

## Tornado Pathlength Forecasts from 2010 to 2011 Using Ensemble Updraft Helicity

ADAM J. CLARK,<sup>\*,+</sup> JIDONG GAO,<sup>+</sup> PATRICK T. MARSH,<sup>\*,+,#</sup> TRAVIS SMITH,<sup>\*,+</sup> JOHN S. KAIN,<sup>+</sup>  
JAMES CORREIA JR.,<sup>\*,@</sup> MING XUE,<sup>#,&</sup> AND FANYOU KONG<sup>&</sup>

<sup>\*</sup> Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

<sup>+</sup> NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

<sup>#</sup> School of Meteorology, University of Oklahoma, Norman, Oklahoma

<sup>@</sup> NOAA/Storm Prediction Center, Norman, Oklahoma

<sup>&</sup> Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma

(Manuscript received 16 April 2012, in final form 14 December 2012)

### ABSTRACT

Examining forecasts from the Storm Scale Ensemble Forecast (SSEF) system run by the Center for Analysis and Prediction of Storms for the 2010 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment, recent research diagnosed a strong relationship between the cumulative pathlengths of simulated rotating storms (measured using a three-dimensional object identification algorithm applied to forecast updraft helicity) and the cumulative pathlengths of tornadoes. This paper updates those results by including data from the 2011 SSEF system, and illustrates forecast examples from three major 2011 tornado outbreaks—16 and 27 April, and 24 May—as well as two forecast failure cases from June 2010. Finally, analysis updraft helicity (UH) from 27 April 2011 is computed using a three-dimensional variational data assimilation system to obtain 1.25-km grid-spacing analyses at 5-min intervals and compared to forecast UH from individual SSEF members.

### 1. Introduction

Recent National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) Spring Forecasting Experiments (SFEs; e.g., Clark et al. 2012a) have begun to address the challenge of providing useful information on explicitly forecast storm attributes for severe weather forecasting. Since 2008, one particularly useful diagnostic produced from convection-allowing models run in support of SFEs is hourly maximum updraft helicity. Updraft helicity (UH), a diagnostic designed for tracking rotation in simulated storms, is computed by taking the integral of the vertical vorticity times the updraft velocity between 2 and 5 km AGL (e.g., Kain et al. 2008). The term “hourly maximum” refers to the maximum value of UH at each grid point at any model time step within the previous hour

[Kain et al. (2010); hereafter UH refers to the hourly maximum quantity].

To test the utility of UH for predicting tornado outbreak severity, Clark et al. (2012b) hypothesized that the cumulative pathlengths of simulated mesocyclones diagnosed using UH could serve as a proxy for cumulative pathlengths of observed tornadoes. This hypothesis was tested by computing Pearson correlation coefficients between observed cumulative tornado pathlengths and simulated cumulative UH pathlengths over 18- and 24-h periods. Simulated UH pathlengths were measured using a three-dimensional (in space and time) object identification algorithm that treats contiguous swaths of UH (which can span multiple forecast hours) exceeding a specified threshold as objects. Simulated UH pathlengths were computed from a 4-km grid-spacing version of the Weather Research and Forecasting Model (WRF; Skamarock et al. 2008) run in real time at the National Severe Storms Laboratory (NSSL) for the period March–August 2008–10, and ensemble UH pathlengths were computed from the Storm Scale Ensemble Forecast (SSEF) system run by the Center for Analysis and

Corresponding author address: Adam J. Clark, National Weather Center, NSSL/FRDD, 120 David L. Boren Blvd., Norman, OK 73072.

E-mail: adam.clark@noaa.gov

Prediction of Storms (CAPS) in support of the 2010 SFE for the period April–June 2010. Clark et al. (2012b) found that to obtain strong UH–tornado cumulative pathlength correlations in both forecast systems, it was necessary to account for whether the environments in which the simulated UH tracks occurred were favorable for tornadoes. When the lengths from UH tracks in environments unfavorable for tornadoes were subtracted from the cumulative pathlengths, the correlation with cumulative tornado pathlengths improved significantly. Unfavorable environments were considered those in which high-based and/or elevated storms (i.e., inflow is drawn from an unstable layer above the surface) would be favored. Also, the correlations between cumulative UH and tornado pathlengths during spring (March–May) were much stronger than during summer (June–August) for the deterministic forecasts. The best results (correlation coefficient = 0.84) were obtained using the ensemble mean cumulative UH pathlengths from members of the SSEF system used in the Advanced Research core of the WRF (WRF-ARW) during the period April–May with a UH threshold of  $100 \text{ m}^2 \text{ s}^{-2}$  and the UH track segments from high-based and/or elevated storms filtered out. Including the June cases, Nonhydrostatic Mesoscale Model (NMM; Rogers et al. 2009) members, UH track segments from high-based/elevated storms, and/or only examining deterministic forecasts resulted in significantly lower correlations. Out of the six UH thresholds tested ( $25\text{--}150 \text{ m}^2 \text{ s}^{-2}$  in increments of  $25 \text{ m}^2 \text{ s}^{-2}$ ), the UH threshold of  $100 \text{ m}^2 \text{ s}^{-2}$  was found to most reliably predict cumulative tornado pathlengths (i.e., UH pathlengths using  $100 \text{ m}^2 \text{ s}^{-2}$  were closest to a one-to-one relationship with tornado pathlengths). There is not a physical reason why the threshold of  $100 \text{ m}^2 \text{ s}^{-2}$  should work best; however, based on our experience with using UH, the threshold of  $100 \text{ m}^2 \text{ s}^{-2}$  is very high and most often delineates only the most intense portion of a simulated mesocyclone. Finally, it is important to note that this method does not provide information on the precise location and/or pathlength of individual tornadoes.

The current study serves as an update to the ensemble results from Clark et al. (2012b) after adding 35 more cases from April to June of 2011 to the 34 cases from April to June 2010 examined in Clark et al. (2012b). Additionally, a simple method for deriving cumulative tornado pathlength exceedance probabilities from the SSEF system UH pathlength forecasts is presented. We also highlight forecasts from three of the major 2011 tornado outbreaks—16 and 27 April and 24 May—in which tornado pathlengths were well

predicted, as well as two other cases—1 and 17 June 2010—in which cumulative tornado pathlengths were not well predicted. Finally, UH tracks computed from storm-scale analyses are compared to forecast UH from selected SSEF system members for the 27 April tornado outbreak. The study is organized as follows. Section 2 contains model specifications and information on the 3D object identification algorithm. Section 3 contains results and section 4 contains a summary and conclusions.

## 2. Model specifications and methodology

Clark et al. (2012b) examined forecasts from a version of the WRF-ARW run in real time at NSSL, as well as ensemble ARW forecasts from the SSEF system run in support of the 2010 SFE. Herein, only SSEF system forecasts are considered. In addition to the 34 cases from April to June 2010 examined by Clark et al. (2012b), the SSEF system dataset is expanded by including 35 more cases (69 total cases) for which the SSEF system was run during April–June 2011 (16 and 27–29 April; 2–6, 9, 11–13, 16–20, 22–28, and 30–31 May; and 1–3 and 6–10 June). These were all of the available cases for which data from all the members used were available. Note that because SFE activities only occurred on weekdays, the SSEF system was not run on weekends. The subset of SSEF members examined for both years is composed of the control ARW member and ARW members with varied physics and perturbed initial and lateral boundary conditions (ICs/LBCs)—there were 10 and 18 of these members for 2010 and 2011, respectively. These members were chosen because they were the most diverse set of ARW members from each year, accounting for both model physics and analysis errors. The diversity in membership was desirable to obtain sufficient ensemble spread and reliable forecasts. Other members in which the physics schemes were varied, but the initial conditions were not varied, were not included. Also, NMM members were not included because Clark et al. (2012b) found that they degraded the UH–tornado pathlength correlations when they were included with the ARW members. Ensemble specifications for the members used in this study are provided in Table 1, and further details on all SSEF members from 2010 and 2011 can be found in Xue et al. (2010) and Kong et al. (2011), respectively. All forecasts had 4-km grid spacing, were initialized at 0000 UTC, and run over a contiguous United States (CONUS) domain (Fig. 1a). Although the 2011 simulations were integrated for 36 h, only 13–30-h forecast periods are considered

TABLE 1. SSEF system configurations for ARW members. All members listed were used during 2011 while only the first 10 members were used during 2010. For 2010 (2011), WRF version 3.1.1 (3.2.1) was used. NAMA and NAMf refer to the NAM analysis and forecast, respectively (12-km grid spacing). ARPSa refers to ARPS 3DVAR and cloud analysis, which uses NAM analyses as the background. Elements in the ICs column started by a + or – indicate SREF member perturbations added to the control member ICs. All members used Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997) shortwave radiation and Goddard (Chou and Suarez 1994) longwave radiation parameterizations. Boundary layer schemes included Mellor–Yamada–Janjić (MYJ; Mellor and Yamada 1982; Janjić 2002); Yonsei University (YSU; Noh et al. 2003); Mellor–Yamada, Nakanishi, and Niino (MYNN; Nakanishi 2000, 2001; Nakanishi and Niino 2004, 2006); quasi-normal scale elimination (QNSE; Sukoriansky et al. 2006); and the Asymmetrical Convective Model version 2 (ACM2; Pleim 2007). Microphysics schemes included Thompson et al. (2004), WRF single-moment six-class (WSM-6; Hong and Lim 2006), WRF double-moment six-class (WDM-6; Lim and Hong 2010), Ferrier et al. (2002), Milbrandt and Yau (2005), and Morrison et al. (2005). Ferrier+ refers to an updated version of Ferrier et al. (2002). Land surface models included Noah (Chen and Dudhia 2001) and Rapid Update Cycle (RUC; Smirnova et al. 1997, 2000).

Ensemble member	ICs	LBCs	Microphysics	Land surface model	Boundary layer
arw_cn	0000 UTC ARPSa	0000 UTC NAMf	Thompson	Noah	MYJ
arw_m4	+ em-p1_pert	em-p1	Morrison	RUC	YSU
arw_m5	+ em-p2_pert	em-p2	Thompson	Noah	QNSE
arw_m6	– nmm-p1_pert	nmm-p1	WSM6	RUC	QNSE
arw_m7	+ nmm-p2_pert	nmm-p2	WDM6	Noah	MYNN
arw_m8	+ rsmSAS-n1_pert	rsmSAS-n1	Ferrier	RUC	YSU
arw_m9	– etaKF-n1_pert	etaKF-n1	Ferrier	Noah	YSU
arw_m10	+ etaKF-p1_pert	etaKF-p1	WDM6	RUC	QNSE
arw_m11	– etaBMJ-n1_pert	etaBMJ-n1	WSM6	Noah	MYNN
arw_m12	+ etaBMJ-p1_pert	etaBMJ-p1	Thompson	RUC	MYNN
arw_m13	+ rsm-p1_pert	rsm-p1	MY	Noah	MYJ
arw_m14	+ em-n1_pert	em-n1	Ferrier+	Noah	YSU
arw_m15	+ em-n2_pert	em-n2	WSM6	Noah	MYNN
arw_m16	+ nmm-n1_pert	nmm-n1	Ferrier+	Noah	QNSE
arw_m17	+ nmm-n2_pert	nmm-n2	Thompson	Noah	ACM2
arw_m18	+ rsm-p2_pert	rsm_p2	WSM6	Noah	MYJ
arw_m19	+ rsm-n1_pert	rsm_n1	Milbrandt–Yau	Noah	MYJ
arw_m20	+ rsm-n2_pert	rsm_n2	MY	RUC	ACM2

herein to match the 30-h forecast length of the 2010 simulations.

For both years, ICs and LBCs (3-h updates) for the control member were from North American Mesoscale Model (NAM; Rogers et al. 2009) analyses and forecasts, respectively. Reflectivity data from up to 140 Weather Surveillance Radar-1988 Doppler (WSR-88Ds) and other traditional data, such as surface observations, rawinsondes, and wind profilers, were assimilated into the ICs of all members using the Advanced Regional Prediction System (ARPS) three-dimensional variational data assimilation (3DVAR) (Xue et al. 2003; Gao et al. 2004) data and cloud analysis (Xue et al. 2003; Hu et al. 2006) system. IC perturbations were derived from evolved (through 3 h) perturbations of 2100 UTC initialized members of the National Centers for Environmental Prediction's (NCEP) Short-Range Ensemble Forecast (SREF) system (Du et al. 2006) and added to the control member ICs. For each perturbed member, the SREF member used for the IC perturbations was also used for the LBCs.

Tornado track length data, as compiled in the National Climatic Data Center publication *Storm Data*, were obtained from the Storm Prediction Center (SPC). Known deficiencies in *Storm Data* are well documented and include inconsistencies in reporting practices and population-dependent reporting density (e.g., Weiss and Vescio 1998; Gallus et al. 2008 and references therein).

Clark et al. (2012b) provides a thorough description of the 3D object identification algorithm including limitations. Basically, to be considered an object, there must be a contiguous region in space and/or time of at least five grid points exceeding a specified UH threshold. The algorithm is applied to the raw UH field without smoothing. Herein, the UH threshold of  $100 \text{ m}^2 \text{ s}^{-2}$  is used to define objects, which was the threshold found to most reliably predict tornado pathlengths in Clark et al. (2012b). To compute the length of UH objects, each object is divided into segments according to each forecast hour the object was present. The latitude–longitude coordinates for the beginning and end points of each segment are

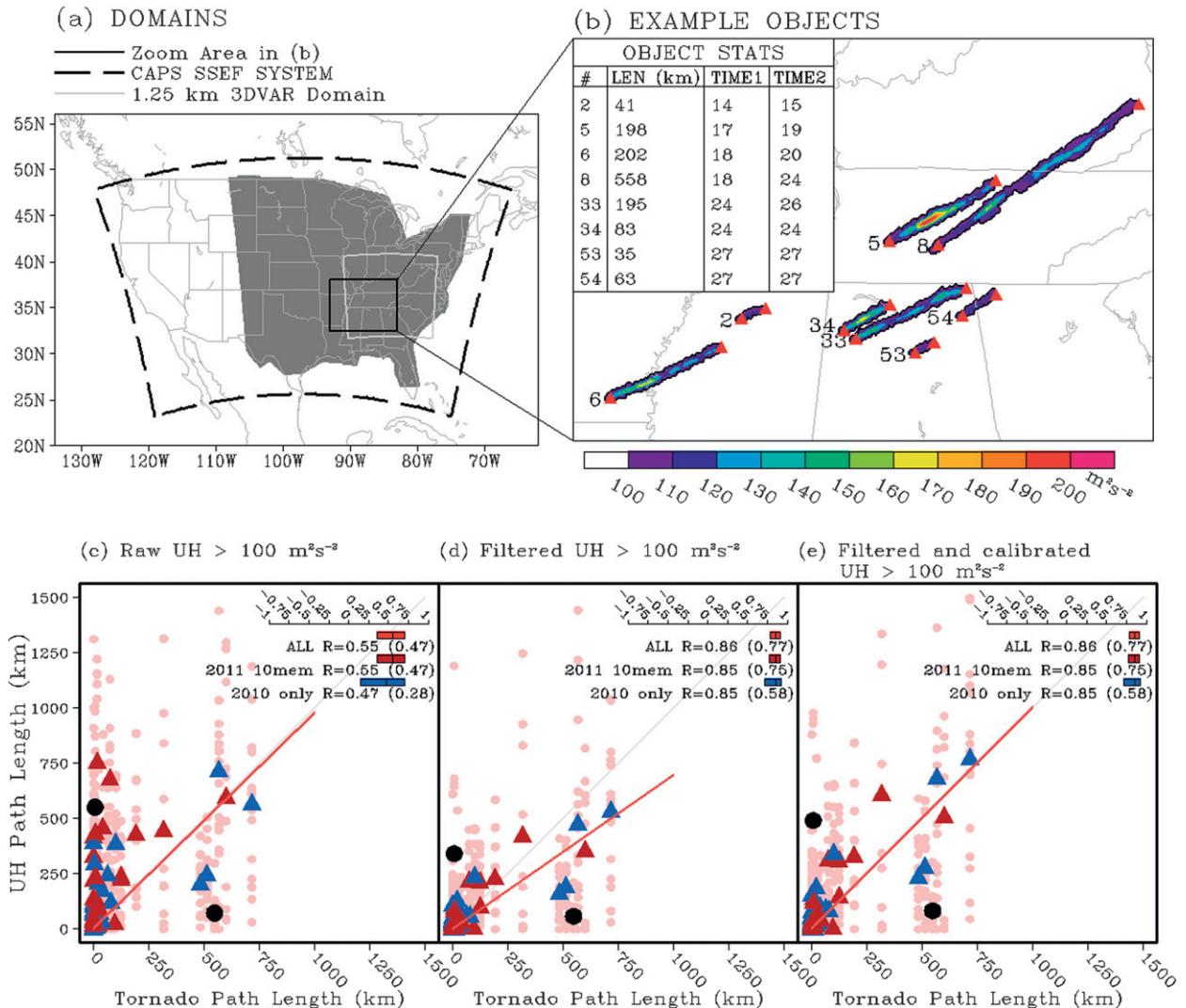


FIG. 1. (a) Model domain for the 2010 and 2011 CAPS SSEF systems, as well as the domain for the 1.25-km grid-spacing 3DVAR analyses created for 27 Apr 2011. The gray-shaded region is the mask over which tornado and UH pathlengths were computed. A legend is provided at the top. (b) Example of selected UH objects (minimum threshold of  $100 \text{ m}^2 \text{ s}^{-2}$ ) identified by the 3D-object algorithm for an SSEF system member initialized 0000 UTC 27 Apr 2011. The red triangles indicate the start and end points of each object as identified by the algorithm. The number to the left of each object denotes the order in which the object was identified by the algorithm; corresponding statistics for each object are provided in the upper left. The “LEN” column is the length between the start and end point of each object; “TIME1” and “TIME2” are the first and last forecast hours, respectively, during which the object was present. (c) Scatterplot of daily forecast UH pathlength using the threshold  $100 \text{ m}^2 \text{ s}^{-2}$  vs cumulative tornado pathlength for the 69 cases covered by SSEF system forecasts during April–June 2010–11. Points are included for the individual members (light pink dots) and mean of the members (red and blue triangles for 2010 and 2011 cases, respectively). The solid black dots are the means for 1 and 17 Jun 2010 (discussed further in the text). Corresponding correlation coefficients are shown in the top-right embedded scale along with 95% confidence intervals: “ALL” is for correlations computed from the mean of all members during 2010 and 2011, “2011 10mem” is only for the set of 10 members that were run in both years, and “2010 only” is for only the 2010 cases. The values inside (outside) parentheses include (exclude) the two cases denoted by black dots. A “best fit” to the ensemble mean points computed using least squares linear regression is shown by the dark red line. (d) As in (c), but here the portions of UH tracks associated with elevated and/or high-based storms are not considered. (e) As in (c), but all UH pathlengths are multiplied by the factor 1.44.

found by searching for the farthest west and east grid points within each segment. Then, segment lengths are summed to find the cumulative 3D object length. Figure 1b illustrates UH objects that were defined for 27 April 2011 in one SSEF system member.

### 3. Results

#### a. Pearson correlation coefficients

To examine how strongly ensemble mean cumulative UH pathlengths and cumulative tornado pathlengths

are correlated, scatterplots of daily ensemble member and ensemble mean UH pathlengths versus tornado pathlengths, along with Pearson correlation coefficients and associated 95% confidence intervals,<sup>1</sup> are shown in Figs. 1c–e. Here, the ensemble mean cumulative UH pathlength is computed by taking the average cumulative pathlength computed from all the members. Note, the scatterplots and associated correlations do not include data from 27 April 2011 (discussed later) because it was such an extreme outlier and could thus have a disproportionate impact on the correlation coefficient. For example, for 27 April the observed cumulative tornado pathlength was greater than 3 times that of the next highest case, and the mean cumulative UH pathlength was also well outside the range of the every other case. In Fig. 1c, correlations derived from raw UH pathlengths are shown, and in Figs. 1d,e, correlations derived from UH pathlengths in which UH path segments from high-based and/or elevated storms have been removed are shown. For determining whether a UH track is produced by a high-based storm, the height of the lifting condensation level (HLCL) is used. The average HLCL over all grid points within a UH object segment present during a particular hour is computed for the previous hour. If  $HLCL > 1500$  m, the UH track is designated as “high based.” To determine if a UH track is produced by an elevated storm, the ratio of surface-based convective available potential energy (SBCAPE) to most unstable convective available potential energy (MUCAPE) is used. If  $SBCAPE/MUCAPE < 0.75$ , the UH track is designated as “elevated.” For further details on discriminating UH tracks from high-based and/or elevated storms, see Clark et al. (2012b).

After filtering out the UH track segments from high-based and/or elevated storms, it was noticed that two of the ensemble mean points in Figs. 1d,e fell quite far from the best-fit line. Subjectively, these points appeared to be outliers and are indicated in Figs. 1c–e by solid black circles (note, unlike the 27 April 2011 case, the cumulative tornado and UH pathlengths for these cases were well within the range of the rest of the dataset). To quantify the impact of these two outliers, Pearson correlation coefficients were computed with and without these two cases. The correlations including the outliers are shown in parentheses in the top right of Figs. 1c–e, next to the correlations excluding the

outliers. Clearly, these two cases have a noticeable negative impact on the correlations. For the 2010-only filtered UH pathlengths (Fig. 1d), including the outliers results in  $R = 0.58$ , while removing them gives  $R = 0.85$ . Including the 2011 cases lessens the impact of the outliers, but there is still a noticeable difference—without the outliers gives  $R = 0.86$  and with the outliers gives  $R = 0.77$ . Interestingly, the outliers fall on either side of the best-fit line. The point nearest the y axis in Figs. 1d,e corresponds to 1 June 2010; for this case, relatively long tornado pathlengths were forecast and did not occur. The point nearest the x axis corresponds to 17 June 2010—for this case, only very short tornado pathlengths were forecast, but a large tornado outbreak occurred with a relatively long cumulative tornado pathlength. These two cases will be examined further in subsequent sections in an attempt to diagnose the reason for forecast failure. Because both of these cases occurred in June 2010, these two cases were likely the main contributors to the improvement in correlations that was seen after excluding the June cases from the ensemble results in Clark et al. (2012b).

Similar to the 2010-only results, the correlations with cumulative tornado track lengths using the filtered UH tracks ( $R = 0.77$  and  $R = 0.86$  with and without the outliers, respectively) are significantly higher than those using the raw UH tracks ( $R = 0.47$  and  $R = 0.55$  with and without the outliers, respectively). Also, addition of the 2011 data has little impact on the correlations relative to the 2010-only cases when the outliers are excluded ( $R = 0.86$  for all cases versus  $R = 0.85$  for 2010 only), but again, the confidence intervals shrink with the addition of more cases. For both the raw (Fig. 1c) and the filtered UH tracks (Fig. 1d), there are only very slight and statistically insignificant differences between the correlations obtained using all 18 members from 2011 versus those obtained using the 10 members from 2011 with the 2010 configurations. The slope of the best-fit line to the ensemble mean points (red lines in Figs. 1c–e) becomes slightly less than 1.0 using the filtered UH tracks (Fig. 1d). Thus, as a simple calibration to bring the slope of the best-fit line back to 1.0, all of the filtered cumulative UH pathlengths are multiplied by 1.44 (Fig. 1e).

#### b. Case studies: Forecast successes

By many metrics (e.g., cumulative tornado pathlengths, fatalities, monetary damage), 2011 had one of the most destructive tornado seasons on record (e.g., Doswell et al. 2012). In fact, using *Storm Data* tornado pathlength records from 1950 to 2011, we identified eight separate 1200–0600 UTC periods during April–May 2011 (periods chosen to match the SSEF forecasts)

<sup>1</sup> Confidence intervals for the correlation coefficients are computed using the `cor.test()` function in the R statistical software package (R Development Core Team 2012), which computes an asymptotic confidence interval based on Fisher's  $Z$  transform.

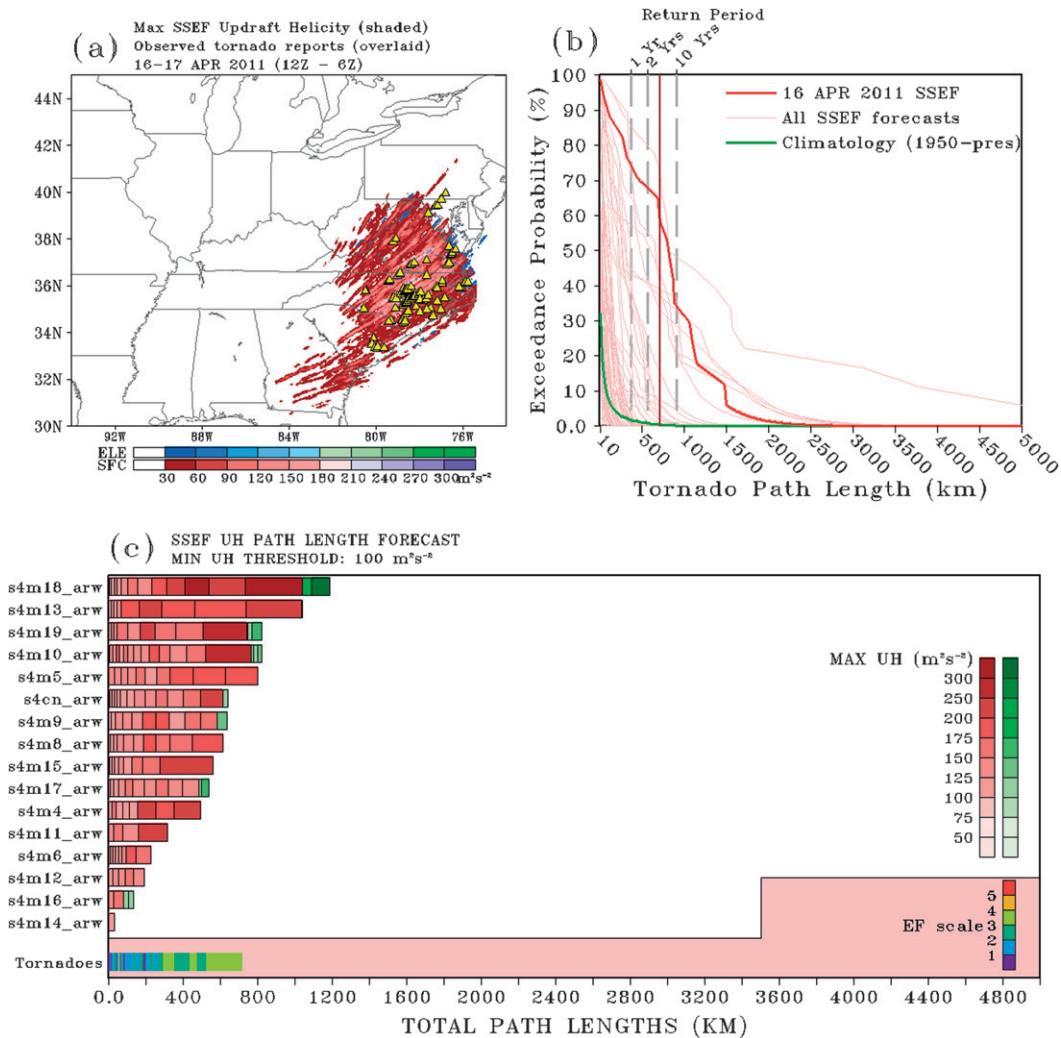


FIG. 2. (a) Maximum UH from any SSEF system member initialized 0000 UTC 16 Apr 2011 for forecast hours 13–30 (valid 1200–0600 UTC 16–17 Apr). The red/purple shading scheme is for UH produced by surface-based storms, while the blue/green shading scheme is for UH produced by elevated and/or high-based storms. Tornado reports (yellow triangles) for the corresponding period are overlaid. (b) Exceedance probabilities as a function of cumulative tornado pathlength computed from the distribution of SSEF member UH pathlength forecasts. The dark red line is for 16 Apr, the light red lines are for the 47 other cases in the dataset, and the green line is for the climatological exceedance probabilities computed from *Storm Data* for the period 1950–2011 (legend provided at top). The dark red vertical line marks the actual cumulative tornado pathlength for 1200–0600 UTC 16–17 Apr. (c) Cumulative length of UH objects for each SSEF member using a minimum threshold of  $100 \text{ m}^2 \text{ s}^{-2}$ . The length of the individual colored bars that comprise each column indicates the length of each UH object for each member. The colors within these bars indicate the maximum value of UH within the corresponding object, with red/pink shades corresponding to objects produced by surface-based storms and green shades to objects produced by elevated and/or high-based storms (color bars provided on the right side). The bars in the bottom column similarly indicate pathlengths and maximum intensities, but for observed tornadoes where maximum intensities correspond to EF ratings.

in which cumulative tornado pathlengths exceeded the 1-yr return period (in chronological order: 15, 16, 25, 26, and 27 April, and 22, 24, and 25 May). SSEF system UH forecasts for the cases on 16 and 27 April, and 24 May, are displayed in Figs. 2–4 (forecasts were also available for 22 and 25 May, but are not shown here). These figures—designed similarly to Figs. 7, 9, and 10 in Clark et al.

(2012b)—present graphical displays to efficiently provide information on key ensemble UH object attributes. Maximum UH from any ensemble member from both surface and non-surface-based storms is displayed along with the length and maximum intensity of individual objects defined using  $\text{UH} \geq 100 \text{ m}^2 \text{ s}^{-2}$  (panel c in Figs. 2–4). Additionally, a new product displays exceedance

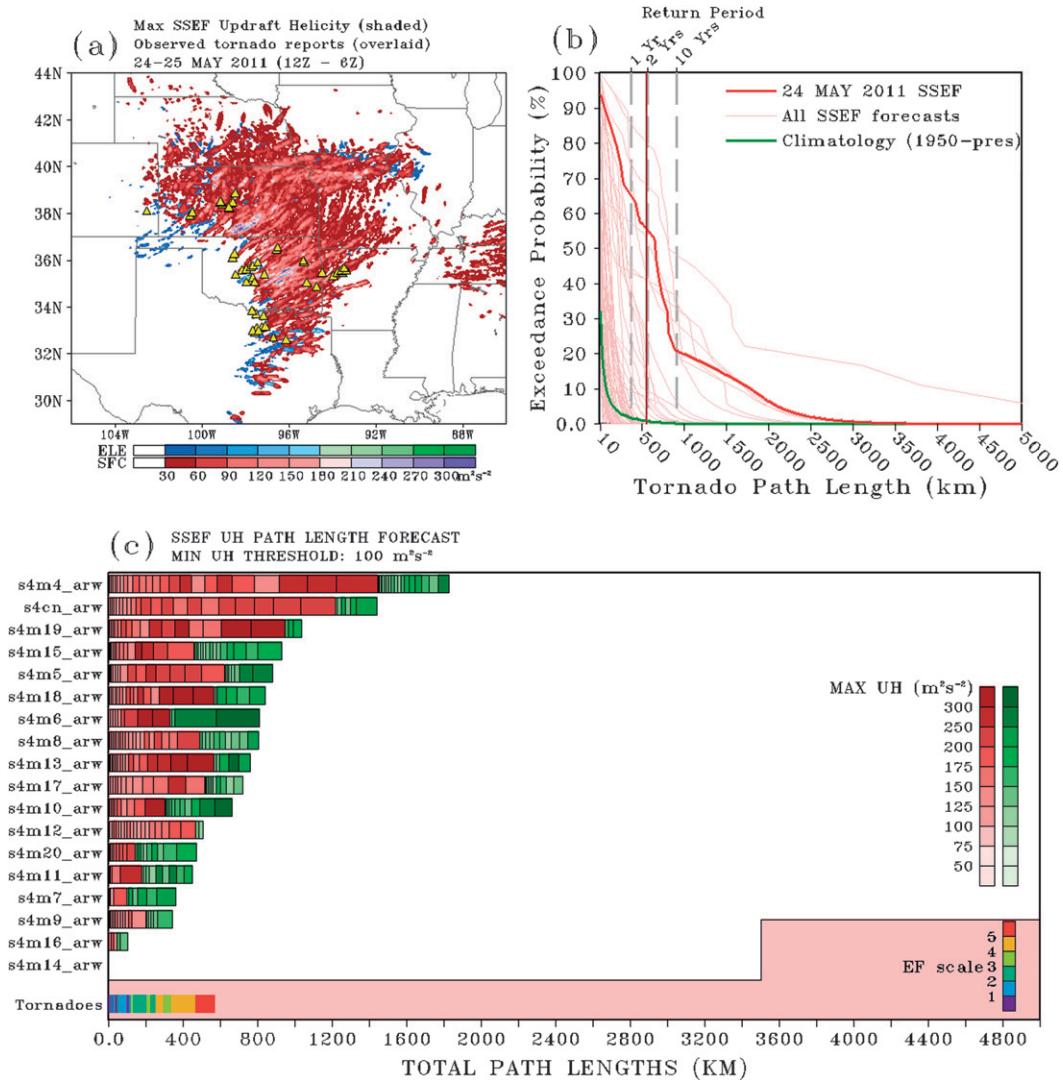


FIG. 3. As in Fig. 2, but for SSEF system forecasts initialized 0000 UTC 24 May 2011.

probabilities for cumulative tornado pathlengths 0–5000 km (panel b in Figs. 2–4). The probabilities are derived from the filtered cumulative UH pathlengths multiplied by 1.44 for each ensemble member (i.e., the calibrated points displayed in Fig. 1e) and computed using the method of Hamill and Colucci (1997, 1998), which involves finding the location of verification thresholds (i.e., tornado pathlengths) within the distribution of ensemble member forecasts. Probabilities for cumulative pathlengths beyond the highest ensemble member forecast are obtained by assuming that the probability distribution function (PDF) in this region follows a Gumbel distribution (Wilks 1995). Exceedance probabilities and return periods based on 1950–2011 tornado pathlength climatology for 1200–0600 UTC periods during April–June are also shown in Figs. 2–4.

These climatological probabilities are based on all days during April–June for 18-h periods corresponding to the forecasts.

For the 16 April 2011 case (Fig. 2), the general region with the highest values of maximum UH from any ensemble member corresponded very well with where the observed tornadoes occurred in Virginia, North Carolina, and South Carolina. The cumulative tornado pathlength for the 1200–0600 UTC period was 716 km. From the plot of forecast UH object attributes for this case (Fig. 2c), it can be seen that almost all the strong UH tracks were being produced by surface-based storms. In addition, the majority of members had cumulative UH pathlengths greater than 400 km, which translated into very high probabilities—greater than 70%—for a cumulative tornado pathlength that exceeded the 1-yr return

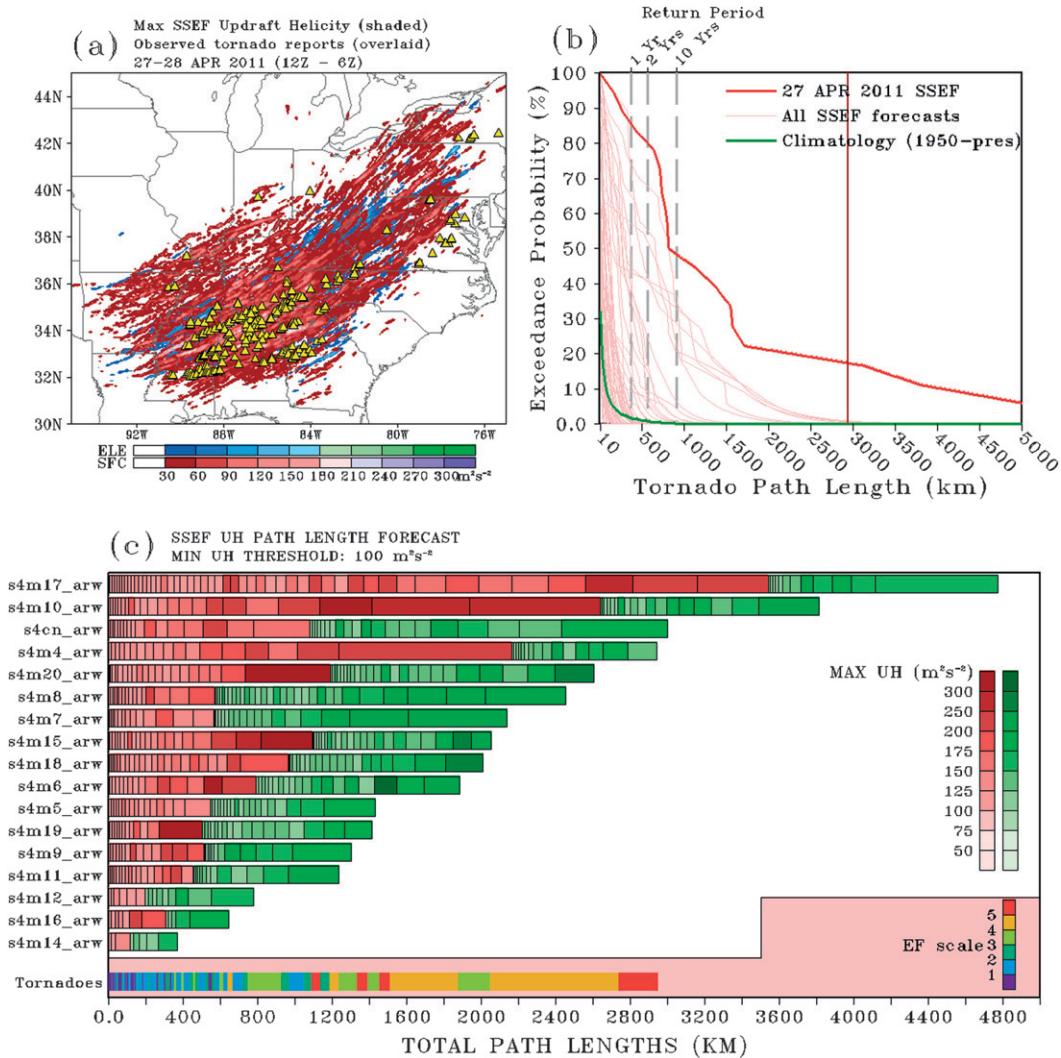


FIG. 4. As in Fig. 2, but for SSEF system forecasts initialized 0000 UTC 27 Apr 2011.

period. The actual cumulative tornado pathlength (indicated by the dark red vertical line in Fig. 2) for this case fell between the 2- and 10-yr return periods and intersected the SSEF probabilities at about 60%.

The 24 May tornado outbreak was another case for which UH track length forecasts indicated a high probability, greater than 60%, for cumulative tornado pathlength exceeding the 1-yr return period (Fig. 3b). Indeed, the SSEF probabilities verified very well with the observed tornado pathlength of 567 km intersecting the SSEF probabilities at about 55%. However, for this particular case, most of the tornadoes, especially those in central Oklahoma and northern Texas, were displaced west of where the highest maximum UH values were forecast by the ensemble (Fig. 3a). Closer examination of the forecasts revealed that most of the members moved a dryline too quickly east over central

Oklahoma by midafternoon, when in reality the dryline was still in far western Oklahoma. Recent work examining forecast dryline position errors has revealed that eastward biases are common in the 4-km grid-spacing WRF configuration run locally at NSSL (Coffer et al. 2013).

The most severe tornado outbreak of 2011, and one of the largest tornado outbreaks in recorded history, occurred on 27 April (e.g., Doswell et al. 2012). During the 18-h period 1200–0600 UTC 27–28 April, the cumulative tornado pathlength was 2949 km with four of the tornadoes rated as being category 5 events (EF5) on the enhanced Fujita scale. The only other comparable outbreak since 1950 was on 3–4 April 1974 and had a cumulative tornado pathlength during the corresponding 1200–0600 UTC period of 3964 km. For 27 April, the area covered by relatively high values of forecast UH from

SSEF members stretched from Mississippi to Pennsylvania (Fig. 4a). The highest UH values were forecast over Mississippi, northern Alabama and Georgia, and Tennessee and matched very well with where most of the tornadoes occurred. For many of the ensemble members, close to 50% of the UH tracks were designated high based and/or elevated (Fig. 4c). These non-surface-based UH tracks mainly emanated from the first of multiple rounds of strong convection that was relatively well forecast by the ensemble members and occurred during the early morning of 27 April before the lower planetary boundary layer began to destabilize. Note, in Fig. 4a the elevated UH tracks that occurred earlier in the day are mostly covered by the surface-based ones that occurred later.

For virtually all tornado pathlengths considered, 27 April exceedance probabilities were highest out of all 69 cases in the dataset. Furthermore, for extreme pathlengths exceeding the 10-yr return period, which corresponds to just less than 1000 km, exceedance probabilities for 27 April were well beyond those of any other case (Fig. 4b). The actual cumulative tornado pathlength of 2949 km intersected the SSEF probabilities at about 19%. For all other cases, exceedance probabilities were very near 0% for this pathlength. This was a very encouraging result, indicating the SSEF system was able to recognize the relative high probability of a very rare event.

### c. Case studies: Forecast failures

For most of the 69 cases covered by the SSEF system forecasts during 2010–11, the method outlined in this study for predicting cumulative tornado pathlengths worked quite well. However, as is made clear in Fig. 1e, there were two cases in which the mean cumulative UH pathlength provided a very poor prediction of the cumulative tornado pathlength. Here, these cases are examined in more detail to find some of the potential causes for forecast failure.

For the first failure case, 1 June 2010, relatively high values of maximum UH from any ensemble member (Fig. 5a) extended from southeast South Dakota and eastern Nebraska into Iowa. The observed cumulative tornado pathlength was 6 km and these tornadoes did occur within the general region where the SSEF members were predicting higher values of UH. However, it is clear from the cumulative UH pathlengths (Fig. 5c) that cumulative tornado pathlengths were overforecast. Four out of 10 SSEF members contained cumulative UH pathlengths from surface-based storms exceeding 500 km (Fig. 5c), with the most extreme member (s4m8\_arw) containing UH pathlengths of about 1200 km. The resulting tornado pathlength exceedance probabilities (Fig. 5b) indicated a relatively high likelihood for a significant tornado outbreak, with the probability for a

once every 10-yr outbreak at about 37%. However, it is also important to note that the overall forecast distribution was bimodal, with six of the other members containing very short UH pathlengths. Thus, from a probabilistic perspective, it could also be argued that this forecast verified well because a majority of members (6 of 10) predicted short pathlengths and short pathlengths were observed.

Considering the synoptic regime in which this case occurred, as well as the forecast and observed convective mode for this case, provides some additional insight. During the day on 1 June, relatively fast mid- to upper-tropospheric zonal flow was present across the north-central United States. In SPC's 1300 UTC day-1 convective outlook it was noted that large-scale forcing for ascent would be aided over a relatively broad region centered over Iowa because this region was in the right-entrance quadrant of an upper-tropospheric jet streak and a subtle short-wave trough embedded in the zonal flow was approaching from the west. Given the degree of instability expected along with favorable vertical shear profiles for supercells, it was expected that supercell storm structures would develop initially that would likely produce large hail and isolated tornadoes. Later in the evening, it was expected that convection would organize upscale into a fast-moving mesoscale convective system, which would be associated with a damaging wind threat. Given that SPC forecasters were only expecting isolated tornadoes over a relatively small area, it seems clear that, although the synoptic regime was favorable for severe weather with isolated tornadoes, the probability of a tornado outbreak exceeding the 1-yr return period was much less than that predicted by the SSEF system cumulative UH pathlengths. Indeed, SPC's forecast verified quite well on this day. There were isolated tornadoes reported in Iowa and Nebraska during the afternoon as well as numerous large hail reports. Later that evening, a strong mesoscale convective system (MCS) formed and tracked across southern Iowa–northern Missouri into Illinois, producing a swath of high wind reports. In Fig. 6a, which displays the observed composite reflectivity valid at 1900 UTC, the scattered nature of the convection with isolated supercellular structures (e.g., in northwest Iowa) is evident, with a transition to a mature MCS 6 h later, as illustrated in Fig. 6d.

Despite the long cumulative UH pathlengths predicted by a few of the SSEF members, all of the members produced a quite skillful prediction of the convective evolution. Figure 6b shows simulated composite reflectivity valid at 1900 UTC for a representative member, s4m8\_arw. Although there are displacement errors that can be seen by comparing the forecast in Fig. 6b to observations in Fig. 6a, the orientation and scattered nature of the convection are very similar.

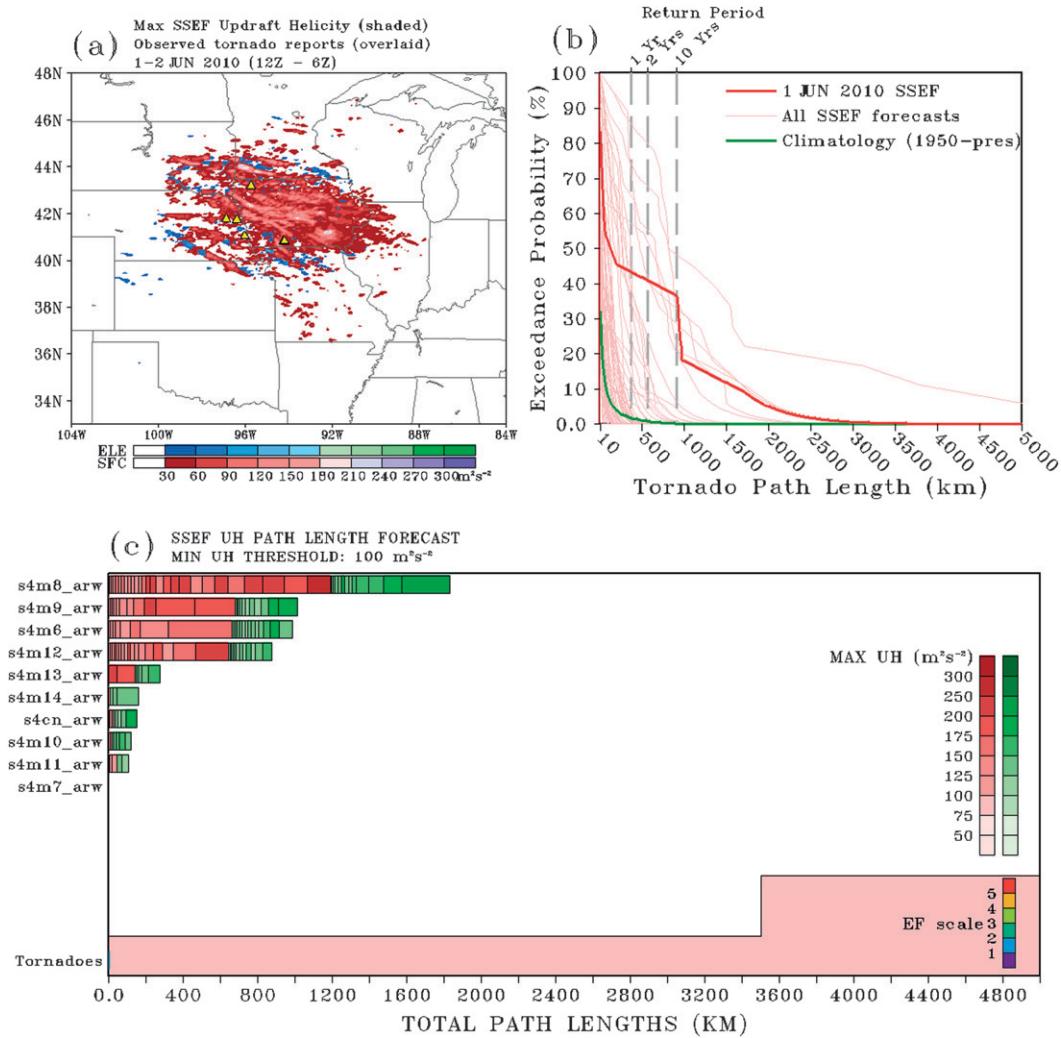


FIG. 5. As in Fig. 2, but for SSEF system forecasts initialized 0000 UTC 1 Jun 2010.

Furthermore, a couple of the forecast storms had supercellular characteristics similar to what was observed. The supercellular characteristics of one simulated storm can clearly be seen from the zoom-in over southwest Minnesota–eastern South Dakota provided in Fig. 6c. Here, the path of a simulated mesocyclone is indicated by the swath of hourly maximum UH  $> 100 m^2 s^{-2}$  (hatched area in Fig. 6c), and the location of the mesocyclone at the time the simulated composite reflectivity is valid is clear from the instantaneous UH<sup>2</sup> (contours in Fig. 6c).

<sup>2</sup> The instantaneous UH is computed at a single model output time using data valid at constant pressure levels rather than constant height as follows:  $UH = w \sum_{z_{850hPa}}^{z_{500hPa}} w \zeta \Delta z = (w \zeta_{850,700} \Delta z_{850,700}) + (w \zeta_{700,600} \Delta z_{700,600}) + (w \zeta_{600,500} \Delta z_{600,500})$ , where the overbar indicates a layer average and the subscripts indicate the bottom and top of the layer (in hPa). Also,  $w$ ,  $z$ , and  $\zeta$  are vertical motion ( $m s^{-1}$ ), height (m), and vorticity ( $s^{-1}$ ), respectively.

Later in the forecast period, all of the SSEF members developed convection upscale and predicted a MCS in Iowa or Illinois by 0300 UTC 2 June, which matched very well the observed convective evolution. In Fig. 6e, the simulated composite reflectivity from member s4m9\_arw valid at 0100 UTC 2 June is shown. Of particular relevance in Fig. 6e are the swaths of UH  $> 100 m^2 s^{-2}$  over western to central Iowa. From the zoom-in over central Iowa in Fig. 6f it is obvious that these UH swaths emanate from circulations at the leading edge of the convective line. Examination of the other ensemble members producing relatively long cumulative UH pathlengths (not shown) revealed similar behavior. In fact, it appeared that the vast majority of the UH pathlengths were associated with linear convective modes. Thus, despite the problems with using UH pathlengths to forecast tornado pathlengths for this case, the forecast convective

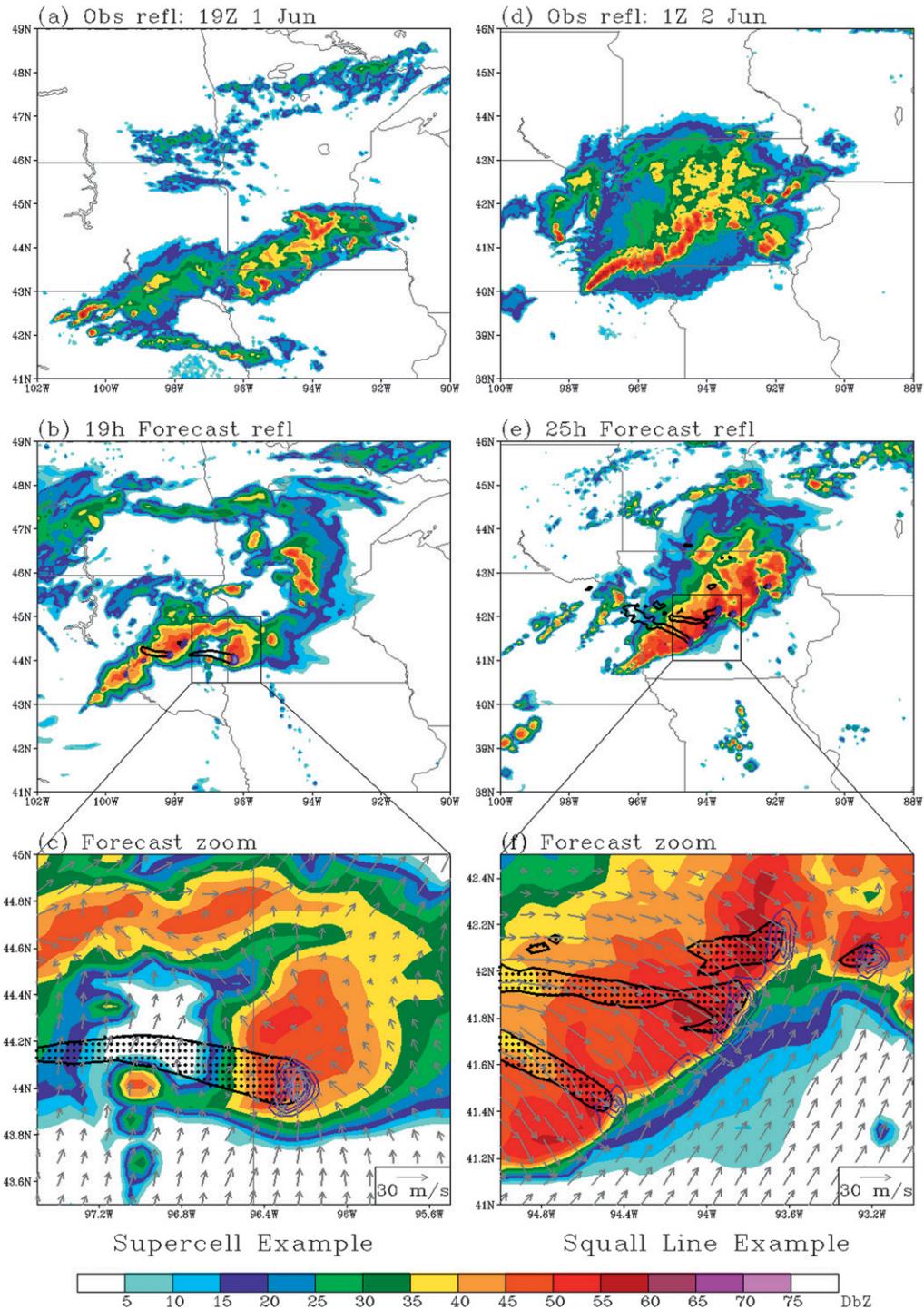


FIG. 6. (a) Observed composite reflectivity valid 1900 UTC 1 Jun 2010. (b) The 19-h forecast composite reflectivity valid 1900 UTC 1 Jun 2010 from member s4m8\_arw (see Table 1 for configuration). The small box shows the area pictured in (c) and black contours show maximum UH greater than  $100 \text{ m}^2 \text{ s}^{-2}$  from forecast hours 18 to 19. The dark to light purple contours, starting at  $50 \text{ m}^2 \text{ s}^{-2}$  and increasing in increments of  $50 \text{ m}^2 \text{ s}^{-2}$ , depict the instantaneous UH valid at 1900 UTC (for details on instantaneous UH computation, see the text). (c) As in (b), but for the zoomed-in region. Also, hatching is added to the region with maximum UH greater than  $100 \text{ m}^2 \text{ s}^{-2}$  and 850-hPa wind vectors are shown (vector scale shown at bottom right). (d)–(f) As in (a)–(c), but the observations and forecasts are valid at 0100 UTC 2 Jun.

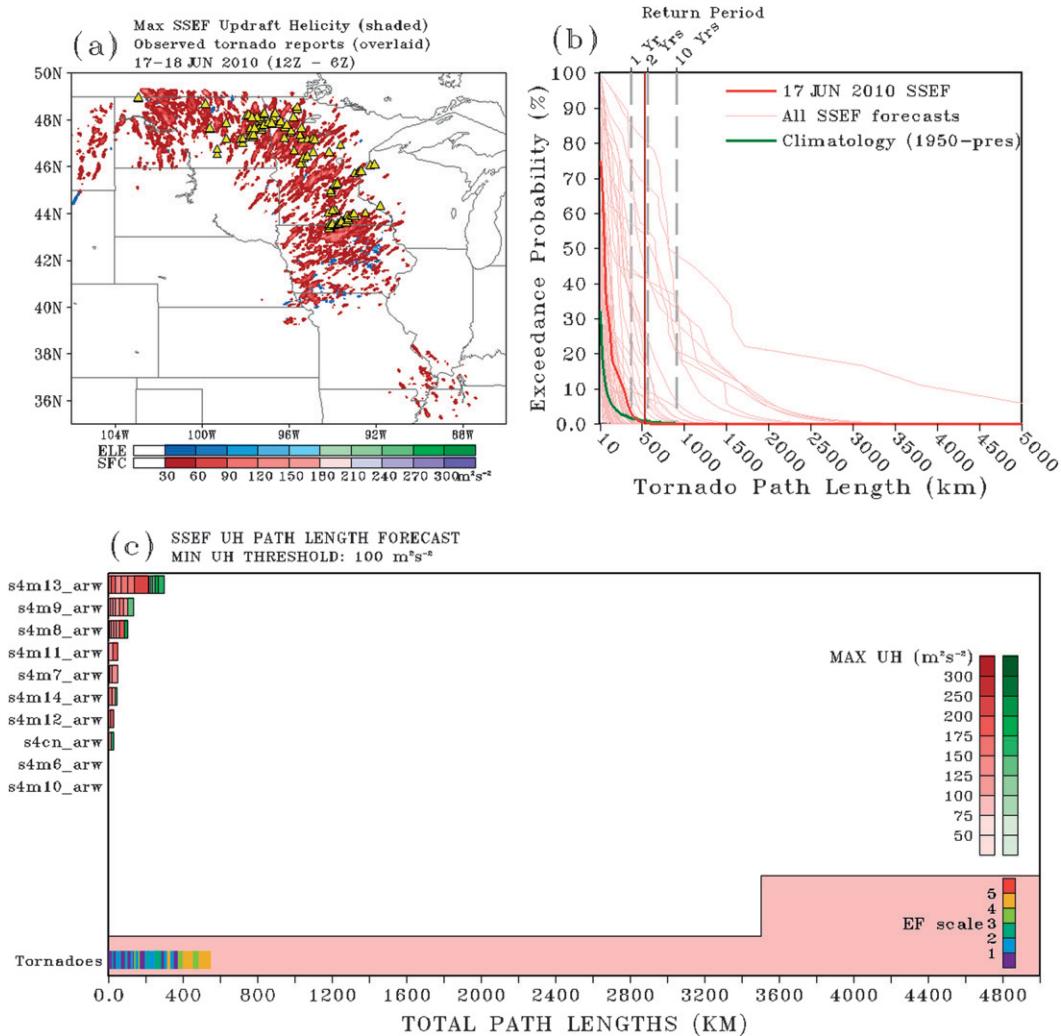


FIG. 7. As in Fig. 2, but for SSEF system forecasts initialized 0000 UTC 17 Jun 2010.

evolution from the SSEF ensemble was actually very accurate. Although tornadoes are not uncommon with linear systems, examining tornadoes that occurred from 1998 to 2000, Trapp et al. (2005b) found that linear modes only accounted for about 18% of all tornadoes, while accounting for a disproportionate number of the weak (EF0 and EF1) tornadoes. Thus, it would seem that to enhance the utility of the cumulative UH-tornado pathlength relationship for the type of case that occurred on 1 June 2010 data mining/image processing algorithms need to be developed that can distinguish the convective mode from which UH emanates.

The other forecast failure occurred on 17 June 2010. For this case, a curved swath of maximum UH from any ensemble member extended from North Dakota to Minnesota and into Iowa (Fig. 7a). However, the maximum values of UH were not particularly high

in any of the members, which resulted in relatively short cumulative UH pathlengths (Fig. 7c). The probability of exceeding a 1-yr return period outbreak on this day was less than 5%, but the observed cumulative tornado pathlength for this day was 547 km (Fig. 7b, vertical red line marks observed cumulative tornado pathlength), which came close to exceeding the threshold for a 2-yr return period outbreak.

Again, consideration of the synoptic regime as well as the forecast and observed convective mode for this case can provide some additional insight. On 17 June, a high-amplitude, negatively tilted upper-level trough and associated midtropospheric jet streak was pivoting northeastward into the north-central United States. The SPC's 1300 UTC day-1 convective outlook indicated they were expecting strong height falls across North and South Dakota to provide dynamic forcing for ascent. Relevant surface features included a surface low

pressure center moving northeast from western South Dakota to central North Dakota with an occluding frontal boundary and cold frontal boundary stretching to the south that was expected to surge eastward into South Dakota, Nebraska, Minnesota, and Iowa. Given the moderate instability that was expected to develop in the warm sector coinciding with very favorable deep and lower-layer shear profiles, the SPC was expecting widespread severe weather with the potential for tornadoes, one or two of which could be strong. In fact, storm survey teams rated four tornadoes EF3 and four EF4.

Given the favorable environmental parameters that were present for a tornado outbreak, it was puzzling why the SSEF system members did not produce higher values of UH. An examination of forecast height and wind fields (not shown) did not reveal any particularly large errors at the synoptic scale, and in our experience with using UH as a forecasting tool for tornadoes, the cases where there is an obvious large-scale signal for a potential tornado outbreak seem to have the most reliable cumulative UH pathlength forecasts. Furthermore, all of the SSEF system members developed convection that included many discrete cells within forecast environments favorable for rotating storms. One such forecast valid 2200 UTC from the control member (s4cn\_arw) is illustrated in Fig. 8b; note the discrete nature of forecast storms, but that these storms were not strongly rotating (Fig. 8c). In fact, none of the storms pictured in Fig. 8c contained  $UH > 100 \text{ m}^2 \text{ s}^{-2}$  and this lack of strong rotation was typical of the other ensemble members, as well.

It was speculated that higher horizontal resolution might be required to adequately resolve mesocyclones in this case. It is not clear why this particular case would require higher resolution while in other cases 4-km grid spacing was adequate to resolve strong mesocyclone-scale rotation, but it happens to be easily testable. During the 2010 SFE, CAPS ran a version of the WRF configured exactly as the control member of the SSEF system except with 1-km grid spacing. The 1-km forecast (Fig. 8d) predicts storms in the same general regions as the 4-km simulation (Fig. 8b), but with finer detail and structure. Also, as evident from the zoom-in over west-central Minnesota (Fig. 8e), many of the 1-km simulated storms have strong mesocyclones and well-defined supercellular structures. One storm in particular, shown in Fig. 8f, has very well-defined supercell characteristics, including a hook-echo feature and a deep rotating updraft within the southeast quadrant of the storm, as indicated by instantaneous UH exceeding  $800 \text{ m}^2 \text{ s}^{-2}$ . The storm even appears to be splitting at this time, with the “left mover” moving away to the north. Note, because the

vorticity and vertical velocity scale with model grid spacing, it is expected that the 1-km simulations will contain higher values of UH. Thus, the difference in UH magnitude alone does not necessarily imply more realism, and without a much more thorough analysis it is not known what threshold of UH from the 1-km simulations would give reliable cumulative tornado pathlength forecasts. On the other hand, the supercellular storm structures that were mainly only seen in the 1-km grid-spacing forecasts do more accurately reflect the type of storms that occurred in this case. Examining in more detail why there were such striking differences between the 1- and 4-km grid-spacing solutions is beyond the scope of this study, but clearly it is a worthwhile topic for future research.

#### *d. Storm-scale 3DVAR analyses of UH for 27 April*

To gain further insight into the 27 April tornado outbreak, a 3DVAR system (Gao et al. 2004) developed for ARPS (Xue et al. 2000, 2001, 2003) is used to construct 5-min, 1.25-km grid-spacing analyses for the period 1500–0300 UTC 27–28 April over an approximately  $1200 \text{ km} \times 1000 \text{ km}$  domain (Fig. 1a) that covered most of the outbreak. The ARPS 3DVAR system was designed especially for storm-scale data assimilation and detection of mesocyclones. The system uses a recursive filter (Purser et al. 2003a,b) with a mass continuity equation and other constraints incorporated into a cost function. Radial velocity data from nearby WSR-88D radars as well as other available surface observations are assimilated with NAM analyses used as the background, and reflectivity data are assimilated using a cloud analysis system (Hu et al. 2006). Because the 3DVAR analyses contain a three-dimensional depiction of the wind field, the same formulation used to compute UH in the SSEF system forecasts can be applied to this analysis dataset. Further details of the 3DVAR system are described in Gao et al. (2009) and references therein.

These 3DVAR analyses were generated with the same system used for the 2010 and 2011 NOAA/HWT Experimental Warning Program (EWP) Real-Time 3D Radar Data Assimilation Experiment (Smith et al. 2010; Gao et al. 2011, 2013). Also, it is important to note that the UH fields computed from the 3DVAR analyses and SSEF members are still far from an “apples to apples” comparison. Among other things, the spatial resolution, modeling systems, and sampling intervals are inconsistent. For example, UH from SSEF system forecasts is the maximum value from any model time step (typically 25 s) within a 1-h period, while UH results in the 3DVAR analyses are instantaneous values computed at 5-min increments. Nevertheless,

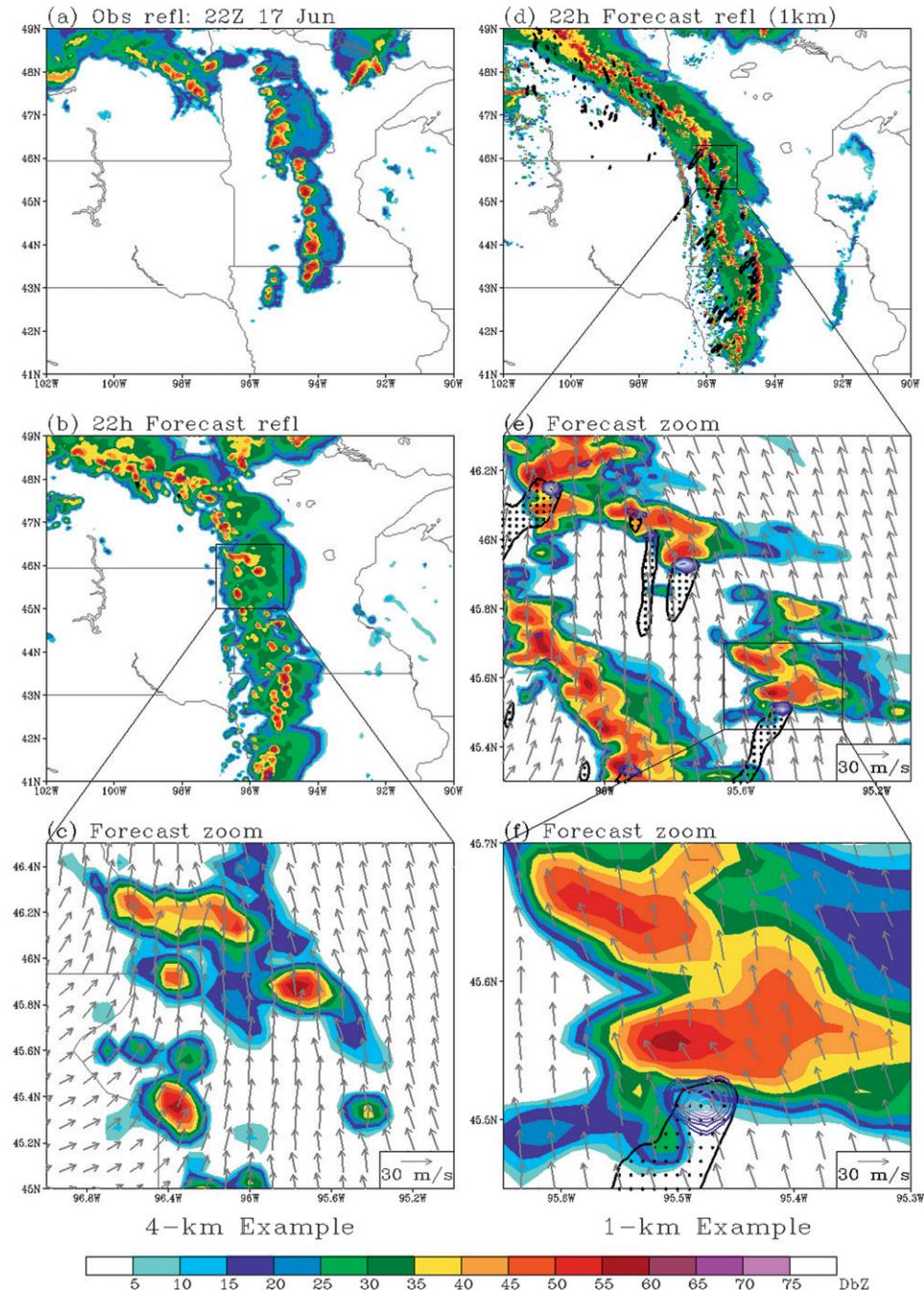


FIG. 8. (a) Observed composite reflectivity valid 2200 UTC 17 Jun 2010. (b) The 22-h forecast composite reflectivity valid 2200 UTC 17 Jun 2010 from member s4cn\_arw (see Table 1 for configuration). The small box shows the area pictured in (c) and black contours show maximum UH greater than  $100 \text{ m}^2 \text{ s}^{-2}$  from forecast hours 21 to 22. The dark to light purple contours, starting at  $50 \text{ m}^2 \text{ s}^{-2}$  and increasing in increments of  $50 \text{ m}^2 \text{ s}^{-2}$ , depict the instantaneous UH valid at 1900 UTC (for details on instantaneous UH computation, see the text). (c) As in (b), but for the zoomed-in region and the lowest model level reflectivity is pictured. Also, hatching is added to the region with maximum UH greater than  $100 \text{ m}^2 \text{ s}^{-2}$  and 850-hPa wind vectors are shown (vector scale shown at bottom right). (d) As in (b), but forecasts are from a 1-km grid-spacing configuration of member s4cn\_arw. The black contours are for UH greater than  $300 \text{ m}^2 \text{ s}^{-2}$  and the contours for instantaneous UH go from 250 to 300 and then up to  $900 \text{ m}^2 \text{ s}^{-2}$  in increments of  $100 \text{ m}^2 \text{ s}^{-2}$ . (e) As in (d), but for a zoomed-in region and the lowest model level reflectivity is pictured. The small box shows the area pictured in (f), hatching is added to the region with maximum UH greater than  $300 \text{ m}^2 \text{ s}^{-2}$ , and 850-hPa wind vectors are shown. (f) As in (e), but for a zoomed-in region.

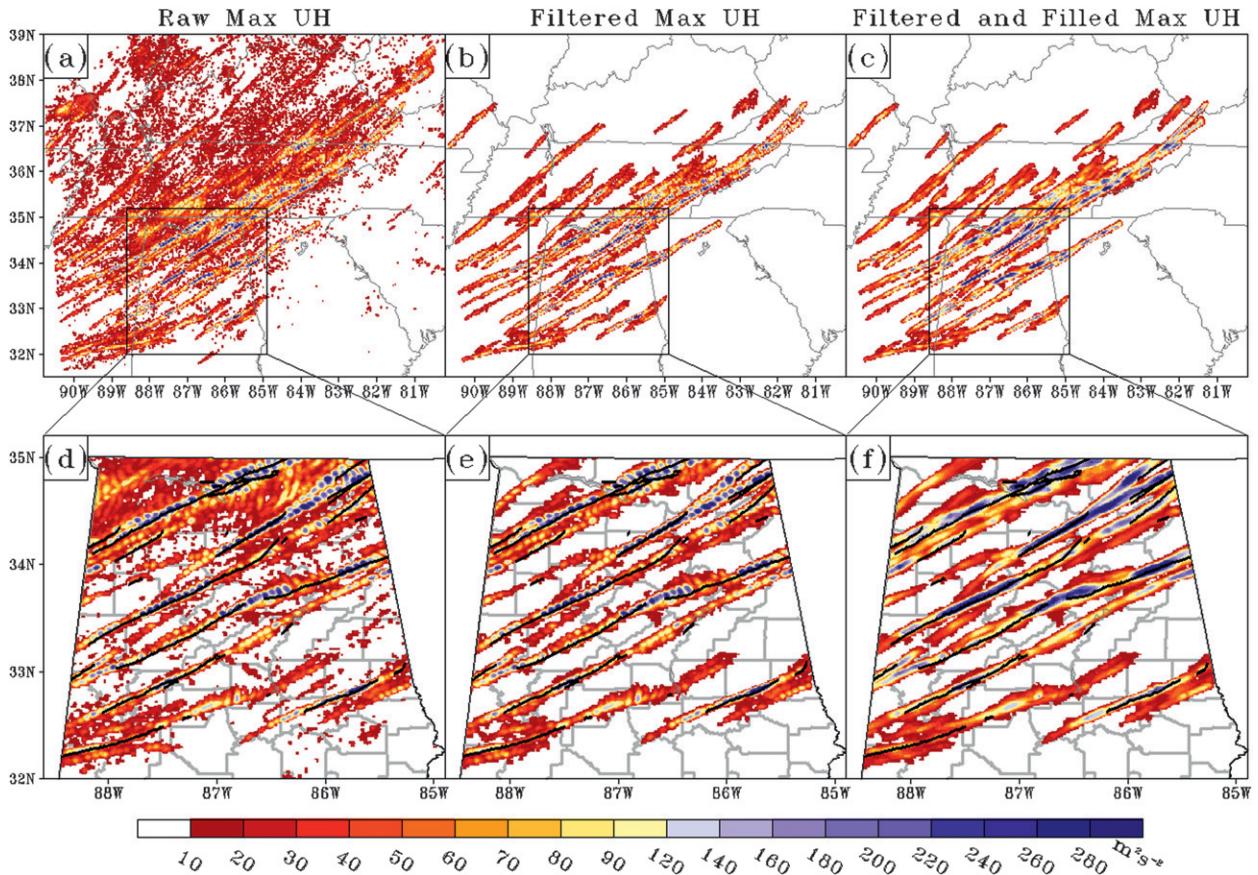


FIG. 9. (a) Maximum UH over the period 1500–0300 UTC 27–28 Apr 2011 computed from a radar-data-assimilating 1.25-km grid-spacing 3DVAR analysis constructed at 5-min intervals. (b) As in (a), but only 3D UH objects are plotted (see text for further information). (c) As in (b), but a “translating smoother” is applied to the 3D UH objects (see text for further information). The small boxes in (a)–(c) show the areas plotted in (d)–(f). (d)–(f) As in (a)–(c), respectively, except for a zoomed-in region and data outside of AL are masked out. The black lines indicate tracks of tornadoes that occurred during the 1500–0300 UTC period.

the analyzed UH gives a clear signal for mesocyclone tracks, as shown in Fig. 9. Furthermore, when comparing the 3DVAR analyses to simple radar-based observations of azimuthal shear for multiple supercell cases, Gao et al. (2009) and Smith et al. (2010) found that mesocyclones diagnosed by the 3DVAR system line up very well with those diagnosed from raw radial velocity data. Thus, the 3DVAR system should be a promising verification tool for high-resolution forecasts. Further work using NSSL’s Warning Decision Support System-Integrated Information (WDSS-II; Lakshmanan 2002) software package to display UH in real time from storm-scale 3DVAR analyses was conducted during the 2012 NOAA/HWT Experimental Warning Program Real-Time 3D Radar Data Assimilation Experiment.

To compute aggregate statistics on analyzed UH tracks, as well as to remove some of the noise in the raw UH fields so that the analyzed mesocyclones could be better visualized, a procedure was devised to

retain (or filter) only the UH from coherent mesocyclone tracks. This was done by defining 3D UH objects using “double thresholding”: if a contiguous area of at least 300 grid points (in space and time) has  $\text{UH} \geq 10 \text{ m}^2 \text{ s}^{-2}$  and anywhere within this area has  $\text{UH} \geq 60 \text{ m}^2 \text{ s}^{-2}$ , the area is considered a 3D object and all the grid points outside objects are set to zero. Double thresholding was necessary for the analyses because the most intense values of UH associated with individual mesocyclones oftentimes would not overlap over consecutive times, but overlap was easily achieved using a very low UH threshold. This behavior was not a problem in the model data because UH is sampled much more frequently (i.e., every model time step). An additional step was taken to manually remove broad and/or irregularly shaped UH objects clearly emanating from linear convective systems rather than mesocyclones. The specific thresholds and minimum number of grid points used for the filter were chosen subjectively based on experimentation with a number of different values. After

the 3D objects were defined, the final step was to apply a “translating smoother.” This involved computing the centroid of a 3D UH object at each time it was present. Centroids for each object at immediately adjacent times were then “connected” by straight-line segments. Along each straight-line segment, points were defined at 1-km increments. At each point along a segment, the two UH objects connected by the segment were translated using their centroid coordinate to the point, and a weighted average (weights were a linear function of the distance of each UH object to the point) of UH was computed between the two translated objects. Because of overlap with previous points, the maximum values of UH were always retained. For example, if the weighted average UH at a grid point from two translated objects was not greater than the previous value of UH at that grid point, the previous value of UH was retained. The translating smoother could be viewed as estimating the spatial distribution of the analyzed UH given continuous sampling, rather than 5 min, and thus allows for a better comparison with UH from the SSEF system members. Also, in this application it worked very well for removing the numerous circle-shaped maxima in UH that composed many of the tracks, which can be seen by comparing Fig. 9e to Fig. 9f.

After applying the filter and translating smoother, a total of 64 UH tracks were identified in the 3DVAR analyses. The longest track was about 728 km and 12 tracks were over 300 km long. The cumulative UH track length was 10 895 km. For the corresponding time period, the cumulative tornado track length was 2882 km. In Figs. 9d–f, tornado tracks over the Birmingham, Alabama, National Weather Service (NWS) County Warning Area (CWA)<sup>3</sup> are displayed along with analyzed UH tracks. Most of the tornado tracks lined up quite well with an analyzed UH track, although there were a couple of exceptions. If it is assumed that all tornado tracks were associated with an analyzed UH track, it can be estimated that roughly 26% of the analyzed UH track length was associated with an observed tornado. Although Trapp et al. (2005a) analyzed mesocyclone *events*, they found a similar percentage of tornadic mesocyclones examining radar data over a 3-yr period.

For a model versus analysis comparison, the maximum values of UH over the 3DVAR analysis domain for the 1500–0300 UTC time period are shown for SSEF system members (Figs. 10c–l; ordered from members that

produced the longest cumulative UH track length to those that produced the shortest) and filtered/smoothed 3DVAR analyses (Figs. 10a,b). Despite some of the previously discussed differences in model versus analyzed UH, the magnitude and distribution of UH values are actually quite similar. This similarity was achieved only after applying the translating smoother; without the smoother there are many fewer higher UH values (e.g., greater than  $200 \text{ m}^2 \text{ s}^{-2}$ ). Thus, the smoother basically serves to spread out the higher values to where they would have occurred in between the 5-min sampling intervals.

Visual comparison of graphics representing observed and simulated UH fields yields several interesting impressions. For example, none of the SSEF members were able to completely encompass the entire region over which the observed UH tracks occurred (blue highlighted region in Figs. 10a and 10c–l). Also, many of the SSEF members appear to have UH tracks that are more closely spaced than observed tracks; for members m17 (Fig. 10c) and m20 (Fig. 10f), this was especially apparent. Finally, the general orientation of the forecast UH tracks matched well with the analyses, and it is also worth noting that, although many of the individual solutions seemed to underpredict the area covered by UH tracks, the ensemble maximum (Fig. 4a) highlighted quite well the region over which the strongest analyzed UH tracks occurred. In addition, neighborhood ensemble probabilities (e.g., Schwartz et al. 2010; Clark et al. 2011) of  $\text{UH} \geq 100 \text{ m}^2 \text{ s}^{-2}$  within 40 km of a point (Fig. 10b, black and white shading) capture very well the main corridor of analyzed UH tracks. In fact, using the neighborhood probabilities as a forecast for an observed UH track (as shown in Fig. 10b), the area under the relative operating characteristic curve (ROC area; e.g., Mason 1982) is 0.92, indicating a very skillful forecast.<sup>4</sup>

#### 4. Summary and discussion

Examining April–June SSEF system forecasts run by CAPS for the 2010 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment, Clark et al. (2012b)

<sup>3</sup> Geographic Information System (GIS) shapefiles containing the tornado track data were obtained online ([http://www.srh.noaa.gov/media/bmx/042711/shapefiles/042711\\_StatewideTracks.zip](http://www.srh.noaa.gov/media/bmx/042711/shapefiles/042711_StatewideTracks.zip)).

<sup>4</sup> Neighborhood probabilities were obtained by computing the total number of members with at least one grid point of  $\text{UH} \geq 100 \text{ m}^2 \text{ s}^{-2}$  within a 40-km radius of each grid point divided by 17 (the total number of members for 27 April 2011). For computation of ROC areas, a grid point was considered a “hit” if it exceeded the specified probability threshold and contained at least one grid point from the analyzed filtered/smoothed UH tracks (i.e., those in Fig. 10b) within a 40-km radius. The ROC areas were computed using the trapezoidal method (Wandishin et al. 2001) and each point on the ROC curve corresponded to 1 of the 19 possible probability values from a 17-member ensemble.

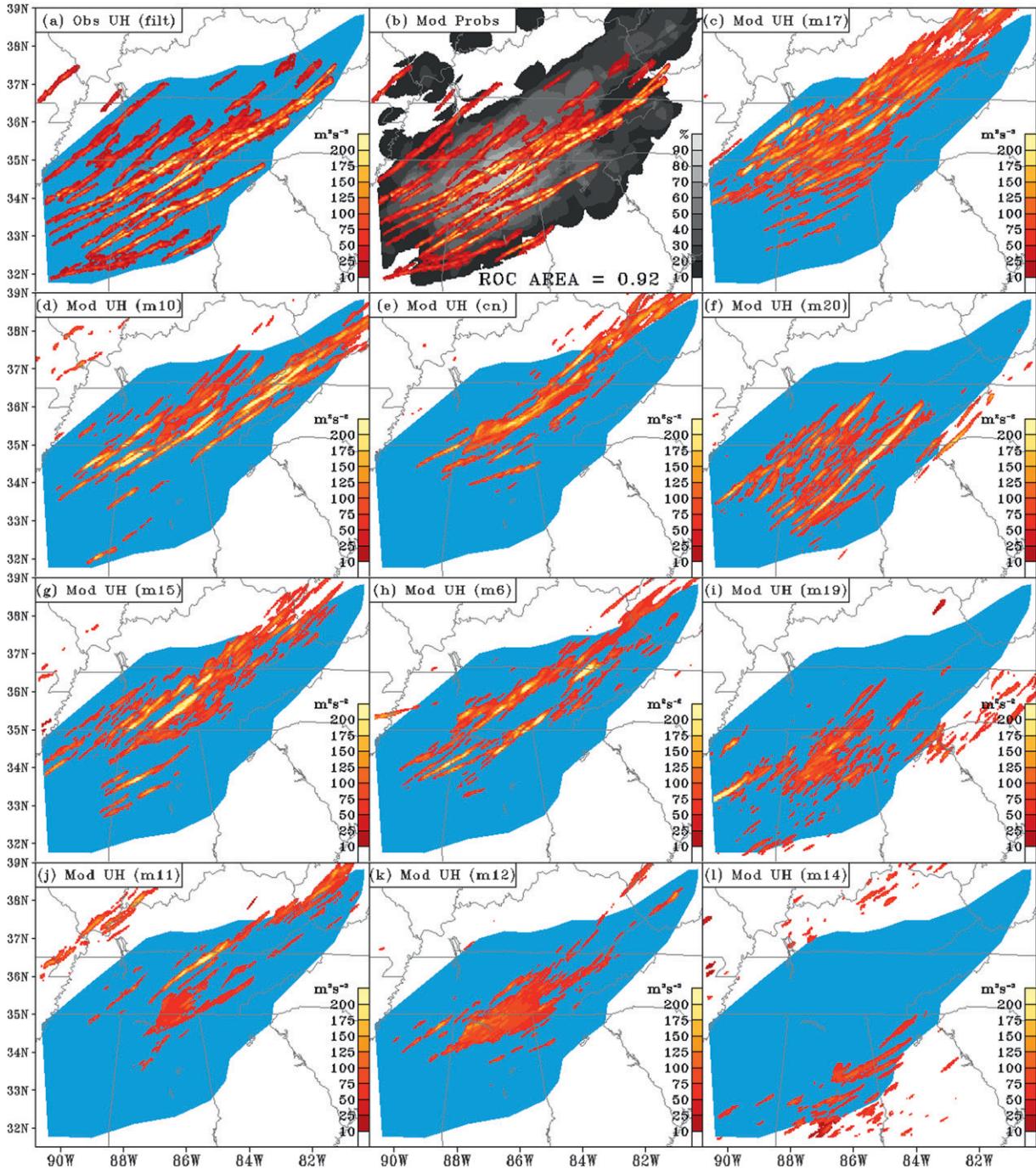


FIG. 10. (a) The 3D UH objects with the translating smoother applied for the period 1500–0300 UTC 27–28 Apr 2011 computed from radar-data-assimilating 1.25-km grid-spacing 3DVAR analyses constructed at 5-min intervals. The blue shading (also shown in other panels) was manually drawn to highlight the main corridor of UH tracks. (b) Neighborhood probabilities of UH  $\geq 100 \text{ m}^2 \text{ s}^{-2}$  within 40 km of a point computed from the SSEF system members valid for the same period as in (a) are shown by the grayscale shading. The analyzed UH objects from (b) are overlaid (red/yellow shading). (c)–(l) As in (a), but for SSEF system members (c) arm\_m17, (d) arw\_m10, (e) arw\_cn, (f) arw\_m20, (g) arw\_m15, (h) arw\_m6, (i) arw\_m19, (j) arw\_m11, (k) arw\_m12, and (l) arw\_m14.

diagnosed a strong relationship between the cumulative pathlengths of simulated rotating storms and the cumulative pathlengths of tornadoes. A three-dimensional object identification algorithm applied to forecast UH was used to measure simulated rotating storm pathlengths. This paper updates those results by including 34 additional April–June 2011 cases (for a total of 69 cases) from the SSEF system run for the 2011 Spring Forecasting Experiment. The main results were that there was little change in the Pearson correlation coefficients between cumulative UH and tornado pathlengths after including the 2011 cases and excluding two outlier cases from June 2010. For example, after filtering UH track segments from elevated and/or high-based storms, correlations went from 0.85 for the 2010-only dataset to 0.86 after including 2011 cases. However, the confidence intervals bounding the correlations noticeably shrank after including more cases.

SSEF system UH forecast examples were illustrated for three major 2011 tornado outbreaks: 16 and 27 April, and 24 May. For each of these cases, which occurred over distinctly different geographic regions, calibrated cumulative tornado pathlength exceedance probabilities were derived from the distribution of SSEF member UH pathlength forecasts. For 16 April and 24 May, these probabilities indicated that cumulative tornado pathlengths exceeding the 1-yr return period were very likely, and for both cases the actual cumulative tornado pathlength fell between the 55th and 60th percentiles of the ensemble and did, in fact, exceed the 1-yr return period. For 27 April, the exceedance probabilities were the most extreme out of all 69 cases in the dataset, especially for very long tornado pathlengths exceeding the 10-yr return period. Indeed, 27 April had the second highest cumulative tornado pathlength over the entire 1950–2011 period, and the 2949-km tornado pathlength that was observed fell at the 19th percentile of the ensemble forecast. There were no other cases for which exceedance probabilities at this pathlength were even above zero, a very encouraging sign that the SSEF system recognized the potential for an extreme event.

SSEF system forecast examples were also illustrated for the two June 2010 outlier (or forecast failure) cases. For the 1 June 2010 case, many of the SSEF system members predicted very long cumulative UH pathlengths, but the observed cumulative tornado pathlength was very short. Closer examination of the ensemble member forecasts revealed that they predicted the convective evolution very well, but that the long UH pathlengths originated from circulations at the leading edge of a linear MCS. Observed MCSs often have circulations similar to those in the simulations, but they do not produce tornadoes nearly as often and as intense as

those associated with supercellular mesocyclones (Trapp et al. 2005b).

For the other failure case that occurred on 17 June 2010, the SSEF system members predicted relatively short UH pathlengths and in actuality a significant tornado outbreak occurred with cumulative tornado pathlengths nearing those for a 2-yr return period outbreak. Comparing a 1-km grid-spacing version of the WRF that was configured identically to the 4-km grid-spacing control member of the SSEF system revealed that the 1-km simulations depicted more intense and realistic supercellular structures than the 4-km simulations. Interestingly, for the other cases, the 4-km simulations appeared to adequately depict supercellular structures/mesocyclones. Thus, without conducting more detailed analyses (which are beyond the scope of this study), it is not clear why the 4-km grid-spacing runs failed for this particular case. Nonetheless, analyses of these two failure cases illustrate two potential avenues for improving the UH–tornado pathlength relationship: 1) accounting for the simulated convective mode that UH emanates from and 2) increased model resolution.

Finally, to gain further insight into the 27 April tornado outbreak, a 3DVAR data assimilation system was used to construct 5-min, 1.25-km grid-spacing analyses from 1500 to 0300 UTC 27–28 April over a domain covering the outbreak. From these analyses, UH was computed using the same formulation as in the model data. It was found that the analyzed UH clearly delineated tracks of observed mesocyclones, especially after a filter based on the 3D object-identifying algorithm was applied as well as a translating smoother that essentially filled in the maximum values of UH that would have occurred between the 5-min sampling intervals. From the analyses, a total of 64 UH tracks were identified. The longest was 728 km and 12 tracks were over 300 km long. From the cumulative tornado and analyzed UH tracks lengths, it was estimated that 26% of the analyzed UH track length was associated with an observed tornado. Comparing spatial plots of analyzed versus forecast UH, it appeared that none of the SSEF members was able to completely encompass the entire region over which observed UH tracks occurred. Also, many of the SSEF members appeared to have more closely spaced UH tracks than in the analyses. Finally, ensemble maximum UH and neighborhood probabilities of  $\text{UH} \geq 100 \text{ m}^2 \text{ s}^{-2}$  worked very well for delineating the main corridor over which the strongest analyzed UH tracks occurred, which was also the general corridor over which the most tornadoes occurred.

Extracting information on explicitly simulated storm attributes for severe weather forecasting is a new and potentially very beneficial area of research. Convection-allowing

models provide much more direct information on convective mode and associated hazards than coarse convection-parameterizing models that can only provide information on characteristics of the forecast severe weather environment. Thus, forecasters using convection-allowing models do not have to infer storm mode from the forecast environment, which can be problematic because of the wide range of storm modes observed in similar regions of severe weather parameter space (e.g., Thompson et al. 2003, 2007; Grams et al. 2012). Aside from forecasting, the UH–tornado pathlength relationship also has potential applications for regional climate modeling and seasonal forecasting. Once regional climate models can be run with sufficient grid spacing to resolve mesocyclone-scale rotation, the cumulative UH pathlengths could prove a useful method for diagnosing the frequency/severity of tornado outbreaks in future climates and/or on seasonal time scales.

Aside from inherent predictability issues, challenges arise because, although convection-allowing simulations are capable of producing realistic storms, the societal hazards associated with these storms are not yet fully resolved at 4-km grid scales. For example, convection-allowing models can clearly reproduce supercell-like storms with strong rotation at the scale of large mesocyclones, but the models cannot resolve the tornadoes embedded within the mesocyclones. In some ways, this problem is analogous to identifying tornadoes using conventional WSR-88Ds. These radars cannot resolve tornadoes, but certain characteristics of radar-observed mesocyclones like strength and altitude (e.g., Trapp et al. 2005a) can help discriminate between tornadic and nontornadic mesocyclones. The same is likely true for simulated storms; certain characteristics of simulated storms should be related to the unresolved hazards these storms would produce if they were “real.” Based on the results from this study, one such characteristic is the simulated cumulative mesocyclone pathlength (measured using UH), which is related to the cumulative tornado pathlength. Ongoing work at NSSL also involves extracting simulated storm attributes to predict maximum hail size and surface wind gusts.

*Acknowledgments.* Funding for AJC, PTM, and JC was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce. CAPS SSEF forecasts were supported by the NOAA Collaborative Science, Technology, and Applied Research (CSTAR) Program with supplementary support from NSF Grant AGS-0802888. M. Xue was supported by NSF Grants OCI-0905040, AGS-0941491, AGS-1046171, and AGS-1046081. CAPS forecasts

were supported by an allocation of advanced computing resources provided by the National Science Foundation. The computations were performed on Athena (a Cray XT4) at the National Institute for Computational Science (<http://www.nics.tennessee.edu/>). CAPS utilized resources from the OU Supercomputing Center for Research and Education for ensemble postprocessing. Many scientists at CAPS, including Kevin Thomas, Yunheng Wang, Keith Brewster, Xuguang Wang, and Jerry Brotzge, contributed to the design and production of the CAPS ensemble forecasts. Christopher Karstens helped with displaying GIS tornado track data, and the comments/suggestions from two anonymous reviewers were very helpful in improving the manuscript.

## REFERENCES

- Chen, F., and J. Dudhia, 2001: Coupling an advanced land-surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model description and implementation. *Mon. Wea. Rev.*, **129**, 569–585.
- Chou, M.-D., and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, Vol. 3, 85 pp.
- Clark, A. J., and Coauthors, 2011: Probabilistic precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble. *Mon. Wea. Rev.*, **139**, 1410–1418.
- , and Coauthors, 2012a: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74.
- , J. S. Kain, P. T. Marsh, J. Correia Jr., M. Xue, and F. Kong, 2012b: Forecasting tornado pathlengths using a three-dimensional object identification algorithm applied to convection-allowing forecasts. *Wea. Forecasting*, **27**, 1090–1113.
- Coffer, B. E., L. C. Maudlin, P. G. Veals, and A. J. Clark, 2013: Dryline position errors in experimental convection-allowing NSSL-WRF model forecasts and the operational NAM. *Wea. Forecasting*, in press.
- Doswell, C. A., III, G. Carbin, and H. E. Brooks, 2012: The tornadoes of spring 2011 in the USA: An historical perspective. *Weather*, **67**, 88–94.
- Du, J., J. McQueen, G. DiMego, Z. Toth, D. Jovic, B. Zhou, and H.-Y. Chuang, 2006: New dimension of NCEP Short-Range Ensemble Forecasting (SREF) system: Inclusion of WRF members. *WMO Expert Team Meeting on Ensemble Prediction Systems*, Exeter, United Kingdom, WMO, 5 pp. [Available online at [http://www.emc.ncep.noaa.gov/mmb/SREF/WMO06\\_full.pdf](http://www.emc.ncep.noaa.gov/mmb/SREF/WMO06_full.pdf).]
- Ferrier, B. S., Y. Jin, Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and rainfall scheme in the NCEP Eta Model. Preprints, *15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., 10.1. [Available online at [https://ams.confex.com/ams/SLS\\_WAF\\_NWP/techprogram/paper\\_47241.htm](https://ams.confex.com/ams/SLS_WAF_NWP/techprogram/paper_47241.htm).]
- Gallus, W. A., Jr., N. A. Snook, and E. V. Johnson, 2008: Spring and summer severe weather reports over the Midwest as a function of convective mode: A preliminary study. *Wea. Forecasting*, **23**, 101–113.
- Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with

- recursive filter for Doppler radars. *J. Atmos. Oceanic Technol.*, **21**, 457–469.
- , D. J. Stensrud, and M. Xue, 2009: Three-dimensional analyses of several thunderstorms observed during VORTEX2 field operations. Preprints, *34th Conf. on Radar Meteorology*, Williamsburg, VA, Amer. Meteor. Soc., P9.8. [Available online at <https://ams.confex.com/ams/34Radar/webprogram/Paper156114.html>.]
- , —, T. Smith, K. Manross, J. Brogden, and K. Kuhlman, 2011: A realtime weather-adaptive 3DVAR analysis system with automatic storm positioning and on-demand capability. Preprints, *35th Conf. on Radar Meteorology*, Pittsburgh, PA, Amer. Meteor. Soc., 115. [Available online at <https://ams.confex.com/ams/35Radar/webprogram/Paper191834.html>.]
- , and Coauthors, 2013: A real-time weather-adaptive 3DVAR analysis system for severe weather detections and warnings. *Wea. Forecasting*, in press.
- Grams, J. S., R. L. Thompson, D. V. Snively, J. A. Prentice, G. M. Hodges, and L. J. Reames, 2012: A climatology and comparison of parameters for significant tornado events in the United States. *Wea. Forecasting*, **27**, 106–123.
- Hamill, T. M., and S. J. Colucci, 1997: Verification of Eta–RSM short-range ensemble forecasts. *Mon. Wea. Rev.*, **125**, 1312–1327.
- , and —, 1998: Evaluation of Eta–RSM ensemble probabilistic precipitation forecasts. *Mon. Wea. Rev.*, **126**, 711–724.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- Hu, M., M. Xue, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698.
- Janjić, Z. I., 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, NOAA/NWS, 61 pp.
- Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **5**, 931–952.
- , S. R. Dembek, S. J. Weiss, J. L. Case, J. J. Levit, and R. A. Sobash, 2010: Extracting unique information from high-resolution forecast models: Monitoring selected fields and phenomena every time step. *Wea. Forecasting*, **25**, 1536–1542.
- Kong, F., and Coauthors, 2011: CAPS storm-scale ensemble forecast in HWT 2011 Spring Experiment: Sensitivity of WRF physics on QPF. *Extended Abstracts, 12th Users' Workshop*, Boulder, CO, NCAR. [Available online at [http://www.mmm.ucar.edu/wrf/users/workshops/WS2011/Power%20Points%202011/6\\_3\\_Kong\\_WRFworkshop\\_11.pdf](http://www.mmm.ucar.edu/wrf/users/workshops/WS2011/Power%20Points%202011/6_3_Kong_WRFworkshop_11.pdf).]
- Lakshmanan, V., 2002: WDSII: An extensible, multi-source meteorological algorithm development interface. Preprints, *21st Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., 134–137.
- Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models. *Mon. Wea. Rev.*, **138**, 1587–1612.
- Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291–303.
- Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851–875.
- Milbrandt, J. A., and M. K. Yau, 2005: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62**, 3051–3064.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102** (D14), 16 663–16 682.
- Morrison, H., J. A. Curry, and V. I. Khvorostyanov, 2005: A new double-moment microphysics parameterization for application in cloud and climate models, Part I: Description. *J. Atmos. Sci.*, **62**, 1665–1677.
- Nakanishi, M., 2000: Large-eddy simulation of radiation fog. *Bound.-Layer Meteor.*, **94**, 461–493.
- , 2001: Improvement of the Mellor–Yamada turbulence closure model based on large-eddy simulation data. *Bound.-Layer Meteor.*, **99**, 349–378.
- , and H. Niino, 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31.
- , and —, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407.
- Noh, Y., W. G. Cheon, S.-Y. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteor.*, **107**, 401–427.
- Pleim, J. E., 2007: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *J. Appl. Meteor. Climatol.*, **46**, 1383–1395.
- Purser, R. J., W.-S. Wu, D. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535.
- , —, —, and —, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548.
- R Development Core Team, cited 2012: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Available online at <http://www.R-project.org/>.]
- Rogers, E., and Coauthors, 2009: The NCEP North American Mesoscale modeling system: Recent changes and future plans. Preprints, *23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction*, Omaha, NE, Amer. Meteor. Soc., 2A.4. [Available online at <http://ams.confex.com/ams/pdfpapers/154114.pdf>.]
- Schwartz, C. S., and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 2. NCAR Tech Note NCAR/TN-475+STR, 113 pp. [Available online at [http://www.mmm.ucar.edu/wrf/users/docs/arw\\_v3.pdf](http://www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf).]
- Smirnova, T. G., J. M. Brown, and S. G. Benjamin, 1997: Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. *Mon. Wea. Rev.*, **125**, 1870–1884.
- , —, —, and D. Kim, 2000: Parameterization of cold-season processes in the MAPS land-surface scheme. *J. Geophys. Res.*, **105** (D3), 4077–4086.

- Smith, T. M., K. M. Kuhlman, K. L. Ortega, K. L. Manross, D. W. Burgess, J. Gao, and D. J. Stensrud, 2010: A survey of real-time 3DVAR analyses conducted during the 2010 Experimental Warning Program Spring Experiment. Preprints, *25th Conf. on Severe Local Storms*, Denver, CO, Amer. Meteor. Soc., P5.7. [Available online at <http://ams.confex.com/ams/pdfpapers/176211.pdf>.]
- Sukoriansky, S., B. Galperin, and V. Perov, 2006: A quasi-normal scale elimination model of turbulence and its application to stably stratified flows. *Nonlinear Processes Geophys.*, **13**, 9–22.
- Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542.
- Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261.
- , C. M. Mead, and R. Edwards, 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102–115.
- Trapp, R. J., G. J. Stumpf, and K. L. Manross, 2005a: A reassessment of the percentage of tornadic mesocyclones. *Wea. Forecasting*, **20**, 680–687.
- , S. A. Tessendorf, E. S. Godfrey, and H. E. Brooks, 2005b: Tornadoes from squall lines and bow echoes. Part I: Climatological distribution. *Wea. Forecasting*, **20**, 23–34.
- Wandishin, M. S., S. L. Mullen, D. J. Stensrud, and H. E. Brooks, 2001: Evaluation of a short-range multimodel ensemble system. *Mon. Wea. Rev.*, **129**, 729–747.
- Weiss, S. J., and M. D. Vescio, 1998: Severe local storm climatology 1955–1996: Analysis of reporting trends and implications for NWS operations. Preprints, *19th Conf. on Severe Local Storms*, Minneapolis, MN, Amer. Meteor. Soc., 536–539.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, 467 pp.
- Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS)—A multiscale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteor. Atmos. Phys.*, **75**, 161–193.
- , and Coauthors, 2001: The Advanced Regional Prediction System (ARPS)—A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76**, 134–165.
- , D. Wang, J. , K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS) storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170.
- , and Coauthors, 2010: CAPS realtime storm scale ensemble and high-resolution forecasts for the NOAA Hazardous Weather Testbed 2010 Spring Experiment. Preprints, *25th Conf. on Severe Local Storms*, Denver, CO, Amer. Meteor. Soc., 7B.3. [Available online at <https://ams.confex.com/ams/25SLS/webprogram/Paper176056.html>.]