Investigation of Near-Storm Environments for Tornado Events and Warnings

ALEXANDRA K. ANDERSON-FREY AND YVETTE P. RICHARDSON
The Pennsylvania State University, University Park, Pennsylvania

ANDREW R. DEAN, RICHARD L. THOMPSON, AND BRYAN T. SMITH
Storm Prediction Center, Norman, Oklahoma

(Manuscript received 8 March 2016, in final form 23 August 2016)

ABSTRACT

In this study, a 13-yr climatology of tornado event and warning environments, including metrics of tornado intensity and storm morphology, is investigated with particular focus on the environments of tornadoes associated with quasi-linear convective systems and right-moving supercells. The regions of the environmental parameter space having poor warning performance in various geographical locations, as well as during different times of the day and year, are highlighted. Kernel density estimations of the tornado report and warning environments are produced for two parameter spaces: mixed-layer convective available potential energy (MLCAPE) versus 0–6-km vector shear magnitude (SHR6), and mixed-layer lifting condensation level (MLLCL) versus 0–1-km storm-relative helicity (SRH1). The warning performance is best in environments characteristic of severe convection (i.e., environments featuring large values of MLCAPE and SHR6). For tornadoes occurring during the early evening transition period, MLCAPE is maximized, MLLCL heights decrease, SHR6 and SRH1 increase, tornadoes rated as 2 or greater on the enhanced Fujita scale (EF2+) are most common, the probability of detection is relatively high, and false alarm ratios are relatively low. Overall, the parameter-space distributions of warnings and events are similar; at least in a broad sense, there is no systematic problem with forecasting that explains the high overall false alarm ratio, which instead seems to stem from the inability to know which storms in a given environment will be tornadic.

1. Introduction

Our understanding of tornadoes, the morphology of their parent storms, and the characteristics of the associated near-storm environment has increased dramatically over the past few decades, as a result of both research into the physical processes governing tornadoes (e.g., Markowski and Richardson 2014; Davies-Jones 2015) and climatological studies of the environments in which tornadoes occur (e.g., Thompson et al. 2003, 2012). Through conferences, publications, and National Weather Service training, the transfer of research to operations has been quite successful in the severe storms community (e.g., Kain et al. 2003). Combined with improvements in technology for the detection of tornadoes (e.g., the installation and operation of Doppler radars), the result has been an increase in tornado warning skill (Brooks 2004).

In the traditional, “textbook” case of a tornado generated by a right-moving supercell thunderstorm, we have a reasonable understanding of many of the environmental parameters associated with an environment conducive to tornadogenesis. We know that the combination of 0–6-km vector shear magnitude (SHR6, the magnitude of the vector difference in winds between the surface and 6-km height) and mixed-layer convective available potential energy (MLCAPE) provides a good first estimate of supercell potential (Brooks et al. 2003), with at least marginal CAPE necessary to fuel the updraft and SHR6 essential for developing the midlevel rotation that is the hallmark of a supercellular storm. In addition to midlevel rotation, tornadogenesis also requires the development of strong vertical vorticity at the surface through a process thought to occur as parcels descend through baroclinic zones in the storm’s cold pool (e.g., Davies-Jones 2015; Dahl et al. 2014; Parker and Dahl 2015). Although these baroclinic zones are
believed to be essential, strong negative buoyancy near the surface in the form of a very intense cold pool may prohibit the final contraction of the near-surface rotation to tornadic strength (Markowski et al. 2003; Markowski and Richardson 2014). This prohibitively intense cold pool can be roughly correlated with small values of environmental humidity at low levels, which in turn are associated with a high mixed-layer lifting condensation level (MLLCL). On the other hand, strong 0–1-km environmental storm-relative helicity (SRH1) appears to aid this final contraction through the development of enhanced vertical perturbation pressure gradients to aid in lifting (and therefore contracting) the air (Markowski and Richardson 2014).

Thus, just as the combination of SHR6 and MLCAPE enables forecasters to effectively discriminate between ordinary and supercellular convection for isolated storms, studies such as Thompson et al. (2012) have found that SRH1 and MLLCL discriminate well between significantly tornadic [rated 2 or greater on the enhanced Fujita scale (EF2+)] and nontornadic supercells, with more marginal (EF0–EF1) tornadoes occupying a nebulous middle ground. These two parameter spaces (MLCAPE–SHR6 and MLLCL–SRH1) have traditionally provided a useful, although not comprehensive, summary of the storm environments that contribute to an enhanced likelihood of tornadoes. These environmental conditions also are contained in the significant tornado parameter (STP; Thompson et al. 2003), which shows skill in discriminating between nontornadic and significantly tornadic environments for isolated storms.

However, despite our advances in understanding, our warning system remains imperfect, with fairly frequent false alarms (i.e., tornado warnings with no reported tornadoes) and some misses (i.e., tornadoes without warning). Even in environments favorable for tornadoogenesis, most supercells will not produce tornadoes, and many tornadoes occur in marginal environmental conditions. Adding to the complexity of the forecasting task, factors such as the diurnal evolution of the near-storm environment, geographical location, time of day, or time of year may impact the types of storm environments associated with tornadoes.

From both a forecasting and a research perspective, it is important to consider storm morphology in any discussion of tornado climatology: in their investigation of 22,901 severe thunderstorm and tornado events between 2003 and 2011, Thompson et al. (2012) used a manually compiled dataset of storm morphology and found that tornadic quasi-linear convective systems (QLCSs) and right-moving supercells (RMSs) were often characterized by different environmental MLCAPE results, especially in the winter months, with QLCSs tending to occur in lower-MLCAPE environments. In contrast, they found that measures of bulk wind difference and storm-relative helicity were relatively similar for tornadic QLCS and RMS storm morphologies. In any ingredients-based approach to identifying a potentially tornadic environment, it is helpful to first acknowledge the environmental differences between the QLCS and RMS storm environments, as well as the seasonal trend of those differences. Following on the work of Thompson et al. (2012), we first examine QLCS and RMS tornadoes separately to quantify and clarify the differences between the environments in which the two storm morphologies produce tornadoes.

In recent years, the availability of reanalysis data to characterize the mesoscale environment every hour has sparked a series of tornado climatology studies with different areas of focus, with some also plotting tornado reports and warnings in the CAPE–SHR6 and MLLCL–SRH1 parameter spaces (e.g., Dean and Schneider 2008; Schneider and Dean 2008; Anderson-Frey et al. 2014). Because of restrictions on data availability, or simply a different area of focus, many past studies have examined either tornado warnings/watches or tornado reports, but rarely both. The storm morphology information used in a recent series of papers (Smith et al. 2012; Thompson et al. 2012; Edwards et al. 2012) has likewise not yet been featured frequently in the literature.

Brotzge et al. (2013) have examined detailed warning performance over two years’ worth of tornado report and tornado warning data (2003–04), sorting the results based on proximity sounding data as well as storm morphology. They found that higher values of CAPE, STP, and wind shear were associated with greater tornado probability of detection (POD) and lead times, and that tornado warnings for nonsupercells had a lower POD and lead time, with supercells in lines showing lower POD/lead time than isolated or clustered supercells. Schneider and Dean (2008) created a tornado climatology by plotting tornado watch performance metrics from 2003 to 2007 in bins of MLCAPE–SHR6 space and found that, given a tornado watch, the greatest conditional probability (given lightning) of tornadoes was associated with MLCAPE > 500 J kg⁻¹ and 0–6-km shear > 25 m s⁻¹ for low CAPE and >15 m s⁻¹ for high CAPE.

To the extent that it has included discussion of tornado environments within the context of climatology and forecasting skill, the literature thus far has focused mainly on the probability of detection and lead time of tornado events (e.g., Brotzge et al. 2013). Still missing are discussions of how false alarm ratios vary by environment and, further, how the environments of tornado
reports, tornado warnings, PODs, and FARs vary in terms of time of day, season, and geographical region. While the FAR cannot be calculated separately for QLCS and RMS storm environments with the current dataset [the Smith et al. (2012) morphology data are limited to tornado events, not tornado warnings], the POD can indeed be examined within the context of storm morphology, using a much larger dataset than the 2-yr subset used by Brotzge et al. (2013).

This study makes use of a unique dataset to identify environmental conditions in which tornadoes occur. We use a full decade of mesoanalysis proximity soundings associated with both tornado reports and tornado warnings and compare the distributions of events and warnings as a function of the environment. The greater sample size used in this study allows for relatively fine-scale two-dimensional binning or kernel density estimation of the data with confidence, in the same vein as Schneider and Dean (2008) or Thompson et al. (2013), rather than having to examine one-dimensional histograms or box-and-whisker plots, as in Brotzge et al. (2013). Studying tornado warnings as well as tornado reports permits a more complete evaluation of warning performance by addressing issues of tornado warning false alarms, as well as the probability of tornado detection. Comparing the distributions of events and warnings can help determine whether there are any obvious forecasting issues (e.g., whether false alarms are occurring because warnings are issued in a different part of the parameter space than the events). Finally, examining QLCS and RMS storms separately permits us to determine the environmental factors unique to each of the storm types.

In the next section of this paper, we describe the dataset and methodology of this study, as well as kernel density estimation, a smoothing procedure replacing the binning techniques of earlier work (e.g., Schneider and Dean 2008; Anderson-Frey et al. 2012, 2014). Section 3 reviews the near-storm environments of tornadoes associated with RMS and QLCS storms, highlighting the unique environments and forecasting considerations associated with these two storm morphologies. Section 3 also brings in the tornado warning dataset, with discussion of tornado warning skill as well as further investigation into particular parts of the parameter space. Finally, section 4 summarizes the results.

2. Methods and data

This study draws information from two major sources. For tornado reports, we use archived mesoanalysis data (Dean et al. 2006) from the Storm Prediction Center (SPC) in the form of proximity soundings obtained from the Rapid Update Cycle model (RUC; Benjamin et al. 2004) for January 2003–April 2012, or the Rapid Refresh model (RAP; Benjamin et al. 2016) for later dates. By filtering county tornado segment data, keeping only the highest EF-scale rating per hour within a 40 × 40 km² area (Smith et al. 2012), 14,814 tornado events were identified for 2003–15 and were subsequently matched with the RUC/RAP proximity soundings in the grid box closest to the reported tornado (Coniglio 2012). In addition, we make use of the Storm Prediction Center’s manual convective mode classification [see Thompson et al. (2012) and Smith et al. (2012) for details], available for the period 2003–12, to examine the differences in tornado environments associated with QLCSs and supercells. Figure 1 depicts the geographical distribution of these tornado reports, split into several subcategories that will be discussed throughout the paper. Ratings on the Fujita scale (F scale) are available for tornado reports ranging from 1 January 2003 through 31 January 2007, after which point EF-scale ratings are available; in subsequent discussion, this paper will group the two rating scales together under the name (E)F-scale.

For warnings, the NWS data contain validation information corresponding to each of the 44,961 tornado warnings issued during the 2003–15 time period. Since a single tornado warning may cover an area that encompasses multiple grid points, RUC/RAP model data are matched to each tornado warning based on the location of maximum STP (Thompson et al. 2003, 2012) contained within the warning area.¹

Several discussions in the literature (e.g., Brooks et al. 1994; Potvin et al. 2010) debate the appropriateness of various measures of “proximity” storm environments. The higher spatial and temporal resolutions available from the RUC analysis provide more accurate representations of storm environments than the rawinsonde dataset, which is sparse in space and time, but it is worth noting that severe thunderstorms sometimes occur in the immediate vicinity of baroclinic zones, so that even a minor error in report placement can result in large errors. However, Thompson et al. (2003, 2012) argue, based on the authors’ extensive forecasting experience, that the very large sample size of this dataset should minimize any bias. This larger dataset does come with a

¹ Note that this selection criterion may bias the tornado warnings toward higher values of any of the variables contained in STP compared to the events. Note also that although these warnings are not filtered in the same way as the tornado events, there is no reason to think that the tornado event filtering process would bias the comparison with the warnings more in any particular section of the parameter space.
Caveat: it is worth emphasizing that statistical significance is more easily attainable with a large dataset, and that there may be an easily missed distinction between a difference that is statistically significant and one that will have practical significance for forecasting and research applications (e.g., Rosen and DeMaria 2012; Daniel 1977).

The limitations of the tornado report dataset have been well known and discussed for several decades (e.g., Doswell and Burgess 1988). The EF-scale has the unenviable task of attempting to condense complicated damage-intensity relationships into a single number; the distinction between an EF0 and EF1 tornado can have more to do with the type of structures (or lack thereof) impacted than any intrinsic meteorological quality of the tornado. In addition, reports of all types of severe weather see fluctuations due to a combination of non-meteorological factors including population density and time of day (for further discussion, see Witt et al. (1998) and Trapp et al. (2005)). A higher ratio of EF1 to EF0 tornadoes at night versus during the day, for instance, may simply reflect the lower likelihood of a marginal tornado being observed and reported at night if it does not cause substantial damage. As Smith et al. (2012) note, it is important to keep these limitations in mind when interpreting this dataset.

Rather than simply plotting the environmental data of the tornado events and warnings using two-dimensional histograms (Schneider and Dean 2008; Anderson-Frey et al. 2012), this work uses a smoothing method known as kernel density estimation (KDE; Zucchini 2003; Peel and Wilson 2008). KDE assigns each data point a shape known as a “kernel,” which is smoother than the rectangles used in the two-dimensional histogram approach; in this case, a Gaussian kernel is used. Each kernel essentially acts as a probability density function centered on the location of its corresponding data point in the parameter space. Since the kernels are not simply points, as in a scatterplot, nor rectangles, as in a histogram, summing the kernels across the entire domain will result in a smooth density estimate.
create smooth transitions between densities (e.g., Anderson-Frey et al. 2014). We choose the KDE approach because, when it comes to interpreting data and locating the maximum density in the parameter space, a histogram, unlike a KDE plot, has a great deal of sensitivity to bin size and placement. Kernel bandwidth can be chosen based on an optimization algorithm, or can be selected in a more qualitative way based on providing the clearest (i.e., least noisy or oversmoothed) at-a-glance representation of the data (Shimazaki and Shinomoto 2010). In what follows, we use an optimization algorithm to gain a starting value of kernel bandwidth, and then make minor adjustments to ease interpretation and clarity of the figures.

As an example of the kernel density estimation approach, Fig. 2a depicts a scatterplot of all 14,814 tornado reports (i.e., QLCS storms, RMS storms, and storms not classified into either of those categories) in the MLCAPE–SHR6 parameter space. Each point on this plot corresponds to the MLCAPE and SHR6 values for the environment corresponding to one particular tornado. KDE functions similarly to a two-dimensional histogram, highlighting regions of high and low density. Figure 2b shows the same information taking advantage of kernel density estimation, essentially smoothing the scatterplot to more clearly display that the maximum density of tornado reports tends to occur at around MLCAPE = 1250 J kg⁻¹ and SHR6 = 25 m s⁻¹. Contours are drawn such that they are centered on the highest density of tornado events, and the contours contain 25%, 50%, 75%, and 90% of the data, moving outward from that center. Figures 2c,d depict an analogous smoothing and contouring of the tornado reports in the MLLCL–SRH1 parameter space. The tornado warning data (green contours) will be discussed in the following section.

To study seasonal influences on the tornado warning and report climatologies, we use the standard meteorological definitions of spring as March–May (MAM), summer as June–August (JJA), fall as...
September–November (SON), and winter as December–February (DJF) (e.g., Thompson et al. 2012). We also examine diurnal influences by grouping all tornado reports and warnings into the following three categories: day (from sunrise until 2 h before sunset), early evening transition (EET; from 2 h before sunset until 2 h after sunset; Acevedo and Fitzjarrald 2001), and night (from 2 h after sunset until sunrise).

3. Results: Tornado events and warnings by environment

To begin, we will describe the characteristics of the entire dataset, and we will then break up the tornado warning and report datasets in terms of time of day, season, and geographical region.

In examining warning skill, probability of detection is defined here as the fraction of tornado reports having positive lead time (i.e., having a warning issued prior to tornado touchdown), and false alarm ratio (Barnes et al. 2007) is defined as the fraction of tornado warnings during which no tornadoes were reported in the warning region. The POD over the entire warning period was 66.4% (i.e., approximately two out of every three tornadoes occurred after a tornado warning was already in effect), and the FAR was 76.3% (i.e., more than three out of every four tornado warnings never had corroborating tornado reports). Figure 3 breaks down the FAR statistics for each weather office. The weather offices in the western parts of the country have tiny circles, reflecting the small number of tornado warnings issued in this region; in contrast, the Gulf Coast, the Midwest, and the Great Plains all show high numbers of tornado warnings issued during the 2003–15 period of interest. With some localized exceptions, FARs are relatively high along the Gulf Coast and into the Ohio River valley and are lower across the central Great Plains. It is important to note that a high FAR could be the result of underreporting rather than overwarning; population biases can certainly come into play here.

The deleterious effects of a missed tornado are obvious; the effects of a false alarm can be more subtle and may occur over a longer time scale (Brotzge et al. 2011). This being said, both FAR and POD should be seen as imperfect metrics given the lack of accounting for “close calls” (i.e., storms that would be perceived as nearly tornadic by the public, or storms that move just out of a warning region before becoming tornadic), as well as the lack of accounting for warnings that occur after initial tornado touchdown but still provide positive lead time for locations farther along the tornado’s path (Barnes et al. 2007).
a. Environments for entire distribution

Kernel density estimates for the 14814 tornado reports and the 44961 warnings are plotted in the MLCAPE–SHR6 (Fig. 2b) and MLLCL–SRH1 (Fig. 2d) parameter spaces. From these plots, it is clear that tornadoes occur in a variety of storm environments, even in typically nonsupercellular low-shear environments. The most common environments in which reported tornadoes occur feature moderate MLCAPE (≈1250 J kg⁻¹), relatively strong SHR6 (≈25 ms⁻¹), relatively low MLLCL heights (≈800 m), and relatively strong SRH1 (≈200 m² s⁻²). In their examination of tornado watches, Schneider and Dean (2008) identified regions of CAPE > 500 J kg⁻¹ and SHR6 > 25 ms⁻¹ for low CAPE and >15 ms⁻¹ for high CAPE as having the highest conditional probability of tornadoes (given lightning and a tornado watch). Note that the highest-density values of SRH1 seen in Fig. 2 exceed Thompson et al.’s (2003) threshold for significant tornadoes of SRH1 = 100 m² s⁻².

Overall, tornado warning distributions match so closely with tornado report distributions (Fig. 2) that a bootstrap test following the approach of Wilks (2011) revealed no statistically significant difference at the 95% confidence level: on the whole, tornado warnings are being issued for the same part of the parameter space as tornado reports. The only areas with visible differences in the distributions are for low SHR6 (<20 m s⁻¹) combined with MLCAPE < 2000 J kg⁻¹ and SRH1 < 200 m² s⁻² combined with MLLCL > 1200 m. To glean additional insight, it is helpful to see plots of FAR and POD for each of the two parameter spaces under investigation. To accomplish this, we bin the data using grids, where each grid square has dimensions of 8 J kg⁻¹ and 2 m s⁻¹, respectively, for the MLCAPE and SHR6 parameter space, and dimensions of 75 m and 30 m² s⁻² for the MLLCL and SRH1 parameter space. Bin sizes were chosen to give the best subjective view of the data. Within each bin, the probability of detection and false alarm ratio are calculated and plotted. In Fig. 4 and the similar figures that follow, the color within each box is a bilinear function of the surrounding values.

To avoid noisiness at the edges of the data, we remove all grid squares with densities (from the KDE analysis) below the 60th percentile for the MLCAPE–SHR6 parameter space and below the 80th percentile for the MLLCL–SRH1 parameter space, and we then remove all bins containing fewer than 10 warnings or reports. Depicted are the POD (Figs. 4a,c) and FAR (Figs. 4b,d) results for the entire dataset. In the MLCAPE–SHR6 parameter space, a best-discriminator line separating above-average and below-average PODs (Fig. 4a) shows that POD has a tendency to increase as MLCAPE and SHR6 increase; Brotzge et al. (2013) also found a positive CAPE–POD relationship within their 2-yr dataset.

Similarly, the FAR (Fig. 4b) shows a very strong tendency to decrease with increasing MLCAPE and SHR6. Thus, as the storm environment becomes more textbook “severe” (Brooks et al. 2003), forecast performance increases as a result. In the MLLCL–SRH1 parameter space, the relationships between POD and the two parameters are less clear-cut: while POD increases with increasing SRH1 for values below 200 m² s⁻², the relationship is less strong between POD and MLLCL or between POD and SRH1 for higher values of storm-relative helicity. In agreement with the optimal range of MLLCL values associated with the higher conditional probability of tornado formation found by Hart and Cohen (2016), probability of detection is lower (for low SRH1 values) and false alarm ratio is higher (for all SRH1 values) when MLLCL < 750 m. Do the lower POD and higher FAR in the low-MLLCL environments merely coincide with events that also happen to have low MLCAPE? Apparently not; repeating the analysis, this relationship between MLLCL and FAR persists even into relatively high-MLCAPE environments (i.e., environments with MLCAPE > 1800 J kg⁻¹).

b. Environments by storm mode

Tornado climatologies often focus on tornadoes produced by supercells (e.g., Alexander 2010; Rasmussen and Blanchard 1998; Rasmussen 2003); this focus reflects the fact that >95% of EF3+ tornadoes are associated with supercells (Smith et al. 2012). Throughout this section, we will examine the 8232 tornado reports associated with right-moving supercells as well as the 1278 tornadoes associated with QLCSs. Figure 1a provides a geographical distribution of the QLCS (blue) and RMS (red) tornadoes. QLCS tornadoes are more apparent along the Gulf Coast, as well as the eastern Midwest (e.g., Illinois and Indiana), which aligns with the regions of higher FAR in Fig. 3. The main environmental difference between the two modes is summarized in Fig. 5. In Fig. 5a, depicting the MLCAPE–SHR6 parameter space, we see that, while there is not much difference in terms of SHR6 (indeed, a bootstrap test does not reveal a statistically significant difference at the 95% confidence level), the QLCS tornadoes tend to occur for significantly lower values of MLCAPE than the RMS tornadoes, which confirms the conclusion found by Thompson et al. (2012) with an earlier version of this dataset. This causes the QLCS tornado distribution to extend into a portion of the parameter space with worse warning performance (cf. Fig. 4). The MLLCL–SRH1
parameter space (Fig. 5b) shows a statistically significant difference in both parameters under consideration: QLCS tornadoes tend to occur for lower MLLCL heights and for greater SRH1 than RMS tornadoes.

In addition, Figs. 5c–f break down the QLCS and RMS tornado reports into hits (black contours; tornadoes that occurred within a tornado warning issued prior to touchdown) and misses (red contours; tornadoes that began with no warning). The distributions of hits and misses show significant overlap for both storm types, perhaps highlighting the difficulty in distinguishing between storms that will be tornadic and those that will not be tornadic in similar environments. For both storm types, a greater percentage of misses than hits occur in low-SHR6 or high-MLLCL environments, while few misses occur for extreme values of SRH1 (>500 m$^2$s$^{-2}$ for supercells and >750 m$^2$s$^{-2}$ for QLCS events). Few misses also occur for the combination of large MLCAPE and SHR6 in the supercell regime or for large MLCAPE in the QLCS regime. Overall, the distribution of misses extends farther (than that of hits) into regions of the parameter space that already have a high false alarm ratio (i.e., regions of low SHR6 or low SRH1 combined with very high or very low MLLCL heights), suggesting that those regions are not simply underwarned.

Unfortunately, EF-scale and morphology information are available only for tornado events, not tornado warnings, so we can determine POD but not FAR. Since the prototypical ingredients-based approaches for forecasting tornadoes are generally based on the ingredients associated with supercells, it is unsurprising that RMS tornadoes tend to have higher PODs than the national average: 78.7% of RMS tornadoes have a warning ahead of tornado occurrence. When it comes to QLCS tornadoes, the POD is only 49.0%; that is, fewer than half of all reported tornadoes from QLCS storms in our sample had positive lead times. This finding is in line with the work of Brotzge et al. (2013), who determined in their analysis of two years’ worth of tornado reports...
that nonsupercellular tornadoes had substantially lower PODs.

RMS storms produce the vast majority of significant (EF2+) tornadoes, but also many EF0 tornadoes. The probability of detection is lower (61.9%) for EF0 tornadoes; these tornadoes often occur in marginal or unexpected storm environments and are far less likely to produce severe damage and loss of life. Can this lower POD for EF0 tornadoes help explain the large discrepancy between QLCS and RMS probabilities of detection? As it turns out, it cannot: when EF0 tornadoes are removed from the dataset, RMS POD rises to 81.1%, but the QLCS POD drops even further to 47.0%.

Fig. 5. KDE plot of tornado events between 2003 and 2012 in the (a) MLCAPE–SHR6 and the (b) MLLCL–SRH1 parameter spaces. The distribution of tornadoes from QLCS storms is depicted in blue, while the distribution of tornadoes from RMSs is in red. Misses (red contours; tornadoes with no prior warning) and hits (black contours; tornadoes with a warning issued ahead of touchdown) are depicted for (c),(d) RMS tornado events and (e),(f) QLCS tornado events, in the (left) MLCAPE–SHR6 and (right) MLLCL–SRH1 parameter spaces. Inner contours are slightly darker than outer contours, for ease of interpretation at a glance.
(i.e., low instability, high shear) environments poses a unique challenge for forecasters, as POD is lower and FAR is higher in those parts of the parameter space. Any relationship between QLCS POD and the storm environment is not apparent, whereas RMS POD clearly increases with both SHR6 and SRH1 (not shown).

c. Diurnal trends

Figure 1b depicts the tornadoes’ geographical distribution by time of day, with daytime in red, the early evening transition in cyan, and night in blue, for the entire dataset. Nocturnal tornadoes are particularly common in a swath that runs through the Mississippi–Alabama–Tennessee region, which is another area of relatively high FAR in Fig. 3. We expect the tornado environments to vary by time of day, given the diurnal evolution of the lowest levels in the atmosphere, including heating at the surface due to insolation, which leads to a warmer and deeper well-mixed layer throughout the day. Once outgoing radiation exceeds incoming radiation, we expect the surface to cool and stabilize, leading to the cessation of convective boundary layer mixing. To examine diurnal effects, local sunrise–sunset has been calculated based on date, latitude, and longitude, in order to accurately place each event into its sunset-relative time frame.

Figure 6 depicts the diurnal variation of several key environmental parameters for all tornado events during the entire time period. The hourly averaged MLCAPE (Fig. 6a) peaks for tornadoes occurring shortly before local sunset, not long after the expected maximum in surface temperature; this is also about when tornado environments have the highest MLLCLs (Fig. 6c). Both of these are consistent with the cycle of boundary layer warming and growth. In contrast, both SHR6 (Fig. 6b) and SRH1 (Fig. 6d) have a maximum value for tornadoes occurring shortly after local sunset (consistent with the cessation of convective boundary layer mixing), and both have their minimum values around 6 h before local sunset. The early evening transition period is thus characterized by the following: while MLCAPE is still relatively large, LCLs are lowering, and SRH1 and SHR6 are increasing; the magenta line in Fig. 6e shows a rapid increase in the STP during the EET (the period between 2 h prior to and 2 h after local sunset), with a peak value shortly after the EET. It is no wonder that the EET has been referred to as “six-o’clock magic” (Coffer and Parker 2015), with storms often becoming tornadic during this period, and a peak in the number of observed EF2+ tornadoes (Mead and Thompson 2011; Fig. 6f).

The fraction of all tornado events that are classified as QLCS (Fig. 6e, red line) ap-

account for 66.3% of all daytime tornadoes, 78.8% of all EET tornadoes, and 67.8% of all nocturnal tornadoes. The fraction of all tornado events classified as QLCS (Fig. 6e, blue line) is lowest near sunset and peaks late at night, when low-level shear and humidity are high and convective instability is relatively low. QLCS tornadoes are only 8.7% of daytime tornadoes and 8.2% of EET tornadoes, but they compose a disproportionately high 24.0% of nocturnal tornadoes, perhaps consistent with upscale growth of afternoon isolated cells into meso-scale convective systems.

Considering diurnal variations in tornado strength, while there is some noise in the data, the percentage of tornadoes that are rated EF2+ (Fig. 6f, black line) appears lowest during the day (~10%), whereas at night, this percentage rises to closer to 20%, even though the actual number of significant tornadoes (Fig. 6f, green line) is low. Thus, although fewer tornadoes are reported at night, a higher percentage are significantly tornadic, perhaps aided by the low-LCL and high-SRH1 conditions. It is possible the higher fraction also is related to reduced reporting of weak tornadoes, which may go unobserved at night.

In terms of warning performance, when averaged over all grid boxes, EET (Fig. 7) has the best warning performance with the highest POD (73.3%) and the lowest FAR (75.0%). Daytime (Fig. 8) has the worst POD (61.7%), while night (Fig. 9) has the worst FAR (79.9%). One reason for daytime’s low POD may be the fact that it includes proportionately more tornadoes in marginal shear and CAPE conditions (SHR6 < 10 m s⁻¹, SRH1 < 100 m² s⁻², MLCAPE < 800 J kg⁻¹) than either of the other two time periods. Indeed, daytime hours are responsible for a whopping 70.5% of low-SRH1 (<100 m² s⁻²) tornado events; in turn, 78.2% of these daytime low-SRH1 tornadoes have a rating of (E)F0. It is also possible that reporting efficiency is higher for daytime tornadoes (Ashley et al. 2008), especially in the case of weak tornadoes.

The black lines in Figs. 7–9 are the best-discriminator lines dividing the areas that tend to have POD and FAR above and below the average value for the entire dataset. From these, it is clear that the EET period has above-average POD over a larger portion of the MLCAPE–SHR6 parameter space compared to the other time periods. Nighttime, on the other hand, has above-average FAR over a larger portion of the MLCAPE–SHR6 parameter space.

All three time periods reflect a similar overall trend in the MLCAPE–SHR6 parameter space (higher POD and lower FAR for higher values of MLCAPE and SHR6), although there are some subtleties. The daytime and EET POD and FAR (Figs. 7a,c and Figs. 8a,c)
depend strongly on both MLCAPE and SHR6. For nocturnal tornadoes (Fig. 9), when 0–6-km shear is below \( \sim 25 \text{ m s}^{-1} \), FAR is high for nearly all values of MLCAPE below \( \sim 2500 \text{ J kg}^{-1} \). Nocturnal FAR also is affected by having more tornadoes in marginal-CAPE parts of the parameter space.

Considering the SRH1–MLLCL parameter space, the Thompson et al. (2003) threshold of 100 m\(^2\) s\(^{-2}\) for SRH1, which is meant to indicate a higher potential for significant tornadic activity, appears to provide a fairly good degree of separation between the parts of the parameter space with relatively low POD and the parts of
the parameter space with relatively high POD at night. The POD with SRH1 < 100 m$^2$s$^{-2}$ is 51.7%, whereas the POD with SRH1 ≥ 100 m$^2$s$^{-2}$ is 72.4%.

For daytime and EET events, POD increases with increasing values of SRH1. For intermediate values of MLLCL (750–1250 m), FAR decreases as SRH1 increases; outside of this MLLCL range, the pattern is less obvious. The high overall FAR in nocturnal tornadoes (79.9%; Fig. 9d) is occurring for marginal (i.e., low SRH1) conditions, but also for environments with relatively high SRH1 and low MLLCL, which is usually a favorable environment for significant tornadoes.

**d. Seasonal effects**

Springtime tornadoes are centered around the Kansas–Missouri–Oklahoma–Arkansas corner, summertime tornadoes become more common moving farther north, and fall and winter tornadoes are most concentrated around the Gulf Coast (Fig. 1c). We would expect the characteristic storm environments of tornadoes to vary from season to season as well; the environment in which tornadoes occur during the spring is considerably different than that which we might expect in winter.

**Figure 10** shows how the MLCAPE–SRH6 parameter space varies from season to season. Spring (Fig. 10a) tornadoes occur most often for moderate values of MLCAPE and fairly high values of SHR6, characteristic of the environments associated with Great Plains spring tornadoes. Fall events (Fig. 10c) tend to be associated with slightly lower MLCAPE and SHR6 than spring events. In the winter (Fig. 10d), low-MLCAPE, high-SHR6 environments tend to dominate. Comparing the distributions of EF0 tornadoes to only those rated EF1+ (i.e., comparing the blue and red lines in Fig. 10), we see that the portion of the distribution extending into low SHR6 contains the EF0 tornadoes, particularly in the spring, but also in the other seasons.

The seasonal breakdown of the MLLCL–SRH1 parameter space is depicted in Fig. 11; removing the EF0 tornadoes results in a slight shift toward higher SRH1 and a decrease in high-LCL events in the spring, summer, and fall. The distributions in both parameter spaces are similar for EF0 winter tornadoes versus EF1+ winter tornadoes, in large part because only 40.0% of winter tornadoes are rated EF0, as compared to 68.0% of summer tornadoes; Thompson et al. (2012) found a similar result when looking at EF2+ winter tornadoes versus all winter tornadoes. Summer tornadoes have markedly smaller SRH1 and higher MLLCLs than the other seasons, while SRH1 is large more often in winter tornadoes.

---

**FIG. 7.** As in Fig. 4, but for tornado reports and warnings occurring during the EET.
In terms of storm morphology, in the winter, RMS storms are responsible for only 64.8% of tornadoes (compared with 76.1% in the spring), whereas QLCS storms make up a disproportionately high 27.9% of winter tornadoes (compared with only 12.0% in the spring). Thus, the winter climatology looks more similar to the QLCS climatology (Figs. 5a,b). These results agree with those of Trapp et al. (2005), who found that QLCS tornadoes across the continental United States during their period of study (1998–2000) were disproportionately common in the cold season.

Differences in POD and FAR are also apparent during different times of year. POD is lower on average during the summer (58.4% overall) than for the other three seasons and is generally highest in the spring (72.1% overall). There is little difference in POD between cold-season QLCS tornadoes (fall and winter; 47.1% POD) and warm-season QLCS tornadoes (spring and summer; 50.0% POD); unfortunately, there are too few data in these season morphology categories to find relationships between POD and parameter space beyond these bulk statistics.

**e. Geographical effects**

The shift in seasonality of tornado environments is, of course, caused by a variety of factors tied to local geography (surface heating, moisture sources) and large-scale shifts in weather patterns (e.g., the location of the midlatitude jet stream). In the following discussion, we make use of the West, Midwest, Great Plains, Northeast, and South regions defined in Fig. 1d. While the dividing lines between regions are subjective and do not strictly follow the NWS regional divisions, they are based on a combination of previous literature (e.g., Brooks et al. 2003) and areas with similar storm types or environmental characteristics. The Southern region features a disproportionately high percentage of QLCS, nocturnal, and winter tornadoes. The Midwest region, on the other hand, tends to have fewer nocturnal tornadoes and more tornadoes near sunset. Spring tornadoes dominate the Great Plains region, whereas the Midwest tends to see their peak in tornadoes during the summer months.

The distributions of tornado events in the parameter spaces for these five regions (blue contours in Figs. 12 and 13) are for the most part unsurprising. The Great Plains and the Midwest, home to a disproportionately high percentage of RMS tornadoes, strongly resemble the distribution of the RMS tornadoes depicted in Figs. 5a,b. The South features more QLCS tornadoes.
than elsewhere in the country, and so tends toward lower MLCAPE and MLLCL values, with higher values of SRH1, similar to the distribution of QLCS tornadoes in Figs. 5a,b. The Northeastern region shows a smaller range of SHR6 values with a maximum occurrence at lower MLCAPE than the Great Plains. The Western region has relatively few tornado reports, and even fewer of EF1+ strength (80.7% of all tornadoes in the Western region are rated EF0).

Geographically, POD is lowest for the Western region (26.2%) and is somewhat lower for the Northeast region as well (52.8%), compared with the Great Plains (69.2%), the Midwest (64.6%), and the South (70.6%). For the Great Plains, the South, and the Midwest regions, FAR values are notably lower for high-MLCAPE, high-SHR6 parts of the parameter space (not shown), especially for the Great Plains (overall FAR of 73.3%) and the Midwest region (overall FAR of 75.4%), which are the two regions with the lowest overall FAR values.

Comparing these tornado events (blue contours in Figs. 12 and 13) with tornado warnings (green contours in Figs. 12 and 13), the tornado warnings are generally centered on the same parts of the parameter space as the tornado reports, but the details of the distributions are different. The biggest differences occur in the Great Plains (where tornado warnings tend to be issued for slightly higher-SHR6, lower-MLLCL, and higher-SRH1 environments than the environments in which tornadoes were actually reported) and the Western region (where tornado warnings tend to be issued for much higher-SHR6, higher-MLCAPE, lower-MLLCL, and higher-SRH1 environments than the environments in which tornadoes were actually reported). As noted before, the low-SHR6, high-MLLCL events tend to be part of the EF0 distribution, which is difficult to distinguish from nontornadic environments (Thompson et al. 2003, 2012).

4. Discussion and summary
To issue timely and accurate tornado warnings, forecasters must make use of data such as numerical modeling output and observations, and must also draw on their own experience and understanding of the situation. Tornado warnings in their present state provide advance warning for fewer than three out of every four tornado reports, and more than three-quarters of all tornado warnings are false alarms. Using a unique dataset
consisting of over 10 years’ worth of tornado warnings and reports, this work has examined warning performance as a function of the MLCAPE–SHR6 and MLLCL–SRH1 parameter spaces, and has also examined in more detail how tornado environments differ based on storm morphology, tornado intensity, time of day, time of year, and geographical region.

Examining the QLCS and RMS storm modes separately, we find that the distribution of QLCS tornadoes is shifted slightly from that of RMS tornadoes toward lower MLCAPE and MLLCL but higher SRH1; they also feature considerably lower probabilities of detection than the more “textbook” RMS results, and the discrepancy cannot solely be explained by the higher percentage of EF0 tornadoes in the QLCS storm mode. These QLCS tornadoes are disproportionately common at night, in the southern United States, and during the winter months. The probability of detection of tornadoes occurring for right-moving supercells increases strongly with SHR6 and SRH1, but the POD of QLCS tornadoes does not increase as consistently with these variables.

Looking at the diurnal variation of tornado environments, early evening transition is a time of maximum MLCAPE, decreasing MLLCL heights, increasing SHR6 and SRH1, and the highest number of EF2+ tornadoes. It is also the time with the highest POD and lowest FAR. At night, SRH1 is a maximum and the largest relative percentage of EF2+ tornadoes is reported. During the day, performance is poor (high FAR, low POD) for SRH1 < 100 m$^2$s$^{-2}$. Our examinations of tornado warning skill by environment note that, for most categories examined in this work, warning skill tends to increase (i.e., POD tends to increase and FAR tends to decrease) with increasing MLCAPE and SHR6. Thus, the region of the parameter space identified by Brooks et al. (2003) as being particularly characteristic of severe convection is also the region for which warning skill is the highest. The relative impacts of MLCAPE and SHR6 on POD and FAR tend to vary somewhat under different conditions. For example, during the night, MLCAPE has a stronger impact on POD than SHR6, whereas SHR6 has a stronger impact on FAR.

![Figure 10](image-url)

Fig. 10. KDE plot of tornado events in the MLCAPE–SHR6 parameter space for (a) spring, (b) summer, (c) fall, and (d) winter from 2003 through 2015. The blue (red) contours depict the environments corresponding to only those tornadoes rated EF0 (EF1+). The EF0 tornadoes make up 51.2%, 69.7%, 56.4%, and 40.9% of all events in each season in (a)–(d), respectively. Inner contours are slightly darker than outer contours, for ease of interpretation at a glance.
We would expect the tornado warning skill in the MLLCL–SRH1 parameter space to be similarly consistent with the literature [i.e., because high-SRH1 environments are characteristic of significantly tornadic environments (Thompson et al. 2012), tornado warning skill should also be higher in these environments], but it is important to note that the relationship between warning skill and MLLCL heights is not a simple one (e.g., Hart and Cohen 2016): in every subcategory investigated by this paper (time of day, time of year, geographical region), the FAR does not change consistently with MLLCL heights. The high FAR of storms in the high-SRH1, low-MLLCL part of the parameter space—which is especially prominent for nocturnal and/or Southern region tornadoes but is present even in higher-CAPE scenarios—warrants further study. This shift in warning skill also applies to POD: SRH1 is substantially more strongly related to POD than are MLLCL heights for events occurring during the EET, spring, and summer, as well as events located in the Great Plains, South, and Midwest regions.

The parts of the parameter space in which they do not match tend to be those dominated by EF0 tornadoes. (The exception to this matching is the Western region, which experiences few tornadoes.) In other words, there is no “smoking gun” with respect to forecasting strategies for identifying potentially tornadic environments when examining the parameter space of the national dataset as a whole; future attempts at using near-storm environmental parameters to improve the detection of tornadoes must at minimum give consideration to the seasonal, diurnal, and geographical impacts on tornado warning skill. Just as one-size-fits-all warning approaches are often inappropriate at the national level, caution must be exercised in applying these broader findings at a local level, given the potential for wide variations in performance from office to office (e.g., Fig. 3).

While POD varies more dramatically across various subsets of the data, the fact that FAR remains consistently high across nearly all of the data subsets indicates that either our knowledge of the environmental controls on tornado formation is incomplete, or that there are factors beyond the environment that determine the differences between tornadic and nontornadic storms.

Fig. 11. As in Fig. 10, but for tornado events in the MLLCL–SRH1 parameter space.
within a given environment. Networks of observational data at higher resolutions than the current space and time scales may breathe new life into the use of atmospheric parameters for false alarm reduction, perhaps with a focus on storm-scale features rather than bulk atmospheric properties. Advances in our understanding of how the public and emergency managers interpret and act upon the risk communicated by tornado warnings—and, perhaps, a shift away from the use of such imperfect metrics as POD and FAR within the forecasting community—also have the potential to provide the next big leap in tornado false-alarm reduction.

This work provides a launching point for future studies aimed at particular problems for which bulk atmospheric parameters can still provide value: interest

FIG. 12. KDE plot of tornado events (blue contours) and warnings (green contours) in the MLCAPE–SHR6 parameter space for the (a) Great Plains, (b) South, (c) Midwest, (d) Northeast, and (e) West regions, as defined by Fig. 1d, from 2003 through 2015. The total POD and FAR for each region is depicted in the top-right corner of each plot. Inner contours are slightly darker than outer contours, for ease of interpretation at a glance.
lies in improving the probability of detection for the QLCS storm mode. One step toward this goal involves assessing the suitability of forecasting indices such as the STP for tornadoes associated with the QLCS storm mode and, if necessary, formulating an improved parameter based on the results of this climatology. We also intend to probe the question of the environments leading to storms that produce multiple and/or long-lived tornadoes, examine warning skill as a function of warning order on a particular storm, and create a climatology of tornado outbreak events.

Acknowledgments. The authors are grateful for the assistance from Brenton MacAloney in obtaining the verification data. Three anonymous reviewers provided suggestions that greatly improved the final manuscript. This work has benefited tremendously from helpful discussions with Martin Tingley, Israel Jirak, Russ Schneider, Richard Grumm, the forecasters at the NWS State College Weather Forecast Office, Steven Weiss, Bill Bunting, Roger Edwards, and Paul Markowski, as well as the mesoscale research group at The Pennsylvania State University. AKAF is supported through

Fig. 13. As in Fig. 12, but for the MLLCL–SRH1 parameter space.
REFERENCES


