2021 Hazard Services – Threats-In-Motion HWT Virtual Experiment

Summary by Greg Stumpf

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 non-uniform 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 a proposed first 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. In addition, TIM provides a continuous history for each storm, which would lead to simplified and consistent messaging for key partners and improving event verification.

The most significant benefits of TIM are from improvements in hazard communication. If implemented, 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. In addition, warnings could be automatically cleared from locations where the threat has passed. This change would result 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).

Efforts have been underway since 2019 to develop the software capability to issue TIM for convective weather warnings (tornadoes, wind, and hail) within the Advanced Weather Interactive Processing System (AWIPS) Hazard Services. This new software, known as HS-TIM, was tested in the HWT in a limited sense in 2019. NWS/MDL, in partnership with NSSL, GSL, WDTD, and U. Akron, carried out another HWT experiment in the summer of 2021. This was the first experiment that focused 100% on TIM, allowing us to explore several ideas to represent realistic challenges currently faced in warning operations in order to focus on workload differences with the current NWS warning system known as AWIPS – WarnGen. Only forecasters participated (no end users), so the resulting feedback is from the operational NWS forecaster perspective. The major differences between WarnGen and HS-TIM that forecasters had to learn and get used to 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.
- Using 2D storm "objects" (versus points and lines) to define and track current threat areas, and to project the future threat areas at 1-minute intervals to create the warning polygon swaths. This includes new 2D drawing tools (polygon, freeform, ellipse) that are currently unavailable in WarnGen.
- The "Persist" feature, which, when enabled, places a warning "in motion", updating the location every minute. For persisted warnings, during subsequent warning updates, the 2D objects can be quickly repositioned and reshaped 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) which allows the warning forecaster to add their warning decision thoughts about why they issued or modified the storm object (e.g., "mid-level rotation is strengthening"). These are thoughts that are 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 lifetime of the storm, the WDD history for the storm is linked from the first time the warning is issued, giving a "story about the storm".

The HS-TIM experiment was carried out for four weeks in the summer of 2021, 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. 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 within the Amazon cloud services (see Fig. 1). 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: There was no in-person social interaction outside working hours; we could not split instances to two monitors; workdays had to be shortened (due to "Zoom fatigue"), resulting in less time for forecasters to participate in hands-on scenarios

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

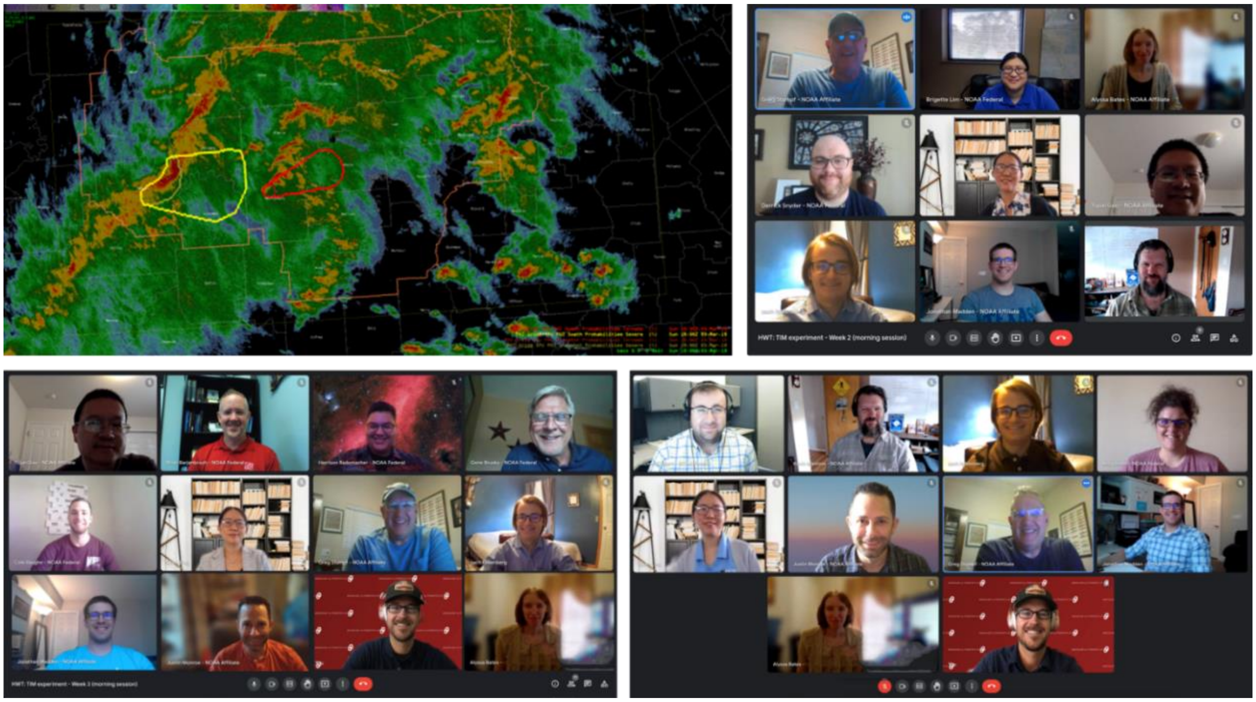


Figure 1. *Images from the HS-TIM experiment. Top left: HS-TIM output from a scenario in southeast Alabama - the yellow (red) polygons show severe thunderstorm (tornado) warnings; the large orange polygon is the county warning area (CWA). The remaining three images are group photos of the virtual meetings during each of the three weeks of the experiment.*

Since its initial limited test in 2019, the HS-TIM software had been robustly hardened, and thus remained stable during the experiment, causing little to no impact on the evaluation of forecaster workload. A number of major new functionalities were added to HS-TIM and evaluated during this experiment:

- "Drag Me To Storm" (a.k.a. "Reset Motion Vector") mode is now on for each update (saves some mouse clicks).
- Added a "Select Hazard Type" label to a new object before the hazard type is selected to prompt the user to take the correct action (saves some mental workload).
- Added a “Latest TIM Frame” button, because sometimes the most-current radar frame is more recent, and the TIM object is not visible.
- Uncertainty values can now be manually set to zero to account for steady-state storm motion situations without greatly fanning out the warning polygons. The default uncertainty values are only used on brand new objects.
- The motion vector is not re-calculated when the object shape is changed. This prevents drastic changes to the motion vector based on the calculation of a new object centroid.
- Changed the duration behavior for non-Persisting warnings (the End Time remains constant and the duration ages off), and Persisting warnings (the duration will reset to the

previous manually-set duration; the duration will not age off and the End Time will increment one minute with each automatic update).
- Improved warning product formatting, with updated IBW tags for Severe Thunderstorm Warnings based on recent NWS directives (Aug 2021).

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

As with previous experiments, this evaluation included human factor experts who recorded video and audio, and administered surveys and interviews to measure mental workload, confidence, and software usability. More information about the human factors analysis is included in Appendix C.

Key Takeaways

- Using 2D objects with Persist to create Threats-In-Motion is a huge workload saver for subsequent warnings
  - The warning updates/follow-ups take much less time with TIM versus WarnGen.
  - At times, forecasters can update warnings more frequently using HS-TIM, which is especially important for high-impact events.
  - The warning updates are more precise.
  - One forecaster's testimonial, "handling these busy cases using WarnGen would have been pretty poor. We would be load shedding in order to keep up with the workload."
- "Persist" should be allowed for *any* long-tracked storm, not just those with observed significant hazards. Should be considered for Tiny TIM.
- Along with the many benefits of TIM from a statistical verification standpoint, the savings in forecaster workload make this a must-need for the NWS.

Future Plans

- Another virtual HWT experiment during Summer 2022.
  - Include new cases
  - The addition of a 3rd NWS forecaster, to act as the mesoanalyst and the conduit to end users.
- Continue software refinement and development of new functionality, including:
  - Hardening code (debugging, refactoring).
  - EDEX Recommender (opening up the possibility to use WES2 for displaced real-time scenarios).
  - Integration of HS-TIM into the operational version of Hazard Services, to take advantage of the many new functionalities in the newest versions of HS.
  - Windshield Wiper Effect (WWE) mitigation, such as a 5-10 minute "cool down" period added to locations removed from the warning.

- Automated county clipping.
- Advanced motion uncertainty features (e.g., splitting right and left direction uncertainty).
- Separate front and back end of the warnings in motion.
- Modify the Hazard Information Dialog to better match HS-Convective and Tiny TIM content.
- Ability to merge, split, and copy threat objects.
- WFO Collaboration domain permission to work with WFO localization.
- Any of the items suggested by forecasters in the Appendices.

Web Presence (each subsequent item goes into more depth on TIM than the previous item)

- 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).

Project Contacts

Greg Stumpf	CIRA and NWS/MDL	HWT Experiment Coordinator / Project Co-PI
Kevin Manross	CIRA and NOAA/GSL	Lead Developer / Project Co-PI
Alyssa Bates	CIWRO and NWS/WDTD	Project Co-PI
Chen Ling	U. Akron	Project Co-PI
Yujun Guo	CIRA and NOAA/GSL	Developer
Pat Hyland	CIWRO and NOAA/NSSL	Project Scientist
Zach Falkenberg	U. Akron	Project Scientist
Justin Monroe	CIWRO and NOAA/NSSL	AWIPS Support
Jonny Madden	CIWRO and NOAA/NSSL	AWIPS Support
Kodi Berry	NOAA/NSSL	NSSL FACETs Program Manager

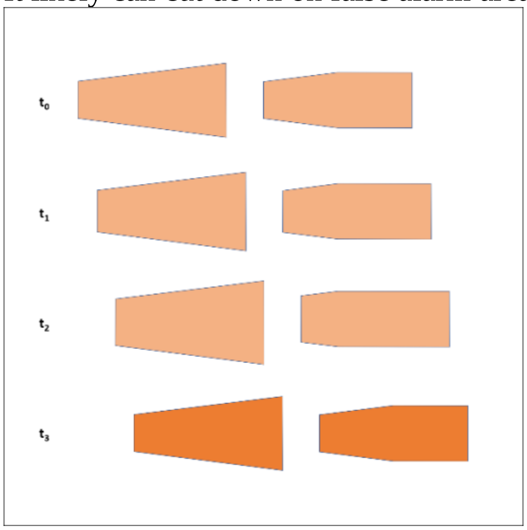
References

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-D-20-0159.1

Appendix A: Software Functionality Suggestions

Warning Output:

- Add the option to have the uncertainty extend to only to 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. Perhaps set the time at which the uncertainty stops expanding. It likely can cut down on false alarm area/time as well.

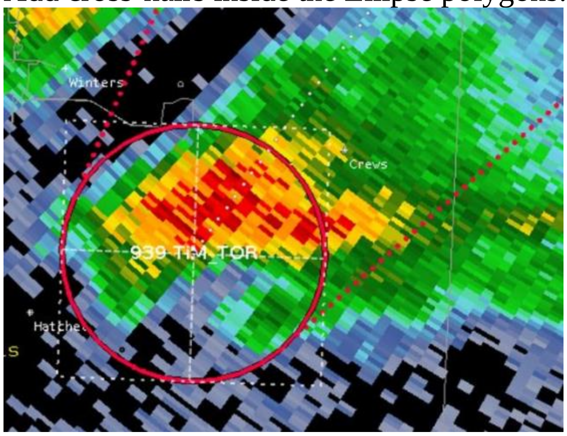


Object Editing:

- For ellipses, use the Ctrl key to make perfect circles. Also add the ability to do a single click for a 5 km wide circle centered on the mouse click.
- Change the workflow to be more like WarnGen. Draw the object, set the motion vector, select the hazard type. Or, to initiate a new object, select New SVR or New TOR, either from a button (see above), or a right-/long-/double- click on the spatial display. Less eye- and mouse-movement are best. Maybe 5 seconds of eye- or mouse-movement is saved, but saves some mental stress that stacks up over a 3-4-hour event.
- New ways to take action on objects to help with workflow:
  - Add Modify and End Object buttons in the object right-click menu.
  - Selecting an object via the spatial display should be with a double click (or the center or right mouse button) to prevent accidental selection when panning and zooming, or when toggling the legend.
- Add a way for forecasters to see the outline of the previous warning swath while they are modifying the current warning. So that they know what areas they are removing. This could be triggered on the “Preview Warning” button, and be shown as a different color.

Spatial Display (visual features):

- Bottom right rotation should be the two arrows around a circle (like PPT). Or, move the rotation to a short stem, not on the box (like Photoshop).
- Add cross-hairs inside the Ellipse polygons.



- Use thicker polygons to denote either 1) persisting warnings, or 2) significant warnings (based on IBW tag).
- 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:

- Remove seconds on the console times to save space.
- The draw controls (which are only 5 buttons: radio:[Polygon, Freeform, Ellipse], radio:[New, Replace]) should be a separate breakaway in a narrow vertical stripe immediately next to the HID, instead of on the Console. That way, the Console could be broken off and put on another monitor as a SA display.
    - Alternatively, grey out Console components that are not used, or make these controls accessible by a right-/long-/double- click on the spatial display.
- Add “triangles” on the timeline to show the persisting portion too. Square would change to a triangle pointing right. If we freeze the back end, something that shows that too, like a back-facing triangle.

HID (General):

- On the HID Details box, have Modify, End Object, Persist, Preview, Reset, Latest TiM Frame, “pinned” into a separate Megawidget at the top (it would add more border boxes).
- Persist button should change to green when depressed. And change the label based on state (“Persist OFF” “Persist ON”).

HID (IBW tags and warning attributes):

- For SVR warning (3 rows, 2 columns [make sure the "Max" are vertically-aligned]):
  - Tornado
  - Hail Source, Max Hail
  - Wind Source, Max Wind
- For TOR warning (1 row; flip the current order):
  - Tornado Source, Tornado Severity

HID (Warning Decision Discussion):

- Add additional fields to WDD output:
  - WFO (so we know when object switched owners), storm motion values, Persist state, tag if geometry changed
- The initial issuance of a warning is the most time-sensitive and urgent, so we shouldn't require a WDD for a new warning. Instead, there should be a way to add a WDD at any time without updating the entire object.
  - This could be done by anyone in ops, including the mesoanalyst.
  - When the new warning is issued, unless the WDD is edited by the forecasters, it should include the phrase, "New warning - More details to come".
  - There should be a pop-up reminder if this isn't done within 5 minutes of new warning issuance.
- Add common WDD phrases (e.g., "no change in status") to a drop-down menu.
  - Include the option to repeat the previous discussion.
- Limit WDD to 280 characters, with a character countdown as you type.

Alerts and Pop-Ups:

- Need alert for expiring warning too like WarnGen. Options include, 1) pop-up, 2) blinking objects, 3) blinking rows in the console, 4) thick versus thin polygons.

Appendix B: Experiment logistics suggestions

Do an inventory of WarnGen functionality and layout to determine differences, and start converging.

Use a Slackbot for our WCM script. Could it be designed to use relative times instead of absolute times, since our reports are from the past?

Add background noise to scenarios: add other people talking not related to weather, phones ringing. Noise and ringing phones ramp up as storms move into the metro area.

Allow the forecasters to continue to "play in the sandbox" for about an additional hour or so after the Scenario 1 training event (keep the instances running, unsupervised).

Future TIM in-person experiment suggestion: when comparing WarnGen to HS-TIM, set up parameters of warnings for a controlled experiment: start at a specific time, have a standard duration (30 min), issue SVRs or updates at a specific time (every 10 minutes), direct them to which storm to warn on (they can decide no warn after analysis).

Procedures:

- Would have liked more time to set up procedures - perhaps on the training day, spend 30-60 min to create all procedures before starting the experiment.
- Or pick 1-2 core analysis procedures from each forecaster. We could also build a handful of pre-made procedures like the RAC procedures. We can alter them for our cases.
    - For example: clockwise from top left: Reflectivity At Lowest Altitude (RALA), Maximum Estimated Size of Hail (MESH) and MESH Tracks (60 min. accum.) image combination, Low-Level (0-2 km AGL) Rotation Tracks (60 min. accum.), and Vertically Integrated Ice (VII).

Should consider having some "pre-made" objects already available in all future scenarios.

Add a 4th DRT scenario to Friday morning (perhaps a back-building case), and have the afternoon for our 2-hour interview?

Appendix C. Human Factors Analysis Survey Results

Participants in the HWT

Six NWS forecasters participated in the 2021 HS-TIM HWT, with two forecasters every week. The NWS work experience ranges from 2.75 years to 32 years, and the average was 9.8 years (standard deviation 11.83). The average NWS warning experience ranges from 2 years to 23 years, with an average of 7.4 years (standard deviation of 8.79).

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, with 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 1 shows the average ratings for the six sub-dimensions and the overall workload for three archived hazardous weather scenarios. The average workload for 2021 Hazard Services TIM HWT across all testing scenarios was 48.5 (out of 100, standard deviation 10.89). 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 2021 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	64.3 (13.81)	63.0 (22.52)	60.5 (19.78)
Physical Demand	29.1 (18.07)	34.0 (18.71)	34.2 (20.25)
Temporal Demand	70.0 (11.40)	51.7 (25.92)	55.3 (25.50)
Effort	70.0 (12.25)	56.7 (23.89)	59.2 (21.45)
Frustration	45.2(22.09)	33.8 (9.34)	37.5 (27.02)
Performance	22.8 (12.02)	26.0 (12.94)	27.2 (13.49)

Overall Workload	50.19 (5.84)	44.11 (11.69)	45.6 (14.07)

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 9-15), Interface Quality (Questions 16-18) and Overall Usability (Questions 1-19). The rating ranges from 1 to 7, with 1 corresponds to low level of usability, 7 to high level of usability, and 4 corresponds to 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 2016 – 2021 HWT (7-point scale)

	2016	2017	2018	2019	2021HS-TIM Mean(std)
Overall Usability	4.62	5.39	4.95	4.92	**5.97 (0.39)**
System Usability	4.96	5.56	5.21	4.88	**6.25 (0.25)**
Information Quality	4.37	5.00	4.37	4.81	**5.71 (0.62)**
Interface Quality	4.72	5.72	5.50	5.18	**5.78 (0.91)**

The overall usability was assessed at 5.97 (on a 7-point scale) for the 2021 HS-TIM testbed, with system usability rated at 6.25, information quality at 5.71, and interface quality at 5.78. It is worth noting that the usability rating has improved from the past years. There are several factors accounting for this. First, the previous four years included a probabilistic component to the software, which increased workload. Second, the software has become more stable throughout the years. Finally, suggestions made by forecasters to improve the workflow have been implemented over the years.