Error modeling of simulated reflectivity observations for ensemble Kalman filter assimilation of convective storms

Ming Xue,[1,2] Youngsun Jung,[1,2] and Guifu Zhang[2]

Received 9 March 2007; revised 16 April 2007; accepted 17 April 2007; published 18 May 2007.

[1] The impact of two different ways of modeling errors in simulated radar reflectivity data for observing system simulation experiments (OSSEs) with an ensemble Kalman filter is investigated. An error model different from the one used in earlier studies is introduced, and it specifies relative Gaussian-distributed errors in the linear domain of the equivalent radar reflectivity factor. This model is consistent with the processes of error propagation in real radar data. When the error variances specified in the filter and in the data are consistently smaller or larger, the analysis is more accurate, but when these values do not match, poorer analyses result. Such behaviors agree with expectation but are not observed when errors are directly added to the reflectivity in the log domain. These results point to the importance of properly modeling observation errors in OSSEs when the observation operator is nonlinear.
Citation: Xue, M., Y. Jung, and G. Zhang (2007), Error modeling of simulated reflectivity observations for ensemble Kalman filter assimilation of convective storms, *Geophys. Res. Lett.*, *34*, L10802, doi:10.1029/2007GL029945.

1. Introduction

[2] The ensemble Kalman filter (EnKF) data assimilation method [*Evensen*, 1994; *Houtekamer and Mitchell*, 1998; *Whitaker and Hamill*, 2002; *Evensen*, 2003] as well as other modern data assimilation methods, including 3DVAR and 4DVAR, are all based on the basic assumptions of Gaussian distributions of both background forecast and observation errors and linear observation operators [*Kalnay*, 2002]. Therefore, an EnKF analysis is optimal only when both forecast and observation error probability density functions (PDFs) are Gaussian and the relations between the model state variables and the observed quantities are linear. In reality, observations, especially those of remote-sensing platforms whose observation operators are nonlinear, often have non-Gaussian error PDFs. This is the case with the radar reflectivity factor in logarithmic form, which is often referred to simply as the reflectivity, or reflectivity in the log domain or logarithmic reflectivity (in units of dBZ). Despite this fact, studies that assimilate simulated reflectivity data directly using the EnKF methods usually assume Gaussian-distributed error in the log domain. For example, *Tong and Xue* [2005] (hereinafter referred to as TX05) and *Xue et al.* [2006] (hereinafter referred to as XTD06) both simulated reflectivity observation errors by adding noise to the simulated reflectivity that has a Gaussian distribution of zero mean and a standard deviation (SD) of 5 dBZ. In the 4DVAR OSSE (Observing System Simulation Experiment) study of *Sun and Crook* [1997], SD errors of 3 dBZ were added to the reflectivity in a sensitivity experiment, while in most other experiments, rainwater mixing ratio derived from reflectivity was assimilated. A larger sensitivity of the 4DVAR analysis to reflectivity errors is reported by them when the errors are added this way. In the OSSE study of *Caya et al.* [2005] that involved warm rain microphysics only, rain water mixing ratio was directly assimilated instead of reflectivity, and no error was added to the simulated observations.

[3] The equivalent radar reflectivity factor, Z_e (with units of $mm^6\ m^{-3}$) from which the logarithmic reflectivity is derived, represents the power return of electromagnetic waves from hydrometeor scatters. The reflectivity factor is estimated from the mean power averaged over a sequence of pulses, containing sampling error. The error PDF approaches Gaussian when there are sufficient independent samples [*Doviak and Zrnic*, 1993, sections 4.3 and 6.3] and its standard deviation is proportional to the expected value. The reflectivity in dBZ value, a logarithmic function of Z_e, however, does not have the above error statistics. Therefore, assuming a Gaussian error distribution in dBZ unit and adding such errors when simulating reflectivity data may not be appropriate.

[4] In this paper, we investigate the effects of different ways of simulating error for reflectivity on the EnKF behavior. This is accomplished by comparing twin experiments in which the errors of Gaussian distribution are specified in the linear or log domain. In section 2, the design of the experiments and the observational error models are introduced. The impact of different ways of modeling error on the EnKF analysis in OSSEs is discussed in section 3 and summarized in section 4.

2. Error Models and Experimental Design

[5] In this study, the 20 May 1977 Del City, Oklahoma supercell storm [*Ray et al.*, 1981] is simulated using the Advanced Regional Prediction System (ARPS [*Xue et al.*, 2000, 2001]) to serve as the truth for OSSEs. The model grid configuration and radar location follow TX05 exactly. The model domain is $64 \times 64 \times 16\ km^3$ with a horizontal spacing of 2 km and a vertical resolution of 0.5 km. The radar is located at the southwest corner of this grid. The simulation of radar data follows XTD06 by using a Gaussian power weighting function in the vertical for observations simulated on radar elevation levels. Fourteen elevation levels are used, using the WSR-88D precipitation scan mode (see XTD06). The ensemble square-root filter

[1]Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma, USA.
[2]School of Meteorology, University of Oklahoma, Norman, Oklahoma, USA.

0094-8276/07/2007GL029945$05.00

(EnSRF [*Whitaker and Hamill*, 2002]) procedure of XTD06 is used as well.

[6] As in TX05 and XTD06, simulated radar observations are assimilated every 5 min. until 100 min. starting from 25 min. of model time. The filter uses 40 ensemble members and a covariance localization radius of 6 km as in XTD06. No covariance inflation is employed because the ensemble spread was found to match the RMS error well and additional inflation did not significantly impact on the results. Only reflectivity is assimilated in this study to allow us to focus on the impact of reflectivity error modeling. For more details on the filter algorithm and experiment configurations, please refer to XTD06 and TX05.

2.1. Error Models

[7] For the OSSEs, simulated observations are created by adding random noise to the error-free observations. The latter are created using the observation operators with the state variables of the truth simulation as the input. In this study, the formula that links the model hydrometeors, including the rainwater, snow and hail mixing ratios, with reflectivity is the observation operator, and is described by Y. Jung et al. (Assimilation of simulated polarimetric radar data for a convective storm using ensemble Kalman filter. Part I: Observation operators for reflectivity and polarimetric variables, submitted to *Monthly Weather Review*, 2007). It is somewhat more sophisticated than that described in TX05 because of the inclusion of Mie scattering and the melting processes for snow and hail but the key components of the formulation are similar. For this study, the exact formulation of the observation operators is less important than the way the observation error is modeled.

[8] A realistic way of modeling the sampling error of reflectivity measurements is to add random errors to uncontaminated reflectivity in the linear domain or the equivalent radar reflectivity factor, Z_e^t, before converting it into the log domain to obtain Z in dBZ:

$$Z^o_{e,lin} = 10\log_{10}\left(Z^t_e + \varepsilon_e\right), \tag{1}$$

where t and o denote the uncontaminated (truth) and error-containing simulated observations, respectively. ε_e represents equivalent reflectivity errors (in units of mm^6m^{-3}) whose SD is proportional toZ_e^t [*Doviak and Zrnic*, 1993, section 6.3] and this relative error is assumed to be Gaussian-distributed with zero mean. Mathematically, $\varepsilon_e = \alpha r Z_e^t$, where α is a specified percentage factor and r is a Gaussian-distributed random number with zero mean and SD of 1. In this study, α values of 25%, 38% and 75% are found to yield effective error SDs of 1.2, 2.0, 3.5 dBZ, respectively, for $Z^o_{e,lin}$, when numerically calculated for $Z^o_{e,lin}$ data exceeding 0 dBZ at all data assimilation times. These are values that we will examine in later experiments (see Table 1). With this model, (relative) errors have Gaussian distributions in the linear domain but become non-Gaussian after they are transformed to the log domain (Figure 1).

[9] With the error model used in earlier studies (e.g., TX05), Gaussian random errors are directly added to the logarithmic reflectivity in dBZ:

$$Z^o_{e,\log} = 10\log_{10}\left(Z^t_e\right) + \varepsilon_{\log}, \tag{2}$$

where $\varepsilon_{\log}$ has a zero mean and an SD of a specified dBZ value.

2.2. Experimental Design

[10] For the purposes discussed earlier, we designed 10 experiments as listed in Table 1. In the experiment names, LG stands for "(Gaussian) error specified in the logarithmic domain of reflectivity", LN for "(Gaussian) relative error specified in the linear domain of reflectivity" and the two-digit number represents the effective error SD of assimilated observations with the decimal point between the digits omitted. 'W' in the names denotes 'wrong value', as will be explained in the next section.

3. Results of Experiments

[11] As shown in Table 1, in LG20, Gaussian errors of 2.0 dBZ SD are added directly to the reflectivity while in LN20 relative errors of 38% SD are added in the linear domain yielding an effective SD of also 2.0 dBZ in Z. Compared to the more realistic error distribution in LN20 (solid curve in Figure 1), the error model of LG20 (dashed curve) yields more large positive errors but fewer large negative errors. In LN20, the error distribution is non-Gaussian in terms of the final assimilated reflectivity. The classic optimal Kalman filter solution assumes that all error distributions are Gaussian, and that all processes, including the observation operators, are linear. The ensemble Kalman filter implementation permits nonlinearity of the prediction model and observation operators, but the under-pinning linear and Gaussian error assumptions used to derive the optimal solution are implicitly present. In the case of LG20, even though the PDF of the added error is Gaussian, because the observation operator is nonlinear, the distribution of the resulting analysis increments (i.e., the effect of observations on analysis variables) is still non-Gaussian. Because all practical data assimilation methods used in the field of meteorology, including variational, Kalman filter and ensemble-based methods, are based on linear and Gaussian assumptions, the behaviors of these assimilation methods for non-Gaussian errors require further investigation. We seek to understand the behaviors of the ensemble Kalman filter in one of such cases and to understand the effects of different error models on the assimilation results.

[12] In each of the experiments LG12W, LG35W, LN12W, and LN35W, we examine the filter behavior by specifying an observational error variance in the assimilation that is different (or wrong) from the effective error variance (squared SD) of the actual data. For instance, in LG12W and LG35W, the effective observation error SDs are 1.2 and 3.5 dBZ respectively, but 3.5 and 1.2 dBZ respectively are specified for the reflectivity observations in the filter. In other experiments, consistent SDs are used. Figures 2 and 3 compare the analysis RMS errors of model state variables from LG12, LG12W, LG35, and LG35W with the corresponding set of LN experiments (see figure captions and TX05 for details on the plots). For Figure 2, it can be seen that the analysis errors from LG12 and LG35W are close while those from LG35 and LG12W show rather similar behaviors even though LG12W errors are smaller than those of LG35 at the end of the assimilation cycles. The specified error SDs for LG12 and LG35W are both

Table 1. List of OSS Experiments

Experiment	Domain in Which Gaussian Error is Specified	Effective Error SD,[a] dBZ	Error SD Used in EnKF, dBZ
LG20	log domain	2.0	2.0
LN20	linear domain	2.0 (38)	2.0
LG12	log domain	1.2	1.2
LG12W	log domain	1.2	3.5
LG35	log domain	3.5	3.5
LG35W	log domain	3.5	1.2
LN12	linear domain	1.2 (25)	1.2
LN12W	linear domain	1.2 (25)	3.5
LN35	linear domain	3.5 (75)	3.5
LN35W	linear domain	3.5 (75)	1.2

[a]The α values in parentheses are given as percentages.

1.2 dBZ, although the effective reflectivity errors for them are 1.2 and 3.5 dBZ, respectively. These two runs produced lower analysis errors overall than LG35 and LG12W, which specified a 3.5 dBZ error SD in the filter while their effective error SDs in Z are 3.5 and 1.2 dBZ, respectively. These results suggest that when the errors are modeled in the log domain, it is the SD specified in the assimilation that to a large extent determines the quality of analysis while the effective error plays a much smaller role. This is obviously not the expected behavior of an optimal filter, where a consistency should exist between the effective error in the assimilated observations and the specified error in the assimilation system. In fact, in LG35W in which the observation error is underestimated (a wrong observation error variance is used), the analyses are actually much better than in LG35. Such a 'wrong' behavior is probably an indication that adding Gaussian errors to the reflectivity in the log domain is not the correct way. We note here that the errors in some of the experiments (e.g., LG35 and LG12W and LN12W in Figure 3) see increase in their magnitude between 60 and 80 min. This is believed to be because that during this period, the truth storm goes through a rapid storm splitting phase [see *Tong and Xue*, 2005, Figure 2] so that the forecast error has a tendency to grow faster. Such a period of error increase is prevented in the experiments in which the filter is better behaved and able to produce more accurate state estimations.

[13] The filter behavior is much more consistent when the errors are modeled in the linear domain (Figure 3). Experiment LN12, in which both effective and specified error SDs are 1.2 dBZ, produces, as expected, the lowest analysis errors (solid black curves in Figure 3) overall. Experiment LN12W, in which the error SD is wrongly specified to be 3.5 dBZ, yields analyses with larger errors (black dotted curves). Overall, the worst analysis is produced by LN35W, in which the effective error SD is large at 3.5 dBZ and the error in assimilation is specified wrongly at 1.2 dBZ; in this case the smaller specified error does not lead to a better analysis as in LG35W. In fact, the underestimated error used in the filer must have over-weighted the observations, causing large increases (instead of the expected decreases) of error by the analysis in many variables during the first two cycles. The remaining errors at the end of cycles are higher than the other cases.

[14] The analysis errors in LN35 at the end of the assimilation cycles are larger than those of LN12 and LN12W apparently because of the larger effective errors, but are smaller than those of LN35W which again agrees with expectations. The errors of LN12W increase during the intermediate cycles (around 70 min. of model time) to the level of those of LN35 and are even larger in q_c even though its effective error is relatively low, which suggests that specifying the wrong error SD in the filter is harmful. Another experiment, named LN27 (not listed in Table 1), produced RMS error curves that lie between those of LN12 and LN35. This behavior further confirms our conclusions. In reality, the Gaussian error added in the log domain does not reflect the right SD and PDF of the reflectivity sampling errors. When we examine the actual analyzed model fields (not shown due to space limit), the cases with larger RMS errors usually show nosier fields. For example, the reflectivity fields are noisier in LN35W than those in LN35 at all analysis times (not shown), apparently because too much weight is given to inaccurate observations. At 100 min., the end of the assimilation window, the analysis in LN35 agrees with the truth in the extent of the low-level cold pool and in the intensity and pattern of reflectivity better than in LN35W, in which the cold pool is too warm but its extent is too large (not shown).

[15] We performed an additional set of experiments assimilating only radial velocity which has an observation operator that is linear. The filter behaves much like the cases of LN, where correctly specified error SD in the filter produces the best results, and the analysis with smaller observation error is more accurate.

4. Summary and Discussion

[16] With OSSEs, one of the challenges is to create observations that have realistic error characteristics. An error model different from the previously used one is examined for radar reflectivity data which adds Gaussian-distributed relative error in the linear domain to equivalent reflectivity factor. This model is physically more accurate than adding Gaussian errors to the reflectivity in the log domain, and allows for larger error variances in strong echo

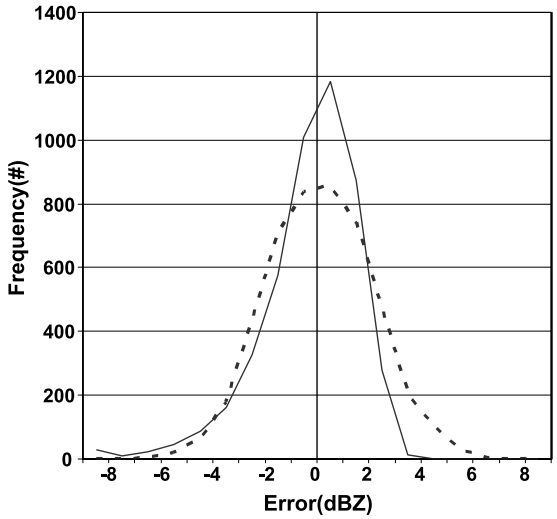

Figure 1. Frequency distributions of effective non-Gaussian reflectivity observation errors in LN20 (solid), and Gaussian reflectivity observation errors in LG20 (dashed), calculated from the data at all observation points and all analysis times.

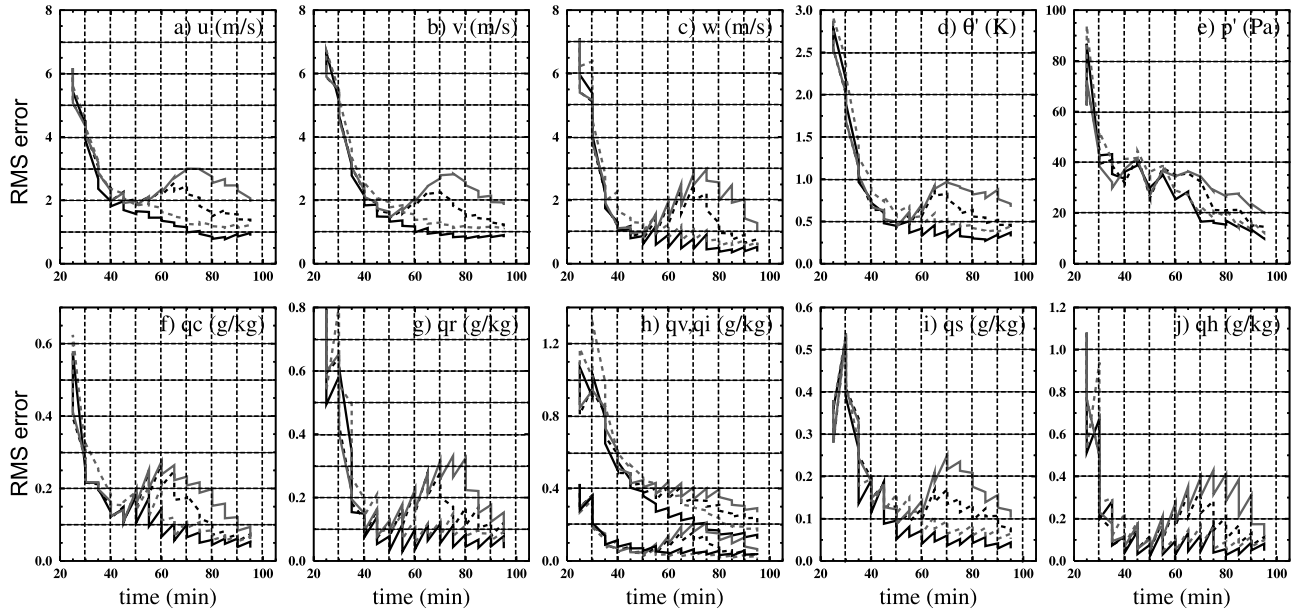

Figure 2. Ensemble-mean forecast and analysis RMS errors averaged over points at which the true reflectivity is greater than 10 dBZ for velocities (a) u, (b) v, (c) w, and (d) perturbation potential temperature θ', (e) perturbation pressure p', (f) cloud water q_c, (g) rainwater q_r, (h) water vapor q_v (larger values), cloud ice q_i (smaller values), (i) snow q_s, and (j) hail q_h; for experiments LG12 (solid black), LG12W (short dashed black), LG35 (solid gray), and LG35W (short dashed gray). Sharp drops in the curves correspond to the reduction (occasionally increase) in RMS errors by the EnKF analysis.

regions and smaller error variances in weak echo regions as expected for real radar sampling error.

[17] It is shown that when the new error model is used and the resultant observations are assimilated using the EnKF method, proper behaviors of the filter are obtained. In such a case, a smaller effective error variance in the reflectivity with a corresponding correct specification of error variance in the filter yields the best analysis and vice versa, while a mismatch between the two error variances leads to larger analysis errors. The same behavior is not found, however, when the error is added to the reflectivity in the logarithmic domain. In that case, a smaller mismatched error variance used in the filter actually results in a lower analysis error than when a correct, larger, value is used. These results point to the importance of properly modeling observation errors for OSSEs when the observation operator is nonlinear. Since the modeling of reflectivity error in the linear domain is more physically correct, we should expect that the errors in the real reflectivity data behave more similar to those simulated by the linear-domain

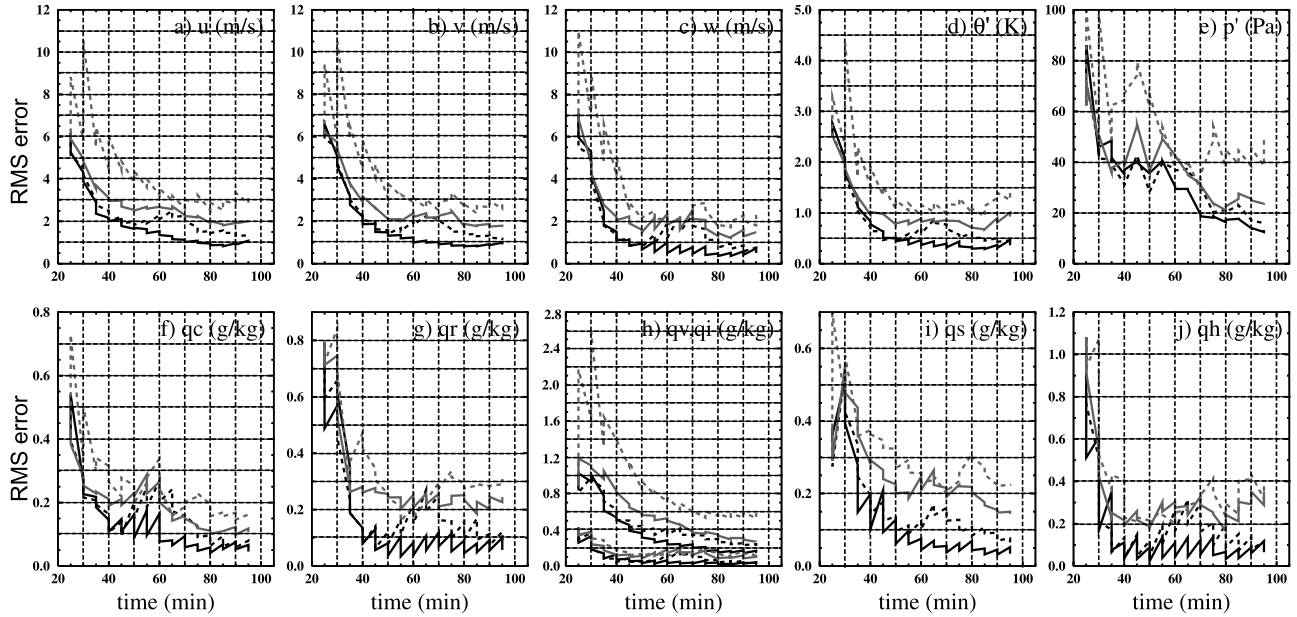

Figure 3. As in Figure 2 but for experiments LN12 (solid black), LN12W (short dashed black), LN35 (solid gray), and LN35W (short dashed gray). Note the difference in the vertical axis scale from Figure 2.

model. The understanding of the expected error distribution and magnitude should provide us guidance with the specification of error variance when we assimilate real reflectivity data.

[18] Finally, we note that the error model proposed in this study is more suitable for representing typical radar sampling errors. Gross errors associated with, e.g., ground clutter and anonymous propagation may require different treatment. Ideally these problems are eliminated through quality control or bias correction procedures before the data enter the assimilation system.

[19] **Acknowledgments.** The first author thanks Mingjing Tong for much help with the initial use of the ARPS EnKF code. William Martin is thanked for proofreading the manuscript and for helpful comments. This work was primarily supported by NSF grants EEC-0313747 and ATM-0608168. Ming Xue was also supported by NSF grants ATM-0530814, ATM-0331594, and ATM-0331756. The computations were performed at the Pittsburgh Supercomputing Center supported by NSF.

References

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

Doviak, R., and D. Zrnic (1993), *Doppler Radar and Weather Observations*, 2nd ed., 562 pp., Elsevier, New York.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10,143–10,162.

Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, *53*, 343–367.

Houtekamer, P. L., and H. L. Mitchell (1998), Data assimilation using an ensemble Kalman filter technique, *Mon. Weather Rev.*, *126*, 796–811.

Kalnay, E. (2002), *Atmospheric Modeling, Data Assimilation, and Predictability*, 341 pp., Cambridge Univ. Press, New York.

Ray, P. S., B. Johnson, K. W. Johnson, J. S. Bradberry, J. J. Stephens, K. K. Wagner, R. B. Wilhelmson, and J. B. Klemp (1981), The morphology of severe tornadic storms on 20 May 1977, *J. Atmos. Sci.*, *38*, 1643–1663.

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

Tong, M., and M. Xue (2005), Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments, *Mon. Weather Rev.*, *133*, 1789–1807.

Whitaker, J. S., and T. M. Hamill (2002), Ensemble data assimilation without perturbed observations, *Mon. Weather Rev.*, *130*, 1913–1924.

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

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

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

Y. Jung and M. Xue, Center for Analysis and Prediction of Storms, University of Oklahoma, NWC Suite 2500, 120 David Boren Boulevard, Norman, OK 73072, USA. (mxue@ou.edu)

G. Zhang, School of Meteorology, University of Oklahoma, 120 David Boren Boulevard, Norman, OK 73072, USA.