multi-model method (Fujita et al., 2007; Meng & Zhang, 2008), and stochastic physics schemes (Houtekamer et al., 2009; Leutbecher et al., 2007; Reynolds et al., 2008). These methods have been examined and intercompared in the maintenance of ensemble spread during DA and in their impacts on the subsequent forecasts in these early studies.

For the storm-scale radar DA, both relaxation to prior perturbation (Zhang et al., 2004) and relaxation to prior spread (RTPS, Johnson et al., 2015; Y. Wang & Wang, 2017, 2020; Whitaker & Hamill, 2012) have been used to relax the excessive ensemble spread reduction during DA. However, both relaxation methods may become ineffective to sustain sufficient ensemble spread when the background ensemble is severely underdispersive, that is, most ensemble members fail to predict the observed reflectivity (Yokota et al., 2018). Alternatively, Dowell and Wicker (2009) (hereafter DW09) proposed an AI method to add univariate random perturbations of conventional variables, including winds (u and v), potential temperature (t), and water vapor mixing ratio (q), to each ensemble member for initializing storm-scale ensemble and for accounting for ensemble underdispersion. Spatial correlation is included during the generation of the random perturbations. Their work showed that AI provides sufficient ensemble spread and causes fewer spurious convective cells than the multiplicative inflation. This technique is easy to implement and commonly adopted for the ensemble-based radar DA (e.g., Dowell et al., 2011; Jung et al., 2012; Sobash & Wicker, 2015; Y. Wang & Wang, 2017, 2020; Yussouf et al., 2013).

Nevertheless, Yokota et al. (2018) mentioned that adding univariate random perturbations for conventional variables (DW09) may not be the most effective in addressing the severe underdispersion for reflectivity, that is, zero spread of reflectivity. Those univariately perturbed conventional atmospheric states in DW09 do not directly correlate with the hydrometeors. Additionally, the univariate approach omits the coherent relationship among variables. Therefore, recent studies (Labriola & Wicker, 2022; Vobig et al., 2021; Yokota et al., 2018) have attempted to generate spatially and physically coherent random perturbations. Yokota et al. (2018) and Vobig et al. (2021) separately proposed an additive covariance inflation (Add_CI) method to inflate the reflectivity spread during DA. Their method constructs the ensemble reflectivity perturbations based on the horizontal cross-correlations of reflectivity with other model variables. However, defects remain in the Add_CI method to parameterize model error. First, the Add_CI method omits the vertical correlations to construct perturbations (Vobig et al., 2021; Yokota et al., 2018). However, broad vertical cross-correlations were found within precipitation areas (Michel et al., 2011; Y. Wang & Wang, 2021). Second, this Add_CI method may become ineffective to provide sufficient spread to assimilate reflectivity when the other model states have a small amount of spread (Vobig et al., 2021). Third, perturbations generated by Add_CI does not provide sampling beyond the space spanned by the existing background ensemble. Labriola and Wicker (2022) drew spatially and physically coherent convective-scale perturbations from the idealized CM1 turbulence simulations at a grid spacing of 250 m initialized with random potential temperature perturbations. While their method obtained promising results in the increase of ensemble spread, application of this method to real-model applications remains to be explored.

In this study, we propose a new multivariate AI approach to increase and maintain the ensemble spread in ensemble-based radar DA. This approach adds multivariate random perturbations to both conventional atmospheric states and variables critical to storm scales (e.g., hydrometeors, vertical velocity) in a dynamically and thermodynamically coherent fashion. Such coherency is achieved through generating random perturbations sampled from a separately constructed convective-scale static BEC matrix (Y. Wang & Wang, 2021). As detailed in Y. Wang and Wang (2021), this static BEC matrix was designed to additionally use vertical velocity, all hydrometeor types and reflectivity as control variables and to include the three-dimensional (3D) cross-correlations between any pair of all control variables. The proposed new AI method can generate perturbations that allow to directly increase and maintain the ensemble spread for variables important to convective scales, allow to modify the background ensemble in a spatially and physically coherent manner, and allow sampling of system deficiency beyond the space spanned by the background ensemble. A similar idea of realizing AI through sampling from static BECs was applied in early studies for large-scale or coarse resolution applications (Houtekamer et al., 2005, 2009; X. Wang et al., 2009; Zeng et al., 2018). The development of static BECs for convection-allowing scales (e.g., Y. Wang & Wang, 2021) allows examining its use for convective-scale ensemble DA in real-data applications.

In this study, the new AI method is first introduced and described. This method is implemented and examined in the two-way coupled Gridpoint Statistical Interpolation (GSI)-based EnVar system (X. Wang et al., 2013) that has the direct radar reflectivity assimilation capability (Y. Wang & Wang, 2017) to address two scientific objectives. The first objective is to examine if the multivariate AI method is more efficient in the increase and maintenance of ensemble spread, compared to the univariate AI method (i.e., DW09). As an apple-to-apple comparison with DW09, only the original conventional variables included in DW09 will first be perturbed. The second objective of this study is to evaluate the impact of including perturbations of hydrometeors and vertical velocity (PHW) in AI. This objective is achieved by comparing experiments with and without the additional inclusion of PHW using both the multivariate and univariate AI methods. Under the second objective, coherently adding hydrometer and vertical velocity will be compared with univariately adding hydrometeor and vertical velocity as well. Such a comparison can also address the first objective. Given that this study is the first to examine this new method for convective-scale ensemble-based DA, we illustrate the new method and perform detailed examinations and diagnostics using a case study of the 8 May 2003 Oklahoma City (OKC) tornadic supercell storm. The diagnostics aims to offer an understanding of the newly proposed approach before systematic testing, which is left for future studies.

The remainder of this paper is organized as follows. Section 2 describes the univariate and the new multivariate AI methods. Section 3 provides an overview of the 8 May 2003 OKC tornadic supercell storm and the experimental design. Section 4 compares the performance of using two AI methods and evaluates the impact of further including PHW in AI. A summary is given in Section 5.

2. Univariate and Multivariate Additive-Inflation Methods

This section introduces the univariate AI method of DW09 and the newly proposed multivariate AI method. Both AI methods parameterize the ensemble system error by adding random perturbations η_i to the ith analysis ensemble member $\mathbf{x}_i^a$,

$$\mathbf{x}_i^a \leftarrow \mathbf{x}_i^a + \eta_i, \tag{1}$$

where $i = 1, 2, \ldots, K$, and K denotes the ensemble size. Following DW09 and many other studies (e.g., Y. Wang & Wang, 2017, 2020; Yussouf et al., 2013), random perturbations are added within high-reflectivity regions only. In this study, perturbations in both methods are only added where the observed reflectivity exceeds 25 dBZ. The differences between the two methods lie in the generation of η_i, and the specific procedure of generating η_i for two methods is described as follows.

2.1. Univariate Additive-Inflation Method (UNI_AI)

In this method, we use the same generation procedure as DW09 to obtain η_i. Briefly, the first step is to generate a random number $\mathbf{r}(x, y, z)$ for each model grid point (x, y, z) within high-reflectivity regions independently or univariately for each model variable and each ensemble member. Such a random number is generated from a normal distribution $N(0, \sigma)$ with a mean of 0 and a standard deviation of σ. All random numbers within high-reflectivity regions consist of a perturbation array $\mathbf{r}$. The second step is to obtain η_i at each grid point (l, m, n) by smoothing the perturbation array $\mathbf{r}$ following Caya et al. (2005):

$$\eta_i(l, m, n) = \sum_{x,y,z} \mathbf{r}(x, y, z) exp\left(-\frac{|x-l|}{l_h} - \frac{|y-m|}{l_h} - \frac{|z-n|}{l_v}\right), \tag{2}$$

where l, m, and n represent the indices of all model grid points; l_h and l_v are the horizontal and vertical smoothing length scales, respectively.

The commonly perturbed variables in UNI_AI include u, v, t, and dewpoint temperature td (e.g., DW09, Dowell et al., 2011; Sobash & Wicker, 2015; Y. Wang & Wang, 2017, 2020; Yussouf et al., 2013). The perturbed td field is further transformed to q. Hydrometeors (cloud water (ql), rainwater (qr), snow (qs), cloud ice (qi), and graupel (qg)) and vertical velocity (w) were not perturbed in these past studies using UNI_AI. As one of the objectives stated in Section 1, this study will explore the impact of additionally including these hydrometeors and vertical velocity perturbations for the storm-scale DA in both the univariate and multivariate AI methods. Therefore, we further develop UNI_AI to include the above-mentioned variables.

2.2. Multivariate Additive-Inflation Method (BEC_AI)

BEC_AI uses a static BEC matrix constructed for convective scales (Y. Wang & Wang, 2021) to sample random perturbations. The detailed construction of this convective scale static BEC matrix and its benefits in improving storm-scale analyses and forecasts have been presented in Y. Wang and Wang (2021). Briefly, such a static BEC matrix was constructed using historical, 3-km ensemble forecast members during May 2019. In the construction of this static BEC matrix, (a) reflectivity, w, and all hydrometeor types are extended as control variables; (b) 3D cross-correlations between any pair of all control variables are established.

The variable-coherent perturbations are generated as follows. For each ensemble member, a random number $\mathbf{r}\mathrm{dBZ}(x, y, z)$ is drawn from a normal distribution $N(0, \sigma_{\mathrm{dBZ}})$ with a mean of 0 and a standard deviation of σ_{dBZ} for each model grid point within observed high-reflectivity (above 25 dBZ) regions. All random numbers within high-reflectivity regions consist of a perturbation array $\mathbf{r}_{\mathrm{dBZ}}$. This array $\mathbf{r}_{\mathrm{dBZ}}$ is smoothed using Equation 2 and then adopted as the perturbation field for the reflectivity control variable. The properly constructed, convective-scale static BEC matrix (Y. Wang & Wang, 2021) $\mathbf{B}$ is then employed to obtain perturbation η_i for u, v, t, q, w, and mixing ratios of hydrometeor (ql, qr, qs, qi, and qg) following

$$\eta_i = \mathbf{B}\mathbf{r}_{\mathrm{dBZ}} = \mathbf{U}_p\mathbf{S}\mathbf{U}_v\mathbf{U}_h\mathbf{U}_h^{\mathrm{T}}\mathbf{U}_v^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{U}_p^{\mathrm{T}}\mathbf{r}_{\mathrm{dBZ}}. \tag{3}$$

The physical transform $\mathbf{U}_p$ denotes the cross-variable correlations between control variables (CVs), as defined in $\mathbf{B}$. This matrix $\mathbf{U}_p$ maps $\mathbf{r}_{\mathrm{dBZ}}$ into perturbations with respect to other CVs, u, v, t, q, w, ql, qr, qs, qi, and qg. The horizontal and vertical transform $\mathbf{U}_h$ and $\mathbf{U}_v$ respectively define the horizontal and vertical correlations for each CV. These two matrices are realized by spatially smoothing the perturbations using the recursive filters (Hayden & Purser, 1995; Purser et al., 2003a, 2003b). The standard deviations $\mathbf{S}$ determine the magnitude of perturbations. Thus, using this approach, the perturbation magnitude, spatial patterns, and relationships with other variables are automatically determined after $\mathbf{r}_{\mathrm{dBZ}}$ is given.

Here the process of obtaining η_i for t is shown as an example with some more details. First, applying $\mathbf{U}_p^{\mathrm{T}}$ on $\mathbf{r}_{\mathrm{dBZ}}$ transforms reflectivity perturbations into perturbations $\mathbf{V}_x$ for each CV using a regression relation,

$$\mathbf{V}_x = p_{x,\mathrm{dbz}}\mathbf{r}_{\mathrm{dBZ}}, \tag{4}$$

where x represents each of u, v, t, q, w, ql, qr, qs, qi, and qg; $p_{x,\mathrm{dbz}}$ consists of $\mathbf{U}_p$ and is the regression to model the cross-variable correlation between the x CV and reflectivity (dbz). Second, $\mathbf{U}_h$, $\mathbf{U}_v$, and $\mathbf{S}$ as well as their transposes are applied on $\mathbf{V}_x$ following Equation 3 to obtain $\hat{\mathbf{V}}_x$ using the recursive filter transforms of GSI. Then, η_i specific for t is calculated by applying $\mathbf{U}_p$ on $\hat{\mathbf{V}}_x$ following

$$\eta_i = \hat{\mathbf{V}}_t + p_{u,t}\hat{\mathbf{V}}_u + p_{v,t}\hat{\mathbf{V}}_v, \tag{5}$$

where $p_{u,t}$ and $p_{v,t}$ represents the regression relations of t with u and v, respectively. To obtain η_i for other variables, Equation 5 is replaced following the $\mathbf{U}_p$ transform of Equation 7 in Y. Wang and Wang (2021).

Per Y. Wang and Wang (2021), the generated perturbations allow to directly increase and maintain ensemble spread for all the variables including those specific to convective scales (e.g., hydrometeors and vertical velocity) and to modify background ensemble in a spatially and physically coherent manner.

Figure 1 shows an example of perturbations respectively produced by using UNI_AI and BEC_AI for the radar reflectivity valid at 21:00 UTC 8 May 2003 for the OKC tornadic supercell. The standard deviations of random perturbations are 0.25 m s^{-1}, 0.25°C, 0.25°C, and 0.25 g kg^{-1} for horizontal winds, t, td, and qr in UNI_AI, and 5.0 dBZ for reflectivity in BEC_AI. The horizontal and vertical length scales of 4 km Equation 2 are used for the perturbation smoothing function in both methods. These parameters here in each AI method have been carefully chosen by trial and error overall for best forecast performances. The perturbations from both methods show some similarities. All perturbations are localized near the observed precipitation areas (Figure 1g), and both methods have similar magnitudes in the t and qr perturbations (Figures 1a–1d). It is noted that BEC_AI has a ~2 m s^{-1} greater maximum of wind perturbations than UNI_AI. Cross-variable correlations of perturbations show significant differences between BEC_AI and UNI_AI. A typical internal storm structure can be revealed by the BEC_AI perturbations due to the cross-correlations of BEC_AI but cannot by UNI_AI. Figures 1e and 1f show the vertical cross-correlations of perturbations between q and qr respectively for UNI_AI and BEC_AI as an example. The

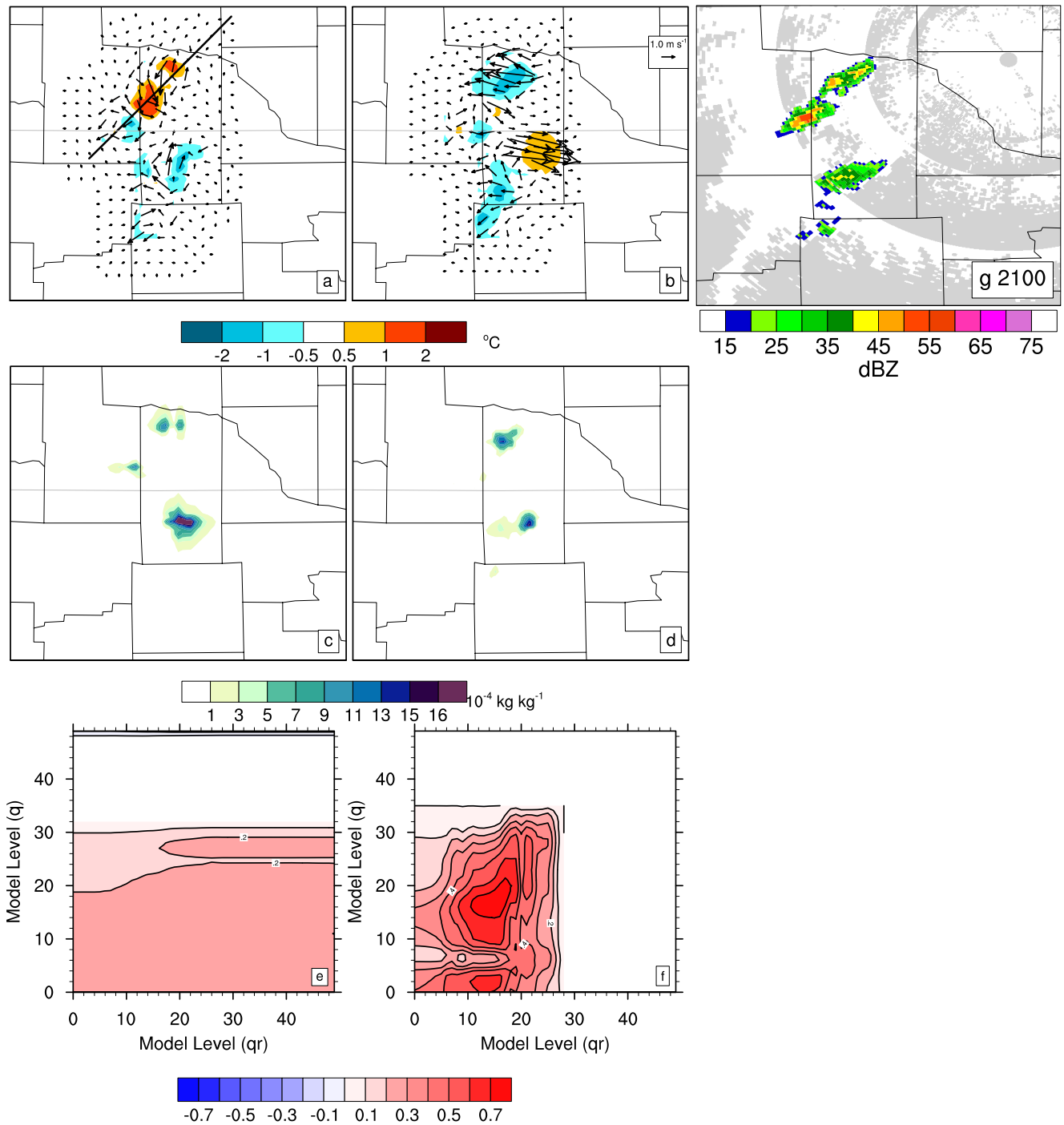

Figure 1. Example of random perturbations in (a, b) temperature (colors; °C) and horizontal winds (vectors; m s^{-1}), and (c, d) rainwater mixing ratio (colors; kg kg^{-1}) at 2 km AGL generated by the use of (a, c) UNI_AI and (b, d) BEC_AI. (g) The corresponding reflectivity observations (colors; dBZ) valid at 21:00 UTC 8 May 2003. Vertical cross-variable correlations between water vapor mixing ratio (q) and rainwater mixing ratio (qr) calculated using the perturbations from (e) UNI_AI and (f) BEC_AI. The vertical cross-variable correlations are spatially averaged over the entire domain. The straight black line in panel (a) indicates the location of the vertical cross section in Figures 4 and 5.

perturbations in BEC_AI are highly correlated with the maxima coefficient greater than 0.7, indicating that the enhanced qr corresponds with the increased q. In contrast, UNI_AI generates perturbations with a correlation value lower than 0.3. We also find that the qr perturbations are added only below the 28th model levels in BEC_AI; however, UNI_AI adds qr perturbations at all vertical levels.

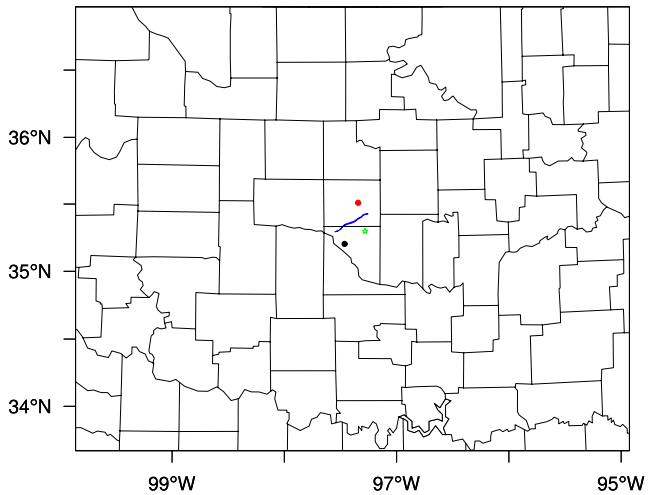

Figure 2. The domain for the storm-scale data assimilation and forecasts. The overlaid blue lines are the observed tornado damage tracks during 22:04–22:38 UTC. Location of the KTLX radar site is indicated by the green star. Locations for the Mesonet sites at Norman and Spencer stations are labeled with black and red dots, respectively.

3. Experimental Design

3.1. Overview of the 8 May 2003 Oklahoma City (OKC) Tornadic Supercell Case

The 8 May 2003 OKC tornadic supercell storm has been introduced and investigated in many studies (e.g., Romine et al., 2008; Y. Wang & Wang, 2017, 2020, 2021; Yussouf et al., 2013). By midafternoon, several storm cells occurred along the dryline across west-central Oklahoma between 20:00 and 21:00 UTC. Among these cells, a storm cell developed into an isolated supercell with a significant hook appendage structure by 22:00 UTC. The supercell propagated northeastward and generated two consecutive east-northeast tornado damage tracks from 22:04 to 22:08 and from 22:10 to 22:38 UTC with track lengths of 3 and 30 km, respectively. This supercell storm started weakening at 23:00 UTC 8 May and finally decayed at ~00:20 UTC 9 May.

3.2. Design of Data Assimilation and Forecast Experiments

This study adopts a single domain (Figure 2) at a 2-km grid spacing, configured as Y. Wang and Wang (2017, 2021). This domain has 226 × 181 horizontal grid points and 50 vertical levels. The Advanced Research Weather Research and Forecasting Model (WRF-ARW) version 3.9.1.1 is used. The employed model physics schemes include the Thompson microphysics scheme (G. Thompson et al., 2008), the Mellor–Yamada–Janjić planetary boundary layer scheme (Janjić, 1990, 1994, 2001; Mellor & Yamada, 1982), the Noah land surface model (Chen & Dudhia, 2001), and the longwave and shortwave radiation schemes of the Rapid Radiative Transfer Model for general circulation models schemes (Iacono et al., 2008).

Most DA configurations in this study follow Y. Wang and Wang (2017, 2021). This study uses a 45-member ensemble at a 2-km grid spacing in the ensemble-based DA. A 45-member mesoscale ensemble at an 18-km grid in Yussouf et al. (2013) is downscaled to serve as initial conditions (ICs) and corresponding lateral boundary conditions (LBCs). The ICs and LBCs for the EnVar control member are provided by the mean of the mesoscale ensemble. A set of 45 samples was drawn from a default BEC in the WRF data assimilation (WRFDA) package and was added to the Global Forecast System analysis at 12:00 UTC 5 May 2003 to form the mesoscale initial ensemble (X. Wang et al., 2008a, 2008b). Please refer to Yussouf et al. (2013) for the details of the mesoscale ensemble. Only radar reflectivity and radial velocity from the Weather Surveillance Radar-1988 Doppler (WSR-88D) in Twin Lakes, Oklahoma (KTLX) are assimilated in this study. These radar observations are obtained by processing the NEXRAD Level-II data retrieved from the National Climatic Data Center. The observation quality control is performed by using the Warning Decision Support System-Integrated Information (Lakshmanan et al., 2007) software. The observed reflectivity weaker than 5 dBZ is treated as the "no-precipitation" observations and is reset to 0 dBZ. Assimilation of these "no-precipitation" observations intends to suppress spurious convections in the model (Aksoy et al., 2009; Duda et al., 2019; Y. Wang & Wang, 2017, 2021; Yussouf et al., 2013). Consistent with past radar DA studies (Dowell et al., 2004; Duda et al., 2019; Johnson et al., 2015; Y. Wang & Wang, 2017, 2020, 2021; Yussouf et al., 2013), the observation errors are defined as 5 dBZ and 2 m s^{-1} for radar reflectivity and radial velocity, respectively.

A two-way coupled GSI-based EnVar method (X. Wang et al., 2013) is applied to compare the impact of utilizing UNI_AI versus BEC_AI on the analyses and forecasts. Figure 3 presents the procedure of radar DA and forecasts with the application of AI. Starting from 21:00 UTC 8 May 2003, the AI technique is applied every 15 min just before the 15-min forecast between assimilation cycles from 21:00 to 21:45 UTC. The purpose of adding random perturbations to ensemble analysis is to establish the flow-dependent structure of these perturbations during the 15-min model integration between two DA cycles. Radar reflectivity and radial velocity are assimilated every 15 min between 21:15 and 22:00 UTC. Free forecasts are launched from every DA cycle time and end at 23:00 UTC for each of the 45-member ensemble analyses and the EnVar control analysis. The control analysis is updated in EnVar with its BEC purely estimated by ingesting background ensemble perturbations. The ensemble analyses are obtained by updating ensemble perturbations using EnSRF and are further recentered around the

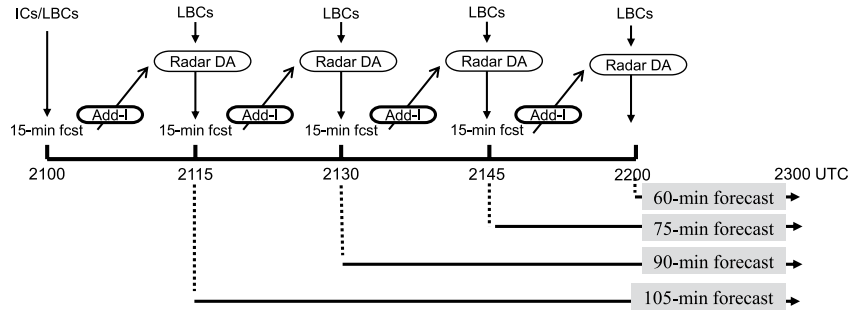

Figure 3. Schematics of data assimilation experiments with the use of additive inflation (AI). The section "Add-I" denotes the application of AI.

control analysis. The same localization scales are used in EnVar and EnSRF with cutoff distances of 12 km in the horizontal and 1.1 scale height (measured in nature log of pressure) in the vertical. Previous studies have shown that the combination of various methods to account for the deficiency of background ensemble spread usually outperforms either method used alone (Bowler et al., 2017; Hamrud et al., 2015; Whitaker & Hamill, 2012). Therefore, this study implements the RTPS inflation method (Whitaker & Hamill, 2012) and the constant inflation, as in Yussouf et al. (2013) and Y. Wang and Wang (2017, 2020), in addition to the AI method. To ensure a homogenous comparison between UNI_AI and BEC_AI, the same RTPS and constant inflation parameters are used. Specifically, in each assimilation cycle, the RTPS relaxes the posterior ensemble spread back to 90% of the prior ensemble spread, and the constant inflation with a coefficient of 1.04 is applied to each ensemble perturbation of the ensemble analyses.

To demonstrate the effectiveness of AI in improving the background ensemble spread, Exp-noAI is first conducted to compare with the remaining experiments in Table 1. The remaining experiments apply AI to each cycle during 21:00–21:45 UTC (hereafter "the AI experiments"). The impact of UNI_AI versus BEC_AI on the convective-scale analyses and forecasts is compared using Exp-UNI and Exp-BEC. Exp-UNI uses UNI_AI with the standard deviations (σ) of 0.25, 0.25 m s^{-1}, 0.25°C, and 0.25°C for u, v, t, and td, respectively, and $l_h = l_v = 4$ km in the perturbation smoothing function of Equation 2. These parameters are similar to those in DW09 and many other studies (e.g., Dowell et al., 2011; Y. Wang & Wang, 2017, 2020; Yussouf et al., 2013). BEC_AI is employed in Exp-BEC with a standard deviation (σ_{dBZ}) of 5.0 dBZ for the reflectivity perturbation. For a fair comparison between Exp-UNI and Exp-BEC, Exp-BEC only adds perturbations of u, v, t, and q to the ensemble. To investigate the impact of adding random perturbations of hydrometeors and w on the increase of reflectivity spread, Exp-UNI_All and Exp-BEC_All are designed. Exp-UNI_All is the same as Exp-UNI except that hydrometeors and w are additionally perturbed with the standard deviations (σ) of 0.25 g kg^{-1}, and 0.25 m s^{-1}, respectively. Exp-BEC_All is the same as Exp-BEC but adds perturbations to all fields, including u, v, t, q, w, qc, qr, qi, qs, and qg. The magnitude of perturbations for each variable in Exp-BEC and Exp-BEC_All is automatically determined, following the description in Section 2.2. Therefore, comparisons between Exp-BEC and Exp-BEC_All and between Exp-UNI and Exp-UNI_All aim to identify the impact of further adding PHW in both AI methods. Note as discussed in Figure 1, the perturbation magnitude parameters for the AI experiments were selected based on sensitivity tests that produce the best forecasts. The perturbation magnitudes in

Table 1
List of Experiments and Their Configurations of Additive Inflation

Experiments	Configurations		
	Additive-inflation method	Perturbed fields	Standard deviation (σ) of random noise
Exp-noAI	\	\	\
Exp-UNI	Univariate	u, v, t, td	0.25 m s^{-1}/°C
Exp-BEC	Multivariate	u, v, t, q	5.0 dBZ
Exp-UNI_All	Univariate	u, v, t, td, w, qc, qr, qi, qs, and qg	0.25 m s^{-1}/°C/g kg^{-1}
Exp-BEC_All	Multivariate	u, v, t, q, w, qc, qr, qi, qs, and qg	5.0 dBZ

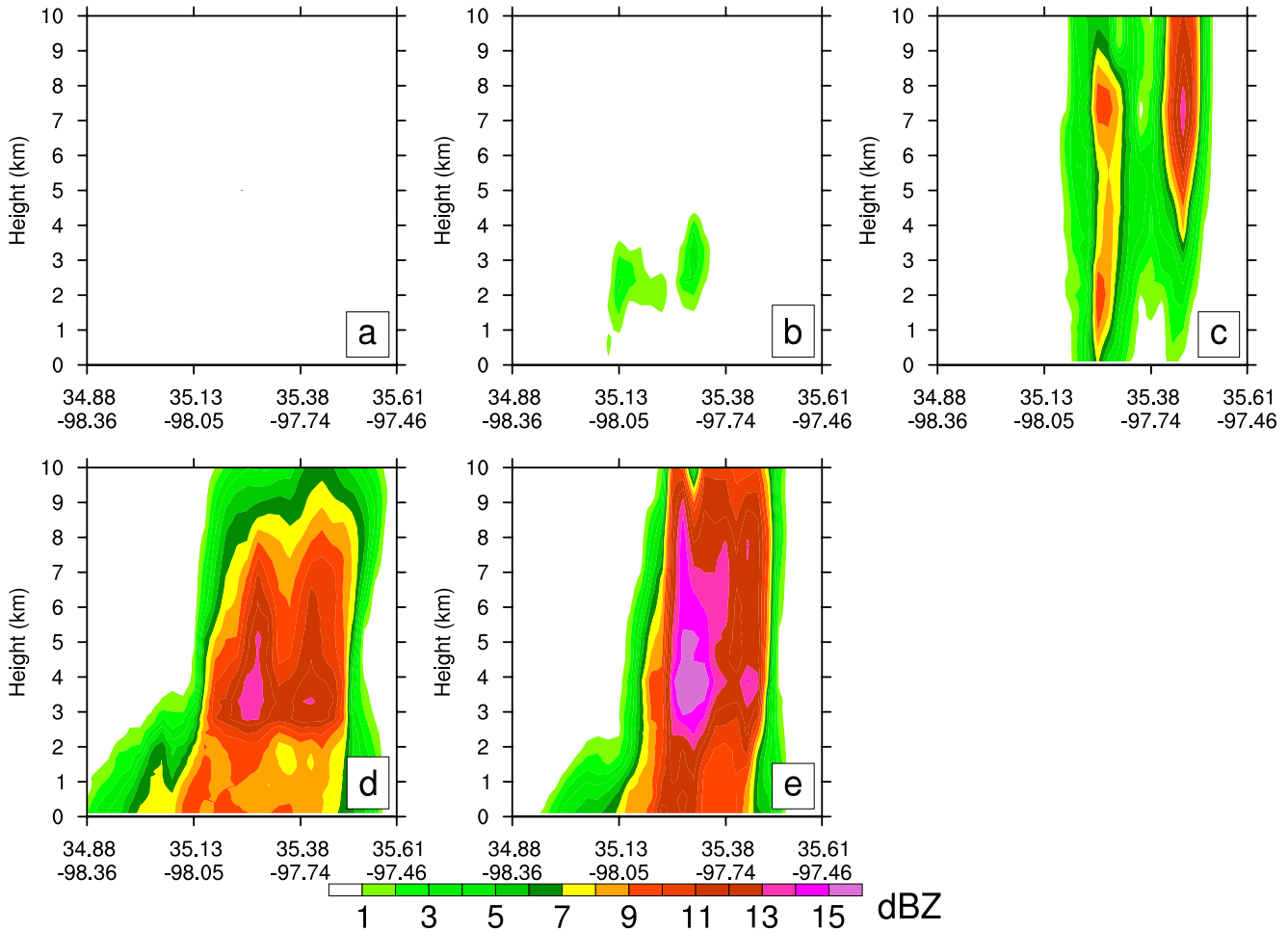

Figure 4. Vertical cross sections along the black line in Figure 1a of reflectivity prior ensemble standard deviations (dBZ) valid at 21:15 UTC for (a) Exp-noAI, (b) Exp-UNI, (c) Exp-BEC, (d) Exp-UNI_All, and (e) Exp-BEC_All.

Exp-BEC_All are determined by the computed **S** Equation 3 in the static BEC matrix in this study. Sensitivity tests by increasing and decreasing the values of **S** with a factor of 50% in Exp-BEC_All showed less sensitivity than by increasing and decreasing σ with the same factor in Exp-UNI_All (not shown). In addition, like the comparisons between Exp-UNI and Exp-BEC, comparisons between Exp-UNI_All and Exp-BEC_All are also performed to examine the impact of multivariate correlations in producing perturbations of all variables. These experiments are compared using both deterministic and ensemble analyses and forecasts to explore the two scientific objectives in Section 1.

4. Results

4.1. Impact on Data Assimilation Cycling

4.1.1. Comparisons in the Increase of Ensemble Spread and the Impact on the First DA Cycle

The direct impact of different AI methods on the increase of ensemble spread can be revealed by comparing the background ensemble of the first DA cycle valid at 21:15 UTC 8 May 2003. The background ensemble spread of reflectivity below 10 km from the first DA cycle for all experiments (Table 1) is presented in Figure 4. Exp-noAI has zero spread as a result of the entire reflectivity distributions missed by all ensemble members (Figure 4a). Compared to Exp-noAI, the AI experiments increase the spread of reflectivity (Figures 4b–4d). These comparisons between Exp-noAI and the AI experiments demonstrate the usefulness of AI in increasing spread. For the comparisons between UNI_AI and BEC_AI, Exp-UNI produces spread up to 4 dBZ below 4 km AGL (Figure 4b), while Exp-BEC has spread 5–13 dBZ larger than Exp-UNI across most vertical levels (Figure 4c). Similarly, the

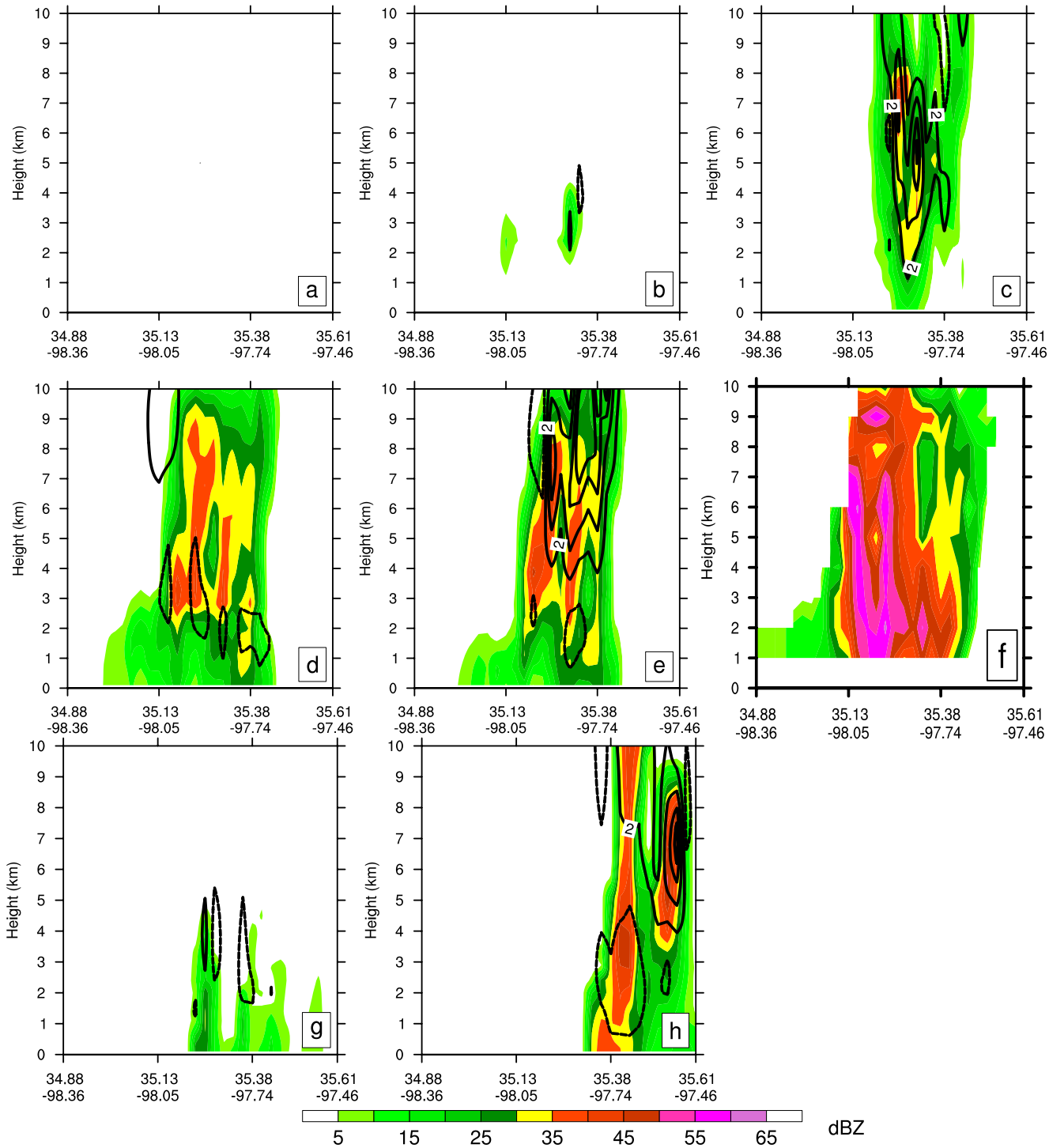

Figure 5. Vertical cross sections along the black line in Figure 1a of reflectivity (colors) and vertical velocity (contours from −4 to 10 m s^{-1} at 3 m s^{-1} interval) from the EnVar control analyses valid at 21:15 UTC for (a) Exp-noAI, (b) Exp-UNI, (c) Exp-BEC, (d) Exp-UNI_All, and (e) Exp-BEC_All and from the EnVar control first guesses (15-min forecasts) valid at 21:30 UTC for (g) Exp-UNI_All and (h) Exp-BEC_All. (f) The corresponding reflectivity observations (colors) at 21:15 UTC.

spread in Exp-BEC_All is 3–5 dBZ greater than that in Exp-UNI_All across all levels (Figures 4d and 4e). These results suggest that BEC_AI is more efficient in increasing spread than UNI_AI. Additionally, larger spread is obtained in Exp-UNI_All (Exp-BEC_All) than in Exp-UNI (Exp-BEC). Such larger spread is associated with the additional inclusion of PHW in both AI methods.

At 21:15 UTC, the first guess control fails to predict all the observed storms (not shown). Consequently, the quality of the analyzed reflectivity highly relies on the magnitude of ensemble spread. Reflectivity is still missing in the analysis of Exp-noAI (Figure 5a) due to its zero spread of reflectivity. Exp-UNI produces some weak reflectivity with the maxima below 15 dBZ between 1 and 4 km AGL (Figure 5b). The analyzed reflectivity of Exp-BEC is significantly improved across all levels, where the largest value exceeds 35 dBZ (Figure 5c). Exp-UNI_All (Figure 5d) and Exp-BEC_All (Figure 5e) share a similar pattern of the analyzed reflectivity distributions, both reaching a maximum value greater than 40 dBZ. Compared to the observation (Figure 5f), the vertical spatial structures of reflectivity in Exp-UNI_All and Exp-BEC_All are closer to reality than in the other experiments. It is noted that all analyzed reflectivity from five experiments is weaker than the observation. Such weaker reflectivity is expected in the first DA cycle and indicates the requirement of a longer spinup time, consistent with previous storm-scale radar DA studies (e.g., Dowell et al., 2011; Y. Wang & Wang, 2017, 2021; Yussouf et al., 2013).

While Exp-UNI_All and Exp-BEC_All have a similar pattern of reflectivity distributions, their analyzed w fields are remarkably different. The strong updraft with the maxima above 8 m s^{-1} in Exp-BEC_All coincides with the high magnitude of the analyzed reflectivity (Figure 5e). In contrast, the analysis in Exp-UNI_All has much weaker updrafts lower than 2 m s^{-1} corresponding to the analyzed reflectivity (Figure 5d), leading to quickly weakened convection in the subsequent 15-min forecast (Figure 5g). Such a consistent relationship between reflectivity and updraft in Exp-BEC_All favors the maintenance and enhancement of great reflectivity in the subsequent forecast (Figure 5h) and thus a fast spinup in the later cycles. In summary, Figures 4 and 5 suggest that BEC_AI provides better sampling of background errors compared to their UNI_AI counterpart.

4.1.2. Observation-Space Diagnostics

The observation-space diagnostic metrics of the root-mean-square fit to observations (RMSF) and the total ensemble spread (total spread, Dowell & Wicker, 2009) are used to quantify the performance of all experiments during DA cycling (Figure 6). The total spread is the square root of the sum of the squared observation error and the observation-space ensemble variance. These two metrics are calculated every 15 min using the assimilated KTLX radar reflectivity and radial velocity observations during 21:15–22:00 UTC. The RMSF of the posterior in Exp-noAI is the largest among five experiments consistently during the entire assimilation period, especially for the reflectivity observations. Compared to the AI experiments, the prior of Exp-noAI fits less to reflectivity during the entire assimilation period and to radial velocity at the last DA cycle. This result can be explained by its barely increased total spread, which only gets slightly enlarged after 21:45 UTC in Exp-noAI. Such smaller total spread of Exp-noAI than the AI experiments further confirms the usefulness of AI in the increase of ensemble spread. Compared to Exp-noAI, the AI experiments present the reduced RMSF of the prior and the posterior and the increased total spread for both radial velocity and reflectivity.

Differences among the AI experiments are more apparent for the reflectivity than for the radial velocity. The differences of RMSF and spread in the reflectivity during the assimilation period (Figure 6b) are consistent with those in the first DA cycle (Figures 4 and 5). Specifically, the total spread from Exp-BEC is larger than that from Exp-UNI by 1–4 dBZ during the assimilation period, especially in the first two cycles. Consequently, the analysis in Exp-BEC fits better to the reflectivity observations in Exp-UNI. The RMSF of the prior from Exp-BEC is smaller than that from Exp-UNI at the first two cycles but their RMSF of the prior becomes similar at the last two cycles. The differences in the total spread and RMSF for reflectivity between Exp-UNI_All and Exp-BEC_All are smaller than those between Exp-UNI and Exp-BEC. The total spread for reflectivity from Exp-BEC_All is slightly larger than that from Exp-UNI_All. Exp-BEC_All produces slightly smaller RMSF of the prior and slightly larger RMSF of the posterior than Exp-UNI_All. Compared to Exp-BEC (Exp-UNI), Exp-BEC_All (Exp-UNI_All) obtains enlarged total spread of reflectivity as well as reduced RMSF of the posterior and prior for reflectivity during the assimilation period.

For radial velocity, the RMSF of the posterior is very similar to each other in the AI experiments. The smallest and largest RMSF of the prior are respectively obtained by Exp-BEC_All and Exp-UNI_All. The RMSFs of Exp-UNI and Exp-BEC are similar; so are the RMSFs between Exp-BEC_All and Exp-UNI_All. The total spread of the prior from Exp-BEC and Exp-UNI are mixed. It is noted that Exp-UNI_All has consistently larger total spread than Exp-BEC_All from 21:15 to 22:00 UTC. Such more quickly enlarged total spread of Exp-UNI_All could be attributed to its more spurious convective cells around the primary storms. The spurious cells are induced by spuriously strong cold pools, which are associated with the added hydrometeor perturbations failing

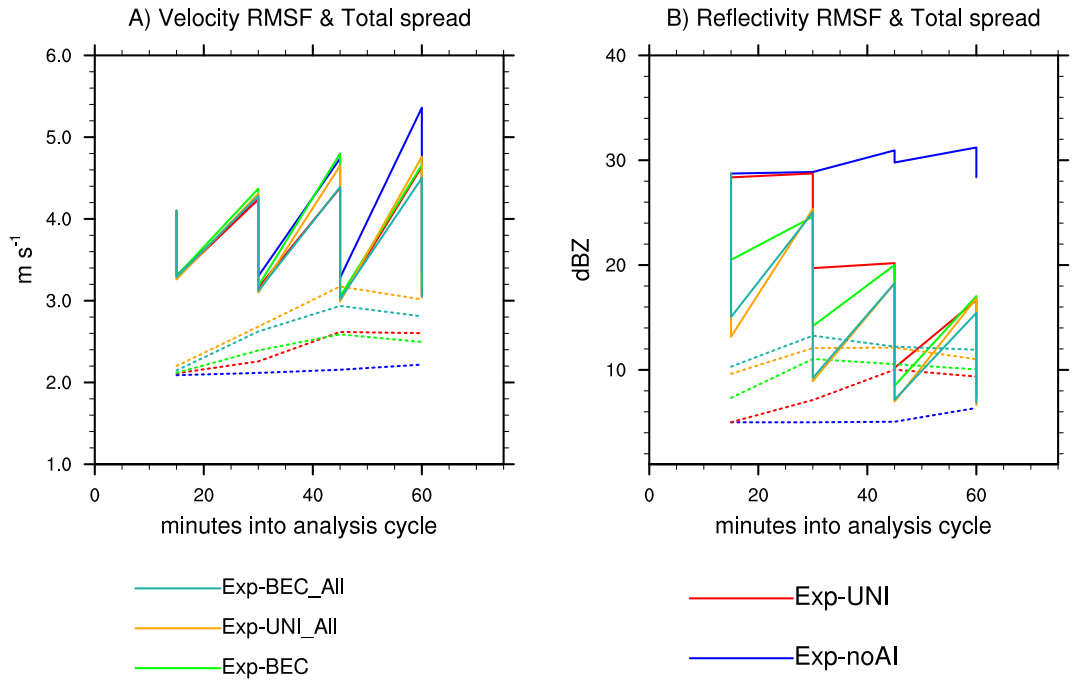

Figure 6. Observation-space diagnostic statistics for the assimilated KTLX (a) radial velocity (m s^{-1}) and (b) reflectivity (dBZ) for Exp-noAI (blue), Exp-UNI (red), Exp-BEC (green), Exp-UNI_All (orange), and Exp-BEC_All (cyan) during the entire data assimilation period. The root-mean-square fit statistics for the first guess control and analyses control shown in the upper points and lower points of the sawtooth patterns, respectively. The dotted lines represent the total spread for the first guess ensemble.

to be supported due to the absence of strong updrafts, as indicated in Figure 5g. Exp-BEC_All (Exp-UNI_All) has a larger total spread value for radial velocity than Exp-BEC (Exp-UNI).

Figure 7 compares the reflectivity analyses from each cycle for all experiments to reveal their impact on the spinup time reduction during DA cycling. Exp-noAI fails to analyze reflectivity until the third DA cycle (21:45 UTC, Figure 7a3) where some spotted reflectivity is produced as a result of its small ensemble spread in Figure 6b. At 22:00 UTC, the reflectivity analysis of Exp-noAI still misses the forward-flank region of the observed storm (Figure 7a4 vs. Figure 7o4). This result indicates that Exp-noAI fails to spinup in 1 hr. In comparison, some weak reflectivity can be analyzed at 21:30 in Exp-UNI (Figure 7b2) and 21:15 UTC in Exp-BEC, Exp-UNI_All, and Exp-BEC_All (Figure 7c1–7e1). The spinup time is remarkably reduced in all AI experiments compared to Exp-noAI. Specifically, Exp-UNI reaches the stable reflectivity value at around 21:45 UTC (Figure 7b3), and Exp-BEC, Exp-UNI_All, and Exp-BEC_All similarly reach it at around 21:30 UTC (Figures 7c2–7e2). The AI experiments share similar reflectivity analyses at the third and fourth DA cycles (21:45 and 22:00 UTC), which are comparable with the observed reflectivity (Figures 7o3 and 7o4). Note that the similar reflectivity analyses in these experiments do not necessarily mean their other state fields are similarly and reasonably analyzed, as suggested by Figure 5d. The quality of analyses from each DA cycle for all experiments will be further evaluated through verifying the corresponding free forecasts in the next subsection.

4.2. Impact on Forecasts

As all experiments fail to predict the primary characteristics of this supercell storm in the forecasts initialized at 21:15 UTC (not shown), this section will present the forecasts initialized from the analyses valid at 21:30, 21:45, and 22:00 UTC. The evaluated forecasts include the ensemble forecasts initialized from the recentered EnSRF ensemble analyses and the deterministic forecast initialized from the EnVar control analysis from each DA cycle.

4.2.1. Ensemble Probabilistic Forecasts of Low-Level Vorticity and Reflectivity

We first verify the ensemble forecasts against the observed tornado damage track to evaluate the forecast performances of five experiments on low-level rotations. A forecast tornado probability is usually used to represent the

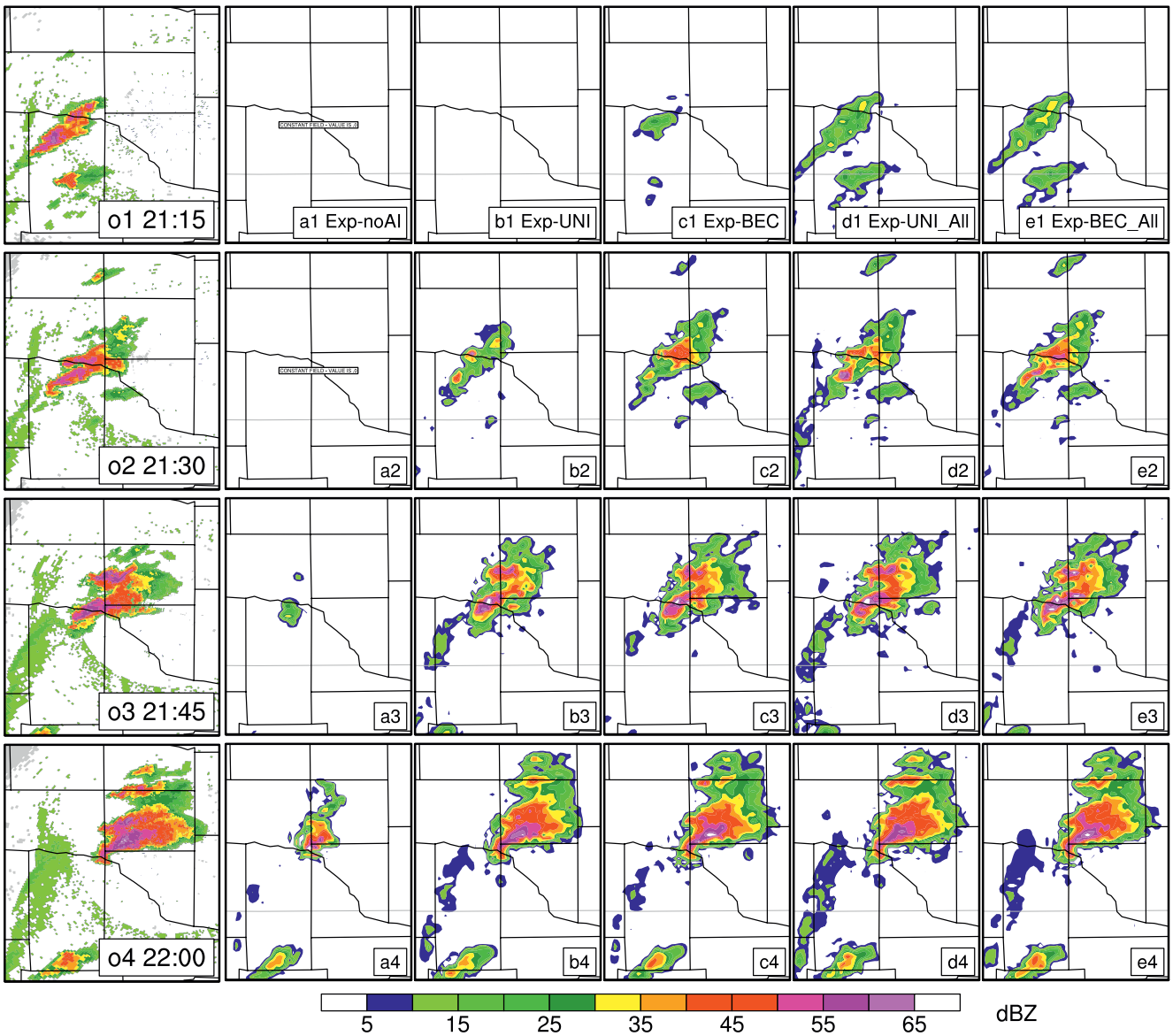

Figure 7. The reflectivity analyses at 1 km AGL for (a1–a4) Exp-noAI, (b1–b4) Exp-UNI, (c1–c4) Exp-BEC, (d1–d4) Exp-UNI_All, and (e1–e4) Exp-BEC_All from each assimilation cycle during the 1-hr data assimilation period. The corresponding observed reflectivity is shown in panels (o1–o4).

existence of low-level rotations (vorticity) because the 2-km grid spacing is too coarse to explicitly predict tornadoes (Stensrud et al., 2013; T. E. Thompson et al., 2015; Y. Wang & Wang, 2017; Yussouf et al., 2013). Consistent with Yussouf et al. (2013) and Y. Wang and Wang (2017), we calculate the forecast probabilities of vorticity exceeding 0.003 s^{-1} at 150 m AGL with a 6-km neighborhood radius at each grid point in Figure 8. As described in Section 3.1, the observed tornado damage track appeared between 22:04 and 22:38 UTC. Therefore, we only present the forecast probabilities of vorticity during 22:00–23:00 UTC for all forecasts initialized from different analysis times. Results indicate that the forecast probabilities of vorticity for all experiments enlarge along with the increase of DA cycles. These results are expected because the ensemble analyses from a later DA cycle are obtained by assimilating more radar observations and closer to the period of interest. Not surprisingly, Exp-noAI only produces some spotted swaths of the weak probability of vorticity (Figures 8a2 and 8a3), which fail to correlate with the observed tornado damage track for all three initialization times considered. In comparison, a non-zero probability of the swaths of significant mesocyclones is predicted by the forecasts initialized as early as 21:30 UTC in Exp-BEC and Exp-BEC_All, and as early as 21:45 UTC in Exp-UNI and Exp-UNI_All. The probabilities of significant low-level mesocyclones for Exp-BEC and Exp-BEC_All are 10%–40%, 40%–80%,

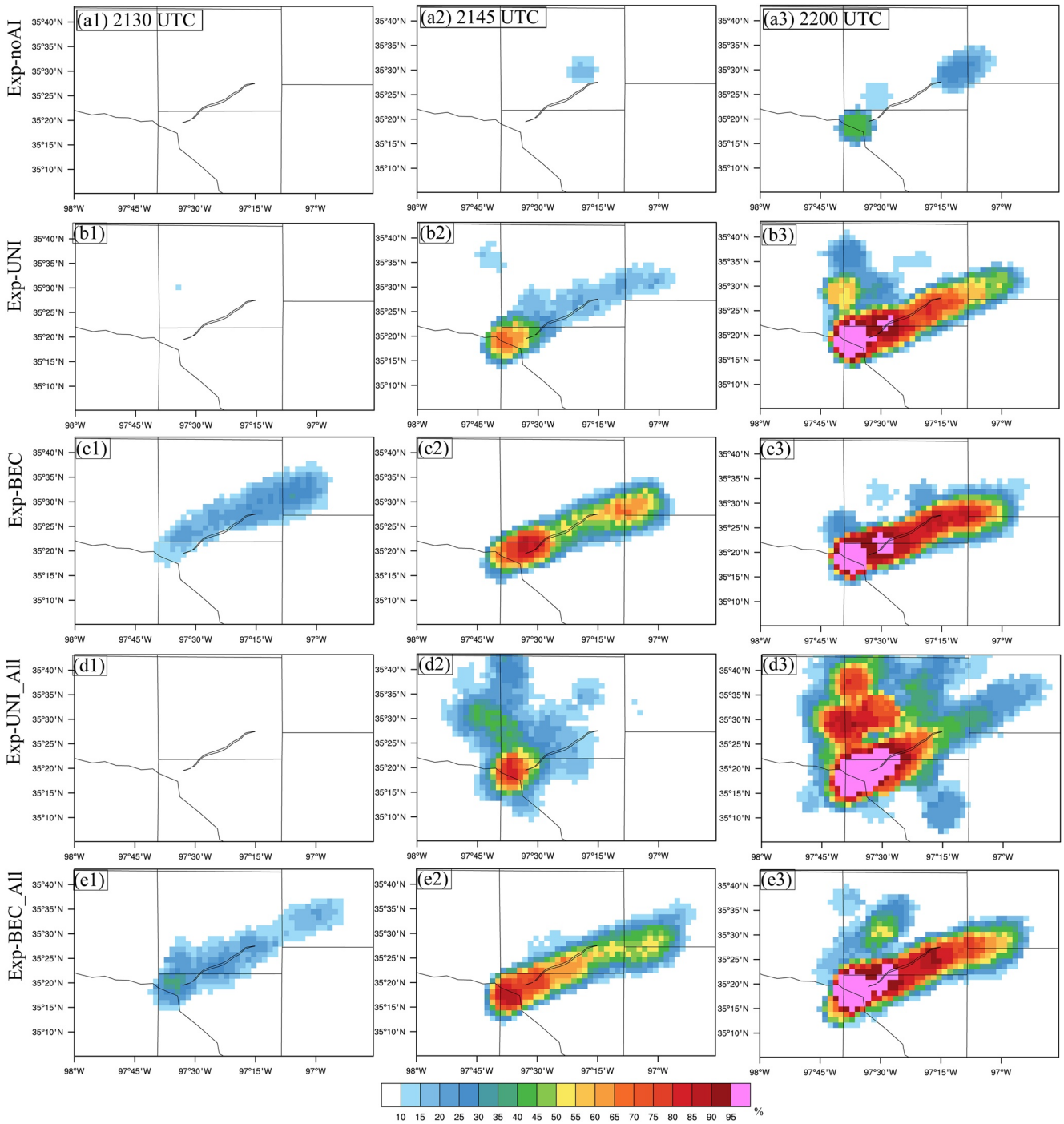

Figure 8. Neighborhood ensemble probability (%) of vorticity exceeding a threshold of 0.003 s^{-1} at 150 m AGL during 22:00–23:00 UTC initialized from the analyses at (a1–e1) 21:30, (a2–e2) 21:45, and (a3–e3) 22:00 UTC, respectively, for (a1–a3) Exp-noAI, (b1–b3) Exp-UNI, (c1–c3) Exp-BEC, (d1–d3) Exp-UNI_All, and (e1–e3) Exp-BEC_All. The overlaid lines in each panel are the observed tornado damage tracks during 22:04–22:38 UTC.

and 40%–90% higher than for Exp-noAI for the ensemble forecasts initialized at 21:30, 21:45, and 22:00 UTC, respectively. The probabilities for Exp-UNI and Exp-UNI_All are 20%–70% and 40%–90% higher than for Exp-noAI for the ensemble forecasts initialized at 21:45 and 22:00 UTC, respectively.

Exp-UNI and Exp-BEC are further compared to reveal the impact of UNI_AI versus BEC_AI on the forecast probabilities of low-level vorticity. Exp-BEC predicts a non-zero probability of the swaths of significant mesocyclones

at a forecast initialization time as early as 21:30 UTC. In comparison, Exp-UNI does not detect such swaths of significant mesocyclones until at a forecast initialization time as early as 21:45 UTC (Figures 8b2 and 8b3). The probabilities of significant mesocyclones of Exp-BEC are ∼20%–45% consistently higher than those of Exp-UNI for the ensemble forecasts initialized at the same DA cycle (Figures 8b1–8b3 vs. Figures 8c1–8c3). The improvement of the forecast probabilities of vorticity of Exp-BEC_All relative to Exp-UNI_All is similar to that of Exp-BEC relative to Exp-UNI. Specifically, all ensemble forecasts in Exp-BEC_All detect the swaths of significant mesocyclones at an initialization time as early as 21:30 UTC (Figures 8e1–8e3). In contrast, Exp-UNI_All fails to produce such a swath for the forecasts initialized at 21:30 UTC (Figure 8d1). The probabilities of low-level vorticity of Exp-BEC_All are ∼20%–40% consistently higher than those of Exp-UNI_All (Figures 8d1–8d3 vs. Figures 8e1–8e3). The selected threshold of vorticity is expected to only appear around the tornado track and its swath can extend beyond the track because this supercell that spawn the tornadoes persisted beyond the tornado track as introduced in Section 3.1. However, we notice that for both the 21:45 and 22:00 UTC initialization times, Exp-UNI_All has spurious vorticity spreading over the northwest of the tornado track for the forecasts valid after 22:45 UTC (Figures 8d2 and 8d3).

In terms of the impact of further including PHW in AI on the forecast probabilities of vorticity, ensemble forecasts from Exp-UNI (Exp-BEC) are compared with those from Exp-UNI_All (Exp-BEC_All). In the later forecast stage, Exp-UNI has less spurious vorticity and 5%–10% higher probabilities of vorticity overlapped with the observed tornado track than Exp-UNI_All for the ensemble forecasts initialized at both 21:45 and 22:00 UTC. In comparison, in the early lead time, Exp-UNI_All has ∼5% higher probabilities than Exp-UNI before 22:05 UTC (Figures 8b2 and 8b3 vs. Figures 8d2 and 8d3). Exp-BEC shares some similarities in the pattern of the swaths of significant mesocyclones with Exp-BEC_All (Figures 8c1–8c3 and 8e1–8e3). The major difference between Exp-BEC and Exp-BEC_All is the value of the probability for the ensemble forecasts initialized at 21:30 and 21:45 UTC (Figures 8c1 and 8c2 vs. Figures 8e1 and 8e2). In the first 50-min forecast period, Exp-BEC has 5%–10% lower probabilities of vorticity than Exp-BEC_All. In the later 10-min forecast period, Exp-BEC has ∼10% higher probabilities than Exp-BEC_All. In other words, during the majority of the forecast lead times, Exp-BEC_All has higher probability than Exp-BEC. These results indicate that for BEC_AI the inclusion of PHW further improves the low-level vorticity forecasts with higher probabilities at most lead times except for the brief degradation (smaller probabilities) in the later forecast stage. At later DA cycles such as 22:00 UTC, Exp-BEC and Exp-BEC_All obtain similar probabilities of low-level vorticity for the ensemble forecasts (Figures 8c3 and 8e3). In comparison, for UNI_AI, the inclusion of PHW improves briefly very short-term forecasts of low-level vorticity with higher probabilities but significantly reduces the forecasts probabilities in the remaining majority of forecast lead times.

Many previous studies have found that the evolution of a low-level mesocyclone highly relies on the characteristics of rear-flank (RF) cold pools (e.g., Lerach & Cotton, 2012; Leslie & Smith, 1978; Markowski et al., 2002, 2003; Snook & Xue, 2008; Y. Wang & Wang, 2020). A tornado-related low-level mesocyclone can be maintained only with a proper strength of RF cold pools. Too strong cold pools may lead to the tilt between near-surface mesocyclone and midlevel updraft, resulting in the demise of the mesocyclone (Marquis et al., 2012; Wakimoto & Martner, 1992; Y. Wang & Wang, 2020; Wicker & Wilhelmson, 1995). Too weak cold pools may not be able to maintain the low-level mesocyclone due to the lack of sufficient stretching (Noda & Niino, 2010; Roberts et al., 2016; Rotunno et al., 2017) and horizontal baroclinic vorticity supply (Klemp & Rotunno, 1983; Markowski et al., 2008). Here we use the cold pool forecasts of member 1 initialized at 21:45 UTC in Figure 9 as an example to understand the differences in the forecast probabilities of low-level vorticity among all experiments. Exp-noAI predicts the warmest cold pools during the 1-hr forecast period (Figures 9a1–9a7), compared to the AI experiments. The weaker cold pools than observations in Exp-noAI may be responsible for its failure in producing the significant low-level mesocyclone (Figure 8b2). The RF cold pools in Exp-BEC are 2°C–5°C weaker than Exp-UNI before 22:20 UTC (Figures 9c1 and 9c2 vs. Figures 9b1 and 9b2) but become ∼3°C stronger than Exp-UNI in the later forecast stage (Figures 9c3–9c7 vs. Figures 9b3–9b7). These stronger cold pools in Exp-BEC better match the observed surface temperature than in Exp-UNI. This result suggests that the greatly decreased probabilities of vorticity in the later forecast stage of Exp-UNI (Figure 8b2) could be attributed to its weakened RF cold pools. Compared to Exp-BEC, the more quickly weakened RF cold pools in Exp-UNI could be caused by its more rapidly decaying storms (indicated by Figures 10 and 11), probably resulting from the inconsistency between hydrometeors and other state fields during the model integration. The negative impact of such a variable-inconsistency is further amplified in Exp-UNI_All by additionally including

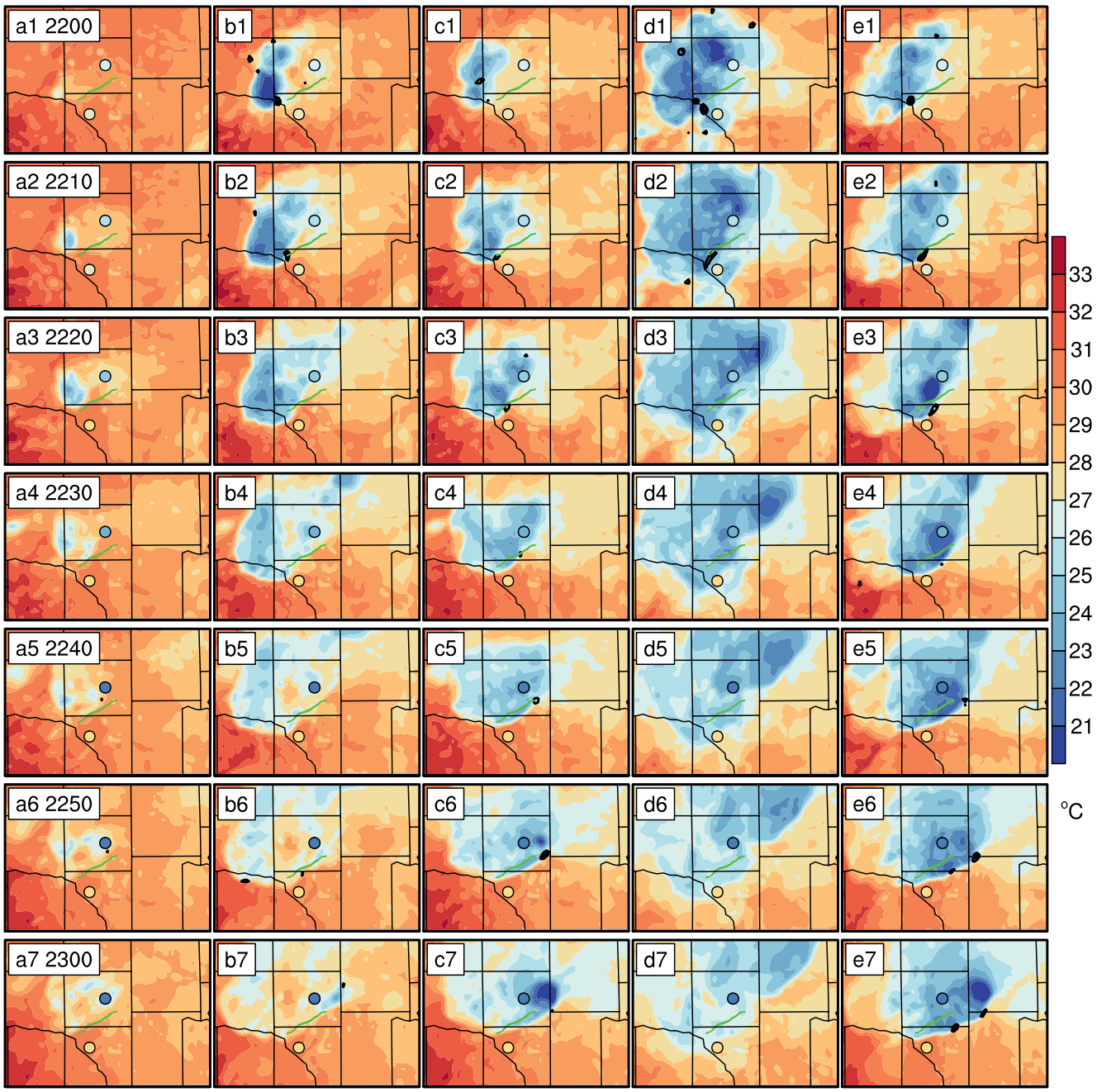

Figure 9. The 2-m temperature (colors; °C) and vertical vorticity (black contours from 0.003 to 0.009 s^{-1} at 0.003 s^{-1} interval) at 150 m AGL from the deterministic forecasts initialized from 21:45 UTC for (a1–a7) Exp-noAI, (b1–b7) Exp-UNI, (c1–c7) Exp-BEC, (d1–d7) Exp-UNI_All, and (e1–e7) Exp-BEC_All during the period from 22:00 to 23:00 UTC. The observed surface temperature from Norman and Spencer stations (labeled in Figure 2) is overlaid in color circles. Corresponding forecast time is indicated in the first column. The green line in each panel is the same as in Figure 8.

PHW, leading to the strongest and the most widespread cold pools before 22:30 UTC (Figures 9d1–9d3) among all experiments. The surface temperature observations suggest that Exp-UNI_All produces spuriously strong cold pools (Figures 9d1–9d3). The strongest cold pools result in the development of new spurious convective cells and the tilt between near-surface mesocyclone and midlevel updraft. Therefore, Exp-UNI_All produces spurious vorticity in the northwest of the tornado track and lower probabilities of vorticity in the later forecast stage (Figures 8d2 and 9d1). Compared to Exp-BEC, Exp-BEC_All produces the enhanced RF cold pools and consequently, the intensified low-level vorticity in the earlier forecast stage (Figures 9e1–9e4). The strength of cold pools in Exp-BEC_All agree well with the observations. In the later forecast stage (Figures 9e5–9e7),

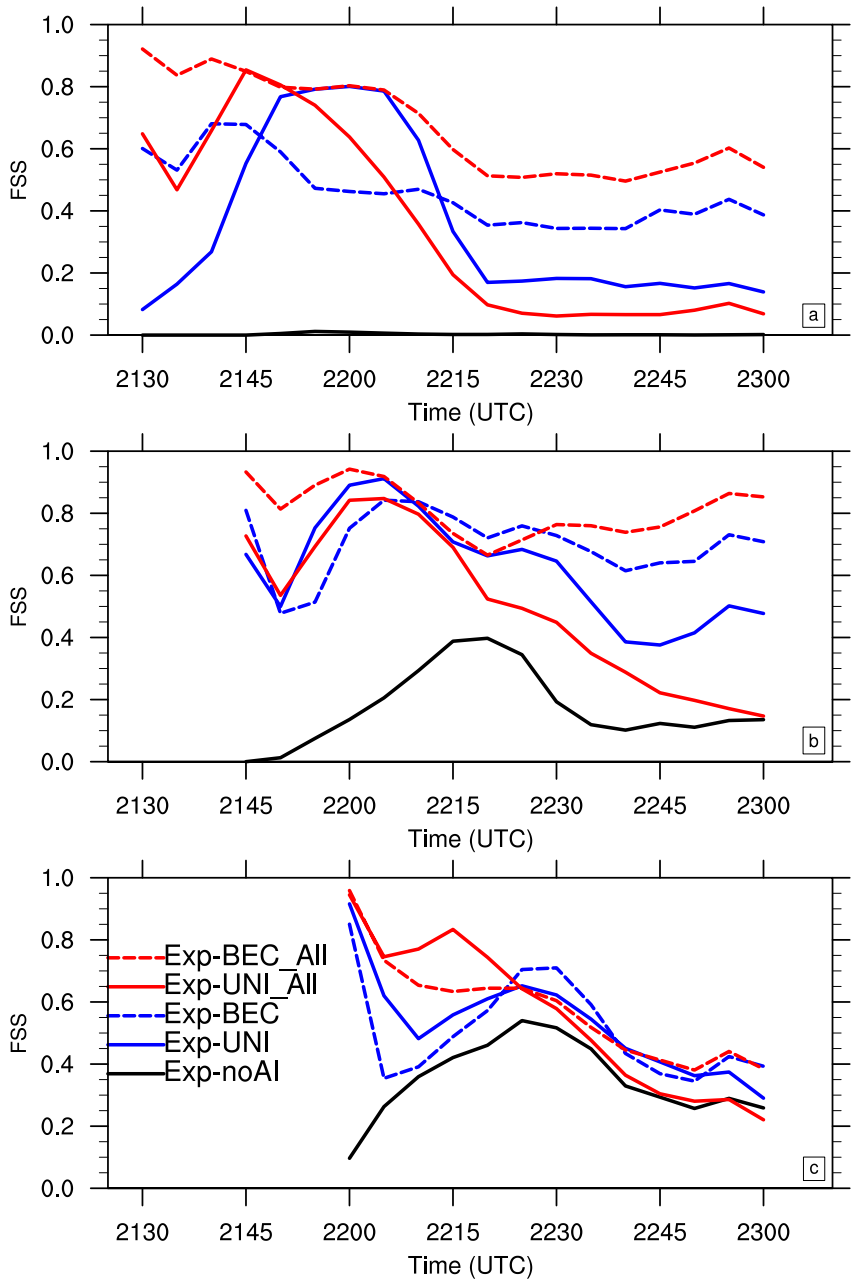

Figure 10. Fractions skill score of reflectivity exceeding 40 dBZ at 1 km AGL from the ensemble forecasts initialized at (a) 21:30, (b) 21:45, and (c) 22:00 UTC for Exp-noAI (solid-black), Exp-UNI (solid-blue), Exp-BEC (dash-blue), Exp-UNI_All (solid-red), and Exp-BEC_All (dash-red).

however, the additionally enhanced cold pools (not shown) of Exp-BEC_All for some members may result in weakened low-level vorticity (Figure 8e2). This result suggests that additional tuning efforts in Exp-BEC_All, that is, reducing the magnitude of hydrometeor and w perturbations, may be required in the future to further improve the forecast probabilities in the later lead times.

Important differences among all experiments also exist in reflectivity forecasts. The fractions skill score (FSS) is computed for the reflectivity above 40 dBZ at 1 km AGL every 5 min using a neighborhood radius of 6 km for the ensemble forecasts initialized at 21:30, 21:45, and 22:00 UTC (Figure 10). The differences in the FSS among all experiments are more pronounced for the ensemble forecasts initialized from the earlier DA cycles than those initialized at 22:00 UTC. Among all experiments, Exp-noAI generally has the lowest forecast skills. Exp-BEC

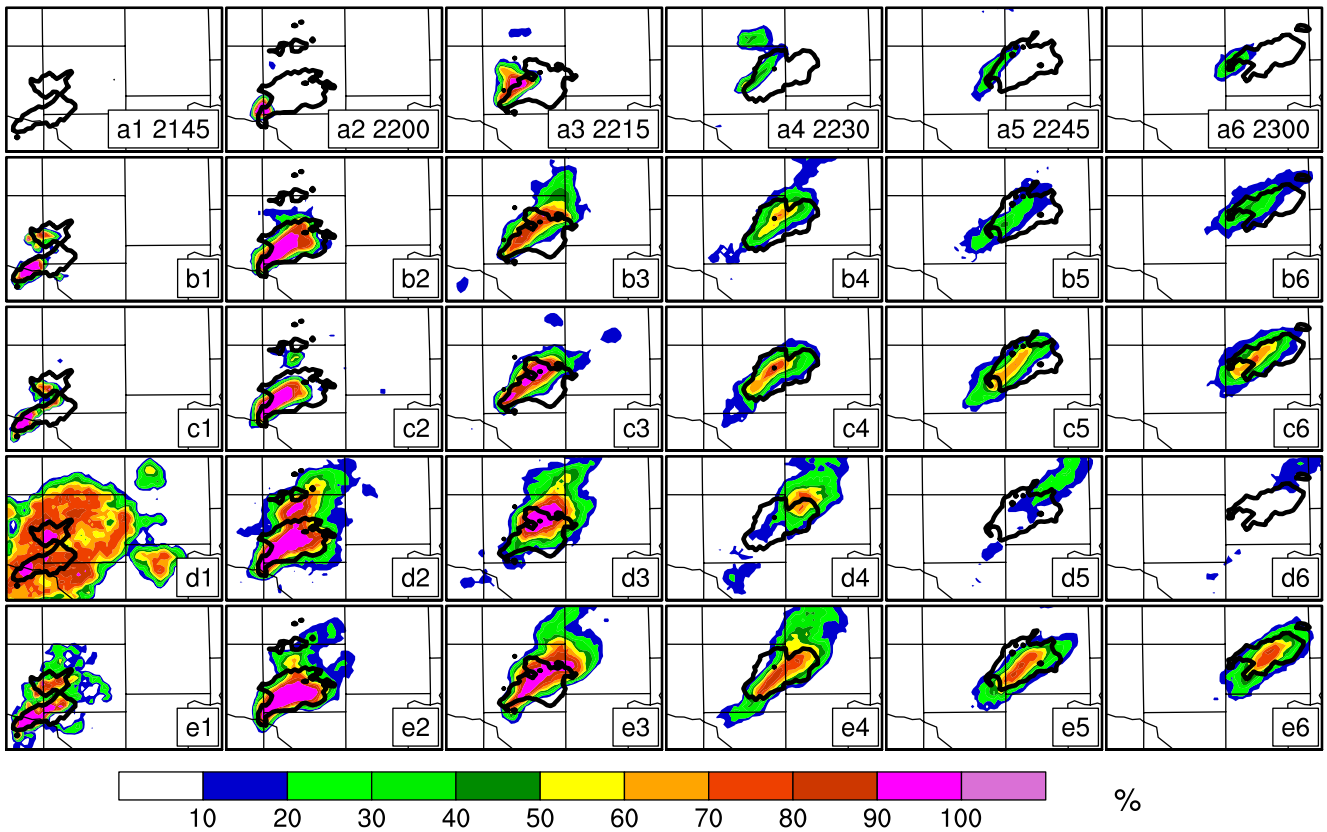

Figure 11. Probability (%) of reflectivity at 1 km AGL exceeding 40 dBZ for (a1–a6) Exp-noAI, (b1–b6) Exp-UNI, (c1–c6) Exp-BEC, (d1–d6) Exp-UNI_All, and (e1–e6) Exp-BEC_All during the period from 21:45 to 23:00 UTC. The first column shows the probability from the analyses at 21:45 UTC; the probabilities from the 15-, 30-, 45-, 60-, and 75-min forecasts are shown in the second to sixth column, respectively. The black line in each panel indicates the observed reflectivity at a threshold of 40 dBZ at the corresponding times.

outperforms Exp-UNI at most lead times except for the period around 22:00 UTC for the forecasts from 21:30 to 21:45 UTC (Figures 10a and 10b). Exp-UNI and Exp-BEC have a mixed performance for the forecasts from 22:00 UTC (Figure 10c). Exp-BEC_All consistently yields higher FSSs than Exp-UNI_All during the entire forecast period for all forecasts, except around 22:15 UTC for the forecasts from 22:00 UTC.

In terms of the impact of further including PHW on the reflectivity forecasts, Exp-UNI_All has higher skills than Exp-UNI in the first 5–25 min forecast periods and lower skills in the remaining 35–70 min lead times for all forecasts. Conversely, Exp-BEC_All produces higher FSSs than Exp-BEC during the entire 90-min forecast period for the forecasts from 21:30 UTC (Figure 10a), during 60 out of 75 forecast minutes for the forecasts from 21:45 UTC (Figure 10b), and during 45 out of 60 forecast minutes for the forecasts from 22:00 UTC (Figure 10c), respectively. Overall, the inclusion of PHW in BEC_AI improves the reflectivity forecasts at most lead times. In comparison, the inclusion of PHW in UNI_AI improves very short-term forecasts of reflectivity but quickly degrades them in the later forecast stage.

As an example, Figure 11 shows the gridpoint-based probability of reflectivity distributions at 1 km AGL exceeding a threshold of 40 dBZ calculated from the ensemble forecasts initialized at 21:45 UTC. Differences in the forecast probabilities of reflectivity distributions shown in Figure 11 are generally consistent with the FSS differences in Figure 10. Exp-noAI fails to predict a large portion of the observed reflectivity in the forward-flank region during the entire 75-min forecast period (Figures 11a1–11a6). Exp-UNI and Exp-BEC produce a similar probability at 21:45 UTC (Figures 11b1 and 11c1). However, Exp-UNI obtains a larger coverage of the high probability above 80% than Exp-BEC at 22:00 UTC (Figure 11b2 vs. Figure 11c2). After that, while the coverage with the high probabilities of 40 dBZ reflectivity in both experiments matches well with the observations, the probabilities in Exp-BEC are ~10% higher than in Exp-UNI (Figures 11b3–11b6 vs. Figures 11c3–11c6). Exp-UNI_All shows a spuriously widespread high probability area at 21:45 UTC (Figure 11d1), which suggests the presence

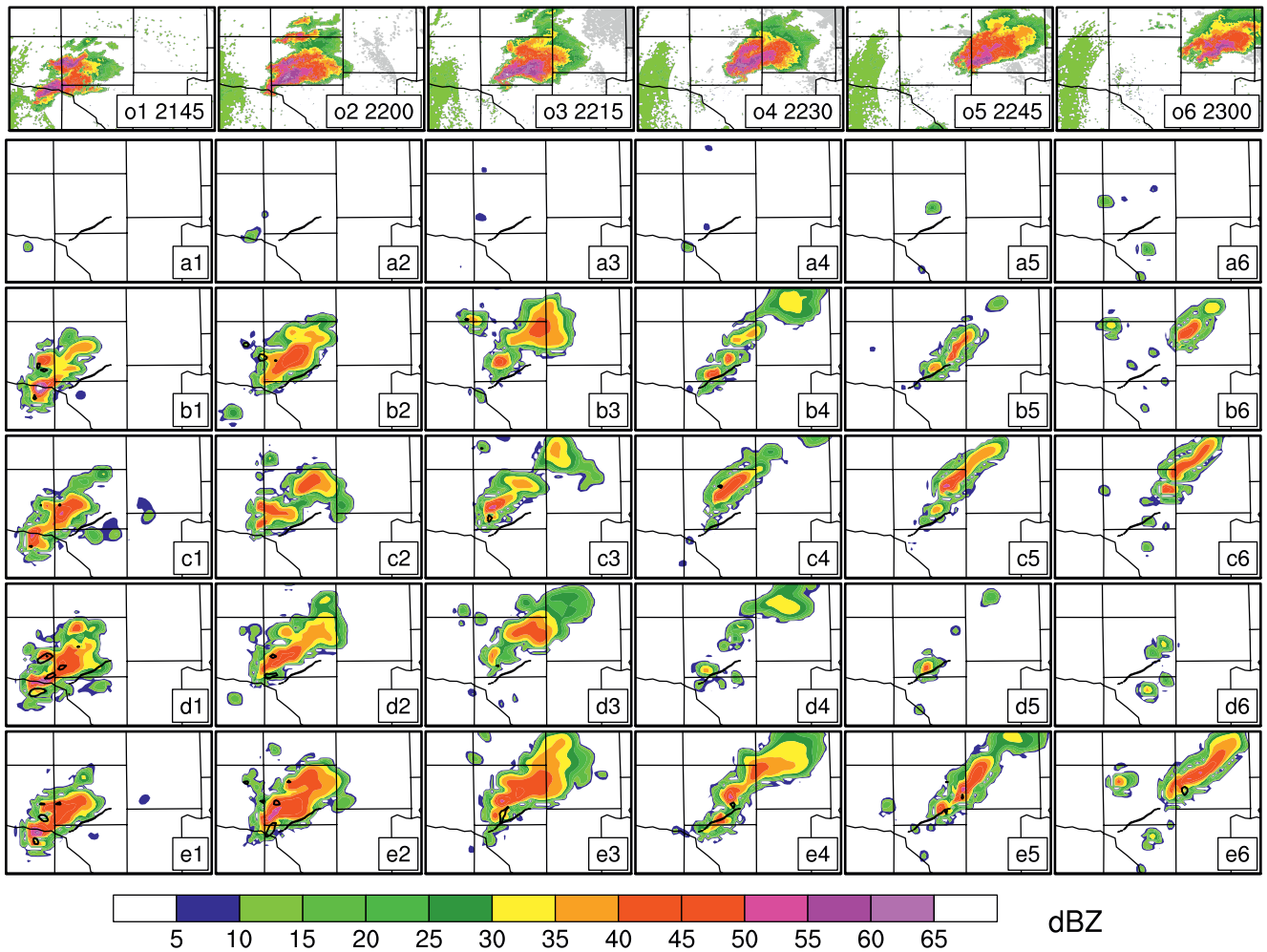

Figure 12. The reflectivity (colors; dBZ) and vertical velocity (contours from 0.003 to 0.009 s^{-1} at 0.003 s^{-1} interval) at 1 km AGL during 21:45–23:00 UTC from the forecasts initialized at 21:30 UTC of (a1–a6) Exp-noAI, (b1–b6) Exp-UNI, (c1–c6) Exp-BEC, (d1–d6) Exp-UNI_All, and (e1–e6) Exp-BEC_All. The observed reflectivity at corresponding times is shown in panels (o1–o6).

of spurious convective cells induced by too strong cold pools (Figure 8d2). In the subsequent forecast, these spurious cells quickly dissipate (Figure 11d2). The primary storm weakens from 22:15 UTC (Figure 11d3) and nearly dissipates at 22:45 UTC (Figure 11d5). The decay of storms in Exp-UNI_All is probably affected by the too-strong cold pools as well (Figures 9d1–9d3). In comparison, Exp-BEC_All predicts the reflectivity distributions with high probabilities matching well with the observations and has the best performance among all experiments (Figures 11e1–11e6).

4.2.2. Deterministic Forecasts of Reflectivity and Midlevel Updraft and Vorticity

As mentioned in Section 3.2, the control analysis is obtained using EnVar with the BECs estimated by background ensemble perturbations from the 15-min free ensemble forecast between two cycles. Therefore, the performance of control analysis and its subsequent forecast heavily depends on the quality of background ensemble perturbations. Exp-noAI fails to add reflectivity at 21:30 UTC (Figure 7a2) and misses the observed storms during the entire 90-min forecast period (Figures 12a1–12a6). In contrast, Exp-UNI analyzes weaker and narrower reflectivity than Exp-BEC at 21:30 UTC (Figures 7b2 and 7c2). Subsequently, compared to the observations (Figures 12o1–12o3), the predicted reflectivity in Exp-UNI is slightly better than that in Exp-BEC before 22:15 UTC (Figures 12b1–12b3 vs. Figures 12c1–12c3). From 22:30 UTC, the forecast storm of Exp-UNI quickly weakens (Figures 12b4–12b6) and produces a smaller reflectivity coverage than that of Exp-BEC (Figures 12c4–12c6). Exp-UNI_All and Exp-BEC_All have comparable analyzed reflectivity distributions (Figures 7d2 and 7e2). However, significantly different reflectivity distributions are produced from 22:15 UTC

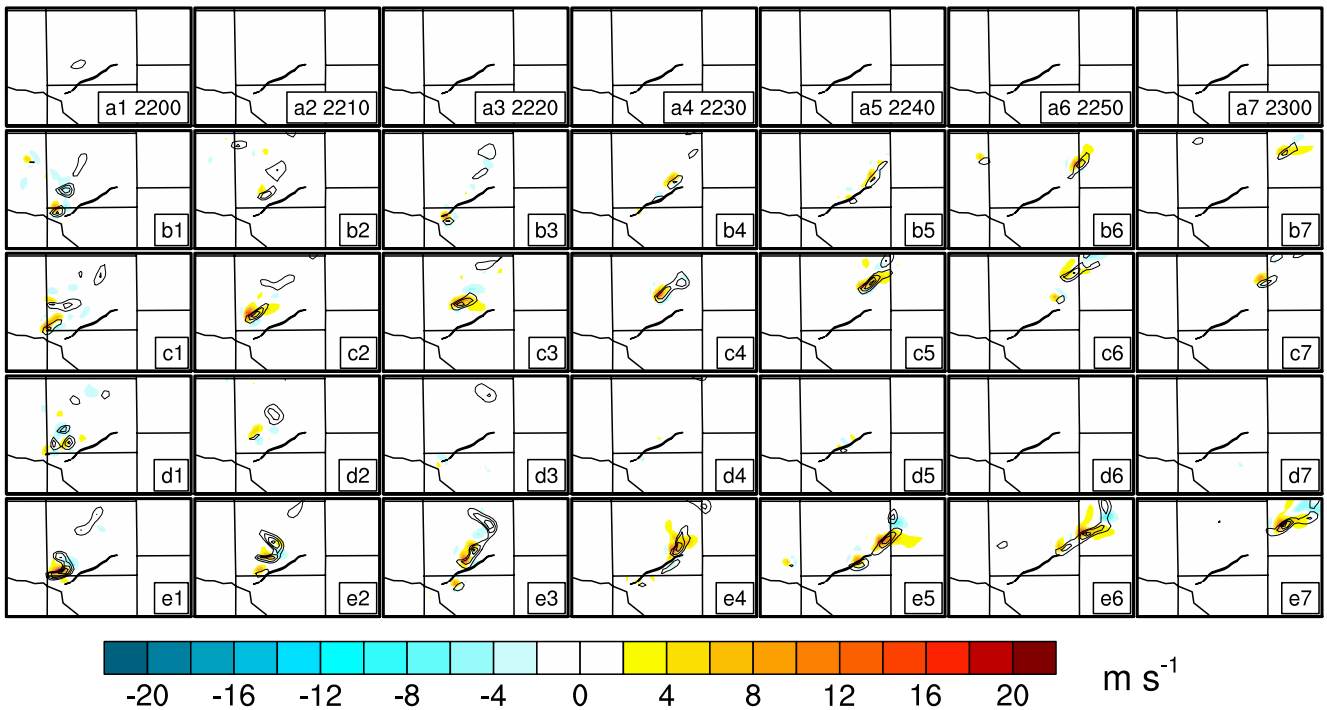

Figure 13. The evolution of vertical velocity (colors; m s^{-1}) and vertical vorticity (contours from 0.002 to 0.02 s^{-1} at 0.002 s^{-1} interval) at 4 km AGL during 22:00–23:00 UTC from the forecasts initialized at 21:30 UTC for (a1–a7) Exp-noAI, (b1–b7) Exp-UNI, (c1–c7) Exp-BEC, (d1–d7) Exp-UNI_All, and (e1–e7) Exp-BEC_All.

between Exp-UNI_All and Exp-BEC_All (Figures 12d3–12d6 vs. Figures 12e3–12e6). The forecast storm in Exp-UNI_All quickly weakens and dissipates at 22:30 UTC (Figure 12d4). Conversely, the predicted reflectivity in Exp-BEC_All maintains the hook-echo structure and the forward flank closest to the observations during the forecast period. Among all the experiments, only Exp-BEC_All has the maximum low-level vorticity above 0.003 s^{-1} near the hook echo during the entire 90-min forecast period. All experiments produce improved reflectivity distributions for the forecasts at later DA cycles (not shown). While Exp-noAI is still less skillful in reflectivity forecast than the AI experiments, differences among the AI experiments become slight.

We further examine the midlevel updraft and vorticity during 22:00–23:00 UTC to explore the performance of all experiments in the maintenance of the supercell storm. For the forecasts initialized at 21:30 UTC (Figure 13), Exp-noAI, Exp-UNI, and Exp-UNI_All fail to maintain the storm with strong updrafts exceeding 10 m s^{-1} and vorticity above 0.004 s^{-1} before 22:30 UTC. The storm in Exp-UNI is better maintained than that in Exp-UNI_All because the storm in Exp-UNI_All dissipates at 22:20 UTC (Figure 13d3). In contrast, the maximum updraft and vorticity remain above 10 m s^{-1} and 0.004 s^{-1}, respectively, before 22:40 UTC in both Exp-BEC and Exp-BEC_All. Particularly, the strong storm in Exp-BEC_All persists beyond 23:00 UTC (Figure 13d7). All experiments have an enhanced storm presented by the updraft for the forecasts from the later DA cycles (not shown). The storm in Exp-noAI is still the weakest among all experiments. Exp-UNI, Exp-BEC, and Exp-BEC_All have similar results and maintain the strong storm during the 1-hr forecast period. In contrast, the strong storm in Exp-UNI_All persists a shorter period partially due to its less balanced fields than the other experiments (not shown).

5. Summary

Ensemble-based convective-scale data assimilation (DA) commonly suffers from ensemble sampling deficiency with an insufficient amount of ensemble spread due to misrepresentation of model error. In this study, a multivariate AI method (BEC_AI) is developed to increase and maintain ensemble spread. In this method, perturbations are generated for all variables including hydrometeors and vertical velocity. Through including spatial and cross-variable correlations, perturbations are generated in a spatially and physically coherent fashion. This new method is achieved by drawing random perturbations from a recently constructed convective-scale static BEC

matrix (Y. Wang & Wang, 2021). This convective-scale static BEC matrix is extended to include hydrometeors, vertical velocity (w), and reflectivity as control variables, and to include 3D cross-correlations between any pair of control variables. The primary objective of this study is to introduce the new method and to understand the effect of this method versus a previously proposed univariate AI (UNI_AI) method (DW09). As a first step of exploring the new approach, the method is demonstrated with a 8 May 2003, OKC tornadic supercell case study.

Five experiments are conducted using the WRF-ARW model and the GSI-based EnVar DA system. Exp-noAI, without AI, serves as the baseline experiment to demonstrate the effectiveness of the AI experiments. The AI experiment named Exp-UNI uses UNI_AI to perturb conventional variables such as winds (u and v), potential temperature (t), and water vapor mixing ratio (q) only, following DW09. Exp-BEC applies BEC_AI to generate perturbations for the same variables as DW09. Exp-UNI_All is the same as Exp-UNI except to further perturb hydrometeors and w univariately. Exp-BEC_All, is the same as Exp-BEC except to further include perturbations of hydrometeors and w in a multivariate fashion. These perturbations are added to the regions with the observed reflectivity exceeding 25 dBZ for each ensemble member every DA cycle for the 8 May 2003 experiments. All experiments assimilate the same radar observations every 15 min from 21:15 to 22:00 UTC. Free forecasts ending at 23:00 UTC are launched from every DA cycle.

During DA cycling, all AI experiments have increased ensemble spread and thus produced better analyses and first guesses during the assimilation period, compared to Exp-noAI. Among the AI experiments, Exp-BEC increases reflectivity spread more efficiently than Exp-UNI. As a result, Exp-BEC has a reduced reflectivity spinup time compared to Exp-UNI. Exp-BEC_All produces larger reflectivity spread than Exp-UNI_All. Exp-BEC_All and Exp-UNI_All obtain enlarged spread for both radial velocity and reflectivity and produce first guesses better fitting to the reflectivity observations, compared to their counterparts, Exp-BEC and Exp-UNI.

Both probabilistic and deterministic forecasts initialized from all DA cycles are evaluated for all experiments. The AI experiments significantly improve the forecasts of low-level rotations, reflectivity distributions and storm maintenance, compared to Exp-noAI. Differences among AI experiments are more apparent for the forecasts initialized from the earlier DA cycles. Exp-BEC and Exp-BEC_All overall yield better forecasts than their counterparts, Exp-UNI and Exp-UNI_All, for the majority of forecast lead times. Exp-BEC_All produces more skillful forecasts for the low-level mesocyclone, reflectivity forecasts, and storm maintenance than Exp-BEC for most lead times. In contrast, compared to Exp-UNI, Exp-UNI_All poses negative impacts on the forecasts for most lead times.

As an initial effort to examine the newly proposed additive-inflation approach, a tornadic supercell case study is adopted to introduce, demonstrate, and understand the method. A systematic study using many cases to assess the robustness of BEC_AI and to determine its optimal configuration needs to be performed in the future. As discussed in Y. Wang and Wang (2021), future efforts are needed to develop the convective-scale static covariances and the AI approach for a variety of storm types.

Data Availability Statement

The WRF model (Skamarock et al., 2008) version 3.9.1.1 used to produce forecasts in this study is licensed under a Creative Commons Attribution-Noncommercial 4.0 International License, available at https://www2.mmm.ucar.edu/wrf/users/download/get_sources.html. The GSI EnVar/EnSRF (Shao et al., 2016) used to perform data assimilation can be downloaded from https://dtcenter.org/community-code/gridpoint-statistical-interpolation-gsi. The Warning Decision Support System-Integrated Information (WDSSII) software (Lakshmanan et al., 2007) can be obtained through further request at http://www.wdssii.org/download.shtml. The radar observations for assimilation and forecast evaluation can be obtained from National Centers for Environmental Information (NCEI, 1995). The surface temperature observations for verifying cold pool strength can be grabbed from the Oklahoma Mesonet (Mesonet, 1994). The Global Forecast System (GFS, 2003) analysis data was used to generate mesoscale initial ensemble.

Acknowledgments
This study was supported by NOAA (NA16OAR4320115, NA19OAR4590231, and NA21OAR4320204). The computation and diagnostic in this study were conducted at the University of Oklahoma Supercomputing Center for Education and Research (OSCER). This work also used the computing resources from the Texas Advanced Computing Center's (TACC) Stampede2 System. We acknowledge Dr. Yue Yang whose comments helped to improve an initial version of the manuscript. Finally, we thank the editor and two anonymous reviewers for their helpful feedback that greatly improved the quality of this study.

References

Aksoy, A., Dowell, D. C., & Snyder, C. (2009). A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Monthly Weather Review*, *137*(6), 1805–1824. https://doi.org/10.1175/2008MWR2691.1

Anderson, J. L., & Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, *127*(12), 2741–2758. https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2

Bowler, N. E., Clayton, A. M., Jardak, M., Lee, E., Lorenc, A. C., Piccolo, C., et al. (2017). Inflation and localization tests in the development of an ensemble of 4D-ensemble variational assimilations. *Quarterly Journal of the Royal Meteorological Society*, *143*(704), 1280–1302. https://doi.org/10.1002/qj.3004

Caya, A., Sun, J., & Snyder, C. (2005). A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Monthly Weather Review*, *133*(11), 3081–3094. https://doi.org/10.1175/MWR3021.1

Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, *129*(4), 569–585. https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Dowell, D. C., & Wicker, L. J. (2009). Additive noise for stormscale ensemble data assimilation. *Journal of Atmospheric and Oceanic Technology*, *26*(5), 911–927. https://doi.org/10.1175/2008JTECHA1156.1

Dowell, D. C., Wicker, L. J., & Snyder, C. (2011). Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Monthly Weather Review*, *139*(1), 272–294. https://doi.org/10.1175/2010MWR3438.1

Dowell, D. C., Zhang, F. L., Wicker, J., Snyder, C., & Crook, N. A. (2004). Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Monthly Weather Review*, *132*(8), 1982–2005. https://doi.org/10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2

Duda, J., Wang, X., Wang, Y., & Carley, J. (2019). Comparing the assimilation of radar reflectivity using the direct GSI based ensemble-variational (EnVar) and indirect cloud analysis methods in convection-allowing forecasts over the continental United States. *Monthly Weather Review*, *147*(5), 1655–1678. https://doi.org/10.1175/MWR-D-18-0171.1

Fujita, T., Stensrud, D. J., & Dowell, D. C. (2007). Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Monthly Weather Review*, *135*(5), 1846–1868. https://doi.org/10.1175/MWR3391.1

Gasperoni, N. A., Wang, X., & Wang, Y. (2022). Using a cost-effective approach to increase background ensemble member size within the GSI-based EnVar system for improved radar analyses and forecasts of convective systems. *Monthly Weather Review*, *150*(3), 667–689. https://doi.org/10.1175/MWR-D-21-0148.1

GFS. (2003). Legacy NCEP NOAA port analysis and forecasts [Dataset]. National Centers for Environmental Information. Retrieved from https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast

Hamrud, M., Bonavita, M., & Isaksen, L. (2015). EnKF and hybrid gain ensemble data assimilation. Part I: EnKF implementation. *Monthly Weather Review*, *143*(12), 4847–4864. https://doi.org/10.1175/MWR-D-14-00333.1

Hayden, C. M., & Purser, R. J. (1995). Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *Journal of Applied Meteorology and Climatology*, *34*(1), 3–15. https://doi.org/10.1175/1520-0450-34.1.3

Houtekamer, P. L., Mitchell, H. L., Buehner, M., Charron, M., Spacek, L., Hansen, M., & Hansen, B. (2005). Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Monthly Weather Review*, *133*(3), 604–620. https://doi.org/10.1175/MWR-2864.1

Houtekamer, P. L., Mitchell, H. L., & Deng, X. (2009). Model error representation in an operational ensemble Kalman filter. *Monthly Weather Review*, *137*(7), 2126–2143. https://doi.org/10.1175/2008MWR2737.1

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shepard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, *113*(D13), D13103. https://doi.org/10.1029/2008JD009944

Janjić, Z. I. (1990). The step-mountain coordinate: Physical package. *Monthly Weather Review*, *118*(7), 1429–1443. https://doi.org/10.1175/1520-0493(1990)118<1429:TSMCPP>2.0.CO;2

Janjić, Z. I. (1994). The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Review*, *122*(5), 927–945. https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2

Janjić, Z. I. (2001). *Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model*. Office Note (Vol. 437, p. 61). National Centers for Environmental Prediction. Retrieved from http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html%232000

Johnson, A., Wang, X., Carley, J. R., Wicker, L. J., & Karstens, C. (2015). A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Monthly Weather Review*, *143*(8), 3087–3108. https://doi.org/10.1175/mwr-d-14-00345.1

Jung, Y., Xue, M., & Tong, M. (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 data. *Monthly Weather Review*, *140*(5), 1457–1475. https://doi.org/10.1175/MWR-D-11-00032.1

Jung, Y., Xue, M., Zhang, G., & Straka, J. M. (2008). Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Monthly Weather Review*, *136*(6), 2246–2260. https://doi.org/10.1175/2007MWR2288.1

Klemp, J. B., & Rotunno, R. (1983). A study of the tornadic region within a supercell thunderstorm. *Journal of the Atmospheric Sciences*, *40*(2), 359–377. https://doi.org/10.1175/1520-0469(1983)040<0359:ASOTTR>2.0.CO;2

Labriola, J. D., & Wicker, L. J. (2022). Creating physically coherent and spatially correlated perturbations to initialize high-resolution ensembles of simulated convection. *Quarterly Journal of the Royal Meteorological Society*, *148*(748), 1–21. https://doi.org/10.1002/qj.4348

Lakshmanan, V., Smith, T., Stumpf, G., & Hondl, K. (2007). The warning decision support system–integrated information [Software]. Weather and Forecasting, *22*(3), 596–612. https://doi.org/10.1175/WAF1009.1

Lei, T., Xue, M., & Yu, T. (2009). Multi-scale analysis and prediction of the 8 May 2003 Oklahoma City tornadic supercell storm assimilating radar and surface network data using EnKF. In *Paper presented at 13th Conf. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*.

Lerach, D. G., & Cotton, W. R. (2012). Comparing aerosol and low-level moisture influences on supercell tornadogenesis: Three-dimensional idealized simulations. *Journal of the Atmospheric Sciences*, *69*(3), 969–987. https://doi.org/10.1175/JAS-D-11-043.1

Leslie, L. M., & Smith, R. K. (1978). The effect of vertical stability on tornadogenesis. *Journal of the Atmospheric Sciences*, *35*(7), 1281–1288. https://doi.org/10.1175/1520-0469(1978)035<1281:TEOVSO>2.0.CO;2

Leutbecher, M., Buizza, R., & Isaksen, L. (2007). Ensemble forecasting and flow-dependent estimates of initial uncertainty. In *Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation* (pp. 185–201).

Markowski, P. M., Richardson, Y., Rasmussen, E., Straka, J., Davies-Jones, R., & Trapp, R. J. (2008). Vortex lines within low-level mesocyclones obtained from pseudo-dual-Doppler radar observations. *Monthly Weather Review*, *136*(9), 3513–3535. https://doi.org/10.1175/2008MWR2315.1

Markowski, P. M., Straka, J. M., & Rasmussen, E. N. (2002). Direct surface thermodynamic observations within the rear-flank downdrafts of nontornadic and tornadic supercells. *Monthly Weather Review*, *130*(7), 1692–1721. https://doi.org/10.1175/1520-0493(2002)130<1692:DSTOWT>2.0.CO;2

Markowski, P. M., Straka, J. M., & Rasmussen, E. N. (2003). Tornadogenesis resulting from the transport of circulation by a downdraft: Idealized numerical simulations. *Journal of the Atmospheric Sciences*, *60*(6), 795–823. https://doi.org/10.1175/1520-0469(2003)060<0795:TRFTTO>2.0.CO;2

Marquis, J., Richardson, Y., Markowski, P., Dowell, D., & Wurman, J. (2012). Tornado maintenance investigated with high-resolution dual-Doppler and EnKF analysis. *Monthly Weather Review*, *140*(1), 3–27. https://doi.org/10.1175/MWR-D-11-00025.1

Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Review of Geophysics*, *20*(4), 851–875. https://doi.org/10.1029/RG020i004p00851

Meng, Z., & Zhang, F. (2008). Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Monthly Weather Review*, *136*(2), 522–540. https://doi.org/10.1175/2007MWR2106.1

Mesonet. (1994). Standard Mesonet observations [Dataset]. The Oklahoma Mesonet. Retrieved from http://www.mesonet.org/index.php/past_data/mesonet_data_files

Michel, Y., Auligné, T., & Montmerle, T. (2011). Heterogeneous convective-scale background error covariances with the inclusion of hydrometeor variables. *Monthly Weather Review*, *139*(9), 2994–3015. https://doi.org/10.1175/2011MWR3632.1

NCEI. (1995). NEXRAD data inventory [Dataset]. National Centers for Environmental Information. Retrieved from https://www.ncdc.noaa.gov/nexradinv/index.jsp

Noda, A. T., & Niino, H. (2010). A numerical investigation of a supercell tornado: Genesis and vorticity budget. *Journal of the Meteorological Society of Japan*, *88*(2), 135–159. https://doi.org/10.2151/jmsj.2010-203

Purser, R. J., Wu, W.-S., Parrish, D. F., & Roberts, N. M. (2003a). Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Monthly Weather Review*, *131*(8), 1524–1535. https://doi.org/10.1175//1520-0493(2003)131<1524:NAOTAO>2.0.CO;2

Purser, R. J., Wu, W.-S., Parrish, D. F., & Roberts, N. M. (2003b). Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Monthly Weather Review*, *131*(8), 1536–1548. https://doi.org/10.1175//2543.1

Reynolds, C. A., Teiseira, J., & Mclay, J. G. (2008). Impact of stochastic convection on the ensemble transform. *Monthly Weather Review*, *136*(11), 4517–4526. https://doi.org/10.1175/2008MWR2453.1

Roberts, B., Xue, M., Schenkman, A. D., & Dawson, D. T. (2016). The role of surface drag in tornadogenesis within an idealized supercell simulation. *Journal of the Atmospheric Sciences*, *73*(9), 3371–3395. https://doi.org/10.1175/JAS-D-15-0332.1

Romine, G. S., Burgess, D. W., & Wilhelmson, R. B. (2008). A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma City area tornadic supercell. *Monthly Weather Review*, *136*, 2849–2870. https://doi.org/10.1175/2008MWR2330.1

Rotunno, R., Markowski, P. M., & Bryan, G. H. (2017). "Near ground" vertical vorticity in supercell thunderstorm models. *Journal of the Atmospheric Sciences*, *74*(6), 1757–1766. https://doi.org/10.1175/JAS-D-16-0288.1

Shao, H., Derber, J., Huang, X.-Y., Hu, M., Newman, K., Stark, D., et al. (2016). Bridging research to operations transitions: Status and plans of community GSI. *Bulletin of the American Meteorological Society*, *97*(8), 1427–1440. https://doi.org/10.1175/BAMS-D-13-00245.1

Skamarock, W. C., Klemp, J. B., Dudhi, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). *A description of the advanced research WRF version 3*. (Technical Report), (pp. 1–113). University Corporation for Atmospheric Research. https://doi.org/10.5065/D6DZ069T

Snook, N., & Xue, M. (2008). Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms. *Geophysical Research Letters*, *35*(24), L24803. https://doi.org/10.1029/2008GL035866

Snyder, C., & Zhang, F. (2003). Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Monthly Weather Review*, *131*(8), 1663–1677. https://doi.org/10.1175//2555.1

Sobash, R. A., & Wicker, L. J. (2015). On the impact of additive noise in storm-scale EnKF experiments. *Monthly Weather Review*, *143*(8), 3067–3086. https://doi.org/10.1175/MWR-D-14-00323.1

Stensrud, D. J., Wicker, L. J., Xue, M., Dawson, D. T., II, Yussouf, N., Wheatley, D. M., et al. (2013). Progress and challenges with warn-on-forecast. *Atmospheric Research*, *123*, 2–16. https://doi.org/10.1016/j.atmosres.2012.04.004

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. R. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, *136*(12), 5095–5115. https://doi.org/10.1175/2008MWR2387.1

Thompson, T. E., Wicker, L. J., Wang, X., & Potvin, C. (2015). A comparison between the local ensemble transform Kalman filter and the ensemble square root filter for the assimilation of radar data in convective-scale models. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1163–1176. https://doi.org/10.1002/qj.2423

Tong, M., & Xue, M. (2005). Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Monthly Weather Review*, *133*(7), 1789–1807. https://doi.org/10.1175/MWR2898.1

Vobig, K., Stephan, K., Blahak, U., Khosravian, K., & Potthast, R. (2021). Targeted covariance inflation for 3D-volume radar reflectivity assimilation with the LETKF. *Quarterly Journal of the Royal Meteorological Society*, *147*(740), 3789–3805. https://doi.org/10.1002/qj.4157

Wakimoto, R. M., & Martner, B. E. (1992). Observations of a Colorado tornado. Part II: Combined photgrammetric and Doppler radar analysis. *Monthly Weather Review*, *120*(4), 522–543. https://doi.org/10.1175/1520-0493(1992)120<0522:OOACTP>2.0.CO;2

Wang, X., Barker, D., Snyder, C., & Hamill, T. M. (2008a). A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Monthly Weather Review*, *136*(12), 5116–5131. https://doi.org/10.1175/2008MWR2444.1

Wang, X., Barker, D., Snyder, C., & Hamill, T. M. (2008b). A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments. *Monthly Weather Review*, *136*(12), 5132–5147. https://doi.org/10.1175/2008MWR2445.1

Wang, X., Hamill, T. M., Whitaker, J. S., & Bishop, C. H. (2009). A comparison of the hybrid and EnSRF analysis schemes in the presence of model error due to unresolved scales. *Monthly Weather Review*, *137*(10), 3219–3232. https://doi.org/10.1175/2009MWR2923.1

Wang, X., Parrish, D., Kleist, D., & Whitaker, J. S. (2013). GSI 3DVar-based ensemble-variational hybrid data assimilation for NCEP global forecast system: Single resolution experiments. *Monthly Weather Review*, *141*(11), 4098–4117. https://doi.org/10.1175/MWR-D-12-00141.1

Wang, Y., & Wang, X. (2017). Direct assimilation of radar reflectivity without tangent linear and adjoint of the nonlinear observation operator in the GSI-based EnVar system: Methodology and experiment with the 8 May 2003 Oklahoma City tornadic supercell. *Monthly Weather Review*, *145*(4), 1447–1471. https://doi.org/10.1175/MWR-D-16-0231.1

Wang, Y., & Wang, X. (2020). Prediction of tornado-like vortex (TLV) embedded in the 8 May 2003 Oklahoma City tornadic supercell initialized from the subkilometer grid spacing analysis produced by the dual-resolution GSI-based EnVar data assimilation system. *Monthly Weather Review*, *148*(7), 2909–2934. https://doi.org/10.1175/MWR-D-19-0179.1

Wang, Y., & Wang, X. (2021). Development of convective-scale static background error covariance within GSI-based hybrid EnVar system for direct radar reflectivity data assimilation. *Monthly Weather Review*, *149*(8), 2713–2736. https://doi.org/10.1175/MWR-D-20-0215.1

Whitaker, J. S., & Hamill, T. M. (2012). Evaluating methods to account for system errors in ensemble data assimilation. *Monthly Weather Review*, *140*(9), 3078–3089. https://doi.org/10.1175/MWR-D-11-00276.1

Whitaker, J. S., Hamill, T. M., Wei, X., Song, Y., & Toth, Z. (2008). Ensemble data assimilation with the NCEP global forecasting system. *Monthly Weather Review*, *136*(2), 463–482. https://doi.org/10.1175/2007MWR2018.1

Wicker, L. J., & Wilhelmson, R. B. (1995). Simulation and analysis of tornado development and decay within a three-dimensional supercell thunderstorm. *Journal of the Atmospheric Sciences*, *52*(15), 2675–2703. https://doi.org/10.1175/1520-0469(1995)052<2675:SAAOTD>2.0.CO;2

Yokota, S., Seko, H., Kunii, M., Yamauchi, H., & Sato, E. (2018). Improving short-term rainfall forecasts by assimilating weather radar reflectivity using additive ensemble perturbations. *Journal of Geophysical Research: Atmosphere*, *123*(17), 9047–9062. https://doi.org/10.1029/2018JD028723

Yussouf, N., Mansell, E. R., Wicker, L. J., Wheatley, D. M., & Stensrud, D. J. (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. *Monthly Weather Review*, *141*(10), 3388–3412. https://doi.org/10.1175/MWR-D-12-00237.1

Zeng, Y., Janjić, T., de Lozar, A., Blahak, U., Reich, H., Keil, C., & Seifert, A. (2018). Representation of model error in convective-scale data assimilation: Additive noise, relaxation methods, and combinations. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2889–2911. https://doi.org/10.1029/2018MS001375

Zeng, Y., Janjić, T., de Lozar, A., Rasp, S., Blahak, U., Seifert, A., & Craig, G. C. (2020). Comparison of methods accounting for subgrid-scale model error in convective-scale data assimilation. *Monthly Weather Review*, *148*(6), 2457–2477. https://doi.org/10.1175/MWR-D-19-0064.1

Zhang, F., Snyder, C., & Sun, J. (2004). Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Monthly Weather Review*, *132*(5), 1238–1253. https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2