WIND AND THERMODYNAMIC RETRIEVALS IN A SUPERCELL THUNDERSTORM:
ENSEMBLE KALMAN FILTER RESULTS

David Dowell[1], Fuqing Zhang[2], Lou Wicker[3], Chris Snyder[1], Bill Skamarock[1], and Andrew Crook[1]
[1] National Center for Atmospheric Research[*], Boulder, CO
[2] Texas A&M University, College Station, TX
[3] National Severe Storms Laboratory, Norman, OK

1. INTRODUCTION

Atmospheric state estimation on the scales of deep, moist convection will be an important component of operational high-resolution numerical weather prediction. For the foreseeable future, radar measurements of Doppler velocity and reflectivity will continue to be the primary source of volumetric observations on these scales. A variety of techniques have been developed for retrieving the wind, temperature, and moisture fields from radar observations of convective storms (e.g., Armijo 1969; Hane et al. 1981; Brandes 1984; Roux 1985; Ziegler 1985; Shapiro et al. 1995; Sun and Crook 1997, 1998; Weygandt et al. 2002a,b). These retrieval techniques range from direct solutions of limited sets of governing equations to data assimilation methods employing a 3D cloud model.

Recently, Snyder et al. (2001) evaluated a relatively new data assimilation method – the ensemble Kalman filter (Evensen 1994; Houtekamer and Mitchell 1998) – for convective scale retrievals and forecasts. In tests on a simulated supercell thunderstorm, the ensemble Kalman filter scheme was stable and was able to reproduce rather accurately the original model fields from a limited number of observations (i.e., perturbed samples from the control simulation) (Snyder et al. 2001). An attractive feature of the ensemble Kalman filter approach is that once a forward model has been developed, relatively little additional coding is necessary to assimilate observations into the model. Therefore, this approach is being considered as an alternative to four-dimensional variational (4DVar) data assimilation, which requires the design and coding of an adjoint model.

In this paper, we describe the first attempts to use the ensemble Kalman filter to retrieve the wind, temperature, and other fields from radar observations of a real thunderstorm. For the retrievals, we selected the 17 May 1981 Arcadia, Oklahoma tornadic supercell case, which includes observations from two 10-cm research Doppler radars that were separated by 40 km and that were 25-55 km from the target storm (Dowell and Bluestein 1997). Although our focus here is on the basic issues of using the ensemble Kalman filter on real data, we also plan to evaluate for the same case the strengths and weaknesses of the ensemble Kalman filter compared to the 4DVar method (Crook et al. 2002) and a single-Doppler retrieval method (Weygandt et al. 2002a, b).

Corresponding author address: David C. Dowell, National Center for Atmospheric Research, Advanced Studies Program, P. O. Box 3000, Boulder, CO 80307-3000; ddowell@ucar.edu

2. DESCRIPTION OF EXPERIMENTS

The data assimilation methodology involves the following steps: generating an ensemble of initial model states, integrating the ensemble members forward in time (i.e., making forecasts), and using the ensemble Kalman filter to update the ensemble members at each time when observations are available. We use the anelastic model of Sun and Crook (1997), which includes a simple warm rain microphysical scheme for moist processes, to produce the forecasts. The model domain in these experiments consists of 51 grid points at 2000-m intervals in each horizontal direction and 34 grid points at 500-m intervals in the vertical.

The base state sounding used in the simulations (Fig. 1) includes a number of modifications to the raw sounding taken from Edmond, Oklahoma at 1430 CST (Dowell and Bluestein 1997). Based on our mesoscale analysis of multiple soundings, instrumented tower data, and winds synthesized from the dual-Doppler observations, we attempted to construct a sounding (Fig. 1) that is more representative of the environment of the mature Arcadia storm.
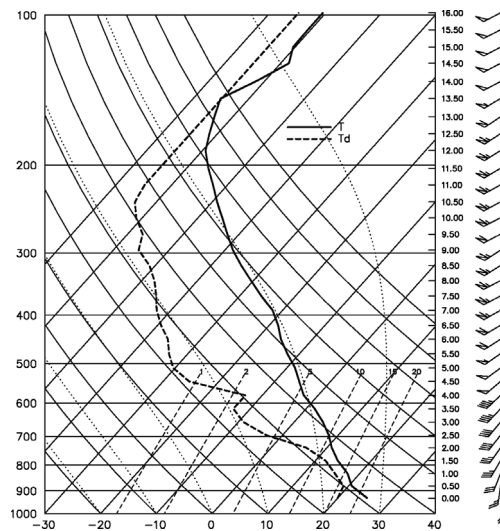

Figure 1. Sounding corresponding to the base state in the simulation. Heights (km) are relative to the lowest model level (250 m AGL).

In a sequential data assimilation scheme such as the ensemble Kalman filter, the retrieved model state after only a few radar volumes have been processed should be rather sensitive to the initial state of the model, especially in the current experiments, which involve assimilation of observations of a mature storm. (An operational data assimilation procedure might be initialized with the fields from a prior forecast. However, the current assimilation of a limited research dataset involves a more ad hoc initialization

procedure.) In the control assimilation experiment, we used the 3-D winds from a dual-Doppler wind synthesis (Dowell and Bluestein 1997) to estimate the model state ("first guess") at the initial time of the assimilation. We initialized the moisture fields in a manner similar to Weygandt et al. (2002b). We estimated the rain water mixing ratio at the initial time with the following empirical relationship (Sun and Crook 1998):

$$q_r = 3.45 \times 10^{-6} Z^{0.571} / \rho \; , \qquad (1)$$

where q_r is the rain water mixing ratio in g g^{-1}, Z is the observed reflectivity factor in mm^6 m^{-3}, and ρ is the density in g kg^{-1}. Since very high magnitudes of reflectivity in the observed storm were associated with hail, but the model does not include ice physics, we truncated the calculation in (1) at 55 dBZ. Throughout most of the model domain, we initialized the total water as the sum of q_r calculated in (1) and the sounding value of the water vapor mixing ratio. However, in regions of $q_r > 0.1$ g kg^{-1}, we increased the relative humidity. In updraft regions (grid points where the vertical velocity from the dual-Doppler synthesis was >5 m s^{-1}), we increased the relative humidity to 100%. Otherwise, we increased the relative humidity to 80% within the precipitation region.

We produced an ensemble of 100 initial model states for the assimilation experiments by adding perturbations to the first guess state. To initialize the state of a particular ensemble member, we added Gaussian noise to the first guess over a limited (40 km wide) region centered on the observed precipitation core of the storm. The standard deviation of the random noise was 5 m s^{-1} for the horizontal velocity components and 5 K for temperature. We linearly varied the standard deviation of the noise in vertical velocity from 0 at the surface and tropopause to 10 m s^{-1} at 7 km AGL. We did not add noise to the first guess estimates of rain water and total water. Although the magnitude of the initial velocity noise was relatively high, by the time that the first observations were assimilated (i.e., after 4 min of model integration), the ensemble standard deviation of radial velocity had decreased to a magnitude comparable to the expected Doppler measurement error (~2 m s^{-1}).

Our initial experiments involved assimilation of only observations of Doppler velocity (although reflectivity information was used for the first guess state). We used a 1000 m Cressman radius of influence to interpolate the raw velocity observations within each sweep to the horizontal locations corresponding to the model scalar grid points. (Each radar collected volumes consisting of sweeps at 12-15 different elevation angles.) Interpolating the observations within each sweep separately, rather than interpolating the entire volume to a standard Cartesian grid, allows one to retain the actual heights of the observations on the conical sweep surfaces; this method of data handling minimizes errors associated with vertical interpolation/extrapolation (Sun and Crook 2001). We assumed that observation errors were uncorrelated; therefore, observations were processed one at a time. The covariances among perturbations in the model states were computed locally rather than globally (Houtekamer and Mitchell 1998). We used a spherical influence region with a radius of 4 km around each observation.

3. RESULTS

The control experiment was initialized, as described in the previous section, at 16:30 CST. Then, interpolated observations from both radars were assimilated at 16:34, 16:38, 16:43, and 16:47. At each analysis time, all of the observations from the Cimarron radar (southwest of the storm) were processed before those from the Norman radar (south of the storm).

Although a comparison of the retrieved state to the truth is not possible for a real data case, one may evaluate the quality of the retrieval by comparing forecasts from the retrieved states to the observations. The differences (the values in the "before" columns in Table 1) between the observed radial velocities and the corresponding velocities in the model after 4-5 min forecasts generally decreased with time, as would be expected in a successful assimilation. However, some of the other ensemble statistics are less encouraging. In particular, the standard deviation in radial velocity among the ensemble members (Table 2) decreased rather quickly and was relatively small compared to both the expected observation error (~2 m s^{-1}) and the computed difference between the forecast and the observations (Table 1). Most of the observations were outside the envelope of forecast states in the ensemble (not shown).

Table 1. RMS difference (m s^{-1}) between the observed radial velocity and the ensemble mean radial velocity, both before the Kalman filter (i.e., after the forecast) and after the Kalman filter, for the control experiment. The differences are with respect to observations from the Cimarron and Norman radars.

	Cimarron before filter	Cimarron after filter	Norman before filter	Norman after filter
16:34	5.2	2.9	6.3	3.6
16:38	5.4	3.9	6.3	4.7
16:43	4.1	2.9	4.9	3.5
16:47	3.9	3.0	4.9	3.6

Table 2. Ensemble standard deviation (m s^{-1}) of radial velocity at the observation points for the control experiment.

	Cimarron before filter	Cimarron after filter	Norman before filter	Norman after filter
16:34	1.7	0.9	1.5	0.9
16:38	1.1	0.7	0.9	0.7
16:43	1.0	0.6	0.8	0.6
16:47	0.8	0.5	0.7	0.5

The minimum temperature perturbation (-3.9 K) in the retrieved cold pool at 0.25 km AGL (Fig. 2a) is comparable in magnitude to the minimum temperature perturbation (-5.5 K) measured by an instrumented tower in the Arcadia storm at that level (Dowell and Bluestein 1997). Since we were curious whether the cold pool developed primarily as a result of evaporation of rain in the forward model integration or as a result of the adjustment of the model states with the Kalman filter, we conducted a second assimilation experiment similar to the control experiment, except that we did not use the same moisture specification step in the initialization at 16:30. Instead, we initialized the model with no rain water and with the magnitude of water vapor in the environmental sounding.

The results of the assimilation that included no rain water in the initial state indicated relatively small temperature perturbations (of opposite sign of those in the control experiment) at low levels (Fig. 2b), even after observations at four times had been assimilated. Therefore, in this case, it is the evaporation of rain in the model, rather than the direct retrieval of temperature from the velocity observations, that is responsible for the low-level cold pool. Since the radar horizon precluded radar scanning near the surface, we suspect that the shallow divergence signature that would have been associated with the cold pool was absent in the observations (Weygandt et al. 2002b). The lowest Doppler velocity measurements in this case were at approximately 300-800 m AGL. Since a typical range from a storm to the closest WSR-88D would be even greater than the range (25-55 km) in this case, we would anticipate similar, or even greater, uncertainty in retrieving the magnitudes of storm-scale cold pools from Doppler velocity observations alone in an operational setting.

Although a cold pool did not develop by 16:47 CST in the assimilation with no rain water in the initial state, a strong, warm, moist updraft aloft had developed aloft by that time (not shown). We produced 30-min forecasts from the retrieved ensemble mean state at 16:47 for both the control experiment (Fig. 3b) and the second experiment just described (Fig. 3c). In each case, the simulated storm was in the approximate location of the observed storm (Fig. 3a). The differences in the initializations were associated with differences in the sub-storm-scale structures (Figs 3b and 3c).

In a third experiment, we assimilated observations from only one radar (Cimarron) into the model. This experiment is more representative of an operational retrieval and prediction, when dual-Doppler observations are not available. We included the reflectivity based estimate of rain water in the first guess state. Otherwise, we initialized the remaining model fields with the base state values in the sounding. The assimilation of observations from only the Cimarron radar from 16:30 to 16:47, and the 30-min forecast from the retrieved ensemble mean state at 16:47, produced a significant storm in the correct location (Fig. 3d).
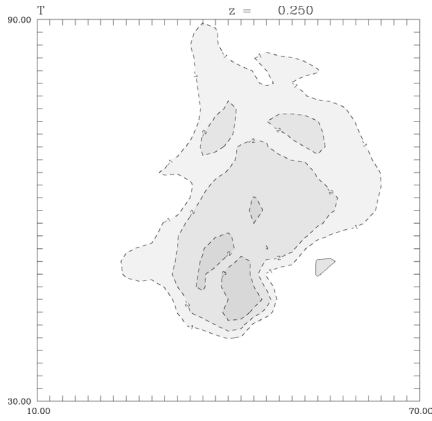
4. FUTURE WORK

The ensemble Kalman filter assimilation of Doppler observations of the 17 May 1981 Arcadia, Oklahoma storm into a numerical model was a stable procedure, and the 30-min forecast from the retrieved model state produced a storm in approximately the correct position. However, a number of challenges must be addressed before an operational implementation of the data assimilation methodology can be considered. The most significant problem appears to be the rapid convergence of the ensemble members. Even though a relatively large ensemble (100 members) was utilized, the ensemble standard deviation of velocity quickly decreased to a magnitude less than the typical observation and forecast errors. Ways for allowing for model error and/or inflating the magnitude of covariance in the ensemble are being considered.

Our experiments with the Arcadia case have just begun. In the near future, we are planning experiments with higher model resolution and with a more appropriate scheme for precipitation microphysics. The results will be evaluated with a detailed comparison of the retrieved states to the observatio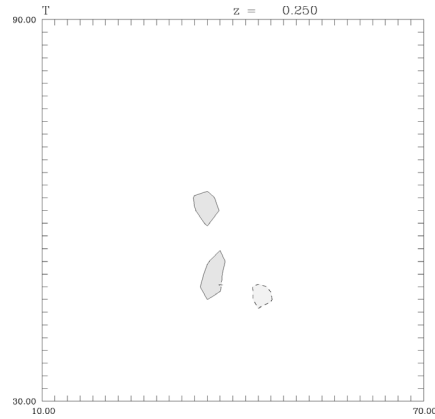ns. We are also planning to compare the results of the ensemble Kalman filter data assimilation method to those of other methods, particularly 4DVar and single-Doppler retrieval.

REFERENCES

Armijo, L., 1969: A theory for the determination of wind and precipitation velocities with Doppler radars. *J. Atmos. Sci.*, **26**, 570-573.

Brandes, E. A., 1984: Relationships between radar-derived thermodynamic variables and tornadogenesis. *Mon. Wea. Rev.*, **112**, 1033-1052.

Crook, A., D. Dowell, and J. Sun, 2002: Wind and thermodynamic retrievals in a supercell thunderstorm: 4DVar results. This volume.

Dowell, D. C., and H. B. Bluestein, 1997: The Arcadia, Oklahoma storm of 17 May 1981: Analysis of a supercell during tornadogenesis. *Mon. Wea. Rev.*, **125**, 2562-2582.

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), 10143-10162.

Hane, C. E., R. B. Wilhelmson, and T. Gal-Chen, 1981: Retrieval of thermodynamic variables within deep convective clouds: Experiments in three dimensions. *Mon. Wea. Rev.*, **109**, 564-576.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796-811.

Roux, F., 1985: Retrieval of thermodynamic fields from multiple-Doppler radar data using the equations of motion and the thermodynamic equation. *Mon. Wea. Rev.*, **113**, 2142-2157.

Shapiro, A., S. Ellis, and J. Shaw, 1995: Single-Doppler velocity retrievals with Phoenix II data: Clear air and microburst wind retrievals in the planetary boundary layer. *J. Atmos. Sci.*, **52**, 1265-1287.

Snyder, C., F. Zhang, J. Sun, and A. Crook, 2001: Tests of an ensemble Kalman filter at convective scales. *Preprints, 14th Conf. on Numerical Weather Prediction*, 444-446.

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

Sun, J., and N. Andrew Crook, 1998: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part II: Retrieval experiments of an observed Florida convective storm. *J. Atmos. Sci.*, **55**, 835-852.

Sun, J., and N. Andrew Crook, 2001: Real-time low-level wind and temperature analysis using single WSR-88D data. *Wea. Forecasting*, **16**, 117-132.

Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002a: Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part I: Single-Doppler velocity retrieval. *Mon. Wea. Rev.*, **130**, 433-453.

Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002b: Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part II: Thermodynamic retrieval and numerical prediction. *Mon. Wea. Rev.*, **130**, 454-476.

Ziegler, C. L., 1985: Retrieval of thermal and microphysical variables in observed convective storms. Part 1: Model development and preliminary testing. *J. Atmos. Sci.*, **42**, 1487-1509.
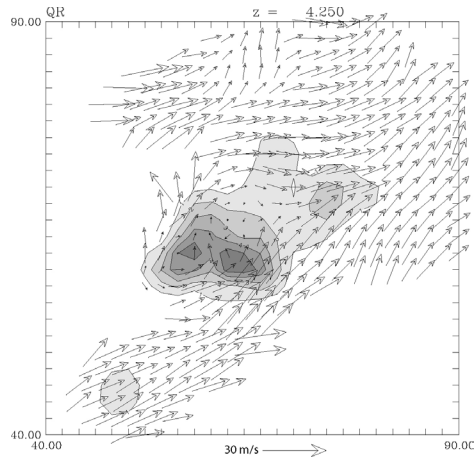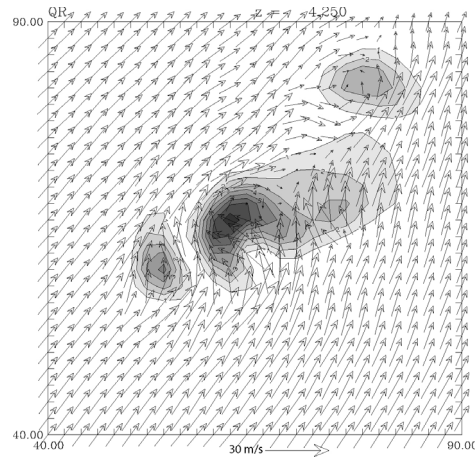
a) with rain in initial state
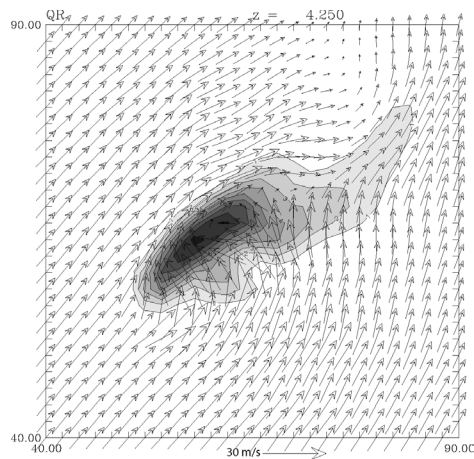

b) without rain in initial state

Figure 2. Ensemble mean of the retrieved temperature perturbation (contours and shading at intervals of 1 K) at 16:47 CST at 0.25 km AGL. A 60 km × 60 km portion of the domain is shown.
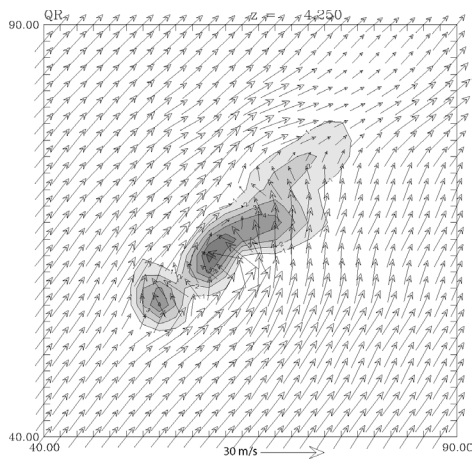

a) Results of a dual-Doppler wind synthesis (rain water mixing ratio estimated from reflectivity).

b) Results of a 30 min forecast initialized with the retrieved ensemble mean state at 16:47 CST.

c) As in (b), except with no rain water in the initial state.

d) As in (b), except for Cimarron radar data only.

Figure 3. Rain water mixing ratio (contours and shading at intervals of 1.0 g kg^{-1}) and horizontal storm-relative winds (vectors) at 4.25 km AGL at 17:17 CST. A 50 km × 50 km portion of the domain is shown.