Week 2 Severe Weather Forecasts: What Should We Predict?

SOPHIE M. MILLER*

National Weather Center Research Experiences for Undergraduates Program
Norman, Oklahoma,
Iowa State University
Ames, Iowa

MICHAEL E. BALDWIN

Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma
NOAA/NWS/Storm Prediction Center
Norman, Oklahoma

KIMBERLY A. HOOGEWIND

Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma
NOAA/OAR/National Severe Storms Laboratory
Norman, Oklahoma

HAROLD E. BROOKS

NOAA/OAR/National Severe Storms Laboratory
Norman, Oklahoma

ABSTRACT

This study explores various ways to quantify the frequency of severe convective storms within weekly periods, aiming to highlight different aspects of the climatology. The number of days per week where the coverage of severe convective storms exceeded a threshold, as defined by the practically perfect hindcast (PPH) framework, were computed for the period 1991–2022. Analyses examined various hazards and thresholds. Monthly climatologies of weekly severe weather day frequency illustrates that the most active weeks occur during late spring and summer. Weeks with a higher frequency of severe weather days tend to have a greater impact, with the most significant day in the week covering a larger area, more frequent occurrences of significant severe weather, and a higher likelihood of a greater number of injuries and fatalities. The frequency of weeks with major tornado days peaks earlier in the year compared to more active weeks with any type of severe weather. Therefore, hazard-specific weekly forecasts may need to be tailored to specific periods within the year. The annual cycle provides insight into when forecasts for the second week should be made, as a higher number of weekly severe weather days occur in late spring and summer, while the cool season experiences far fewer events over a week.

1. Introduction

Severe convective storms (SCSs), defined by storms with winds $\geq$ 58 mph, hail $\geq$ 1 inch in diameter, and/or a tornado, pose a major threat to life and property each year in the United States. From 1980–2024, the United States had 195 $1+ billion dollar disaster events caused by SCSs totaling $485.2 billion in estimated damages. The most recent period from 2011–2023 is notable with 128 total $1+ billion dollar disaster events, producing $326.3 billion in damages. In 2023, there were 19 billion-dollar disaster events caused by SCSs, totaling $55.2 billion, making it the costliest year on record for billion-dollar SCS events (NOAA National Centers for Environmental Information (NCEI) 2024). The uptick in recent years of billion-dollar SCS events highlights the importance of studying new ways to improve the forecasting of SCS events, as longer event lead times could give people more time to prepare for SCS events, leading to fewer deaths, injuries, and damages from SCSs.

The Storm Prediction Center (SPC) forecasts the likelihood of severe weather within 25-mi of a location for the

Corresponding author address: Sophie Miller, Iowa State University, 3410 Beardshear Hall, Ames, IA 50011
E-mail: sophiemiller9474@gmail.com

1–8 day time frame. Beyond day 8, no operational forecast product for severe weather exists within the National Weather Service. The SPC has historically been less likely to issue a high probabillity SCS forecast for days 4–8 due to decreasing forecasting confidence and skill in numerical weather prediction (NWP) with lead time (Gensini et al. 2020a). However, recent advances in NWP and machine learning leading to the development of probabilistic prediction forecast models (Dyer et al. 2016; Morrison et al. 2020; Hill et al. 2023), and enhanced understanding of SCSs have improved forecasting skills through the day 4–8 time frame (Wurman et al. 2012).

Recent research has shown that there may be skill in forecasting SCSs throughout the 8-14 day time frame. Wang et al. (2021) developed a dynamical-hybrid model to forecast the spatial anomalies of SCS reports using the Global Ensemble Forecast System (GEFS). The weekly mean Supercell Composite Parameter (SCP) anomaly was used as a predictor using a linear regression model on a 0.5° latitude-longitude grid. The model had low skill for forecasting for week 2 to a 0.5° grid area. However, increasing the grid spacing to a 5° latitude-longitude grid improved the forecasting reliability for week 2. In another study by Miller et al. (2020), a dynamical-hybrid SCS prediction model was developed from the European Centre for Medium-Range Weather Forecast (ECMWF) S2S model to predict weekly tornado frequency. Daily synoptic-scale patterns (or "weather regimes") were identified using k-means clustering into 5 common patterns of 500mb geopotential height anomalies, and the weekly frequency of patterns were aggregated and related to the frequency of tornadoes and tornado day occurrence during the month of May. Persistent patterns were found to account for more than 75% of tornado outbreak days, with over 40% characterized by a persistent pattern of troughing in the western U.S. and ridging over the eastern U.S. The model predicts above/below normal frequency of weekly tornado days based on the weather regime frequency within each ensemble member weather, and forecasts were found to be skillful out to week 3 relative to climatology. On a daily level, Gensini and Tippett (2019) used forecasts from the Global Ensemble Forecast System to forecast tornado and hail frequencies. The ensemble mean SCP was used as a proxy for tornado and hail occurrence, and evaluation over the spring months in 2016 and 2017 demonstrated skill out to day 9 for tornadoes and day 12 for hail.

In prior studies, the Madden-Julian Oscillation (MJO) and Global Wind Oscillation (GWO) have been related to tornado and hail frequency anomalies over the CONUS. In Miller and Gensini (2023), the MJO and GWO were used to predict tercile categories of weekly CONUS tornado counts relative to climatology for the week 2-3 time frame. Heidke skill scores (HSS) were examined at the accuracy of the week 2 tornado counts forecast. Every HSS for week 2 was positive, indicating that skill exists with the forecasts. Identification of favorable MJO and GWO patterns for SCS development were identified up to 4 weeks in advance during the second half of May 2019, which is one of the most active periods of severe weather in modern times (Gensini et al. 2019). Using the initial state as the MJO as a predictor, weekly values of above-normal or below-normal tornadic events, hail events, convective available potential energy (CAPE), and storm-relative helicity (SRH) were forecasted for in regions of the Great Plains and American South East in the week 2 to subseasonal time frame during MAMJ when the outgoing longwave radiation-based MJO Index (OMI) ≥ 1. Positive HSS results extended into the week 2-5 time period (Baggett et al. 2018). In evaluating the difference of the MJO between high and low-skill day 10 tornado forecasts, the forecasts with high skill exhibited a strong MJO 15-20 days prior to the tornadic event, leading to a significant modulation of the atmospheric angular momentum (AAM) (Miller and Gensini 2023). Previous research (Gensini and Tippett 2019; Gensini et al. 2020a; Miller et al. 2020) suggests that a significant modulation of the AAM will occur when exhibiting a strong MJO. Low-skill day 10 tornado forecasts will exhibit colder than normal SSTs in the eastern Pacific, a weaker/slower MJO, and no significant modulation in the AAM (Miller and Gensini 2023).

This research is motivated by the question of how to measure an active or inactive severe weather week. Previous studies have measured such weeks by considering spatial anomalies of all severe weather reports(Wang et al. 2021), the weekly frequency of tornado days (Miller et al. 2020), both tornado and hail frequencies (Gensini et al. 2019), and CONUS tornado counts compared to climatological norms (Gensini and Tippett 2019; Miller and Gensini 2023). Despite some agreement that the greatest forecast skill is associated with tornado and hail parameters (Lepore et al. 2017), there is no consensus on how severe weather should be measured or if a specific hazard should be used (e.g., looking at all severe weather, tornadoes, days versus local storm report counts, etc.) or if geography should be accounted for. This study explores climatological characteristics of weekly CONUS severe weather days using various definitions to quantify severe weather occurrence.

2. Data and methodology

SCS reports obtained from the SPC storm database archive for the years 1991–2022 (available online at: `https://www.spc.noaa.gov/archive/`). It is important to note that SCS observations are sensitive to spatial and non-meteorological biases relating to variations in population, measurement estimation, and other discontinuities when interpreting multiple years of reports (Verbout et al. 2006; Potvin et al. 2019; Gensini et al. 2020b) All tornado reports are still utilized, despite the adjustments made in determining damage indicators of tornadoes during the transition from the F-scale to the EF-scale in 2007. Additionally, in 2010, the criteria for severe hail changed from 0.75 to 1 inch in diameter. Hail reports are classified as SCS events if they meet the NWS-defined criteria, regardless of whether the hail met the post-2010 criteria. Injury and fatality counts are included as part of an SCS report from the SPC storm database archive.

All SCS reports are aggregated on daily 1200–1200 UTC 24–h time intervals and are used to calculate PPH probabilities of SCS hazards to a Lambert-conformal grid using 80-km horizontal grid spacing (NCEP grid 211), which is approximately equivalent to the spatial scale of SPC probabilistic forecasts of SCSs within 25-mi of a point. The goal of PPH is to show what a perfect SPC forecast might look like, and thus makes it a useful quantity for verification purposes (Gensini and Tippett 2019). Each storm report is gridded via a nearest-neighbor method and then transformed to a binary field of 1's and 0's. If at least one report falls within a grid box, it is represented by a 1, otherwise it is set to 0. A two dimensional Gaussian filter is applied to smooth the gridded points. The PPH value calculated is set a threshold percentage based on similar thresholds used by the SPC to convert probabilistic forecasts to risk categories (Hitchens et al. 2013) (30% all severe = enhanced risk, 45% all severe = moderate risk, 10% tornado = enhanced risk, 15% tornado = moderate, 30% tornado = high), where all severe is defined as all tornado, hail, and wind reports. The 30% probability corresponds roughly to an enhanced risk, while the 45% probability corresponds to a moderate risk. More details on the conversion of SPC probabilistic forecasts to their categories can be found at `https://www.spc.noaa.gov/misc/SPC_probotlk_info.html`. Significant severe is assessed using a threshold of 15%.

Figure 1 demonstrates on how PPH may differ from a forecast issued by the SPC by overlaying the convective day 1 outlook on March 31st, 2023 issued at 1630z for tornado probabilities compared to the calculated tornado PPH. There are 188 tornado reports from the day resulting in a max PPH of 0.8 or 80%. To evaluate weekly severe weather based on daily frequencies, the daily PPH, after thresholding, is summed over the rolling seven day window the max PPH over the CONUS is greater than or equal to the threshold set. Specifically, this work will focus on different tornado and all severe PPH. Additionally, their significant severe counterparts will also be examined. Finally,the climatology and distributions of severe weather characteristics based on weekly frequencies will be examined. To examine weekly frequency in the context of climatology address this, tercile categories are assigned each

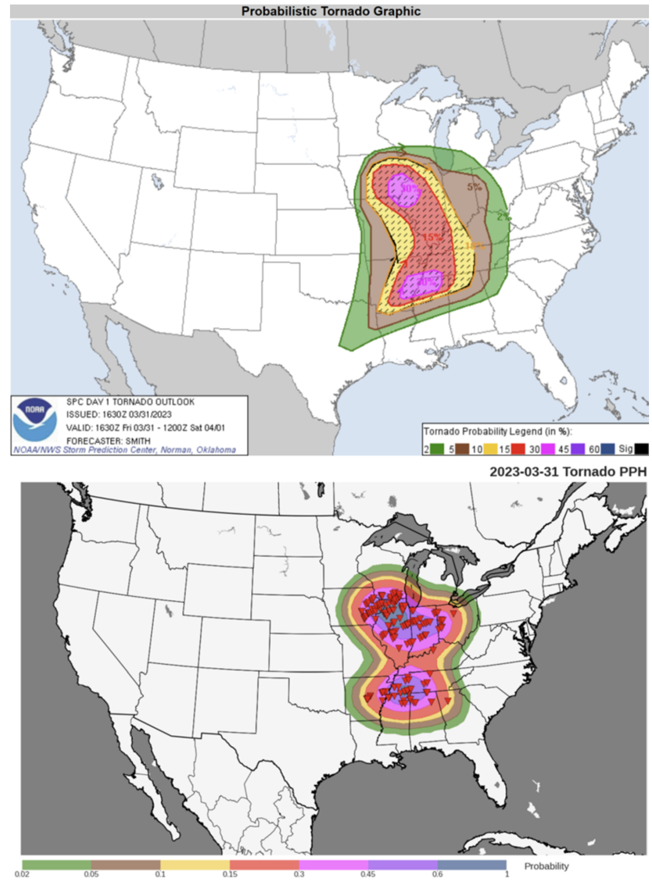

FIG. 1. The 1630 UTC SPC outlook issued on March 31, 2023, for tornado probabilities (top) and tornado PPH (bottom)

rolling 7-day time window a number based on the climatological distribution of SCS weekly frequency activity. -1 (0–33% percentile) is assigned to below average SCS activity weeks, 0 for average (33–66% percentile) SCS activity, and 1 (>66% percentile) for active SCS activity. Categorizing SCS activity based on how active it is relative to climatology helps mitigate bias when comparing the number of days in a week in each month relative to the climatology expectations. Once the terciles are categorized, proper analysis can be done to compare different PPH thresholds and variables.

3. Results

Understanding weekly SCS days relative to climatology can help highlight periods during the year in which SCS is most common, and when it might be more feasible to forecast for week 2. To examine weekly all severe day frequency, the 30% and 45% PPH thresholds were chosen as they approximate an SPC enhanced risk. Figure 2 illustrates the annual cycle of weekly SCS days indicates limited to low activity from November to January. For both

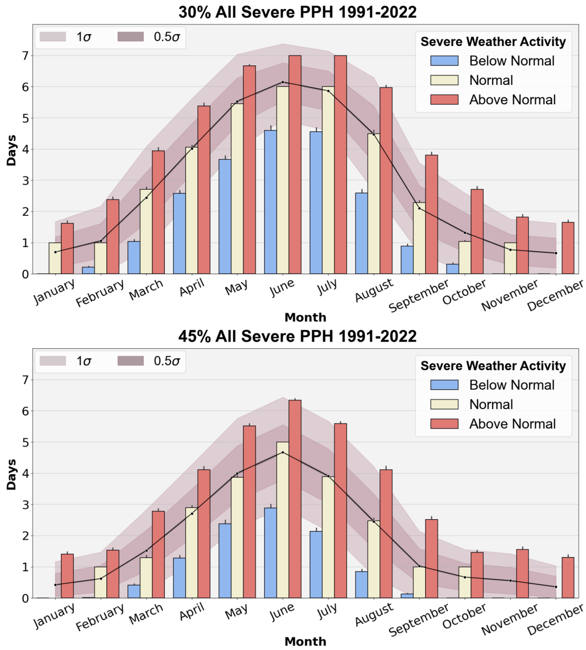

FIG. 2. Number of days (1200-1200 UTC) in a given week per month from 1991 to 2022 that meet or exceed a PPH probability of (a) 30% for all severe (b) 45% all severe. The black line represents the mean count of days. The bars represent the days that are expected to meet or exceed the PPH threshold in above normal (red), normal (yellow), and below normal (blue) SCS activity. The dark and light-shaded regions represent 0.5 and 1 standard deviation (σ) from the climatological mean, respectively

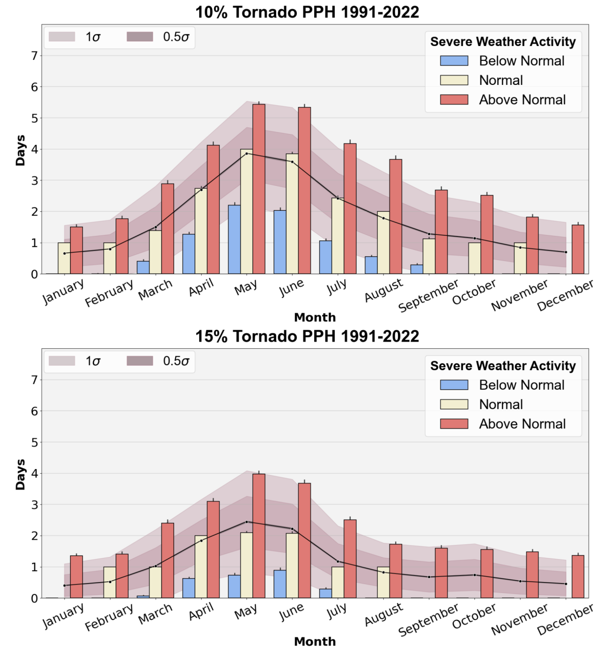

FIG. 3. How many days (1200-1200 UTC) in a given week per month from 1991 to 2022 that meet or exceed a PPH probability of (a) 10% for tornado (b) 15% for tornado. The black line represents the mean count of days. The bars represent the days that are expected to meet or exceed the PPH threshold in above normal (red), normal (yellow), and below normaln (blue) SCS activity. The dark and light-shaded regions represent 0.5 and 1 standard deviation (σ) from the climatological mean, respectively.

thresholds, activity increases every month in late winter and spring, peaking in June and July, followed by a decrease in activity during the late summer and fall months. In June and July, 30% of severe PPH cases show normal SCS activity peaking at 6 out of 7 days, while 45% all severe PPH show the peak occurring in June with 5 out of 7 days meeting the threshold. During late fall/winter, normal SCS activity shows 1 day a week meeting the 30% all severe PPH threshold, except in December when 0 days meet the threshold, compared to 0 days with 45% severe PPH November–February. For 30% of severe PPH cases when SCS activity is below normal, 4–5 days out of 7 days a week meet the threshold, while 45% all severe PPH show the peak occurring in June with 3 out of 7 days meeting the threshold. Below normal SCS activity in fall/winter shows 0 days meeting both thresholds. When SCS activity is above normal, at least one day a week is expected to meet both thresholds (Figure 2).

With the normal and above normal terciles during peak SCS activity for 30% all severe PPH showing 6-7 days a week meeting the thresholds, this could make it difficult to determine an active versus normal week. Compare this with 45% all severe PPH with normal SCS activity showing 5 days a week the threshold is met and above normal SCS activity showing 6–7 the threshold is met. This suggests setting a higher threshold would allow for more clarity of the forecasts. Consequently, during the late fall/winter months, the normal terciles for 30% all severe PPH show 1 day a week in which the threshold is met, except for December when normal SCS activity equals 0 days, compared with 1–2 days a week during above normal SCS activity. During this time period, the 45% severe PPH terciles indicate that the threshold is met 0 days a week, in contrast to 1—2 days a week during above normal SCS activity. This suggests that the threshold used might need to vary seasonally as the number of days in a week cannot be larger than 7 and no less than 1 for multiple terciles (Fig. 2).

In addition to all severe, we can apply the same analysis to 10% and 15% tornado PPH days, which roughly correspond to a tornado-driven SPC enhanced risk. Climatologically, the peak frequency of weekly 10% and 15% tornado PPH days occurs in May, which is earlier than the June peak of 30 and 40% all severe days (Figure 3). The weekly below-average tercile only exists from March–September and April –July. This results from the fact that the upper tail of below-average tercile is zero. Both the weekly 10% and 15% tornado days highlight that a single tornado day quantifies a week as above-average during late fall and winter. Using 10% tornado days as a potential predictand for week 2 forecasts would allow for a longer potential period of the year for forecasts which have all 3 tercile categories as compared to 15% tornado days.

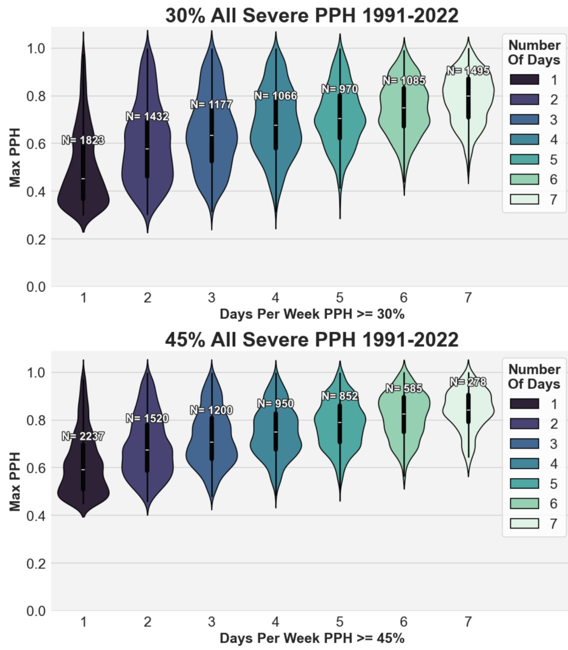

FIG. 4. Distribution of the weekly daily maximum PPH (“Max PPH”) grouped by the number of days per week that meet or exceed a PPH probability of (top) 30% all severe (bottom) 45% all severe. N represents the number of samples from the 1991–2022 period.

Additionally, a larger SCS coverage area leads to a greater potential impact on lives and property, making it crucial to understand how to predict coverage area.The maximum daily PPH value for a given week provides a measure of SCS coverage from the “worst” day of that week. A higher “max PPH” value indicates a greater number of SCS reports and therefore larger areal extent of severe weather. Both the 30% all severe and 45% all severe thresholds show similar trends, where an increased number of days in a given week meeting or exceeding the threshold results in a higher maximum PPH for the week, subsequently expanding the coverage area. A comparison of the maximum PPH distributions between the 30% and 45% all severe thresholds reveals that the 45% all severe threshold results in a higher frequency of larger maximum PPH values. It is important to note the smaller sample size of the 45% all severe, designated from the N value, compared to 30% all severe (Figure 4). As the PPH threshold increases, the sample size from the data will decrease because a higher number of SCS reports will be required to confirm the higher PPH thresholds. In summary, higher PPH thresholds will lead to increased coverage during a week, and meeting or exceeding a certain PPH threshold

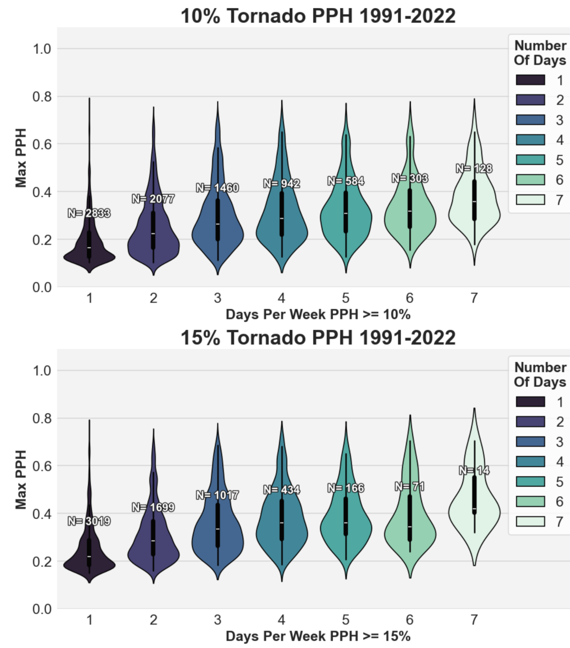

FIG. 5. As in Fig. 4, but for (top) 10% tornado days and (bottom) 15% tornado days.

on more days in a week will result in a higher concentration of severe weather on the ”worst” day of the week.

As assessed previously for all severe, the relationship between the numbers of days and the maximum daily PPH within the week is analyzed. The distributions relay a similar message as Figure 2 for all severe, illustrating the tendency for the highest PPH day of the week to have a higher magnitude, and thus a larger areal coverage of tornadoes. A comparison of the distribution of the 10% tornado and 15% tornado days reveal similar relationships, but it is noted that 15% tornado days occur less frequently (Figure 5).

Given the disproportionate risk to life and property of significant severe convective storms (winds $\geq$ 75 mph, hail $\geq$ 1 inch in diameter, and/or an EF2+ tornado), the relative frequency of significant SCS within a given week is of particular interest. Figure 6 shows that as the number of days meeting or exceeding the 30% or 45% PPH thresholds for any severe SCS increases, so does the number of days when the 15% significant severe PPH threshold is met or exceeded. This indicates that more frequent severe weather days in a given week are associated with an increase in the occurrence of significant severe weather events. Despite there being fewer days that meet the 45% threshold, it is more likely for 15% significant severe weather days to occur when at least one day exceeds the 45% threshold compared to when at least one day meets or exceeds the 30% threshold. This suggests that more frequen days of any severe weather is more

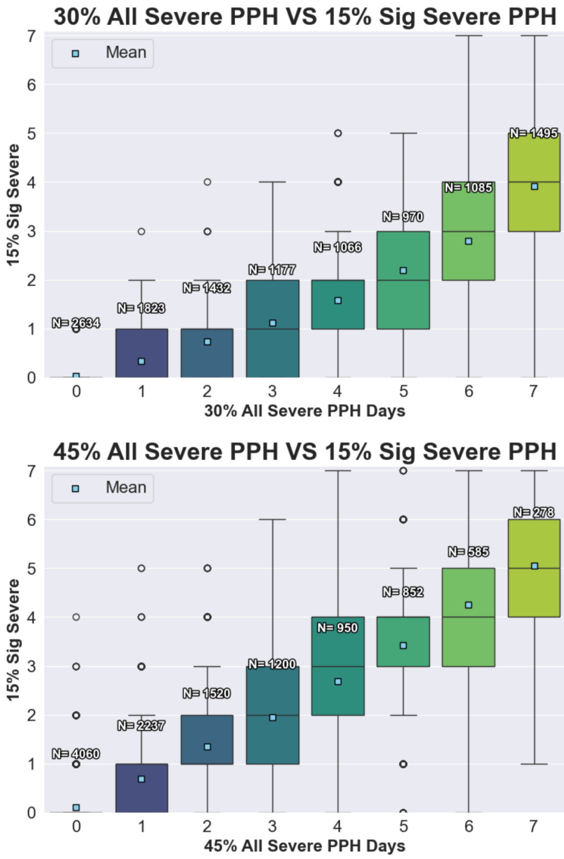

FIG. 6. Distribution of 15% significant severe PPH days relative to weekly all severe (top) 30% PPH days and (bottom) 45% PPH days. N represents the number of samples from the 1991–2022 period.

likely to also have more significant severe weather days within the same period, indicating a greater likelihood of more intense storms.

A similar analysis can be performed with tornado days. Figure 7 relates the weekly frequency of 10%, 15%, and 30% tornado PPH days to the frequency of 10% significant tornado days. When analyzing different tornado probability percentage thresholds, it was observed that meeting the 10% tornado PPH for multiple days in a week does not necessarily lead to an increase in the number of days meeting 10% significant tornado PPH. Even when the 10% tornado threshold is met for all 7 days in a week, only 1-2 days would meet the 10% significant tornado threshold. However, increasing the coverage area from 10% to 15% tornado PPH leads to 1-2 additional days per week where 10% significant tornado PPH is met. Furthermore, when at least one day in a week meets or exceeds 30% tornado PPH, on average, one day will meet 10% significant tornado PPH. As the number of days meeting or exceeding 30% tornado PPH increases, so does the likelihood of meeting 10% significant tornado PPH. When the 30% tornado PPH happens 6 out of 7 days in a week, 5 out of 6 of those days also met the 10% significant tornado PPH, indicating an increase in intense tornado coverage. This means that meeting the 30% tornado PPH threshold is associated with a higher likelihood of intense tornadoes, making it an essential parameter for identifying intense tornado outbreaks.

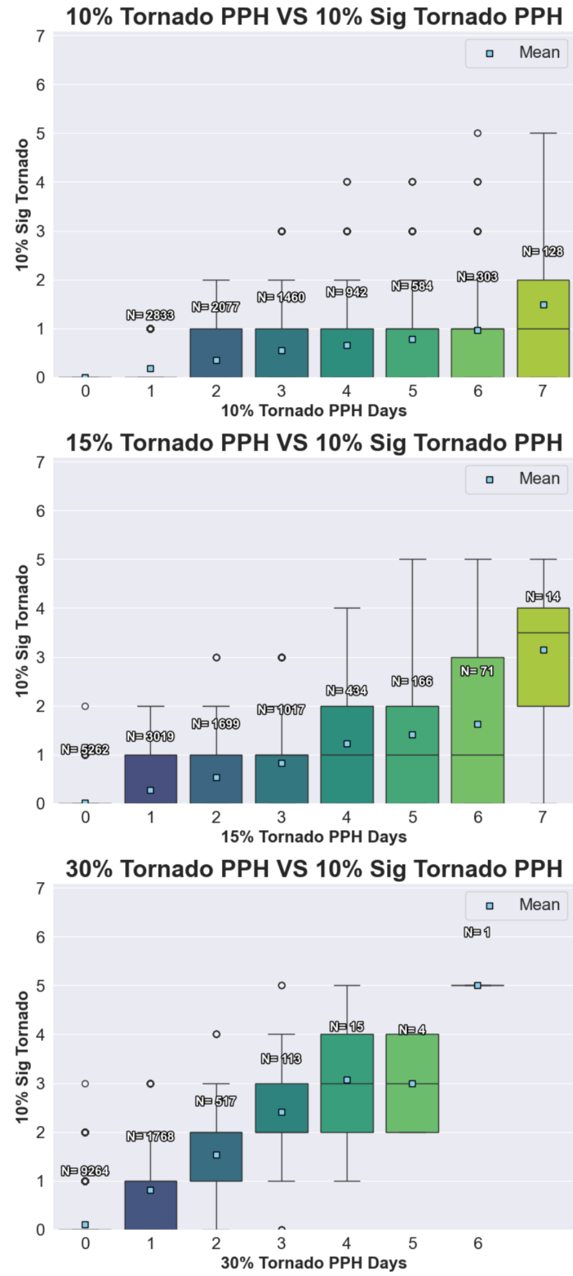

FIG. 7. Distribution of weekly significant severe tornado days relative to weekly (top) 10% tornado days (center) 15% tornado days (bottom) 30% tornado days. N represents the sample size for each boxplot.

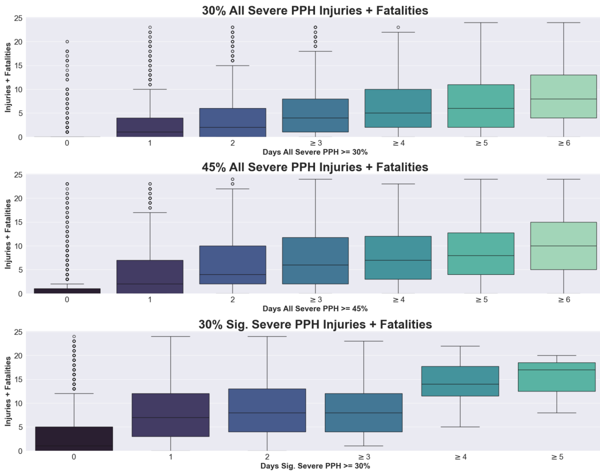

FIG. 8. Distribution of weekly SCS-driven fatalities and injuries relative to the weekly number of (top) 30% all severe, (middle) 45% all severe, and (bottom) 30% significant severe days.

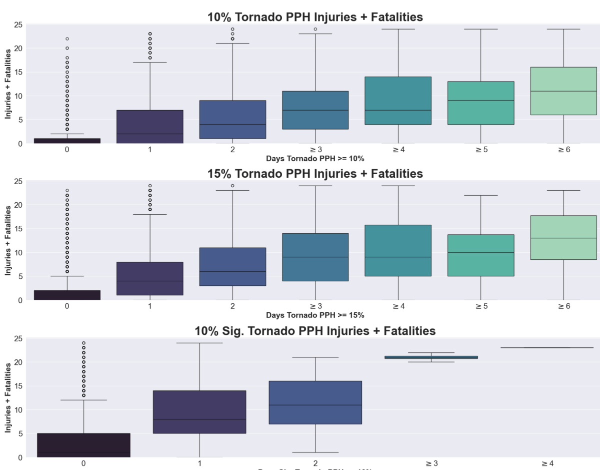

FIG. 9. Distribution of weekly SCS-driven fatalities and injuries relative to the weekly number of (a) 10% tornado (b) 15% tornado (c) 10% significant tornado days over the period 1991–2022.

The coverage area and intensity of SCSs may help forecasters anticipate the types of SCS risks more likely to result in injuries and fatalities. Figure 8 compares weekly severe weather days to the weekly total fatality and injury counts due to SCSs. This comparison was done to determine if an increase in storm coverage, storm intensity, or both play a role in fatalities and injuries. The fatality and injury counts increase the more days in a given week in which the threshold is met or exceeded. When comparing the 30% all severe to 45% all severe, injury and fatality counts remain fairly similar for days in which the threshold is met or exceeded, with a marginal increase exhibited by the 45% all severe. Comparing 30% all severe PPH to 30% significant severe PPH demonstrates that there are substantially higher injury and fatality counts for the 30% significant severe PPH cases (Figure 8). Overall, SCS intensity plays a key role in storm-related fatalities and injuries.

As assessed previously will all severe PPH, tornado PPH values were evaluated to determine if an increase in storm coverage and storm intensity plays a role in fatality and injury counts. When comparing 10% tornado PPH and 15% tornado PPH, it was observed that injury and fatality counts remain fairly similar for both thresholds, with a marginal increase from the 15% tornado PPH. In contrast, comparing the 10% tornado and 10% significant tornado PPH showed that due to the more strict criteria for a significant tornado, it is nearly impossible to obtain a week where all seven days verify as a significant threshold greater than slight. Interestingly, when comparing 10% tornado PPH and 10% significant tornado PPH, for the days in a week that exceed the 10% significant tornado PPH, there is a significantly higher injury and fatality count compared to the 10% tornado PPH versus 15% tornado PPH (Figure 9). This demonstrates that intensity plays a greater role in fatalities and injuries when compared to just coverage area.

4. Summary and discussion

Quantifying severe weather frequency on a weekly basis can take on a number of forms. For example, several studies in the literature have explored predicting anomalous severe weather frequency on a weekly timescale but with different choices of predictand varying from total CONUS tornado and hail count anomalies, tornado days, or even all severe weather report spatial anomalies. Motivated by the fact that an active severe weather week may not be especially active over the same regions, and that local storm reports are significantly impacted by nonmeteorological biases, this study opted to explore weekly severe weather days that incorporate a spatial component criteria as a potential method to assess activity. Rolling seven-day frequencies of all severe (tornado, hail, wind) and tornado days exceeding a specified PPH threshold were analyzed.

The findings suggest that a larger coverage area, as indicated by identifying the maximum PPH over a rolling 7-day period, coupled with the frequency of significant severe weather events (tornadoes and all severe), can serve as an effective method for predicting the overall coverage and intensity of SCSs. However, certain limitations were acknowledged, including the inability to compare hail PPH thresholds, wind PPH thresholds, significant hail PPH thresholds, and significant wind PPH thresholds with the tornado and all severe thresholds.

The study emphasized the challenges of working with smaller sample sizes for specific PPH thresholds, which makes it difficult to fully understand trends. It is also noteworthy that validating the 30% tornado PPH is challenging, as there are only 1 or 2 days per year that meet this threshold. Furthermore, there are only a few instances where 5 or more days in the week met the 30% tornado PPH criteria. The study noted that severe weather is more frequent in late spring and summer and that more frequent severe weather days in a week are associated with a greater impact, covering a larger area and posing a higher risk of causing injuries and fatalities.

This study indicates that weeks with major tornado occurrences typically happen earlier in the year than weeks with heightened severe weather activities. This suggests the importance of creating hazard-specific weekly forecasts tailored to different periods within the year. The study found that there is a greater number of weekly severe weather occurrences during late spring and summer, compared to notably fewer events during the cool season. It also suggests that forecasting during the warm season is more accurate in distinguishing between different levels of activity, while forecasting during the cool season presents more challenges.

5. Acknowledgements

We would like to personally thank Alex Marmo and Dr. Daphne LaDue for doing such an incredible job at running the NWC REU and making this work possible. The corresponding author would like to give a special thanks to Dr. Kimberly Hoogewind, Dr. Michael Baldwin, and Dr. Harold Brooks for their mentoring on the topics relating to this research.

This material is based upon work supported by the National Science Foundation under Grant No. AGS–2050267, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, NOAA, or the U.S. Department of Commerce.

References

Dyer, J., C. Zarzar, P. Amburn, R. Dumais, J. Raby, and J. A. Smith, 2016: Defining the influence of horizontal grid spacing on ensemble uncertainty within a regional modeling framework. *Weather and Forecasting*, **31 (6)**, 1997 – 2017, doi:10.1175/WAF-D-16-0030.1, URL https://journals.ametsoc.org/view/journals/wefo/31/6/waf-d-16-0030_1.xml.

Gensini, V. A., B. S. Barrett, J. T. Allen, D. Gold, and P. Sirvatka, 2020a: The extended-range tornado activity forecast (ertaf) project. *Bulletin of the American Meteorological Society*, **101 (6)**, E700 – E709, doi:10.1175/BAMS-D-19-0188.1, URL https://journals.ametsoc.org/view/journals/bams/101/6/BAMS-D-19-0188.1.xml.

Gensini, V. A., D. Gold, J. T. Allen, and B. S. Barrett, 2019: Extended u.s. tornado outbreak during late may 2019: A forecast of opportunity. *Geophysical Research Letters*, **46 (16)**, 10 150–10 158, doi:https://doi.org/10.1029/2019GL084470, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL084470.

Gensini, V. A., A. M. Haberlie, and P. T. Marsh, 2020b: Practically perfect hindcasts of severe convective storms. *Bulletin of the American Meteorological Society*, **101 (8)**, E1259 – E1278, doi:10.1175/BAMS-D-19-0321.1, URL https://journals.ametsoc.org/view/journals/bams/101/8/bamsD190321.xml.

Gensini, V. A., and M. K. Tippett, 2019: Global ensemble forecast system (gefs) predictions of days 1–15 u.s. tornado and hail frequencies. *Geophysical Research Letters*, **46 (5)**, 2922–2930, doi:https://doi.org/10.1029/2018GL081724, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081724.

Hill, A. J., R. S. Schumacher, and I. L. Jirak, 2023: A new paradigm for medium-range severe weather forecasts: Probabilistic random forest–based predictions. *Weather and Forecasting*, **38 (2)**, 251 – 272, doi:10.1175/WAF-D-22-0143.1, URL https://journals.ametsoc.org/view/journals/wefo/38/2/WAF-D-22-0143.1.xml.

Hitchens, N. M., H. E. Brooks, and M. P. Kay, 2013: Objective limits on forecasting skill of rare events. *Weather and Forecasting*, **28 (2)**, 525 – 534, doi:10.1175/WAF-D-12-00113.1, URL https://journals.ametsoc.org/view/journals/wefo/28/2/waf-d-12-00113_1.xml.

Lepore, C., M. K. Tippett, and J. T. Allen, 2017: Enso-based probabilistic forecasts of march–may u.s. tornado and hail activity. *Geophysical Research Letters*, **44 (17)**, 9093–9101, doi:https://doi.org/10.1002/2017GL074781, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL074781.

Miller, D. E., and V. A. Gensini, 2023: Gefsv12 high- and low-skill day-10 tornado forecasts. *Weather and Forecasting*, **38 (7)**, 1195 – 1207, doi:10.1175/WAF-D-22-0122.1, URL https://journals.ametsoc.org/view/journals/wefo/38/7/WAF-D-22-0122.1.xml.

Miller, D. E., Z. Wang, R. J. Trapp, and D. S. Harnos, 2020: Hybrid prediction of weekly tornado activity out to week 3: Utilizing weather regimes. *Geophysical Research Letters*, **47 (9)**, e2020GL087 253, doi:https://doi.org/10.1029/2020GL087253, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL087253.

Morrison, H., and Coauthors, 2020: Confronting the challenge of modeling cloud and precipitation microphysics. *Journal of Advances in Modeling Earth Systems*, **12 (8)**, e2019MS001 689, doi:https://doi.org/10.1029/2019MS001689, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001689.

NOAA National Centers for Environmental Information (NCEI), 2024: U.S. Billion-Dollar Weather and Climate Disasters (2024). URL https://www.ncei.noaa.gov/access/billions/, doi:10.25921/stkw-7w73.

Potvin, C. K., and Coauthors, 2019: Systematic comparison of convection-allowing models during the 2017 noaa hwt spring forecasting experiment. *Weather and Forecasting*, **34 (5)**, 1395 – 1416, doi:10.1175/WAF-D-19-0056.1, URL https://journals.ametsoc.org/view/journals/wefo/34/5/waf-d-19-0056_1.xml.

Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz, 2006: Evolution of the u.s. tornado database: 1954–2003. *Weather and Forecasting*, **21 (1)**, 86 – 93, doi:10.1175/WAF910.1, URL https://journals.ametsoc.org/view/journals/wefo/21/1/waf910_1.xml.

Wang, H., A. Kumar, A. Diawara, D. DeWitt, and J. Gottschalck, 2021: Dynamical–statistical prediction of week-2 severe weather for the united states. *Weather and Forecasting*, **36 (1)**, 109 – 125, doi:10.1175/WAF-D-20-0009.1, URL https://journals.ametsoc.org/view/journals/wefo/36/1/waf-d-20-0009.1.xml.

Wurman, J., D. Dowell, Y. Richardson, P. Markowski, E. Rasmussen, D. Burgess, L. Wicker, and H. B. Bluestein, 2012: The second verification of the origins of rotation in tornadoes experiment: Vortex2. *Bulletin of the American Meteorological Society*, **93 (8)**, 1147 – 1170, doi:10.1175/BAMS-D-11-00010.1, URL https://journals.ametsoc.org/view/journals/bams/93/8/bams-d-11-00010.1.xml.