UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

# TORNADO WATCH PROBABILITIES IN PRACTICE:
# PERFORMANCE, TRUST, AND USE AMONG EMERGENCY MANAGERS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By VANESSA DUNHAM

Norman, Oklahoma

2026

TORNADO WATCH PROBABILITIES IN PRACTICE:

PERFORMANCE, TRUST, AND USE AMONG EMERGENCY MANAGERS

A THESIS APPROVED FOR THE

SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Makenzie Krocak (Chair)

Dr. Joe Ripberger

Dr. Austin Harris

Dr. Stacey Hitchcock

# Acknowledgements

First and foremost, I would like to deeply thank my advisor, Dr. Makenzie Krocak. Her unwavering guidance, support, and encouragement throughout the last few years has been invaluable to my thesis, and to my growth as a researcher. I would also like to thank my committee members, Drs Joe Ripberger, Austin Harris, and Stacey Hitchcock, as well as Drs Harold Brooks and Matthew Flounory for their time, thoughtful feedback, and valuable insights that immensely helped improve this research. I am sincerely grateful to everyone at IPPRA for many lunch chats that always raised my spirits, and for helping me celebrate the little (and big) accomplishments along the way. And to my NSF SOARS family, Marissa Vara, Vanessa Vincente, Robert Prestly, and all of the Proteges, for allowing me to be a part of such an amazing community and for being my home away from home, away from home. Of course, I have to thank all of the students in the School of Meteorology, especially my cohort. From endlessly working on statistics assignments, to celebrating milestones on Main Street, your support and friendship have made this journey both meaningful and memorable. A special thank you goes to Chris Reeves, for inspiring me to become a meteorologist and telling me OU was *the* place to do it (I really ran with that one). For that I will be forever grateful. I also must thank Dr. Joseph Trujillo-Falcón, for granting me the opportunity that changed the path of my career, and for always providing graduate school and general life guidance whenever I needed it. Thank you for always being in my corner. To all of my family back home in Missouri of whom there are too many to list, thank you for cheering me on and believing in me from the start. To my Mom and Dad, thank you for believing I could accomplish anything I set my mind to, and for instilling a stubborn determination in me that has gotten me to where I am today. Finally, to my fiancé Zane, thank you for your unwavering support whenever I needed it, for always being up for an adventure, and being with me every step of the way, whether we were six hours apart or in the same room.

# Table of Contents

# List of Tables

# List of Figures

# Abstract

Probabilistic information in forecasts was long thought to be too confusing for the public and other key stakeholders. As such, it was not included in many forecasts. However, research over the last twenty years has shown that including probabilistic information increases trust and understanding in forecast information, especially among emergency managers who use probabilistic information as thresholds for decision making in their operations.

Since April 28th, 2006, the Storm Prediction Center (SPC) has officially included two probabilities in tornado watches: the probability of two or more tornadoes occurring in a tornado watch, and the probability of one or more strong (EF2+) tornadoes occurring in a tornado watch. Despite these probabilities being in place for almost twenty years, little to no research has been conducted to understand how these probabilities verify, or how stakeholders, such as emergency managers, use these probabilities in their operations. Utilizing tornado watch and tornado report data, we show that the tornado probabilities associated with tornado watches are overall very reliable, with some slight variations in time and area. Through the Extreme Weather and Emergency Management survey, we find that emergency managers use these probabilities in their operations, and many take increased action as tornado probability values increase. Additionally, while most emergency managers correctly identified the reference class for tornado probabilities within tornado watches, many confused this reference class with the Storm Prediction Centers convective outlook reference class. Severe weather communicators should ensure that the reference class for each product is explicitly defined, and the difference between commonly used products is stressed when communicating this information with stakeholders such as emergency managers.

# Chapter 1 - Introduction

## 1.1 Motivation

Tornado watches are one of the oldest communication methods to inform the public of the potential for tornadoes to occur, dating back to the 1880s in the US and having been consistently issued since 1952 (Coleman et al. 2011). The term 'tornado watch' was changed from 'tornado forecast' in 1965 after a survey team from the National Weather Service (NWS) (then known as the Weather Bureau) determined that the public was not aware of the Weather Bureau's ability to forecast tornadoes (NWS Heritage 2019). After internally testing tornado probabilities within tornado watches in 1997 and 1998 (Vescio and Thompson 2001), the Storm Prediction Center (SPC) officially began issuing seven probabilities in tornado watches on April 28th, 2006: two for tornadoes, two for strong wind, two for hail, and one for combined strong wind and hail. The two tornado probabilities associated with tornado watches are the probability of two or more tornadoes within the watch area, and the probability of one or more strong (EF2+) tornadoes within the watch area. Assessing how well tornado watches and their associated probabilities correspond to observations, or the quality of the probabilities, is a crucial component of understanding their overall 'goodness' (Murphy 1993).While many studies have investigated different components of the quality (Murphy 1993) of tornado warnings and their interactions with tornado watches (Davies and Fischer 2009; Chamberlain et al. 2023; Anthony and Leftwich 1992; Krocak and Brooks 2018, 2021; Krocak et al. 2021; Anderson-Frey et al. 2016), there is a lack of research surrounding the quality of tornado watches, and even less so for the tornado probabilities associated with tornado watches.

While forecasts like tornado watches and their associated probabilities inherently do not hold any value, value can be produced by the decisions made by the forecast users while using the information included in the forecast product, and is another way to measure the goodness of a forecast (Murphy 1993). For example, previous research has investigated how emergency managers use uncertainty information in their decision making process and the kinds of information they deem important throughout the severe weather event timeline (Cross and LaDue 2021; Ernst et al. 2018; Hogg et al. 2026; Maciag et al. 2025; Olson et al. 2023; Wanless et al. 2023), with results indicating that different types of information are valuable at different points leading up to a severe weather event. However, there is a lack of understanding of how emergency managers incorporate the tornado probabilities from tornado watches into their decision making process. This study seeks to begin to fill this knowledge gap by simultaneously investigating the quality of tornado probabilities in tornado watches, and how emergency managers use these probabilities in their operations. The results of this study may further stipulate how to improve the design of the probabilities associated with tornado watches to better communicate threats to key stakeholders, potentially increasing the self-efficacy of emergency managers to protect their jurisdiction.

## 1.2 Probabilities/Uncertainty in Forecasts and Decision Making

Every severe weather forecast inherently contains uncertainty, but how to best communicate this uncertainty while making sure members of the public can make appropriate protective-action decisions remains challenging. Many communicators remain hesitant to include numeric probability information in their forecasts for fear that it could be too confusing and difficult to understand for members of the public, thus forecasters and other weather communicators often do not include this information in severe weather forecasts (Ripberger et al. 2022). For probability of precipitation (PoP) forecasts, which have been around for decades, the majority of the public do not fully understand the true meaning of a PoP forecast, mainly due to the lack of explanation for what a PoP forecast means (Morss et al. 2008). However, research has shown that explicit mentions of uncertainty and probabilistic information in forecasts, such as probability of occurrence or exceedance, or probabilistic range forecasts, allow members of the public to have more trust and understanding in the information provided and make decisions better suited to their needs (Joslyn and LeClerc 2012, 2013; Morss et al. 2008; Ripberger et al. 2022). Additionally, members of the public prefer forecasts that include uncertainty information, even when compared to deterministic information in forecasts (Morss et al. 2008; Rosen et al. 2021). Furthermore, including explicit numeric uncertainty information, such as a probability, allows for better decision making under uncertainty, especially when combined with deterministic information (Joslyn and LeClerc 2012, 2013). In order for probabilistic information to be processed quickly, it needs to be presented in a succinct and descriptive way, ensuring that no "information overload" occurs (Ripberger et al. 2022). However, in general, the inclusion of probabilistic information in forecasts allows members of the public to better understand the information provided and make decisions that are more beneficial to them.

Warnings for severe convective storms have provided a useful context for the study of probabilistic weather forecasts. For example, when analyzing the difference in public sheltering decisions with probabilistic and deterministic tornado warning polygons, probabilistic polygons had higher rates of understanding, and better sheltering decisions (Qin et al. 2024). However, when comparing deterministic polygons and gradient polygons (where the center of the polygon is a darker color than the outside of the polygon), gradient polygons were not more beneficial for decision making than deterministic polygons (Jon et al. 2018). In the case of temperature and precipitation probabilistic outlooks, using a discrete legend for probabilities increases trust in and understanding of the outlooks by professional users and the general public (Gerst et al. 2020). Even in the context of convective storm warnings, probabilistic information can improve the public's trust and understanding in the forecast, so long as probabilistic information is well-communicated.

## 1.3 Tornado Probabilities in Watches

While there is an abundance of literature regarding tornado warning verification, little research has been conducted regarding tornado watch verification. A study conducted in 2021 by Krocak and Brooks analyzed how watch type can impact the quality of tornado warnings, and found that as watch type increases in severity from tornado watch to a potentially dangerous situation, the probability of detection, false alarm ratio, and lead time for tornado warnings all improve. Additionally, tornado warnings that were issued outside of a watch had a mean lead time of 5.5 minutes, and the warnings inside a particularly dangerous situation tornado watch had a mean lead time of 15.1 minutes (Krocak and Brooks 2021).

Outside of Krocak and Brooks 2021, there is a lack of research surrounding tornado watches and the two tornado probabilities associated with them: the probability of two or more tornadoes and the probability of one or more strong tornadoes within a tornado watch. In 1997 and 1998, the SPC conducted an internal test to see if forecasters could accurately issue tornado probabilities within tornado watches (Vescio and Thompson 2001). They focused on issuing two probabilities in watches: the probability of one or more tornadoes and three or more tornadoes occurring in the watch area. The results of this internal test showed that these probabilities were extremely reliable, and from there the SPC decided to move forward with issuing probabilities in watches (Vescio and Thompson 2001). The reliability analyzed in Vescio and Thompson 2001 is one component of a forecasts' quality, and shows the correspondence between the conditional mean observation and conditioning forecast, averaged over all forecasts (Murphy 1993). Therefore, reliability is a beneficial way to assess the quality of the probabilities in severe watches due to the binary nature of the probabilities (Wilks 2011).

On April 28th 2006 and April 25th 2006, the Storm Prediction Center (SPC) began officially issuing tornado watches and severe thunderstorm watches with seven probabilities regarding tornadoes, severe hail, and wind events, with the two associated tornado probabilities being: the probability of two or more tornadoes and the probability of one or more strong tornadoes (EF2+) within the watch area. While the probabilities regarding hail and wind have also not been analyzed, tornadoes are the main threat in a tornado watch, and thus were selected to be analyzed for this study. It is important to note that there are different definitions of EF2+ strength tornadoes, however since the probability is explicitly 'one or more strong tornadoes (EF2+)', we will be referring to them as 'strong' tornadoes in this study. While these probabilities are different from the initial internal test at SPC, they provide more information regarding the potential strength and number of tornadoes within the watch area. Given the lack of overall verification surrounding tornado watches, and more specifically the tornado probabilities associated with them, this study examines the quality of tornado probabilities in tornado watches by analyzing the reliability and accuracy of them over time, season, time of day, and geographic region, in addition to investigating four cases where tornado watches under or over verified.

## 1.4 Emergency Managers and Severe Weather Information

Since forecasts, such as the tornado probabilities in tornado watches, hold no value until users make decisions based on the forecast information, understanding how users apply this information in their decision making process is crucial to know if value is assigned to these probabilities. While members of the public and EMs both make decisions based on severe weather information to keep themselves and others safe, EMs have the added responsibility of ensuring their jurisdictions are as prepared as possible when severe weather occurs. Additionally, since the tornado probabilities associated with tornado watches are only available on the SPC's website and are not typically included in information sent to the public, EMs may be the main user and source of value for these probabilities.

EMs operate on a variety of different levels, such as city, county, state, regional, or national, and ultimately serve as a focal point of decision making for their jurisdiction. Much of our knowledge of the kinds of information EMs deem important and where they receive weather information from comes from the Extreme Weather and Emergency Management Survey (WxEM), which routinely fields surveys to a nationwide panel of EMs across a variety of jurisdiction sizes and includes questions that pertain to NWS operations (Wanless et al. 2023). Specifically for severe weather, past WxWM surveys have asked EMs to rank information attributes in order of importance, where EMs deemed severity and protective action information increasingly important leading up to a severe weather event. The need for severity information becomes the top piece of information at 4 hours prior to a severe weather event, which is generally the time frame watches are issued (Wanless et al. 2023). Another study conducted with WxEM found that EM's heavily rely on their local NWS Weather Forecast Offices to receive severe weather information, but also seek information from a wide variety of sources, such as local TV stations and private forecasting companies (Hogg et al. 2026). EMs were also reported to primarily use social media to disseminate information to the public, while email was primarily used for internal communication.

Another study that conducted interviews with EMs found that as the potential severe weather event approached in time, EMs searched for increased forecast detail, and would seek more details about the event from their local NWS Weather Forecast Office (WFO) (Ernst et al. 2018). The EMs in the study described using uncertainty information as thresholds for response action, and noted that not all the information they received had uncertainty information which would have been useful to note when certain probability thresholds had been passed (Ernst et al. 2018). In addition, EMs would rather have uncertainty information about the event even if nothing occurs, so they can be prepared for every scenario, especially having information related to ‘worst case’ scenarios (Cross and LaDue 2021). Based on EMs need for increased severity information, forecast detail, and uncertainty information leading up to severe weather events, the tornado probabilities associated with tornado watches could be of great importance to EMs given they include all of this information. As such, this study uses a nation-wide panel of EMs to understand if, and how they use the tornado probabilities associated with tornado watches in their operations.

# Chapter 2 - Tornado Probabilities in Watches

## 2.1 Tornado Watch Verification Data and Methods

All analysis for tornado watch verification was completed in R, utilizing the tidyverse, sf, lubridate, units, maps, plotly, suncalc, and tigris packages (R Core Team 2024). Tornado watch data, including PDS tornado watches, for this study came from Iowa State University's Iowa

Environmental Mesonet (IEM 2025), and included three datasets with data across the continental US. Two spatial datasets were used to represent the watches: watch *polygon* shapefiles and watch *county outline* shapefiles. While county outlines are the official watch shape used by the Storm Prediction Center, many platforms still utilize polygon outlines to visualize the extent of tornado watches. Thus, we used both spatial representations in this study. Due to the size and shape differences between polygon and county outlines, there is a potential for a tornado to occur in the polygon but not in the county outline and vice versa. For this analysis, the largest geographic area was utilized, as seen in Figure 2.1. For example, if the polygon extended beyond the county area, the area utilized in the analysis included the extension of the polygon area beyond the county area.

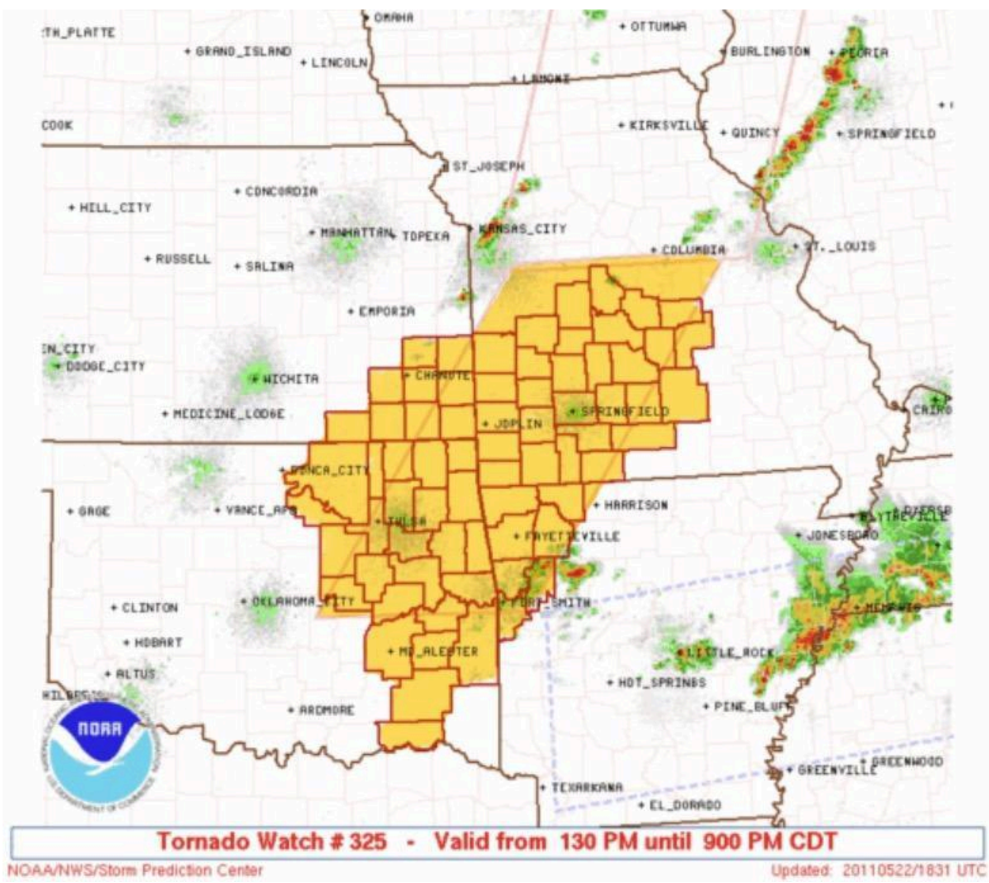


Figure 2.1: Tornado watch valid 05/22/2011 from 1:30pm to 9pm CDT. Yellow region indicates area used for tornado watch.

Neither spatial dataset included probability information, and therefore the last dataset, watch *county CSV file*, was merged with the *county outline* shapefiles to ensure probabilities were included in this dataset. A total of 4,084 tornado watches were included in the county data. All watch data included an issue and expiration time, watch geometry, and the initial watch number, in addition to other information. Issue and expiration times were converted to UTC, and a unique identifier for each watch was created for all files. Four tornado watches out of the 4,084 were removed at this stage because they were not included in the polygon data, and two watches were removed due to them being duplicates, for a total of 4,078 tornado watches.

Tornado report data came from the SPC SVRGIS tornado report database (Marsh and Elliott 2025), and included information about the starting and ending location, time, magnitude of the tornado, state, injuries and fatalities, etc. These reports were used as a proxy for actual occurrence of tornadoes. It is important to note that there is the potential for population density to impact report density, in addition to general reporting errors (Anderson et al. 2007, Elsner et al. 2013, Potvin et al. 2022). However, this discrepancy has decreased over time, especially between the 1970s and mid 2000s (Elsner et al. 2013). All tornado watch and tornado report data were filtered from April 28th, 2006 (when probabilities were officially implemented by SPC in tornado watches) to December 31st, 2024. After finalizing the watch datasets, confirmed tornado reports from the SPC SVRGIS tornado report database were intersected with tornado watches spatially and temporally. A new column was added to the merged dataframes to distinguish whether a confirmed tornado occurred within a watch area while the watch was in effect. Overall, 10,632 of 23,192, or 45.8% of tornado reports did not occur within a tornado watch while a watch was in effect, and 1,637 of 4,078, or 40.1% of tornado watches had no confirmed tornado report(s) occur while they were in effect.

To conduct the verification of tornado probabilities in tornado watches, 64 tornado watches were removed from the dataset due to not including probabilities, leaving 4,014 tornado watches with probabilities, and the number of tornadoes and strong tornadoes per watch were counted. The tornado counts were then merged back into the watch dataframes, where the watches were then grouped by probability and the number of verified watches per bin was calculated for the "two or more tornadoes" probability and the "one or more strong tornado" probability. The probabilities that can be issued by forecasters are 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 95 percent. Additionally, these probabilities are placed into a categorical rating of low (<2% to 20% for the probability of one or more strong tornadoes, 5% to 20% for the probability of two or more tornadoes), moderate (30% to 60% for both probabilities), and high (70% to greater than 95%). For tornado watches, the most commonly issued probabilities are 40% for the probability of two or more tornadoes, and 20% for the probability of one or more strong tornadoes, with 1443 tornado watches having these probabilities together, as seen in Figure 2.2. This is reflected in the categories each tornado watch probability is placed in, and is represented spatially in Figure 2.3. Additionally, the climatology and uncertainty values for probability values are 0.44 and 0.25 for the probability of two or more tornadoes and is 0.21 and 0.16 for the probability of one or more strong tornadoes respectively.

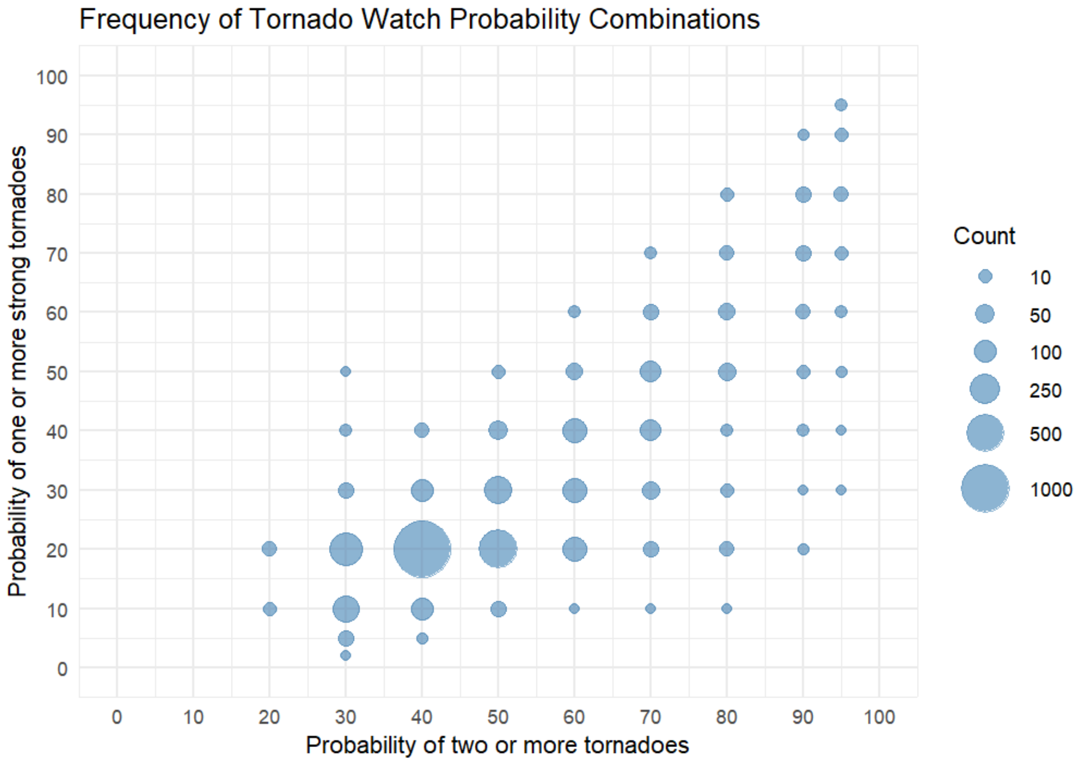



Figure 2.2: Frequency of tornado watch probability combinations for all 4,014 tornado watches.

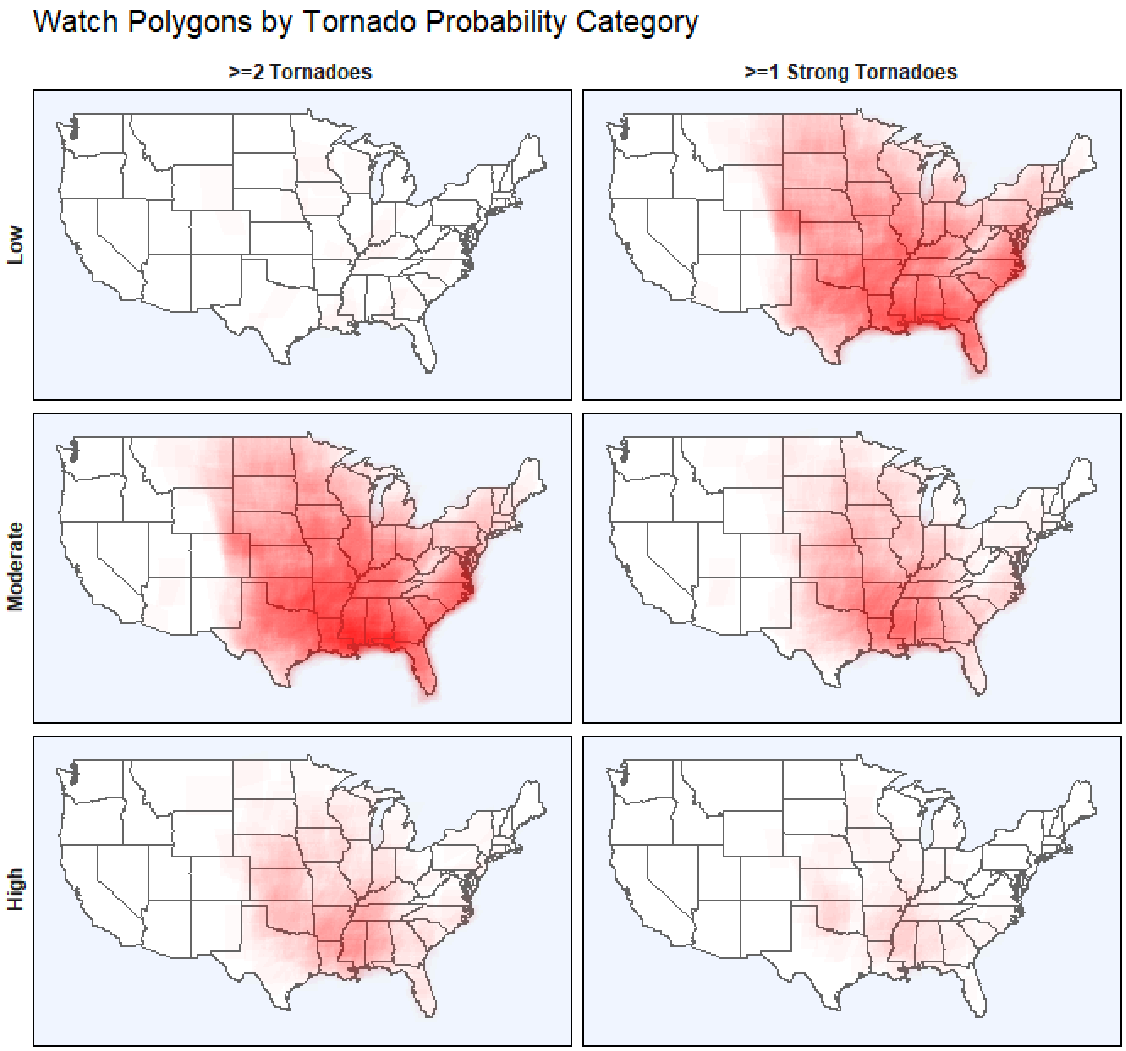



Figure 2.3: Heat map of all tornado watch polygons across the continental US by probability and category.

After the initial analysis was complete, data was also stratified by year, time of year, time of day, and geographic location. Any probability that had less than 20 watches was removed from all plots to ensure sample sizes were large enough to interpret. To stratify data over time by years, the data was divided as evenly as possible over time to ensure large enough sample sizes for each year group. Thus, the 19 year period was divided into three groups, 2006 to 2011, 2012 to 2017, and 2018 to 2024. To assess probability performance by time of year, watches were grouped by meteorological seasons, or December, January, and February for winter, March, May, and April for spring, June, July, and August for summer, and September, October, and November for fall. To stratify data by time of day, the issuance time and centroid point of each watch was used to determine the local sunrise/sunset time. Then, each watch was assigned as either occurring during the day or at night following the local sunrise and sunset time at the centroid of each watch. For the geographic region analysis, the continental US was split into five regions based on environmental characteristics as seen in Figure 2.4 (adapted from Anderson Frey et al. 2016) and each watch was assigned to a region based on the centroid point.

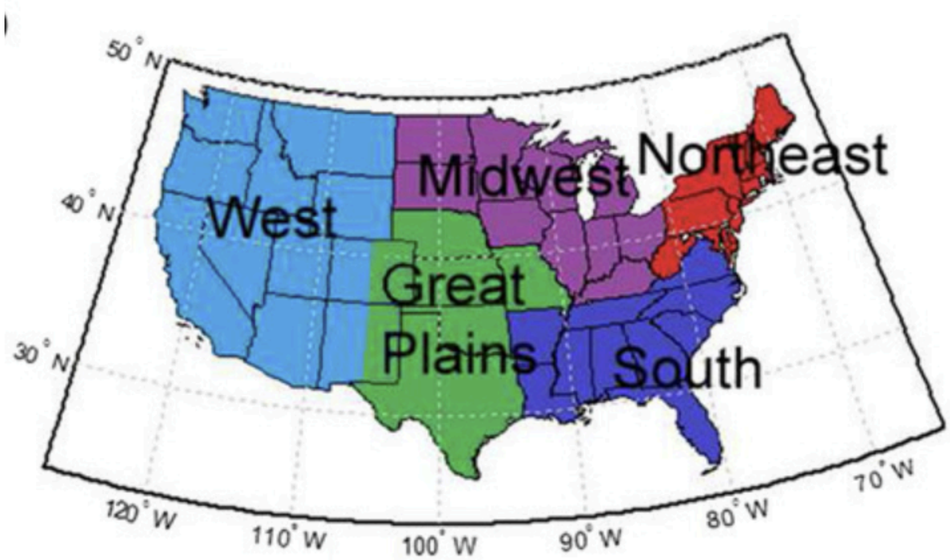



Figure 2.4: Five geographic regions across the continental US, with Colorado and New Mexico split at the 105th meridian. Adapted from Anderson-Frey et al. (2016). Their Figure 1d.

## 2.2 Overall Tornado Watch Performance Results

Prior to analyzing tornado watch probabilities specifically, overall tornado watch performance was evaluated to provide more context of the tornado watches being analyzed.

Figure 2.5a shows that from 2006-2024, the percent of tornado reports within tornado watches became less variable, hovering around 50% as time increases. Figure 2.5b illustrates that tornado reports are more likely to be in tornado watches during the winter and fall months, compared to the summer months (when overall tornado reports are lower, (Krocak and Brooks 2018)). Looking at time of day, tornado reports that occur in the evening and early morning are more likely to be in a tornado watch compared to midday, as seen in Figure 2.5c. Figure 2.5d shows that as tornadoes increase in strength on the EF scale, they are more likely to be within a tornado watch, with EF5 tornadoes occurring 100% of the time within a tornado watch. In conjunction with tornado strength, Figures 2.5e and 2.5f show that as the number of fatalities and injuries increase, the associated tornadoes are more likely to be in a tornado watch. These results indicate that the strongest tornadoes that cause the most impacts are more likely to be in watches compared to weaker tornadoes.

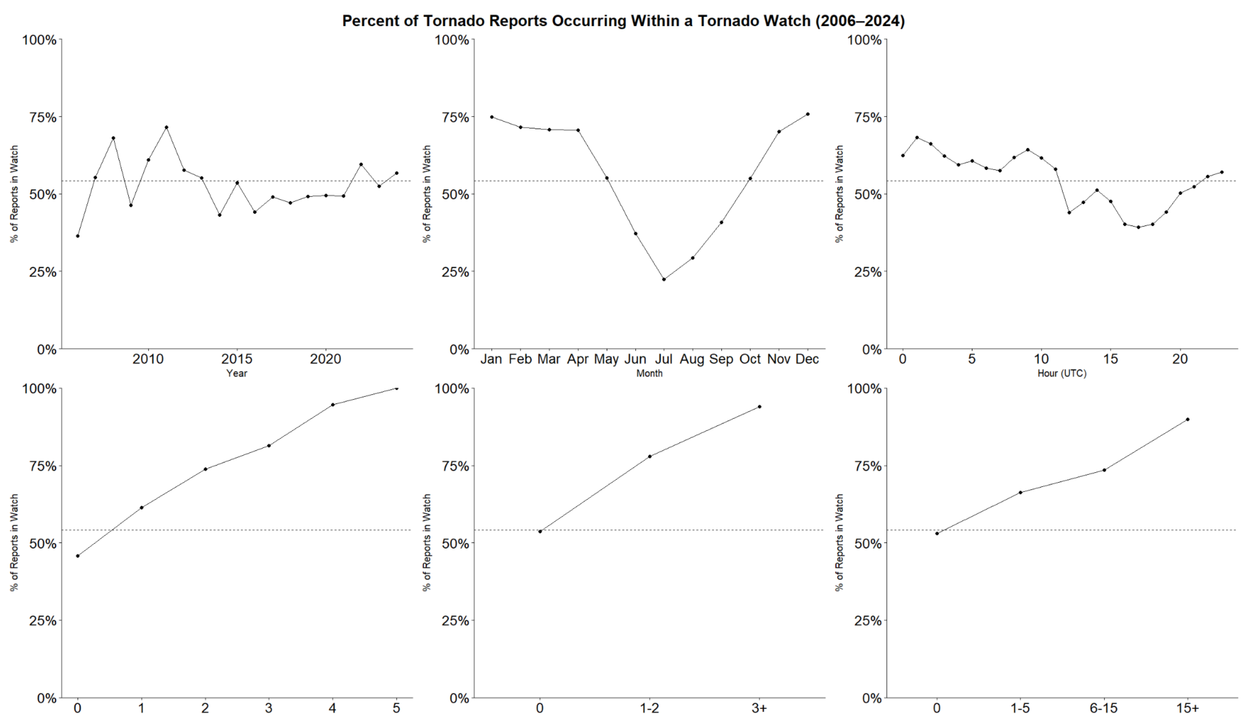


Figure 2.5: 2.5a, percent of tornado reports within a tornado watch by year, 2.5b, percent of tornado reports within a tornado watch by month, 2.5c, percent of tornado reports within a tornado watch by time of day, 2.5d, percent of tornado reports in a tornado watch by tornado strength on the EF scale, 2.5e, percent of tornado reports within a tornado watch by number of fatalities associated with a tornado, 2.5f, percent of tornado reports within a tornado watch by number of injuries associated with a tornado.

## 2.3 Tornado Watch Probability Verification Results

Next, we evaluated the two tornado probabilities associated with tornado watches: the probability of two or more tornadoes (of any strength) occurring in a tornado watch, and the probability of one or more strong (EF2+) tornadoes occurring in a tornado watch. In addition to the overall analysis, the results were also stratified by year, season, day versus night, and geographic region. The analysis of the reliability of tornado probabilities associated with tornado watches are all displayed as reliability diagrams. Reliability itself is one of the many ways used to evaluate the quality of forecasts, and is defined as the correspondence between the conditional mean observation (the observed probabilities) and the conditioning forecast (the forecast probability) (Murphy 1993). Reliability diagrams are specifically used to show the calibration of probability forecasts for binary events, with the X-axis representing the forecast probability, and the Y-axis representing the observed probability (Wilks 2011). This allows for 'perfect' reliability to be the 1:1 line diagonally down the middle of the plot. When the forecast probability is the same or very close to the observed probability, this is considered to be 'good' reliability. When the observed probability is not close to the forecast probability, this is considered poor reliability. Poor reliability can either be 'under-forecast', when the observed probability is higher than the forecast probability, or 'over-forecast', when the observed probability is lower than the forecast probability. The probabilities referenced in this section refer to the *forecasted* probabilities in each tornado watch.

Additionally, Brier scores (BS) and Brier skill scores (BSS) were calculated for each forecast probability to provide extra insight to the overall reliability diagram for both tornado probabilities. The Brier score is used to measure the error in probabilistic forecasts, typically with binary outcomes, with a score of zero representing perfect calibration and score of one being the worst possible score (Brier 1950; Wilks 2011). The Brier score for *n* forecasts is defined as

$$\mathrm{BS} = \frac{1}{n}\sum_{i=1}^{n}(f_i - o_i)^2$$

where $f_i$ is the forecast probability, and $o_i$ is the outcome, with $o_i = 1$ if the event does occur, and $o_i = 0$ if it does not (Mason 2004). In the context of the tornado probabilities associated with tornado watches, the event does occur if two or more tornadoes, or one or more strong tornadoes occur within the watch area. The BS equation can be broken down into this decomposition

$$\mathrm{BS} = \overline{o}(1 - \overline{o}) + \frac{1}{n}\sum_{k=1}^{m} n_k(f_k - \overline{o}_k)^2 - \frac{1}{n}\sum_{k=1}^{m} n_k(\overline{o}_k - \overline{o})^2$$

where *o* is the climatological probability of the event, and *m* represents the number of probability bins of the forecast (Murphy 1973). The forecast probability of bin *k* is $f_k$, and the observed frequency of occurrence of the event for $f_k$ is $o_k$. These three terms can also be written as

BS = UNC + REL - RES

where UNC is uncertainty, REL is reliability, and RES is resolution. The uncertainty term is not reliant on the forecast probability, and is only a function of the uncertainty of the event that is forecasted. Thus, resolution and reliability are the only terms that illustrate the performance of the forecast (Mason 2004).

The Brier skill score (BSS), is written as

$$BSS = 1 - \frac{BS}{BS_{ref}}$$

with $BS_{ref}$ being the reference forecast (Wilks 2011), or in this case, the Brier score of the climatological probability for the probability of two or more tornadoes and the probability of one or more strong tornadoes within the watch area. The BSS shows the relative accuracy of a forecast compared to a reference forecast (in this case, climatology), and measures improvement from this reference, with 1.0 being a perfect forecast, 0.0 showing no improvement from the reference, and any negative value being worse than the reference. Given the equation for the Brier score and Brier skill score, the climatological BSS equation can be written as

$$BSS_{clim} = \frac{RES - REL}{UNC}$$

meaning skill is demonstrated whenever RES is greater than REL (Mason 2004). These calculations were not performed on any stratifications of the tornado watches, as sample size rapidly decreases when parsing the data.

### 2.3.1 Overall Reliability and Performance of Tornado Probabilities in Tornado Watches

Figures 2.6 and 2.7, and Tables A.1 and A.2 indicate that overall, the probabilities associated with tornado watches are generally very reliable. Certain percents, such as 30%, 40%, and 90% for the probability of two or more tornadoes, and 5%, 10%, 20%, 30%, 50%, and 80% for the probability of one or more strong tornadoes are slightly under the perfect reliability line, indicating a slight over-forecast of these probabilities. Other percentages, such as 20%, 70%, and 95% for the probability of two or more tornadoes, and 60% and 70% for the probability of one or more strong tornadoes are slightly above the perfect reliability line, indicating that these percentages have been slightly under-forecast over the dataset. However, it's important to note that overall these probabilities are well-calibrated and very reliable.

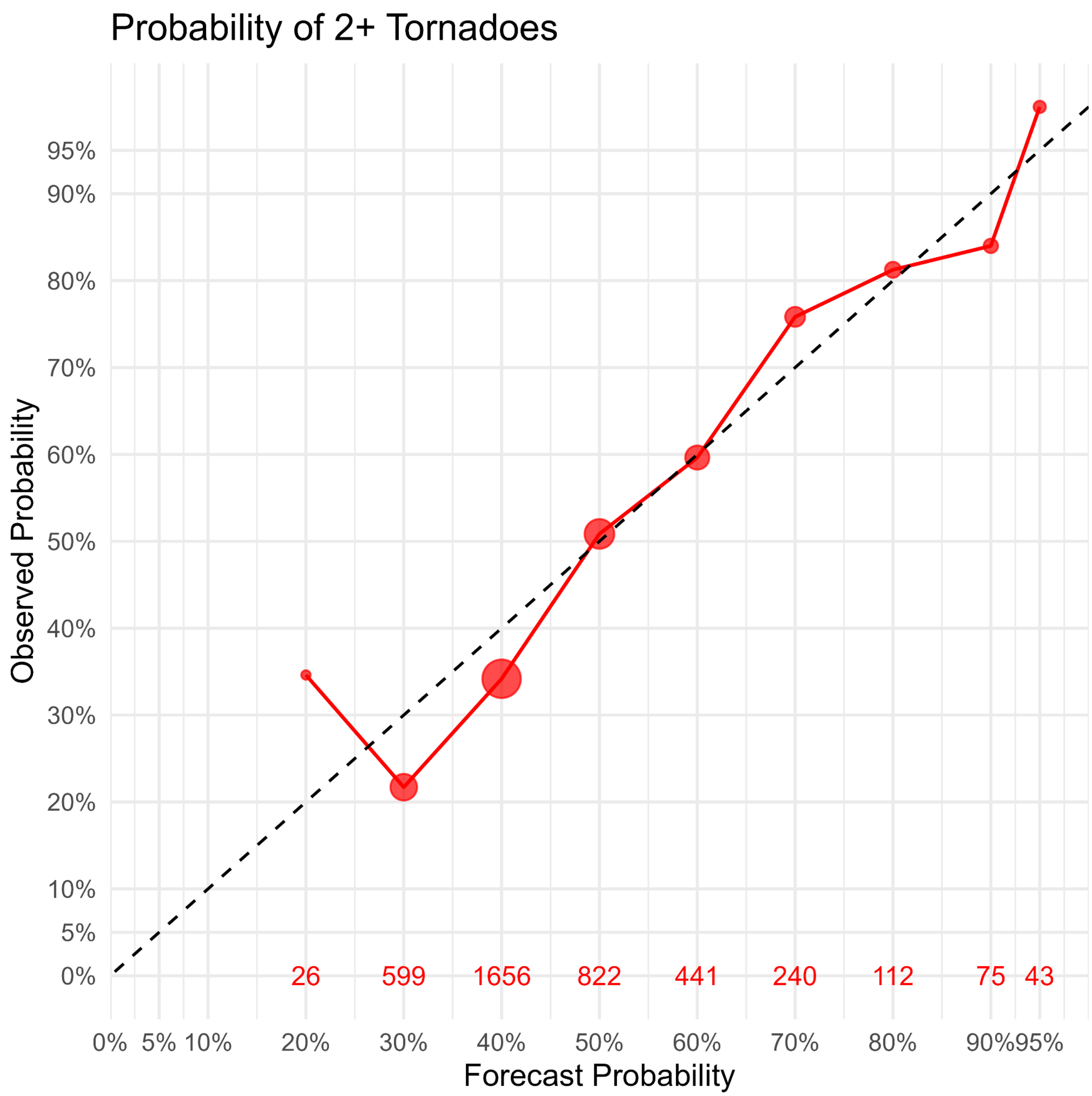


Figure 2.6: Reliability diagram of the probability of two or more tornadoes occurring within a tornado watch. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

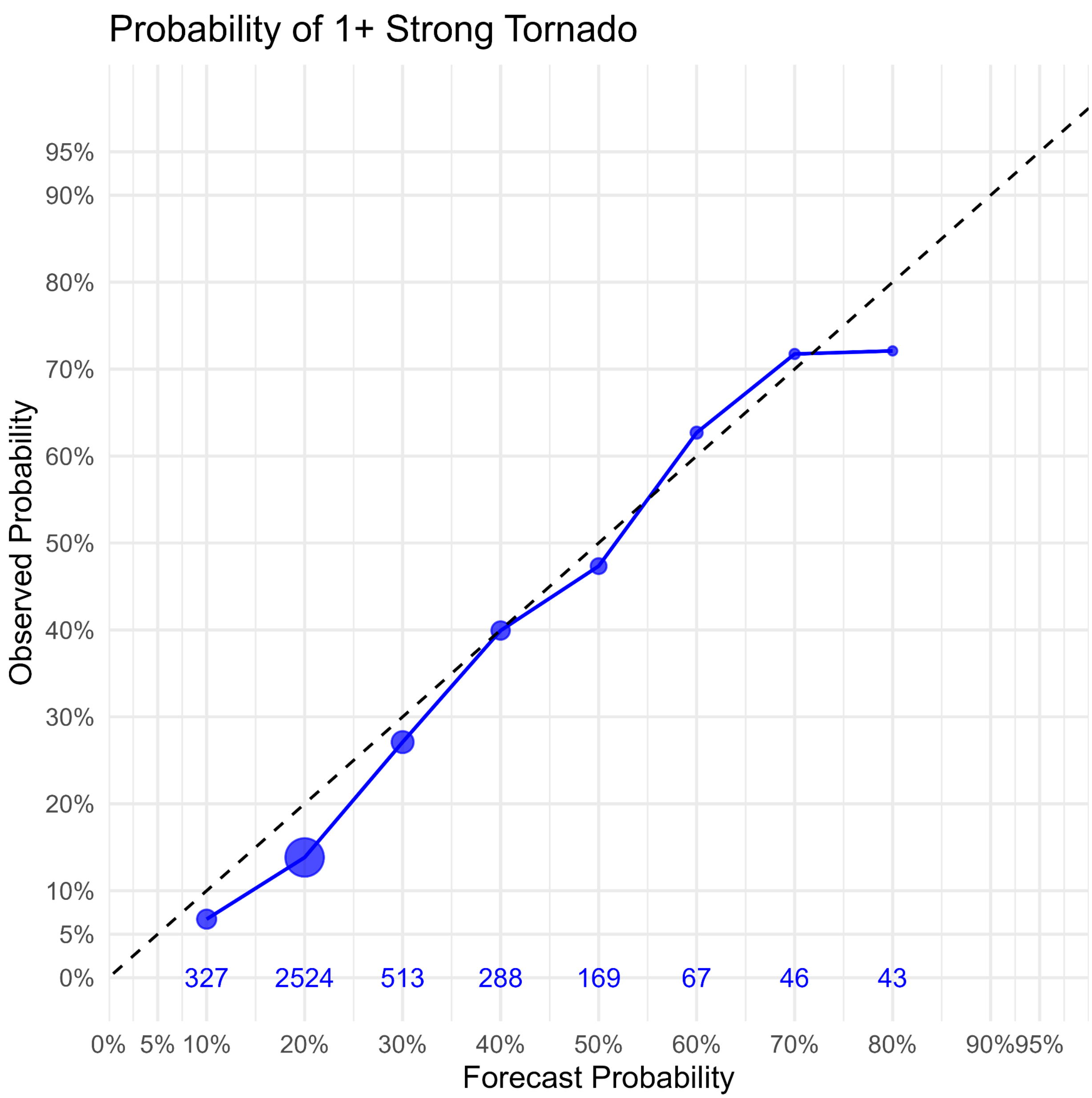



Figure 2.7: Reliability diagram of the probability of one or more strong (EF2+) tornadoes occurring within a tornado watch. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

Beyond the overall reliability diagrams, Tables A.1 and A.2 also show the Brier scores and Brier skill scores for both tornado probabilities. The Brier scores for the probability of two or more tornadoes decreases from 0.25 to 0.0025 from 50% to 95%, indicating increasing accuracy as the forecasted probability increases. Below 50%, the Brier scores are all below 0.25, with 30% having the lowest Brier score below 50%. The Brier skill score for the probability of two or more tornadoes is above zero for all forecasted probabilities with the exception of 20% and 50%. However, these negative values are very close to zero, indicating that while they are less accurate than a reference forecast that always predicts the climatological event frequency, they are not too far from being as accurate as the climatological event frequency. The remaining Brier skill scores are positive, and generally become higher as the forecasted probability value increases, indicating the forecast is more accurate than the climatological event frequency.

For the probability of one or more tornadoes, the Brier scores increase from the forecasted probabilities of 2% to 50%, before slowly decreasing from 60% to 80%, and then sharply decreasing for 90% and 95%, indicating that the accuracy of the forecasted probabilities decreases, and then increases. The Brier skill scores are very close to being perfectly accurate for 2% and 5%, before decreasing but still staying above zero for 10% and 20%, meaning all of these forecasted probabilities are more accurate than a forecast that always predicts the climatological event frequency. Every Brier skill score for 30% to 80% is below zero and 90% and 95% both have Brier skill scores well above zero, with 95% being very close to one.

## 2.3.2 Reliability by Year

After performing verification on these probabilities overall, we decided to stratify these probabilities by year to assess whether or not they became more or less reliable over time. Figures 2.8 and 2.9 demonstrate that tornado watch probabilities are typically very reliable across all years in the dataset, with moderate probabilities generally becoming more reliable over time for the probability of two or more tornadoes. However, while the probability of two more tornadoes was becoming more reliable and hit almost perfect reliability in 2012-2017, watches in 2018-2024 were under-forecast for moderate to high probabilities (50% to 70%), whereas in 2006-2011, probabilities from 30% to 60% were over-forecast, meaning that the forecasted probabilities are higher than the observed probabilities. This general trend can be seen in both Figure 2.8 and 2.9, but is slightly more prevalent in Figure 8 for the probability of two or more tornadoes. For the probability of one or more strong tornadoes (Figure 2.9), we see that reliability has much less variability over the years in comparison to Figure 2.8. Tables A.3 and A.4 illustrate the lack of higher probabilities for the probability of one or more strong tornadoes, with most probabilities being on the lower end, whereas the probability of two or more tornadoes is more clustered around the middle probabilities, with more probabilities on the higher end in comparison to the probability of one or more strong tornadoes.

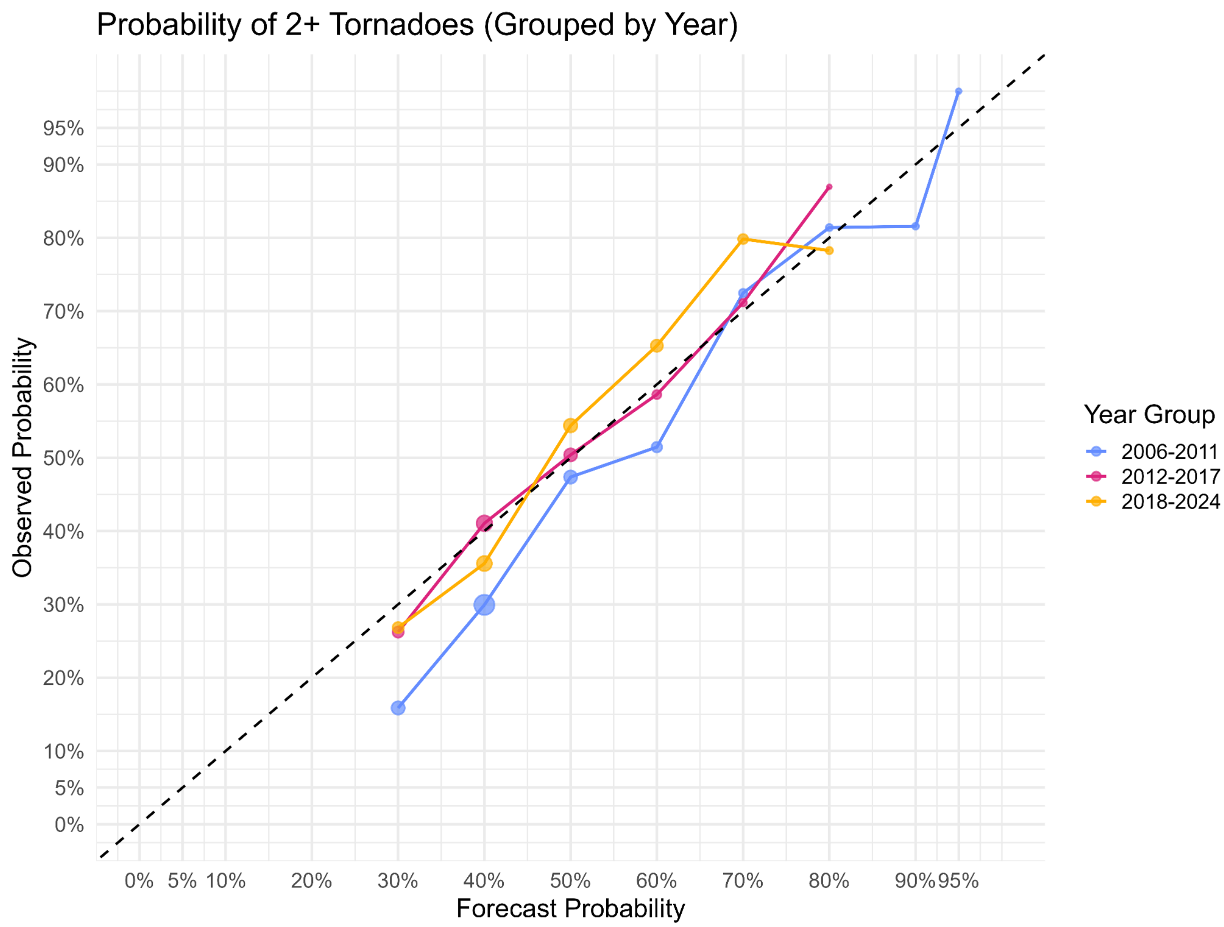



Figure 2.8: Reliability diagram of the probability of two or more tornadoes occurring within a tornado watch, stratified by years (2006-2011, 2012-2017, 2018-2024). Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

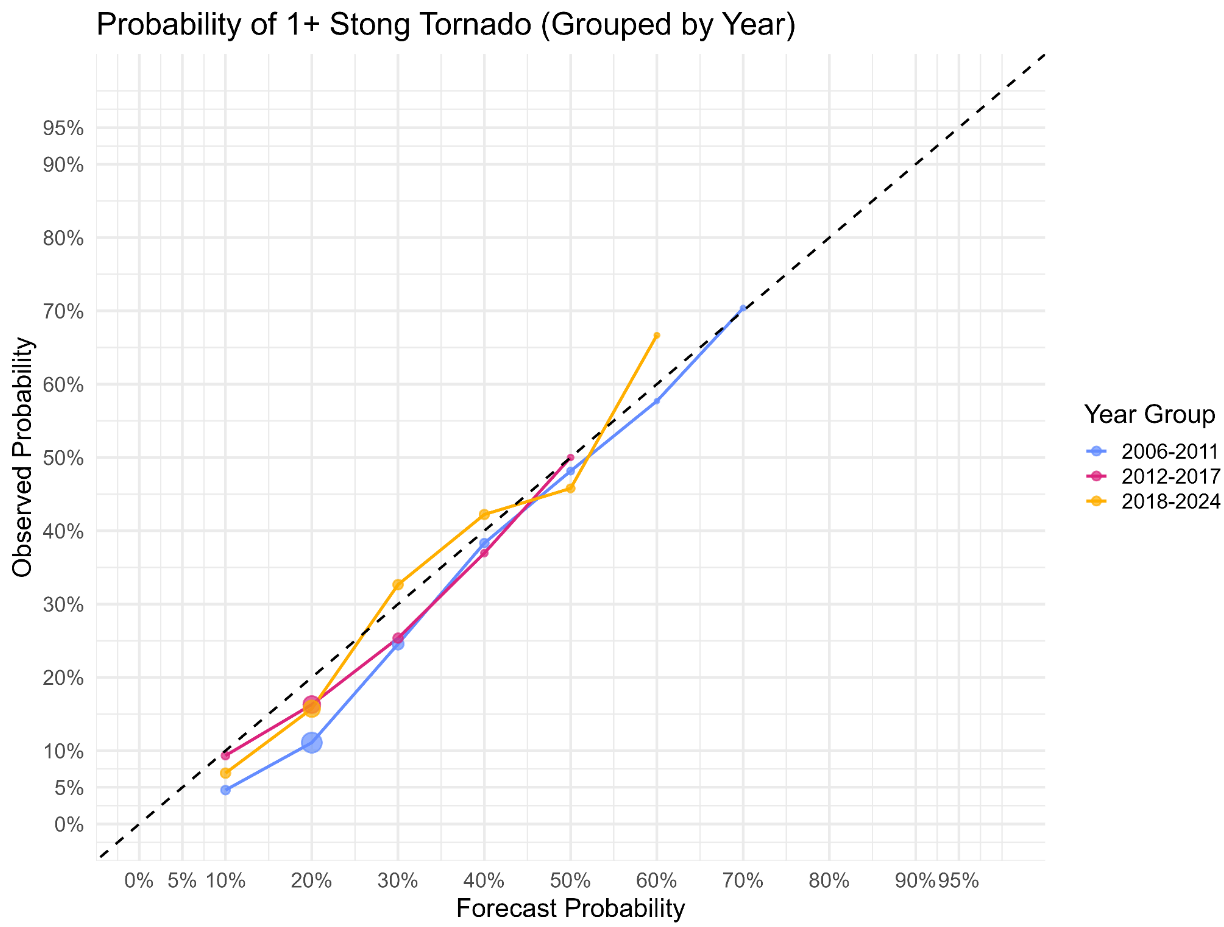



Figure 2.9: Reliability diagram of the probability of one or more strong tornadoes (EF2+) occurring within a tornado watch, stratified by years (2006-2011, 2012-2017, 2018-2024). Dot size corresponds to the number of watches issued for each probability. The dashed line represents perfect reliability.

### 2.3.3 Reliability by Season

Due to the difference in typical environments in different times of the year, we deemed stratifying the data by meteorological season as the next step in this analysis. We can see spring is the season with the most tornado watches, with winter and fall being the seasons with the least (Tables A.5 and A.6). We can also see that the probability of two or more tornadoes is again overall very reliable (Figure 2.10, Table A.5), but it does vary slightly by season. Generally, probabilities in fall tornado watches (September-October-November) are slightly over-forecasted, while winter (December-January-February) and summer (June-July-August) watch probabilities are both fairly close to being perfectly reliable. Spring (March-April-May) is slightly over-forecast for the lower percentages, and slightly under-forecast for higher percentages, meaning the forecasted probabilities are lower than the observed probabilities. Analyzing the probability of one or more strong tornadoes (Figure 2.11), fall and summer tornado watch probabilities only range from 10% to 40%, as there are few or no watches with higher probabilities of one or more strong tornadoes. Probabilities for summer watches are over-forecast, while fall watches are over-forecast for the lowest probabilities (10-20%), then slightly under-forecast for higher probabilities (30-40%). Winter and spring tornado watches show near perfect reliability, with the exception of lower probabilities being slightly over-forecast for spring tornado watches.

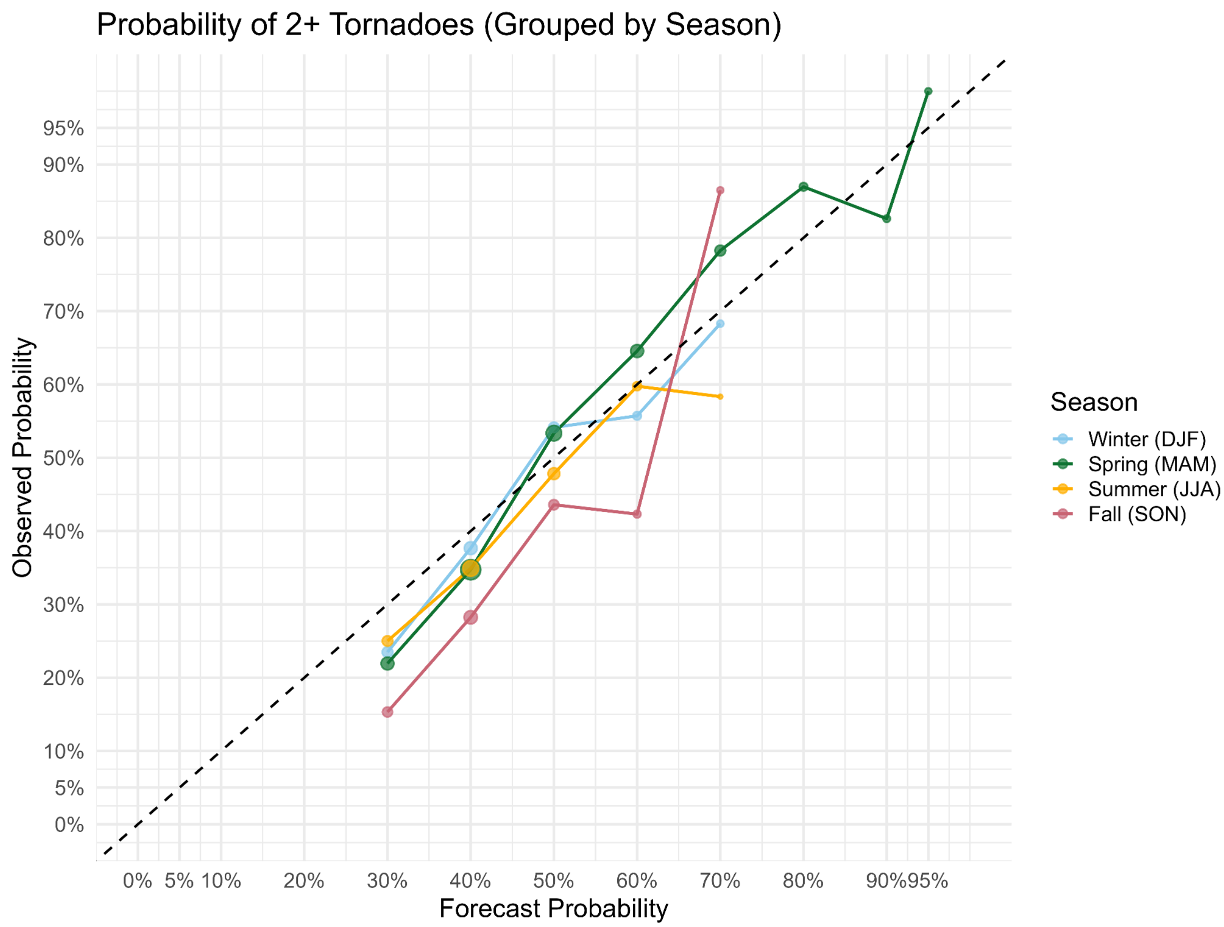



Figure 2.10: Reliability diagram of the probability of two or more tornadoes occurring within a tornado watch, stratified by season. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

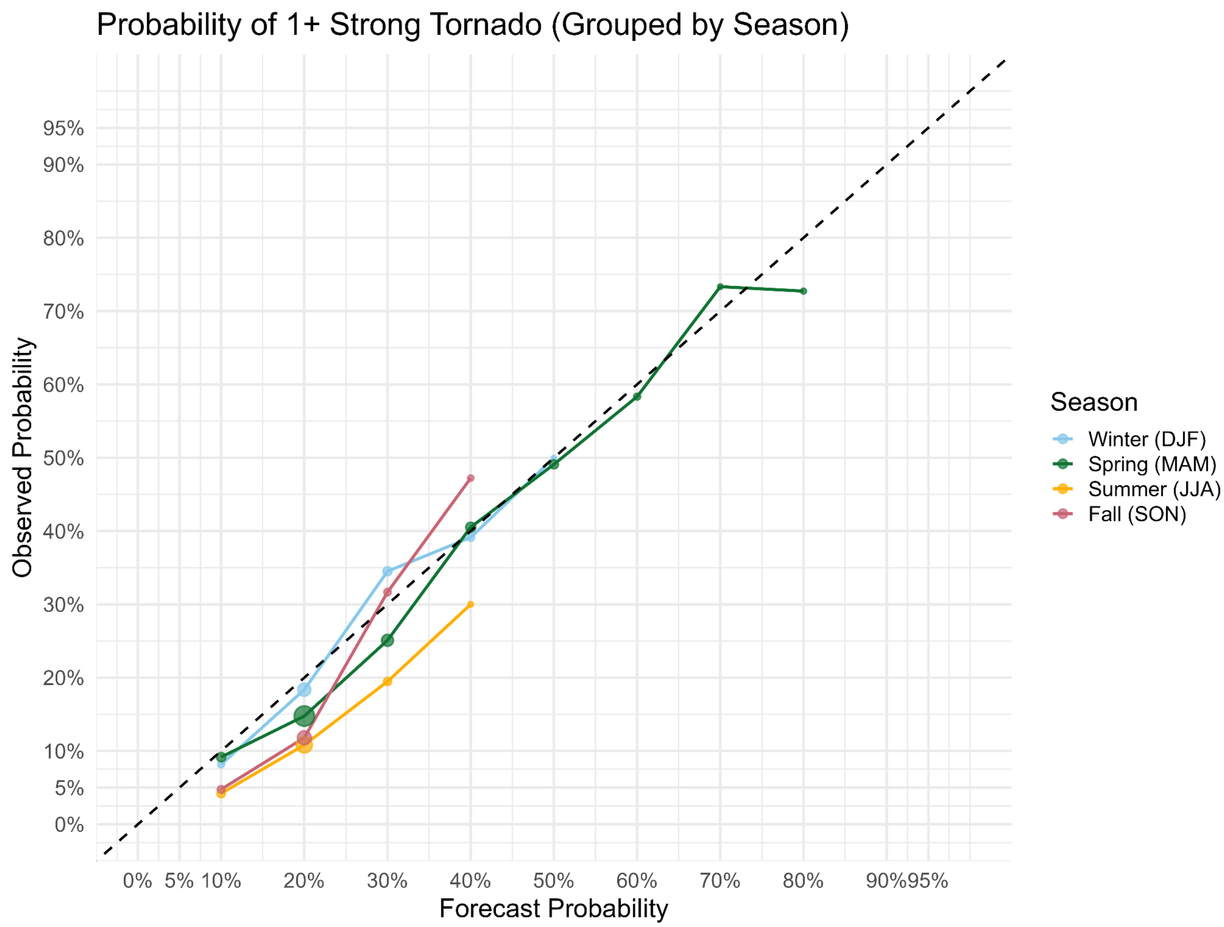



Figure 2.11: Reliability diagram of the probability of one or more strong tornadoes (EF2+) occurring within a tornado watch, stratified by season. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

## 2.3.4 Reliability by Day and Night

While more tornadoes occur during the day compared to at night, the majority of fatalities from tornadoes occur between sunset and sunrise. Due to this disproportionate amount of fatalities from the inherent vulnerabilities that come with nighttime tornadoes (Ashley et al. 2008), the data was then stratified by local day and nighttime. When stratifying the tornado probabilities associated with tornado watches by their local day and night times, we can see that overall more tornado watches were issued during the day compared to nighttime for both probabilities, as seen in Tables A.7 and A.8. For the probability of two or more tornadoes in Figure 2.12, nighttime watch probabilities are generally over-forecast across 30% to 50%, meaning the forecasted probabilities are higher than the observed probabilities for two or more tornadoes, especially when compared to the daytime watches. For the probability of one or more strong tornadoes in Figure 2.13, we see that day and nighttime watches were both slightly over-forecast for the lower percentages of 10% to 30%, but nighttime watches then flip to being under-forecast for 40% and 50%.

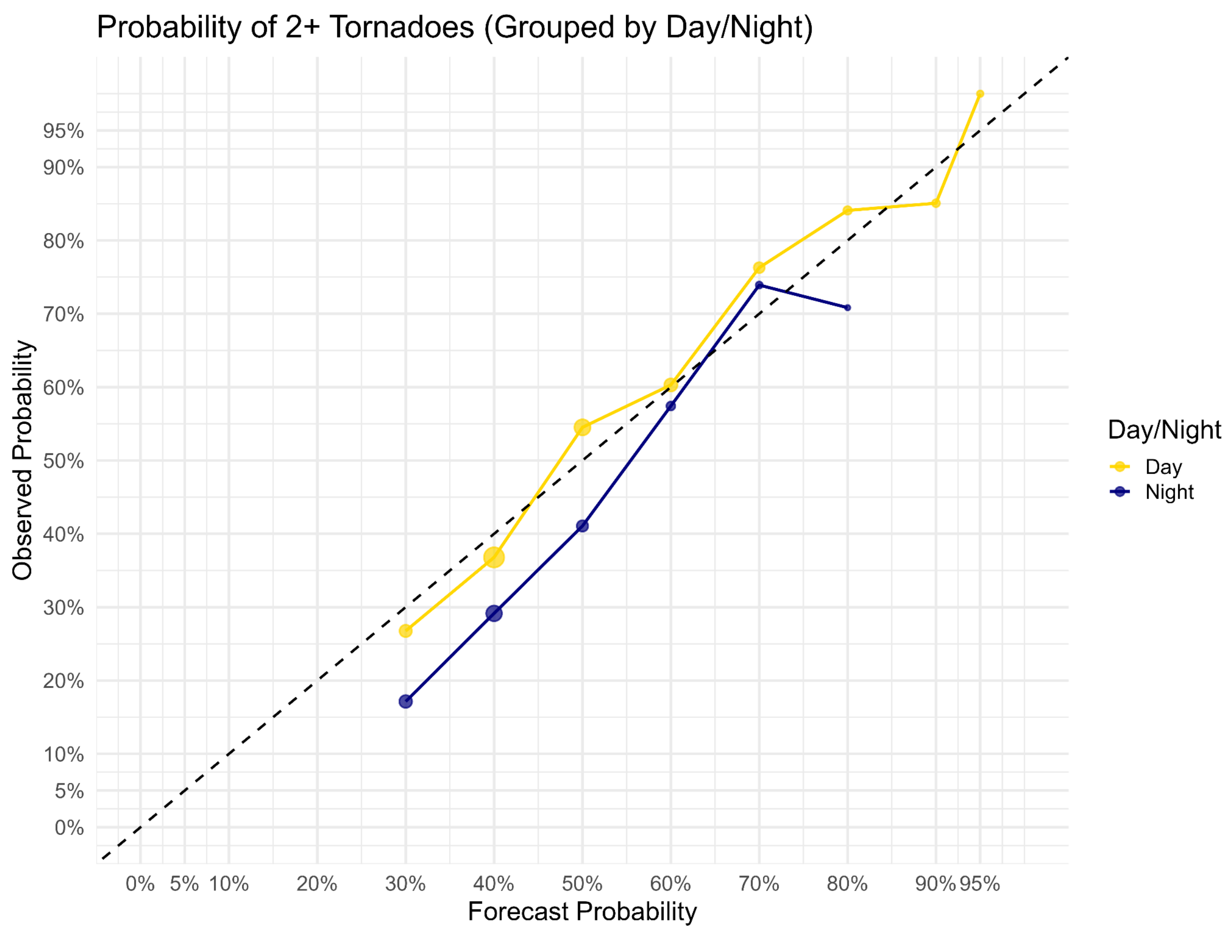



Figure 2.12: Reliability diagram of the probability of two or more tornadoes occurring within a tornado watch, stratified by time of day. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

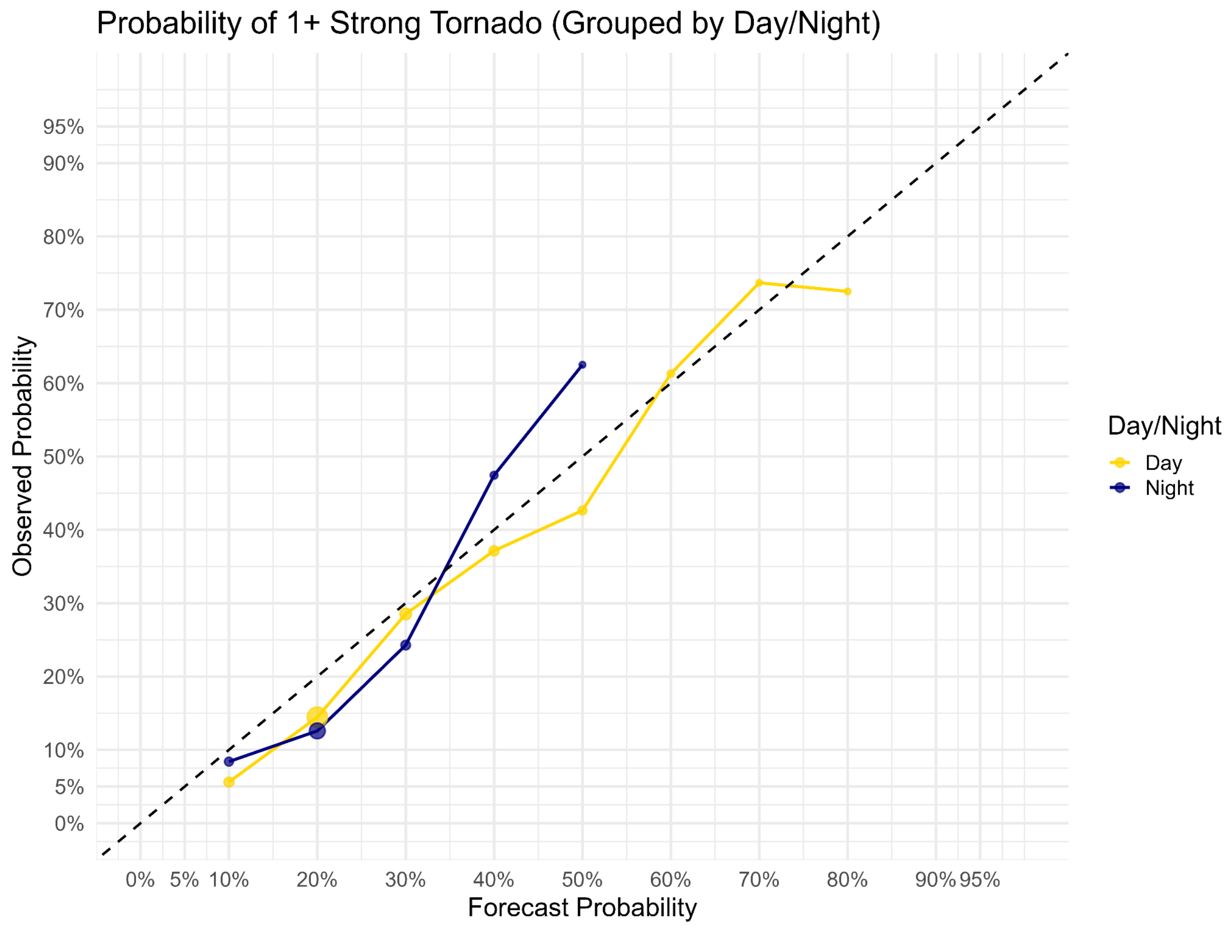



Figure 2.13: Reliability diagram of the probability of one or more strong tornadoes (EF2+) occurring within a tornado watch, stratified by time of day. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

## 2.3.5 Reliability by Geographic Region

Dividing the tornado watches into geographic regions, we can see from Tables A.9 and A.10 that the West and Northeast regions have a very small number of tornado watches over the 19 year period. This is also represented in Figures 2.14 and 2.15, where the West and Northeast only have enough watches to calculate the reliability for 40% and 20% respectively.

Additionally, Tables A.9 and A.10 also illustrate that the South had the most tornado watches over the 19 year period, followed by the Great Plains and the Midwest. For the probability of two or more tornadoes in Figure 2.14, we see that the Great Plains is the closest to perfect reliability for the lower percentages, and then becomes slightly under-forecast for higher percentages. The Midwest region generally follows the same trend, but with lower percentages being over-forecast. The South has the greatest difference from perfect reliability for the lower percentages, with most probabilities over-forecasted. The probability of one or more strong tornadoes in Figure 2.15 is less variable than the probability of two or more tornadoes, as the Midwest, South, and Great Plains are grouped closer together throughout most of the percentages. All regions have more watches on the lower and middle end of percentages, and do not have many high percentages for the probability of one or more strong tornadoes. All three of these regions are over-forecast for low probabilities (10% to 30%). However, it's still important to note that these differences from perfect reliability are small, and in general tornado watch probabilities across these geographic regions are very reliable.

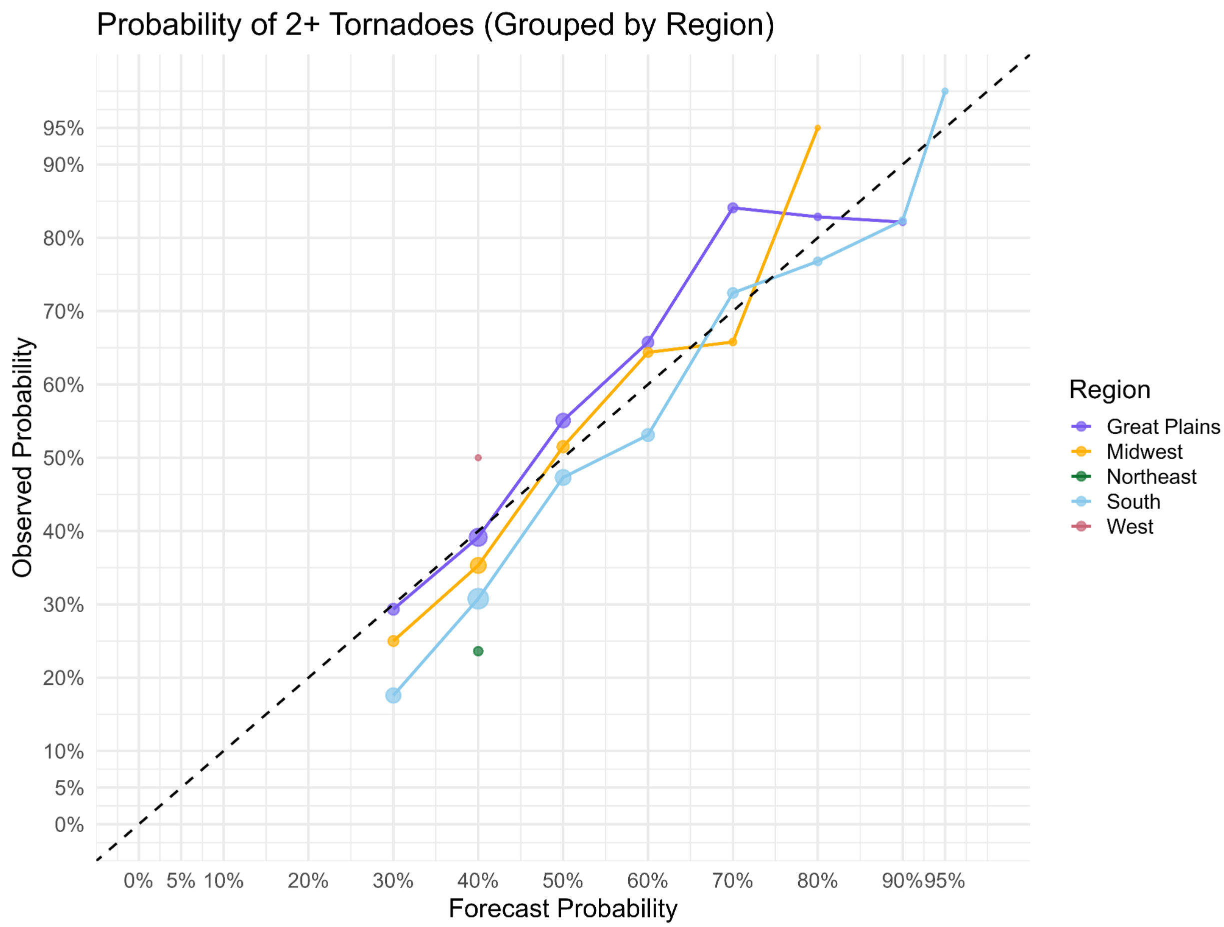



Figure 2.14: Reliability diagram of the probability of two or more tornadoes occurring within a tornado watch, stratified by geographic region. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

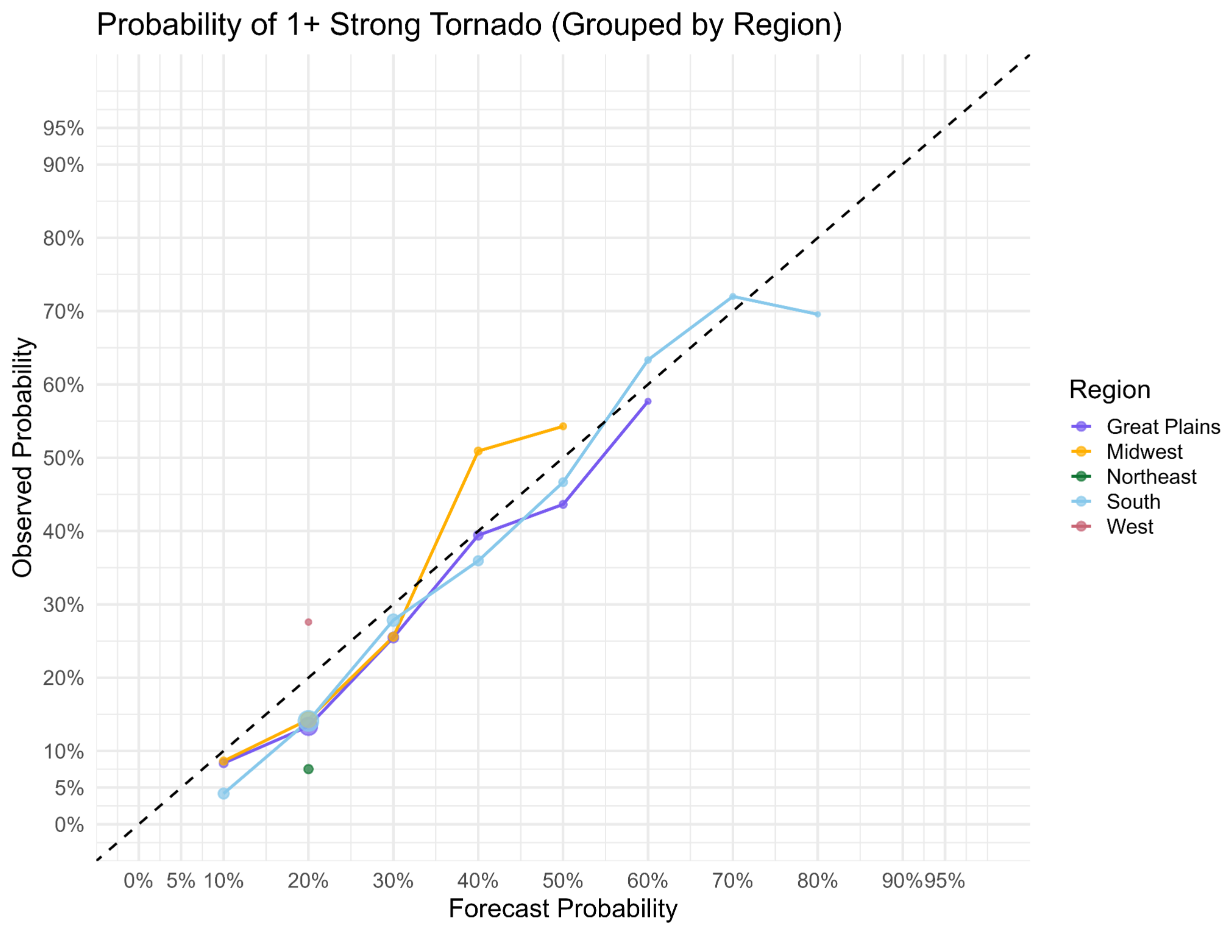



Figure 2.15: Reliability diagram of the probability of one or more strong tornadoes (EF2+) occurring within a tornado watch, stratified by geographic region. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

# 2.4 Tornado Watch Case Studies Methods

While on average the tornado probabilities associated with tornado watches generally perform as they should, there were a small number of tornado watches that greatly over or under performed based on the probabilities for the watch. When looking at individual watches, the probabilities for each watch can only be verified in a binary manner. For example, regardless of the percentage given to the probability of two or more tornadoes occurring in a watch, the watch will either have two or more tornadoes occur in it, or it will not. Thus, in order to gain a better understanding of which watches over or under performed, the risk categories associated with each probability were used (low, <2% to 20% for the probability of one or more strong tornadoes, 5% to 20% for the probability of two or more tornadoes, moderate, >=30% and<=60%, and high, >=70% and <= 95%). If both tornado probabilities in a watch were categorized as 'low' and both probabilities verified, meaning two or more tornadoes and one or more strong tornadoes occurred in the tornado watch, the watch was flagged to be analyzed. This process was also done for watches where both probabilities were categorized as 'high', and both probabilities did not verify, meaning less than two tornadoes and no strong tornadoes occurred in the watch.

From this analysis, ten watches were found with both probabilities in the 'high' category that did not verify, of which three watches were chosen for further analysis. These three watches were chosen for further analysis because of the variation in different scenarios that unfolded. For instance, tornado watch #22 of 2017 was chosen due to the slight uncertainty in the watch location, tornado watch #195 of 2009 was chosen due to the seemingly favorable conditions in the watch area for tornadoes, though tornadoes did not occur, and tornado watch #559 of 2021 was chosen due to the potential influence of ongoing tornadoes on the watch's probabilities. One watch (tornado watch #554 of 2013) was found where both probabilities were in the 'low' category and did verify. Thus, these four tornado watches were analyzed to try to understand why the watch did or did not verify based on the environment and circumstances leading up to and the day the watch was issued. To complete these analyses, the Storm Prediction Center's Severe Weather Events Archive (Storm Prediction Center 2026) was used to gather information on convective outlooks, mesoscale discussions (MDs), severe thunderstorm and tornado watches, and tornado reports.

## 2.4.1 Tornado Watch #22 01/22/2017

Tornado watch #22 of 2017 was issued at 21:25 UTC (4:25pm EST) January 22nd and expired at 00:55 UTC January 23rd (7:55pm January 22nd) (Figure 2.16). Prior to this watch being issued, the day of January 22nd, 2017 was slated to be a severe weather day for much of the Gulf Coast. The event was first forecasted by SPC on January 16th, as the day 6 convective outlook showed a 15% over the southern Gulf Coast from Louisiana to Florida (Figure 2.17). At the time, an upper level shortwave trough south of Alaska was predicted to move across the continental US and develop into a deep surface level low over the next week. The morning of January 17th, the already mature low pressure system was over the Oklahoma/Arkansas border, and in just six hours progressed to be positioned over Alabama, ultimately deepening from 1000mb to 989mb. During this time, the SPC elevated the prior issued moderate risk from the 0600 UTC (1am EST) convective outlook (Figure 2.18) to a high risk in the 1300 UTC (8am EST) convective outlook update (Figure 2.19). The majority of this high risk was later placed under tornado watch #21 as a Particularly Dangerous Situation (PDS) watch, which was valid from 12:45pm to 8:00pm EST and all associated probabilities were listed at 90% (Figure 2.20).

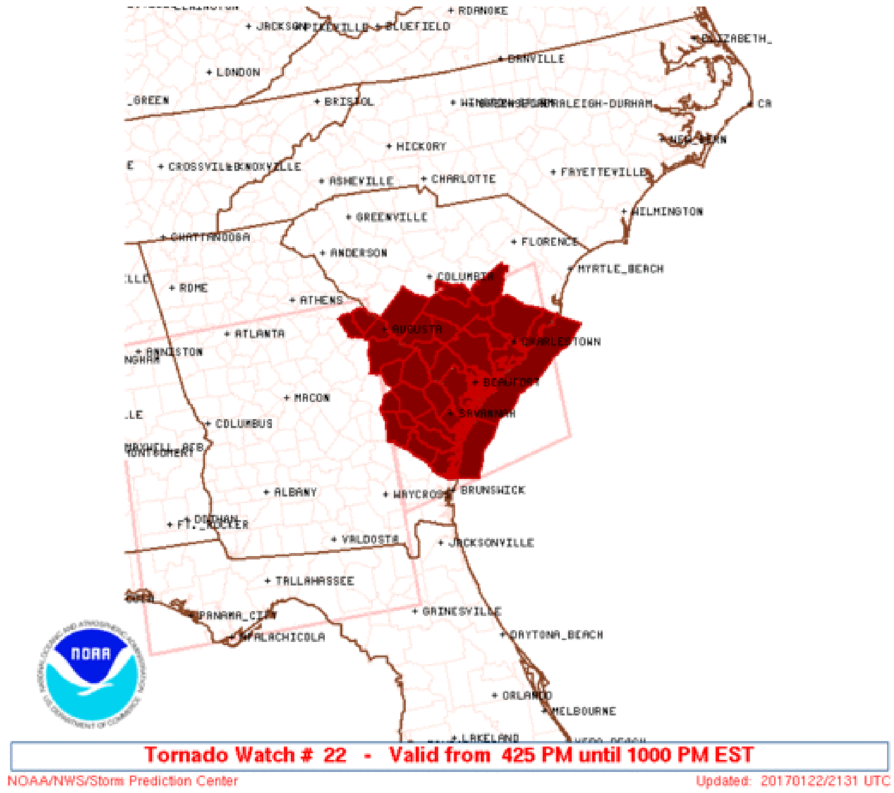



Figure 2.16: Tornado watch #22 of 2017, issued on 21:25 UTC January 22nd and expired at 00:55 UTC January 23rd. Red counties indicate the official watch area, with the eastern most pink box indicating the watch polygon. Note: The expiration time shown on the Figure (10:00 PM EST) is the initial expiration time. The watch was fully canceled at 7:55pm EST.


Figure 2.17: Day 6 of the SPC convective outlook, issued on 0957 UTC January 17th, 2017. A 15% risk is located over the gulf coast from Louisiana to the panhandle of Florida.

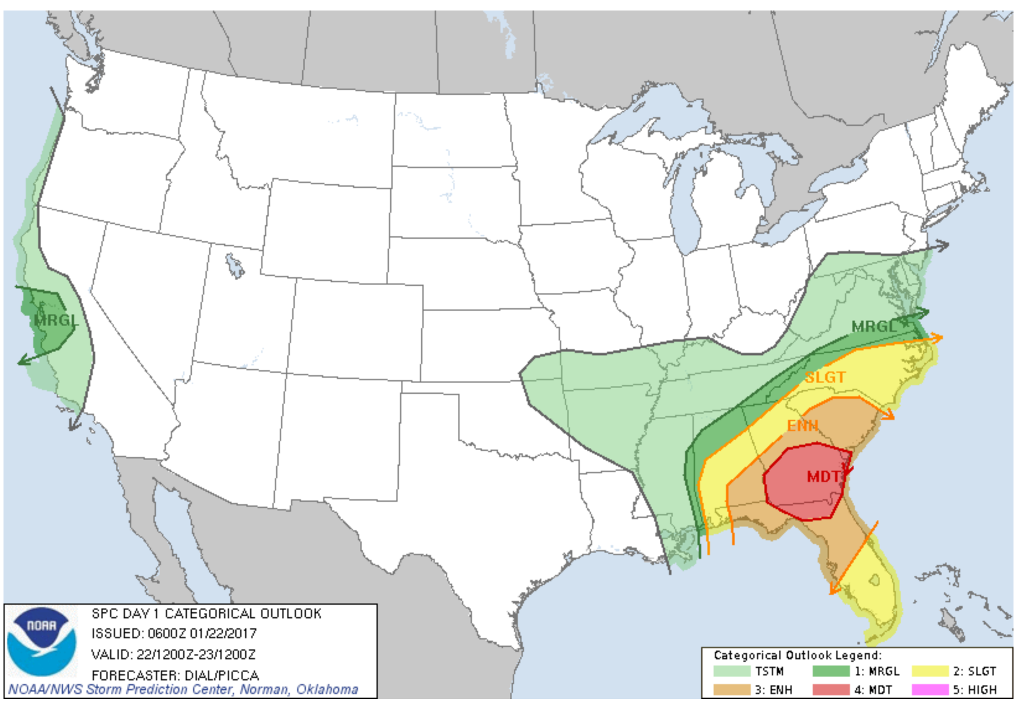



Figure 2.18: Day 1 of the SPC convective outlook, issued at 06:00 UTC on January 22nd, 2017. A moderate risk is currently located over southern Georgia and northern Florida.

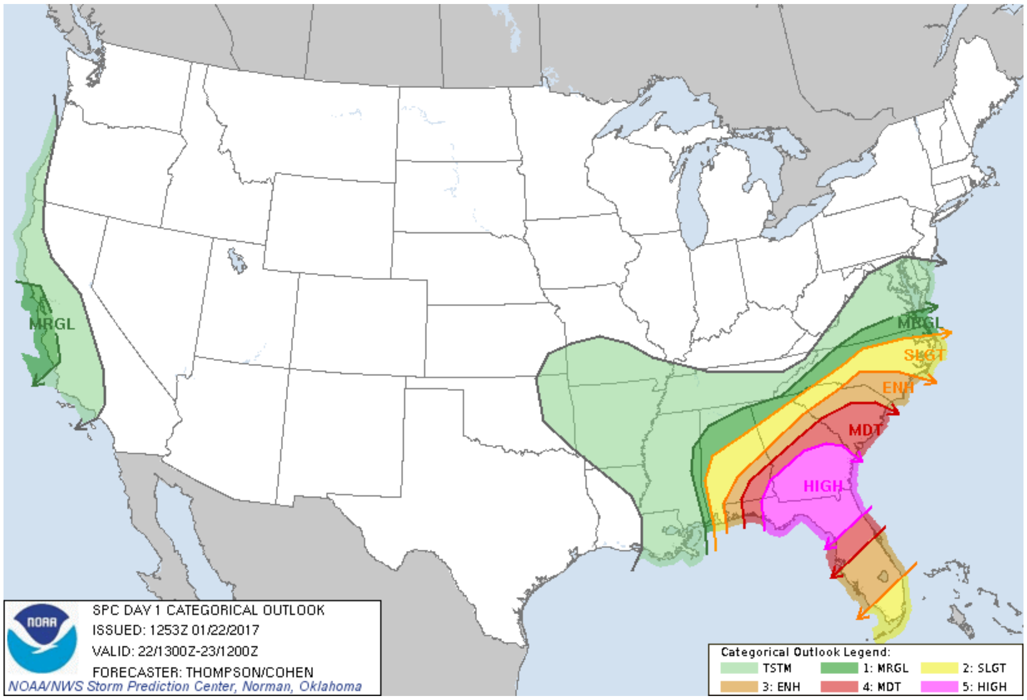



Figure 2.19: Day 1 of the SPC convective outlook, issued at 12:35 UTC on January 22nd, 2017. The moderate risk in southern Georgia and northern Florida has been upgraded to a high risk.

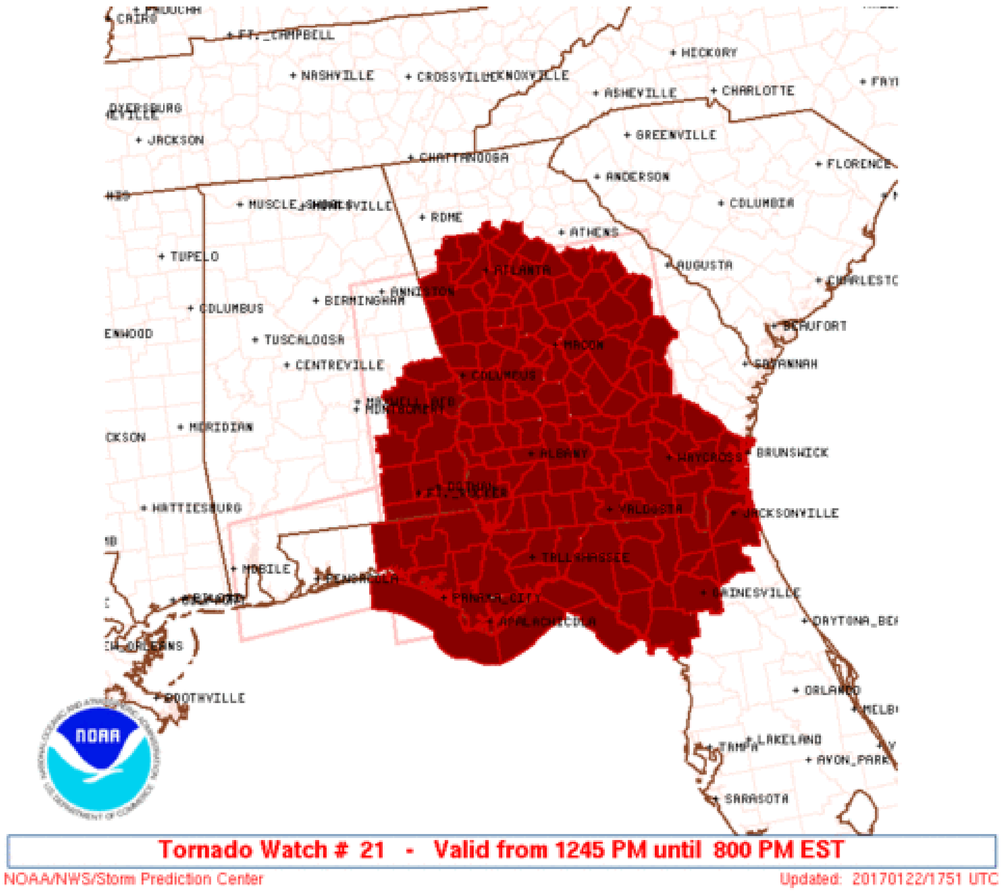
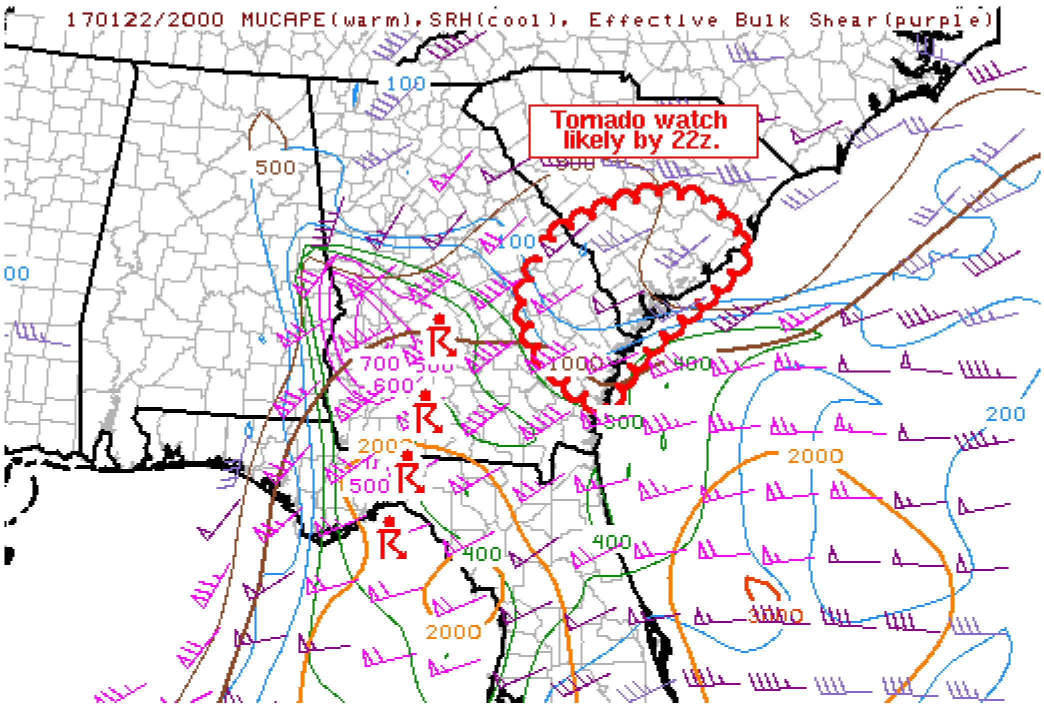



Figure 2.20: Tornado watch #21, valid from 12:45PM EST to 8:00PM EST on January 22nd, 2017. Red counties indicate the official watch area, with the eastern most pink box indicating the watch polygon.

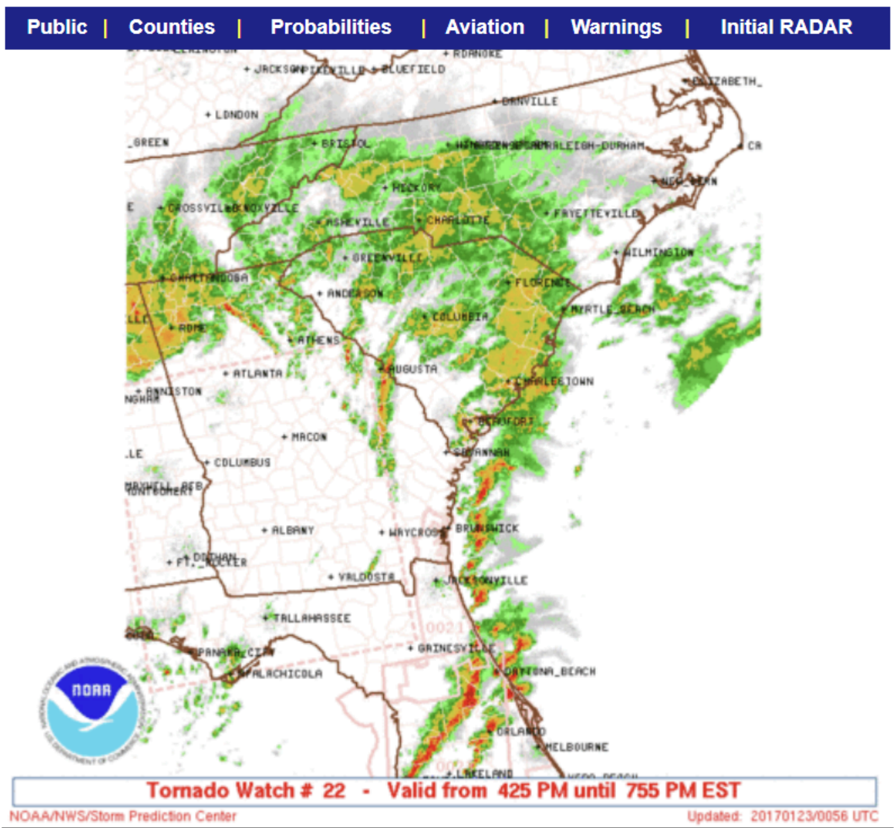
While confidence in severe weather developing in southern Georgia and northern Florida was high, this was not the case for northern Georgia and the Carolinas. The 0600 UTC Convective outlook on January 22nd (Figure 2.18) noted that the early morning convection may have impacted the destabilization potential, but was still under a severe weather threat due to the expected northward movement of the warm front. In addition, MD #119 (Figure 2.21), which was issued at 4:02pm EST, stated that there was still uncertainty as to how far into northern Georgia and South Carolina the already destructive storms would track due to limited heating and ample rainfall. But spottiness in clouds had allowed dewpoints and temperatures to rebound, and forecasters expected the warm front to move through the region and the low level shear to improve later in the evening. The discussion stated that a tornado watch would likely be issued soon for the area.




SPC MCD #0119

Figure 2.21: Mesoscale discussion #119 of 2017, issued at 4:02PM EST on January 22nd, 2017.

Tornado watch #22 of 2017 (Figure 2.16) was issued 23 minutes after MD #119 (Figure 2.21), at 4:25pm on January 22nd as a PDS watch. This was after multiple tornadoes already occurred in south Georgia and north Florida in tornado watch #21 (Figure 2.20). Despite the uncertainty around severe development in the watch region, both tornado probabilities were set at 80%, which falls into the ‘high’ category (Figure 2.22). However, the hail and wind probabilities all decreased by a significant amount in comparison to tornado watch #21, with only the probability of 10 or more severe wind events staying within the ‘high’ category. This may have signified that forecasters expected storms to weaken as they moved to the east. As the system began to move through the watch area, the only severe thunderstorm and tornado warnings that occurred were in the southwest area of the watch, which carried over from tornado watch #21. MD #123 (Figure 2.23) was issued at 2345 UTC (6:45pm EST), and explained that no severe events had occurred thus far due to the warm front not yet moving into the cooler air mass in South Carolina. The warm front was attempting to advance northward, however there was only a brief period of time before the convective system moved offshore.




| Hazard | Tornadoes | EF2+ Tornadoes | Severe Wind | 65 kt+ Wind | Severe Hail | 2"+ Hail |
|---|---|---|---|---|---|---|
| Likelihood | High | High | High | Moderate | Moderate | Very Low |

**Note:** Click for Complete Product Text.

**Tornadoes**

| | |
|---|---|
| Probability of 2 or more tornadoes | High (80%) |
| Probability of 1 or more strong (F2-F5) tornadoes | High (80%) |

Figure 2.22: Tornado watch #22 of 2017, issued at 4:25PM EST, and canceled at 7:55PM EST January 22nd, 2017. Tornado probabilities associated with this watch are issued at 80% (high) for both probabilities.

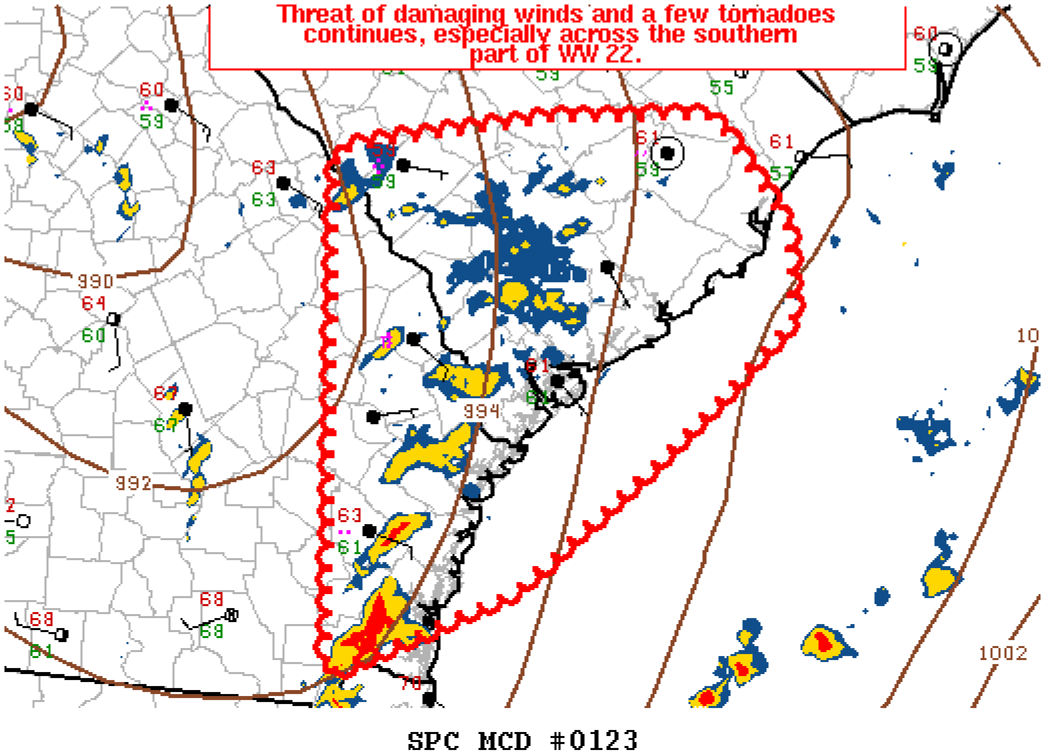



Figure 2.23: Mesoscale discussion #123 of 2017, issued at 6:45PM EST, January 22nd, 2017.

Ultimately, the warm front did not reach the convective system before it moved offshore, and no tornadoes (or any other severe reports) occurred inside the watch during this event, as seen in Figure 2.23 and 2.24. The watch was initially set to expire at 10pm EST, but was ultimately canceled at 7:55pm EST due to the lack of tornado risk. Compared to the high tornado probabilities, the watch did not ‘verify’, but the uncertainty of the watch was well discussed in various products throughout the day. The high probabilities were also in line with the probabilities in the PDS watch issued to the west, such that risk messaging remained consistent between both watches. The forecasters likely issued this watch with high probabilities to ensure that if the warm front *did* happen to reach the convective system before it moved offshore, local officials would have had the proper measures in place prior to a strong tornado event occurring.

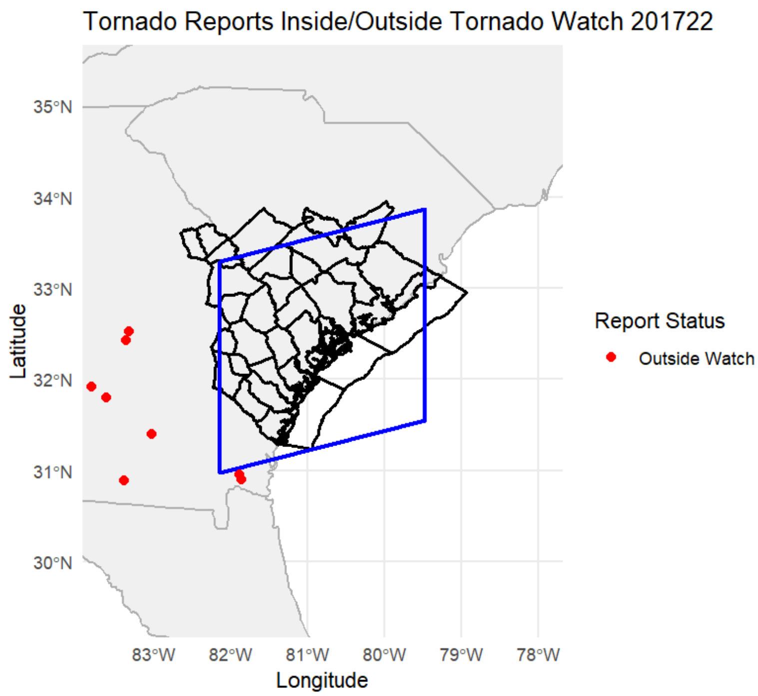



Figure 2.23: Confirmed tornado reports inside/outside tornado watch #22 of 2017. The black county outlines indicate the official watch area, with the blue box representing the tornado watch polygon.

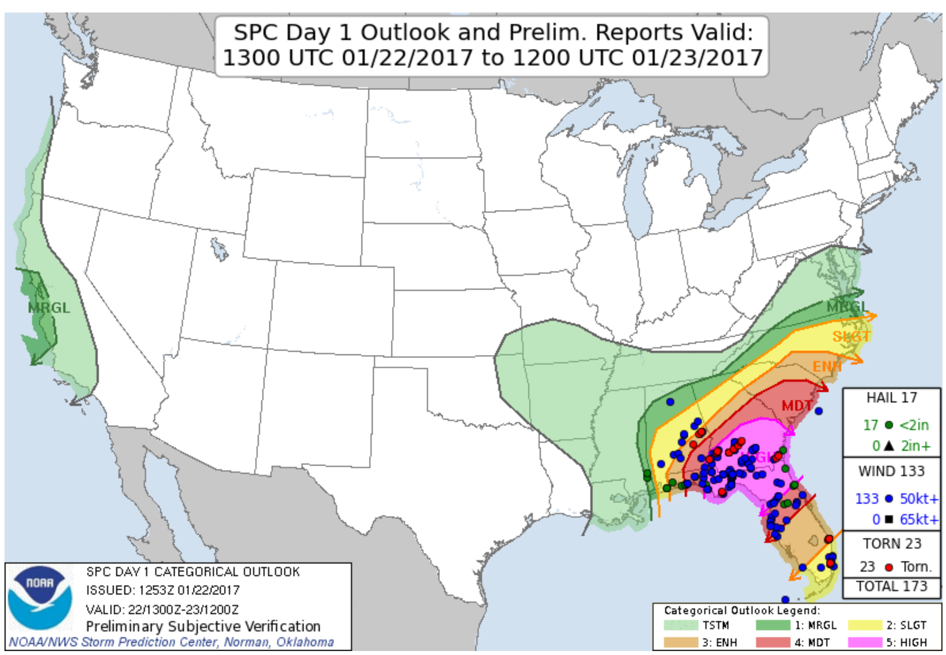


Figure 2.24: Preliminary storm reports during 13:00 UTC January 22nd to 12:00 UTC January 23rd, 2017, overlaid onto the SPC's day 1 convective outlook.

## 2.4.2 Tornado Watch #195 04/26/2009

Tornado watch #195 of 2009 was issued at 23:15 UTC (6:15 PM CDT) on April 26th, and was canceled at 07:13 UTC (2:13 AM CDT) on April 27th. Prior to the watch being issued, the SPC's day 5 convective outlook included a 30% risk for April 25th-27th over the watch area, as seen in Figure 2.25. At the time, the emergence of one significant short wave from an amplified trough located over the western US was expected to occur in the following four to five days. This shortwave was expected to bring a significant frontal wave across the regions shown in Figure 2.25. This in combination with the ideal conditions existing over this region from gulf flow, severe storms, including discrete supercells, were predicted to occur. However, it is important to note that from here until the day before the event, forecasters note model disagreement as to where the most severe storms would occur due to differences in location of the shortwave.

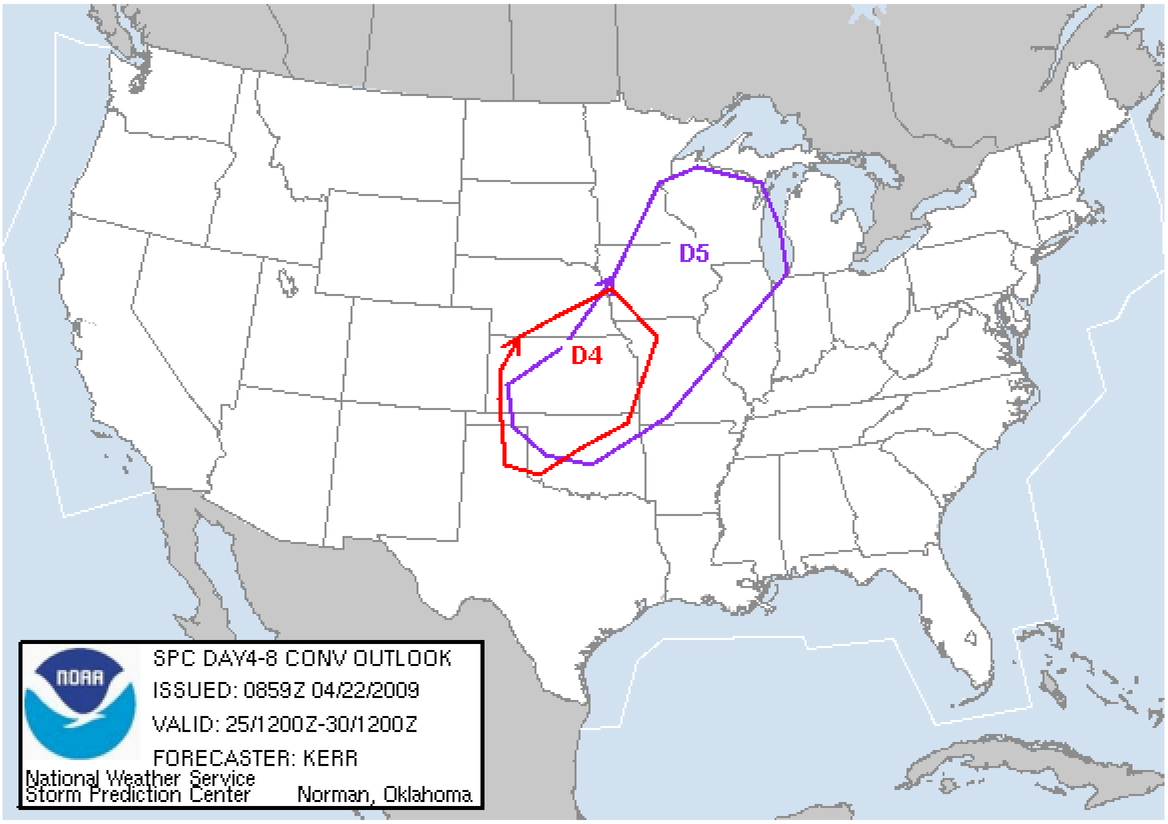



Figure 2.25: SPC's day 4-8 convective outlook issued on day 5. The purple and red outline indicates the area of 30% or higher probability for severe thunderstorms within 25 miles of a point for day 5 and 4 respectively.

On the morning of April 26th, the shortwave trough and an associated 70 knot mid level jet lifted from the Central Rockies into the Northern Plains, in addition to a shortwave perturbation that moved from New Mexico into west Texas. At the surface, a dryline extended from west Texas to southeast New Mexico, a surface low was deepening in the Oklahoma panhandle, while a warm front was draped across the US from Michigan to Kansas. Just before 6 AM CDT, severe storms began to fire off along the warm front in central Kansas. The SPC's 1300 UTC day 1 convective outlook expected the southern impulse to trigger storms along the dryline in the afternoon in western Oklahoma and the Texas Panhandle. As severe storms began to pop up in western Oklahoma, the 1630 UTC day 1 convective outlook upgraded the event from a moderate risk to a high risk over western Oklahoma and south central Kansas, as seen in Figure 2.26. At 11:30 AM CDT, tornado watch #190 is issued as a PDS watch, with the probability of two or more tornadoes placed at >95% (high), and the probability of one or more strong tornadoes placed at 60% (moderate) (Figure 2.27).

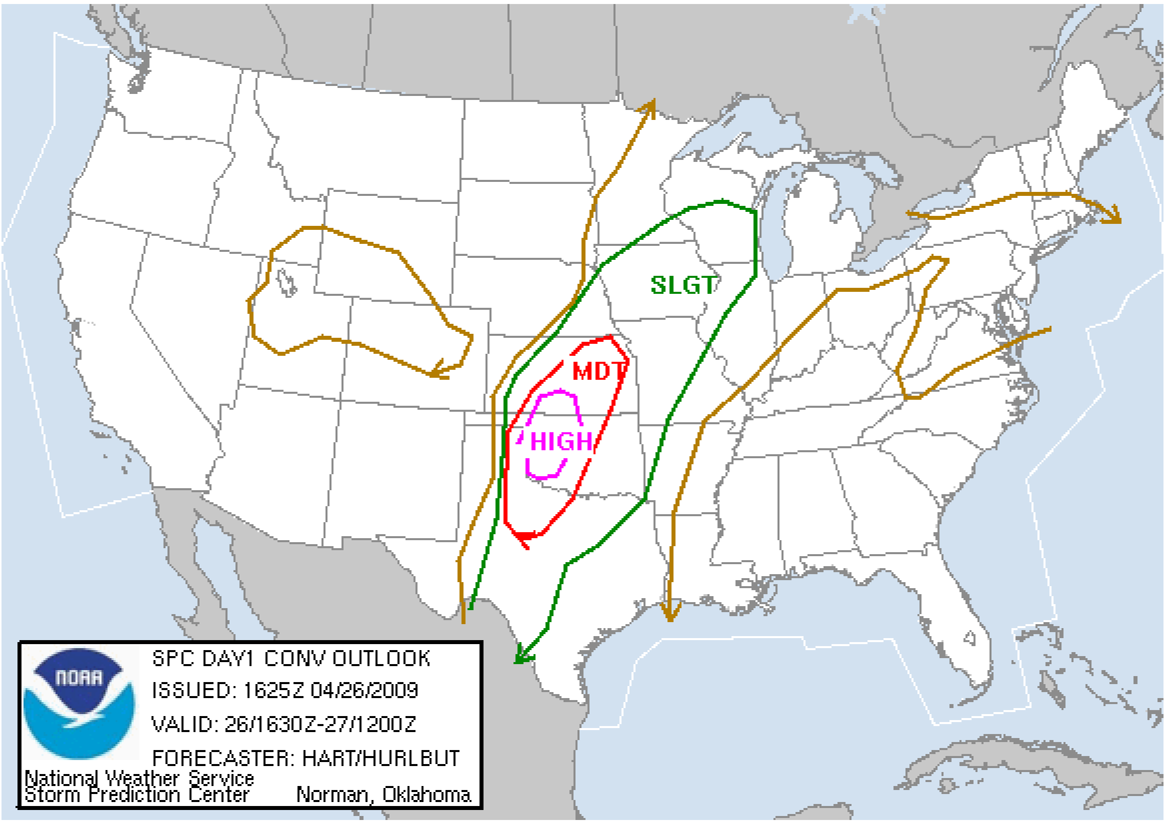



Figure 2.26: 16:30 UTC Day 1 Convective Outlook, valid April 26th 2009 16:30 UTC to April 27th 2009 12:00 UTC. A high risk is located over western Oklahoma and south central Kansas.

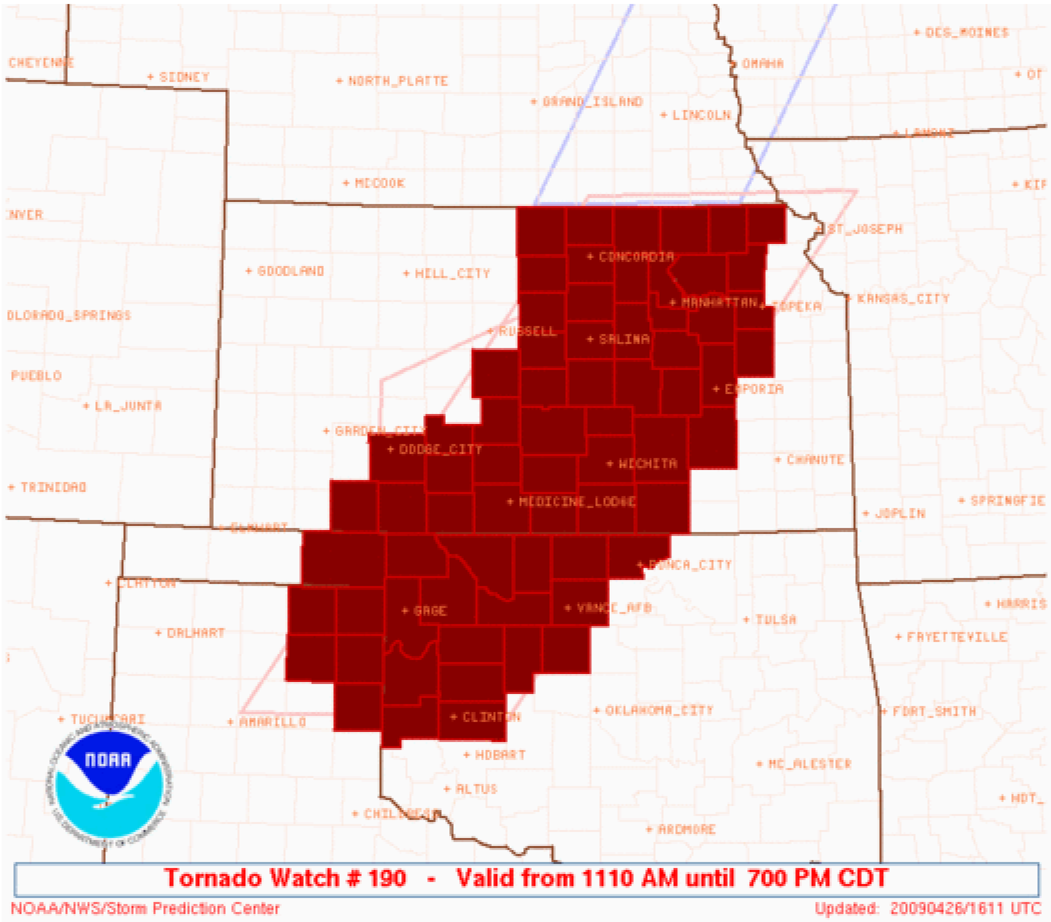



Figure 2.27: Tornado watch #190, issued 11:10 AM and canceled at 6:15 PM. Red counties indicate the official watch area, with the pink box indicating the watch polygon.

As the dryline continued to move east, storms began to pop up south of tornado watch #190. MD #0582 noted that at first these storms will be mainly hail driven, but with increased heating and more storms developing along the dryline, all aspects of severe weather will be likely. Tornado watch #192 was issued as a PDS watch roughly an hour after MD #0582, once tornado potential had increased (Figure 2.28). The associated tornado probabilities were 90% (high) for the probability of two or more tornadoes, and 60% (moderate) for the probability of one or more strong tornadoes. In MD #0585 issued at 3:34 PM CDT, the increase in boundary layer moisture and continued surface heating have caused MLCAPE values to increase to 3000

J/Kg ahead of storms. With these environmental conditions, storms were expected to continue intensifying throughout the evening. As storms continued to strengthen and grow, MD #0587 issued at 5:44 PM CDT (Figure 2.29) concluded that since the tornado and hail threat would continue past the expiration times of both tornado watches #190 and #192, these watches would be canceled and a new watch would be issued encompassing both watch areas. It should also be noted that forecasters mentioned a damaging tornado had already been observed in tornado watch #190, and they did not expect conditions to degrade in the remainder of the evening.

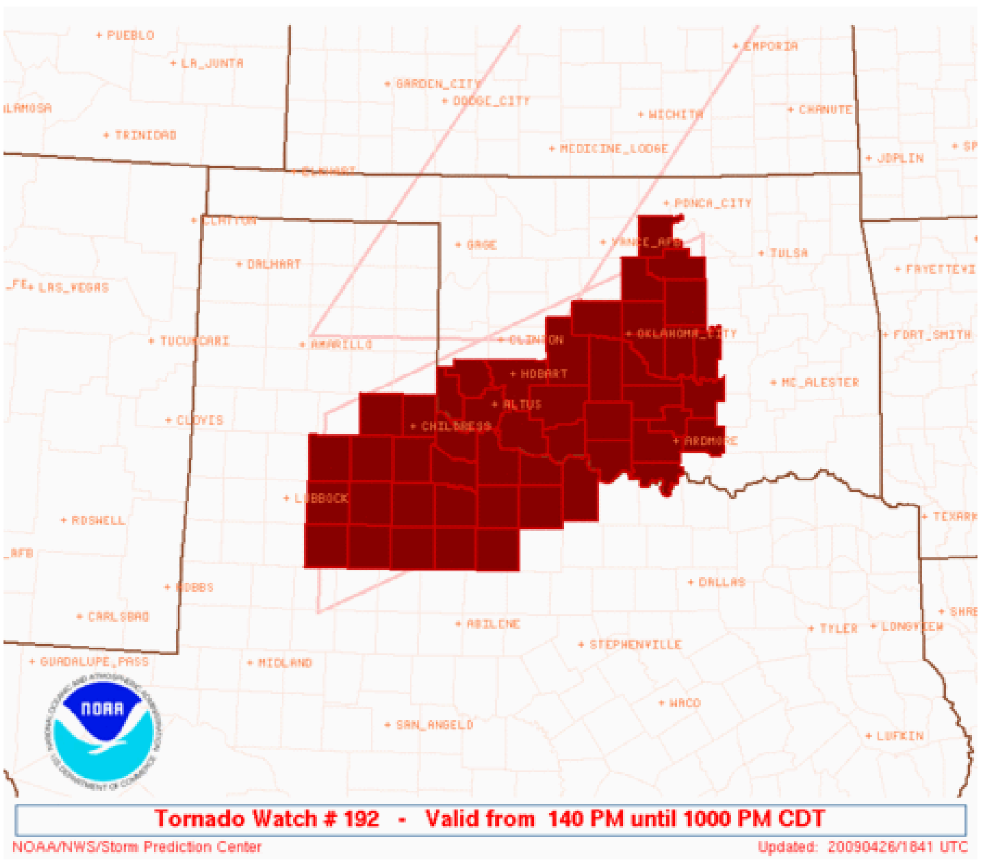



Figure 2.28: Tornado watch #192, valid from 1:40PM CDT to 10:00PM CDT on April 26th, 2009.

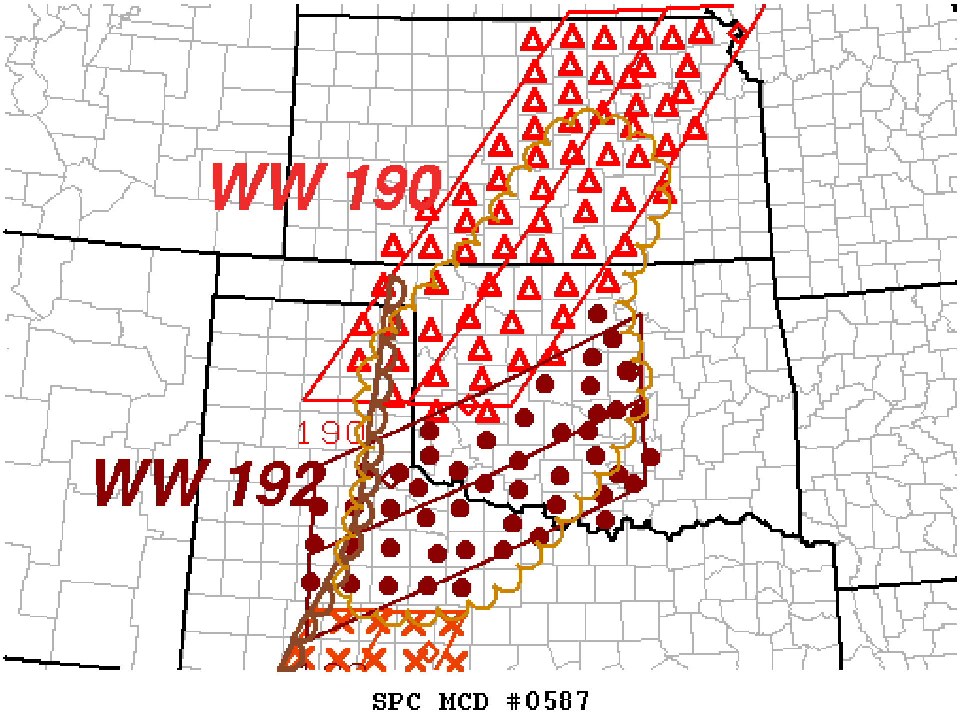



Figure 2.29: Mesoscale Discussion #0587, issued at 5:44PM CDT on April 26th, 2009.

At 6:15 PM CDT, PDS tornado watch #195 was issued, replacing tornado watches #190 and #192, as shown in Figure 2.30. Differing from these two watches, the tornado probabilities in watch #195 were 90% (high) for the probability of two or more tornadoes, and 70% (high) for the probability of one or more strong tornadoes. The latter probability was increased from the previous two watches, while the former was decreased slightly from watch #190, and unchanged from watch #192. Once watch #195 was issued, multiple tornado warnings continued to be extended in the watch's area, but after 7:20 PM CDT, no more tornado warnings were issued. The remaining discrete cells merged with one of the lines of storms moving through the region. Although environmental conditions supportive of significant tornadoes persisted, the overall storm mode transitioning from discrete supercells to a QLCS limited tornado production. A handful of tornadoes occurred on April 26th, but they were either in the previous tornado watch #190, or were in a different watch entirely (Figure 2.31). Despite this tornado watch not ‘verifying’ in the traditional sense, conditions were favorable for tornadoes in the watch area for a significant amount of time, they simply did not occur.

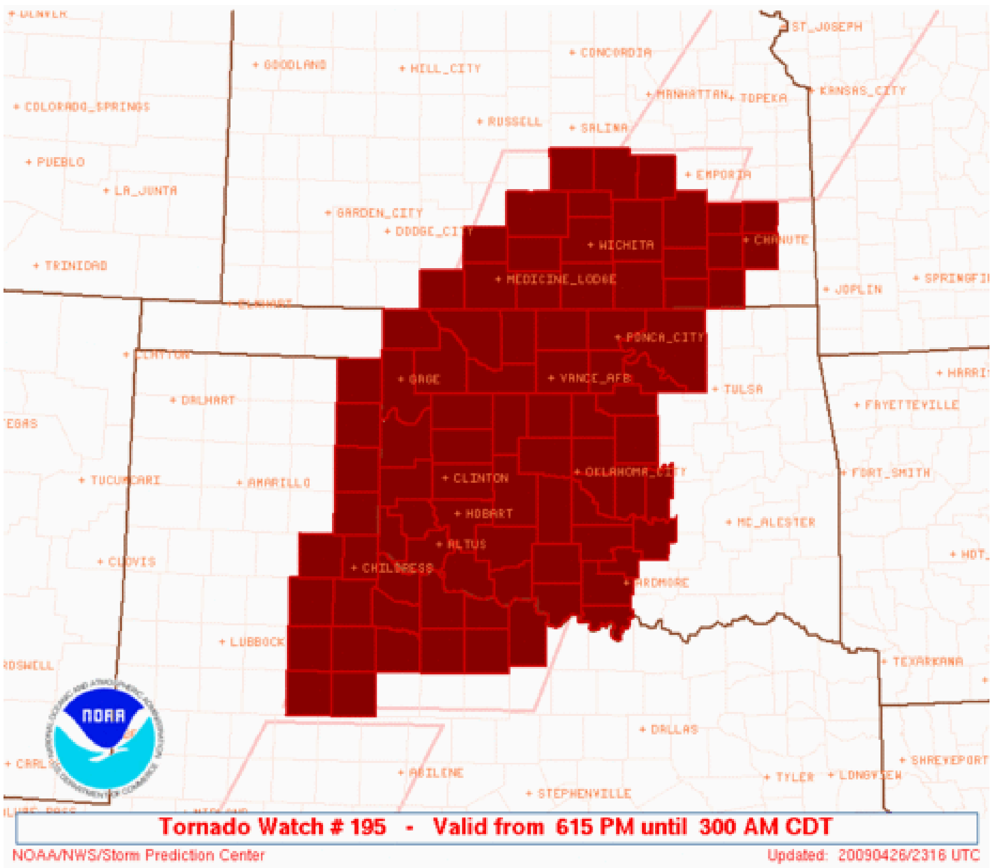



Figure 2.30: Tornado watch #195 of 2009, issued at 6:15PM CDT on April 26th, 2009, and expired at 3:00AM CDT April 27th, 2009. Red counties indicate the official tornado watch area, with the central pink box indicating the watch polygon.

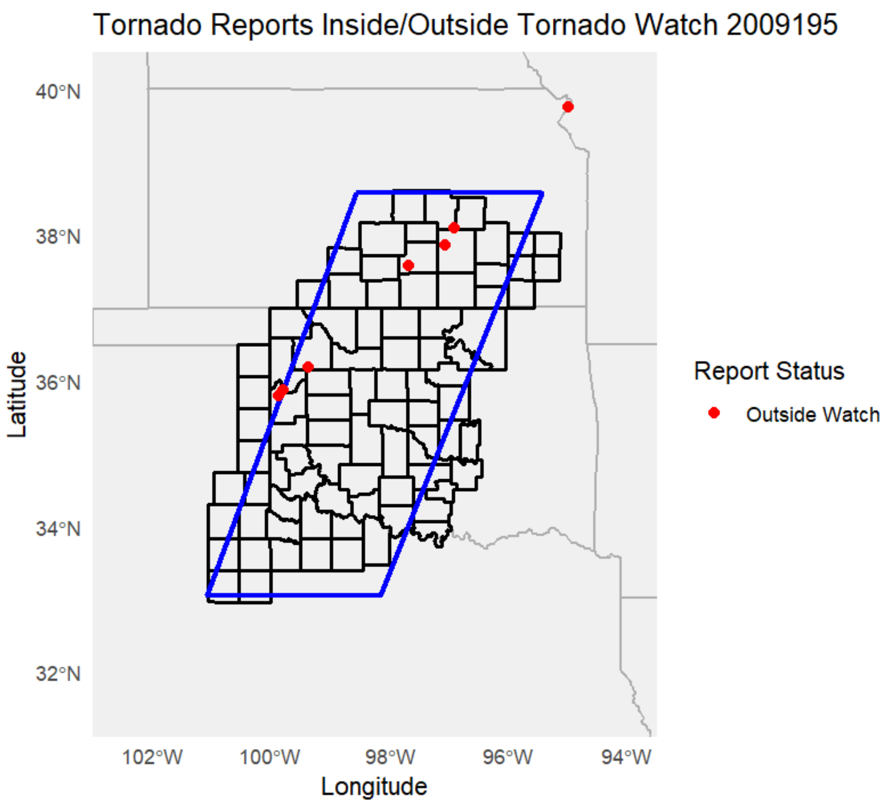



Figure 2.31: Confirmed tornado reports inside/outside tornado watch #195 of 2009. The black county outlines indicate the official watch area, with the blue box representing the tornado watch polygon.

## 2.4.3 Tornado Watch #559 12/10/2021

Tornado watch #559 for 2021 was issued on December 11th at 0500 UTC (11:00 PM CST December 10th) and expired on December 11th at 0952 UTC (3:52 AM CST). In the days leading up to the watch and severe weather, the SPC did not issue a 30% or 15% on any of the day 4-8 convective outlooks due to the uncertainty surrounding the location and development of the upper level trough across the US. By day 4, there were still multiple limiting factors for not issuing a 15% over the US, such as the positive tilt of the trough and the uncertainty around convection occurring early in the day inhibiting the chance for more severe storms in the warm sector later in the day. Day 2 and 3 of the convective outlook continued to discuss some uncertainty, but a slight and an enhanced were issued over the lower/mid Mississippi river valley on day 3 and day 2 (Figure 2.32) respectively. By the first day 1 convective outlook issuance (0600 UTC), much of the uncertainty around the severity of the event had disappeared, with the exception of the timing and spatial distribution of convection after sunset. The discussion stated that even with nocturnal initiation of storms, strong tornadoes with supercells were still expected, and “serious consideration was given to higher severe probabilities” in the enhanced area. As conditions over the next ten hours became more favorable for strong tornadoes and severe weather, the 1630 UTC day 1 convective outlook upgraded the enhanced to a moderate, with a 15% hatched risk for tornadoes, as seen in Figure 2.33.

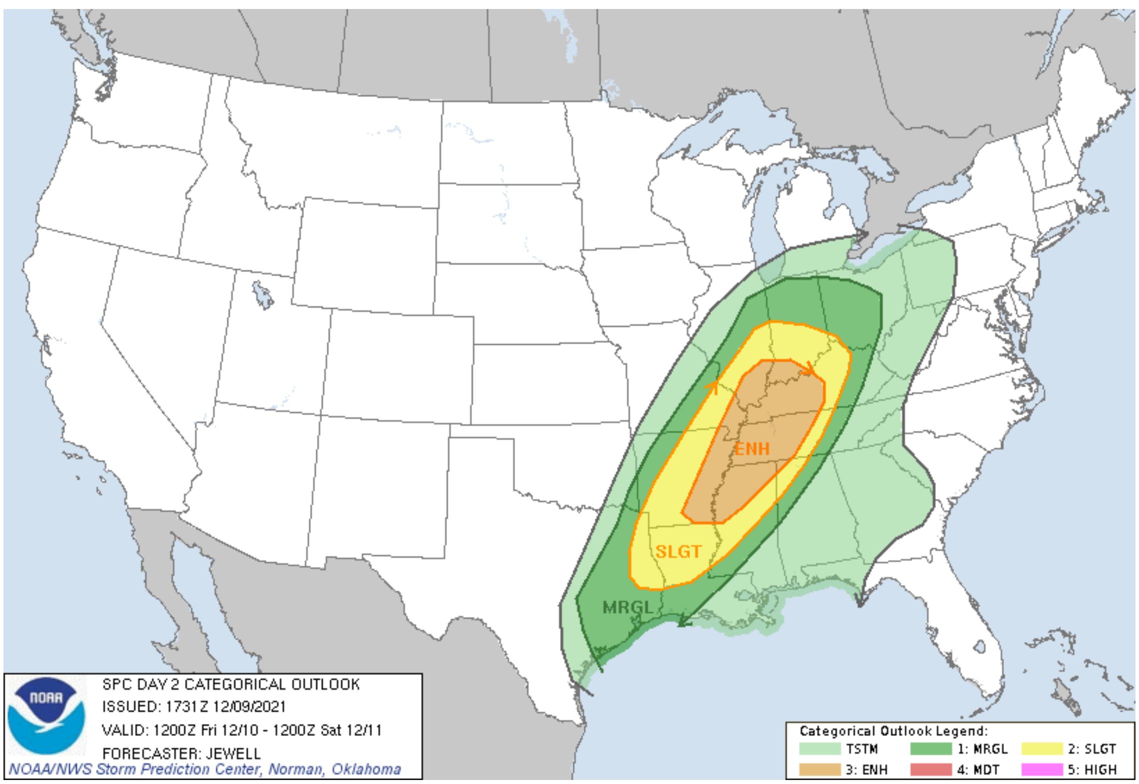



Figure 2.32: The SPC’s Day 2 Convective Outlook, issued December 9th, 2021 at 17:30 UTC. An enhanced risk is located over the central Mississippi river valley, extending from southern Arkansas to southern Indiana.

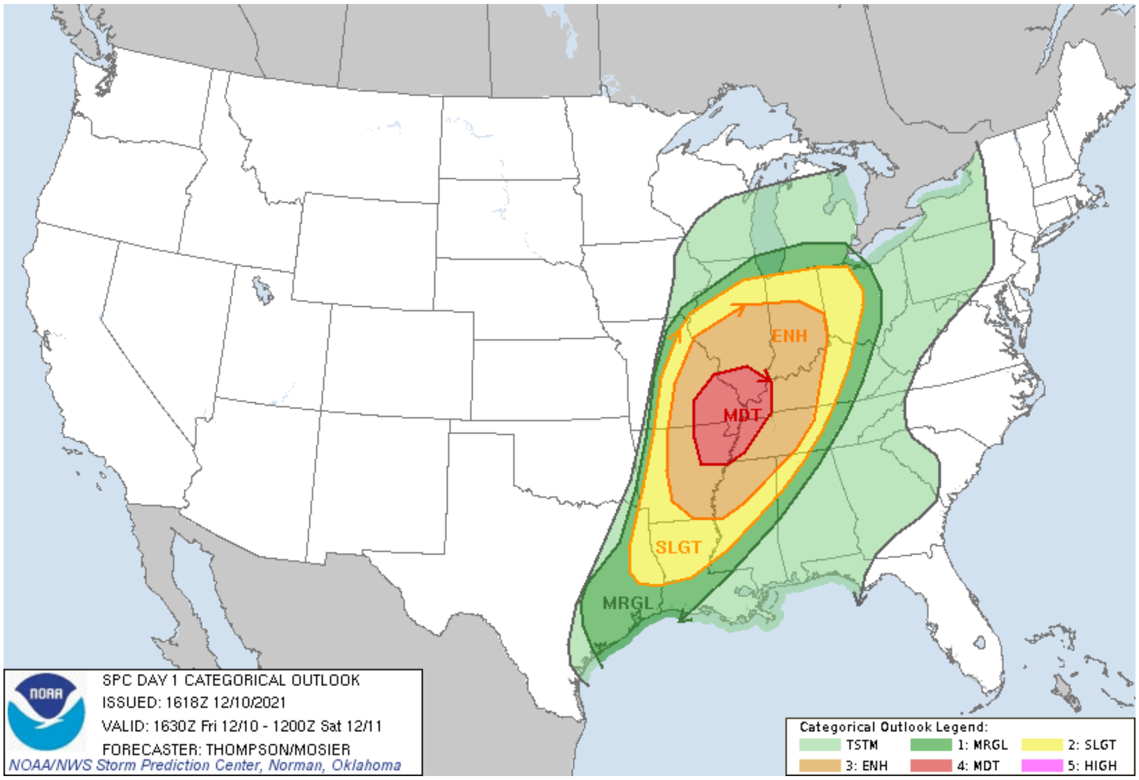



Figure 2.33: The SPC's Day 1 Convective Outlook, issued December 10th, 2021 at 16:18 UTC. The previous enhanced risk has been upgraded to a moderate risk over the central Mississippi river valley, centered on the Missouri bootheel.

At this point in time, the midlevel trough was positioned over the high plains, with a surface low pressure system developing in Kansas. Both were expected to move eastward overnight, with the surface low deepening as it does so. With convective inhibition expected to weaken by the evening, storms were expected to develop in the warm sector in front of the cold front. At 1:35 PM CST, MD #1978 is issued, stating that the risk for supercell development was increasing, and these supercells could potentially produce strong tornadoes by 4-6 PM CST (Figure 2.34). The first tornado watch is issued at 3:00 PM CST, with an expiration time at 11:00 PM CST (Figure 2.35). Both tornado probabilities for watch #552 were set at the high category, with the probability of two or more tornadoes at 90% and the probability of one or more strong tornadoes at 70%. At roughly 5:00 PM CST, individual cells begin to form in central Arkansas to western Kentucky, with tornado warnings quickly following development. Storms in central Missouri quickly begin to become severe between 5:00 and 6:00 PM CST, with a second round of initiation following shortly after. Multiple other tornado watches are issued all around tornado watch #552 as multiple areas of storms begin to strengthen. However, it is important to note that none of these watches have tornado probabilities as high as watch #552, including watch #554 (Figure 2.36), which was issued at 6:20 PM CST and extended to the east of watch #552.

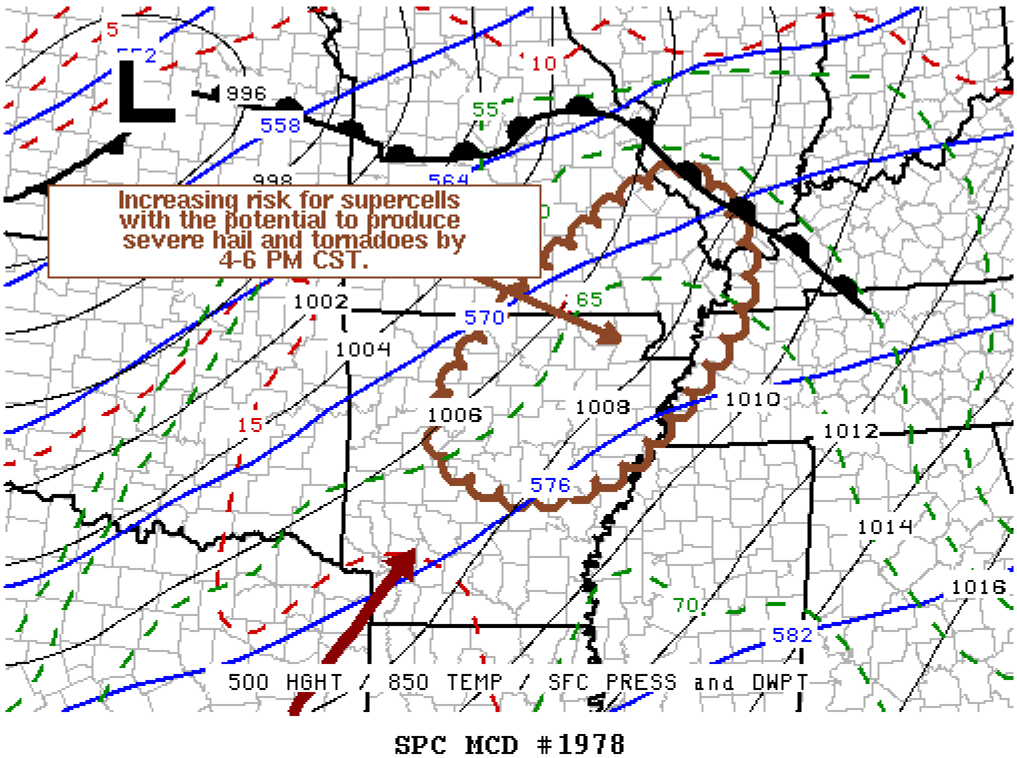



Figure 2.34: Mesoscale discussion #1978, issued at 1:35PM CST on December 10th, 2021.

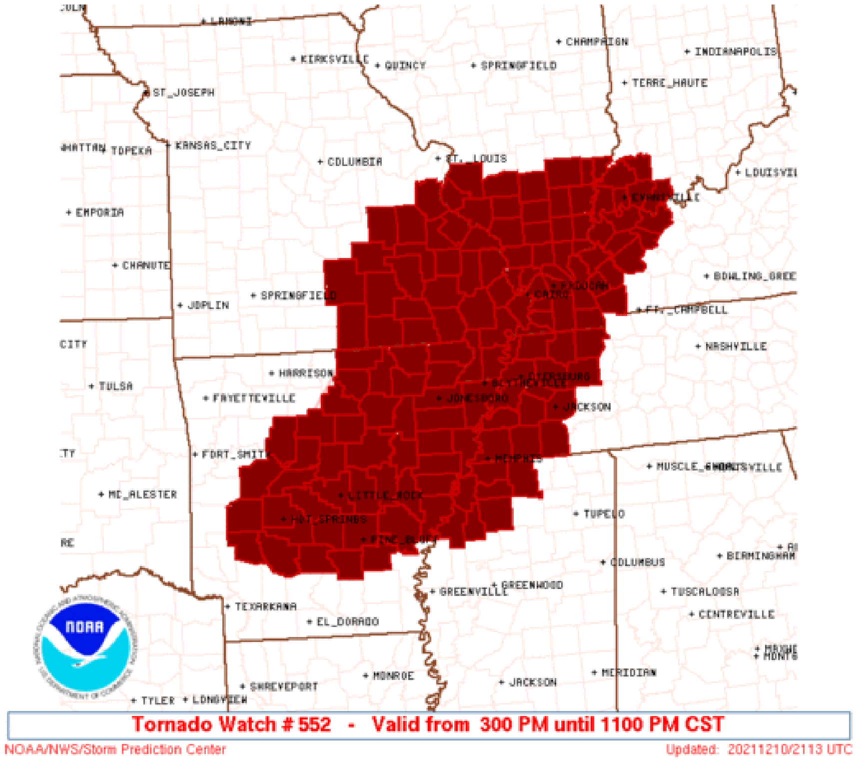



Figure 2.35: Tornado watch #552 of 2021, issued at 3:00PM CST, and expired at 11:00PM CST on December 10th, 2021. Red counties indicate the official tornado watch area.

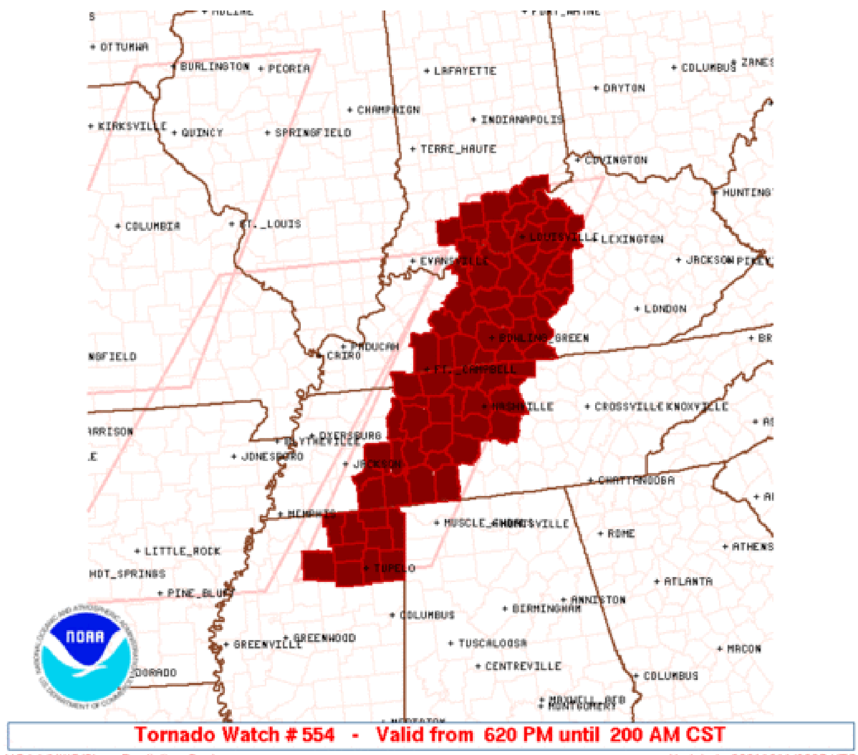


Figure 2.36: Tornado watch #554 of 2021, issued at 6:20PM CST on December 10th, 2021, and expired at 2:00AM CST on December 11th, 2021. Red counties indicate the official tornado watch area, with the central pink box indicating the watch polygon.

MD #1998 (Figure 2.37), issued at 9:40 PM, stated that there was some uncertainty if and where there was an area of higher tornado potential, but concluded that if there was such a region, it would be western Tennessee. After MD #1998 and before 11 PM CST, two other MDs (#1999, #2003) were issued detailing a strong and potentially violent long track tornado that was expected to continue for an extended period of time through western Kentucky. As tornado watch #552 was set to expire at 11:00 PM CST, tornado watch #558 was issued at 10:45 PM CST extending the previous watch to the east (Figure 2.38). Watch #558 had only a 70% (high) for the probability of 2 or more tornadoes, and only 50% (moderate) for the probability of one or more strong tornadoes. At this point in time, the strong, long track tornado mentioned in MD #1999 and #2003 was in western Kentucky, which was not included in tornado watch #558. However, the first note of a tornadic debris signature from a different long track tornado in western Tennessee was at 10:37 PM CST, with the first report of downed trees and power lines at 10:47 PM CST. MD #2006, issued at 11:36 PM CST, mentioned this tornado, but only stated that this cell had produced tornadoes in the past and was now in an extremely favorable position for a strong tornado(s).

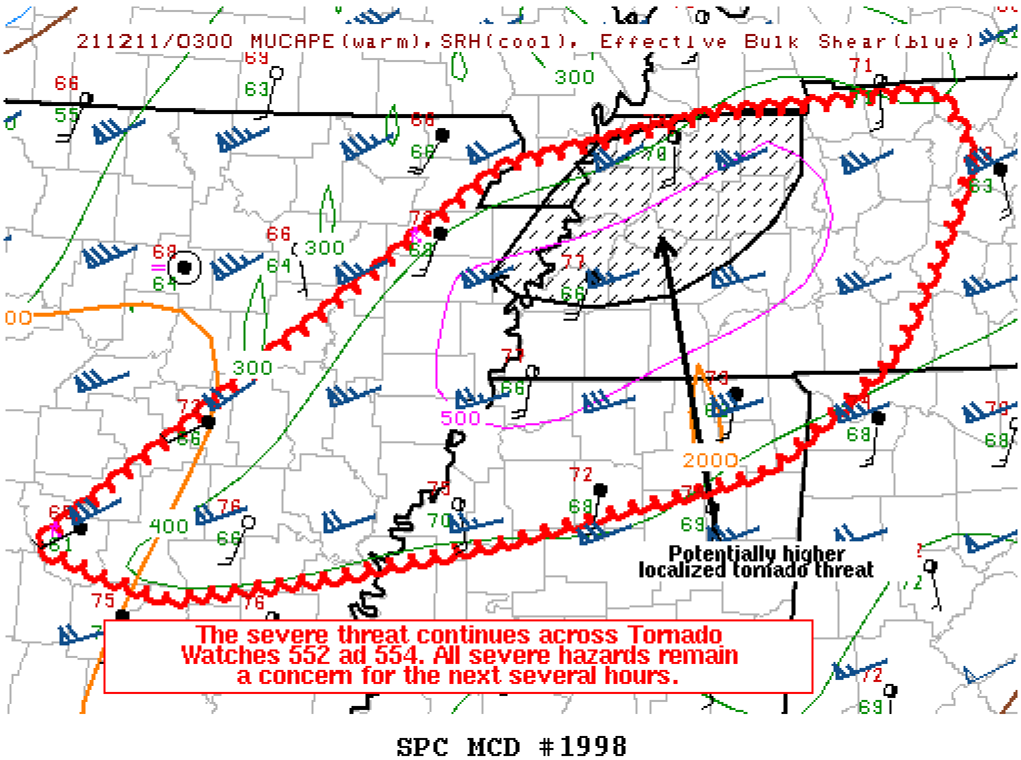



Figure 2.37: Mesoscale discussion #1998 of 2021, issued at 9:40PM CST on December 10th, 2021.

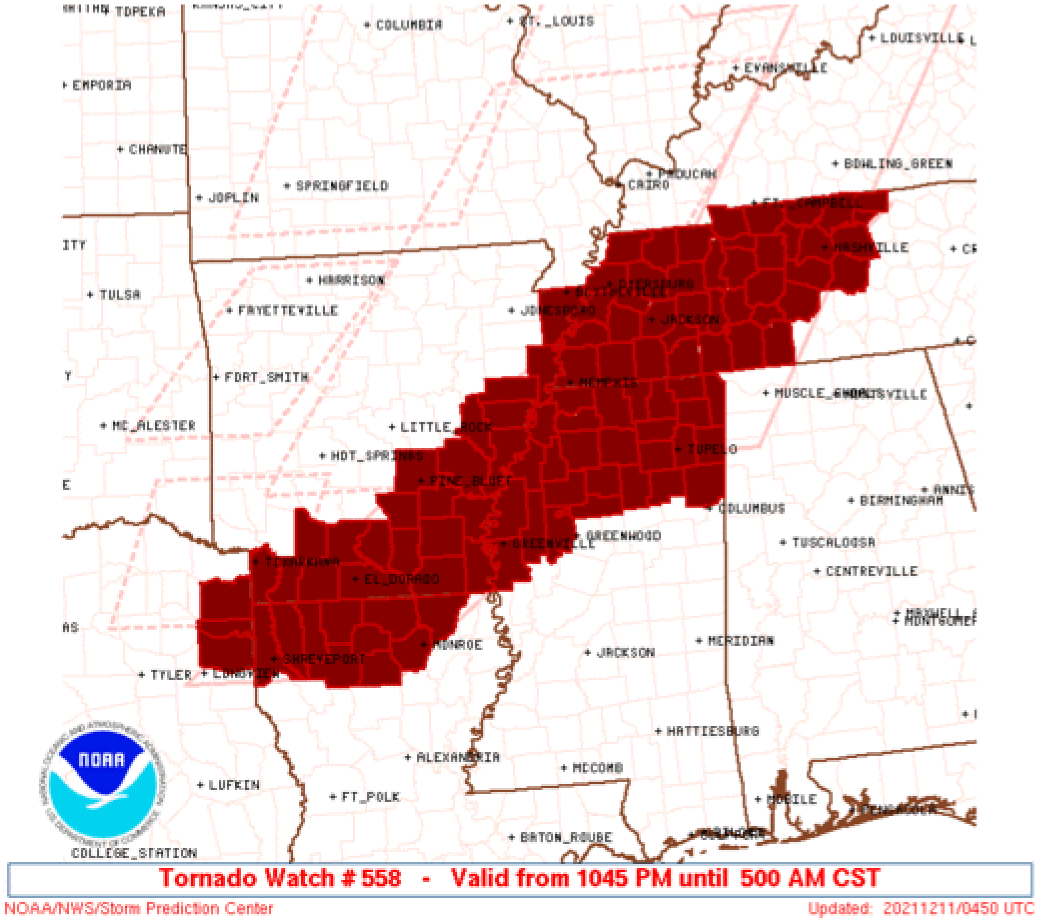



Figure 2.38: Tornado watch #558, issued at 10:45PM CST on December 10th, 2021, and expired at 5:00AM CST on December 11th, 2021. Red counties indicate the official tornado watch area, with the central pink box indicating the watch polygon.

At 11:00 PM CST, tornado watch #559, as seen in Figure 2.39, was issued as a replacement and expansion for tornado watch #552. Both tornado probabilities were in the high category, with the probability of two or more tornadoes at 80%, and the probability of one or more strong tornadoes at 70%. Despite being issued with both tornado probabilities in the high category, no tornadoes occurred in the watch while it was active (Figure 2.40). The first tornado mentioned above, a violent, long-track EF4 tornado, exited the watch area just as the watch was issued at 11:00 PM CST. However, aside from tornado watch #552, this was the only tornado watch of the day issued with both probabilities in the high category. The second tornado mentioned above, turned out to be a long track EF3 tornado, occurring in watch #552 before passing into watch #558. Since watch #558 was issued very close in time to the start of the EF3 tornado, the lower probabilities in watch #558 could be indicative of forecasters expecting the environment to not be as favorable as it was for the EF4 tornado in this new watch area, and the higher probabilities in watch #559 could be reflective of this difference.

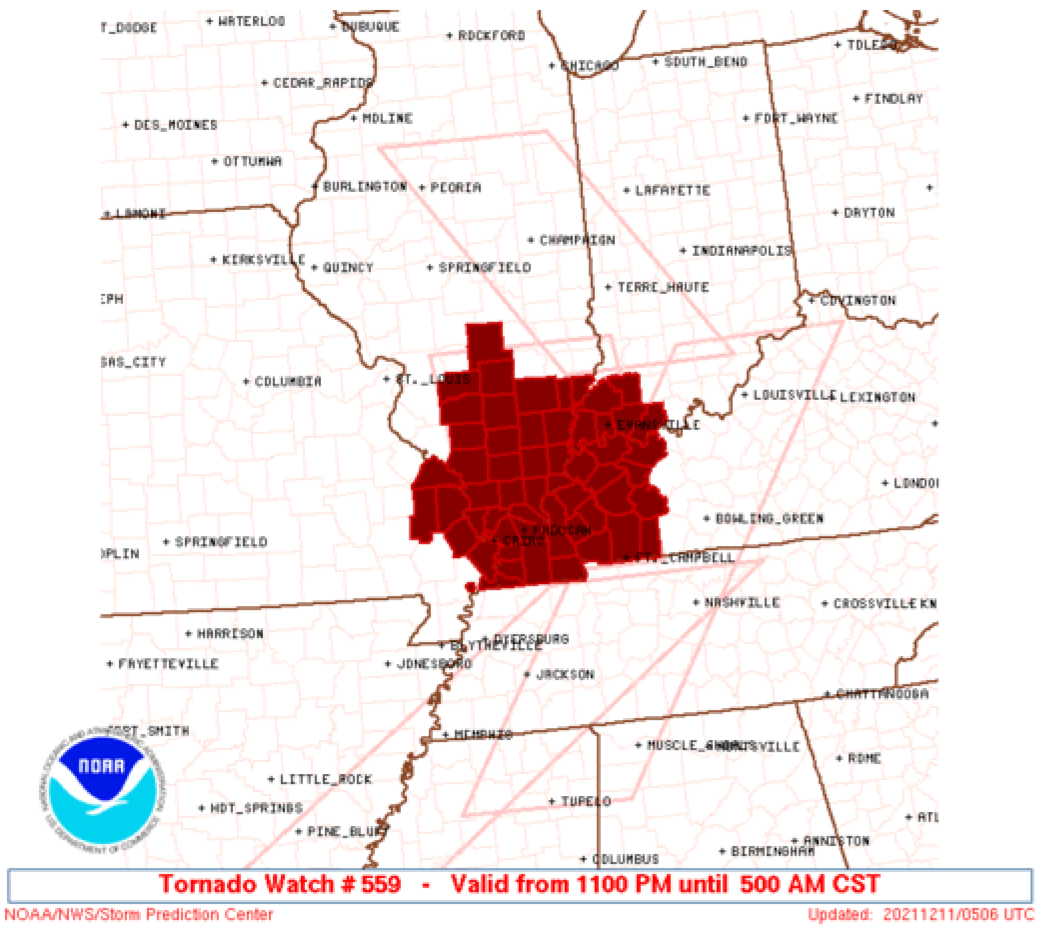



Figure 2.39: Tornado watch #559 of 2021, issued at 11:00PM CST December 10th, 2021, and expired at 5:00AM CST on December 11th, 2021. Red counties indicate the official tornado watch area, with the central pink box indicating the watch polygon.

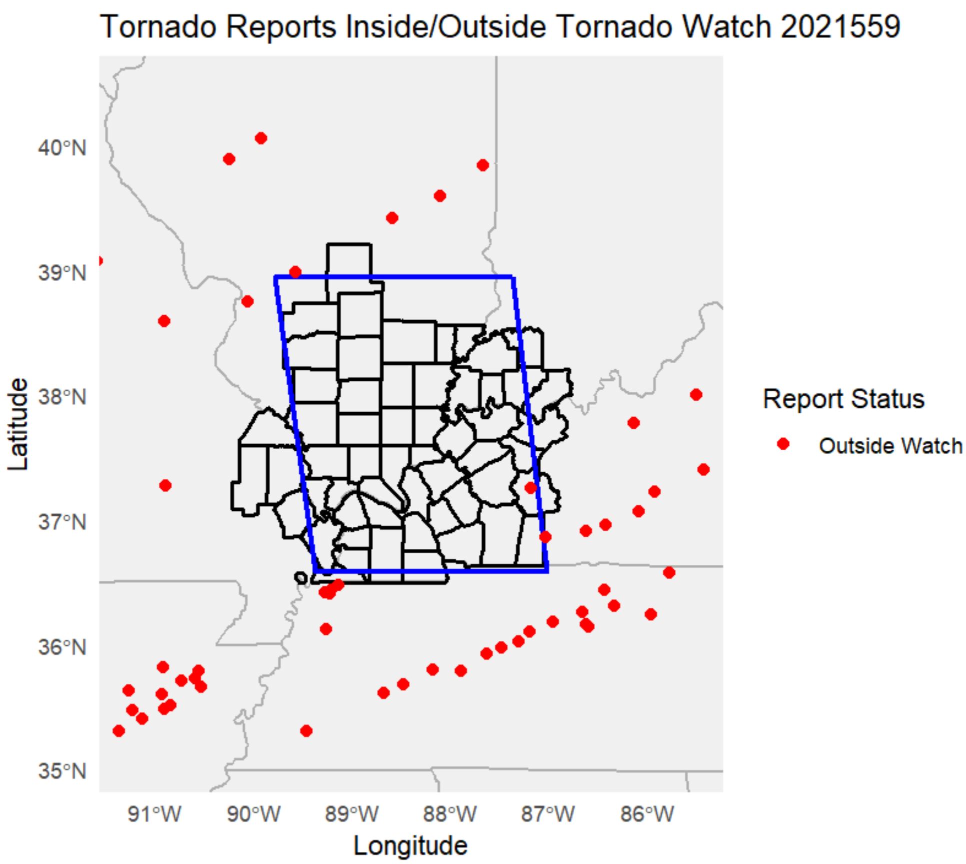



Figure 2.40: Confirmed tornado reports inside/outside tornado watch #559 of 2021. The black county outlines indicate the official watch area, with the blue box representing the tornado watch polygon.

## 2.4.4 Tornado Watch #554 10/31/2013

Tornado watch #554 of 2013 was issued at 22:25 UTC (5:25PM CDT) on October 31st, and expired at 04:55 UTC (11:55PM CDT) on November 1st (October 31st). Prior to this watch being issued, the SPC's day 4-8 convective outlook issued on October 27th, 2013 indicated a 30% chance of severe thunderstorms within 25 miles of a point for the lower Mississippi River valley, as seen in Figure 2.41. During this time frame, it was predicted that a considerable upper trough and associated cold front would intercept a moist low-level airmass, bringing severe thunderstorms to the central CONUS in the coming five days. By the day 3 convective outlook, a slight was issued from eastern Texas to western Ohio (Figure 2.42), with forecasters expecting convection to occur ahead of the cold front and potentially complicating where the most severe weather would occur.

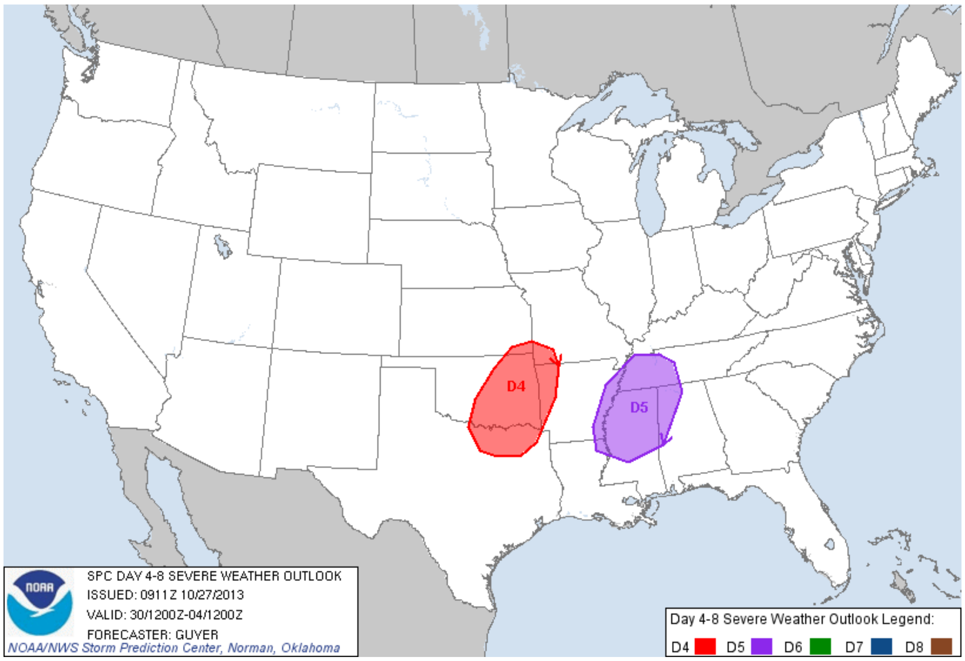



Figure 2.41: SPC's day 4-8 convective outlook, issued on October 27th 2013 at 09:11 UTC. Day 4 and day 5 outlines are red and purple respectively.

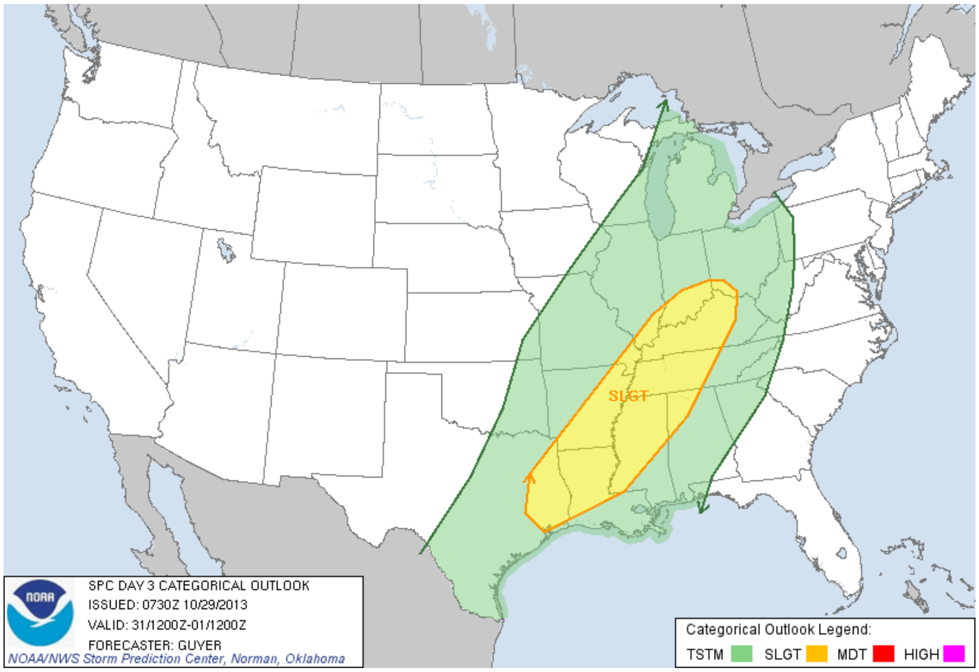



Figure 2.42: SPC day 3 convective outlook issued October 29th, 2013 at 07:30 UTC. A slight risk is located over the south central US, from eastern Texas to south western Ohio.

The early morning of October 31st was particularly rainy, with precipitation not moving out of the region of the impending tornado watch until 11:00AM. At 11:30AM, the updated SPC day 1 convective outlook included an upgraded 5% tornado area over the southeast portion of Missouri into the southwestern tip of Indiana (Figure 2.43). Forecasters expected localized parts of this region to experience modest heating, which combined with the strong large-scale ascent would allow for low-topped cells capable of producing a couple of tornadoes if sustained updrafts ended up forming.

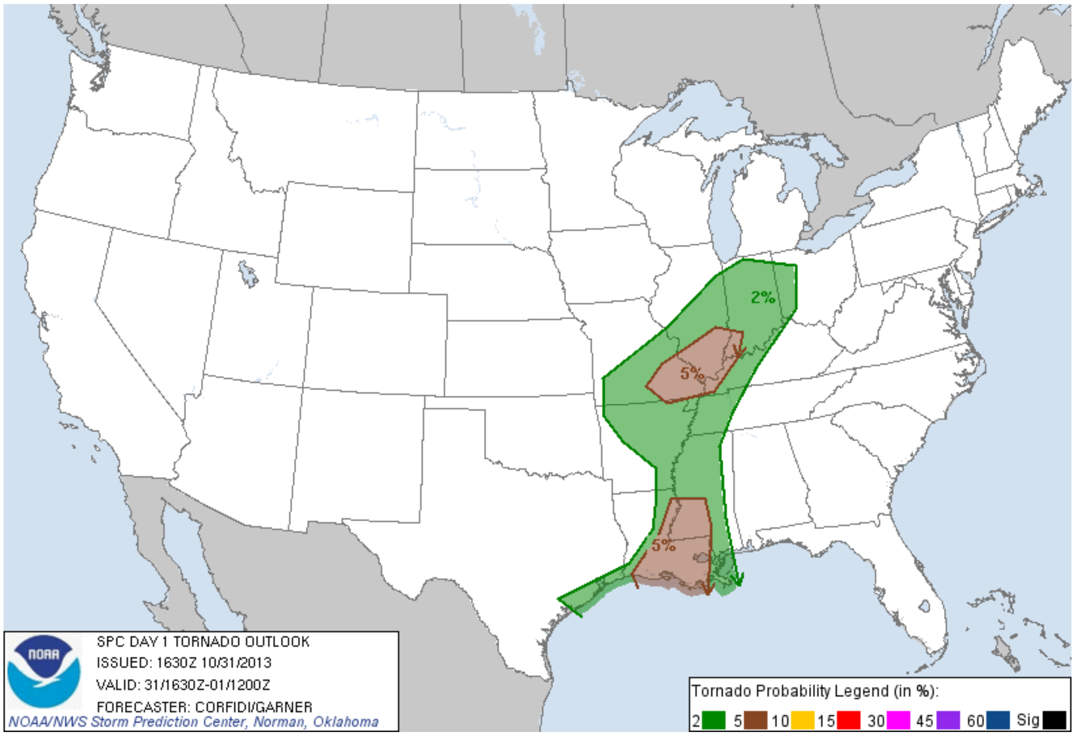



Figure 2.43: SPC day 1 convective outlook probabilistic tornado graphic. Issued at 16:30 UTC on October 31st, 2013. A 5% tornado risk was issued over south eastern Missouri to south western Indiana.

Within an hour of the updated SPC convective outlook, MD #1933 (Figure 2.44) was issued, and detailed the increasing severe storm potential over the Ozarks with the probability of issuing a severe thunderstorm watch in the coming hours at 60%. Forecasters expected these storms to develop and strengthen as they moved to the east across the region, with damaging wind being the main threat and a risk for a couple of tornadoes. As isolated multicells and supercells began to develop over south central Missouri, severe thunderstorm watch #553 was issued at 2:55PM CDT (Figure 2.45), with the associated tornado probabilities at 10% and <2% for the probability of 2 or more tornadoes and the probability of one or more strong tornadoes occurring in the watch, respectively.

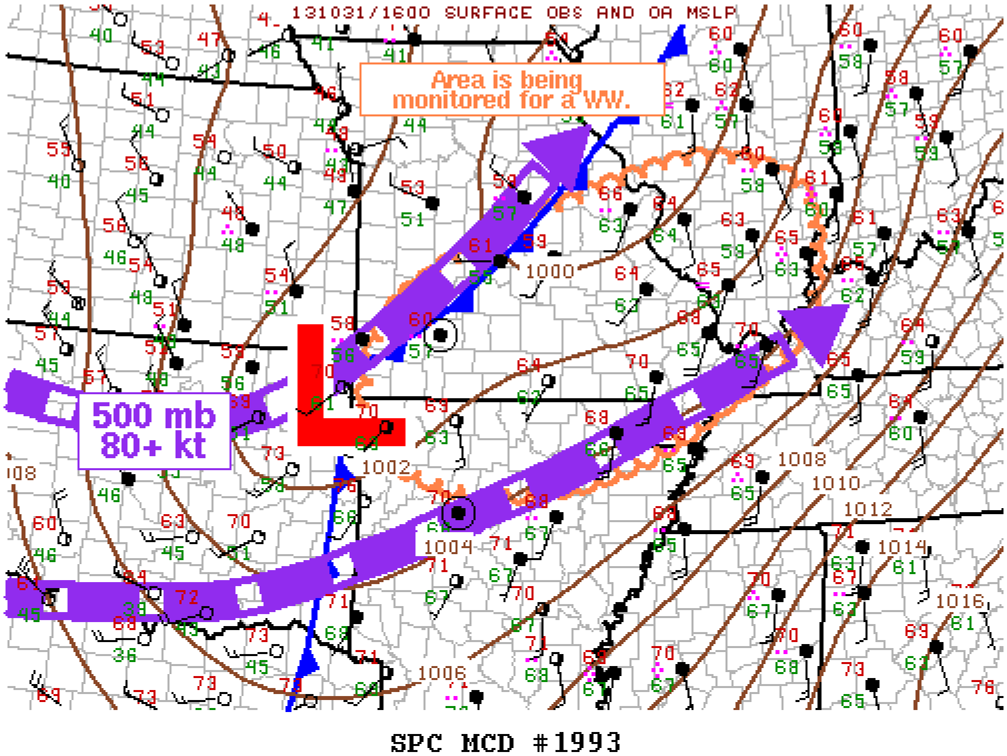



Figure 2.44: Mesoscale Discussion #1933 issued 12:31PM CDT October 31st, 2013.

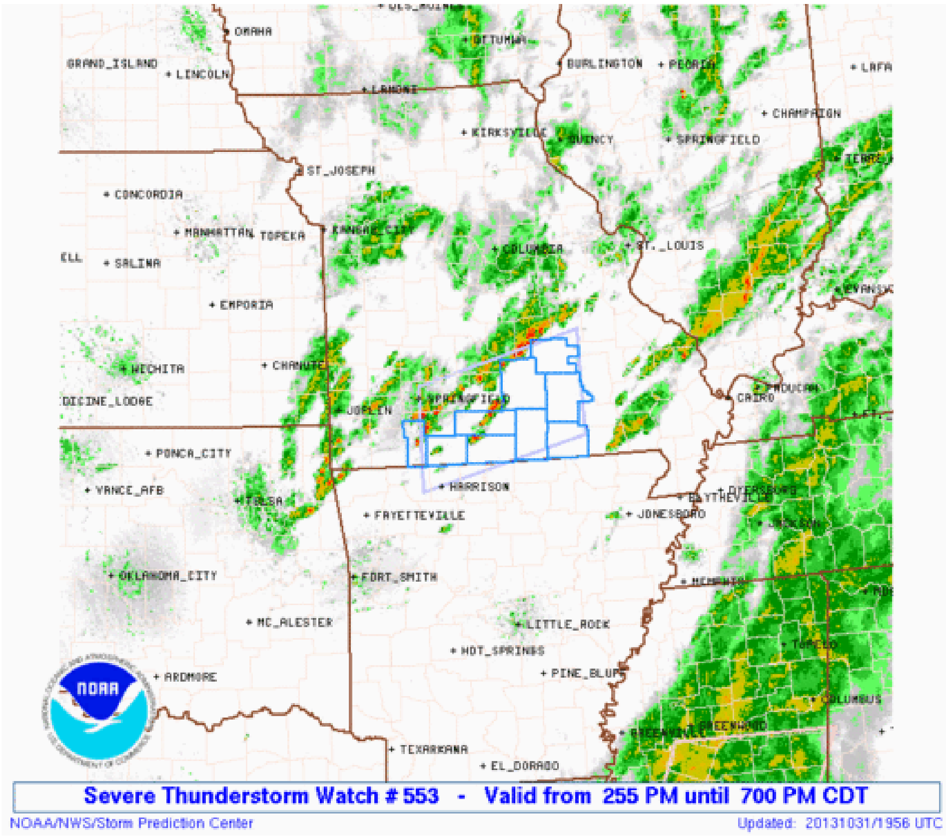



Figure 2.45: Severe thunderstorm watch #553 issued at 2:55PM and canceled at 6:17PM on October 31st, 2013.

As the cold front continued to sweep across the region and these multicells continued to develop, MD #1995 was issued at 5:01PM CDT (Figure 2.46). This discussion described the extremely strong shear and wind profiles ahead of the cold front and the uncertainty around if this strong wind shear would be effective in strengthening the storms since instability was relatively weak. Furthermore, with poor lapse rates aloft, only low topped convection would be possible in these conditions. However, forecasters expected a line of storms to form from the cells moving to the east due to the strong forcing with the cold front and the moist low level profiles. The main threat was expected to be damaging winds or an isolated tornado, with any tornado that would occur to be brief, but the possibility remained for it to be strong.

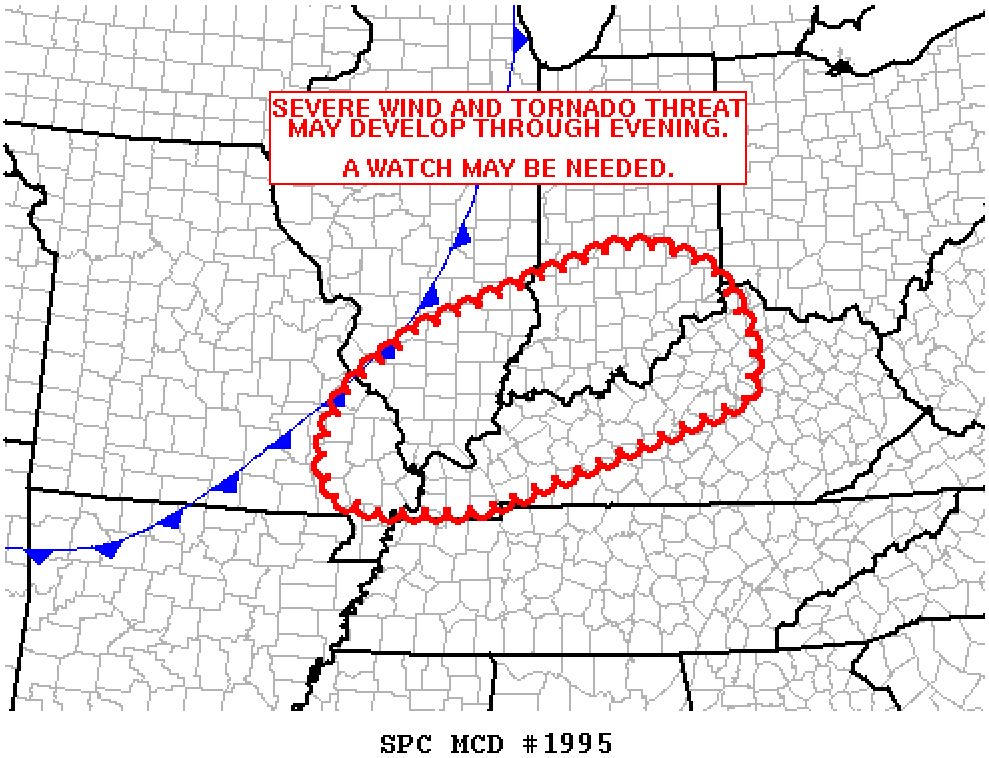



Figure 2.46: Mesoscale discussion #1995 issued at 5:01PM CDT on October 31st, 2013.

Shortly following MD #1995, tornado watch #554 was issued for the same region the MD described, at 5:25PM CDT on October 31st, 2013 (Figure 2.47). Both tornado probabilities were issued at 20%, an increase for both probabilities from the previously issued severe thunderstorm watch #553 earlier in the day. Within thirty minutes of the tornado watch being issued, a line of storms formed from the already existing cells, just as the previous MD had described. Multiple severe thunderstorm warnings were issued in the first two hours of the tornado watch being active, with multiple tornado warnings starting around 7:15PM CDT. With the extension of the tornadic storm line farther south than the tornado watch covered, tornado watch #555 (Figure 2.48) was issued at 7:55PM to cover the counties to the south of tornado watch #554.

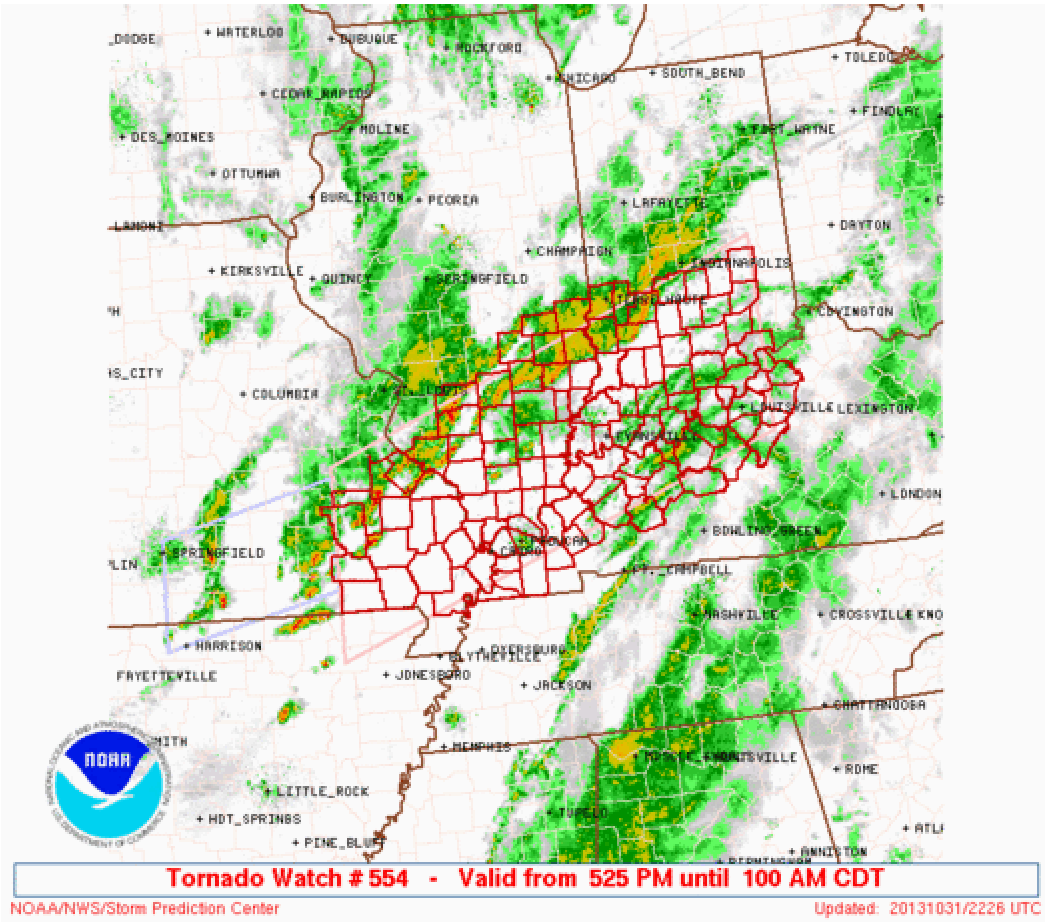



Figure 2.47: Tornado watch #554 issued at 5:25PM and canceled at 11:55PM on October 31st, 2013.

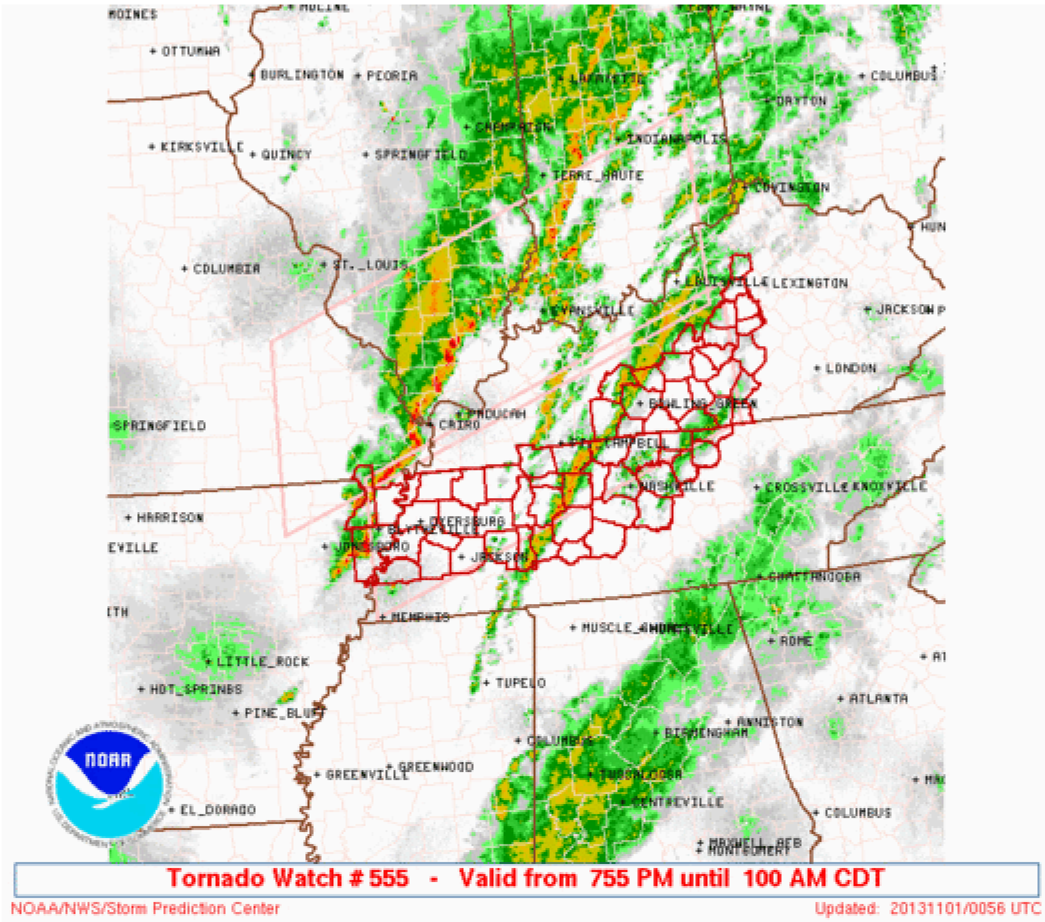



Figure 2.48: Tornado watch #555 issued at 7:55PM CDT on October 31st, 2013 and expired at 1:00AM on November 1st, 2013.

Throughout the evening of October 31st, 2013, a multitude of tornado and severe thunderstorm warnings were issued, with a total of 18 confirmed tornadoes occurring, the highest rating being an EF2 for one tornado. In terms of the tornado probabilities associated with tornado watch #554, both verified with the number of confirmed tornadoes. For the probability of one or more tornadoes within the watch area, it may seem that the 20% issued in tornado watch #554 'over-verified', with 18 tornadoes occurring in the watch. One may question why this probability was not placed higher than 20%. However, when analyzing the reports of these tornadoes in Figure 2.49, we see that the majority of the confirmed tornado reports are in sequence with two main tracks. The environment in this region was not conducive to producing strong, long track tornadoes. Instead, the environment favored brief, less powerful tornadoes. Based on the environmental conditions for this watch, this environment was HSLC, which forecasters view as particularly difficult to forecast in. Thus, while the line of storms did produce 18 separate tornadoes, many of these occurred from the same two segments of this system, and likely could have been singular tornadoes had conditions been more favorable. Additionally, the uncertainty of the strength of this storm system was communicated thoroughly before the event occurred, and 17 of the 18 tornadoes within the watch were rated EF1 or lower.

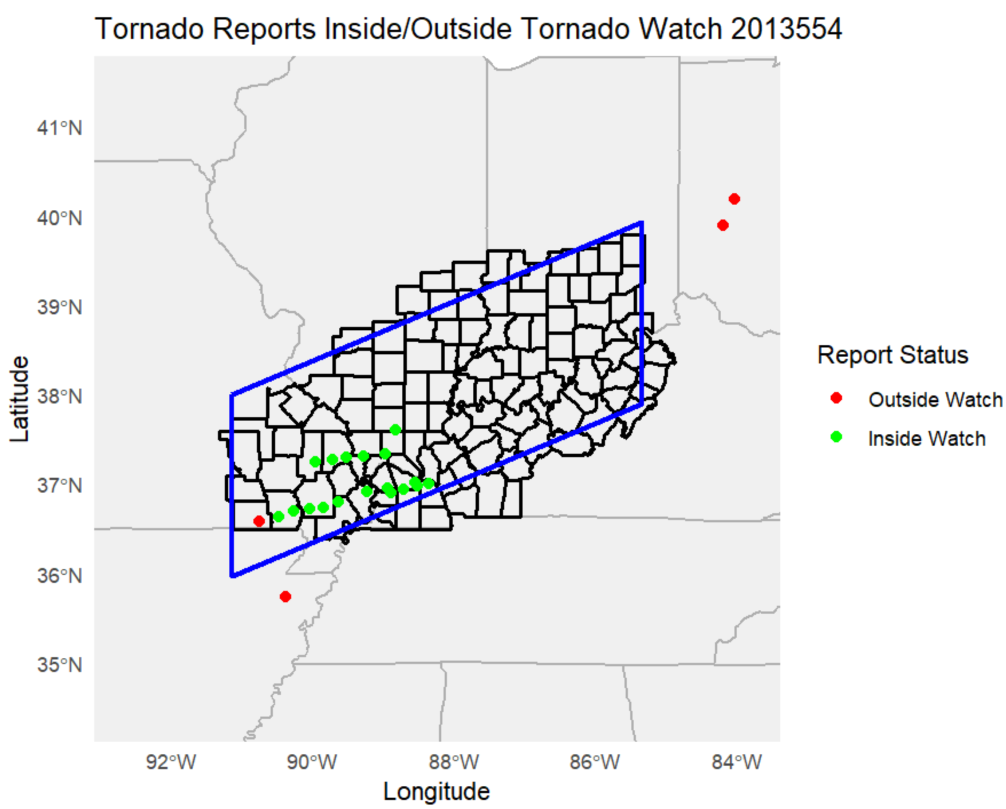



Figure 2.49: Confirmed tornado reports inside/outside tornado watch #554 of 2013. The black county outlines indicate the official watch area, with the blue box representing the tornado watch polygon.

### 2.4.5 Tornado Watch Case Study Discussion

The four analyzed tornado watch case studies discussed above give four different examples of the endless situations forecasters have to evaluate. While these case studies may only give us a glimpse into some of the scenarios forecasters encounter, they do provide some perspective on why some tornado probabilities may seem like they do or do not verify in the traditional sense. Of the three case studies that were forecasted with high tornado probabilities in tornado watches and did not ‘verify’, the first shows that tornadoes were contingent on the timing and environmental conditions in a specific place, with the third case having favorable conditions for tornadoes to occur, but the storm mode wasn’t the most conducive for them to form. The second case showed the complex dynamics of forecasters potentially being impacted by strong tornadoes occurring in another location. For all three of these scenarios, significant tornadoes occurred prior to the tornado watch of interest being issued, either in the watch area or outside of it. With all three tornado watches having high tornado probabilities, forecasters may have had an increased confidence in tornado occurrence after these tornadoes occurred, issued high probabilities in these watches, and significant tornadoes within the watch time and area did not materialize. Additionally, the tornado watches of the second and third cases functioned as follow up or extension tornado watches later in the event timeline. This suggests that a better understanding of when tornado events tend to downtrend would be very beneficial for real-time forecasting.

The last case, which was the only example in the dataset where a tornado watch was issued with low tornado probabilities and both probabilities verified, illustrates that while 18 tornadoes did occur within the watch, the majority of these tornadoes came from the same segment of the QLCS and simply did not stay in contact with the ground for the duration of the system. It is also worth noting that anticipating brief but potentially impactful QLCS tornadoes remains a forecast challenge, particularly in HSLC environments such as this one.

Overall, while it may seem as though certain cases ‘bust’ or ‘boom’, it is always crucial to keep in mind the hundreds of factors that go into a forecast being issued, and the decisions that must be made by forecasters in high-stress situations. On the surface we can deem one specific forecast ‘good’ or ‘bad’, but it is always more complicated than how the forecast is taken at face-value. It is how stakeholders use these forecasts to make their decisions that assigns value to a forecast, as stakeholders decisions are ultimately the deciding factor in the actions taken to protect life and property (Murphy 1993).

## 2.5 Severe Thunderstorm Watch Analysis

After analyzing the tornado watches and seeing that 45.8% of tornado reports did *not* occur in a tornado watch, and 40.1% tornado watches had no confirmed tornado report, questions remained as to how often tornadoes were occurring in other types of watches. Severe thunderstorm watches also contain probabilities of tornado occurrence, and have been issued in these watches since April 25th, 2006. In order to analyze potential tornado reports and the tornado probabilities in severe thunderstorm watches, the same process was used as for tornado watches. All severe thunderstorm watch data came from Iowa State University's Iowa Environmental Mesonet (IEM 2025), and included the same three datasets as the tornado watches but with severe thunderstorm watch data. The same process was used for overlaying county and polygon watch outlines, and was then the dataset was intersected with the SPC’s SVRGIS tornado report database (Marsh and Elliott 2025) spatially and temporally.

After removing 34 severe thunderstorm watches due to them not being included in the polygon data, 8,468 severe thunderstorm watches were left over the 19 year period. Of these

8,468 watches, 6,862, or 81%, had no confirmed tornado report(s) occur while the watches were in effect. There were 23,192 tornado reports that occurred between April 28th, 2006 and December 31st, 2024. Of those, 3,618 occurred in a severe thunderstorm watch, 12,303 occurred in a tornado watch, and 257 occurred in both a tornado and severe thunderstorm watch while both were active, as seen in Table 2.1. Overall, 16,178 (69.8%) tornado reports occurred within at least one watch. Additionally, Table 2.1 also shows that of the tornadoes with a rating of EF2 or higher, the majority were inside just a tornado watch (n = 1933, or 8.3% of all reports), with 270 reports, or 1.2% occurring in a severe thunderstorm watch, 29 reports, or 0.1% occurring in both tornado and severe thunderstorm watches, and 345 reports, or 1.5% not occurring in any kind of watch.

| Watch Type | Number of Reports (strong tornadoes EF2+) | Percent of Reports (strong tornadoes EF2+) |
|---|---|---|
| No Watch | 7014 (345) | 30.2% (1.5%) |
| Tornado Watch Only | 12303 (1933) | 53.0% (8.3%) |
| Severe Thunderstorm Watch Only | 3618 (270) | 15.6% (1.2%) |
| Both Watches | 257 (29) | 1.1% (0.1%) |

Table 2.1: Total number of tornado reports and percent of total reports by watch type, with the number and percent of strong tornadoes (EF2+) in parenthesis.

To conduct the verification of tornado probabilities in severe thunderstorm watches, an additional 121 severe thunderstorm watches were removed for not having probabilities, leaving 8,347 watches. Following the same process used for the tornado watches, we see in Figure 2.50 that the most commonly issued combination of severe thunderstorm watch tornado probabilities is 20% for the probability of two or more tornadoes, and 5% for the probability of one or more strong tornadoes, with 2,963 of 8,347, or 35.5%, watches having this combination. There were 10 severe thunderstorm watches with the probability of two or more tornadoes being greater than 20%, at 30%, 40%, 50%, and 60%. These watches started in 2015, and became more frequent getting closer to 2024. These 10 watches were all issued as PDS watches, with some verifying and some not. The probability of one or more strong tornadoes was issued at 10% or 20% in all ten of these watches, and no severe thunderstorm watch out of the 8,347 had a probability of one or more strong tornadoes above 20%. Overall, there were 24 PDS severe thunderstorm watches over the 19 year period, but prior to 2015, both tornado probabilities were 20% in these watches, indicating those watches were mainly hail or wind driven PDS severe thunderstorm watches.

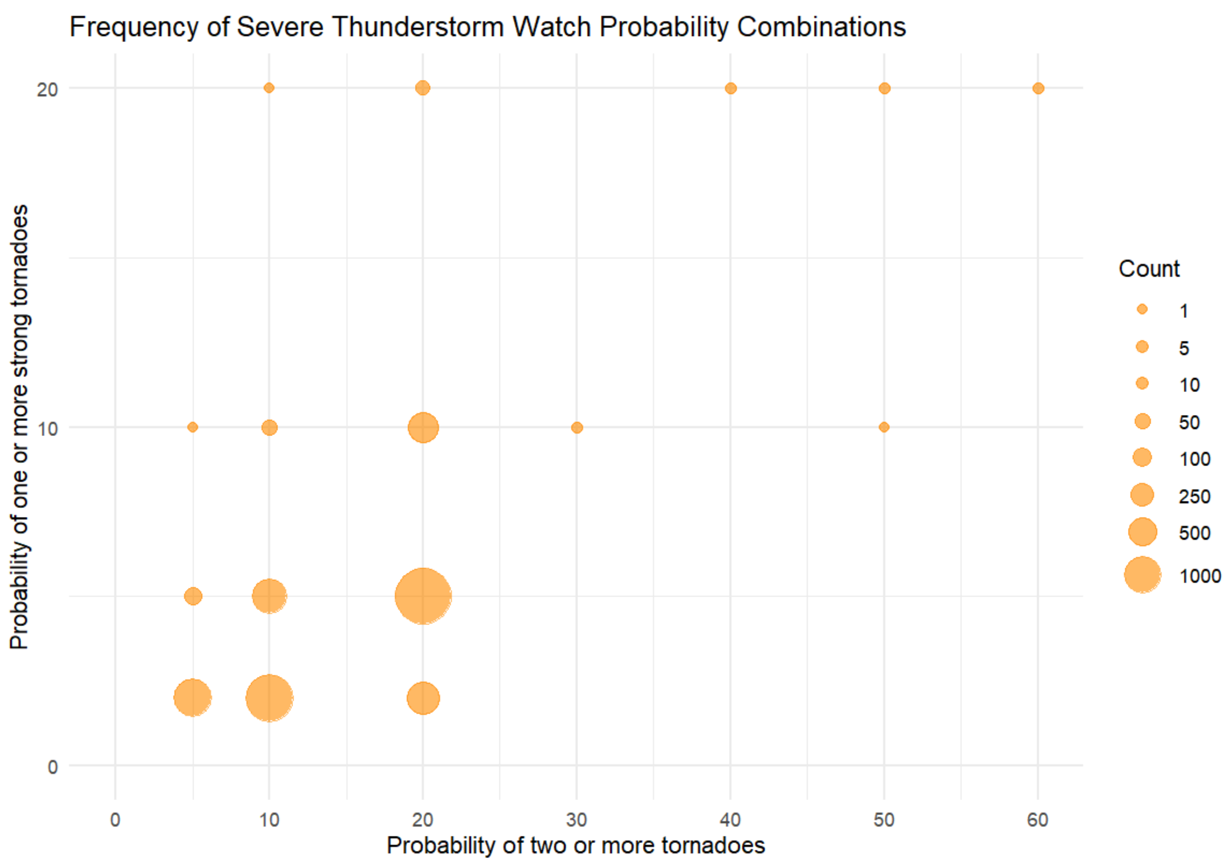



Figure 2.50: Frequency of tornado watch probability combinations for all 8,347 severe thunderstorm watches.

In terms of the performance of the tornado probabilities in severe thunderstorm watches, Figure 2.51 shows that for the probability of two tornadoes, all the probabilities are over-forecasted. The same can be seen in Figure 2.52 for the probability of one or more strong tornadoes, where all probabilities are under the dashed line. However, given that only 15.6% of tornado reports occurred in severe thunderstorm watches (Table 2.1), it is not surprising that both of these probabilities are slightly over-forecasted. While tornadoes do occur in severe thunderstorm watches, tornado watches exist with the context of being issued for a higher risk of tornadoes in comparison to severe thunderstorm watches. This is evident in Tables A.11 and A.12, where we can see the breakdown of the number of watches that had two or more tornadoes, or one or more strong tornadoes occurring in a watch in comparison to how many severe thunderstorm watches were issued.

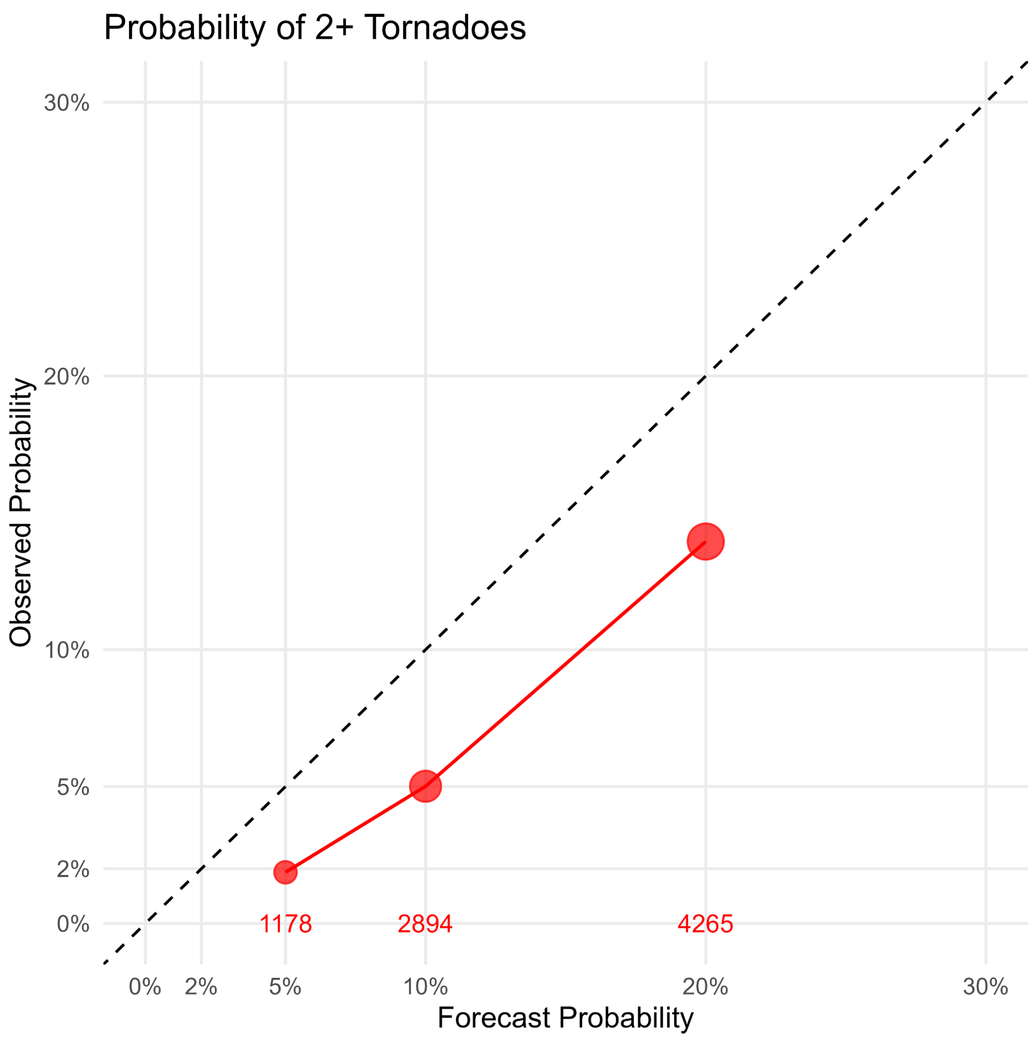



Figure 2.51: Reliability diagram of the probability of two or more tornadoes occurring within a severe thunderstorm watch. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

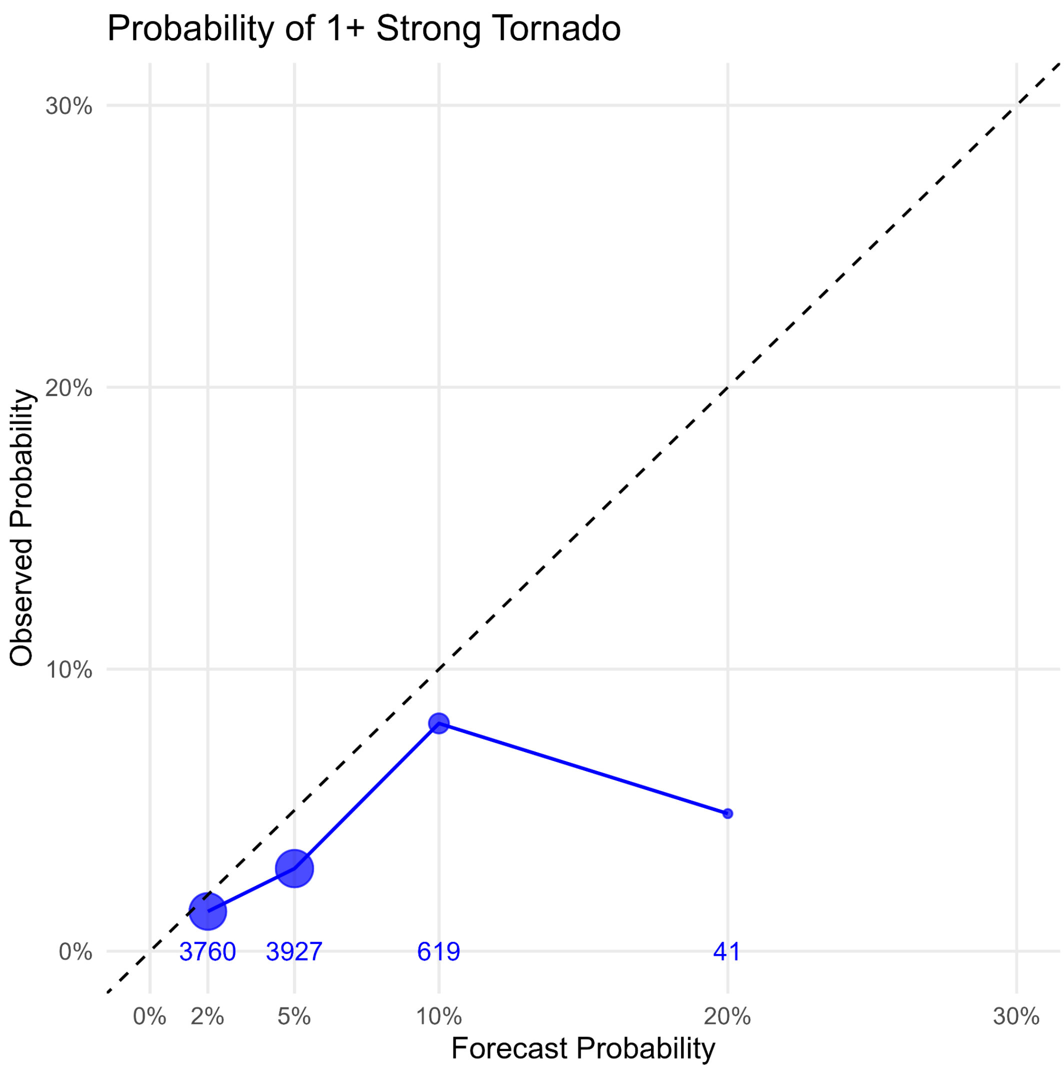



Figure 2.52: Reliability diagram of the probability of one or more strong tornadoes occurring within a severe thunderstorm watch. Dot size corresponds to the number of watches issued with each probability. The dashed line represents perfect reliability.

## 2.6 Tornado Probabilities in Tornado and Severe Thunderstorm Watches Discussion

Over the 19 year period from April 28th, 2006 to December 31st, 2024, tornado probabilities in tornado watches are very reliable, with some slight variations across time and geographic region. As seen across all figures and tables, 40% for the probability of two or more tornadoes and 20% for the probability of one or more strong tornadoes have the highest number of watches. These two percentages are seen as the 'default' percentages used by forecasters at the SPC when issuing tornado watches (M. Flournoy, personal communication, June 2025), which may account for the large amount of watches for these two percentages. Specifically for the probability of two or more tornadoes, there are no watches with any percent under 20%. This may be due to the fact that if a tornado watch would have under 20% for the probability of two or more tornadoes, forecasters may instead issue a severe thunderstorm watch in place of a tornado watch. This is reflected in the severe thunderstorm watch results, as the probability of two or more tornadoes is rarely issued above 20% for severe thunderstorm watches. Since 20% for the probability of two or more tornadoes has been issued for both tornado and severe thunderstorm watches, other factors such as the probabilistic graphics in the SPC convective outlook, the probabilities within watches for other hazards, coordination with WFO's, and other environmental indicators may play a role in what kind of severe watch is issued. However, it is important to note that the severity of the watch that is issued impacts the overall POD, performance, and lead time of the tornado warnings issued in these watches, with all of these factors increasing as severity increases from severe thunderstorm watches, tornado watches, and PDS tornado watches (Krocak and Brooks 2021).

The overall reliability of tornado probabilities in tornado watches is very high, with only slight over-forecasting occurring around the low and lower moderate percentages, as seen in Figure 2.6 and 2.7. However, when analyzing the Brier scores and Brier skill scores for the two tornado probabilities, a handful of important results emerge. The best Brier scores and Brier Skill scores occurred for the forecasted probabilities of 95% for the probability of two tornadoes, and

2%, 5%, and 95% for the probability of one or more strong tornadoes, typically indicating that these probabilities are most accurate and have the highest levels of improvement from a forecast that always predicts the climatological event frequency. However, the sample sizes of these forecast probabilities are extremely small, especially for the probability of one or more strong tornadoes. Thus, they should be regarded with caution, as it is difficult to draw conclusions from such a small amount of forecasts and observations. Additionally, while the Brier scores for the remaining forecast probabilities are all at or below 0.25, there is room for these probabilities to be more accurate. This is also true for the forecast values that are considered 'default' by forecasters (40% for the probability of two or more tornadoes and 20% for the probability of one or more strong tornadoes), as they are essentially the climatology or base rate of both tornado probabilities (0.44 and 0.21). Furthermore, all of the Brier skill scores from 30% to 80% for the probability of one or more strong tornadoes are negative, indicating that these probabilities are all less accurate than a forecast that always forecasts the climatological event frequency. A potential reason for these negative Brier skill scores is the large difference in the number of watches issued with 20% (2,524) to the number of watches issued with a 30% to 80% (1,126) for the probability of one or more strong tornadoes. While it can be difficult to ensure a probability outside the climatology will verify in the moment of forecasting it, issuing probabilities above and below the 'default' 40% and 20% for both tornado probabilities would increase the accuracy of 40% and 20%, and thus increase the accuracy of the other forecast values as well. This should only be done, of course, when forecasters are confident in the watch's ability to verify outside of the climatological event frequency.

Once stratified by time and geographic areas, some differences emerge. Typically across these stratifications, we see the low probabilities being over-forecast, and moderate probabilities tend to be very close to the perfect reliability line, but are slightly under-forecast for the probability of two or more tornadoes within a watch. This may be due to the difference in storm environments for low and high probabilities, as forecasters may feel more confident issuing forecasts where the tornado signal is more “obvious.” For instance, the majority of probabilities for one or more strong tornadoes is below 30%, while the probability of two or more tornadoes is much more dispersed across all probabilities. This cluster of probabilities for one or more strong tornadoes may be due to forecasters being more certain about the probability of two or more tornadoes occurring in a tornado watch in comparison to the probability of one or more tornadoes occurring.

When stratified over years, Figure 2.8 shows a slight underforecast for moderate to high probabilities (50-70%) for tornado watches issued between 2018-2024, which changed from previous years. Over time, it seems forecasts may have been slightly over-corrected, going from over-forecast in 2006-2011, to almost perfect reliability in 2012-2017, to then under-forecast in 2018-2024. This may be due to the fact that forecasters may have known these probabilities were slightly overforecast for 2006-2011, and in an attempt to improve reliability or other components of quality, ended up under-forecasting these probabilities. Stratifying by time of day, Figures 2.12 and 2.13, and Tables A.7 and A.8 show that the number of daytime tornado watches was considerably more than the number of nighttime tornado watches, in addition to daytime tornado watches being closer to perfect reliability. Daytime watches also have higher probabilities of one or more strong tornadoes occurring within the watch area. Day time environments are typically more favorable for strong, long track tornadoes compared to at night, hence the higher probabilities of one or more strong tornadoes, in addition to the number of daytime tornadoes being considerably higher than nighttime tornadoes (Davies and Fischer 2009). Forecasters may be more confident in daytime watches in comparison to nighttime, as storm environments during the day may be cleaner and easier to interpret. In addition, nighttime tornado environments are slightly different than daytime environments, as significant nighttime tornadoes typically happen with smaller instability but with substantially larger storm-relative helicity (Davies and Fischer 2009). Similar results are seen for springtime tornado watches in comparison to other seasons (Figure 2.10 and 2.11). The high probabilities for both two or more tornadoes and one or more strong tornadoes is most likely due to more favorable environments for tornadoes in spring compared to other seasons (Anderson-Frey et al. 2016). Forecasters may feel these environments are closer to 'textbook' setups, and have higher confidence in issuing these high probabilities, and thus could explain the near perfect reliability for spring in Figure 2.11.

As for geographic regions, the West and Northeast had the smallest number of tornado watches overall, which is consistent with how infrequently these regions receive severe weather (Anderson-Frey et al. 2016). The South had the largest number of tornado watches, followed by the Great Plains and the Midwest. While the Great Plains is regarded as tornado alley and thus is typically thought of as the area that receives the most tornadoes especially during springtime months (Anderson-Frey et al. 2019; Krocak and Brooks 2018), since 1979 negative tendencies of tornado occurrence have occurred in the Great Plains, with positive trends occurring in the South and Midwest (Gensini and Brooks 2018). Thus, there may be more tornado watches being issued in the South due to this shift in where tornadoes are more frequently occurring. Additionally, the number of tornadoes in a tornado watch and the size of tornado watches can vary greatly, so it may be that the tornado watches issued in the South are smaller than those issued in the Great Plains due to the variability of environments in the South, and thus more tornado watches have to be issued over time to cover a larger geographic area.

In terms of reliability, the South had the largest difference from perfect reliability for the lower-to-moderate probabilities (30% & 40% for the probability of two or more tornadoes) in comparison to the other geographic regions, with these probabilities being over-forecasted. The over-forecasting of these probabilities may be due to the uncertainty forecasters have when forecasting for high shear low-CAPE (HSLC) environments (Davis and Parker 2014; Sherburn and Parker 2014; Sherburn et al. 2016) in the Southeast, in comparison to the textbook high shear high-CAPE environments in the Great Plains. HSLC environments create a challenge for forecasters, especially in the Southeast CONUS. On average, tornadoes in the Southeast have higher shear and lower CAPE than tornadoes in the rest of the CONUS, and probability of detection for this region is better and the false alarm ratio is worse for tornadoes in comparison to the CONUS as a whole (Anderson-Frey et al. 2019). Additionally, the Southeast has almost double the CONUS percentage of deadly tornadoes due to a number of factors, such as a high frequency of nocturnal tornadoes, a disproportionate percentage of long-track tornadoes, low visibility of tornadoes in forested regions, (Coleman and Dixon 2014; Davis and Parker 2014; Sherburn and Parker 2014; Anderson-Frey et al. 2016; Sherburn et al. 2016; Krocak and Brooks 2018; Anderson-Frey et al. 2019), and social risk factors such as a high population of elderly residents, and high concentration of mobile homes (Ashley et al. 2008; Sherman-Morris 2013; Wallace et al. 2015). However, those tornadoes with the lowest POD are, on average, the tornadoes that are the weakest and least likely to be deadly (Anderson-Frey et al. 2019). While the South region utilized in this analysis is larger than the Southeast region used in Anderson-Frey et al. 2019, the South does include the Southeastern states, therefore the HSLC environments play a large role in this region. Additionally, the South region includes coastal states, such as Florida, where tropical cyclones frequently occur. This adds a further layer of complexity to forecasting tornadoes in the South, and could also explain over forecasting of the low-to-moderate probabilities. The combination of HSLC environments, tropical cyclones, and other factors mentioned above likely make forecasting the probability of tornadoes in watches a larger challenge in the South than other regions, especially with lower-end storms such as those being issued 30% and 40% for the probability of two or more tornadoes in a tornado watch.

Despite small variations across time and area, the tornado probabilities issued with tornado watches are very reliable and generally accurate, and are a good indication of the chance of two or more tornadoes and one or more strong tornadoes occurring within a tornado watch. While there are some cases within the 19 year timeframe that over or under verify, watch probabilities are on the whole very reliable. However, forecasters should consider attempting to stray away from relying on 'default' probabilities, and use the other probability values to try and improve the accuracy of the probability of one or more strong tornadoes. Additionally, the methods used to perform verification in this study is just one of the ways to verify the probabilities in tornado watches. Other methods, such as calculating the probability of detection and success ratio, should be used to continue to evaluate the quality of probabilities inside tornado watches, in addition to furthering this analysis by stratifying other variables, such as lead time, varying storm environments, and watch type (PDS versus non-PDS).

# Chapter 3 - Emergency Managers & Tornado Watch Probabilities

## 3.1 WxEM Survey Data & Methods

### 3.1.1 Survey Data Collection

Data was collected through Wave 9 of the Extreme Weather and Emergency Management Survey (WxEM), using a nationwide panel of EMs. WxEM was developed by the University of Oklahoma's Institute for Public Policy Research and Analysis (IPPRA) to field routine surveys to understand how EMs receive, process and use forecast information, with a specific focus on NWS operations and forecasts (Wanless et al. 2023). Wave 9 of WxEM was administered to the nationwide panel of EMs through email, which was fielded from November 3rd to December 5th, 2025. Unlike other surveys administered to the public, WxEM does not collect a representative sample of participants, as there is no complete list of EM demographic information to collect a representative sample. Information regarding each EM is collected as they join the panel that WxEM is administered to, however this information is not included in the data or analysis to keep each EMs responses anonymous. A total of 236 EMs across the US completed Wave 9 of WxEM (Figure 3.1).

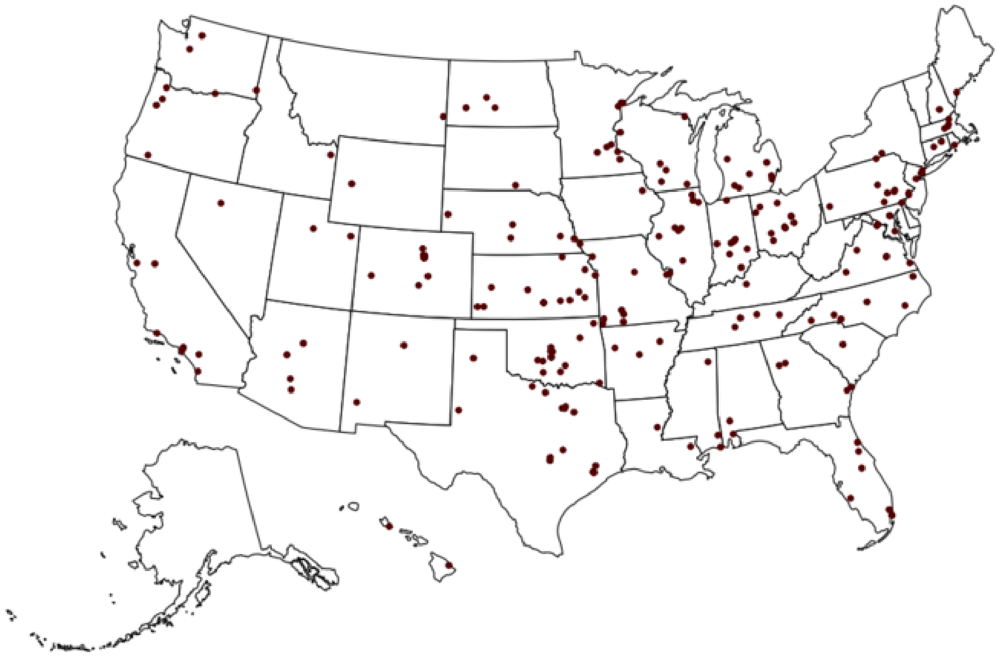


Figure 3.1: Nationwide distribution of emergency managers that completed wave 9 of the WxEM.

Wave 9 of the WxEM included three overarching themes of questions: the day 4 through 8 SPC convective outlook, tornado probabilities associated with tornado watches, and the experimental tornado warning system, Threats In Motion. For the purpose of this study, only questions regarding the tornado probabilities in tornado watches will be used for analysis. These questions were developed to ultimately gain an understanding of if, and how EMs use these probabilities in their operations. As shown in Appendix B, participants were first shown text detailing that the SPC issues tornado watches in addition to convective outlooks, and then described three pieces of tornado watches: the graphical component, probabilistic information, and the text narrative. Along with this information, EMs were shown a graphic with these three components to match the description given. The initial questions in this section asked participants how frequently they used tornado watches in their operations, and on average how frequently they received tornado watches in a year. These questions were used to understand which EMs were in areas that did not regularly receive tornado watches, and thus would not use the probabilistic information in tornado watches. If participants stated they used tornado watches more than ‘not at all’, they were then asked which sources and channels they use to receive tornado watch information, and were then shown a graphic of three components of tornado watches: the graphical, textual, and probabilistic information component. Participants were then asked how frequently they used each of these components to understand if EMs are using the probabilistic information in tornado watches, and how that compared with the graphical and textual components.

Participants that stated they did not use tornado watches at all in their operations were then asked why they did not use probabilistic information, and if they stated they were unaware of the probabilities prior to the survey, were asked how likely they were to use that information in the future. These questions were included to understand why EMs did not use probabilistic information in tornado watches, and if they would see this information as helpful moving forward.

Additional questions were asked to understand how important and accurate probabilistic information in tornado watches is to EMs, and to understand how this information is used in their operations. After asking how they use probabilistic information in tornado watches, participants were asked both how high the probability of two or more tornadoes and one or more strong tornadoes would have to be to take the actions they described previously, and if one probability was used more frequently than the other to take action.

Next, participants' subjective and objective understanding of tornado probabilities in tornado watches was collected. For the subjective measure, participants were asked to rate their understanding of the probabilistic information in tornado watches on a 5-point Likert scale from poor to excellent. For the objective measure, participants were first shown one of three randomized images of a tornado watch with both tornado probabilities at Low (20%), Moderate (50%), or High (80%), then asked to choose the correct definition, or reference class, for probabilities in tornado watches. Participants were also asked which actions they would take if the randomized tornado watch and associated probabilities were issued for their jurisdiction.

### 3.1.2 Survey Data Analysis

After data for wave 9 of WxEM was collected, an analysis was conducted using the statistical programming language R to create graphics and gain an understanding of the data. The initial analysis included understanding general sample sizes, and creating bar plots of each question to see which questions yielded interesting results. Comparison plots were also created between questions that were in relation to each other, such as the questions that asked how frequently EMs used tornado watches in their operations, and how frequently their jurisdictions received tornado watches. A mean reliance plot was created for the questions relating to which components of tornado watches EMs used in their operations to understand which pieces of tornado watches EMs used the most, and how much other information was used in relation to probabilistic information. The experiment questions were analyzed and plotted by the randomized tornado probabilities each participant received, to understand if actions and objective understanding of the probabilities varied by threat level.

Participants were then asked an open-ended question about how they use the probabilistic information in tornado watches. To analyze these data, a directed content analysis (Hsieh and Shannon 2005; Krippendorff 2019) was performed on responses. This process was done deductively, where responses were manually sorted into categories based on the main topics of the response posed by the participants. Categories were chosen following the response options from previous WxEM surveys: increase weather monitoring, activate the EOC, share this information with partners, share this information with the public, increase staffing, contact my local National Weather Service office, seek out information from other weather sources, and other. After an initial review of the responses, multiple other categories were added, including: understanding the impacts and/or protective actions, awareness/preparedness of resources, probabilistic thresholds, no actions, and increase situational awareness. A description of each category and an example response are shown in Table 3.1.

| Category title | Category description | Example responses |
|---|---|---|
| Increase weather monitoring | EM pays more attention to the weather, such as watching radar, or looking for more information | "With the issuance of a tornado watch, monitoring of regional radar begins and closer monitoring of NWS Chat for local offices…"<br>"Makes me weather aware and I start watching all the products I have access to." |
| Activate/Update EOC | EM upgrades the EOC level, or fully activates the EOC | "The EOC goes to a level 3 status and put out a code red to the awareness level people"<br>"I utilize it as a gauge compared to previous tornado watch probabilistic information to determine the possibility of activating our EOC." |
| Share information with the public | EM shares the probabilistic information, in some way, with the public | "It provides more information so I can make decisions on our protective posture and messaging to the public."<br>"Depending on the level, I would notify the community of the potential for a storm and send some graphics that |

| | | |
|---|---|---|
| | | include safety measures” |
| Share information with partners | EM shares the probabilistic information, in some way, with their partners | “We forward the information through our channels to local responders and the community”<br>“Depends on any events we might be having, I would watch more closely to this and relay the information to any event organizers.<br>For staff I would make sure the front line staff is aware. Especially public works” |
| Contact local NWS office | EM reaches out to their local NWS office through any means | “Escalete EOC to "Monitoring" level, actively follow NWSChat, and products from local NWS office.”<br>“With the issuance of a tornado watch, monitoring of regional radar begins and closer monitoring of NWS Chat for local offices.” |
| Seek out information from other weather sources | EM looks for more information from other weather sources besides the tornado watch | “ I spend more time looking at area forecast discussions and hazardous weather outlook.”<br>“Makes me weather aware and I start watching all the products I have access to.” |
| Increase staffing | EM decides to increase staffing, or prepare to increase staffing | “I do, I make sure I have an extra person available at the 911 Center. Depending on how high the probability is, I may have two extra persons on duty.”<br>“... Having this information starts a set of triggers in our daily staffing and response levels.” |
| Understanding impacts/protective actions | EM tries to understand the potential impacts to their | “We share with leadership to prep for staff impacts, |

| | | |
|---|---|---|
| | jurisdiction or the necessary protective actions to be taken | building and grounds impacts and to assure they've taken precautions to protect their homes and families”<br>“In short, these probabilities help us gauge not just where severe weather might occur, but how likely significant impacts are allowing us to tailor our preparedness and response accordingly.” |
| Awareness/Preparedness of resources | EM takes an inventory of the resources available to them, or prepares their resources for the potential of a tornado | “...We can also prepare county resources to assist community in case of an emergency.”<br>“...if we need to predeploy/request specific equipment” |
| Probabilistic thresholds | EM takes different actions depending on how high the tornado probabilities are | “As the probability level increases so does my EOC response, IE: increase monitoring, potential pre activation, additional community preparedness messaging ahead of the storm system”<br>“I use this information to convey the severity of my message on social media. The higher the probability, the greater the tone of my messaging has to be. I may also open an incident in our damage reporting software and post that on social media, as well.” |
| No actions | EM does not take any action, or previously determined action, with the probabilistic information | “no change I would amplify messages from NWS”<br>“I do not take any previously taken actions based on the probabilistic information, but it does help with decision-making about |

| | | |
|---|---|---|
| | | staying in the office later in the evenings.” |
| Increase situational awareness | EM heightens their situational awareness | “Situational awareness greatly increase and a more intense focus on weather monitoring starts to occur.” “No real change...alerts and Hyper vigilance” |
| Other | Response was too vague or unclear, or did not fit into another category | “Specific actions and forecasting” “Determining who to pass the information to, promotion of actions to take” “I take probabilistic information in account when looking at seasonal weather patterns and historical weather events. If things match very .heavily, then I should look at making the proper preparations for same conditions” |

Table 3.1: The thirteen categories used in this study, including the category title (left), a short description (middle), and an example of responses (right).

# 3.2 WxEM Survey Results

## 3.2.1 Descriptive Question Results

First, EMs were asked to answer if they currently use tornado watches in their operations, and were given a likert scale from “not at all” to “extensively”. Results show 42.7% of respondents said they used tornado watches “very much” in their operations, and 22.0% said they use watches “extensively” (Figure 3.2). Additionally, 17.2%, 11.6%, and 6.5% answered “somewhat”, “not very much”, and “not at all” as their responses, respectively. EMs were then asked to select the number of tornado watches their jurisdiction receives in an average year, with

28.4% selecting “between 3 and 5”, 23.3% selecting “between 6 and 10”, and 19.8% selecting “more than 10”. Only 10.8% of EMs selected “0” tornado watches, and 17.7% selected “between 1 and 2” (Figure 3.3). To compare the results between these two questions, Figure 3.4 shows that of all the EMs that said they do not use tornado watches at all in their operations, 100% of them said they receive zero tornado watches in an average year within their jurisdiction. Conversely, of the number of EMs that said they use tornado watches extensively in their operations, 33% said they receive 10 or more tornado watches in an average year. Additionally, the percentage of EMs that receive 10 or more tornado watches a year increases from “not very much” to “extensively”, from 11% to 33%. The percentage of EMs that selected “between 1 and 2” tornado watches decreased from “not very much” to “extensively”, from 41% to 14%. Of the EMs that said they received “between 3 and 5”, and “6 and 10” tornado watches in their jurisdiction, the highest percentage for these groups were “somewhat” with 38%, and “very much” with 33%, respectively. Overall, those participants that indicated they used tornado watches frequently also receive tornado watches frequently within their jurisdiction.

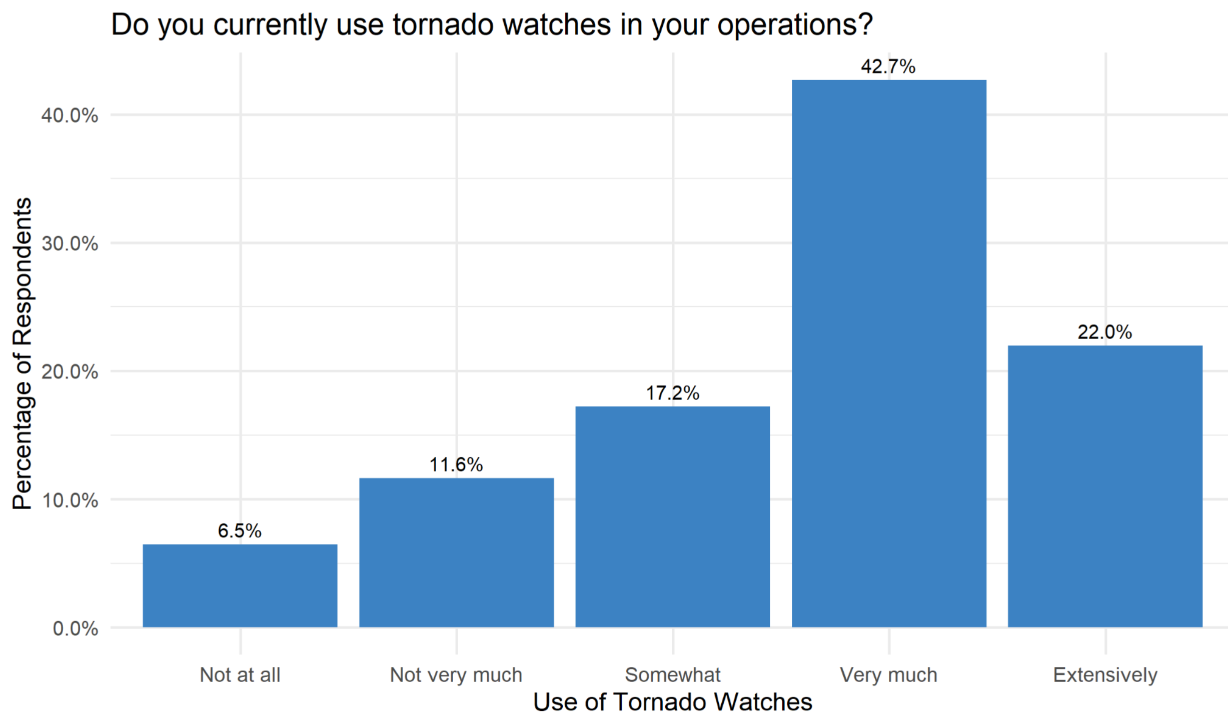



Figure 3.2: Plot of the use of tornado watches by EMs in their operations

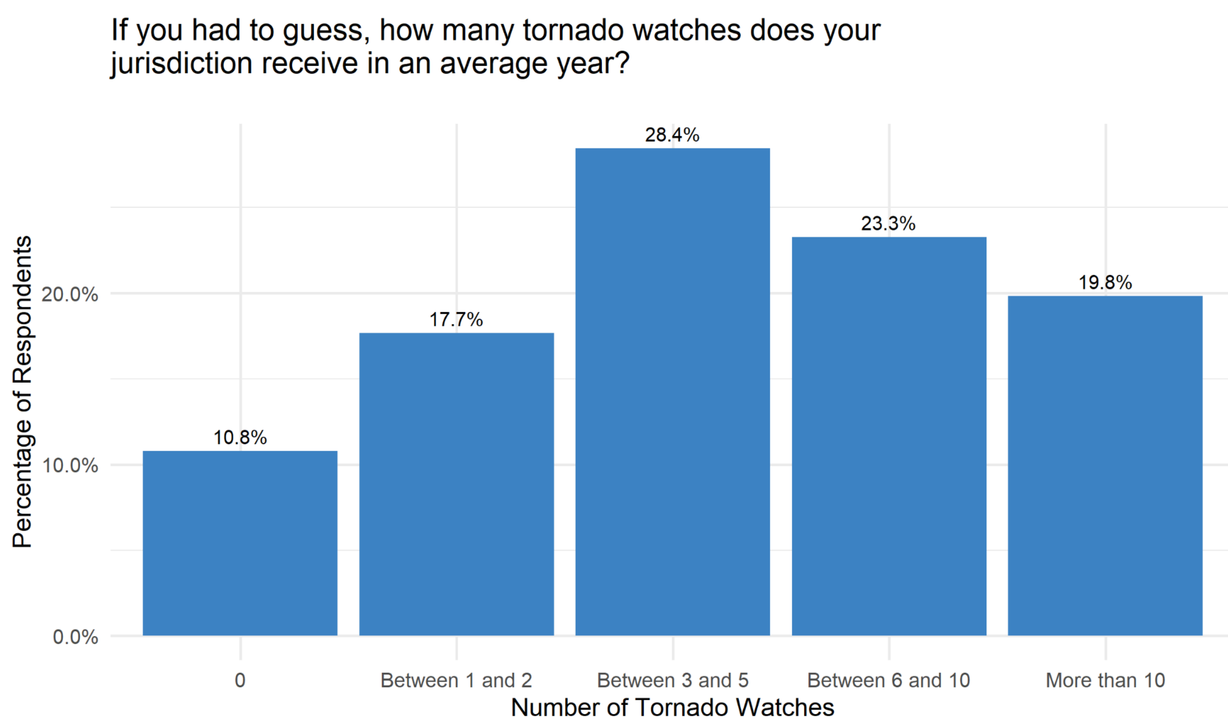


Figure 3.3: Plot of how many tornado watches EMs receive in their jurisdiction in an average year

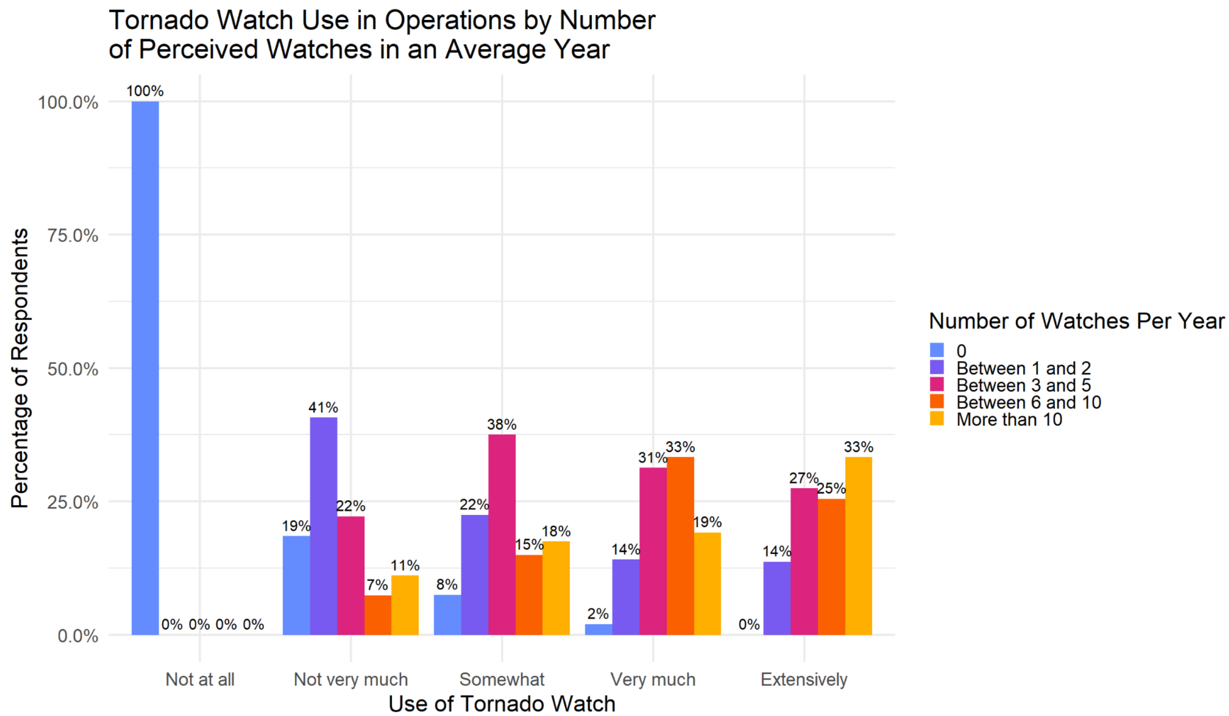



Figure 3.4: Plot of tornado watch use in EM operations by the number of perceived tornado watches in an average year

Survey participants were then asked to identify the primary sources they use for receiving tornado watches. Respondents were given seven options to choose from, as well as an "other" option, and were asked to select all that applied. The most common selection by far was "local NWS offices", with 90.2% of EMs selecting this option as their source (Figure 3.5). Dropping roughly 35%, The "SPC" was the next most common source (56.2%), followed by "local TV stations" (34.5%), and "private companies" (26.0%). 15.3% of participants selected "national TV stations", and 11.9% of EMs indicated they use "other emergency managers" as a source for tornado watches. 9.4% of EMs selected "something else", with many of the responses being either iNWS or FEMA applications. Finally, 7.7% of respondents selected an “in-house meteorologist” as their source.

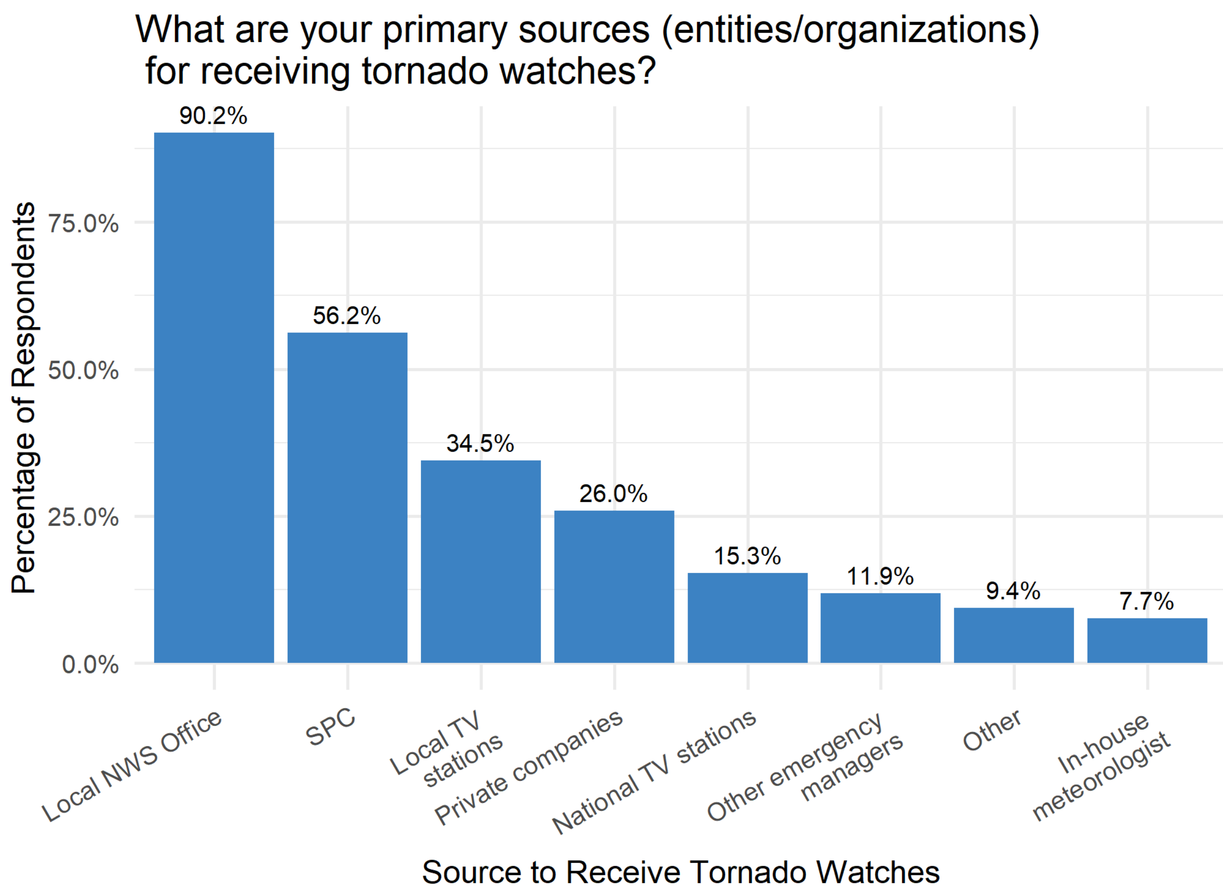



Figure 3.5: Plot of the primary sources (entities/organizations) EMs use for receiving tornado watches

Following sources for receiving tornado watches, EMs were then asked to identify all channels they use to receive tornado watches given eleven options to choose from, including a “something else” option. As seen in Figure 3.6, the most common channel that EMs use to receive tornado watches is “NWS Chat” (76.2%), followed by “cellphone applications” (59.1%) and “weather focused websites”, such as weather.gov, or weather.com (57.0%). After the top three most common channels, 47.2% of participants selected “weather radio”, followed by a drop to 30.6% for “social media”, 29.4% for “television”, and 16.6% for “word-of-mouth”. 11.1% of participants chose “something else”, with many of these responses indicating they use iNWS or other mass notification systems. Finally, “raw weather observations or model data” (9.4%), and “amateur/ham radio” (3.8%) were the least two common channels selected by EMs as a way of receiving tornado watch information.

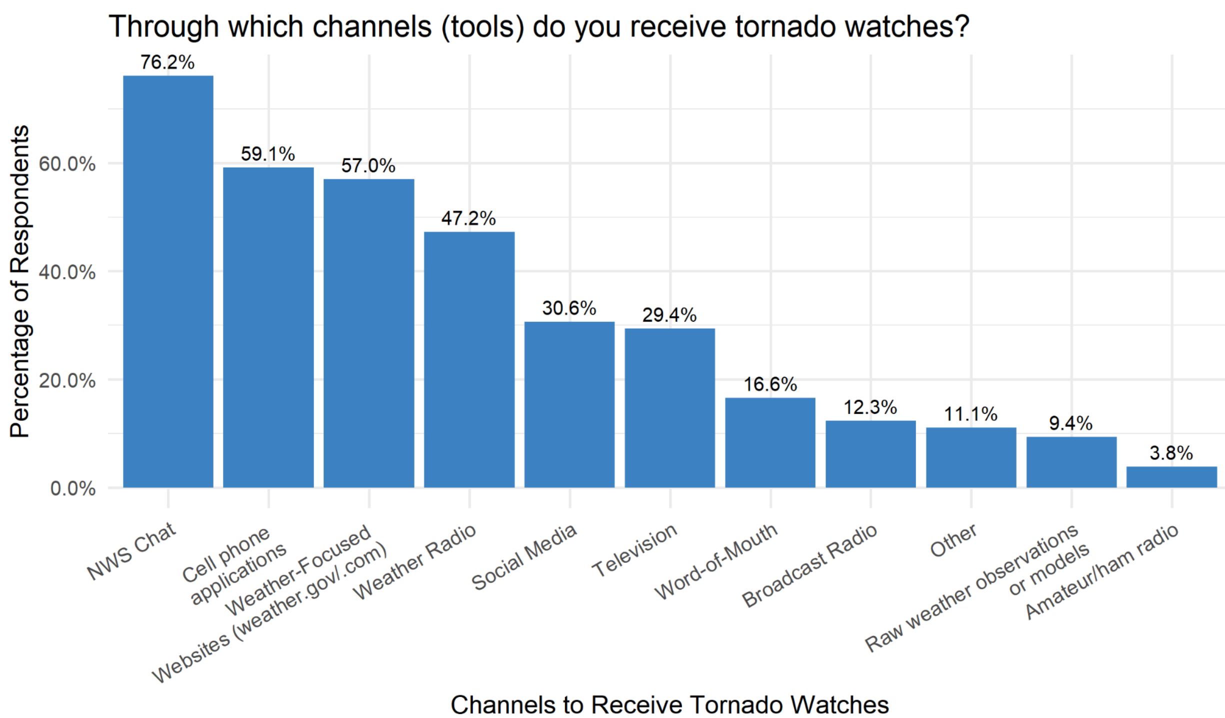



Figure 3.6: Plot of the channels (tools) EMs use for receiving tornado watches

Participants were then asked how much they use the graphical component, probabilistic information, and text narrative of tornado watches, as seen in Appendix B. Comparing the mean rating of these three components in Figure 3.7, tornado watch graphics were the most used, with a mean rating of 4.12. The graphical component was followed by the text narrative with a mean of 3.77, and finally probabilistic information, with a mean rating of 3.59. If participants indicated that they did not use probabilistic information at all, they were directed to answer why they did not use this information. Of the eight participants that selected this option, four said they were hard to interpret, three said they were not aware they existed, and one said they did not know where to find them.

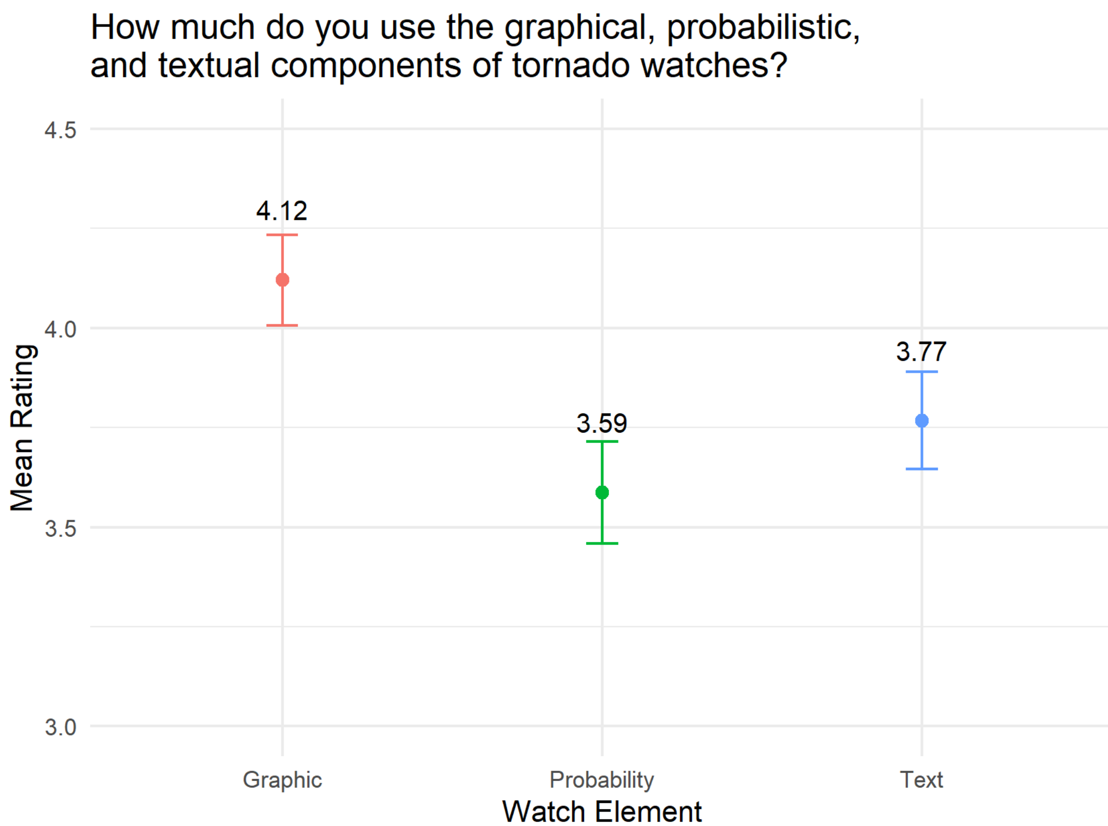



Figure 3.7: Mean use of the graphical, probabilistic, and textual components of tornado watches.

The participants that indicated that they use the probabilistic information in tornado watches were then asked how high the probability of 2 or more tornadoes would have to be for them to take action. As seen in Figure 3.8, 16.9% said they did not have a threshold, and 15.9% were "not sure", or it would depend on the situation. For those participants that chose a numeric threshold, 6.5% selected "20%", 10.9% selected "30%", 12.4% selected "40%", and 13.4% selected "50%". After "50%" of the probability of 2 or more tornadoes, the percentage of respondents dropped to 5.0% for "60%", before increasing to 6.0% for "70%", and then dropping back to 4.5% for "80%", 1.0% for "90%" and 1.5% for "95%". In comparison to the probability of one or more strong tornadoes in Figure 3.9, the “no threshold” and the “not sure/it depends” selections were at 17.3% and 14.4% respectively. Instead of steadily increasing to “50%” and then dropping after, “20%” jumps to 11.4% before dropping to 7.4% for “30%”, and increasing up to 10.9% for “40%” and 13.4% for “50%”. “60%” was also higher in Figure 3.9 compared to Figure 3.9, with 7.9% of respondents selecting “60%” as their threshold to act.

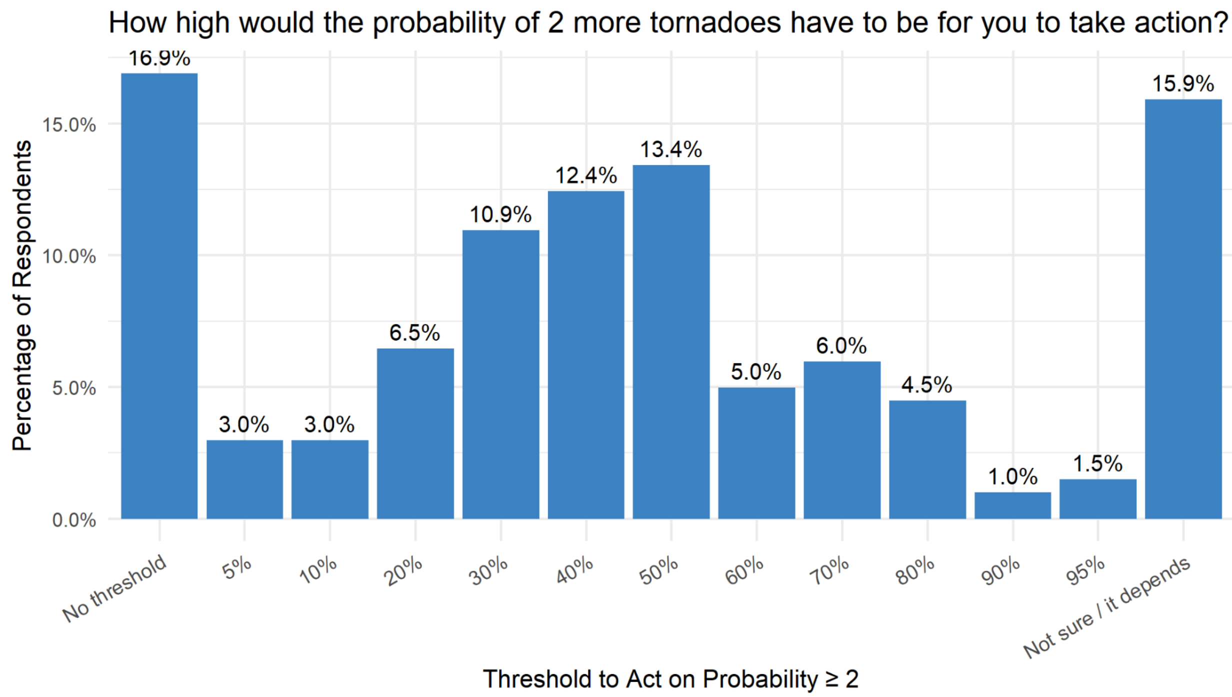



Figure 3.8: The threshold an EM would use before taking the action for the probability of 2 or more tornadoes within a tornado watch.

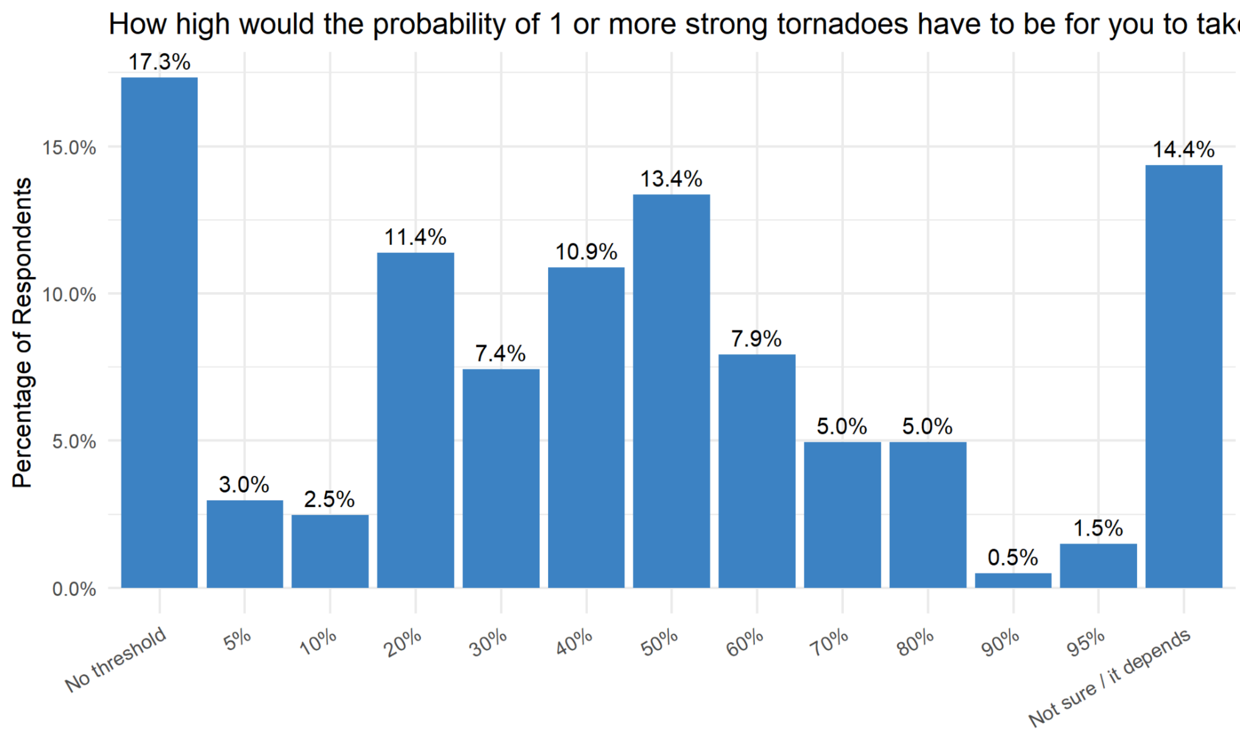



Figure 3.9: The threshold an EM would use before taking the actions for the probability of 1 or more strong tornadoes within a tornado watch.

EMs were also asked which tornado probability they use more frequently, the probability of two or more tornadoes, or the probability of one or more strong tornadoes. As seen in Figure 3.10, the majority of EMs stated they use both probabilities equally (61.4%), with 9.2% indicating they were not sure. Of those respondents that indicated they used one probability more than the other, 8.7% stated they used the probability of two or more tornadoes more, and 20.8% indicated they used the probability of one or more strong tornadoes more. Additionally, participants were asked to rate their understanding of the probabilistic information included in tornado watches, or their subjective understanding. Figure 3.11 illustrates a near perfect bell curve of responses, with 7.8% of respondents rating their understanding as “poor”, 23.8%

selecting “fair”, and 35.9% choosing “good”. 26.4% of EMs said their understanding was “very good”, with 6.1% saying their understanding was “excellent”.

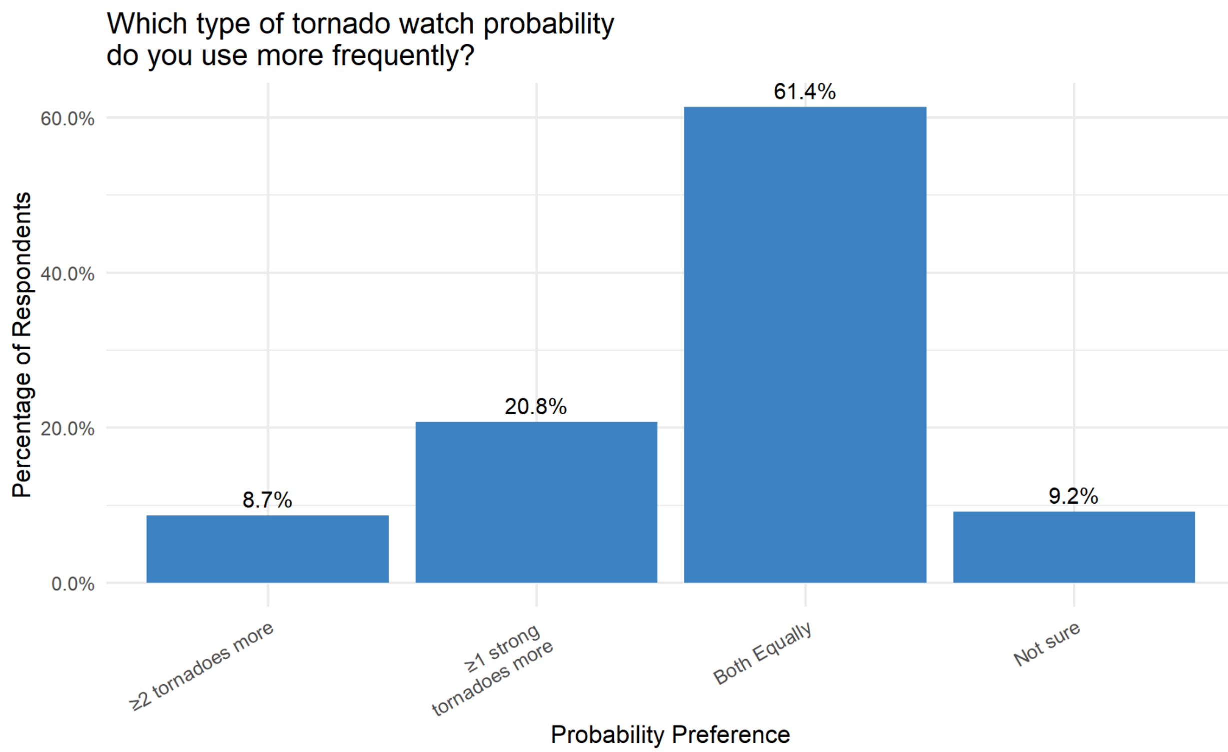



Figure 3.10: EMs preference for the two tornado probabilities in tornado watches.

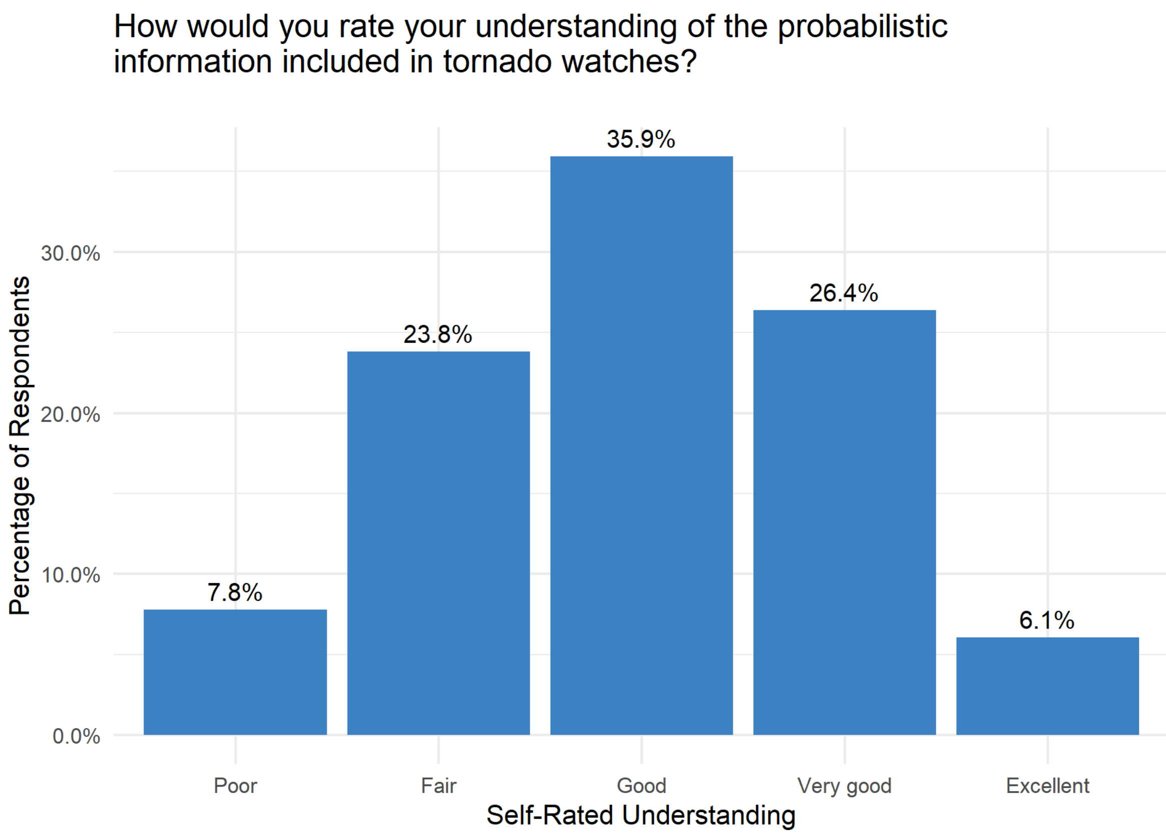



Figure 3.11: EMs subjective understanding of the probabilistic information included in tornado watches.

Of the EMs that indicated they receive one or more tornado watches in the jurisdiction in an average year, they were asked about their perceived importance of the probabilistic information for protecting life and property in their jurisdiction. As seen in Figure 3.12, only 2.2% of respondents indicated they were "not at all important", while 14.3% said they were "slightly important". There was a large increase from "slightly important" to "moderately important" (33.8%) and "very important" (37.7%), with a drop back down to 12.1% for "extremely important". Finally, EMs were asked to indicate perceived accuracy of the probabilistic information included in tornado watches for their area. 20.9% of EMs indicated they were "not sure" what the accuracy of probabilistic information in tornado watches was for their area, and only 0.5% said they were "not at all accurate" (Figure 3.13). 17.0% indicated they were "slightly accurate", with almost the majority of EMs saying they were "moderately accurate" (48.1%). "Very accurate" dropped back down to 12.6%, and only 1.0% of EMs said that probabilistic information in tornado watches were "extremely accurate" for their area.

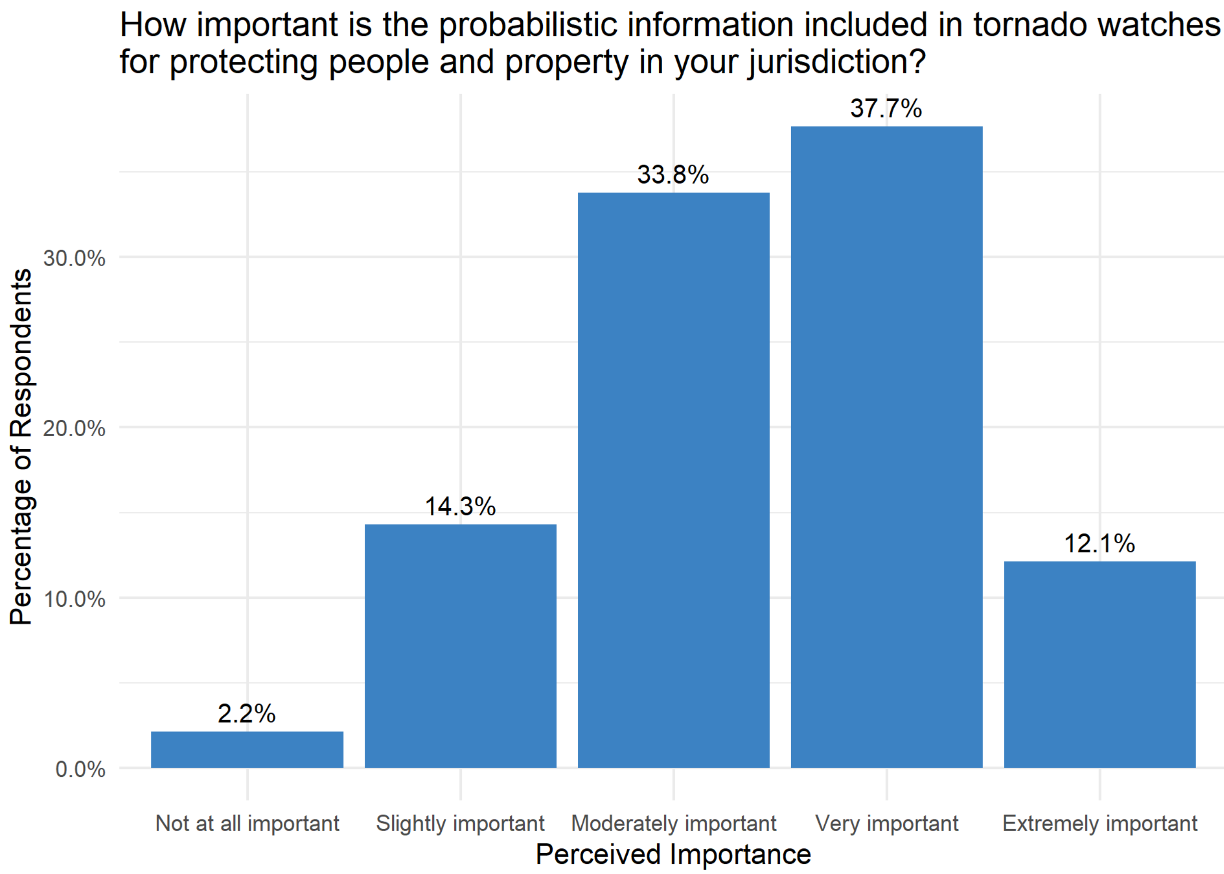



Figure 3.12: EMs perceived importance of the probabilistic information included in tornado watches to protect life and property in their jurisdiction.

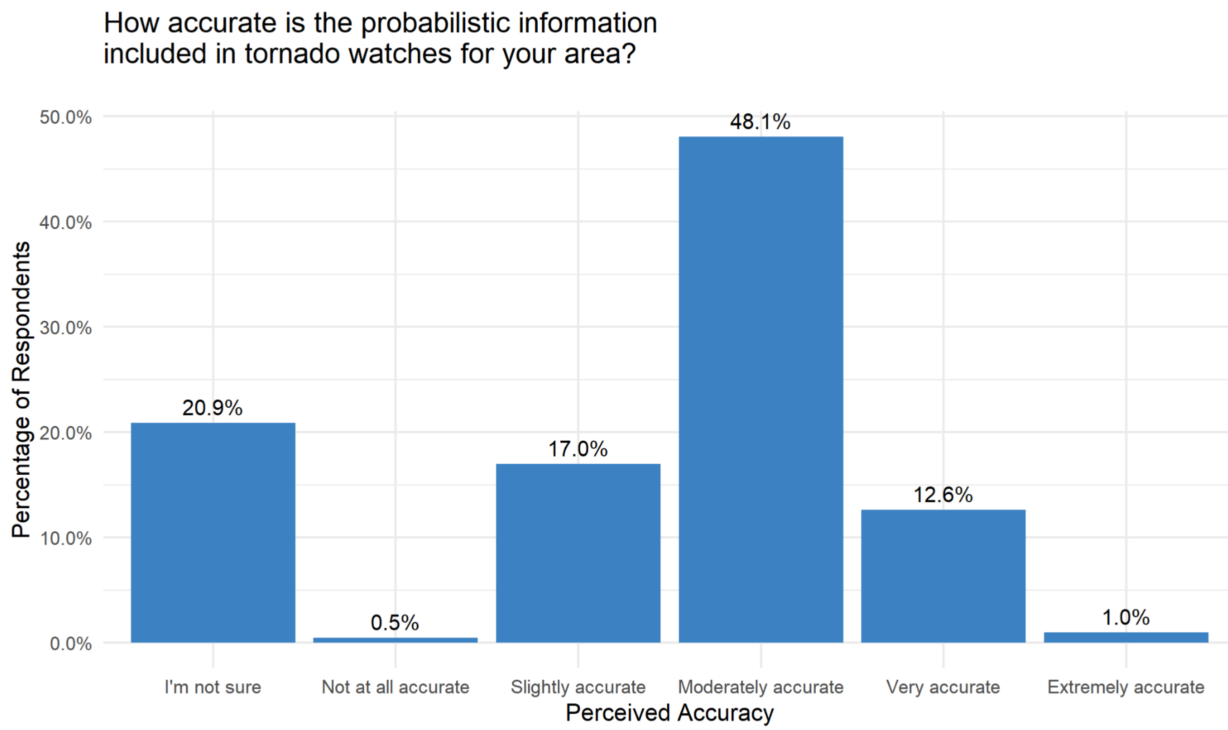



Figure 3.13: EMs perceived accuracy of the probabilistic information included in tornado watches for their area.

### 3.2.2 Verbatim Question Results

After participants indicated that they use probabilistic information in tornado watches, they were then asked how they use this information, such as taking specific actions in response to different probability thresholds. Participants were given an open text box to input their answers, and then researchers categorized the text in the responses. Figure 3.14 shows these categories, where 30.6% of respondents indicated that they share this probabilistic information with their partners within their jurisdiction. Many of these responses mentioned sharing this information with city and county staff, event organizers, and first responders (such as firefighters, police, and 911 centers). Many respondents (24.7%) also specified sharing the probabilistic information with the public, through social media or other local channels. Some participants indicated the information they shared to their partners or the public was not in the original format, such as "...We also share much of the information, in some form, with the public."or " I forward this information along to all partners in language they can all understand." These responses sometimes mention that the public does not understand probability, and thus the information is altered for the public to understand the message, such as "We relay the information and try to make it digestible. The general public has difficulty understanding probability…" or "...Occasionally I will use that language in social media posts or stakeholder briefings, though sometimes probabilistic language can be confusing to the general public.".

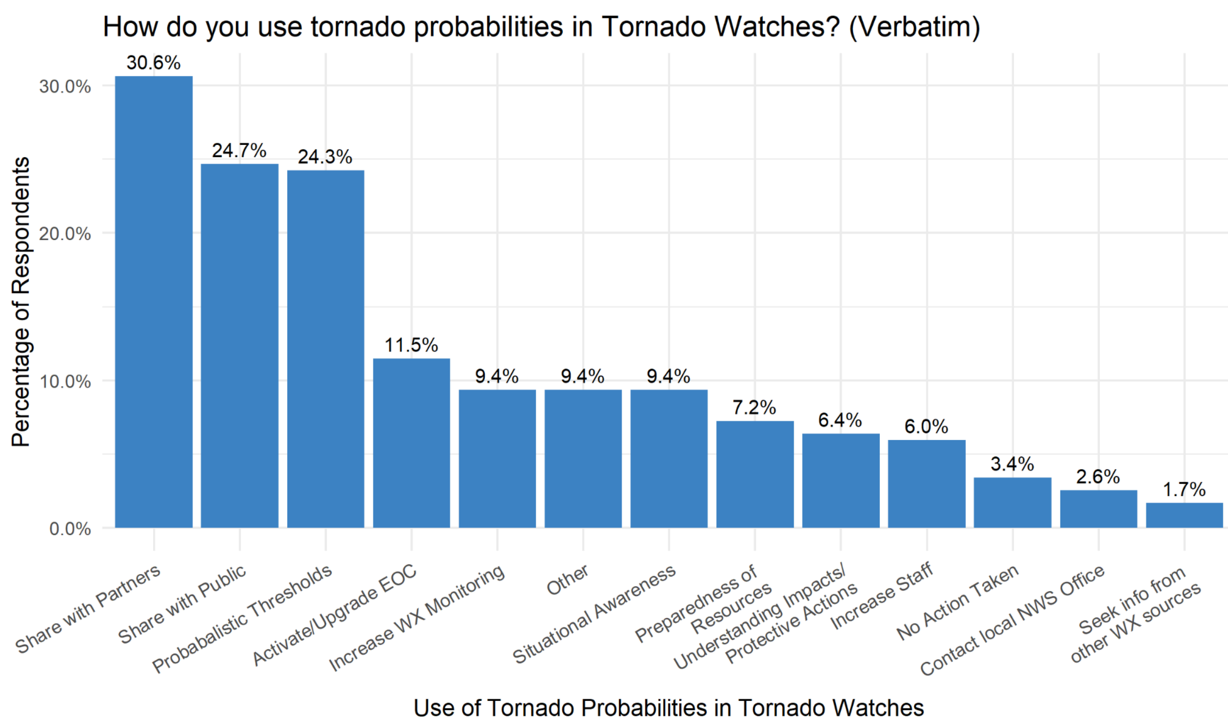



Figure 3.14: EMs use of tornado probabilities in tornado watches in their operations by verbatim responses.

Following sharing probabilistic information with the public and with partners, 24.3% of EMs indicated they would take different actions based on the severity of the probabilities (Figure 3.14). Several participants mentioned having different responses for lower or higher probabilities, such as “I use this information to convey the severity of my message on social media. The higher the probability, the greater the tone of my messaging has to be.” or “We use the probabilistic information included in tornado watches to better assess the level of risk within the watch area. Specifically, we look at the probability values (such as the likelihood of two or more tornadoes, or the chance of a strong EF2+ tornado) to help prioritize monitoring and response readiness. Higher probabilities prompt closer coordination with local agencies, increased staffing or alertness in operations centers, and more frequent communication with field personnel.”. Many of these responses mentioned taking actions for higher probabilities, but did not include actions for lower probabilities.

After the 24.3% of respondents who used probabilistic thresholds in their operations, 11.5% of EMs indicated they use the probabilistic information to potentially upgrade or activate their EOC (Figure 3.14). Several of these responses also mentioned the varying probability levels, with one EM stating “Based upon the different probability levels, we activate the EOC and send out storm spotters.”, and another saying “Higher probability levels increase my chances of distributing educational content about the upcoming weather events. They also impact my decision of how intensively to monitor the weather and whether to open the EOC.”. Additionally, 9.4% of EMs indicated they use the probabilistic information to increase their weather monitoring or increase their situational awareness, or were placed into the ‘other’ category. Many of the responses for the ‘other’ category included responses that were too vague, such as “Specific actions and forecasting”, or “Probability helps to dictate my response.” Increasing weather monitoring had many straightforward responses, such as “Makes me weather aware and I start watching all the products I have access to.”, with some responses including weather monitoring as an action for higher probabilities being issued. This was also the case for those that indicated they used the probabilities for situational awareness, as many responses said they generally use the probabilities just for situational awareness.

7.2% of respondents indicated they will use the probabilistic information to take an inventory of resources, with many responses mentioning having resources prepared at this stage. Following inventory of resources, 6.4% and 6.0% of EMs said they will use probabilistic information to understand the potential for impacts and necessary protective actions for their community, and increase staffing, respectively. Additionally, 3.4% of EMs said they would not take any action based on the probabilities, 2.6% said they would contact their local NWS office for additional information, and 1.7% said they would look for additional information from other weather sources.

### 3.2.3 Experiment Question Results

Following the descriptive and verbatim survey questions, two experiment questions were asked to participants to understand their objective understanding of tornado probabilities in tornado watches, and to further understand the actions EMs take with this probabilistic information. For these two questions, participants were shown one of three randomized tornado watches, with the probabilistic tornado information either being 20% (Low), 50% (Moderate), or 80% (High). As seen in Figure 3.15, the majority of EMs selected the correct reference class for the probabilistic information, "20/50/80% of two or more tornadoes occurring within the watch". However, 41% of respondents that were given 20% for the probability of two or more tornadoes selected the reference class of the SPC convective outlook, which is "the percent chance of the event occurring within 25 miles of a point". This was also seen for the 50% and 80% questions, with 43% of respondents selecting this option for 50%, and 37% for 80%. Additionally, looking at the actions EMs said they would take based on the assigned probability, "monitoring the weather" and "sharing the information with partners" increases from 29% to 30% to 31%, and 27%, to 28% to 29% for 20%, 50% , and 80% of the probability of two or more tornadoes respectively (Figure 3.16). "Contacting their local NWS office" and "seeking information from other weather sources" also slightly increases from 20% to 50% to 80% for the probability of two or more tornadoes, with these actions increasing from 12% and 13%, to 15% and to 17%. "Activating the EOC" increases more than these actions, with 20% starting at 5% and 50% going to 6% of EMs indicating they would take this action, and then jumping to 10% for 80%. Similarly, "increasing staffing" doubles from 4% of respondents for 20% and 50%, to 8% for an 80% chance of two or more tornadoes occurring within the tornado watch.

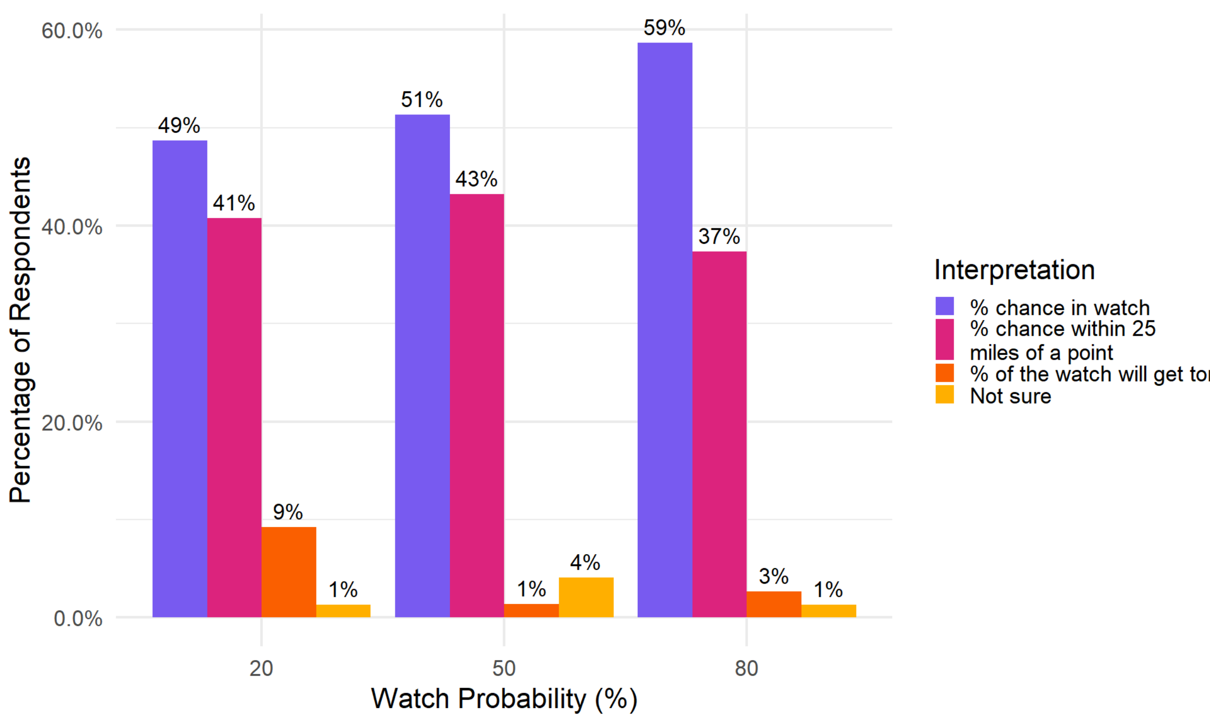



Figure 3.15: EMs objective understanding of tornado probabilities in tornado watches.

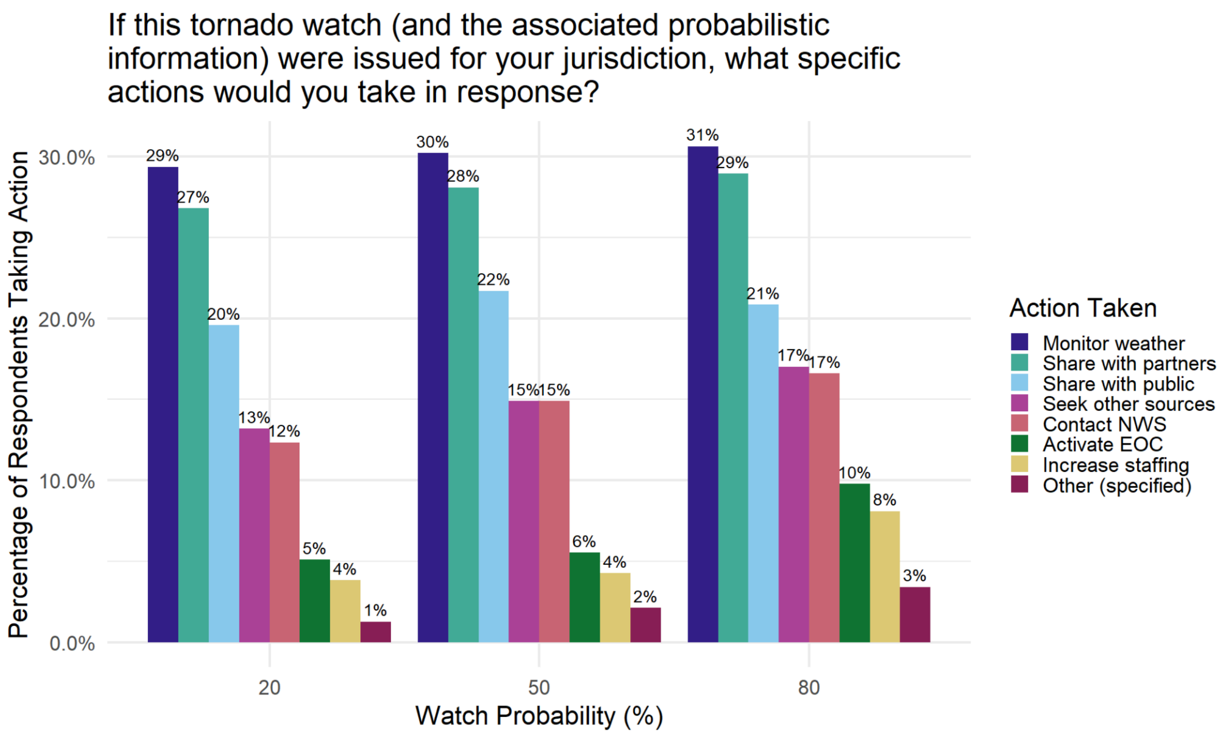



Figure 3.16: EMs actions taken based on a tornado watch and the associated probabilistic information being issued for their jurisdiction.

## 3.3 Survey Discussion

Based on the responses from the WxEM survey, the majority of EMs use tornado watches in their operations, and the more tornado watches EMs said they receive in a year, the more frequently they used them in their operations. EMs most frequently relied on their local NWS office to receive tornado watch information, which is also consistent with the most relied upon source for hazardous weather in general (Hogg et al. 2026). Additionally, NWSChat, cellphone applications, and weather focused websites were the channels EMs use to obtain tornado watch information most frequently. While the order of the remainder of sources and all of the channels used by EMs vary in comparison to Hogg et al. 2026,  this does show the nuance of the different sources and channels EMs rely upon for differing extreme weather events. Additionally, these results were consistent with previous studies that show that rather than relying on one source or channel for forecast information, EMs use a broad array of sources and channels to receive this information (Hogg et al. 2026; Olson et al. 2023; Ernst et al. 2018).

Out of the three components of a tornado watch that were tested, EMs indicated they relied on graphical information the most, followed by the text component and then probabilistic information. Given that tornado watches are a spatial product, the graphical component is commonly used to portray these watches, and thus EMs using this component the most is not surprising. Despite the probabilistic information being the least used out of the three components tested, it was not far behind the text information, and is still regularly being used by EMs. Additionally, EMs are not just using these probabilities in their operations, but the majority (83.6%) of EMs deemed the probabilistic information included in tornado watches “moderately important” or higher for protecting people and property within their jurisdiction, further illustrating the importance of including probabilistic information in forecasts for EMs.

Of the EMs that indicated they used this probabilistic information in their operations, a handful of responses stated that the general public does not understand probabilistic information overall, and thus EMs altered the forecast information before sharing it with the public. While this was once thought to be true, research in the recent past has shown that including probabilistic information in forecasts can improve trust and understanding of the forecast in the public, as long as the reference class is communicated (Morss et al. 2008; Joslyn and LeClerc 2013, 2012; Ripberger et al. 2022). Thus, continuing to advocate for specific information about the uncertainty of a forecast will continue is essential to improve understanding and decision making.

Furthermore, many EMs in the verbatim question specifically mentioned taking more or different actions when the probability values in tornado watches were higher, even if those thresholds were not previously determined. EMs using probabilistic information as thresholds for decision making has been previously shown in Ernst et al. 2018, where probabilistic information helped EMs “identify when threat levels passed either their institutional or personal thresholds for when to begin preparatory actions”. Thus, continuing to include probabilistic information in forecast information sent to EMs will allow them to have a better understanding of the forecasters confidence in a specific event, and therefore allow EMs to make more informed decisions prior to a severe weather event. Additionally, when comparing the two questions that asked how EMs use the probabilistic information from tornado watches in their operations, both results show that EMs take higher level actions as the probability value increases. This increase in action further shows the value of probabilistic information to EMs, allowing them to take greater action as they feel the severe weather will be more impactful to their jurisdiction.

For EMs that indicated they used a threshold for taking action from the probabilistic information, the threshold values are much lower for the probability of one or more strong tornadoes in comparison to the probability of two or more tornadoes. This may be due to EMs considering the probability of one or more strong tornadoes to be more impactful to their jurisdiction than the probability of two or more tornadoes, especially since EMs are increasingly searching for severity information leading up to severe weather events (Wanless et al. 2023). This result is also shown when EMs were asked if they used one probability more than the other in their operations. While the majority indicated they use both equally, more EMs selected that they use the probability of one more strong tornado more than the probability of two or more strong tornadoes.

However, this difference in EMs choosing a lower threshold for probability of one or more strong tornadoes may be due to some EMs confusing the different reference classes between these probabilities in tornado watches and the SPC convective outlook probabilistic tornado graphics for day 1 and 2. This was very prevalent in the objective understanding portion of the experiment questions, despite the majority of EMs rating their own understanding of the probabilistic information in tornado watches as good or higher. In the experiment question, the majority of EMs increasingly chose the 'correct' reference class from 20% to 80% for the probability of two or more tornadoes. However, many EMs still selected the reference class for the SPC convective outlook. The 80% randomization saw a decrease in the number of respondents that selected this option, and an increase in the correct reference class. This may be due to EMs understanding that the SPC convective outlook does not go above 60% for the tornado probabilistic graphic, and thus used the process of elimination to select another response. Overall, these results stress the importance of *explicitly* including the reference class in any probabilistic information included in a forecast to allow users to fully understand and contextualize the information.

# Chapter 4 - Discussion

## 4.1 Review of Major Findings

The overarching results of this study show that the two tornado probabilities associated with tornado watches are very reliable and generally accurate, and EMs regularly use the probabilities in their operations for threshold decision making. Over the course of the almost 19 year dataset, 69.8% of tornado reports occurred in at least one watch, either tornado or severe thunderstorm, and 98.5% of strong (EF2+) tornadoes occurred in a watch. Many of the Brier skill scores for the probability of one or more strong tornadoes were negative, indicating less accuracy than the climatological event frequency. Additionally, the South showed the largest difference from perfect reliability for low to moderate probabilities for two or more tornadoes in a watch area, in comparison to other regions. Daytime tornado watch probabilities were in general more reliable than nighttime watch probabilities, with these nighttime probabilities being more over-forecast than daytime, though the differences are small. Of the four case studies analyzed, three of them had significant tornadoes occur prior to the tornado watch of interest being issued, either in the watch area or outside of it.

Of the EMs we surveyed, almost all of them indicated they receive tornado watch information from their local NWS office, many of which received that information through NWS Chat. Of the EMs that indicated they use the tornado probabilities associated with tornado watches in their operations, many mentioned taking increased action for higher probabilities in comparison to lower probabilities. The majority of EMs selected the reference class attributed to tornado watch probabilities, however some EMs are confusing the SPC convective outlook reference class with tornado watch reference class.

## 4.2 Implications & Future Research

The results for the reliability verification of the tornado probabilities associated with tornado watches illustrate some of the current forecast challenges, specifically in the South and with overnight storms. While the Southeast US was not an explicit region used in this study, the South region included the states that make up the SE US, thus, play an important role in the interpretation of the South's results. The Southeast has the largest amount of High Shear Low CAPE tornado environments (Anderson-Frey et al. 2019), which are considered to be more difficult to forecast given that these environments evolve rapidly and tornadoes in these environments are typically quick spin ups within QLCSs (Goodnight et al. 2022; Lovell and Parker 2022). This coincides with overnight storms, as QLCSs are much more common in overnight hours compared to supercells. These difficult-to-forecast setups are slightly more overforecast for tornado watch probabilities in comparison to other regions of the CONUS and to daytime tornado watch probabilities. However, the differences in reliability (i.e., the slight overforecast) are relatively small, especially when considering the challenges that come with providing adequate time to seek shelter for overnight and storms in the South. While the goal for verification is generally to be as close to perfect as possible, the impact of these probabilities being slightly higher than they should have been is likely minimal compared to issuing a tornado watch, allowing the public to take the necessary steps to keep themselves and their families safe. This is especially true for those that do not have adequate shelter within their own home (i.e. a manufactured home), and have to relocate to a more secure shelter. Additionally, the four case studies analyzed illustrate some of the numerous situations forecasters encounter when issuing tornado watches and show that while verification may provide information regarding how forecasts perform overall, the true value of a forecast comes from how stakeholders use these forecasts to inform their decision making (Murphy 1993).

While the difference in slightly overforecasting tornado probabilities associated with tornado watches may not greatly impact the public, they may impact EMs and their decision making processes. Many EMs indicated they take increased action with higher tornado probabilities and some EMs confused the probability reference class within tornado watches with the SPC's convective outlook reference class. Thus, the overforecast of these probabilities combined with the confusion of different probabilistic reference classes, some EMs may take more increased action for a 40% chance of two or more tornadoes and a 20% chance of one or more strong tornadoes than necessary. With this being said, it is important to note that EMs prefer to know the worst case scenario for a severe weather event (Ernst et al. 2018; Cross and LaDue 2021; Hogg et al. 2026). Thus, the largest implication for EMs and other end users is the importance of clarifying the reference class for each probabilistic forecast and product to ensure they are being interpreted correctly by any user.

Given how little published research has been conducted on the probabilities within tornado watches, this analysis was meant to serve as a stepping stone for further analysis. The verification portion of this analysis focused specifically on the reliability of the tornado probabilities associated with tornado watches, and did not analyze other aspects of verification, inherently simplifying the verification process, but also masking other forms of verification (Wilks 2013). Future work on this topic should perform other verification techniques on these probabilities, such as analyzing probability of detection, success ratio, and the false alarm rate, to gain more insight about how they perform and how they influence the quality of tornado warnings (Krocak and Brooks 2021). Additionally, a significant aspect of tornado watches that was not included in this study is the impact of varying watch sizes on the interpretation of the tornado watch probabilities. Since these probabilities are within the watch area and not within 25 miles of a point, a large tornado watch inherently has a larger forecasted chance of receiving more tornadoes than a smaller tornado watch with the same probabilities. Therefore, an analysis of how these probabilities perform by watch size is crucial to understand how the probabilities are being treated from a forecaster's perspective. On the same note, it would also be beneficial to analyze the number of tornadoes that occur by forecasted probability and to perform verification at different thresholds of environmental variables, such as the Significant Tornado Parameter or HSLC environments. Extending this analysis into the hail and wind probabilities in tornado and severe thunderstorm watches could also provide more insight into how probabilities perform across different hazards in watches.

Furthermore, the WxEM survey results are the beginning of understanding how stakeholders use the tornado probabilities associated with tornado watches. Thus, continuing this research by conducting interviews with the EMs who completed the survey would be beneficial to understand how exactly these probabilities fit into their decision making processes, and if they believe there is a better way to present this information. Additionally, since EMs were not given the option to explain their choices on the experiment questions of this survey, discussing these choices in an interview setting would be useful to understand why many EMs confused the reference class of these probabilities with the SPC convective outlook. While other surveys of the general public are able to take a representative sample of the US population, there is no full list of every EM in the US, thus it is not possible to take a representative sample of EMs or to weigh these results to match a representative sample. Outside of EMs, conducting a survey with members of the public to understand if they know these probabilities exist, if they use them, and if including these probabilities in public messaging heightens their risk perceptions could potentially benefit EMs and forecasters to know if they should include this information in their messaging to the public.

## 4.3 Conclusions

In summary, the two tornado probabilities associated with tornado watches are used by emergency managers in their operations, and are generally very reliable and accurate. Negative Brier skill scores for the probability of one or more strong tornadoes and slight overforecasts for the South and nighttime tornado watch probabilities reflect current operational forecasting challenges, but the value of issuing a tornado watch with specific probabilities may outweigh not issuing a tornado watch in a scenario where members of the public need ample time to shelter from severe weather. However, these probabilities may impact stakeholders, such as EMs, who use these probabilities in their decision making process. While the majority of EMs objectively understand the reference class of these probabilities, many confuse the reference class with the SPC convective outlook, indicating the need to explicitly state the reference class of these two products when communicating this information to EMs, and continue working with stakeholders to understand how they use these products to protect their jurisdictions.

# Appendix A

| Probability of two or more tornadoes ($f_k$) | Number of Watches (n) | Watches with 2 or more tornadoes | Observed frequency of the event ($o_k$) | Brier Score | Brier Skill Score | REL | RES |
|---|---|---|---|---|---|---|---|
| 20 | 26 | 9 | 0.346 | 0.248 | -0.005 | 1.384e-04 | 5.670e-05 |
| 30 | 599 | 130 | 0.217 | 0.177 | 0.282 | 1.027e-03 | 7.400e-03 |
| 40 | 1656 | 566 | 0.342 | 0.228 | 0.073 | 1.398e-03 | 3.956e-03 |
| 50 | 822 | 418 | 0.509 | 0.250 | -0.015 | 1.485e-05 | 9.695e-04 |
| 60 | 441 | 263 | 0.596 | 0.241 | 0.023 | 1.446e-06 | 2.696e-03 |
| 70 | 240 | 182 | 0.758 | 0.187 | 0.242 | 2.035e-06 | 6.070e-03 |
| 80 | 112 | 91 | 0.813 | 0.153 | 0.381 | 4.360e-06 | 3.878e-03 |
| 90 | 75 | 63 | 0.840 | 0.138 | 0.440 | 6.726e-05 | 2.993e-03 |
| 95 | 43 | 43 | 1.00 | 0.003 | 0.990 | 2.678e-05 | 3.339e-03 |

Table A.1: Table of the number of tornado watches, tornado watches with two or more tornadoes, the observed frequency of two or more tornadoes occurring, Brier score, Brier skill score, Reliability value, and Resolution value for each forecasted probability. Yellow boxes indicate less than 20 watches.

| Probability of one or more strong tornadoes ($f_k$) | Number of Watches (n) | Watches with 2 or more tornadoes | Observed frequency of the event ($o_k$) | Brier Score | Brier Skill Score | REL | RES |
|---|---|---|---|---|---|---|---|
| 2 | 1 | 0.0 | 0.0 | 4.0e-04 | 0.998 | 9.965e-08 | 1.055e-05 |
| 5 | 20 | 0.0 | 0.0 | 2.500e-03 | 0.985 | 1.246e-05 | 2.110e-04 |
| 10 | 327 | 22 | 0.067 | 0.064 | 0.609 | 8.723e-05 | 1.563e-03 |
| 20 | 2524 | 349 | 0.138 | 0.123 | 0.248 | 2.396e-03 | 2.866e-03 |
| 30 | 513 | 139 | 0.271 | 0.198 | -0.214 | 1.078e-04 | 5.429e-04 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 40 | 288 | 115 | 0.399 | 0.240 | -0.468 | 3.460e-08 | 2.688e-03 |
| 50 | 169 | 80 | 0.473 | 0.250 | -0.530 | 2.985e-05 | 3.015e-03 |
| 60 | 67 | 42 | 0.627 | 0.235 | -0.436 | 1.205e-05 | 2.960e-03 |
| 70 | 46 | 33 | 0.717 | 0.203 | -0.242 | 3.466e-06 | 3.000e-03 |
| 80 | 43 | 31 | 0.721 | 0.207 | -0.269 | 6.697e-05 | 2.842e-03 |
| 90 | 11 | 10 | 0.909 | 0.083 | 0.494 | 2.265e-07 | 1.356e-03 |
| 95 | 5 | 5 | 1.000 | 0.003 | 0.985 | 3.114e-06 | 7.857e-04 |

Table A.2: Table of the number of tornado watches, tornado watches with one or more strong tornadoes (EF2+), the observed frequency of one or more strong tornadoes occurring, Brier score, Brier skill score, Reliability value, and Resolution value for each forecasted probability. Yellow boxes indicate less than 20 watches.

| Probability of two or more tornadoes | 2006-2011 | 2012-2017 | 2018-2024 |
|---|---|---|---|
| 5 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 |
| 20 | 12 | 14 | 0 |
| 30 | 271 | 164 | 164 |
| 40 | 832 | 436 | 388 |
| 50 | 266 | 260 | 296 |
| 60 | 138 | 87 | 216 |
| 70 | 69 | 52 | 119 |
| 80 | 43 | 23 | 46 |
| 90 | 38 | 17 | 20 |
| 95 | 26 | 6 | 11 |

Table A.3: The number of tornado watches per forecast probability of two or more tornadoes stratified by years. Yellow boxes indicate less than 20 watches

| Probability of one or more strong tornadoes | 2006-2011 | 2012-2017 | 2018-2024 |
|---|---|---|---|

| | | | |
|---|---|---|---|
| 2 | 0 | 0 | 1 |
| 5 | 7 | 4 | 9 |
| 10 | 108 | 75 | 144 |
| 20 | 1126 | 724 | 674 |
| 30 | 228 | 138 | 147 |
| 40 | 107 | 46 | 135 |
| 50 | 54 | 32 | 83 |
| 60 | 26 | 14 | 27 |
| 70 | 27 | 6 | 13 |
| 80 | 7 | 18 | 18 |
| 90 | 3 | 1 | 7 |
| 95 | 2 | 1 | 2 |

Table A.4: The number of tornado watches per forecast probability of one or more strong tornadoes (EF2+) stratified by years. Yellow boxes indicate less than 20 watches.

| Probability of two or more tornadoes | Winter (DJF) | Spring (MAM) | Summer (JJA) | Fall (SON) |
|---|---|---|---|---|
| 2 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 |
| 20 | 2 | 13 | 4 | 7 |
| 30 | 132 | 228 | 128 | 111 |
| 40 | 231 | 766 | 404 | 255 |
| 50 | 133 | 388 | 184 | 117 |
| 60 | 70 | 237 | 82 | 52 |
| 70 | 41 | 138 | 24 | 37 |
| 80 | 19 | 69 | 7 | 17 |

| | | | | |
|---|---|---|---|---|
| 90 | 12 | 46 | 11 | 6 |
| 95 | 3 | 36 | 1 | 3 |

Table A.5: The number tornado watches per forecast probability of two or more tornadoes stratified by season. Yellow boxes indicate less than 20 watches.

| Probability of one or more strong tornadoes | Winter (DJF) | Spring (MAM) | Summer (JJA) | Fall (SON) |
|---|---|---|---|---|
| 2 | 0 | 0 | 0 | 1 |
| 5 | 4 | 5 | 8 | 3 |
| 10 | 49 | 120 | 95 | 63 |
| 20 | 354 | 1145 | 610 | 415 |
| 30 | 113 | 263 | 77 | 60 |
| 40 | 69 | 153 | 30 | 36 |
| 50 | 28 | 110 | 16 | 15 |
| 60 | 11 | 48 | 6 | 2 |
| 70 | 7 | 30 | 1 | 8 |
| 80 | 7 | 33 | 2 | 1 |
| 90 | 1 | 9 | 0 | 1 |
| 95 | 0 | 5 | 0 | 0 |

Table A.6: The number of tornado watches per forecast probability of one or more strong tornadoes (EF2+) stratified by season. Yellow boxes indicate less than 20 watches.

| Probability of two or more tornadoes | Day | Night |
|---|---|---|
| 5 | 0 | 0 |
| 10 | 0 | 0 |
| 20 | 19 | 7 |

| | | |
|---|---|---|
| 30 | 284 | 315 |
| 40 | 1090 | 566 |
| 50 | 598 | 224 |
| 60 | 340 | 101 |
| 70 | 194 | 46 |
| 80 | 88 | 24 |
| 90 | 67 | 8 |
| 95 | 38 | 5 |

Table A.7: The number of tornado watches per forecast probability of two or more tornadoes stratified by time of day. Yellow boxes indicate less than 20 watches.

| Probability of one or more strong tornadoes | Day | Night |
|---|---|---|
| 2 | 0 | 1 |
| 5 | 13 | 7 |
| 10 | 196 | 131 |
| 20 | 1674 | 850 |
| 30 | 340 | 173 |
| 40 | 210 | 78 |
| 50 | 129 | 40 |
| 60 | 62 | 5 |
| 70 | 38 | 8 |
| 80 | 40 | 3 |
| 90 | 11 | 0 |
| 95 | 5 | 0 |

Table A.8: The number of tornado watches per forecast probability of one or more strong tornadoes (EF2+) stratified by time of day. Yellow boxes indicate less than 20 watches.

| Probability of two or more tornadoes | Great Plains | Midwest | Northeast | South | West |
|---|---|---|---|---|---|
| 5 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 |
| 20 | 5 | 10 | 0 | 10 | 1 |
| 30 | 150 | 116 | 18 | 313 | 2 |
| 40 | 493 | 354 | 72 | 715 | 22 |
| 50 | 276 | 165 | 17 | 353 | 11 |
| 60 | 146 | 87 | 8 | 194 | 6 |
| 70 | 88 | 38 | 5 | 109 | 0 |
| 80 | 35 | 20 | 0 | 56 | 1 |
| 90 | 28 | 11 | 1 | 34 | 1 |
| 95 | 15 | 5 | 0 | 23 | 0 |

Table A.9: The number of tornado watches per forecast probability of two or more tornadoes stratified by geographic region. Yellow boxes indicate less than 20 watches.

| Probability of one or more strong tornadoes | Great Plains | Midwest | Northeast | South | West |
|---|---|---|---|---|---|
| 2 | 0 | 0 | 1 | 0 | 0 |
| 5 | 5 | 3 | 2 | 9 | 1 |
| 10 | 84 | 58 | 11 | 167 | 7 |
| 20 | 771 | 546 | 93 | 1085 | 29 |
| 30 | 157 | 86 | 11 | 255 | 4 |
| 40 | 104 | 55 | 0 | 128 | 1 |
| 50 | 55 | 35 | 3 | 75 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| 60 | 26 | 11 | 0 | 30 | 0 |
| 70 | 16 | 5 | 0 | 25 | 0 |
| 80 | 13 | 6 | 0 | 23 | 1 |
| 90 | 3 | 1 | 0 | 7 | 0 |
| 95 | 2 | 0 | 0 | 3 | 0 |

Table A.10: The number of tornado watches per forecast probability of one or more strong tornadoes (EF2+) stratified by geographic region. Yellow boxes indicate less than 20 watches.

# Appendix B

In addition to Convective Outlooks, the NWS (Storm Prediction Center) issues tornado watches to indicate favorable conditions for tornadoes to form. Tornado Watches include three components:

- A graphical component that indicates the geographic area of the tornado watch
- Probabilistic information that indicates: (1) the probability of two or more tornadoes within the watch area and (2) the probability of one or more 'strong' (EF2+) tornadoes within the watch area
- A text narrative that summarizes the tornado watch using a plain-language summary and a more in-depth scientific discussion

[SHOW tor_watch_ex]

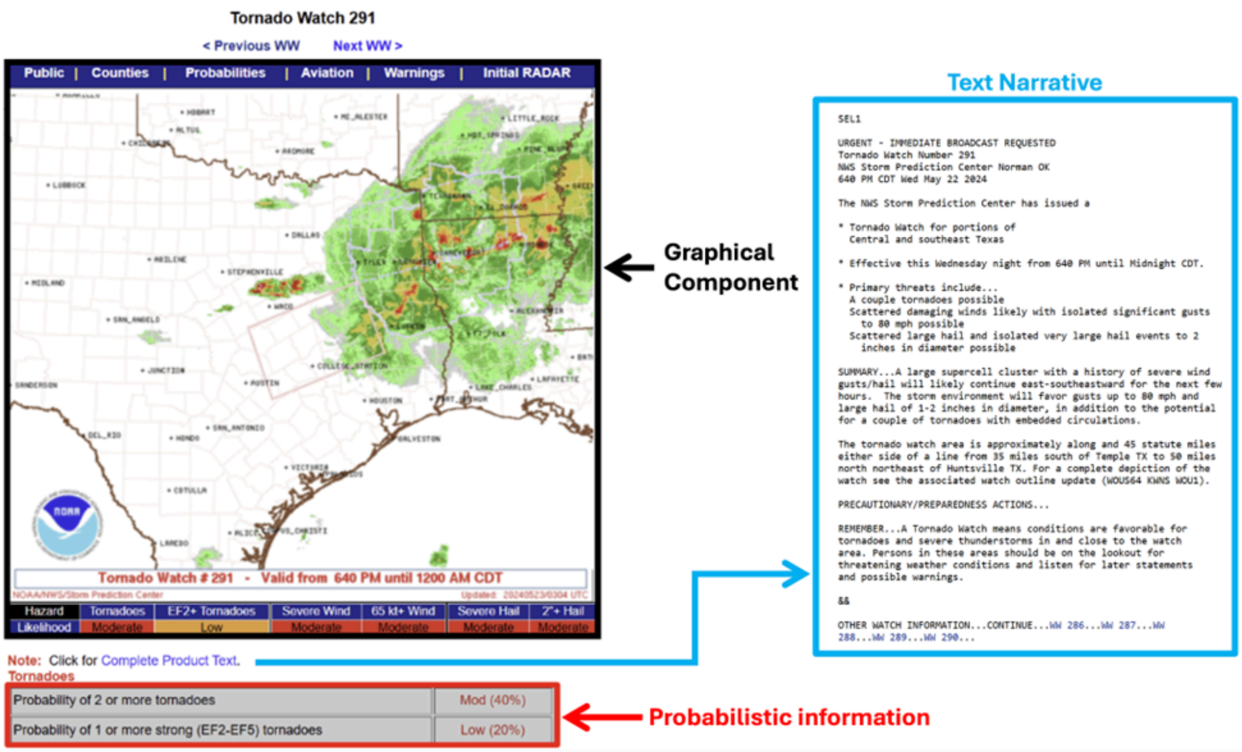



Figure A.1: Graphic shown to emergency managers while taking the WxEM survey to show the three components included in tornado watches.

use_torwtch: Do you currently use tornado watches in your operations?
1 - Not at all
2 - Not very much
3 - Somewhat
4 - Very much
5 - Extensively

num_torwtch: If you had to guess, how many tornado watches does your area (jurisdiction) receive in an average year?
1 - 0
2 - Between 1 and 2
3 - Between 3 and 5
4 - Between 6 and 10
5 - More than 10

---------------------------- page break ---------------------------

[SHOW IF use_torwtch > 1]

watch_source: What are your primary sources (entities/organizations) for receiving tornado watches? Please select all that apply. [RANDOM ORDER]
watch_source_nws: National Weather Service (local forecast office)
watch_source_spc: The Storm Prediction Center
watch_source_loctv: Local TV stations (for example: ABC, NBC, CBS, or FOX affiliates)
watch_source_nattv: National TV stations (for example: The Weather Channel)
watch_source_em: Other emergency managers
watch_source_priv: Private companies (for example: AccuWeather, Earth Networks, or The Weather Company)
watch_source_hse: In-house meteorologist
watch_source_oth: Something else (Please specify) [VERBATIM]

watch_chan: Through which channels (tools) do you receive tornado watches? Please select all that apply. [RANDOM ORDER]
watch_chan_brad: Broadcast radio
watch_chan_wxrad: Weather radio (NOAA/NWS radio)
watch_chan_tv: Television
watch_chan_web: Internet websites focused on weather forecasts (such as weather.gov or weather.com)
watch_chan_soc: Social media
watch_chan_wom: Word-of-mouth (including telephone calls, texts, or emails) from co-workers, contacts, etc.
watch_chan_phone: Cell phone applications
watch_chan_ham: Amateur/ham radio
watch_chan_data: Raw weather observations or models
watch_chan_chat: National Weather Service Chat (NWS Chat, now via Slack)
watch_chan_oth: Something else (Please specify) [VERBATIM]

---------------------------- page break ---------------------------

[SHOW IF use_torwtch >1]

[SHOW tor_watch_ex]

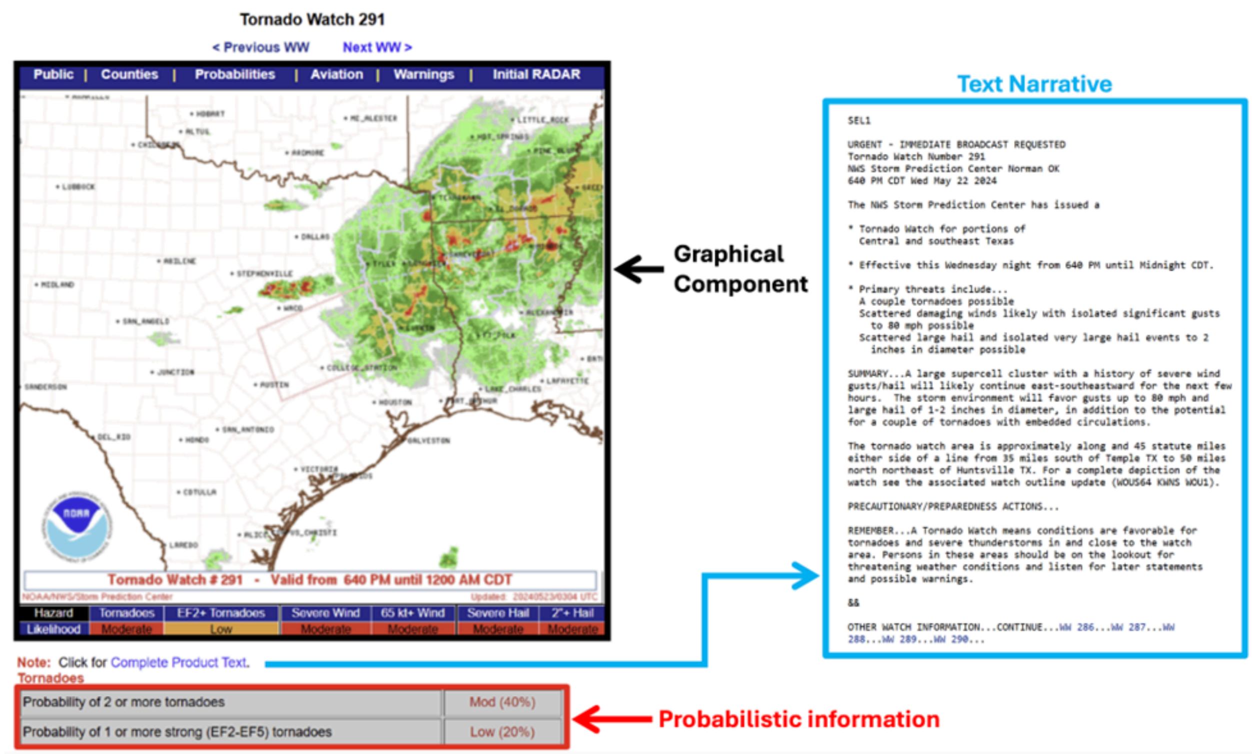



Figure A.1: Graphic shown to emergency managers while taking the WxEM survey to show the three components included in tornado watches.

How much do you use the following components of tornado watches? [SHOW IN TABLE; RANDOMIZE]

wtch_graph: Graphical component
1 - Not at all
2 - Not very much
3 - Somewhat
4 - Very much
5 - Extensively

wtch_prob: Probabilistic information
1 - Not at all
2 - Not very much
3 - Somewhat
4 - Very much
5 - Extensively

wtch_text: Text narrative
1 - Not at all
2 - Not very much

3 - Somewhat
4 - Very much
5 - Extensively

---------------------------- page break ---------------------------

[SHOW IF wtch_prob = 1]

The next few questions focus specifically on the probabilistic information included in tornado watches. As a reminder, these probabilities indicate: (1) the probability of two or more tornadoes within the watch area and (2) the probability of one or more 'strong' (EF2+) tornadoes within the watch area.

On the previous page, you indicated that you do not use the probabilistic information included in tornado watches. Can you tell us why? Please select all that apply.
no_use_where - I don't know where to find tornado watch probabilities.
no_use_inacc - I believe tornado watch probabilities are too inaccurate or unreliable.
no_use_relv - Tornado watch probabilities are not relevant or useful to my Emergency Management duties.
no_use_inter - I find it hard to interpret tornado watch probabilities.
no_use_aware - I was not aware of tornado watch probabilities before taking this survey.
no_use_other - Other (please specify) [VERBATIM]

---------------------------- page break ---------------------------

[SHOW IF no_use_aware = 1]

no_use_unaware: It sounds like you were not aware of the probabilistic information included in tornado watches before taking this survey. Now that you know about this information, how likely are you to use it to inform your operational decisions in the future?
1 - Very unlikely
2 - Somewhat unlikely
3 - Neither likely nor unlikely
4 - Somewhat likely
5 - Very likely

---------------------------- page break ---------------------------

[SHOW IF wtch_prob > 1]

The next few questions focus specifically on the probabilistic information included in tornado watches. As a reminder, these probabilities indicate: (1) the probability of two or more tornadoes within the watch area and (2) the probability of two or more 'strong' (EF2+) tornadoes within the watch area.

yes_use_torwtch: On the previous page, you indicated that you use the probabilistic information included in tornado watches. Can you briefly describe how you use it? (For example, do you take specific actions in response to different probability levels?) [VERBATIM]

thres_two_tor: How high would the probability of 2 more tornadoes have to be for you to take the actions and decisions you described in the previous question? [DROPDOWN, No threshold, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, Not sure / it depends (codebook: 999)]

thres_strong_tor: How high would the probability of 1 or more strong tornadoes have to be for you to take the actions and decisions you described in the previous question? [DROPDOWN, No threshold, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, Not sure / it depends (codebook: 999)]

more_prob: When you take the actions you described above, which type of tornado watch probability do you use more frequently, the probability of 2 or more tornadoes OR the probability of 1 or more strong tornadoes?
1 - I use the probability of 2 or more tornadoes more
2 - I use the probability of 1 or more strong tornadoes more
3 - I use them both equally
4 - Not sure

--------------------------- page break --------------------------

self_und_torwtch: How would you rate your understanding of the probabilistic information included in tornado watches?
1 - Poor
2 - Fair
3 - Good
4 - Very good
5 - Excellent

protect_torwtch: In your opinion, how important is the probabilistic information included in tornado watches for protecting people and property in your jurisdiction?
1 - Not at all important
2 - Slightly important
3 - Moderately important

4 - Very important
5 - Extremely important

---------------------------- page break ---------------------------

[SHOW IF num_torwtch  > 0]

acc_torwtch: In your experience, how accurate is the probabilistic information included in tornado watches for your area?
0 - I'm not sure
1 - Not at all accurate
2 - Slightly accurate
3 - Moderately accurate
4 - Very accurate
5 - Extremely accurate

---------------------------- page break ---------------------------

[RANDOMIZE:
watch_prob = 20, watch_days = 2 OR
watch_prob = 50, watch_days = 5 OR
watch_prob = 80, watch_days = 8]

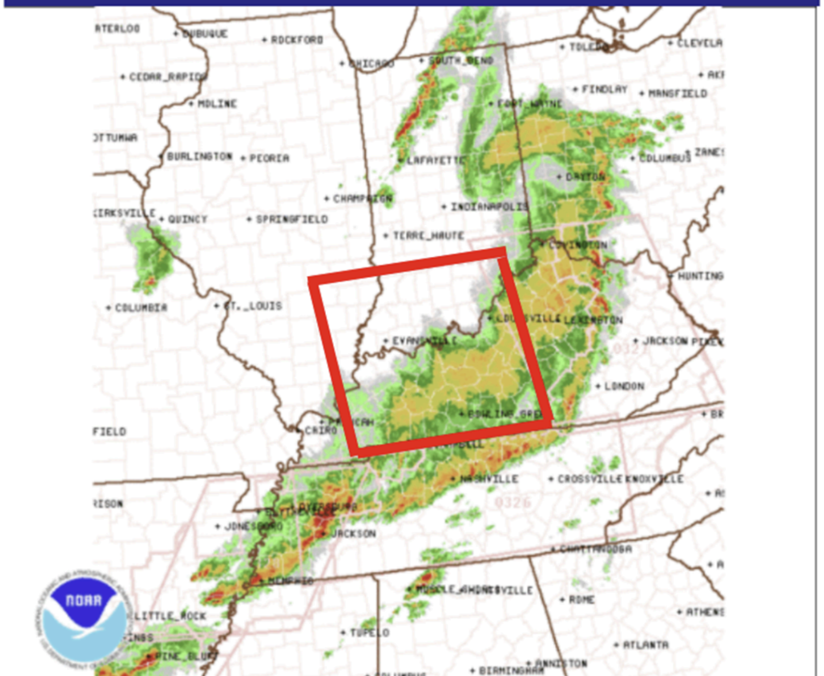
Imagine you see the following tornado watch graphic and tornado probabilities:

[SHOW watch_prob_graphic]

**Tornado Watch 324**

< Previous WW     Next WW >

Public | Counties | Probabilities | Aviation | Warnings | Initial RADAR | Related MD




Tornado Watch # 324 - Valid from 650 PM until 100 AM CDT

NOAA/NWS/Storm Prediction Center     Updated: 20240527/0544 UTC

| Hazard | Tornadoes | EF2+ Tornadoes | Severe Wind | 65 kt+ Wind | Severe Hail | 2"+ Hail |
|---|---|---|---|---|---|---|
| Likelihood | High | High | High | High | Moderate | Moderate |

Note: Click for Complete Product Text.

Tornadoes

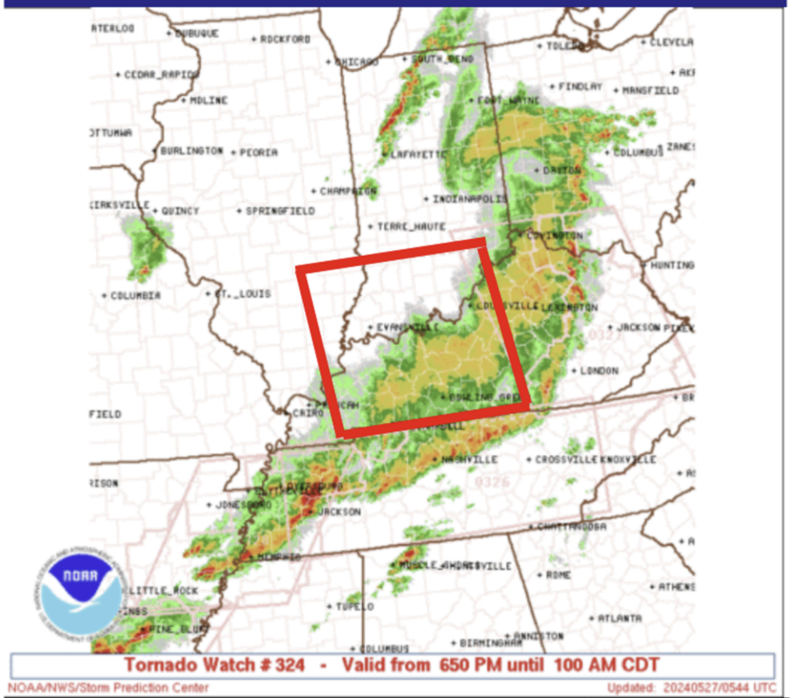
| | |
|---|---|
| Probability of 2 or more tornadoes | High (80%) |
| Probability of 1 or more strong (EF2-EF5) tornadoes | High (80%) |

Figure A.2: (watch_prob_graphic_80) One of three randomized tornado watch graphics and probabilities shown to emergency managers while taking the WxEM survey. Tornado probabilities are High (80%).

**Tornado Watch 324**

< Previous WW    Next WW >

Public | Counties | Probabilities | Aviation | Warnings | Initial RADAR | Related MD




| Hazard | Tornadoes | EF2+ Tornadoes | Severe Wind | 65 kt+ Wind | Severe Hail | 2"+ Hail |
|---|---|---|---|---|---|---|
| Likelihood | Moderate | Moderate | High | High | Moderate | Moderate |

Note: Click for Complete Product Text.

Tornadoes

| | |
|---|---|
| Probability of 2 or more tornadoes | Mod (50%) |
| Probability of 1 or more strong (EF2-EF5) tornadoes | Mod (50%) |

Figure A.3: (watch_prob_graphic_50) One of three randomized tornado watch graphics and probabilities shown to emergency managers while taking the WxEM survey. Tornado probabilities are Moderate (50%).

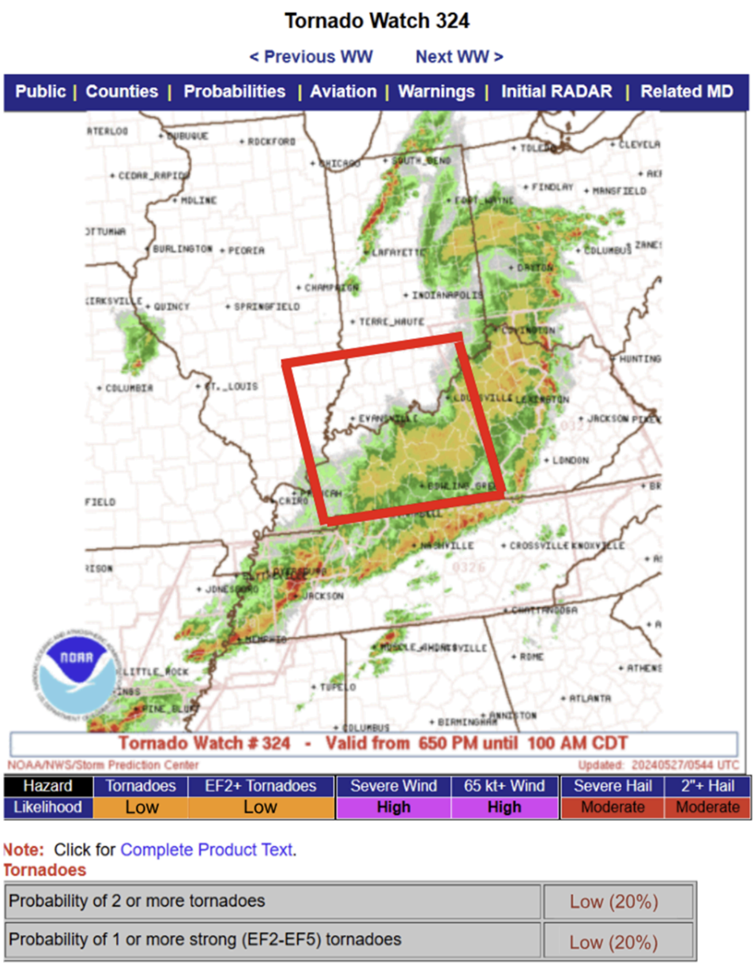



Figure A.4: (watch_prob_graphic_20) One of three randomized tornado watch graphics and probabilities shown to emergency managers while taking the WxEM survey. Tornado probabilities are low (20%).

und_torwtch: The tornado watch probabilities indicate a [watch_prob]% chance of 2 or more tornadoes. This information is telling you that: [RANDOMIZE]
1 -There is an [watch_prob]% chance that two or more tornadoes will occur somewhere within the watch area.
2 - There is an [watch_prob]% chance that two or more tornadoes will occur within about 25 miles of any location inside the watch area.
3 - About [watch_prob]% of the watch area would experience two or more tornadoes.
4 - I'm not sure

If this tornado watch (and the associated probabilistic information) were issued for your jurisdiction, what specific actions would you take in response? Please select all that apply.
act_torwtch_mon: Increase weather monitoring
act_torwtch_eoc: Activate the EOC

act_torwtch_part: Share this information with partners
act_torwtch_pub: Share this information with members of the public
act_torwtch_staff: Increase staffing
act_torwtch_nws: Contact my local National Weather Service office
act_torwtch_seek: Seek out information from other weather sources
act_torwtch_othr: Other (please specify)

# Reference List

Anderson, C. J., C. K. Wikle, Q. Zhou, and J. A. Royle, 2007: Population influences on tornado reports in the United States. *Wea. Forecasting*, **22**, 571–579, https://doi.org/10.1175/WAF997.1.

Anderson-Frey, A. K., Y. P. Richardson, A. R. Dean, R. L. Thompson, and B. T. Smith, 2016: Investigation of near-storm environments for tornado events and warnings. *Wea. Forecasting*, **31**, 1771–1790, https://doi.org/10.1175/WAF-D-16-0046.1.

——, ——, ——, ——, and ——, 2019: Characteristics of tornado events and warnings in the southeastern United States. *Wea. Forecasting*, **34**, 1521–1536, https://doi.org/10.1175/WAF-D-18-0211.1.

Anthony, R. W., and P. W. Leftwich, 1992: *Verification of severe local storms forecasts issued by the National Severe Storms Forecast Center: 1991*. NOAA Tech. Memo. NWS NSSFC-34, 18 pp.

Ashley, W. S., A. J. Krmenec, and R. Schwantes, 2008: Vulnerability due to nocturnal tornadoes. *Wea. Forecasting*, **23**, 795–807, https://doi.org/10.1175/2008WAF2222132.1.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.

Chamberlain, J. U., M. D. Flournoy, M. J. Krocak, H. E. Brooks, and A. K. Anderson-Frey, 2023: Analyzing tornado warning performance during individual storm life cycles. *Wea. Forecasting*, **38**, 927–944, https://doi.org/10.1175/WAF-D-22-0142.1.

Coleman, T. A., and P. G. Dixon, 2014: An objective analysis of tornado risk in the United States. *Wea. Forecasting*, **29**, 366–376, https://doi.org/10.1175/WAF-D-13-00057.1.

——, K. R. Knupp, J. Spann, J. B. Elliott, and B. E. Peters, 2011: The history (and future) of tornado warning dissemination in the United States. *Bull. Amer. Meteor. Soc.*, **92**, 567–582, https://doi.org/10.1175/2010BAMS3062.1.

Cross, R. N., and D. S. LaDue, 2021: When uncertainty is certain: A nuanced trust between emergency managers and forecast information in the southeastern United States. *Wea. Climate Soc.*, **13**, 3–17, https://doi.org/10.1175/WCAS-D-20-0017.1.

Davies, J. M., and A. Fischer, 2009: Environmental characteristics associated with nighttime tornadoes. *Electron. J. Oper. Meteor.*, **2009-EJ3**, 1–22.

Davis, J. M., and M. D. Parker, 2014: Radar climatology of tornadic and nontornadic vortices in high-shear, low-CAPE environments in the Mid-Atlantic and southeastern United States. *Wea. Forecasting*, **29**, 828–853, https://doi.org/10.1175/WAF-D-13-00127.1.

Elsner, J. B., L. E. Michaels, K. N. Scheitlin, and I. J. Elsner, 2013: The decreasing population bias in tornado reports across the central Plains. *Wea. Climate Soc.*, **5**, 221–232, https://doi.org/10.1175/WCAS-D-12-00040.1.

Ernst, S., D. LaDue, and A. Gerard, 2018: Understanding emergency manager forecast use in severe weather events. *J. Oper. Meteor.*, **6**, 95–105, https://doi.org/10.15191/nwajom.2018.0609.

Gensini, V. A., and H. E. Brooks, 2018: Spatial trends in United States tornado frequency. *npj Climate Atmos. Sci.*, **1**, 38, https://doi.org/10.1038/s41612-018-0048-2.

Gerst, M. D., and Coauthors, 2020: Using visualization science to improve expert and public understanding of probabilistic temperature and precipitation outlooks. *Wea. Climate Soc.*, **12**, 117–133, https://doi.org/10.1175/WCAS-D-18-0094.1.

Goodnight, J. S., D. A. Chehak, and R. J. Trapp, 2022: Quantification of QLCS tornadogenesis, associated characteristics, and environments across a large sample. *Wea. Forecasting*, **37**, 1521–1538, https://doi.org/10.1175/WAF-D-22-0016.1.

Hogg, D., A. Wanless, S. Stormer, E. Meister, J. Ripberger, and M. J. Krocak, 2026: Examining the role of emergency managers as conduits of weather information. *J. Oper. Meteor.*, **12**, 33–48, https://doi.org/10.15191/nwajom.2026.1404.

Hsieh, H.-F., and S. E. Shannon, 2005: Three approaches to qualitative content analysis. *Qual. Health Res.*, **15**, 1277–1288, https://doi.org/10.1177/1049732305276687.

IEM, 2025: IEM :: RADAR & NWS Warnings. Iowa Environmental Mesonet. Accessed 2 May 2025, https://mesonet.agron.iastate.edu/GIS/apps/rview/warnings.phtml.

Jon, I., S.-K. Huang, and M. K. Lindell, 2018: Perceptions and reactions to tornado warning polygons: Would a gradient polygon be useful? *Int. J. Disaster Risk Reduct.*, **30**, 132–144, https://doi.org/10.1016/j.ijdrr.2018.01.035.

Joslyn, S., and J. LeClerc, 2013: Decisions with uncertainty: The glass half full. *Curr. Dir. Psychol. Sci.*, **22**, 308–315, https://doi.org/10.1177/0963721413481473.

Joslyn, S. L., and J. E. LeClerc, 2012: Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *J. Exp. Psychol. Appl.*, **18**, 126–140, https://doi.org/10.1037/a0025185.

Krippendorff, K., 2019: *Content analysis: An introduction to its methodology.* 4th ed. SAGE Publications, 472 pp., https://doi.org/10.4135/9781071878781.

Krocak, M. J., and H. E. Brooks, 2018: Climatological estimates of hourly tornado probability for the United States. *Wea. Forecasting*, **33**, 59–69, https://doi.org/10.1175/WAF-D-17-0123.1.

——, and ——, 2021: The influence of weather watch type on the quality of tornado warnings and its implications for future forecasting systems. *Wea. Forecasting*, **36**, 1883–1898, https://doi.org/10.1175/WAF-D-21-0052.1.

——, M. D. Flournoy, and H. E. Brooks, 2021: Examining subdaily tornado warning performance and associated environmental characteristics. *Wea. Forecasting*, **36**, 1907–1924, https://doi.org/10.1175/WAF-D-21-0097.1.

Lovell, L. T., and M. D. Parker, 2022: Simulated QLCS vortices in a high-shear, low-CAPE environment. *Wea. Forecasting*, **37**, 1539–1557, https://doi.org/10.1175/WAF-D-21-0133.1.

Maciag, T. A., H. Obermeier, K. Berry, M. J. Krocak, K. K. McClain, and D. Hogg, 2025: Emergency manager preferences for rapidly updating severe weather warnings. *Wea. Forecasting*, **40**, 112–130, https://doi.org/10.1175/WAF-D-24-0029.1.

Marsh, P., and M. Elliott, 2025: Storm Prediction Center Severe Weather GIS (SVRGIS) Page. Accessed 17 June 2025, https://www.spc.noaa.gov/gis/svrgis/.

Mason, S. J., 2004: On using "climatology" as a reference strategy in the Brier and ranked probability skill scores. *Mon. Wea. Rev.*, **132**, 1891–1895, https://doi.org/10.1175/1520-0493(2004)132<1891:OUCAAR>2.0.CO;2.

Morss, R. E., J. L. Demuth, and J. K. Lazo, 2008: Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Wea. Forecasting*, **23**, 974–991, https://doi.org/10.1175/2008WAF2007088.1.

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595–600.

——, 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281–293.

NWS Heritage, 2019: The start of tornado forecasts. Accessed 14 January 2026, https://vlab.noaa.gov/web/nws-heritage/-/the-start-of-tornado-forecasts.

Olson, J. R., C. M. Doyle, D. S. LaDue, and A. N. Marmo, 2023: End-user threat perception: Building confidence to make decisions ahead of severe weather. *J. Oper. Meteor.*, **11**, 95–109, https://doi.org/10.15191/nwajom.2023.1108.

Potvin, C. K., C. Broyles, P. S. Skinner, and H. E. Brooks, 2022: Improving estimates of U.S. tornado frequency by accounting for unreported and underrated tornadoes. *J. Appl. Meteor. Climatol.*, **61**, 909–930, https://doi.org/10.1175/JAMC-D-21-0225.1.

Qin, C., S. Joslyn, S. Savelli, J. Demuth, R. Morss, and K. Ash, 2024: The impact of probabilistic tornado warnings on risk perceptions and responses. *J. Exp. Psychol. Appl.*, **30**, 206–239, https://doi.org/10.1037/xap0000486.

R Core Team, 2024: *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Ripberger, J., A. Bell, A. Fox, A. Forney, W. Livingston, C. Gaddie, C. Silva, and H. Jenkins-Smith, 2022: Communicating probability information in weather forecasts: Findings and recommendations from a living systematic review of the research literature. *Wea. Climate Soc.*, **14**, 703–722, https://doi.org/10.1175/WCAS-D-21-0034.1.

Rosen, Z., M. J. Krocak, J. T. Ripberger, R. Cross, E. Lenhardt, C. L. Silva, and H. C. Jenkins-Smith, 2021: Communicating probability information in hurricane forecasts: Assessing statements that forecasters use on social media and implications for public assessments of reliability. *J. Oper. Meteor.*, **9**, 89–101, https://doi.org/10.15191/nwajom.2021.0907.

Sherburn, K. D., and M. D. Parker, 2014: Climatology and ingredients of significant severe convection in high-shear, low-CAPE environments. *Wea. Forecasting*, **29**, 854–877, https://doi.org/10.1175/WAF-D-13-00041.1.

——, ——, J. R. King, and G. M. Lackmann, 2016: Composite environments of severe and nonsevere high-shear, low-CAPE convective events. *Wea. Forecasting*, **31**, 1899–1920, https://doi.org/10.1175/WAF-D-16-0086.1.

Sherman-Morris, K., 2013: The public response to hazardous weather events: 25 years of research. *Geogr. Compass*, **7**, 669–685, https://doi.org/10.1111/gec3.12076.

Storm Prediction Center, 2026: SPC Severe Weather Events Archive. Accessed 9 February 2026, https://www.spc.noaa.gov/exper/archive/.

Vescio, M. D., and R. L. Thompson, 2001: Forecaster's forum: Subjective tornado probability forecasts in severe weather watches. *Wea. Forecasting*, **16**, 192–195.

Wallace, Z. C., L. Keys-Mathews, and A. A. Hill, 2015: The role of experience in defining tornado risk perceptions: A case from the 27 April 2011 outbreak in rural Alabama. *Southeast. Geogr.*, **55**, 400–416.

Wanless, A., and Coauthors, 2023: The Extreme Weather and Emergency Management Survey. *Wea. Climate Soc.*, **15**, 1–18, https://doi.org/10.1175/WCAS-D-23-0085.1.

Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences.* 3rd ed. Academic Press, 676 pp.

——, 2013: The calibration simplex: A generalization of the reliability diagram for three-category probability forecasts. *Wea. Forecasting*, **28**, 594–605, https://doi.org/10.1175/WAF-D-13-00027.1.