Low-level Polarimetric Radar Signatures in EnKF Analyses and Forecasts of the 8 May 2003 Oklahoma City Tornadic Supercell: Impact of Multi-moment Microphysics and Comparisons with Observations

Daniel T. Dawson II¹, Louis J. Wicker³, Edward R. Mansell³, Youngsun Jung¹ and Ming Xue¹²

Center for Analysis and Prediction of Storms¹ and School of Meteorology²
University of Oklahoma, Norman Oklahoma 73072

NOAA/National Severe Storms Laboratory³
National Weather Center
120 David L. Boren Blvd.
Norman, OK 73072

May 2013

Submitted to Advances in Meteorology

Corresponding author address:
Daniel T. Dawson II
Center for Analysis and Prediction of Storms
University of Oklahoma,
120 David L. Boren Blvd, Norman OK 73072
ddawson@ou.edu
Abstract

The impact from increasing the number of predicted moments in a multi-moment bulk microphysics scheme is investigated using ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm and the analyses are validated using dual-polarization radar observations. The triple-moment version of the microphysics scheme exhibits the best performance, relative to the single- and double-moment versions, in reproducing the low-$Z_{DR}$ hail core and high-$Z_{DR}$ arc, as well as an improved probabilistic track forecast of the mesocyclone. A comparison of the impact from the improved microphysical scheme on probabilistic forecasts of the mesocyclone track with the observed tornado track is also discussed.
1. Introduction

The assimilation of radar data into storm scale models using the Ensemble Kalman Filter (EnKF)[1] approach has proven to be an extremely useful tool for the analysis and prediction of convective storms in recent years. There have been many recent successful uses of this approach for both analyses [2-7] and short-range forecasts[8-10] based on these analyses. In general, these studies have focused on improving techniques for assimilation of radar data, on the design of the overall data assimilation system, or on the impact of initial and boundary conditions. High-resolution numerical weather prediction has progressed during the past decade such that prediction of the dynamics of individual convective storms is now routinely attempted. One substantial challenge is the improvement and validation of the microphysics parameterization and the associated impacts on storm structure and behavior (e.g., through the development of the cold pool). Errors from the model’s microphysical parameterization can significantly impact forecasts of these storms. Polarimetric radar observations offer a rich source of data to validate the output of such schemes within this context.

Several storm-scale simulation studies have shown that the microphysics parameterization has a profound impact on simulated storm structure and behavior [11-17], and even on tornadic potential[18]. Here, we restrict our discussion to bulk microphysics schemes, which assume \textit{a priori} a certain functional form for the underlying drop or particle size distribution (DSD/PSD) for several hydrometeor categories. Typically, one or more moments of the PSD for a given category are explicitly predicted within a scheme, with single-moment schemes that predict the mass mixing ratio (proportional to the 3\textsuperscript{rd} moment) being the most common. Double-moment schemes that typically predict total number concentration (0\textsuperscript{th} moment) in addition to mixing ratio are becoming increasingly common[19-22], and triple-
moment schemes that additionally predict the 6\textsuperscript{th} moment (proportional to radar reflectivity factor) are also available[23]. Milbrandt and Yau [11] and Dawson et al. [12] found that a multi-moment bulk microphysics parameterization, as opposed to the more traditional single-moment schemes, better simulated the reflectivity structure and cold pool intensity as compared with observations, in their respective supercell storm simulation experiments. Dawson et al. [8] briefly investigated the impact of decreasing the complexity of the microphysics parameterization (from double- to single-moment) on their EnKF-based ensemble forecasts of the 4 May 2007 Greensburg, KS tornadic supercell and found that doing so substantially degraded the forecast track of the simulated mesocyclone across the ensemble. Similarly, Yussouf et al. [10] showed that the double-moment schemes they considered in their EnKF-based forecast experiments generally performed better than the single-moment schemes in predicting the forecast track of the 8 May 2003 Oklahoma City tornadic supercell.

Polarimetric radar offers an unprecedented amount of information about microphysical characteristics within convective storms[24-33]. Jung et al. [34] utilized a polarimetric emulator to successfully evaluate the performance of a single- and double-moment microphysics scheme. They found that the double-moment scheme performed substantially better than its single-moment counterpart in simulating commonly observed polarimetric radar signatures in supercell storms, such as the $Z_{DR}$ arc and the low-$Z_{DR}$ hail signature[35], due at least in part to its ability to parameterize the important process of gravitational size sorting of hydrometeors.

In the current study, we investigate the impact of systematically increasing the microphysics complexity on EnKF analyses and ensemble forecasts of the 8 May 2003 Oklahoma City tornadic supercell, the subject of several other recent studies[3, 10, 36-38]. We emphasize in particular the benefits of systematically increasing the number of PSD moments
predicted from one to three, pertaining to the improved simulation of size sorting. Milbrandt and Yau [39] showed that, for idealized hydrometeor sedimentation, a single-moment scheme is unable to simulate size sorting at all, a double-moment scheme can vastly overestimate size sorting, and a triple-moment scheme closely approximates an analytical bin solution. In this study, we show that accurately representing size sorting is very important for accurate simulation of low-level polarimetric signatures, focusing on the $Z_{DR}$ arc and low-$Z_{DR}$ hail core. Additionally, we demonstrate an improved probabilistic mesocyclone track prediction for the double- and triple-moment microphysics schemes.

This paper is organized as follows. In section 2, the methodology of the EnKF assimilation and forecast experiments and an improved version of the Jung et al. [34] polarimetric emulator are briefly described. Section 3 describes the results of the EnKF analysis and forecast experiments. Summary and conclusions are given in section 4.

2. Methodology

a. EnKF analysis and forecast experiments

All experiments use the NSSL Collaborative Model for Multiscale Atmospheric Simulation (COMMAS)[20, 40, 41] and its associated EnKF radar data assimilation system[3, 38]. The model is run using a horizontally homogeneous background environment representative of the inflow conditions during the mature phase of the 8 May 2003 tornadic supercell storm (not shown). No surface or radiation physics are included in the simulations, and the bottom and top boundaries are free slip. We use a horizontal domain size of 100 km in each direction. The grid spacings are 1 km in the horizontal and utilize a stretched vertical grid with 60 total levels with the lowest 10 levels set to 150-m grid spacing, stretched thereafter to 600 m at the model top at 26 km AGL. This storm produced a long-track F4 tornado in the city of
Moore, Oklahoma and the reader is referred to Romine et al. [37] for a thorough overview of the storm. While some previous studies of this storm assimilate both radar reflectivity (Z) and radial velocity (V_r) data from nearby S-band radars[3, 10], we choose to assimilate only V_r data from the KOUN polarimetric S-band radar. However, reflectivity data were used to define the regions for additive noise and thermal bubbles, in order to spin up the analysis and increase ensemble spread[38]. These data were first quality-controlled, dealiased, and objectively-analyzed to a regular 2 km horizontal grid, but left on the original conical radar sweep surfaces[42]. The V_r observations are then aggregated into 2-min bins and assimilated every two minutes from 2040 to 2208 UTC 8 May 2003. This period covers the time from the beginning of the first echoes associated with the storm to just before tornadogenesis[38]. Furthermore, we restrict the covariance updates of the model state variables by the assimilated V_r to only the three wind components (u,v,w), potential temperature θ, and water vapor mixing ratio q_v. The goal of this “minimalist” or “quasi-kinematic” EnKF update strategy is to allow the microphysics state variables as much freedom as possible to evolve according to the microphysics scheme itself, without any direct updating from the EnKF portion of the data assimilation system. In this regard, our study differs from the recent study of Jung et al. [6] who also verified their analyzed storm against polarimetric observations, but directly assimilated reflectivity throughout the analysis period.

A total of four assimilation experiments with 30 members each are performed, each utilizing an increasing number of predicted moments using the NSSL multi-moment microphysics scheme[20, 43], from one (1M) to two (2M and 2MSC) to three (3M). In 1M, all fixed intercept parameter values for rain, snow, and hail are set as in Dawson et al. [12] for their “LINB” experiment, with the intercept parameter for rain n_0, reduced from its default value of
8.0 \times 10^6 \text{ m}^{-4} \text{ to } 4.0 \times 10^5 \text{ m}^{-4}, to reduce the overall strength of the cold pool. A reduction in \( n_{0r} \) by a factor of 10 reduces the evaporation rate by roughly half. The single-moment graupel category has an intercept of $4.0 \times 10^5 \text{ m}^{-4}$ and particle density of 500 kg m$^{-3}$. For the double-moment configuration, two separate experiments are performed, one without correction to the sedimentation (2M), and one using the sedimentation correction method (I+II, 2MSC) of Mansell[44], which compensates for the excessive size sorting inherent in double-moment sedimentation[39, 44-46]. Finally, from the 2200 UTC analyses of each experiment, 1 hr forecasts are launched for each ensemble member, from which vorticity probabilities are computed (section 3b).

Excessive size sorting is characterized by increased reflectivity (unreasonably large particles) when the mass moment (q) falls too much faster than the number concentration (N), particularly at the downward leading edge of a precipitation shaft, and is more pronounced for small fixed values of the shape parameter [e.g., 38]. Briefly, the correction scheme of Mansell[44] creates a temporary field of reflectivity (Z) moments, and then all three moments sediment using the corresponding moment-weighted fall speeds. The sedimented mass and reflectivity moments (q' and Z') are then used to generate a number concentration (N''), which is compared to the sedimented number concentration (N'). The final number concentration is the maximum of N' and N'' at each point. The result is to artificially increase N to prevent Z from increasing [43]. In a 3-moment scheme, however, the size distribution shape parameter increases automatically in response to size sorting, which decreases the differences between the moment-weighted fall speeds and thus decelerates further size sorting.
b. Microphysics and Polarimetric radar emulator

We use an updated version of the multi-moment microphysics scheme described in Mansell et al. [20] and originally based on Ziegler [43]. The updated scheme predicts up to three moments (mixing ratio, total number concentration, and radar reflectivity factor – proportional to the 3rd, 0th, and 6th moments) of the assumed three-parameter gamma size distribution[47] for the hydrometeor categories of rain, graupel and hail, and up to two moments for cloud, cloud ice, and snow. The 6th moment closure mainly follows Milbrandt and Yau [23] and is described in detail in Dawson et al. [48]. Additionally, as described in Mansell et al. [20] and Mansell and Ziegler [49], the bulk densities of graupel and hail are predicted for the double- and triple-moment versions of the scheme, but are held fixed at 500 kg m\(^{-3}\) and 900 kg m\(^{-3}\), respectively, for the single-moment scheme.

To derive polarimetric variables from the predicted model microphysics state variables, we utilize an updated version of the polarimetric emulator described in Jung et al. [34]. The emulator makes use of the T-matrix method[50-52] to compute scattering amplitudes across size bins for each precipitating hydrometeor category, where the size bins have been discretized from the model predicted PSDs at each grid point. From these scattering amplitudes, the standard polarimetric variables of reflectivity at horizontal and vertical polarization \(Z_H, Z_V\), differential reflectivity \(Z_{DR}\), specific differential phase \(K_{DP}\), and cross-correlation coefficient \(\rho_{HV}\), are computed.

The updates primarily apply to how the water fraction on wet graupel and hail is diagnosed. Like Jung et al. [34], we diagnose a water fraction at each grid point where both ice and rain are present, where the ice can be any or all of the snow, graupel, and hail categories, by borrowing a portion of the available water from the rain field. However, whereas Jung et al. [34]
assumed that the computed water fraction was applied evenly across the sizes in a given graupel or hail distribution, we make use of the empirical relationship of critical water fraction derived by Rasmussen et al. [53], which allows for a varying water fraction across the size distribution (possibly leading to completely melted hail or graupel that is smaller than 8 mm). This updated method provides more realistic polarimetric signatures in regions where partially melted hail and/or graupel are present. A complete description of the diagnostic water fraction technique is provided in Dawson et al. [48].

3. Results

a. Low-level polarimetric signatures

From each of the four assimilation experiments, we choose the final analysis time (2208 UTC) and compute the polarimetric variables for the prior analysis (i.e. the 2-min mean forecast from the previous EnKF analysis update) using the emulator. The variables are then vertically interpolated to the 2.5° elevation sweep surface corresponding to the KOUN radar, to facilitate direct comparison with the objectively-analyzed observations (Fig. 1). For brevity, we focus only on the \( Z \), \( Z_{DR} \) and \( \rho_{HV} \) fields in this paper. The observed \( Z_{DR} \) signatures (Fig. 1b) feature a well-defined \( Z_{DR} \) “arc” near the southeast edge of the forward flank and a low-\( Z_{DR} \) (and high \( Z \)) hail core between the hook echo region and the \( Z_{DR} \) arc [35, 37]. The \( Z_{DR} \) arc has been hypothesized to be a result of size sorting of large rain drops toward the upwind (in a storm-relative sense) side of the forward flank in the presence of low-level environmental wind shear by Kumjian [54]. In \( \rho_{HV} \), the region of hail near the core is highlighted by relatively low values (~0.90-0.96, Fig. 2a), commensurate with the presence of mixed rain and tumbling hail.

Substantial differences in the simulated polarimetric fields are seen across each of the four experiments (Fig. 1c-j,2b-e). In general, the triple-moment experiment 3M performs the
best of all in qualitatively reproducing the $Z_{DR}$ arc signature and does reasonably well for the $\rho_{HV}$ pattern (Fig. 1j,2e), though it tends to over-predict both the magnitude and coverage of high reflectivity in the core (Fig. 1i). In contrast, the single-moment experiment 1M performs poorly in $Z_{DR}$ (Fig. 1d,2b) and $\rho_{HV}$, as does the double-moment without sedimentation correction (2M, Fig. 1f,2c), but on the other hand has better $Z$ in the hook region of the storm. In 1M, the forward flank region is oriented differently than the observations (from southwest to northeast, Fig. 1c) and the $Z_{DR}$ arc and low-$Z_{DR}$ and $\rho_{HV}$ hail core are entirely absent (Fig. 1d,2b). The pattern of mean volume diameter $D_m$\textsuperscript{1} of rain and hail\textsuperscript{2} for 1M (Fig. 3a,b) mirrors the $Z_{DR}$ pattern; $D_{mr}$ and $D_{mh}$ are both relatively small across the storm (1 mm or less). In 2M, the $Z_{DR}$ arc is almost completely absent due to very large predicted mean hail diameters (~20 mm) in the forward flank region where the $Z_{DR}$ arc would otherwise be (Fig. 2d), despite the presence of large ($D_{mr}$ ~ 6 mm) rain here (Fig. 2c). Additionally, the reflectivity values in the core are much too high (>70 dBZ, Fig. 1e), and the $\rho_{HV}$ magnitudes much too low (~0.90-0.96) over a broad region in the forward flank (Fig. 2c). The double-moment experiment with sedimentation correction (2MSC, Fig. 1g-h) is intermediate between 2M (Fig. 1,3e-f, Fig. 2d) and 3M (Fig. 1,3i-j, Fig. 2e), and in fact 2MSC appears to perform slightly better than 3M in the magnitude and coverage of the low-$\rho_{HV}$ region (compare Fig. 2d,e with Fig. 2a), although this may come at the expense of a muted $Z_{DR}$ arc.

\textsuperscript{1}The mean volume diameter $D_m$ is the diameter of the (spherical) particle with a volume equal to the mean volume of the distribution.

\textsuperscript{2}As stated previously, separate graupel and hail categories are predicted in the double- and triple-moment versions of the scheme. The graupel in the low levels, however, has hail-like characteristics due to a high predicted bulk density (not shown), and so is referred to as “hail” here. The separate hail category shows similar behavior (not shown), and so is omitted for the sake of brevity.
The patterns of mean volume diameter of both rain and hail in the double- and triple-moment experiments (Fig. 3c-h) reflect the action of size sorting in these experiments, with the largest hail and rain falling out closer to the updraft and south edge of the forward flank in each case[54, 55], while the smaller rain and hail are advected further downstream in a storm-relative sense (i.e., to the north and northeast). The somewhat smaller $D_{mr}$ in the forward flank in 2MSC relative to 3M (compare Fig. 3e,g) may be a result of the correction scheme increasing total number concentration $N_{tr}$ to prevent spurious growth in $Z$ via size sorting, whereas in 3M, the shape parameter of the gamma distribution is also allowed to increase, allowing larger $D_{mr}$ for the same $Z$[44]. For rain in 3M (Fig. 3g), this leads to larger $Z_{DR}$ values in the $Z_{DR}$ arc region than elsewhere in the storm, due to the well-known increase in $Z_{DR}$ with increasing drop oblateness, and the fact that the hail sizes in this region are relatively small ($D_{mh} < 10$ mm) and relatively wet (Fig. 4). The region of larger hail sizes in 3M (Fig. 3h) corresponds well with the low-$Z_{DR}$ hail core, as expected, due to the assumed tumbling characteristics for large hail in the emulator, which, combined with less diagnosed water fraction (Fig. 4) tends to drive $Z_{DR}$ towards lower values [34].

These results demonstrate that a triple-moment bulk microphysics scheme, in which hydrometeor size sorting is “fully” parameterized (i.e. all three parameters of the assumed gamma size distribution can vary independently), performs well in reproducing the salient polarimetric signatures in an analyzed supercell thunderstorm. In contrast, microphysics schemes that do not allow size sorting (single-moment schemes), or parameterize it only partially (double-moment schemes without some sort of correction mechanism to take into account the effects of the narrowing of the distribution during size sorting) exhibit degraded polarimetric signatures as compared with observations [39, 44, 45]. The seemingly better $Z$ in the hook region
in the 1M case is not typical of pure simulation tests, and the 15 min forecast (below) shows a broadened reflectivity region more similar to the other cases. The 1M ensemble members produce much less hail and have graupel that tends to be smaller and lower density than the multimoment cases, both of which contribute to the lower reflectivity in the hook region. These results are consistent with Dawson et al. [48], who similarly investigated the impact of size sorting on the low-level polarimetric signatures in idealized simulations of a different supercell thunderstorm; they also found that the triple-moment scheme, by virtue of its complete parameterization of size sorting, performed better than the single-moment scheme as compared with observations.

b. Forecast probabilistic vorticity swaths

To further evaluate the impact of the microphysics scheme complexity, we investigate four ensemble forecast experiments, each of which is launched from the 2200 UTC prior ensemble state of the above four EnKF experiments, and run out to 2300 UTC. Stensrud and Gao [56], Dawson et al. [8] and Yussouf et al. [10] made use of probabilistic vorticity swaths to evaluate the ability of their ensemble supercell forecasts to predict the overall track of the mesocyclone. We follow the approach of these studies and compute the ensemble probability of vertical vorticity $\zeta$ at the surface and 1 km AGL exceeding a given threshold (0.01 s$^{-1}$) for each of the four experiments, at any time over the 1 hr forecast period. The procedure is on a grid point-by-grid point basis for each level. For example, at a given grid point, if 3 out of the 30 ensemble members exceed the given threshold of $\zeta$, the ensemble probability at that point is 10%. In this manner, a probabilistic vorticity “swath” is derived. The 1km level is used to indicate low-level mesocyclones, whereas the surface (lowest model level) yields an indication of how well the model can maintain the enhanced vorticity imposed by the data assimilation. The results
(Fig. 5) show that, while each experiment shows an overall southward bias as compared with the observed tornado track (also seen in Yussouf et al. [10]), especially at the surface, nevertheless the double- and triple-moment experiments outperform the single-moment experiment, with the largest improvement seen for the surface mesocyclone. Investigation of the surface cold pool at the 16-min forecast time for a select member for each experiment (Fig. 6) reveals one possible reason for the differences in the mesocyclone track forecast at the surface: the cold pool strength in the double- and triple-moment schemes is somewhat larger than in the corresponding single-moment scheme. This allows stronger rotation at the surface to persist, while it is suppressed in the single-moment scheme due possibly to lack of surface convergence. These results are overall consistent with the results of Dawson et al. [8] who also found an improved track forecast for the double-moment scheme they investigated. We hasten to point out, however, that whether the single-moment scheme will exhibit stronger or weaker cold pools than the multi-moment schemes depends strongly on the choice of tunable parameters in the scheme, particularly the intercept parameter for the various hydrometeor categories [8, 14, 18]; it may be difficult a priori to choose reasonable values, whereas the multi-moment schemes offer substantially more flexibility in allowing these parameters to vary in time and space. In any case, it appears that in the current study, the cause of the stronger cold pool in the multi-moment schemes relative to that of the single-moment scheme is tied to the presence of more hydrometeor mass in the low levels (not shown) which in turn results in more evaporation and melting. Further investigation of these differences is left for future work.

---

3 The presence of high probabilities of significant rotation at the surface is used as a proxy for the tornado track, since grid spacings of 1 km are far too coarse to explicitly resolve the tornadic circulations themselves [8, 10].
4. Conclusions

The goals of this study were to investigate the impact of varying the number of moments predicted in a multi-moment bulk microphysics scheme on EnKF analyses and ensemble forecasts of a tornadic supercell thunderstorm. We focused on two main aspects: 1) the ability of the microphysics to qualitatively reproduce the low-level polarimetric (specifically $Z$, $Z_{DR}$, and $\rho_{HV}$) signatures seen in the observed storm, and 2) the impact on probabilistic vorticity swath forecasts (as a proxy for the tornado track). For the former aspect, the triple-moment scheme, at least in part by virtue of its complete parameterization of size sorting[48], was able to reproduce well the overall placement and magnitudes of the $Z_{DR}$ arc and low-$Z_{DR}$ and $\rho_{HV}$ hail core. For the latter aspect, both the double- and triple-moment experiments performed much better than the single-moment scheme in predicting the overall track of the surface mesocyclone out to 1 hour, as compared with the observed tornado track. Though all experiments had a south-of-track bias, the single-moment scheme failed to maintain significant rotation near the surface during the duration of the forecast, while both the double- and triple-moment schemes maintained rotation. The weaker cold pool in the single-moment experiment likely led to the suppression of significant near-ground rotation due to weaker surface convergence. We stress, however, that this reported sensitivity is not general to single-moment schemes, but rather depends substantially on the \textit{a priori} choices of fixed parameters in such schemes (typically the intercept parameters of the assumed exponential size distributions).

Our overall results indicate that multi-moment schemes, in broad agreement with several recent studies, generally outperform their single-moment counterparts when various methods of analysis and forecast performance are assessed. In the current study, the simulated polarimetric signatures and ensemble vorticity forecasts were investigated. For future work, a more robust
examination of the impact of microphysics might be undertaken across multiple cases. Additionally, sensitivity to various microphysical processes that become important in a multi-moment context, such as the rain drop breakup parameterization[17] is needed. The results shown here and in related work demonstrates the large sensitivity of forecast output to the choice of microphysical parameters and/or parameterization that will need to be addressed prior to any operational implementation to the Warn-on-Forecast paradigm[57].

**Acknowledgement:** This work was primarily supported by the National Research Council and National Science Foundation Postdoctoral Fellowships (AGS-1137702) awarded to the first author, the NOAA National Severe Storms Laboratory, and was partially supported by NOAA/OAR under NOAA-University of Oklahoma Cooperative Agreement NA11OAR4320072, NSF grant AGS-1046171, and by a research grant from the Weather Radar Center, Korea Meteorological Administration. Graphics were generated using Matplotlib [58].
References


List of figures

Fig. 1. Radar reflectivity $Z$ (left column, color fill, dBZ) and differential reflectivity $Z_{DR}$ (right column, color fill, dB) for (a-b) objective analysis of the 2.5 degree elevation scan from KOUN at approximately 2208 UTC 8 May 2003, (c-d) EnKF prior analysis mean for experiment 1M, (e-f) 2M, (g-h) 2MSC, and (i-j) 3M. For reference in the right column, $Z$ is overlaid in black contours in 20 dBZ increments, starting at 10 dBZ. ........................................................................................................... 22

Fig. 2. As in Fig. 1 but for cross-correlation coefficient $\rho_{HV}$. .......................................... 23

Fig. 3. As in Fig. 1, but for (left column) rain mean volume diameter $D_{mr}$ (color fill, mm), and (right column), hail mean volume diameter $D_{mh}$ (color fill, mm) for (a-b) 1M, (c-d) 2M, (e-f) 2MSC, and (g-h) 3M. In each panel, reflectivity $Z$ is overlaid in 20 dBZ increments, starting at 10 dBZ (thin black contours). Also shown for reference are the 1.5 dB (bold black dashed) and 4.5 dB (bold black solid) $Z_{DR}$ contours, denoting the low-$Z_{DR}$ hail core and $Z_{DR}$ arc, respectively. ....................................................................................................................................................... 25

Fig. 4. As in Fig. 3, but for the diagnosed water fraction on hail (color fill) for 3M. ..... 26

Fig. 5. Ensemble probability of vertical vorticity $\zeta$ exceeding 0.01 s$^{-1}$ (color fill) for the forecast period 2200-2300 UTC at (left column) the surface and (right column) 1 km AGL for each of the four ensemble forecast experiments: (a-b) 1M, (c-d) 2M, (e-f) 2MSC, and (g-h) 3M. Also overlaid in each panel is the observed 8 May 2003 F4 tornado track (bold line) and Oklahoma county boundaries (thin gray lines). .......................................................................................................................... 27

Fig. 6. Surface potential temperature perturbation (defined relative to the horizontally-homogeneous base state) $\theta'$ (color fill, K), radar reflectivity (thick black contours, 5 dBZ increment), surface and 1 km AGL vertical vorticity $\zeta$ (red and green contours, respectively, 0.00375 s$^{-1}$ increment, starting at 0.00375 s$^{-1}$), and storm-relative wind vectors plotted every 3
km (key in lower right) for ensemble member 13 at 2216 UTC (960 s forecast) for (a) 1M, (b) 2M, (c) 2MSC, and (d) 3M.
Fig. 1. Radar reflectivity $Z$ (left column, color fill, dBZ) and differential reflectivity $Z_{DR}$ (right column, color fill, dB) for (a-b) objective analysis of the 2.5 degree elevation scan from KOUN at approximately 2208 UTC 8 May 2003, (c-d) EnKF prior analysis mean for experiment 1M, (e-f) 2M, (g-h) 2MSC, and (i-j) 3M. For reference in the right column, $Z$ is overlaid in black contours in 20 dBZ increments, starting at 10 dBZ.
Fig. 2. As in Fig. 1 but for cross-correlation coefficient $\rho_{HV}$. 
Fig. 3. As in Fig. 1, but for (left column) rain mean volume diameter $D_{mr}$ (color fill, mm), and (right column), hail mean volume diameter $D_{mh}$ (color fill, mm) for (a-b) 1M, (c-d) 2M, (e-f) 2MSC, and (g-h) 3M. In each panel, reflectivity $Z$ is overlaid in 20 dBZ increments, starting at 10 dBZ (thin black contours). Also shown for reference are the 1.5 dB (bold black dashed) and 4.5 dB (bold black solid) $Z_{DR}$ contours, denoting the low-$Z_{DR}$ hail core and $Z_{DR}$ arc, respectively.
Fig. 4. As in Fig. 3, but for the diagnosed water fraction on hail (color fill) for 3M.
Fig. 5. Ensemble probability of vertical vorticity $\zeta$ exceeding 0.01 s$^{-1}$ (color fill) for the forecast period 2200-2300 UTC at (left column) the surface and (right column) 1 km AGL for each of the four ensemble forecast experiments: (a-b) 1M, (c-d) 2M, (e-f)
2MSC, and (g-h) 3M. Also overlaid in each panel is the observed 8 May 2003 F4 tornado track (bold line) and Oklahoma county boundaries (thin gray lines).
Fig. 6. Surface potential temperature perturbation (defined relative to the horizontally-homogeneous base state) $\theta'$ (color fill, K), radar reflectivity (thick black contours, 5 dBZ increment), surface and 1 km AGL vertical vorticity $\zeta$ (red and green contours, respectively, 0.00375 s$^{-1}$ increment, starting at 0.00375 s$^{-1}$), and storm-relative wind vectors plotted every 3 km (key in lower right) for ensemble member 13 at 2216 UTC (960 s forecast) for (a) 1M, (b) 2M, (c) 2MSC, and (d) 3M.