National Weather Association, *Electronic Journal of Operational Meteorology,* 2009-EJ7

Warning Success Rate: Increasing the Convective Warning's Role in Protecting Life and Property

PETER L. WOLF
National Weather Service, Jacksonville, Florida

(Manuscript received 19 August 2009; in final form 10 November 2009)

ABSTRACT

Warning Success Rate, defined as the percentage of the warned population that receives, correctly interprets, and appropriately responds to a given warning message, is a new way to look at the potential impact of warnings on the effort of protecting life and property. Whereas present day efforts to evaluate warning success focus on accuracy and timeliness statistics, Warning Success Rate focuses on warning utilization, with emphasis on three key measures: reception, interpretation, and response. Is an accurate and timely warning meaningful if it is received, correctly interpreted, and utilized by only a small fraction of the warned population?

1. Introduction

A key aspect of the National Weather Service (NWS) mission, to protect life and property from weather-related hazards, is the issuance of accurate and timely forecasts and warnings that provide people with information needed to protect themselves, their families, and their property. Severe thunderstorms and tornadoes are among these hazards, and are the focus of this paper. Weather-related injuries and fatalities in severe thunderstorms, primarily due to higher-impact hazards, have decreased over the past several decades, as illustrated by Fig. 1 for tornado-related fatalities (Brooks and Doswell 2002). During the same period, warning accuracy (Fig. 2) and timeliness improved, and public education efforts increased.

Internal and external measures of "warning success" are evaluated in an effort to document the role of warning accuracy vs. public education in the reduction of weather-related

---

Corresponding author address: Peter L. Wolf, National Weather Service, 13701 Fang Drive, Jacksonville, Florida 32218
E-mail: peter.wolf@noaa.gov

casualties. Will further reductions in such casualties come from increased warning accuracy or from efforts to improve warning reception, interpretation and response? Which direction maximizes warning effectiveness when it comes to achieving the NWS mission? This document focuses on the concept of Warning Success Rate (WSR), a new warning utilization measure that indicates the percentage of the warned population that receives, correctly interprets, and appropriately responds to a given warning message.

2. "Warning Success"…Internal Perspective

Currently in the NWS, the success of the warning program is measured solely by the accuracy and timeliness of warning issuances, relative to event occurrences. Warning verification statistics are the primary mechanism for judging warning success. Improvement goals focus on these statistics.

a. Warning Accuracy and Timeliness

In October 2007, the NWS switched from county-based to storm-based (or threat-based) warnings in an effort to focus on perceived imminent threats, rather than geopolitical boundaries (Fig. 3). Through an active first two years of the program, convective warning polygon size averaged around 600 square miles (mi^2), half of that for earlier county-based warnings (from the NWS Performance Management Website, NWS 2009). Warning lead times and critical success index scores continued to show a slight increase.

One of the primary issues associated with NWS verification statistics is the focus on scientific accuracy over customer receipt, reaction and effective utilization of warnings. Current verification statistics do not take into account human responses to warning information. Warning accuracy and timeliness statistics are not meaningful if the products are not received and utilized by the public in a timely manner.

b. Impact Probabilities within Warning Polygons

A warning program feature rarely studied is the actual probability of being struck by severe weather conditions while within a warning. Given the focus here on high-impact severe weather, which verify a minority of weather warnings (e.g., tornado warning false alarm rate is 0.75), let's assume that 1 out of 4 convective warnings verifies with higher-impact occurrences.

This assumption seems to suggest rather high false alarm rates, and yet warning-based false alarm rates pale in comparison to location-based false alarm rates. As illustrated by Fig. 4, the probability of being struck by, or within a square mile grid block of, high-impact severe weather at any given location within a warning polygon, while not zero, is extremely low. The illustration shows probabilities based on average warning polygon size (x-axis), average event area (y-axis), and the assumption that 1 out of 4 warnings verifies with a high-impact report, (based on data from NWS 2009).

NWS statistics suggest the average warning polygon size is near 600 mi^2. Assuming the area affected by, or within a square mile grid block of, tornadoes and other high-impact severe weather occurrences averages 3 mi^2 (likely less for tornadoes, but can be greater for other event types), the resultant probability of being struck by (or close to) high-impact severe weather, while within a warning area, computes to **1/8th of 1%.** For the net probability, over the course of all warnings, to accumulate to 10%, one would need to receive **80 convective warnings**. In many areas, it may take a decade or longer to receive that many warnings.

How much would the probability of being struck (or within a square mile grid block of a damage path), while within a warning polygon, increase if either polygon average size was cut in half to 300 mi^2, or warning accuracy improved such that 1 out of 2 polygons verified with a high-impact report? Amazingly, either of these challenging improvements would increase the probability to just **$1/4^{th}$ of 1%.** This result questions the benefit gained from further warning precision efforts, from a customer perspective, especially given the added risk of missed events as a result from such efforts. Is there a difference, to the customer, between $1/8^{th}$ of 1% and $1/4^{th}$ of 1% probability? Both are non-zero, but extremely low.

3. "Warning Success"…External Perspective

One benefit to the credibility of NWS warnings is the utilization of different "verification" criteria by the public. Many public reports relate thunderstorm severity to the amount of lightning produced, or the existence of blinding rain and flooding, small hail, or strong (albeit non-severe) winds. These hazards occur over a much greater area. For the same 600 mi^2 average warning area, let's assume that one-fourth of that area, on average, observes one of these conditions during each warning. The resultant probability of observing one of these conditions while within a warning polygon is around 25%, much higher than high-impact severe weather impact probability calculations. However, a negative impact on warning response is likely if the public does not anticipate life-threatening weather conditions during a given warning.

Is there any relationship between warning accuracy and warning utilization? No matter how accurate and timely a convective warning may be, it is essentially a meaningless product if it is not received, understood, and properly responded to by the target audience. A large number of studies (e.g., Gruntfest 1987, Liu et al. 1996, Balluz et al. 2000, Golden and Adams 2000) have examined the social factors that influence warning reception, interpretation and response.

Several studies [e.g., Mileti and Sorenson (1988, 1990), Lindell and Perry (1992, 2004), Mileti 1995] refer to a number of cognitive and behavioral steps taken by the public from warning reception to response, including: **1) hearing** the warning; **2) understanding** the warning; **3) believing** the warning is credible; **4) personalizing** the threat; **5) making a decision** to act; and **6) actually carrying out** the decision. Here, these steps are combined into three external measures of warning success: **reception; interpretation;** and **response.** The NWS has no means to evaluate them, and yet if they are not addressed, improvements to warning accuracy and timeliness may have little impact.

a. Warning Reception

The role of warnings in protecting life and property is minimal if people in the warned area do not receive the warnings in a timely manner. Prior studies (e.g., Balluz et al. 2000, Ewald and Guyer 2002, Paul et al. 2003) have discovered that local commercial TV is the primary source of warning information, often reaching a majority of people within the warned area, with sirens serving as a substantial source in communities that have them. Although the Internet is a rapidly growing source of real-time weather information, it is not a key source of warning information. Weather radio ownership is believed to be in the 5-10 percent range.

The percent of the population in a warning area receiving a certain warning message is believed to vary greatly with time and location, and may also vary based on demographics of the warned population, including size of a given community (Gruntfest 1987, Golden and Adams 2000, Cross 2001). For example, percentages are likely greater in urban areas where local television meteorologists focus more on threat information, and where greater percentages of people have local TV stations, weather radios, etc. compared to rural communities. Several studies (e.g., Balluz et al. 2000, Ewald and Guyer 2002, Paul et al. 2003) have found that, for significant tornado events where there is considerable TV coverage, warning reception may be 70% or greater. For smaller-scale events, non-tornado events, and events in the middle of the night, for which there have been fewer studies, the percentage is expected to be much lower.

b. Warning Interpretation

Warning information is supplied through lengthy text products online, lengthy audio broadcasts, and occasional scrolls on local TV. During rare, high-impact events, local TV stations may interrupt broadcasts to provide detailed warning information, and sirens may be utilized in towns that have them.

Without such broadcast interruptions or sirens, the public is left to interpret the text products, audio broadcasts and scrolls. Good reading/listening skills, and an understanding of the local geography, are needed to properly interpret warning messages (e.g., determine if your location is in danger). How does a person in a given town interpret a warning received in which that town is not mentioned as being in danger, but other towns are mentioned?

Lengthy warning messages and weather radio broadcast cycles, brief TV scrolls that display only a portion of the warning message, and the average person's listening (or reading) skills and geographic knowledge are obstacles to effective public interpretation of warning messages. Although few studies address warning message interpretation, it is reasonable to estimate that based on the communication obstacles that exist, less than 50% of those who receive the warning may correctly interpret the message, understand where the threat is, what the threat is, and what the recommended actions are (for any given location, not just along the severe thunderstorm's path).

The percentage is reduced even further when considering how many of those who understand the message go on to believe and personalize the threat, especially after recently receiving false alarms. Figure 4 shows that the probability of being struck by, or close to, high impact severe weather while under a warning, is extremely low, likely less than ¼ of 1%. The high false alarm rate received locally, from the customer viewpoint, negatively impacts believability and personalization of the threat, which ultimately negatively affects response.

c. Warning Response

While considerable public awareness efforts have been ongoing for many years to improve response to warnings, **desensitization** serves to diminish public response by impacting threat credibility and personalization. Given the extremely low probability of being struck by life-threatening convective weather (Fig. 4), and the number of warnings one must receive for the net probability (over the course of all warnings issued) to reach just 10%, desensitization may reduce the percentage of people responding to the warning to well below 50% of those receiving and correctly interpreting it.

A number of prior studies have documented the impact of desensitization on public response to emergency messages. For example, Breznitz (1984) demonstrated that the human response to false alarms included reduced probability of engaging in protective measures, increased latency in taking protective measures, and reduced level of protective measures taken. Other studies (e.g., Zabyshny and Ragland 2003) related the negative impact of false alarms on response to traffic warning systems.

Even so, recent research (e.g., Barnes et. al. 2006) suggests the public has a high false alarm tolerance. Only a few convective warnings are received in a given location each year, rather than one every day. This, and the criteria used by the public to define a thunderstorm as "bad" or "severe," may make desensitization due to false alarms less substantial than what the results of desensitization studies might otherwise suggest. However, such high tolerance can be quickly eroded by a few late night warning alerts in areas not in the path of the warned storms, which can lead people to turn off weather radios or otherwise ignore future weather threats.

d. Warning Success Rate

Warning Success Rate (WSR) is defined here as *the percentage of the warned population that receives, correctly interprets, and appropriately responds to a given warning message*. Precise WSR numbers are unknown since reliable percentages for warning reception, accurate interpretation, and appropriate response are very difficult to acquire. Given the reasonable estimate that, on average, 50-75% of the warned population receives the warning message, less than 50% of this group correctly interpreting the message, and less than 50% appropriately responds to it (thus 25% or less for correct interpretation and response combined), the resultant WSR is likely to average as low as 10-15%. This suggests that 85-90% of the population within a given warning is not successfully reached. **To what extent can a warning message protect life and property if 90% of the warned population does not receive it, understand it, or appropriately respond to it?**

Figure 5 shows WSR results based on chosen assumptions for warning reception, interpretation, and response percentages. It illustrates the challenge of achieving WSR values of at least 50%. Large percentages of warning reception, interpretation, and response are needed just to get the WSR above 20%. Even so, the warning would not be "successful" for a large majority of the warned population. However, improvements in reception, interpretation, and response percentages that yield WSR changes from less than 5% to greater than 20% would still be substantial, indicating a greater role for the warning in protecting life and property.

4. Discussion

What makes a warning "successful?" Internal to the NWS, an accurate warning that offers sufficient lead time gets such a designation. However, externally, WSR may be more important. A high (low) WSR suggests the warning likely plays a key (minimal) role in the protection of life and property for a given event, regardless of accuracy and timeliness.

Further study is needed to better understand the typical percentages of a warned population that receive, understand, and appropriately respond to a given weather warning message, as well as to come up with a precise percentage criterion for labeling a warning as "successful." Available evidence suggests that WSR scores are likely to be low, perhaps 10% or less, suggesting that the large majority of a given warned population may not receive, understand, or respond to a given warning message. Is an accurate, timely warning effective if less than 10% of the warned population receive, understand, and respond to it?

A successful warning requires all three warning utilization measures described (reception, interpretation, response). Failure in just one of these measures dramatically reduces WSR (Fig. 5). Efforts to improve these measures, given existing technologies, stand to substantially increase the role of warnings in fulfilling the NWS mission at a smaller price tag compared to that for the science and technology improvements needed to yield even small additional gains in warning accuracy and timeliness.

Increased warning reception is possible by focusing dissemination efforts on the technologies accessed by a majority of the public, including the Internet, telephone, and television. Several private sector companies provide NWS warning information through telephone or online alerts.

Improved threat interpretation could be achieved by shifting focus from text format to graphical displays to more easily define the areas in danger and the time frame of the threat. Graphical displays can relay more information faster and more precisely compared to reading, or listening to, multiple paragraphs of text information. Probability displays that define areas of greatest risk, threat confirmation via plotted storm reports, and real-time radar imagery are examples of technology that can enhance threat credibility and personalization in the greatest threat area within a warning polygon.

Improved warning response could be achieved through a greater focus on public education efforts. A considerable increase in such efforts has been observed nationwide over the past decade. Continued efforts should focus not only on improved response to weather warnings, but also to threatening weather conditions in situations when warnings cannot be received.

As suggested on Fig. 5, efforts that double warning reception rates, accurate interpretation percentages, and appropriate response percentages combine to increase a warning's effectiveness in protecting life and property by a factor of eight. Would a less accurate or timely warning, that successfully reaches 20% of the warned population, play a greater role in protecting life and property than a more accurate or timely warning with a success rate of less than 5%? Efforts that increase WSR may achieve a greater "*bang for the buck*" when it comes to increasing the role of convective warnings in achieving the key NWS mission compared to efforts primarily focused on increasing warning accuracy and timeliness.

Acknowledgements - The author wishes to thank the four reviewers, over the course of two separate reviews, whose comments helped improve the manuscript.

REFERENCES

Balluz, L. et al. 2000. Predictors for people's response to a tornado warning: Arkansas, 1 March 1997. *Disaster.* **24 (1)**, 71-77.

Barnes, L.R. et. al. 2006: False alarms and close calls: A conceptual model of warning accuracy. *Wea. Forecasting*, **22**, 1140-1147.

Breznitz, S. 1984: *Cry Wolf: The Psychology of False Alarms*. Hillsdale, New Jersey, Lawrence Erlbaum Associates.

Brooks, H.E. and C.A. Doswell III 2002: Deaths in the 3 May 1999 Oklahoma City tornado from a historical perspective. *Wea. Forecasting*, **17**, 354-361.

Cross, J.A. 2001: Megacities and small towns: Different perspectives on hazard vulnerability. *Environmental Hazards*, **3**, 63-80.

Ewald, R. and J. Guyer 2002: *The Ideal Lead Time for Tornado Warnings – A Look from the Customer's Perspective.* U.S. Dept of Commerce. 4pp.

Golden, J.H. and C.R. Adams 2000: The tornado problem: Forecast, warning, and response. *National Hazards Review*, **1 (2)**, 107-118.

Gruntfest, E. 1987: Warning Dissemination and Response with Short Lead-Times. *Hazard Management – British and International Perspectives.* J. Handmer, Ed., Norwich, Geo Books, 191-202.

Lindell, M.K. and R.W. Perry 1992: *Behavioral Foundations of Community Emergency*

Planning. Washington DC, Hemisphere Publishing Corp.

Lindell, M.K. and R.W. Perry 2004: *Communicating Environment Risk to Multi-ethnic Communities*. Thousand Oaks CA, Sage.

Liu, S.L. et al. 1996. Assessment of a severe weather warning system and disaster preparedness, Calhoun County, Alabama, 1994. *American Journal of Public Health*. **86**, 87-89.

Mileti, D. and Sorensen, J. 1988: Planning and implementing warning systems. *Mental Health Care In Mass Emergencies: Theory And Practice*, M. Lystad, Ed., New York, Brunner/Mazel Psychological Stress Series, 321-345.

Mileti, D.S. and J.H. Sorenson 1990: *Communication of Emergency Public Warnings: A Social Science Perspective and State-of-the-Art Assessment*. Washington D.C.: FEMA.

Mileti, D.S. 1995: *Factors Related to Flood Warning Response.* U.S.-Italy Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods**.** 17 pp.

NWS 2009: *Performance Management Website*. https://verification.nws.noaa.gov. Available via internal access only.

Paul, B.K. et al. 2003: *Public Response to Tornado Warnings: A Comparative Study of the May 4, 2003 Tornadoes in Kansas, Missouri, and Tennessee*. Quick Response Report #165. Univ. of Colorado.

Zabyshny, A.A and D.R. Ragland 2003: *False Alarms and Human-Machine Warning Systems*. U.C. Berkeley Traffic Safety Center, Institute of Transportation Studies. 13pp.

FIGURES

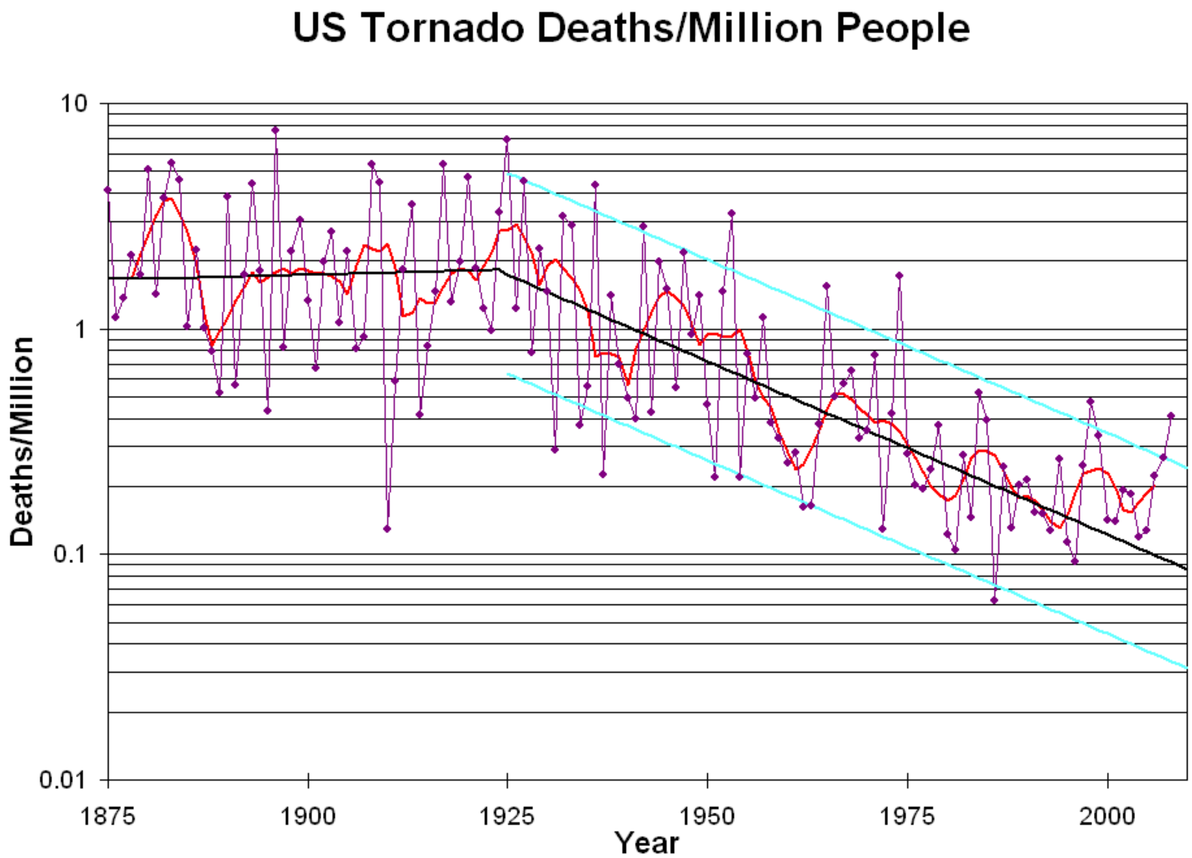

Figure 1. Normalized tornado-related fatalities across the United States, from ***Deaths in the 3 May 1999 Oklahoma City Tornado from a Historical Perspective*** (Brooks and Doswell, 2002). Image was updated by Brooks to add additional data from 2000-2008.

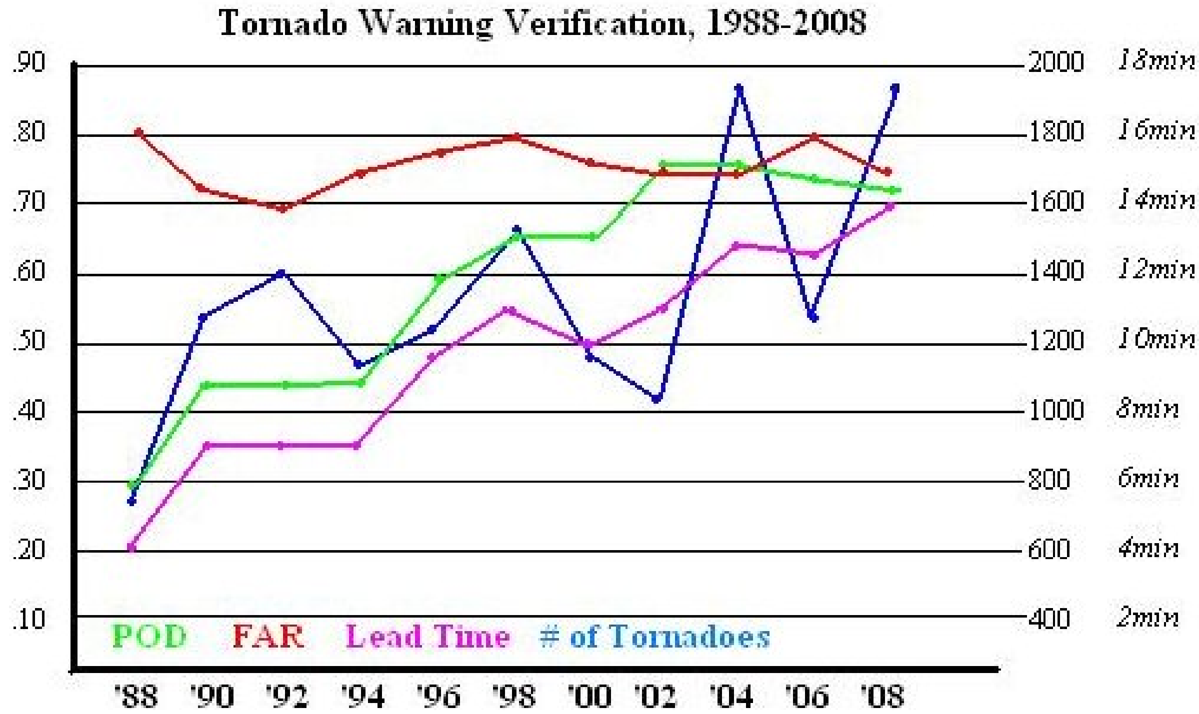

Figure 2. Trends of national verification scores for tornado warnings, based on data plotted every 2 years from 1988 through 2008, based on data from the NWS verification website.

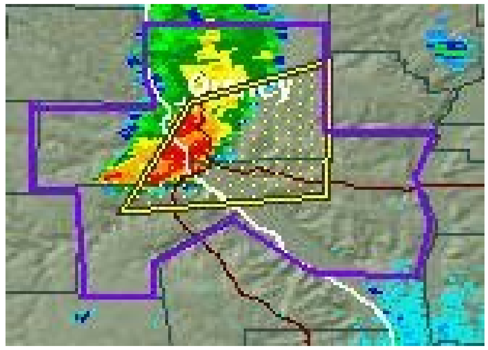

Figure 3. Example of a convective warning polygon (yellow), compared to the much larger county-based warning area (purple).

Warning Area (in square miles)

Area (in square miles) affected by (or close to) high-impact severe weather	100	200	300	400	500	600	700	800	900	1000
1	1/4th	1/8th	1/12th	1/16th	1/20th	1/24th	1/28th	1/32nd	1/36th	1/40th
3	3/4th	3/8th	1/4th	3/16th	3/20th	1/8th	3/28th	3/32nd	1/12th	3/40th
5	1.25	5/8th	5/12th	5/16th	5/20th	5/24th	5/28th	5/32nd	5/36th	1/8th
10	2.5	1.25	5/6th	5/8th	1/2	5/12th	5/14th	5/16th	5/18th	1/4th
20	5.0	2.5	1.67	1.25	1.0	5/6th	5/7th	5/8th	5/9th	1/2
30	7.5	3.75	2.5	1.87	1.5	1.25	1.0	15/16th	7/8th	3/4th
40	10.0	5.0	3.35	2.20	2.0	1.67	1.43	1.25	1.11	1.0
50	12.5	6.25	4.2	3.13	2.5	2.0	1.80	1.56	1.39	1.25

Severe Weather Strike Probabilities (in percent)

Figure 4. Probability of being within a square mile grid block of high-impact severe weather, based on average warning polygon size (X-axis), average area affected by high-impact severe weather (Y-axis), and the assumption that approximately one of four warnings verifies with a tornado or other high-impact severe weather occurrence.

Correct Warning Interpretation and Response

Warning Received	1%	10%	20%	30%	40%	50%	60%
1%	.01	.10	.20	.30	.40	.50	.60
10%	.10	1.0	2.0	3.0	4.0	5.0	6.0
20%	.20	2.0	4.0	6.0	8.0	10.0	12.0
30%	.30	3.0	6.0	9.0	12.0	15.0	18.0
40%	.40	4.0	8.0	12.0	16.0	20.0	24.0
50%	.50	5.0	10.0	15.0	20.0	25.0	30.0
60%	.60	6.0	12.0	18.0	24.0	30.0	36.0
70%	.70	7.0	14.0	21.0	28.0	35.0	42.0
80%	.80	8.0	16.0	24.0	32.0	40.0	48.0

Warning Success Rate (WSR)

Figure 5. Warning Success Rate - Percentage of warned population that receives, correctly interprets, and properly responds to a given convective warning message, as related to warning success. Green percentages indicate a successful warning issuance, while red and orange values indicate marginal utilization of the warning message. Note that there are a number of combinations of interpretation and response that can yield the percentages on the x-axis. For example, if 50% interpret the warning correctly, and 50% of this group responds appropriately, the x-axis value would be 25%. If 70% interpret correctly, and 70% of this group responds appropriately, the x-axis value would be just below 50%.