# The Utility of ProbSevere v2.0 for Predicting Pulse Severe Thunderstorms

THOMAS L. GARD,[a] HENRY E. FUELBERG,[a] AND JOHN L. CINTINEO[b]

[a] *Department of Earth, Ocean, and Atmospheric Science, Florida State University, Tallahassee, Florida*
[b] *Cooperative Institute of Meteorological Satellite Studies, University of Wisconsin–Madison, Madison, Wisconsin*

(Manuscript received 12 January 2022, in final form 27 June 2022)

ABSTRACT: Pulse severe storms are single-cell thunderstorms that produce severe wind and/or severe hail for a brief period of time. These storms pose a major warm season forecasting problem since forecasters presently do not have sufficient guidance to know which, if any, of the cells that are observed will become severe. The empirical Probability of Severe (ProbSevere) model, developed by the Cooperative Institute for Meteorological Satellite Studies (CIMSS), fuses real-time data to produce short-term (0–60 min), statistically derived probabilistic forecasts of thunderstorm intensity. This study evaluates the ability of ProbSevere to predict pulse severe storms in the southeast United States. ProbSevere objects fitting the usual definition of a pulse severe environment were matched with severe events from *Storm Data* to create a dataset of ProbSevere objects that corresponded to pulse severe thunderstorms. A null dataset consisted of objects in pulse severe environments that did not match with a severe event. Results reveal that ProbSevere's probabilities are small to moderate at the times corresponding to pulse severe events. While probabilities of nonsevere storms are generally smaller, there are a large number of outliers. Lightning flash rate is the only predictor relevant to this study that correlates strongly with increasingly favorable pulse storm probabilities. We conclude that ProbSevere provides forecasters only limited guidance as to whether a pulse severe event will soon occur. Developing a version of ProbSevere specifically for pulse severe storms would likely lead to better predictability for this mode of convection.

KEYWORDS: Forecast verification/skill; Mesoscale forecasting; Operational forecasting; Severe storms

## 1. Introduction

### a. Pulse severe storms

Deep moist convection is often categorized into three modes—single cells, multicells, and supercells, with vertical wind shear being the greatest environmental factor determining which mode will occur (Markowski and Richardson 2010b). Single-cell storms tend to occur in environments whose 0–6-km vertical shear is less than ~10 m $s^{-1}$; multicells when 0–6-km vertical shear is between ~10 and ~20 m $s^{-1}$; and supercells when 0–6-km vertical shear exceeds ~20 m $s^{-1}$. A single-cell storm contains a single updraft and does not initiate subsequent convection in any organized way. Since synoptic scale forcing is weak, the storms often occur near the time of maximum diurnal heating and dissipate shortly after sunset. Although single-cell environments lack sufficient shear to initiate organized new cell development (as does multicellular convection), they often evolve into disorganized clusters.

A small percentage of single-cell storms produce strong wind gusts and/or hail that meet the National Weather Service (NWS) requirements for the issuance of a special weather statement or a severe thunderstorm warning. Although the environmental 0–6-km wind shear of pulse severe storms is weak (<10 m $s^{-1}$), the convective available potential energy (CAPE) is large, being ≥2000 J $kg^{-1}$ (Markowski and Richardson 2010b). If severe weather does occur, it is of the "pulse" variety, i.e., short lived and only marginally severe, but still able to produce damage. Severe weather can be hail ≥ 2.54 cm (1 in.) in diameter, but more often presents as wind gusts ≥ 25.72 m $s^{-1}$ (50 kt), sometimes in the form of microbursts. The type of microburst produced is regionally dependent. In the southeast United States, pulse severe storms often produce wet microbursts, which are characterized by heavy rainfall and high reflectivities. Dry microbursts, which are distinguished by their smaller to no rainfall amounts reaching the surface and low reflectivities, occur in the High Plains (Atkins and Wakimoto 1991; Markowski and Richardson 2010a). Regardless of type, microbursts have been associated with numerous aviation accidents, and pulse thunderstorms in general are responsible for more deaths by lightning strikes than any other storm type (Ashley and Gilson 2009).

The meanings of the terms "pulse" and "pulse severe" have evolved over time. The original definition of "pulse" was a multicellular storm that mostly was nonsevere, but with one cell occasionally becoming severe (Wilk et al. 1979). Cerniglia and Snyder (2002) expanded the term "pulse" to include nonsevere thunderstorms. The term "pulse severe" also was beginning to be used (Miller and Mote 2017b). The NWS currently defines a "pulse storm" as "a thunderstorm within which a brief period (pulse) a strong updraft occurs, during and immediately after which the storm produces a short episode of severe weather," while a "pulse severe thunderstorm" is defined as a "single cell thunderstorm which produces brief periods of severe weather" (NWS Internet Services Team 2009). These two NWS definitions do not provide a clear distinction since both contain the words "brief" or "short" and "severe." Miller and Mote (2017b) recently proposed new definitions: "single cells," often misidentified as multicells, should be renamed "Byers–Braham cells," or "weakly forced thunderstorms," and those that become severe should be denoted "severe weakly forced thunderstorms." Clearly there is

*Corresponding author*: Thomas L. Gard, tlg14c@my.fsu.edu

DOI: 10.1175/WAF-D-22-0003.1

disagreement regarding how these storms are defined. However, for the purposes of this study, "pulse severe storms" will refer to weakly forced thunderstorms that become severe for a brief period and are characterized by an environmental 0–6-km wind shear $< 10$ m s$^{-1}$ and CAPE $\geq 2000$ J kg$^{-1}$. By "severe" we mean that the pulse severe storm was associated with a report contained in *Storm Data*. Additional details are provided in the following section.

Miller and Mote (2017a) developed a warm season (May–September) 15-yr climatology of pulse severe storms from 30 WSR-88D radar sites in the Southeast. While pulse severe storms were found to be most common near the Appalachians, they occurred throughout their study area. Although pulse severe storms are less common in the Northeast (Miller and Mote 2017a), radar-derived warning criteria have been developed for them (Cerniglia and Snyder 2002). In actuality, brief but damaging pulse storms are a concern to the great majority of NWS Forecast Offices (WFOs) around the country.

Local NWS WFOs are responsible for issuing severe storm warnings for their areas of responsibility—their County Warning Areas (CWAs). On a day with weak shear but large CAPE, numerous cells can develop within a CWA. Even though organized severe thunderstorms are not expected, weakly forced cells with a short life cycle could still reach severe criteria. However, forecasters presently do not have sufficient guidance to distinguish between a cell that soon will become severe and the many cells—perhaps all—that will not (Guillot et al. 2008). The timing of that sudden transition also is problematic.

Most severe storm research has focused on multicells and supercells because they can produce relatively large areas of injuries and damage. In contrast, little research has examined the structure and evolution of pulse severe storms and how they differ from nonsevere single cell storms. This paucity of information is a major limitation to the appropriate issuance of warnings. By the time a forecaster realizes that a cell has the likelihood of becoming severe and issues a warning, the severe phase often is about to end or already has ended. Thus, the lead time for issuing warnings is considerably less than for other types of convection. Compared to other storm modes (e.g., multicells, supercells), false alarm ratios (FARs) are larger, and probabilities of detection (PODs) are smaller for warnings issued on pulse thunderstorms (Guillot et al. 2008).

### b. *ProbSevere*

Since weather phenomena cannot be forecast with certainty, it is more realistic to make probabilistic forecasts that express the likelihood that an event will occur. The NWS is exploring a new approach to its advisory/watch/warning products whereby severe weather watches and warnings and other hazards are disseminated in a gridded, probabilistic manner (Rothfusz et al. 2014, 2018). The empirical Probability of Severe (ProbSevere) models were developed by NOAA and the Cooperative Institute for Meteorological Satellite Studies (CIMSS) to further that objective.

ProbSevere version 1 (PSv1) fused real-time information from several data sources to produce short-term, statistically derived probabilistic forecasts of thunderstorm intensity (Cintineo et al. 2014, hereafter C14; Cintineo et al. 2018, hereafter C18). More specifically, its goal was to probabilistically predict whether a storm will produce severe weather within the next 60 min. Its data sources included output from Multi-Radar Multi-Sensor products (MRMS; Smith et al. 2016), the Earth Networks Total Lightning Network (ENTLN; Thompson et al. 2014), Rapid Refresh Model (RAP; Benjamin et al. 2016), and Geostationary Observational Environmental Satellites (GOES; Schmit et al. 2017). The naïve Bayesian classifier model was used to compute the conditional probability of a storm becoming severe based on parameters from the aforementioned sources. PSv1 did not give separate probabilities for severe wind, severe hail, or tornadoes, but instead provided a single probability that combined these hazards together. Verification studies showed that PSv1 compared to NWS forecasters produced a smaller POD and critical success index (CSI) but a greater FAR. However, its median lead time was more than twice that of NWS forecasters (C18).

Based on feedback from experiments at the Hazardous Weather Testbed (HWT) and the NOAA Operations Proving Ground between 2013 and 2016, a major upgrade to PSv1 was released during spring 2017 (Cintineo et al. 2020, hereafter C20). This upgrade, denoted PSv2, utilizes the same data sources and major statistical concepts as PSv1, but produces separate probabilities for severe hail (diameter $\geq 2.54$ cm), severe straight-line convective winds (gusts $\geq 25.72$ m s$^{-1}$), and tornadoes. These separate probability classifiers are denoted ProbHail, ProbWind, and ProbTor, respectively. A fourth probability (called probSevere) is simply the maximum value of the three aforementioned classifiers. C20 provides a detailed description of how PSv2 was developed. The ProbHail and ProbWind models are of most interest to the present research since pulse severe thunderstorms generate severe straight-line winds and/or severe hail, but rarely landspouts which *Storm Data* classifies as tornado events (Murphy 2021). While quantifiable data on how PSv2 guidance impacts lead time statistics for NWS forecasters currently is unavailable, between 60% and 70% of forecasters at the HWT indicated that PSv2 does increase lead time for severe hazards. It is important to note that all three PSv2 models were trained and evaluated on storm reports of the corresponding type (e.g., hail reports for ProbHail, etc.) and consisted of all modes of severe thunderstorms. How well the models specifically predict pulse severe thunderstorms is not known.

### c. *Objectives*

This research quantifies how well PSv2 (specifically its ProbHail and ProbWind components) predicts pulse severe storm occurrence in the southeast United States. In addition, we describe the temporal characteristics of pulse severe objects, evaluate how well the NWS warns on pulse severe storms, and determine how much information about pulse severe storms is provided by each of their input component parameters.

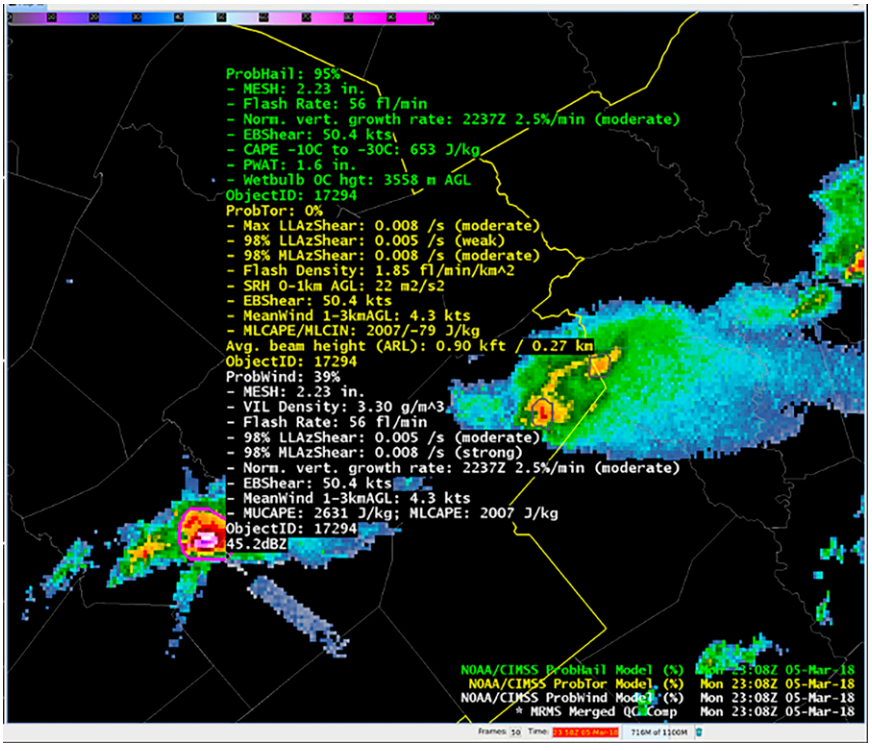


FIG. 1. Example of PSv2 object information as it appeared on AWIPS-II on 5 Mar 2018. Probabilities are ProbHail = 95%, ProbTor = 0%, and ProbWind = 39% (https://cimss.ssec.wisc.edu/severe_conv/training/training.html).

## 2. Data and methodology

### a. *Development of PSv2*

The development of PSv1 and PSv2 has been described in detail by C14, C18, and C20. Therefore, PSv2 is only briefly summarized below. Both versions of ProbSevere utilize a multisensory storm identification and tracking system that combines satellite and radar data to produce "objects" that indicate the areal outlines of storms. Radar objects created from 3D composite reflectivity fields are generated and spatially grown from local reflectivity maxima using the Warning Decision Support System-Integrated Information "w2segmotionII" algorithm (WDSS-II) (Lakshmanan et al. 2007). Corresponding satellite objects are created and tracked using the identification and tracking system developed by Sieglaff et al. (2013). The satellite-derived objects are collections of pixels having a local cloud-related maximum in 11-$\mu$m emissivity ($\varepsilon_{\text{tot}}$). The $\varepsilon_{\text{tot}}$ is the emissivity that a cloud would have if its effective cloud top were at the tropopause (C14). For clouds with a large optical depth (such as pulse severe thunderstorms), it is a measure of how close a cloud's radiative center is to the altitude of the tropopause. Unlike brightness temperature, which traditionally is used to track convection, $\varepsilon_{\text{tot}}$ is not influenced by season or latitude (Sieglaff et al. 2013).

The radar and satellite-derived objects are merged to create a single object that ideally corresponds to a single storm (C14). Included in the data describing the object and its environment are the object's ID, the latitude/longitude of its centroid, the time in coordinated universal time (UTC), the spatial area of the object, predictor inputs, and probability outputs (Fig. 1). An important characteristic is that this information is updated every 2 min.

To select the predictors used in PSv2, 46 potential parameters from the four previously mentioned sources were tested statistically (MRMS, ENTLN, RAP, and *GOES-16*). The capability of a naïve Bayesian classifier deteriorates if too many correlated predictors are used; therefore, each PSv2 model (Hail, Wind, and Tor) was limited to 1) only one reflectivity-based MRMS parameter, 2) only one ENTLN lightning-based parameter, and 3) only one instability-based RAP parameter. Correlated parameters were paired into 2D predictors if appropriate. ProbWind is subdivided into two classifiers—a model for cellular storms on the mesogamma scale (2–20 km), and a model for linear storms on the mesobeta scale (20–200 km) and the lower end of the mesoalpha scale (200–500 km). Because pulse severe thunderstorms are cellular by nature, only the cellular ProbWind model (hereafter denoted ProbWind) is relevant to the present study. Table 1 lists the predictors that were statistically selected for the relevant ProbWind and Prob Hail models (C20).

ProbSevere objects sometimes do not perfectly correspond to individual thunderstorm cells. One object can merge with an adjacent object or split into two separate objects. Figure 2 shows an example of an object splitting into two. Although

TABLE 1. PSv2 model predictors. All acronyms have their standard meteorological meanings (after Cintineo et al. 2020).

| Model | Predictors |
|---|---|
| ProbHail | A priori: Climatology of severe hail producing thunderstorms from database<br>Max of maximum expected size of hail (MESH)/the height of the wet bulb 0°C<br>Flash rate (Earth Networks Total Lightning)/effective bulk shear (EBSHEAR)<br>CAPE in layer from −10° to −30°C/ precipitable water<br>Normalized satellite growth rate |
| Cellular ProbWind | A priori: MUCAPE/EBSHEAR<br>Max MESH<br>Flash rate/EBSHEAR<br>Normalized satellite growth rate |

the tracking system developed by Sieglaff et al. (2013) attempts to mitigate these issues, it does not completely eliminate them (C14). The result is that if two objects merge into one, one of the object's IDs could be retained, or an entirely new ID be created for both, or one of the splitters could retain its original ID. This introduces uncertainty in some of our assessments, especially object lifetime and area.

### b. *Methodology to evaluate PSv2 for pulse severe storms*

Pulse severe wind events sometimes are attributed to microbursts. Since the type of microburst that can be produced by pulse severe storms differs between regions, the area of study was the southeast United States (Alabama, Florida, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee) where wet microbursts are most common. Our study period was June, July, and August 2019. We used *Storm Data*, a National Centers for Environmental Information (NCEI) publication, to create a dataset of severe weather events. While numerous studies have demonstrated the shortcomings of *Storm Data* (Witt et al. 1998; Goodman et al. 2005; Trapp et al. 2006; Schultz et al. 2009; Smith et al. 2012), it remains the best source of severe storm events. As previously stated, pulse severe storms produce severe hail (diameter $\geq 2.54$ cm) and/or severe wind (gusts $\geq 25.72$ m s$^{-1}$). Therefore, every *Storm Data* event during the study period and in the domain that met these criteria was recorded. While *Storm Data* classifies storms that produce qualifying damage as severe even if the reported hail/wind is below severe thresholds, those events were not included here since pulse severe storms are strictly defined by severe hail and wind criteria. Our resulting dataset at this step consisted of 1931 severe wind events and 181 severe hail events.

The next step was to determine which of the events could be attributed to pulse severe storms. That determination was based on Markowski and Richardson (2010b) who described pulse severe environments as having CAPE $\geq 2000$ J kg$^{-1}$ and 0–6 km AGL wind shear $< 10$ m s$^{-1}$. However, neither of these parameters is a predictor in ProbWind or ProbHail (Table 1), and neither is included in the PSv2 data files. Instead, PSv2 utilizes most unstable convective available potential energy (MUCAPE) which we used as the substitute for the CAPE criterion.

PSv2's closest equivalent to 0–6 km AGL wind shear is EBSHEAR. EBSHEAR is "the maximum bulk shear of the "most unstable" lifted parcel level upward to 40%–60% of the equilibrium level (EL) height" (Thompson et al. 2007). Although conceptually similar to 0–6 km bulk shear, EBSHEAR accounts for storm depth instead of the fixed 0–6-km layer (Thompson et al. 2007). However, its results are similar to 0–6 km AGL bulk shear for most surface-based storms with an EL of 11–13 km AGL (R. Thompson 2020, personal communication). We used EBSHEAR as a suitable alternative to 0–6-km shear for identifying pulse severe storm environments.

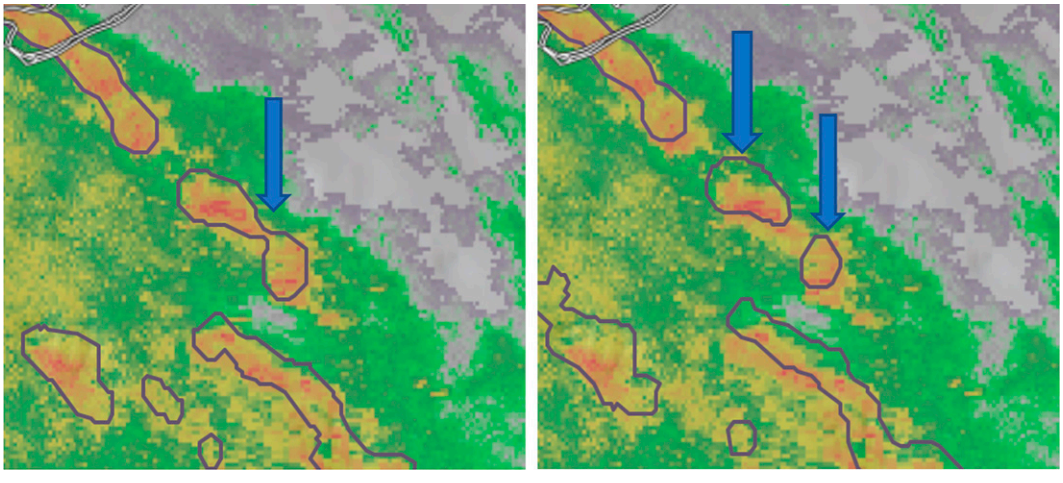

FIG. 2. Example of the splitting of one PSv2 object into two separate objects (blue arrows). The object of note is near the center of the left panel at 1354 UTC 15 Sep 2020. The object then splits into two objects 2 min later (right panel) at 1356 UTC.

There were originally 309 400 ProbSevere objects within the study area. These objects were spatially and temporally matched with the severe events in *Storm Data*. We tested numerous matching thresholds, including various spatial (0.1°, 0.15°, etc.) and temporal bounds (1 or 2 min) around the severe event. Results from the Thunderstorm Project (Byers and Braham 1949) indicated that a single-cell thunderstorm has a maximum diameter of 10 km, which corresponds to approximately 0.1° latitude and longitude in the southeast United States. The centroid of an object was used in calculating the distance to a severe event, since the exact boundaries of an object were not available. Therefore, we matched an object to the closest *Storm Data* report if the object's centroid was within ±0.1° latitude and longitude and within ±2 min of the severe event. If objects are assumed to be circular, the median radius of the PSv2 objects at the time of the severe event was 8.33 km. The median distance between the centroid of an object and its corresponding *Storm Data* event was 4.82 km for the 0.1° sort, with a maximum distance of 12.59 km. Although use of a 0.15° threshold was investigated to account for possible errors in *Storm Data*, it was decided that the larger search radius was excessive, since mismatches became more frequent. The time window of ±2 min was consistent with the methodology of C18 and C20. If a severe event matched with more than one object, only the object with the nearest centroid was kept. Similarly, if an object matched with more than one severe event, the first reported event was used since it provided the first indication that storms were becoming severe.

Objects occasionally exhibited very short durations, most likely due to merging or splitting. Such short lifetimes would be insufficient for forecasters to evaluate the storm and decide whether to issue a warning or special weather statement. To remove these occurrences, an object had to exist for at least 4 time steps (8 min). The final step in creating the pulse severe dataset was to apply the pulse severe environmental thresholds of MUCAPE $\geq$ 2000 J kg$^{-1}$ and EBSHEAR $<$ 10 m s$^{-1}$ (19.44 kt) at the creation time of an object (not the very beginning of the actual storm); 365 objects remained.

To verify that the selected ProbSevere objects were indeed pulse severe storms, Level 2 radar data (Amazon 2021) of randomly selected storms were scrutinized using the Gibson Ridge Level II Analyst software. Approximately one-third of our pulse severe storm selections also were confirmed with the aid of forecasters at the Tallahassee NWS forecast office. Our final pulse severe dataset consisted of 365 storm objects during June, July, and August of 2019: 336 of them produced severe winds, while only 29 produced severe hail. This ratio of severe windstorms to severe hailstorms is similar to the ratio of total severe wind events to total severe hail events in *Storm Data*.

A null dataset was created to include storms located in pulse severe environments (MUCAPE $\geq$ 2000 J kg$^{-1}$ and EBSHEAR $<$ 10 m s$^{-1}$) within the same study period and area but not matching with a severe event in *Storm Data*. Once the duration criteria of 4 time steps was applied, the final null storm dataset totaled 77 800 objects.

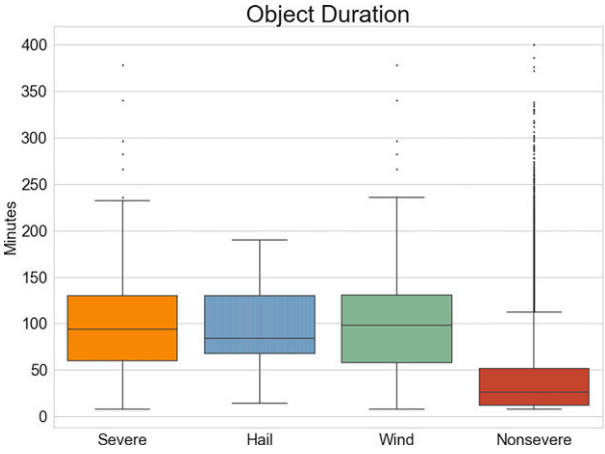


FIG. 3. Duration of storm objects in minutes. Shown from left to right are all pulse severe storms, hailstorms, windstorms, and nonsevere storms. Horizontal lines indicate the medians. Edges of the box indicate the first and third quartiles. The range is denoted by whiskers that extend to the farthest point within 1.5 × IQR from the edges of the box, where (IQR = Q3 − Q1). Outliers are plotted as separate points.

## 3. Results

Three aspects of the relation between PSv2 objects and pulse severe storms were examined: 1) the temporal characteristics of pulse severe objects compared to nonsevere objects, 2) how well ProbHail and ProbWind predict observed pulse severe events, and 3) how well individual predictors discriminate between pulse severe and nonsevere objects. In addition, how well the NWS warns on pulse severe storms was evaluated.

Most results are quantified using traditional box-and-whisker plots, e.g., Fig. 3. Median values on the plots are denoted as horizontal lines within the boxes. The top and bottom of each box represent the interquartile range (IQR), while the whiskers extend from the edges of the box to the data point whose value extends the farthest within 1.5 × IQR. The whiskers will be used to represent the range of the data. Finally, outliers are plotted as separate points beyond the whiskers and will be considered outside the defined range. The pulse severe objects will be analyzed as a single group unless stated that they have been subdivided into hailstorms and windstorms.

### *a. Temporal characteristics of storm objects*

The median lifetime of pulse severe objects (Fig. 3) is 94 min (47 time steps of 2-min duration). Median values for hailstorm and windstorm objects separately are 84 and 98 min, respectively. Nonsevere objects (Fig. 3) exhibit much shorter durations, with a median of 26 min. While the duration of most nonsevere objects is short, the large number of outliers (Fig. 3) suggests that some can persist for long periods. However, the large number could be a result of the settings used for identifying objects in the WDSS-II code. It also may denote a limitation in PSv2's ability to track individual thunderstorm cells, especially the weaker ones. Specifically, unusually

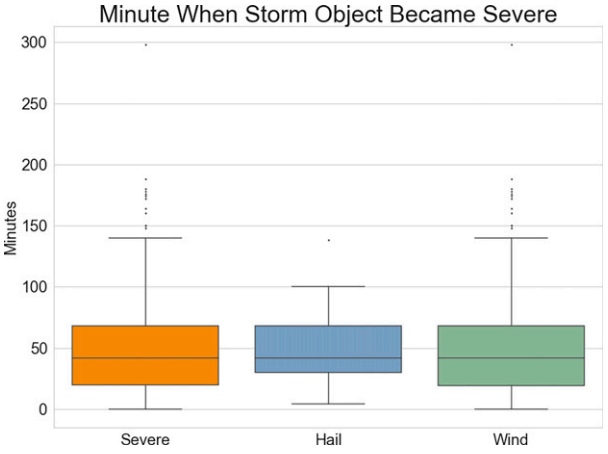


FIG. 4. The minute after an object is created when the storm object became severe. Shown from left to right are all pulse severe storms, severe hailstorms, and severe windstorms.

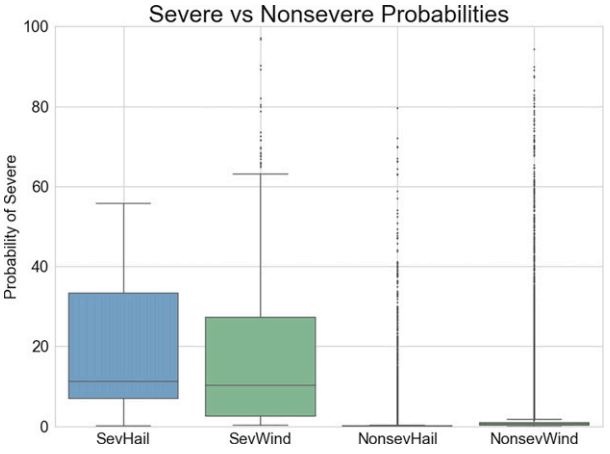


FIG. 5. Shown from left to right are ProbHail probabilities at the severe time step for severe hail producing storms, ProbWind probabilities at the severe time step for severe wind producing storms, ProbHail probabilities at the proxy severe time step for nonsevere storms, and ProbWind probabilities at the proxy severe time step for nonsevere storms.

long object lifespans may be due to objects merging and how the associated IDs are handled. This potential problem is exacerbated by the commonly observed clustering of single cell thunderstorms. The differences in duration between nonsevere and severe objects also may be partially due to the nonsevere dataset being two orders of magnitude larger than the severe dataset. These unusually long and short duration cases were retained in the dataset.

The median time when pulse severe objects produced a severe event (the severe time step) is 42 min after their creation (Fig. 4). However, a few objects (12 out of 365 severe cases) matched with a severe event at their first time step which would provide forecasters no time to evaluate them and issue warnings. These objects also were retained in the dataset. Severe wind and hail objects exhibit a similar timing as to when they produce severe weather.

### b. *Ability of PSv2 storm objects to forecast pulse severe events*

The main goal of this research is to assess how well the ProbHail and ProbWind models of PSv2 forecast pulse severe thunderstorms in the Southeast. We first evaluate both models at their object's first time step and then at their severe time step to determine the trends in guidance that occur. While not shown, the values of ProbHail and ProbWind at the first time step are very small, with medians of less than one percent. Although the smaller number of hailstorms must be considered, both models expectedly provide virtually no guidance at the objects' initial creation about future pulse severity.

The time of the severe event ideally represents the peak in storm intensity when the probability of severe weather is expected to be greatest, although severe weather can precede or follow peak intensity. Guidance at the severe time step (Fig. 5) reveals that the median ProbHail value is 11.27%, with an IQR of 7.04%–33.27%. The median ProbWind value is 10.23%, with an IQR from 2.55% to 27.33%. Although both models exhibit major increases in probability between the first (<1%) and severe time steps, values remain small. Hence, ProbSevere's operational utility to help forecast whether an object will produce a pulse severe event is limited. This finding is not surprising since PSv2 was not designed specifically for pulse severe storms, which are difficult to forecast (Guillot et al. 2008).

Nonsevere objects by definition do not match with severe events; therefore, they do not have a "severe time step." To provide further comparison between nonsevere objects and severe objects, a crude proxy for the severe time step was created for the nonsevere storm objects. The median time when storms became severe was found to be 51% through their objects' lifetime. We used this ratio to estimate the time step when the nonsevere objects were at the same stage of their life cycle. It will be denoted the "proxy severe time step." While this proxy is very simplistic, it does allow a gross comparison between nonsevere and pulse severe storms.

Probabilities of nonsevere objects at their proxy severe time step are shown in Fig. 5. Results show that ProbHail exhibits a median value of 0.11%, an IQR between 0.09% and 0.14%, with 11 209 of the 77 800 objects being outliers. Similarly, ProbWind exhibits a median value of 0.54%, an IQR between 0.35% and 0.94%, with 8171 of the 77 800 objects being outliers. Thus, probabilities for nonsevere objects at their proxy severe time step are much smaller than the probabilities for severe objects at their severe time step (Fig. 5)—a desirable characteristic. However, the large number of outliers within the nonsevere object dataset is indicative of PSv2's difficulty discriminating between severe and nonsevere environments.

To further explore the utility of PSv2 for pulse severe storm forecasting, the maximum probability preceding the severe time step for severe objects is examined (Fig. 6). The 12 objects whose severe time step occurred at the first time step are not included. These maximum probabilities for severe objects

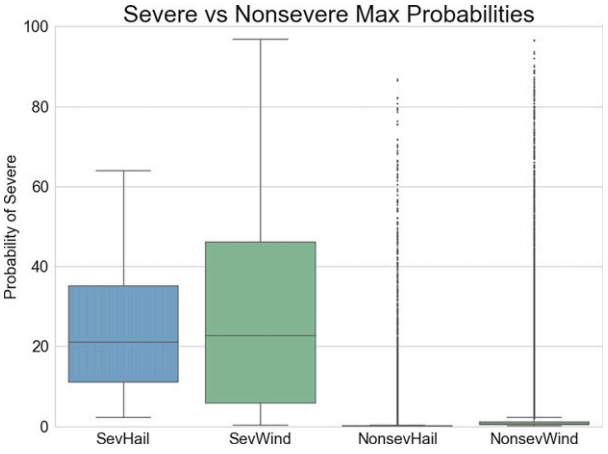


FIG. 6. Shown from left to right are maximum ProbHail probabilities before the severe time step for severe hail producing storms, maximum ProbWind probabilities before the severe time step for severe wind producing storms, maximum ProbHail probabilities for nonsevere storms, and maximum ProbWind probabilities for nonsevere storms.

are greater than those at the severe time step, with medians of 21.12% and 22.74% for hail and wind objects, respectively. Likewise, both quartiles for hail and wind objects are shifted positively, with IQRs between 11.02% and 35.18%, and between 5.91% and 46.10%, respectively. Similar differences are found for nonsevere objects, with maximum probabilities for ProbHail and ProbWind at any time step plotted in Fig. 6. Results show median values of 0.12% (ProbHail) and 0.62% (ProbWind), with 11 986 and 8900 outliers. Given the similarities between Figs. 5 and 6, these results again show that PSv2's operational utility is limited.

Figure 7 and Table 2 provide an alternative comparison of forecast probabilities for pulse severe objects (hail and wind combined) at their time of severity and nonsevere objects at their proxy severe time step. The probability guidance was

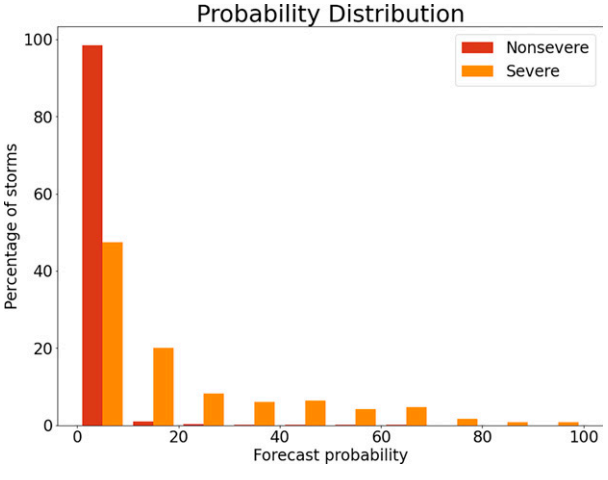


FIG. 7. The percentage of severe (orange) and nonsevere (red) storms in 10 equally spaced probability bins.

divided into 10 equally spaced bins ranging from 0% to 100%, with the percentage of storm objects in each bin being plotted. For nonsevere objects, the larger value of ProbHail and ProbWind is used. Results show that over 98% of the nonsevere objects exhibit probabilities between 0% and 10%, and less than 2% have greater probabilities. Severe storm objects are quite different, with 47.40% in the 0%–10% category, and 12.05% having probabilities > 50%. While a larger percentage of severe storms have probabilities > 10%, nonsevere storms still outnumber severe storms in all but the highest probability bin.

A reliability diagram serves as our final evaluation metric for PSv2 (Fig. 8), again using 10 equally spaced bins from 0% to 100%. The diagram expresses the ratio of severe storm objects to all storm objects as a function of the average probability of pulse severe weather per bin (Wilks 2011). Probabilities for severe objects were taken at the severe time step, and probabilities for nonsevere objects were taken at the proxy severe time step. The 1:1 line represents a perfectly calibrated model. Results reveal that PSv2 tends to under forecast pulse severity. This under forecasting is similar to that of figures in C14, C18, and C20 who examined all types of severe storms over the United States, not just pulse severe storms. The large degree of under forecasting at probabilities > 65% is partly due to the decreasing number of storm objects, with only four cases in the 90th–100th percentile (Table 2). C14 suggested that errors in radar data, storm identification and tracking, and deficiencies in storm reports may contribute to decreased model performance. More cases will be needed to fairly evaluate the skill of PSv2 for pulse severe storms.

To summarize, PSv2 does exhibit some skill in discriminating between pulse severe and nonsevere storms (Figs. 5, 6, 7, and 8). However, all of the metrics described above suggest that the skill is very limited.

### c. *NWS performance*

A component of C20's methodology was to determine which range of PSv2 probabilities produces the "best" metrics of POD, FAR, and CSI. Results showed that the probabilities producing the highest CSI varied geographically (Fig. 15 of C20), ranging from 40% to 70% in the Southeast. C20 found

TABLE 2. The number of objects in each probability bin. Bins, total number of storm objects, severe storm objects, and nonsevere storm objects, shown from left to right.

| Bins | All storms | Severe | Nonsevere |
|---|---|---|---|
| All bins | 78 165 | 365 | 77 800 |
| 0%–10% | 76 715 | 173 | 76 542 |
| 10%–20% | 770 | 73 | 697 |
| 20%–30% | 268 | 30 | 238 |
| 30%–40% | 141 | 22 | 119 |
| 40%–50% | 95 | 23 | 72 |
| 50%–60% | 74 | 15 | 59 |
| 60%–70% | 55 | 17 | 38 |
| 70%–80% | 29 | 6 | 23 |
| 80%–90% | 14 | 3 | 11 |
| 90%–100% | 4 | 3 | 1 |

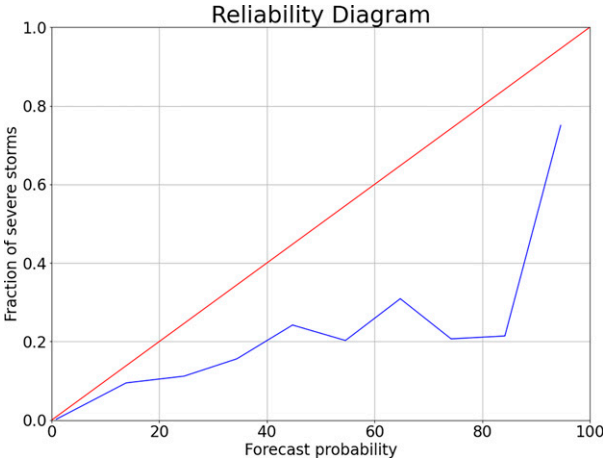


FIG. 8. Reliability diagram of PSv2 to forecast hail and wind pulse severe storms. That is, the probability that a storm will produce pulse severe weather as a function of the ratio of severe storms to all storms (blue line). The 1:1 line represents perfect model reliability (red line).

the Southeast to exhibit low CSIs between 0.25 and 0.30, due mostly to high FARs. Their small CSIs and high FARs were attributed to quickly developing storms, lightning rich storms, erroneously large values of MESH, and an inflated number of severe wind reports. These qualities are all associated with pulse severe storms.

To evaluate how well the NWS predicts pulse severe storms, skill scores were calculated. Using archived warning data from the Iowa State Mesonet (Iowa State University 2021), severe event data from *Storm Data*, and the pulse severe storm and nonsevere storm datasets, the number of warned events, unwarned events, verified warnings, and unverified warnings was calculated. There were 1987 warnings issued by WFOs within the study region during the summer of 2019. Warnings were matched with severe events if the event/object occurred within the warning area. Likewise, nonsevere objects were matched with warnings in a similar fashion to create a list of unverified warnings. Any warning that did not match with a ProbSevere object was removed from consideration. There were a total of 246 warned events, 119 unwarned events, 240 verified warnings, and 445 unverified warnings. Only warnings that corresponded to events with severe hail/wind criteria were used.

Results show that the POD for NWS, the number of warned events divided by the number of warned and unwarned events, is 67.40%. The FAR for NWS, the number of unverified warnings divided by the sum of unverified and verified warnings, is 64.96%. The CSI, a function of POD and FAR, is 29.96%. A FAR of this magnitude is consistent with previous results, given the large number of nonsevere storms compared to severe storms. These findings further support the notion that it is difficult to assess which of many nonsevere thunderstorms will become pulse severe (Guillot et al. 2008).

Lead times of the NWS warnings for the pulse severe storms were calculated in two ways: the NWS method where missed warnings equal zero lead time and calculating the lead time based only on verified warnings (Bieringer and Ray 1996). Using the NWS method, the average lead time for pulse severe storms is 9.58 min. The verified warning method yields an average lead time of 14.21 min. These lead times are shorter than those calculated by Guillot et al. (2008) for pulse severe storms. Such results reinforce the need to forecast pulse severe storms in a timelier manner.

### d. Predictor data

Pulse severe environments were defined as having MUCAPE $\geq 2000$ J kg$^{-1}$ and EBSHEAR $< 10$ m s$^{-1}$ at the time of an objects' creation. Therefore, it is informative to quantify whether values of these two parameters sufficiently differentiate between objects that will be associated with pulse severe events within the next 60 min and those that will not.

Scatter diagrams of MUCAPE and corresponding ProbWind probabilities for objects associated with severe wind at the severe time step, and for nonsevere objects at the proxy severe time step are shown in Fig. 9a. The threshold of $\geq$2000 J kg$^{-1}$ is closely maintained between the objects' initiation (not shown) and the severe time step (Fig. 9a). The median value of severe wind MUCAPE is 2910.5 J kg$^{-1}$, while the median nonsevere MUCAPE value is 2877.6 J kg$^{-1}$, only slightly less than the severe value. In addition, the two distributions appear similar, and neither the pulse severe objects nor the nonsevere objects exhibit a strong, positive correlation between values of MUCAPE and the probability of severe wind (0.15 and 0.11, respectively). Thus, MUCAPE alone poorly distinguishes between environments that will or will not produce pulse severe wind.

Figure 9b is similar to Fig. 9a, showing ProbHail values for hailstorms. Although there are relatively few cases of severe hail, the distributions of MUCAPE for the severe hail producing and nonsevere objects exhibit no obvious differences, and the threshold of $\geq$2000 J kg$^{-1}$ is largely conserved. The median MUCAPE value for these severe hailstorms is 2823.4 J kg$^{-1}$, and the nonsevere value is 2877.6 J kg$^{-1}$ (the same as in Fig. 9a). Neither the pulse severe objects nor the nonsevere objects exhibit a strong correlation between MUCAPE and the probability of severe hail ($-0.12$ and 0.05, respectively).

The pulse severe storm environment also was partially defined by EBSHEAR $< 10$ m s$^{-1}$. However, the scatter diagram of EBSHEAR for objects associated with severe wind events at the severe time step (Fig. 10a) differs little from that of objects not associated with severe wind. The median value of EBSHEAR is 7.31 m s$^{-1}$ for the severe wind objects, with 304 of the 336 values $< 10$ m s$^{-1}$. The median value for the nonsevere objects is 6.76 m s$^{-1}$, with 75 531 of the 77 800 values being $<10$ m s$^{-1}$. Once again, neither the severe wind objects nor the nonsevere objects exhibit a useful correlation between EBSHEAR and the probability of pulse severe wind (0.18 and 0.13, respectively).

Scatter diagrams of EBSHEAR and corresponding values of ProbHail for objects associated with severe hail at the severe time step and for nonsevere objects at the proxy severe time step are shown in Fig. 10b. Similar to the severe wind

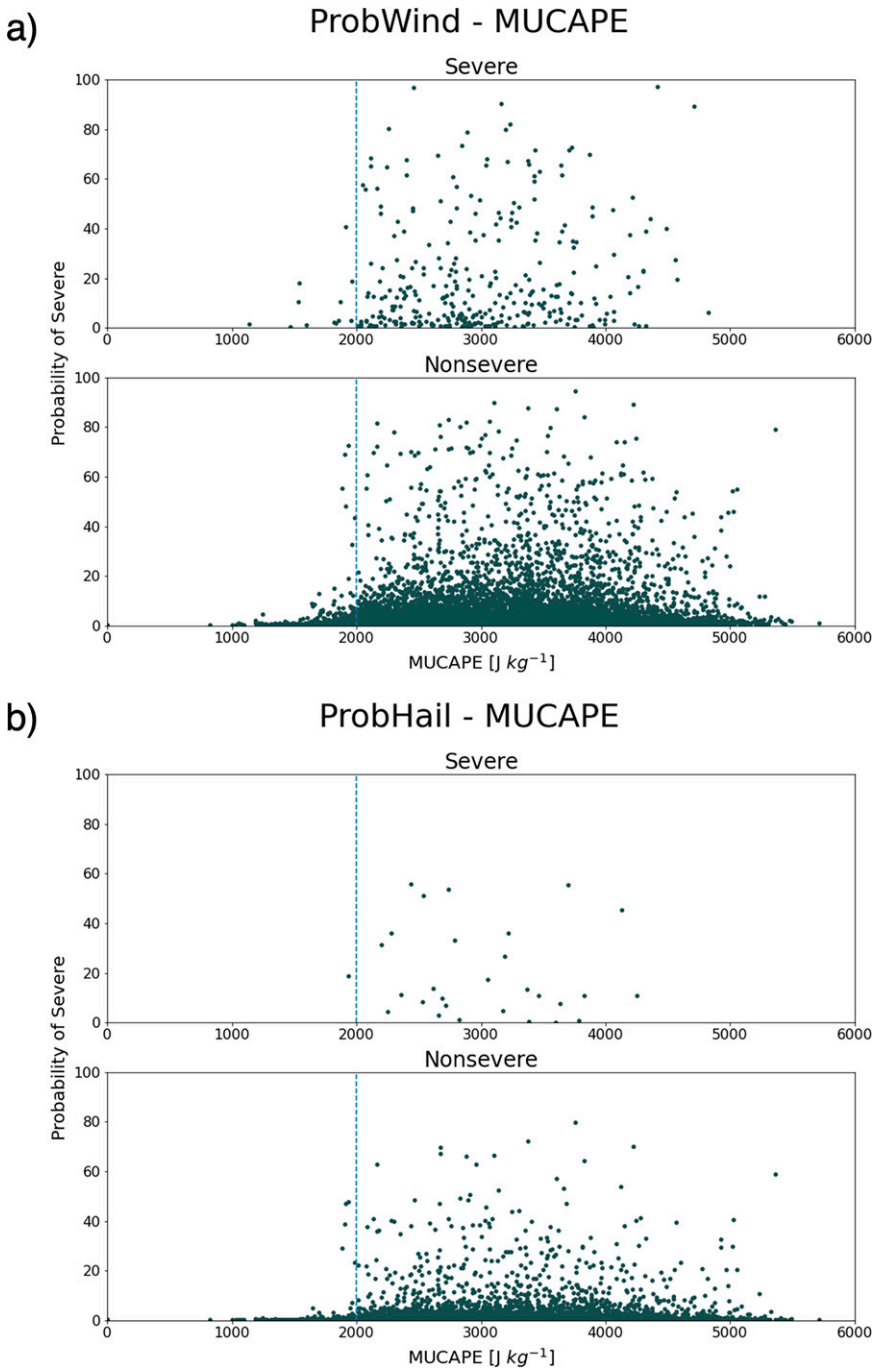


FIG. 9. (a) Values of ProbWind (%) as a function of MUCAPE (J kg$^{-1}$) for severe wind producing storms at the severe time step (at top) and for nonsevere storms at the proxy severe time step (at bottom). (b) Values of ProbHail (%) as a function of MUCAPE (J kg$^{-1}$) for severe hail producing storms at the severe time step (at top) and for nonsevere storms at the proxy severe time step (at bottom).

cases (Fig. 10a), there is no strong correlation between values of EBSHEAR and the probability of pulse severe hail for severe hail objects or nonsevere objects (0.42 and 0.04, respectively). A median EBSHEAR value of 6.50 m s$^{-1}$ is found for severe cases, and only one object is considered an outlier. Furthermore, there is no discernable difference in the distribution of EBSHEAR and ProbHail between severe and nonsevere objects.

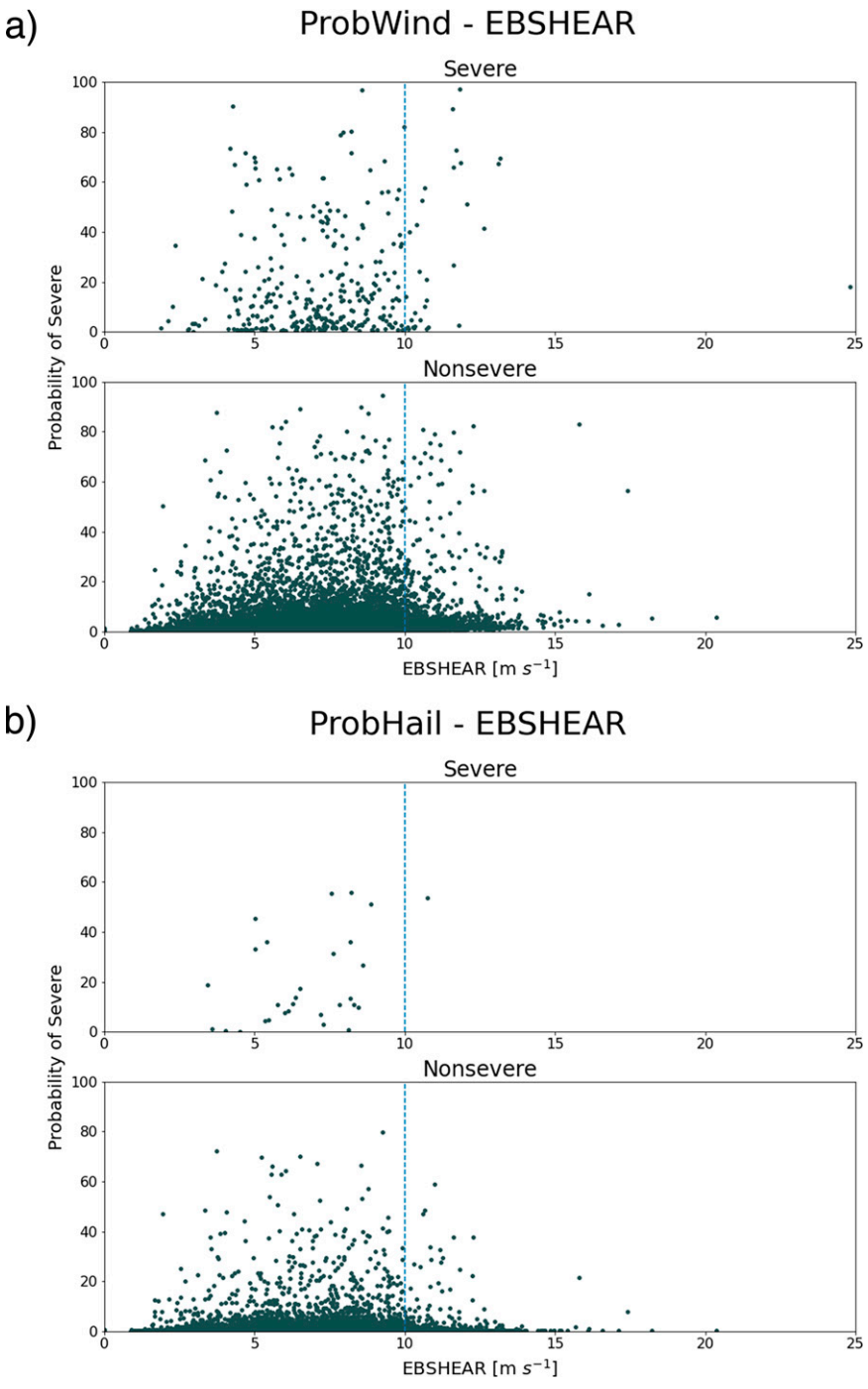


FIG. 10. (a) Values of ProbWind (%) as a function of EBSHEAR (m $s^{-1}$) for severe wind producing storms at the severe time step (at top) andfor nonsevere storms at the proxy severe time step (at bottom). (b) Values of ProbHail (%) as a function of EBSHEAR (m $s^{-1}$) for severe hail producing storms at the severe time step (at top) and for nonsevere storms at the proxy severe time step (at bottom).

To summarize, Figs. 9 and 10 reveal virtually no correlation between values of MUCAPE or EBSHEAR and the increasing probability of pulse severe events. The distributions also differ little between pulse severe storm objects and nonsevere objects. Thus, the thresholds of EBSHEAR and MUCAPE provide little insight into whether a pulse severe storm actually will occur, only that the environment is favorable for them.

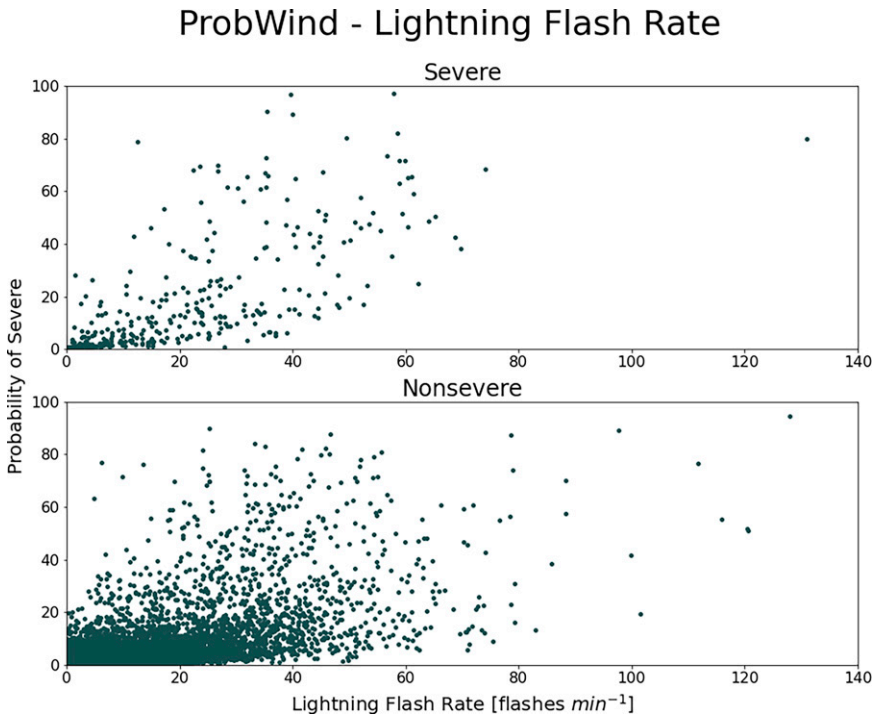


FIG. 11. Values of ProbWind (%) as a function of lightning flash rate (flashes per minute) (top) for severe wind producing storms at the severe time step and (bottom) for nonsevere storms at the proxy severe time step.

Each of the other parameters used in ProbWind and ProbHail (Table 1) was investigated for its utility to differentiate between environments that produce pulse severe events and those that do not. The procedure described above for MUCAPE and EBSHEAR was employed on each. Results show that only ENTLN-derived lightning flash rate used in calculating ProbWind exhibits a reasonably strong positive correlation between predictor value and probability (Fig. 11). Severe wind objects exhibit a correlation of 0.69, and nonsevere objects a correlation of 0.65. Smaller (greater) values of flash rate are associated with smaller (greater) probabilities of severe wind. The median value of flash rate is 18.95 flashes per minute for severe objects, and 1.4 flashes per minute for nonsevere objects. Although there is a reasonable correlation between values of this predictor and probability, it is important to note that there is no clear difference in the distributions of severe and nonsevere cases. Lightning jump algorithms (e.g., Schultz et al. 2009, 2016) have been shown to be useful in nowcasting when storms are about to become severe. Our future research will investigate whether lightning flash rate or lightning jumps can be used to help predict pulse severe thunderstorms and investigate whether GLM-derived flash rates are a useful substitute or supplement to ground based lightning detection.

## 4. Conclusions

Pulse severe thunderstorms are relatively common over the southeast United States during the warm season. However, successfully forecasting their occurrence is a major operational challenge. There is insufficient guidance to distinguish between the few (or no) single cell storms that will become pulse severe on a given day and the great majority (or all) that will not. We investigated whether the PSv2 statistical model developed by NOAA and CIMMS would provide a partial solution to this problem. However, PSv2 was not designed specifically for pulse severe storms, but for the combination of all modes of convection. Although PSv2 has been shown to increase forecaster confidence and increase the lead-time for issuing warnings, its utility for warning specifically on pulse severe storms had not been tested. This research has sought to quantify how well PSv2 performs for pulse severe storms.

We matched PSv2 objects whose pre-storm environment fit the commonly accepted thresholds for pulse severe thunderstorms with severe hail and severe wind events from *Storm Data* and labeled these pulse severe objects. Conversely, any object occurring in a pulse severe environment that did not match with a severe event was labeled a nonsevere object. The total dataset consisted of 365 severe objects—of which 29 matched with severe hail events and 336 matched with severe wind events—and 77 800 nonsevere objects.

Results show that PSv2 pulse severe storm objects persisted between 60 and 130 min and became severe between 20 and 68 min after their creation, i.e., approximately 51% through their duration. The possibility for PSv2 objects to merge or split can extend or reduce the length of their duration. These occurrences may be exacerbated by the tendency of single cell thunderstorms to occur in clusters.

PSv2 objects that corresponded to pulse severe storms exhibited forecast probabilities between 2.55% and 33.27% at the severe time step, and only rarely exceeded 63.04%. However, the severe time step did not always correspond to the greatest probabilities. Maximum probabilities preceding the severe time step were generally greater. The large difference in the number of single cell thunderstorms that do become severe compared to those that do not, the outlier probability values of nonsevere objects, and the low to moderate probabilities found for pulse severe objects indicate that PSv2 provides only limited guidance for forecasting pulse severe storms.

Forecast skill scores were calculated to evaluate how well the NWS predicts pulse severe storms. The POD, FAR, and CSI were found to be 67.40%, 64.96%, and 29.96%, respectively. The average lead time for NWS warnings was calculated to be 9.58 min or 14.21 min, depending on which method was used. These results confirm that pulse severe storms are difficult to forecast in a timely manner.

Pulse severe storm environments have been defined by CAPE $\geq$ 2000 J kg$^{-1}$ and 0–6-km shear $<$ 10 m s$^{-1}$ (Markowski and Richardson 2010b). However, their substitutes MUCAPE and EBSHEAR were found to exhibit only weak correlations with the increasing probability for pulse severe events. Other parameters used in PSv2 (Table 1) exhibited equally small correlation coefficients and no discernable difference in the distribution of probabilities as a function of predictor value. Only ground based lightning flash rate showed a stronger correlation with the probability of severe wind.

These results demonstrate PSV2's limited ability to represent and predict pulse severe thunderstorms in the southeast United States. Our future research will seek to develop a new version of the PSv2 that is designed specifically for pulse severe storms. That is, it will describe the probability that a storm will produce pulse severe wind or hail within the next 60 min, with the goal of improving our ability to predict them. Current results suggest that a beginning point for that research could be a detailed analysis of including GLM data in a future algorithm. PSv2 object data tables already contain GLM data, but they are not used in any of its component probability models. Additional parameters that may prove useful in predicting pulse severe storms include dual-polarization radar products. As a by-product of the future research, we will gain information about the evolution and structure of pulse severe storms, and the environments conducive for them so that forecasters have a greater knowledge base when deciding whether or not to warn.

*Acknowledgments.* Initial support for the lead author was provided by a teaching assistantship at Florida State University. Later support was provided by NOAA CSTAR contract NA20NWS468004 to Florida State University.

*Data availability statement.* Level II radar data are publicly available at https://s3.amazonaws.com/noaa-nexrad-level2/index.html (Amazon 2021). *Storm Data* is publicly available at https://www.ncdc.noaa.gov/stormevents. ProbSevere data were obtained from John Cintineo at the Cooperative Institute of Meteorological Satellite Studies.

## REFERENCES

Amazon, 2021: AWS S3 Explorer, NOAA-Nexrad-Level2. Accessed 14 July 2021, https://s3.amazonaws.com/noaa-nexrad-level2/index.html.

Ashley, W. S., and C. W. Gilson, 2009: A reassessment of U.S. lightning mortality. *Bull. Amer. Meteor. Soc.*, **90**, 1501–1518, https://doi.org/10.1175/2009BAMS2765.1.

Atkins, N. T., and R. M. Wakimoto, 1991: Wet microburst activity over the Southeastern United States: Implications for forecasting. *Wea. Forecasting*, **6**, 470–482, https://doi.org/10.1175/1520-0434(1991)006<0470:WMAOTS>2.0.CO;2.

Benjamin, S., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Bieringer, P., and P. S. Ray, 1996: A comparison of tornado warning lead times with and without NEXRAD Doppler radar. *Wea. Forecasting*, **11**, 47–52, https://doi.org/10.1175/1520-0434(1996)011<0047:ACOTWL>2.0.CO;2.

Byers, H. R., and R. R. Braham Jr., 1949: *The Thunderstorm*. U.S. Government Printing Office, 287 pp.

Cerniglia, C. S., and W. R. Snyder, 2002: Development of warning criteria for severe pulse thunderstorms in the Northeastern United States using the WSR-88D. National Weather Service Eastern Region Tech. Attachment 2002-03, 14 pp., https://www.weather.gov/media/erh/ta2002-03.pdf.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and D. T. Lindsey, 2014: An empirical model for assessing the severe weather potential of developing convection. *Wea. Forecasting*, **29**, 639–653, https://doi.org/10.1175/WAF-D-13-00113.1.

——, and Coauthors, 2018: The NOAA/CIMSS ProbSevere model: Incorporation of total lightning and validation. *Wea. Forecasting*, **33**, 331–345, https://doi.org/10.1175/WAF-D-17-0099.1.

——, M. J. Pavolonis, J. M. Sieglaff, L. Cronce, and J. Brunner, 2020: NOAA ProbSevere v2.0—ProbHail, ProbWind, and ProbTor. *Wea. Forecasting*, **35**, 1523–1543, https://doi.org/10.1175/WAF-D-19-0242.1.

Goodman, S. J., and Coauthors, 2005: The North Alabama lightning mapping array: Recent severe storm observations and future prospects. *Atmos. Res.*, **76**, 423–437, https://doi.org/10.1016/j.atmosres.2004.11.035.

Guillot, E. M., T. M. Smith, V. Lakshmanan, K. L. Elmore, D. W. Burgess, and G. J. Stumpf, 2008: Tornado and severe thunderstorm warning forecast skill and its relationship to storm type. *24th Int. Conf. Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, New Orleans, LA, Amer. Meteor. Soc., 4A.3, https://ams.confex.com/ams/88Annual/techprogram/paper_132244.htm.

Iowa State University, 2021: Iowa Environmental Mesonet Cow (NWS Storm-Based Warning Verification). Iowa State University, accessed 11 October 2021, https://mesonet.agron.iastate.edu/cow/.

Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The Warning Decision Support System–Integrated Information. *Wea. Forecasting*, **22**, 596–612, https://doi.org/10.1175/WAF1009.1.

Markowski, P., and Y. Richardson, Eds., 2010a: Hazards associated with deep moist convection. *Mesoscale Meteorology in Midlatitudes*, Wiley-Blackwell, 292–306.

——, and ——, Eds., 2010b: Organization of isolated convection. *Mesoscale Meteorology in Midlatitudes*, Wiley-Blackwell, 201–213.

Miller, P. W., and T. L. Mote, 2017a: A climatology of weakly forced and pulse thunderstorms in the Southeast United States. *J. Appl. Meteor. Climatol.*, **56**, 3017–3033, https://doi.org/10.1175/JAMC-D-17-0005.1.

——, and ——, 2017b: Standardizing the definition of a "pulse" thunderstorm. *Bull. Amer. Meteor. Soc.*, **98**, 905–913, https://doi.org/10.1175/BAMS-D-16-0064.1.

Murphy, J. D., 2021: Storm data preparation. National Weather Service Instruction 10-1605, NOAA/NWS, 110 pp., https://www.nws.noaa.gov/directives/sym/pd01016005curr.pdf.

NWS Internet Services Team, 2009: National Oceanic and Atmospheric Administration's National Weather Service glossary. NOAA, accessed 14 August 2020, https://w1.weather.gov/glossary/index.php?letter=p.

Rothfusz, L. P., P. T. Schlatter, E. Jacks, and T. M. Smith, 2014: A future warning concept: Forecasting a Continuum of Environmental Threats (FACETs). *Second Symp. on Building a Weather-Ready Nation*, Atlanta, GA, Amer. Meteor. Soc., 2.1, https://ams.confex.com/ams/94Annual/webprogram/Paper232407.html.

——, R. Schneider, D. Novak, K. Klockow-McClain, A. E. Gerard, C. Karstens, G. J. Stumpf, and T. M. Smith, 2018: FACETs: A proposed next-generation paradigm for high-impact weather forecasting. *Bull. Amer. Meteor. Soc.*, **99**, 2025–2043, https://doi.org/10.1175/BAMS-D-16-0100.1.

Schmit, T. J., P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, 2017: A closer look at the ABI on the GOES-R series. *Bull. Amer. Meteor. Soc.*, **98**, 681–698, https://doi.org/10.1175/BAMS-D-15-00230.1.

Schultz, C. J., W. A. Petersen, and L. D. Carey, 2009: Preliminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather. *J. Appl. Meteor. Climatol.*, **48**, 2543–2563, https://doi.org/10.1175/2009JAMC2237.1.

Schultz, E. V., C. J. Schultz, L. D. Carey, D. J. Cecil, and M. Bateman, 2016: Automated storm tracking and the lightning jump algorithm using GOES-R Geostationary Lightning Mapper (GLM) proxy data. *J. Oper. Meteor.*, **4**, 92–107, https://doi.org/10.15191/nwajom.2016.0407.

Sieglaff, J. M., D. C. Hartung, W. F. Feltz, L. M. Cronce, and V. Lakshmanan, 2013: A satellite-based convective cloud object tracking and multipurpose data fusion tool with application to developing convection. *J. Atmos. Oceanic Technol.*, **30**, 510–525, https://doi.org/10.1175/JTECH-D-12-00114.1.

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, https://doi.org/10.1175/WAF-D-11-00115.1.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 1617–1630, https://doi.org/10.1175/BAMS-D-14-00173.1.

Thompson, K. B., M. G. Bateman, and L. D. Carey, 2014: A comparison of two ground-based lightning detection networks against the satellite-based Lightning Imaging Sensor (LIS). *J. Atmos. Oceanic Technol.*, **31**, 2191–2205, https://doi.org/10.1175/JTECH-D-13-00186.1.

Thompson, R., C. M. Mead, and R. Edwards, 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102–115, https://doi.org/10.1175/WAF969.1.

Trapp, R. J., D. M. Wheatley, N. T. Atkins, R. W. Przybylinski, and R. Wolf, 2006: Buyer beware: Some words of caution on the use of severe wind reports in postevent assessment and research. *Wea. Forecasting*, **21**, 408–415, https://doi.org/10.1175/WAF925.1.

Wilk, K. E., L. R. Lemon, and D. W. Burgess, 1979: Interpretation of radar echoes from severe thunderstorms: A series of illustrations with extended captions. Prepared for training of FAA ARTCC Coordinators, National Severe Storms Laboratory, 55 pp.

Wilks, D. S., Ed., 2011: Forecast verification. *Statistical Methods in the Atmospheric Sciences,* Academic Press, 334–340.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. W. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286–303, https://doi.org/10.1175/1520-0434(1998)013<0286:AEHDAF>2.0.CO;2.