Composite Environments of Severe and Nonsevere High-Shear, Low-CAPE Convective Events

KEITH D. SHERBURN, MATTHEW D. PARKER, JESSICA R. KING, AND GARY M. LACKMANN

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina

(Manuscript received 1 May 2016, in final form 13 September 2016)

ABSTRACT

Severe convection occurring in environments characterized by large amounts of vertical wind shear and limited instability (high-shear, low-CAPE, or “HSLC,” environments) represents a considerable forecasting and nowcasting challenge. Of particular concern, NWS products associated with HSLC convection have low probability of detection and high false alarm rates. Past studies of HSLC convection have examined features associated with single cases; the present work, through composites of numerous cases, illustrates the attributes of “typical” HSLC severe and nonsevere events and identifies features that discriminate between the two. HSLC severe events across the eastern United States typically occur in moist boundary layers within the warm sector or along the cold front of a strong surface cyclone, while those in the western United States have drier boundary layers and more typically occur in the vicinity of a surface triple point or in an upslope regime. The mean HSLC severe event is shown to exhibit stronger forcing for ascent at all levels than its nonsevere counterpart. The majority of EF1 or greater HSLC tornadoes are shown to occur in the southeastern United States, so this region is subjected to the most detailed statistical analysis. Beyond the documented forecasting skill of environmental lapse rates and low-level shear vector magnitude, it is shown that a proxy for the release of potential instability further enhances skill when attempting to identify potentially severe HSLC events. This enhancement is likely associated with the local, in situ CAPE generation provided by this mechanism. Modified forecast parameters including this proxy show considerably improved spatial focusing of the forecast severe threat when compared to existing metrics.

1. Introduction

Environments characterized by limited CAPE and large vertical wind shear (colloquially referred to as high-shear, low-CAPE, or “HSLC,” environments) are common across the contiguous United States (Dean and Schneider 2008) and can be associated with severe convection, particularly during the cool season and overnight (e.g., Johns et al. 1993; Sherburn and Parker 2014). HSLC severe weather, including significant (EF2 or greater on the enhanced Fujita scale) tornadoes, occurs worldwide and is recognized as a considerable forecasting challenge (e.g., Hanstrum et al. 2002; Tyrrell 2007; Romero et al. 2007; Clark 2009, 2011; Gatzen et al. 2011; Inoue et al. 2011; Clark and Parker 2014; Mulder and Schultz 2015). To date, the peer-reviewed literature on HSLC convection within the United States is largely confined to case studies (e.g., Kennedy et al. 1993; Markowski and Straka 2000) and climatologies (Evans 2010; Davis and Parker 2014; Sherburn and Parker 2014), with a few related modeling studies primarily focused on tropical storm environments (McCaul and Weisman 1996, 2001). The goal of this work is to build upon the methods and metrics of recent HSLC studies [e.g., Sherburn and Parker (2014), explored in further detail shortly] and to identify features discriminating between severe and nonsevere HSLC episodes, thereby improving existing forecasting techniques.

Across the continental United States, HSLC environments are present considerably more frequently than “high-shear, high-CAPE” (HSHC) environments (Dean and Schneider 2008, their Fig. 1; Schneider and Dean 2008, their Fig. 3a). However, given the occurrence of lightning, the probability of severe weather is nearly an order of magnitude lower in HSLC environments than in environments with comparable deep-layer shear vector magnitudes but 2500 J kg$^{-1}$ of most unstable CAPE (MUCAPE; Dean et al. 2009, their Fig. 4). Accordingly, the skill of NOAA/Storm Prediction Center’s tornado watches is substantially lower in HSLC environments, with the false alarm area and fraction of missed tornadoes

Corresponding author e-mail: Keith Sherburn, kdsherbu@ncsu.edu

DOI: 10.1175/WAF-D-16-0086.1

© 2016 American Meteorological Society
maximized in this zone of the parameter space (Dean and Schneider 2012).

A large fraction\\(^1\) of U.S. tornadoes (and significant tornadoes) occur in HSLC regimes (e.g., Schneider et al. 2006; Guyer and Dean 2010; Dean and Schneider 2012). HSLC significant tornadoes tend to be concentrated in the Mississippi, Tennessee, and Ohio valleys (e.g., Sherburn and Parker 2014). These locations coincide with the typical locations of significant cool season and overnight tornadoes and tornado outbreaks (e.g., Vescio and Parker 2014). Additionally, given the compact spatial [tornadic vortex depths and diameters of approximately 2–4 km; Davis and Parker (2014)] and temporal dimensions of HSLC convection (e.g., McCaul 1987), resolving rotation may be challenging, delayed, or even impossible, particularly far from radar (e.g., McCaul 1993; Lane and Moore 2006; Davis and Parker 2014). Thus, the climatology of HSLC environments producing severe convection also exhibits considerable overlap with time periods known for poor tornado warning skill and correspondingly high normalized fatality totals (e.g., Ashley et al. 2008; Brotzge and Erickson 2010; Brotzge et al. 2014). Further, recent work highlights the low probability of detection (POD) and high false alarm rate (FAR) specifically in association with HSLC tornado warnings (Anderson-Frey et al. 2016).

Given the aforementioned operational challenges, there has been some emphasis on the large-scale environments accompanying HSLC severe weather. Figure 1 provides an example of the 4 March 2008 HSLC event, illustrating several features identified in prior case studies. HSLC environments producing severe convection tend to be associated with potent upper-level troughs, surface cyclones, and cold fronts (e.g., McAvoy et al. 2000; Cope 2004; Lane and Moore 2006; Wasula et al. 2008). Moisture is typically abundant in the low levels and through much of the troposphere, contributing to low lifted condensation levels (LCLs; as in Schneider et al. 2006 and Lane 2008; Fig. 1d). Winds over the depth of the troposphere tend to be intense, with low-level jets (LLJs) and deep-layer shear vector magnitudes commonly reaching upward of 25–30 m s\(^{-1}\) [e.g., Lane and Moore (2006); Konarik and Nelson (2008); note the 70–75-kt (1 kt = 0.51 m s\(^{-1}\)) 850-hPa jet in Fig. 1c]. Numerous case studies have also noted dry intrusions aloft coincident with the development or strengthening of convection (Kennedy et al. 1993; Lane and Moore 2006; Wasula et al. 2008; Evans 2010; Gatzen et al. 2011). These features could be indicative of a cold front aloft [e.g., Hobbs et al. (1990); note strong 850-hPa cold front in Fig. 1c] or an approaching upper-level jet streak (Gatzen et al. 2011; Fig. 1a). Either way, dry air aloft suggests the presence of potential instability, which could be released by ascent, thus supplementing the apparently low ambient CAPE values.

Because CAPE is by definition low in HSLC environments, forecasting parameters that include CAPE as a constituent [including the significant tornado parameter (STP; Thompson et al. 2003, 2004, 2012), energy helicity index (EHI; e.g., Davies 1993), vorticity generation parameter (VGP; e.g., Rasmussen and Blanchard 1998), and Craven–Brooks significant severe parameter (Craven and Brooks 2004)] struggle to adequately represent the risk for severe hazards in HSLC environments (Guyer and Dean 2010; Sherburn and Parker 2014). In fact, Sherburn and Parker (2014) showed that all of the aforementioned indices had their maximum skill values well below those values recommended as operational guidance (e.g., Rasmussen and Blanchard 1998; Craven and Brooks 2004; Thompson et al. 2004).

To address these concerns, Sherburn and Parker (2014) developed the severe hazards in environments with reduced buoyancy (SHERB) parameters. This combination of lapse rates (0–3 km and 700–500 hPa in the original formulation) and shear vector magnitudes (particularly 0–3 km and effective bulk shear magnitudes) was statistically shown to improve the discrimination between significant severe convection and nonsevere convection. However, because it does not directly rely on CAPE, SHERB is subject to large false alarm areas, as values often climb above its severe guidance threshold of 1 in environments where convection is not expected (Sherburn and Parker 2014; R. Thompson 2014, personal communication). This is of particular concern at locations where low-level lapse rates are climatologically steeper than those in the Southeast (for which the parameter was developed), such as in the plains, or in elevated terrain where the two lapse rate layers overlap. Additionally, values are often observed to rise above the threshold behind cold fronts. Other parameters, such as the combination of the lifted index, low-level shear vector magnitude, and near-surface convergence suggested by Hanstrum et al. (2002) or the TQ index (Henry 2000)—a modified version of the total totals index using the 700-hPa temperature rather than the 500-hPa temperature—have not been subject to rigorous statistical tests but have shown practical operational utility in other regions.

---

1 Guyer and Dean (2010) found that approximately 28% of all 2003–09 tornadoes occurred in environments characterized by MLCAPE \(\leq 500 \text{ J kg}^{-1}\). Schneider et al. (2006) showed that 54% of EF2 or greater tornadoes in 2004–05 were associated with MLCAPE \(\leq 1000 \text{ J kg}^{-1}\) and 0–6-km shear vector magnitude \(\geq 18 \text{ m s}^{-1}\).
Despite a recent uptick in the number of studies investigating HSLC environments, several gaps in the knowledge base associated with the dynamics and predictability (both long and short term) of severe HSLC convection remain. As a first step toward addressing these knowledge gaps, we seek an understanding of the climatology of severe HSLC convection within the scope of the entire severe weather spectrum. Next, we explore the range of synoptic patterns conducive to severe HSLC convection. We find that considerable information and context is added by focusing on spatial patterns of fields instead of individual gridpoint values (i.e., as explored by Sherburn and Parker 2014). Thorough interrogation of 3D reanalysis fields suggests a highly effective forecast parameter for the prediction of HSLC convection. A novel aspect of this parameter is a proxy measure for the in situ generation of instability through the release of potential instability. Results indicate that this parameter is a useful discriminator between severe events and nulls (the latter are defined in section 2b).
2. Data and methods

a. Definition of events

Nationwide EF1 or greater tornado reports and significant wind reports (wind gusts \( \geq 65 \text{ kt or } 33.44 \text{ m s}^{-1} \)) from 2006 to 2011 were gathered from the Storm Prediction Center (SPC). For each report, archived SPC Surface Objective Analysis [SFCOA; i.e., Mesoscale Analysis or Mesoanalysis; Bothwell et al. (2002)] data for several fields at the nearest grid point and previous hour were available. For this work, the SFCOA data were only utilized to determine if the report occurred within an HSLC environment [i.e., did it meet our HSLC criteria, as defined in Sherburn and Parker (2014): surface-based CAPE (SBCAPE) \( \geq 500 \text{ J kg}^{-1} \), MUCAPE \( \leq 1000 \text{ J kg}^{-1} \), and 0–6-km bulk wind difference \( \geq 18 \text{ m s}^{-1} \)]. As will be discussed in section 2c, the majority of analysis within this work is performed utilizing North American Regional Reanalysis (NARR; Mesinger et al. 2006) data. Comparisons between SFCOA and NARR data suggest the two datasets can vary considerably for a given event, particularly for parameters relevant to this study [see and compare scatterplots for SBCAPE and mean sea level pressure (PMSL) in Fig. 2]. Detailed verification of the representativeness of SFCOA versus NARR is beyond the scope of this work (given that there are few, if any, “ground truth” measurements of thermodynamic and wind profiles in HSLC environments). However, given the aforementioned differences and the observation that NARR CAPE tended to be higher than that of the SFCOA [consistent with the positive bias noted by Gensini et al. (2014)], we used an additional constraint requiring NARR SBCAPE and mixed-layer CAPE (MLCAPE) \( \geq 1000 \text{ J kg}^{-1} \) to ensure that truly high-CAPE environments were excluded. Reports that met the above criteria were included in our HSLC dataset.

We chose to exclude significant hail reports, which composed a small percentage of the overall dataset, because of concerns that some of these may have been associated with elevated convection that bypassed our HSLC thresholds (as discussed in Sherburn and Parker 2014). However, analysis including significant hail reports showed no appreciable differences. In addition, EF0 tornado reports and nonsignificant wind reports were excluded from our analyses. Generally, these reports are considered to be less reliable because of the nonmeteorological factors affecting their inclusion in the official Storm Data database (e.g., Brooks et al. 2003;
Smith et al. 2013), including diurnal occurrence, population density (e.g., Doswell et al. 2009), land-use dependencies (Weiss et al. 2002; Trapp et al. 2006), and poorly estimated wind speeds (Doswell et al. 2005). Further discussion on these caveats can be found in Smith et al. (2012) and Sherburn and Parker (2014).

The authors acknowledge that this approach excludes a portion of the convective spectrum, including measured, nonsignificant wind reports; the present work focuses on high-end tornado and wind events that ultimately have the largest impact on life and property. Hereafter, the 1447 reports meeting these criteria (Fig. 3a; separated by season in Figs. 3b–e) will be referred to as “events.”

b. Definition of nulls

Because binary forecasting statistics can be sensitive to how a null case is defined, two different null datasets (here, generally referring to nonsevere convection) were utilized in this study for comparison. The first set of nulls is defined as in Sherburn and Parker (2014): “the initial latitude–longitude point [i.e., the subjectively defined location of the storm’s area of interest, as determined by the NWS warning meteorologist] of a severe thunderstorm or tornado warning that was issued in an HSLC environment [as defined in section 2a] when there were no severe reports from the Storm Data archives in the corresponding CWA [NWS county warning area] throughout that convective day (1200–1200 UTC).” As in that work, the null dataset spans October 2006–April 2011 and encompasses the contiguous United States. The SFCOA and NARR CAPE and shear constraints were again used to determine which nulls occurred within an HSLC environment. Using these criteria, 925 nulls were identified, and these are hereafter referred to as “warning nulls.”

The primary concern in basing null identifications on warnings is their inherent subjectivity; the severity of a given cell is judged by the forecaster on shift, and he or she may have a warning threshold that differs from that of a peer. To address such concerns of objectivity, the second set of nulls is defined based upon the WSR-88D’s reports and excluding the kernel density estimation. (c) As in (b), but for summertime reports. (d) As in (b), but for autumn reports. (e) As in (b), but for wintertime reports. Insets show pie charts for percentages of reports that fall in the three subjectively defined regions, as demarcated by the green lines in (a): Southeast (SE), Northeast (NE), and West (W). Counts are shown at the bottom right of each panel.
storm cell identification and tracking algorithm (SCIT; Johnson et al. 1998), which identifies cells based upon given spatial and temporal thresholds of radar reflectivity and predicts their subsequent motion. The purpose of this alternative null dataset is to identify nonsevere convection without the requirement of warning issuance. To this end, only the diagnostic portion of the SCIT (i.e., the identification of cells) was utilized, and a minimum reflectivity threshold of 45 dBZ was enforced. SCIT data were gathered for January–April and October–December in 2011, encompassing the typical period of HSLC convective activity across the southeastern United States. Only radars in the Southeast (Fig. 4) and cells within 5 min of a NARR data time (e.g., for 0300 UTC, data from 0255 to 0305 UTC) were utilized for this null subset.

To ensure a clean radar-based null population, only cells clearly lacking severe weather reports (no reports in their CWA or any adjacent CWA during the given convective day) were retained for analysis. Then, the aforementioned SFCOA and NARR filters were again utilized to ensure the detected cells were within HSLC environments. Additionally, to ensure that non-convective cells were not included, minimal buoyancy (i.e., at least 10 J kg\(^{-1}\) of CAPE in SFCOA or NARR) was required. Finally, only the cell with the highest reflectivity for a given NARR data time was retained to prevent oversampling of a given environment. Using this definition, 1301 nulls were identified. These are hereafter referred to as “radar-based nulls.”

c. Composites

Composite maps and soundings were generated using NARR data from the National Centers for Environmental Information (NCEI; formerly NCDC). The NARR has a horizontal grid spacing of approximately 32 km with 29 vertical levels and is available every 3 h. Though its resolution is not ideal, recent studies (e.g., Gensini and Ashley 2011; Walters et al. 2014) suggest that the NARR is practical for climatological analyses, even for mesoscale phenomena. Further, the NARR has several research benefits over the SFCOA, including its full three-dimensionality and the ability to derive additional variables from its grids. For each report or null, a 20\(^\circ\) latitude \(\times\) 20\(^\circ\) longitude box centered on the report or null was created using the nearest NARR data time. Data within these boxes were then averaged over the number of sample times to create report- or null-relative composites.

Regional, seasonal, and diurnal subsets of the HSLC population of events and warning nulls were created to further assess potential differences. Regions were defined by latitude and longitude, as shown in Fig. 3a. While these regions were defined subjectively, it is shown later that the majority of the HSLC activity occurs in the Southeast subset, while cases in the West exhibit markedly dissimilar synoptic-scale characteristics to those in the East. Spring, summer, autumn, and winter were defined as March–May, June–August, September–November, and December–February, respectively. Daytime cases were those occurring between 0800 and 1700 LST, while nighttime cases were limited to the time period of 2000–0500 LST. Compound subsets (e.g., nighttime winter events) were also created to determine the seasonal differences in diurnal variability.

3. HSLC severe climatology and composites

Previous studies have noted that HSLC events peak in the cool season and overnight (e.g., Evans 2010; Guyer and Dean 2010; Sherburn and Parker 2014). Our data further show that the majority of significant wind reports and EF1 or greater tornadoes from winter through midspring occur within HSLC environments, particularly during the overnight and morning hours (Fig. 5). As was shown by Sherburn and Parker (2014), HSLC significant wind reports and EF1 or greater tornadoes are most common in the Ohio, Tennessee, and Mississippi valleys but can occur in any region (Fig. 3a). A clear
annual shift in this distribution is apparent, with an increasing number of reports in northern and western portions of the country during the warm season (Figs. 3b–e). However, the majority of HSLC events occur in the Southeast during the winter or spring (Figs. 3b,e).

Comparing the three regions via typical synoptic scale fields (Fig. 6), it becomes apparent that western HSLC events are fundamentally dissimilar from those in the Southeast and Northeast. The composite event in the West is characterized by weaker upper-level flow (cf. Figs. 6a,b, 6d,e, and 6g,h) and appears closer to a surface triple-point or upslope setup than the warm sector/cold front-driven pattern common in the East (cf. Figs. 6c,f,i). This is likely attributable to both the relative climatology of events in the West (biased toward the warm season) and greater case-to-case variability in the region. As noted in section 1, lapse rates are climatologically steeper in the West than the East; however, lapse rates across the West are also far less useful indicators than in the East (Fig. 7), with steeper lapse rates observed in the null environments than in the events (Figs. 7g–i). This suggests that techniques developed to improve the forecasting of HSLC environments utilizing lapse rates as ingredients [e.g., Sherburn and Parker’s (2014) SHERB parameters, reviewed in section 1] may lack utility across the West and may be more subject to large false alarm areas in this region. Additionally, although compositing vertical profiles smooths features considerably, western soundings (interpolated to the composite center) are characterized by much lower relative humidity in the lower troposphere than in the East (Fig. 8), leading to higher LCLs and larger downdraft CAPE. Hodographs in the West are also relatively straight compared to those in the East (Fig. 8). These sounding features corroborate the comparatively higher percentage of straight-line winds in the western U.S. sample of HSLC significant severe reports (approximately 50% of eastern U.S. events are straight-line wind reports compared to 83% of western U.S. events).

Given the apparent differences between western and eastern HSLC events and our desire to focus on the cool season phenomenon that appears to be relatively absent from the West, we exclude western cases from our subsequent analysis. Additionally, because Southeast and Northeast events appear to be fundamentally rather similar (cf. soundings in Fig. 8 and respective panels in Figs. 6 and 7), we chose to focus on the Southeast subset, which encompasses the majority of the density “bull’s-eye” exhibited in Fig. 3a and contains the vast majority of significant tornadoes. Focusing the discussion on a single region allows for easier segmentation of the data and a clearer depiction of what processes and features show operational utility at discriminating between events and nulls. The Northeast subset returns later as an independent verification dataset following the development of new forecasting techniques using the Southeast subset.

---

*4 Recall that the composite maps are presented within a report- or null-relative framework. Background maps are primarily meant to provide a sense of scale, though they also show the mean report or null location for each subset.*
Fig. 6. Regional comparisons of (a), (d), (g) 300-hPa winds (shaded; kt), wind barbs (kt), and geopotential heights (black contours, every 120 m); (b), (e), (h) 500-hPa absolute vorticity (shaded; $10^{-5}$ s$^{-1}$), wind barbs (kt), geopotential heights (black contours, every 60 m), and 700-hPa omega (blue contours; μbar s$^{-1}$); and (c), (f), (i) 2-m dewpoint (shaded; °C), mean sea level pressure (black contours, every 2 hPa), and 10-m wind barbs (kt). Plots are shown for the (top) Southeast, (middle) Northeast, and (bottom) West regions. Maps are shown for a reference scale, with the white dot depicting the event-relative composite center point and the average latitude and longitude of each subset. The number of times in each composite is shown in (a), (d), and (g) at bottom right.
Fig. 7. Regional comparisons of 0–3-km lapse rates (K km$^{-1}$) for (a),(d),(g) severe events, defined as EF1 or greater tornadoes and significant wind reports (wind gusts ≥ 65 kt); (b),(e),(h) warning nulls; and (c),(f),(i) the difference field between events and nulls. All else is as in Fig. 6.
There are several recurring themes in the Southeast cases. First, events are characterized by especially strong synoptic-scale forcing for ascent, including intense upper-level divergence and low-level convergence, a potent upstream vorticity maximum at 500 hPa, strong low-level warm-air advection, and a surface cyclone centered just north of the composite center (Figs. 9a,d,g,j). A coupled jet feature aloft is also noted, with the composite report occurring near the right-entrance region of a northern jet streak and the exit region of a southern jet streak. The warning null composite—centered approximately 100 km south of the event composite—reveals similar features (e.g., a coupled jet feature aloft, an upstream trough at all levels, upper-level divergence atop low-level convergence, a closed surface cyclone), but these are displaced slightly northwest with respect to the composite center and have substantially weaker magnitudes, particularly in the lower levels (Figs. 9b,e,h,k). On the other hand, radar-based nulls show a similar jet streak focused north of the composite center, but the upper-level divergence, low-level convergence, and surface cyclone are much more benign (Figs. 9c,f,i,l). Radar-based nulls further lack any hint of a southern upper-level jet streak and, compared with the warning nulls, show a considerably less amplified lower- and upper-level trough. Combined, these features contribute to much weaker synoptic-scale forcing for ascent in the radar-based nulls, as reflected in 700-hPa $v$ (vertical velocity with respect to pressure) fields. Overall, on the mean, there is generally a decrease in forcing strength from severe events to warning nulls to radar-based nulls.

Instability is inherently limited in both the events and nulls\(^5\) (Figs. 9j–l); however, a tongue of enhanced low-level instability is apparent in 0–3-km surface-based CAPE, lapse rate, and lifted index fields in the events (Figs. 10a–i). While the zone of enhanced lapse rates extends well upstream of the composite event center (Figs. 10d–f), relatively large magnitudes of low-level CAPE are confined to areas near and just upstream of the event center, with a local maximum apparently

\(^5\) Note that SBCAPE near the composite report is actually above the HSCL threshold of 500 J kg\(^{-1}\), representing differences between the NARR and SFCOA, as shown in Fig. 2.
FIG. 9. Mean (a)–(c) 300-hPa divergence (shaded; $10^{-5}$ s$^{-1}$), geopotential heights (black contours, every 120 m), isolachys (purple contours, every 5 kt), and wind barbs (kt); (d)–(f) 500-hPa absolute vorticity (shaded; $10^{-5}$ s$^{-1}$), geopotential heights (black contours, every 60 m), wind barbs (kt), and 700-hPa omega (blue contours; µbar s$^{-1}$); (g)–(i) 850-hPa divergence (shaded; $10^{-5}$ s$^{-1}$), geopotential heights (black contours, every 30 m), temperature advection (blue and red contours; K s$^{-1}$), and wind barbs (kt); and (j)–(l) SBCAPE (shaded; J kg$^{-1}$), mean sea level pressure (black contours, every 2 hPa), and 10-m wind barbs for the Southeast (left) EF1 or greater tornadoes and significant wind reports, (center) warning nulls, and (right) radar-based nulls. Maps are shown for a reference scale, with the white dot depicting the event-relative composite center point and the average latitude and longitude of each subset. The number of times in each composite is shown in (a)–(c) at bottom right.
FIG. 10. Mean of (a)–(c) 0–3-km SBCAPE (J kg$^{-1}$), (d)–(f) 0–3-km lapse rate (K km$^{-1}$), (g)–(i) 3-km surface-based lifted index (K), and (j)–(l) 3–5-km lapse rate (K km$^{-1}$) for Southeast (left) EF1 or greater tornadoes and significant wind reports, (center) warning nulls, and (right) radar-based nulls. All else is as in Fig. 9.
detached from the region of higher CAPE to the south (Figs. 10a–c). This zone is oriented roughly linearly along the mean surface trough, suggesting a narrow unstable sector ahead of an approaching cold front. Large interquartile ranges relative to mean values (not shown), however, suggest that the distribution of low-level CAPE values (in particular) within HSLC environments is wide, meaning that this feature—though noteworthy on the mean—may not be apparent in all events and may be present in some nulls, as well.

Low-level lapse rate distributions have relatively small interquartile ranges (not shown), and the lapse rates are overall less variable than low-level CAPE. Given that low-level lapse rates tend to be steeper in events than in nulls (Fig. 10), values of SHERB utilizing the 0–3-km bulk wind difference (SHERBS3) and effective bulk wind difference (SHERBE) are accordingly higher in events (Fig. 11). Compared to the other two SHERB ingredients, the midlevel lapse rate term (using the 3–5-km layer) appears to be a less practical discriminator between events and nulls (Figs. 10j–l); in fact, in several subsets, this midlevel lapse rate is higher in nulls than events. This suggests that 1) the midlevel lapse rate’s skill is conditional, perhaps only in a portion of the parameter space, and 2) a modified version of the SHERB parameters with new ingredients replacing or supplementing the midlevel lapse rate may be more operationally advantageous.

The orientation of locally enhanced low-level CAPE relative to a zone of enhanced synoptic-scale forcing for ascent suggests the possibility that potential instability is being released in situ as forcing arrives in a given location. To address this hypothesis, a product of $\frac{d\theta_e}{dz}$ (i.e., the change of equivalent potential temperature with height, indicating potential instability when negative) and $\omega$ was calculated and plotted over several layers. As shown in Fig. 12, there is a robust discriminatory signal in this product, particularly when comparing events and radar-based nulls. Additionally, there is reasonable overlap between the maxima in these fields and the enhanced area of low-level CAPE, implying that the release of potential instability could be the mechanism responsible for this feature. The practical utility of these features and ingredients in discriminating between
FIG. 12. Mean of (a)–(c) product of 3-km $\omega$ and $0$–$3$-km $d\theta_v/\text{dz}$ (K Pa km$^{-1}$ s$^{-1}$), (d)–(f) product of 4.5-km $\omega$ and $0$–$4.5$-km $d\theta_v/\text{dz}$ (K Pa km$^{-1}$ s$^{-1}$), (g)–(i) product of 6-km $\omega$ and $0$–$6$-km $d\theta_v/\text{dz}$ (K Pa km$^{-1}$ s$^{-1}$), and (j)–(l) the maximum $\omega$ and $d\theta_v/\text{dz}$ product from $0$–$2$ km through $0$–$6$ km (K Pa km$^{-1}$ s$^{-1}$) for Southeast (left) EF1 or greater tornadoes and significant wind reports, (center) warning nulls, and (right) radar-based nulls. All else is as in Fig. 9.
events and nulls will be assessed within a more rigorous quantitative framework in section 4.

Individual ingredients show substantial variability through the diurnal cycle. There is clear evidence of steeper lapse rates during the day and stronger shear vector magnitude at night, regardless of whether the event or null dataset is being investigated (Fig. 13). There is also a variability (albeit weaker) in both parameters on the seasonal cycle. This implies that a single value or threshold of lapse rates, for example, is unlikely to be equally skillful for nighttime versus daytime cases. However, forecasting indices that include several individual ingredients are somewhat more resistant to seasonal and diurnal variability. One such combination is the STP (Thompson et al. 2012). A “reservoir” of enhanced effective STP values (albeit still below the operational guidance value of 1) is observed upstream of the composite event in all six diurnal/annual cycle event subsets, while such a feature is practically absent in all null subsets (Fig. 14). These findings suggest that multi-ingredient forecast indices have a higher potential to be consistently discriminatory regardless of season or time of day.

4. Skill score tests

In addition to the creation of report-relative composites, NARR data were used in a series of skill score tests to determine the most statistically skillful combination of ingredients among a selection of fields included within NARR grids and additional parameters calculated from them. A set of over 250 potential ingredients was assessed. Here, each ingredient or combination of ingredients was tested independently for its ability to discriminate between HSLC events and warning nulls within the Southeast. Four possible formulations of each ingredient \( X \) were tested:

1) \( X - a \), or the ingredient minus some constant value \( a \);
2) \( X \), or the ingredient itself;
3) \( b - X \), or the ingredient subtracted from some constant value \( b \); and
4) \( X^2 \), or the ingredient squared.

This builds upon the work of Sherburn and Parker (2014), who only used formulation 2 and who applied a tiered skill score test that first identified the most skillful ingredient, then explored the conditional skill of additional ingredients.

Herein, only data from the Southeast subset were utilized, and skill was defined by correctly discriminating between events and warning nulls. This analysis was also performed on radar-based nulls. However, comparing radar-based nulls with events may be of less practical utility, given that the former were not necessarily associated with convection that subjectively appeared severe. The maximum (or minimum, depending on the particular ingredient) gridpoint value in an approximate 160 km \( \times \) 160 km box surrounding the composite center was utilized in the skill score tests in an attempt to capture the most representative fields available given that small-scale evolution at a point may not be captured by the relatively coarse temporal resolution of the NARR. The two compound metrics used to determine skill were the Heidke skill score (HSS), calculated by

\[
HSS = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)},
\]

and true skill statistic (TSS; Wilks 1995), given by

\[
TSS = \frac{(ad - bc)}{(a + c)(b + d)},
\]

where \( a \) is a correct forecast of an event, \( b \) is a false alarm, \( c \) is a missed event, and \( d \) is a correct null. Preference was given to the HSS based upon its consistency regardless of event rarity (Doswell et al. 1990). Traditional forecasting metrics such as POD and FAR were also calculated.

Initial tests determined the skill of individual ingredients using the entire set of parameters. Subsequent tests then assessed the skill of multi-ingredient combinations of parameters consisting of those that had exhibited skill on their own or in conjunction with one other ingredient. Tests with multiple ingredients allowed for the inclusion of more than one term of a given type but were subjectively checked to avoid the redundancy of ingredients. It is possible that some ingredients or combinations of parameters exhibit statistical skill but little practical skill. For example, PMSL may be statistically skillful, as events tend to be associated with deeper surface cyclones than nulls; however, in practice, this would imply that the threat would increase along a cold front as one travels closer to the cyclone center, which is not necessarily true. Thus, only those composite parameters judged to have a robust physical basis were considered for further analysis.

As shown in Table 1, the 0–3-km lapse rate was the most skillful individual thermodynamic ingredient,
Fig. 13. The 0–3-km lapse rate (shaded; K km$^{-1}$) and 0–3-km bulk shear vector (barbs; kt) for (left) autumn, (center) winter, and (right) spring (a)–(c) daytime events, as defined in the text; (d)–(f) daytime warning nulls; (g)–(i) nighttime events; and (j)–(l) nighttime warning nulls within the Southeast subset. Counts for the respective subsets are provided in the bottom right of each panel.
FIG. 14. As in Fig. 13, but showing STP. Black contour denotes an STP value of 0.25.
consistent with findings from Sherburn and Parker (2014). Most of the other independent terms that exhibited skill were kinematic fields, including the effective shear vector magnitude [also consistent with Sherburn and Parker (2014)] and its zonal component, as well as the meridional component and magnitude of the 10-m wind speed (perhaps highlighting the importance of low-level $u_e$ advection). Finally, frontogenesis within the 750–700-hPa layer was also skillful, perhaps suggesting that the depth of linear forcing or the presence of a cold front aloft may be important for determining the potential severity of an event.

The SHERBE and SHERBS3 parameters developed by Sherburn and Parker (2014) (here using the recommended 3–5-km lapse rate instead of the 700–500-hPa lapse rate) showed the most skill among existing combination indices [a group that also included STP, the supercell composite parameter (Thompson et al. 2003, 2004), the Craven–Brooks significant severe parameter, 0–3-km EHI, and 0–3-km VGP], followed by several low-level turbulent kinetic energy (TKE) terms (Table 1). Although statistically skillful, TKE terms are dependent on model configuration and parameterizations (here, these terms are outputted from the NARR grid, not calculated manually). As a result, these terms were excluded from further investigation. However, in general, large low-level TKE would result from large low-level lapse rates and low-level vertical shear (e.g., low Richardson numbers). In this sense, it provides rather similar information to SHERBS3 and SHERBE. Products of $\omega$ and $d\theta_e/\partial z$, meant to approximate the release of potential instability, also showed relatively high skill, consistent with the spatial footprint noted in Fig. 10. While $\omega$ is also dependent on model configuration (primarily grid spacing), it is an explicitly represented

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Max HSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zonal component of effective bulk shear vector</td>
<td>0.329</td>
</tr>
<tr>
<td>Zonal component of 0–2-km shear vector</td>
<td>0.302</td>
</tr>
<tr>
<td>Planetary boundary layer height</td>
<td>0.294</td>
</tr>
<tr>
<td>725-hPa frontogenesis</td>
<td>0.292</td>
</tr>
<tr>
<td>Effective shear vector magnitude</td>
<td>0.291</td>
</tr>
<tr>
<td>750-hPa frontogenesis</td>
<td>0.288</td>
</tr>
<tr>
<td>10-m wind magnitude</td>
<td>0.277</td>
</tr>
<tr>
<td>Meridional component of 10-m wind vector</td>
<td>0.275</td>
</tr>
<tr>
<td>700-hPa frontogenesis</td>
<td>0.273</td>
</tr>
<tr>
<td>0–3-km lapse rate</td>
<td>0.262</td>
</tr>
</tbody>
</table>

Combination index

<table>
<thead>
<tr>
<th>Combination index</th>
<th>Max HSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHERBE</td>
<td>0.374</td>
</tr>
<tr>
<td>SHERBS3</td>
<td>0.366</td>
</tr>
<tr>
<td>950-hPa TKE</td>
<td>0.347</td>
</tr>
<tr>
<td>900-hPa TKE</td>
<td>0.335</td>
</tr>
<tr>
<td>925-hPa TKE</td>
<td>0.332</td>
</tr>
<tr>
<td>875-hPa TKE</td>
<td>0.284</td>
</tr>
<tr>
<td>700-hPa $\omega \times 1000–700$-hPa $d\theta_e/\partial z$</td>
<td>0.257</td>
</tr>
<tr>
<td>4.5-km $\omega \times 0–4.5$-km $d\theta_e/\partial z$</td>
<td>0.248</td>
</tr>
<tr>
<td>3-km $\omega \times 0–3$-km $d\theta_e/\partial z$</td>
<td>0.247</td>
</tr>
<tr>
<td>500-hPa $\omega \times 1000–500$-hPa $d\theta_e/\partial z$</td>
<td>0.246</td>
</tr>
</tbody>
</table>

Fig. 15. Box-and-whisker plots for maximum (within a 160 km × 160 km box surrounding the respective report or null) LLLR, S15MG, ESHR, and MAXTEVV term contributions within MOSHE (as defined in Eq. (3); e.g., [(LLLR – 4$^2$)/4 K$^2$ km$^{-2}$] across Southeast events, warning nulls, and radar-based nulls. Purple line shows the median value, while the blue box encompasses the 25th–75th (q1–q3) percentile values. Black whiskers extend to values in the range [q1 – 1.5(q3 – q1), q3 + 1.5(q3 – q1)]. Outliers are excluded.
field (unlike TKE, which is parameterized). Thus, it was retained for the multi-ingredient tests.

Low-level lapse rates and bulk wind differences appear in nearly every combination of skillful ingredients, suggesting they are crucial in discriminating between HSLC events and nulls. Steep low-level lapse rates can promote downward momentum transfer of strong flow aloft (e.g., Johns and Hirt 1987; Johns 1993; Evans and Doswell 2001) or the strengthening of low-level vortices (e.g., Parker 2012). Meanwhile, larger low-level shear vector magnitudes also promote the strengthening of low-level mesocyclones (e.g., Markowski and Richardson 2014; Coffer and Parker 2015) or mesovortices (e.g., Weisman and Trapp 2003) and are representative of stronger environmental flow that could potentially be transported to the surface via convective downdrafts. Combining the two terms, rather than looking at each in isolation, helps preserve their discriminatory skill while diminishing each term’s respective diurnal cycle (as shown in Fig. 13). In particular, the effective shear vector magnitude appeared in many skillful combinations. The effective layer is approximately defined as the lower half of the vertical layer exhibiting CAPE (Thompson et al. 2007); the inclusion of the effective bulk shear magnitude therefore adds an inherent CAPE mask that may be a useful contributor to an index lacking an explicit buoyancy term.

### Table 2: Skill scores and forecast skill metrics for SHERBS3, SHERBE, MOSH, and MOSHE at discriminating between HSLC events and warning nulls within the entire Southeast dataset.

<table>
<thead>
<tr>
<th>Index</th>
<th>Max HSS</th>
<th>Max TSS</th>
<th>POD at max HSS</th>
<th>FAR at max HSS</th>
<th>Max HSS threshold</th>
<th>HSS at threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHERBS3</td>
<td>0.366</td>
<td>0.363</td>
<td>0.861</td>
<td>0.515</td>
<td>0.90</td>
<td>0.343</td>
</tr>
<tr>
<td>SHERBE</td>
<td>0.374</td>
<td>0.415</td>
<td>0.690</td>
<td>0.275</td>
<td>0.80</td>
<td>0.222</td>
</tr>
<tr>
<td>MOSH</td>
<td>0.402</td>
<td>0.446</td>
<td>0.799</td>
<td>0.392</td>
<td>1.00</td>
<td>0.402</td>
</tr>
<tr>
<td>MOSHE</td>
<td>0.377</td>
<td>0.412</td>
<td>0.709</td>
<td>0.297</td>
<td>1.00</td>
<td>0.377</td>
</tr>
</tbody>
</table>

**Fig. 16.** Mean (a)–(c) MOSH and (d)–(f) MOSHE for Southeast HSLC (left) events, (center) warning nulls, and (right) radar-based nulls. All else is as in Fig. 9.
When combined with instability and shear ingredients, the most skillful synoptic forcing ingredient was typically either a product of $\omega$ and $d\theta_e/dz$ or the near-surface wind magnitude, showing that potential instability or low-level advection may be critical in some cases. This idea is supported by ongoing modeling work (e.g., King and Parker 2015) and will be discussed further in section 5. Midlevel lapse rates, despite not showing skill on their own, do exhibit skill when combined with other ingredients, as found by Sherburn and Parker (2014); this suggests potential nonlinear interaction with other ingredients or compensation in some portion of the parameter space.

After testing all combinations of up to five ingredients, an index including the 0–3-km lapse rate [low-level lapse rate (LLLR)], 0–1.5-km bulk shear vector magnitude (S15MG), the maximum $d\theta_e/dz \times \omega$ product calculated from the 0–2-km layer through the 0–6-km layer at 0.5-km intervals (MAXTEVV), and (optionally, as discussed shortly) the effective bulk shear magnitude (ESHR) was found to outperform all other combinations. The most skillful formulation including the effective bulk shear magnitude [hereafter referred to as modified SHERB (effective version), or MOSHE] is

$$\text{MOSHE} = \frac{(\text{LLLR} - 4 \text{K km}^{-1})^2}{4 \text{K}^2 \text{km}^{-2}} \times \frac{(\text{S15MG} - 8 \text{m s}^{-1})}{10 \text{m s}^{-1}} \times \frac{(\text{ESHR} - 8 \text{m s}^{-1})}{10 \text{m s}^{-1}} \times \frac{(\text{MAXTEVV} + 10 \text{K Pa km}^{-1} \text{s}^{-1})}{9 \text{K Pa km}^{-1} \text{s}^{-1}},$$

where LLLR is in degrees kelvin per kilometer (a positive value indicates decreasing temperature with height), S15MG and ESHR are in meters per second, and MAXTEVV (K Pa km$^{-1}$ s$^{-1}$) is signed such that $\theta_e$ decreasing with height multiplied by upward motion yields a positive value. If any ingredient is negative (e.g., LLLR $< 4$ K km$^{-1}$), MOSHE is set to 0. MOSHE was set to missing when the environment lacked an effective layer (e.g., there was little or zero diagnosed MUCAPE). The given denominators were chosen such that the most skillful discriminating value of MOSHE within the Southeast HSLC events subset would approximately equal 1 (using the NARR dataset). Further, the subtracted constants and squared lapse rate term result in the most skillful formulation of MOSHE, while maintaining approximately equivalent interquartile ranges and contributions from each term (Fig. 15). Physically, in addition to the MUCAPE consideration above, the subtracted constants lead to a MOSHE value of 0 when the low levels are especially stable (i.e., when LLLR $\leq 4$ K km$^{-1}$) or when the low-level shear is especially weak (i.e., when S15MG $\leq 8$ m s$^{-1}$), neither of which is conducive for the production of severe weather in HSLC environments. The version of this index excluding the effective bulk shear vector magnitude [hereafter referred to as simply the modified SHERB (MOSH)] is given by

$$\text{MOSH} = \frac{(\text{LLLR} - 4 \text{K km}^{-1})^2}{4 \text{K}^2 \text{km}^{-2}} \times \frac{(\text{S15MG} - 8 \text{m s}^{-1})}{10 \text{m s}^{-1}} \times \frac{(\text{MAXTEVV} + 10 \text{K Pa km}^{-1} \text{s}^{-1})}{9 \text{K Pa km}^{-1} \text{s}^{-1}},$$

FIG. 17. As in Fig. 15, but for the SHERBS3, SHERBE, MOSH, and MOSHE value distributions.
As with MOSHE, if any ingredient is negative, MOSH is set to 0. Likewise, denominators were determined in a similar fashion. MOSH and MOSHE have comparable maximum HSS and TSS. MOSHE may be desired to further reduce false alarm area where convection is not anticipated (i.e., locations where CAPE is near zero), and MOSH may especially be a desired aid when the analyzed or forecast CAPE is suspected to be too low, which could lead to an underestimate of the severe threat by MOSHE (as shown by POD and FAR in Table 2).

Spatially, the MOSH composite maps reveal a much clearer delineation of the threat area in events compared with the rather broad SHERB signals (cf. Figs. 11 and 16), largely attributable to the inclusion of the MAXTEVV term. In addition, MOSH parameters show clear separation between events and nulls because of their larger fundamental range of values as compared with the SHERB parameters (Fig. 17). Overall, the MOSH parameters show an improvement over the SHERB parameters, both in statistical skill (Table 2) and spatial accuracy, while maintaining the core emphasis on low-level lapse rates and shear vector magnitude.

MOSH and MOSHE maximum skill scores are comparable to those of SHERBS3 and SHERBE, with MOSH showing a lower FAR than SHERBS3 (Table 2). Monte Carlo simulations (Wilks 1995) utilizing 1000 iterations of 50 random events and warning nulls within the Southeast subset showed that MOSH consistently outperformed all other combination forecast indices at their traditional guidance values (Table 3). Generally similar maximum skill scores were calculated using the independent Northeast subset, though the value where the maximum skill occurred varied (Table 4). As a final verification of the robustness of the MOSH parameters, independent 2012–14 event and null datasets were generated for the Southeast region, subject to the same criteria as the original 2006–11 datasets. Parameter distributions (Fig. 18) and calculated skill scores (Table 5) reveal similar characteristics of the MOSH parameters within the 2012–14 dataset, though in a similar fashion to the Northeast subset, the parameter value associated with maximum skill varied. This suggests that the guidance value of 1 should not be used as a hard threshold; rather, it can be expected that the conditional threat for HSLC severe hazards will increase as MOSH progressively exceeds 1.

Consistent with the SHERB parameters (Sherburn and Parker 2014), formulations of MOSH including the midlevel lapse rate showed a modest increase in skill over the formulations listed above (maximum HSSs of 0.410 and 0.377, respectively, for MOSH and MOSHE with the midlevel lapse rate term included). However, results of a principal component analysis (not shown) revealed that the midlevel lapse rate contributed far less variance than the other MOSH ingredients, implying that its role is minor compared to other terms. Additionally, the practical importance of the midlevel lapse rate remains in question, given the concerns posed in section 3. Finally, the MOSH and MOSHE formulations excluding the midlevel lapse rate still exhibit the key improvements over SHERBS3 and SHERBE: a spatial

<table>
<thead>
<tr>
<th>Index</th>
<th>Max HSS</th>
<th>Max TSS</th>
<th>POD at max HSS</th>
<th>FAR at max HSS</th>
<th>Max HSS threshold</th>
<th>HSS at threshold = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHERBS3</td>
<td>0.426</td>
<td>0.424</td>
<td>0.809</td>
<td>0.385</td>
<td>0.95</td>
<td>0.418</td>
</tr>
<tr>
<td>SHERBE</td>
<td>0.377</td>
<td>0.379</td>
<td>0.588</td>
<td>0.209</td>
<td>0.80</td>
<td>0.207</td>
</tr>
<tr>
<td>MOSH</td>
<td>0.396</td>
<td>0.398</td>
<td>0.638</td>
<td>0.241</td>
<td>1.45</td>
<td>0.366</td>
</tr>
<tr>
<td>MOSHE</td>
<td>0.390</td>
<td>0.394</td>
<td>0.528</td>
<td>0.134</td>
<td>1.30</td>
<td>0.331</td>
</tr>
</tbody>
</table>

TABLE 4. As in Table 2, but for the Northeast subset.
focusing of enhanced values near the location of reports in addition to an increase in statistical skill. Thus, for simplicity, the midlevel lapse rate was excluded from the presented version of the MOSH and MOSHE indices herein.

We have heretofore addressed the skill of parameters in the reference frame of gridded analyses, but it is worth considering their value in numerical weather prediction forecasts. Additional tests will be necessary to determine the sensitivity of MOSH ingredients—particularly the MAXTEVV field—to model setup, including grid spacing. For example, $\omega$ values will locally be considerably larger within high-resolution models because of their ability to resolve convective updrafts. Thus, the most skillful values of MOSH and MOSHE may be sensitive to horizontal grid spacing. Future work could examine the potential of “capping” the MAXTEVV contribution (e.g., as done with ESHR, among other ingredients, in the STP) on higher-resolution grids or working with spatially smoothed vertical motion fields (to capture synoptic or frontal-scale updrafts rather than convective updrafts) to alleviate this concern. Philosophically, the use of the MOSH and MOSHE fields is most practical in numerical weather prediction models with coarser grid spacing; after all, these indices are designed to diagnose a favorable environment for severe HSLC convection. In contrast, convection-allowing models should theoretically be capable of resolving the severe convection itself. This is corroborated by ongoing model simulations of HSLC severe and nonsevere convection (King and Parker 2015), while previous studies have noted improved forecasts of convection in strongly forced regimes (e.g., Duda and Gallus 2013), which are present in most HSLC cases.

As a preliminary step toward examining the sensitivity of MOSH fields to model grid spacing, MOSH and MOSHE values were calculated using archived forecasts from the Global Forecast System (GFS; 0.5° horizontal grid spacing), and the North American Mesoscale (NAM; 12-km horizontal grid spacing), and Rapid Refresh (RAP; 13-km horizontal grid spacing) models, for an HSLC event on 30 January 2013. For this case, 6-h forecasts valid at 0600 UTC 30 January 2013 show enhanced MOSH values focused over a smaller area than SHERBS3 (cf. Figs. 19 and 20), consistent with the NARR event composites. The corridors of enhanced MOSH values do vary somewhat more than those of the SHERBS3 values based upon the chosen analysis; however, it is difficult to determine if these distinctions are artifacts of the model setup or simply due to differing degrees of forecast and analysis accuracy. By breaking MOSH down into its respective components (Fig. 21), it

<table>
<thead>
<tr>
<th>Index</th>
<th>Max HSS</th>
<th>Max TSS</th>
<th>POD at max HSS</th>
<th>FAR at max HSS</th>
<th>Max HSS value</th>
<th>HSS at index = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHERBS3</td>
<td>0.398</td>
<td>0.401</td>
<td>0.709</td>
<td>0.310</td>
<td>1.15</td>
<td>0.266</td>
</tr>
<tr>
<td>SHERBE</td>
<td>0.440</td>
<td>0.445</td>
<td>0.619</td>
<td>0.173</td>
<td>1.00</td>
<td>0.400</td>
</tr>
<tr>
<td>MOSH</td>
<td>0.451</td>
<td>0.450</td>
<td>0.760</td>
<td>0.310</td>
<td>1.50</td>
<td>0.407</td>
</tr>
<tr>
<td>MOSHE</td>
<td>0.461</td>
<td>0.464</td>
<td>0.691</td>
<td>0.227</td>
<td>1.85</td>
<td>0.425</td>
</tr>
</tbody>
</table>

Fig. 18. As in Fig. 17, but for the independent 2012–14 Southeast dataset.
appears that the contribution from the low-level shear vector magnitude term is rather consistent among all forecasts and analyses; the low-level lapse rate term and, especially, the potential instability term appear to be responsible for the varying modeled threat areas. For example, RAP appears to underestimate the threat across the northern portions of the report area (Fig. 19b); this is coincident with a relatively low potential instability term contribution in RAP (Fig. 21f). Overall, however, the forecasts of MOSH depict the region where severe weather occurred in the surrounding 6 h with reasonable accuracy. A more robust evaluation of the sensitivity of MOSH and MOSHE to grid spacing and model configuration will be left for future work.

5. Discussion and conclusions
High-shear, low-CAPE (HSLC) severe convection remains a forecasting challenge across time scales ranging from days to minutes in advance. These environments compose a majority of the EF1 or greater tornado and significant wind report population during the cool season, especially during the overnight hours.
This population corresponds to time periods when climatological conditions are generally unfavorable for severe convection and, accordingly, the performance of severe weather watches and warnings is notably decreased. HSLC severe convection occurs across all regions of the contiguous United States; however, the majority of significant tornadoes and winds occur within the Southeast (Fig. 3). Across the West, HSLC environments are characterized by a drier lower troposphere and a surface triple-point or upslope setup, which suggests they may be fundamentally dissimilar from HSLC events in the East that are typically associated with low lifted condensation levels and occur in the warm sector or along a cold front (Figs. 6 and 8).

However, middle- and upper-tropospheric synoptic-scale features across all regions are fairly similar, with strong forcing for ascent exhibited at all levels (Fig. 6).

Much of the analysis herein focused on the Southeast subset, given that it comprises a large fraction of the severe HSLC convection climatology. Through the analysis of composite maps and skill score tests, several features were shown to discriminate between events and nulls, including (but not limited to) intense upper-level divergence and low-level convergence (Fig. 9); a localized maximum of low-level CAPE (Fig. 10); conditions for the release of potential instability, which could supplement low values of ambient CAPE (Fig. 12); and enhanced near-surface flow (Fig. 9). Additionally, plan-view...
FIG. 21. Contributions from the (left) low-level lapse rate term, (center) 0–1.5-km shear vector magnitude term, and (right) potential instability term [all as defined in Eq. (3)] for (a)–(c) NARR analysis valid at 0600 UTC 30 Jan 2013, (d)–(f) RAP 6-h forecast, (g)–(i) GFS 6-h forecast, and (j)–(l) NAM 6-h forecast, all valid at 0600 UTC 30 Jan 2013.
and maps showed that events tended to occur on the northern nose of a surface-based unstable sector, where SBCAPE values exceeded the 500 J kg\(^{-1}\) threshold (Fig. 9) and STP values were marginally enhanced to the near south (Fig. 14). Overall, it appears that both the magnitude and relative positioning of the aforementioned features is important in determining whether an HSLC setup will produce severe or nonsevere convection, with stronger, more closely collocated features conducive to severe HSLC events. Future work could examine how these features differ in location and intensity with convective mode \(\text{e.g., as in work conducted by the Sterling, Virginia, NWS Weather Forecast Office (WFO); M. Kramar (2016, personal communication), particularly given the propensity of HSLC tornadoes occurring within quasi-linear convective systems (e.g., Thompson et al. 2012).}

The SHERB parameters (Sherburn and Parker 2014) continue to exhibit skill at discriminating between Southeast HSLC events and nulls (Table 2; Figs. 11 and 17; note that the 3–5-km lapse rate is used and recommended to replace the 700–500-hPa lapse rate). However, this study provides evidence that skill could be further improved when interrogating additional fields or incorporating other ingredients. Shear vector magnitudes over shallower layers than those utilized in SHERBS\(^3\) exhibit higher skill than the original formulation. Additionally, skill score tests suggest that including a term meant to represent the release of potential instability further improves skill while also focusing the spatial footprint of parameter maximum values. These ingredients were combined into a modified SHERB, or MOSH, as introduced in section 4 (Table 2; Figs. 16 and 17). Another version of this index including the effective bulk shear magnitude (MOSHE, also introduced in section 4) provides similar skill while alleviating concerns of enhanced values where there is little or no chance for convection. While the spatial (Fig. 16) and statistical (Tables 2–5) skill of these indices are robust, future work must address the sensitivity of these new combination parameters to model setup, particularly grid resolution, to determine if the distributions of values presented herein are consistent across all platforms. Using an ensemble approach (e.g., assessing the ensemble mean MOSH or MOSHE) may ultimately be preferable, given the potential sensitivity of the ingredients on model physics choices and resolution.

Our ongoing work (e.g., King and Parker 2015) has identified similar features in simulations of HSLC events and nonevents using the Advanced Research version of the Weather Research and Forecasting Model (WRF-ARW; Skamarock et al. 2008). It seems that simulations on a 3-km horizontal grid can reasonably discriminate between HSLC events and nulls. In simulated HSLC events, destabilization may occur rapidly in the few hours prior to convection (and potentially on scales not resolved by 3-h NARR data). Simulated nonevents do exhibit some destabilization, but not on the order of that seen in the events. In all simulated events, temperatures near the surface increase modestly leading up to convection, resulting in steeper lapse rates and increased CAPE in the low levels. Some simulated events also exhibit characteristics of destabilization by the release of potential instability, during which low-to-midlevel cooling occurs as a result of large-scale layer lifting. Neither of these features tends to be as prominent in the simulated nonevents. These results corroborate the present NARR-based findings and will be presented in a later article.

We are continuing to examine the roles of upper-level troughs and jets, as well as various drivers of warm, moist-air advection in low levels. Both may account for the rapid destabilization in some cases (e.g., Lackmann 2002; Trier et al. 2006; Tuttle and Davis 2006), potentially on temporal and spatial scales unresolvable by conventional observations. Though the results herein address some of the forecasting challenges associated with HSLC environments, the sources and predictability of rapid environmental evolution ahead of HSLC convection remain open to investigation. Detailed studies into these items are appropriate next steps toward continuing to improve our understanding of severe weather in HSLC environments.

Acknowledgments. The authors would like to acknowledge collaborators on the Collaborative Science, Technology, and Applied Research (CSTAR) program grant that made this research possible, both in NOAA and at North Carolina State University (NCSU), for their constructive suggestions. In particular, we thank Steve Weiss and Rich Thompson for providing useful feedback on earlier versions of this manuscript. We are also grateful to Alex Anderson-Frey; Matthew Kramar from NWSFO Pittsburgh, Pennsylvania; Brice Coffer from NCSU’s Convective Storms Group; Lamont Bain from NWSFO Fort Worth, Texas; and an anonymous reviewer, whose constructive comments and suggestions greatly improved the manuscript. Further, we thank Andy Dean at SPC for providing much of the data utilized in this work. We also thank Brandon Vincent from WFO Raleigh for his suggestion to calculate the maximum potential instability product over the 0–6-km layer. We would finally like to acknowledge NOAA Award NA14NWS4680013, the aforementioned CSTAR grant, which provided funding for this research.
REFERENCES


