THE EXPERIMENTAL WARNING PROGRAM

2022–2023 EXPERIMENT SUMMARY

NOAA Hazardous Weather Testbed, Norman, OK

2 April 2024

Kodi Berry[1], Anthony Lyza[1], Kristin Calhoun[1], Abby Bitterman[2], Sean Ernst[2,3,4], Jack Friedman[5], Burkely Gallo[3,4], K. Klockow-McClain[3,1], Makenzie Krocak[1], Daphne LaDue[6], C. Ling[7], Eric Loken[3,1], Holly Obermeier[3,1], Joe Ripberger[2], Thea Sandmæl[3,1], Michelle Saunders[6], Miranda Silcott[3,1], Brandon Smith[3,1], Greg Stumpf [8,9], Kevin Thiel[3,4], Anna Wanless[2], and Katie Wilson[3,1]

[1]NOAA/OAR/National Severe Storms Laboratory
[2]University of Oklahoma Institute for Public Policy Research and Analysis
[3]University of Oklahoma/Cooperative Institute for Severe and High-Impact Weather Research and Operations
[4]NOAA/NWS/NCEP Storm Prediction Center
[5]University of Oklahoma/Center for Applied Social Research
[6]University of Oklahoma/Center for Analysis and Prediction of Storms
[7]University of Akron
[8]Colorado State University/Cooperative Institute for Research in the Atmosphere
[9]NOAA/OAR/Global Systems Laboratory

TABLE OF CONTENTS

1. INTRODUCTION

The NOAA Hazardous Weather Testbed (HWT) is a joint project of the National Weather Service (NWS) and the National Severe Storms Laboratory (NSSL). The HWT provides a conceptual framework and a physical space to foster collaboration between research and operations to test and evaluate emerging technologies and science for NWS operations. The HWT emerged from the "Spring Program" which, for more than a decade, tested and evaluated new forecast models, techniques, and products to support NWS Storm Prediction Center (SPC) forecast operations. Now, the HWT consists of two primary programs: the original Spring Program, which is part of the Experimental Forecast Program (EFP), and the Experimental Warning Program (EWP).

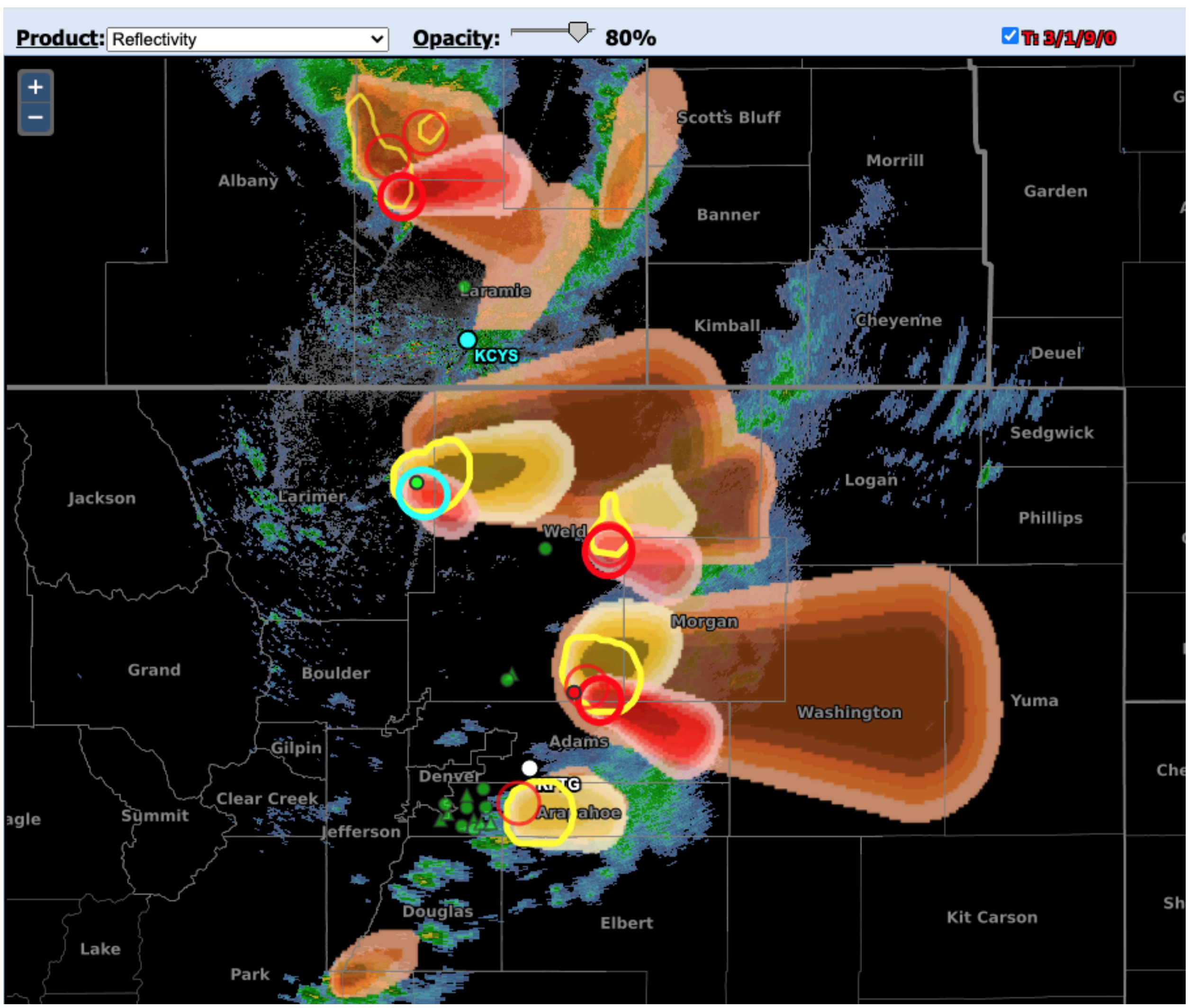


Figure 1. A forecaster's view of the PHI Prototype Tool as it is used to manage storm objects and probability plumes for an archived case.

The EWP tests and evaluates new applications, techniques, and products to support Weather Forecast Office (WFO) severe convective weather warning operations. This was the fourteenth year for warning activities in the HWT. Feedback was gathered from NWS operational meteorologists. The experiment participants issued experimental warnings, published live blogs, engaged in focus groups, and completed surveys. User comments were also collected during shifts, which have been used to inform product development. This input is vital to improving the NWS warning process, which ultimately leads to saved lives and property.

2. OVERVIEW

Due to the COVID-19 pandemic, all in-person HWT activities were conducted virtually in 2022. In 2023, EWP experiments returned to a mix of in-person and virtual experiences. The HWT EWP hosted seven experiments over 20 calendar weeks in 2022 and five experiments over 14 weeks in 2023 to improve NWS severe weather forecasts and warnings.

Cloud services were procured using NOAA's cloud contract vehicle. This contract allows for the investigation and implementation of remote HWT experiments using AWIPS in a cloud environment. With NOAA's inclusiveness of remote working employees and partners we feel that investigating methods to improve the opportunity and collaboration with as many NOAA employees and partners for the HWT is paramount.

Table 1. Details for the 2022–2023 Experimental Warning Program.

EWP Experiment	Dates	Length	Number of Participants
Brief Vulnerability Overview Tool	10–14 Jan 2022 28 Feb–4 Mar 2022 7–11 Mar 2022 19–23 Sep 2022 3–7 Oct 2022	5 weeks	29 forecasters 36 emergency managers
Tiny Threats-in-Motion	14–18 Mar 2022 21–25 Mar 2022 13–17 Feb 2023 27 Feb–3 Mar 2023	4 weeks	10 forecasters
PHI Prototype	25–29 Apr 2022 2–6 May 2022 9–13 May 2022 1–5 May 2023 8–12 May 2023 15–19 May 2023	6 weeks	37 forecasters
Satellite Convective Applications	1–4 Jun 2022 7–11 Jun 2022 14–18 Jun 2022 22–26 May 2023 5–9 Jun 2023 12–16 Jun 2023	6 weeks	42 forecasters

Hazard Services - Threats-in-Motion	19–23 Jul 2022 2–6 Aug 2022 30 Aug–3 Sep 2022 17–21 Apr 2023 1–5 May 2023 8–12 May 2023	6 weeks	15 forecasters
End-User Decisions Over Time	8–11 Aug 2022 15–18 Aug 2022	2 weeks	15 broadcast meteorologists 10 emergency managers
Convective Outlook Innovations	11–13 Oct 2022 18–20 Oct 2022	2 weeks	18 forecasters 11 emergency managers 8 public
Watch-to-Warning Experiment	14–17 Aug 2023 21–24 Aug 2023 5–8 Sep 2023	3 weeks	9 WFO forecasters 5 SPC forecasters

3. PROJECT DETAILS AND RESULTS

Brief Vulnerability Overview Tool Experiment

Summary by Jack R. Friedman, Michelle E. Saunders, and Daphne S. LaDue

Figure 1. BVOT research team convened to conduct a virtual HWT experiment.

Overview

This HWT project applied and integrated relevant social and behavioral science methodologies to assesses WFO forecasters' and end-users' abilities to assess, understand, and respond effectively to 1) experimental, short-term (through Day 1) forecasts for convective weather hazards and 2) a tool that will enhance their awareness of vulnerabilities within their County Warning Area (CWA). Specifically, we assessed how 1) new Storm Prediction Center's ability to *identify enhanced threat corridors* within Day 1 Probabilistic Convective Outlooks and *communicate those in event-driven Outlook updates in temporally disaggregated Day 1 Outlooks* and 2) Friedman/LaDue's *Brief Vulnerability Overview Tool (BVOT) will impact WFO forecaster product issuance and messaging* with deep core partners (e.g., Emergency Managers) as well as how EMs interpret these products. This project tested the impact of increased vulnerability knowledge and awareness of CWA-based vulnerabilities on how 1) WFO products are issued and 2) EMs and other deep core partners interact with WFOs and (independently) interpret new temporally fine-grained SPC Outlooks in a simulated operational environment. This project assessed impacts on WFO forecaster and EM behavior by simulating end-to-end severe weather communication — SPC to WFO to EMs — through realistic experimental scenarios involving SPC and WFO forecasters and EMs.

We hypothesized: 1) that there would be a measurable difference in the content, timing, and frequency of messaging to deep core partners (e.g., EMs) due to consideration of spatially explicit vulnerability data (via the BVOT); 2) that there would be a difference in the content of forecasts, messaging to the public, and formal briefings provided to deep core partners associated with including new Storm Prediction Center Severe Timing Guidance products; 3) that the impact of both tools being used simultaneously would have an additive effect on changes in messaging and formal product issuance; and 4) that these changes in messaging and briefings would be positively perceived by Emergency Managers and will result in improved decision making.

The SPC Severe Timing Guidance Product

Our HWT experiment included testing the Storm Prediction Center's (SPC) new prototype graphics to provide automated timing information. These graphics were created using the HREF and SREF models and provide hourly updated guidance using a rolling 4-hour window of time. The time frame of these graphics is valid for the traditional SPC Day 1 severe weather outlook. We tested seven graphics in total, five static images and two animations (Fig. 2 & Fig. 3). The SPCs motivation is to create a nationally consistent database for severe weather timing information that can be used both internally at the WFO and to aid in messaging their core partners. These timing graphics begin to help communicate the timing of severe weather threats while maintaining consistency across spatial boundaries.

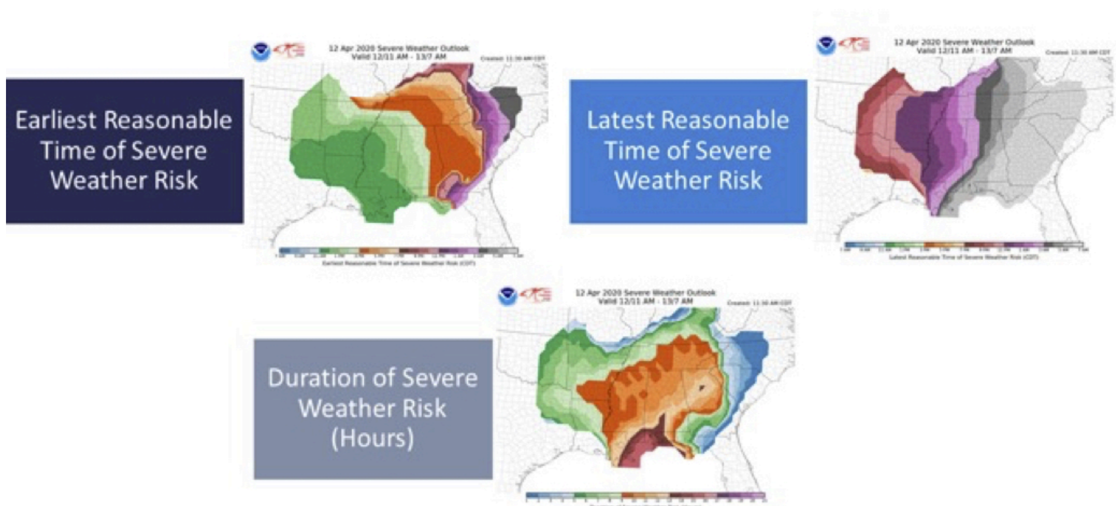


Figure 2. Three of seven experimental SPC timing graphics tested in our HWT experiment.

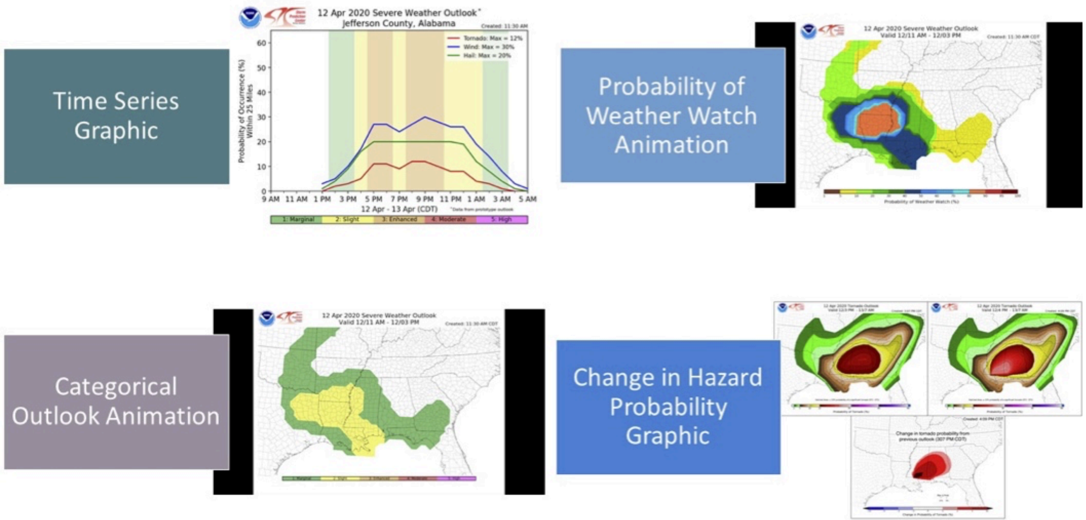


Figure 3. The remaining four of seven experimental SPC timing graphics tested in our HWT experiment.

Our experimental design uses several methods and data collection techniques including participant observation, post case debriefings, and end of day and week whole group discussions. We utilize two surveys, the System Usability Scale (SUS) survey to measure how the usability of the graphics changes from the beginning of the experiment week to the end of the week. Finally, we use a post-experiment survey to gather any final information forecaster participants want to share about each of the seven graphics.

The Brief Vulnerability Overview Tool (BVOT)

The Brief Vulnerability Overview Tool (BVOT) is a GIS-based, graphical map overlay that can display specific, place-based, known vulnerabilities across a National Weather Service (NWS) Weather Forecasting Office's (WFO) County Warning Areas (CWA). The goal of the development of the BVOT is that it will provide additional spatial situational awareness to NWS WFO meteorologists by allowing them to quickly assess whether a weather hazard will be directly threatening a specific vulnerability. This will provide NWS meteorologists with a visual display that will allow them to quickly assess whether they will need to provide enhanced or specifically-tailored messaging to their partners (e.g., emergency managers, law enforcement, etc.).

The NWS has supported research to create a more generalizable and standardizable way of operationalizing vulnerability information for NWS meteorologists (Fig. 4). Many of these efforts focus on converting vulnerability data that was originally collected for non-meteorological purposes — various Social Vulnerability Indexes — into a format that can be used to inform NWS meteorologists' IDSS efforts. Many of these efforts take existing databases like the CDC's SVI and develop ways of mapping and visualizing census-derived vulnerability data — like households in poverty, population without a high school education, or percentage of residences that are mobile or manufactured housing — in order to provide spatial awareness of these vulnerabilities to operational meteorologists. However, because of the aggregated nature of these vulnerability data — aggregated spatially across census tracts and thematically by bringing together multiple vulnerabilities to develop social vulnerability "themes" — it is more difficult to translate these hypothesis-derived vulnerability data (HDVD) into actionable decision support for WFO-level, operational meteorologists who need to communicate Decision Support Service-oriented messaging to their core partners.

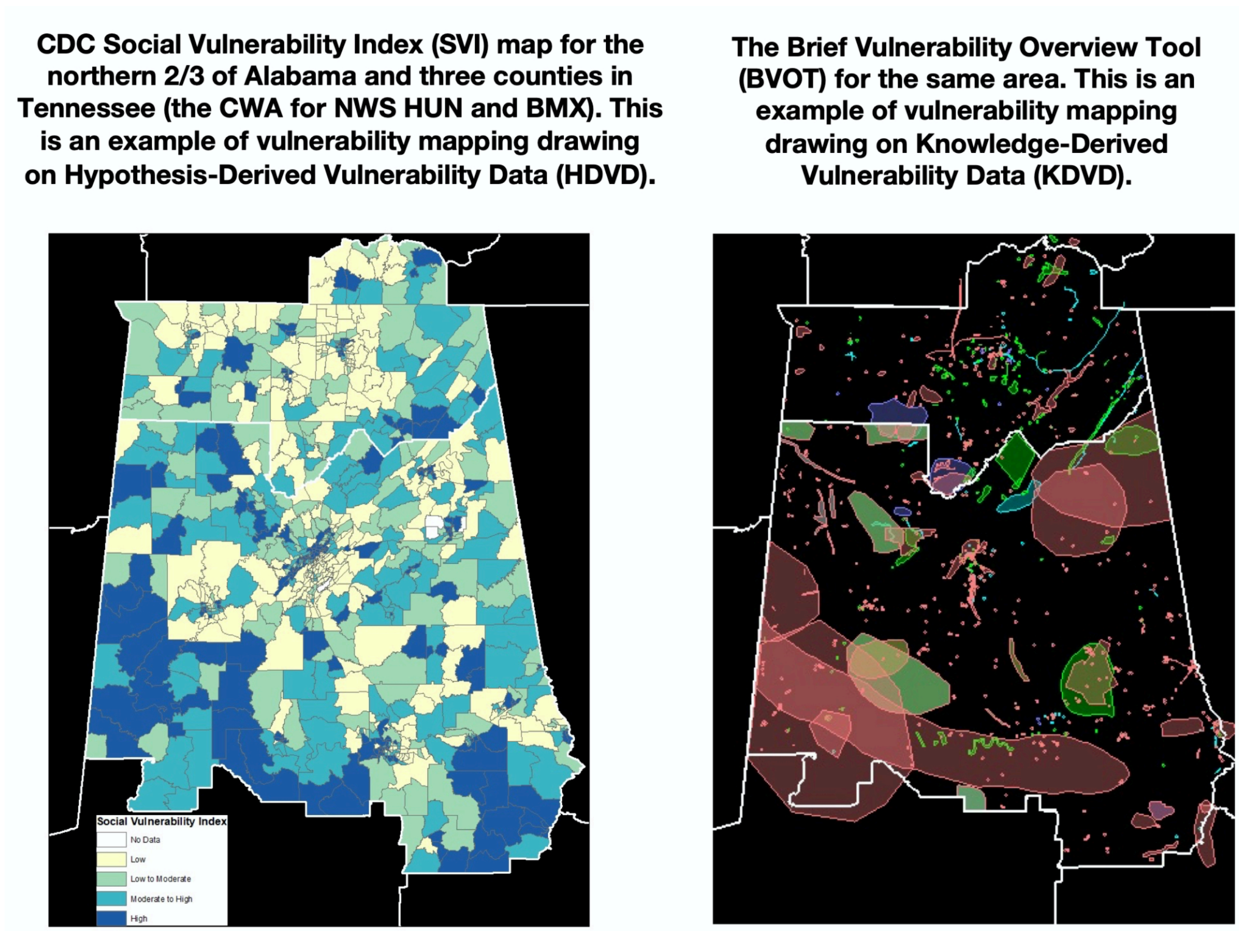


Figure 4. Comparison of the types of visualizations of different types of vulnerability data, in this case, comparing the CDC SVI with the Brief Vulnerability Overview Tool (BVOT).

The Brief Vulnerability Overview Tool, or BVOT, was developed as a way of balancing the strengths of the HDVD-approaches (like those that derive vulnerabilities from the CDC's SVI) with the specificity of something like the NWS's Impacts Catalog, while addressing weaknesses that have been identified by NWS meteorologists. The BVOT is composed of discrete, knowledge-derived vulnerability data (KDVD). This can be contrasted with general, hypothesis-derived vulnerability data (HDVD) that, as noted above, aggregates census tract-level data. In other words, what distinguishes KDVD from HDVD is that KDVD is more spatially fine-grained and the vulnerabilities that have been identified are known by experts to be at risk of being impacted by specific weather hazards. In our HWT experiments, we have sought to understand how, when, and if added vulnerability awareness changes the 1) decision-making, 2) messaging, and/or 3) product issuance of NWS WFO-level meteorologists.

Experiment Details and Results

The experiment involved the participation of both NWS WFO-level meteorologists and emergency managers from a range of different backgrounds (e.g., county, municipal, higher education, military/federal, etc.). Each week of the experiment (3 in 2021, 5 in this report period) involved 6 NWS meteorologists (18 total in 2021) and 6-12 EMs recruited (30 total in 2021) from around the country. Each week was divided into 3 sections: 1) the first half of Day 1 involved orientation to the project, the SPC tool, and the BVOT; 2) the second half of Day 1 through the first half of Day 5 involved participants in decisions making, messaging, and product issuance based on 8 recorded (WES) severe weather cases (more details below); and 3) the second half of Day 5 was used to conduct focus groups to collect data on the overall impressions, usability, value, and concerns about the SPC and BVOT tools.

Experimentally, the study was designed around the NWS meteorologists who are divided into 3 teams of two meteorologists, with each team rotating through different experimental conditions through the week (Fig. 5):

- Condition A: BOTH the SPC Experimental Severe Timing Graphics AND the BVOT
- Condition B: Only the BVOT
- Condition C: Only the SPC Experimental Severe Timing Graphics

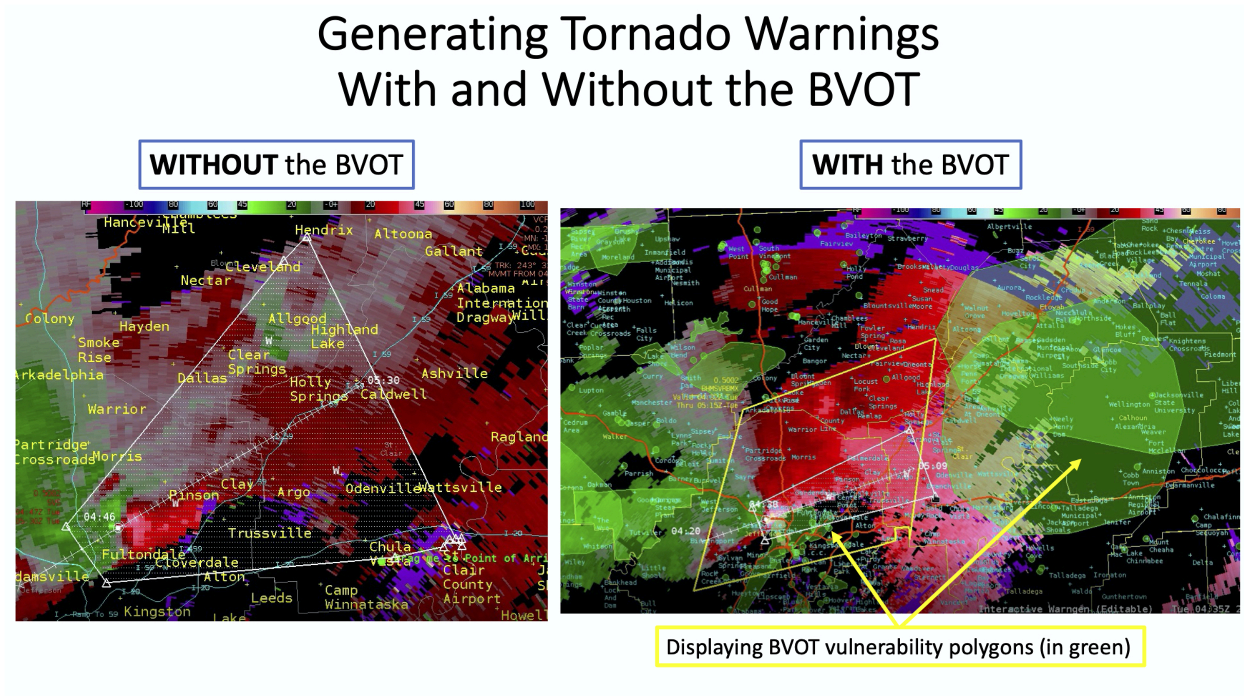


Figure 5. Example screenshots of the difference between a participants' AWIPS screens during the storm-on-the-ground period without (left) and with (right) the BVOT display.

Because we were assessing when these tools might be used or of value to NWS meteorologists in providing added or more focused messaging to their partners, each of the teams of meteorologists was tasked with making decisions based on different time periods in the evolution of a severe weather threat. These time periods were:

- 24-48 hours before the event
- 4-12 hours before the event
- Storm-on-the-ground (a 45-minute period during which active severe weather is on-the-ground, requiring critical engagement with AWIPS-in-the-cloud instances that are hosted on Amazon Web Services (AWS) and managed through the HWT)

Specifically, in each of the three time periods during each of the 8 cases, each team was given 45 minutes to review briefing material and then provide briefing packets, provide other forms of messaging to partners/publics (the project uses Slack to simulate both NWSChat and social media issuances), and, when relevant, issue formal NWS warning products. At the start of each period, teams were provided a briefing packet that contains SPC outlooks, SPC discussions, as well as the local NWS WFO's area forecast discussions (AFDs), and, during "storm-on-the-ground" periods, a radar loop showing the evolution of the event in order to provide the NWS meteorologists (and the EMs) realistic situational awareness on each event as it unfolds. While all EMs received the most detailed briefing packets for each of the events, those NWS teams that were in Condition B will NOT receive the SPC Experimental Severe Timing

Graphics, while those NWS teams that were in Condition C were NOT able to access the BVOT vulnerability map layers. In this way, we were able to record any differences that we were seeing in the decision making, messaging, and/or product issuance correlated with different Conditions/access to these experimental tools.

Data collection included:

- Recordings of all meteorologist decision-making through interactions (via recorded Google Meets) between meteorologists
- Recording all AWIPS instances (via screencapture) for each meteorologist throughout all case during all periods
- Recording all EM discussions and decision-making (via Google Meets) at each stage of the experiment
- Recording all interaction between meteorologists and EMs (including the briefing material/packets provided, any live webinar-style briefings (via Google Meets), and any text-based messaging or interaction (via Slack, representing both NWSChat and social media)
- The Post Study System Usability Questionnaire (start and end of the week)
- The NASA Task Load Index Survey (after each of the 8 cases)
- The Confidence Continuum Survey (after each of the 8 cases)
- The Secondary Traumatic Stress Survey (modified) (start and end of the week)
- Debriefing interviews with each team of meteorologists after each case period (24 total)
- Post-case debriefings at the end of each day involving both the meteorologists and the EMs
- Post-study survey gathering additional information about the usability and preferences regarding the SPC Severe Timing Guidance Products (for the NWS meteorologists)

SPC Severe Timing Graphics — Early Results

After preliminary analyses, we found that the time series graphics (see Fig. 6) are valued the most out of the seven graphics. The time series graphics display the probability of occurrence (%) within 25 miles for tornadoes, wind, and hail over time, along with the categorical outlook. Participants stated that they would feel comfortable sharing the time series graphics with core partners (Fig. 7).

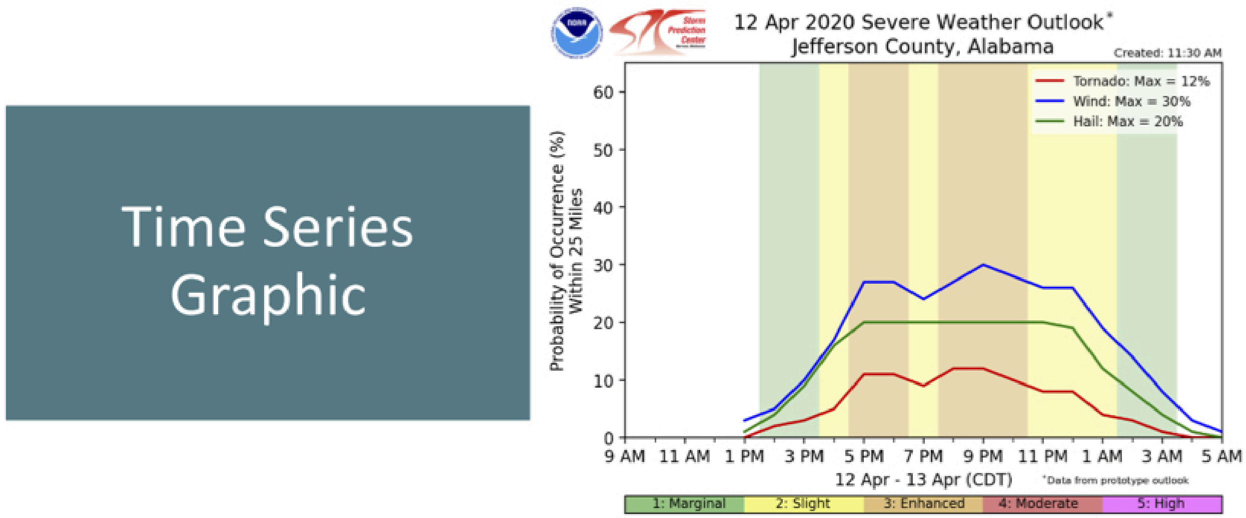


Figure 6. Example of a point-based time series graphic from the SPC's Severe Timing Graphic tool.

The time series graphics were also shared the most with EMs during our experimental cases. The earliest reasonable timing of severe weather risk graphic was the next most valued graphic by both forecaster and EM participants. Forecasters liked the graphical representation of severe weather timing and EMs stated that this graphic could be important for scheduling staffing and preparing internally before a severe weather event. So far, most forecaster and EM participants do not find the probability of a weather watch graphic (Fig. 3) to be as useful, especially compared to the other timing graphics being tested. One forecaster concluded their thoughts on this graphic, stating that, "It's good information to have, but we can draw conclusions based on everything else."

Participants described wanting to be able to customize the SPC timing graphics before sharing them with their partners and provided examples for how they would do this. Each group of forecasters have expressed interest in wanting to see the timing guidance for the maximum risk category for a weather event. At present, several graphics were designed to show the timing for the slight risk category.

Finally, the SUS scale (Fig. 8) indicated that there was an increase in inconsistency when using the timing graphics from the beginning to the end of the week. This inconsistency is most likely due to the peak categorical risk not appearing in some of the timing graphics, which was highlighted by participants during debriefings, group discussions, and in the post-experiment survey.

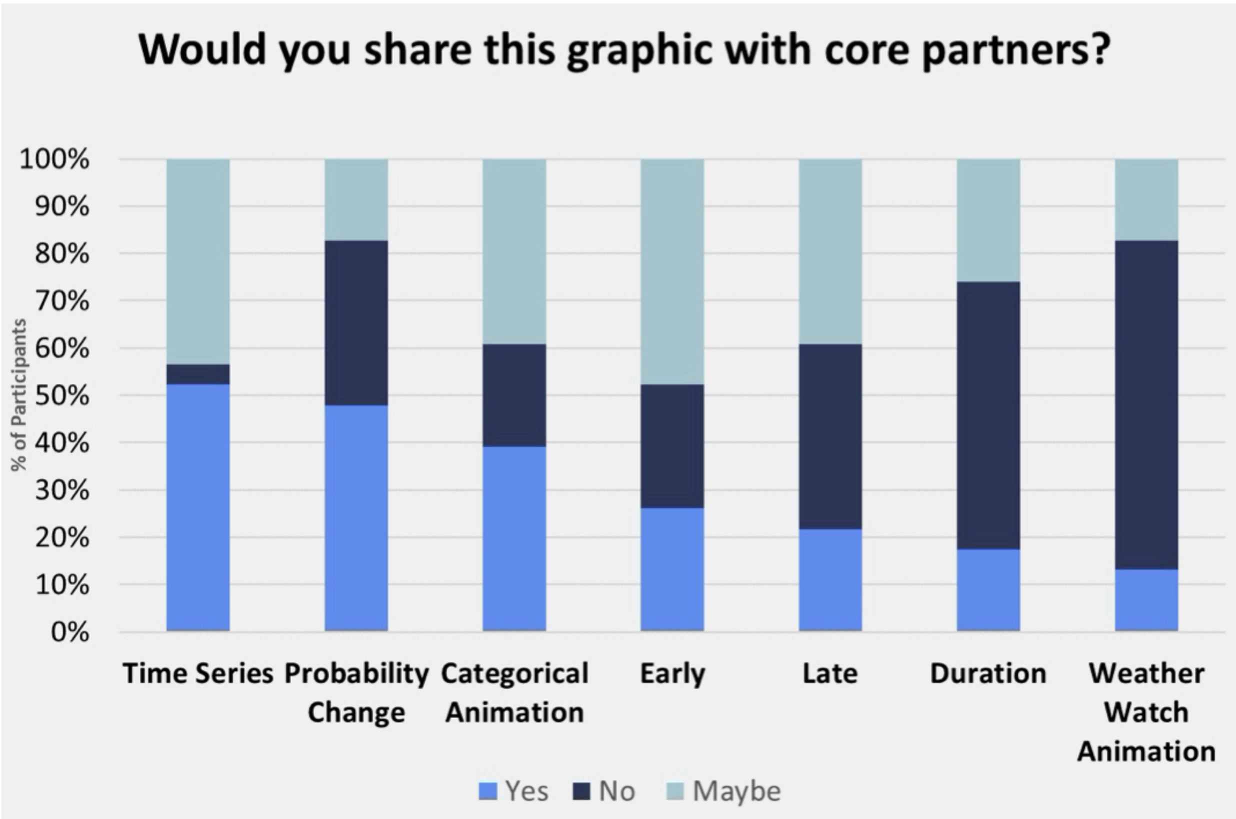


Figure 7. Post-experiment survey data from forecaster participants on whether they would share each experimental graphic with their core partners.

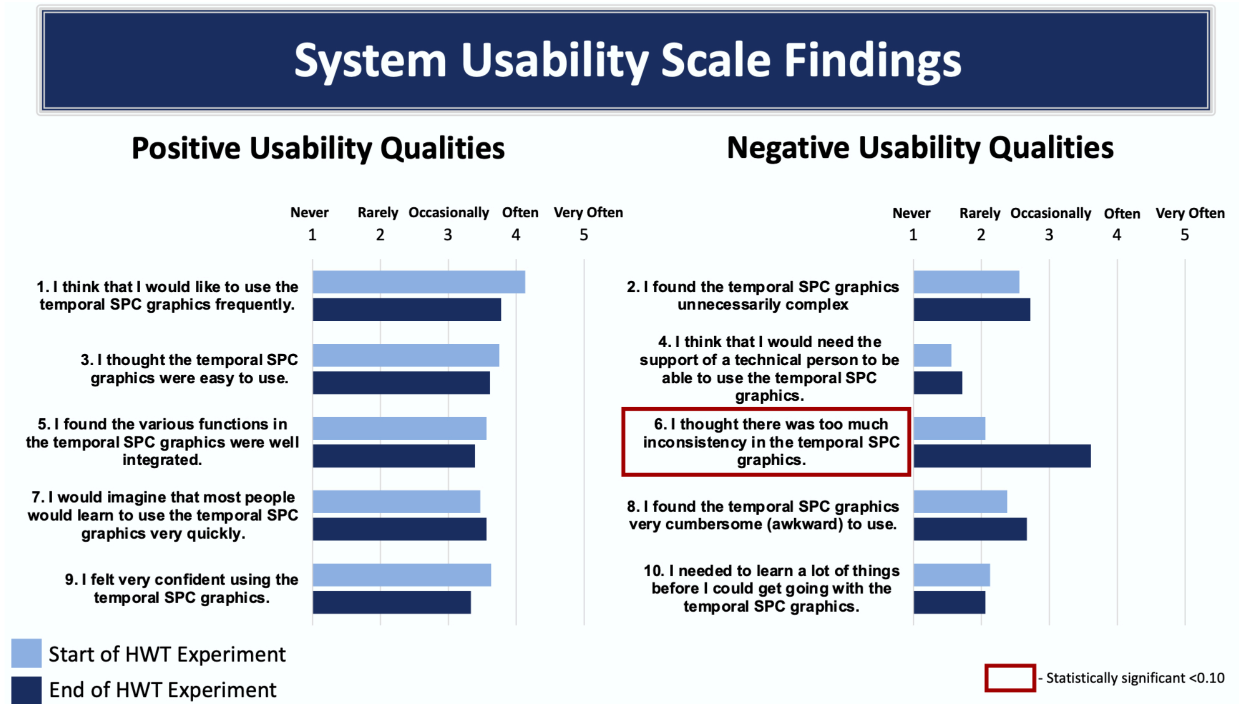


Figure 8. Summary of positive and negative usability qualities reported by NWS meteorologists regarding SPC Severe Timing Graphics.

Brief Vulnerability Overview Tool (BVOT) — Early Results

We have been evaluating (ongoing) the first three weeks of HWT experiments, focusing on four critical questions regarding the impact of increased awareness of vulnerabilities on NWS meteorologist operations:

- What impact does the BVOT have on the nature of messaging to/with core partners?
- What impact does the BVOT have on NWS meteorologists' understandings of the impacts of a severe weather event?
- What impact does the BVOT have on the practices that occur before, during, and after severe weather operations?
- What impact does the BVOT have on the issuance of formal NWS products (watches, warnings, etc.)?

We are in the middle of analyzing these data, so, while we have significant anecdotal evidence to begin to answer some of these questions, we will focus on discussing the last of these questions — What impact does the BVOT have on the issuance of formal NWS products (watches, warnings, etc.)? — since answering this question is critical for a number of reasons. First, we did not design the BVOT to strongly impact the decision to issue or not to issue a formal NWS product. Rather, the BVOT is designed to provide added awareness regarding the potential impacts of weather hazards on discrete, vulnerable people, places, or things. As such, we did not expect to see a significant impact on the decision to or the timing of formal product issuance. In general, this is what we have found, so far, in this experiment (Fig. 9). For instance, with our Case 4, a recording of the data from a 13-14 April 2019, overnight severe weather event that impacted NWS HUN, and spawned multiple verified EF-0 and EF-1 tornadoes, one can see that all of the NWS meteorologist teams, regardless of experimental condition, issued tornado warnings at or near the exact same times.

Timing of Issuance of Tornado Warnings: Example from Experimental Case 4, 13-14 April 2019, HUN WFO, Multiple EF-0 and EF-1 Verified

	Condition A (w/BVOT & SPC Timing Guidance)	Condition B (w/BVOT only)	Condition C (w/SPC Timing Only, NO BVOT DATA)
Week 1	11:58	11:57	11:58 12:20 (reissue)
Week 2	11:56 12:21 (reissue)	11:56	11:58
Week 3	11:53 12:10 (reissue)	11:54 12:22 (reissue)	11:58 12:01 (reissue)

Figure 9. Timing of all tornado warnings issued for Case 4 in the HWT experiment. All warnings were issued within 4-5 minutes of each other across the first three weeks of the experiment.

While the decision to issue/not to issue a formal warning on a storm was not impacted by increased vulnerability awareness (i.e., those NWS meteorologists under Condition A or Condition B, who had access to the BVOT), we *did* find that increased awareness of vulnerabilities could impact the *tags* that were associated with tornado warnings. We found that, in several of the severe weather cases, the added awareness of specific vulnerabilities *combined with* the radar-indicated severity of the storm (e.g., a clear and significant TDS) prompted meteorologists in Condition A and Condition (those with access to the BVOT) to issue Severe Weather Statements along with their tornado warnings that indicated the Particularly Dangerous Storm (PDS) tag. Event Case 8 was a particularly telling example (Fig. 10) because it was a sudden and quick-spin-up tornado that was in an area that had few AWIPS base-map locations but had several BVOT-indicated vulnerabilities. In this case, while all of the meteorologists all of the weeks (regardless of experimental Condition) issued tornado warnings, *only* those meteorologists with access to the BVOT added the PDS tag, recognizing the potential for significant risk to life and property due to their awareness of on-the-ground vulnerabilities.

Timing of Issuance of Tornado Warnings:
Example from Experimental Case 8, 25 January 2021 (Fultondale, AL), HUN BMX, EF-3 Verified

	Condition A (w/BVOT & SPC Timing Guidance)	Condition B (w/BVOT only)	Condition C (w/SPC Timing Only, NO BVOT DATA)
Week 1	10:32 Svr TS 10:40 TOR 10:46 SWS (PDS) 10:55 SWS (PDS)	10:33 Svr TS 10:36 TOR 10:40 SWS 10:46 SWS (PDS) 10:59 TOR (reissue)	10:39 TOR 10:47 SWS
Week 2	10:32 TOR 10:38 SWS (PDS) 10:46 SWS (PDS) 10:52 SWS (PDS)	10:37 Svr TS 10:39 TOR 10:44 SWS 10:51 SWS (PDS)	10:36 Svr TS 10:43 TOR 10:49 SWS
Week 3	10:33 TOR 10:38 TOR 10:50 SWS (PDS) 10:52 SWS	10:32 SWS (Svr TS) 10:37 TOR 10:42 SWS (BVOT-points) 10:47 SWS (PDS)	10:33 TOR 10:37 SWS 10:43 SWS 10:45 TOR (reissue) 10:47 SWS 10:51 SWS 10:52 SWS

Figure 9. Showing the timing of a number of products issued for experimental Case 8 during the first three weeks of the HWT study. While all teams of NWS meteorologists — regardless of experimental Conditions — issued tornado warnings, *only* the teams with access to the BVOT issued PDS Severe Weather Statements (SWS). In addition, in one case, the SWS contained a specific reference to a BVOT-only point, meant to convey additional impact-based messaging.

Project Contacts

Jack R. Friedman	Univ. of Oklahoma, CASR	Project PI
Daphne LaDue	Univ. of Oklahoma, CAPS	Project Co-PI
Michelle Saunders	Univ. of Oklahoma, CASR* **now with Mississippi State University*	Postdoctoral Research Scientist
Alex Marmo	Univ. of Oklahoma, CAPS	Research Associate
Justin Monroe	CIWRO and NOAA/NSSL	EWP Technical Lead

Tiny – Threats-In-Motion

Summary by Greg Stumpf

2022 Executive Summary

Threats-In-Motion (TIM) is a proposed warning decision and dissemination approach that would enable the NWS to upgrade severe thunderstorm and tornado warnings from the current static polygon system to continuously-updating warning polygons that move with the storm. TIM can improve average warning lead time as well as provide more equitable (uniform) lead times for those in the path of long-tracked severe storms. A short-term solution is being explored to take initial steps in adopting the TIM concept to existing convective warnings by allowing current polygon-based warnings to be extended in area and/or time by exploiting existing product extension VTEC action codes which are used for other products, but are presently not used for severe thunderstorm or tornado warnings. The software capability for this interim solution, known as "**Tiny TIM**", was developed within the AWIPS Hazard Services (HS) - Convective workflow. The Tiny TIM software and warning concepts were tested within the NOAA Hazardous Weather Testbed in the winter of 2022 with participating NWS forecasters.

The Tiny TIM software includes "warning extension recommenders" which allow forecasters to automatically extend the polygon area and time prior to issuance. Forecasters agreed that these recommenders need to be enhanced with additional features in order to improve workflow. The software also includes an optional capability to track hazards using a 2D footprint, which can more-precisely describe the 2D hazard than either the point or line tracking tools which are presently available. Finally, it was recommended that the software adds some enhancements to help handle dissemination issues that can remove locations from warnings and put them back in just several minutes later. These improvements, as well as some others, should be implemented and tested in the NWS Operations Proving Ground later this year. Additional half-day functional assessment tests with forecasters are planned to occur during the remainder of the year.

The biggest takeaways from the experiment are: a) Tiny TIM offers a reduction in forecaster workload and mental demand by eliminating the need to issue a new warning once a storm has exited a polygon, 2) Tiny TIM offers an improvement in messaging by keeping the same Event Tracking Number (ETN) for the lifecycle of the hazard and a cleaner output field by eliminating overlapping warnings, and 3) warning extensions should be allowed for any long-tracked storm, not just those with observed significant hazards.

2022 Detailed Overview

Current severe weather warnings (tornadoes, wind, and hail) require the forecaster to issue multiple sequential warnings for long-tracked storms because the current policy prohibits extending a warning's area and time during updates. This situation frequently results in nonuniform lead times for users on the downstream border of a warning polygon. For example, nearly adjacent locations can have dramatically different lead times if one location is just outside the upstream warning.

Threats-In-Motion (TIM) is a proposed warning decision and dissemination approach that would enable the NWS to upgrade severe thunderstorm and tornado warnings from the current static polygon system to continuously-updating warning polygons that move with the storm. TIM is meant to be an evolutionary step of Forecasting A Continuum of Environmental Threats (FACETs) for the convective weather warning scale. With TIM, the forecaster would only need to issue a single warning, updated regularly as workload permits, embodying a “one storm-one story” concept. This approach would reduce forecaster workload because downstream warning issuance would be replaced by a less time-consuming warning update.

This “one storm-one story” concept is the most significant benefit of TIM, as it also offers improvements in hazard communication. TIM provides a continuous history for each storm, versus times when two or more separate overlapping warnings may be in effect for the same hazard when a new downstream warning is issued before the previous warning can be canceled. These improvements could lead to more simplified and consistent messaging for key partners and improved event verification.

TIM can provide more equitable (uniform) lead times for those in the path of long-tracked severe storms because these storms remain continually tracked and warned. As such, TIM mitigates gaps in warning coverage and improves the handling of storm motion changes. This change also results in greater average lead times and decreased average time spent in a warning relative to today’s warnings, with little impact on average false alarm time. This impact is particularly noteworthy for storms expected to live longer than the average warning duration (30 – 45 minutes), such as the long-tracked supercells seen during violent severe weather outbreaks. A robust statistical analysis of TIM’s scientific benefits is available in Stumpf and Gerard (2021).

While the aforementioned considerations are being examined as a long-term solution, a short-term solution is being explored to take initial steps in adopting the TIM concept to existing convective warnings by allowing current polygon-based warnings to be extended in area and/or time. Warnings are issued with a Valid Time Event Code (VTEC), which enables weather providers and vendors to automate and tailor the product stream delivered to their clients.

Within the VTEC are two fields germane to this new functionality. First is the Event Tracking Number (ETN), which is a unique ID that is attached to each warning type (Severe Thunderstorm Warning or Tornado Warning). Presently, a warning cannot be extended in time or area for the same ETN. Each subsequent warning issued on a hazard has a different ETN assigned to it. Once the warning has expired or is canceled, the ETN is not reused.

Presently, the only way to update an existing warning is to issue a Severe Weather Statement, or warning continuation statement. The second field within the VTEC used for this new functionality is the Action Code. A warning continuation statement uses the action code of CON. With each CON update, the warning polygon is constrained to its original area – forecasters are only allowed to reduce the area during a warning update. And the expiration time of the warning cannot be extended.

In order to permit a single storm to maintain the same ETN throughout its lifecycle, we are proposing the allowance to extend the area and time during warning updates by exploiting these existing product-extension VTEC action codes. These codes are used for other products, but are presently not used for severe thunderstorm or tornado warnings:

- VTEC = EXA: An extension of the area (can enlarge area) of the warning polygon
- VTEC = EXT: An extension of the expiration time of the warning polygon
- VTEC = EXB: An extension of BOTH the area (can enlarge area) and the expiration time of the warning polygon

With TIM, these product-extension VTEC action codes are meant to be used when a hazard is expected to last beyond the duration of a typical severe thunderstorm or tornado warning (e.g., 30- or 45-minutes). For short-lived hazards, or for long-lived hazards nearing the end of their lifecycle, the traditional use of the warning continuance (VTEC = CON) is used instead until the warning has expired or is canceled early. Examples of a series of ETNs and action codes for today's warnings and for TIM warnings are shown in Table 1.

Time (UTC)	NWS VTEC	ETN	TIM VTEC	ETN
1900	NEW	1	NEW	1
1910	CON	1	EXB	1
1920	CON	1	EXB	1
1930	NEW	2	EXB	1
1940	CON	2	EXB	1
1950	CON	2	EXB	1
2000	NEW	3	EXB	1
2010	CON	3	EXB	1
2020	CON	3	EXB	1
2030	NEW	4	EXB	1
2040	CON	4	EXB	1
2050	CON	4	EXB	1
2100	NEW	5	EXB	1
*2110	CON	5	CON	1
*2120	CON	5	CON	1
*2130	EXP	5	EXP	1

Table 1. Warning decision times, VTEC action codes, and ETNs for a hypothetical storm case. Today's warnings use the blue columns on the left. TIM uses the green columns on the right.

A graphical representation of the current warning method and TIM is shown in Fig. 1. In this example, warning updates are issued every 15 minutes. For the current warning methodology, the warning polygon can only be made smaller with each update, and the warning expiration time remains constant. After 60 minutes, the warning is replaced with a new warning having a new polygon, ETN, and expiration time. For the TIM methodology, the warning can continue downstream at each update, with new areas added to the warning, some areas removed from the warning, and the expiration time extended. The TIM warning maintains the same ETN throughout its lifecycle.

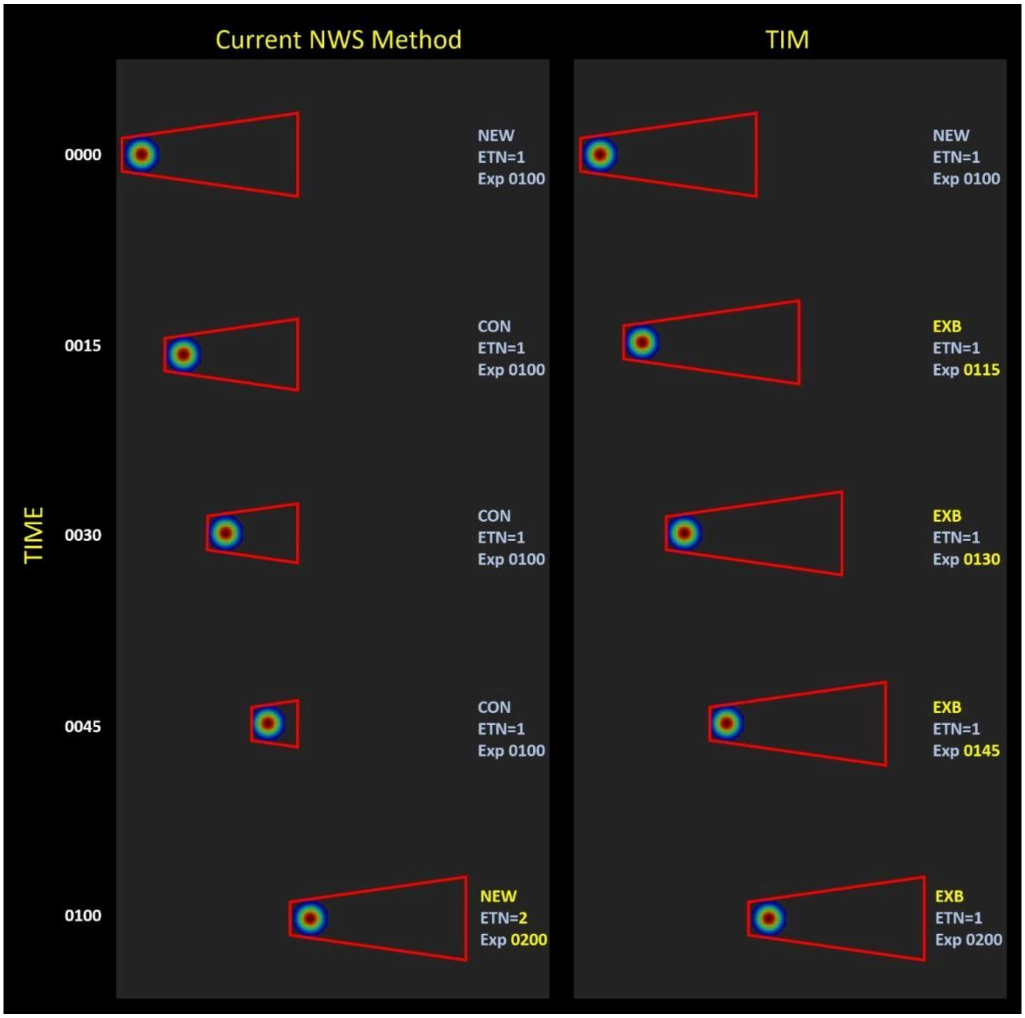


Fig. 1. *Comparison of (left) current NWS warning practice using separate polygons, (right) to TIM. Images are shown at 15-min intervals. For each time interval, the VTEC action code, ETN, and warning expiration time are shown. The colored circle represents a hypothetical hazard.*

This short-term interim solution described above is affectionately known as "**Tiny TIM**". The software capability was developed within the Advanced Weather Interactive Processing System (AWIPS) Hazard Services (HS). In particular, the Tiny TIM capabilities were built upon the HS-Convective workflow, which is designed to be the modern replacement for the AWIPS Warning Generation (WarnGen) software.

NWS/MDL, in collaboration with NSSL, GSL, and WDTD, carried out this NOAA Hazardous Weather Testbed (HWT) experiment in the winter of 2022. This experiment was the first to test the Tiny TIM software and concept. It follows three other experiments of the full TIM capability (one-minute automation between forecaster updates). This experiment allowed the collaborators to explore several ideas to represent realistic challenges currently faced in warning operations in order to focus on workload and workflow differences with the traditional method for issuing warnings today.

Because Tiny TIM is built upon the HS-Convective software that has yet to be fielded, the main differences between WarnGen and HS-Convective that forecasters had to learn and understand are summarized. Note that HS-Convective is not operationally ready, so some of the issues discovered during testing may still be under development:

- The Hazard Services screen layout (Spatial Display, Console, Hazard Information Dialog). This layout is already available in other HS perspectives (e.g., HS-Hydro), so some of the forecasters already had some experience.
- A change in the motion vector by dragging the tracking dots results in immediate change to polygon (versus requiring an additional Redraw Box button press).
- For traditional warning updates when area extensions are prohibited, the polygon will shrink to only cover the previously issued hatched area with each tracking point click or polygon vertex drag (versus requiring an additional Redraw Box button press).
- When updating a warning, the user needs only to select the warning from the console or spatial display (versus having to find the specific warning from a drop-down menu).
- The Hazard Information Dialog (HID) “Details” selections are easier to navigate, and automatically select appropriate Impact Based Warning (IBW) tags.

The major differences between HS-Convective and Tiny TIM capabilities specifically being tested are:

1. Warnings can be extended in area and/or time:
    a. Recommenders for automatic warning extension (via drop down menu):
        i. <u>Fixed</u>
            1. For hazards not expected to continue beyond warning duration (i.e., short-lived).
            2. Will not replace the current polygon with a “default” polygon derived from the storm track.
            3. Does not allow extensions in area (can only shrink the polygon).
            4. Keeps the same end time.
            5. Results in a CON (same as today’s Severe Weather Statement (SVS)).
        ii. <u>Update Back</u>
            1. For hazards not expected to continue beyond warning duration (i.e., short-lived).

2. Replaces the current polygon with a "default" polygon derived from the storm track and automatically moves *only the back* of the warning polygon forward (removing area).
3. Does not allow extensions in area (can only shrink the polygon).
4. Keeps the same end time.
5. Results in a CON (same as today's SVS).

iii. Update Both

1. For hazards expected to continue beyond warning duration (i.e., long-lived or long-tracked).
2. Replaces the current polygon with a "default" polygon derived from the storm track and automatically moves *the front and back* of the warning polygon forward.
3. Allows extensions in area.
4. Automatically extends end time based on previous warning duration.
5. Usually results in EXB (on rare occasions, can result in EXA or EXT).

iv. Update Front

1. For back-building hazards expected to continue beyond warning duration.
2. Replaces the current polygon with a "default" polygon derived from the storm track and automatically moves *only the front* of the warning polygon forward.
3. Allows extensions in area.
4. Automatically extends end time based on previous warning duration.
5. Usually results in EXB (on rare occasions, can result in EXA or EXT).

b. The warning area and/or time can also be extended manually:

i. Extending warning end time (via a drop-down menu).
ii. Extending warning area (by editing polygon vertices or dragging the entire polygon to a new location).

2. Along with point and line tracking tools (Fig. 2: left), a new 2D Footprint tracking tool has been added (Fig.2: right).

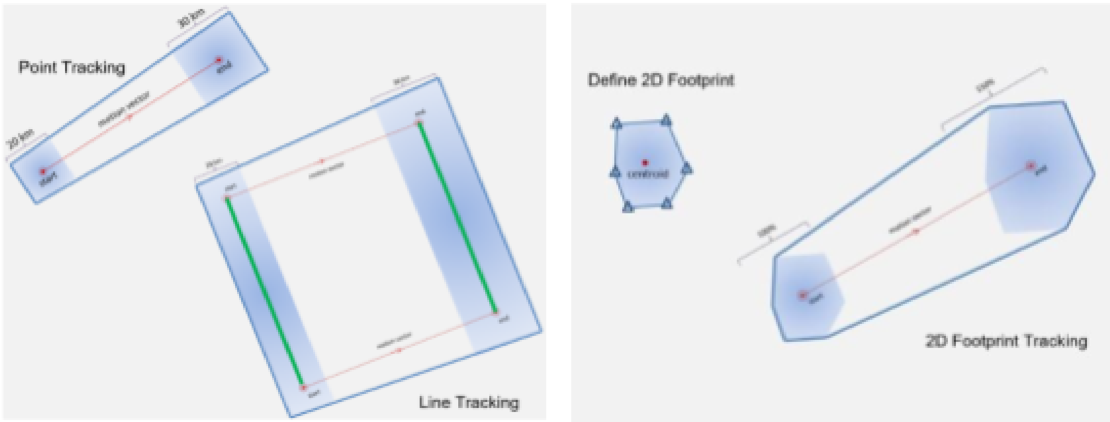


Fig. 2. *Left: The default warning polygon that is produced by AWIPS WarnGen for point and line tracking; right: the warning polygon resulting in the tracking of a 2D footprint.*

For this experiment, only forecasters participated (no end users), so the resulting feedback is from the operational NWS forecaster perspective. The first week of the Tiny TIM HWT experiment was a "shakedown" of the system using two "test" forecasters. The remaining operational weeks included a pair of participants from several WFOs nationwide. Each week saw forecasters learning how to use the software via a guided, hands-on training exercise on the first day, and displaced real-time (DRT) severe weather scenarios on Days 2 through 4. The final day was spent conducting a 2-hour guided interview of the forecasters on their experience during the experiment week.

Because of the continuing COVID-19 pandemic restrictions, the experiment was conducted virtually, using a version of the AWIPS software hosted by the NOAA Virtual Laboratory (VLab) within the Amazon Cloud Services (see Fig. 3). There were some pros and cons to this approach:

- Pros: Developers had quick and convenient browser access to the cloud systems from anywhere, so that quick software tests could be performed; forecasters who typically cannot travel to Norman for a variety of reasons can now participate in an HWT experiment; developers could quickly diagnose problems, without having to ask forecasters to leave their workstations; each participant had a close up, high-resolution view of the workstations (rather than looking over shoulders or on mounted TV screens).
- Cons: Workdays had to be shortened (due to "Zoom fatigue"), resulting in less time for forecasters to participate in hands-on scenarios and less feedback for developers; there was no in-person social interaction outside working hours; we could not split instances to two monitors.

Archived data events were used to perform the evaluations; including the training case, there were four cases from different locations nationwide. Each case had unique domains and represented a variety of different severe storm types (e.g., squall lines, long-tracked tornadoes, etc.).

The aforementioned "shakedown" week was used to identify any issues needing attention before the operational weeks of the experiment began. These are the software changes made before the operational experiment weeks commenced:

- Warning extensions are proposed to be used only if the IBW damage threat tags are "Considerable", "Destructive", or "Catastrophic". For this experiment, we disabled the IBW tag requirement for extending warnings so that the feature can be applied to any long-tracked hazard.
- Currently, warnings go into the "Ending" status with 5 minutes until expiration, and there is no way to extend the warning after this (can only expire the warning). This threshold was changed to 1 minute, to allow warnings to be extended until the final minute.
- The default warning time alert thresholds were changed
  - Yellow alert: old (halfway to expiration), new (15 minutes from expiration)
  - Red alert: old (10 minutes from expiration), new (5 minutes from expiration)
- A new warning grid display was developed, as the current method for displaying warnings in AWIPS (under the "Obs" menu) could not display certain warnings with the EXB, EXA, or EXT VTEC action codes. Severe thunderstorm (tornado) warnings were displayed as yellow (red), with a 30-second update rate. Data sampling capabilities were extended to include the VTEC action code and ETN in the sampling output.

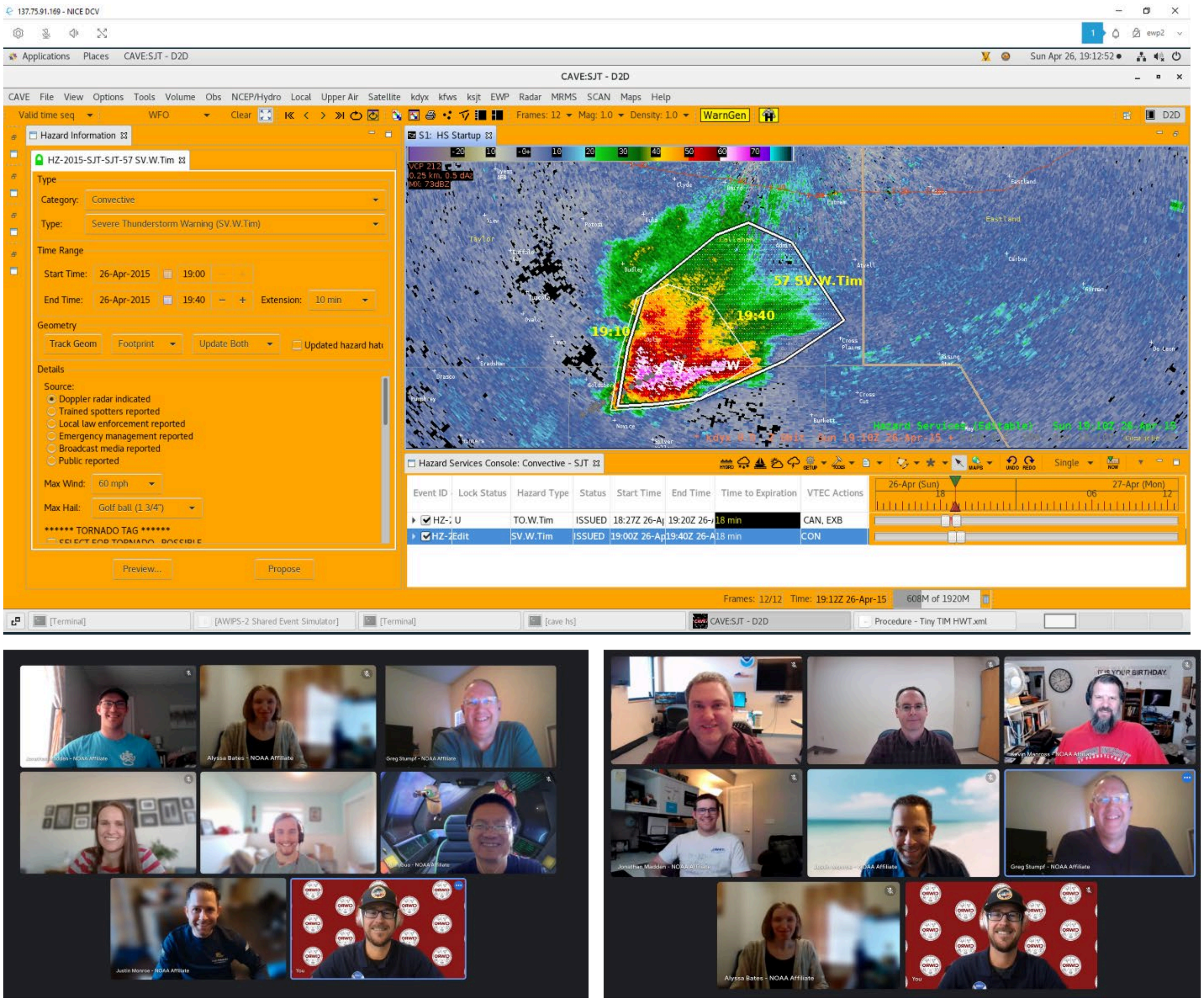


Fig. 3. Images from the Tiny TIM experiment. Top left: Tiny TIM output from a scenario in southeast Alabama - the yellow (red) polygons show severe thunderstorm (tornado) warnings; the large orange polygon is the county warning area (CWA). The remaining images are group photos of the virtual meetings during each week of the experiment.

A major issue with Hazard Services as a whole (not confined to the Tiny TIM additions) resulted in the software becoming very sluggish after a while as more hazard events were created and updated. This necessitated some adjustment to the experiment to overcome the issue. We had the forecasters limit their warning issuances to only 2 warnings at a time. However, even with this limit, subsequent warning updates, or new warnings created after previous warnings expired, added individual entries to the HS “registry” (database) per event issuances. As these events accumulated in the registry, the software became more sluggish. Another way we overcame this was to have the forecasters reload the HS layer after each warning issuance. This provided about 5-10 minutes before the software started becoming sluggish again.

Meanwhile, HS developers at GSL (external to the Tiny TIM project) determined that the sluggishness was the result of a possible memory leak in the third-party graphics code used to draw the HS Console timeline. Just before the final scenario on the second week of the experiment, we disabled the HS Console timeline, and the entire scenario was executed without any noticeable sluggishness.

There were also several times where the VLab cloud instances were all terminated unexpectedly as a result of security patching by the VLab administrative team, who were unaware that we were conducting the HWT experiment at the time. This required a complete restart of the exercises and caused delays in our experiment schedule. Thankfully, the experiment participants were flexible enough to allow adjustments to our experiment schedule to overcome the delay. The HWT administrative team was also added to the email list alerting us of any future security patches to the virtual systems so that we could coordinate alternate times for patching outside the experiment hours.

During the two operations weeks, there were only very few software bugs discovered. In addition, some minor enhancements to functionality were also recommended. All of these minor issues were addressed by simple local software changes and were in place before the next scenario began. Issues recommended or discovered during the operational weeks were:

- Made changes to the default HS Console columns:
  - Replaced Event ID (a lengthy string that took up valuable space on the console) with ETN (the optimal way to track the warning event history).
  - Replaced warning Start Time (which remains unchanged until the warning is canceled or expired) with warning Issue Time (this time changes with each warning update, and can be used to sort columns to determine which warnings have gone the longest without an update).
- Fixed an issue where sometimes the unedited warnings in the Spatial Display were white instead of the color they were meant to be (SVR=yellow; TOR=red). Warnings should only be displayed as white when they are being edited.

The biggest outcome from the experiment was related to the workflows that the forecasters employed to create and update warnings. There were several reasons for this:

- It was never entirely clear when the forecaster should hit the “Track Geom” button while composing or updating a warning. Forecasters suggested “emphasizing” the button when it should be clicked (e.g., a dark border, which would not be present if the current polygon is already the default).

- Hitting Track Geom on a polygon that had been manually edited or county-clicked removed all of the manual changes and replaced the polygon with a default polygon tied to the storm track. Forecasters would like the option to track manually-edited polygons as well.
    - County-clicking should be treated differently when advecting a polygon forward, as the geo-political outline should not be retained when they are no longer coincident with the actual boundary.
- The differences between WarnGen and HS-Convective caught the forecasters by surprise.
- The Fixed and Update Back options did not work as intended (and they worked differently from HS-Convective). Forecasters could manually edit the polygon or manually extend the end time of the warning and break the rule that these two options disallow extensions in area and time. While this should be changed to match HSConvective behavior, forecasters also wanted to be able to override this restriction via a toggle button.
- Forecasters did not like the feature in which the default polygons would auto-update (i. e., move forward) with each radar frame update. They would prefer to make this optional.
- The buttons in the Geometry section of the HID placed the Track Geom button at the top left, even though it is usually the last button clicked in the workflow.

Several other suggestions for improved software functionality were offered by the forecasters. Some of these suggestions will be incorporated into a future version (as funding and budgets permit) in order to make the software more robust. These suggestions are listed in 2022 Appendix A.

In addition, the forecasters made some suggestions to improve future experiment logistics. These are listed in 2022 Appendix B.

This experiment also included a simple human factors analysis in the form of online surveys designed to measure mental workload, confidence, and software usability. More information about the human factors analysis is included in 2022 Appendix C.

2022 Key Takeaways

- Tiny TIM greatly reduces forecaster workload and mental demand when it is needed (for long-tracked hazards)
    - Warning updates can be made much faster and are more spatially accurate.
    - Because warnings can be extended, forecasters do not have to make a decision on whether to issue a separate new downstream warning.

    - Because warnings are constantly being updated, storms will rarely exit the polygons.
    - Adjustments can be made if the previous warning polygon was not perfect.
- Not having overlapping warnings on a single hazard cleans up the output field and should improve messaging. There is no need to issue a brand new warning and then manage two separate warnings for the same hazard if a hazard looks to be exiting the polygon. Tiny TIM fixes these situations:
    - When a new downstream warning is issued, it sometimes overlaps the previous warning. When a previous warning is subsequently canceled for the same locations of the new warning, this can send a confusing message.
    - When there are overlapping warnings, two warnings have to be playing on NOAA Weather Radio and television crawls, which can be confusing.
- The new 2D footprint tool for drawing warnings received favorable feedback.
    - Some forecasters prefer the resultant warning shape that results from using the new "Footprint" option, as the downstream edges can be more rounded.
    - The footprint tool is a better alternative to the line tracking tool, as there is no need to drag the original polygon back to cover the storm hazard when tracking for the leading gust front. It can also be used to better capture severe rear-inflow winds.
- Every forecaster agreed that warning extension should be allowed for any long-tracked storm, not just those with observed significant hazards.
    - It may take some time for a tornado to occur with a long-tracked storm, and waiting to extend the warning until a report is received will not add lead time to the start of the tornado. For example, in the second scenario (Lee County, Alabama, 3/3/2019), the first report of a tornado was not until one minute after the town of Beauregard was hit and 19 people died.

Testimonials

"I absolutely love where this is headed. Cleaner graphics, better/clearer messaging, increased/equitable lead time and reduced false alarm time are excellent reasons to get this software operational asap. I definitely want to see it in use with other types of events, but what is here is a great start!"

"I think this will be a great addition to NWS warning operations to help improve lead times, decrease warning messaging confusion, and aid workload management for warning forecasters."

"The ability to extend warnings in both area and time is the shining star of Tiny TIM, and they would be massively valuable additions to a forecaster's toolbox. They aren't always needed, and in fact may not be needed in the majority of warning scenarios, but when they are needed, they are awesome."

2022 Future Plans

- Continue software refinement and development of new functionality prior to future experimentation with forecasters, including:
  - Hardening code (debugging, refactoring).
  - Regular integration of the Tiny TIM code into the operational version of Hazard Services, to remain up-to-date on HS-Convective development.
  - Incorporate any specific functionalities/modifications to address issues with dissemination (there are two suggestions pertinent to this in 2022 Appendix A).
  - Adding storm longevity guidance so that forecasters can have an idea on whether or not to extend warnings.
  - Incorporate refinements suggested by forecasters, including the issues detailed above regarding improving workflow, as well as those listed in the Appendices, as developer funding and time permits.
- Host half-day functional assessment tests with former Tiny TIM experiment participants, to get feedback on new functionalities developed since the HWT experiment.
- Execute an experiment at the NWS Operations Proving Ground in September 2022.
  - Takes advantage of a fully operational AWIPS system, allowing localization to any WFO and access to real-time data.

2023 Executive Summary

Threats-In-Motion (TIM) is a proposed approach for warning decisions and dissemination that aims to upgrade the current static polygon system used by the National Weather Service (NWS) for severe thunderstorm and tornado warnings. The main objective of TIM is to create warning polygons that dynamically update and move along with the storm, allowing for improved average warning lead time and more equitable (uniform) lead times for individuals in the path of long-tracked severe storms. Currently, severe thunderstorm or tornado warnings cannot be extended in area or time without the issuance of a separate new warning. To address this issue, a short-term solution called "**Tiny TIM**" has been developed, which incorporates the TIM concept by allowing current polygon-based warnings to be extended in area and/or time. Tiny TIM has been developed as an "add-on" feature to the Advanced Weather Interactive Processing System (AWIPS) Hazard Services (HS) - Convective workflow. It was tested in person at the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) for two weeks in winter 2023 with participating NWS forecasters. This was the third Tiny TIM experiment, following a virtual HWT experiment and an in-person NWS Operations Proving Ground (OPG) experiment, both in 2022.

The Tiny TIM software adds functionality that allows forecasters to manually extend the warning polygon's area and time during warning updates when a hazard is expected to be long tracked. When updating a warning, forecasters can choose to use "warning extension shortcuts" that automatically extend the polygon area and time. Additionally, the software offers an optional feature to track hazards using a 2D footprint tool, which provides a more precise description of the hazard area than the currently available point or line tracking tools. The new "dynamic" line tool offers the capability to utilize individual motion vectors for each vertex, which proves to be valuable when dealing with bowing line segments. The software explicitly prevents the "accordion effect" where downstream locations are removed from the warning if the storm's motion vector speed decreases, temporarily making the warning shorter until the locations are placed back within the warning during the next update as the storm continues to move forward. There are options to track manually-edited polygons as well as default polygons, and an option to auto-update the polygon immediately after changes are made to the track or geometry or when the radar frame updates. Forecasters can also issue warnings using the traditional process and choose not to extend the expiration time, only maintaining or shrinking the warning area.

The main findings from the 2023 HWT experiment are twofold: 1) Tiny TIM reduces the workload and mental demand on forecasters by eliminating the need to issue a new warning once a storm has exited a polygon, and 2) Tiny TIM improves messaging by maintaining the same Event Tracking Number (ETN) throughout the hazard's lifecycle and generating cleaner output fields by eliminating overlapping warnings.

2023 Detailed Overview

Current severe weather warnings (tornadoes, wind, and hail) require the forecaster to issue multiple sequential warnings for long-tracked storms because the current policy prohibits extending a warning's area and time during updates. This situation frequently results in nonuniform lead times for users on the downstream border of a warning polygon. For instance, nearly adjacent locations can have dramatically different lead times if one location is just outside the upstream warning.

Threats-In-Motion (TIM) is a proposed warning decision and dissemination approach that would enable the National Weather Service (NWS) to upgrade severe thunderstorm and tornado warnings from the current static polygon system to continuously-updating warning polygons that move with the storm. Specifically, a warning polygon is attached to the threat and advances along with it. Warnings are automatically cleared from locations where the threat has passed. TIM represents an evolutionary step of the Forecasting A Continuum of Environmental Threats (FACETs) initiative for the convective weather warning scale. With TIM, forecasters would only need to issue a single warning, updated regularly as workload permits, aligning with a "one storm-one story" principle. This approach has the potential to alleviate forecaster workload, as the time-consuming process of issuing downstream warnings would be replaced by a less time consuming warning update.

The "one storm-one story" concept is the primary advantage of TIM, as it also improves hazard communication. TIM ensures a continuous history for each storm, avoiding situations where multiple separate overlapping warnings may be active for the same hazard due to the issuance of a new downstream warning before the cancellation of the previous one. These enhancements can result in simplified and consistent messaging for key partners and improved event verification.

TIM can offer more equitable (uniform) lead times for individuals in the path of long-tracked severe storms because these storms are continuously tracked and warned. TIM can mitigate gaps in warning coverage and can improve the handling of changes in storm motion. Additionally, this change can lead to greater average lead times and reduced average time spent in a warning compared to the current warning system, with minimal impact on average false alarm time. This impact is particularly significant for storms that are expected to persist longer than the average warning duration (30-45 minutes), such as long-tracked supercells observed during major severe weather outbreaks. A comprehensive statistical analysis of the scientific benefits of TIM is available in Stumpf and Gerard (2021).

This promising and innovative approach is currently being considered for implementation in NWS operations. While the "full implementation" aspects are being evaluated as a long-term solution, a short-term approach is being explored to begin incorporating the TIM concept into existing convective warnings. This short-term interim solution, referred to as "**Tiny TIM**," enables the extension of current polygon-based warnings by area and/or expiration time. The software capability is integrated within a development version of the Advanced Weather Interactive Processing System (AWIPS) Hazard Services (HS). Specifically, the Tiny TIM capabilities are built upon the HS-Convective workflow, which serves as a modern replacement for the AWIPS Warning Generation (WarnGen) software.

NWS convective warnings are accompanied by a Valid Time Event Code (VTEC), which enables weather providers and vendors to automate and customize the product stream delivered to their clients. Within the VTEC, there are two relevant fields for this new functionality. The first is the Event Tracking Number (ETN), which serves as a unique identifier attached to each warning type (Severe Thunderstorm Warning or Tornado Warning). Currently, it is not possible to extend a warning in terms of time or area while maintaining the same ETN. Each subsequent warning issued for a specific hazard is assigned a different ETN. Once a warning has expired or been canceled, the ETN is not reused (NWS, 2020).

Presently, the only way to update an existing warning is by issuing a Severe Weather Statement – a warning continuation statement. The second field within the VTEC that is relevant to this new functionality is the Action Code. A warning continuation statement uses the action code "CON." With each "CON" update, the warning polygon is constrained to its original area, and forecasters are only allowed to reduce the area, without the ability to extend the expiration time of the warning.

To permit a single storm to retain the same ETN throughout its lifecycle, we propose allowing the extension of the area and time during warning updates by utilizing these existing product extension VTEC action codes. These codes are currently used for other products but are not presently utilized for severe thunderstorm or tornado warnings:

- VTEC = EXA: Extension of the warning polygon's area (can enlarge the area).
- VTEC = EXT: Extension of the expiration time of the warning polygon.
- VTEC = EXB: Extension of BOTH the warning polygon's area (can enlarge the area) and expiration time.

With TIM, these product-extension VTEC action codes are intended to be used when a hazard is expected to persist beyond the typical duration of a severe thunderstorm or tornado warning (e.g., 30 or 45 minutes). For short-lived hazards or long-lived hazards nearing the end of their lifecycle, the traditional warning continuance (VTEC = CON) is used until the warning expires or is canceled early. Examples of a series of ETNs and action codes for current warnings and TIM warnings are illustrated in Table 2.

Figure 4 depicts a graphical representation of the current warning method and TIM. In this example, warning updates are issued every 15 minutes. Under the current warning methodology, the warning polygon can only be reduced in size with each update, and the expiration time remains constant. After 60 minutes, a new warning is issued, featuring a new polygon, ETN, and expiration time. With the TIM methodology, the warning can continue downstream with each update, allowing for the addition and removal of areas from the warning, as well as an extension of the expiration time. The TIM warning maintains the same ETN throughout its lifecycle.

Time (UTC)	NWS VTEC	ETN	TIM VTEC	ETN
1900	NEW	1	NEW	1
1910	CON	1	EXB	1
1920	CON	1	EXB	1
1930	NEW	2	EXB	1
1940	CON	2	EXB	1
1950	CON	2	EXB	1
2000	NEW	3	EXB	1
2010	CON	3	EXB	1
2020	CON	3	EXB	1
2030	NEW	4	EXB	1
2040	CON	4	EXB	1
2050	CON	4	EXB	1
2100	NEW	5	EXB	1
*2110	CON	5	CON	1
*2120	CON	5	CON	1
*2130	EXP	5	EXP	1

Table 2. *Warning decision times, VTEC action codes, and ETNs for a hypothetical storm case. Today's warnings use the blue columns on the left. TIM uses the green columns on the right.*

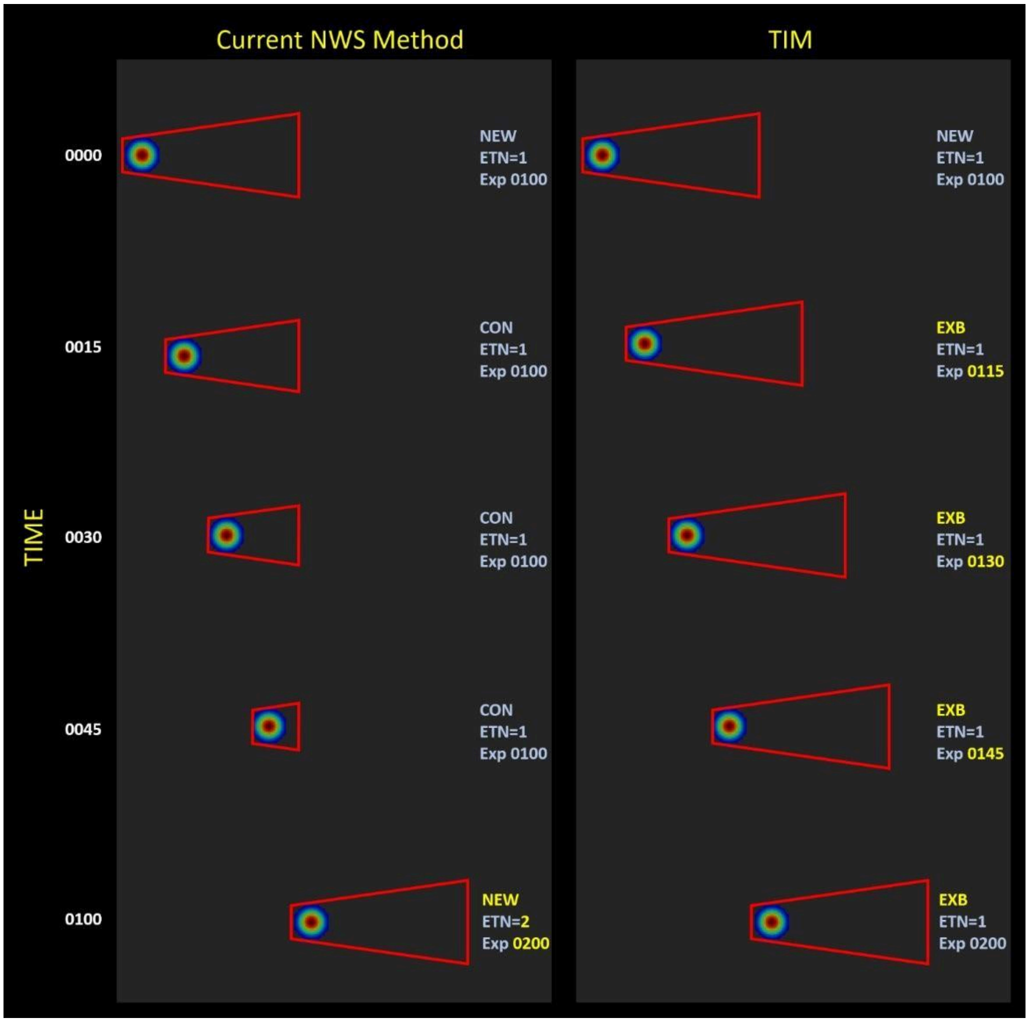


Fig. 4. *Comparison of (left) current NWS warning practice using separate polygons, (right) to TIM. Images are shown at 15-min intervals. For each time interval, the VTEC action code, ETN, and warning expiration time are shown. The colored circle represents a hypothetical hazard.*

Software Details

Because Tiny TIM is based on the yet-to-be-fielded HS-Convective software, this summary list highlights the main differences between WarnGen and HS-Convective that forecasters had to learn and understand. It should be noted that HS-Convective is not currently operational, and therefore, some of the issues discovered during testing may still be in the process of being resolved.

- The HS screen layout (Fig. 5; Spatial Display, Console, Hazard Information Dialog) differs from WarnGen. This layout is already available in other HS perspectives (e.g., HS-Hydro), so some forecasters already have some experience with it.
- When updating a warning, the user only needs to left-click the warning from the Console or spatial display (rather than having to find the specific warning from a drop-down menu with WarnGen).
- Dragging the warning polygon to a new position using the left mouse also requires pressing the Alt key.
- The HS Hazard Information Dialog (HID) "Details" selections are easier to navigate and, unlike WarnGen, automatically select appropriate Impact Based Warning (IBW) tags.

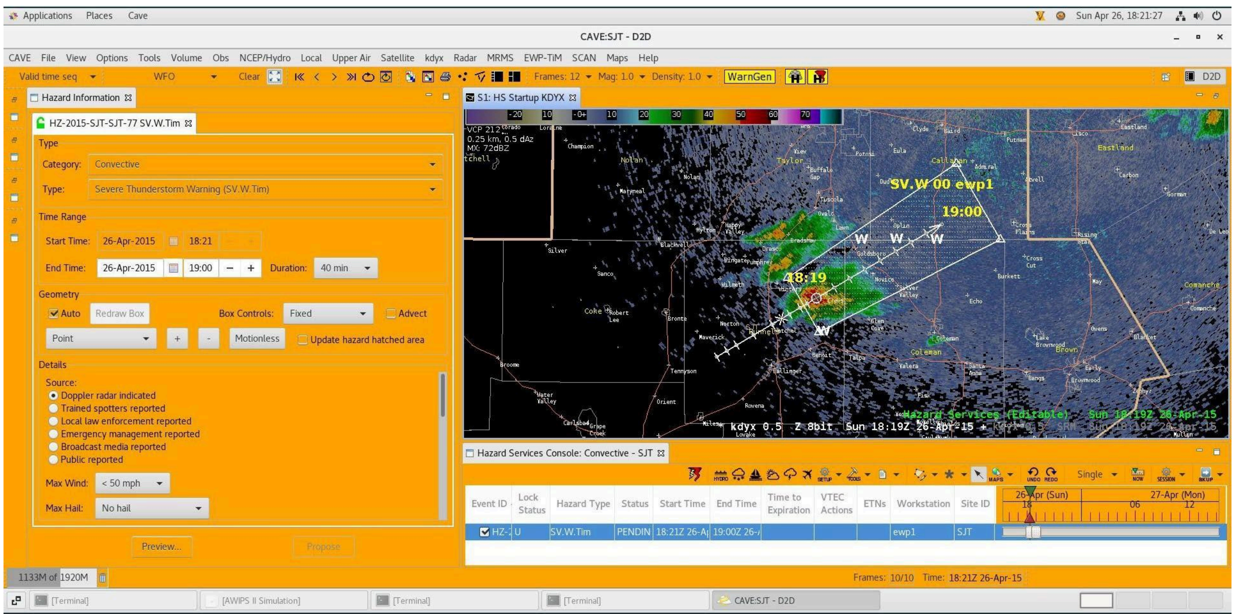


Fig. 5. *Tiny TIM output in AWIPS from the training scenario in Central Texas.*

The following changes were also made to the baseline HS-Convective to aid in the use of the new Tiny TIM functionality:

- Currently, warnings transition to the "Ending" status with 5 minutes until expiration, and there is no way to extend the warning after this point (only expiration is possible). This threshold was changed to 1 minute, allowing warnings to be extended until the final minute.
- The default alert thresholds for time to expiration were modified as follows:
  - Yellow alert: old (halfway to expiration), new (15 minutes from expiration)
  - Red alert (includes a pop-up alert): old (10 minutes from expiration), new (5 minutes from expiration)

- New Tiny TIM default console columns:
  - Event ID, ETNs, Lock Status, Hazard Type, Status, Start Time, End Time, Time To Expiration, VTEC Actions, Workstation, Site ID [County Warning Area (CWA)]
- The local warning overlay in AWIPS (under the "Obs" menu) does not display warnings with partial county cancellations (CAN) combined with warnings featuring the new EXB, EXA, or EXT VTEC action codes. Consequently, an experimental warning grid overlay was developed to display warnings with these new VTEC combinations. This experimental warning overlay also offers improved data sampling capabilities, including the VTEC action code, ETN, and IBW tags, which are not available in the local warning overlay. High-end SVR and TOR warnings with Destructive and Catastrophic IBW intensity tags are depicted with thick lines.

The major new capabilities added to HS-Convective to specifically support Tiny TIM functionality are as follows:

1. Along with point and line tracking tools (Fig. 6: left), a new 2D Footprint tracking tool has been added (Fig. 6: right).

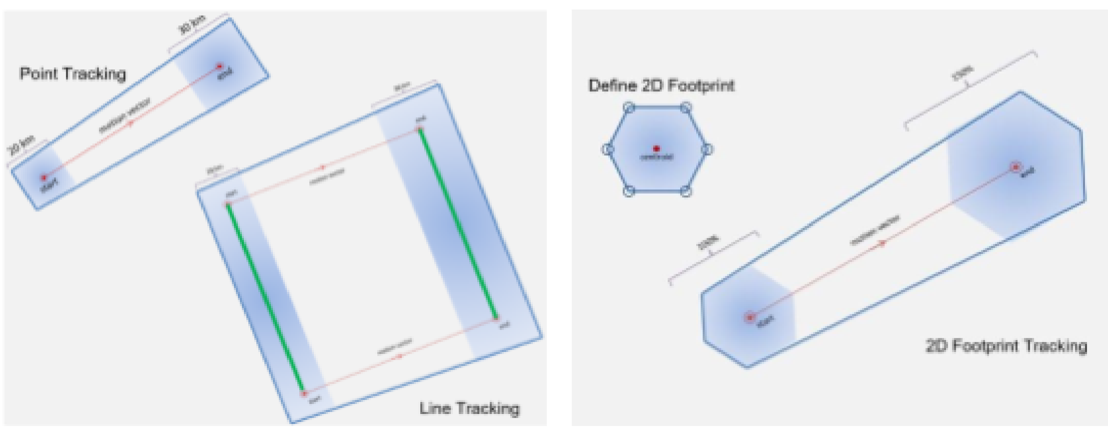


Fig. 6. *Left - The default warning polygon produced by AWIPS WarnGen for point and line tracking; Right - The warning polygon resulting from tracking a 2D footprint.*

2. Warnings can be extended in area and/or time automatically or manually:
   a. Automatic Extension: Utilizes shortcuts known as "**Box Controls**" available through a drop-down menu:
      i. <u>Fixed</u>
         1. For hazards not expected to persist beyond the warning duration (i.e., short-lived)

2. The polygon's position remains fixed without any movement

ii. Update Rear (the default for "**CON**" – see below)

1. For hazards not expected to persist beyond the warning duration (i.e., short-lived)
2. Automatically moves **only the Rear** of the warning polygon forward (subtracting area) based on point/line/footprint and the motion vector
3. Keeps the same expiration time, with the duration aging off

iii. Update Both

1. For hazards expected to persist beyond the warning duration (i.e., long-lived or long-tracked)
2. Automatically moves **both the Front and Back** of the warning polygon forward based on point/line/footprint and the motion vector
3. Automatically extends expiration time based on previous warning duration, with the duration remaining constant
4. Results in **EXB**
5. This is the default for the first warning update, after which the previous setting is used for subsequent updates.

iv. Update Front

1. For back-building hazards expected to continue beyond the warning duration
2. Automatically moves **only the Front** of the warning polygon forward based on point/line/footprint and the motion vector
3. Automatically extends expiration time based on previous warning duration, with the duration remaining constant
4. Results in **EXB**

b. Manual extension:

i. Extending warning expiration time by increasing the duration via a dropdown menu.

ii. Extending warning area by editing polygon vertices or dragging the entire polygon to a new location that includes locations outside the previous warning area.

Additional improvements have been made to the software following Tiny TIM experiments at the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) and NWS Operations Proving Ground (OPG), both in 2022. Many of these improvements were suggested by forecasters in these experiments and can be found in those experiment summaries (Stumpf 2022a, 2022b). These include:

- Redesigned the functionality to track manually-edited polygons versus default polygons, as well as enabling auto-updates when manual changes are made to the motion vector, duration, or geometry, or when the radar frame updates. The Compute-AdvectComputing-Advecting choices (Stumpf 2022b) have been replaced with the following options:
  - Auto:
    - ON: Any change to the motion vector, duration, or point/line/footprint tool geometry, or when the radar frame updates, the warning polygon will immediately redraw for the new conditions.
    - OFF: The warning polygon will NOT automatically redraw. Pressing the "Redraw Box" button is required to redraw a polygon for any of the above new conditions.
  - Advect:
    - OFF: The warning polygon is drawn using the “default” shapes corresponding to the point, line, or 2D footprint tool (see Fig 3).
    - ON: This setting will automatically be enabled if the warning polygon is edited in any way (e.g., moving, adding, or deleting a vertex, or dragging the entire polygon). The edited polygon will be "advected" when there is a change to the motion vector or position of the tool.
  - Notes:
    - NEW warnings will have Auto turned on and Advect turned off by default.
    - The chosen settings will be saved for the next warning update.
- The End Time Extension feature for warning updates has been replaced with the same Duration selection used for new warnings. Forecasters were having difficulty making mental calculations of the new duration. Unless the warning is moving forward with the hazard, the duration will now "age off".
  - We have also added the capability to decrease the expiration time on a warning update if the forecaster believes the storm will not last as long as previously thought, but they are not ready to cancel the warning yet.
- Added the option to restrict warning updates to the current WarnGen policy (cannot extend the expiration time and can only shrink the warning area) using a "CON" toggle (where the VTEC action code of CON is used in the warning). If the forecaster attempts to redraw the polygon outside the previous area, it will automatically snap back to only include previously-warned areas. The forecaster also cannot extend the End Time when this condition is met.
- Developed a better method for selecting vertices when a footprint vertex overlaps a polygon vertex. The polygon was slightly expanded (1 km) beyond the edge of the footprint to prevent point overlap.

- The Line and Footprint tracking tools currently use a single motion vector for the entire tracking object. For the 2022 OPG experiment, new versions were introduced with individual motion vectors assigned to each vertex on the tracking tool. These updated versions are known as the Line* and Footprint* tracking tools. o Based on feedback from forecasters, it was suggested to rename Line* to Dynamic Line and Footprint* to Dynamic Footprint.
- Implemented a "buffer" functionality for polygons, allowing users to conveniently increase or decrease the size of the warning by 1km increments using the + and – buttons.
- During the OPG, it was discovered that when a warning was selected for viewing, if the radar frame updated before the forecaster had a chance to edit the warning, the warning would automatically enter Edit mode and become locked from being edited by another forecaster. To address this issue, we introduced a 2-minute "grace period" during which the warning would not enter Edit mode, allowing the forecaster to deselect it before the grace period expired. However, it was found that the 2-minute grace period occurred too frequently. As a result, we decided to eliminate the grace period entirely and instead require forecasters to make a change to the warning, either by editing a value on the HID or adjusting the geometry on the display, before the warning enters Edit mode.
    - When a warning is updated and the Auto mode was previously enabled, the autoupdate feature remains temporarily inactive until the "Redraw Box" button is pressed.
- Incorporated specific functionalities/modifications to address dissemination issues:
    - When extending a warning, a dashed “former” polygon should be shown in the spatial display when a warning is being updated so that forecasters can see what they are changing their polygon from. This helps alleviate the potential slight addition to the workload when the user has to decide how much to extend the warning, to avoid the windshield wiper and accordion effects.
    - Explicitly prevent the "accordion effect" where downstream locations are removed from the warning if the storm's motion vector speed decreases, temporarily making the warning shorter until the locations are placed back within the warning during the next update as the storm continues to move forward (Fig. 7).

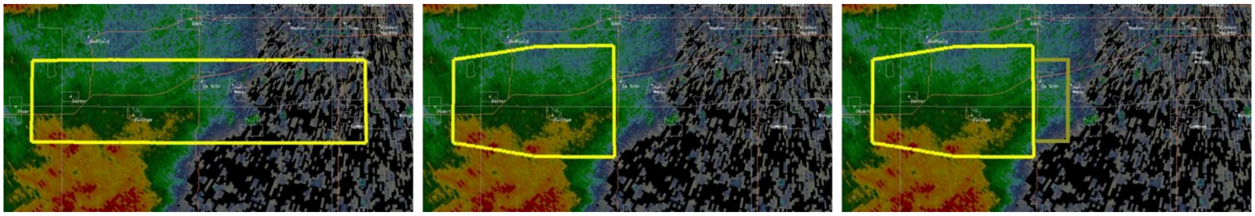


Fig. 7. *Left: A New warning; Center: An updated warning with a shorter polygon; Right: the darker yellow indicates the portion of the New warning (left) that has been added to the updated warning (center), mitigating the "accordion effect".*

2023 Experiment Details

The NWS Meteorological Development Laboratory (MDL), in collaboration with the National Severe Storms Laboratory (NSSL), the Global Systems Laboratory (GSL), and the NWS Warning Decision Training Division (WDTD), conducted this experiment in person at the HWT for one shakedown (dry run) week and two operations weeks in 2023:

- Shakedown Week: Jan 30 – Feb 3
- Week 1: Feb 13 – Feb 17
- Week 2: Feb 27 – Mar 3

This was the first in-person experiment conducted at the HWT in about three years. This HWT experiment allowed the collaborators to explore several ideas to represent realistic challenges currently faced in warning operations with a broad range of forecasters. The primary focus was on workload and workflow differences compared to the traditional method of issuing warnings today. The aim was to assess functionality, identify any impacts on forecaster workload, identify any inconsistencies across CWAs/products/services, examine dissemination challenges, examine ideal timing updates, and suggest improvements and new functionality.

Each week, the experiment included three participants: two warning forecasters and one warning coordinator/mesoanalyst (WC/MesoA) (Figs. 8 and 9). For each scenario, a different participant would rotate into the WC/MesoA position. The two forecasters worked on a series of displaced real-time events with varying storm type/evolution scenarios. The WC/MesoA aided by providing information about the near-storm environment and its impact on storm evolution, as well as keeping track of all the warnings and storm reports and monitoring the forecasters' workload. The WC/MesoA would step in as a third warning forecaster when needed if the workload was too high for the other two.

For this experiment, only forecasters participated (no end users), so the resulting feedback is from the operational NWS forecaster perspective. The first week of the Tiny TIM HWT experiment was a "shakedown" of the system, involving three "test" forecasters from the Norman, OK, NWS Weather Forecast Office (WFO). The shakedown week was used to identify and fix any issues (software and logistical) that needed attention before the operational weeks of the experiment began.

On the first day (Monday), after an introductory presentation and software demo, forecasters learned how to use the software through a job sheet-guided, hands-on training exercise. Since the last experiment, we extended our training session by an additional hour and provided forecasters with more time to practice on their own.

On Days 2, 3, and 4 (Tuesday, Wednesday, and Thursday), forecasters issued Tiny TIM warnings for five different displaced real-time (DRT) severe weather scenarios (two scenarios per day). Each DRT case had unique domains and represented a variety of severe storm types, such as long-tracked tornadoes, squall lines with quasi-linear convective system (QLCS) tornadoes, right-turning supercells, and left-hooking occluding tornadoes (Table 3). Two new archive cases were added, in addition to those used in 2022, including an event spanning 5 ½ hours without breaks (forecasters ate lunch during operations). Storm reports, photographs, video clips, and Storm Prediction Center products were injected into a Slack channel using an automated "bot" and made available to the forecasters in real-time. For some scenarios, original television broadcasts were played, synchronized in time with the event.

	Type	Storm Modes	Length
Scenario 1	Training	Isolated supercells	2 hours
Scenario 3	Operations	Long-tracked supercells	2 hours 15 minutes
Scenario 7	Operations	Squall line with QCLS tornadoes	1 hour 45 minutes
Scenario 4	Operations	Multicell clusters and right-turning supercells	2 hours
Scenario 5	Operations	Long-tracked supercells	2 hours 15 minutes
Scenario 6	Operations	Multiple supercell areas; left-hooking tornadoes	5 hours 30 minutes

Table 3. *Scenario Descriptions.*

After each scenario, the forecasters filled out an online workload survey. Next, the entire group gathered for a 45-minute discussion of that specific scenario. At the end of all the scenarios on Thursday, the forecasters filled out an online usability survey. The final day (Friday), a half day, was spent conducting a 3-hour guided interview with the forecasters to gather their feedback and experiences during the experiment week.

Before the COVID-19 pandemic, we conducted in-person HWT experiments using physical workstations on site. During the pandemic, we conducted our experiments virtually using a version of the AWIPS software hosted by the NOAA Virtual Laboratory (VLab) within Amazon Web Services. The cloud-based software also allowed us to carry out software testing in a common environment, where developers could connect simultaneously in a collaborative, yet virtual, environment. We decided to use the virtual system within the HWT to conduct the experiment instead of maintaining the latest software builds on two systems (the physical and virtual systems). We learned to split a virtual instance across two monitors, expanding screen real estate, which enabled us to run one Common AWIPS Visualization Environment (CAVE) instance per monitor. One monitor was dedicated solely to Hazard Services, and the other could be used for one (or more) CAVEs to peruse and analyze any other data. There were some pros and cons to this approach:

- Pros: Developers had quick and convenient browser access to the cloud systems from anywhere, enabling quick software tests. Developers could diagnose problems quickly without having to ask forecasters to leave their workstations. Additionally, because we were operating with strict COVID protocols, if any of our participants fell ill with little or no symptoms, they could still participate virtually.
- Cons: The virtual instances ran slightly slower than the physical systems. Hopefully, this will improve as cloud computing technology advances.

Fig. 8. *Photos from the Tiny TIM experiment in action at the HWT.*

Fig. 9. *Group photos from Weeks 1 and 2.*

These were the specific objectives of the experiment:

- Technology: Evaluate HS-Tiny TIM components and performance to improve the software before operational implementation.

- Human Factors: Measure forecaster workload using HS-Tiny TIM, including ease of use and graphical design.

- Methodology: Assess how forecasters adopt their legacy warning methodology into the HS-Tiny TIM environment as they create and manage continuous, feature-following warnings.

- Conceptology: Collect and analyze data on forecasters' thoughts on the paradigm change from "static" warnings to continuously-updating warnings.

We gathered participant feedback on user experience, operational applications, and workflow considerations. The data collected will be used to further refine both the software and the concepts of operations of TIM. The following methods were used (those indicated with an asterisk were also recorded, and speech-to-text transcripts were created):

- Pre-operations online survey
- Discussions during events with meteorologists and developers
- Post-event online workload surveys
- Post-event discussion*
- End-of-week online usability survey
- End-of-week interview*

2023 Results

The Tiny TIM software was relatively stable throughout the experiment, but a few issues were discovered that will be addressed post-experiment:

- Expired storms (EXP) were not disappearing from the experimental warning overlay.
- For certain durations (30 min), the accordion effect mitigation technique was adding too much area.
- When a warning crosses a CWA border, sometimes the warning edge is composed of many points rather than just two points.
- A "No projected tracking points found within polygon" pop-up sometimes appears when the footprint centroid is outside or nearly outside the current CWA.

- It is not possible to switch the storm track tool type in the HID for Tornado Warnings.
- Several times when expiring a warning, we received a warning text with two "EXP" sections instead of one.

Some other software changes were made during the experiment based on feedback from the forecasters. These included:

- The "CON" feature was initially designed to handle situations where certain warnings couldn't be extended in time, such as lower-end severe warnings. However, all forecasters unanimously agreed that any warning, regardless of intensity, should have the option for extension. As a result, the feature has been disabled and is unlikely to be included in the final software.
- The Dynamic Footprint option was deemed unnecessary and has been removed due to the absence of practical use cases.
- After a few scenarios, the Accordion Effect mitigation technique was disabled. This decision was made because the output was confusing to forecasters. They had difficulty distinguishing which parts of the warnings were being retained due to the mitigation, and whether the warning extension box controls were functioning as intended.

The forecasters made several notable software suggestions. For the Accordion mitigation feature, they recommended adding an alert when it is activated. This can be achieved by displaying a small accordion icon next to a toggle (ON) on the HID. Users should have the option to toggle accordion areas off and differentiate the new areas using "alternate" hatching or a different color shade.

Regarding the HID, the forecasters proposed replacing the Advect toggle with a "Reset Shape" button. This button should only become active or visible after a polygon is edited. Clicking the button would reset the shape to the default polygon. Additionally, the shape should automatically reset when switching Tool Type or editing footprint or line tool vertices (excluding when dragging the tool).

Regarding the Storm Track Tool, they suggested adding a centroid point to the footprint tool to assist with motion tracking. They also recommended that the buffer controls should adjust the size of the polygon even if the default polygon is not used. Additionally, they proposed implementing an automated method to "sense" the CWA border and suggest reducing the duration. A pop-up notification should appear if the warning extends outside the CWA, providing an option to decrease the duration or to press a button to do this automatically.

For the Console, the forecasters suggested adding a "Time Since Last Modified" column. They also proposed changing the default sorting of console columns to be based on the username, prioritizing the workstation's user as the first entry. The secondary sort should be in ascending order based on the expiration time. Furthermore, they recommended including a button to reset the default sorting if it has been modified.

2023 Appendix A contains a comprehensive list of suggestions for improving the software functionality provided by the forecasters, including those mentioned above. Some of these suggestions will be considered for incorporation into future versions of the software, subject to funding and budget availability, in order to enhance its robustness. Certain enhancements fall within the broader scope of HS - Convective as a whole, rather than being specific to Tiny TIM, and are identified as such. Furthermore, the forecasters also made suggestions to enhance the logistics of future experiments, which are listed in 2023 Appendix B.

This experiment also incorporated a basic human factors analysis through online surveys, aiming to assess mental workload, confidence levels, and software usability. Further details regarding the human factors analysis can be found in 2023 Appendix C.

Tiny TIM offers the advantage of faster updates, with the forecasters suggesting an update rate of approximately every 10 minutes per warning, taking into account workload considerations. However, for high-impact events, more frequent updates, at least every 5 minutes, can be made. It is important to note that intervals of less than 5 minutes would significantly increase workload and diminish the overall impact. An exception to this general rule would be when a report is received, in which case the warning should be updated as soon as possible. If a forecaster is currently editing a warning but receives a tornado report for a different unwarned storm, they can "pause" the editing process for the current warning update and urgently address the tornado report by issuing a new TOR. This flexibility is enabled by Hazard Services, which allows for the simultaneous editing of multiple warnings (unlike WarnGen). Once the more-urgent warning is issued, the forecaster can then return to finish editing the original warning. However, there is a potential downside to the ability to conduct faster updates. Forecasters might feel compelled to update warnings more frequently than necessary, which could lead to a loss of focus on other storms within the CWA.

During the experiment, several discussions and considerations arose regarding the handling of QLCS vortices and the issuance of warnings. One question was whether to issue separate warnings per vortex or apply a blanket warning for the entire line. Additionally, there was a need to clarify the approach for Tornado Emergency (TOR-E) warnings. While it was customary to create a new warning for TOR-E (at least in Central Region), it was emphasized that with Tiny TIM, this practice was unnecessary, and the same warning could be utilized. It was suggested that dissemination systems should be designed to issue new alerts for the entire polygon whenever there is an upgrade to a TOR-E. Finally, there were discussions about the duration of a TOR-E (Tornado Emergency) warning, with a suggestion that it should only apply to the initial 15-20 minutes of the warning, rather than encompassing the entire duration. In this context, the use of a CAP-formatted warning was considered beneficial. It was proposed that an "inner" polygon could be implemented to specifically target the TOR-E area, ensuring that only those areas within the designated region would receive a new warning alert if they were already under an existing warning.

One of the experiment scenarios involved a classic left-hooking occluding tornado cyclone (4-5 May 2022 near Lockett, TX), which posed a challenge for the original warning forecaster due to the existing paradigm where warnings cannot be extended. Figure 10 illustrates the combined tornado warnings issued during the 0130-0230 UTC one-hour period, overlaid on the "rotation tracks" product depicting the left-hook path of the tornado cyclone. On the left side are the original NWS warnings for the event, while on the right side are Tiny TIM warnings created by the same NWS warning forecaster who issued the original warnings. Figure 11 presents a time series of the same warnings for this event. It is noteworthy that in some time frames, there were two, three, and occasionally ***four*** separate warnings in effect for the same single hazard. With Tiny TIM, there was only one warning in effect for the hazard at any given time. Throughout different time periods, that warning would be extended in time and area, following the path of the hazard.

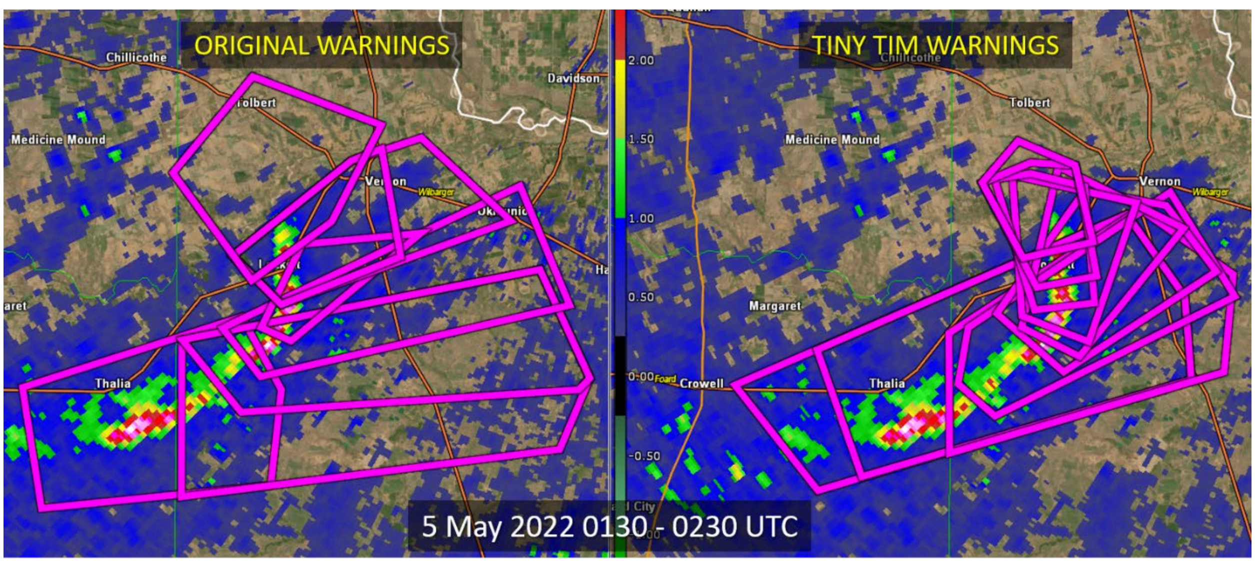


Fig. 10. *Rotation tracks for the Lockett, TX, tornado event of 4-5 May 2022. On the left is a composite of all the tornado warnings issued by the NWS Norman. On the right is a composite of all the Tiny TIM tornado warnings issued by the original NWS forecaster but within the HWT experiment setting.*

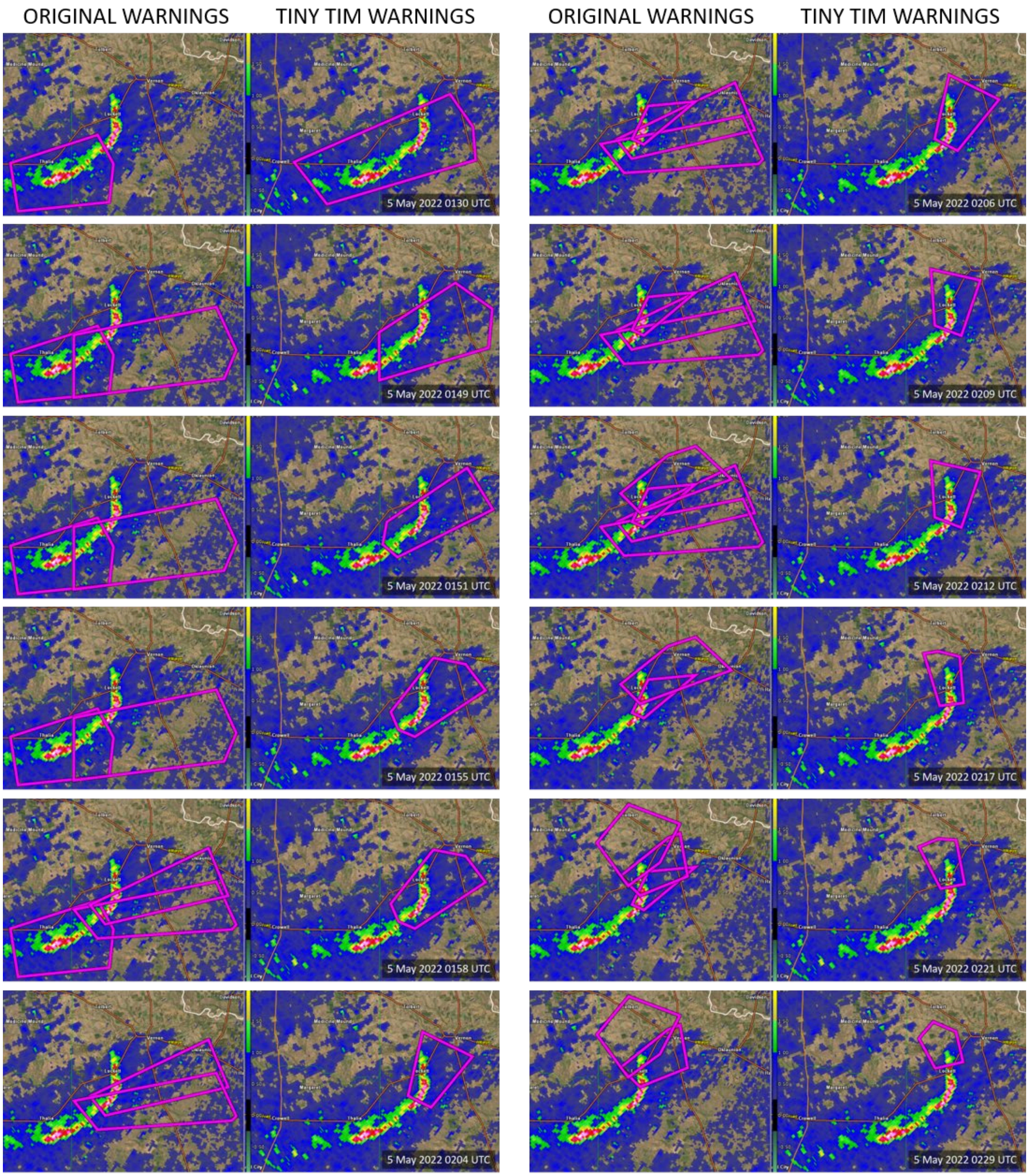

ORIGINAL WARNINGS
TINY TIM WARNINGS
ORIGINAL WARNINGS
TINY TIM WARNINGS
5 May 2022 0130 UTC
5 May 2022 0149 UTC
5 May 2022 0151 UTC
5 May 2022 0155 UTC
5 May 2022 0158 UTC
5 May 2022 0204 UTC
5 May 2022 0206 UTC
5 May 2022 0209 UTC
5 May 2022 0212 UTC
5 May 2022 0217 UTC
5 May 2022 0221 UTC
5 May 2022 0229 UTC

Fig. 11. Time mosaic of the original NWS warnings (displayed on the left side of each sub-figure) and the Tiny TIM warnings (displayed on the right side) from 0130 to 0230 UTC on 5 May 2022.

There are only a few remaining changes necessary to bring the Tiny TIM software to operational Readiness Level 7. The development team has identified the most important bug fixes and new features to be implemented within the remaining timeframe of the FY23 grant (September 30, 2023). Following FY23, the focus will shift towards advancing the infrastructure component to facilitate dissemination.

Emphasis will be on utilizing the Common Alert Protocol (CAP) Formatter instead of VTEC to enable warning extensions. CAP offers greater flexibility and aligns with the future dissemination path of the NWS. It allows for the inclusion of new fields using XML tags, which can be considered "optional" and utilized by compatible dissemination systems. For instance, multiple polygons can be defined for different portions of the warning: 1) areas that are added (triggering an alert), 2) areas that remain warned (no alert unless upgraded to a Tornado Emergency), and 3) areas that are removed (triggering potential "All Clear" alerts). Additionally, higher-resolution polygons can be incorporated, such as a 100-point maximum for enhanced precision, as well as "inner" TOR-E polygons. With this framework in place, future software experimentation will likely be conducted at the OPG, where a comprehensive dissemination test involving AWIPS Network Control Facility test system (NCF-Test), NOAA Weather Radio (NWR), Wireless Emergency Alerts (WEA), and private industry partners should be planned.

Finally, several functionalities have been developed for Tiny TIM which could be implemented between the initial operating capability of HS - Convective and the deployment of Tiny TIM. These functionalities can be utilized within the current warning paradigm and include the following features: Auto-Update, Reset Shape (replacing Advect), Update Rear (without extending the area or time), Buffer (-/+) controls, Dynamic Line Tool, and 2D Footprint Tool.

2023 Testimonials

> “Less time spent on adjusting polygons/the polygonology, more time available to spend interrogating storm/environment data. Warning operations are always fairly complex, but the functionality of Tiny TIM removes a lot of the time sink that is tweaking vertices to perfection (especially as someone with minimal experience in warning ops in general).”

"Overall, I felt I was doing minimal hands-on edits to my polygons... the software did almost everything for me (and with enough success that I didn't feel the need to make more than minor edits)."

"Cannot understate how much of the time sink that is tweaking individual vertices/having to completely readjust polygons given storm motions Tiny TIM *completely* removes."

"With TIM, I just readjusted my footprint and everything was righted -- a whole new set of warnings would've been necessary for our current system."

"In my opinion, it is much simpler to use this, even when problems arise, than WarnGen on its best day."

"The one very area that saves a lot of time is allowing the warning to shift slightly to account for changes in the storm motion, rather than issuing a brand-new warning with takes more time and causes dissemination/communication issues."

"Using Tiny TIM left more time for me to interrogate storms and make warning decisions rather than fiddle with new warnings."

"Tiny TIM creates less confusion because there are fewer polygons (cleaner warning display)."

"This is one of the first times that felt I had enough time to issue separate tornado and severe warnings on supercells that had separated hazards without much additional workload overall."

"Cannot underestimate how much having the Update Rear/Front/Both box controls improves workload and are way more efficient than WarnGen, and it makes me incredibly sad that I have to go back to regular warning ops in WarnGen where I have to fight with vertices in between updates."

"The 2D Footprint Tool latches onto mesocyclones or hail cores extremely well, and the "customized" polygons that result from the shape of the footprint require little to no updates. The 2D Footprint Tool is a game changer in terms of accurately capturing the hazard."

"I think the great majority of warning forecasters would be able to understand Tiny TIM without a complete paradigm shift in training."

"After the initial learning curve, I was able to get very comfortable using this system. It was similar enough to WarnGen, but the added tools made it even better. It allowed me to warn for storms in a more intuitive and people-centric manner."

"I think the learning curve could be a little steep for some forecasters, especially the OG ("old guard"), so having thorough training will be crucial to making the transition as smooth as it can be."

"I was able to learn this system easily, but I also think using it for four days made it easier to learn how to use it. I think the field will have a little more growing pains than us since they'll be taking Commerce Learning Center modules and testing the software at random and infrequent times. It will take longer for the field to feel comfortable with and learn to use the system."

"It is important to test every kind of possibility if this is going to be our new warning paradigm. For example, short-lived pulse storms with weak steering aloft won't take much advantage of the Tiny TIM features, but even these could move out of polygon and require area extensions."

"I am very satisfied with Tiny TIM and am sad to have to go back to using WarnGen when I go back to my office. I think Tiny TIM is going to be a big leap forward for our warning operations and will greatly decrease workload and frustration with polygons, and I hope that it is implemented as soon as possible!"

"Please let me take this system home with me. I dread issuing warnings with WarnGen when I know that this technology exists."

2023 Key Takeaways

- Tiny TIM significantly reduces forecaster workload and mental demand, particularly for long-tracked hazards.
    - Warning updates can be executed faster and with improved spatial accuracy.
    - The option to extend warnings eliminates the need for separate downstream warnings, reducing decision-making complexity.
    - Timely updates ensure that storms rarely exit the warning polygons, allowing for adjustments if the previous polygon is not optimal.
- The elimination of overlapping warnings for a single hazard results in a cleaner output field and improved messaging. There is no longer a need to issue a completely new

warning and manage two separate warnings (the new warning and the imminent cancellation of the previous one) for the same hazard when it appears to be leaving the polygon. Tiny TIM addresses these situations:
    - When a new downstream warning overlaps with the previous one, canceling the previous warning can create confusion.
    - Overlapping warnings require the simultaneous broadcast of two warnings on NWR and television crawls, leading to potential confusion.
- There are only a few remaining software capabilities that need to be added, and these will be completed during the remainder of FY23. This will allow future efforts to focus on advancing the infrastructure component to facilitate dissemination.
- There is unanimous agreement among forecasters that warning extensions should be allowed for **any** long-tracked storm, regardless of the presence of observed significant hazards.

2023 Future Plans

- Document, analyze, and prioritize newly identified and forecaster-suggested functionalities, modifications, and other software issues specific to Tiny TIM, including those listed in the Appendices.
    - Complete the remaining software work for FY23.
    - Regularly integrate the Tiny TIM code into the operational version of HS to stay current with HS-Convective development.
- Continue familiarizing forecasters with Tiny TIM and work toward starting a preliminary training plan for Tiny TIM that spans multiple divisions of the Office of the Chief Learning Officer within the NWS.
- Propose a multi-year NWS collaboration between the Office of Science and Technology Integration (STI), the Analyze, Forecast and Support Office (AFS), and the Office of Dissemination (DIS) to advance the infrastructure component to facilitate dissemination. The focus will be on adopting the CAP Formatter instead of VTEC to enable warning extensions. This effort includes planning a comprehensive dissemination test at the OPG involving NCF-Test, NWR, WEA, and private industry partners.
- Update the TIM transition plan for Tiny TIM based on the aforementioned points.

Personnel

Greg Stumpf	CIRA* and NWS/MDL	Experiment Coordinator / Project Co-PI
Kevin Manross	CIRA and NOAA/GSL	Lead Developer / Project Co-PI

Alyssa Bates	CIWRO** and NWS/WDTD	Project Scientist / Project Co-PI
Jim Ramer	CIRA and NOAA/GSL	Developer
Yujun Guo	CIRA and NOAA/GSL	Developer
Emily Schlie	CIRA and NOAA/GSL	Developer
Pat Hyland	CIWRO and NOAA/NSSL	Project Scientist
Justin Monroe	CIWRO and NOAA/NSSL	HWT AWIPS Support
Jonny Madden	CIWRO and NOAA/NSSL	HWT AWIPS Support
Kodi Berry	NOAA/NSSL	NSSL FACETs Program Manager

*CIRA - Cooperative Institute for Research in the Atmosphere
**CIWRO - Cooperative Institute for Severe and High-Impact Weather Research and Operations

Disclosure
ChatGPT was used to help proofread portions of this summary.

References
NWS, 2020a: WFO severe weather products specification. National Weather Service Instruction 10-511, 64 pp., http://www.nws.noaa.gov/directives/sym/pd01005011curr.pdf.

Stumpf, G. J., and A. E. Gerard, 2021: National Weather Service severe weather warnings as Threats-in-Motion (TIM). *Wea. Forecasting*, **36**, 627-643. https://doi.org/10.1175/WAF-D20-0159.1.

Stumpf, G. J., 2022a: 2022 Tiny – Threats-In-Motion (Tiny TIM) HWT Experiment Summary.

Stumpf, G. J., 2022b: 2022 Tiny – Threats-In-Motion (Tiny TIM) OPG Experiment Summary.

Web Presence

- NSSL Bite-Sized Science 3-minute video on Threats-In-Motion
- A Blog summary about Tiny TIM (use NOAA credentials to log in)
- The TIM Weather and Forecasting journal article (Stumpf and Gerard, 2021).

2022 Appendix A: Software Functionality Suggestions

Warning Output:

- Disallow the "accordion effect": Update Back/Front/Both should never automatically remove areas on the front end of the warning if the motion vector speed is decreased. *This falls under the category of incorporating any specific functionalities/modifications to address issues with dissemination.*

Object Editing:

- A dashed "former" polygon should be shown in the spatial display when a warning is being updated, so that forecasters can see what they are changing their polygon from. This helps alleviate the potential slight addition to the workload when the user has to decide which update options and how much to extend the warning time, in order to avoid the windshield wiper and accordion effects. *This falls under the category of incorporating any specific functionalities/modifications to address issues with dissemination.*
- Add a Tiny TIM storm track tool that tracks with a 5-km circle.
- Be able to draw a footprint from scratch (polygon, freeform, ellipse), and also use a "replace" feature during updates, like in HS-TIM.

Spatial Display:

- Warning labels should include ETN and user/workstation.
- Add the ability to change the warning grid overlay to dashed, blinking, different colors.

Console:

- Add workstation/username to the default Console columns.
- Add IBW tags to Console columns.
- The Console time bar should be toggleable.

HID:

- The "Details" portion of the HID should start scrolled to the top when a warning update is initiated.
- Change the name of Update Back to Update Rear, because "Both" and "Back" begin with the letter "B" and this can be confusing.
- Add the capability to decrease the end time. This would be useful if storms begin to weaken quicker than expected, or if storms are moving out of the warning area.

2022 Appendix B: Experiment Logistics Suggestions

Remain on the leading edge of HS-Convective development by integrating with the operational branch often. Be cognizant of the current operational software issues and what is going to be fixed before HS-Convective becomes operationally "ready". To this end, add a toggle to switch Tiny TIM on and off to test differences using a localization override.

Consider having half-day functional assessment tests (FAT) with forecasters to get feedback on new functionalities developed since the experiment. Invite former Tiny TIM HWT experiment participants or HS-Convective FAT participants due to their familiarity with the software.

Tracking is currently based on the centroid of the footprint. Need to determine if the tracking can be based on the leading edge.

Establish use cases to determine what development is needed for specific scenarios, including impacts to dissemination. Use cases should include figures of how the software would be used for those kinds of cases:

- Create use cases for EXA and EXT. For example:
    - EXA:
        - Issue warning, but forgot to include additional locations. Want to extend the area but keep the end time constant. Would be used with Fixed or Update Back.
    - EXT:
        - Issue warning and want to correct the duration to make it longer.
        - A stationary storm that sticks around longer than expected, the area does not need to be extended but the end time does need an extension.

- Create a use case for a squall line. Currently, the best practice taught by NWS/WDTD is to place the line tool along the leading edge of the hazard. As an extra step, the back of the polygon is then dragged to cover the whole storm. Because the update recommenders currently replace custom polygons with a default tied to the track, this process would need to be repeated for each warning update. With Tiny TIM, the footprint tool could be used to remove the extra step.

- Create a use case for a storm moving out of the side of the warning polygon and compare how these would be handled with WarnGen and with Tiny TIM. The Dayton, OH, tornado on 5/28/19 makes a good example.

- Create use cases with training storms, with varying distances between hazards. Should warning overlap be avoided, and if so, at what distance threshold? If the separation is small, should warnings be overlapped anyway, and paired with the use of messaging tools to lessen confusion?

2022 Appendix C. Human Factors Analysis Survey Results

HWT Participant Data

Four NWS forecasters participated in the 2022 Tiny TIM Virtual HWT, with two forecasters each week. The forecasters' work experience ranged from 3.5 years to 22 years, and the average was 12.3 years (standard deviation 9.6). Their average NWS warning experience ranged from 3 years to 17 years, with an average of 9.9 years (standard deviation of 7.1).

Mental Workload (NASA TLX) Survey

The NASA-TLX (Hart & Staveland, 1988; Hart, 2006) workload index is a questionnaire-based workload rating tool. The tool encompasses 6 aspects of workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. The raw scores of the mental workload ranges from 1 to 100, where 1 stands for extremely low workload and 100 stands for extremely high. The ratings were averaged from all the sessions for each of the 6 aspects of workload.

Table A1 shows the average ratings for the six sub-dimensions and the overall workload for three archived hazardous weather scenarios. The average workload for the 2022 Tiny TIM Virtual HWT experiment across all testing scenarios was 52.4 (out of 100, standard deviation 24.30). Each of three scenarios was chosen with an increasing level of difficulty (more storm coverage) throughout the course of the week, and the workload ratings increase accordingly.

Table A1. 2022 Tiny TIM Virtual NASA-TLX Mental Workload Rating for Three Test Scenarios. Scenario 1 was a training scenario and was not included here.

	Scenario 2 Mean (std)	Scenario 3 Mean (std)	Scenario 4 Mean (std)
Mental Demand	63.3 (10.01)	72.5 (8.89)	84.5 (7.00)
Physical Demand	47.5 (29.35)	56.8 (19.6)	61.3 (14.71)
Temporal Demand	43.3 (23.64)	65.8 (10.84)	70.8 (11.73)

Effort	61.0 (14.14)	72.8 (11.90)	67.8 (13.38)
Frustration	15.0 (5.23)	29.5 (12.71)	24.3 (7.63)
Performance	32.0 (21.68)	43.8 (29.20)	32.5 (24.24)
Overall Workload	43.7 (24.03)	56.8 (22.06)	56.8 (25.27)

PSSUQ Usability Questionnaire

The Post Study System Usability Questionnaire (PSSUQ; Lewis, 2002) is a survey tool designed to evaluate usability of a computer system. The tool is designed with 19 usability questions to assess 4 different areas of System Usefulness (Questions 1-8), Information Quality (Questions 915), Interface Quality (Questions 16-18) and Overall Usability (Questions 1-19). The rating ranges from 1 to 7, where 1 corresponds to low level of usability, 7 to high level of usability, and 4 corresponds to a neutral level of usability.

The PSSUQ questionnaire was filled out by the participants after they completed all test scenarios in the testbed. Table A2 shows the average responses for each of the 4 categories: System Usability, Information Quality, Interface Quality, and Overall Usability.

Table A2. Usability Ratings based on the PSSUQ for 2016 – 2022 HS-PHI, TIM, and Tiny TIM HWT Experiments (7-point scale)

	2016	2017	2018	2019	2021	2022
System Usability	4.96	5.56	5.21	4.88	6.25	**6.03**
Information Quality	4.37	5.00	4.37	4.81	5.71	**6.23**
Interface Quality	4.72	5.72	5.50	5.18	5.78	**5.92**
Overall Usability	4.62	5.39	4.95	4.92	5.97	**6.09**

The overall usability was assessed at 6.09 (on a 7-point scale) for the 2022 Tiny TIM Virtual experiment, with system usability rated at 6.03, information quality at 6.23, and interface quality at 5.92. It is worth noting that the usability rating has improved between 2016-2022. There are several factors accounting for this. First, the first four years included a probabilistic component to the software, which increased workload. Second, in 2022, we only tested the new Tiny TIM software. Third, the software has become more stable throughout the years. Finally, suggestions made by forecasters to improve the workflow have been implemented over the years.

References

Hart, S. G., and L. E. Staveland, 1988: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, **52**, 139–183, https://doi.org/10.1016/S0166-4115(08)62386-9.

____ , 2006: NASA-Task Load Index (NASA-TLX); 20 years later. *Proc. 50th Annual Meeting, Human Factors and Ergonomics Society*, San Francisco, CA, 904–908, https://doi.org/10.1177/154193120605000909.

Lewis, J., 2002: Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction*, 14 (3-4), 463-488, https://doi.org/10.1080/10447318.2002.9669130.

2023 Appendix A: Software Functionality Suggestions

NOTE: Enhancements that fall within the broader scope of HS - Convective as a whole (and not specific to Tiny TIM) are indicated with an "(HS-C)" prefix.

Hazard Information Dialog (HID):

- Replace the Advect toggle with a “Reset Shape” button.
    - It should be inactive/invisible until a polygon is edited. Clicking it would reset the shape to the default polygon.
    - Also, the shape should automatically reset when:
        - Switching Tool Type.
        - Editing footprint or line tool vertices (excluding when dragging the tool).
- Add editable fields for motion vector speed and direction.
- Change the names of the Box Control options. Some suggestions include:

Current Name	**Suggestion A**	**Suggestion B**
Fixed	Fixed	Anchor Both
Update Rear	Remove Rear	Anchor Front
Update Both	Update Both	Anchor None
Update Front	Extend Front	Anchor Rear

- If the extension time does not equal 0 minutes, the label should turn red.
- (HS-C) Enclose the Tornado Emergency selection within a magenta-colored box to prevent accidental clicking.
- (HS-C) Add a pop-up alert if Preview is clicked before any edits are made.
- (HS-C) Add multiple "first-guess" locations to the DESTRUCTIVE tag's specific locations field.
- (HS-C) Add a previous text button for the DESTR/CATAS locations field.
- (HS-C) Make sure the highest IBW tag is “greater than” for hail and wind.

Object Editing (Storm Track Tool):
- Add a centroid point to the footprint tool to assist with motion tracking.
- The buffer controls should adjust the size of the polygon even if not using the default polygon.
- Implement an automated method to "sense" the CWA border and suggest a reduction in duration.
    - Include a pop-up if the warning extends outside the CWA, suggesting decreasing the duration or providing a button to do it automatically.
- The initial symmetric hexagon footprint for TOR should be half the size of an SVR, similar to what the point tool does for TOR.
- Introduce a "QuickWarn" feature to issue TOR warnings in less than 30 seconds using the hexagon footprint.
- For the Line and Dynamic Line Tools, provide separate controls for the front and rear buffers.
- Add the ability to set the default number of vertices for the 2D Footprint Tool (currently set to 6).
- Enable independent movement of the 2D Footprint centroid to the location of the greatest hazard (e.g., biggest hail) to emphasize that location in the warning text.
- Allow drawing a footprint from scratch (polygon, freeform, ellipse) instead of starting with a perfect hexagon.
- (HS-C) When initiating the STT on a past radar frame, reset the display to the latest radar frame.
- (HS-C) Include the ability to replace an SVR with a TOR (and vice versa) with a single action. The older warning should automatically be canceled (CAN) with a message stating that it has been replaced by the newer warning. The newer warning should include a message indicating that it replaces the older warning. This capability already exists in HS-Hydro.
- (HS-C) Allow "linking" (and "unlinking") of an SVR and TOR on the same storm, so that when updating the TOR, the SVR automatically updates (moves the polygon). Updating the maximum hail on the SVR would also update it on the TOR.
- (HS-C) Add a "snap to another polygon" feature.

Spatial Display:
- (HS-C) Add storm motion values to the StormTrackTool decorations.
- (HS-C) Require a Ctrl-Click to select a warning from the display, to prevent accidental selection while panning.

Console:
- Add a "Time Since Last Modified" column to the console.
- The default sorting of console columns should be by username (e.g., "my warnings"), prioritizing the workstation's user as the first entry. The secondary sort should be in ascending order based on the expiration time. ○ Include a button to reset the default sorting if it has been changed.
- (HS-C) Change default warning threshold alerts.

	Alert	**Time**	**Text Color**	**Background Color**	**Popup**	**Blink**
Current	1st	15 min	yellow	black	no	no
	2nd	5 min	red	white	yes	no
Proposed	1st	15 min	yellow	grey	no	no
	2nd	5 min	red	black	yes	yes

- (HS-C) Include a filter for warning type (TOR or SVR).
- (HS-C) Automatically re-enable some warnings in the Console after 5 minutes when they have been clicked off.

Accordion:
- Accordion mitigation should provide an alert when it becomes activated:
  - Display a small accordion icon next to a toggle (ON) on the HID. Users can toggle accordion areas off.
  - Use "alternate" hatching or a different color shade to distinguish the new areas.

2023 Appendix B: Experiment Logistics Suggestions

- Best practices to add to training:
  - When handling warnings exiting CWAs, continue using EXB/Update Both, but reduce the duration so that the arrowhead is just outside the CWA. Then, on the next update, start using Update Rear.
  - When dragging to a storm on a past frame, do not step forward the same number of frames. Instead, use the "latest frame" hotkey to save time.

  - When there are many active warnings, toggle off the other warnings to make it easier to see your warning. Just remember to toggle them back on later.
  - If a forecaster is editing a warning and receives a report of a tornado for a different unwarned storm, inform users that they can "pause" the editing on the current warning update and move to the location of the tornado report to issue a new TOR. Afterward, they can go back and finish editing the first warning.
  - For most situations, use the following:
    - Footprint Tool for non-linear SVR warnings.
    - Line Tool for linear SVR warnings.
    - Point Tool for TOR.
  - Emphasize that even though updates are faster, it does not mean that forecasters should feel compelled to update more frequently.
- New scenario suggestions:
  - Include a scenario with marginally-severe pulse convection to challenge forecasters on whether or not to extend warnings.
  - Include a scenario with a simultaneous flash flood threat to increase workload and incorporate the use of HS-Hydro to handle flood hazards.
  - Consider using the system with real-time weather, where isolated long-tracked storms are not guaranteed, unlike in our archive scenarios.
- Experiment logistics:
  - Obtain Warn on Forecast System (WoFS) archive data for the MesoA.
  - Provide a quick reference guide for frequently encountered issues, questions, or best practices.
  - Include local county/city maps for each scenario since the forecasters are likely unfamiliar with those WFOs.
  - Include a recorded demo before the HWT visit.
  - Allow experiment non-participants to learn and practice with Tiny TIM via the virtual cloud instance and job sheet exercises.
  - Inform participants ahead of time that they will be starting from scratch regarding procedures, as their customized local procedures do not align well with the HWT's AWIPS system, which involves transitioning from WFO to WFO. In the initial email, add an action item to consider which procedures they would use for the warning desk and the MesoA desk.
- Outreach:
  - Arrange for one of the PIs to give a VLab Forum seminar.
  - Publish a journal article about Tiny TIM.

2023 Appendix C. Human Factors Analysis Survey Results

HWT Participant Data

Six NWS forecasters participated in the 2023 Tiny TIM HWT experiment. The forecasters' work experience ranged from 2 years to 14 years, with an average of 7.0 years (standard deviation of 5.5). Their average NWS warning experience also ranged from 1 year to 13 years, with an average of 4.2 years (standard deviation of 4.5).

Mental Workload (NASA TLX) Survey

The NASA-TLX (Hart & Staveland, 1988; Hart, 2006) workload index is a questionnaire-based workload rating tool. The tool encompasses six aspects of workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. The raw scores for the mental workload range from 1 to 100, where 1 represents an extremely low workload and 100 represents an extremely high workload. The ratings were averaged from all sessions for each of the six aspects of workload.

Table 1 shows the average ratings for the six sub-dimensions and the overall workload for five archived hazardous weather scenarios. Forecasters who solely performed WC/MesoA duties without issuing warnings were not included in the count. The average workload for the 2023 Tiny TIM HWT experiment across all testing scenarios was 44.9 (out of 100, with a standard deviation of 24.2). Each of the five scenarios was chosen with an increasing level of difficulty (more storm coverage) throughout the course of the week. But note that after the middle scenario (#4), the workload ratings started to decrease as the forecasters became more accustomed to using the software.

Table 1. 2023 Tiny TIM NASA-TLX Mental Workload Ratings for Five Test Scenarios. Scenario 1, a training scenario, is not included in this table. The order of the scenarios in the table corresponds to the order in which each scenario was worked throughout the week (Tuesday: S3, S7; Wednesday: S4, S5; Thursday: S6).

	Scenario 3	Scenario 7	Scenario 4	Scenario 5	Scenario 6
	Mean (std)				
Mental Demand	55.2 (21.3)	58.3 (25.0)	64.2 (29.3)	46.3 (15.7)	54.2 (19.6)
Physical Demand	38.8 (14.8)	28.0 (9.1)	50.6 (31.3)	39.0 (25.8)	41.3 (16.3)

Temporal Demand	46.0 (19.7)	43.8 (34.7)	62.0 (30.7)	44.0 (23.1)	37.3 (23.0)
Effort	58.2 (23.2)	59.0 (28.1)	67.8 (28.0)	54.8 (21.8)	53.2 (26.6)
Frustration	30.2 (30.9)	42.8 (19.6)	58.8 (27.1)	28.5 (15.8)	23.8 (12.7)
Performance	40.6 (25.9)	40.5 (6.8)	40.8 (24.8)	33.2 (23.7)	18.3 (9.7)
Overall Workload	**44.8 (23.2)**	**45.4 (23.0)**	**57.4 (27.6)**	**41.0 (21.6)**	**38.0 (22.2)**

PSSUQ Usability Questionnaire

The Post Study System Usability Questionnaire (PSSUQ; Lewis, 2002) is a survey tool designed to evaluate the usability of a computer system. The tool consists of 19 usability questions that assess four different areas: System Usefulness (Questions 1–8), Information Quality (Questions 9–15), Interface Quality (Questions 16–18), and Overall Usability (Questions 1–19). The rating scale ranges from 1 to 7, where 1 corresponds to a low level of usability, 7 to a high level of usability, and 4 represents a neutral level of usability.

The participants completed the PSSUQ questionnaire after finishing all test scenarios in the testbed. Table 2 displays the average responses for each of the four categories: System Usability, Information Quality, Interface Quality, and Overall Usability.

Table 2. PSSUQ Usability Ratings for the three Tiny TIM experiments (7-point scale).

	2022 Tiny TIM HWT	2022 Tiny TIM OPG	2023 Tiny TIM HWT
System Usability	6.04	6.41	6.46
Information Quality	5.51	6.16	5.93
Interface Quality	5.67	6.67	6.44
Overall Usability	**5.79**	**6.36**	**6.25**

The overall usability was assessed at 6.25 (on a 7-point scale) for the 2023 Tiny TIM HWT experiment, with system usability rated at 6.46, information quality at 5.93, and interface quality at 6.44. It is worth noting that the usability rating improved from the first to the second Tiny TIM experiment and then remained relatively steady for the third experiment. This trend suggests that software improvements made between each experiment may be reaching a plateau, indicating that we are nearing the final version of the software before the first operational integration.

References

Hart, S. G., and L. E. Staveland, 1988: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, **52**, 139–183, https://doi.org/10.1016/S0166-4115(08)62386-9.

____ , 2006: NASA-Task Load Index (NASA-TLX); 20 years later. *Proc. 50th Annual Meeting, Human Factors and Ergonomics Society*, San Francisco, CA, 904–908, https://doi.org/10.1177/154193120605000909.

Lewis, J., 2002: Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction*, 14 (3-4), 463-488, https://doi.org/10.1080/10447318.2002.9669130.

Probabilistic Hazard Information (PHI) Prototype Experiment

Summary by Kristin M. Calhoun

Overview

During the 2022-2023 Probabilistic Hazard Information (PHI) Prototype experiments NWS forecasters issued PHI in combination with storm-based warnings using a cloud-based web platform. Both spring experiments evaluated PHI creation and use across multiple levels of the warning decision process with a focus on addressing the feasibility of continuous forecaster generation of PHI and Threats-In-Motion (TIM) warnings simultaneously.

In 2022, the experiment was designed to examine using PHI: 1) as guidance for warning decisions (where one forecaster was responsible for PHI creation and others could use it to issue warnings within the current warning paradigm or 2) in tandem with warning decisions (one forecaster was responsible for creating both PHI and a warning polygon following the TIM paradigm). Additionally, forecasters provided general feedback on the management and use of multiple automated objects and the machine learning algorithms providing the initial probabilistic guidance for the creation of PHI for both severe and tornado hazards.

In 2023, a new relatively unobtrusive notification system was introduced to alert forecasters to actions which may require their attention, such as rapidly increasing hazard probabilities suggested by automated guidance or hazard information which has not been recently updated. Significant upgrades to the stability and quality control of the automated guidance were also integrated into the experiment, and their subsequent impact on forecaster workload was measured. Additionally, communication systems and designs for partners were tested using NWSChat via Slack and social media via a hidden Twitter account. Feedback gathered during the 2023 experiment has provided inspiration and direction for further development of PHI and TIM-related research, as well as guidance for which concepts might successfully progress towards operations through Hazard Services, the next-generation forecast and warning software. This summary will focus on both results from the most recent experiments as well as discuss the potential impacts for future communication of severe storms hazards and warnings.

Experiment Design

Both experiments were conducted virtually with a total of 27 forecasters participating in activities across both years. The virtual design of the experiment allowed for a larger demonstration of the software and algorithm changes as well as provided flexibility for forecaster participation. In 2023, a mixture of forecasters were purposefully chosen to bring back participants that had previously participated in a PHI and/or TIM evaluation to provide a measure of the changes in the system in recent years. Additional forecasters were chosen from all NWS regions with varying backgrounds and expertise in severe weather forecasting and warnings in order to provide diverse feedback. Additionally, in 2023, prior to their participation, forecasters were asked to complete a training exercise using a cloud-based version of the tool for an event (9 May 2016: a supercell and single cell thunderstorms over central Oklahoma) in displaced real-time to gain familiarity with the tools and algorithms. Forecasters that did not previously have experience with the prototype PHI tool found this additional training case to be extremely helpful in providing the fundamental background needed to feel comfortable with the tool on day 1.

Both years, Monday of each week focused on training of the system and additional background on the tools and concepts. Both discussion and a hands-on archive case (21 Oct 2019: hail-producing supercell storms in the Dallas-Ft. Worth and north Texas area) were used to ensure everyone's comfort in using the new system to create PHI and TIM warnings.
In 2022, forecasters participated in two case studies daily Tues-Thurs. This typically consisted of an archive event in the morning and live weather in the afternoon. However, in the absence of live severe weather a second archive case was substituted that day. In 2023, the experiment focused solely on one event daily – either an archive or live weather event depending on the expected likelihood of severe weather that day.

Forecasters could choose between multiple tracked objects and manipulate the speed and motion uncertainty relative to each PHI threat-in-motion. We will also be exploring the ability to produce quick updates to individual aspects of PHI, such as modifying storm motion or location alone without issuing a new product. In addition to automated guidance available within the PHI-prototype system, forecasters will have access to real-time and archived data within the AWIPS-II cloud platform for storm interrogation and analysis

2022 Experiment Results

Archive and Live Weather Case Descriptions

Case	Region & Responsibility	Meteorological Description	Weeks Used
26 May 2019 2130 - 2330 UTC **Archive**	Central Plains across Kansas (groups 1 and 2) and Texas panhandle (group 3). Forecasters were divided into three pairs to cover three different regions and storm types. One forecaster in each team created PHI in the prototype tool while the other issued traditional warnings in AWIPS (potentially using PHI as a recommender).	Ongoing convection in the warm sector in central KS (group 1) with developing supercell storms and increasing tornado potential near the low in northwest KS (group 2). Potential for isolated storm development along dryline extending southward in the TX panhandle region (group 3).	Week 1 (Tues a.m.) Week 2 (Tues a.m.) Week 3 (Tues a.m.)
6 May 2021 2000 - 2200 UTC **Archive**	S. Illinois (group 1); W. Kentucky / Indiana (group 2), W. Tennessee (group 3) Forecasters again separated into three pairs to cover three different regions. Forecasters handled PHI and warnings together for both severe and tornado (split workload by area within domain).	SPC Slight risk over TN and W. KY, marginal risk farther north. Discrete (low-topped) supercell and multicell storms ahead of a growing line of storms associated with a cold front moving southeast. Primary concern was hail and damaging wind gusts further south.	Week 1 (Tues p.m.)
11 Aug 2021 1715-1915 UTC **Archive**	Great Lakes (Wisconsin / Michigan / Indiana). Forecasters were split as three teams of two forecasters each (by state). Each pair had a separate location and decided internally how to handle the workload split (e.g., regionally or by hazard).	Ongoing clusters of strong-to-severe storms with potential for upscale growth to MCS over the MI/IN region. Potential for supercellular and multicellular convection farther west in WI in the warm sector ahead of advancing cold front with CAPE > 3000 J Kg^{-1} and 0-1 km SRH of 100-200 m^2s^{-2}.	Week 1 (Wed a.m.) Week 2 (Wed a.m.) Week 3 (Wed a.m.)
25 March 2021 0000-0200 UTC **Archive**	Southeast US. Forecasters paired in three different NWSFO county areas: Nashville, Birmingham, and Atlanta. One forecaster using the PHI tool and the second of each pair in AWIPS-WARNGEN for warning creation.	Short wave moving into the southeast US. SPC Moderate to High Risk in the area of the surface warm front and across the warm sector. Potential for long track supercell storms with multiple tornadoes.	Week 1 (Wed p.m.)
15 Dec 2021 **Archive**	Central Plains (E. Nebraska / W. Iowa). Forecasters split into two groups of 3, one team north and one	Widespread potential for severe (>80 mph) winds and tornadoes due to developing QLCS associated with a	Week 1 (Thurs a.m.)

	team south. One forecaster handled severe threats, another tornado, and the third on each team simulated communication aspects during a high-end event through Slack/NWS-chat with Emergency Management and a Public "Twitter-like" interface.	midlevel shortwave trough and strong/deepening surface cyclone.	Week 2 (Thurs a.m.) Week 3 (Thurs a.m.)
3 May 2021 1600-1800 UTC **Archive**	Georgia, South Carolina (NWS Offices Atlanta and Greenville/ Spartanburg). Two groups of three forecasters with one forecaster responsible for PHI creation, one issuing warnings via WARNGEN, and one handling social media / NWS chat using both PHI and Warnings.	SPC marginal risk for severe storms with short wave trough ejecting across the Ohio valley area. Ongoing convection across the region is expected to continue with increased tornado threat across the case period. Continued moderate CAPE values (1000-2000 J/Kg) and increasing effective bulk shear of >40 kts as low level jet increases during the event.	Week 1 (Thurs p.m.)
3 May 2022 1945 - 2150 UTC **Live**	Louisville, KY / Cleveland, OH. Forecast teams sectorized by area with forecasters creating PHI with warning decisions together in the prototype tool.	Scattered storms with the possibility of some supercell storms. Primary threat hail/wind as storms develop and group upscale. Enhanced tornado risk with any supercell storms that develop.	Week 2 (Tues p.m.)
4 May 2022 1918 - 2200 UTC **Live**	Texas panhandle and W. Oklahoma (Forecast pairs in separate NWSFO county areas: 1. LBB, 2. AMA, and 3. OUN). Forecasters created PHI and made warning decisions within the PHI tool. Each forecast team decided internally how to handle the workload split (e.g., regionally or by hazard).	SPC moderate risk across OK and southeastern TX panhandle. Warm front moving northward across the region. Capping inversion in warm sector, but decreasing over period with surface heating. Discrete supercell storms with extremely large hail likely and isolated tornadoes possible.	Week 2 (Wed p.m.)
5 May 2022 2008 - 2200 UTC **Live**	Group 1: Arkansas and Memphis, TN. Group 2: Shreveport, LA and Houston, TX. Two teams of three forecasters (PHI with warnings and one forecaster focused on communication)	SPC enhanced risk as mid-level trough associated with convection on previous day moves across the area. Convection already ongoing at the start of the case. Increasing coverage and upscale growth expected with potential for QLCS and embedded tornadoes.	Week 2 (Thurs p.m.)
10 May 2022 2005-2210 UTC **Live**	Group 1: Wisconsin. Group 2: Lubbock, TX, Midland, TX and Mexico. Two teams of three forecasters. All forecaster covered a separate storms (and all individual hazards for that storm)	Enhanced potential for development of storms along the dryline in west TX (wind/hail primary concern). Slight risk in WI with stationary front and surface low pressure. Scattered storms likely with potential for isolated supercell storms.	Week 3 (Tues p.m.)

11 May 2022 2010-2204 UTC **Live**	Group 1. Minneapolis, MN and Des Moines, IA. Group 2: Sioux City, IA. Two teams of three forecasters. The northern team sectorized by hazard and the southern team sectorized by area.	Enhanced risk area over the upper midwest as warm front moves northward across the region. Potential for supercell storms (including tornadoes) with large lapse rates and abundant low-level moisture.	Week 3 (Wed p.m.)
12 May 2022 1935 - 2215 UTC **Live**	Group 1: Eastern South Dakota and southern Minnesota.. Group 2: Central and Eastern Nebraska. Two teams of three forecasters. Northern team sectorized by area (warnings with PHI together for all hazards) and southern team sectorized by activity (warnings with PHI, mesoanalysis, and communications)	Moderate risk of severe weather across SD and western MN with enhanced risk covering eastern NE. Discrete storms including long track tornado potential, particularly in SD/MN region. Possible evolution into linear structure (derecho) later in the period. Widespread reports of hail/wind and tornadoes expected.	Week 3 (Wed p.m.)

Table 1: *Summary of cases used during 2022 experiment.*

2022 Summary Statistics

Forecasters were responsible for more individual storms when covering a single hazard (Table 1). Forecasters working this single hazard maintained more objects on average, typically with a lower number of updates per object times and more time between updates. In particular, lightning received significantly more attention when a forecaster was responsible for that as a single hazard. When maintaining multiple hazards forecasters typically triaged their responsibilities focusing first on the tornado threat, followed by severe, then updating lightning if time allowed. Most of the time, forecasters in this situation were comfortable allowing the automation to maintain the lightning coverage.

On average, forecasters had a lower initial probability to create tornado PHI objects compared to both severe and lightning. While this specific probability varied by case, median initial tornado probabilities ranged from 30-60% whereas initial median severe probabilities were 80% and above (except for the Iowa archive case [DMX] which was lower due to lower likelihood of severe hail and wind). Median initial lightning probabilities were >90% across all the cases evaluated in the experiment. Forecasters typically increased probabilities for both severe and tornado PHI over the initial probabilities derived from the machine learning algorithms (Fig. 1).

Hazard type	Severe	Tornado	Lightning
No. of objects	8.67 (2.94)	7.75 (4.06)	10.33 (2.62)
No. of updates	14.67 (7.35)	15.75 (10.0)	22.67 (4.69)
Updates per object	1.69 (2.5)	2.03 (2.47)	2.19 (1.79)
Avg time per update (s)	196.64 (105.29)	139.22 (113.39)	105.96 (94.57)
Freq of update (min)	28.34 (18.09)	18.15 (10.81)	16.63 (19.12)

Table 2. *Task analysis when working single (multiple) hazards for cases during the experiment. "No. of objects" provides the average number of unique objects the forecasters controlled during a single case. "No. of updates" is the average number of updates issued by the forecaster for both manually created or automatically created objects. "Updates per object" is the average number of updates for each unique object. "Avg time per update" is the average duration (in seconds) it took a forecaster between clicking to update an object and clicking the "issue" button. "Freq of update" is the average time between updates for each unique object in minutes.*

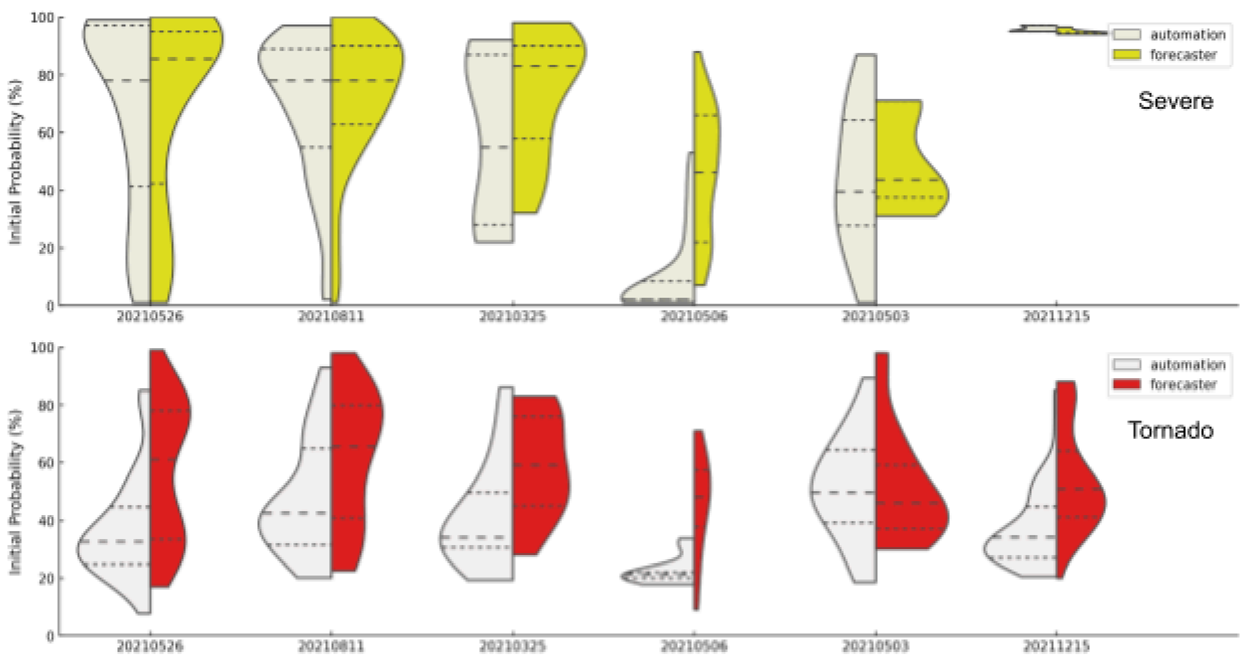


Figure 1. *Initial probability of objects issued by forecasters (both automated and manual) for each of the archive cases. Forecaster probabilities were generally greater than automated probabilities for both severe (top) and tornado PHI (bottom).*

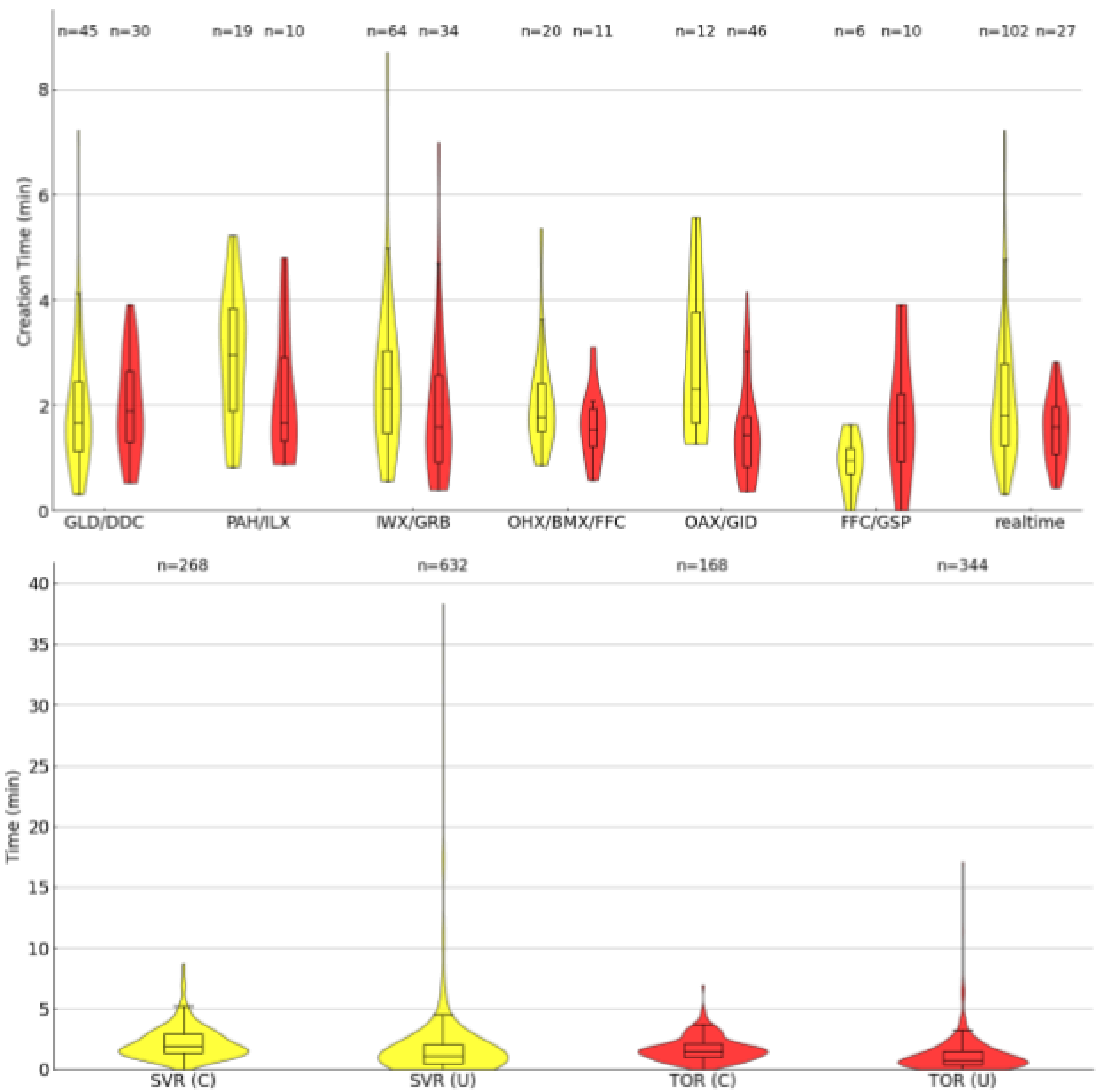


Figure 2. *Top: Amount of time taken to create the severe (yellow) and tornado (red) PHI for each of the cases during the experiment. Bottom: Time taken to create (C) vs update (U) a PHI object or all cases during the experiment. Counts for each are labeled at above the violin plot (n = count).*

2023 Experiment and Results

Two major human factors updates were the primary focus of the 2023 spring evaluation: 1) increased stability for storm tracking and motion and 2) forecaster storm notifications through banners and storm-based icons.

Stability features were added to help reduce issues with motion direction and speed noted by forecasters in previous experiments. These errors resulted in plumes and associated warnings that could expand and contract in length or add and remove regions from the edges of the plumes over short periods of time. First, a confidence interval was created using the original tracked storm vertices. A moving average was then used to smooth the original storm objects, resulting in objects that were elliptical and had fewer drastic changes over a given time step (Fig 3a). Next, a Kalman filter was applied to improve the steadiness of the motion over time which allowed the storm motion to adapt to persistent changes while ignoring transient changes and mitigated these “accordion” and “windshield-wiper” effects (Fig 3b).

Notifications were tested to address previous forecaster comments of difficulty managing multiple storms and hazards simultaneously. Based on this feedback, notifications were created to call attention to: 1) storms with a rapid increase in the automated probability, 2) when the forecaster created probabilities had diverged from automated probabilities, 3) when automated tracking stopped following a storm, and 4) when a forecaster had not interacted with a storm after a given period of time (default of 20 min). The goal of these notifications was to get the forecasters’ attention without disturbing the current activity. Both banners in the top right corner that lasted for 10 seconds and icons that stayed with the storm object were tested (Fig. 4). Forecasters could choose banners, icons, or both as well as modify the default

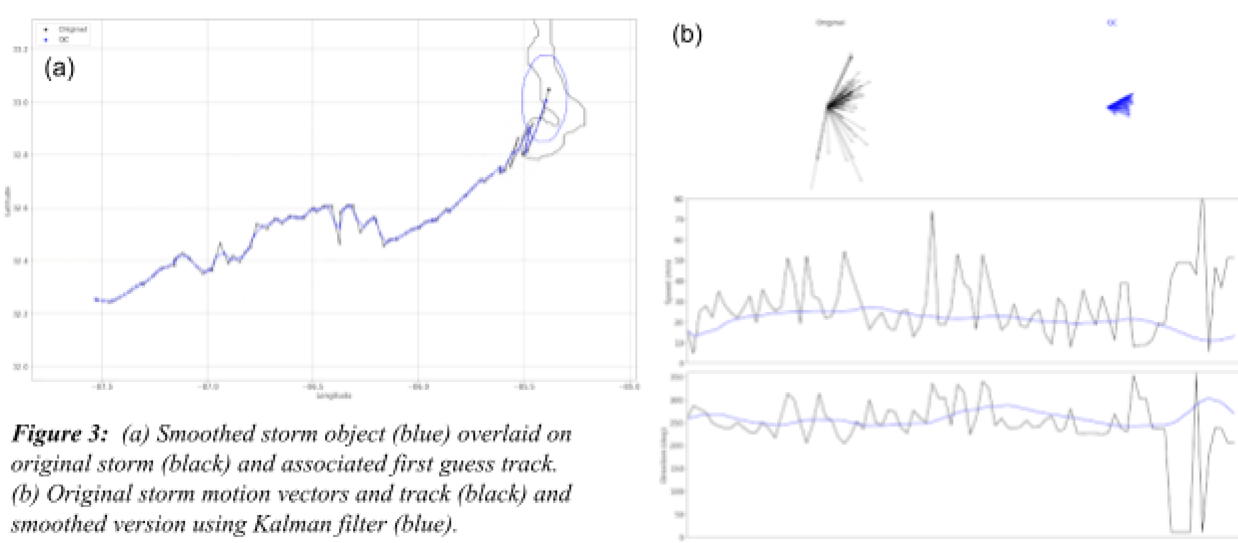


Figure 3: *(a) Smoothed storm object (blue) overlaid on original storm (black) and associated first guess track. (b) Original storm motion vectors and track (black) and smoothed version using Kalman filter (blue).*

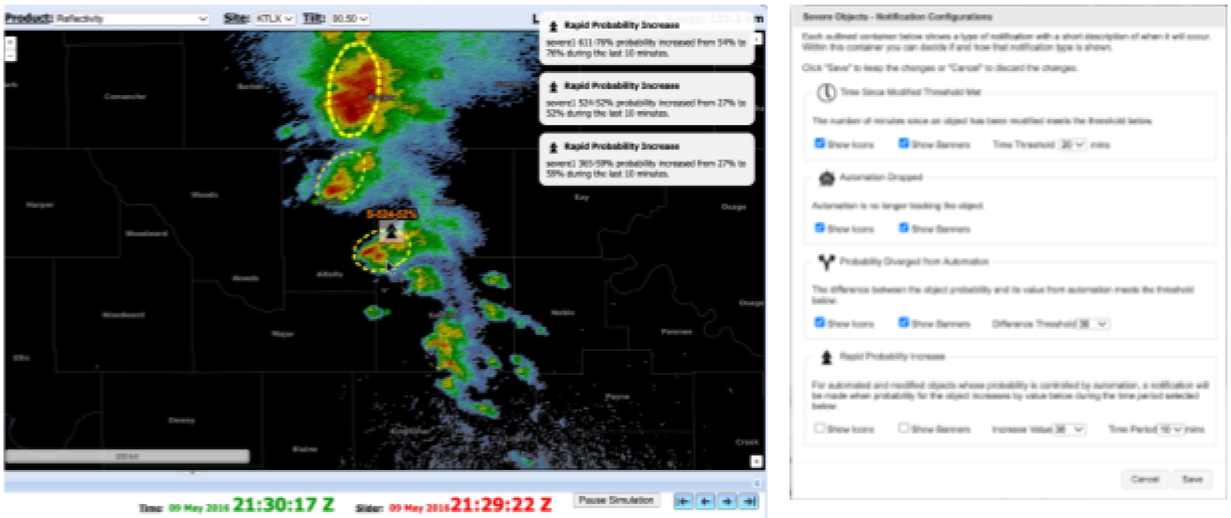


Figure 4: *Example of notifications in the PHI tool. Left: A rapid probability increase is shown shown for multiple storms as banners at the top right corner; clicking a banner will take the forecaster to that particular storm by centering the storm on the screen (icon also shown on storm object). Right: Customization options for the different types of notifications.*

thresholds for time and probability. Clicking on the icon allowed the forecaster to then take an action for that notification which included more details about the notification and the choice to acknowledge with no changes or immediately update the storm configuration and/or warning.

2023 Results

Forecaster feedback across all three weeks was overwhelmingly positive. Previous participants consistently commented how much easier it was to work with the automation due to the updates on storm tracking and motion stability. Overwhelmingly (for >80% of the forecasts, Fig. 5), forecasters used a combination of the automated objects as a first guess and slightly modified individual parts such as the base probability, storm coverage, or storm motion. This trust and use of the automation is beneficial in two ways: (1) the process of using automation as a first guess greatly sped up the creation time of PHI plumes and warnings compared to when forecasters had to draw new objects without automated guidance and (2) the automated probabilities provided calibration across different forecasters working the same archive case.

The NASA Task Load Index (TLX) has consistently been used in HWT PHI experiments to provide assessment of the self-perceived forecaster workload. These TLX scores for returning forecasters show a notable decrease in workload for year 2 (Fig.6). Discussions with forecasters, both during events and in post-case evaluations, indicate this was likely a result of multiple factors. In addition to the increased stability of the automation during 2023, returning forecasters had additional familiarity with the tool and the underlying functionality. One of the returning forecasters noted during the discussion: *"I have a different perspective having done this last year. It's a night-and-day difference compared to last year. I*

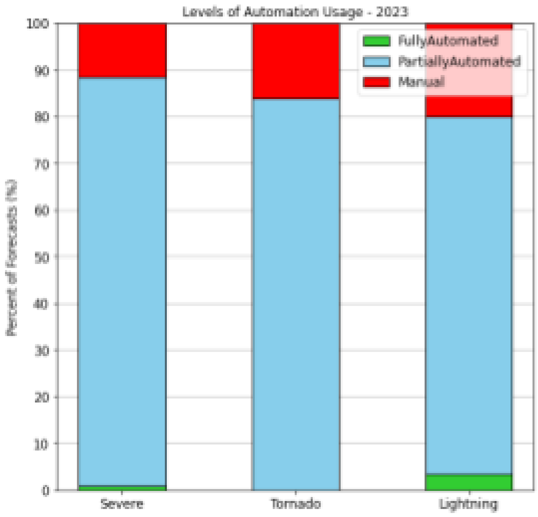


Figure 5: *Use of automation by forecasters during the 2023 experiment separated by hazard (severe, tornado, and lightning). Forecasters most frequently used a combination of automation with some modification by forecasters (blue) and rarely used the automation completely on its own (green) or fully manually drew their own objects (red).*

was mentally and physically just as worn out because of managing and maintaining objects and plumes, but the stability in QC stuff made a huge difference in workload. I spent much more time doing data analysis and looking at the actual radar data versus trying to manage everything."

Correspondingly, the TLX scores for forecasters new to the PHI system in 2023 were similar to those of the forecasters first participating in 2022 (Fig. 7). This likely points to the familiarity with the tool as the primary driver of the lower scores. Similarly, it is possible that there is a self-selection bias, as forecasters who chose to participate in the experiment again were likely to be those that enjoyed their prior experience. As seen in past experiments, the mental aspect (how much mental and perceptual activity was required), overall effort (how hard was the work mentally/physically to accomplish expected level of performance), and performance (how successful/satisfied someone is with their performance) factors carried the highest workload weights.

Overall, forecasters found the notifications to be a positive addition to the system in 2023. In particular, they liked how they were unobtrusive and didn't distract from the current tasks. Specifically, forecasters noted in discussions that the icons were "super helpful" and they allowed them to "refocus on things that I may not be seeing". However, some forecasters commented that they were at times not obtrusive enough to grab their attention. Comments from the discussion clarified that much of the time they did not even notice the banners in the background as they were gone before they could take a moment away from their current task. Due to this, a majority preferred and used the icons as they were co-located with the storm on the screen and did not disappear unless they took an action of either modifying the object or acknowledging the notification. However, the end-of-week survey results highlight a desire to keep both the icons and the banners, particularly for rapid probability increases and when the automated probability differs significantly from the probability set by the forecaster. In addition to the banners lasting longer, forecasters also suggested that a table of notifications or an addition to the hazard services console could be useful. Additionally, many forecasters requested the ability to add custom notifications (e.g., for a

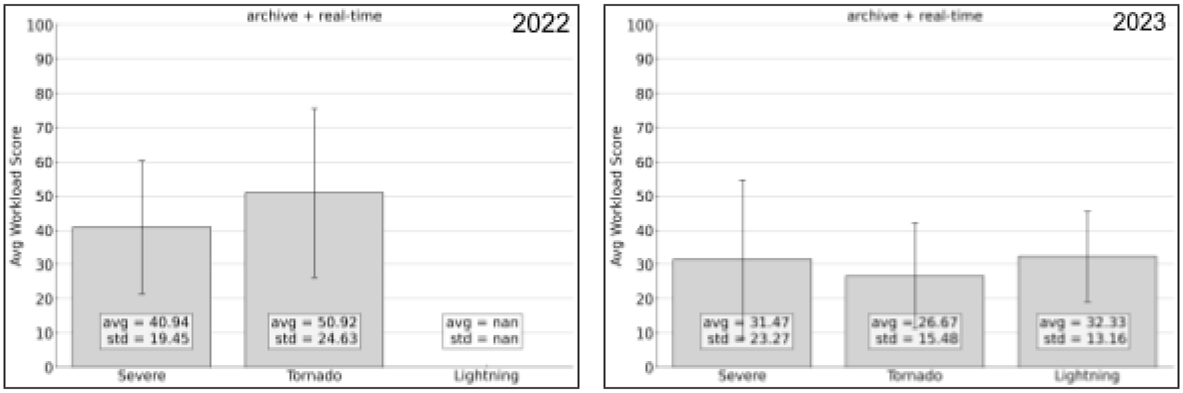


***Figure 6:** Returning forecaster NASA TLX scores separated by hazard (severe, tornado, lightning) for both archive and real-time (live) events during the 2022 (left) and 2023 (right) experiments.*

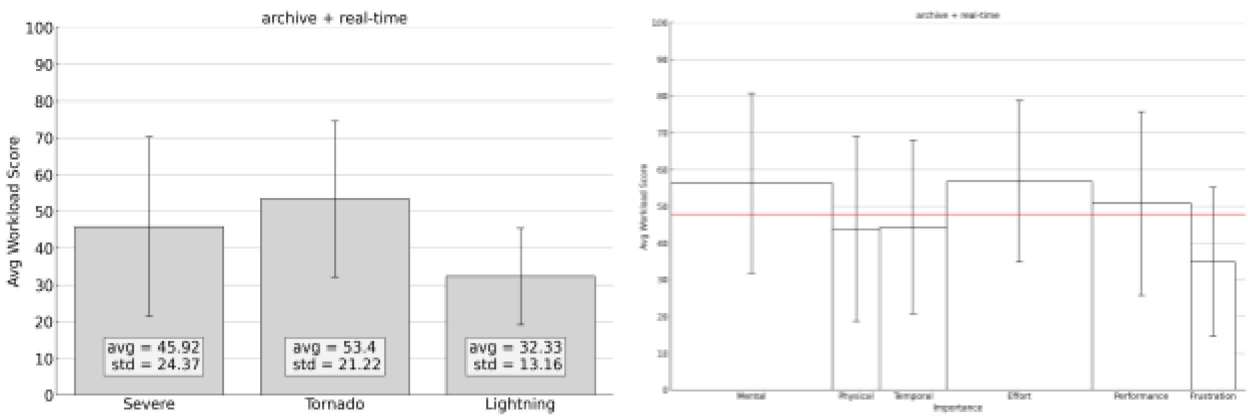


***Figure 7:** NASA TLX scores for all forecasters in 2023 separated by hazard (severe, tornado, lightning) for both archive and real-time (live) events (left). NASA TLX subscales weighted by self-judged importance of each factor shown as the width of the each of the bar graphs (right).*

rapid increase in MESH values associated with a storm) and strongly desired notifications regarding warning expirations within 5 or 10 min.

In addition to creating PHI, forecasters were asked to add warnings to storms when appropriate. In the prototype PHI tool, warnings follow the edge of the generated PHI plume out to a time set by the forecaster (e.g., 45-min warning, 60-min PHI plume). By default, the length of the two will match. Forecasters had the choice of issuing a static warning (similar to the system available today), partial TIM (the area behind the storm is cleared automatically as the storm moves), and full TIM (the warning area is both advanced forward and cleared as the storm moves). Typically, forecasters would use static or partial TIM earlier in the week and then, with additional experience, would switch to using solely partial and full TIM. Once a storm was expected to last more than an hour, forecasters would use full TIM for warnings. In post-event and end-of-week discussion, some forecasters even wondered "*is the Static button necessary if we have Partial and Full available?*".

Initial guidance for forecasters was that PHI was designed to come out prior to warnings. However, no guidance was given as to specific probability thresholds that should be used for each. The expectation was that forecasters would continue to make warning decisions at the same relative thresholds they do today. In general, forecasters considered initially issuing PHI for probabilities between 20-50% for severe storms and 10-40% for tornadoes (Fig. 8). As expected, warnings had a higher probability threshold than PHI alone, though there was some overlap. As with PHI, forecasters thought that tornado warnings should be issued at a lower probability threshold than severe with 25-60% probability used for tornado warnings and 40-75% probability for severe thunderstorms (Fig. 8).

Future Plans

The PHI Prototype experiment in 2024 will be combined with the Watch-to-Warning experiment, which combines PHI storm-based probabilities with Warn-on-Forecast probabilities. The goals of this experiment are to better understand how the new products, algorithms, and tools could enhance communication and product creation between when a watch is issued and the warnings are created. Feedback from the 2023 experiment will be used to create and test new notifications as well as provide guidance for the forecasters on timing and probabilities of PHI creation prior to warning decisions. Additionally, following the success of the stability and quality control additions to the system, a fully-automated version of the PHI plumes are under development for initial testing with Science Operations Officers (SOOs) and local NWS Weather Forecast Offices as part of a wider research-to-operations initiative. This additional test will allow us to understand the impact of automated PHI on a wider range of events and locations than an individual testbed experiment alone.

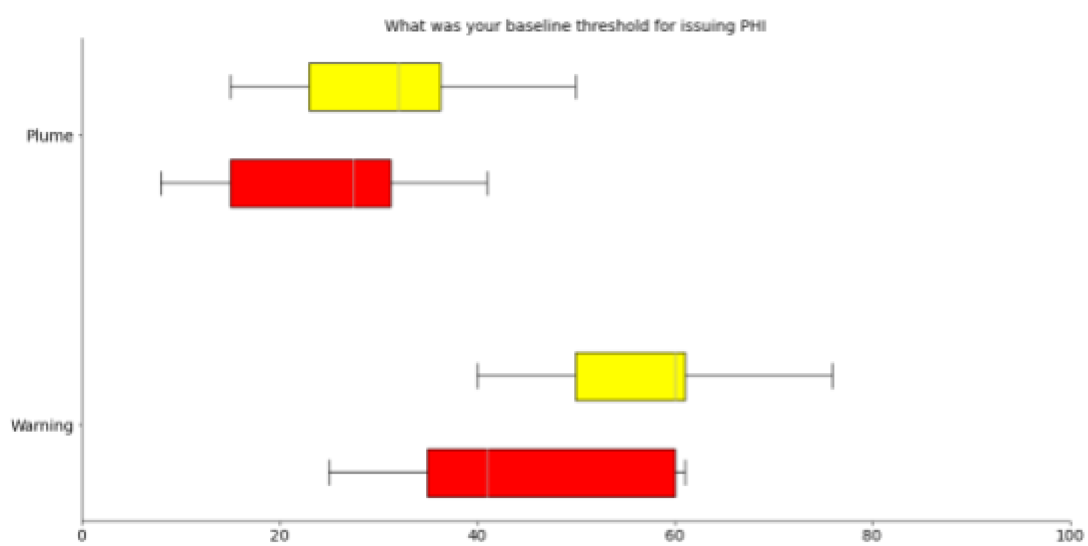


Figure 8: *Forecaster consensus on initial probability for issuing PHI plumes (top) and warnings (bottom) for both severe (yellow) and tornadic (red) hazards.*

Project Contacts

Kristin Calhoun	NOAA/NSSL	Experiment Lead, Project PI
Adrian Campbell	CIWRO and NOAA/NSSL	Systems and Operations Design
Rebecca Steeves	CIWRO and NOAA/NSSL	PHI Tool Design and Support
Thea Sandmael	CIWRO and NOAA/NSSL	PHItor PI and Experiment Scientist
Clarice Satrio	CIWRO and NOAA/NSSL	Experiment Scientist
Patrick Hyland	CIWRO and NOAA/NSSL	Experiment Scientist
Justin Monroe	CIWRO and NOAA/NSSL	AWIPS Support
Jonny Madden	CIWRO and NOAA/NSSL	AWIPS Support

References:

Calhoun, K. M. , and Coauthors, 2018: Cloud-to-ground lightning probabilities and warnings within an integrated warning team. AMS Annual Meeting. *Special Symp. on Impact-Based Decision Support Service*s, Austin, TX, Amer. Meteor. Soc., 4.4, https://ams.confex.com/ams/98Annual/meetingapp.cgi/Paper/329888

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, L. Cronce, and J. Brunner, 2020: NOAA ProbSevere v2.0—ProbHail, ProbWind, and ProbTor. *Wea. and Forecasting*, 35, 1523-1543

Karstens, C. D., J. Correia Jr., D. S. LaDue, J. Wolfe, T. C. Meyer, D. R. Harrison, J. L. Cintineo, K. M. Calhoun, T. M. Smith, A. E. Gerard, and L. P. Rothfusz, 2018: Development of a Human–Machine Mix for Forecasting Severe Convective Events, *Wea. and Forecasting*, 33(3), 715-737.

Sandmael, T., R. B. Steeves, P. A. Campbell, K. M. Calhoun, Z. A. Cooper, 2021: The Probabilistic Hazard Information Tornado (PHItor) Product: Using Machine Learning to Provide Probability Guidance for Tornado Nowcasting. *11th Conference on Transition of Research to Operations. AMS Annual Meeting*. Virtual. https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/380062

Satellite Convective Applications Experiment

Summary by Kevin Thiel

Overview

Satellite Proving Ground demonstrations in the HWT have provided users with a glimpse into the capabilities, products and algorithms that are and will be available with new updates, technology, and products available on both geostationary and polar-orbiting satellites. The education and training received by participants in the HWT fosters interest and excitement for new satellite data and helps to promote readiness for the use of satellite data and products. The HWT provides a unique opportunity to enhance research-to-operations and operations-to-research (R2O2R) by enabling product developers to interact directly with operational forecasters, and to observe the satellite-based algorithms being used alongside standard observational and forecast products in a simulated operational forecast and warning environment. This interaction helps the developer to understand how forecasters use the product and what improvements might increase the product utility in NWS operations. Feedback received from participants in the HWT has proven invaluable to the continued development and refinement of GOES-R and JPSS algorithms. Furthermore, the EWP facilitates the testing of satellite-based products in the AWIPS-II data processing and visualization system currently used at NWS Weather Forecast Offices (WFOs).

Experiment Details and Results

Due to the ongoing COVID-19 Pandemic in 2022, all 2022 GOES-R/JPSS Proving Ground activities were conducted in a virtual environment during the weeks of 23 May, 6 June, and 13 June. Building upon feedback from the 2021 and 2022 experiments regarding the virtual experiments and the subsiding COVID-19 pandemic in 2023, the first week of the testbed (22-26 May) was conducted in-person in the HWT's Development Lab. The second and third weeks of the experiment (5-9 June and 12-16 June) were conducted virtually using Google Meet and Slack. Seven to eight NWS forecasters volunteered each week to evaluate the products, providing 21 participants in 2022 and 22 participants in 2023.

Before the testbed user guides, PowerPoint presentations, and online learning modules were shared with all participants for each of the products demonstrated through Google Drive. The Monday of each week began with one hour of introductions, orientations and product summaries from developers, with the second hour devoted to setting up forecasters with their cloud-based AWIPS instances. For these activities all participants were in a single video conference. After a brief forecast discussion, forecasters were placed into operations. All subsequent days began with a discussion of the previous day's operations involving questions from developers and feedback from forecasters, followed by a forecast discussion, operations, and daily surveys. End of week surveys were then sent to participants the Friday of each week.

During the virtual weeks in 2022 and 2023 activities began at 1 pm CDT and ended at 6pm CDT Monday through Thursday, and forecasters spent approximately four hours in operations. During the one in-person week in 2023, activities began on Monday at 11am CDT (16 Z) and ended at 7 pm CDT (0 Z). Tuesday through Thursday also followed this schedule, but these days could have begun any time between 10am CDT (15 Z) and 1 pm CDT (18 Z) based on the convective environment and was decided on the previous day by the Satellite Liaison. No operations occurred Friday for any of the experiment weeks, as forecasters filled out the weekly survey and held their weekly debrief with the developers. Additionally, forecasters participating in the one in-person week were encouraged to participate in the Tales from the Testbed webinar hosted by the NWS Warning Decision Training Division (WDTD). **Figure 1** provides an approximate schedule of operations for both the virtual and in-person (on-site) demonstrations in 2022 and 2023.

Virtual Demonstartion Schedule

Time (CDT)	13:00	13:30	14:00	14:30	15:00	15:30	16:00	16:30	17:00	17:30
Monday	Orientation			Product Training/Operations						Daily Survey
Tuesday	Discussion/ Forecast		Operations							Daily Survey
Wednesday	Discussion/ Forecast		Operations							Daily Survey
Thursday	Discussion/ Forecast		Operations							Daily Survey
Friday	Weekly Survey	Weekly Debrief								

On-Site Demonstration Schedule

Hour (From start)	0:00	0:30	1:00	1:30	2:00	2:30	3:00	3:30	4:00	4:30	5:00	5:30	6:00	6:30	7:00	7:30
Monday	Orientation				Product Training				Operations							Daily Survey
Tuesday	Discussion/ Forecast		Operations													Daily Survey
Wednesday	Discussion/ Forecast		Operations													Daily Survey
Thursday	Discussion/ Forecast		Operations											Webinar Prep		Daily Survey
Friday	Weekly Survey	Weekly Debrief		Webinar												

Figure 1: An approximate weekly schedule of the virtual (above) and in-person (below) GOES-R/JPSS Proving Ground HWT Experiment, outlining the major activities from each day.

Typical feedback included suggestions for improving the algorithms, ideas for making the displays more effective for information transfer to forecasters, best practices for product use, suggestions for training, and situations in which the tools worked well and not so well. Most of the products evaluated in 2022 and 2023 were advancements of previous product iterations from the previous experiments. The products demonstrated each year of the experiment can be found in **Table 1**. All products except for Optical Flow Winds were available in the AWIPS-II CAVE, and all products were available online.

Year	Products Demonstrated
2022	● Geostationary Lightning Mapper (GLM) ● NUCAPS Temperature and Moisture Profiles ● Optical Flow Winds ● PHSnMWnABI Model ● Probability of Severe (ProbSevere) LightningCast Model ● Probability of Severe (ProbSevere) Model - Version 3
2023	● NUCAPS Temperature and Moisture Profiles ● OCTANE Speed and Direction Sandwiches ● PHS Model ● Probability of Severe (ProbSevere) LightningCast Model ● Probability of Severe (ProbSevere) Model - Version 3

Table 1: A list of the products demonstrated in 2022 and 2023 Satellite Proving Ground HWT experiments.

Within operations forecasters had several tasks, such as building procedures to integrate experimental products with the ones they currently use, issuing warnings and advisories, having discussions with the subject matter experts, and writing blog posts. New in 2022, and continuing in 2023, all simulated forecast groups were also provided a decision support service (DSS) event to communicate with. This was created to adapt to the growing mission of the NWS in 2021. Discussions between forecasters and developers often involved questions from both groups concerning best display practices and applications, along with feedback from forecasters of what they were observing in real-time. Forecasters also had the opportunity to create blog posts by filling out a template through Google Drive. The PI would then use the templates to create blog posts for the HWT EWP Blog (https://inside.nssl.noaa.gov/ewp/), publishing them the next day. Feedback was summarized from the 2022 and 2023 GOES-R/JPSS Proving Ground experiments through the end of day surveys, end of week surveys, blog posts, public graphics, and daily debrief discussions, and placed into Final Reports for 2022 and 2023 respectively. Recommendations were provided for each product, and the experiment itself, with the categories of ‘recommended’, ‘strongly recommended’ and ‘highly recommended’ in an ascending order of significance from the forecasters. Below or the summarized recommendations:

2022 Recommendations

GLM

- It is strongly recommended that the gridded GLM Minimum Flash Area product be added as a baseline (Level 2) product, and that the Total Optical Energy product not be added at this time.

- It is strongly recommended that the NWS continue to emphasize GLM training related to applications and limitations of the GLM gridded products for warning operations and decision support services.
- It is recommended that the default maximum value displayed for the Flash Extent Density product be reduced from 260 flashes per 5 minutes.

<u>NUCAPS</u>

- It is highly recommended that, when available, NUCAPS data from additional satellites in low-Earth orbit be integrated into AWIPS to improve its spatiotemporal coverage, and increase the ability to assess temporal changes in thermodynamic profiles.
- It is strongly recommended that improvements continue to be made in the display of NUCAPS data.
- It is recommended that comparisons between NUCAPS and other sources of thermodynamic data continue to be emphasized in training to build forecaster confidence.
- It is recommended that the integration of wind information with NUCAPS profiles and gridded products be considered as a product enhancement.

<u>Optical Flow Winds</u>

- It is highly recommended that display of the Optical Flow Winds product be improved, such that forecasters can more readily identify features of importance within the wind analysis.
- It is strongly recommended that NWS forecasters be trained on the use of satellite derived wind field in short-term convective forecasting, such that the physical basis of the Optical Flow Winds product be well established.

<u>PHSnMWnABI Model</u>

- It is strongly recommended that training efforts and future model development continue to emphasize the utility of the PHSnMWnABI model to analyze near term (less than four hours), rapidly updating mesoscale environments.
- It is strongly recommended that the web display of PHSnMWnABI model output improve to aid in future integration and testing for NWS forecasters.
- It is recommended that future demonstrations of the PHSnMWnABI model to NWS forecasters include additional fields in AWIPS-II, and direct comparisons to high-resolution NWP currently in operations.

<u>ProbSevere LightningCast</u>

- It is highly recommended that the ProbSevere LightningCast model be integrated into NWS operations, for the purposes of decision support service efforts, along with situational awareness for initiating convection or thunderstorm maintenance.
- It is strongly recommended that NWS forecaster training efforts focus on the applications of the ProbSevere LightningCast model for decision support services and using trends in probability outputs to provide actionable lead times of lightning.
- It is recommended that future development efforts concerning the ProbSevere LightningCast model focus on the potential integration of radar data into the model, along with a point-based time series tool to effectively display probability trends.

<u>ProbSevere v3</u>

- It is highly recommended that the ProbSevere v3 model be integrated into NWS operations and replace ProbSevere v2.
- It is strongly recommended that forecaster training of the ProbSevere v3 model incorporate recalibration of forecasters from the ProbSevere v2 model, best display practices, and how to leverage the time series tool.
- It is recommended that future development efforts of the ProbSevere v3 model include improvements to the storm object identification algorithm, more continuous integration of modeled environmental field, and increased flexibility of the ProbSevere v3 display in AWIPS-II if possible.

Experiment

- It is strongly recommended that future HWT Satellite Proving Ground Experiments consider in-person and virtual setups for participants, developers, and visiting scientists.
- It is strongly recommended that simulated DSS events continue as a portion of the testbed.

2023 Recommendations

NUCAPS

- It is strongly recommended that Gridded NUCAPS and NUCAPS-Forecast datasets be evaluated against commonly used datasets and fields by NWS forecasters to increase confidence in the products.
- It is strongly recommended that exploration of a NUCAPS-Forecast product with half-hourly output out to six hours be evaluated in future testbeds.
- It is recommended that data quality parameters for Gridded NUCAPS and NUCAPS-Forecast continue to be explored to increase forecaster confidence in the available thermodynamic information from NUCAPS.
- It is recommended that the integration of wind information into NUCAPS-Forecast fields be explored to increase the utility of these data when interpreting storm mode. Potential avenues of development may include the use of wind data from models (e.g. RAP, HRRR, GFS, etc.) or adjacent RAOBs to create additional convective indices.

OCTANE

- It is highly recommended that forecasters leverage OCTANE Motions and Speed Sandwich to assess cloud top divergence trends from convection, convective initiation, and perform mesoanalysis of the ambient kinematic environment.
- It is strongly recommended that OCTANE training materials include, along with a description of the product and its applications, the physical basis for using cloud top winds with high spatiotemporal resolution in context with radar, lightning, and satellite imagery datasets.
- It is recommended that a cloud-top divergence field be explored to synthesize the OCTANE speed and direction products.

PHS Model

- It is strongly recommended that product fields be developed which show how the data assimilation technique employed by the PHS model adds value when compared to a control or similar model run.
- It is strongly recommended that improvements to the PHS model be made such that data are more readily available in AWIPS-II with decreased latency, in addition to an improved web display for this experimental model.
- It is recommended that PHS model data be used to perform mesoanalysis regarding convection initiation timing and location, storm mode, and potential convective hazards based upon the environmental and reflectivity fields provided.

ProbSevere LightningCast

- It is highly recommended that the ProbSevere LightningCast model be integrated into NWS operations for the purposes of situational awareness, anticipating the initiation of convection and associated lightning activity, and monitoring the advection of lightning activity from mature convection.
- It is recommended that ProbSevere LightningCast training incorporate the previously mentioned applications, along with model limitations such as lower than expected probabilities in dense cirrus and stratiform precipitation.
- It is recommended that the ProbSevere LightningCast meteogram tool use parallax-corrected model output to improve the interpretation of these probabilities and align with corresponding lightning information.

ProbSevere v3

- It is highly recommended that ProbSevere v3 be implemented within NWS operations and replace ProbSevere v2.
- It is highly recommended that forecasters leverage the ProbSevere v3 time series tool within convective warning operations to efficiently display trends in thunderstorm intensity and identify convective hazards.
- It is recommended that training regarding the ProbSevere v3 model include the calibration of probabilities from the ProbSevere v2 model, the products input into ProbSevere v3, the definition of the individual probabilities, and examples depicting the expected model performance.

Experiment

- It is strongly recommended that both in-person and virtual weeks of the HWT Satellite Proving Ground experiment continue.
- It is recommended that, prior to the testbed, online versions of products be made available to the forecasters alongside their training materials when possible, along with sharing the science questions from each developer group.
- It is recommended that flexible start times and longer operational periods for the virtual weeks, along with archived cases for both formats, be considered to improve the quality of the evaluation.

Results from the 2021 experiment signaled the ability to evaluate products with respect to their DSS capabilities. Therefore, nearly all forecaster groups were provided with simulated DSS events within operations, as the testbed grew with the expanded NWS mission. These events were successfully used to demonstrate decision support capabilities in the 2022 and 2023 experiments, and provide numerous evaluation benefits for products which may fall outside the operational warning paradigm. Participants sent messages using a Google Form both years, and also described what products and signals influenced their messaging of hazardous weather information. Additionally, forecasters could create 'public graphics' that they would theoretically issue to the public or their partners. Including both written and graphical means to communicate experimental hazardous weather information provides a new dimension to the holistic evaluation process used within the experiment.

To facilitate a more efficient experiment, and act on feedback from the 2021 experiment, pre-made AWIPS procedures were created in collaboration with each product developer group and considered their R2O objectives. These were well received by the forecasters, especially amongst the virtual participants, as it allowed them to quickly become familiar with the products and leverage the best display practices from their training materials. Similarly, pre-built displays helped forecasters more directly address several key research questions whether it was validation with similar products/features or comparisons between different display techniques. Forecasters also modified these procedures to create their own display, and provided several fruitful conversations with developers about the optimal ways to display and interpret their products.

The reemerging capacity of in-person experiments between 2022 and 2023 provided the opportunity to directly compare between the in-person and virtual formats, and to assess their benefits. Similar to the 2021 experiment, forecasters and developers in the virtual components of the 2022 and 2023 experiments largely felt the setting provided a valuable opportunity to demonstrate new and developing products. The developers who participated in both in-person and virtual formats in 2023 stated the additional time in operations and face-to-face interactions with the forecasters provided a more thorough evaluation of their product, and allowed the developers to build a rapport with the participants throughout the week. When asked, developers expressed the desire to increase the number of in-person weeks, even if this means fewer forecasters each week. Additionally in 2022 and 2023, a growing number of virtual participants stated that they were unable to participate in-person due to office staffing limitations or personal commitments that made extended travel from their home area difficult. Under these scenarios, a testbed with more remote elements may allow for a more diverse pool of applicants across the NWS or other sectors than previously observed. Lastly, in 2023 both in-person and virtual formats successfully used cloud-based instances of AWIPS-II, and the technical support provided by Justin Monroe and Jonathan Madden made the process nearly identical between both formats.

Future Plans

Based upon the successes and lessons learned from the 2022 and 2023 Satellite Convective Applications Experiment, the 2024 experiment looks to further refine the optimal balance between its in-person and virtual formats. Three full weeks of evaluations are planned during May and June of 2024, with two weeks leveraging the in-person format and one week for the virtual format. The two in-person weeks will feature fewer forecasters (approximately four per week) compared to the virtual week (approximately 8).

Web Presence

- GOES-R/JPSS Satellite Proving Ground Final Report (2022/2023)
- GOES-R Proving Ground Overview (webpage)
- EWP Blog Posts (2022/2023)
- WDTD Tales from the Testbed (27 May 2023)
- NWS Satellite Book Club Webinar (22 September 2022/11 January 2024)

Project Contacts

Kevin Thiel	CIWRO and NOAA/SPC	Project PI
Justin Monroe	CIWRO and NOAA/NSSL	EWP Technical Lead
Jonny Madden	CIWRO and NOAA/NSSL	EWP Technical Support

Hazard Services – Threats-In-Motion HWT Experiment

2022 Executive Summary

Threats-In-Motion (TIM) is a proposed warning decision and dissemination approach that would enable the National Weather Service (NWS) to upgrade severe thunderstorm and tornado warnings from the current static polygon system to continuously-updating warning polygons that move with the storm. TIM can improve average warning lead time as well as provide more equitable (uniform) lead times for those in the path of long-tracked severe storms. The software capability for this solution was developed within AWIPS Hazard Services (HS) and is known as HS-TIM. The TIM software and warning concepts were tested within the NOAA Hazardous Weather Testbed for three weeks in the summer of 2022 with participating NWS forecasters.

After a year since the previous iteration of this HWT experiment, there were many software updates to add new features and improve performance. Some of these modifications resulted in slight disruptions in the workflow and workload for the forecasters; however, forecasters were able to look beyond those issues and provide valuable feedback to further improve the software and add new functionality.

Besides the improvements that TIM provides with lead time, the experiment revealed these noteworthy findings: 1) TIM has the potential to reduce forecaster workload and mental demand by eliminating the need to issue a new warning once a storm has exited a polygon; and, 2) TIM has the potential to improve messaging by keeping the same Event Tracking Number (ETN) for the lifecycle of the hazard and a cleaner output field by eliminating overlapping warnings on individual hazards. In addition, the forecasters unanimously agreed that warning extensions should be allowed for any long-tracked storm, not just those with observed significant hazards.

2022 Detailed Overview

Current severe weather warnings (tornadoes, wind, and hail) require the forecaster to issue multiple sequential warnings for long-tracked storms because the current policy prohibits extending a warning's area and time during updates. This situation frequently results in nonuniform lead times for users on the downstream border of a warning polygon. For example, nearly adjacent locations can have dramatically different lead times if one location is just outside the upstream warning (Fig. 1).

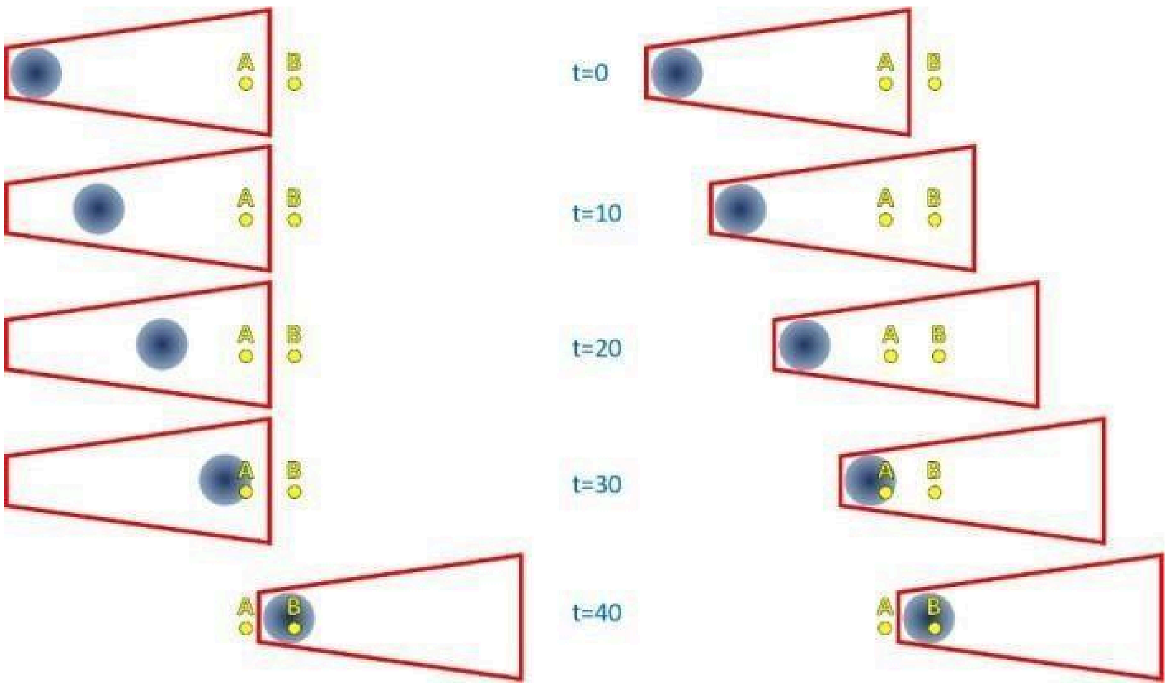


Figure 1. *Comparison of (left) current NWS warning practice using separate polygons and (right) using TIM. The positions of two user locations are shown as A and B. Images are shown at 10-min intervals; the intermediate 1-min TIM polygons are not shown. The blue-gray blob represents a hypothetical storm moving west to east.*

Threats-In-Motion (TIM) is a proposed warning decision and dissemination approach that would enable the NWS to upgrade severe thunderstorm and tornado warnings from the current static polygon system to continuously-updating warning polygons that move with the storm. More specifically, a warning polygon is attached to the threat and moves forward along with it. Warnings are automatically cleared from locations where the threat has passed. TIM is meant to be an evolutionary step of the Forecasting A Continuum of Environmental Threats (FACETs) initiative for the convective weather warning scale.

TIM offers several benefits over the current warning methodology. With TIM, the forecaster would only need to issue a single warning, updated regularly as workload permits, embodying a "one storm-one story" concept. This approach has the potential to reduce forecaster workload because downstream warning issuance would be replaced by a less time-consuming warning update. Additionally, TIM mitigates warning coverage gaps and improves the handling of storm motion changes. It also offers improvements in hazard communication. TIM provides a continuous history for each storm, versus times when two or more separate overlapping warnings may be in effect for the same hazard (e.g., when a new downstream warning is issued before the previous warning can be canceled). These improvements can lead to more simplified and consistent messaging for key partners and improved event verification.

Perhaps the most significant benefit of TIM is that it can provide more equitable (uniform) lead times for those in the path of long-tracked severe storms because these storms remain continually tracked and warned. This change also results in greater average lead times and decreased average time spent in a warning relative to today's warnings, with little impact on average false alarm time. This impact is particularly noteworthy for storms expected to live longer than the average warning duration (30–45 minutes), such as the long-tracked supercells seen during violent severe weather outbreaks. A robust statistical analysis of TIM's scientific benefits is available in Stumpf and Gerard (2021).

This promising, innovative approach is under consideration for transition to NWS operations. If TIM is approved via the NWS governance process, the transition to operations is expected to occur in these two phases:

1. Tiny TIM: adds the capability to extend warnings in time and area, so that long-tracked storms do not need to be re-warned along their path, resulting in fewer warning polygons (one hazard has one event tracking number [ETN]). Summaries of recent Tiny TIM experiments are available upon request.

2. Taller TIM: includes Tiny TIM capability, but also fills in the periods between warning updates by automatically-translating warnings at one-minute intervals. Taller TIM would provide an even greater improvement in average lead time and lead time equitability. This is the longer-term "full implementation" of TIM, and the subject of this experiment.

Efforts have been underway since 2019 to develop the software capability for Taller TIM for convective weather warnings (tornadoes, wind, and hail) within the Advanced Weather Interactive Processing System (AWIPS) Hazard Services (Fig. 2). This software is known as Hazard Services – Threats-In-Motion, or HS-TIM. This software has been tested within the NOAA Hazardous Weather Testbed (HWT), and this summary details the most-recent experiment, held in the summer of 2022.

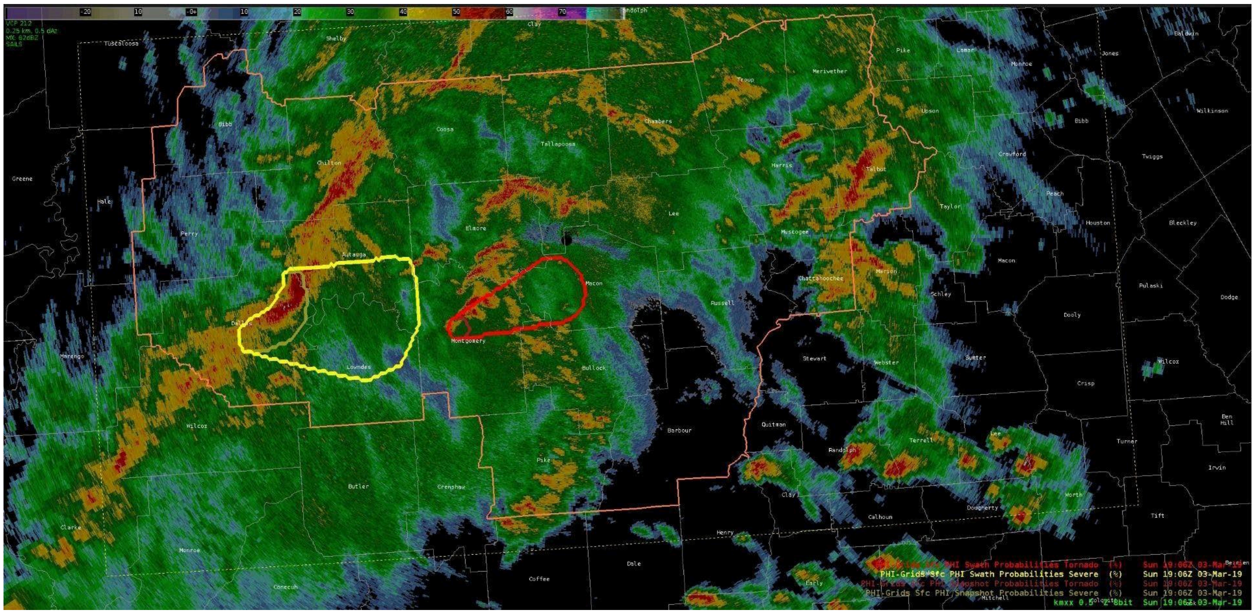


Figure 2. *HS-TIM output from a scenario in southeast Alabama - the yellow (red) polygons show severe thunderstorm (tornado) warnings; the large orange polygon is the county warning area (CWA).*

2022 Software Details

Differences with current warning software

The main differences between WarnGen and HS-TIM are:

- The Hazard Services screen layout (Spatial Display, Console, Hazard Information Dialog). This will be the layout for the upcoming WarnGen replacement (HS-Convective) and is already available in other HS perspectives (e.g., HS-Hydro), so some of the forecasters already had some experience with the new layout.
- The use of 2D storm "objects" (versus points and lines) to define and forecast current threat areas and to create the warning swath. These 2D storm objects are created using any of these drawing tools: polygon, freeform, or ellipse. The attributes of the object – the geometry (shape and location), duration, storm motion, and motion uncertainty – are used to create forecast objects at 1-minute intervals. These 1-min forecast objects are then combined to create the warning polygon swath. Forecasters cannot edit the polygon swaths directly and can only change the swaths by editing the object attributes.
- Defining the motion vector is done using the "drag me to hazard" process. The forecaster steps back a few radar frames in time, and then drags the 2D object over the location of

the threat area in the past. The centroids of the 2D object at the current and past locations are used to calculate the motion vector (speed and direction).

- Along with the motion vector, the forecaster can also define the storm motion uncertainty. A larger value of speed (motion) uncertainty will cause the warning polygon swath to be stretched more in the direction parallel (perpendicular) to the motion vector.
- The "Persist" feature, which, when enabled, places a warning "in motion", and automatically moves the warning swath forward every minute. Both the front and the rear end of the warning remain in motion while Persist is on.
    - Conversely, for any warning where the Persist feature is off, the rear end of the warning automatically clears out at one-minute intervals.
- During subsequent warning updates (to be issued at the same frequency as today), the 2D objects can be quickly repositioned and reshaped, and the warning attributes updated, to continue the warnings indefinitely or until the storm dissipates. This process consumes far less workload than re-issuing brand new warnings each time a warning expires.
- A Warning Decision Discussion (WDD) allows the warning forecaster to add their thoughts about why they issued or modified the warning (e.g., "mid-level rotation is strengthening"). This is information that is typically not included in the actual text of a warning product but may be relayed via NWSchat or similar end-user communication software. Because TIM warnings use the same ID throughout the storm's lifetime, the WDD history is linked from the first time the warning is issued.

Forecaster Workflow Details

The workflow to create a new warning is as follows (* is optional):

1. Draw a 2D object around the current threat area (a.k.a. "geometry").
2. Select the warning type (SVR or TOR).
3. Define duration.
4. Choose whether or not to Persist the warning.
5. Define the motion vector using "drag me to hazard".
6. *Define the motion uncertainty.
7. Select warning attributes (Max Wind/Hail/Tornado, Impact-Based Warning (IBW) tags, Calls-To-Action statements, etc.).
8. *Write a Warning Decision Discussion (WDD)
9. Preview the warning text and edit if needed.
10. Issue the warning.

The workflow to update a warning is as follows (* is optional):

1. *Drag the 2D object over the current threat area if the storm motion has changed
2. *Redraw (edit or replace) the 2D object geometry for the current threat area if the storm shape has changed.
3. *Re-define duration.
4. Choose whether or not to Persist the warning.
5. Re-define the motion vector
6. *Re-define the motion uncertainty.
7. *Change warning attributes (Max Wind/Hail/Tornado, Impact-Based Warning (IBW) tags, Calls-To-Action statements, etc.).
8. *Write a Warning Decision Discussion (WDD)
9. Preview the warning text and edit if needed.
10. Issue the warning update.

As the storm moves with time, so does the 2D object. Unless the storm motion has changed, the 2D object will remain over the storm with time. If the storm motion has changed (e.g., a rightturning supercell), the forecaster simply has to drag the 2D storm object back over the current threat area.

Persisting vs. Non-Persisting Warning Details

To decide whether to Persist or not Persist the warning, these general guidelines are followed:

- For a storm <u>not expected</u> to last longer than the warning duration (short-lived), <u>do not persist</u> the warning.
- For a storm <u>expected</u> to last longer than the warning duration (long-track), <u>persist</u> the warning.
- For a <u>dissipating</u> long-tracked storm (no longer expected to last longer than the warning duration), <u>do not persist</u> the warning.

Persisting warnings:

- The front side of the polygon continuously moves downstream at 1-minute intervals, providing equitable lead time.
- The rear side of the polygon continuously moves downstream at 1-minute intervals, allowing for automated All Clear statements.
- VTEC action code = EXB (EXtend Both area and time)
- The Expiration Time updates every minute.

Non-Persisting warnings:

- The front side of the polygon remains stationary.
- The rear side of the polygon continuously moves downstream at 1-minute intervals, allowing for automated All Clear statements.
- VTEC action code = CON (continuation)
- The Expiration Time does not update every minute.
- When the Expiration Time is reached, the warning will expire (VTEC = EXP)
- The forecaster can also cancel warning (VTEC = CAN) prior to warning expiration

Warning Dissemination Details

NWS convective warnings are issued with an Event Tracking Number (ETN), which is a unique ID that is attached to each warning type (Severe Thunderstorm Warning or Tornado Warning). At present, each subsequent warning issued on a hazard has a different ETN assigned to it. Once the warning has expired or is canceled, the ETN is not reused (NWS, 2020). With TIM, the same ETN is used throughout the lifecycle of the hazard. The updated warning retains the same ETN because the forecaster does not have to issue a new warning when the old warning expires. Examples of a series of ETNs and action codes for today's warnings and TIM warnings are shown in Table 1.

Time (UTC)	NWS VTEC	ETN	TIM VTEC	ETN
1900	NEW	1	NEW	1
1910	CON	1	EXB	1
1920	CON	1	EXB	1
1930	NEW	2	EXB	1
1940	CON	2	EXB	1
1950	CON	2	EXB	1
2000	NEW	3	EXB	1
2010	CON	3	EXB	1
2020	CON	3	EXB	1
2030	NEW	4	EXB	1
2040	CON	4	EXB	1
2050	CON	4	EXB	1
2100	NEW	5	EXB	1
*2110	CON	5	CON	1
*2120	CON	5	CON	1
*2130	EXP	5	EXP	1

Table 1. *Warning decision times, VTEC action codes, and ETNs for a hypothetical storm case. Today's warnings use the blue columns on the left. TIM uses the green columns on the right.*

A graphical representation of the current warning method and TIM is shown in Fig. 3. In this example, warning updates are issued every 15 minutes. For the current warning methodology, the warning polygon can only be made smaller with each update, and the warning expiration time remains constant. After 60 minutes, the warning is replaced with a new warning having a new polygon, ETN, and expiration time. For the TIM methodology, the warning can continue downstream at each update, with new areas added to the warning, some areas removed from the warning, and the expiration time extended. The TIM warning maintains the same ETN throughout its lifecycle.

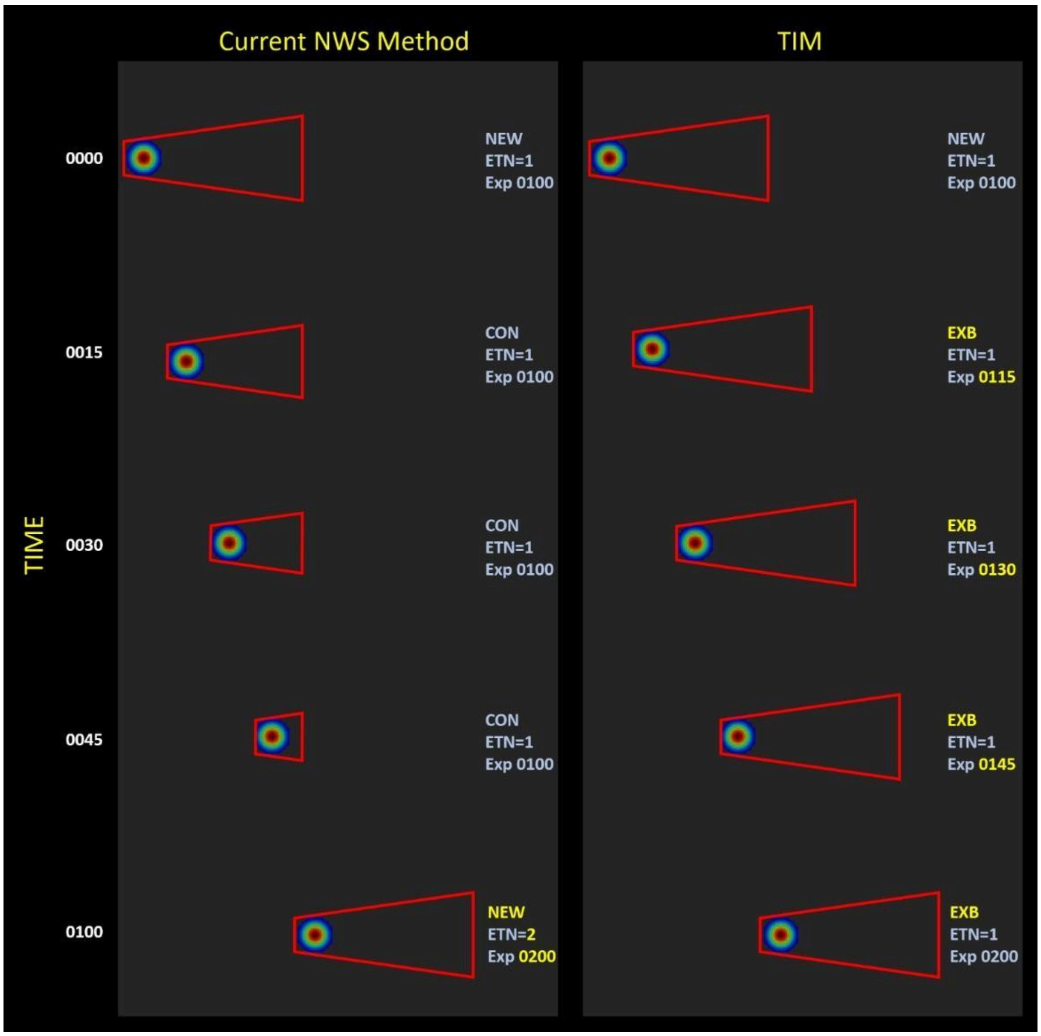


Figure 3. *Comparison of (left) current NWS warning practice using separate polygons, (right) to TIM. Images are shown at 15-min intervals. The VTEC action code, ETN, and warning expiration time are shown for each time interval. The colored circle represents a hypothetical hazard.*

2022 Recent Software Improvements

Since the 2021 TIM HWT experiment, several more improvements have been made to the software. Many of these improvements were suggested by forecasters in that HWT experiment. The items denoted with an asterisk were not completely ready for the experiment. Those with two asterisks were completed after the experiment concluded.

- Developed an EDEX recommender, which meant that the software was no longer tied to having to run a separate central processor instance, streamlining the software and having it run like baseline, operational Hazard Services. Similarly, HS-TIM was fully integrated into AWIPS version 20.3.2, which allowed the software to take advantage of the latest features within Hazard Services.
- A “QuickWarn” tool for tornado warnings that need to be issued quickly for unwarned storms that have received a tornado report. The tool places a 5 km circle over the userselected mesocyclone centroid, and automatically fills out the rest of the warning with details on an observed tornado. The forecaster merely needs to define the motion vector before issuing the warning. This process usually takes less than 30 seconds.
- The addition of two advanced motion uncertainty techniques:
    - Asymmetric direction uncertainty: This “flares” the warning polygon on either side. Useful for right-turning supercell storms, or left-hooking, occluding tornadoes.
    - *Future storm turning: the forecasters can specify the amount of turning in degrees, the direction of the turn (right or left), and the start and end time of the turn such that on either side of these times, the motion vector is constant.
- **A “Back building” feature for a Persisted warning which will keep the rear end of the warning polygon stationary, only moving the front end of the polygon.
- *Object Split and Merge capabilities.
- Incorporating specific functionalities/modifications to address issues with dissemination:
    - Allowing for a “cooldown” period for locations that have been just removed from the warning after a warning update. This prevents the “windshield wiper effect” where locations can be removed and then immediately placed back into the warning. The default cooldown period was set to 5 minutes, which could be adjusted on the fly (or turned off with a value of 0).
    - **Explicitly disallow the “accordion effect”, in which locations on the downstream portion of the warning are removed if the motion vector speed is decreased, thus shortening the length of the warning. Typically, these locations will be placed back into a warning upon the next warning update, as the storm continues to move forward. The default accordion mitigation time was set to 15 minutes, which could be adjusted on the fly (or turned off with a value of 0). o **An automated county clipping feature that allows forecasters to choose a

threshold for county overlap percentage and whether to clip just the sides or also the front of the warning as a moving warning crosses county boundaries.

- Changed the order and placement of some of the buttons in the Hazard Information Dialog, to improve workflow.
- Modified the "Details" section of the Hazard Information Dialog (HID) to better match HS-Convective and Tiny TIM content. In addition, the "Details" portion of the HID now starts scrolled to the top when a warning update is initiated.
- The extended area of a persisted warning update now uses the VTEC action code EXB.
- Added ETN to object labels and in the console event entries.
- A new experimental warning overlay that displayed both SVR and TOR warning polygons as yellow or red outlines, respectively. This replaced an earlier version that displayed the warnings on a grid. Having the warnings displayed as polygons allowed for improved data sampling, which includes warning type, VTEC action code, ETN, and IBW tags. This new version also can distinguish the cooldown and accordion effect mitigation features with different color shading.
- A memory leak in the third-party graphics code used to draw the HS Console timeline (which affected Hazard Services as a whole, not just HS-TIM) resulted in the software becoming more and more sluggish as the events wore on, to the point of the software being nearly unusable by the end of the 2-hour scenarios. The bug was fixed before this 2022 experiment, removing the sluggishness.

2022 Experiment Details

HS-TIM was tested in the HWT in a limited sense in 2019 and 2020. The first experiment to solely focus on TIM was in the summer of 2021. NWS/MDL, in partnership with NSSL, GSL, and WDTD, carried out another TIM-focused HWT experiment in the summer of 2022. This experiment allowed the collaborators to explore several ideas to represent realistic challenges currently faced in warning operations, focusing on workload and workflow differences with the traditional method for issuing warnings today. Our primary goal was to conduct testing with a broad range of forecasters to assess functionality, identify any impacts on forecaster workload, identify any new dissemination challenges, and suggest improvements and new functionality. End users (e.g., emergency managers, and broadcasters) did not participate in this experiment, so the resulting feedback is only from the operational NWS forecaster perspective.

This HS-TIM experiment was carried out for four weeks in the summer of 2022, with the first week being a “shakedown” of the system using two “test” forecasters. The remaining operational weeks included a pair of participants from several WFOs nationwide. Each week saw forecasters learning how to use the software via a guided training exercise on the first day, and displaced real-time (DRT) severe weather scenarios on Days 2 through 4. Each DRT case had unique domains and represented a variety of different severe storm types (e.g., long-tracked tornadoes, clusters, squall lines, etc.). Storm reports, photographs, video clips, and SPC products were injected using an automated “bot” into a Slack channel and made available to the forecasters in real time. After each scenario, the group gathered for an hour-long discussion of that scenario. The final day was spent conducting a 2-hour guided interview of the forecasters on their experience during the experiment week.

These were the specific objectives of the experiment:

- Technology: Evaluate HS-TIM components and performance to improve the software before operational implementation.
- Human Factors: Measure forecaster workload using HS-TIM, including ease of use and graphical design.
- Methodology: Assess how forecasters adopt their legacy warning methodology into the HS-TIM environment as they screen, rank, analyze, and decide to create, issue, modify, and manage continuous, feature-following objects.
- Conceptology: Collect and analyze data on forecasters’ thoughts on the paradigm change from “static” warnings to continuously-updating warnings.

We gauged participant feedback on user experience, operational applications, and workflow considerations. The data collected will be used to further refine both the software and the concepts of operations of TIM. The following methods were used (those indicated with an asterisk were also recorded and speech-to-text transcripts created):

- Pre-operations online survey
- Discussions during events with meteorologists and developers
- Post-event online surveys
- Post-event discussion*
- End-of-week online survey
- End-of-week interview*

Because of the continuing COVID-19 pandemic restrictions, the experiment was conducted virtually, using a version of the AWIPS software hosted within the Amazon Cloud Services (see Fig. 4). There were some pros and cons to this approach:

- Pros: Developers had quick and convenient browser access to the cloud systems from anywhere so that quick software tests could be performed; forecasters who typically cannot travel to Norman for a variety of reasons could now participate in an HWT experiment; developers could quickly diagnose problems without having to ask forecasters to leave their workstations; each participant had a close-up, high-resolution view of the workstations (rather than looking over shoulders or on mounted TV screens).

- Cons: Workdays had to be shortened (due to "Zoom fatigue"), resulting in less time for forecasters to participate in hands-on scenarios; there was no in-person social interaction outside working hours; we could not split instances to two monitors, meaning forecasters' displays did not match their operational ones

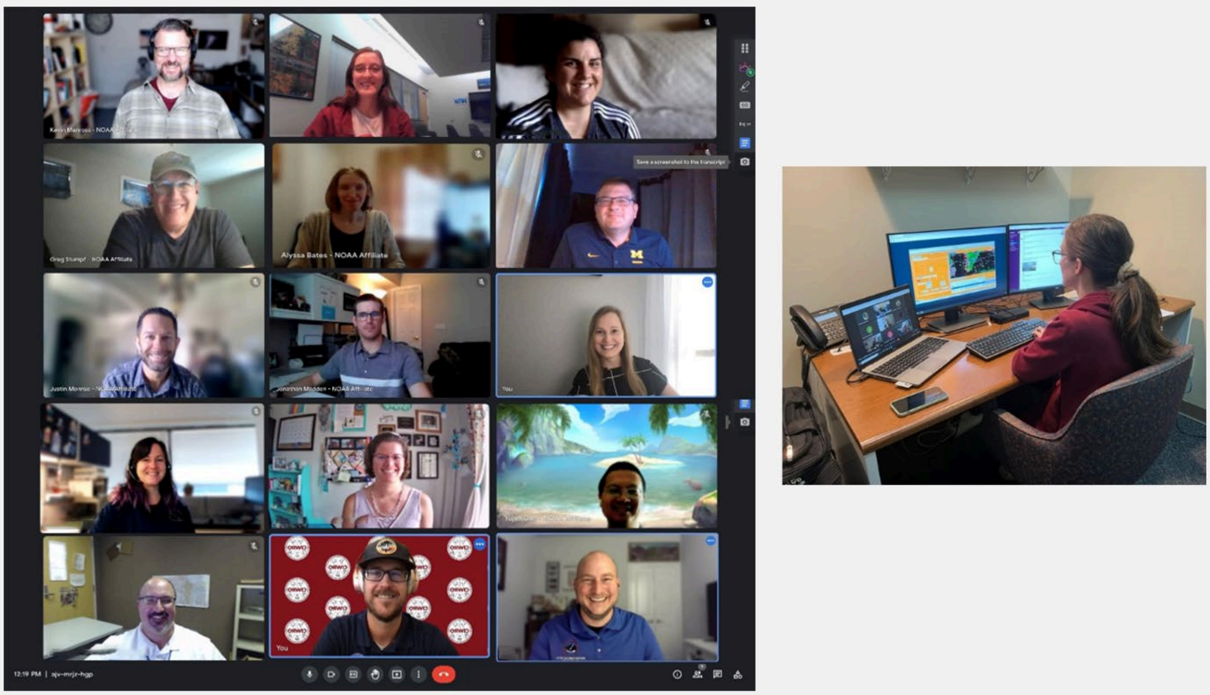


Figure 4. *Photographs from the HS-TIM virtual experiment. Left: A combined group photo representing every participant from the virtual meetings during all three weeks of the experiment. Right: The same forecaster on the top center panel at their workstation during the experiment.*

Given that a year had passed since the previous HWT experiment, and there were many software updates to add new features and improve performance, this experiment was affected by some issues that disrupted workflow and added workload for the forecasters. These issues included:

- Numerous times, warning objects were being "locked", seemingly randomly. This meant that forecasters were unable to update warnings until someone used the "Break Hazard Lock" feature. When this happened, it broke the workflow of the forecasters and caused

some warnings to take longer to issue. We are still investigating this issue after the experiment.
- We changed the warning expiration threshold from 5 minutes to 1 minute to allow the forecaster up to the final minute to extend warnings. Unfortunately, at some point during the experiment, this was no longer possible. Warnings with 1-5 minutes left until expiration were being forced to expire, compelling the forecaster to create a new warning to cover the moving threat. This problem added workload to the forecaster and was contrary to the concept that a single warning could track a single long-tracked hazard.
- Our Scenario 4 which took place in a fabricated CWA that covered portions of the reallife Mobile, Jackson, Birmingham, and Slidell CWAs was plagued by sluggishness in the Shakedown Week and Weeks 1 and 2 of the experiment. This scenario was our busiest and required the highest number of warnings. We traced the problems down to two issues. First, a purge process that cleared out any Ended warnings from the registry was not working correctly. We fixed this after Week 2. Second, and perhaps most importantly, there was an underlying bug within HS that caused the software to become extremely sluggish when any hazard event intersected with the real-life Slidell CWA. Our solution during Week 3 involved asking the forecaster to not create any hazards that intersected the Slidell CWA. This scenario worked without sluggishness for the final week. The cause for this issue was not readily apparent, but it plagues Hazard Services in general, and also happens for the Juneau CWA. This issue may be related to the intricate coastline borders in various shapefiles and other geographic information used within the product generator to determine warning coverage. As of the writing of this report, developers are working on a holistic solution to handle differences in underlying shapefiles that led to the problems discovered in the Slidell and Juneau CWAs.

Here is a summary of the most-notable feedback and suggestions from the forecasters:

- The software does pose a slight learning curve with initial experience due to processes that must be completed in a certain order compared to the current framework as well as new functionality and options. This unfamiliarity with the software may initially slow down the workflow for users until they gain additional experience with the software. Many forecasters agreed that once they became familiar with the software after practice, it was nearly as easy to use as WarnGen. Becoming highly skilled with this system will require continued training and practice. It is recommended that HS-TIM is implemented in the WES before it becomes operational.
- The fact that the warnings follow the storms leads to less worrying about storms exiting warning polygons or making decisions on new warnings downstream. Using the same warning and ETN takes away the laborious aspect of managing hazards with multiple warning re-issuances as storms move forward.

- The automated removal of the back side of warning polygons is a huge time saver since this is done automatically with TIM. Many warnings today are not trimmed properly from where storms have moved out. There is also the added benefit of potential future “All Clear” messages for the public. However, forecasters need to be careful about setting the motion vector too fast, or the back of the warning polygon may clear out too quickly, leaving the storm unwarned and falsely in the “all clear”.
- As with other TIM experiments, the forecasters noted the issue with the Windshield Wiper Effect (WWE) and Accordion Effect, where areas taken out of a polygon may be put back in on the next update. While we had the cooldown and accordion mitigation features available, the forecasters were sometimes confused when they were turned on. It would be nice for the forecasters to be able to toggle on/off which areas will be in the “cooldown” and “accordion mitigation” areas. In addition, forecasters felt that every new warning should have a 5-min cooldown at the start of a new warning because some locations might get an all-clear one minute after a warning is issued. Also, for warning updates, the cooldown should not affect the rear of the warning, only the front and sides. The forecaster should be able to decide if those areas should be immediately removed on the update because if not, it will delay the all-clear.
- The hazard event list in the Console needs some work to improve warning situational awareness. New eye-catching Console display options need to be added to decipher between which objects are Persisting versus those that are not, which warnings are owned by the warning forecaster on that workstation, and which warnings need attention because they are expiring soon.
- Workflow improvements are needed for the Hazard Information Dialog (HID). For example, some buttons should remain visible even after the HID is scrolled into the “Details” section to select impact-based warning tags and calls-to-action. The Modify, Latest TIM Frame, Persist, and Warning Preview buttons should stay “frozen” at the top.
- The Warning Decision Discussion (WDD) feature within HS-TIM, where forecasters can add to a running history of their “thoughts” about why they are issuing and updating a warning, should not be a requirement for warning issuance. It adds workload to the forecaster’s warning-decision-making process and can delay warning issuance if the forecaster gets hung up with spelling and grammar errors as they are typing the discussion. In real operations, the warning forecasters are rarely the persons using NWSchat during warning operations. Therefore, writing the WDD should be a separate process from warning issuance, and anyone in the WFO should be allowed to write it (e.g., the warning coordinator). In addition, the WDD section should be moved to the bottom of the HID since it is an optional attribute. Finally, in the future, the WDD feature might include automatic “injects” into the NWSchat communication channel.
- Need better ways to allow the forecasters to manage the warning swaths. Some felt that the focus was too much on tracking a storm object and much less emphasis on who is warned. One suggestion was to develop tools on the spatial display that would allow

forecasters to edit the warning swath in a way that tied any changes to the object properties. For example, if they made the swath wider, it would automatically increase the motion uncertainty. Or, if they dragged the swath farther downstream from the storm, it would automatically increase the warning duration. Or if they clicked a downstream county off, it would automatically decrease the warning duration.

- While specific hazards no longer are covered by multiple warnings, there can be at times significant overlap of the sides of warnings for hazards that are adjacent (such as a line of supercells). This overlap needs to be better handled unless the capability to edit the polygon swaths directly is allowed.
- Some of the forecasters noted that HS-TIM requires them to first select a warning, and then also have to click the Modify button to update/edit a warning, which adds some workload. This is not required in WarnGen - once a warning is selected, all that is needed to put the warning in Edit mode is to make a change to the warning attributes. Note the reason there is a Modify button in HS-TIM is that once a storm is in Modify mode, if there is a radar frame update, the warning is placed into Edit mode and automatically updates and moves the storm object along the motion vector to remain aligned with the hazard. Once a warning is in Edit mode, it is locked out from any other forecaster. This design is meant to avoid warnings being accidentally placed into Edit mode just by selecting them. One possible solution is to change the way storms are selected for editing, perhaps by selecting with a long right click and choosing Modify from a menu. Or by double-clicking or using a Shift/Ctrl/Alt key with the click.
- Currently, a motion vector can be defined by more than one upstream point. If 3 or more points are used to calculate the motion vector, the variation of the point locations from an average vector contributes to increasing the motion uncertainty. Sometimes the swath widens too much in these cases. Therefore, it was suggested to remove the capability to define motion vectors by more than one upstream point. This was implemented for the 3rd week of the experiment. Beyond the experiment, we will investigate why the swath widening can be very large, and ways to “temper” the value.
- Forecasters had a very favorable opinion of the QuickWarn feature for TOR warnings, so much so that they wanted to use it for all TORs. It saves a lot of clicks and workload. However, the feature is presently designed to be used when a tornado report is received and a warning needs to be immediately issued, and is thus pre-populated for an observed tornado in the IBW tags and calls-to-action. The forecasters suggested adding a second QuickWarn option for radar-indicated tornadoes. One other comment was that the 5 km diameter circle was a bit too small (~3 miles), and they suggested making it a little larger (e.g., 8 km, which is ~5 miles in diameter).
- HS-TIM does not have a Line Tool like WarnGen, and it was suggested to add one that allows the forecaster to set separate buffers on the front and back of the line (the Line Tool in WarnGen uses a 10 km buffer on both sides).

- From a technological point of view, when running the experiment virtually using AWIPS in the Cloud, we are limited to only one monitor per virtual machine instance. Forecasters are used to having 2 or 3 monitors for a workstation desktop, with the ability to tear off the HID and Consoles to other monitors, giving them a larger area for the spatial display. Although we investigated a solution that allows the virtual workstation to be used on multiple monitors, it had some major issues that could not be workable for the experiment. We will investigate further if virtual experiments continue.

Many suggestions for improved software functionality were offered by the forecasters. Some of these suggestions will be incorporated into a future version (as funding and budgets permit) to make the software more robust. These suggestions are listed in 2022 Appendix A. In addition, the forecasters made some suggestions to improve future experiment logistics. These are listed in 2022 Appendix B.

This experiment also included a simple human factors analysis in the form of online surveys designed to measure mental workload, confidence, and software usability. More information about the human factors analysis is included in 2022 Appendix C.

2022 Key Takeaways

- TIM greatly reduces forecaster workload and mental demand when it is needed (for longtracked hazards)
    - Warning updates can be made much faster and are more spatially accurate.
    - Because warnings are in motion, forecasters do not have to spend additional time to decide whether to issue a separate new downstream warning.
    - At times, forecasters can update warnings more frequently, which is especially important for high-impact events.
    - Because warnings are constantly being updated, storms will rarely exit the polygons.
    - Adjustments can be made if the previous warning polygon is not perfect.
- Not having overlapping warnings on a single hazard cleans up the output field and should improve messaging. There is no need to issue a brand-new warning and then manage two separate warnings (the new warning and the warning about to be canceled) for the same hazard if a hazard looks to be exiting the polygon. TIM fixes these situations:
    - When a new downstream warning is issued, it sometimes overlaps the previous warning. When a previous warning is subsequently canceled for the same locations as the new warning, this can send a confusing message.

  - When there are overlapping warnings, two warnings have to be playing on NOAA Weather Radio and television crawls, which can be confusing.
- The forecasters unanimously agreed that the Persist option to put warnings in motion should be available for *any* long-tracked storm, and not just those storms with confirmed or even more-significant severe hazards.

Testimonials

- "The quality and ease with which the NWS will be able to issue high impact warnings has a very bright future and will better achieve the NWS mission on a time and area scale with TIM."
- "The outcomes of this experiment will result in a much better-informed public, and a much better ability for NWS to save lives and property."
- "A great way to bring the NWS into the future, into what today's technology has to offer."
- "I think this will be a great revolutionary shift to the way we do things now - for the better. Conceptually it makes more sense to have warning polygons that move with the storm. It also better serves the public in numerous ways. Yay TIM!"
- "It will be a paradigm shift in how we think about and create warnings, but I think it is a much-needed shift that will greatly improve our warnings to the public when utilized correctly. There will be some growing pains, but I think this technology is easy to learn and understand and will allow forecasters to better manage their workload during severe weather events."
- "I'm very satisfied with the ability to focus on the hazard threat here and now when generating warnings using HS-TIM, and not worry as much about the downstream/future hazard as I would in WarnGen."
- "TIM is the future of the NWS!"

Future Plans

- The HS-TIM software will continue to be developed and improved throughout the course of this project, using an agile methodology with frequent iterations and testing:
  - Complete the development of a real-time capability, including system-wide stress testing and upscaling of domain size to ensure a robust inter-office collaboration capability between multiple WFOs for the CONUS domain.
  - Ensure that warning output meets NWS dissemination requirements – including the need to mitigate rapid shifts in warning location and/or size and deal with

    - county clipping – as highlighted during the TIM dissemination workshops in 2019 and 2021.
    - Fix bugs, and harden/refactor code as necessary.
- Document, analyze, and prioritize newly-identified and forecaster-suggested TIM-specific functionalities, modifications, and other software issues, including those listed in the Appendices.
    - Conduct the above software development.
    - Carry out regular integration of the HS-TIM code into the operational version of HS, to remain up-to-date on HS-Convective development.
- Conduct several virtual ½ day functional assessment tests (FAT) as new functionality is added, using the AWIPS cloud system with forecasters from previous TIM HWT experiments, as situations permit.
- Conduct in-person HWT experiments in the springs of 2023 and 2024 to assess new functionality, and identify any additional impacts on forecaster workload.
    - Include new cases
    - The addition of a 3rd NWS forecaster, to act as the mesoanalyst and the conduit to end users.
- Investigate how to carry out a full dissemination test and update the TIM transition plan with this info, so that we can execute such a test in 2025, perhaps in the OPG. This should include dissemination testing with NCF-Test, NWR, WEA, and private industry partners (e.g., using simulated output files, etc.).
- Social science:
    - Develop and deploy surveys for broadcast meteorologists and emergency managers based on HWT experiment data.
    - Analyze social science survey data, and publish a NOAA Technical Memorandum of the results, to include recommendations for TIM implementation that meets user needs.
- Training:
    - From the most-previous HWT experiment, identify best practices and concepts of operation for future forecaster operational training.
    - Continue to coordinate with the Office of the Chief Learning Officer personnel to create a full-scale forecaster training plan for TIM.
    - Coordinate with Warning Decision Training Division personnel to explore the creation of TIM forecaster training simulations in the WES.
- Update the TIM transition plan based on experimental findings and dissemination investigation by FY-23Q4.

2023 Executive Summary

Threats-In-Motion (TIM) is a proposed approach for warning decisions and dissemination that aims to upgrade the current static polygon system used by the National Weather Service (NWS) for severe thunderstorm and tornado warnings. The main objective of TIM is to create warning polygons that continuously update and move along with the storm at rapid intervals, allowing for improved average warning lead time and more equitable (uniform) lead times for individuals in the path of long-tracked severe storms. The software capability for this solution was developed within AWIPS Hazard Services (HS) and is known as HS-TIM. It was tested *in person* at the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) in Norman, Oklahoma, for three weeks in spring 2023 with participating NWS forecasters. This was the third HS-TIM experiment, following virtual HWT experiments in 2021 and 2022. This was also the first TIM experiment to use a new real-time weather data capability. Real-time cases can be considered an "uncontrolled" test and offer new challenges to the software and concept of operations, as the outcome is not known in advance, unlike when using archive cases.

The proposed concept of HS-TIM represents a paradigm shift in the way warnings could be constructed by forecasters. Instead of creating a static warning polygon swath, continuously updating polygons can be derived from two-dimensional object tracking. Forecasters would have the option of putting the warnings in motion using a feature called "Persist." Persisted warnings would move along the motion vector at one-minute intervals until the forecaster intervenes with another warning update (usually every 10-15 minutes). During an update, forecasters could adjust the motion vector and object geometry and decide whether to continue or discontinue warning persistence. TIM has the potential to reduce forecaster workload by eliminating the need to issue a new warning once a storm has exited a polygon. Additionally, TIM can improve messaging by keeping the same Event Tracking Number (ETN) for the lifecycle of the hazard and providing a cleaner output field by eliminating overlapping warnings on individual hazards.

Since the previous iteration of this HWT experiment, there have been many software updates to add new features and improve performance. However, there were also several new issues with hazard event locking and mouse focus that hampered workload during the 2023 experiment. These undesirable issues, combined with the use of several challenging real-time cases, led to a slight degradation in overall forecaster system usability. Nevertheless, this experiment yielded novel feedback that had never been captured in previous experiments, which could potentially guide new directions in the software, particularly regarding the evolution of adjacent polygons over time during particularly widespread severe weather events.

2023 Detailed Overview

Current severe weather warnings (tornadoes, wind, and hail) require the forecaster to issue multiple sequential warnings for long-tracked storms because the current policy prohibits extending a warning's area and time during updates. This situation often leads to non-uniform lead times for users on the downstream border of a warning polygon. For example, nearly adjacent locations can have significantly different lead times if one location is just outside the upstream warning (Fig. 5).

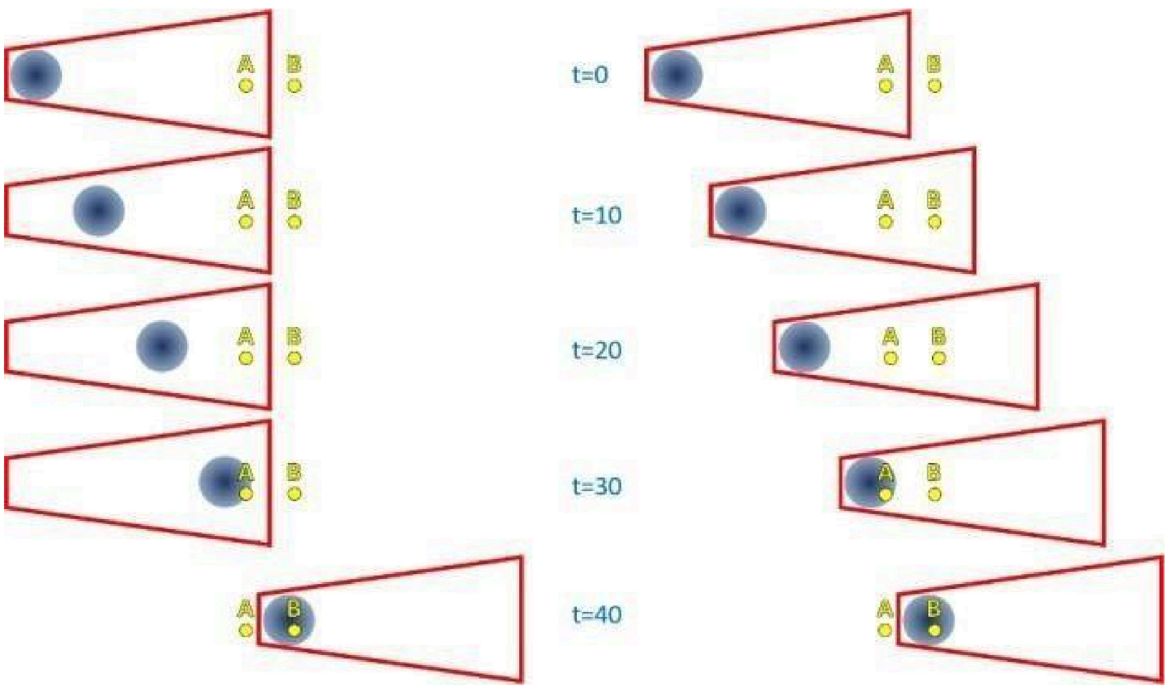


Figure 5. *Comparison of (left) current NWS warning practice using separate polygons and (right) using TIM. The positions of two user locations are shown as A and B. Images are shown at 10-minute intervals; the intermediate 1-minute TIM polygons are not shown. The blue-gray circle represents a hypothetical storm moving from west to east.*

Threats-In-Motion (TIM) is a proposed warning decision and dissemination approach that would enable the NWS to upgrade severe thunderstorm and tornado warnings from the current static polygon system to continuously-updating warning polygons that move with the storm. Specifically, a warning polygon would be attached to the threat and would move forward along with it. Warnings would automatically be cleared from locations where the threat has passed. TIM represents an evolutionary step of the Forecasting A Continuum of Environmental Threats (FACETs) initiative for the convective weather warning scale.

The most significant benefit of TIM is that it can provide more equitable (uniform) lead times for those in the path of long-tracked severe storms because these storms remain continually tracked and warned. This change also results in greater average lead times and decreased average time spent in a warning relative to today's warnings, with little impact on the average false alarm time. This impact is particularly noteworthy for storms expected to last longer than the average warning duration (30-45 minutes), such as the long-tracked supercells observed during violent severe weather outbreaks. A robust statistical analysis of TIM's scientific benefits is available in Stumpf and Gerard (2021).

TIM offers several additional benefits over the current warning methodology. With TIM, the forecaster would only need to issue a single warning, updated regularly as workload permits, embodying a "one storm-one story" concept. This approach has the potential to reduce forecaster workload because downstream warning issuance would be replaced by a less time-consuming warning update. Additionally, TIM would mitigate warning coverage gaps and improve the handling of storm motion changes. TIM would also enhance hazard communication by providing a continuous history for each storm, as opposed to situations where two or more separate overlapping warnings may be in effect for the same hazard (e.g., when a new downstream warning is issued before the previous warning can be canceled). These proposed improvements can lead to more simplified and consistent messaging for key partners and improved event verification.

This promising, innovative approach is currently under consideration for transition to NWS operations. If TIM is approved through the NWS governance process, the transition to operations is expected to occur in two phases:

1. <u>Tiny TIM</u>: This phase adds the capability to extend warnings in time and area, eliminating the need to re-warn long-tracked storms along their path. This results in fewer warning polygons and ensures that one hazard has one event tracking number (ETN). For more information, refer to the summaries of the Tiny TIM experiments (Stumpf 2022a, Stumpf 2022c, Stumpf 2023).

2. <u>TIM (also referred to as "Taller TIM")</u>: This phase includes the capabilities of Tiny TIM and further enhances the system by automatically updating warnings at one-minute intervals, filling in the periods between warning updates. Taller TIM provides an even greater improvement in average lead time and lead time equitability. This phase represents the longer-term "full implementation" of TIM and is the subject of this experiment.

Efforts have been underway since 2019 to develop the software capability for Taller TIM for convective weather warnings (tornadoes, wind, and hail) within the Advanced Weather Interactive Processing System (AWIPS) Hazard Services (Fig. 6). This software is known as Hazard Services - Threats-In-Motion, or HS-TIM. The software has been tested within the NOAA Hazardous Weather Testbed (HWT), and this summary provides details of the most recent experiment held in the spring of 2023 in person at the National Weather Center (NWC) in Norman, Oklahoma.

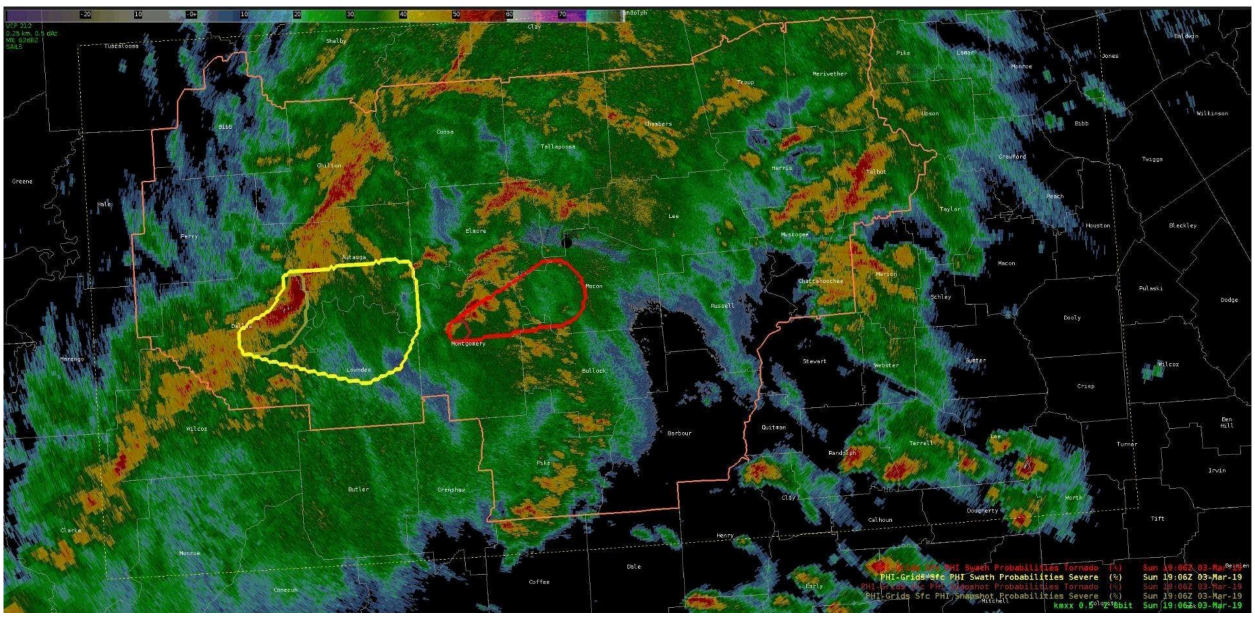


Figure 6. *HS-TIM output from a scenario in southeast Alabama - the yellow (red) polygons show severe thunderstorm (tornado) warnings; the large orange polygon represents the county warning area (CWA).*

2023 Software Details

Differences with current warning software

This summary list highlights the main differences between WarnGen and HS-TIM that forecasters had to learn and understand.

- The HS screen layout (Fig. 7; Spatial Display, Console, Hazard Information Dialog) differs from WarnGen. This layout is already available in other HS perspectives (e.g., HS-Hydro), so some forecasters may already have some experience with it.

- When updating a warning, the user only needs to left-click the warning from the Console or spatial display, eliminating the need to find the specific warning from a drop-down menu as in WarnGen.
- The HS Hazard Information Dialog (HID) "Details" selections are easier to navigate and automatically select appropriate Impact Based Warning (IBW) tags, unlike WarnGen.
- HS-TIM uses 2D storm "objects" (versus points and lines) to define and forecast current threat areas and create the warning swath. These objects are created using drawing tools such as polygon, freeform, or ellipse. The attributes of the objects, including geometry (shape and location), duration, storm motion, and motion uncertainty, are used to create forecast objects at 1-minute intervals, which are then combined to form the warning polygon swath. Forecasters cannot directly edit the polygon swaths; they can only change the swaths by editing the object attributes.
- Defining the motion vector is done using the "Drag Me To Storm" process similar to WarnGen, where forecasters step back a few radar frames in time and drag the 2D object over the location of the threat area in the past. The centroids of the 2D object at the current and past locations are used to calculate the motion vector (speed and direction).
- The forecaster can define the storm motion uncertainty, unlike in WarnGen. A larger value of direction (speed) uncertainty will cause the warning polygon swath to be stretched more in the direction perpendicular (parallel) to the motion vector.
- The "Persist" feature in HS-TIM, when enabled, places a warning "in motion" and automatically moves the warning swath forward every minute between warning updates. Both the front and rear ends of the warning remain in motion while Persist is on. Conversely, for any warning where the Persist feature is off, the rear end of the warning automatically clears out at one-minute intervals.
- During subsequent warning updates (issued at the same frequency as today), the 2D objects can be quickly repositioned and reshaped, and the warning attributes updated to extend the warnings indefinitely or until the storm dissipates. This process in HS-TIM consumes far less workload than with WarnGen, where a new warning needs to be issued each time a warning expires.
- A Warning Decision Discussion (WDD) in HS-TIM allows the warning forecaster to add their thoughts about why they issued or modified the warning (e.g., "mid-level rotation is strengthening"). This information is typically not included in the actual text of a warning product but may be relayed via NWSChat or similar end-user communication software. Since TIM warnings use the same ETN throughout the storm's lifetime, the WDD history is linked from the first time the warning is issued.

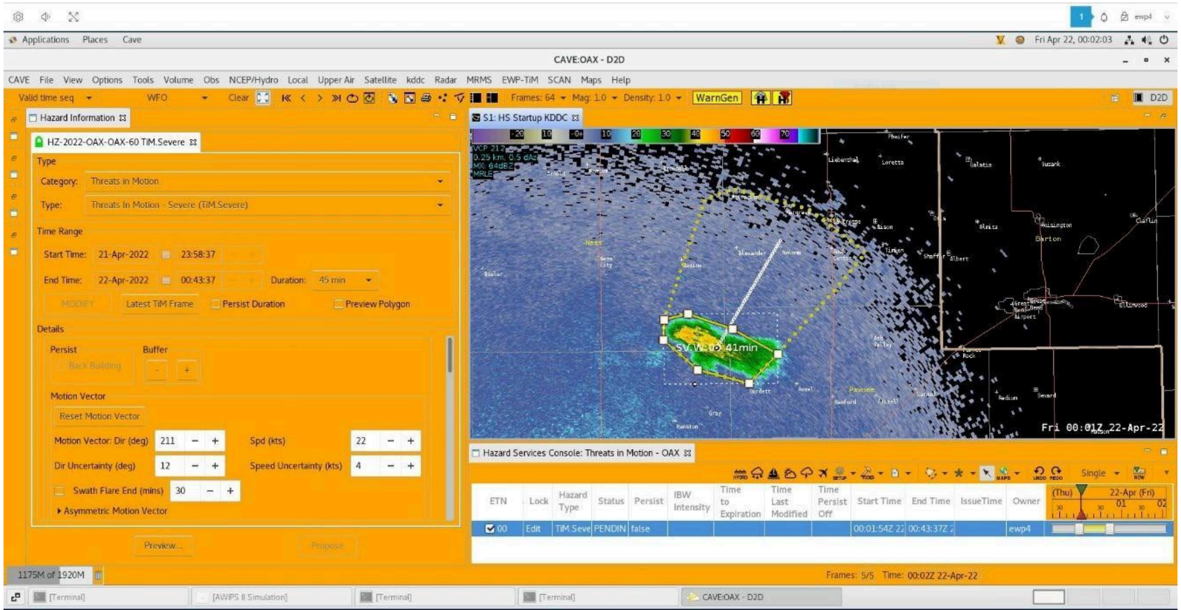


Figure *7. HS-TIM output in AWIPS from the training scenario in Central Kansas.*

The following changes were made to aid in the use of HS-TIM functionality:

- Currently, warnings transition to the "Ending" status with 5 minutes until expiration, and there is no way to extend the warning after this point (only expiration is possible). This threshold was changed to 1 minute, allowing warnings to be extended until the final minute.
- The default alert thresholds for time to expiration were modified as follows:
  - Yellow alert: old (halfway to expiration), new (15 minutes from expiration)
  - Red alert (includes a pop-up alert): old (10 minutes from expiration), new (5 minutes from expiration)
- New HS-TIM default console columns:
  - ETN : Lock : Hazard Type : Status : Persist : IBW : Time to Exp : Time Pers Off : Time Last Mod : Start Time : End Time : Issue Time : Owner.
- Currently, the local warning overlay in AWIPS (under the "Obs" menu) does not display warnings with partial county cancellations (CAN) combined with warnings featuring the new EXB, EXA, or EXT VTEC action codes. Consequently, an experimental warning grid overlay was developed to display warnings with these new VTEC combinations. This experimental warning overlay also offers improved data sampling capabilities, including the VTEC action code, ETN, and IBW tags, which are not available in the local warning overlay. High-end SVR and TOR warnings with Destructive and Catastrophic IBW intensity tags are depicted with thicker lines.

Forecaster Workflow Details

The workflow to create a new warning is as follows (* is optional):

1. Draw a 2D object around the current threat area (also known as "geometry").
2. Select the warning type (SVR or TOR).
3. Define the motion vector using the "drag me to hazard" method.
4. Define the duration.
5. Choose whether or not to Persist the warning.
6. If the warning is persisted, choose whether or not to use Back Building (the rear of the warning will not clear out with time).
7. *Define the motion uncertainty.
8. Select warning attributes (Max Wind/Hail/Tornado, Impact-Based Warning [IBW] tags, Call-To-Action statements, etc.).
9. *Write a Warning Decision Discussion (WDD).
10. Preview the warning text and edit if needed.
11. Issue the warning.

The workflow to update a warning is as follows (* is optional):

1. *Drag the 2D object over the current threat area if the storm motion has changed.
2. *Redraw (edit or replace) the 2D object geometry for the current threat area if the storm shape has changed.
3. Re-define the motion vector.
4. *Re-define the duration.
5. Choose whether or not to Persist the warning.
6. If the warning is persisted, choose whether or not to use Back Building (the rear of the warning will not clear out with time).
7. *Re-define the motion uncertainty.
8. *Change warning attributes (Max Wind/Hail/Tornado, Impact-Based Warning [IBW] tags, Call-To-Action statements, etc.).
9. *Write a Warning Decision Discussion (WDD).
10. Preview the warning text and edit if needed.
11. Issue the warning update.

As the storm moves over time, the 2D object also moves accordingly. Unless the storm motion changes, the 2D object will remain positioned over the storm as time progresses. However, if the storm motion has changed (e.g., in the case of a right-turning supercell), the forecaster can easily drag the 2D storm object back over the current threat area.

Warning Issuance Details

NWS convective warnings are issued with an Event Tracking Number (ETN), which is a unique ID attached to each warning type (Severe Thunderstorm Warning or Tornado Warning). Currently, each subsequent warning issued for a hazard has a different ETN assigned to it. Once the warning expires or is canceled, the ETN is not reused (NWS, 2020). However, with TIM, the same ETN is used throughout the lifecycle of the hazard. The updated warning retains the same ETN because the forecaster does not need to issue a new warning when the previous warning expires. Examples of a series of ETNs and action codes for both current warnings and TIM warnings are shown in Table 3.

Time (UTC)	NWS VTEC	ETN	TIM VTEC	ETN
1900	NEW	1	NEW	1
1910	CON	1	EXB	1
1920	CON	1	EXB	1
1930	NEW	2	EXB	1
1940	CON	2	EXB	1
1950	CON	2	EXB	1
2000	NEW	3	EXB	1
2010	CON	3	EXB	1
2020	CON	3	EXB	1
2030	NEW	4	EXB	1
2040	CON	4	EXB	1
2050	CON	4	EXB	1
2100	NEW	5	EXB	1
*2110	CON	5	CON	1
*2120	CON	5	CON	1
*2130	EXP	5	EXP	1

Table 3. *Warning decision times, VTEC action codes, and ETNs for a hypothetical storm case. Today's warnings are indicated in the blue columns on the left, while TIM warnings are represented in the green columns on the right.*

A graphical representation of the current warning method and TIM is shown in Fig. 8. In this example, warning updates are issued every 15 minutes. Under the current warning methodology, the warning polygon can only be reduced in size with each update, and the warning expiration time remains unchanged. After 60 minutes, the warning is replaced with a new warning that has a new polygon, ETN, and expiration time. In contrast, under the TIM methodology, the warning can progress downstream at each update, with new areas added to the warning, some areas removed from the warning, and the expiration time extended. The TIM warning maintains the same ETN throughout its lifecycle.

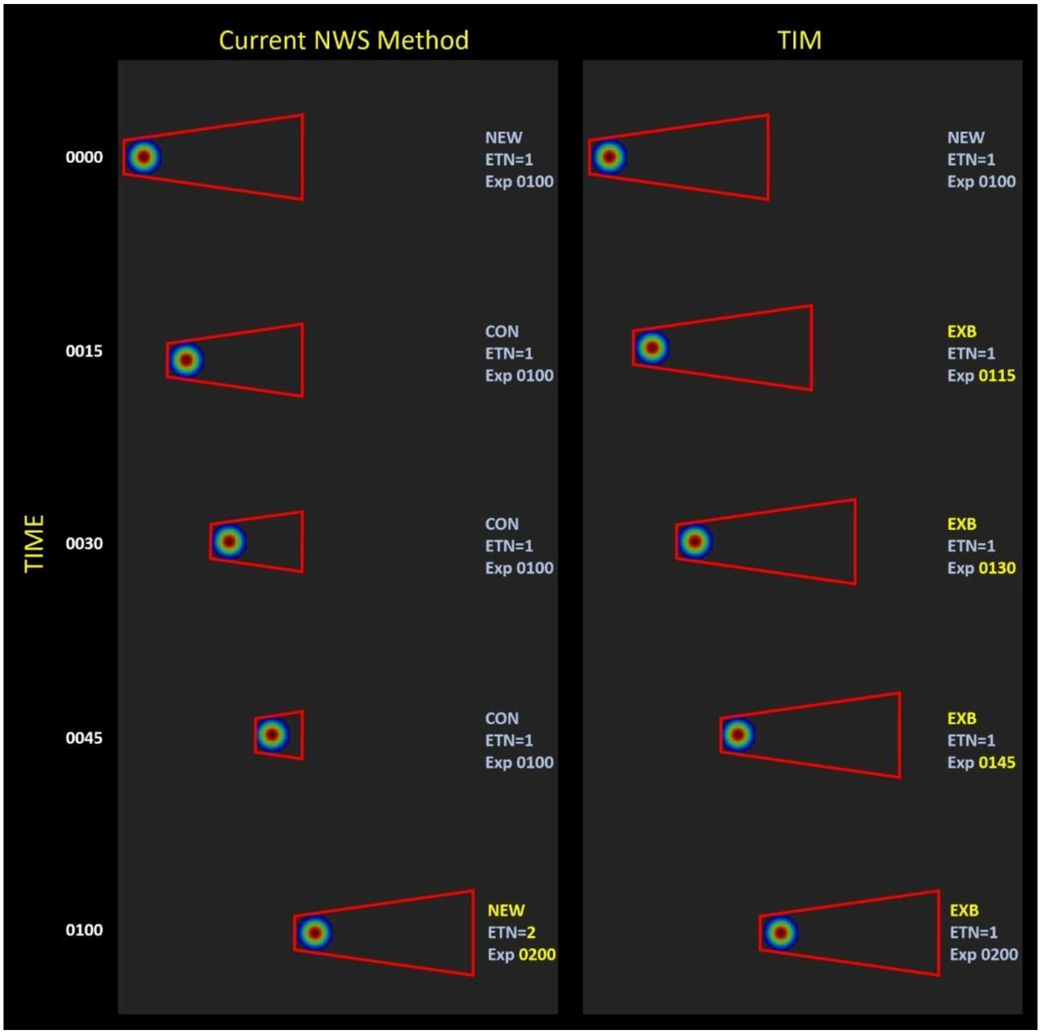


Figure 8. *Comparison of (left) current NWS warning practice using separate polygons and (right) TIM. Images are shown at 15-minute intervals. The VTEC action code, ETN, and warning expiration time are displayed for each time interval. The colored circle represents a hypothetical hazard.*

Persisting vs. Non-Persisting Warning Details

To determine whether to Persist or not Persist the warning, the following general guidelines are used:

- For a storm that is not expected to last longer than the warning duration (short-lived), the warning should not be persisted.
- For a storm that is expected to last longer than the warning duration (long-track), the warning should be persisted.
- For a dissipating long-tracked storm that is no longer expected to last longer than the warning duration, the warning should not be persisted.

Persisting warnings:

- The front side of the polygon continuously moves downstream at 1-minute intervals, providing equitable lead time.
- The rear side of the polygon continuously moves downstream at 1-minute intervals, allowing for automated All Clear statements.
    - If Back-Building is chosen, the rear side of the polygon remains stationary.
- VTEC action code = EXB (EXtend Both area and time).
- The Expiration Time updates every minute.

Non-Persisting warnings:

- The front side of the polygon remains stationary.
- The rear side of the polygon continuously moves downstream at 1-minute intervals, allowing for automated All Clear statements.
- VTEC action code = CON (CONtinuation).
- The Expiration Time does not update every minute.
- When the Expiration Time is reached, the warning will expire (VTEC = EXP).
- The forecaster can also cancel the warning (VTEC = CAN) before the warning expiration.

2023 Recent Software Improvements

Updates have been made to the software for the 2023 HWT experiment, following the HS-TIM experiments at the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) in 2021 and 2022. Many of these improvements were suggested by forecasters in these experiments and are detailed in the experiment summaries (Stumpf 2021, Stumpf 2022b). They include:

- Real-time capability for any CONUS WFO.
- The full functionality for "cooldown" and "accordion" mitigation is now available.
- An additional "cooldown" is implemented on the back edge of all new warnings, so that locations recently added to a new warning remain under warning for a user-adaptable time (default 5 minutes) instead of being immediately removed after the first warning.
- The capability to update the Warning Decision Discussion (WDD) on any warning without issuing a warning update product. This allows someone other than the warning forecaster, such as the warning coordinator, to input this information. The hazard event remains locked during this process to prevent others from overwriting the information.

QuickWarn Tornado Warning Tool:
  - Two options are now available: 1) tornado observed and 2) tornado not observed.
  - The tool now uses an 8 km (5 mi) circle instead of a 5 km circle.
- Hazard Information Dialog (HID):
  - Improved layout of the warning time and geometry controls (Modify, Latest TIM Frame, Persist, and Preview Polygon), so that they will remain "pinned" on top and won't scroll away when editing the warning Impacts Based Warnings (IBW) and Calls To Action (CTA).
  - The IBW tag selection has been improved.
    - Adjacent to CTAs.
    - SVR: Three rows: 1) Hail Size and Hail Source, 2) Wind Speed and Wind Source, 3) Tornado Possible. Hail/Wind defaults to blank.
    - TOR: Two rows: 1) Intensity, Source, 2) Hail Size (this is on a second row since it can sometimes be missed).
  - The Warning Decision Discussion (WDD) dialog has been moved to the bottom of the HID.
  - Advanced motion vector menu:
    - The asymmetric Right and Left labels, slider, and buttons are now on one line each.
    - The +/- on the Advanced MV sliders increment by 1 (for fine-tuning) instead of 5.
  - Unchecking Preview Polygon automatically resets to the Latest TIM Frame.
- Console:
  - A "Time Last Modified" column has been added, allowing forecasters to sort and prioritize warnings that require the highest attention.
  - A Persist column has been included, with false showing no color and true showing blue.

  - A "Time Persist Off" feature has been introduced, displaying a countdown to when Persist will automatically turn off if a warning is not updated within that time.
  - IBW intensity tags (CONS = Considerable, CATA = Catastrophic, DEST = Destructive) have been added.
  - The default column order has been updated to:
    - ETN : Lock : Hazard Type : Status : Persist : IBW : Time to Exp : Time Pers Off : Time Last Mod : Start Time : End Time : Issue Time : Owner.
  - For a Severe Thunderstorm Warning (SVR) with a "tornado possible" tag, a "t" suffix is now added to the ETN.
- Spatial Display:
  - Time to Expiration (minutes) has been added to the warning display label.
- Experimental Warning Overlay:
  - Warning "Emergencies" (SVR [Destructive]; TOR [Catastrophic]) now have thicker polygons.
  - IBW tags have been included in the data readout.

Swath Tool:

  - During a warning update, the current warning polygon is displayed as a dashed shape, enabling the forecaster to visualize the updates they are making to the warning geometry. This helps alleviate the potential addition to the workload when the user has to decide which update options and how much to extend the warning time, to avoid the windshield wiper and accordion effects.
  - Advanced motion uncertainty track turn needs to properly age off. During warning updates, the new motion vector should come from the predicted turn.
  - Warning vertex point controls have been changed to squares to differentiate them from the circular "bounding box" controls.
  - A "buffer" functionality for objects has been added, allowing the forecaster to incrementally adjust the size of the warning by +/- 1 kilometer.
  - Warnings with a right or left turn built into their future track now age off properly when selected for a warning update.
  - A "Swath Flare End" functionality has been introduced (Fig. 9), which terminates the flaring of the polygon swath due to direction uncertainty after a specified amount of time (default is one-half of the duration). This is because the warning will likely be modified by the time we get 50% through the duration. This helps reduce unnecessary false alarm area (FAA).

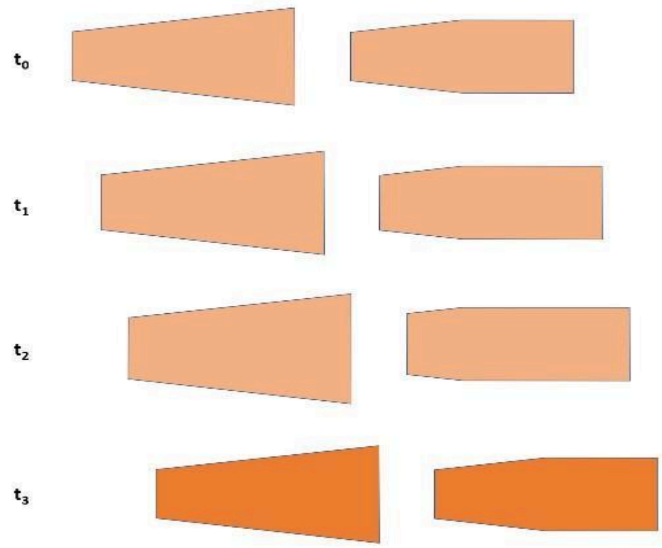


Figure 9. *Illustration of the "Swath Flare End" functionality across three times (*t_0, t_1, t_2, t_3*). The final time (*t_3*) is when the forecaster updates the warning, "resetting" the polygon until the next warning update. The left column demonstrates the current TIM functionality, showing a persisting polygon swath across three times and then "resetting" during the warning update at the fourth time. The right column illustrates how the flaring of the polygon swath, caused by direction uncertainty, terminates halfway through the duration, and then "resets" during the warning update at the fourth time.*

2023 Experiment Details

The NWS Meteorological Development Laboratory (MDL), in collaboration with the National Severe Storms Laboratory (NSSL), the Global Systems Laboratory (GSL), and the NWS Warning Decision Training Division (WDTD), conducted this experiment in person at the HWT for one shakedown (dry run) week and three operational weeks in 2023:

- Shakedown Week: Apr 3 – Apr 7
- Week 1: Apr 17 – Apr 21
- Week 2: May 1 – May 5
- Week 3: May 8 – May 12

This was the second in-person experiment conducted at the HWT in three years (following an inperson Tiny TIM experiment earlier in 2023), and the third HWT experiment focused on HSTIM, following experiments in 2021 and 2022 (Stumpf, 2021; Stumpf, 2022b).

This HWT experiment allowed the collaborators to explore several ideas to represent realistic challenges currently faced in warning operations with a broad range of forecasters. The primary focus was on workload and workflow differences compared to the traditional method of issuing warnings today. The aim was to assess functionality, identify any impacts on forecaster workload, uncover any inconsistencies across CWAs/products/services, examine dissemination challenges, analyze ideal timing updates, and suggest improvements and new functionality.

Real-time capabilities were available for the first time in any HS-TIM experiment (or Tiny TIM or HS-PHI experiment for that matter). This introduced some unique challenges, as the cases could not be curated in advance of the experiment, except for determining the area and time of the event in advance based on model data and the Storm Prediction Center (SPC) Day 1 convective outlook. Fortunately, there were severe storms during each real-time event, encompassing a variety of storm types and coverage.

Each week, the experiment included three participants: two warning forecasters and one warning coordinator/mesoanalyst (WC/MesoA) (Figs. 10 and 11). For each scenario, a different participant would rotate into the WC/MesoA position. The WC/MesoA aided by providing information about the near-storm environment and its impact on storm evolution, as well as keeping track of all the warnings and storm reports, and monitoring the forecasters' workload. The WC/MesoA would step in as a third warning forecaster when needed if the workload became too high for the other two.

For this experiment, only forecasters participated (no end users), so the resulting feedback is from the operational NWS forecaster perspective. The first week of the HS-TIM HWT experiment was a "shakedown" of the system, involving three "test" forecasters from the Norman, OK, NWS Weather Forecast Office (WFO). The shakedown week was used to identify and resolve any issues (software and logistical) that required attention before the operational weeks of the experiment commenced.

On the first day (Monday), after an introductory presentation and software demo, forecasters were instructed on how to use the software through a jobsheet-guided, hands-on training exercise. The original 2015 training case was replaced by a more contemporary case from 2022 that included the latest data types (e.g., GOES-R satellite data) to facilitate the creation of AWIPS procedures for the rest of the week. Since the last experiment, we extended the training session by an additional hour and allotted forecasters more time to practice independently or as a group. Additionally, we incorporated the advanced features into the training itself, rather than introducing them later during event scenarios. Lastly, we developed a concise quick reference guide that provided crucial information and instructions to assist forecasters with software functionality, workflow shortcuts, and troubleshooting guidance.

Fig. 10. *Photos from the HS-TIM experiment in action at the HWT.*

On Days 2, 3, and 4 (Tuesday, Wednesday, and Thursday), forecasters issued TIM warnings for either archive displaced real-time (DRT) severe weather scenarios or real-time (RT) scenarios. Each DRT case had unique domains and represented a variety of severe storm types, such as long-tracked tornadoes, squall lines with quasi-linear convective system (QLCS) tornadoes, right-turning supercells, and left-hooking occluding tornadoes (Table 2). Two new archive cases were added, in addition to one used in 2022, including an event spanning 5 ½ hours without breaks. Typically, there were two scenarios per shift unless Scenario 6 was used. In two cases, a real-time event spanned the entire shift. Scenarios 3 and 7 were used for all three weeks, and Scenario 6 was used for two of the three weeks. Storm reports, photographs, video clips, and Storm Prediction Center products were injected into a Slack channel using an automated "bot" and made available to the forecasters as they had occurred when the event happened. For some of the archive scenarios, original television broadcasts were played, synchronized in time with the event. The remaining time was spent on real-time scenarios (Table 4).

Fig. 11. *Group photos from Weeks 1, 2, and 3 (clockwise).*

For the RT scenarios, at the end of each Monday, Tuesday, and Wednesday shift, we used the SPC Day 2 Convective Outlook to determine the shift time for the following day. If no real-time events were to be worked, the shift time would be 9 am – 5 pm with lunch eaten midway through. If severe weather was expected, the shift time was chosen from three options (12 – 8 pm, 1 – 9 pm, 2 – 10 pm), with dinner eaten midway through. Storm reports and other information from monitoring Twitter and NWSChat were manually entered into the Slack channel and made available to the forecasters in real time. Live television was also played during real-time events, where available.

Scenario	Type	Storm Modes	Length
S1*	Training	Isolated supercells	2 hours
S3	Operations	Long-tracked supercells	2 hours 15 minutes
S6*	Operations	Multiple supercell areas; left-hooking tornadoes	5 hours 30 minutes
S7*	Operations	Squall line with QCLS tornadoes	1 hour 45 minutes

Table 3. *Description of archived displaced real-time (DRT) scenarios. An asterisk indicates a new scenario introduced in 2023. Scenarios 2, 4, and 5, which were developed before 2023, were not utilized in this experiment.*

Scenario	Date	WFO	Storm Modes	Length
RT1	18 April	GID	Marginally-severe cells	2 hours 30 minutes
RT2†	19 April	OUN	Long-track supercells; left-hooking tornadoes	3 hours 30 minutes
RT3	4 May	OUN	Long-track supercells; no tornadoes	6 hours 15 minutes
RT4	9 May	EAX	Marginally-severe cells and lines	2 hours 45 minutes
RT5	10 May	BOU	Multicell clusters and boundary interactions	4 hours 45 minutes
RT6†	11 May	OUN	Long-track supercells; left-hooking tornadoes	6 hours 15 minutes

Table 4. *Dates, WFOs (Weather Forecast Offices), descriptions, and durations of real-time (RT) scenarios. A "dagger" (†) denotes an event where the location of our experiment (the National Weather Center) was at risk of a nearby tornado.*

After each scenario, the forecasters completed an online workload survey. Following that, the entire group convened for a 45-minute discussion focused on that particular scenario. Upon completion of all the scenarios on Thursday, the forecasters participated in an online software usability survey. The final day (Friday), which was a half day, was dedicated to conducting a 3hour guided interview with the forecasters to gather their feedback and experiences from the experiment week.

These were the specific objectives of the experiment:

- Technology: Evaluate the components and performance of HS-TIM to enhance the software before its operational implementation.

- Human Factors: Measure forecaster workload using HS-TIM, including assessing ease of use and graphical design.

- Methodology: Assess how forecasters adapt their existing warning methodology to the HS-TIM environment while creating and managing continuous, feature-following warnings.

- Conceptology: Collect and analyze data on forecasters' perspectives regarding the paradigm shift from "static" warnings to continuously-updating warnings.

We gathered participant feedback on user experience, operational applications, and workflow considerations. The data collected will be utilized to further refine both the software and the concepts of operations of TIM. The following methods were employed (those marked with an asterisk were also recorded, and speech-to-text transcripts were generated):

- Pre-operations online survey
- Discussions during events with meteorologists and developers
- Post-event online workload surveys
- Post-event discussion*
- End-of-week online usability survey
- End-of-week interview*

Before the COVID-19 pandemic, we conducted in-person HWT experiments using physical workstations on-site. However, during the pandemic, we shifted to conducting our experiments virtually using a version of the AWIPS software hosted by the NOAA Virtual Laboratory (VLab) on Amazon Web Services. This cloud-based software allowed us to carry out software testing in a shared environment, where developers could collaborate simultaneously. We decided to use the virtual system within the HWT for the experiment instead of maintaining the latest software builds on both physical and virtual systems. We also learned how to split a virtual instance across two monitors, which provided us with additional screen space. This setup allowed us to run one Common AWIPS Visualization Environment (CAVE) instance per monitor, dedicating one monitor exclusively to Hazard Services and the other for viewing and analyzing additional data using one or more CAVEs. This approach had its advantages and disadvantages:

- Advantages: Developers enjoyed the convenience of quick browser access to the cloud systems from any location, facilitating rapid software tests. They could diagnose issues promptly without disrupting the forecasters' workflow. Moreover, given the strict COVID protocols in place, ill participants who experienced mild or no symptoms could still actively take part in the experiment through virtual means.
- Disadvantages: The virtual instances operated at a slightly slower pace compared to the physical systems. However, it is anticipated that this performance limitation will improve as cloud computing technology continues to advance.

2023 Results

The HS-TIM software was affected this year by some new issues that disrupted workflow and added workload for the forecasters. Some of these issues were inherent in the baseline Hazard

Services and not necessarily due to TIM. These issues will require immediate attention from the developers to ensure they do not affect future experiments. The issues included:

- Locking errors: Numerous times, warning objects were being "locked" while not being edited, seemingly randomly. This meant that forecasters were unable to update warnings until someone used the "Break Hazard Lock" feature. Some warnings were also becoming randomly unlocked while being edited. Many warnings were throwing "interoperability errors" as well when this happened. There were other side effects: some warnings elapsed on their own when they should not, some continuing warnings had their ETNs reassigned resulting in mismatching between display labels and Console entries, and there were multiple AlertViz banners that needed to be acknowledged. All of these issues disrupted the workflow of the forecasters and caused some warnings to take much longer to issue than usual.
- Mouse Focus: Occasionally, when drawing an object, the mouse loses focus, causing the drawing tool to disengage. As a result, forecasters had to restart drawing the object from scratch, adding unnecessary workload. This issue occurred when the radar frame updated, the Swath Tool updated with each edit, or when the Timer Alert or any other tool ran. To mitigate this within HS-TIM, we attempted to limit the number of tool runs by having the Swath Tool update only while an object is being edited on the editing workstation and scheduling all Timer Alert Tool runs at the top of the minute. This issue was not present before the 2023 experiment and was introduced within the baseline HS in the past year. Efforts are underway to address this problem in the baseline HS.
- Window Focus: The window focus does not always follow the mouse. If a user wants to use keyboard shortcuts to increase or decrease the radar frame time, an additional mouse click is required within the spatial display to regain focus after editing the HID. This adds to the workload. This is not an individual workstation or CAVE preference setting, but rather a software setting, as this problem was not present in previous experiments.
- Cloud System Speed: AWIPS in the cloud appeared to be slower compared to previous experiments and physical systems. This was the first time we expanded the cloud instances across two monitors, which may have contributed to the slowness. For future experiments, we need to improve the response speed. One potential solution is to investigate ways to synchronize cloud systems with the physical HWT systems. This would allow the system to be remotely monitored when necessary while still enabling forecasters to utilize the faster physical systems.

Some additional changes were made before the first week of the experiment based on feedback from the forecasters during the shakedown week. These changes included:

- Made improvements to the training job sheet and the Quick Reference Guide.

- Changed the default SVR duration from 45 minutes to 30 minutes.
- Added bookmarks for the SPC Meso Analysis data, Pivotal Weather model data, the NSSL Warn on Forecast System (WoFS), and upper-air soundings to the corner computer used by the Mesoanalyst.
- Implemented a script in the real-time system to purge the warning data from the last realtime run.
- Adjusted the dashed outline of previous warnings to a brighter gray.
- Changed the county clipping default to "Update All Sides."
- Set the cooldown and accordion defaults to 0 minutes (turned off), and only activated them (one or the other) for specific scenarios after the forecasters had already become accustomed to how warnings persist without these two features.

Here is a summary of the most-notable experiment findings and suggestions from the forecasters:

- Throughout each operational week, the forecasters' comfort level with the software increased as they became more familiar with it, although some felt that one week was insufficient to grasp all the new advanced features. It was agreed that a complete refresh of warning decision-making training was necessary, requiring ALL warning forecasters to undergo a new in-person or cloud-based week-long course similar to WDTD's Radar and Applications Course (RAC) workshop. Annual seasonal readiness virtual training on TIM would need to be incorporated into the existing RAC course. The courses would need to be tailored to accommodate varying technical skill levels among the service.

- Workload varied greatly depending on the case, with some real-time events often demanding a higher workload. Certain cases during the experiment, including the April 19 and May 11 events in the Norman CWA, added additional stress due to significant tornadoes threatening the area. Unusual storm evolutions during these cases presented unique challenges. The April 19 event is considered a strong candidate for future experiments.

- The workload was also influenced by the above negative system issues. While the Persist feature saved time by eliminating the need to manage expiring warnings, dealing with system issues sometimes offset the time saved. Overall, when the system worked well, forecasters felt that TIM greatly reduced workload and stress, allowing them to keep up with warning updates. Forecasters were less concerned about storms exiting the polygon, as it was easy to update the warning. It was also much easier to recover from mistakes compared to using WarnGen.

- Maintaining multiple warnings required some effort, but the Persist feature reduced stress by alleviating concerns about warnings expiring. The Persist option made it easier to focus on storm geometry and motion without the constant need for warning re-issuance. Reducing redundancies with downstream warnings increased forecaster productivity.

- The hazard event list in the Console was very useful, especially with the new "Time Last Modified" addition, which helped forecasters quickly sort warnings that required immediate attention. Forecasters expressed a desire for a way to determine quickly which warnings are owned by the warning forecaster on a workstation. The suggestion to display "my warnings" has been made for all flavors of HS (not just TIM) and will be added to the baseline software soon.

- The Hazard Information Dialog (HID) was generally well-organized, but it may now contain too many options for warnings, including geometry drawing options, uncertainty, buffer, asymmetric flare, and turning tracks. Some forecasters felt the need to use all these options, even though the latter of these options were meant to be used sparingly. Some thought that the system may have gone too far with too many new options.

- Regarding advanced motion uncertainty options, the consensus was that the asymmetric flare would be more useful for depicting potential right or left-turning hazards compared to the turn feature, as it is more challenging trying to predict when and how much the hazard will turn. They suggested combining the direction uncertainty control with the asymmetric flare control, using a left slider, right slider, and a "lock to both" button similar to the NSSL PHI Prototype. This would eliminate the need for a separate asymmetric direction uncertainty control.

- Having a separate Warning Decision Discussion (WDD) Tool gained favorable feedback. It allowed any forecaster to add to the warning's running history of forecaster "thoughts" without updating the entire warning, saving workload and removing delays to warning issuance. The warning coordinator usually issued these discussions. In the future, the WDD feature might include automatic "injects" into the NWSChat communication channel.

- Forecasters had a very favorable opinion of the QuickWarn feature for TOR warnings and wanted to use it for all TORs. To improve the capability and speed up warning issuance, the software should import a default storm motion vector to eliminate the need for manual computation when time is of the essence. They also suggested adding a QuickWarn for SVR that starts with a symmetric 6-sided polygon and pre-filled HID fields.

- Forecasters preferred the 2D object method for outlining hazards rather than focusing on the exact centroid using the WarnGen point tool. However, they also felt that HS-TIM needs a line tool similar to WarnGen for proper storm motion representation, especially for large or linear objects. They suggested adding a Line Tool that allows separate buffers on the front and back of the line and provides a unique motion vector for each linesegment vertex, similar to the Dynamic Line Tool in Tiny TIM.

- Some forecasters suggested improving the swath decorations, such as centroids, object outlines, proposed polygon outlines, and current polygon outlines, with brighter colors and a choice of line type (solid versus dashed). This should be a configurable option.

- There is unanimous agreement among forecasters that warning extensions should be allowed for **any** long-tracked storm, regardless of the presence of observed significant hazards. Confirmation can be challenging, especially at night, and a long-tracked event may not produce a tornado immediately, resulting in zero lead time if confirmation is required to start using Persist. One forecaster commented that “handcuffing a useful tool to only be used during significant severe events was seen as unnecessary.”

- Mitigation functionalities like cooldown and accordion effect received positive feedback. Forecasters requested the ability to preview these effects before issuing warnings and to toggle them on/off on a per-warning basis. The automated removal of the back side of warning polygons was seen as a time saver, but adjustments were suggested to prevent unwarned storms.

- The automated removal of the back side of warning polygons was seen as a huge time saver. However, forecasters noted that sometimes they set the motion vector too fast, causing the back of the warning polygon to clear out too quickly, leaving the storm unwarned. They suggested replacing the Back-Building feature with an "All Clear" setting that offers three different speed choices: "full" (clears out as quickly as the polygon moves forward), "half" (clears out at half the speed of the forward edge), and "off" (no clearing for back-building hazards – this capability was tested in the NSSL PHI Prototype). This functionality would be available regardless of the Persist setting.

- Forecasters expressed the need for manual edits to warning polygon swaths, as is done today in WarnGen, rather than only being able to shape the warning by adjusting the duration, speed, direction, and motion uncertainty values. However, the capability to derive moving swath and location-specific times of arrival and departure would be compromised. The development team is urged to consider adding capabilities to directly edit the polygon swath while keeping it tied to objects. For example, 1) Interacting with the centerline to change direction and speed, 2) Controls to flare the swath on either or both sides and 3) Controls to add a curve to the centerline. We would also need to determine how this would affect the dissemination/end-user side of things.

- Gaps between polygons over time and overlapping adjacent warnings were identified as issues that needed better handling. The ability to preview all polygons, not just the one being worked on, was suggested to ensure there are no future gaps or overlaps.

- While TIM allows for faster forecaster updates, caution was advised to avoid excessively frequent updates that may compromise situational awareness. Updates should only occur when conditions change, such as when the first observations of a tornado are received, when the intensity ramps up, if the motion changes, etc. However, situations may arise where very frequent updates are warranted, including high-impact events affecting a population center, as long as there is sufficient WFO staffing to handle other hazards in the CWA.

- Exploiting the ProbSevere storm object identification algorithm and incorporating PHI Prototype probabilistic grids as guidance products were also proposed as potential improvements.

- The phased implementation of TIM capabilities, starting with Tiny TIM before the full TIM capability, received unanimous agreement. Some capabilities within Tiny TIM which don't require warning extension capabilities, such as the 2D Footprint Tool, were suggested to be phased in earlier.

2023 Appendix A contains a comprehensive list of suggestions for improving the software functionality provided by the forecasters, including those mentioned above. Some of these suggestions will be considered for incorporation into future versions of the software, subject to funding and budget availability, to enhance its robustness. Furthermore, the forecasters also made suggestions to enhance the logistics of future experiments, which are listed in 2023 Appendix B.

This experiment also incorporated a basic human factors analysis through online surveys, aiming to assess mental workload, confidence levels, and software usability. Further details regarding the human factors analysis can be found in 2023 Appendix C.

2023 Testimonials

"I hope that TIM is the future of NWS warnings because the concept is something we should be embracing as meteorologists and communicators."

"We should have been doing this years ago. We are still in the old paradigm 16 years later. TIM is a big plus for the customers."

"Weather, especially severe thunderstorms/tornadoes, is not static. Why are we using static warnings for something that is moving? It's easy to make mistakes using static polygons, so this gives us a chance to adjust."

"The system allowed more time to interrogate, analyze, and issue warnings, ultimately making me more productive and decreasing the physical demands of issuing products. I strongly believe this allowed me to have more confidence when issuing my warnings because it allowed easier customizations and quick tools."

"I think Taller TIM has great promise and conceptually would be a significant upgrade from the current methodology of NWS convective warnings. However, I think some of the applications of the system still need to be improved to be more user-friendly and reduce the amount of time needed to edit warning attributes."

"Having the Persist function is a game-changer, especially in our area with smaller counties and more warnings. I had more time for storm interrogation than reissuing downstream again."

"Not having to think about my polygonology as much makes life much better. With the HSTIM console showing update times, owners, events, etc., it was easy to note which objects needed attention. With objects being drawn directly on radar overlays, it's also easy to see when things are going wrong. Both of those save time, which gives you more time to look at the actual radar data!"

"I think this was one of the better-run experiments I've been a part of. Work with what you've got. Well done, well thought out."

"I think the system is very powerful! There are a lot of great tools and functions with this system that I think will do well in a real-time operational setting. The ability to expand a warning in both time and space is something that I've wished I could do since issuing my first warning. I will undoubtedly go back to my home office and perform warning ops thinking about how much better this could be with Threats-in-Motion. It's easy to create updates, add information about a storm using the WDD tool, trim warnings, expand warnings, etc."

"I think it's about as close to perfect as you can make a 4.5-day experiment last. I wish it could've been longer, but I never found myself thinking 'Why are we doing this now?' or 'Why didn't we do that some other time?' I think it was great!"

"If I hear any comments from management, it is polygonology. We avoid those issues with the new paradigm. From a broadcast perspective, it would easily enhance their messaging."

"It's coming along, and I can't wait to see it in operations!"

"I am excited to see how this evolves!"

"Excited for what this will be, but the speed at which it will get here is not fast enough."

2023 Key Takeaways

- HS-TIM has the potential to significantly reduce forecaster workload and mental demand, particularly for long-tracked hazards.
    - Warning updates can be executed faster and with improved spatial accuracy. o The option to extend warnings eliminates the need for separate downstream warnings, reducing decision-making complexity.
    - Timely updates ensure that storms rarely exit the warning polygons, allowing for adjustments if the previous polygon is not optimal.
- The elimination of overlapping warnings for a single hazard has the potential to result in a cleaner output field and improved messaging. There is no longer a need to issue a completely new warning and manage two separate warnings (the new warning and the imminent cancellation of the previous one) for the same hazard when it appears to be leaving the polygon. HS-TIM addresses these situations:
    - When a new downstream warning overlaps with the previous one, canceling the previous warning can create confusion.
    - Overlapping warnings require the simultaneous broadcast of two warnings on NWR and television crawls, leading to potential confusion.
- The HS-TIM software was affected this year by some new issues that disrupted workflow and added workload for the forecasters, such as hazard locking and mouse focus issues. Some of these issues were inherent in the baseline Hazard Services and not necessarily

due to TIM. Software developers will immediately focus on these limitations to ensure they do not affect future experiments.

- With the introduction of real-time events, the workload varied greatly from the controlled archive case experiments done in the past. Testing the software and concept on as many use cases as possible is needed before any operational implementation.

2023 Future Plans

- The HS-TIM software will continue to be developed and improved throughout the course of this project, using an agile methodology with frequent iterations and testing:
    - Complete the development of real-time capability, including system-wide stress testing and upscaling of domain size, to ensure a robust inter-office collaboration capability between multiple WFOs for the CONUS domain.
    - Ensure that warning output meets NWS dissemination requirements, including the need to mitigate rapid shifts in warning location and/or size and deal with county clipping, as highlighted during the TIM dissemination workshops in 2019 and 2021.
    - Fix bugs and harden/refactor code as necessary.
- Document, analyze, and prioritize newly identified and forecaster-suggested TIMspecific functionalities, modifications, and other software issues, including those listed in the Appendices:
    - Conduct the new software development. o Regularly integrate the HS-TIM code into the operational version of HS to stay current with HS-Convective development.
    - Merge workflows for Tiny TIM and Taller TIM (HS-TIM) into a single codebase.
- Conduct several virtual half-day functional assessment tests (FAT) as new functionality is added, using the AWIPS cloud system with forecasters from previous TIM HWT experiments, as situations permit.
- Conduct in-person HWT experiments in the spring of 2024 to assess new functionality and identify any additional impacts on forecaster workload:
    - Include new cases.
- The NWS is investing in the infrastructure to enhance CAP messaging to enable moving polygons for all hazards. As part of this effort, the TIM transition plan needs to be updated, and a full dissemination test should be executed in 2025 at the OPG. This should include dissemination testing with NCF-Test, NWR, WEA, and private industry partners (e.g., using simulated output files, etc.).
- Social science:

    - Develop and deploy surveys for broadcast meteorologists and emergency managers based on HWT experiment data.
    - Analyze social science survey data and publish a NOAA Technical Memorandum of the results, including recommendations for TIM implementation that meets user needs.
- Training:
    - Identify best practices and concepts of operation for future forecaster operational training based on the most recent HWT experiments.
    - Continue familiarizing forecasters with HS-TIM and work toward starting a preliminary training plan for HS-TIM that spans multiple divisions of the Office of the Chief Learning Officer within the NWS.
    - Coordinate with Warning Decision Training Division personnel to explore the creation of TIM forecaster training simulations in the WES.
- Update the TIM transition plan based on experimental findings and dissemination investigation when appropriate.

Web Presence

- NSSL Bite-Sized Science 3-minute video on Threats-In-Motion
- A Blog summary about TIM (use NOAA credentials to log in)
- The TIM Weather and Forecasting journal article (Stumpf and Gerard, 2021).

Disclosure

ChatGPT was used to assist in proofreading sections of this summary.

Personnel

Greg Stumpf	CIRA* and NWS/MDL	Experiment Coordinator / Project Co-PI
Kevin Manross	CIRA and NOAA/GSL	Lead Developer / Project Co-PI
Alyssa Bates	CIWRO** and NWS/WDTD	Project Scientist / Project Co-PI
Yujun Guo	CIRA and NOAA/GSL	Developer
Pat Hyland	CIWRO and NOAA/NSSL	Project Scientist
Justin Monroe	CIWRO and NOAA/NSSL	HWT AWIPS Support

Jonny Madden	CIWRO and NOAA/NSSL	HWT AWIPS Support
Kodi Berry	NOAA/NSSL	NSSL FACETs Program Manager

*CIRA - Cooperative Institute for Research in the Atmosphere
**CIWRO - Cooperative Institute for Severe and High-Impact Weather Research and Operations

References

Stumpf, G. J., and A. E. Gerard, 2021: National Weather Service severe weather warnings as Threats-in-Motion (TIM). *Wea. Forecasting*, **36**, 627-643. https://doi.org/10.1175/WAF-D20-0159.1

Stumpf, G. J., 2021: 2021 Hazard Services – Threats-In-Motion (HS-TIM) HWT Experiment Summary.

Stumpf, G. J., 2022a: 2022 Tiny – Threats-In-Motion (Tiny TIM) HWT Experiment Summary.

Stumpf, G. J., 2022b: 2022 Hazard Services – Threats-In-Motion (HS-TIM) HWT Experiment Summary.

Stumpf, G. J., 2022c: 2022 Tiny – Threats-In-Motion (Tiny TIM) OPG Experiment Summary.

Stumpf, G. J., 2023: 2023 Tiny – Threats-In-Motion (Tiny TIM) OPG Experiment Summary.

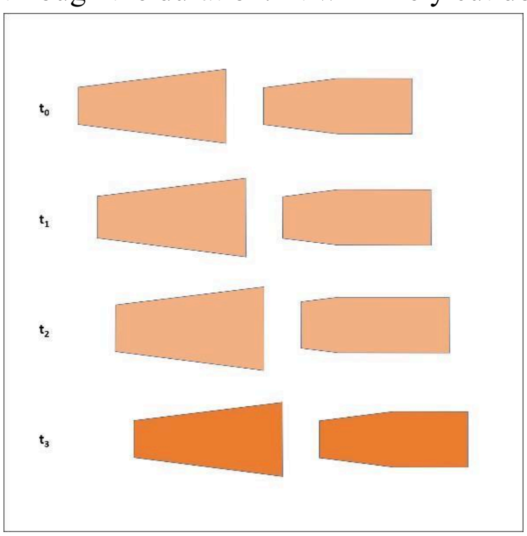
2022 Appendix A: Software Functionality Suggestions

NOTE: Enhancements that are within the broader scope of HS – Convective as a whole (and not specific to HS-TIM) are indicated with a "(HS-C)" prefix.

Warning Output:

- Cooldown:
  - For new warnings, add a 5-min cooldown at the start of a new warning, because some locations might get an all-clear one minute after a warning is issued.
  - For warning updates, cooldown should not affect the rear of the warning, only the front and sides. The forecaster should be able to decide if those areas should be immediately removed on the update because if not, it will delay the all-clear. o If the storm speed is too fast, the warning could clear out before the hazard departs

the area. Perhaps for fast-moving storms, there is a slower cooldown on the back.
    - Ability to control cooldown on a per-warning basis.
- Experimental warning overlay:
    - Separate SVR and TOR warnings, as is done with the local warning overlay. o Add two more experimental warning layers [Warnings (without Cooldown/ Accordion delineated as dimmed layers); Warnings (with Cooldown/Accordion delineated as dimmed layers); Snapshot (Threat Area)]. This would help forecasters see the effects of each independently.
    - Warning "Emergencies" (SVR [Destructive]; TOR [Catastrophic]) should have thicker polygons.
    - Persisting polygons/objects should be delineated differently, perhaps with a dashed-dot-dashed pattern instead of a solid pattern.
- Output should be smart enough to prevent warnings that are overlapping on the edges. When one warning ends, the other warning should fill in the overlapped space.
- Add Time Of Arrival and Time Of Departure output grids.
- Add the option to have the uncertainty extend to only 50% of the length of the track, and then go straight. Because the warning will likely be modified by the time we get 50% through the duration. It will likely cut down on the false alarm area/time.



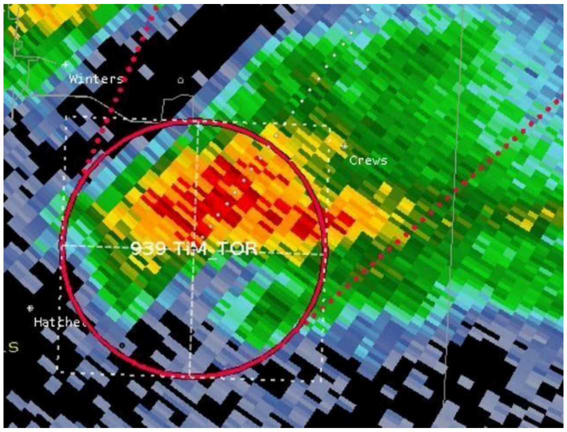
Object Creation/Editing:

- Advanced motion uncertainty track turn needs to properly age off. During warning updates, the new motion vector should come from the predicted turn.
- The centerline of the swath should not move if the direction uncertainty is made asymmetric.

- Make the bounding box slightly larger than the 2D threat area, to avoid vertex overlap.
- Use different shapes or color dots on the bounding box.
- When reshaping using the bounding box controls, it should only stretch one side unless you hold Shift.
- Display the "former" polygon when a warning is being updated, so that forecasters can see what they are changing their polygon from. This helps alleviate the potential slight addition to the workload when the user has to decide which update options and how much to extend the warning time, to avoid the windshield wiper and accordion effects.
- QuickWarn TOR: ○ Should use an 8 km (5 mi) circle instead of 5 km (which is too small). ○ Have two tools, one for Observed and one for Not Observed. ○ Right-click the center of circulation rather than dragging a dot.
    - Selection should be under drawing tools, and/or have its own Console button(s).
- Make the swath dashed outline a little more prominent. Thicker, or add black borders.
- Replicate the PHI Prototype's "buffer" functionality for objects, where the + and - buttons increase/decrease the size by 1km increments. Similar to dragging the corner of the bounding box, this new functionality would allow the forecaster to do this exactly one 1 km at a time.
- Add a segmented/freeform Line Tool like in WarnGen. It puts a buffer around the line to create the threat area. Have a way to change the buffer size, with separate forward and rearward buffers.
- Add a motionless polygon option. Either the polygon doesn't move at all (front or back) or expands radially.
- Selecting an object to Modify via the spatial display should be done with a double click (or the center or right mouse button) to prevent accidental selection when panning and zooming, or when toggling the legend.
- Investigate why the swath widening can be very large when using 3 or more upstream centroids to define the storm motion, and ways to "temper" the value.

Spatial Display (visual features):

- Remove Hazard Type and add Time to Expiration (minutes) to the TIM warning label.
- The current owner's warnings should be bright, while the others are dim.
- Add cross-hairs inside the Ellipse tool to help with aligning it to the center of circulation.



- New motion controls on swath decoration. Dragging the final point will change speed and direction. And ways to fan the swath wider to right and left to change direction uncertainty. These could be controls right on the swath, like the controls on the bounding box. All changes would reflect back on changes to initial conditions, and the upstream polygons and centroids would respond to these changes accordingly.

Console:

- For a SVR with a TOR PSBL tag, add a "t" suffix to the Object ID.
- The unused buttons and selections should be hidden for HS-TIM to avoid accidentally clicking on them.
- New console column features:
    - Replace Event ID with Object ID (which is just a number). ○ Add Persist ('false'=no color; 'true'=blue) ○ Add Time Since Last Modified (minutes), and use color shading. ○ Add Time Until Persist Off (minutes), and use color shading.
    - (HS-C) Add IBW intensity tags (cons=Considerable, CATA=Catastrophic, DEST=Destructive).
- When Modifying an object, the Draw option should default to Replace. Sometimes the user forgets to hit Replace and starts drawing a new object.
- After clicking Replace, it should already select either the polygon or ellipse drawing tool depending on what was previously used. And you have the option to switch to the other only if you want to.
- Add a Latest TIM Frame button in the Console time bar. It should be available when no objects are selected.
- Replace should be its own button in the Console.
- (HS-C) The computer should make a sound when warnings are expiring.
- (HS-C) Console scroll is opposite than usual (scrolls up to go down and scrolls down to go up).

- (HS-C) Add a checkbox on top of all warnings in the Console, that will allow one to Deselect/Select All (like in Gmail).

Hazard Information Dialog (HID):

- Move the Modify, Persist, and Latest TIM Frame selections into the HID Time Range area. Because they are time-related, they will remain "pinned" on top and won't scroll away when editing the warning IBWs and CTAs.
- The HID tabs and the expiration pop-ups should have the ETN instead of the Object ID.
- Hide the Start Time field and the End Time field in the HID Time Range area to save space. The only current HID Time Range field needed is Duration.
- Unchecking Preview Polygon should automatically reset to the Latest TIM Frame.
- Add Loop Preview Polygon button. Next to Preview Polygon, add a button that will play the "future loop" one time, dwell on the last frame for 1 second, and then return to the Latest TIM Frame. This avoids using the Selection Time on the console, which can be tricky.
- Hide the disabled Warn button (because this is only needed for HS-PHI).
- Advanced motion vector menu:
    - Put the asymmetric Right and Left label, slider, and buttons on one line each. o The +/- on the Advanced MV sliders should fine-tune and go by 1 instead of 5.
- Improved IBW selection layout:
    - Adjacent to CTAs.
    - SVR: Three rows: 1) Max Hail and Hail Source, 2) Max Wind and Wind Source, 3) Tornado Possible. Max Hail/Wind should default to blank, like in HS-C.
    - TOR: Two rows: 1) Intensity, Source, 2) Max Hail (this is on a second row since it can sometimes be missed).
    - Remove the word "Max" from the Hail/Wind labels.
- For Persisting warnings, the End Time should instead say "Persist" or "Until Further Notice" ("UFN").
- (HS-C) You won't change the Category or Type after you've issued a new warning. Therefore, for warning updates, this portion of the HID should be hidden.
- (HS-C) Cancel warning HID layout: the selections should be a list of radio buttons instead of a drop-down box.
- (HS-C) Add a pop-up GUI to select the appropriate selection for the CAN portion.

Warning Decision Discussion (WDD):

- Move WDD to the bottom of HID; hide the WDD history box on a new warning.
- WDD should be writable at any time without issuing a warning update, by any forecaster (e.g., warning coordinator), to speed up issuance. Perhaps trigger with a hotkey or rightclick menu option?

- Have a feature (e.g., a bot) to push WDD into NWSchat/Slack or Twitter each time a warning is issued/modified.
- Add canned statements for the WDD, some based on IBW tags.
- Add spell check and profanity check capability to WDD.

Miscellaneous:

- Configure the Product History (available from right click in Console) for HS-TIM.
- The SRM "from WarnGen" should come from HS-TIM.
- Manual county clipping. Need to determine how this would work with persisting warnings.

2022 Appendix B: Experiment logistics suggestions

- Include a recorded demo in the pre-operations information.
- Schedule:
  - Increase the training period to two hours.
  - For virtual experiments, shrink the lunch hour from 2 hours to perhaps 90 minutes.
- Training:
  - Throw in some more social science in the training videos. Mention how people are benefiting from TIM because it's improving equitable lead time. That should be a big motivational boost for NWS forecasters and gives more purpose to what we do because we are serving the public better.
  - Add more of the advanced features to the training. For example: ▪ TOR QuickWarn
    - How storm motion is calculated differently.
    - How Persist will turn off after a certain amount of time
    - How clicking a different object will change your HID even if already editing something else
  - Replace S1 with a more modern case, and include all the data so that procedures can be built for the week.
  - Do a better job in training to point out that the Replace feature is usually preferable to editing all the vertices and using the bounding box to manually adjust the object geometry.
- Cases:
  - More marginal cases.
  - The 2021 Iowa derecho event (was not tornadic).
- Create a troubleshooting FAQ and/or quick reference guide to help explain what to do when things go wrong (e.g., click this, select this, do this).
- For virtual experiments, determine a more-optimal method to have the desktop across two or more monitors on the virtual cloud instances.

2022 Appendix C. Human Factors Analysis Survey Results

Participants in the HWT

Six NWS forecasters participated in the 2022 HS-TIM HWT, with two forecasters every week. The NWS work experience ranges from 2 years to 24 years, and the average was 7.5 years (standard deviation 9.3). The average NWS warning experience ranges from 1 year to 24 years, with an average of 6.8 years (standard deviation of 9.7).

Mental Workload (NASA TLX) Survey

The NASA-TLX (Hart & Staveland, 1988; Hart, 2006) workload index is a questionnaire-based workload rating tool. The tool encompasses 6 aspects of workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. The raw scores of the mental workload range from 1 to 100, where 1 corresponds to an extremely low workload and 100 corresponds to an extremely high workload. The ratings were averaged from all the sessions for each of the 6 aspects of workload.

Table 1 shows the average ratings for the six sub-dimensions and the overall workload for three archived hazardous weather scenarios. The average workload for 2022 Hazard Services TIM HWT across all testing scenarios was 57.8 (out of 100, standard deviation 22.8). Each of these scenarios was chosen with an increasing level of difficulty (more storm coverage) throughout the course of the week.

Table 1. HS-TIM 2022 Testbed NASA-TLX Mental Workload Rating for Three Test Scenarios. Scenario 1 was a training scenario and was not included here.

	Scenario 2 Mean (std)	Scenario 3 Mean (std)	Scenario 4 Mean (std)
Mental Demand	64.2 (17.8)	64.5 (14.4)	61.8 (19.4)
Physical Demand	45.8 (30.3)	43.8 (19.4)	46.3 (15.9)
Temporal Demand	50.5 (13.2)	59.7 (26.6)	55.3 (22.0)
Effort	61.8 (23.3)	75.2 (19.0)	66.7 (27.8)
Performance	69.3 (11.0)	73.2 (16.5)	74.7 (12.9)
Frustration	29.8 (20.0)	47.0 (25.7)	51.5 (30.1)

Overall Workload	**53.6 (23.1)**	**60.6 (22.7)**	**59.4 (22.7)**

PSSUQ Usability Questionnaire

The Post Study System Usability Questionnaire (PSSUQ; Lewis, 2002) is a survey tool designed to evaluate the usability of a computer system. The tool is designed with 19 usability questions to assess 4 different areas of System Usefulness (Questions 1-8), Information Quality (Questions 915), Interface Quality (Questions 16-18), and Overall Usability (Questions 1-19). The rating ranges from 1 to 7, where 1 corresponds to a low level of usability, 7 to a high level of usability, and 4 corresponds to a neutral level of usability.

The PSSUQ questionnaire was filled out by the participants after they completed all test scenarios in the testbed. Table 2 shows the average responses for each of the 4 categories: Overall Usability, System Usability, Information Quality, and Interface Quality.

Table 2. Usability Ratings based on the PSSUQ for the two HS-TIM Experiments (7-point scale).

	2021 HS-TIM HWT	2022 HS-TIM HWT
System Usability	6.25	6.04
Information Quality	5.71	5.51
Interface Quality	5.78	5.67
Overall Usability	**5.97**	**5.79**

The overall usability was assessed at 5.79 (on a 7-point scale) for the 2022 HS-TIM HWT experiment, with system usability rated at 6.04, information quality at 5.51, and interface quality at 5.67. It is worth noting that the usability rating has slightly declined from the first to the second HS-TIM experiment. This was likely due to the software issues reported in the main body of this report such as the unexplained hazard locking issue and system sluggishness.

References

Hart, S. G., and L. E. Staveland, 1988: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, **52**, 139–183, https://doi.org/10.1016/S0166-4115(08)62386-9.

____, 2006: NASA-Task Load Index (NASA-TLX); 20 years later. Proc. *50th Annual Meeting, Human Factors and Ergonomics Society*, San Francisco, CA, 904–908, https://doi.org/10.1177/154193120605000909.

Lewis, J., 2002: Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction*, **14 (3-4)**, 463-488, https://doi.org/10.1080/10447318.2002.9669130.

2023 Appendix A: Software Functionality Suggestions

HID:

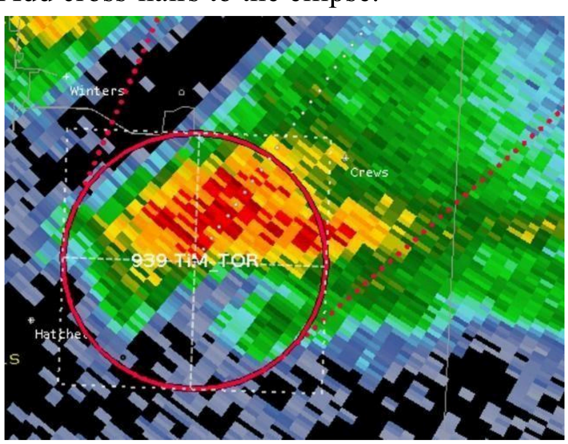
- Need a way to preview all polygons, not just the one you are working on, to ensure that there are no future gaps or overlaps between adjacent warnings.
    - Add Loop Preview Polygon button. Two options: play the current warning, or play all warnings.
- Replace Back-Building with an "All Clear" setting with three choices (off/half/full). This would work regardless of the Persist setting.
- Hide Swath Flare in the Advanced Motion Uncertainty submenu
- The Latest TIM Frame button should highlight when not on the latest frame
- Preview polygon should highlight areas where adjacent warnings will overlap in time
- Add forecasted speed changes to Advanced Motion Uncertainty
- Back-Building shouldn't auto-turn off with each update, because you may intend to keep it on.
- The tabs should show ETN and not Object ID.
- Move the Back-Building button next to the Persist button.
- The boxes to check for Persist should be a large button, and it turns on the check mark or changes color (blue to match the console).
- Move county clipping above CTAs.
- The warning preview should highlight the future overlaps from adjacent warnings.
- The HID should not populate until you set storm motion, to force it to be set. Unless storm motion is from a default (STI or PTT algorithm).

Advanced Motion Uncertainty:

- Motion direction uncertainty: left slider, right slider, and a "lock to both" button like in the Prototype. Then we wouldn't need a separate asymmetric direction uncertainty control.
    - Left should be on top because when storms move to the east, the left is on the "top" of the map.
- The motion turn should not be additive, but just the actual value.
- Hide the motion vector numbers in the drop-down with the advanced motion uncertainty functionality.
- The storm track turning tool default should start with a turn angle of 0°, whereby you can opt to increase it. The default start time and end time of the turn should be 1/3 and 2/3 of the warning duration. The values should age off and not reset to the new start time on a warning update.
- Rename "Swath Flare End" to "Limit Flare".
- Add a "tornado occlusion" button that is based on the hodograph from an hourly model sounding (Cameron Nixon deviant left).

Swath Tool:

- Improved swath decorations (match those in HS-C).
- Increase the size of the centroid dots.
- First guess at storm motion from STI or the new Prototype motion algorithm.
- Stretching the bounding box should only be symmetric when holding Ctrl.
- Add cross-hairs to the ellipse.



- Reword "Drag Me To Storm" to "Set Motion Vector" and update its properties with the following rules:
    - If you drag the current object without already dragging an upstream polygon, the label should remain the same.

  - The label will only be replaced with the ID, etc., once you've done an upstream drag.
- Add point, line, and dynamic line drawing tools
- Make the bounding box a little larger than the object, and/or use different color dots. [Re-opened old ticket]
- The order of operations should mimic HS-C and Tiny TIM: 1) launch tool, 2) select the warning type, 3a) select the tool type (point, line, dynamic line, polygon, freehand, ellipse), 3b) draw object (if applicable), 4) set motion vector, 5) make HID selections.
- Add toggleable time of arrival markings on the centerline.
- Add owner to the object labels.
- Fade the swath decorations when using the Replace geometry feature.
- When editing a warning, the other objects should disappear, but the warnings remain in the background.
- Add a "spin arrow" around the rotate dot.
- Configurable color and line schemes for the swath decorations.
- A way to further reduce the number of vertices on a freehand polygon to the standard 20 vertices. Perhaps hit a button to reduce by another 5 each time (20, 15, 10, minimum of 5).
- The bounding box should not remain rotated when updating a warning.

Polygon Editing:

- Add capabilities to directly edit the polygon swath and have it directly tied to objects so that they remain coupled. For example:
  - Interact with the centerline to change the direction and speed.
  - Controls to flare the swath on either side or both sides.
  - Controls to add a curve to the centerline.

Console:

- New columns:
  - Back-Building (or All Clear status: none, half, full).
  - Wind speed and hail size
- Color code the hazard type column (like in Tiny TIM).
- Replace "TiM.Severe" and "TiM.Tornado" with SVR and TOR.
- QuickWarn should be a separate button and not in the Tools dropdown.
- Options to show the user's/workstation's warnings (a.k.a. "my warnings"):
  - Add a control (e.g., button) to put "my warnings" on the top of the list.
  - Have two horizontal sections with "my" warnings in the top section and all other warnings in the bottom section. Each section has a separate vertical scroll bar.

  - Shade the rows containing "my" warnings (or vice versa).
- Remove the date from the date columns. Just show the times to save space.
- Tab warnings by WFO.

Cooldown/Accordion:

- Allow Cooldown and Accordion values per warning. The default should match what is set in the PHI Config Tool, but they could set it to 0 and require an "opt-in" for each warning.
- Make sure auto-county clipping and cooldown play well together to avoid in-out-in situations

Timer Alerts:

- Add a "snooze alerts" feature to the reminder timer alerts · Change the warning expiration timer alert from 5 to 10 minutes.
- The Timer Alert pop-ups should use ETN instead of Object ID.
- Allow the option for the Persist on too long (30 minutes) for it just to be a pop-up reminder instead of turning it off automatically. Or have a third warning:
  - 15 minutes: Reminder to update the warning.
  - 30 minutes: Reminder that Persist has been on for 30 minutes without an update.
  - 45 minutes: Automatically turn off Persist

Quick TOR/SVR:

- Add a QuickWarn SVR Tool. Use a rectangle (instead of an ellipse).
- Use the scroll wheel to change the ellipse size for a QuickWarn TOR

Experimental Warning Overlay:

- Add ETN (or "TO.W.0002") labels to the TIM Warning Overlay.
- Warnings should appear in the experimental warning overlay as soon as they are issued (instead of at the top of the minute)
- Advanced experimental warning overlays o For all products:
    - SVR and TOR should be separate layers that can be toggled on/off individually, as is done with the Obs > Warning overlay. SVR=yellow, TOR=red
    - DEST/CATA are drawn with thicker lines.
  - Three different products:
    - A default warning product should show the combined effect of cooldown and accordion all with no delineation of brighter and dimmer colors (what was seen during the experiment).

        - A warning product delineating Cooldown/Accordion as dimmed and bright layers (you've done this already but instead, this is a separate product).
- The accordion and cooldown are separate layers that can be toggled on/off individually.
        - A Snapshot (Threat Area) product (showing the object only - no swath).
    - Additionally, we could also delineate IBW this way:
        - SVR:
            - BASE: thin yellow
            - CONS: thick yellow
            - DEST: thick orange
        - TOR:
            - BASE: thin red
            - CONS: thick red
            - CATA: thick magenta

County Clipping:

- Once a county auto-clip is turned off, it should not turn back on until there is a warning update.
- Add auto county clipping based on the absolute area of the county, like in WarnGen.
- County clipping should include a 3rd option: Update Sides Only · Use a 5% default for automated county clipping (instead of 10%).
- Add manual county off/on toggling like in HS-C.
    - Add a warning pop-up when clicking off counties on the front of the warning that says that Persist will no longer have any effect.

WDD Tool:

- Perform a WDD update from a right-click on the event in the console
- WDD Tool storm ID should be warning type and ETN (e.g., TO.W.0002)

Miscellaneous:

- Swath Flare End should always be turned on for Persisted warnings, using the same amount of time as the duration. The flare should end at the end of the warning duration (e.g. 30 min). Can opt to turn it off.
- Create a WDD bot for NWSChat (and our Slack channel). Should include a link to the ETN's WDD history.
- Complete development of a national real-time capability using a federated registry
- When editing an object, the Latest Radar Frame functionally should switch under the hood to the Latest TIM Frame functionally

- Add a SCAN-like table to the warning overlay to show the list of warnings and IBW tags
- Have TIM for Significant Weather Advisories.
    - Ability to convert an SPS to an SVR or TOR, and vice versa, so that there is continuity from one warning type to another on the same storm.
- Add the ability to link SVR and TOR on the same storm, perhaps sharing the same motion vector.
- Add a failsafe to prevent warnings from going too fast and outrunning the storm. Perhaps use MRMS for this, and have a pop-up that says, "Object may be moving too fast".

Appendix B: Experiment logistics suggestions

Before the operations week:

- Include YouTube demos in the pre-operation information.
- Send out the job sheets in advance.
- Some of the forecasters suggested that we move more of the Monday activities (presentations, training) into the time before their travel (via supernumerary shifts) to free up time for cases. However, we tried this in the past and learned that some forecasters didn't have time before their travel to do this, thus ending up being "behind the curve".
- Add some virtual participation to practice on cases before bringing forecasters into the HWT. RAC does this – but it requires live support if there are issues.
- Provide the very detailed logistics email from the operations coordinator, which is normally sent out a few weeks before the start of the experiment, closer to the time of the initial selection email.
- The forecasters would like to be able to import procedures and color tables from their home offices before the experiment. However, we've tried that in the past, and because the HWT AWIPS systems are not customized like their home office systems, and because the HWT AWIPS system is designed to use for any CWA, the procedure frequently did not work. They also suggested some pre-built procedures and perspectives, however, when we've done this, they are typically not used.

Training:

- The first day's documentation helped with the basics; the more advanced material did need some further explanation. Might consider splitting training into basic and advanced?

Operations:

- Reconsider how to use the pre-recorded television simulcasts during the displaced realtime scenario. Many offices tend to keep the volume down on these or have one

dedicated person listening on headphones. This was particularly noteworthy for the OKC events (archive and real-time) where the local meteorologists tend to greatly exaggerate.
- I think the mesoanalyst would benefit from having more dedicated space on a PC.
- It may be interesting to go through a short scenario (or very similar scenarios) back-toback - once in WarnGen (or better yet, in HS – Convective), and once in HS-TIM, just to make those comparisons a bit easier.
- It was suggested to have the experiment facilitators sit closer to the forecasters so that there is more interaction. Normally, we have operations in the HWT Operations Area where this is more feasible. However, for this experiment, we were in the NSSL Development Laboratory, and there is less room for the facilitators to sit close. Also, during the previous Tiny TIM experiment, it was suggested to move the Mesoanalyst to an adjacent workstation instead of behind the forecasters. This took up space normally used for facilitators.

Equipment:
- Add an external mouse and keyboard on the mesoanalyst laptop.
- We typically use CAVE in Practice Mode if we need to Reset Events (clear out all warnings) when there is system instability. But the forecasters said the orange background was hindering. Therefore, we should consider not using CAVE in Practice Mode unless we have to Reset Events – we can bring up another CAVE in Practice Mode to do that.

Suggestions for new cases:
- Intermountain-west dry microbursts.
- Gust front outrunning echoes, how to deal with this?
- Back-Building event.
- Tropical rain band (quickly evolving tornado threat, deviant motions).

Summer pulse/quasi-stationary convection, to test WDM fatigue (whether or not to Persist), and dealing with polygon overlaps.

Appendix C. Human Factors Analysis Survey Results

<u>Participants in the HWT</u>

Nine NWS forecasters participated in the 2023 HS-TIM HWT, with three forecasters each week. The NWS work experience ranged from 2 years to 24 years, with an average of 7.5 years (standard deviation of 9.3). The average NWS warning experience ranges from 1 year to 24 years, with an average of 6.8 years (standard deviation of 9.7).

Mental Workload (NASA TLX) Survey

The NASA-TLX (Hart & Staveland, 1988; Hart, 2006) workload index is a questionnaire-based workload rating tool. The tool encompasses six aspects of workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. The raw scores for the mental workload range from 1 to 100, where 1 represents an extremely low workload and 100 represents an extremely high workload. The ratings were averaged from all sessions for each of the six aspects of workload.

Table 1 shows the average ratings for the six sub-dimensions and the overall workload for the three archived hazardous weather scenarios. Table 2 shows the ratings for the six real-time scenarios. Forecasters who solely performed WC/MesoA duties without issuing warnings were not included in the count. The average workload for the 2023 Hazard Services TIM HWT experiment across all testing scenarios was 42.3 (out of 100), with a standard deviation of 26.3. The workload varied on a scenario-by-scenario basis, with the highest workload observed during the two very challenging real-time events (RT2 OUN, RT5 BOU), and the lowest workload during one archive case and two marginally severe events (S7 GID, RT1 GID, RT3 OUN). Only one archive case was worked by all forecasters in 2022 and 2023, and the overall workload for S3 decreased from 2022 to 2023 (Table 3).

	Scenario 3	Scenario 7	Scenario 6
	Mean (std)		
Mental Demand	65.8 (17.7)	50.0 (12.7)	62.9 (18.5)
Physical Demand	25.0 (10.5)	22.0 (16.4)	40.0 (17.6)
Temporal Demand	50.0 (33.9)	20.0 (14.6)	54.3 (23.7)
Effort	21.7 (12.9)	20.0 (13.7)	42.9 (22.5)
Frustration	65.8 (22.0)	40.0 (23.7)	57.1 (22.3)
Performance	45.8 (22.2)	22.0 (15.2)	40.0 (32.7)
Overall Workload	**45.6 (26.5)**	**29.0 (28.1)**	**49.5 (20.8)**

Table 1. 2023 Hazard Services TIM NASA-TLX Mental Workload Ratings for the three archive scenarios. Scenario 1, a training scenario, is not included in this table. The order of the archive scenarios in the table corresponds to the order in which each scenario was worked throughout the week.

	RT1 (GID)	RT2 (OUN)	RT3 (OUN)	RT4 (EAX)	RT5 (BOU)	RT6 (OUN)
	Mean (std)					
Mental Demand	22.5 (10.6)	88.3 (16.1)	26.7 (17.6)	47.5 (10.6)	90.0 (7.1)	50.0 (42.4)
Physical Demand	10.0 (7.1)	76.7 (10.4)	21.7 (14.4)	30.0 (0.0)	62.5 (17.7)	37.5 (38.9)
Temporal Demand	10.0 (7.1)	81.7 (15.3)	16.7 (10.4)	22.5 (10.6)	50.0 (21.2)	55.0 (49.5)
Effort	10.0 (7.1)	26.7 (10.4)	16.7 (11.5)	40.0 (14.1)	27.5 (3.5)	25.0 (0.0)
Frustration	32.5 (3.5)	78.3 (29.3)	28.3 (16.1)	30.0 (0.0)	65.0 (7.1)	37.5 (31.8)
Performance	15.0 (7.1)	48.3 (25.2)	20.0 (8.7)	52.5 (3.5)	85.0 (7.1)	52.5 (24.7)
Overall Workload	**17.3 (15.9)**	**66.7 (23.0)**	**21.7 (31.0)**	**37.1 (11.0)**	**63.3 (28.6)**	**42.9 (32.1)**

Table 2. 2023 Hazard Services TIM NASA-TLX Mental Workload Ratings for the six real-time (RT) scenarios.

	Scenario 3 (2022)	Scenario 3 (2023)
	Mean (std)	
Mental Demand	64.5 (14.4)	65.8 (17.7)
Physical Demand	43.8 (19.4)	25.0 (10.5)
Temporal Demand	59.7 (26.6)	50.0 (33.9)
Effort	75.2 (19.0)	21.7 (12.9)
Frustration	73.2 (16.5)	65.8 (22.0)
Performance	47.0 (25.7)	45.8 (22.2)
Overall Workload	**60.6 (22.7)**	**45.6 (26.5)**

Table 3. 2023 Hazard Services TIM NASA-TLX Mental Workload Ratings for archive Scenario Three (S3), the March 5, 2022, long-tracked tornado event.

<u>PSSUQ Usability Questionnaire</u>

The Post Study System Usability Questionnaire (PSSUQ; Lewis, 2002) is a survey tool designed to evaluate the usability of a computer system. The tool consists of 19 usability questions that assess four different areas: System Usefulness (Questions 1–8), Information Quality (Questions 9–15), Interface Quality (Questions 16–18), and Overall Usability (Questions 1–19). The rating scale ranges from 1 to 7, where 1 corresponds to a low level of usability, 7 to a high level of usability, and 4 represents a neutral level of usability.

The participants completed the PSSUQ questionnaire after finishing all test scenarios in the testbed. Table 4 displays the average responses for each of the four categories: System Usability, Information Quality, Interface Quality, and Overall Usability across three experiment years.

	2021 HS-TIM HWT	2022 HS-TIM HWT	2023 HS-TIM HWT
System Usability	6.25	6.04	5.11
Information Quality	5.69	5.51	5.23
Interface Quality	6.00	5.67	5.48
Overall Usability	**5.97**	**5.79**	**5.22**

Table 4. PSSUQ Usability Ratings for the three Hazard Service TIM experiments (7-point scale).

The overall usability was assessed at 5.22 (on a 7-point scale) for the 2023 HS-TIM HWT experiment, with system usability rated at 5.11, information quality at 5.23, and interface quality at 5.48. It is worth noting that the usability rating has decreased over time, with a slightly-more significant decrease in 2023. This is likely due to the higher level of software issues, the longer duration of case time per week in an in-person experiment compared to a virtual experiment, and the addition of real-time cases.

References

Hart, S. G., and L. E. Staveland, 1988: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, **52**, 139–183, https://doi.org/10.1016/S0166-4115(08)62386-9.

____, 2006: NASA-Task Load Index (NASA-TLX); 20 years later. Proc. *50th Annual Meeting, Human Factors and Ergonomics Society*, San Francisco, CA, 904–908, https://doi.org/10.1177/154193120605000909.

Lewis, J., 2002: Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction*, **14 (3-4)**, 463-488, https://doi.org/10.1080/10447318.2002.9669130.

End-User Decisions Over Time

Summary by Holly Obermeier and Kodi Berry

Broadcast meteorologists and emergency managers serve as primary communicators of hazardous weather information to the public. As experimental products evolve, researchers continue to prioritize user research techniques by integrating broadcasters and emergency managers into the development process. While research has been done to understand how broadcasters and emergency managers could use experimental products *during* a severe weather event, less is known about their forecasting and communication processes in the days and hours *leading up* to an event. In addition, researchers have questioned if experimental products could potentially fill any informational gaps that might exist during these timeframes.

To address these time periods, researchers conducted a two-week virtual experiment in NOAA's Hazardous Weather Testbed in the summer of 2022. A total of 10 broadcast meteorologists and 15 emergency managers participated in two-day exercises, during which researchers implemented timeline and ranking methods using a modified card sorting technique. Participants completed their timelines by ranking cards with the types of weather information and NWS products they currently use, while complementing this information with their decision types. If any gaps in their informational needs existed, researchers asked participants to note the gap. Once timelines were completed, participants were introduced to developmental products including Warn-on-Forecast System guidance, storm-scale probabilistic hazard information, Threats-in-Motion, dynamically-evolving watches, and timing guidance from the Storm Prediction Center. After seeing these products, participants were asked to add cards with this experimental information into their timeline, if they felt the information was useful to them. Specifically, researchers hoped to assess if the new products filled any of the information gaps that participants had initially identified. With the completed timelines, researchers conducted post-activity focus groups. A brief summary of results with broadcast meteorologists can be found at https://inside.nssl.noaa.gov/kphi/2023/01/decisions-over-time/.

Project Contacts

Katie Wilson (CIWRO/NSSL)	Holly Obermeier (CIWRO/NSSL)
Kenzie Krocak (OU IPPRA/SPC)	Kodi Berry (NSSL)
Kim Klockow-McClain (NSSL)	Joseph Trujillo (CIWRO/NSSL)
David Hogg (CIWRO/NSSL)	Sean Ernst (OU SoM/IPPRA)
Taylor DeWinter (CIWRO/NSSL)	

Hazardous Weather Testbed Convective Outlook Innovations Experiment

Summary Report

Makenzie Krocak[1,2,3], Anna Wanless[1], Abby Bitterman[1], Burkely Gallo[2,3], Sean Ernst[1,2,3], and Joe Ripberger[1]

[1]University of Oklahoma Institute for Public Policy Research and Analysis
[2]Cooperative Institute for Severe and High-Impact Weather Research and Operations
[3]NOAA/NWS Storm Prediction Center

The Storm Prediction Center is currently evaluating new innovations to provide additional information in the convective outlook, one of their flagship products. These innovations include converting the existing discrete probability lines (e.g., 2%, 5%, 10%, etc. for tornadoes) into a continuous probability field and adding conditional intensity information. This intensity information is envisioned as a complement to the coverage probabilities; highlighting when there is a heightened threat of significant severe weather reports *given that a storm forms*. Although these innovations have been studied previously from a meteorological perspective, less work has been done to understand if and how these innovations would change communication of severe weather events at the WFO level, partner (e.g., emergency managers, EMs) workflow, and how members of the public would interpret these forecasts.

This report summarizes the activities and findings of a virtual Hazardous Weather Testbed (HWT) experiment designed to study how NWS forecasters, EMs, and members of the public might interpret and use these two new innovations (continuous probabilities and conditional intensity information). The experiment ran for two weeks for three days each week. Participants included 18 NWS forecasters (primarily from WFOs), seven EMs, and seven public evaluators. Forecasters participated all three days, while EMs participated on day 2 and public evaluators participated on day 3. This experiment was the first to include members of the public in official HWT activities. Both the researchers and the forecasters found the feedback from even a limited number of public evaluators extremely valuable, and the researchers recommend that future experiments consider including public evaluators.

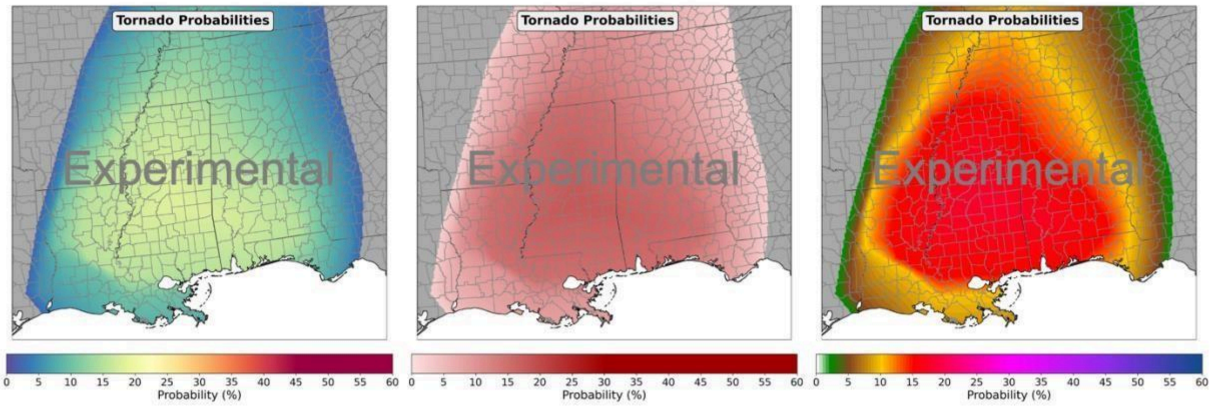


Figure 1: Examples of the continuous probability visualizations used in the Convective Outlook Innovations experiment. These examples are for tornado probabilities for the "high end" event in the Southeast US. Forecasters were also given continuous probabilities for hail and wind.

On day 1 of the experiment, forecasters were introduced to the new innovations and then they participated in focus groups on the probability and the intensity information. The focus groups were designed to elicit information about the benefits, challenges, and workload concerns that NWS WFO forecasters face. Prototype visualizations of the continuous probabilities are shown in Figure 1. Forecasters were given multiple color table options. Prototype visualizations for the conditional intensity forecasts are shown in Figure 2. Researchers chose to display the conditional intensity information separately from the probabilities for this experiment to test forecaster and user opinions about having multiple graphics instead of a single graphic with probability and conditional intensity information overlaid. However, the topic of single graphics vs. multiple graphics for display was discussed in focus groups.

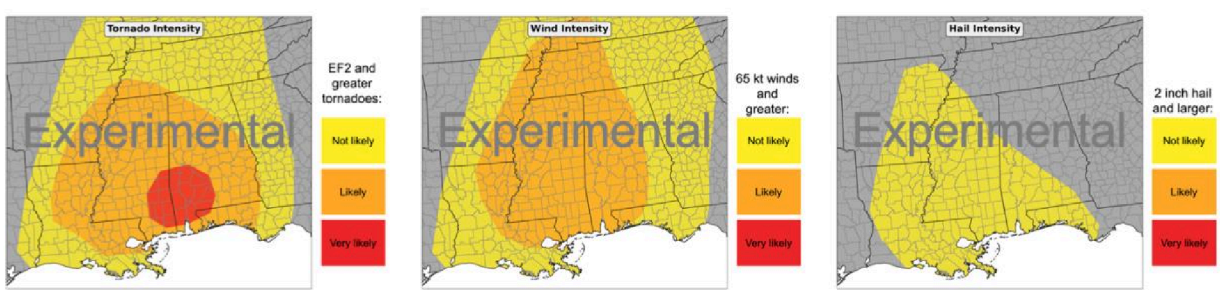


Figure 2: Examples of the conditional intensity visualizations used in the Convective Outlook Innovations Experiment. These examples are from the "high end" event in the Southeast US.

On day 2 of the experiment, EMs were introduced to the facilitators and forecasters, then forecasters spent the morning creating simulated communication products (in the form of a partner briefing and social media posts) for two different severe weather events. Half of the forecasters used information for a "medium end" event in Iowa, while the other half used information from a "high end" event in the Southeast US. Forecasters were given the current categorical outlook, the continuous probability fields (Figure 1), the conditional intensity forecasts (Figure 2), and a simulated convective outlook discussion created by a current SPC forecaster using the new innovations.

While forecasters were creating their communication and briefing products, EMs participated in a focus group on their current use of the SPC convective outlook and then were introduced to the new innovations. In the afternoon, forecasters presented the simulated briefings to EMs and then EMs had the opportunity to ask questions. To close the day, EMs participated in a focus group designed to get feedback on their interpretation, use, and the benefits and challenges they saw with the new innovations. Forecasters observed this focus group and were able to propose questions to facilitators to see what feedback the EMs had on the visualization and communication of the new innovations.

Day 3 of the experiment included public evaluators and forecasters. After introductions, the facilitators walked the public evaluators through the severe weather scenario, some information about the location, and expectations for the activity. Then the public evaluators were shown the social media posts created by the forecasters the day before and asked to provide feedback. This feedback included interpretation, possible response actions, and suggestions when something did not make sense. Forecasters observed this activity and asked questions to spark discussion.

At the end of the experiment, each group (forecasters, EMs, and public evaluators) were given an opportunity to share feedback through an anonymous exit survey. Everyone was asked if they thought this experiment was useful to them personally, with response options including "Definitely not", "Probably not", "Not sure", "Probably yes", and "Definitely yes". All of the forecaster participants answered "Probably" or "Definitely yes", all of the EM participants answered "Definitely yes", and 6 of the 7 members of the public answered "Probably" or "Definitely yes". Forecaster participants thought the experiment was well structured, facilitated well, and most mentioned that they especially enjoyed the interactions with EMs and members of the public. Similarly, EM participants thought that the experience was very useful for their severe weather operations and had little feedback in the way of improvements. Public evaluators said they learned a lot and thought it was a safe environment to share feedback.

Opportunities for improvement mentioned in the exit surveys mostly consisted of having more time to interact with each other and more EM and public participants. Another specific comment was to have EM and public evaluators provide initial thoughts individually and then head into discussion such that other participant comments do not bias personal feedback. Finally, all participants would recommend a friend (public) or colleague (all groups) participate in an experiment like this one.

Forecaster feedback on continuous probabilities

To begin this activity, forecasters were asked to fill out a short survey exploring different visualizations of continuous probabilities (see Figure 3 for examples of the visualizations). First, respondents saw a single color table (i.e., one row in Figure 3) and were asked what the risk was for people in location A. Response options included a 5-point scale with labels “No risk”, “Low risk”, “Medium risk”, “High risk”, and “Extreme risk”. Although differences in risk rating were modest, the SPC color table (top row of Figure 3) had the highest risk ratings (1 “Extreme” and 4 “High”) of the three options. The spectral color table (middle row of Figure 3) had the second highest risk ratings (5 “High” and 1 “Medium”) and the monochromatic color table (bottom row of Figure 3) had the lowest risk ratings (4 “High” and 3 “Medium”). These results align well with results from the HWT Spring Forecasting Experiment, which also had the highest risk ratings for the SPC color table, followed by the spectral color table and then the monochromatic color table.

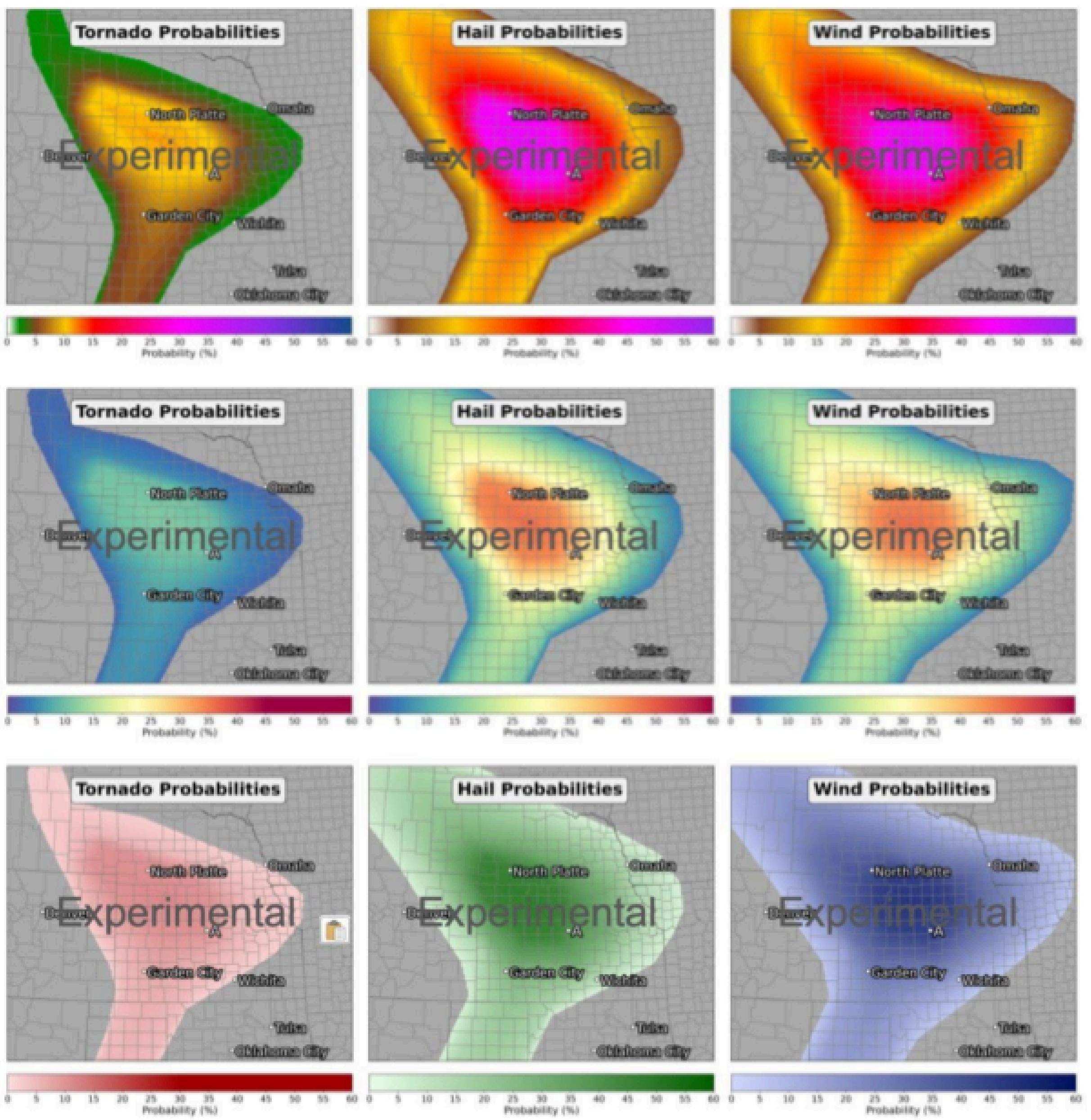


Figure 3: Continuous probability visualizations included in the introduction survey for forecasters. Respondents initially saw one of the color tables (i.e., one row of the figure) and were then shown all three to compare side-by-side.

Next, forecasters were shown all three visualizations side-by-side to compare them. In general, respondents thought that the SPC color table (row 1 in Figure 3) was attention-grabbing, which they saw as a good thing. Many also commented that small gradients in probability were easier to identify with the SPC color table or the spectral color table (rows 1 and 2, respectively in Figure 3). Another benefit participants saw with the spectral color table was the "cool" colors to "warm" colors with red being the highest, or "worst", color. However, many forecaster participants mentioned blue being an odd color to represent severe weather, especially since it is commonly used for winter weather scenarios. Participants also liked the idea of a single, consistent color table across all three hazards, like what was shown in the spectral color table. Finally, very few respondents liked the monochromatic color table (row 3 in Figure 3), saying it "washed out" or "faded" the probabilities too much, and noting that it is hard to identify exactly what the probability is for individual locations.

During the discussion period of this activity, most forecaster participants said that they currently do not directly communicate the hazard probabilities to partners or the public very often. Instead, those probabilities are used as event context and situational awareness. When they do pass on the probabilities, it is to highlight the areas of greatest threat during significant events, mostly internally or to sophisticated partners. Part of the concern about passing on the probabilities is that the participating forecasters thought their audiences may not understand the significance of, for example, a 2% tornado risk. Many forecaster participants said that the continuous probabilities would likely not change these operational procedures, instead they would likely be used in a similar manner to how the discrete probabilities are used now; mostly internally. Overall, the opinion of the product was somewhat split, with some saying they would use it more than the current product, highlighting which hazard was of highest risk, while others wouldn't. Many participants stated that training for WFO forecasters from the SPC would help contextualize and provide strategies for communicating the probabilities (both discrete and continuous) with partners.

Beyond general use, most of the forecaster participants agreed that the continuous probability product would help those located on or near the lines of the current discrete probabilities. The color gradient may help audiences contextualize their risk a bit better than the abrupt change between categories. The gradient may also help add context to the product when zoomed into a CWA or state, instead of just seeing one solid color (if the entire CWA is in the same categorial risk, for example). Similarly, a continuous field would also avoid splitting jurisdictions so definitively, with the eastern half being yellow (10%) and the western half being red (15%), for example. However, there were some concerns that eliminating the definitive borders would overcomplicate the product and that emergency managers may have plans based on the distinct probabilities and categories. One suggestion to alleviate these concerns was to overlay the stairstep percentage contours on top of the continuous probability field.

On appearance, there was strong consensus that the monochromatic color scheme was problematic. Most forecaster participants did not like how "washed out" the probabilities looked with the single color. There was also concern about users not being able to tell exactly what the probability was for their location. The other color schemes were better received, with some forecaster participants leaning towards the SPC color scale since that is what they are accustomed to. Most forecaster participants pushed for consistency between hazards, using the same color and scale for tornado, wind, and hail probabilities. This sentiment was also echoed by EM and public participants throughout the experiment.

Overall, participating forecasters' opinions of the continuous probability graphic were mixed. While no forecaster was against changing the probabilities to a continuous field, there was concern that these new products may overwhelm partners, the public, and forecasters. Since this information already exists, many participating forecasters thought it would be a reasonable change after making some adjustments, like adding discrete contours over the gradient, picking a consistent color scheme, and using a consistent scale across all three hazards. Further feedback should be solicited after these adjustments are made to the prototype visualizations.

Forecaster feedback on conditional intensity forecasts

One of the first themes to emerge during these conversations was related to the separation of coverage vs. intensity information. Most of these responses were mixed, with forecaster participants from areas that see less severe weather indicating it was more important to highlight coverage than intensity, but others said the opposite (especially those in areas that get severe weather more often). Many forecaster participants noted that the importance of both attributes often dependent on geography, recency of impactful severe weather, if there is a special event occurring (e.g., concert/festival) and the anticipated hazard. Overall, groups concluded that both coverage and intensity information were important to include in the message.

There were many questions from forecaster participants about how to interpret and message this information, since the conditionality of this product is different from other products. The verbiage used on the experimental graphic in the legend was "Not Likely", "Likely", and "Very Likely" (Figure 2). Many forecaster participants disliked this verbiage. Some thought it suggested forecaster confidence rather than conditional intensity (e.g., similar to current precipitation forecasts), and others thought that "not likely" indicated too low of a threat (since this area is still in the tornado probabilities). Many participants liked the idea of a "Low/Medium/High" legend, indicating the likelihood of seeing significant reports, but they also recognized that this might be cause confusion with the "High Risk" category in the categorical outlook.

All forecaster participants emphasized the need for clear guidance on how to use and communicate this product, especially given the conditional nature of it. Similarly, internal consistency was emphasized, with many forecaster participants highlighting a need for forecaster training, consistent messaging, and interpretation. Emphasizing the *conditionality* of this product to all users (forecasters, EMs, broadcasters, and the public) was one of the top concerns raised. Some forecasters indicated that they would likely emphasize this information to partners rather than the public, particularly in the context of IDSS briefings. These results in aggregate highlight the need for formal training on these new products for *all* forecasters.

Participants were also asked whether they preferred the conditional intensity and coverage probability information to be displayed on one graphic, or whether they wanted that information to be displayed on two separate graphics. While many of the participants indicated that the conditional intensity information should be provided with context, most also indicated that they wanted the conditional intensity and coverage probability information to be provided on separate graphics. This division allows them to tailor their message more easily than a combined graphic would. Having both sets of information on the same graphic also results in a lot of information on a single graphic, more information than might be relevant to some people's needs. Forecaster participants noted that using both sets of information would be most useful when the area of highest coverage probabilities and highest conditional intensity do not overlap, which would distinguish the information from the current hatched area paradigm. Collectively, this feedback suggests that conditional intensity might be a product that forecasters use sparingly, when it provides a clear storyline above and beyond what the coverage probabilities display.

Emergency manager use of the current convective outlook

Emergency managers (EMs) joined the experiment on day 2. Week one included four EMs and week two included three EMs. Every EM participant said that the outlook plays a role in their weather operations, but the extent to which it is used to make decisions varied. Some EMs said that they use it extensively, while others said they use it as more of an awareness tool. Many said that certain levels of the outlook dictated how often (or if) they would check the future forecasts. While Enhanced (level 3 of 5) was the most common level at which their weather awareness/operations changed, some said that they would not send any additional communications to stakeholders until Moderate (level 4 of 5) was in the forecast. All EM participants also said that the current categorical labelling system was familiar to them, but not intuitive to others. They reported that they often have to explain what the labels mean (and the order) or that they use level numbers (1 through 5) in place of the labels.

To gather information about upcoming severe weather, the participating EMs indicated that they look at the outlook a couple times (or more) per week. They mostly do this through Twitter, with multiple EMs noting that the outlook shows up faster on the social media website than on the official SPC website. They will then navigate to the SPC website through the link on Twitter. Those who do not go to Twitter said they usually get the information from their local WFO. Once they see the official outlook, most EM participants said that they wait for a briefing from the WFO or contact the office directly with specific questions (especially if the WFO does not do a full email or video briefing). For example, all participating EMs noted that timing information is critical to their operations but is not included in the SPC products, so they rely on their WFO for that information.

Although the participating EMs said that the outlook contains most of the information they need, they all said that an interactive map display would be enormously helpful. Capabilities to zoom in, see local landmarks, and layer the map with other data (like traffic) were all mentioned as capabilities that would be useful. Their number one request was a zoom function. This was partially due to the desire to identify exactly where the lines delineating the categories fall. While the EM participants know that storms do not care about exact lines, being in or near a certain level means something to them. One EM even said that while they do not activate their Emergency Operations Center until an Enhanced risk is issued, they would probably open it with limited staff if they were close to the Slight/Enhanced line. They are inferring an increased risk near the line, and less risk farther from the line.

Emergency manager feedback on Convective Outlook Innovations

Emergency managers are largely supportive of most prototype weather products, and this experiment was no different. However, there were places where participating EMs had suggestions to improve the visualization of the products. Specific to the continuous probabilities, most EM participants did not like the green-brown-yellow transition, noting that brown doesn't make sense in the color order. Like forecaster feedback, EM participants exhibited a preference for using the same color and probability scales for hail, wind, and tornadoes.

While EM participants said that they use the probabilities for their own context, they expressed some hesitation about showing probabilistic information to the public – they want simplified, attention grabbing information for public purposes. Some participants made comments about probabilities being difficult to interpret for members of the public, specifically that they are worried members of the public would "blow off" a 10% because it is so low. When shown the new continuous probability field, many EMs thought it looked too "muddy", with the fuzzy field being more difficult to interpret than the discrete probability areas. While they appreciate the idea that the field is continuous, they thought that having distinct notation of certain lines (like every 10% interval) would be useful.

EM participants were generally enthusiastic about the intensity information. Many said they would share it with high-level users and decision makers to drive home the possibility of highend events. However, there were mixed feelings about the verbiage used in the experimental graphics. Most EM participants did not like "Not likely", and preferred edited versions using other words.

There were also split feelings about whether to use simplified hazard descriptions (like strong tornado) or specific descriptions (like EF2+ tornado). Like forecaster participants, the EM participants generally preferred the probability and intensity information presented on separate graphics, but side-by-side. Finally, one question that was brought up regularly was what to tell people to do differently in these different areas. While EMs generally want to highlight the areas of greatest concern, all EM participants said that people in the yellow/orange area (Figure 2) should be preparing in the same way as the red area because "all tornadoes are serious."

There were also a few general themes that all the EM participants mentioned as being useful during severe weather operations. First, they consistently requested an interactive map that NWS products are displayed on with the ability to zoom in and out, add town and city names, add county boundaries, add interstates, etc. This idea was mentioned during both experiment weeks with all EM participants in agreement that this type of tool would be enormously helpful. Furthermore, they said that an interactive map with zoom capabilities would become more vital if or when SPC moves to more granular forecast information. Second, participating EMs were very interested in more timing information, both for the onset of storms as well as when storms will end. This information would be particularly useful as it relates to watch and warning issuance – many EMs said that like to know when watches will be issued for them and when areas will be removed from watches. This information is very important to their response process and operations. Finally, EM participants consistently reiterated the importance of direct communication with the NWS in addition to graphics and products.

Public evaluator feedback on Convective Outlook Innovations

During the final day of the experiment, public evaluators were asked to give feedback on social media posts created by forecasters using the prototype convective outlook innovations. In general, forecaster participants showed a strong preference toward using the SPC's categorical outlook with members of the public. It was the first, and sometimes only graphic the forecasters used in their posts. Public participants also seemed to have the most positive reaction to the categorical outlook. Some examples of how forecasters used this graphic in combination with other information are provided in Figures 4 and 5. Public evaluators generally thought the other information (like timing, impacts, actions to take, etc.) combined with the categorical graphic provided all the information they needed to make decisions. Furthermore, this small sample of evaluators consistently emphasized the need for a "one stop shop" post that was straightforward and easy to interpret. Most people said that do not spend a lot of time reading posts, so one short message with all the important information they need is helpful.

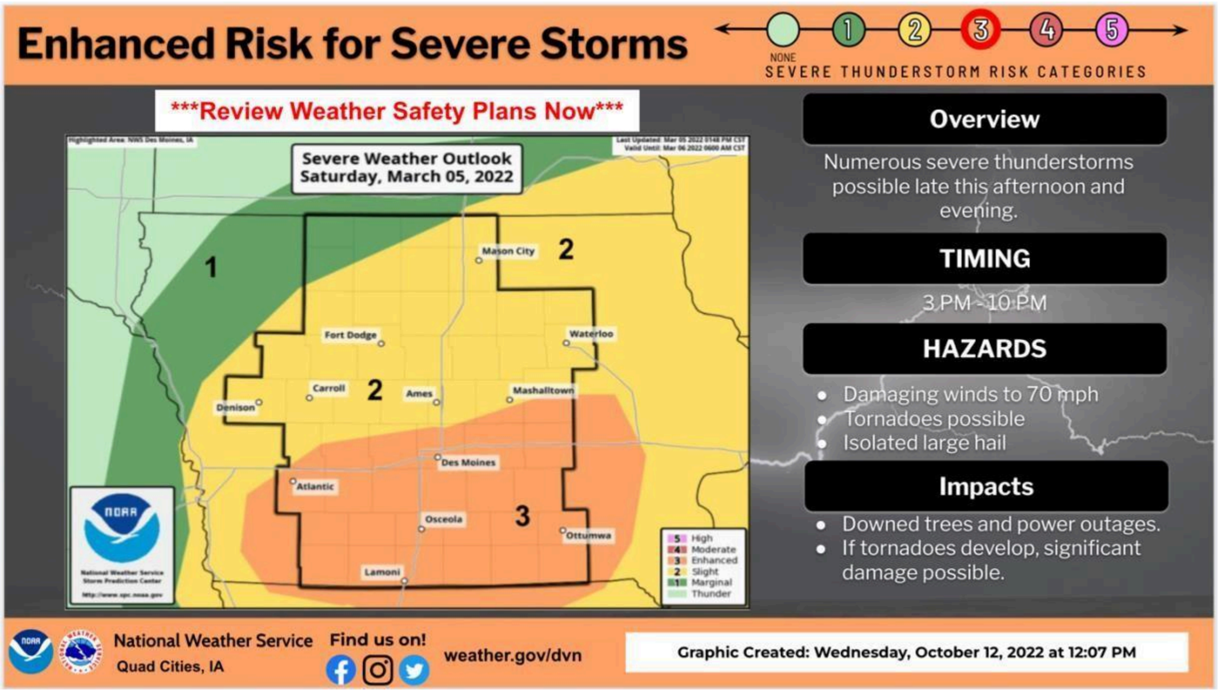


Figure 4: An example social media post created by forecasters for public evaluators in the experiment using the categorical outlook for the Iowa event.

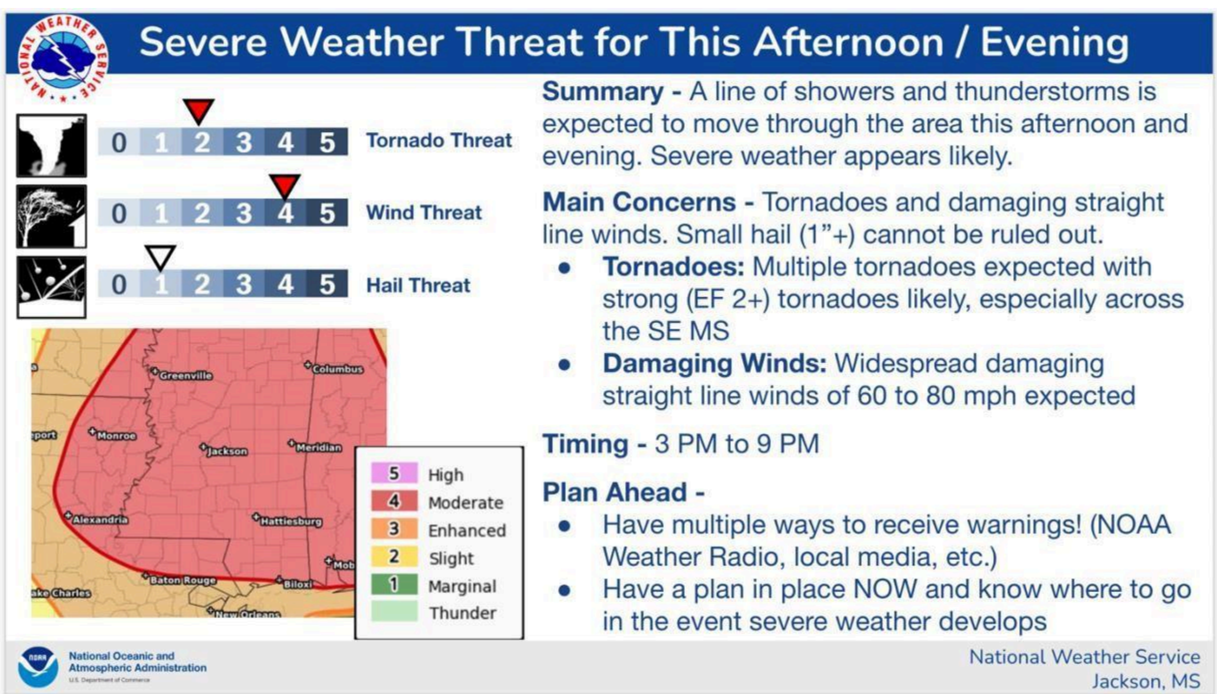


Figure 5: An example social media post created by forecasters for public evaluators in the experiment using the categorical outlook for the Southeast US event.

Of the two innovations, the public evaluators had the strongest preference for the conditional intensity graphics. They generally liked the intensity visualization and said it would get their attention and spur more of a reaction if they were in the higher areas (like the orange and red areas). However, they said that some details still need to be ironed out. Like the forecaster and EM participants, the public evaluators did not like the terms "Not likely/Likely". Some of the evaluators went as far as to say that "Not likely" would indicate that they would not do anything if they were in the yellow area. Evaluators were more enthusiastic about "Low/Medium/High" verbiage.

One interesting comment that came from the public evaluators and was echoed in conversations with EM participants was the question of what should be done differently with strong or not strong tornadoes. One evaluator said, "I don't like weak vs strong tornadoes… aren't all tornadoes strong?" This sentiment was also brought up by some of the EM participants, who were concerned that people in the yellow area may think they are in the clear. Further work needs to be done to study the potential decrease in risk perceptions in the "lower end" areas of the conditional intensity forecasts, and how forecasters can communicate the appropriate risk for the entire forecast area.

During this experiment, the continuous probabilities were used far less with public evaluators than the conditional intensity graphics. Evaluators who did see probability graphics did not have much feedback for them. The feedback that was given expressed lower levels of concern since the probabilities were low, and a desire for context. One forecaster explained to evaluators that the 15% tornado probabilities only happen a few times a year and the evaluators said that context should be included in the graphic. Otherwise, most public evaluators did not have many comments on the color scales except to say that the blue color scale reminded them of winter weather.

Summary

This experiment explored forecaster, EM, and public communication, interpretation, and possible use of two innovations to the SPC convective outlook: continuous probabilities and conditional intensity information. It was also the first experiment to include members of the public, which sparked useful discussions between facilitators, forecasters, and the public evaluators. A summary of the most common pieces of feedback is provided below.

Forecaster feedback

- Forecaster participants said they use but rarely distribute SPC probability graphics, and they did not see that changing with the continuous probabilities.
- The continuous probabilities could help add context when a location is on the line in the current discrete probability product.

- It may be important to include discrete probability lines on the continuous probability graphic to delineate differences in colors and help with numeric interpretation of the probabilities.
- Participants thought the monochromatic color scheme does not do as good of a job highlighting the probability gradient compared to other color schemes.
- Many forecaster participants said the same color scheme and scale should be used for all three hazards.
- Conditional intensity information is best displayed separately from the probabilities (though side-by-side is likely useful).
- Training for forecasters is needed for both products, but particularly related to the use and communication of the conditional intensity information. There is concern about how to explain the conditionality of the intensity information.

EM feedback

- EM participants said the outlook is used regularly for their severe weather operations.
- Many participating EMs get the product from social media channels and then go to their local office for more information.
- Participants said the categorical labels are not intuitive if you do not use them very often. EMs participants often explain the order, or they use the level numbers.
- Participants said the same color scheme and scale should be used for all three hazards.
- Putting discrete lines (like 10%, 20%, etc.) on top of the continuous field would help with quick interpretation and contextualization of the probabilities. EM participants said the continuous field is too muddy without some lines on it.
- EM participants said conditional intensity information is best displayed separately from the probabilities.
- EM participants strongly desire an interactive SPC product page with the ability to zoom in and add geographic markers (like cities, counties, roads, etc.) This was seen as vital to use the continuous probabilities.

Public feedback

- A comprehensive graphic with the categorical outlook and additional "What, When, Where, What should I do" information was the most popular with participating public evaluators.
- The intensity graphic captured their attention and may increase the likelihood that they would stay weather aware if they were in the higher areas.
- Public evaluators said the probability graphics were less useful, requiring too much time to interpret. Context about how often the hazard probabilities occur would be helpful.

Watch-to-Warning Experiment

Summary by Eric Loken and Miranda Silcott

Overview

Forecasting and communicating about the severe weather threat between watches and warnings can be challenging for multiple reasons: at these lead times, the responsibilities of national center and local office forecasters can overlap, information from numerical weather prediction (NWP) models and observations must be effectively synthesized, and knowing how to message the threat can be difficult. Nevertheless, better prediction and communication of severe weather between watches and warnings has the potential to improve the timeliness and usefulness of severe weather warnings, contributing to a more Weather Ready Nation (Rothfusz et al. 2018).

To that end, the 2023 Watch-to-Warning Experiment (W2W) sought to investigate: 1) how Storm Prediction Center (SPC) and local weather forecasting office (WFO) forecasters communicate amongst themselves and to the public about the severe weather threat between watches and warnings, and 2) how forecasters use new and existing tools to help predict severe weather at these times. Building on the first W2W held in 2021 (Wilson et al. 2024), the 2023 W2W allowed participants to issue rapidly-updating, probabilistic products as envisioned by the Forecasting a Continuum of Environmental Threats (FACETS; Rothfusz et al. 2018) paradigm, and it provided participants with a suite of new and existing tools for evaluation in addition to the Warn-On-Forecast System (WoFS; e.g., Jones et al. 2020). The tools evaluated in the 2023 W2W included: the National Severe Storm Laboratory's (NSSL's) prototype Probabilistic Hazard Information (PHI) Tool; Tornado Probability Algorithm (TORP; Sandmael et al. 2023); WoFS; ProbSevere Version 2 (PS2; Cintineo et al. 2020); and a new machine learning algorithm, called WoFS-PHI, designed to predict the likelihood of severe weather at different spatial radii in 30-minute windows out to 3 hours of lead time. During each simulated case, forecasters used these tools to issue 2-h watches, mesoscale discussions, warnings, PHI plumes, and impact-based-decision support services (IDSS) graphics. Surveys and formal discussions after each simulated case captured forecasters' experiences with the tools as well as their thoughts on the overall watch-to-warning forecasting and communication process.

Experiment Details and Results

3 WFO and 1-2 SPC forecasters participated in the four displaced real-time forecasting case simulations per week. These simulations increased in complexity as the week progressed, with the first featuring a single supercell and the fourth containing multiple hazards and storm modes as well as greater event uncertainty. Forecasters were interviewed and surveyed daily after each case. In total, 12 interviews were conducted and 56 survey responses were collected over the three weeks of the experiment. Information was gathered on how forecasters communicated with each other (especially across organizations) and what they thought of the various tools provided.

Communication

In the post-case surveys, 54 of the 56 respondents indicated that they engaged in cross-organization (e.g., SPC-WFO or vice-versa) communication. Forecasters indicated that they discussed aspects of the mesoscale environment, how they anticipated the event unfolding, threat messaging strategies, planned issuance of watches and mesoscale discussions, hand-offs of PHI plumes and warnings between WFOs (when multiple WFOs were involved), and occasionally radar analysis. The vast majority of respondents indicated they found the external communication either extremely or somewhat helpful (Fig. 1). The WFO forecasters were more likely than SPC forecasters to indicate that they found the external communication to be “extremely” helpful, with SPC forecasters more likely to indicate they found the communication to be “somewhat” helpful. This makes sense given that SPC typically issues their products first, with predictability increasing at finer spatiotemporal scales only at shorter lead times. Participants said they appreciated having a second opinion from another office, which increased their confidence in their forecasts and products. WFO forecasters liked that they could leverage SPC expertise, and all forecasters appreciated knowing what members of the other organizations were thinking because it allowed them to present more consistent public-facing products and messaging across spatiotemporal scales.

Most forecasters also indicated that inter-organization communication also either substantially or somewhat impacted their forecasting thinking and products, with WFO forecasters much more likely than SPC forecasters to note a “substantial” impact (Fig. 2). In particular, WFO forecasters mentioned that their public-facing IDSS graphics were most impacted–and sometimes even driven–by communication with SPC. Both SPC and WFO forecasters expressed a desire to make the messaging in their products as consistent as possible.

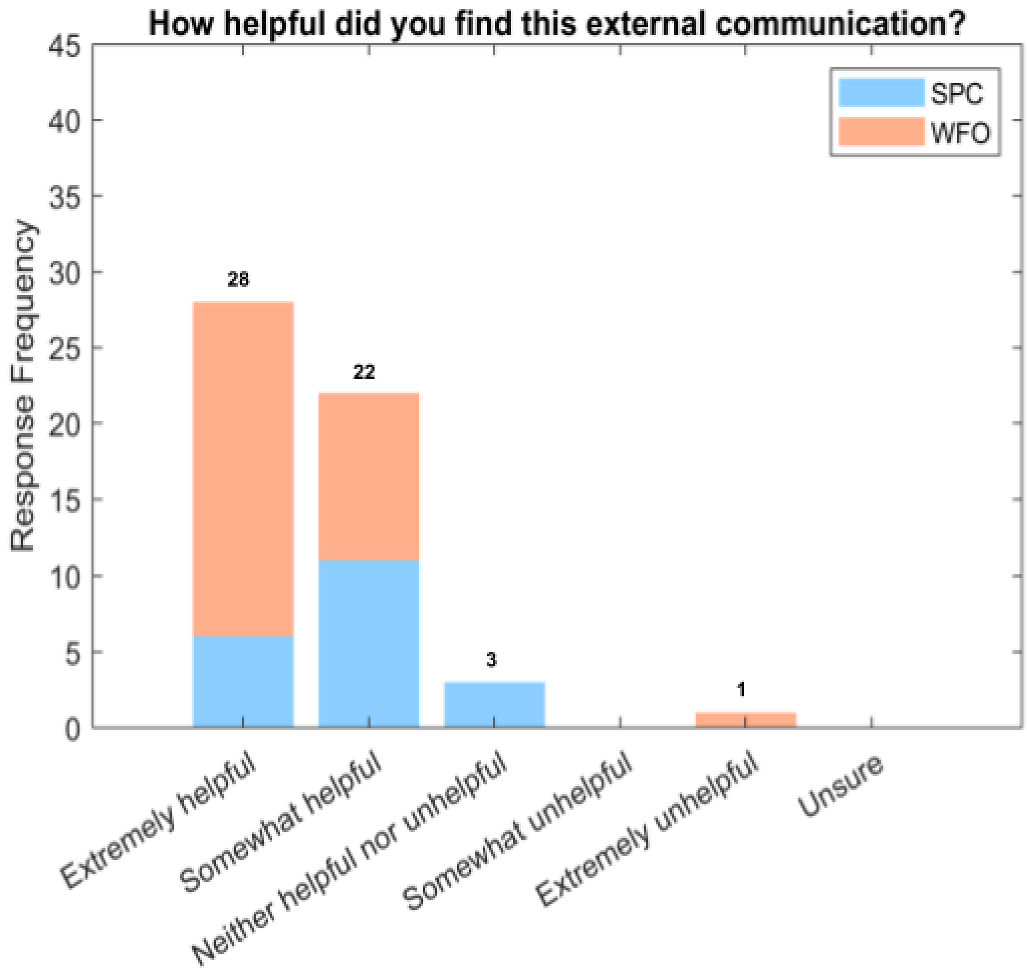


Figure 1: Survey responses from SPC (blue) and WFO (salmon) forecasters showing how helpful they found communication from the other agency during displaced-real-time case simulations, aggregated over all cases and weeks of the 2023 W2W.

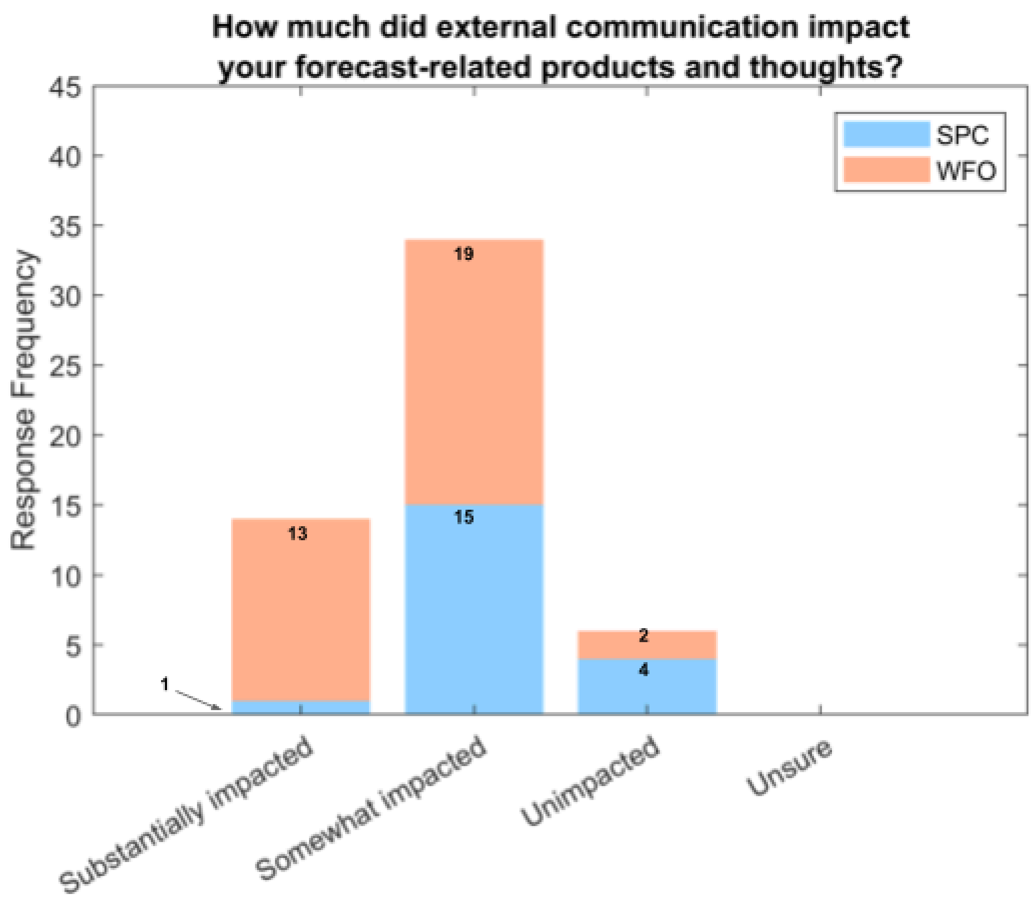


Figure 2: Survey results from SPC (blue) and WFO (salmon) forecasters showing how much they felt inter-organization communication influenced their forecast-related thoughts and products during displaced-real-time case simulations, aggregated over all cases and weeks of the 2023 W2W.

Tools

The W2W featured the development and evaluation of several new tools to facilitate the issuance of probabilistic, rapidly-updating forecast products. First, the PHI Tool, which was initially designed to help with warning and PHI plume issuance, was modified to enable the issuance of 2-h watches as well as mesoscale discussions with associated hazard probabilities over time. With these modifications, the PHI Tool served as a centralized location for the issuance of all forecast products (except IDSS graphics), which enabled both SPC and WFO forecasters to see all past products as well as those currently being worked on.

The W2W also evaluated multiple versions of WoFS-PHI, a new machine learning tool designed to predict severe weather likelihood at different time and space scales based on the combination of WoFS and ProbSevere Version 2. WoFS-PHI was implemented in both the WoFS Viewer and the PHI Tool. In the WoFS Viewer, WoFS-PHI “forecast mode” (Fig. 3a) enabled users to view 30-minute probabilities out to 3-hours of lead time at 7.5-, 15-, 30-, and 39-km radii. Forecast mode was updated every 30 minutes (i.e., every time a new WoFS initialization forecast was released). Meanwhile, “warning mode” (Fig. 3b) allowed forecasters to view severe weather probabilities over the next 90 minutes within 15- or 30-km and was updated every 5 minutes based on the newest ProbSevere and WoFS data. WoFS-PHI forecast mode, but not

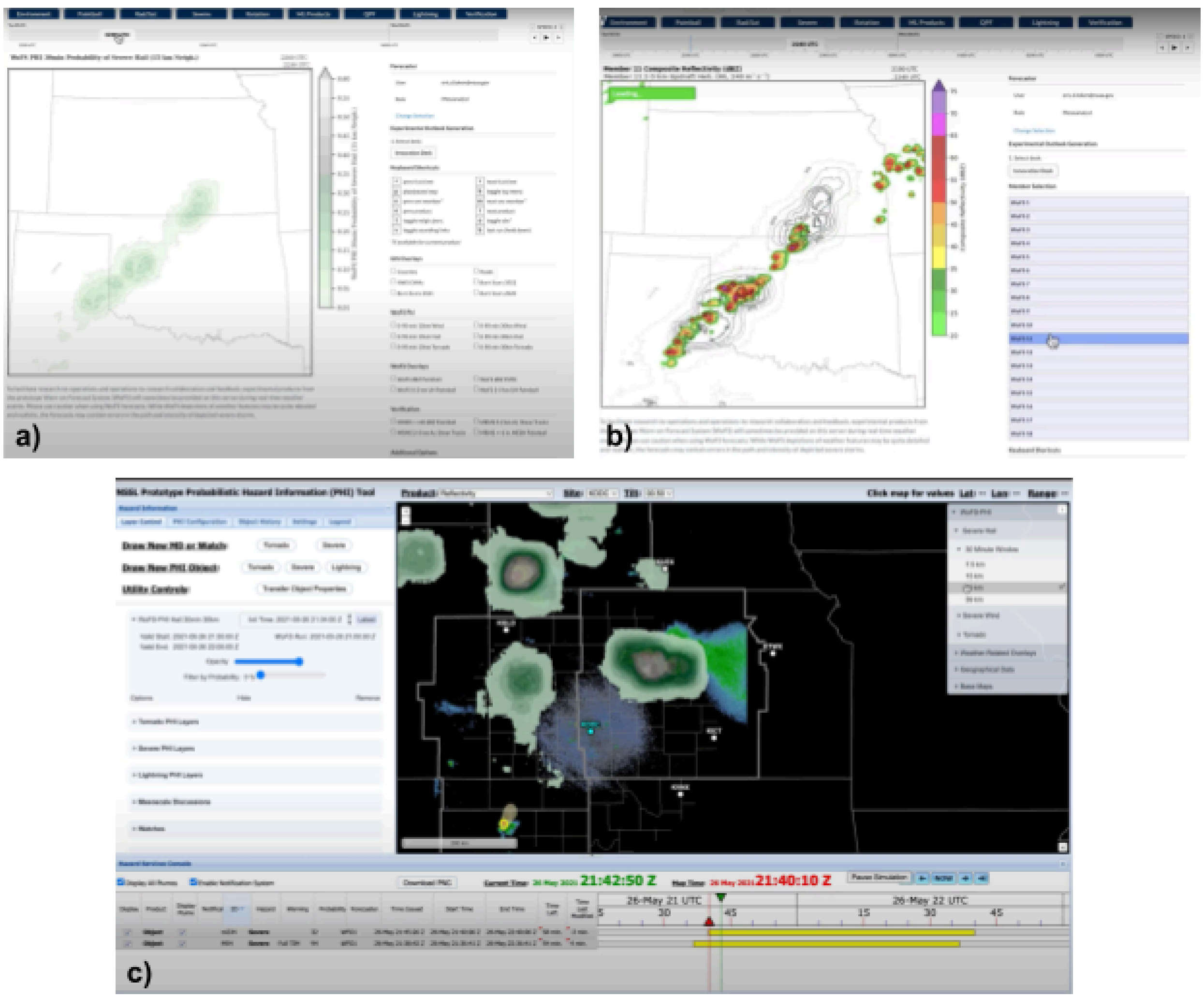


Figure 3: WoFS-PHI (a) 30-minute, 15-km forecast mode hail probabilities, (b) 0-90-minute 30-km warning mode hail probabilities (contours) overlaid with simulated reflectivity from WoFS member 11 (shaded), and (c) 30-minute 30-km forecast mode hail probabilities in the PHI Tool (shaded).

warning mode, was implemented into the PHI Tool as well (Fig. 3c). Thus, the PHI Tool contained versions of WoFS-PHI, TORP, ProbSevere, as well as observed radar data.

In post-case interviews, forecasters generally noted that they liked the usability and capabilities of the PHI Tool. They especially liked that the PHI Tool provided a centralized location where they could quickly view all products issued by multiple organizations, noting that such a tool would be very valuable in real-world operations. They also offered numerous suggestions for improvement, some of which were implemented in later cases of the W2W. Many of these suggestions focused on improving the readability of PHI Tool forecasts, since during complex cases the number of products displayed could be overwhelming. Additionally, a number of participants expressed a desire for the PHI Tool to automatically generate a social media graphic after a product was issued. Plans are in place to implement this capability in future versions of the PHI Tool.

To better understand how different versions of WoFS-PHI were used, participants were asked to identify which versions of WoFS-PHI were most helpful at different lead times. Overall, the 30-minute, 30- and 15-km radii forecast mode probabilities were seen as most helpful, followed by the 0-90 minute, 30-km warning mode probabilities (Fig. 4). In general, as lead time increased, there was a greater preference toward larger-radii products, which makes sense given that predictability at the smallest spatial scales decreases with lead time. However, SPC and WFO forecasters tended to prefer different versions of WoFS-PHI. SPC forecasters generally found the 90-minute, 30-km warning mode probabilities most helpful, while WFO forecasters

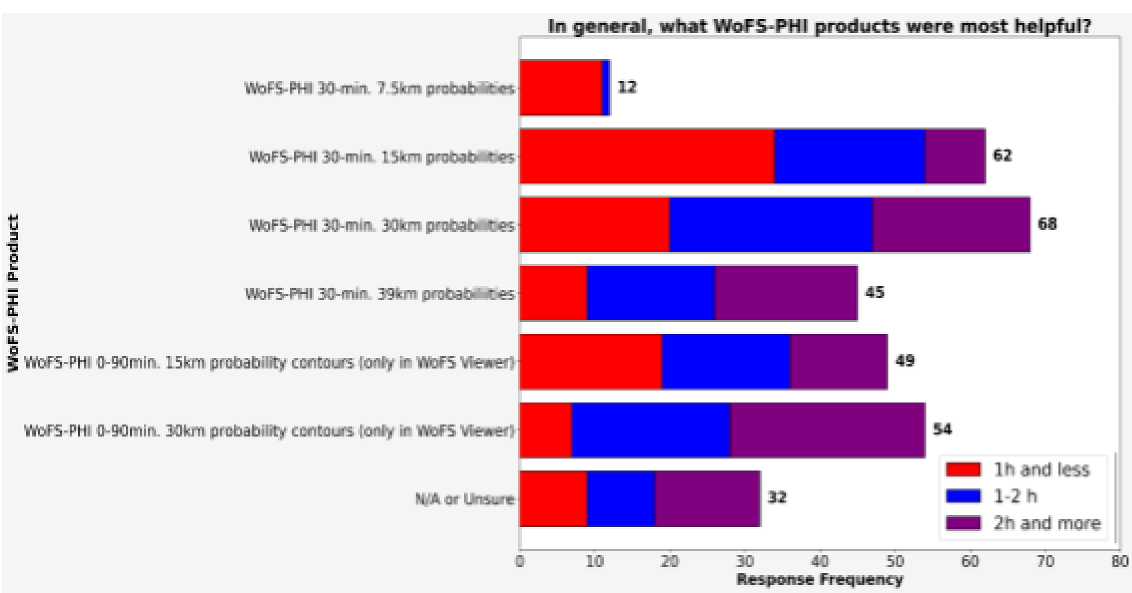


Figure 4: Number of forecasters indicating that a particular WoFS-PHI version was helpful at lead times of 1 h and less (red), 1-2 h (blue), and 2 h and more (purple). Forecasters could select up to 3 WoFS-PHI versions for each lead time. Results are aggregated over all forecasters and all simulated cases.

preferred the 30-minute, 15-km forecast mode probabilities (not shown). These preferences likely reflect the different primary duties (and spatiotemporal scales analyzed) by the SPC and local WFOs.

All forecasters were additionally asked to rate, on a 100-point scale, how helpful WoFS-PHI was for predicting severe weather hazard spatial coverage, likelihood, and timing, as well as for creating public-facing graphics. Overall mean participant ratings were around 60/100 for each category (Fig. 5). Interestingly, mean WFO ratings were higher than mean SPC ratings for WoFS-PHI's usefulness for predicting spatial coverage and timing. The ratings for WoFS-PHI's usefulness for creating public-facing graphics had the greatest variance. One likely explanation is that the creation of public-facing graphics was the first task to be abandoned during periods of heavy workload. This was most notable on the SPC side, as most SPC

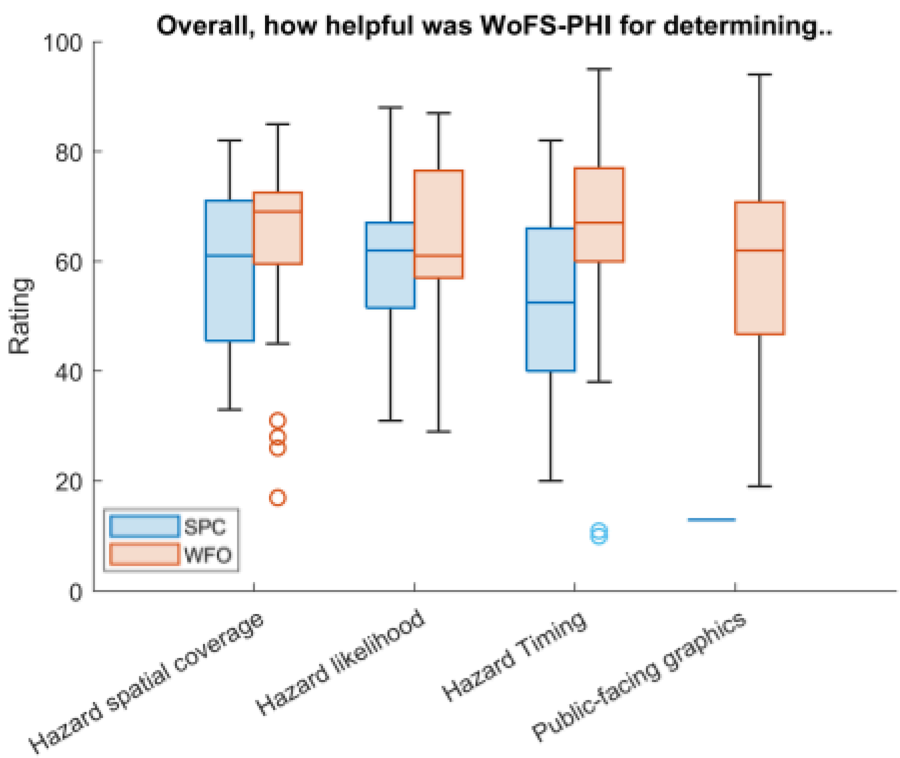


Figure 5: Box-and-whisker plots showing the distribution of SPC (blue) and WFO (salmon) ratings for WoFS-PHI's usefulness for predicting hazard spatial coverage, likelihood, timing, and for creating public-facing graphics. Plots consider ratings from all forecasters over all simulated cases.

forecasters did not have the capacity to create watches, mesoscale discussions, *and* public-facing graphics.

Overall, forecasters shared that they felt WoFS-PHI was most useful for situational awareness (i.e., monitoring the storms/areas with the greatest threat), identifying the timing and spatial extent of hazards, and making warning decisions. Forecasters mentioned that WoFS-PHI probability magnitudes routinely felt too low but highlighted reasonable spatial areas. In general, forecasters' biggest challenge with WoFS-PHI was understanding the meaning and calibration of typical probabilities. For example, one forecaster expressed a desire to know, "...at what loose [probability] thresholds should I become concerned about an area, if I'm not already?" Future versions of WoFS-PHI will attempt to put probability values in better context for forecasters. Other challenges mentioned by forecasters included knowing when to best utilize WoFS-PHI, understanding the technical details of the system ("knobology"), and interpreting WoFS and WoFS-PHI output.

FACETs Paradigm

Products evaluated in the 2023 W2W supported the FACETs paradigm since they tended to be probabilistic in nature and were updated frequently. Overall, SPC participants provided positive feedback on the use of 2-h watches, feeling that these shorter-duration watches allowed them to be more precise with their messaging and ultimately provided more consistent watch lead times. Further, they liked being able to control watch updates, which they currently are unable to do in operations. Meanwhile, WFO forecasters tended to appreciate the ability to issue PHI plumes without warnings. However, all forecasters struggled with knowing how to translate probabilistic hazard forecasts into warnings and public-facing messages. While most of the case simulations involved a single WFO, participants noted that the paradigm used in the W2W could be challenging to implement in real operations with multiple WFOs, as it could be unclear how to handle storms (and PHI plumes) crossing WFO boundaries. The PHI Tool was valuable in that regard by providing a centralized location for storing and issuing forecast products and information. Still, forecasters noted that the paradigm tested in the W2W would likely require more communication than is currently done operationally, which would almost certainly require additional staffing. Participants felt that additional training would also be needed on mesoscale analysis and the probabilistic nature of the new tools. Although participants generally acknowledged the enormity of the paradigm shift relative to current operations, they expressed enthusiasm for the products and services that the new paradigm would provide.

Future Plans

Results from the 2023 W2W will inform future product development and HWT experiment design. For example, the next iteration of WoFS-PHI will incorporate ProbSevere Version 3 and TORP predictors, and different ways to contextualize output probabilities are currently being explored. A real-time forecasting activity involving WoFS-PHI is planned as part of the 2024 Spring Forecasting Experiment to assess the value of the product for predicting severe weather in real-time. Finally, a second W2W is planned for fall 2024. Major goals include evaluating updated versions of the tools used in the 2023 W2W and studying how emergency managers use probabilistic forecasting products from the SPC and WFOs to make actionable decisions.

Web Presence

- Cloud-based WoFS Viewer: https://cbwofs.nssl.noaa.gov/
- NSSL WoFS-PHI Github Repository: https://github.com/NOAA-National-Severe-Storms-Laboratory/frdd-wofs-phi

Project Contacts

Eric Loken	OU CIWRO	Project PI
Kristin Calhoun	NSSL	Project Co-PI
Thea Sandmael	OU CIWRO	Project Co-PI
Miranda Silcott	OU CIWRO	Research Associate

References

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

Jones, T. A., and Coauthors, 2020: Assimilation of *GOES-16* radiances and retrievals into the Warn-on-Forecast System. *Mon. Wea. Rev.*, 148, 1829–1859, https://doi.org/10.1175/MWR-D-19-0379.1.

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

Sandmæl, T. N., B. R. Smith, A. E. Reinhart, I. M. Schick, M. C. Ake, and J. G. Madden, 2023: The tornado probability algorithm: A probabilistic machine learning tornadic circulation detection algorithm. *Weather and Forecasting*, 38, 445–466, https://doi.org/10.1175/WAF-D-22-0123.1.

Wilson, K. A., P. C. Burke, B. T. Gallo, P. S. Skinner, T. T. Lindley, C. Gravelle, S. W. Bieda III, J. G. Madden, J. W. Monroe, J. E. Guerra, and D. A. Morris, 2024: Collaborative exploration of storm-scale probabilistic guidance for NWS forecast operations. *Wea. Forecasting*, 39, 387-402, https://doi.org/10.1175/WAF-D-23-0174.1.

4. PERSONNEL

Tony Lyza
HWT Executive Officer anthony.lyza@noaa.gov

Alan Gerard
2022–2023 Acting HWT Co-Executive Officer alan.e.gerard@noaa.gov

Kodi Berry
2022–2023 Acting HWT Co-Executive Officer kodi.berry@noaa.gov

Kristin Calhoun
2022–2023 Acting HWT Co-Executive Officer kristin.calhoun@noaa.gov

Justin Monroe
EWP Technical Lead justin.monroe@noaa.gov

Jonathan Madden
EWP Technical Advisor jonathan.madden@noaa.gov

5. ACKNOWLEDGMENTS

We would like to acknowledge the contributions of the following individuals: Tanya Riley, Tracy Reinke, Keli Pirtle, Emily Summars, James Murnan, Karen Cooper, Carrie Langston, Vicki Farmer, and Alan Gerard.

HWT Infrastructure Support was provided, in part, by the U.S. Weather Research Program and Joint Technology Transfer Initiative. Research was funded by the U.S. Weather Research Program, Joint Technology Transfer Initiative, National Severe Storms Laboratory, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA16OAR4320115, U.S. Department of Commerce.