Forecasting Tornado Pathlengths Using a Three-Dimensional Object Identification Algorithm Applied to Convection-Allowing Forecasts

Adam J. Clark,[*,+] John S. Kain,[+] Patrick T. Marsh,[*,+,#] James Correia Jr.,[*,@] Ming Xue,[#,&] and Fanyou Kong[&]

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma
[+]*NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*
[#]*School of Meteorology, University of Oklahoma, Norman, Oklahoma*
[@]*NOAA/Storm Prediction Center, Norman, Oklahoma*
[&]*Center for Analysis and Prediction of Storms, Norman, Oklahoma*

ABSTRACT

A three-dimensional (in space and time) object identification algorithm is applied to high-resolution forecasts of hourly maximum updraft helicity (UH)—a diagnostic that identifies simulated rotating storms—with the goal of diagnosing the relationship between forecast UH objects and observed tornado pathlengths. UH objects are contiguous swaths of UH exceeding a specified threshold. Including time allows tracks to span multiple hours and entire life cycles of simulated rotating storms. The object algorithm is applied to 3 yr of 36-h forecasts initialized daily from a 4-km grid-spacing version of the Weather Research and Forecasting Model (WRF) run in real time at the National Severe Storms Laboratory (NSSL), and 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. Methods for visualizing UH object attributes are presented, and the relationship between pathlengths of UH objects and tornadoes for corresponding 18- or 24-h periods is examined. For deterministic NSSL-WRF UH forecasts, the relationship of UH pathlengths to tornadoes was much stronger during spring (March–May) than in summer (June–August). Filtering UH track segments produced by high-based and/or elevated storms improved the UH–tornado pathlength correlations. The best ensemble results were obtained after filtering high-based and/or elevated UH track segments for the 20 cases in April–May 2010, during which correlation coefficients were as high as 0.91. The results indicate that forecast UH pathlengths during spring could be a very skillful predictor for the severity of tornado outbreaks as measured by total pathlength.

1. Introduction

Recent National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) Spring Forecasting Experiments (e.g., Clark et al. 2012) in which convection-allowing[1] modeling systems have been tested for severe weather forecasting have made it increasingly clear that fully exploiting the information provided by such systems requires new and innovative model diagnostics, verification, and visualization strategies. This need arises because existing strategies utilize relatively coarse convection-parameterizing modeling systems that can only provide information on characteristics of the forecast severe weather environment. In contrast, convection-allowing models explicitly simulate convection, providing direct information on attributes and related hazards from explicitly simulated storms. Accurately predicting the environment remains important (e.g., Coniglio et al. 2010), but, given 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), explicit information on forecast storm attributes adds significant value to the forecast.

---

[1] The term convection allowing refers to simulations using the maximum grid spacing (or below) at which convection can be treated explicitly and midlatitude MCSs can be adequately resolved, which is generally thought to be ~4 km (Weisman et al. 1997).

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

One challenge in providing useful information on forecast storm attributes is that convective storms evolve on time scales of order minutes, which are much shorter than typical model output frequencies of 1–3 h. To monitor fields between model output times, Kain et al. (2010) developed a strategy in which individual gridpoint temporal-maximum values from each hour are stored in two-dimensional arrays and saved at hourly output intervals. These hourly maximum fields (HMFs) provide useful information on peak intensity and track of simulated storm features without storing all model fields every time step. One particularly valuable HMF designed to track rotation in simulated storms is updraft helicity [Kain et al. (2008); hereafter, UH denotes the hourly maximum quantity], which is computed by taking the integral of the vertical vorticity times the updraft velocity between 2 and 5 km AGL. Participants in recent NOAA/HWT Spring Forecasting Experiments have noticed that, oftentimes, the appearance of relatively intense swaths of UH in convection-allowing forecasts is a good indicator of supercells and associated hazards. However, participants have also noted that the perceived utility of UH was limited simply because of their unfamiliarity with the typical range of UH values and the lack of a documented relationship to observed severe weather.

Sobash et al. (2011) began addressing these issues using forecasts from a 4-km grid-spacing version of the Weather Research and Forecasting Model [WRF; using the Advanced Research WRF (ARW) dynamic core (Skamarock et al. 2008)] run in real time by the National Severe Storms Laboratory (NSSL) during spring 2008. Sobash et al. (2011) developed a simple method that involved identifying extreme values of UH and treating these as "surrogate" severe weather reports from which forecast severe weather probabilities were computed using kernel density estimation with a Gaussian kernel. These forecast probabilities were verified against observed storm reports and were found to be quite skillful, with the combination of best skill and reliability obtained by defining surrogate severe reports using a UH threshold of about $35 \ m^2 \ s^{-2}$ and a sigma value, or bandwidth, for the Gaussian kernel of about 200 km.

This study aims to build on previous work by further examining the statistical properties of forecast UH and its relationship to observed severe weather. Furthermore, methods for visualizing forecast attributes of UH from deterministic and ensemble prediction systems are presented that quickly convey useful information to forecasters. These aims are accomplished through the application of a three-dimensional (in space and time) object identification algorithm to forecast UH. Objects identified by the algorithm are contiguous regions, or tracks, of UH exceeding a specified threshold. Once UH objects are identified, attributes like length and maximum intensity are computed for each object.

The concept of "object based" forecast verification has become increasingly popular for verifying high-resolution forecasts because it addresses shortcomings in traditional methods that rely on exact matches of paired model–observation grid points (e.g., Ebert and McBride 2000; Davis et al. 2006a,b). These traditional methods harshly penalize high-resolution forecasts containing high-amplitude features with sharp gradients (e.g., Baldwin et al. 2001; Mass et al. 2002; Gilleland et al. 2009; Clark et al. 2010), provide limited diagnostic information on forecast errors, and are often inconsistent with subjective evaluations (Kain et al. 2003). In contrast, object-based methods provide meaningful diagnostic information on forecasts errors like displacement, orientation, and intensity, and they are designed to more closely mimic subjective evaluation approaches (Davis et al. 2006a). Until recently, object-based approaches have only considered two-dimensional spatial objects. However, recent work using the Method for Object-based Diagnostic Evaluation (MODE; Davis et al. 2009) has incorporated the time dimension. Thus, "3D objects" are contiguous regions of grid points encompassing both space and time. Incorporating time results in a much more powerful diagnostic tool that provides important information on aspects of phenomena attributes like timing, evolution, and translation speed, which would not be available by only considering the spatial dimension. In fact, we believe that a 3D object-based algorithm will be a particularly powerful tool for warn-on-forecast applications in which very short-range/high-resolution ensemble forecasts will be used to forecast the evolution and timing of hazardous convective weather (Stensrud et al. 2009). The 3D object algorithm applied herein was developed in house at NSSL, and was specifically designed for the current application to UH. Including the time dimension allows UH tracks to span multiple forecast hours and encompass entire life cycles of simulated rotating storms. As of this writing, the next community release of the Developmental Testbed Center's Model Evaluation Tools (MET; current version available online at http://www.dtcenter.org/met/users/downloads/) is planned to have 3D capabilities in MODE.

In this study, the relationship of forecast UH pathlengths to observed tornado pathlengths is examined. Compared to other tornado attributes (e.g., number and intensity), pathlengths are a preferred metric because, as noted by Edwards et al. (2004), pathlengths can be more precisely measured during storm surveys, and cumulative pathlength is a very good indicator for the overall severity of a tornado outbreak. Thus, a strong relationship

between total tornado and forecast UH pathlengths would mean that forecast UH could be used as a skillful predictor for the overall severity of a tornado outbreak. Ideally, forecast UH paths would be verified against observed mesocyclone paths (i.e., "observed UH" or vertical vorticity). Work is on going at NSSL to develop such a product for verification purposes using the Vortex Detection and Diagnosis Algorithm (VDDA; Stumpf et al. 2004) within the NSSL Warning Decision Support System–Integrated Information (WDSS-II; Lakshmanan 2002) software package. Future applications of the 3D object algorithm will use this WDSS-II-derived dataset with a strategy similar to that used herein.

Convection-allowing models can reproduce supercell-like storms with strong rotation at the scale of large mesocyclones. Thus, we expect at least some degree of correspondence between forecast UH and observed tornado tracks, because most tornadoes (especially long-track ones) are produced by supercells. However, it is not clear how strong the relationship will be because UH tracks in convection-allowing simulations often emanate from simulated storms that are elevated [i.e., inflow is drawn from an unstable layer above the surface; Colman (1990)] and/or high based and, thus, would have a low chance of producing a tornado (e.g., Trapp et al. 2005a). In addition, even supercells that produce tornadoes often do so for only a fraction of their lifetime. Furthermore, strong/long-lived UH tracks sometimes originate from features in simulated organized convective systems, like line-end vortices. Finally, the leading edges of strong gust fronts where there are strongly sheared updrafts can also result in high values of UH (e.g., Sobash et al. 2011). Tornadoes occurring with such MCS-related features do happen, but they are much less frequent and usually weaker and shorter lived than tornadoes from supercells (e.g., Trapp et al. 2005b).

Despite the potential for "false alarms" using UH tracks to predict tornadoes, increased knowledge of the relationship with tornadoes should be beneficial to forecasters. The remainder of this study is organized as follows. Section 2 contains information on the models examined, the tornado report database, and the 3D object algorithm. Section 3 contains results from an analysis of deterministic and ensemble forecasts, and section 4 contains a summary and conclusions.

2. Model specifications and methodology

Deterministic forecasts are examined using a version of the WRF-ARW run in real time at NSSL (hereafter, referred to as NSSL-WRF; Kain et al. 2010). The analysis period covers 31 January 2008, which is when NSSL began archiving UH fields, to 31 December 2010 (1016

cases). The forecasts are initialized daily at 0000 UTC and integrated for 36 h. The 3D object algorithm is applied to forecast hours 13–36, corresponding to the 24-h period 1200–1200 UTC. This 24-h period is ideal because it matches the period over which the Storm Prediction Center (SPC) issues their initial day-1 severe weather outlooks, it avoids the first 12 h of the forecast when model "spinup" occurs, and it results in non-overlapping daily analysis periods. Before 9 June 2009, WRF version 2.2 was used with a domain encompassing most of the United States except for portions of the west (Fig. 1a). After 9 June 2009, the domain was expanded to encompass the entire CONUS and the model version was updated to version 3.1.1. Other model specifications during the 3-yr analysis period were unchanged. Physics parameterizations are listed in Table 1. Initial conditions (ICs) and lateral boundary conditions (LBCs; 3-h updates) are from National Centers for Environmental Prediction's (NCEP) North American Mesoscale (NAM; Rogers et al. 2009) model. The 3D object algorithm is applied over a masked subregion of the NSSL-WRF domain (Fig. 1a), which was chosen so that mostly land areas within the United States away from the lateral boundaries and covered by the Next Generation Weather Radar (NEXRAD) network were included.

Ensemble forecasts examined were from the Storm Scale Ensemble Forecast (SSEF) system run by the Center for Analysis and Prediction of Storms (CAPS) for the 2010 NOAA Hazardous Weather Testbed Spring Forecasting Experiment (Xue et al. 2010; Clark et al. 2012). During 2010, the SSEF system had 26 members with 4-km grid spacing that were initialized on weekdays at 0000 UTC and integrated 30 h over a CONUS domain (Fig. 1a) from late April to mid-June. The 3D object algorithm is applied to forecast hours 13–30, corresponding to the 18-h period 1200–0600 UTC. The WRF-ARW was used for 19 SSEF members, the Nonhydrostatic Mesoscale Model (NMM; Rogers et al. 2009) dynamic core was used for 5 members, and the Advanced Regional Prediction System (ARPS; Xue et al. 2003) was used for 2 members. In this study, a 14-member subset of the ensemble comprising the ARW and NMM control members and members with varied physics and perturbed ICs/LBCs are examined.[2] Ensemble specifications for these members are provided in Table 1 and further details on all members can be found in Xue et al. (2010). ICs and LBCs (3-h updates) for the control member were taken from NAM analyses and forecasts, respectively. Radial velocity and reflectivity

---

[2] ARPS members were not examined because hourly maximum UH was not an available diagnostic from these members.
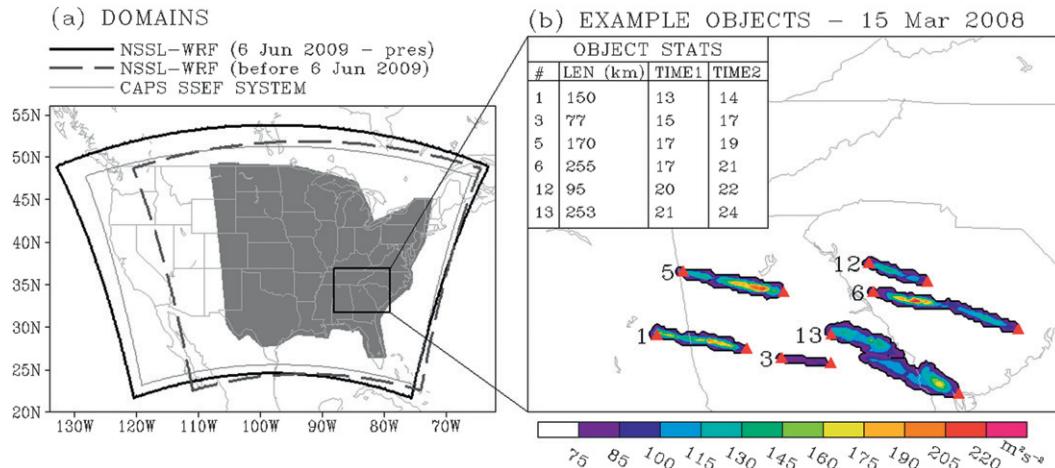
FIG. 1. (a) Model domains for NSSL-WRF before and after 9 Jun 2009, and the domain for the 2011 CAPS SSEF system. The gray-shaded region is the mask over which all analyses were conducted. A legend is provided at the top. (b) Example of selected UH objects (minimum threshold of 75 m^2 s^{-2}) identified by the 3D object algorithm for NSSL-WRF initialized at 0000 UTC 15 Mar 2008. 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 top 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, in which the object was present.

data from up to 140 Weather Surveillance Radar-1988 Doppler (WSR-88D) and other high-resolution observations were assimilated into the ICs using the 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; Xue et al. 2008) system. IC perturbations were derived from evolved (through 3 h) perturbations of 2100 UTC initialized members of NCEP's 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 reports, as compiled in the National Climatic Data Center (NCDC) publication *Storm Data*, were obtained from the SPC web site (http://www.spc.noaa.gov/wcm/). Known deficiencies in *Storm Data* include inconsistency in reporting practices among County Warning Areas and population-dependent reporting density (Weiss and Vescio 1998; Gallus et al. 2008 and references therein).

The 3D object algorithm is applied to the raw UH fields (i.e., no smoothing) for UH thresholds of 50, 75, 100, 125, and 150 m^2 s^{-2}. These thresholds were chosen based on typical ranges of UH observed from daily examinations of the NSSL-WRF forecasts. The identification criteria are very simple: the objects must contain at least five contiguous grid points in space and/or time. Smoothing is undesirable because UH swaths or tracks are very finescale features with high amplitudes and

sharp gradients. Smoothing, especially when multiple swaths are closely spaced, can result in unwanted merging of individual tracks. Note, even without smoothing, merging can occur for closely spaced tracks.

A different type of problem occurs for splitting or merging rotating storms that result in a single UH track that branches into multiple cases, or multiple UH tracks that merge into one. When splitting or merging occurs, all branches get lumped into one 3D object, so there is not a decision process by which a particular branch is chosen, as is the case for other more sophisticated radar storm tracking algorithms [e.g., Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN; Dixon and Wiener (1993); Han (2009)]. Not accounting for splits and/or mergers should result in an underestimation of total UH track lengths. However, based on subjective examinations of the UH forecasts, merging and/or splitting rarely occurs at the higher thresholds examined (e.g., UH $\geq$ 100 m^2 s^{-2}). Because these high thresholds most reliably predict total tornado pathlengths (discussed later), merging likely does not significantly impact the more important results. Furthermore, systematic underestimation of UH pathlengths should not affect the correlation of UH track lengths to tornado track lengths, which is the focus of subsequent analyses.

To compute the length of UH objects, the objects are divided into segments according to each forecast hour they were present. The latitude–longitude coordinates

TABLE 1. SSEF system and NSSL-WRF configurations. All SSEF system WRF members used version 3.1.1, while NSSL-WRF used version 3.1.1 (2.2) after (before) 9 Jun 2009. Here, NAMa and NAMf refer to the NAM analysis and forecast, respectively. The 12- and 40-km NAM grids were used in the SSEF system and NSSL-WRF, respectively. ARPS 3DVAR and cloud analysis is referred to by ARPSa. Elements in the ICs column followed by a plus (+) or minus (−) indicate SREF member perturbations added to the control member ICs. Boundary layer schemes included MYJ (Mellor and Yamada 1982; Janjić 2002), YSU (Noh et al. 2003), MYNN (Nakanishi 2000, 2001; Nakanishi and Niino 2004, 2006), and QNSE (Sukoriansky et al. 2006). 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), Purdue Lin (Chen and Sun 2002 and references therein), and Morrison et al. (2005). Radiation schemes included the Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997), Dudhia (1989), and Geophysical Fluid Dynamics Laboratory (GFDL; Lacis and Hansen 1974) short wave, and Goddard (Chou and Suarez 1994), RRTM (Mlawer et al. 1997), GFDL (Fels and Schwarzkopf 1975; Schwarzkopf and Fels 1991), and Dudhia (1989) long wave. Land surface models included Noah (Chen and Dudhia 2001), and Rapid Update Cycle (RUC; Smirnova et al. 1997, 2000) options.

Ensemble member	ICs	LBCs	Radar data	Microphysics	Shortwave radiation	Longwave radiation	Land surface model	Boundary layer
1) arw_cn	00Z ARPSa	00Z NAMf	Yes	Thompson	Goddard	RRTM	Noah	MYJ
2) arw_m6	+ em-p1_pert	em-p1	Yes	Morrison	Goddard	RRTM	RUC	YSU
3) arw_m7	+ em-p2_pert	em-p2	Yes	Thompson	Goddard	RRTM	Noah	QNSE
4) arw_m8	− nmm-p1_pert	nmm-p1	Yes	WSM6	Goddard	RRTM	RUC	QNSE
5) arw_m9	+ nmm-p2_pert	nmm-p2	Yes	WDM6	Goddard	RRTM	Noah	MYNN
6) arw_m10	+ rsmSAS-n1_pert	rsmSAS-n1	Yes	Ferrier	Goddard	RRTM	RUC	YSU
7) arw_m11	− etaKF-n1_pert	etaKF-n1	Yes	Ferrier	Goddard	RRTM	Noah	YSU
8) arw_m12	+ etaKF-p1_pert	etaKF-p1	Yes	WDM6	Goddard	RRTM	RUC	QNSE
9) arw_m13	− etaBMJ-n1_pert	etaBMJ-n1	Yes	WSM6	Goddard	RRTM	Noah	MYNN
10) arw_m14	+ etaBMJ-p1_pert	etaBMJ-p1	Yes	Thompson	Goddard	RRTM	RUC	MYNN
11) nmm_cn	00Z ARPSa	00Z NAMf	Yes	Ferrier	GFDL	GFDL	Noah	MYJ
12) nmm_m3	+ nmm-n1_pert	nmm-n1	Yes	Thompson	RRTM	Dudhia	Noah	MYJ
13) nmm_m4	+ nmm-n2_pert	nmm-n2	Yes	WSM6	RRTM	Dudhia	RUC	MYJ
14) nmm_m5	+ em-n1_pert	em-n1	Yes	Ferrier	GFDL	GFDL	RUC	MYJ
NSSL-WRF configuration								
NSSL-WRF	00Z NAMa (40-km)	00Z NAMf (40-km)	No	WSM6	Dudhia	RRTM	Noah	MYJ

for the beginning and end points of each segment are found by searching for the farthest west and farthest east grid points within each segment. Then, the total object length is found by summing the lengths of the segments [computed using the haversine formula (Sinnott 1984)—a common geometric method for calculating distances]. A subjective examination of the beginning and end points identified by the algorithm found that it appeared to work very well, which is likely related to UH objects typically having very large aspect ratios (i.e., they are very long and skinny). Figure 1b illustrates a case for which the algorithm was applied using a threshold of 75 m^2 s^{-2}.

3. Results

a. Example deterministic forecasts: NSSL-WRF

For usefulness as a forecasting tool, information on the number, length, and maximum intensity of UH objects should be presented in a way that quickly and intuitively conveys the key object attributes. Figure 2 graphically presents information on forecast UH objects in such a manner (see Fig. 2 caption for details). NSSL-WRF forecasts are shown for five cases in which

relatively long total UH pathlengths were present, but the corresponding total tornado pathlengths were highly variable. For example, on 5 February 2008 (Fig. 2a), UH > 75 m^2 s^{-2} worked best for predicting total tornado pathlength; on 23 May 2008 (Fig. 2b), UH > 100 m^2 s^{-2} worked best; on 5 June 2010 (Fig. 2c), UH > 125 m^2 s^{-2} worked best; and for 7 June 2010 (Fig. 2d) and 15 June 2009 (Fig. 2e), all UH thresholds overpredicted total tornado pathlength, as only a few tornadoes with short pathlengths occurred on these two days. Generally, total UH pathlengths decrease quite rapidly with increasing UH threshold, and the longer UH tracks have stronger maximum values of UH.

According to various criteria (e.g., Edwards et al. 2004), the three cases in Figs. 2a–c could be classified as tornado outbreaks; on each of these days at least one violent tornado (i.e., enhanced Fujita scale rating ≥ EF4) occurred and total tornado track lengths were greater than 350 km. The 5 February 2008 case, popularly known as the "Super Tuesday" outbreak, had the longest total tornado track length within the analysis period (1032 km). The 23 May 2008 case had the second longest, and the 5 June 2010 case was in the top 20. For each of these outbreaks, forecast UH tracks
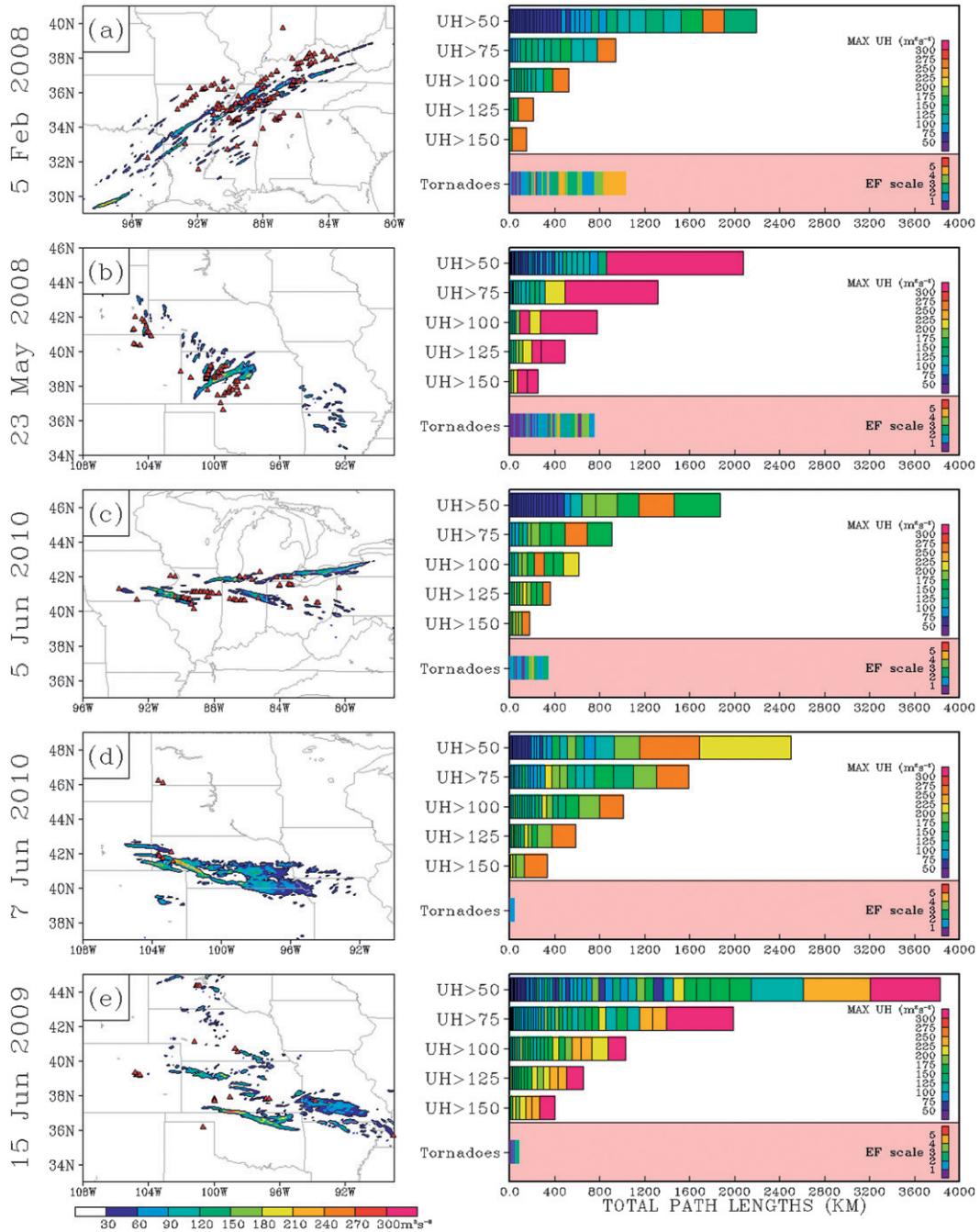
FIG. 2. (a) Maximum UH forecast by NSSL-WRF initialized at 0000 UTC 5 Feb 2008 for forecast hours 13–36 (valid 1200–1200 UTC 5–6 Feb). Tornado reports (red triangles) for the corresponding time period are overlaid. Right panel shows total length of UH objects for each threshold. The length of the individual bars composing each column indicates the length of the individual UH objects sorted shortest to longest and color coded according to the maximum intensity of UH in each object. The bars in the bottom column similarly indicate pathlengths and maximum intensities, but for observed tornadoes where maximum intensities correspond to EF ratings. (b)–(e) As in (a), butt for forecasts initialized at 0000 UTC on 23 May 2008, 5 Jun 2010, 7 Jun 2010, and 15 Jun 2009, respectively.

corresponded very well with the general regions of observed tornadoes. Additionally, areas of forecast UH were mainly characterized by long, skinny, and coherent tracks that would have suggested long-track mesocyclones.

In contrast to the cases in Figs. 2a–c, those in Figs. 2d,e were associated with only a few weak tornadoes with short pathlengths. However, many of the corresponding UH tracks for these cases were very long with relatively intense maximum values. For the 7 June 2010 case (Fig. 2d), inspection of observed composite reflectivity data (Fig. 3) revealed that the observed convective mode quickly transitioned from discrete cells (some supercellular) to a strong linear mesoscale convective system (MCS). Consistent with this quick transition to a linear convective mode, reported severe hazards transitioned from mainly large hail and a few tornadoes to high winds (Fig. 3). Although not obvious from the UH object attributes (rhs of Fig. 2d), the spatial plots of UH reveal a clear difference in the character of the forecast UH. Specifically, the westernmost tracks (e.g., southeast Wyoming–western Nebraska in Fig. 2d) tend to be long and coherent, whereas the easternmost tracks (e.g., southeast Nebraska in Fig. 2d) are more diffuse without the more intense maximum UH values. NSSL-WRF-simulated composite reflectivity (Fig. 3) revealed that the transition in UH character reflected a transition in simulated convective mode from strong isolated cells to a linear MCS, quite accurately reflecting the observed evolution of convective mode. In this case, the easternmost UH tracks associated with the simulated MCS were likely not indicative of strong rotation; rather, these areas of UH probably reflected strongly sheared updrafts, a characteristic feature of linear systems like squall lines. Sobash et al. (2011) also discussed how relatively high values of UH can be generated by linear convective systems; however, for their purposes, association of high UH with such systems was beneficial because they used UH not only as a surrogate for tornadoes, but also for wind and hail—hazards commonly associated with linear systems. For this study, which focuses on tornadoes, the 7 June 2010 case illustrates that further work is needed to more reliably discriminate between specific types of potential severe weather threats. For example, development of methods to distinguish UH tracks produced by simulated phenomena that typically produce tornadoes (i.e., supercells) versus those that do not (i.e., MCSs) would clearly be useful.

Finally, similar to Fig. 2d, the westernmost UH tracks in Nebraska–Kansas–Oklahoma for the 15 June 2009 case (Fig. 2e) were long and coherent whereas the easternmost UH swath in southern Missouri was more diffuse. However, unlike 7 June 2010, when the convective evolution was very well forecast, NSSL-WRF forecasts for 15 June 2009 featured a spurious morning MCS that produced the UH swath in southern Missouri. In addition, the long UH tracks in Kansas and northern Oklahoma, some of which contained very intense maximum values, would have indicated intense/long-track mesocyclones, which did not actually occur.

In summary, these examples illustrate the variety of solutions that can be obtained for different cases when using total UH pathlength to predict total tornado pathlength. Clearly, in some cases, the spatial UH plots as well as the diagrams of object attributes would have provided useful information; in one case they would have been a "false alarm" even though convective evolution was well forecast, and in another case they would have been a false alarm because the convective evolution was poorly forecast. In the following section, a more robust statistical analysis is presented using the entire dataset to derive the relationship between forecast UH and tornado pathlengths.

b. Statistical analysis of NSSL-WRF deterministic forecasts

To compare general characteristics of the total forecast UH and tornado pathlength climatology for the 3-yr analysis period, Fig. 4 shows frequency histograms, time series, and daily pathlengths ordered from shortest to longest (see Fig. 4 caption for more details). The characteristics of the distribution of total UH pathlengths computed using the highest UH thresholds—125 and 150 m^2 s^{-2} (Figs. 4e and 4f, respectively)—most closely matches that of the total tornado pathlengths. Using the lower UH thresholds (Figs. 4b–d), pathlengths are generally longer and the longer pathlengths occur more frequently so that the frequency distributions are not as right skewed as for total tornado pathlengths. In addition, comparing the time series of daily UH and tornado pathlengths (top inside plot in Fig. 4) reveals that the longest UH pathlengths tend to occur during June and later, while for tornadoes the longest pathlengths generally occur before June.

The tendency for longer UH pathlengths to occur after June is very clearly seen from the monthly averages of forecast UH pathlengths compared to pathlengths of tornadoes (Fig. 5a). During March–May, the increase in average UH pathlengths follows the increase in average tornado pathlengths, with the 125 m^2 s^{-2} UH threshold matching tornadoes most closely. However, average tornado pathlengths decrease in June from their peak in May, while forecast UH pathlengths increase sharply from May to June. In fact, even the highest UH threshold examined (150 m^2 s^{-2}) overpredicts the average tornado pathlengths during June and July. It is
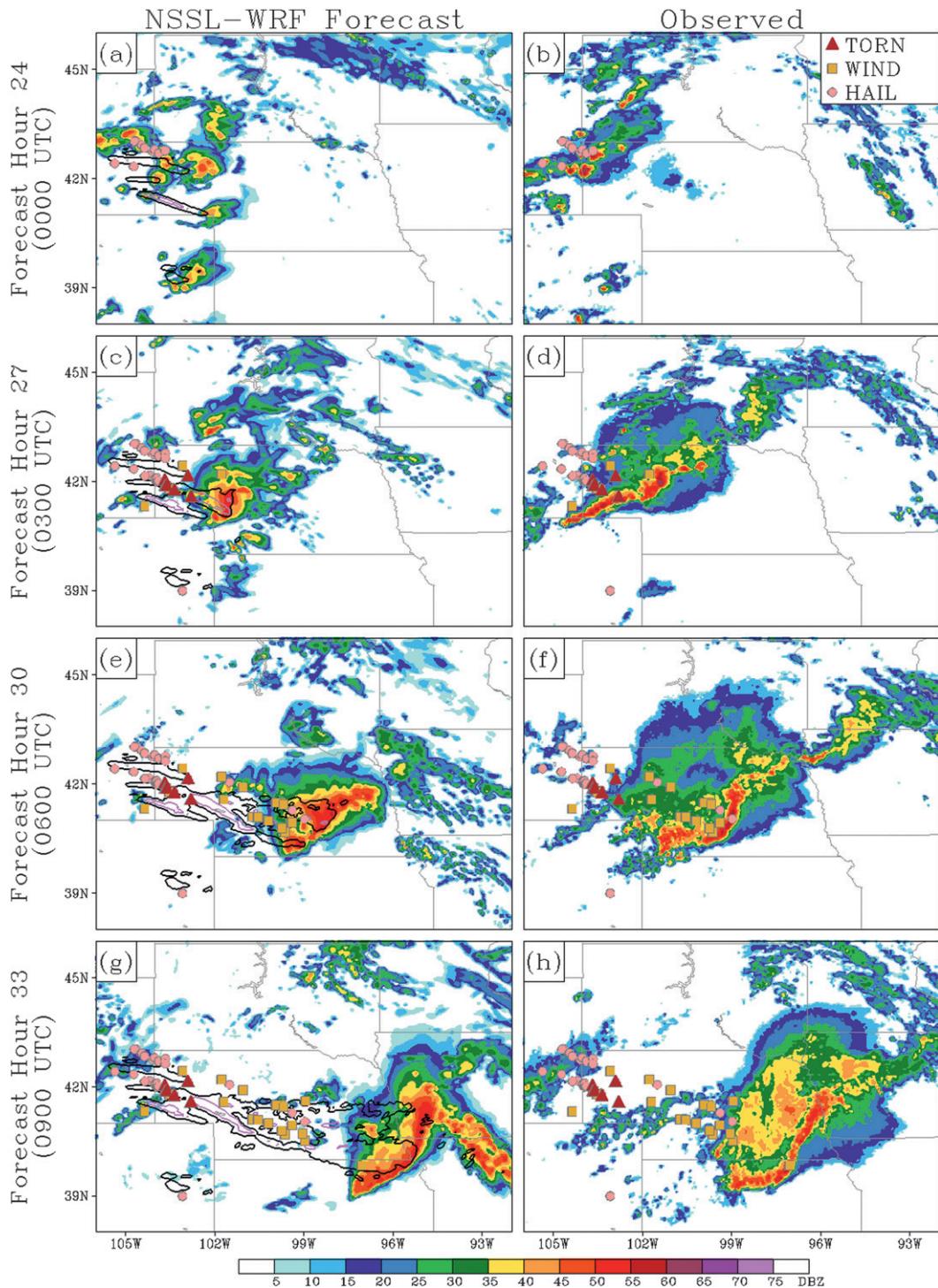
FIG. 3. NSSL-WRF-simulated composite reflectivity (shaded) for forecasts initialized at 0000 UTC 7 Jun 2010, and corresponding maximum UH (black and red contours for UH greater than 30 and 125 $m^2 s^{-2}$, respectively) from forecast hour 13 to the forecast hour for which the reflectivity images are valid, which includes forecast hours (a) 24, (c) 27, (e) 30, and (g) 33. Observed composite reflectivity corresponding to forecast hours (b) 24, (d) 27, (f) 30, and (h) 33. Observed storm reports [legend in top-right of (b)] up to valid time overlaid in all panels.
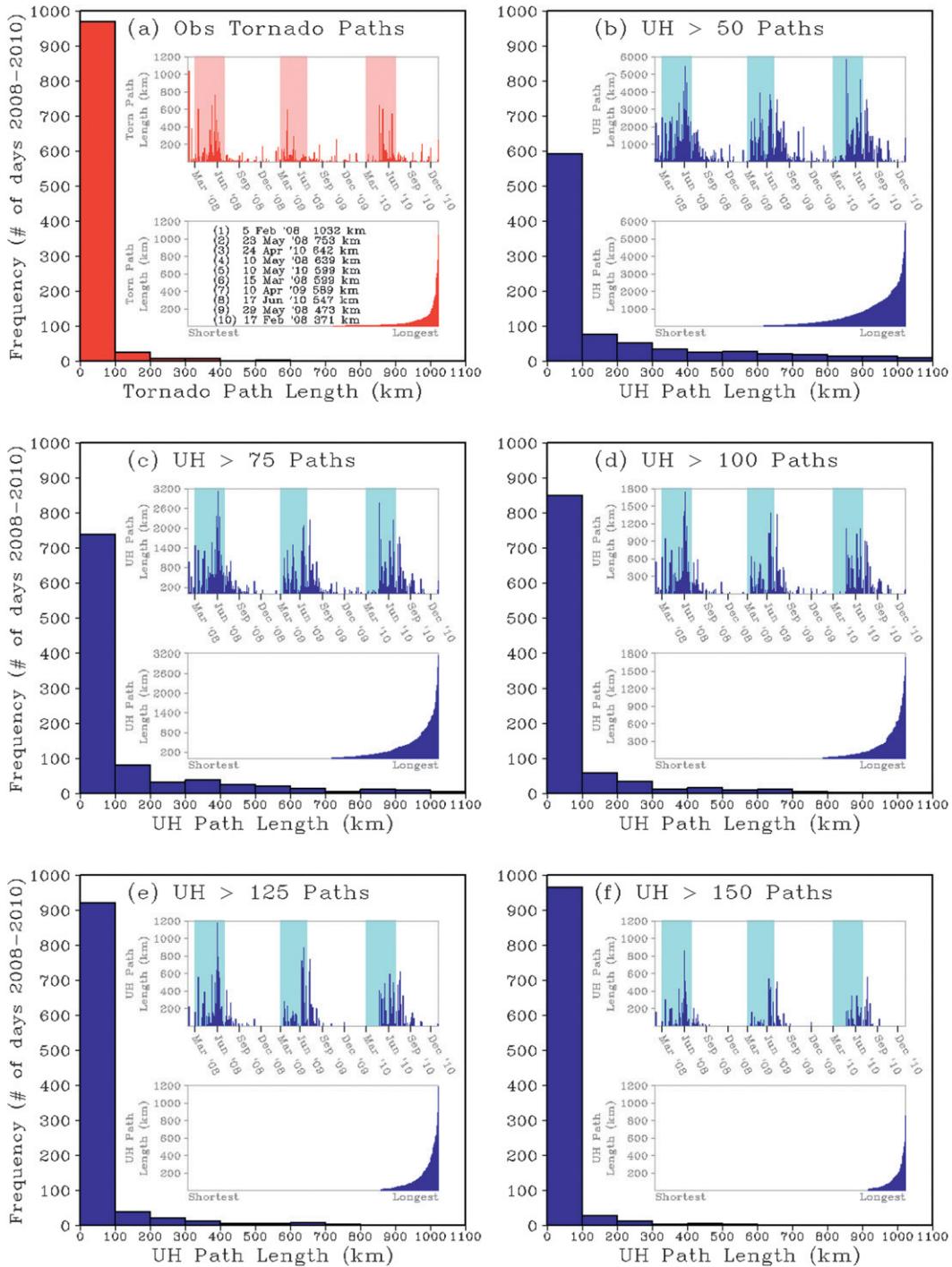
FIG. 4. (a) Frequency histogram of total daily tornado pathlengths for the period 2008–10. The top inside plot shows tornado pathlengths for each day with the pink-shaded regions highlighting the months March–June. The bottom inside plot shows the daily tornado pathlengths ordered from shortest to longest and a top 10 ranking of the days with the longest total tornado pathlengths. (b) As in (a), but for UH pathlengths (minimum threshold of 50 m² s⁻²) identified by the 3D object algorithm. Note, the *y* axis is scaled differently than in (a). (c)–(f) As in (b), but for minimum UH thresholds of (c) 75, (d) 100, (e) 125, and (f) 150 m² s⁻². Again, note the differently scaled *y* axes in the inside plots.
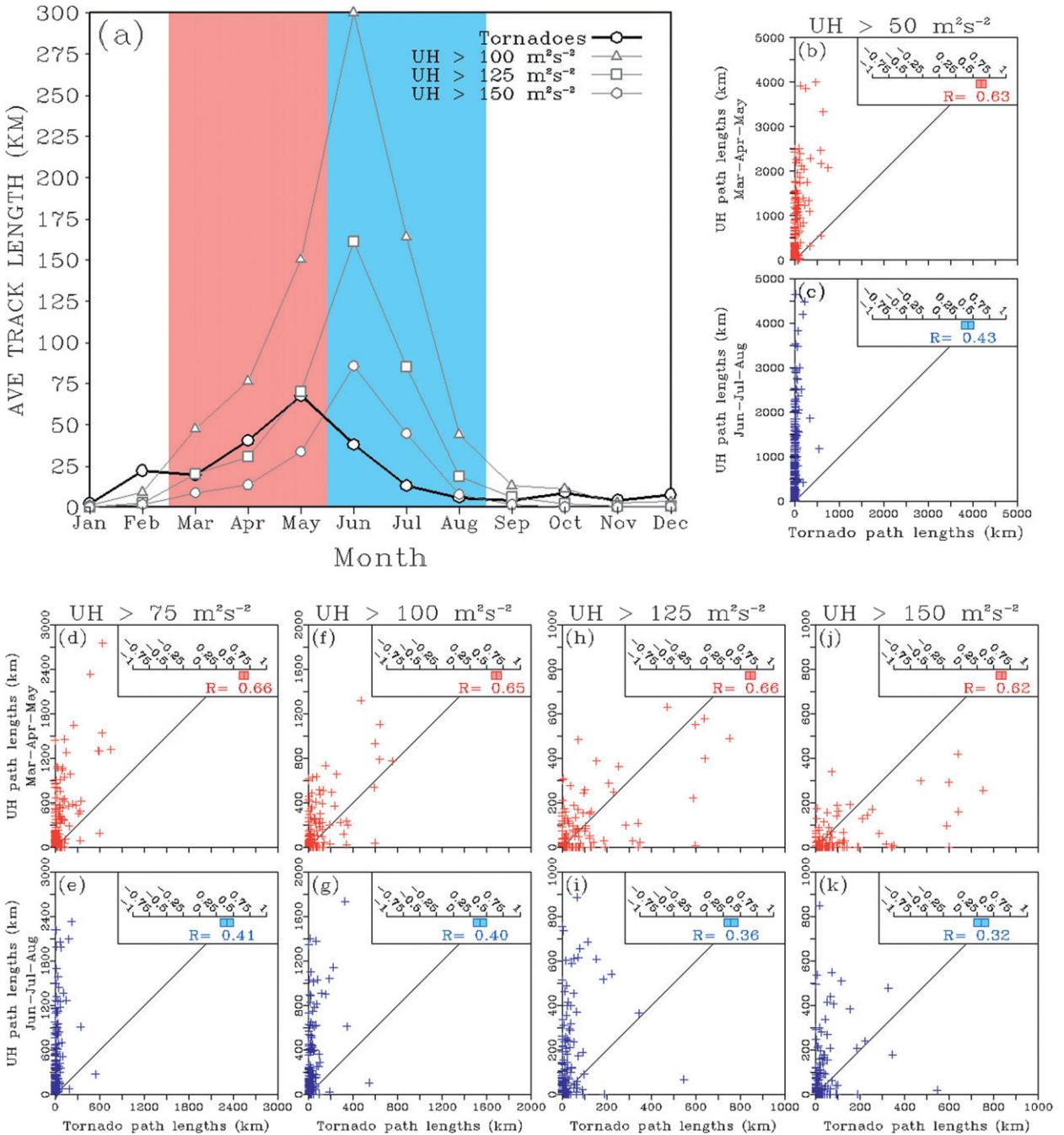
FIG. 5. (a) Average monthly tornado pathlengths and UH pathlengths for thresholds 100, 125, and 150 $m^2 s^{-2}$ over the 3-yr analysis period. The pink shading highlights the months March–May and the light blue shading highlights June–August. (b) Scatterplot of daily forecast UH pathlength using the threshold 50 $m^2 s^{-2}$ vs the total tornado pathlength for the months March–May. The corresponding correlation coefficient is shown in the top-right embedded scale along with 95% confidence intervals. (c) As in (b), but for the months June–August. Rest of panels as in (b),(c), but for UH thresholds of (d),(e) 75, (f),(g) 100, (h),(i) 125, and (j),(k) 150 $m^2 s^{-2}$, respectively.

suspected that UH produced by MCS-related phenomenon in NSSL-WRF forecasts may largely explain the peak forecast UH pathlengths during June. Because UH is often associated with nontornadic severe weather (e.g., Sobash et al. 2011), the June peak is consistent with

the climatological peak of observed nontornadic severe weather in June (e.g., Kelly et al. 1985; Clark et al. 2009). Further work, beyond the scope of this study, is needed to verify the cause for such high relative forecast UH pathlengths during June. Note, for the 3-yr period considered

in this study, February contains a small relative maxima because the Super Tuesday outbreak 5 February 2008 was such an extreme outlier.

To examine how strongly daily total UH and tornado pathlengths were correlated, Pearson correlation coefficients were computed for daily pairs of UH and tornado pathlengths for each UH threshold (Figs. 5b–k). The correlation coefficients were computed separately for the periods March–May and June–August because these two periods were distinctly different in terms of the relationship between monthly average UH and tornado pathlength (Fig. 5a). Furthermore, these two periods separate spring and summer and cover the months with the highest climatological tornado frequencies. The 95% confidence intervals for the correlation coefficients are computed using the cor.test(·) function in the R statistical software package (R Development Core Team 2011). This function computes an asymptotic confidence interval based on Fisher's Z transform. For testing whether statistically significant differences exist between various dependent correlation coefficients, the r.test(·) function is used. This function follows procedures recommended by Steiger (1980) for comparing elements of a correlation matrix.

For March–May (Figs. 5b, 5d, 5f, 5h, and 5j), correlation coefficients were between 0.62 and 0.66 for the five UH thresholds considered. The scatterplots of daily UH versus tornado pathlengths show that the UH threshold 150 m^2 s^{-2} generally underestimates total tornado pathlengths, while UH thresholds below 125 m^2 s^{-2} generally overestimate them. For June–August (Figs. 5c, 5e, 5g, 5i, and 5k), correlation coefficients are between 0.32 and 0.43. For each UH threshold, the differences in correlations between March–May and June–August are highly significant. Thus, forecast UH pathlengths are a more skillful predictor of tornado pathlengths during spring compared to summer.

To potentially improve the correlations, a method was devised to distinguish UH produced by simulated storms that are likely high based and/or elevated and, thus, have a lower probability of producing tornadoes (e.g., Rasmussen and Blanchard 1998; Grant 1995; Horgan et al. 2007). For determining whether a UH track is produced by a high-based storm, the height of the lifting condensation level (HLCL) computed using surface-based parcels is used. Using storm proximity soundings, Rasmussen and Blanchard (1998) found that about 90% of significant tornadoes [i.e., those rated as category 2 on the enhanced Fujita scale (EF2) and stronger] were associated with HLCL ≤ 1500 m. Later studies (e.g., Thompson et al. 2003) have verified HLCL as a strong discriminator between tornadic and nontornadic supercells. To determine an average value of HLCL

associated with UH objects, the average HLCL over all the grid points within a UH object present during a particular hour is computed for the previous hour. Thus, we basically sample the model environment covered by the UH object 0–1 h before the UH object is present. Individual hours are considered separately to allow for simulated storms that transition between high based and non–high based during their lifetime. A value of HLCL > 1500 m is used to discriminate high-based storms. Tracks or portions of tracks associated with high-based storms based on this criteria are excluded from the computation of the total UH track length.

To discriminate UH tracks produced by elevated storms, the ratio of surface-based convective available potential energy (SBCAPE) to most unstable convective available potential energy (MUCAPE) is used. If SBCAPE is much less than MUCAPE, it is assumed that a stable near-surface layer exists and the effective storm inflow layer is elevated. This assumption is often used by SPC forecasters in determining whether convection is likely to be elevated and works quite well (S. Weiss 2011, personal communication). A value of SBCAPE/MUCAPE < 0.75 is used to discriminate UH tracks associated with elevated storms. This value was chosen somewhat arbitrarily, but appeared to effectively distinguish environments for a few selected cases with elevated instability north of warm fronts in which simulated elevated rotating storms were present. The average SBCAPE/MUCAPE ratio was computed the in same way as for HLCL and, similar to HLCL, tracks or portions of tracks associated with elevated storms are excluded from the computation of the total UH track length.

To examine whether filtering UH tracks associated with high-based and/or elevated simulated storms improved the correlation of total UH and tornado track lengths, the analysis of monthly statistics and correlations in Fig. 5 is repeated in Fig. 6, but the elevated and/or high-based portions of UH tracks are excluded from the computation of total UH track lengths. The filter reduces monthly average UH pathlengths (Fig. 6a), and results in the UH threshold 100 m^2 s^{-2} matching best with total tornado pathlengths during spring, rather than 125 m^2 s^{-2}, which matched best during spring without the filter (Fig. 5a). Pearson correlation coefficients were slightly improved during spring for the UH thresholds that on average matched best with total tornado pathlengths: 75, 100, and 125 m^2 s^{-2} (Figs. 6d, 6f, and 6h). However, the improvements in correlations were significant (just barely) only for the 125 m^2 s^{-2} threshold. Furthermore, correlations were slightly *reduced* during summer (Figs. 6c, 6e, 6g, 6i, and 6k) after applying the filter, but none of the differences during summer were significant.
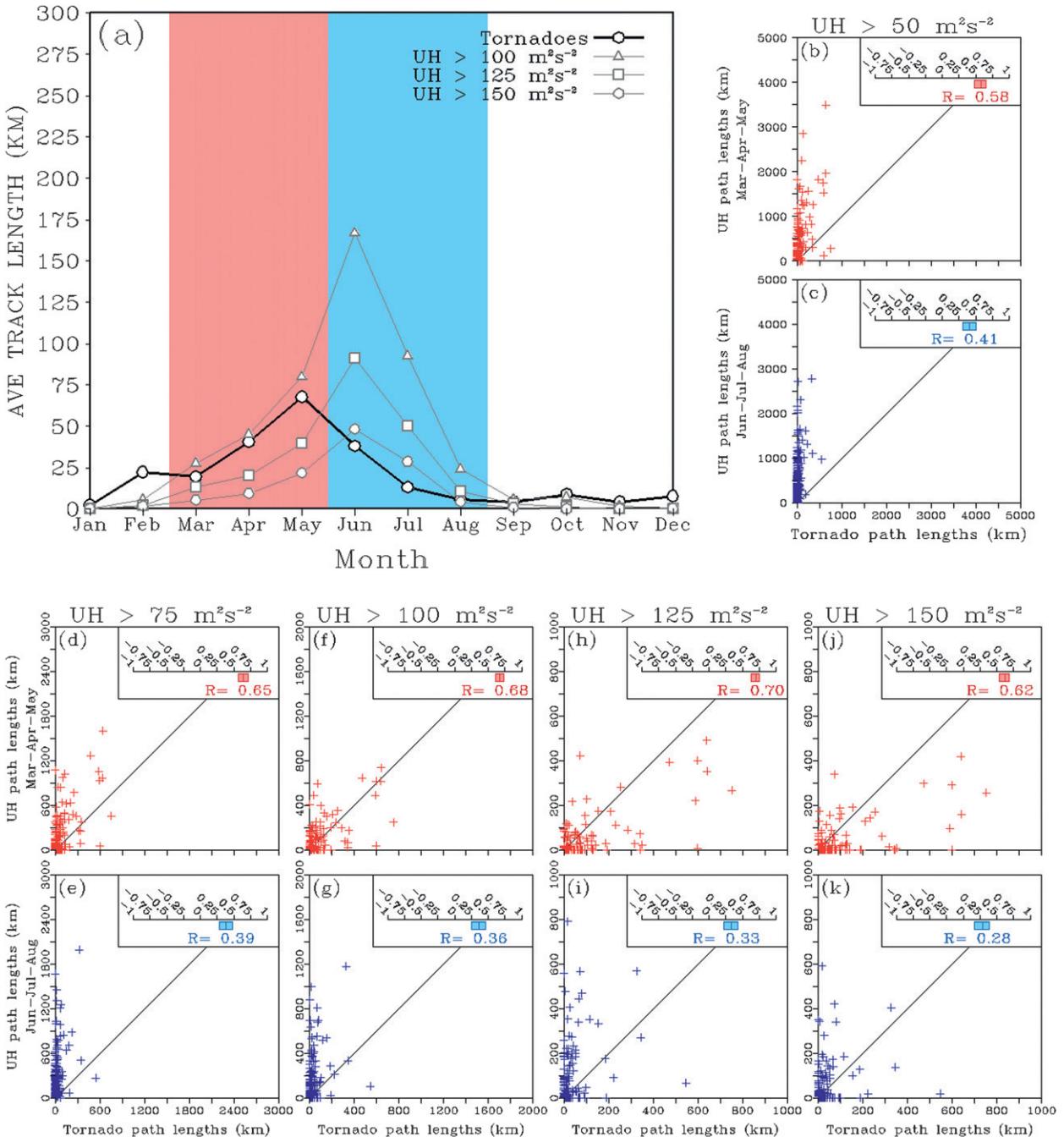
FIG. 6. As in Fig. 5, but for UH paths associated with elevated and/or high-based storms are not considered.

It is not clear why filtering the UH tracks produced by high-based and/or elevated simulated storms did not result in more significant improvements in the relationship with tornado pathlengths. Perhaps NSSL-WRF improperly simulated the severe weather environment in some cases, or perhaps non–supercell storm modes in environments supportive of surface-based convection with low HLCLs produced many long UH tracks, which

would not have been filtered out by the algorithm. Clearly, more detailed analyses are needed to explore this issue further.

c. Example ensemble forecasts

Similar to deterministic forecasts, for the ensemble UH object attributes to be useful to forecasters, information on the key object attributes should be

presented intuitively for efficient interpretation. For ensembles, effective graphical displays of object attributes are more challenging because of the extra dimension (i.e., multiple members) relative to deterministic forecasts. We present object attributes from each ensemble member for SSEF system forecasts initialized at 0000 UTC on 10, 19, and 6 May 2011 (Figs. 7, 9, and 10 respectively; see figure captions for details). In addition to the added dimension of ensemble members, these figures also distinguish UH tracks produced by high-based and/or elevated storms according to the criteria outlined in the previous section. These three cases were chosen because they depict a range of distinctly different severe weather events involving rotating storms.

1) 10 MAY 2010

The 10 May 2010 case had the longest total tornado pathlength during the analysis period for the SSEF system (599 km). The synoptic regime reflected a classic pattern for a southern plains tornado outbreak with a high-amplitude, negatively tilted, midlevel short-wave trough quickly moving from southwest to northeast over the region, with an associated sharp dryline arcing from southwestern Kansas through central Oklahoma and into west-central Texas (not shown). The strong environmental indicators led the SPC to issue a "high risk" for severe weather. Indeed, numerous long-track tornadoes occurred, mainly over central and eastern Oklahoma, with the strongest rated EF4. In Fig. 7a the maximum UH from any ensemble member over the 13–30-h forecast period is shown, with different shading schemes separating UH tracks produced by high-based and/or elevated storms and observed tornado locations overlaid. The ensemble UH tracks highlighted the main region where tornadoes occurred and indicated that storms producing UH in this region would not be high based and/or elevated.

Interestingly, the distribution of UH pathlengths for this case was bimodal, which was especially apparent for UH thresholds of 75 m^2 s^{-2} and greater (Figs. 7c–f). One group of members forecast multiple long and intense UH tracks, while another group forecast a few short and weak UH tracks. The NSSL-WRF forecast for this case (included in Figs. 7b–f) was similar to the members with short and weak UH tracks. The members with the short/weak UH tracks did not produce strong/long-lived storms within the unstable and strongly sheared environment most favorable for supercells, while the members with the long/intense UH tracks did produce strongly rotating storms in this favorable region. It is not yet clear why some of the members did not produce strongly rotating storms. The predicted synoptic features were very similar among all the members

with respect to the intensity and location of the dryline and midlevel short-wave trough. Additionally, even though all members depicted a strong elevated mixed layer that capped surface-based convection during morning and early afternoon, all members eventually contained large areas in which the cap had completely eroded by late afternoon immediately ahead of the dryline.

The boundary layer parameterization seemed to be the aspect of the ensemble member configurations most closely related to whether convection initiated. Out of the eight members that did not produce strongly rotating storms, seven used either the Mellor–Yamada–Janjić (MYJ) or quasi-normal scale elimination (QNSE) parameterization. Note, QNSE defaults to MYJ in unstably stratified boundary layers; thus, QNSE can be largely considered "MYJ driven." Out of the six members that did produce strongly rotating storms, five used either the Mellor–Yamada Nakanishi and Niino (MYNN) or Yonsei University (YSU) schemes. Overlaying 2000 UTC forecast soundings for the point 35°N, −97°W with a Verification of the Origins of Rotation in Tornadoes Experiment (VORTEX2; Wurman et al. 2010) sounding (Fig. 8) taken at the same time and only a few miles away, clearly shows a systematic difference in the depiction of the boundary layer between these two sets of members. The MYJ/QNSE members have boundary layers too cool and shallow relative to the observations, while the MYNN/YSU members have boundary layers that are generally deeper and warmer than the MYJ/QNSE members.[3] It is not clear how (or even if) the different boundary layer characteristics were that determined whether convection initiated, but the large differences are rather striking. Further work is planned to address the characteristics of different boundary layer schemes in convection-allowing simulations, but it is beyond the scope of this study.

For the members that produced strongly rotating storms, the total pathlength of tornadoes for this event matched very well with the portion of total UH pathlengths that were not elevated/high based using the UH threshold of 100 m^2 s^{-2} (red shaded portion of bars in Fig. 7d). This correspondence was encouraging because the 100 m^2 s^{-2} threshold was, on average, the most reliable predictor of total tornado pathlengths for NSSL-WRF during spring (Fig. 6a). The uncertainty information inherent in the ensemble would have suggested

---

[3] The IC–LBC perturbations and land surface models also contribute to differences in the boundary layer structure, but the systematic differences appear to emanate from the PBL schemes. This systematic influence is supported by an analysis of four SSEF members with only different PBL schemes (not shown), which basically mirror the behavior of the different subsets in Fig. 8.
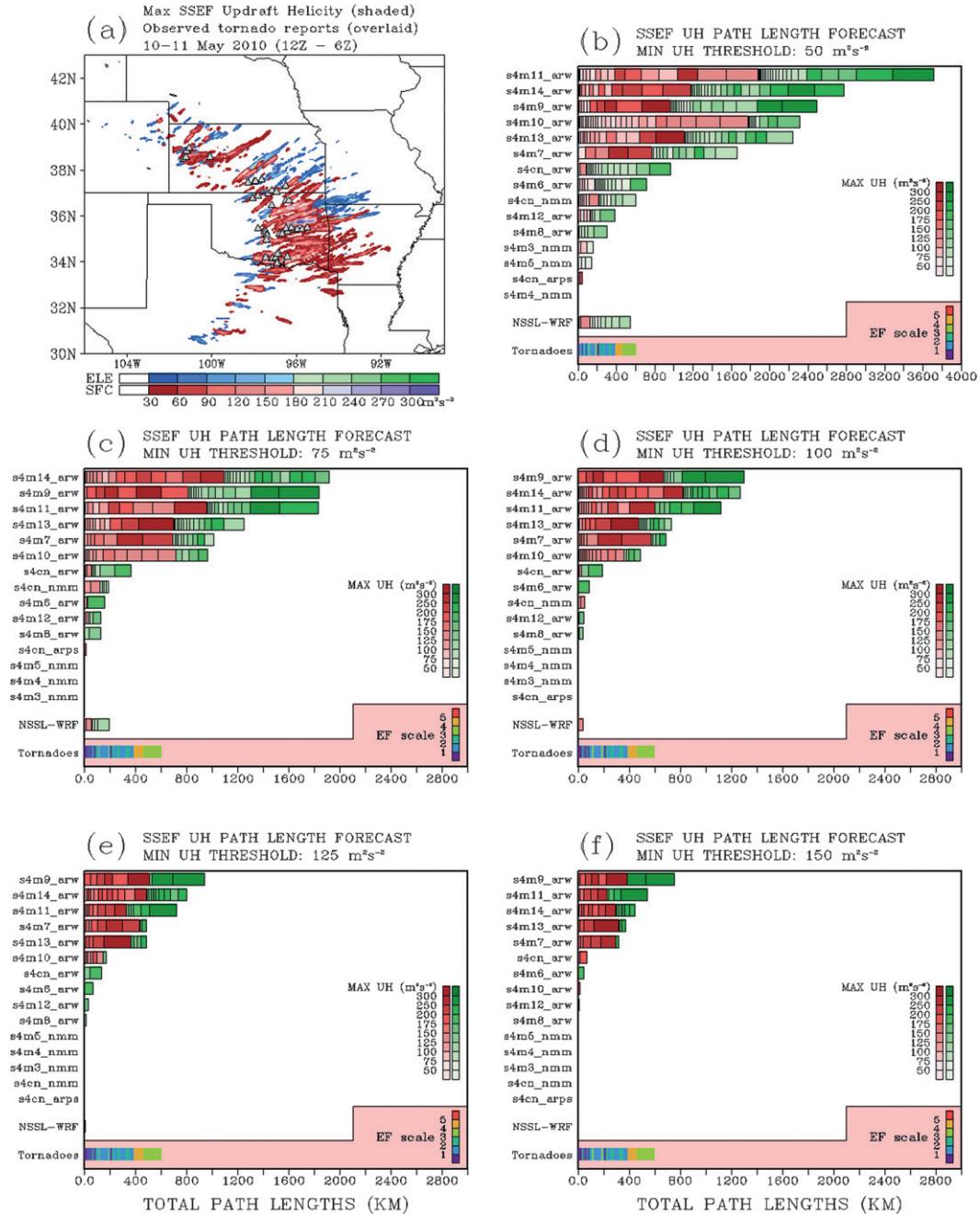
FIG. 7. (a) Maximum UH from any SSEF system member and NSSL-WRF initialized at 0000 UTC 10 May 2010 for forecast hours 13–30 (valid 1200–0600 UTC 10–11 May). The red/purple shading scheme is for UH produced by surface-based storms, while the blue/green shading scheme is for UH produced by elevated and/or high-based storms. Tornado reports (gray triangles) for the corresponding period are overlaid. (b) Total length of UH objects for each SSEF member using a minimum threshold of 50 m^2 s^{-2}. The length of the individual colored bars that compose each column indicate 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 surfaced-based storms and green shades to objects produced by elevated and/or high-based storms (color bars provided on 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. (c)–(f) As in (b), but for UH thresholds (c) 75, (d) 100, (e) 125, and (f) 150 m^2 s^{-2}.
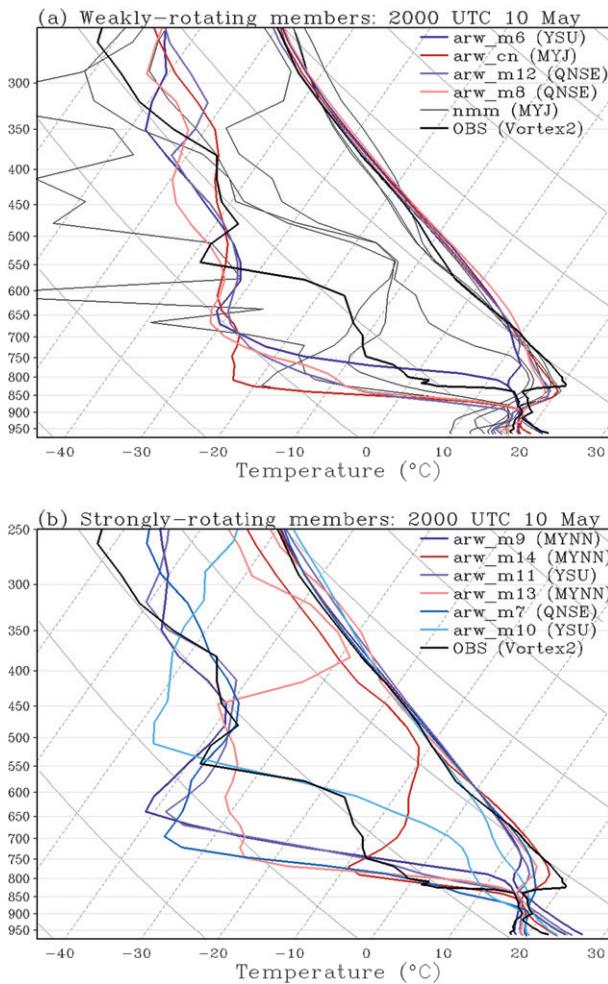
FIG. 8. (a) Forecast SSEF member soundings at 2000 UTC (20-h forecast) 10 May 2010 for the members that did not produce strongly rotating storms, and an observed sounding taken for VORTEX2 at the corresponding time just a few miles away. (b) As in (a), but for the SSEF members that did produce strongly rotating storms. Locations for the observed and forecasts soundings are 35.435°N, −97.001°W and 35.000°N, −97.000°W, respectively, in central OK.

a strong conditional risk for rotating storms. In other words, if storms were to be initiated and sustained, they would have intense rotation and long tracks.

2) 19 MAY 2010

In some ways, the case on 19 May 2010 was similar to 10 May, but with slightly less extreme environmental conditions for severe weather. Nonetheless, this was also a case in which the SPC issued a high risk for severe weather over Oklahoma. The synoptic set up featured a short-wave trough moving from west to east that helped to intensify a dryline over western Oklahoma during the afternoon. Rotating storms formed near and ahead of the

dryline and near a remnant outflow boundary in central and western Oklahoma. Numerous tornadoes occurred, but most were relatively weak and short lived. The most intense were rated EF1 and the total track length was about 100 km.

The maximum UH from each member during the 13–30-h forecast period (Fig. 9a) shows a concentrated region of UH tracks centered over eastern Oklahoma. Most of the observed tornadoes occurred within this region, except for a cluster of weak tornadoes in southwest Kansas associated with the cold-core center of an upper low. In contrast to 10 May, the distribution of ensemble UH pathlengths was relatively uniform. For UH < 100 m^2 s^{-2} (Figs. 9b,c), virtually all members produced UH paths of variable lengths from non-high-based/nonelevated storms, suggesting rotating surface-based storms would certainly form and a capping inversion would not prevent convective initiation. For UH $\geq$ 100 m^2 s^{-2} (Figs. 9d–f), pathlengths were relatively short, especially compared to the members with strongly rotating storms on 10 May. For the UH threshold of 100 m^2 s^{-2} (Fig. 9d), the total tornado pathlength fell within the bottom half of the distribution of ensemble members.

3) 6 MAY 2010

This case featured a west–east-oriented warm front that developed and gradually lifted northward throughout the day from northern Oklahoma to central Kansas. South of the warm front, moisture was relatively limited and a strong capping inversion above the boundary layer along with a lack of dynamic forcing for ascent in the warm sector prevented convection from developing. North of the warm front in northern Kansas, a combination of isentropic ascent and height falls from a midlevel short-wave trough approaching from the west led to a series of strong elevated supercells and storm clusters that moved from west to east across northern Kansas from 0000 to 0600 UTC 7 May. These storms mainly produced large hail and severe wind gusts and one brief tornado was reported in eastern Kansas.

The main corridor of severe reports coincided very well with the maximum UH from any ensemble member for the 13–30-h forecast period (Fig. 10a). Furthermore, the object attributes (Figs. 10b–f) indicated that many of the members contained one or two very long UH tracks with intense maximum values. However, almost all of the UH tracks were marked as produced by storms that were high based and/or elevated, which matched very well with what actually occurred for this case. The high-based/elevated nature of the UH tracks would have strongly reinforced the idea that hail/winds would be the main severe threats for this case, rather than tornadoes.
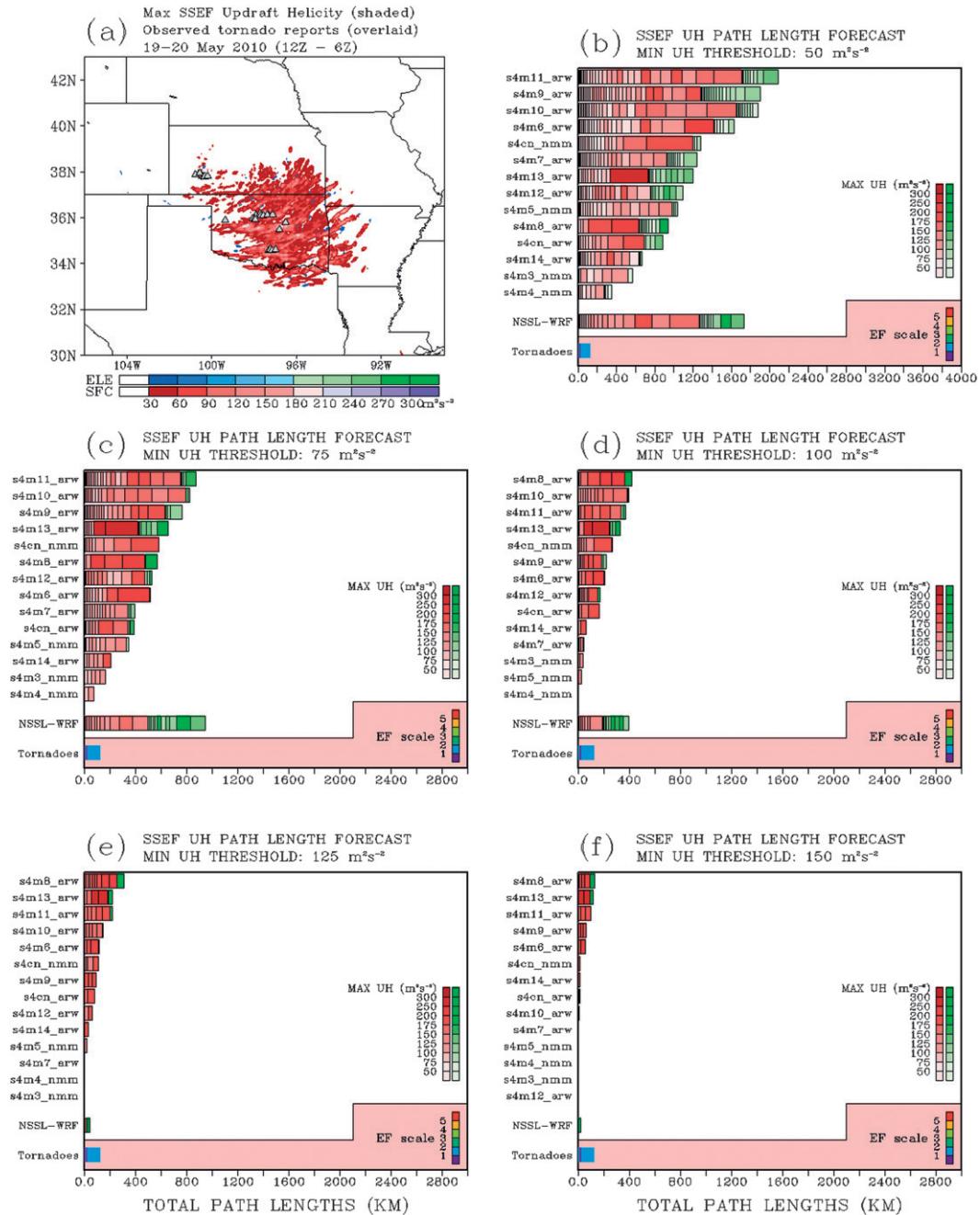
FIG. 9. As in Fig. 7, but for forecasts initialized at 0000 UTC 19 May 2010.

In summary, the ensemble UH track attributes provided unique information for each highlighted case. For 10 May, a strong conditional risk for severe weather was suggested for strong/long-track surface-based rotating storms; for 19 May, high certainty was implied, with most members suggesting relatively short tracks for surface-based rotating storms; and for 6 May 2010, the ensemble implied high certainty for intense/long-track rotating elevated storms. In the following section, a statistical analysis is presented covering all days during 2010 when the SSEF system was run.

d. Statistical analysis of 2010 SSEF system forecasts

Average daily UH pathlengths over all 34 cases during 2010 are computed for the mean of ARW and NMM members, as well as NSSL-WRF (Fig. 11a). For each threshold, the ARW members have the longest average UH pathlengths. The NSSL-WRF falls within the lower
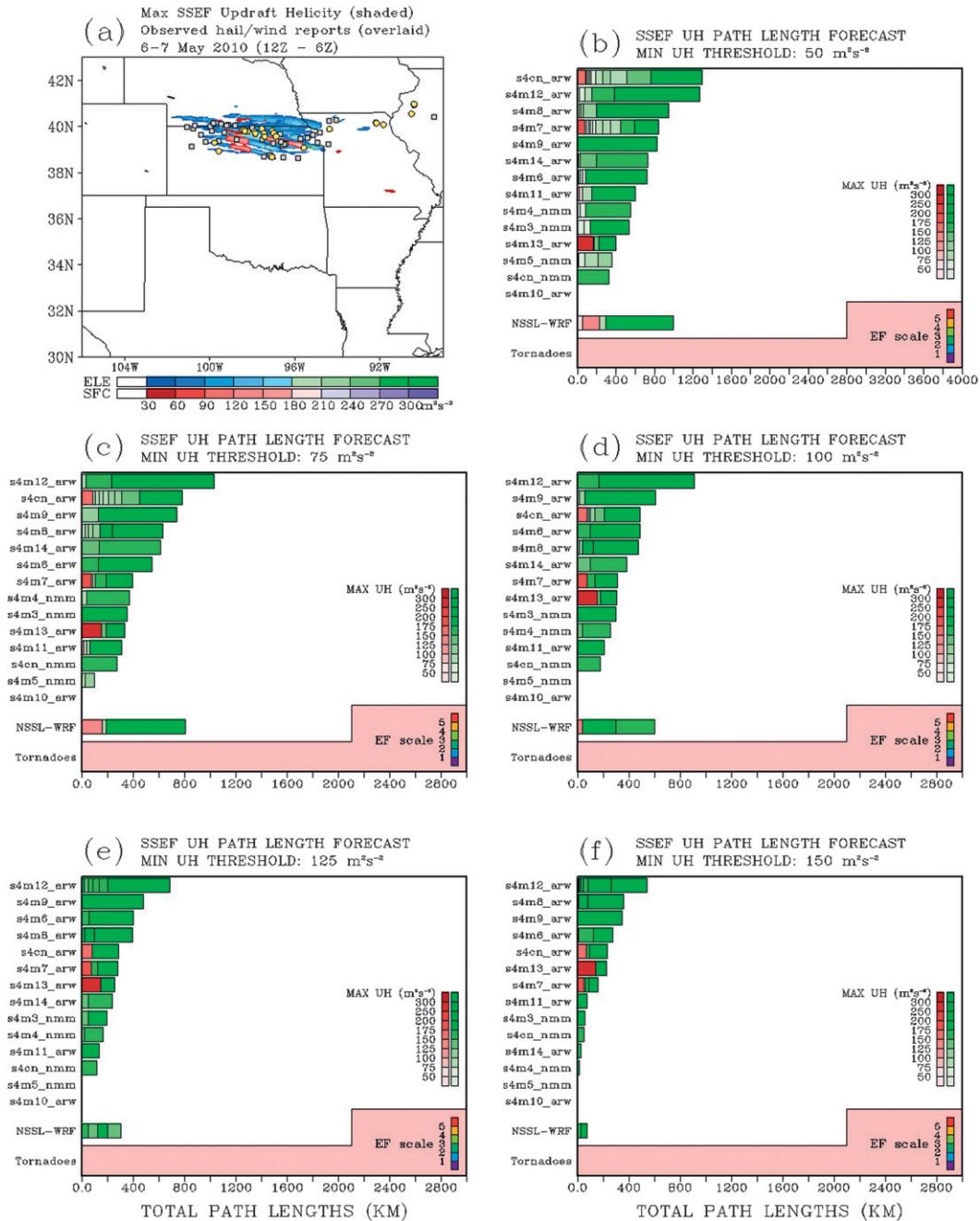
FIG. 10. As in Fig. 7, but for forecast initialized at 0000 UTC 6 May 2010.

range of ARW SSEF members (not shown), which is expected because NSSL-WRF also uses the ARW dynamics core. The NMM members have average pathlengths about half as long as the ARW members. After filtering out high-based and/or elevated storms, average pathlengths for ARW SSEF members and NSSL-WRF decrease by about 50%. The percent decrease for NMM members after applying the filter is less than for the ARW members, especially for UH ≥ 100 m^2 s^{-2}.

Similar to the NSSL-WRF climatology for March–May, the average filtered UH pathlengths for SSEF ARW members match the average tornado pathlengths most closely using a UH threshold of 100 m^2 s^{-2}.

To examine the relationship between ensemble UH and tornado pathlengths, Pearson correlation coefficients are computed (Figs. 11b–k). The NMM members are excluded from these analyses because they actually have a negative impact on the ensemble correlations
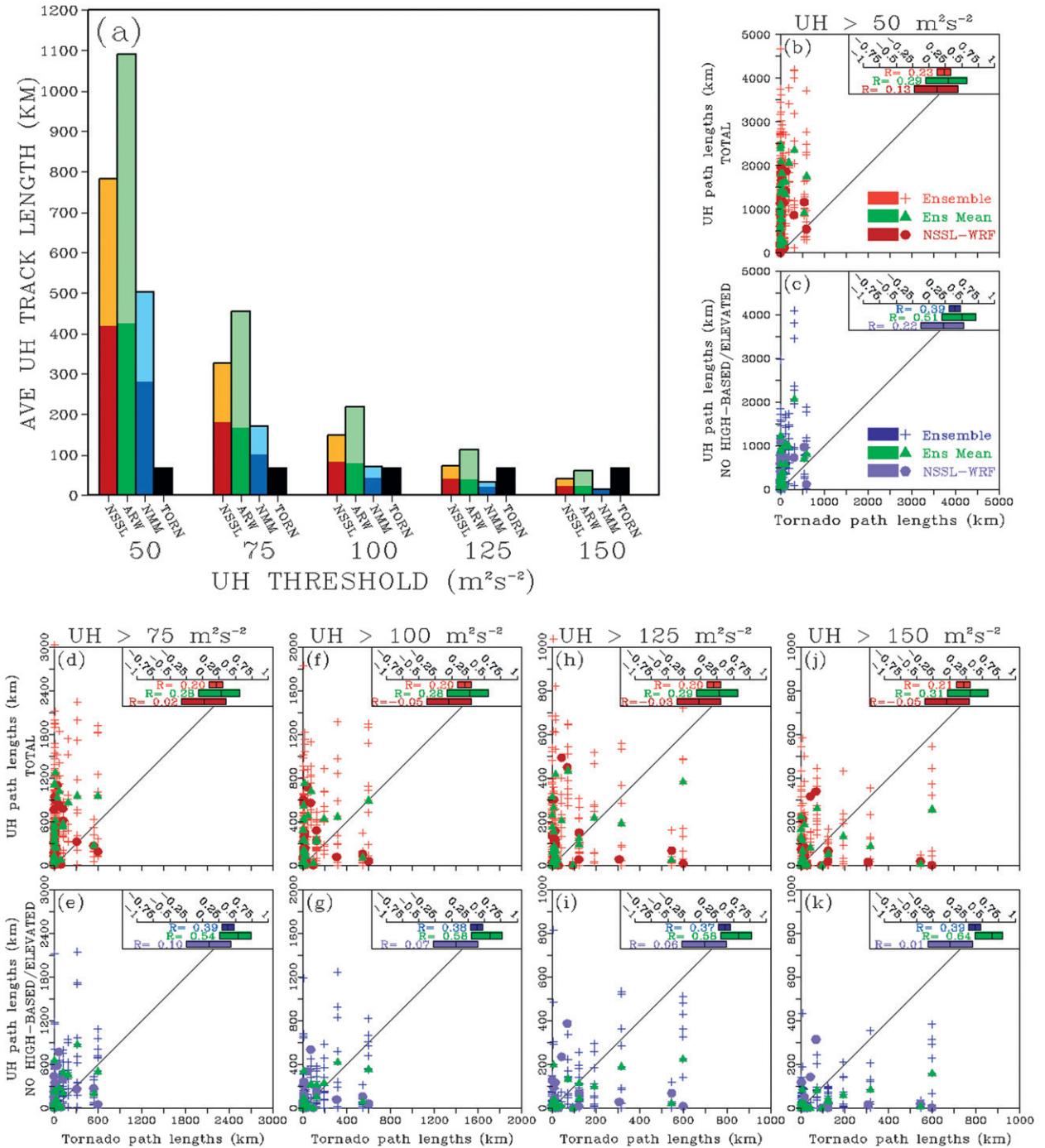
FIG. 11. (a) Average daily UH pathlengths for 50, 75, 100, 125, and 150 m² s⁻² thresholds from the NSSL-WRF (red/orange bars), and means of the SSEF system ARW members (green bars) and NMM members (blue bars). Average tornado pathlengths are indicated by the black histogram bars. The averages cover the 34 cases during April–June 2010 for which the SSEF system forecasts were available. For each model or group of models, the total height of each histogram bar includes portions of UH tracks identified as high based and/or elevated, while the height of the dark-colored bars does not. (b) Scatterplot of daily forecast UH pathlength using the 50 m² s⁻² threshold vs total tornado pathlength for all 34 cases covered by the SSEF system forecasts. Points are included for individual ARW SSEF members, the mean of the ARW members, and NSSL-WRF. A legend is provided in the bottom right. Corresponding correlation coefficients are shown in the top-right embedded scale along with 95% confidence intervals. (c) As in (b), but for UH pathlengths where portions of tracks high-based and/or elevated are excluded. Rest of panels as in (b),(c), but for UH thresholds of (d),(e) 75, (f),(g) 100, (h),(i) 125, and (j),(k) 150 m² s⁻², respectively.

(not shown).[4] Correlations are computed by considering each pair of ensemble member–observation values (e.g., for 34 cases and 10 ARW ensemble members, $34 \times 10 = 340$ observation–model pairs of values are used). Correlations are also computed using ensemble mean UH pathlength, resulting in one pair of observation–model values for each case, and for the NSSL-WRF forecasts that cover the same 34 cases (Figs. 11b–k). The results are presented separately for total UH pathlengths (top row of panels), and UH pathlengths computed after filtering out track segments from high-based and/or elevated storms (bottom row of panels).

For each UH threshold, the ensemble mean correlation with total tornado pathlengths is higher than that computed using each ensemble member–observation pair. For example, at the $125 \text{ m}^2 \text{ s}^{-2}$ UH threshold (Fig. 11h), the ensemble correlation is 0.20 while the ensemble mean correlation is 0.29. The improvement of the ensemble mean correlation over that of the ensemble is more dramatic after filtering high-based and/or elevated portions of UH track segments. Again, considering the $125 \text{ m}^2 \text{ s}^{-2}$ UH threshold (Fig. 11i), the ensemble correlation is 0.37 while the ensemble mean correlation is 0.58. The improvement from using the ensemble mean is reassuring and implies that more uncertain aspects of the ensemble member forecasts are being "filtered out" by taking the mean (e.g., Kalnay 2003).

Additional improvements in correlation with total tornado pathlengths, especially for the ensemble mean, are achieved through filtering out the high-based/elevated portions of UH track segments. For example, after filtering high-based/elevated UH track segments using the $125 \text{ m}^2 \text{ s}^{-2}$ threshold, ensemble mean correlations improve from 0.29 to 0.58 (Figs. 11h,i). The differences between the ensemble mean filtered versus nonfiltered correlations were significant for UH thresholds $> 50 \text{ m}^2 \text{ s}^{-2}$. The UH versus tornado pathlength scatterplots for total and filtered UH pathlengths suggest that the improvement in correlation for the filtered UH pathlengths results from a downward shift in the location of points in the far-left portion of the scatterplots. In other words, for the cases in which the total tornado pathlengths were very short but many of the members forecast long UH pathlengths, the forecast UH pathlengths were dramatically reduced by filtering out the tracks from high-based/elevated storms.

Another notable result is the apparent poor relative performance of NSSL-WRF. Even after filtering UH track segments from high-based/elevated storms, the NSSL-WRF correlations were very low, ranging from 0.01 to 0.22 for the different thresholds examined. For the filtered UH track segments, the ensemble mean correlations were higher with statistical significance than NSSL-WRF for each threshold examined except $50 \text{ m}^2 \text{ s}^{-2}$. Compared to the correlations from the much larger set of NSSL-WRF cases examined in section 3b, these correlations are very low, so it is clear that this particular set of cases was not representative of typical NSSL-WRF performance. Furthermore, because the set of cases for which the SSEF system was run is relatively small and covers only a few days with very long tornado pathlengths, a bad forecast for one of the long tornado pathlength days could have a large negative impact on the overall correlations. Indeed, for the tornado outbreak on 10 May 2010, which had a total tornado pathlength of 599 km, NSSL-WRF had a total UH pathlength near zero for the $125 \text{ m}^2 \text{ s}^{-2}$ UH threshold. Similarly, for the case with the third highest total tornado pathlength (30 April 2010), the NSSL-WRF UH pathlengths were very short. For both of these cases, NSSL-WRF UH pathlengths were within the range of SSEF ARW members, but near the lower end of the distribution. Furthermore, there were two or three (depending on the threshold examined) ARW SSEF members with lower overall correlations than NSSL-WRF (not shown). Thus, the poor relative performance of NSSL-WRF should be interpreted with caution and is more of an indication of the advantages of using the ensemble.

Because the analyses of deterministic forecasts from NSSL-WRF found significant differences in the relationship between total UH and tornado pathlength for the March–May period compared to June–August, the analysis for the ensemble is repeated including only the April–May cases and excluding the 14 June cases (Fig. 12). The comparisons of average daily UH pathlengths for the different sets of members (Fig. 12a) yielded results similar to those including all cases. However, the correlations for the ensemble and ensemble mean were much higher for only the April–May cases, and the differences were more dramatic after filtering the UH track segments from high-based/elevated storms. For example, for total UH pathlengths at the $100 \text{ m}^2 \text{ s}^{-2}$ UH threshold, ensemble mean correlations were 0.28 (Fig. 11f) and improved to 0.46 after excluding June cases (Fig. 12f). For the filtered UH pathlengths at the same threshold, ensemble mean correlations were 0.58 (Fig. 11g) and improved to 0.84 after excluding June cases (Fig. 12g). For the set of cases excluding June, the ensemble mean correlations were surprisingly high—up to 0.91 using the filtered UH track

---

[4] The negative impact is likely at least partially a result of systematic differences in average UH pathlengths (Fig. 11a). However, examination of ARW and NMM members separately also revealed that NMM members tended to have weaker correlation to tornado pathlengths.
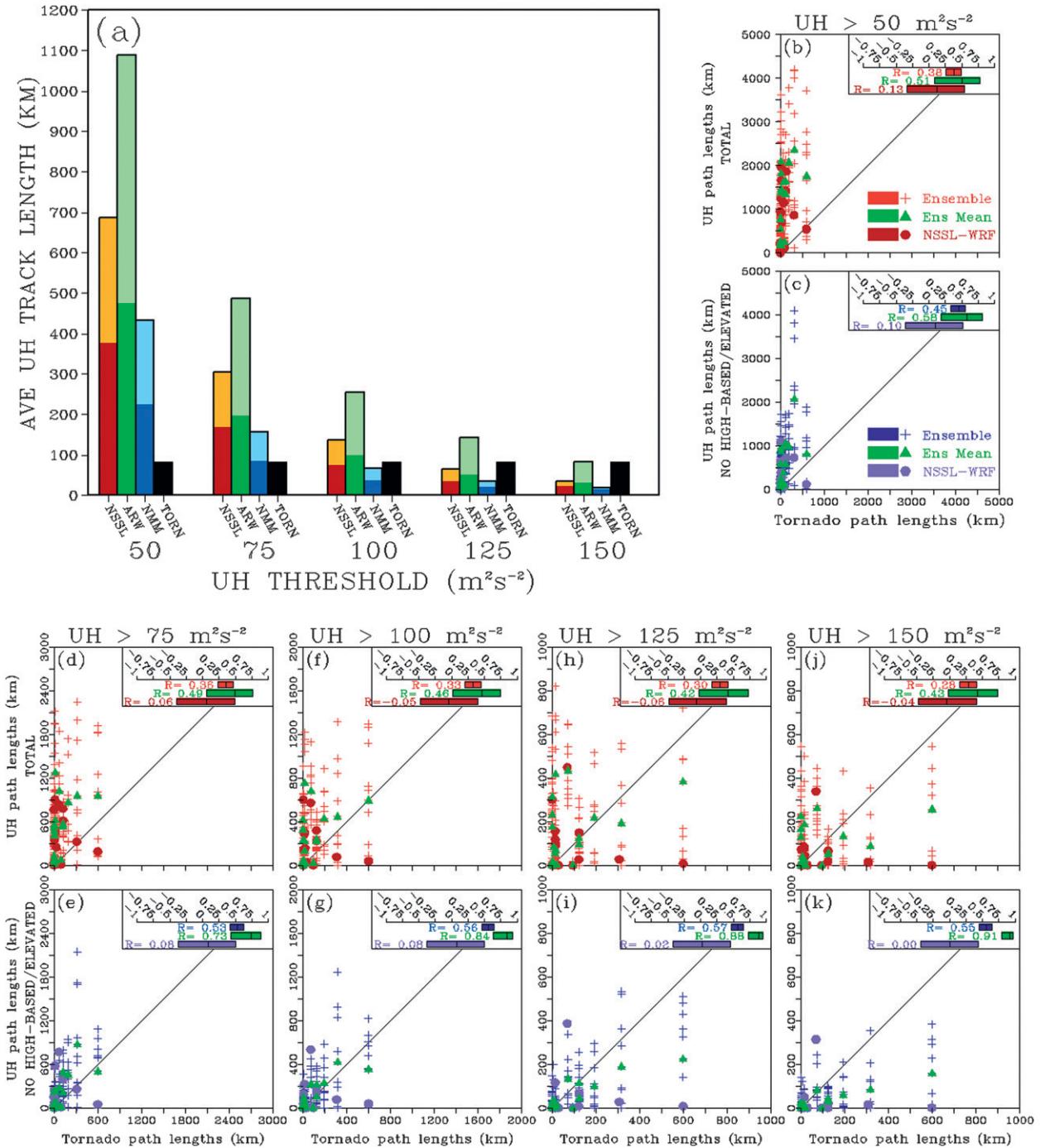
FIG. 12. As in Fig. 11, but only for the 20 cases covered by SSEF system forecasts during April and May 2010 (i.e., the 14 cases in June are excluded).

lengths for the 150 m² s⁻² threshold. These results suggest the ensemble mean UH pathlengths could potentially be a very skillful predictor for the overall severity of a tornado outbreak as measured by the total tornado pathlength. Certainly, it will be desirable to verify these results with a much larger set of cases.

4. Summary and conclusions

A 3D object identification algorithm is used to measure the track length of simulated rotating storms in convection-allowing forecast models. Simulated rotation is quantified in terms of hourly maximum updraft

helicity (UH), and various UH thresholds are used to identify objects. Inclusion of the time dimension in the object algorithm allows tracks to span multiple forecast hours and encompass entire life cycles of simulated rotating storms. Methods for visualizing UH object attributes from the convection-allowing forecasts that quickly convey useful information to forecasters are presented. In addition, the relationship of total UH track length to total tornado pathlengths over 24-h periods is documented using Pearson correlation coefficients. Deterministic UH forecasts from a 4-km grid-spacing version of the WRF-ARW (NSSL-WRF) for the time period 2008–10 are examined, as are ensemble UH forecasts from the SSEF system run by CAPS for the 2010 NOAA Hazardous Weather Testbed Spring Forecasting Experiment.

Analysis of NSSL-WRF forecasts during 2008–10 revealed significant differences in the strength of the relationship between total UH and tornado pathlengths during spring (March–May) compared to summer (June–August), with the relationship during spring being much stronger. To improve the correlations, a method was devised to filter out segments of UH tracks likely produced by simulated storms that are high based and/or elevated, with the idea that these storms would be unlikely to produce tornadoes. For the deterministic forecasts during spring, this filter slightly improved the correlations with only one UH threshold having significantly different correlations relative to the nonfiltered UH tracks. For the summer, applying the filter did not result in any significant differences.

For the ensemble, three cases were highlighted that illustrate some of the potentially useful information provided by forecast object attributes. For the 10 May 2010 case, a bimodal distribution of total UH pathlengths implied a conditional risk for a tornado outbreak (i.e., if storms formed, they would be long track and strongly rotating); for 19 May 2010, the distribution of total UH pathlengths was quite uniform, indicating a high degree of certainty that there would be rotating storms, but the pathlengths were fairly short; and for 6 May 2010 there was a high degree of certainty for long-track strongly rotating, but elevated storms. Each of these cases verified very well.

For the statistical analyses covering all 34 ensemble cases in 2010, the correlations with total tornado tracks lengths computed using the ensemble mean improved over correlations computed using each ensemble member. Additionally, in contrast to the analysis of deterministic forecasts, filtering out high-based/elevated storms significantly improved the correlations. Consistent with the results from the deterministic forecasts, inclusion of the 14 cases from June in the ensemble

analyses was found to negatively impact the correlations. The best results were obtained for the April–May cases after filtering out UH paths from high-based/elevated simulated storms. Pearson correlation coefficients of the ensemble mean UH pathlengths with total tornado pathlengths were as high as 0.91 for these cases.

These results are very encouraging and illustrate the potential usefulness of 3D object identification algorithms within a forecasting context. However, more analyses of a much larger sample of cases are needed to verify some of the very high correlations that were obtained from the ensemble. Furthermore, it is acknowledged that the UH pathlength versus tornado comparisons are not ''apples to apples.'' Ideally, UH paths would be verified against observed mesocyclone paths, because not all rotating storms produce tornadoes. It is possible that the discrepancy in datasets explains the dropoff in the strength of the relationship between total UH pathlength and total tornado pathlength during summer—perhaps the summer months truly do have more storms that produce longer average rotation tracks—but the storms do not produce tornadoes as efficiently as springtime storms. Work is ongoing at NSSL to develop a product for verifying mesocyclone paths using the Vortex Detection and Diagnosis Algorithm (VDDA) within the NSSL WDSS-II software package (Lakshmanan 2002).

Future work will involve testing the displays of 3D object attributes during NOAA/HWT Spring Forecasting Experiments, as well as warn-on-forecast applications. We also plan to extend the application of 3D object algorithms to other diagnostics like hourly maximum vertically integrated graupel to document the relationship with observed hail swaths. In addition, because the 3D algorithm incorporates time, it is ideally suited for diagnosing convective initiation timing, which is becoming an emphasis in forthcoming NOAA/HWT Spring Forecasting Experiments.

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.

REFERENCES

Baldwin, M. E., S. Lakshmivarahan, and J. S. Kain, 2001: Verification of mesoscale features in NWP models. Preprints, *Ninth Conf. on Mesoscale Processes,* Fort Lauderdale, FL, Amer. Meteor. Soc., 255–258.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.,* **129,** 569–585.

Chen, S.-H., and W.-Y. Sun, 2002: A one-dimensional time dependent cloud model. *J. Meteor. Soc. Japan,* **80,** 99–118.

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., C. J. Schaffer, W. A. Gallus Jr., and K. Johnson-O'Mara, 2009: Climatology of storm reports relative to upper-level jet streaks. *Wea. Forecasting,* **24,** 1032–1051.

——, W. A. Gallus Jr., and M. L. Weisman, 2010: Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF model simulations and the operational NAM. *Wea. Forecasting,* **25,** 1495–1509.

——, and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.,* **93,** 55–74.

Colman, B. R., 1990: Thunderstorms above frontal surfaces in environments without positive CAPE. Part I: A climatology. *Mon. Wea. Rev.,* **118,** 1103–1122.

Coniglio, M. C., K. L. Elmore, J. S. Kain, S. J. Weiss, M. Xue, and M. L. Weisman, 2010: Evaluation of WRF model output for severe-weather forecasting from the 2008 NOAA Hazardous Weather Testbed Spring Experiment. *Wea. Forecasting,* **25,** 408–427.

Davis, C. A., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.,* **134,** 1772–1784.

——, ——, and ——, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.,* **134,** 1785–1795.

——, ——, and ——, 2009: Spatial and temporal object-based evaluation of numerical precipitation forecasts. Preprints, *23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction,* Omaha, NE, Amer. Meteor. Soc., 5A.4. [Available online at http://ams.confex.com/ams/pdfpapers/154156.pdf.]

Dixon, M., and G. Wiener, 1993: TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting—A radar-based methodology. *J. Atmos. Oceanic Technol.,* **10,** 785–797.

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.]

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

Ebert, E. E., and J. L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.,* **239,** 179–202.

Edwards, R., R. L. Thompson, K. C. Brosbie, J. A. Hart, and C. A. Doswell III, 2004: Proposals for modernizing the definitions of tornado and severe thunderstorm outbreaks. Preprints, *22nd on Conf. Severe Local Storms,* Hyannis, MA, Amer. Meteor. Soc., 7B.2. [Available online at http://ams.confex.com/ams/pdfpapers/81342.pdf.]

Fels, S. B., and M. D. Schwarzkopf, 1975: The simplified exchange approximation: A new method for radiative transfer calculations. *J. Atmos. Sci.,* **32,** 1475–1488.

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., 280–283

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.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting,* **24,** 1416–1430.

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.

Grant, B. N., 1995: Elevated cold-sector severe thunderstorms: A preliminary study. *Natl. Wea. Dig.,* **19** (4), 25–31.

Han, L., 2009: 3D convective storm identification, tracking and forecasting—An enhanced TITAN algorithm. *J. Atmos. Oceanic Technol.,* **26,** 719–732.

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.

Horgan, K. L., D. M. Schultz, J. E. Hales Jr., S. F. Corfidi, and R. H. Johns, 2007: A five-year climatology of elevated severe convective storms in the United States east of the Rocky Mountains. *Wea. Forecasting,* **22,** 1031–1044.

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, 61 pp.

Kain, J. S., P. R. Janish, S. J. Weiss, M. E. Baldwin, R. S. Schneider, and H. E. Brooks, 2003: Collaboration between forecasters and research scientists at the NSSL and SPC: The Spring Program. *Bull. Amer. Meteor. Soc.,* **84,** 1797–1806.

——, and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting,* **23,** 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.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability.* Cambridge University Press, 341 pp.

Kelly, D. L., J. T. Schaefer, and C. A. Doswell, 1985: Climatology of nontornadic severe thunderstorm events in the United States. *Mon. Wea. Rev.,* **113,** 1997–2014.

Lacis, A. A., and J. E. Hansen, 1974: A parameterization for the absorption of solar radiation in the earth's atmosphere. *J. Atmos. Sci.,* **31,** 118–133.

Lakshmanan, V., 2002: WDSSII: An extensible, multi-source meteorological algorithm development interface. Preprints, *21st Conf. on Severe Local Storms,* San Antonio, TX, Amer. Meteor. Soc., 134–137. [Available online at http://ams.confex.com/ams/pdfpapers/47541.pdf.]

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.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.,* **83,** 407–430.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.,* **20,** 851–875.

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.

Rasmussen, E. N., and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting,* **13,** 1148–1164.

R Development Core Team, cited 2011: 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.]

Schwarzkopf, M. D., and S. B. Fels, 1991: The simplified exchange method revisited—An accurate, rapid method for computation of infrared cooling rates and fluxes. *J. Geophys. Res.,* **96** (D5), 9075–9096.

Sinnott, R. W., 1984: Virtues of the haversine. *Sky Telesc.,* **68** (2), 159.

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.]

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.

Sobash, R. A., J. S. Kain, D. R. Bright, A. R. Dean, M. C. Coniglio, and S. J. Weiss, 2011: Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. *Wea. Forecasting,* **26,** 714–728.

Steiger, J. H., 1980: Tests for comparing elements of a correlation matrix. *Psychol. Bull.,* **87,** 245–251.

Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system. *Bull. Amer. Meteor. Soc.,* **90,** 1487–1499.

Stumpf, G. J., T. M. Smith, V. Lakshmanan, K. L. Manross, and K. D. Hondl, 2004: Status of multiple-sensor severe weather application development at NSSL. Preprints, *20th Int. Conf. on Interactive Information and Processing Systems for Meteor., Oceanography, and Hydrology,* Seattle, WA, Amer. Meteor. Soc., 12.3A [Available online at http://ams.confex.com/ams/pdfpapers/88063.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.

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.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.,* **125,** 527–548.

Wurman, J., L. Wicker, Y. Richardson, P. Markowski, D. Dowell, D. Burgess, and H. Bluestein, 2010: VORTEX2: The Verification

of the Origins of Rotation in Tornadoes Experiment. Preprints, *25th Conf. on Severe Local Storms,* Denver, CO, Amer. Meteor. Soc., 5.1. [Available online at http://ams.confex.com/ams/pdf/papers/176068.pdf.]

Xue, M., D. Wang, J. Gao, 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, 2008: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment. Preprints, *24th Conf. on Severe Local Storms,* Savannah, GA, Amer. Meteor. Soc., 12.2. [Available online at http://ams.confex.com/ams/pdfpapers/142036.pdf.]

——, 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.]