National Weather Service Warning Performance Associated With Watches

Jessica Ram

National Weather Center Research Experiences for Undergraduates, and
The Pennsylvania State University, University Park, Pennsylvania

Robert Johns

NOAA/NWS Storm Prediction Center (retired), Norman, Oklahoma

Elizabeth Quoetone

NOAA/NWS Warning Decision Training Branch, Norman, Oklahoma

Paul Schlatter

NOAA/NWS Warning Decision Training Branch, Norman, Oklahoma

Research Experiences for Undergraduates Final Project
Last Revised August 3, 2007

ABSTRACT

The National Oceanic and Atmospheric Administration's (NOAA) National Weather Service (NWS) is responsible for alerting the public to the threat of severe weather by issuing severe weather watches and warnings. The NWS Storm Prediction Center can issue severe thunderstorm watches, tornado watches, or particularly dangerous situation tornado watches while Weather Forecast Offices (WFO) issue severe thunderstorm warnings and tornado warnings. It is vital for these warnings to be accurate and illicit an effective response from those likely to be affected. Although many factors affect the warning decision process, it isolating and examining each factor is an important step towards improving the process. Data were collected using online archives for information on all of the watches, warnings, and events that occurred between January 1, 2006 and April 19, 2006. Several WFOs were given surveys to help determine some of the human factors that might lead to an association between watches and warning issuance. When combining all of the information, watches and watch type appear to be correlated with warning performance in a positive way. The forecasters who issue warnings have also indicated through the surveys that watches influence how forecasters make decisions, especially while issuing warnings.

1. Introduction

One of the most important goals of the National Oceanic and Atmospheric Administration's (NOAA) National Weather Service (NWS) is to alert the public to the threat of severe weather by issuing watches and warnings. The quality and timeliness of severe weather warnings are crucial so that the warnings can illicit an effective response from those likely to be affected. Many meteorological and non-meteorological factors affect the type and number of warnings that are issued by a given NWS Weather Forecast Office (WFO). The purpose of this study is to examine if severe weather watches influence warning performance, and if so, how and in what ways is warning performance affected?

The National Weather Service is required to issue a watch when "the risk of a hazardous weather or hydrologic event has increased significantly, but its occurrence, location, and/or timing is still uncertain. It is intended to provide enough lead time so that those who need to set their plans in motion can do so" (DOC/NOAA/NWS, 2007). A weather warning is required to be issued when "a hazardous weather or hydrologic event is occurring, is imminent, or has a very high probability of occurring. A warning is used for conditions posing a threat to life or property" (DOC/NOAA/NWS, 2007). The Storm Prediction Center (SPC), located in Norman, Oklahoma, is responsible for issuing severe weather watches across the United States. Watches cover an average area of about 25,000 square miles (about 65,000 km^2) (Corfidi, 1999). On the other hand, warnings are issued by local WFOs, which are spread out in 124 locations across the country. The average size of a warning is thought to be about 600 square miles (about 1,500 km^2) (P. Schlatter, personal communication).

Although the SPC and the WFO forecasters have a great amount of communication between each other prior to the issuance of a weather watch, local WFO forecasters may also be subconsciously influenced by the SPC's products. Previous studies (i.e. Polger et al., 1994; Andra et al., 2001; Brooks, 2004; Glahn, 2005) have examined factors that effect warning performance, however no studies to date have incorporated the potential effects of watches and watch type.

Three possible hypotheses are: watches have a positive effect, a negative effect, or no effect on warning performance. Whichever is the case, it is extremely important to find out what kind of influence watches have. By isolating some of the valuable or invaluable factors that influence warning performance, steps can be taken to improve the warning process.

Section 2 will discuss the methods of gathering, organizing, and analyzing the data pertaining to the watches, warnings, and events that were studied. A summary of the survey that was used for the project is also explained. Section 3 will discuss the results of the warning statistics and surveys, followed by the conclusions drawn from the study in section 4. In some cases, the results led to more questions and pointed to further areas for research.

2. Data and Methodology

The dataset used for this project consists of the severe weather watches issued between January 1, 2006 and 0700 UTC on April 19, 2006. Of the 200 watches that were issued, 109 were tornado watches, 78 were severe thunderstorm watches, and 13 were particularly dangerous situation (PDS) tornado watches. Although a PDS can be associated with a severe thunderstorm watch, they are rarely issued and did not exist in the data. The size of the dataset is a somewhat limiting factor for this project. Conclusions from less than 4 months of data from one year may not represent the rest of the year's 892 total number of watches, nor is it likely to be representative of any other year.

During the time period of this dataset, thousands of warnings were issued and thousands of events occurred. Information gathered from each watch issued during this time period was obtained using the online SPC Product and Report Archives (http://www.spc.noaa.gov/archive/). This website listed the time, date, and type of the watch, along with the Watch Outline Update (WOU), which lists every county initially covered by the watch. The date and counties included in the watch were then plugged into The National Weather Service Verification/Storm Data website (http://verification.nws.noaa.gov/verification/severe/svrverif.html) to produce a list of all of the warnings and events that occurred on those dates in the specified counties.

Entering this information into a spreadsheet made it easy to manage the data and calculate the results. Although the NWS verification website calculates its own warning statistics, we computed the statistics somewhat differently. For example, the website verifies a tornado warning based on either a wind or hail event or a tornado report. As discussed later, a tornado warning for our purposes was only verified by a tornado report.

Once a warning is issued, there are 3 types of events that can verify if severe weather occurred: hail with a diameter of at least ¾ of an inch, damaging convective winds estimated or measured as 50 knots or greater, or wind damage, or a report of a tornado or tornado damage. In this study, a tornado warning was verified only by a tornado report, while a severe thunderstorm warning was verified by any of wind, hail, or tornado reports. A tornado event was only considered to be warned for if it was preceded by a tornado warning for that county, while a wind or hail event was considered warned for if either a tornado warning or a severe thunderstorm warning had been issued beforehand. Once it was determined how to verify warnings, focus shifted to computing statistics. Formulas were entered into the spreadsheet in order to find the probability of detection (POD), false alarm rate (FAR), and critical success index (CSI), and provided in Appendix A. The usage of CSI and 2x2 contingency tables to evaluate forecasting skill has been well established in previous studies (i.e. Donaldson et al., 1975; Doswell, 1990; Schaefer, 1990).

One of the downfalls to our method of calculating the statistics is that we looked at the combined numbers and not the numbers for individual watches. Any outlying watches would not be seen in our calculations. Also, multiple verifying events occurring in the same warned county were all included in the calculations, thereby inflating the POD statistics.

Another problem could exist with this data because of the verification process that the NWS uses in order to validate a warning. The events used in this study are taken at face value from the website. Unfortunately, verifying warnings is not as easy or as simple as the website would show. Events are only reported when either the NWS is able to contact a witness to the event or a witness calls to report an event. Because of this, warnings in sparsely populated areas are hard to validate. In addition, a single report may not represent the severity of the storm or the entirety of the affected areas (Doswell et al., 2005).

Along with looking at the statistics from every watch, we sent survey forms to several select WFOs in an attempt to identify their perceptions or how watch type might affect the warning decision process. The questions that were asked in this survey are provided in Appendix B. The survey asked the forecasters about their views on different types of watches and what they thought the SPC's expectations for a watch were. The formal criteria for the SPC's watches are shown in Table 1.

Watch Type		Criteria
	Severe Thunderstorm	Forecast of 6 or more wind/hail events
	Tornado	Forecast of multiple weak tornadoes or any tornado that could produce EF2 or greater damage
	PDS Tornado	Same as Tornado but with a likelihood of multiple strong (damage of EF2 or EF3) or violent (damage of EF4 or EF5) tornadoes

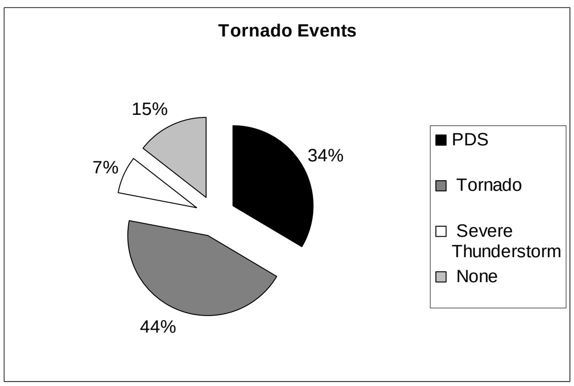
Table 1 – A chart explaining the SPC's criteria for issuing the different types of watches; obtained from the National Weather Service Instruction 10-512.

The survey also asked forecasters to assess their levels of confidence when issuing warnings while under different types of watches. Lastly, the forecaster was asked to assess how confident they thought their coworkers were in the same situations. This survey was sent to 12 WFOs across the contiguous U.S., consisting of 3 offices located in each of the 4 NWS regions in order to obtain a geographically diverse sample. The calculations that have been made from these surveys came from 54 forecasters who were located in 11 different forecast offices.

3. Results

a. Events

Over the time period of the dataset, there were 509 tornadoes and 5805 wind or hail events. Figure 1 shows under which types of watches these events occurred. Over 75% of tornado occurrences happened within either a PDS watch or a tornado watch, and about 75% of wind or hail events were within some sort of watch. Within the watches, it was important to find out if the events had been properly warned. Table 2 displays the percentage of tornadoes and wind or hail events that were warned by either a tornado or severe thunderstorm warning within the different types of watches. One of the more important points to notice about these numbers is that the highest percentage for warned tornadoes was within a PDS tornado watch, and the highest percentage for warned severe events was within a severe thunderstorm watch. Tornadoes with a tornado warning were significantly lowest outside of any type of watch. Although it might seem like the percentage of wind or hail events warned with a severe thunderstorm warning are lower for PDS and tornado watches, this is because the wind or hail event could be warned by a tornado warning instead. All of this shows that given the proper watch type, the detection and proper warning for tornadoes is likely to increase.



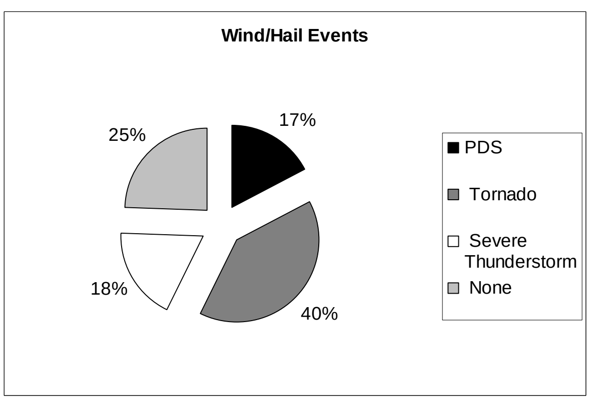


Figure 1 – The percentage of tornado events that occurred within the different types of watches on top, and wind and hail events on the bottom.

Tornado Events in Tornado Warnings

PDS Watch	TOR Watch	SVR Watch	No Watch
71%	48%	46%	26%

Wind/Hail Events in Severe Tstm or Tornado Warnings

	PDS Watch	TOR Watch	SVR Watch	No Watch
TSTM warning	65%	76%	84%	77%
TORN warning	24%	12%	5%	3%

Table 2 – The first table above displays the percentage of tornadoes within each watch type that occurred within a tornado warning. The next table shows the percentage of wind/hail events within each watch type that occurred within either a severe thunderstorm warning or a tornado warning.

As expected, often times a watch or a warning is issued without any ground truth to verify. The chart below in Figure 2 displays the percentages of these “false alarms” that occur based on watch type. Out of the 4094 false alarms, 3379 of them were severe thunderstorm warnings and the rest were tornado warnings. Most of these false alarms occurred within tornado watches. Also, within a PDS watch, there were twice as many false alarms for tornadoes than there was for wind or hail events.

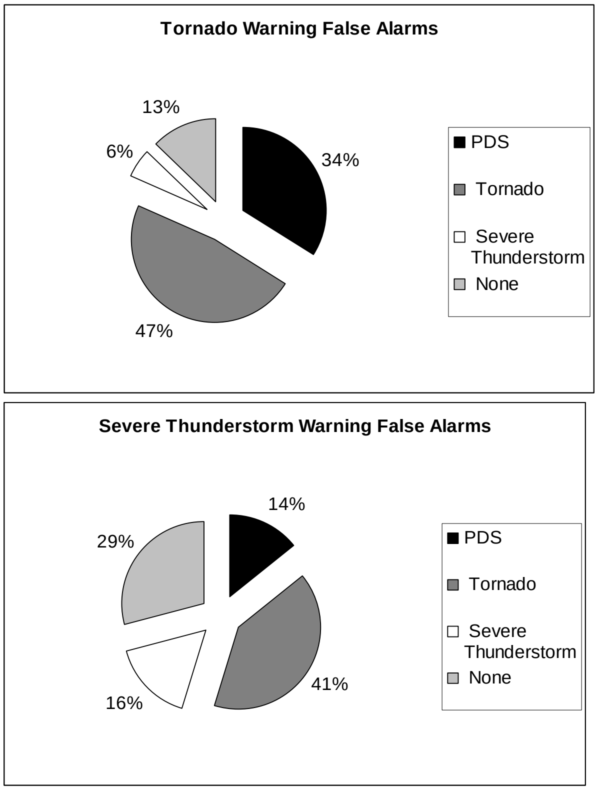


Figure 2 – The amount of false alarms for tornadoes that occurred in each watch type on top, and for wind/hail events on the bottom.

b. Warnings vs. Watches

The first table in Table 3 shows a broad overview of FAR, POD, and CSI based on the different types of watches. This table has combined the calculations for both tornado and severe thunderstorm warnings. Although the PDS watches have the highest FAR and the lowest CSI, they also have the highest POD. Out of all of the watch types, the no watch category has the lowest POD. Considering a potentially large threat tornadoes pose to life and property, tornado warning performance was isolated in Table 4. The large amount of wind/hail events masks the statistics for tornado events. The FAR for tornadoes is high, but so is the POD. The resulting CSI is very low, primarily because of the strong dependence of CSI to FAR.

		Watch Type			
		PDS	**TOR**	**SVR**	**NONE**
Statistic	**FAR**	.595	.555	.449	.555
	POD	.864	.854	.744	.733
	CSI	.381	.414	.463	.383

Table 3 – A table of statistics based on the type of watch for both tornado and severe thunderstorm warnings combined.

		Watch Type			
		PDS	**TOR**	**SVR**	**NONE**
Statistic	**FAR**	.800	.853	.814	.808
	POD	.909	.861	.861	.494
	CSI	.196	.143	.181	.161

Table 4 – The statistics for tornado warning performance only.

As previously stated, most (about 75%) of the tornado events in this dataset occurred within a tornado or PDS watch. Not only did most of the tornadoes occur within these two types of watches, but this is also where most of the tornado warnings were issued. While the PDS and tornado watches had 1449 tornado warnings combined, the severe thunderstorm watches only had 102 tornado warnings, while 130 tornado warnings were issued outside of any type of watch. Figure 3 shows the charts for the events and warnings that occurred in the different types of watches. Under the PDS watches, only 5 tornadoes occurred with no type of warning, compared to 18 tornadoes with no warning in tornado watches, and 14 tornadoes with no warning in the cases where no watches were in effect. Even though there were a lot of false alarms with severe thunderstorm warnings, these graphs show that they are nicely paired with severe thunderstorm events. All of the watch types have a fair amount of unwarned wind and hail events.

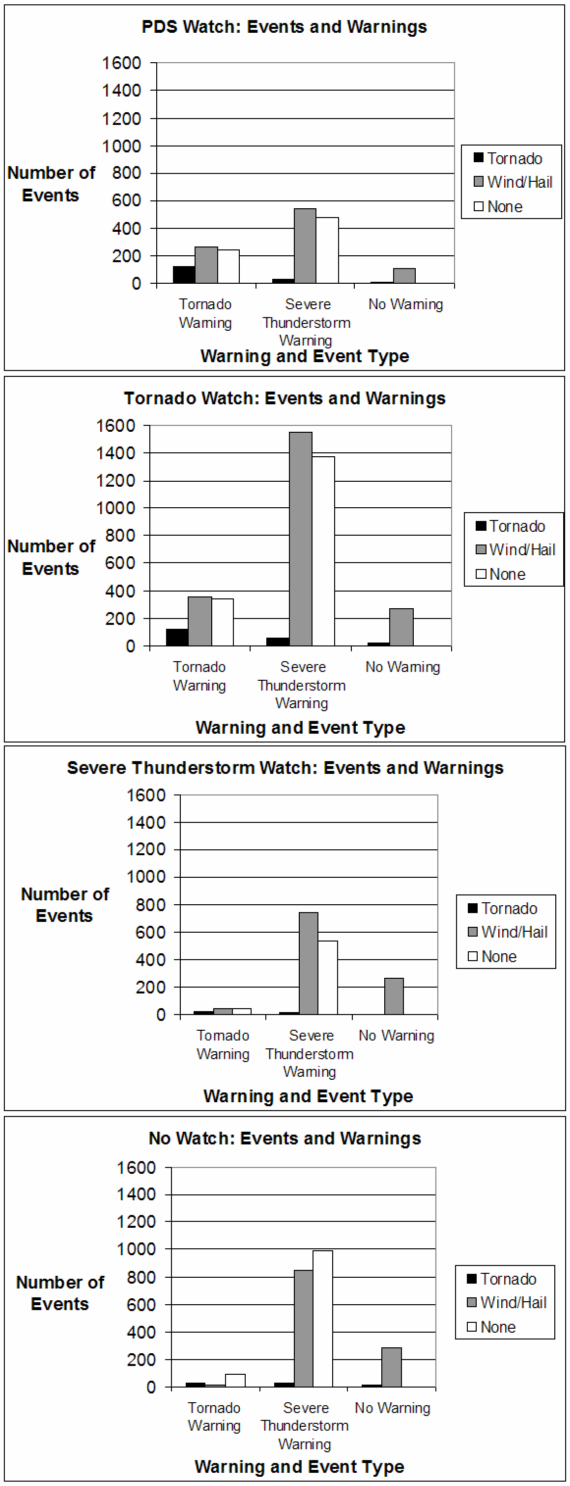


Figure 3 – Charts of events and warnings for different watch types.

c. Survey – Perceptions

The survey sent to WFO forecasters is an important element of this study. When asked about their interpretation of the SPC's expectations for watches, forecasters' responses varied widely. However, most forecasters had the same general expectations for a certain type of watch. For severe thunderstorm watches, the most prevalent expectations were for severe thunderstorms with several severe weather reports. A large number of forecasters also mentioned that the severe thunderstorms would be organized and not just pulse or isolated.

Forecasters responded with more variation when asked about tornado watches. Most of the forecasters agreed that numerous severe weather reports could be expected, but they commented differently on the number and intensity of possible tornadoes. Many forecasters thought a tornado watch meant that there was potential for tornadoes. Some thought that at least one isolated tornado was expected, while others thought the potential was for at least a few weak tornadoes or one significant tornado (EF2 or stronger).

The consensus expectation for a PDS tornado watch was for a greater chance of severe weather along with more significant tornadoes. Most of the forecasters agreed that the SPC was predicting multiple, violent, and long-lived tornadoes within a PDS tornado watch.

When asked about their thoughts on watches in general, 13 forecasters responded by saying that watches did not affect their warning decision process, and really felt they shouldn't. They stated that forecasters should already be aware of the situation without the SPC calling attention to it. On the other hand, many forecasters did admit that watches raise awareness to what is going on, both for their office and the public. Additionally, forecasters felt more confident that an event would occur if a watch was out, and would sometimes lead to a quicker or easier decision to issue a warning. A few forecasters also specifically commented that they appreciated the conference calls with the SPC so that they could talk about the situation and listen to another viewpoint.

An important factor regarding these surveys is that many of these forecasters work in various locations around the country. Forecasters in some offices stated that they had little experience with PDS tornado watches or tornado watches, and sometimes even severe thunderstorm watches. Some felt that the SPC may be looking at synoptic processes for severe convection, while some of these county warning areas are dominated by mesoscale processes on a smaller or less organized scale than the average watch.

d. Survey – Confidence Levels

The average confidence level needed by a forecaster in order to issue a warning is shown in Table 5. These results show that as the severity of a watch decreases, forecasters need to feel more confident a severe event will occur before issuing a warning. This is not surprising because in theory the local WFO forecaster would assess the environment in the local area and potentially arrive at the same conclusion as the SPC.

		Watch Type			
		PDS	**TOR**	**SVR**	**NONE**
Warning Type	**TSTM**	59%	63%	66%	71%
	TORN	54%	60%	67%	72%

Table 5 – A table showing the average confidence levels of forecasters for issuing the different types of warnings within the different types of watches.

Figure 4 shows the percentage of forecasters who thought that others in their office had a lower, higher, or the same confidence threshold for issuing tornado warnings. Although these numbers are slightly different depending on the type of watch, they do not change significantly and so the averages of all watches are displayed. These numbers are also very similar and representative of the numbers for severe thunderstorm warnings. As the charts show, a majority of the forecasters thought that either their coworkers issued warnings with about the same confidence level as they did, or that their coworkers were more apt to issue warnings. Interestingly, not many (only 13%) thought that their coworkers had a higher confidence for issuing warnings.

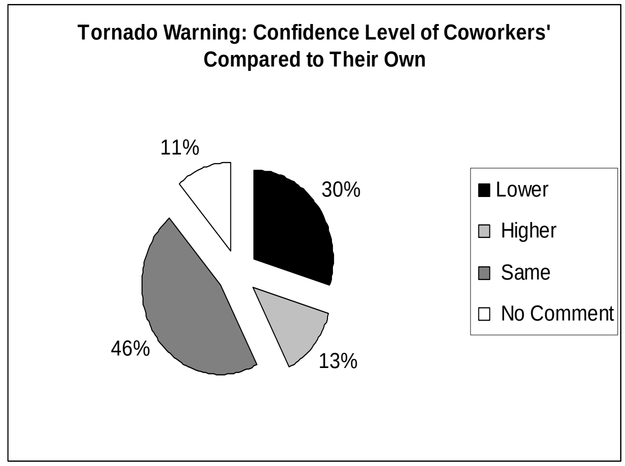


Figure 4 – Charts comparing the forecaster's levels of confidence with their opinion of their coworkers' levels of confidence under all watch types for issuing a tornado warning.

4. Conclusions

The results have shown that watches can be associated with changes in the warning decision process. When it comes to verification statistics, this change is associated with improving warning performance. Regarding human factors, the forecasters are also affected by watches, but this influence cannot be determined as good or bad. Cause and effect cannot be seen from these results alone, only an association.

Based on the percentages given earlier, watches cover a majority of severe events. In particular, 75% of all tornadoes occurred in either a tornado watch or a PDS tornado watch, and tornado warning performance in terms of POD is maximized for PDS tornado watches. Even though there were quite a few false alarms that occurred throughout the time period, these false alarms are usually associated with just as many warnings that did verify.

Although warning performance did, for the most part, improve in the appropriate watches, we cannot conclude that warning performance improved because of the watches. Even as many of the forecasters mentioned in their surveys, they should already be aware of the potential hazardous weather threats, and could have the same expectations as the SPC. In recent years, forecasters have increased access to the same datasets the SPC uses, so their expectation may be formed independently of the SPC watch. However, this is confirmation that watches are associated with an improved warning system for both the forecasting offices and the public.

Acknowledgements. This material is based upon work supported by the National Science Foundation under Grant No. ATM-0648566. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We greatly appreciate the help we have received from Harold Brooks and John Hart.

References

Andra D. L. Jr., E. M. Quoetone, and W. F. Bunting, 2002: Warning decision making: The relative roles of conceptual models, technology, strategy, and forecaster expertise on 3 May 1999. *Wea. Forecasting*, **17**, 559–566.

Brooks H. E., 2004: Tornado-warning performance in the past and future: A perspective from signal detection theory. *Bull. Amer. Meteor. Soc*, **85**, 837–843.

Corfidi, S. F., 1999. The birth and early years of the Storm Prediction Center, *Wea. Forecasting*, **14**, 507–525.

DOC/NOAA/NWS, 2007: National Severe Weather Products Specification. National Weather Service Instruction 10-512, May 2007, 65 pp.

DOC/NOAA/NWS, cited 2007: National Weather Service Glossary. [Available online at http://www.nws.noaa.gov/glossary/.]

DOC/NOAA/NWS, cited 2007: National Weather Service Verification/Storm Data. [Available online at https://verification.nws.noaa.gov/verification/severe/svrverif.html.]

DOC/NOAA/NWS, cited 2007: Storm Prediction Center Product and Report Archives. [Available online at http://www.spc.noaa.gov/archive/.]

Donaldson, R.J., R. M. Dyer and M. J. Krauss, 1975: An objective evaluator of techniques for predicting severe weather events. Preprints: *9th Conf. Severe Local Storms,* Norman, Oklahoma, Amer. Meteor. Soc., 321-326.

Doswell C. A. III, H. E. Brooks, and M. P. Kay, 2005: Climatological estimates of daily local nontornadic severe thunderstorm probability for the United States. *Wea. Forecasting*, **20**, 577–595.

Doswell C. A. III, R. P. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576–585.

Glahn, B., 2005: Tornado-warning performance in the past and future – Another perspective. *Bull. Amer. Meteor. Soc.,* **86**, 1135-1141.

Polger, P. D., B. S. Goldsmith, R. C. Przywarty, and J. R. Bocchieri, 1994: National Weather Service warning performance based on the WSR-88D. *Bull. Amer. Meteor. Soc.,* **75,** 203–214.

Schaefer J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575

Appendix A:

Following Doswell et al. (1990), we used the following 2x2 contingency table for statistical analysis of the warnings inside and outside of watches.

		observed	
		Yes	No
forecasted	Yes	a	b
	No	c	d

The table was populated from our dataset, with the exception of ‘d’. ‘d’ states that no warning was issued and no event occurred. Estimating ‘d’ is difficult at best given an ideal dataset. However, to compute False Alarm Rate (FAR), Probability of Detection (POD, and Critical Success Index (CSI), ‘d’ is not needed. The following equations were used given the variables from the 2x2 contingency table:

$$\text{False Alarm Rate (FAR)} = \frac{b}{a + b}$$

$$\text{Probability of Detection (POD)} = \frac{a}{a + c}$$

$$\text{Critical Success Index (CSI)} = \frac{1}{\frac{1}{(1\text{-FAR})} + \frac{1}{\text{POD}} - 1}$$

Appendix B:

Questionnaire on Warning Decisions

Questions:

1. What CWA do you warn for? (ex. OUN) __________

If you are not a SOO, please skip ahead to #4.

2. (SOO only) When an SPC outlook forecasts potential severe weather for your CWA, how does your office prepare for warning operations prior to severe weather development?

3. (SOO only) How does your office adjust for warning operations once a watch is issued? If you prepare differently for different types of watches, please specify.

4. A. In your own words, when the SPC issues a Severe Thunderstorm Watch, what do you perceive are their expectations for that Watch? Use the space provided below:

B. Same as A, but for a Tornado Watch:

C. Same as A and B, but for a Tornado Watch with a "Particularly Dangerous Situation":

5. What must your level of confidence be before you issue a Severe Thunderstorm warning while working a shift when deep convection in your CWA is also within a

Enter Percentage, 0-100%

PDS Tornado Watch	Tornado Watch	Severe TSTM Watch	No Watch	No Outlook

6. What level of confidence do you think others in your office must have before issuing a Severe Thunderstorm Warning while working a shift when deep convection in your CWA is also within a

Enter Percentage, 0-100%

PDS Tornado Watch	Tornado Watch	Severe TSTM Watch	No Watch	No Outlook

7. What must <u>your</u> level of confidence be before <u>you</u> issue a Tornado Warning while working a shift when deep convection in your CWA is also within a

Enter Percentage, 0-100%

PDS Tornado Watch	Tornado Watch	Severe TSTM Watch	No Watch	No Outlook

8. What level of confidence do you think <u>others in your office</u> must have before issuing a Tornado Warning while working a shift when deep convection in your CWA is also within a

Enter Percentage, 0-100%

PDS Tornado Watch	Tornado Watch	Severe TSTM Watch	No Watch	No Outlook

Please add any comments on Severe Thunderstorm Watches and subsequent warning decisions in the space provided below:

Please add any comments on Tornado Watches and/or PDS Tornado Watches and subsequent warning decisions in the space provided below or on the backside of this page: