Application of a WRF Mesoscale Data Assimilation System to Springtime Severe Weather Events 2007–09

DUSTAN M. WHEATLEY
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

DAVID J. STENSRUD
NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

DAVID C. DOWELL
NOAA/OAR/Earth System Research Laboratory/Global Systems Division, Boulder, Colorado

NUSRAT YUSSOUF
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 2 May 2011, in final form 10 November 2011)

ABSTRACT

An ensemble-based data assimilation system using the Weather Research and Forecasting Model (WRF) has been used to initialize forecasts of prolific severe weather events from springs 2007 to 2009. These experiments build on previous work that has shown the ability of ensemble Kalman filter (EnKF) data assimilation to produce realistic mesoscale features, such as drylines and convectively driven cold pools, which often play an important role in future convective development. For each event in this study, severe weather parameters are calculated from an experimental ensemble forecast started from EnKF analyses, and then compared to a control ensemble forecast in which no ensemble-based data assimilation is performed. Root-mean-square errors for surface observations averaged across all events are generally smaller for the experimental ensemble over the 0–6-h forecast period. At model grid points nearest to tornado reports, the ensemble-mean significant tornado parameter (STP) and the probability that STP > 1 are often greater in the experimental 0–6-h ensemble forecasts than in the control forecasts. Likewise, the probability of mesoscale convective system (MCS) maintenance probability (MMP) is often greater with the experimental ensemble at model grid points nearest to wind reports. Severe weather forecasts can be sharpened by coupling the respective severe weather parameter with the probability of measurable rainfall at model grid points. The differences between the two ensembles are found to be significant at the 95% level, suggesting that even a short period of ensemble data assimilation can yield improved forecast guidance for severe weather events.

1. Introduction

Numerical weather prediction (NWP) is an important component of operational severe convective storm forecasting. The forecaster relies heavily upon a combination of mesoscale observations and NWP products to identify the severe convective weather threat accurately (Johns and Doswell 1992; Stensrud et al. 2003). Still, the initial conditions of NWP models are often devoid of important mesoscale features, such as drylines and cold pools (Stensrud et al. 1999). This problem can degrade forecast representations of preconvective and near-storm environments, which can then produce timing and location errors in convective forecasts, particularly in the 0–6-h forecast period. The present study investigates the role of
data assimilation in improving short-term forecasts of mesoscale environments supportive of tornadoes and convectively driven high-wind events.

The ensemble Kalman filter (EnKF; Evensen 1994; Houtekamer and Mitchell 1998) is an increasingly utilized technique for data assimilation. Similar to advanced variational data assimilation techniques, the EnKF assimilates observations using flow-dependent covariances, but in this case estimated from the ensemble. The EnKF approach can be used in conjunction with NWP models, such as the Weather Research and Forecasting Model (WRF; Skamarock et al. 2007), to provide ensemble analyses and to initialize ensemble forecasts of severe weather events.

Recent studies have demonstrated the ability of the EnKF to produce realistic mesoscale structures in ensemble-mean fields. Hacker and Snyder (2005) demonstrated the potential for EnKF surface data assimilation to influence the entire planetary boundary layer, even during overnight hours. In their four-part study, Zhang et al. (2006) and Meng and Zhang (2007, 2008a,b) demonstrated in increasingly realistic environments the ability of mesoscale data assimilation to yield analyses leading to more accurate forecasts of an East Coast snowstorm event, a warm-season mesoscale convective vortex (MCV) event, and a more prolonged warm-season period (in June 2003). Similarly, Fujita et al. (2007) showed that EnKF surface data assimilation improved the location and intensity of the dryline, PBL height and structure, and rainbands that formed during simulations of two warm-season days. More recently, Stensrud et al. (2009) and Wheatley and Stensrud (2010) have shown EnKF’s ability to produce realistic mesoscale temperature and pressure patterns associated with several mesoscale convective systems (MCSs), as well as attendant circulations. Collectively, these studies show that the EnKF data assimilation technique can produce a reasonable environment within the lower troposphere. The importance of this result is underscored by the work of Thompson et al. (2002), who found that vertical shear and moisture within 1 km of the ground could differentiate between supercell thunderstorm classes.

With the exception of Fujita et al. (2007), these data assimilation studies focused mainly on the relative accuracy of EnKF mesoscale analyses. But, the EnKF approach also has potential to improve short-term mesoscale ensemble forecasts, which can assist forecasters in anticipating severe weather events. Thus, through analysis of EnKF forecasts of 24 prolific severe events from springs 2007 to 2009, the objective of this research is to document changes and any improvements in ensemble forecast products relevant to tornadoic storms and convectively driven high-wind events. Particular emphasis is given to the impact of the assimilation on multi-ingredient severe weather parameters that combine measures of buoyancy, vertical shear, and low-level moisture, namely the significant tornado parameter (STP; Thompson et al. 2002) and MCS maintenance probability (MMP; Coniglio et al. 2007).

Section 2 provides a discussion of the ensemble design and EnKF formulation employed in this study. The assimilation results are presented in section 3. Finally, in section 4, the results of this study and their implications are summarized, and directions for future research on this topic are offered.

2. Methodology
a. Ensemble design

The Advanced Research Weather Research and Forecasting Model (WRF-ARW core; Skamarock et al. 2007) is employed in this study. The model domain covers the eastern two-thirds of the contiguous United States (Fig. 1), with a horizontal gridpoint spacing of 20 km. There are 35 vertical grid levels, spaced from less than 100 m apart near the surface to over 1 km apart at the model top, the 50-hPa-pressure surface. The prognostic variables include the three wind components, perturbation potential temperature, perturbation geopotential, perturbation surface pressure of dry air, as well as water vapor and hydrometeors. Diagnostic variables of interest include the 10-m wind fields, 2-m temperature and water vapor fields, and total surface pressure, which are diagnosed by the surface and boundary layer schemes using prognostic variables on the model grid. Both prognostic and diagnostic variables (which are used in calculations from the surface and boundary layer schemes) are included in the model state vector and thus updated in the assimilation scheme described below.

Two 30-member ensemble forecasts—one control and one experimental—are created and compared to each other and observations for each of the 24 selected severe weather events. All ensemble forecasts (i.e., a control forecast and an experimental forecast) start at 1800 UTC and extend out 12 h to 0600 UTC the following day. The control ensemble (CNTRL) uses the 1800 UTC North American Mesoscale (NAM, 12 km)¹ model analysis.

¹ The NAM model is initialized with a 12-h run of the NAM data assimilation system [beginning from downscaled Global Forecast System (GFS) data], which runs a sequence of four 3-h WRF-Nonhydrostatic Mesoscale Model (NMM) forecasts and gridpoint statistical interpolation (GSI; Wu et al. 2002) analyses. The GSI analysis is a 3DVAR technique with spatially inhomogeneous and anisotropic covariances.
and forecasts, obtained from the National Oceanic and Atmospheric Administration (NOAA) National Operational Model Archive and Distribution System (NOMADS; http://nomads.ncdc.noaa.gov), for its ensemble-mean initial and boundary conditions. Random samples from a default background error covariance file estimated by the National Meteorological Center’s (NMC, now known as the National Centers for Environmental Prediction) method (Parrish and Derber 1992) are generated by the WRF data assimilation (WRFDA) software and then added to each ensemble member to account for uncertainties in the initial and boundary conditions (Torn et al. 2006). The horizontal scales of these perturbations are “mesoscale” in nature, ranging from several tens to several hundreds of kilometers. The perturbation magnitudes for the horizontal components of wind, water vapor mixing ratio, and temperature are in the range of 5–10 m s$^{-1}$, 2–4 g kg$^{-1}$, and 2–4 K, respectively. The values of initial ensemble spread are in the range of 2.00–2.50 K for temperature and dew-point (considered later in this study), and in the range of 2.00–2.50 m s$^{-1}$ for the horizontal components of wind.

The WRF physics options are varied among the ensemble members to address uncertainties in model physics. Previous experience with this ensemble-based data assimilation system as well as other studies (e.g., Stensrud et al. 2000; Fujita et al. 2007) indicates that ensemble designs that incorporate initial condition and model physics uncertainties generally outperform those designs that use only initial condition perturbations. This diversity is introduced into the physics categories as follows (see also Table 1):

- microphysics: Thompson (Thompson et al. 2004), WRF Single-Moment 5-Class Microphysics scheme (WSM5; Hong et al. 2004).

Thus, CNTRL accounts for initial condition, boundary condition, and model physics uncertainties.

The second ensemble forecast produced is an experimental ensemble (EXP) that uses the 1200 UTC NAM analysis and forecasts for the ensemble-mean initial and boundary conditions. The ensemble member initial and boundary conditions and physics are perturbed exactly as in CNTRL. It is important to note that the NAM model products used to construct EXP are 6 h older than those utilized by CNTRL. The difference is that the EXP assimilates hourly surface, rawinsonde (i.e., any special soundings launched in anticipation of a severe weather event), and aircraft observations from 1300 to 1800 UTC before producing a 12-h ensemble forecast. Thus, the EXP is used to explore the potential benefits of a short 6-h EnKF data assimilation period to short-range ensemble forecasts. During this 6-h window, the NAM data assimilation system uses more observations—tens of thousands—in its three-dimensional variational data assimilation (3DVAR; which yields the 1800 UTC NAM model products used to construct CNTRL) than are assimilated in the EnKF of the EXP—less than 10 000 observations. Thus, there is not necessarily a clear advantage of EXP over CNTRL in producing more accurate forecasts.

b. EnKF formulation

The EXP assimilates observations prior to launching an ensemble forecast at 1800 UTC. Routinely available observations (of altimeter setting, temperature, dewpoint, and horizontal wind components) from land and marine surface stations, rawinsondes, and aircraft are assimilated using an EnKF at hourly intervals for the period 1300–1800 UTC. An assimilation cycle begins with the forecast step, whereby a 1-h forecast is made from each ensemble member and then an ensemble-
mean forecast is formed by averaging all members. In the analysis step that follows, the available observations are assimilated using the parallel version of the Data Assimilation Research Testbed (DART) software (Anderson and Collins 2007; Anderson et al. 2009), which is based upon the ensemble adjustment Kalman filter described in Anderson (2001) [a variant of the ensemble square root filter (Whitaker and Hamill 2002)]. This second step produces the EnKF update of the ensemble-mean forecast, and then also updates each ensemble member from its prior forecast state. Then, as part of the forecast step of the next assimilation cycle, a 1-h forecast (until the time of the next available observations) is made from each of the updated ensemble members, and the above procedure is repeated.

A spatial localization function—the fifth-order piece-wise rational function of Gaspari and Cohn [1999, their Eq. (4.10)]—is used to lessen the (negative) impact of spurious correlations at great distances from observations. The shape of this function is Gaussian in nature, but goes to zero at separation distances greater than twice the half-radius. The horizontal half-radius is 150 km, while the vertical half-radius is 4 km.

The magnitudes of error standard deviations for wind component, temperature, and dewpoint observations are set to 2 m s$^{-1}$, 2 K, and 2 K, respectively, based on the work of Zapotocny et al. (2000). Based on earlier work (Wheatley and Stensrud 2010), the magnitude of error standard deviation for the altimeter setting is set to 1 hPa. For all observation types, data are not assimilated if the absolute difference between the model and actual terrain height exceeds 300 m.

c. Forecast diagnostics

The goal of the present study is to assess the value of ensemble data assimilation to short-term ensemble forecasts of mesoscale environments in regions where tornadoes and convectively driven high-wind events form. As such, for each event in this study, the 12-h EXP forecasts are compared to the 12-h CNTRL forecasts (in which no ensemble-based data assimilation is performed) initialized at the same time. The analysis of the next section makes use of two forecast diagnostics commonly used to identify severe weather environments in observations or model forecast fields. A short description of these diagnostics and their application to this study follows below.

The first forecast diagnostic, the significant tornado parameter (Thompson et al. 2002), helps discriminate between significantly tornadic (F2 or greater damage) and nontornadic supercell environments. The STP equation is a function of 0–6-km-vector vertical shear magnitude (SHR), 0–1-km storm-relative helicity (SREH), convective available potential energy (CAPE), convective inhibition (CIN), and lifting condensation level (LCL), and is defined as

$$\text{STP} = \frac{\text{CAPE}}{1000 \, \text{J kg}^{-1}} \times \frac{\text{SHR}}{20 \, \text{m s}^{-1}} \times \frac{\text{SREH}}{100 \, \text{m}^2 \, \text{s}^{-2}} \times \frac{(2000-\text{m LCL})}{1500 \, \text{m}} \times \frac{(150 \, \text{J kg}^{-1} + \text{CIN})}{125 \, \text{J kg}^{-1}}$$

Values of $\text{STP} > 1$ are shown by Thompson et al. (2002) to be associated with tornadic storms. The second forecast diagnostic, MCS maintenance probability (Coniglio et al. 2007), expresses the conditional probability of a long-lived forward-propagating MCS, which will often produce damaging surface winds. The MMP equation is a function of maximum deep vertical shear (maxshear), 3–8-km lapse rate (lr), CAPE, and 3–12-km mean wind (mw) speed, and is defined as

$$\text{MMP} = \frac{1}{1 + \exp\{a_0 + [a_1(\text{maxshear})] + [a_2(\text{lr})] + [a_3(\text{CAPE})] + [a_4(\text{mw})]\}}$$

for CAPE $\geq 100$ J kg$^{-1}$ and

$$\text{MMP} = 0$$

for CAPE $< 100$ J kg$^{-1}$. The regression coefficients are $a_0 = 13.0$ (dimensionless), $a_1 = -4.59 \times 10^{-2}$ m$^{-1}$ s$^{-1}$, $a_2 = -1.16$ K$^{-1}$ km, $a_3 = -6.17 \times 10^{-4}$ J$^{-1}$ kg, and $a_4 = -0.17$ m$^{-1}$ s$^{-1}$. Values of MMP greater than 70% are considered high. In both the STP and MMP equations, the kinematic and thermodynamic input variables represent various aspects of severe weather environments that can be strongly influenced by the data assimilation process.

For each of the 24 events (Table 2), tornado or wind reports are collected for the period from 1800 UTC on day 1 to 0000 UTC on day 2. For tornado (wind) events, the ensemble-mean STP (MMP) for the 0–6-h forecast period is calculated for the control and experimental ensembles by averaging all available forecasts at 2100 UTC (day 1) and 0000 UTC (day 2). Then, based on the primary mode of severe weather (i.e., tornado or wind), the relevant severe weather parameter is extracted from
Table 2. List of 24 severe thunderstorm events considered as part of this study. For event type, “T” indicates a predominantly tornadic event, while “W” indicates a predominantly convectively driven high-wind event.

<table>
<thead>
<tr>
<th>Event</th>
<th>Type</th>
<th>Event</th>
<th>Type</th>
<th>Event</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>23 Mar</td>
<td>T</td>
<td>5 Feb</td>
<td>T</td>
<td>23 Mar</td>
<td>T</td>
</tr>
<tr>
<td>28 Mar</td>
<td>T</td>
<td>24 Apr</td>
<td>T</td>
<td>8 May</td>
<td>W</td>
</tr>
<tr>
<td>21 Apr</td>
<td>T</td>
<td>1 May</td>
<td>T</td>
<td>13 May</td>
<td>T</td>
</tr>
<tr>
<td>23 Apr</td>
<td>T</td>
<td>22 May</td>
<td>T</td>
<td>5 Jun</td>
<td>T</td>
</tr>
<tr>
<td>4 May</td>
<td>T</td>
<td>23 May</td>
<td>T</td>
<td>11 Jun</td>
<td>W</td>
</tr>
<tr>
<td>5 May</td>
<td>T</td>
<td>29 May</td>
<td>T</td>
<td>12 Jun</td>
<td>W</td>
</tr>
<tr>
<td>19 Jun</td>
<td>W</td>
<td>4 Jun</td>
<td>W</td>
<td>18 Jun</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 Jun</td>
<td>T</td>
<td>24 Jun</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11 Jun</td>
<td>T</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Results

a. Overall EnKF system performance

EnKF system performance is first evaluated in an average sense by calculating domain-averaged root-mean-square (RMS) innovations (i.e., the difference between an observational value and the interpolated ensemble-mean value) for the ensemble-mean fields at each analysis and forecast time for CNTRL and EXP, and then averaging these results across events. Surface observations compose a supermajority (more than three-fourths) of the observations ingested into the ensemble, so the diagnostics associated with these fields provide a strong indication of the performance of a predominant facet of the system. Over the assimilation period (1300–1800 UTC on day 1) for surface and all other observed variables, the consistency ratios (Dowell et al. 2004) of ensemble variance to the sum of the squared ensemble RMS difference and observational error variance generally fall between 0.5 and 1.5, which suggests that the ensemble spread is sufficient (not shown). Over the forecast period (from 1800 UTC day 1 to 0600 UTC day 2), smaller RMS innovations are evident for approximately the 0–6-h forecast period in the EXP-derived temperature and dewpoint fields compared to CNTRL (Figs. 2b,c). A bootstrap technique (Efron and Tibshirani 1993) utilizing 10 000 resamples of the original dataset found that differences in RMS innovations obtained for the 0–6-h forecast period are statistically significant at the 95% level. The RMS innovation plots further suggest that the EnKF-initialized forecast fields of the altimeter setting may possess smaller errors than CNTRL beyond 12 h (Fig. 2a).

For the wind fields, ensemble forecast innovations from EXP are smaller than those from CNTRL for approximately the 0–3-h forecast period (Figs. 2d,e). The application of a bootstrap technique also found the differences in RMS innovations for altimeter setting and the wind fields obtained for the 0–11- and 0–3-h forecast periods, respectively, to be statistically significant at the 95% level. In some instances, errors associated with EnKF-initialized forecast fields grow larger than those of CNTRL, possibly owing to the older boundary conditions used in the EnKF experiments and the accumulation of model error with time. With that said, overall, the results show a benefit to short-range forecasting from ensemble data assimilation.

While informative, the above diagnostics are of limited functional use to operational severe weather forecasting. The following subsections investigate the potential benefits of data assimilation to ensemble forecast products relevant to tornadic storms and convectively driven high-wind events.

b. Short-term ensemble forecasts of STP

Ensemble-mean STP values derived from EXP are greater than those of CNTRL for a clear majority of the grid points nearest to tornado reports from all events (Fig. 3a). At model grid points where EXP is more suggestive of a significant tornadic environment, the absolute maximum difference between EXP and CNTRL is approximately 5. At the remaining model grid points (where CNTRL is more suggestive of a significant tornadic environment), the absolute maximum difference between EXP and CNTRL is smaller (approximately 2). In 4 of the 24 events, ensemble-mean STP values derived from CNTRL are greater than those of EXP at...
FIG. 2. Domain-averaged RMS differences between the ensemble mean and observations of (a) altimeter setting, (b) 2-m $T$, (c) 2-m $T_d$, (d) 10-m $u$, and (e) 10-m $v$. Values for RMS difference represent an average across all events. Solid lines represent the control ensemble, while dashed lines represent the experimental ensemble.
a majority of the model grid points nearest to tornado reports. In two of these four events, tornadic MCSs (as opposed to supercells) were responsible for the tornado reports of that day. The STP is not necessarily applicable to these cases because the thermodynamic and dynamic variables used as input in the STP equation are chosen specifically with significant tornadic supercell environments in mind.

An analysis of the probability of STP > 1 indicates that the probability values derived from EXP also are greater than those of CNTRL for a clear majority of the tornado reports from all events (Fig. 3b). The absolute maximum difference EXP − CNTRL is approximately 0.5, regardless of whether the greater probability occurs at a model grid point in EXP or CNTRL. Distinctions in the analysis occur, however, when the probability of STP > 1 is coupled with the probability of precipitation to arrive at a total probability (Fig. 3c). While the absolute maximum difference EXP − CNTRL remains unchanged at approximately 0.5, the difference changes sign from negative to positive at a number of model grid points nearest to tornado reports. In fact, for the 11 June 2008 tornado event, the difference EXP − CNTRL changes sign in favor of EXP (i.e., to positive) at nearly every model grid point near a tornado report. Furthermore, the difference EXP − CNTRL becomes more strongly positive at a number of other model grid points. At the remaining model grid points where the total probability derived from CNTRL is greater, the absolute difference EXP − CNTRL is typically small, falling within the range of 0.00–0.25. A Wilcoxon signed-rank test (Wilks 1995) was used on the paired datasets considered in Fig. 3, and in each dataset the differences were found to be significant at the 95% level.

For all events, the total number of model grid points with nonzero STP total probability in EXP is 27,328 less than the total number of grid points in CNTRL (Fig. 4a). This result suggests that the predominantly positive differences of EXP − CNTRL above are not the result of simply increasing the area of nonzero total probability in EXP. In addition, EXP produces more model grid points with total probability greater than 0.4 compared to CNTRL. These results suggest that EXP is...
focusing the area of threat compared to CNTRL. Indeed, when only considering total probability values at model grid points nearest to tornado reports, it is seen that EXP produces fewer “misses” (zero total probability counts) than CNTRL, while also tending to increase the total probabilities near tornado reports (Fig. 4b).

In the next two subsections, ensemble-mean STP and probability fields for two tornado events are examined in greater detail to better understand the above results.

1) THE 5 MAY 2007 TORNADO EVENT

The 5 May 2007 tornado outbreak occurred over the central United States (Figs. 5a–d) in association with a low pressure area (not shown) over the high plains. The outbreak was part of a larger two-day event (4–5 May), and on 5 May, it resulted in tens of tornado reports in a north–south corridor from South Dakota to western Oklahoma.

An ensemble-mean STP > 1 for CNTRL and EXP exhibits a similar spatial extent, encompassing a north–south corridor over the central United States from South Dakota to north Texas (Figs. 6a,b). The primary difference in the two forecasts lies in the relative amplitude. In CNTRL, two local maxima of STP > 5 are located over southeastern South Dakota/western Nebraska and southwestern Oklahoma, deemphasizing (in a relative sense) the portion of significant outbreak over central Kansas (Fig. 6a). While in EXP, STP > 5 can be found in a nearly continuous corridor from South Dakota to north Texas (Fig. 6b).
FIG. 5. Radar images at 2100 UTC (day 1) and 0000 UTC (day 2) of the severe weather events of interest to this study, including the (a)–(d) 5–6 May 2007 tornado event, (e),(f) 23–24 May tornado event, (g),(h) 4–5 Jun 2008 MCS event, and (i),(j) 8–9 May 2009 MCS event.
FIG. 6. For the 6-h period from 1800 UTC 5 May to 0000 UTC 6 May 2007, ensemble-mean STP derived from (a) CNTRL and (b) EXP, probability of STP > 1 derived from (c) CNTRL and (d) EXP, and the total probability of STP > 1 and precipitation from (e) CNTRL and (f) EXP. Black dots represent tornado reports from the 6-h forecast period.
When one alternatively considers the probability of STP > 1, it is found that the spatial distribution of these fields is quite similar to the ensemble-mean fields (Figs. 6c,d). But more notably, the STP probability (and ensemble-mean STP) fields outline a much larger areal threat of tornadoes than was observed for this event. The probability of STP > 1, though, can be coupled with the probability of precipitation to form a total probability field that further refines those threatened regions (Figs. 6e,f). In EXP, total probability values as high as 0.9 reside within a narrow north–south corridor from South Dakota to central Kansas that collocates well with storm reports (Fig. 6f). In CNTRL, the west–east extent of the region is greater and total probability values tend to be smaller (Fig. 6e). The latter finding is particularly evident over central Kansas, where EXP-derived total probability values exceed those of CNTRL by as much as 0.2. Also in CNTRL, a local maximum in the total probability field occurs over southwest Oklahoma, which is out of phase with ongoing convection over western Oklahoma around 0000 UTC 6 May 2007 (and a cluster of tornado reports, beginning at 0023 UTC 6 May 2007, from the same region). Several local maxima can be found across southern Oklahoma in EXP, only the westernmost of which is collocated with any tornado reports for this event.

Differences in the total probability fields of EXP and CNTRL can largely be explained through consideration of the probability of precipitation fields. In CNTRL, a broader spatial distribution of rain probabilities over the upper plains states results in the greater west–east extent of the total probability field over the same region (Fig. 7a). Furthermore, very low rain probabilities over the central plains states in CNTRL result in relatively small (as compared to EXP) total probability values over central Kansas. The rain probability fields can be similarly used to explain the positions of the local maxima in EXP over the southern half of Oklahoma (Fig. 7b).

2) The 23 May 2008 Tornado Event

On 23 May 2008, tens of tornadoes occurred over the western half of Kansas (Figs. 5e,f) in association with a slow-moving meridional trough (not shown) approaching the high plains. A smaller cluster of tornadoes also occurred over the extreme southeast corner of Wyoming.

Similar to the analysis for the 5 May 2007 tornado event, ensemble-mean STP > 1 for CNTRL and EXP exhibits a similar spatial extent, with STP values greater than 0.5 present in both ensembles along an axis extending southeastward from the Colorado/Wyoming border to south-central Kansas (Figs. 8a,b). In CNTRL, values of STP > 1 extend as far north as extreme western Nebraska, while in EXP, an area of STP > 1 extends into southeast Wyoming. It is also noted that ensemble-mean STP values calculated from EXP are greater than one over much of central and eastern Oklahoma, where convective activity was minimal.

Probabilities fall below 0.1 over much of Oklahoma when the probability of STP > 1 and probability of precipitation are coupled (Figs. 8c,d). Two local maxima can be identified in each of the total probability fields derived from CNTRL and EXP. In CNTRL, the first maximum of 0.7 is located over extreme western Nebraska, slightly out of phase with the tornado occurrence over extreme southeastern Wyoming (Fig. 8c). The second maximum of 0.8 is centered over the northwest quadrant of Kansas. Probabilities are significantly smaller over south-central Kansas, owing to the small probability of precipitation over this region (not shown). While no tornado reports originated from this region before 0000 UTC 24 May 2008, isolated supercell thunderstorms produced large hail and tornadoes across this region in the following 1–2 h. In EXP, an elongated maximum of 0.9 stretches southeastward from extreme northwest to south-central Kansas (Fig. 8d). The second maximum of 0.7 in the EXP-derived total probability field is centered over the location of the tornado occurrence over extreme southeastern Wyoming.

These two cases solidify our interpretation of the general results seen when examining all 24 cases (Figs. 3, 4). The EXP clearly focuses on the area of tornadic thunderstorm threat and slightly elevates the threat probability compared to CNTRL, and the region of threat in EXP better agrees with the location of tornado reports than the region of threat in CNTRL. Thus, the short period of ensemble data assimilation leads to improved ensemble forecasts of STP-related parameters used in forecasting tornadic thunderstorms.

c. Short-term forecasts of MCS MMP

Ensemble-mean MMP values from EXP are greater than those from CNTRL for a majority of the wind reports in four of the six high-wind days considered (Fig. 9a). At model grid points where EXP is more supportive of a strong, long-lived MCS, the absolute maximum difference between EXP and CNTRL is approximately 0.25. Similar magnitudes are seen in the reverse scenario where CNTRL produces the more favorable mesoscale environment. Even for events where EXP initially produces larger ensemble-mean MMP than CNTRL, there is a distinct trend for the difference EXP – CNTRL to reverse sign from positive to negative in the 3–6-h forecast period (not shown). Such a result was not seen in the above analysis of STP forecasts. A possible explanation for this result may lie in the fact that the input predictors for STP are weighted heavily.
toward thermodynamical measures for the lowest 1 km above ground level, whereas the input predictors for MMP are weighted heavily toward dynamical measures calculated over a significant depth of the troposphere, where observations available to the ensemble data assimilation system (and possibly the benefits of assimilation) are more sparse.

As with the STP forecasts, the MMP forecasts are improved by coupling the MMP probability with the probability of precipitation to arrive at a total probability (Fig. 9b). The absolute maximum difference between EXP and CNTRL increases to approximately 0.25–0.50, regardless of whether the greater probability occurs at a model grid point in EXP or CNTRL. Again,
as with the STP forecasts, the difference EXP – CNTRL changes sign from negative to positive at a number of model grid points nearest to wind reports. A Wilcoxon signed-rank test was used on the paired datasets considered in Fig. 9, and in each dataset the differences were found to be significant at the 95% level.

Similar to the results for STP, the total number of model grid points (for all events) with nonzero total probability in EXP is less than that of CNTRL by 19 789 (Fig. 4c). While EXP reduces the coverage of nonzero total probability, the experimental ensemble again produces more model grid points with total probability greater than 0.4. These results suggest that EXP is focusing the area of threat compared to CNTRL. Considering only total probability values at model grid points nearest to wind reports, it is seen that EXP...
produces fewer misses (zero total probability counts) than CNTRL. More wind reports fall within the total probability ranges of 0.3–0.7 and 0.9–1.0 in EXP compared to CNTRL (Fig. 4d).

In the next two subsections, ensemble-mean MMP and probability fields for two wind events are examined in greater detail to better understand the above results.

1) THE 4 JUNE 2008 MCS EVENT

A series of mesoscale convective systems moved across the lower Ohio Valley and lower mid-Atlantic states during 4 June 2008 (Figs. 5g,h). The trajectories of the MCSs followed the northern periphery of a ridge building northward from the northern Gulf of Mexico across to the southeastern states (not shown). This active severe weather pattern resulted in nearly 200 wind reports in the affected region.

In CNTRL and EXP, a broad west–east corridor of ensemble-mean MMP extends eastward across the Ohio Valley and mid-Atlantic states (Figs. 10a,b). Probability values of 0.3–0.4 across the mid-Atlantic states are commonplace in both ensembles, with slightly higher values of 0.5 present across the lower Ohio Valley in EXP. Notably, in EXP and CNTRL, the highest values of MMP are found across the central and southern plains states, and no storm reports originated from the southern plains for this event.

As with the STP forecasts, the total probability fields (which couple the MMP probability with the probability of precipitation) provide sharper overall guidance for this event (Figs. 10c,d). In EXP and CNTRL, distinct corridors of MMP can be found across the lower Ohio Valley and lower mid-Atlantic states. This corridor is more sharply defined in EXP, yet still collocates well with wind reports for this event. Furthermore, EXP-derived MMP values over the lower mid-Atlantic states are as much as 0.3 higher than those derived from CNTRL. Also, in the total probability fields for both ensembles, MMP values over the southern plains are near zero. High values remain in both ensembles over the central plains, where intense convection and numerous storm reports were observed.

2) THE 8 MAY 2009 MCS EVENT

Before the 1800–0000 UTC forecast period of interest here, the 8 May 2009 MCS event evolved from intense multicellular convection over northeastern Colorado and northwestern Kansas into a highly organized bow-echo system over southeastern Kansas (not shown). As it moved over southern Missouri, southern Illinois, and all of Kentucky, this system produced tens of tornado reports and over 150 wind reports (not shown).

During the period from 1800 UTC 8 May to 0000 UTC 9 May (Figs. 5i,j), this bow-echo system moved from southern Illinois and western Kentucky to parts of extreme eastern Kentucky. An ensemble-mean MMP of 0.6–0.8 can be found across this region in EXP and CNTRL, with the highest values found in the control ensemble (Figs. 11a,b). While the affected region is better delineated in CNTRL, forecast guidance from both CNTRL and EXP shows MMP values greater than 0.7 extending well beyond the affected region. A pronounced maximum in ensemble-mean MMP can be
found across Iowa in EXP, with values in excess of 0.9. An area of mesoscale convection moved across Iowa during the 6-h forecast period of interest, producing several wind reports over southeast Iowa. Smaller values of ensemble-mean MMP (around 0.5 or less) can be found across the same region of Iowa in CNTRL.

Coupling the MMP probability with the probability of precipitation provides the sharper forecast guidance. In both ensembles, total probability fields show a more isolated region of relatively high MMP total probability, with a local maximum of 0.9, centered over much of Kentucky (Figs. 11c,d). In CNTRL, significant total probability can be found over southern Indiana and northern Kentucky, while total probability values do not exceed 0.3 over the same regions in EXP. The latter finding is more consistent with radar imagery and wind reports for this event, as the most intense convection (dBZ $>35$) moved over the southern half of Kentucky and no wind reports originated from the northern part of the state. Also, in the total probability field derived from EXP, a pronounced local maximum can again be found across Iowa, with probability values around 0.9 or less. In CNTRL, a much smaller maximum can also be found across parts of Iowa, with probability values around 0.5 or less.

4. Summary and discussion

An ensemble-based data assimilation system using WRF-ARW has been used to provide forecasts of prolific severe weather events from springs 2007 to 2009, with the primary objective of documenting changes and any improvements in ensemble forecast products relevant to tornadic storms and convectively driven high-wind events. Routinely available observations from land and marine surface stations, rawinsondes, and aircraft are assimilated into the ensemble. For each event in this study, severe weather parameters are calculated from an experimental ensemble forecast started from EnKF analyses, and then compared to a control ensemble forecast in which no ensemble-based data assimilation is performed.

Root-mean-square innovations for surface observations (of altimeter setting, temperature, dewpoint, and horizontal wind components) averaged across all events are generally smaller for the experimental ensemble over the 0–6-h forecast period, suggesting at least in the short term that forecast products derived from the experimental ensemble can be of better use to the operational severe weather forecaster, as compared to those derived from the control ensemble. Over this period, the
ensemble-mean STP, as well as the probability of STP (>1), derived from the experimental ensemble forecasts is greater at model grid points nearest to tornado reports. With less generality, the probability of MCS maintenance probability (MMP) is greater with the experimental ensemble at model grid points nearest to wind reports.

Sharper forecast guidance is provided when the probability of STP (>1) or MMP at a model grid point is combined with the probability of measurable precipitation. Unlike ensemble-mean STP or MMP, the total probability fields generally deemphasize regions where some severe weather ingredients are present, but no organized convection materializes. For all events, the total number of model grid points with nonzero total probability is significantly smaller in EXP than CNTRL. Still, a greater number of model grid points in EXP falls within the probability range of 0.5–1.0 as compared to CNTRL. Considering only total probability values at model grid points nearest tornado or wind reports, it is seen that EXP also produces fewer missed events than CNTRL. Thus, the EXP clearly focuses the area of severe threat compared to CNTRL, slightly elevates the threat probability compared to CNTRL, and the region of threat in EXP better agrees with the location of severe reports than the region of threat in CNTRL. Thus, the short 6-h period of ensemble data assimilation leads to improved 6-h ensemble forecasts of STP-related parameters used in forecasting tornadic thunderstorms and damaging winds.

![Graphs showing comparison of CNTRL and EXP for STP and MMP probabilities](https://example.com/fig11.png)
In future studies, it is possible that short-term mesoscale ensemble forecasts from the above type of system can be improved by cycling the ensemble for one or more days prior to the first data assimilation step (as opposed to a cold start). The inclusion of other observation types, such as satellite data, might also further improve the realism of simulated mesoscale features and the larger-scale environment in which they are embedded. Expanding the number of test cases beyond the list of 24 severe thunderstorm events considered as part of this study will also help generalize the findings presented here.

Acknowledgments. The authors appreciate the help obtained from the Information Technology Services group at NSSL in maintaining the computer systems used to conduct this research. The authors would also like to thank staff from the Data Assimilation Research Section (DAReS) and the National Center for Atmospheric Research (NCAR) for their user support. Funding was provided by the NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA08OAR4320904, U.S. Department of Commerce.

REFERENCES


——, 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP meso model. NCEP Office Note 437, 61 pp.


