# Explicit Forecasts of Low-Level Rotation from Convection-Allowing Models for Next-Day Tornado Prediction

RYAN A. SOBASH, GLEN S. ROMINE, AND CRAIG S. SCHWARTZ

*National Center for Atmospheric Research, Boulder, Colorado*

DAVID J. GAGNE II

*National Center for Atmospheric Research, Boulder, Colorado, and Center for Analysis and Prediction of Storms/School of Meteorology, University of Oklahoma, Norman, Oklahoma*

MORRIS L. WEISMAN

*National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 13 April 2016, in final form 19 July 2016)

ABSTRACT

Three diagnostic fields were examined to assess their ability to act as surrogates for tornadoes in a convection-allowing ensemble system run during the spring of 2015. The diagnostics included midlevel (2–5 km AGL) updraft helicity (UH25), low-level (0–3 km AGL) updraft helicity (UH03), and low-level (1 km AGL) vertical relative vorticity (RVORT1). RVORT1 was used as a direct measure of low-level rotation strength. Each storm's RVORT1 magnitude and near-storm environment properties were extracted from each hour's forecasts using an object-based approach. The near-storm environments of storm objects with large magnitudes of RVORT1 were very similar to the environments identified as conducive for the development of tornadic supercells in previous proximity sounding-based studies (e.g., low lifted condensation levels and strong low-level shear). This motivated the use of RVORT1 as a direct surrogate for tornadoes, without the need to filter forecasts with environmental information. The relationship between UH25 and UH03 was also explored among the simulated storms; UH03 only exceeded UH25 in storms occurring within low-CAPE/high-shear environments, while UH03 rarely exceeded UH25 in traditional supercell environments. Next-day ensemble surrogate severe probability forecasts (E-SSPFs) for tornadoes were generated using these diagnostics for 92 forecasts, with thresholds based on the number of observed tornado reports. E-SSPFs for tornadoes using RVORT1 and UH03 were more skillful than E-SSPFs using UH25. The UH25 E-SSPFs possessed little skill, regardless of threshold or smoothing length scale. All E-SSPFs suffered from poor sharpness at skillful scales, with few forecast probabilities greater than 40%.

## 1. Introduction

Convection-allowing models (CAMs) have permitted the development of novel forms of forecast guidance summarizing the potential for severe convective weather based on their explicit predictions of convective storms (e.g., Kain et al. 2008; Sobash et al. 2011, 2016; Clark et al. 2012, 2013). Identifying supercells in CAM forecasts has garnered particular interest, as a result of the disproportionate share of severe weather reports produced by these intense, damaging storms (Gallus et al. 2008). Kain et al. (2008) proposed computing updraft helicity (UH), the product of vertical velocity and vertical vorticity over a specified depth, to identify midlevel (2–5 km AGL) mesocyclones, and thus supercells, in CAM output. Later work has successfully used midlevel UH as a surrogate for the occurrence of severe weather hazards emanating from supercells [e.g., strong winds, large hail, and tornadoes; Sobash et al. (2011, 2016); Clark et al. (2013)] and further verified the ability of midlevel UH to reliably identify supercells in CAM output (e.g., Naylor et al. 2012).

While this work has demonstrated the usefulness of midlevel UH as a diagnostic for identifying simulated

*Corresponding author address*: Dr. Ryan A. Sobash, NCAR/MMM, P.O. Box 3000, Boulder, CO 80307.
E-mail: sobash@ucar.edu

DOI: 10.1175/WAF-D-16-0073.1

supercells, discriminating tornadic from nontornadic events using present-day CAMs presents a more substantial, and so far unaddressed, challenge. Observational studies of supercells routinely indicate that the presence of a midlevel mesocyclone, which midlevel UH was initially designed to detect, is not a sufficient condition for tornado occurrence (Trapp et al. 1999, 2005; Wakimoto et al. 2004). In fact, Trapp et al. (2005) estimated that only 15% of storms with midlevel mesocyclones produce tornadoes. Assuming modeled storms possess a similar relationship, then midlevel UH alone is not an appropriate surrogate for tornado occurrence, and an additional surrogate that is more closely related to tornadogenesis is needed. While directly sampling CAM output for the presence of tornado-like vortices[1] would be preferable to relying on surrogates, present-day CAMs have horizontal grid spacings between 1 and 4 km, where only storm-scale circulations, such as mesocyclones, are close to being resolved. Resolving tornado-like vortices requires CAMs with grid spacings well below 1 km (e.g., Schenkman et al. 2014), which will not be available in operational NWP systems for many years or, possibly, decades.

Despite the obstacles involved with using midlevel UH as a surrogate for tornado occurrence in CAM output, several recent studies have presented encouraging results using midlevel UH to anticipate next-day tornado potential. For example, Clark et al. (2012, 2013) used the total time- and space-integrated midlevel UH swath length from a CAM ensemble as a surrogate for observed tornado track length over a collection of cases in 2011 and 2012. A positive correlation existed between the forecast and observed tornado track length after removing midlevel UH tracks that were produced by elevated or high-based convection, as determined by CAM forecasts of lifted condensation level (LCL) and convective inhibition (CIN). While the forecasts examined in Clark et al. (2013) were for aggregated tornado pathlength on a given day, Jirak et al. (2014) and Gallo et al. (2016) used forecasts of midlevel UH, combined with similar environmental information (e.g., CAPE, LCL, significant tornado parameter) as in Clark et al. (2013), to produce probabilistic guidance akin to Storm Prediction Center probabilistic forecasts that was superior to guidance based strictly on midlevel UH. These results suggest that CAMs appear capable of providing skillful guidance for the next-day tornado threat, *only when forecasts of midlevel UH are combined with environmental parameters that skillfully discriminate tornado-prone environments.*

Filtering CAM forecasts of midlevel UH with environmental information reduces the overforecasting bias that plagues guidance solely based on midlevel UH as a surrogate for tornadoes (Gallo et al. 2016). The choices of environmental fields to use for filtering are informed by previous work that used proximity soundings to reveal relationships between characteristics of a storm's inflow environment and its propensity to produce low-level rotation and tornadoes (e.g., Brooks et al. 1994; Rasmussen and Blanchard 1998; Rasmussen 2003; Thompson et al. 2003, hereafter T03; Thompson et al. 2012). These studies largely omit tornadic storms occurring in nontraditional environments (e.g., low CAPE/high shear) or tornadoes that develop in quasi-linear convective systems (QLCSs) or tropical cyclones. For example, the T03 dataset excluded tropical cyclone supercells from their dataset. Having a surrogate that is directly related to the process of tornadogenesis would remove the need for environmental filters that may not perform equitably across different environments and storm types.

Toward this end, we have chosen to interrogate CAMs for low-level (<3 km AGL) rotation that can identify the presence and intensity of *low-level* mesocyclones in simulated supercells. Compared to midlevel rotation (i.e., 2–5 km AGL UH), low-level mesocyclone strength is more strongly associated with tornado occurrence (Trapp et al. 2005) and with tornado intensity (Smith et al. 2015). Further, field studies have directly tied the behavior and strength of low-level mesocyclones to tornadogenesis (e.g., Markowski et al. 2012; Skinner et al. 2014). Finally, using low-level rotation as a surrogate for tornadoes may improve forecasts of tornadoes occurring in storms where midlevel rotation (>3 km AGL), and thus midlevel UH, is reduced, for example, mesovortices in QLCSs (e.g., Trapp and Weisman 2003) and shallow mesocyclones in miniature supercells associated with landfalling tropical cyclones (e.g., McCaul 1991; McCaul and Weisman 1996; Eastin and Link 2009; Edwards et al. 2012).

However, a prerequisite step to using explicit forecasts of low-level rotation as a surrogate for tornadoes is to verify the model's ability to produce low-level rotation within environments that are favorable for tornadoes in observed supercells. In other words, an additional goal of this work is to confirm the model can mimic results obtained by proximity sounding studies (e.g., T03). If the magnitude of low-level rotation in simulated storms depends on similar parameters to

[1] Here, we use the term tornado-like vortex to describe a near-surface vortex that occurs in association with a parent thunderstorm using CAMs run at resolutions that partially, but not fully, resolve tornado-scale circulations [$O(100)$ m]. The term tornado-like vortex has also been used in the literature to describe an idealized simulation of a tornado that is not associated with a parent thunderstorm, regardless of model resolution (e.g., Rotunno 1979).

those identified in T03 and other proximity-sounding based studies, it would inspire confidence that the model is capable of producing low-level rotation for the right reasons [e.g., as a result of interaction of the storm's updraft with environment- and storm-generated vorticity; Markowski et al. (2012)]. Additionally, it would lend support to the use of explicit forecasts of low-level rotation as a surrogate for tornadoes *without the need to filter based on environmental parameters*.

In some sense, this goal is similar to the goals of the pioneering work with cloud-scale models that examined the parameter space of environmental conditions controlling convective mode and behavior (e.g., Weisman and Klemp 1982, 1984), with a focus here on the development of low-level rotation in simulated storms. Low-level rotation has been explored as a surrogate for tornadoes in recent storm-scale data assimilation studies in a small collection of cases (e.g., Stensrud et al. 2009; Dawson et al. 2012; Yussouf et al. 2015; Wheatley et al. 2015). Yet, it remains unclear whether present-day CAMs with horizontal grid spacings near 3 km are capable of reproducing intense low-level rotation across a broader spectrum of environments that support tornadoes.

To investigate the relationship between model environment and low-level rotation, we use a collection of ensemble forecasts from a 3-km horizontal grid-spacing CAM ensemble run during spring 2015 (Schwartz et al. 2015). From this dataset, we extract three surrogate fields related to rotation within simulated convective storms, two of which are directly related to low-level (<3 km AGL) rotation that have not previously been used as surrogate fields in next-day CAM forecasts. The first, 2–5 km AGL UH (UH25), is the traditional UH diagnostic introduced by Kain et al. (2008) and examined in Clark et al. (2013) and Gallo et al. (2016). The second, 0–3 km AGL UH (UH03), is the same as UH25, except UH is calculated using fields between the surface and 3 km AGL. The third, 1 km AGL relative vertical vorticity (RVORT1), is used as a measure of the rotation at a fixed level AGL, rather than over an integrated depth as in UH25 and UH03.

In section 2, details of the ensemble forecasts and forecast guidance generation procedure are presented, followed by an examination of the model low-level rotation climatology in section 3. Section 3 also examines the environments in which intense low-level rotation occurs in the forecasts compared to results previously identified as conducive for tornadic supercells in proximity soundings (e.g., T03). Based on these results, next-day forecast guidance for tornadoes using the low-level rotation surrogates is presented and verified in the final subsection of section 3. A summary and discussion of the findings are presented in section 4.

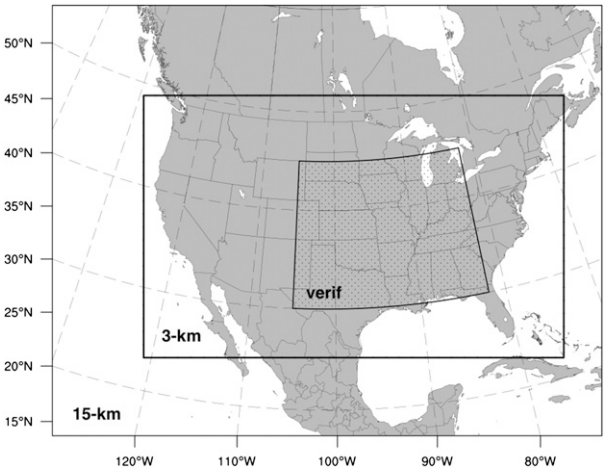


FIG. 1. Nested forecast domains for the 2015 NCAR real-time ensemble system. Verification of E-SSPFs is performed only over the speckled region. [Adapted from Schwartz et al. (2015).]

## 2. Methodology

### a. Ensemble forecasts

Ensemble forecasts were generated from a continuously cycled ensemble Kalman filter (EnKF) based ensemble system currently run in real time at NCAR using the Advanced Research version of the Weather Research and Forecasting Model (WRF; Skamarock et al. 2008) and the Data Assimilation Research Testbed (DART; Anderson et al. 2009). The cycled system produces a set of 50 analyses at 15-km horizontal grid spacing every 6 h by assimilating a variety of surface and upper-air observations using the ensemble adjustment Kalman filter (Anderson 2001, 2003) within DART and integrating the forecasts to the next assimilation time with WRF. Following the assimilation step at 0000 UTC, the first 10 members from the set of 50 analyses from 0000 UTC are downscaled to 3 km over a continental United States (CONUS)-spanning domain (Fig. 1) and integrated to 48 h using WRF. Further details about the analysis and ensemble forecast system are found in Schwartz et al. (2015). Here, we use the set of 92 ten-member ensemble forecasts initialized daily at 0000 UTC from 30 April through 30 July 2015. Although this is a larger collection of forecasts relative to other studies (e.g., Schwartz et al. 2015), because of the rarity of tornado events this remains a fairly small sample, and the results are only valid over the warm season.

### b. Rotation diagnostics

The UH25, UH03, and RVORT1 diagnostics were implemented in WRF as "hourly maximum" fields (Kain et al. 2010). These fields track the maximum value

TABLE 1. List of settings for the enhanced watershed algorithm, defined using the terminology in Lakshmanan et al. (2009). The saliency size was chosen to identify longer, more coherent, UH25 swaths, while reducing the number of smaller, transient, UH25 features (smaller saliency thresholds resulted in long UH25 swaths being broken into multiple objects). No smoothing was applied to the UH25 field during the object-finding process, as a nine-point average is applied each time step when computing the UH25 field in the WRF Model.

| Field | Min ($a$) | Max ($b$) | Increment ($\delta$) | Saliency (km$^2$) | Smoothing |
|---|---|---|---|---|---|
| UH25 | 25 m$^2$ s$^{-2}$ | 250 m$^2$ s$^{-2}$ | 5 m$^2$ s$^{-2}$ | 900 | None |

of the diagnostic at each grid point that occurs at any time step within each hour of the forecast. As in Kain et al. (2008), UH25 was computed as

$$\text{UH25} = \sum_{2\,\text{km}}^{5\,\text{km}} w\zeta\, dz,$$

where $w$ is the vertical velocity and $\zeta$ is the vertical vorticity at a grid point. The value of UH03 was computed in an analogous manner, with the integration occurring between 0 and 3 km AGL. Finally, RVORT1 was computed using vertical vorticity computed as $dv/dx - du/dy$, where $u$ and $v$ are the zonal and meridional wind components, respectively.

### c. Storm object and environment identification

Storm objects and their environments were extracted from the ensemble forecasts valid between 1200 and 1200 UTC the following day (i.e., forecast hours 12–36). The enhanced watershed algorithm (Lakshmanan et al. 2009; Lakshmanan 2012), as implemented in the hagelslag processing package (Gagne et al. 2016), was used to identify objects in the model output. The watershed algorithm, as applied here, takes a two-dimensional model field and identifies maxima to begin the object-finding process. The objects initiated at these maxima grow until the saliency is reached (Table 1) or points are encountered that are associated with other objects. While object size is partially controlled by the saliency parameter, an additional minimum size criterion is also applied to remove objects smaller than 90 km$^2$ [to remove objects smaller than the effective resolution of the model, as in Davis et al. (2006) and Johnson et al. (2011)]. The benefit of using the watershed object-finding algorithm is that it does not depend on a predefined threshold above which objects are created; here, regions are grown beginning at field maxima. The reader is encouraged to consult Lakshmanan et al. (2009) for full details of the enhanced watershed algorithm.

For this work, objects were identified using the UH25 field, with the minimum UH25 threshold set to 25 m$^2$ s$^{-2}$. Other watershed algorithm settings are provided in Table 1. We chose UH25 to identify objects since the present work focuses on the behavior of low-level and midlevel rotation within storms possessing some degree of midlevel rotation; other fields used as input into the object identification algorithm to identify storm objects (e.g., hourly maximum vertically integrated graupel) produced similar results. The watershed algorithm, as used in this work, does not identify storms per se, but rather the associated hourly maximum UH25 swaths.

Storm and environmental variables were extracted from the forecasts at the collection of points composing each object (Table 2). The storm variables were

TABLE 2. List of storm and environmental properties extracted from the ensemble forecasts using the enhanced watershed object-finding algorithm. For the storm properties, the maximum value within each object is extracted, while for the environmental properties, the mean value within each object is extracted.

| Field | Name | Storm or environment |
|---|---|---|
| UH25 | Hourly max updraft helicity within the 2–5 km AGL layer | Storm |
| UH03 | Hourly max updraft helicity within the 0–3 km AGL layer | Storm |
| RVORT1 | Hourly max 1 km AGL relative vorticity | Storm |
| SBCAPE | Surface-based parcel convective available potential energy | Environment |
| SBCIN | Surface-based parcel convective inhibition | Environment |
| SBLCL | Surface-based parcel lifted condensation level | Environment |
| SRHEL01 | 0–1 km AGL storm-relative helicity (computed using Bunkers storm motion) | Environment |
| SRHEL03 | 0–3 km AGL storm-relative helicity (computed using Bunkers storm motion) | Environment |
| SHR01 | 0–1 km AGL shear vector magnitude | Environment |
| SHR06 | 0–6 km AGL shear vector magnitude | Environment |
| STP | Fixed-layer significant tornado parameter, defined as STP = (SBCAPE/1500 J kg$^{-1}$) × [(2000 − SBLCL)/1000 m] × (SRHEL01/150 m$^2$ s$^{-2}$) × (SHR06/20 m s$^{-1}$), where the SBLCL term can only vary between 0.0 and 1.0, and the SHR06 term can only vary between 0.625 and 1.5 | Environment |

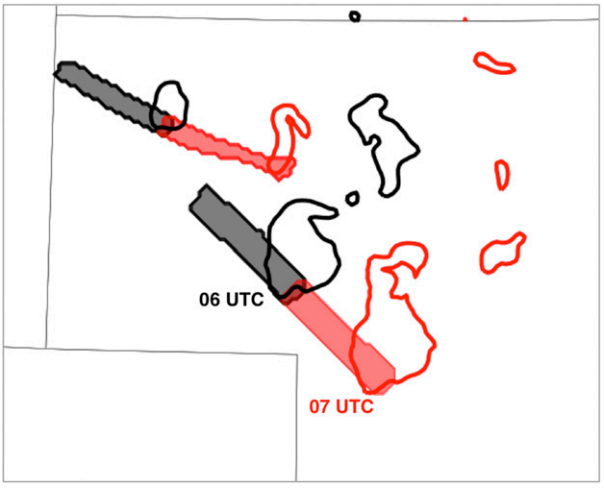


FIG. 2. Example of UH25 objects associated with forecast convection in western NE on 21 Jun 2015 at two forecast hours (0600 and 0700 UTC). Maximum column reflectivity (contours) and UH25 objects (filled contours) are plotted together from both the 0600 UTC (black) and 0700 UTC (red) forecasts. The algorithm uses the set of grid points within each object to extract 1) storm properties from the same hour's forecast (e.g., for the 0700 UTC object in red, the 0700 UTC hourly maximum fields are used) and 2) environmental properties from the previous hour's forecast (e.g., the environment for the 0700 UTC object is taken from the collection of points comprising the 0700 UTC object but using the 0600 UTC forecast).

extracted from the hourly maximum fields valid at the same forecast hour as the UH25 object, while the environmental variables used the forecast valid at the hour prior to the time the object was defined, similar to the method used by Clark et al. (2012). (Figure 2 provides some additional detail into this process.) The maximum value of each field within the set of grid points defining the object was stored at each forecast hour for the storm properties, while the mean value was used for the environment properties. Using the mean value for the environmental properties reduces the influence of grid points that may be convectively contaminated near the storm.

The final output of the object identification process was a collection of ~192 000 objects and their associated maximum within-storm and mean environmental properties, facilitating a comparison between attributes of the storm (e.g., low-level rotation) and the environment (e.g., low-level shear). The object and environment extraction process is applied individually to each ensemble member with the results aggregated across all ensemble members.

### *d. Creation of next-day tornado guidance*

In addition to the creation of storm objects, the three rotation fields were used to produce next-day (1200–1200 UTC, forecast hours 12–36) tornado guidance following the procedure of Sobash et al. (2011) and extended to ensemble forecasts in Sobash et al. (2016). This procedure does not use the previously created objects, but applies thresholds to each hourly max rotation field from each member. This produces a binary field of ones and zeros (Table 3) that are referred to as surrogate severe reports (SSRs). The SSRs were mapped onto an 80-km grid,[2] where each 80-km grid point could only be flagged as a "hit" once per 24-h period. The final grid of SSRs for each 1200–1200 UTC period was then smoothed with a Gaussian kernel to produce a 24-h surrogate severe probability forecast (SSPF) for each ensemble member. A Gaussian smoothing kernel width, defined by $\sigma$, ranged between $\sigma = 20$ and 200 km, in 20-km increments. Each ensemble member's SSPF was averaged together to create an ensemble SSPF (E-SSPF) tornado forecast for each day.

For verification, tornado reports were retrieved from the Storm Prediction Center's (SPC) archive, mapped, and smoothed onto the same 80-km grid as the SSRs to produce an analogous field of observed storm reports (OSRs) and a smoothed observation field [termed an observed severe probability field (OSPF), as in Sobash et al. (2011)]. The set of SSR thresholds in Table 3 was selected using the total number of OSRs over the 92-day period as a benchmark (285 OSRs); that is, the range of SSR thresholds used here produced SSR biases centered on 1 (Table 3).

The E-SSPFs created from the SSRs were verified against the OSPFs (each with the same smoothing length scale) using the fractions skill score (FSS; Roberts and Lean 2008) over a domain consisting of most of the central and eastern United States (Fig. 1). The E-SSPFs with different $\sigma$ values were used within the FSS to identify the spatial scales over which the forecasts were skillful. As described in Roberts and Lean (2008), when the fractional coverage of the event is small, the spatial scales where FSS $\geq 0.5$ is deemed the lower limit of the scales that are believable. In addition to the FSS, E-SSPFs were evaluated using metrics such as reliability, resolution, and the area under the relative operating characteristic (AUC; Wilks 2006).

[2] The 80-km grid used here is the NCEP 211 grid (http://www.nco.ncep.noaa.gov/pmb/docs/on388/tableb.html#GRID211). Using this grid produces probabilities that are consistent with SPC's probabilistic convective outlooks. Also, this coarse grid reduces some of the reporting biases present in the observed severe weather report database (e.g., Weiss and Vescio et al. 1998). The E-SSPFs created using the native model grid, by "flagging" all grid points within 40 km of each 3-km grid point, produced qualitatively similar results.

TABLE 3. Thresholds, SSR biases, and numbers of SSRs for UH25, UH03, and RVORT1. The boldface row indicates the thresholds that produce an SSR bias closest to 1, when compared to the number of OSRs (285).

| UH25 ($m^2 s^{-2}$) | | | UH03 ($m^2 s^{-2}$) | | | RVORT1 ($s^{-1}$) | | |
|---|---|---|---|---|---|---|---|---|
| Threshold | Bias | No. | Threshold | Bias | No. | Threshold | Bias | No. |
| 25 | 35 | 9975 | 20 | 23.87 | 6804 | 0.005 | 27.7 | 7894 |
| 50 | 16.64 | 4743 | 30 | 15.17 | 4323 | 0.006 | 17.18 | 4895 |
| 75 | 9.15 | 2609 | 40 | 10.27 | 2926 | 0.007 | 10.56 | 3010 |
| 100 | 5.45 | 1552 | 50 | 7.26 | 2068 | 0.008 | 6.46 | 1842 |
| 125 | 3.42 | 975 | 75 | 3.38 | 964 | 0.009 | 3.99 | 1136 |
| 150 | 2.2 | 628 | 100 | 1.7 | 484 | 0.01 | 2.4 | 685 |
| 175 | 1.45 | 413 | 110 | 1.33 | 378 | 0.0112 | 1.29 | 369 |
| **200** | **0.99** | **281** | **120** | **1.04** | **296** | **0.0117** | **1.01** | **287** |
| | | | 130 | 0.84 | 238 | 0.012 | 0.87 | 247 |

## 3. Results

### a. *Model climatology of low-level and midlevel rotation*

To develop a sense for the behavior of the two new low-level surrogates (UH03 and RVORT1) compared with UH25, frequency distributions were computed for one member of the ensemble for the entire 92-day period for each forecast hour using the native model grid points (Fig. 3). Distributions using other members produced similar results. Three percentiles (99.99th, 99.995th, and 99.999th) were used to identify values associated with "extremes" in the model climatology. After stabilizing following the model spinup period, UH03 values associated with the three percentiles were approximately half of the corresponding UH25 values (Fig. 3). While the specific UH25 and UH03 values varied with forecast hour, the ratio between UH25 and UH03 was not sensitive to the time of day and generally remained near 0.5 (i.e., UH03/UH25 ~ 0.5). The UH25

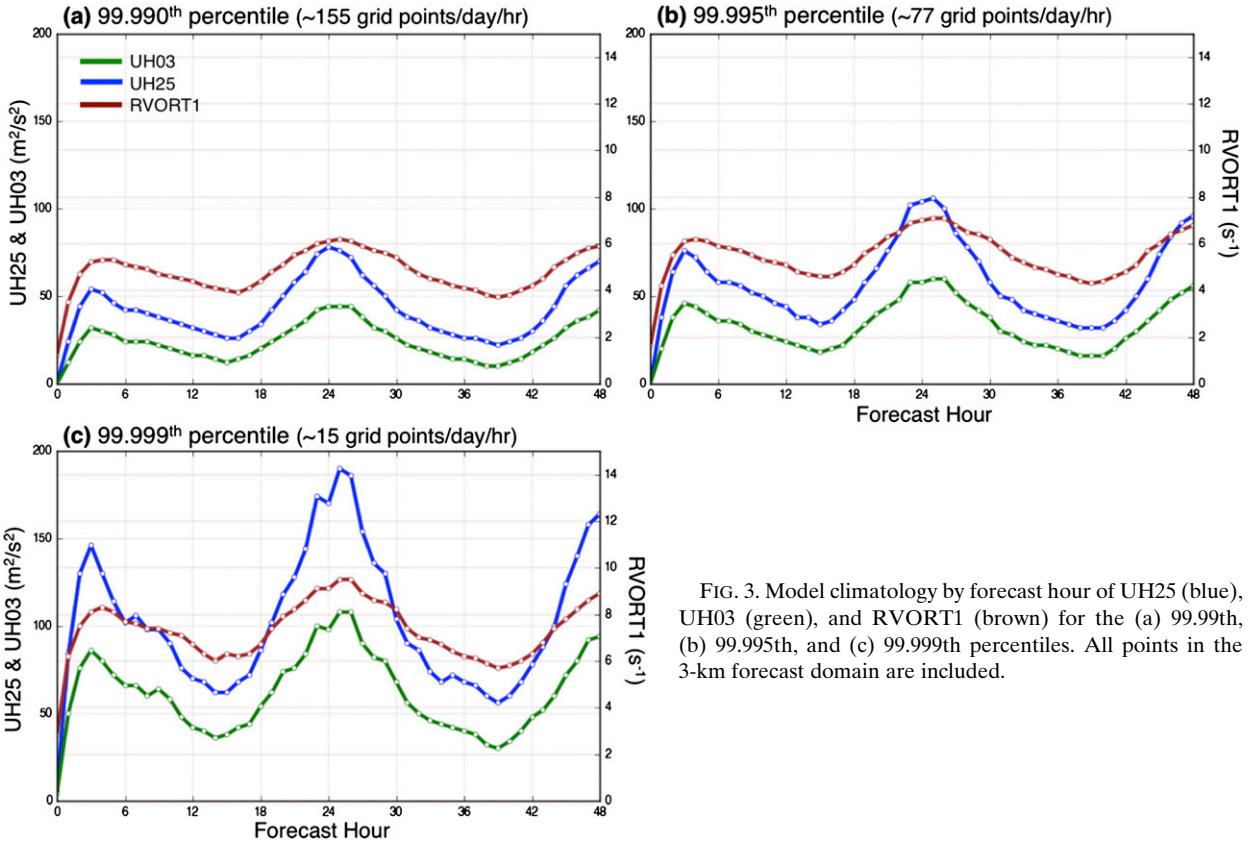


FIG. 3. Model climatology by forecast hour of UH25 (blue), UH03 (green), and RVORT1 (brown) for the (a) 99.99th, (b) 99.995th, and (c) 99.999th percentiles. All points in the 3-km forecast domain are included.

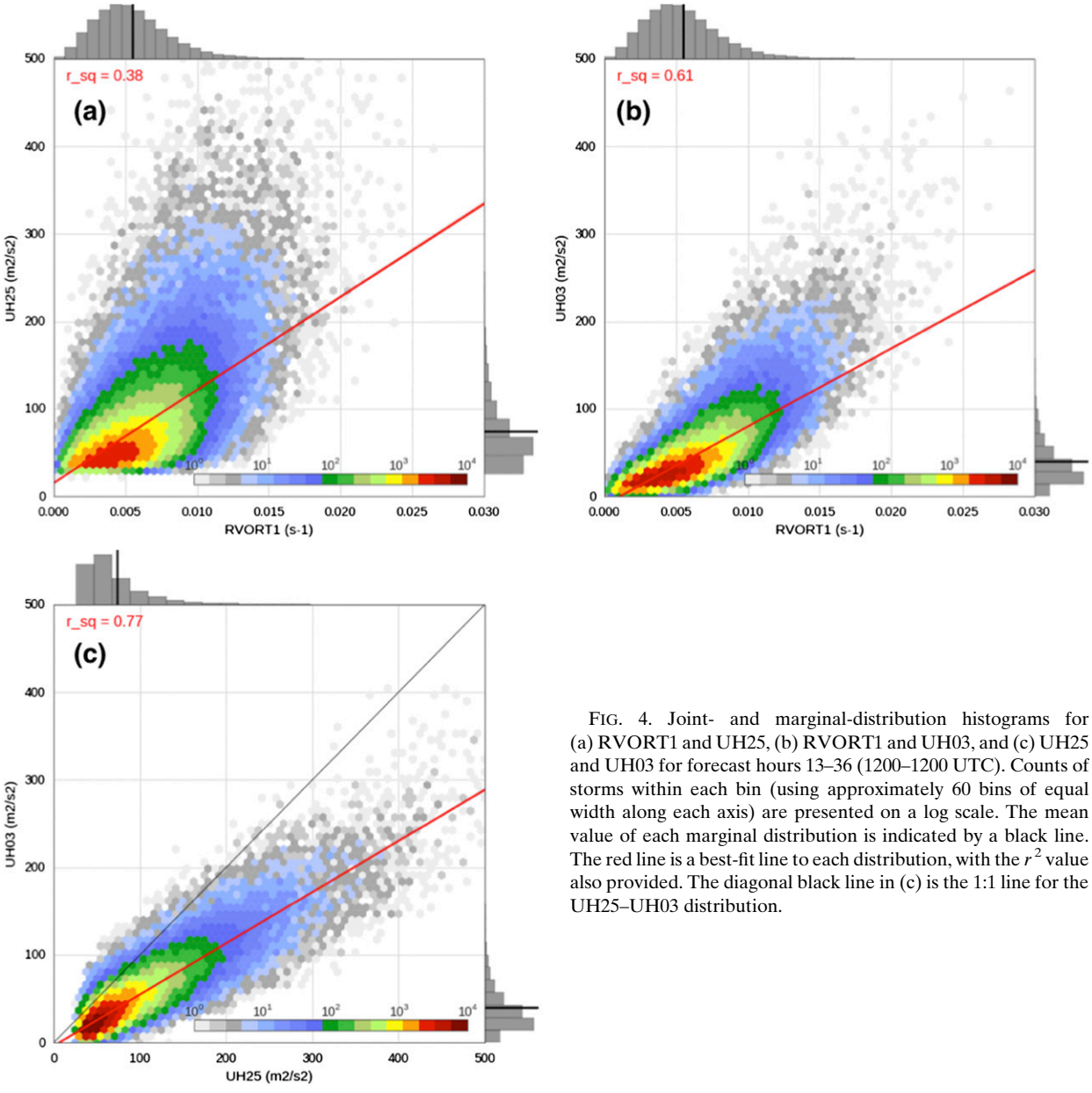


FIG. 4. Joint- and marginal-distribution histograms for (a) RVORT1 and UH25, (b) RVORT1 and UH03, and (c) UH25 and UH03 for forecast hours 13–36 (1200–1200 UTC). Counts of storms within each bin (using approximately 60 bins of equal width along each axis) are presented on a log scale. The mean value of each marginal distribution is indicated by a black line. The red line is a best-fit line to each distribution, with the $r^2$ value also provided. The diagonal black line in (c) is the 1:1 line for the UH25–UH03 distribution.

and UH03 magnitudes peaked near forecast hour 24 (i.e., 0000 UTC), when convective coverage and intensity reached a maximum, and decreased during the overnight convective minimum, as convective intensity waned. RVORT1 magnitudes for all three percentiles exhibited similar behavior, with maxima and minima during similar forecast hours as the UH25 and UH03 distributions.

In addition to the frequency distributions using the native model grid points, a storm-based climatology was performed to compare UH25, UH03, and RVORT1 values within storms at each forecast hour over the entire ensemble. Two-dimensional histograms of storm counts were created to examine the relationship between these three surrogate fields (Fig. 4). Both UH25 and UH03 are positively correlated with RVORT1 (Figs. 4a,b), but the strength of the relationship is much greater between UH03 and RVORT1, which is expected, since RVORT1 is a component of the UH03 computation. UH25 and UH03 had a strong positive correlation, with a best-fit line to this distribution producing a similar scaling between the two fields (i.e., UH03 $\sim$ UH25 $\times$ 0.5) to what was determined based on the grid-based frequency distributions (Fig. 4c). Yet,

variability exists in this relationship with many storms possessing UH03 greater or less than 0.5 × UH25. For example, there exists a subset of storms that possess UH03 values greater than UH25 values (the collection of points above the diagonal in Fig. 4c). These occur most frequently for storms with UH25 values $< 100\,m^2\,s^{-2}$.

### b. *Environments of simulated supercells possessing low-level rotation*

An extensive amount of research has examined the environmental conditions conducive for supercell development, strong low-level rotation within supercells, and tornadoes (e.g., Fawbush and Miller 1954; Brooks et al. 1994; Rasmussen and Blanchard 1998; Rasmussen 2003; T03, Thompson et al. 2012). These studies used estimates of the near-storm environment taken from environmental soundings or model analysis and forecast soundings (e.g., Rapid Update Cycle analyses as in T03). Collectively, these studies have identified environmental characteristics that discriminate well between tornadic and nontornadic supercells, and the significant tornado parameter (STP) was developed to combine those environmental predictors that most strongly discriminate between tornadic and nontornadic supercells (SHR06, SRHEL01, SBCAPE, and SBLCL; see Table 2 for definitions of predictors) into a composite index (Thompson et al. 2002; T03).

Here, we verify the model's ability to reproduce the results obtained in these proximity-sounding studies by directly sampling the magnitude of low-level rotation produced within supercells (using RVORT1) and relating it to the storm's environment. To do so, the environments of all simulated storm objects were extracted at each forecast hour, as described in section 2c, which resulted in a collection of ~192 000 storm-environment pairs over the 92-day forecast period from the 10-member CAM ensemble. The list of environmental fields sampled from the model is provided in Table 2. The choice of environmental parameters was guided by preexisting parameters whose discrimination potential has already been proven in previous work (e.g., STP). For the analyses in section 3b, we explore exclusively the environments of simulated supercells,[3] defined as the subset of the ~192 000 storm objects containing UH25 $> 75\,m^2\,s^{-2}$ based on the threshold used in Sobash et al. (2016). This resulted in a set of ~63 000 simulated supercell objects.

#### 1) THERMODYNAMIC ENVIRONMENTS

The thermodynamic environments of the ~63 000 simulated supercells were examined in aggregate across three distributions: those possessing 1) weak values of RVORT1 ($<0.005\,s^{-1}$), 2) moderate values of RVORT1 ($\geq 0.005\,s^{-1}$, but $<0.015\,s^{-1}$), and 3) strong values of RVORT1 ($\geq 0.015\,s^{-1}$). These thresholds were chosen in order to correspond with the three categories in T03 (nontornadic, weakly tornadic, and significantly tornadic).

Overall, supercells possessing strong values of RVORT1 tended to occur in thermodynamic environments with lower SBLCLs, smaller SBCIN, and slightly larger SBCAPE than storms with smaller values of RVORT1. For example, storms with RVORT1 $< 0.005\,s^{-1}$ occurred within environments characterized by a broad range of SBLCL magnitudes, with a mean SBLCL of ~894 m (Fig. 5). As RVORT1 increased, the range of SBLCL values decreased; supercells with RVORT1 $> 0.015\,s^{-1}$ had a mean SBLCL of ~479 m, with a much smaller interquartile range (~436 m) than supercells possessing weak RVORT1 (interquartile range ~1012 m) (Fig. 5). The dependence of low-level rotation intensity on SBLCL has been documented in other studies using proximity soundings (e.g., Rasmussen and Blanchard 1998; T03) and supports in situ observations of warmer downdrafts within tornadic supercells due to increased boundary layer relative humidity (Markowski et al. 2002).

The RVORT1 magnitude also exhibited sensitivity to the amount of environmental instability (Fig. 6). Supercells with stronger low-level rotation tended to have larger mean values of SBCAPE (~1024 J kg$^{-1}$ for storms with weak RVORT1 and ~2070 J kg$^{-1}$ for storms with strong RVORT1). Yet, there was considerable overlap among the three SBCAPE distributions (Fig. 6a). T03 noted a similar result using MLCAPE, with a large overlap between MLCAPE distributions in significantly tornadic and nontornadic supercells (e.g., their Fig. 6). While the mean SBCIN for supercells with weak RVORT1 is similar across all three distributions (roughly 25 J kg$^{-1}$), the range of environments supporting strong RVORT1 is much narrower (Fig. 6b). For example, the 10th percentile for supercells with weak RVORT1 is ~155 J kg$^{-1}$, but for supercells with strong RVORT1 it is ~85 J kg$^{-1}$, suggesting supercells with strong RVORT1 rarely occur in environments with SBCIN > 85 J kg$^{-1}$. It is likely that the supercells with weak RVORT1 and large values of SBCIN are elevated, preventing appreciable low-level rotation from developing.

#### 2) KINEMATIC ENVIRONMENTS

The strength of RVORT1 does not appear to be sensitive to deep-layer shear magnitude (SHR06), but is

[3] While referred to as "simulated supercells," the storm mode of each object studied in section 3b was not identified. UH25 $> 75\,m^2\,s^{-2}$ can occur within other convective modes (e.g., intense squall lines).

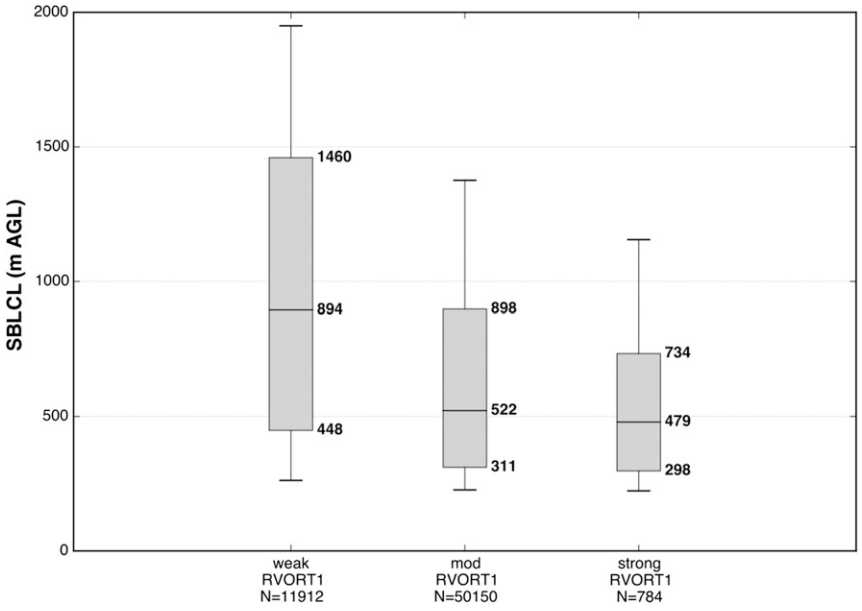


FIG. 5. Boxplots of mean SBLCL (m AGL) for the set of UH25 objects with UH25 $> 75\,m^2\,s^{-2}$ (~63 000 objects) stratified by RVORT1 magnitude (weak, RVORT1 $< 0.005\,s^{-1}$; moderate, RVORT1 between 0.005 and $0.015\,s^{-1}$; strong, RVORT1 $> 0.015\,s^{-1}$). Magnitudes for the 25th, 50th (median), and 75th percentiles are shown in the boxplots, and whiskers extend to the 10th and 90th percentiles.

sensitive to the low-level shear (SHR01; Fig. 7a). For example, the mean values of SHR06 range between ~21 and $24\,m\,s^{-1}$ among the three distributions, which are nearly identical to the mean SHR06 values obtained by T03 in their three groups of supercells. On the other hand, SHR01 discriminated well between the weak and strong RVORT1 supercell distributions (Fig. 7a), with little overlap of the weak and strong distributions' interquartile ranges. Furthermore, the mean values for each distribution of SHR01 are very similar to those identified in T03. In addition to SHR01, SRHEL01 and SRHEL03 discriminate well between supercells with weak and strong RVORT1 (Fig. 7b). A substantial portion (~75%) of the increase in the mean SRHEL03 values between the weak RVORT1 and strong RVORT1 distributions occurs in the 0–1 km AGL layer, as was also the case in Rasmussen (2003) and T03 (Fig. 7b).

### 3) COMBINED PARAMETER FIELDS AND STP

Based on their ability to discriminate between nontornadic and significantly tornadic supercells in T03, SHR01 and SBLCL have been combined, along with SHR06 and SBCAPE, into composite indices such as the STP (T03). While the range of STP values among the supercells with strong RVORT1 is generally smaller here than in T03 (Fig. 8), which is likely due to differences in the specific STP formulation, STP remains an effective discriminator between supercells that produce weak and strong low-level rotation in the present dataset (Fig. 8).

In addition to STP, the discriminatory ability of a combination of SHR01 and SBLCL fields can be further illustrated by looking at the SHR01–SBLCL parameter space for supercells with weak versus strong magnitudes of RVORT1, as has been done in Craven and Brooks (2004) and Brooks (2006) (Fig. 9). For the supercells (i.e., simulated storms with UH25 $> 75\,m^2\,s^{-2}$; see Figs. 9a,b), those with weak low-level rotation occur across a wide range of the parameter space (Fig. 9a), while those with strong low-level rotation (Fig. 9b) occur only where SBLCLs are low (<~1000 m AGL) and SHR01 magnitudes are relatively large (>~$10\,m\,s^{-1}$). Many of the storms in the overlap region of these two distributions (Figs. 9a,b), that is, those storms that possess weak low-level rotation, but occur in environments conducive to strong low-level rotation, occur in environments with large SBCIN, suggesting they are elevated supercells (e.g., those occurring overnight) or could be associated with convective modes other than supercells (not shown).

The overlap in the two distributions is reduced further when considering only intense supercells (UH25 > $150\,m^2\,s^{-2}$; Figs. 9c,d). The intense supercells with weak

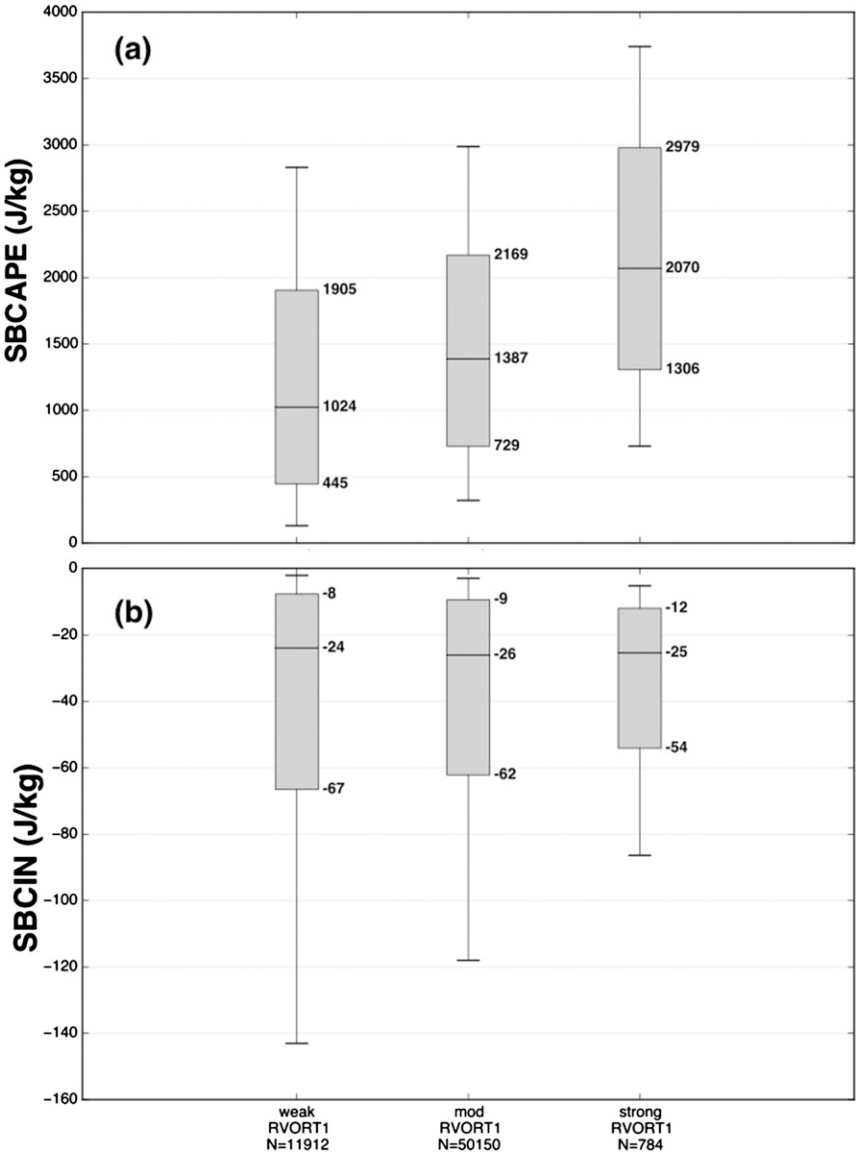


FIG. 6. As in Fig. 5, but for (a) mean SBCAPE (J kg$^{-1}$) and (b) mean SBCIN (J kg$^{-1}$).

low-level rotation occupy the parameter space of ''high based'' environments with small SHR01 magnitudes ($<\sim$10 m s$^{-1}$) and large SBLCLs ($>\sim$1000 m AGL; Fig. 9c), while those with strong low-level rotation occur in environments nearly identical to the distribution in Fig. 9b, that is, environments with large SHR01 and low SBLCLs (Figs. 9b and 9d are similar since most storms with strong low-level rotation have UH25 magnitudes $>$ 150 m$^2$ s$^{-2}$; see Fig. 4a). The distribution of the storms in Figs. 9c,d within the SHR01–SBLCL parameter space is strikingly similar to that found with proximity soundings in Brooks (2006), with a similar separation between the two groups using a linear discriminant analysis (Fig. 10). LCLs are smaller in Fig. 10a likely because of the use of MLLCL in Brooks (2006) instead of SBLCL in the present work. Again, the model appears broadly capable of reproducing the relationships deduced from observational studies of tornadic storms.

Figure 9 also illustrates a key reason for the inability of UH25 to act as an effective surrogate for tornadoes. Identifying potentially tornadic storms based on UH25 alone, even when choosing a higher threshold such as UH25 $>$ 150 m$^2$ s$^{-2}$, will inevitably include a fraction of storms occurring within environments hostile to tornadogenesis (e.g., the set of storms in Fig. 9c when using UH25 $>$ 150 m$^2$ s$^{-2}$). On the other hand, the identification of storms within the favorable part of the parameter space for tornadoes (i.e., low LCLs and large SHR01) can be

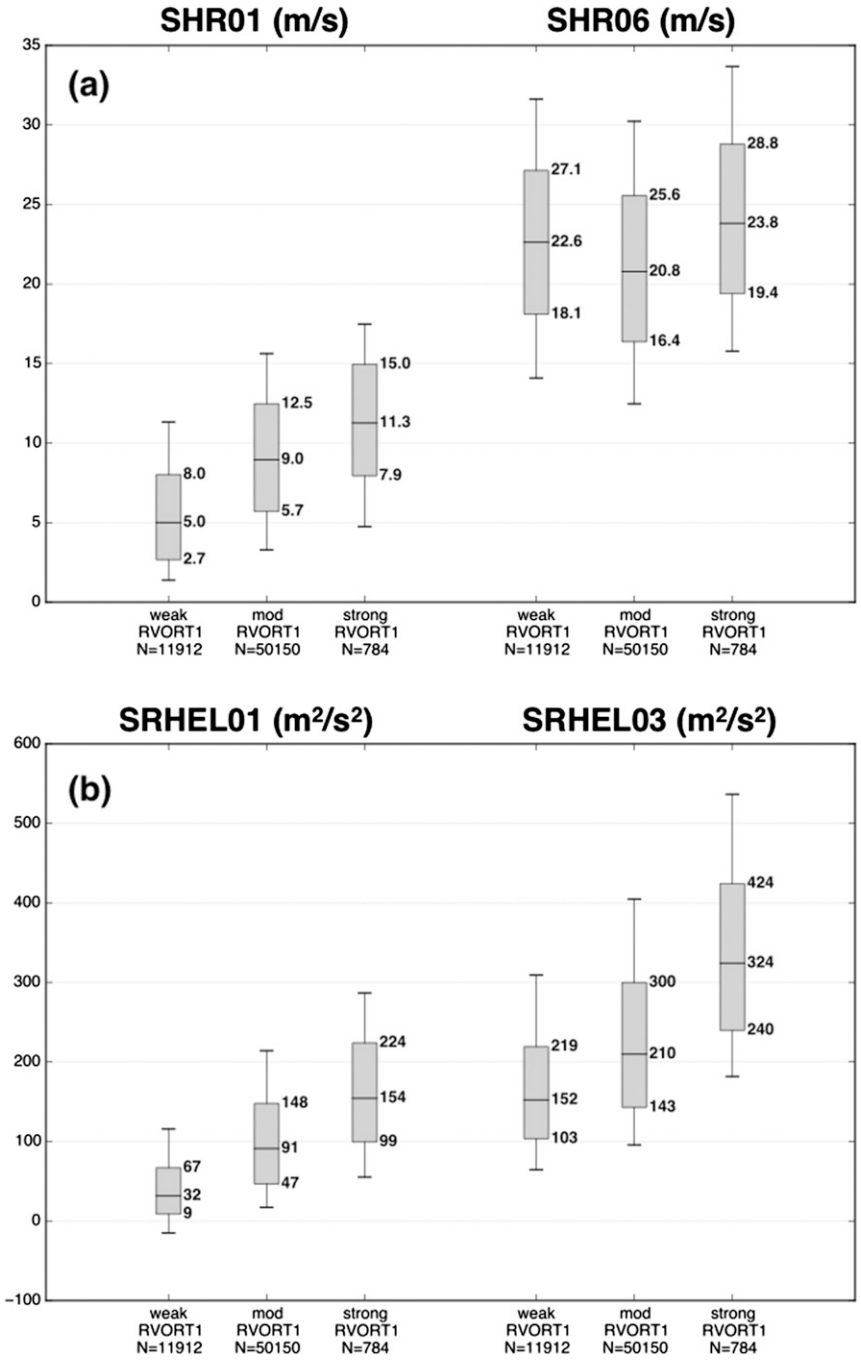


FIG. 7. As in Fig. 5, but for (a) mean SHR01 ($m\,s^{-1}$) and SHR06 ($m\,s^{-1}$) and (b) mean SRHEL01 ($m^2\,s^{-2}$) and SRHEL03 ($m^2\,s^{-2}$).

identified using RVORT1, without the use of UH25 or supplementary environmental parameters.

### c. *Behavior of UH25 and UH03 within simulated storms*

In this section, we further examine how the midlevel (UH25) and low-level rotation (UH03 and RVORT1) diagnostics are related within simulated convective storms. For this, we return to using all storm objects ($UH25 > 25\,m^2\,s^{-2}$), without limiting the analysis to a particular subset. We pay particular attention to the characteristics of storms that possess large UH03 and RVORT1 magnitudes, without the presence of appreciable midlevel rotation (i.e., $UH25 < 75\,m^2\,s^{-2}$). Many

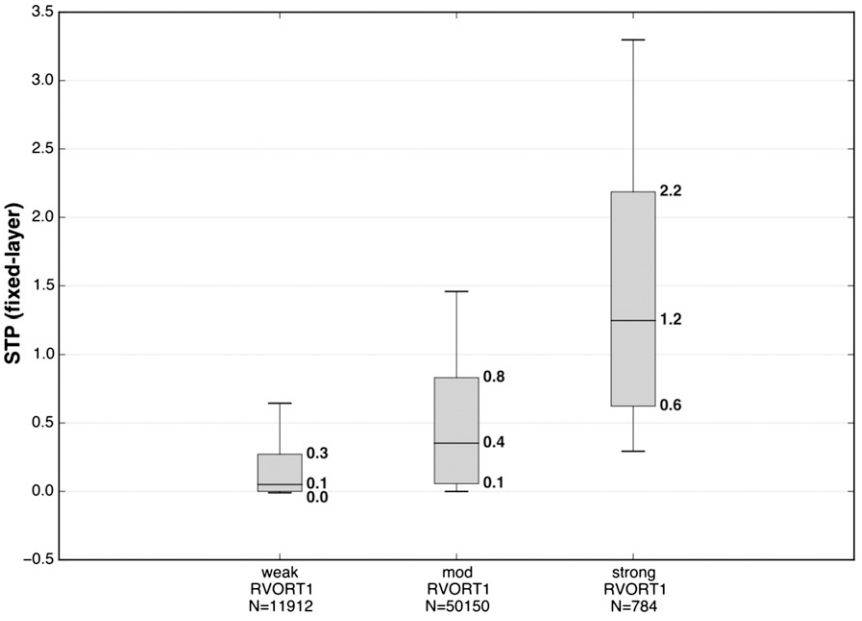


FIG. 8. As in Fig. 5, but for mean fixed-layer STP.

of these storms possess UH03 > UH25 (Fig. 4c), which implies that low-level rotation is particularly strong, relative to midlevel rotation (assuming that vertical velocity, the other component in the UH computation, is greater aloft than near the surface, which alone would result in UH25 > UH03).

To do so, the UH ratio within each simulated storm was computed as

$$\text{UHratio} = \text{UH03}/\text{UH25},$$

using the maximum UH25 and UH03 magnitude within each object. This metric allows for an evaluation of how midlevel and low-level rotation is related within each storm at any given time. Overall, a majority of storms possess UH ratios between 0.3 and 0.7, with a maximum in the number of storms near 0.5 (Fig. 11), agreeing with the relationship identified between UH25 and UH03 in the grid-scale climatology (Fig. 3). For the vast majority of storms, those with UH ratios closer to 1 tend to possess larger RVORT1 magnitudes. Two sets of storms exist that deviate from this relationship, especially for large magnitudes of RVORT1 ($>0.015\,s^{-1}$) and large UH ratios (>1). Although the overall number of storms is small within these parts of the parameter space, we are interested in examining these two subsets of storms, specifically 1) storms with RVORT1 $> 0.015\,s^{-1}$, which were identified in the section 3c, and 2) storms with UH ratios greater than 1, where UH03 exceeds UH25.

For storms with RVORT1 magnitudes $> 0.015\,s^{-1}$, the UH ratios are often between 0.6 and 1.0, with a mean of ~0.8 (Fig. 11). Very few storms in this regime possess UH ratios greater than 1, indicating that UH03 rarely exceeds UH25 in simulated storms that possess RVORT1 $> 0.015\,s^{-1}$. Further, the UH ratio appears to be largely independent of RVORT1 magnitude for storms in this regime, suggesting UH25 increases linearly with UH03 among these storms, such that increased midlevel rotation is associated with a corresponding increase in the strength of the low-level rotation. These UH25 and UH03 magnitudes for these storms are large, well above the UH thresholds typically used to identify supercells (Fig. 12).

For storms with UH ratios > 1, intense low-level rotation (i.e., RVORT1 $> 0.015\,s^{-1}$) rarely occurs (Fig. 11). RVORT1 ranges from ~0.005 to ~$0.015\,s^{-1}$ within this subset of storms, with the range centered around $0.01\,s^{-1}$. The UH25 and UH03 magnitudes for these storms are quite small, with most storms possessing UH25 magnitudes smaller than $75\,m^2\,s^{-2}$ (Fig. 12). The behavior of these storms is distinct from the behavior of the storms with strong low-level rotation in subset 1. Among the storms in subset 2, those with larger UH ratios, the average RVORT1 magnitude stays relatively constant, near $0.01\,s^{-1}$. Since RVORT1 is strongly correlated with UH03, this suggests that within this subset, those storms with larger UH ratios are primarily due to a relative *decrease* in the magnitude of UH25 (rather than a large increase in the magnitude of UH03, while UH25 stays constant).

The environments of the storms in subsets 1 and 2 have different distributions of CAPE and SBLCL. The mean

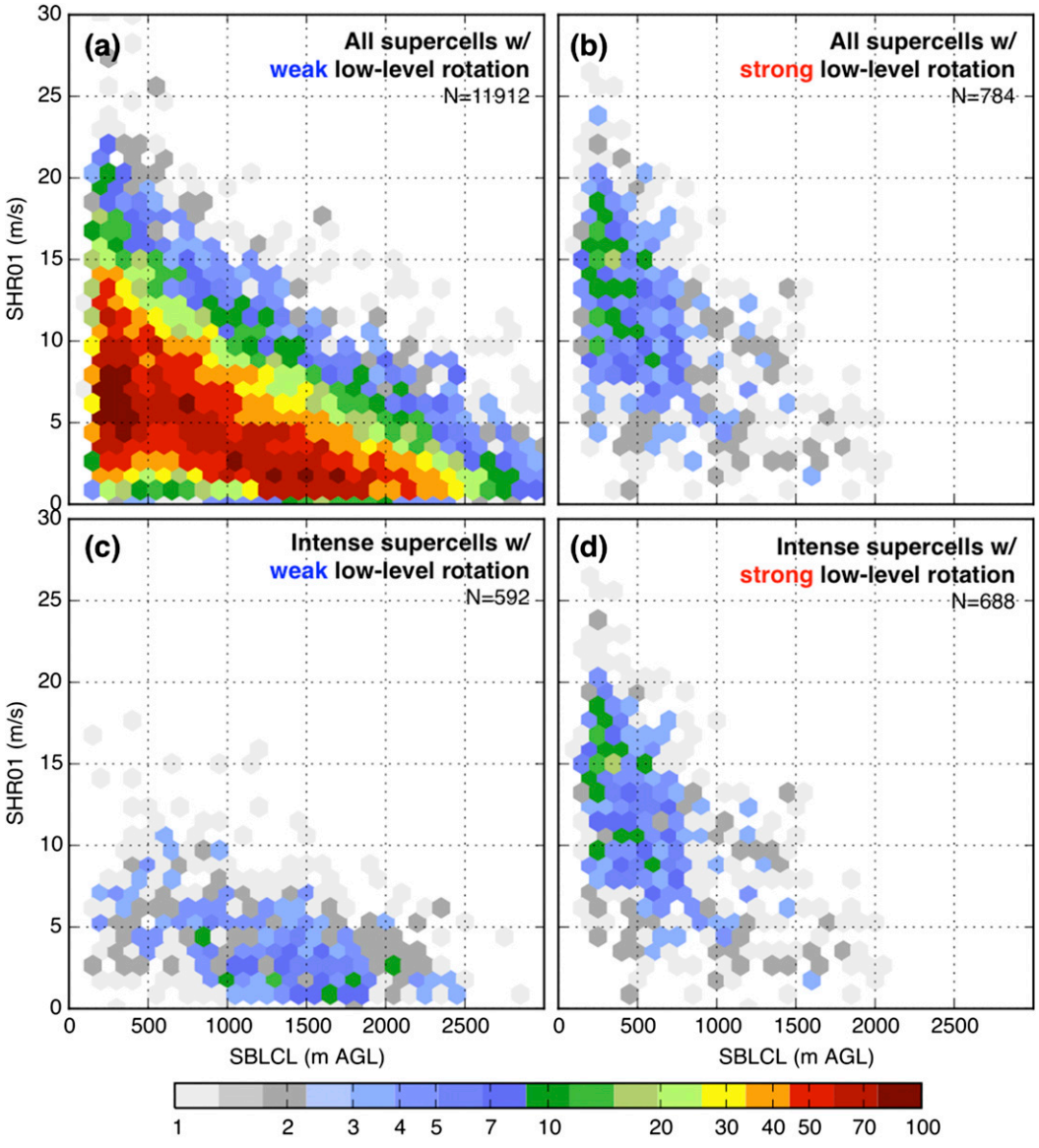


FIG. 9. As in Fig. 4, but for the joint distribution of SHR01–SBLCL for (top) all supercells, defined as UH25 > 75 $m^2 s^{-2}$, and (bottom) intense supercells, defined as UH25 > 150 $m^2 s^{-2}$. Those storms are further stratified by (a),(c) weak low-level rotation, defined as RVORT1 < 0.005 $s^{-1}$, and (b),(d) strong low-level rotation, defined as RVORT1 > 0.015 $s^{-1}$.

SBCAPE among storms in subset 1 is ~1600 J $kg^{-1}$ larger than those in subset 2, with little overlap between the two distributions (Fig. 13). These differences are also present for MLCAPE and MUCAPE (not shown). A smaller difference is present for SBLCL, with the mean SBLCL ~150 m less in the environments for the subset 2 storms (not shown). Among the kinematic variables (e.g., SHR01, SHR06, and SRHEL01), little differences are noted in the two distributions, suggesting that both subset 1 and 2 storms occur in environments of strong low-level and deep-layer shear (not shown). Finally, the storm objects in subset 1 tend to be much larger in size than those in subset 2 (Fig. 14).

The storms in subset 1 possess the traits of ''bona fide'' simulated supercells, that is, storms containing large magnitudes of UH25 and UH03, large values of environmental instability, and large object areas. On the contrary, those in subset 2 have weak magnitudes of UH25 and UH03, are less persistent (leading to smaller object areas), and occur in environments of relatively weak environmental instability. Yet, those storms in subset 2 are still capable of producing

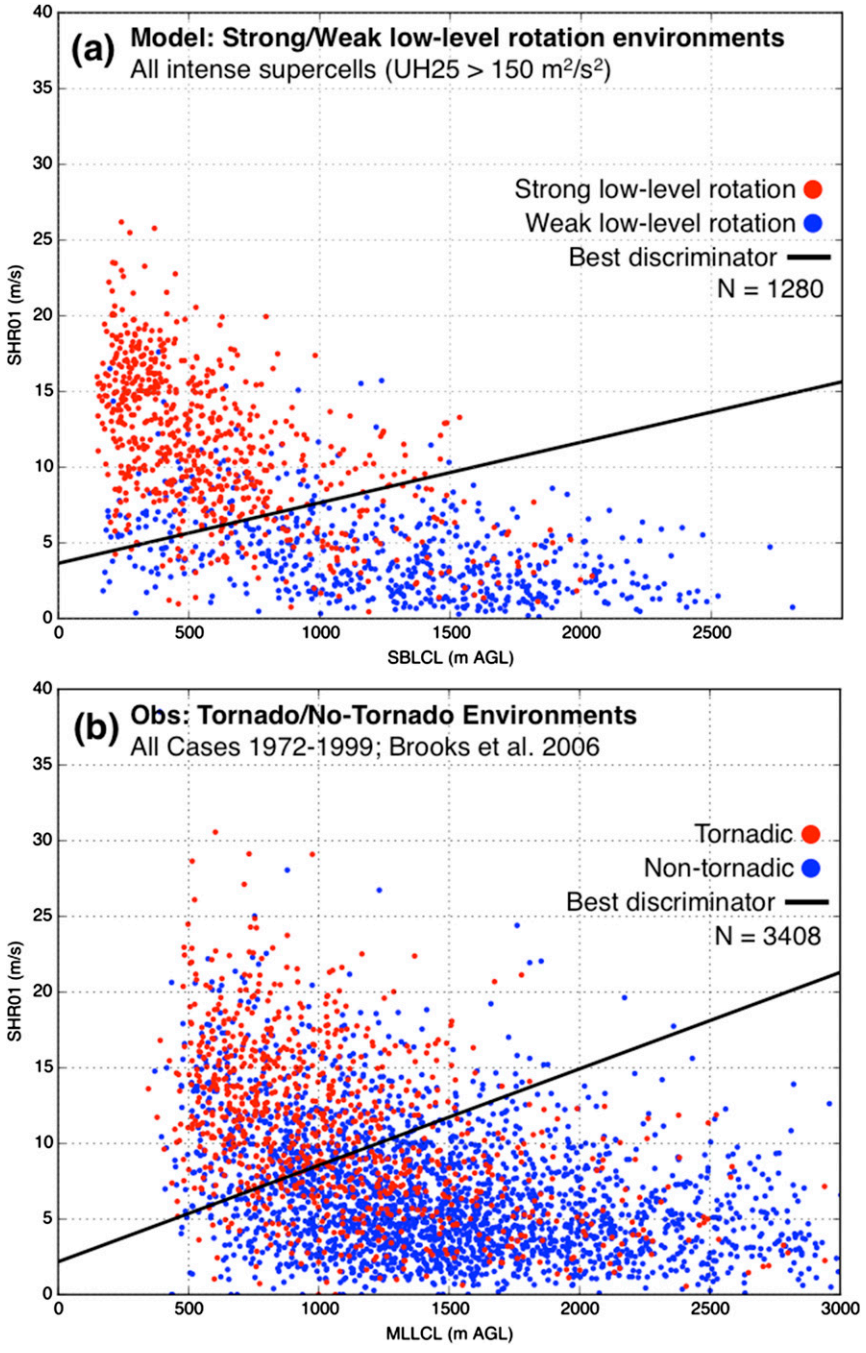


FIG. 10. (a) Scatterplot of UH25 objects from Figs. 11c and 11d (i.e., intense supercells with maximum UH25 $> 150\,m^2\,s^{-2}$) color coded by magnitude of low-level rotation (red, RVORT1 $> 0.015\,s^{-1}$; blue, RVORT1 $< 0.005\,s^{-1}$). Best linear discriminator is plotted as a black line. (b) As in (a), but for SHR01–MLLCL values from observed proximity soundings (from 1972 to 1999) associated with observed tornadic (red) and nontornadic (blue) storms [Adapted from Brooks (2006).]

nonnegligible amounts of low-level rotation (often greater than $0.01\,s^{-1}$; Fig. 11). While the simulated storms in subset 2 do not appear to be traditional supercells (i.e., those with UH25 $> 75\,m^2\,s^{-2}$), they share traits of severe storms that occur in low-CAPE/high-shear environments, or with shallow supercells that are more poorly resolved than those in subset 2 because of their smaller size.

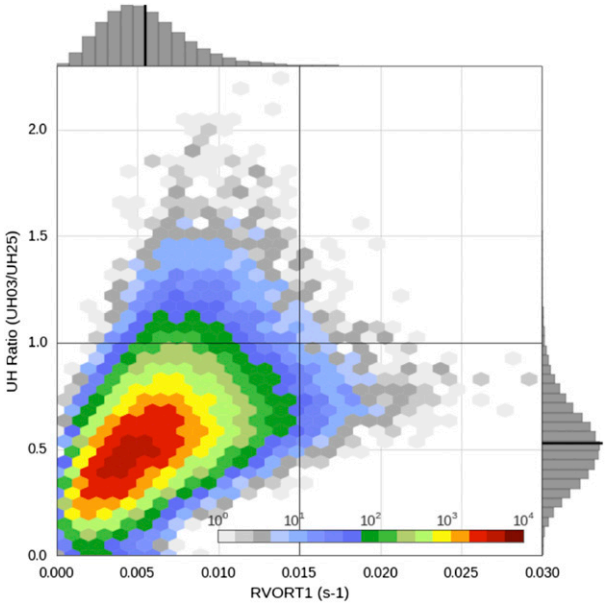


FIG. 11. As in Fig. 4, but for UH ratio (UH03/UH25) and RVORT1. Black lines denote thresholds used to stratify storms (UH ratio $> 1$ and RVORT1 $> 0.015\,s^{-1}$) in text and Figs. 12–14.

### d. *Next-day tornado forecasts using low-level rotation surrogates*

So far, the results have provided confidence that low-level rotation is produced in simulated supercells that occur in environments similar to those conducive for observed tornadoes. These results motivate the development of next-day tornado guidance that uses low-level rotation diagnostics as surrogates for tornadoes without the need for environmental filtering. Such forecasts should be capable of providing an indication of the potential for tornadoes and should be more skillful than using UH25 alone. To test these hypotheses, next-day probabilistic forecasts of tornadoes are generated and verified following the approach outlined in section 2d. As described in section 2d, this method uses the diagnostic fields on the native model grid and not the storm objects used in sections 3b and 3c.

As in Sobash et al. (2016), E-SSPFs are generated over a range of thresholds that produce a number of SSRs over the 92-day period that is similar to the number of observed storm reports, after being mapped to the 80-km forecast grid (Table 3). This ensures that the number of forecast tornado "hits" is similar to the number of observed tornado "hits," reducing the effects of forecast bias when using scores such as the FSS (Mittermaier and Roberts 2010). During the 92-day forecast period, 285 grid points were flagged where a tornado report occurred at some point within the 24-h period of interest (i.e., 285 OSRs). For UH25, UH03, and RVORT1, thresholds of approximately $200\,m^2\,s^{-2}$, $120\,m^2\,s^{-2}$, and $1.17 \times 10^{-2}\,s^{-1}$ produced SSR biases closest to 1.0, respectively (Table 3).

E-SSPFs for two cases are presented here to demonstrate the potential advantages of using UH03 and RVORT1 as surrogates for tornadoes. The first case, 16 May 2015, includes a large tornado outbreak across

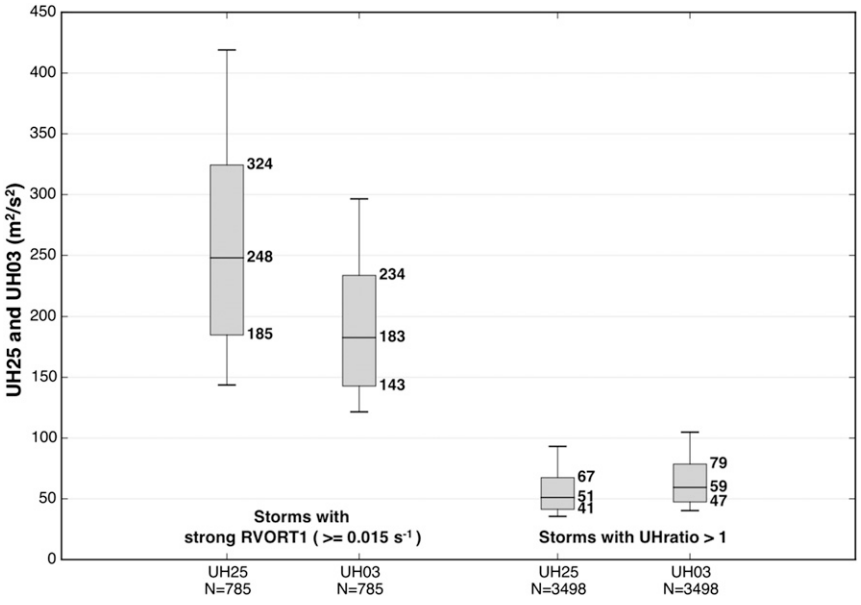


FIG. 12. As in Fig. 5, but for both the UH25 and UH03 distributions within storms that possess (left) RVORT1 $> 0.015\,s^{-1}$ and (right) UH ratio $> 1$ (indicated by black lines in Fig. 11). Sample size for each distribution is provided along the $x$ axis.

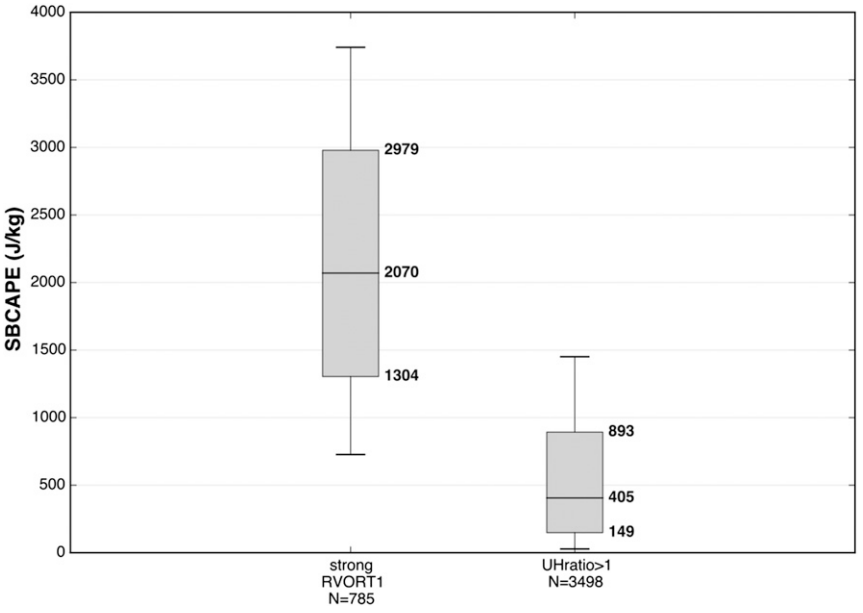


FIG. 13. As in Fig. 12, but for SBCAPE (J kg$^{-1}$).

the central United States, with 25 OSRs occurring within the United States, the largest 1-day total during the 92-day period. The E-SSPFs for this day created using the UH25, UH03, and RVORT1 thresholds where the bias is closest to one are shown in Fig. 15. On this day, the UH25 E-SSPF produced the lowest probabilities, since a relatively small number of storms exceeded the high UH25 threshold (Fig. 15a). Both the UH03 (Fig. 15b) and RVORT1 E-SSPFs (Fig. 15c) produced higher probabilities, closer to the values produced in the smoothed OSR field (Fig. 15d). The spatial patterns of the three E-SSPFs are similar across Kansas and Oklahoma, although the maximum probabilities were positioned in different regions (cf. Figs. 15b,c). In Minnesota, the RVORT1 E-SSPF has the highest probabilities where a cluster of tornadoes were reported, while UH03 E-SSPF probabilities are reduced, and are close to zero for the UH25 E-SSPFs.

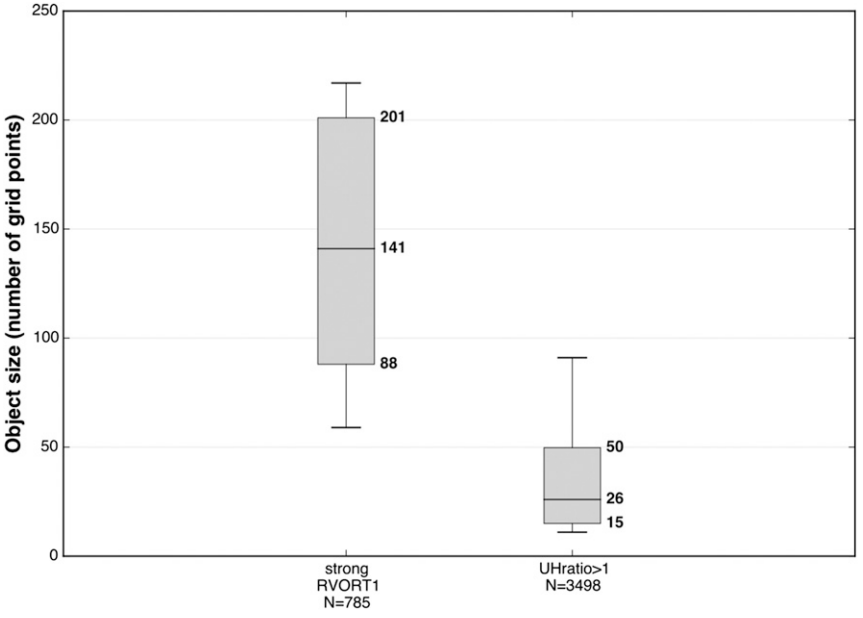


FIG. 14. As in Fig. 12, but for object size (number of grid points).

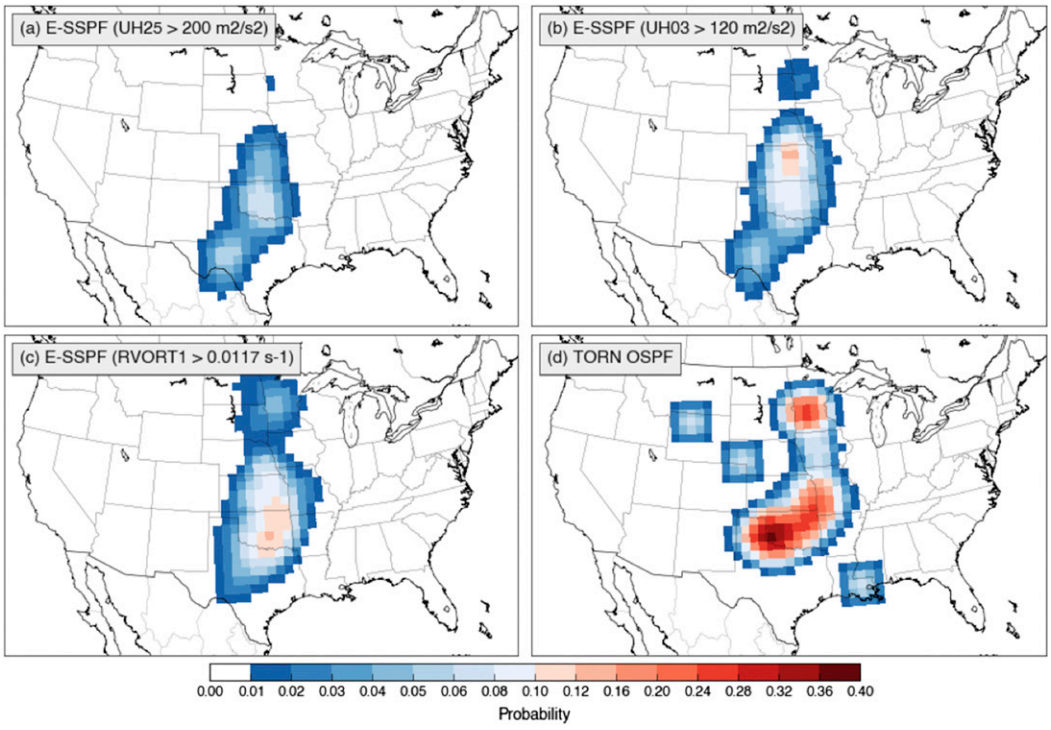


FIG. 15. The 24-h E-SSPFs for tornadoes valid 1200 UTC 16 May–1200 UTC 17 May 2015 produced from the NCAR ensemble forecast initialized at 0000 UTC 16 May 2015 for (a) UH25 $> 200\,m^2\,s^{-2}$, (b) UH03 $> 120\,m^2\,s^{-2}$, and (c) RVORT1 $> 0.0117\,s^{-1}$ with a Gaussian smoother ($\sigma = 120$ km). (d) Tornado OSPF, smoothed with a Gaussian smoother ($\sigma = 120$ km). The E-SSPFs shown here use the SSR thresholds that produce SSR biases closest to 1 (Table 3).

Inspection of the forecast storms in Minnesota revealed that these storms were shallow, with rotation restricted to the lowest levels, hence, the improved E-SSPFs for these storms using low-level rotation surrogates. Also, a probability maximum exists in the UH25 E-SSPF in SW Texas, but does not in the RVORT1 E-SSPF. Several ensemble members correctly forecasted the development of isolated, intense, high-based, supercells in this region, but no tornadoes were observed. This is a scenario where appreciable low-level rotation did not develop, reducing probabilities in the RVORT1 E-SSPF.

The second case was a primarily nontornadic event that occurred across the high plains of Montana, Wyoming, and South Dakota (Fig. 16). Similar to forecasts of high-based intense supercells in the previous case, the storms on this day were high based and mainly produced severe wind gusts and large hail (only one tornado report in this region; Fig. 16d). The use of the RVORT1 surrogate (Fig. 16c) reduced the magnitudes of the E-SSPF probabilities over the UH25 (Fig. 16a) and UH03 E-SSPFs (Fig. 16b), improving the E-SSPFs. The three surrogate fields were not useful in identifying tornadic convection across southern Illinois and Indiana, even though convection was predicted in this area. The storms in southern Illinois were similar to the subset 2 storms discussed in section 3d and did not exceed the thresholds used for the surrogate fields.

For UH25 E-SSPFs, FSS values $\geq 0.5$ (the FSS threshold indicating a skillful forecast) only occur when using the largest UH25 thresholds, where the forecast bias is closest to 1 (Fig. 17a). Even so, the FSS values here barely exceed 0.5 and only occur at the largest smoothing length scales (i.e., $\sigma > 180$ km), with a maximum FSS over all UH25 thresholds of 0.521 at $\sigma = 200$ km. At $\sigma = 200$ km, the E-SSPFs suffer from very poor resolution since large probability values (e.g., >25%) are rarely produced (Fig. 19b). Compared with the UH25 E-SSPFs, the UH03 and RVORT1 E-SSPFs produce FSS $\geq 0.5$ at smaller smoothing length scales (i.e., $\sigma \geq 120$ km), with a maximum FSS approaching 0.64 at $\sigma = 200$ km (Figs. 17b,c). Compared to the UH25 E-SSPFs, UH03 E-SSPFs produced larger FSS values at smaller smoothing length scales for SSR thresholds with a forecast bias closest to 1. RVORT1 E-SSPFs produced similar FSSs to the UH03 E-SSPFs, with a small increase in FSS noted at thresholds where the surrogate bias is closest to 1, although the FSS differences between the UH03 E-SSPFs and

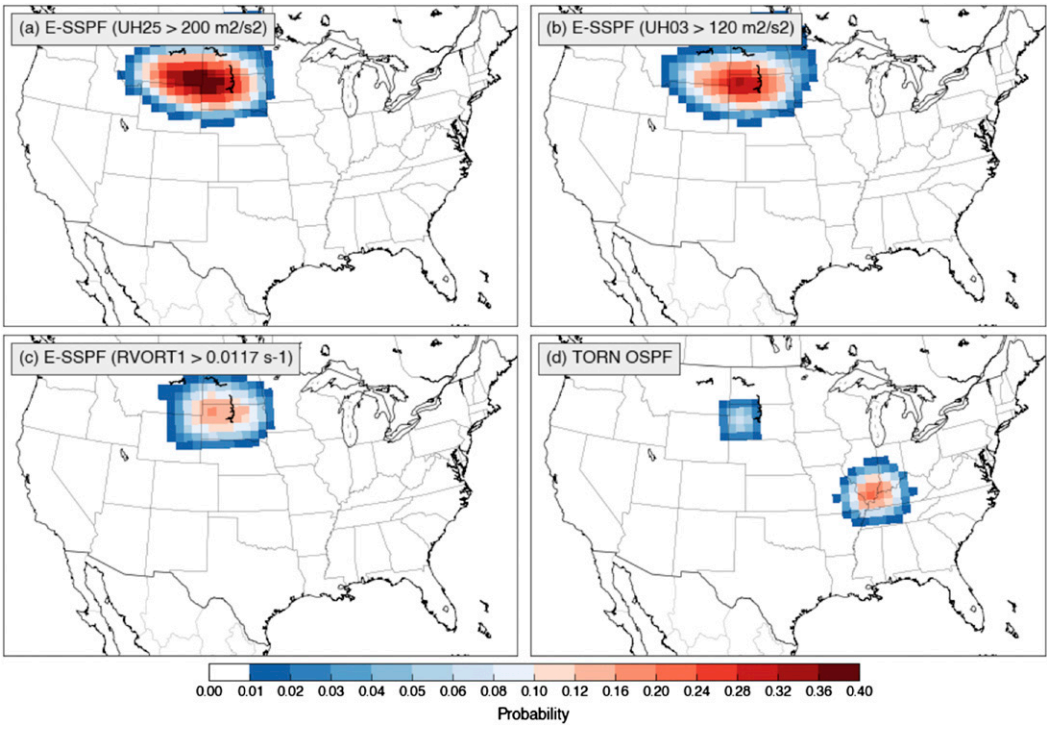


FIG. 16. As in Fig. 15, but for E-SSPFs valid 1200 UTC 19 Jun–1200 UTC 20 Jun 2015 from the ensemble forecast initialized at 0000 UTC 19 Jun 2015.

RVORT1 E-SSPFs are not statistically significant at these thresholds, using a bootstrap resampling between FSS differences, as in Sobash et al. (2016). In addition to the FSS gains for UH03 and RVORT1 E-SSPFs, receiver operating characteristic (ROC) areas for these forecasts were larger than the UH25 E-SSPFs at all smoothing length scales (Fig. 18).

It is worth noting that for all E-SSPFs, FSSs increase when the surrogate threshold is chosen to be slightly lower than the threshold that produces an SSR bias closest to 1 (Fig. 17). This result implies that overforecasting the number of SSRs (i.e., an SSR bias $> 1$) has a positive impact on forecast skill. This behavior appears to occur up to SSR biases of ~1.5 (Table 3), with FSS values decreasing for biases larger than 1.5. Given the small overall number of SSRs for these thresholds (generally 200–400 for biases between 1 and 1.5), the number of unique storms identified is likely very sensitive to the SSR threshold. Because of the small number of SSRs, a slight decrease in the UH25, UH03, or RVORT1 threshold could lead to small improvements in FSS by producing probabilities in new areas that were not identified using a higher threshold. This effect was not observed in Sobash et al. (2016), likely because of a larger overall number of SSRs and, thus, less sensitivity to small changes in threshold (their work used all severe reports rather than just tornado reports, resulting in a much larger number of SSRs).

In addition to gains in FSS and ROC area, the reliability of the E-SSPFs, shown in Fig. 19a for E-SSPFs using the bias = 1 SSR thresholds (Table 3) and $\sigma = 120$ km, also improves. These differences in reliability are most certain for probabilities $< 20\%$, where sample sizes are generally larger than 100 (Fig. 19b). For these forecasts, RVORT1 E-SSPFs produce fairly reliable probabilities, while the UH25 E-SSPFs suffer from overforecasting and hover near the "no skill" line. For E-SSPF probabilities $> 20\%$, the results are noisier, reducing confidence in the reliability results in this range, although there is some suggestion of improvement in reliability for the UH03 and RVORT1 E-SSPFs. For reference, the reliability of E-SSPFs created using UH25 $> 75\,m^2\,s^{-2}$, the traditional threshold for identifying supercells in the present forecast dataset, is provided in Fig. 19a. These forecasts are substantially overconfident and produce small FSSs.

The lack of sharpness in the present set of E-SSPFs is not ideal, but may not be a significant issue, since operational probabilistic forecasts of tornadoes (e.g., SPC probabilistic outlooks) also do not issue probabilities $>$ 30%. While the filtered UH25 forecasts produced in Gallo et al. (2016) did produce probabilities up to 80%

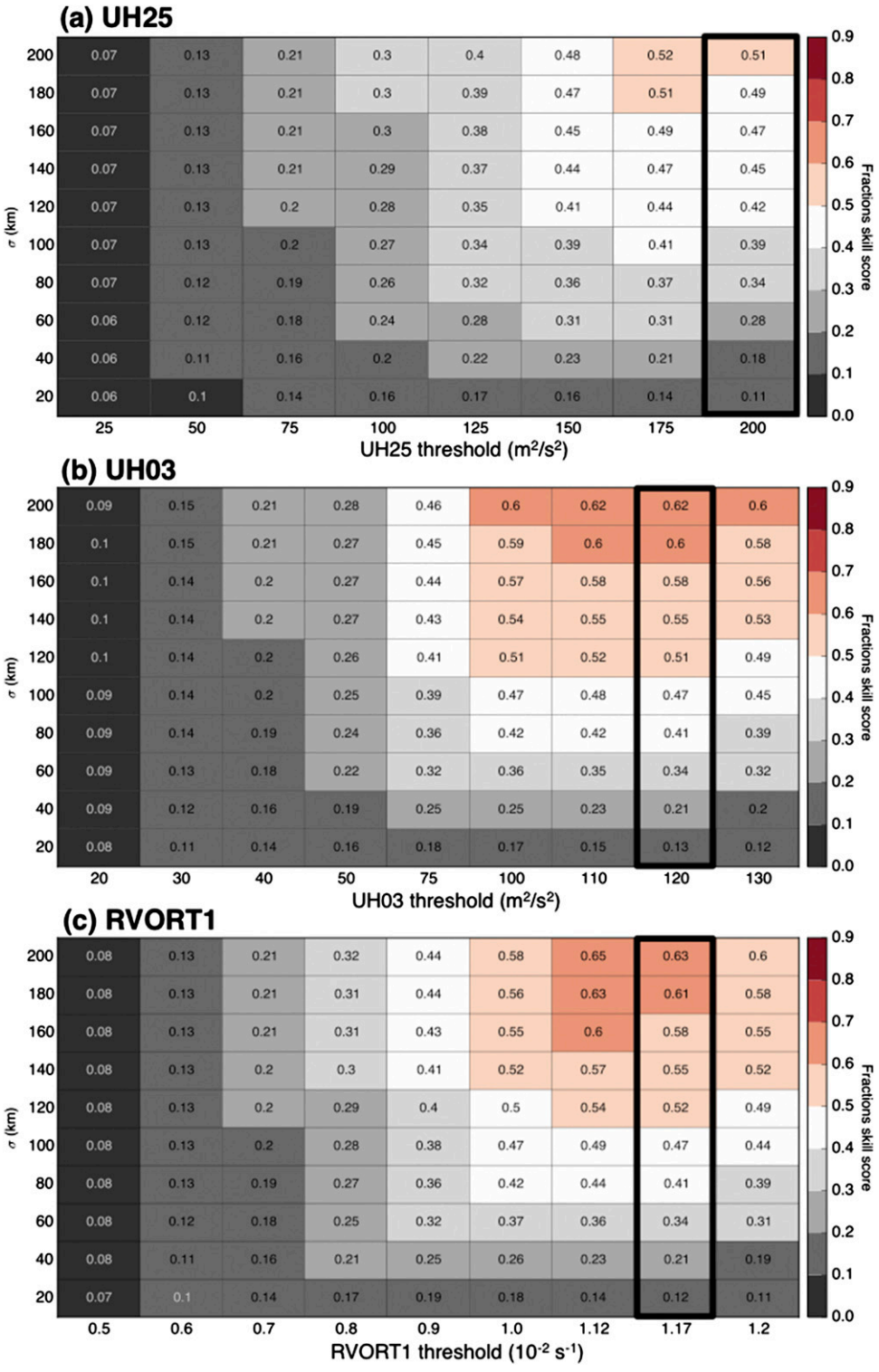


FIG. 17. FSSs for 24-h E-SSPFs (valid 1200–1200 UTC) over the 92-day forecast period (using forecasts initialized daily from 0000 UTC 30 Apr through 0000 UTC 30 Jul 2015) as a function of SSR threshold and smoothing length scale for (a) UH25, (b) UH03, and (c) RVORT1. Boldface boxes indicate the E-SSPFs that use the SSR thresholds that produce SSR biases closest to 1 (Table 3).

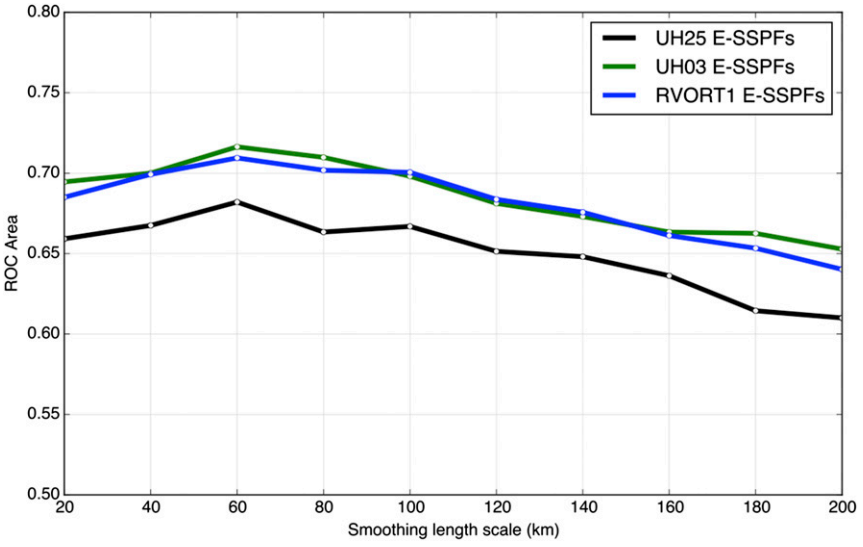


FIG. 18. AUC curves for UH25 E-SSPFs (black), UH03 E-SSPFs (green), and RVORT1 E-SSPFs (blue), as a function of smoothing length scale $\sigma$ (km). The E-SSPFs shown here use the SSR thresholds that produce SSR biases closest to 1 (Table 3).

(although with very small sample sizes), the calibrated probabilities of Jirak et al. (2014) did not exceed 40%, similar to this work. An additional season or two of data should reveal more about the ability of the forecast system to produce sharper E-SSPFs for tornadoes.

## 4. Summary and discussion

The goal of this work was to gauge the ability of CAMs, run with 3-km horizontal grid spacing and initialized at 0000 UTC, to provide useful information concerning the next-day (i.e., forecast hours 12–36) threat for tornadoes by directly sampling the CAM's forecasts of low-level rotation within intense convection. To do so, two diagnostic fields related to low-level rotation, 0–3 km AGL updraft helicity (UH03) and 1 km AGL relative vorticity (RVORT1), were extracted from CAM output in a manner analogous to other diagnostic fields designed to detect instances of convective extremes in CAMs (e.g., Kain et al. 2010). Using the CAM predictions of *low-level* rotation differs from prior work that relied on CAM predictions of *midlevel* rotation and environmental parameters (e.g., STP) to create tornado forecast guidance (e.g., Clark et al. 2012; Jirak et al. 2014; Gallo et al. 2016).

The potential for UH03 and RVORT1 to act as surrogates for tornadoes, without the need for additional environmental information, was established by directly examining the near-storm environments in which supercells (i.e., defined as storm objects with object-maximum UH25 $> 75\,m^2\,s^{-2}$) with appreciable low-level rotation developed. The forecasts were quite capable of producing large RVORT1 and UH03 magnitudes associated with supercells occurring within environments of large SHR01 and SRHEL01, as well as low SBLCLs. The aggregate environmental statistics of storms with weak versus strong low-level rotation presented here generally reproduced the differences in environments of tornadic and nontornadic supercells deduced in studies using observed or model-based proximity soundings (e.g., Rasmussen and Blanchard 1998; Brooks 2006). The STP was also a successful discriminator between simulated supercells with weak and strong low-level rotation, just as was observed in T03 for observed tornadic versus nontornadic storms.

While UH03 magnitudes were typically around half as large as UH25 magnitudes, the UH ratio (UH03/UH25) was used to isolate storms that deviated from this climatology. Supercells with the strongest low-level rotation (RVORT1 $> 0.015\,s^{-1}$) possessed UH ratios between 0.6 and 1.0, but rarely exceeded 1 (i.e., UH03 rarely exceeded UH25). Among these storms, UH25 and UH03 magnitudes often exceeded $150\,m^2\,s^{-2}$, and occurred within environments of strong instability (mean SBCAPE $> 2000\,J\,kg^{-1}$) and large magnitudes of shear. Those storms that did possess UH ratios $> 1$ had UH25 and UH03 magnitudes below $75\,m^2\,s^{-2}$, yet they often produced moderate to strong amounts of low-level rotation (RVORT1 $> 0.01\,s^{-1}$). These storms occurred in environments of weak instability (mean SBCAPE $< 500\,J\,kg^{-1}$), but with shear magnitudes similar to the environments of the traditional supercells containing

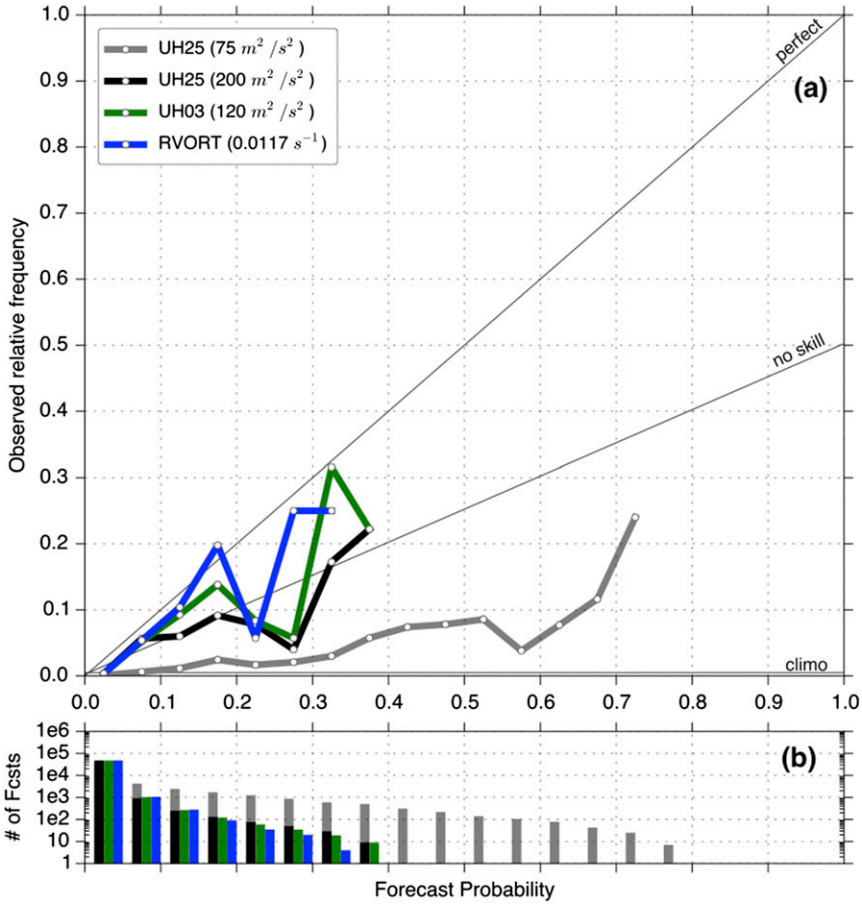


FIG. 19. (a) Attributes diagram for E-SSPFs for tornadoes using $\sigma = 120$ km for UH25 $> 75\,m^2\,s^{-2}$ (gray), UH25 $> 200\,m^2\,s^{-2}$ (black), UH03 $> 121\,m^2\,s^{-2}$ (green), and RVORT1 $> 0.017\,s^{-1}$ (blue). (b) Counts of forecasts per bin. Other than the E-SSPF using UH25 $> 75\,m^2\,s^{-2}$, the E-SSPFs shown here use the SSR thresholds that produce SSR biases closest to 1 (Table 3).

the strongest low-level rotation. The objects composing the storms with UH ratios $> 1$ were also smaller in size, suggesting the features were not persistent and likely were poorly resolved. Given the likelihood that these storms are overlooked by methods that use fixed UH25 thresholds to identify supercells (e.g., Sobash et al. 2011), future work should examine storms within the high-shear/low-CAPE/low-UH25 parameter space to determine if they are an important part of the next-day tornado prediction problem, and the potential usage of UH03 as a better diagnostic to identify these storms compared to UH25.

Finally, probabilistic tornado forecasts (E-SSPFs) were produced using the low-level rotation diagnostics as surrogates for tornadoes. A set of surrogate thresholds was computed using the methodology outlined in Sobash et al. (2016) by comparing SSRs to the number of observed tornado reports. Tornado E-SSPFs produced using UH03 and RVORT1 possessed higher FSSs, larger ROC areas, and were more reliable than E-SSPFs using UH25. The limited sample size of E-SSPFs above 20%, due in part to the smoothing necessary to produce large FSSs, precludes a definitive statement on the reliability of these forecasts until a larger sample of forecasts is produced.

The FSSs at which the tornado E-SSPFs were skillful were restricted to the largest scales examined ($\sigma \geq 160$ km), suggesting that CAMs have little to no skill at anticipating tornadoes on smaller scales, or that the low-level rotation surrogates are inherently less skillful than, say, those using UH25 as a surrogate to predict all severe weather hazards (e.g., Sobash et al. 2016), resulting in a need for additional smoothing to improve reliability. From a practical perspective, given the similar E-SSPF skill using both UH03 and RVORT1, it does not appear that the inclusion of the RVORT1 diagnostic in model output should be preferred over UH03, at least in models with similar horizontal grid spacing (i.e., 3 km).

We hope this work spurs several important areas of research. First, the low-level rotation diagnostics should be embedded within CAMs run with finer grid spacing than 3 km (e.g., 1 km) to compare with the present set of results. It is hypothesized that higher resolution will be needed to resolve near-surface rotation associated with mesovortices in systems such as QLCSs and may also improve the ability of the surrogates to anticipate tornado occurrence in traditional supercells. Second, the E-SSPFs need to be verified against radar-derived verification datasets, such as multiradar multisensor (MRMS) rotation tracks (Zhang et al. 2016), in addition to observed tornado reports. Using radar-derived rotation tracks would provide an important "apples to apples" comparison between model-predicted and observed low-level rotation. This comparison could be used as a baseline to distinguish between the ability of the forecasts to correctly place storms in space and time and the ability of the low-level rotation diagnostics to be used as surrogates for tornadoes.

Finally, although not examined here, any biases that exist in the simulated storm environment will inevitably lead to errors in forecasts of low-level rotation. These errors also impart limits on the predictability of storm mode, placement, and timing. Thus, efforts to improve the representation of the forecast near-storm environment should translate into gains in skill for E-SSPFs of tornadoes and all severe weather hazards.

*Acknowledgments.* This work was partially supported by National Oceanic and Atmospheric Administration Grant NA15OAR4590191. The real-time ensemble forecasts were run on Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. Most of the data analysis and visualization was accomplished using the packages within the scientific Python ecosystem (e.g. numpy, scipy, and matplotlib); we thank the many contributors to those projects for their efforts. The hagelslag software package was developed with funding from the National Science Foundation (Grant AGS-1261776). We also thank Dr. Harold Brooks for graciously providing the data from Brooks (2006) and used in Fig. 10.

## REFERENCES

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.

Brooks, H. E., 2006: A global view of severe thunderstorms: Estimating the current distribution and possible future changes. Preprints, *Symp. on the Challenges of Severe Convective Storms*, Atlanta, GA, Amer. Meteor. Soc., J4.2. [Available online at https://ams.confex.com/ams/Annual2006/techprogram/paper_102202.htm.]

——, C. A. Doswell III, and J. Cooper, 1994: On the environments of tornadic and nontornadic mesocyclones. *Wea. Forecasting*, **9**, 606–618, doi:10.1175/1520-0434(1994)009<0606:OTEOTA>2.0.CO;2.

Clark, A. J., J. S. Kain, P. T. Marsh, J. Correia Jr., M. Xue, and F. Kong, 2012: Forecasting tornado pathlengths using a three-dimensional object identification algorithm applied to convection-allowing forecasts. *Wea. Forecasting*, **27**, 1090–1113, doi:10.1175/WAF-D-11-00147.1.

——, J. Gao, P. Marsh, T. Smith, J. Kain, J. Correia, M. Xue, and F. Kong, 2013: Tornado pathlength forecasts from 2010 to 2011 using ensemble updraft helicity. *Wea. Forecasting*, **28**, 387–407, doi:10.1175/WAF-D-12-00038.1.

Craven, J. P., and H. E. Brooks, 2004: Baseline climatology of sounding derived parameters associated with deep, moist convection. *Natl. Wea. Dig.*, **28**, 13–24.

Davis, C., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, doi:10.1175/MWR3145.1.

Dawson, D. T., II, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140**, 696–716, doi:10.1175/MWR-D-11-00008.1.

Eastin, M. D., and M. C. Link, 2009: Miniature supercells in an offshore outer rainband of Hurricane Ivan (2004). *Mon. Wea. Rev.*, **137**, 2081–2104, doi:10.1175/2009MWR2753.1.

Edwards, R., A. R. Dean, R. L. Thompson, and B. T. Smith, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part III: Tropical cyclone tornadoes. *Wea. Forecasting*, **27**, 1507–1519, doi:10.1175/WAF-D-11-00117.1.

Fawbush, E. J., and R. C. Miller, 1954: The types of air masses in which North American tornadoes form. *Bull. Amer. Meteor. Soc.*, **35**, 154–166.

Gagne, D. J., II, A. McGovern, N. Snook, R. Sobash, J. Labriola, J. K. Williams, S. E. Haupt, and M. Xue, 2016: Hagelslag: Scalable object-based severe weather analysis and forecasting. *Proc. Sixth Symp. on Advances in Modeling and Analysis Using Python*, New Orleans, LA, Amer. Meteor. Soc., 447. [Available online at https://ams.confex.com/ams/96Annual/webprogram/Paper280723.html.]

Gallo, B. T., A. J. Clark, and S. R. Dembek, 2016: Forecasting tornadoes using convection-permitting ensembles. *Wea. Forecasting*, **31**, 273–295, doi:10.1175/WAF-D-15-0134.1.

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, doi:10.1175/2007WAF2006120.1.

Jirak, I. L., C. J. Melick, and S. J. Weiss, 2014: Combining probabilistic ensemble information from the environment with simulated storm attributes to generate calibrated probabilities of severe weather hazards. *Proc. 27th Conf. on Severe Local Storms*, Madison, WI, Amer. Meteor. Soc., 2.5. [Available online at https://ams.confex.com/ams/27SLS/webprogram/Paper254649.html.]

Johnson, A., X. Wang, F. Kong, and M. Xue, 2011: Hierarchical cluster analysis of a convection-allowing ensemble during the Hazardous Weather Testbed 2009 Spring Experiment. Part I: Development of the object-oriented cluster analysis method for precipitation fields. *Mon. Wea. Rev.*, **139**, 3673–3693, doi:10.1175/MWR-D-11-00015.1.

Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, doi:10.1175/WAF2007106.1.

——, 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, doi:10.1175/2010WAF2222430.1.

Lakshmanan, V., 2012: *Automating the Analysis of Spatial Grids: A Practical Guide to Data Mining Geospatial Images for Human and Environmental Applications*. Geotechnologies and the Environment, Vol. 6, Springer, 320 pp.

——, K. Hondl, and R. Rabin, 2009: An efficient, general-purpose technique for identifying storm cells in geospatial images. *J. Atmos. Oceanic Technol.*, **26**, 523–537, doi:10.1175/2008JTECHA1153.1.

Markowski, P. M., J. M. Straka, and E. N. Rasmussen, 2002: Direct surface thermodynamic observations within the rear-flank downdrafts of nontornadic and tornadic supercells. *Mon. Wea. Rev.*, **130**, 1692–1721, doi:10.1175/1520-0493(2002)130<1692:DSTOWT>2.0.CO;2.

——, and Coauthors, 2012: The pretornadic phase of the Goshen County, Wyoming, supercell of 5 June 2009 intercepted by VORTEX2. Part II: Intensification of low-level rotation. *Mon. Wea. Rev.*, **140**, 2916–2938, doi:10.1175/MWR-D-11-00337.1.

McCaul, E. W., Jr., 1991: Buoyancy and shear characteristics of hurricane-tornado environments. *Mon. Wea. Rev.*, **119**, 1954–1978, doi:10.1175/1520-0493(1991)119<1954:BASCOH>2.0.CO;2.

——, and M. L. Weisman, 1996: Simulations of shallow supercell storms in landfalling hurricane environments. *Mon. Wea. Rev.*, **124**, 408–429, doi:10.1175/1520-0493(1996)124<0408:SOSSSI>2.0.CO;2.

Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. *Wea. Forecasting*, **25**, 343–354, doi:10.1175/2009WAF2222260.1.

Naylor, J., M. S. Gilmore, R. L. Thompson, R. Edwards, and R. B. Wilhelmson, 2012: Comparison of objective supercell identification techniques using an idealized cloud model. *Mon. Wea. Rev.*, **140**, 2090–2102, doi:10.1175/MWR-D-11-00209.1.

Rasmussen, E. N., 2003: Refined supercell and tornado forecast parameters. *Wea. Forecasting*, **18**, 530–535, doi:10.1175/1520-0434(2003)18<530:RSATFP>2.0.CO;2.

——, and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148–1164, doi:10.1175/1520-0434(1998)013<1148:ABCOSD>2.0.CO;2.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, doi:10.1175/2007MWR2123.1.

Rotunno, R., 1979: A study in tornado-like vortex dynamics. *J. Atmos. Sci.*, **36**, 140–155, doi:10.1175/1520-0469(1979)036<0140:ASITLV>2.0.CO;2.

Schenkman, A. D., M. Xue, and M. Hu, 2014: Tornadogenesis in a high-resolution simulation of the 8 May 2003 Oklahoma City supercell. *J. Atmos. Sci.*, **71**, 130–154, doi:10.1175/JAS-D-13-073.1.

Schwartz, C. S., G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2015: NCAR's experimental real-time convection-allowing ensemble prediction system. *Wea. Forecasting*, **30**, 1645–1654, doi:10.1175/WAF-D-15-0103.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., doi:10.5065/D6DZ069T.

Skinner, P. S., C. C. Weiss, M. M. French, H. B. Bluestein, P. M. Markowski, and Y. P. Richardson, 2014: VORTEX2 observations of a low-level mesocyclone with multiple internal rear-flank downdraft momentum surges in the 18 May 2010 Dumas, Texas, supercell. *Mon. Wea. Rev.*, **142**, 2935–2960, doi:10.1175/MWR-D-13-00240.1.

Smith, B. T., R. L. Thompson, A. R. Dean, and P. T. Marsh, 2015: Diagnosing the conditional probability of tornado damage rating using environmental and radar attributes. *Wea. Forecasting*, **30**, 914–932, doi:10.1175/WAF-D-14-00122.1.

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, doi:10.1175/WAF-D-10-05046.1.

——, C. S. Schwartz, G. S. Romine, K. R. Fossell, and M. L. Weisman, 2016: Severe weather prediction using storm surrogates from an ensemble forecasting system. *Wea. Forecasting*, **31**, 255–271, doi:10.1175/WAF-D-15-0138.1.

Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499, doi:10.1175/2009BAMS2795.1.

Thompson, R. L., R. Edwards, and J. A. Hart, 2002: Evaluation and interpretation of the supercell composite and significant tornado parameters at the Storm Prediction Center. Preprints, *21st Conf. on Severe Local Storms/19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., J3.2. [Available online at https://ams.confex.com/ams/pdfpapers/46942.pdf.]

——, ——, ——, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261, doi:10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2.

——, B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, doi:10.1175/WAF-D-11-00116.1.

Trapp, R. J., and M. L. Weisman, 2003: Low-level mesovortices within squall lines and bow echoes. Part II: Their genesis and implications. *Mon. Wea. Rev.*, **131**, 2804–2823, doi:10.1175/1520-0493(2003)131<2804:LMWSLA>2.0.CO;2.

——, E. D. Mitchell, G. A. Tipton, D. W. Effertz, A. I. Watson, D. L. Andra Jr., and M. A. Magsig, 1999: Descending and nondescending tornadic vortex signatures detected by WSR-88Ds. *Wea. Forecasting*, **14**, 625–639, doi:10.1175/1520-0434(1999)014<0625:DANTVS>2.0.CO;2.

——, G. J. Stumpf, and K. L. Manross, 2005: A reassessment of the percentage of tornadic mesocyclones. *Wea. Forecasting*, **20**, 680–687, doi:10.1175/WAF864.1.

Wakimoto, R. M., H. Cai, and H. V. Murphey, 2004: The Superior, Nebraska, supercell during BAMEX. *Bull. Amer. Meteor. Soc.*, **85**, 1095–1106, doi:10.1175/BAMS-85-8-1095.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520, doi:10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2.

——, and ——, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, **112**, 2479–2498, doi:10.1175/1520-0493(1984)112<2479:TSACON>2.0.CO;2.

Weiss, S. J., and M. D. Vescio, 1998: Severe local storm climatology 1955–1996: Analysis of reporting trends and implications for NWS operations. Preprints, *18th Conf. on Severe Local Storms*, Minneapolis, MN, Amer. Meteor. Soc., 536–539.

Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, doi:10.1175/WAF-D-15-0043.1.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences: An Introduction*. 2nd ed. Academic Press, 627 pp.

Yussouf, N., D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**, 3044–3066, doi:10.1175/MWR-D-14-00268.1.

Zhang, J., and Coauthors, 2016: Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, doi:10.1175/BAMS-D-14-00174.1.