GOES-R and JPSS Proving Ground Demonstration at the Hazardous Weather Testbed 2022 Spring Experiment Final Evaluation

Project Title: GOES-R and JPSS Proving Ground Demonstration at the 2022 Spring Experiment – Experimental Warning Program (EWP)

Organization: NOAA Hazardous Weather Testbed (HWT)

Evaluators: National Weather Service (NWS) Forecasters, Storm Prediction Center (SPC), National Severe Storms Laboratory (NSSL), University of Oklahoma (OU), Cooperative Institute for Severe and High-Impact Weather Research and Operations (CIWRO)

Duration of Evaluation: 23 May 2022 – 17 June 2022

Prepared By: Kevin Thiel (OU/CIWRO and NOAA/SPC)

Submitted Date: 10 September 2022

Table of Contents:

1. Executive Summary

This report summarizes the activities and results from the Geostationary Operational Environmental Satellite R-Series (GOES-R) and Joint Polar Satellite System (JPSS) Proving Ground demonstration at the 2022 Spring Experiment, which took place virtually at the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) in Norman, Oklahoma from 23 May to 17 June 2022. This year featured 19 participants in the EWP experiment. 17 of the participants were National Weather Service (NWS) forecasters from Weather Forecast Offices (WFOs) in five NWS regions and one Center Weather Service Unit. Additionally, there was one participant from the Forecast Decision Training Division and another from the NWS Pacific Region Headquarters. This group evaluated six major baseline, future capability, and experimental GOES-R and JPSS products in the real-time simulated short-term forecast, decision support service (DSS), and warning environment of the Experimental Warning Program (EWP). Additionally, they used cloud-based instances of the second-generation Advanced Weather Interactive Processing System (AWIPS-II) and web-based interfaces to interact with the products.

Forecaster feedback during the evaluation was collected through daily and weekly surveys, daily and weekly debriefs, blog posts, a warning and DSS reporting form, public forecast graphics, and informal conversations during the testbed. Typical feedback included suggestions for improving the algorithms, display techniques, product use, training, and awareness of product applications or limitations. Most of the products evaluated in 2022 were advancements of previous product iterations from the 2021 GOES-R/JPSS Proving Ground. This included data from the Geostationary Lightning Mapper (GLM), the NOAA Unique Combined Atmospheric Processing Systems (NUCAPS), the Optical Flow Winds product, and the Probability of Severe (ProbSevere) model – Version 3. The Polar Hyperspectral Sounder and Microwave Imagery in the Advanced Baseline Imager (PHSnMWnABI) model and the ProbSevere LightningCast model were evaluated in the HWT for the first time (Table 1).

Demonstrated Product	**Category**
Geostationary Lightning Mapper	GOES-R Baseline & National Weather Service
NUCAPS Temperature and Moisture Profiles	JPSS Baseline
Optical Flow Winds	GOES-R Risk Reduction
PHSnMWnABI Model	GOES-R Risk Reduction
ProbSevere LightningCast	GOES-R Risk Reduction
ProbSevere Version 3	GOES-R Risk Reduction

Category Definitions:

- GOES-R Baseline Products – GOES-R Level 1 Requirement products that are funded for operational implementation.
- GOES-R Risk Reduction – New or enhanced GOES-R applications that explore possibilities for improving Algorithm Working Group (AWG) products. These products may use the individual GOES-R sensors alone or combine with data from other in-situ and satellite observing systems or NWP models with GOES-R.
- National Weather Service – Products created within AWIPS-II.
- JPSS Baseline – Products funded through the JPSS program.

Table 1: List of GOES-R and JPSS products demonstrated within the HWT/EWP 2022 Spring Experiment

Over 20 visiting scientists attended the EWP over the three weeks to provide additional product expertise and interact directly with operational forecasters. Organizations represented by those individuals included four NOAA Cooperative Institutes, five federal partners, and three external partners (Table 2). The Storm Prediction Center (SPC) and HWT Satellite Liaison Kevin Thiel (OU/CIWRO and NOAA/SPC) provided overall project management and subject matter expertise for the HWT Satellite Proving Ground efforts in the HWT. Technical support for AWIPS-II were provided by Jonny Madden and Justin Monroe (OU CIWRO and NOAA-NSSL).

Category	**Organization Name**
NOAA Cooperative Institutes	University of Oklahoma Cooperative Institute for Severe and High-Impact Weather Research and Operations (OU/CIWRO)
	University of Wisconsin Cooperative Institute for Meteorological Satellite Studies (UW/CIMSS)
	Colorado State University Cooperative Institute for Research in the Atmosphere (CSU/CIRA)
	University of Maryland-College Park Cooperative Institute for Satellite Earth System Studies (UMD/CISESS)
Federal Partners	NOAA's National Severe Storms Laboratory (NSSL)
	NOAA's National Environmental Satellite, Data, and Information Service (NESDIS)
	NOAA's National Weather Service (NWS)
	NASA's Jet Propulsion Laboratory (JPL)
	NWS Operations Proving Ground (OPG)
External Partners	Science and Technology Corporation (STC)
	Howard University
	Meteorological Service of Canada Storm Prediction Centre

Table 2: List of affiliations from all visiting scientists during the 2022 HWT/EWP 2022 HWT Satellite Proving Ground Experiment.

2. Introduction

GOES-R 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 the GOES-R satellite series, beginning with GOES-16 which launched in November 2016. 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 GOES-R data and products. Additional demonstration of JPSS products introduces and familiarizes users with advanced satellite data that are already available. 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 (Calhoun et al. 2021) facilitates the testing of satellite-based products in the AWIPS-II data processing and visualization system currently used at NWS Weather Forecast Offices (WFOs).

Due to the ongoing COVID-19 Pandemic, 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. Six to seven NWS forecasters volunteered each week to evaluate this year's products. Before the testbed user guides, PowerPoint presentations, and online learning modules were shared with all forecasters through Google Drive for each of the products demonstrated. The Monday of each week began with one hour of introductions and product summaries from developers, with the second hour devoted to familiarizing forecasters with their cloud-based AWIPS instances. Tuesday, Wednesday, and Thursday began with a one-hour discussion of the previous day's operations involving questions from developers and feedback from forecasters. After a brief forecast discussion each day, forecasters were placed into operations at their simulated WFOs and assigned DSS events. Two to three NWS WFOs were selected to maximize the probability of severe thunderstorms. Forecasters were then divided into groups for each simulated WFO, with all groups in their own video conference. Organizers and product developers could then move between the different video conference rooms to interact with the forecasters at each simulated WFO. Expanding on an initiative from the 2021 experiment, mock-Decision Support Service (DSS) events were created for a majority of simulated WFOs to investigate how the experimental products could also be utilized in communicating hazards to NWS partners. Forecasters were encouraged to fill out a short online form after submitting a convective warning or advisory, along with any forecaster who wished to issue DSS for their assigned event. Responses from this form were then examined to identify how the experimental products were incorporated into the communication of convective hazards by the participants. Additionally, forecasters had the option to create social media graphics using the experimental products, further showcasing their ability to be interpreted and transmitted to the public.

Forecasters viewed GLM, NUCAPS, PHSnMWnABI, ProbSevere LightningCast, and ProbSevere Version 3 data in the cloud-based instances of AWIPS. The Optical Flow Winds product was available in a web-based interface, along with additional model outputs from PHSnMWnABI.

Prior to the testbed AWIPS procedures were built by the Satellite Liaison for each product in AWIPS, so forecasters could quickly access the products and leverage best display practices as described in the training. Within operations forecasters had several tasks, such as building procedures of their own to integrate experimental products with the ones they currently use, having discussions with the subject matter experts, writing blog posts, and issuing warnings, advisories, and DSS guidance. 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 document through Google Drive, which were published online later to the HWT EWP Blog (https://inside.nssl.noaa.gov/ewp/) and the GOES-R HWT Satellite Proving Ground Blog (http://goesrhwt.blogspot.com/)

At the end of each operations period, forecasters were given a daily survey regarding product performance and utility during the day's events. Activities Monday through Thursday began at 1 pm CDT (18 Z) and ended at 6pm CDT (23 Z), approximately following the schedule outlined in Figure 1. On Friday, an end of week survey was sent to the participants in the morning, followed by a two-hour final discussion with developers, observers, and SMEs to summarize the week's events and encapsulate key product themes. The schedule of this year's experiment in a virtual format differs from iterations of the experiment prior to the COVID-19 Pandemic, with a condensed and static schedule designed to accommodate the virtual format of the 2022 experiment. However, the timing for operations was selected to offer sufficient opportunities for analysis in pre-convective and post-convective initiation environments. The collective feedback from the 2022 GOES-R/JPSS Proving Ground (daily surveys, weekly surveys, blog posts, and daily debrief discussions) are summarized in this report. Product recommendations are listed at the end of each section as 'recommended', 'strongly recommended' and 'highly recommended' in an ascending order of significance from the forecasters.

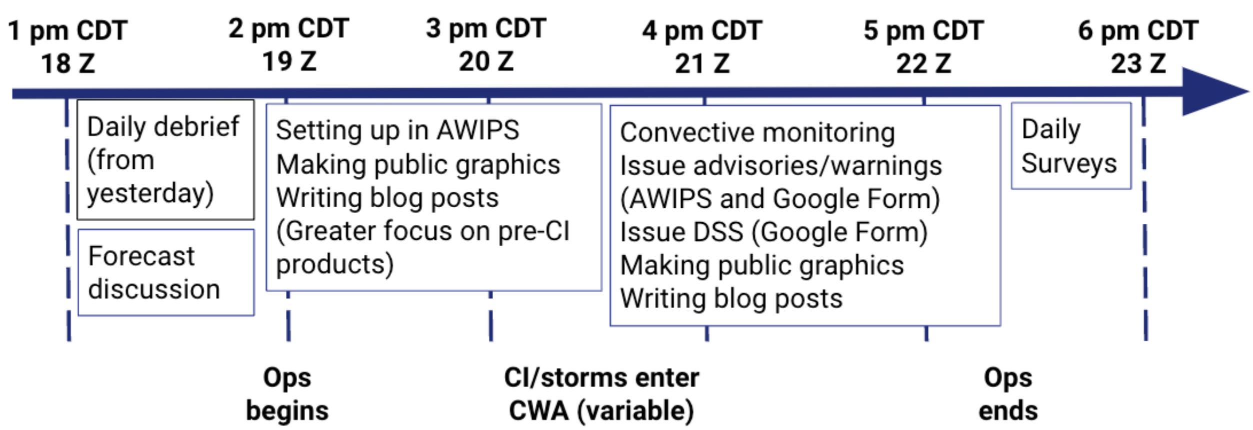

Figure 1: A timeline of the average day in the GOES-R/JPSS Proving Ground, outlining the major activities. On Monday the daily debrief section is replaced by the product introductions by each of the developers. On Friday a two-hour weekly debrief session took place from 1 to 3 pm CDT.

3. Products Evaluated

3.1 Geostationary Lightning Mapper (GLM)

Gridded flash products from the GLM (Goodman et al. 2013; Bruning et al. 2019) were demonstrated again in the testbed, building upon lessons learned from the 2021 evaluation. The 2022 evaluation included Flash Extent Density (FED), Minimum Flash Area (MFA), and Total Optical Energy (TOE) with updates based on feedback from prior testbeds along with testbed-driven research (Thiel et al. 2020). FED displays the total number of flashes that propagate through a single GLM pixel over a given time summation. Similarly, MFA and TOE show the smallest flash area and accumulated optical energy in each GLM pixel, respectively. One- and five-minute summations, updating every minute, were provided from each product. NWS forecasters can currently view the FED product in their home offices from GOES-16 and GOES-17. Testbed goals for the GLM demonstration included development of necessary training and best practices guidance, determining the added forecast value of MFA and TOE, and creating the best display practices for FED in various convective scenarios. Additionally, the increased DSS focus of this year's demonstration, where lightning has a higher priority, provided an opportunity to demonstrate another area where gridded GLM products may be useful.

GLM training materials for these products were available through GOES-R QuickGuides and an online learning module created by the Warning Decision Training Division. On the first day of the testbed, a subject-matter expert briefed forecasters on the GLM products, parallax correction, best display practices, and its differences with ground-based networks. Forecasters were also provided AWIPS procedures which leveraged all three gridded products at once, comparisons between GLM data from GOES-16 and GOES-17, and comparisons of GOES-16 FED color bars with the maximum values of 260 (default), 130, and 65 flashes per five minutes. GLM and lightning subject-matter experts were available throughout the week to answer questions and for discussion with forecasters across a variety of severe and convective weather situations. Overall, forecasters showed the greatest confidence and utility in FED followed closely by MFA, with TOE showing the lowest confidence and fewer applications. Specific applications of GLM data relative to warning decisions, DSS, parallax correction, FED display practices, and sensor limitations are discussed below.

Use of GLM in the HWT

Consistent with the 2021 experiment, the gridded GLM products were evaluated for their utility in various convective scenarios. More specifically, to assess the value added from MFA and TOE for potential operational implementation alongside FED. In the daily surveys, forecasters indicated high confidence in FED and MFA, with over 90% and 74% of all responses indicating confidence values of 'High' or 'Very High' respectively. In contrast, over 67% of forecasters answered that their confidence of TOE was 'Low' or 'Medium'. When asked how useful the participants found each of the three products, over 95% and 85% of responses rated FED and MFA as 'Moderately Useful' or greater respectively, the highest three categories. Approximately 70% of TOE responses were rated 'Not at all Useful' or 'Somewhat Useful', the lowest two categories. The forecasters consistently cited prior experience with FED as a means for its higher confidence and utility, while training and discussion with GLM subject-matter experts were cited for MFA.

‘Confidence is based on amount of time using the product. Feel confident with FED the most and feel as though MFA was explained well in the training. Still not sure how or when to use TOE and that is why the confidence is lower there.’
Forecaster – End of Day Survey

In operations, forecasters most often used changes in local flash rates and area from FED and MFA to assess updraft strength when monitoring convection. Increasing flash rates and decreasing flash areas were often used to identify storms with an increasing severity potential. Additionally, smaller flashes from the MFA were referenced when identifying new thunderstorms near ongoing convection. Of the 82 warnings reported by forecasters during the experiment, the GLM was cited 31 times (38%) as a contributing factor to their warning decision. Increases in FED were mentioned 11 times explicitly, and signals from FED and MFA together were mentioned 5 times. Correspondence with trends in the ProbSevere v3 model output also helped to increase forecaster confidence in these warning decisions.

‘With the Buffalo CWA being a long and narrow forecast area, there are areas where the Buffalo Radar doesn’t provide good coverage. On a day like today where their second radar (KTYX) isn’t providing any data, the entire eastern half of the CWA has the Buffalo radar hitting storms at 20 kft or higher. This meant that we had to more heavily rely on Satellite data for warning operations... Being able to see the upward trends in time with the prob severe over time and then the increasing GLM was a big confidence boost to put out a warning. [Figure 2]’
16 June 2022, Blog Post: *GLM & ProbSevere in Low Radar Coverage*
https://inside.nssl.noaa.gov/ewp/2022/06/24/glm-prob-severe-in-low-radar-coverage/

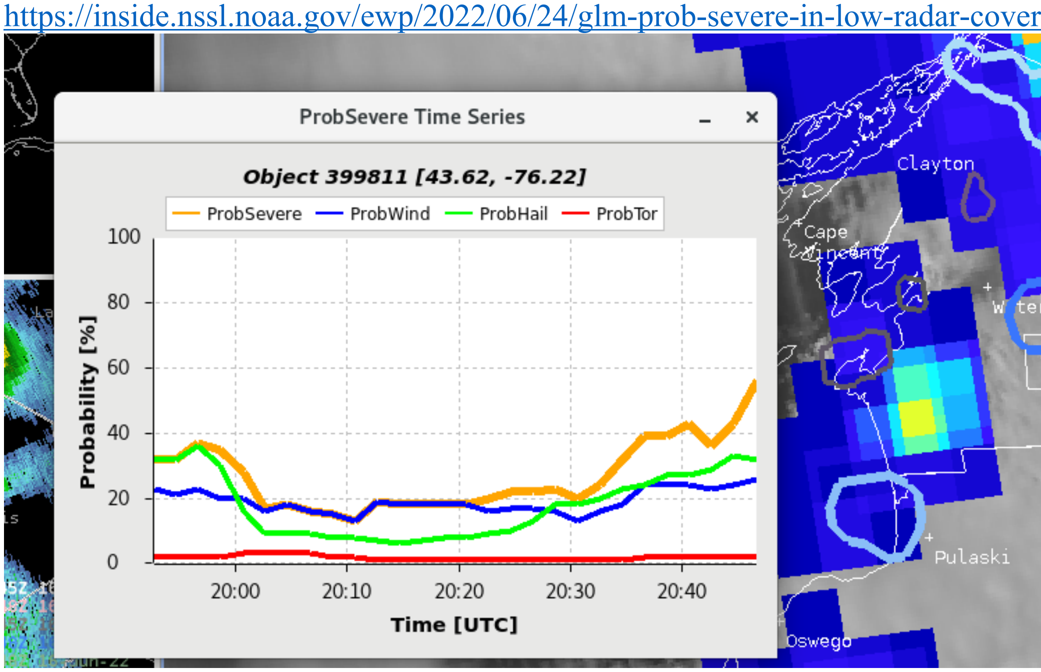

Figure 2: A comparison of the ProbSevere v3 time series with GLM FED in the Buffalo, NY CWA.

‘FED and MFA were confidence boosters for me when deciding whether to issue or hang onto headlines, especially as storms were fairly marginal. Even if reflectivity looked

"spooky" at times, I hadn't noticed much change in the GLM trends and that left me feeling more comfortable with letting a headline expire or not issuing at all.'
Forecaster – End of Day Survey

'Around 21Z, I noticed a jump in GLM FED for the area of storms in the northwest part of the CWA [Figure 3]. Alongside this, the GLM TOE also increased, along with a decrease in MFA with the same storm cell. This area corresponded with increased flash rates in the EarthNetworks.'
26 May 2022, Blog Post: *ILX Ramblings*
https://inside.nssl.noaa.gov/ewp/2022/06/24/ilx-ramblings/

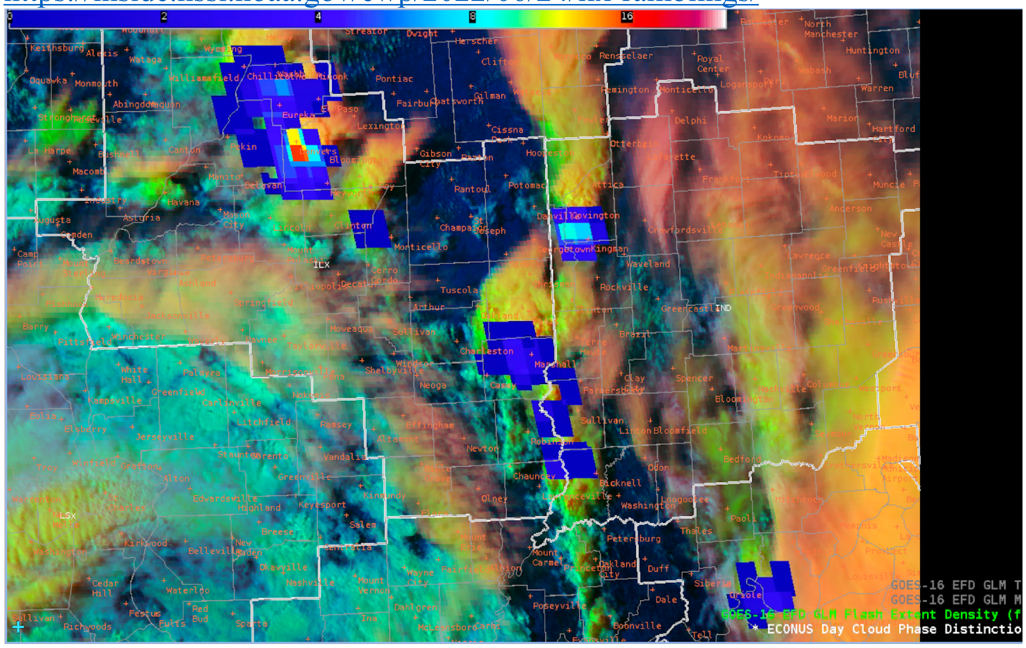

Figure 3:GLM FED (5 minute total) overlaid on the Day Cloud Phase Distinction RGB.

During DSS events, the GLM was frequently used to monitor for lightning activity often in tandem with the ProbSevere LightningCast model. Forecasters were also encouraged to pair the GLM data with cloud-to-ground lightning information from the ground-based networks. Viewing the areal extent of lightning flashes with the GLM provided a key source of developing and incoming convective hazards. From the 52 total DSS messages issued, 35 (67%) cited GLM data as a contributing factor to their messaging. One forecaster also used the GLM as a DSS-related graphic to show an increasing lightning threat as flashes were propagating ahead of the thunderstorm direction and towards the venue (Figure 4). One mentioned limitation of GLM data was the impact of using non-parallax corrected data when attempting to spatially reference the GLM products. This was especially true if the forecasters were unfamiliar with the amount of parallax correction in their daily localization, or if there were multiple storms near each other.

'The weaknesses of the GLM data was the parallax compared to the ENTLN data. The GLM data was typically placed roughly a county north of the surface based lightning data. This was troublesome when providing DSS if only using the GLM data. Would have provided the incorrect timing and location of the strongest part of the storm.'
Forecaster – End of Day Survey

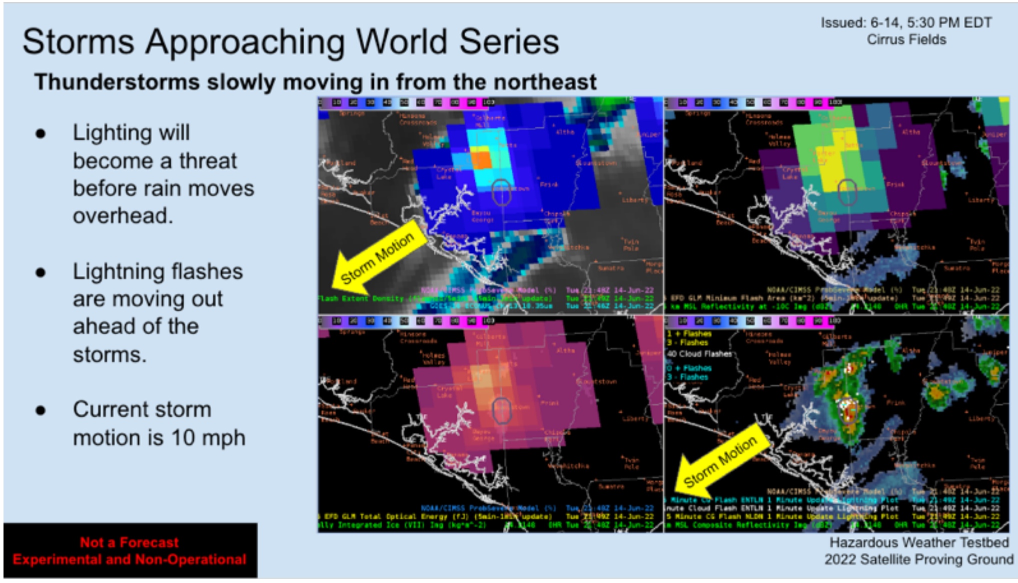

Figure 4: An example of a public information graphic from a simulated DSS event using GLM data.

> 'In dense cloud cover, MFA and FED were able to recognize an area of convection pulsing once it left my DSS event area. This could be useful in a similar situation where the organizer of an event recognizes increased lightning in their area and is looking to return to operations …I would be able to use this as a tool to see that lightning near the event is a continued concern because of potential backbuilding around a strengthening storm.'
> 26 May 2022, Blog Post: *Louisville, KY WFO Observations During DSS*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/louisville-ky-wfo-observations-during-dss/

One new research question brought to the testbed this year by the GLM developers was if the current default color bar range was appropriate for forecasters monitoring intensifying and ongoing convection. By using the AWIPS-II procedure with the adjusted FED color bar maxima as previously explained, forecasters were able to view the same data through three different options. Additionally, forecasters were encouraged in training and within operations to adjust the color bar when needed. In the daily surveys, forecasters were asked which maximum FED value they found most useful in operations. From the 60 responses, 9 (15%) reported that 260 flashes were the most useful, in comparison to 19 (32%) for 130 flashes, 21 (35%) for 65 flashes, and 8 (13%) for 50 flashes or less. Consistently in discussions forecasters stated that lowering the maximum FED value also allowed them to quickly identify initiating convection near ongoing severe thunderstorms, improving situational awareness. To address the need for lowering the default maximum on the FED color bar, forecasters proposed creating FED climatologies or building procedures that used different FED color bar maxima depending on the season or convective scenario.

> 'There should either be a few products with a lower data ranges for FED or a way to give people in ops a gentle reminder that the data ranges can easily be changed. Most

forecasters tend to find what works and then stick with it, without even thinking on how they can change the color curves or data ranges to help themselves out.'
Forecaster – End of Week Survey

'Lowering max FED values effectively highlights the area of strongest thunderstorm activity as seen in this example. This makes the strongest two thunderstorms (circled) [Figure 5] more distinguishable from other thunderstorms, especially on the northern circled thunderstorm.'
8 June 2022, Blog Post: *GLM maxFED Comparisons*
https://inside.nssl.noaa.gov/ewp/2022/06/24/glm-maxfed-comparisons/

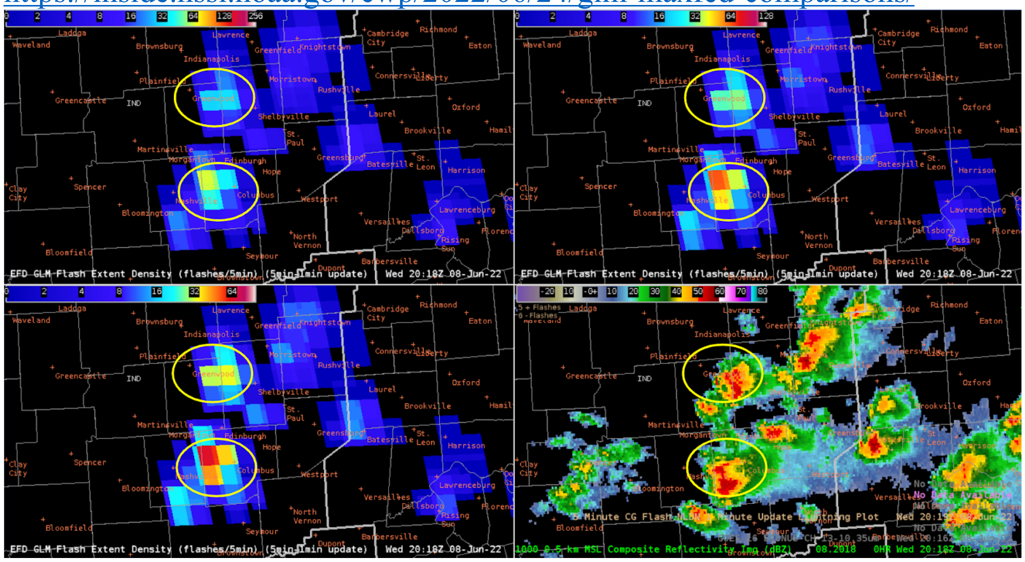

Figure 5: A four-panel display in AWIPS-II showing the GOES-16 GLM Flash Extent Density with maximum values of 260 flashes per 5 minutes (upper left), 130 flashes per 5 minutes (upper right), and 65 flashes per 5 minutes (lower left). MRMS Composite Reflectivity is shown in the lower right corner. The two circles note the storms of interest.

During the 2022 experiment, severe weather potential often positioned forecasters in regions of sensor overlap where they could use observations from both the GOES-East and GOES-West GLMs, mostly across the central United States. When the simulated WFOs were placed in these areas, forecasters were encouraged to use an AWIPS-II procedure that compared the gridded products from each GLM. Oftentimes forecasters reported that the trends from each GLM would match, while their relative values would vary. One feature observed in the testbed was periods where the GOES-West GLM would present greater FED values than the GOES-East GLM, even though the GOES-East GLM was closer and on average has greater detection efficiencies. This was driven by the GOES-West GLM having a more direct view of flashes initiating near the thunderstorm's updraft, especially in cases of deep layer shear displacing precipitation and the anvil to the east. When asked during the daily debrief sections, forecasters were consistently supportive of training to help identify when each GLM should be used, along with the potential development of products that combined flash information from both GLMs.

'GOES-17 had the better angle to see lightning activity in these supercells developing over the high plains of WY and NE. Whereas GOES-16's perspective from further east had to punch through spreading anvils downstream of the main updraft that likely obscured the light emanating from the lightning, GOES-17 had a more side-on view of the updraft with less to no obscurations of light emanating from lightning occurring in the updraft.[Figure 6]'

7 June 2022, Blog Post: *GLM GOES-16 vs GOES-17*

https://inside.nssl.noaa.gov/ewp/2022/06/24/glm-goes-16-vs-goes-17/

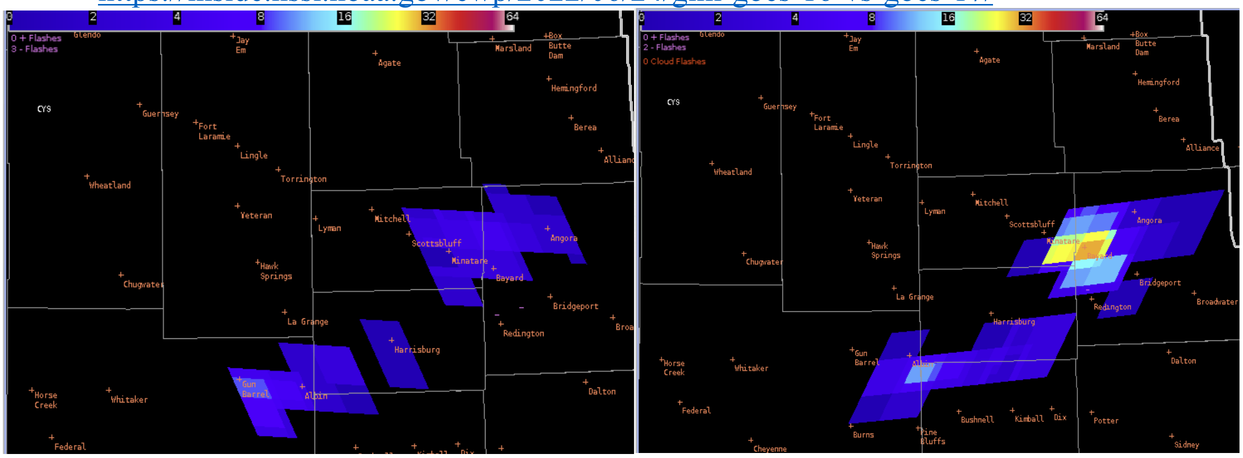

Figure 6: GLM Flash Extent Density (5 minute total) from the perspective of GOES-16 (left) and GOES-17 (right).

Recommendations for Operational Implementation

Based upon the evaluation of GLM products in the 2022 HWT Satellite Proving Ground, the following items have been recommended:

- **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.** Further investigations into the use of optical energy from the GLM to assess data quality or other machine learning applications are encouraged.
- **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.** Training should also emphasize using these gridded products in coordination with radar data, satellite imagery, and ground-based lightning networks. Additional topics may include parallax correction, the impacts of viewing angle and optical depth on GLM flash detection efficiency, and best practices for leveraging the GOES-East and GOES-West GLM in regions of sensor overlap.
- **It is recommended that the default maximum value displayed for the Flash Extent Density product be reduced from 260 flashes per 5 minutes.** The most recommended value is 130 flashes per 5 minutes, however additional methods for determining a situationally useful value may also be explored. Additionally, NWS training efforts may emphasize the need for forecasters to reduce this maximum value, especially in cases of marginally severe convection.

3.2 NOAA Unique Combined Atmospheric Processing System (NUCAPS) Temperature and Moisture Profiles

The NOAA Unique Combined Atmospheric Processing System (NUCAPS; Barnet et al. 2021) was evaluated in another GOES-R/JPSS HWT Satellite Proving Ground (Esmaili et al. 2020; Berndt et al. 2020). Temperature and moisture profiles from NUCAPS were used to provide three products for the 2022 experiment: Operational NUCAPS Profiles, Modified NUCAPS Profiles, and Gridded NUCAPS (Berndt et al. 2020). NUCAPS Profiles in the 2022 HWT were generated using select channels from a mix of infrared and microwave sounders in low-Earth orbit onboard four satellites (Table 3). NUCAPS from NOAA-20 is operationally available on the SBN but was delivered over to the HWT via LDM while other satellites were delivered from a direct broadcast link. Regardless of source, NUCAPS had a latency of under 60 minutes.

Satellite	**Infrared Sounder**	**Microwave Sounder(s)**	**Local Overpass Time (Approx.)**
NOAA-20	Cross-track Infrared Sounder (CrIS)	Advanced Technology Microwave Sounder (ATMS)	1:30 PM
MetOp-B/C	Infrared Atmospheric Sounding Interferometer (IASI)	Advanced Microwave Sounding Unit (AMSU)/ Microwave Humidity Sounder (MHS)	9:30 AM
Aqua	Atmospheric Infrared Sounder (AIRS)	Advanced Microwave Sounding Unit (AMSU)	1:30 PM

Table 3: Information regarding the infrared and microwave sounders on all the satellites providing NUCAPS retrievals for the 2022 HWT.

NUCAPS Profiles were viewed in the National Skew-T and Hodograph Analysis and Research Program (NSHARP) application in AWIPS, along with the addition of the Skew-T and Hodograph Analysis and Research Program-python (SHARPpy) application (Blumberg et al. 2017) for the 2022 experiment. Modified NUCAPS Profiles, available only in NSHARP, provided an adjustment to the boundary layer by using Real-Time Mesoscale Analysis (RTMA) and GOES data over the previous eight hours. Along with viewing the individual soundings from each overpass, forecasters had the ability to view NUCAPS data on a constant surface, layer, or vertical cross-section using the Gridded NUCAPS product.

All three NUCAPS products were evaluated in the 2022 experiment to determine their utility in pre-convective environments including product accuracy, display quality, forecaster understanding, and readiness in operations. Although NOAA-20 NUCAPS soundings and Gridded NUCAPS are baseline AWIPS products, the return of NUCAPS to the testbed was motivated by the need to increasing forecaster awareness that these products operationally are available at their home office and to determine the value of multiple satellites (e.g., increased temporal/spatial coverage of NUCAPS products). Importantly, multi-satellite assessments provide information on the value of the future use of NOAA-20 and NOAA-21, morning MetOp observations, and the potential value of the increased spatial/temporal coverage of a sounder on the next series of geostationary satellites, Geo-XO. Additionally, continued display improvements to Gridded

NUCAPS, the utility of Modified NUCAPS as a basis for future research and development, and the applicability of open source tools such as SHARPpy continue to motivate NUCAPS assessment at HWT. Prior to the testbed forecasters were provided with several JPSS QuickGuides for each product, a document with the timing of overpasses from each satellite, a video on NUCAPS soundings from the 2021 HWT, and a NASA SPoRT training module with case studies featuring NUCAPS. Throughout the demonstration NUCAPS data availability, on a per situation basis, appeared to directly impact the utility of these products in various warning and DSS scenarios. Feedback regarding sounding quality in comparison to other platforms, the display of Gridded NUCAPS, applications to mesoscale analysis and situational awareness, and suggested product improvements are described in the following section.

Use of NUCAPS in the HWT

Forecasters most often used NUCAPS when performing mesoanalysis prior to, or in the early stages of, severe weather entering their CWA, along with intercomparisons with SPC mesoanalysis (RTMA) and PHSnMWnABI model fields. Comparisons of NUCAPS observations with modeled data provided additional verification and confidence of the current mesoscale environment, and was mentioned consistently by forecasters in operations, daily debrief sessions, surveys, and blog posts. If these fields did not match spatially, temporally, or by magnitude, forecasters became aware of potential uncertainties and features to investigate further to improve their conceptual model of the mesoscale environment. Forecasters also used NUCAPS to provide additional insight for where convection may form, along with how ongoing thunderstorms may develop throughout the forecast period.

The operational NUCAPS Profiles were most often used in the 2022 experiment, as 49 (70%) of the daily survey responses (70) indicated that they were included in their daily analysis. The daily responses also indicated that Modified NUCAPS Profiles and Gridded NUCAPS were used in 37 of 65 (57%) and 37 of 67 (55%) responses respectively. When asked which convective parameters and indices were used each day, convective available potential energy (CAPE: surface-based, mixed-layer, and most unstable) was mentioned in 47 of the 60 responses, followed by convective inhibition (CIN) listed 18 times, lifted condensation level (LCL) with 9 mentions, and lapse rates listed 7 times. The following quotes from forecasters highlight these comparisons and the impact of NUCAPS data on their mesoscale analysis.

> ‘I used the modified NUCAPS profiles in AWIPS and compared them to the observed analysis. It also helped me get a sense of where severe storms may form and for convective initiation. I viewed several sites in our southern CWA, where skies were generally clear. It was consistent with SPC mesoanalysis, showing instability over 1500 J/kg and high DCAPE in our region.’
> *Forecaster – End of Day Survey*

> ‘Since the SPC Meso-a page starts with a RAP model background field, the ability to QC check this data will be helpful in gauging the accuracy of hourly RAP and HRRR model fields…Excluding the likely unreliable data in the region of lingering cloud cover across central Virginia, the NUCAPS data roughly ranged from 6.5C to 7.6C/km across the RNK CWA, which is fairly close to the SPC Meso-A 700-500 mb lapse rates.’ [Figure 7]

14 June 2022, Blog Post: *The Utility of Satellite Derived Data in Mesoanalysis & Near Term Convective Forecasting*
https://inside.nssl.noaa.gov/ewp/2022/06/24/the-utility-of-satellite-derived-data-in-mesoanalysis-near-term-convective-forecasting/

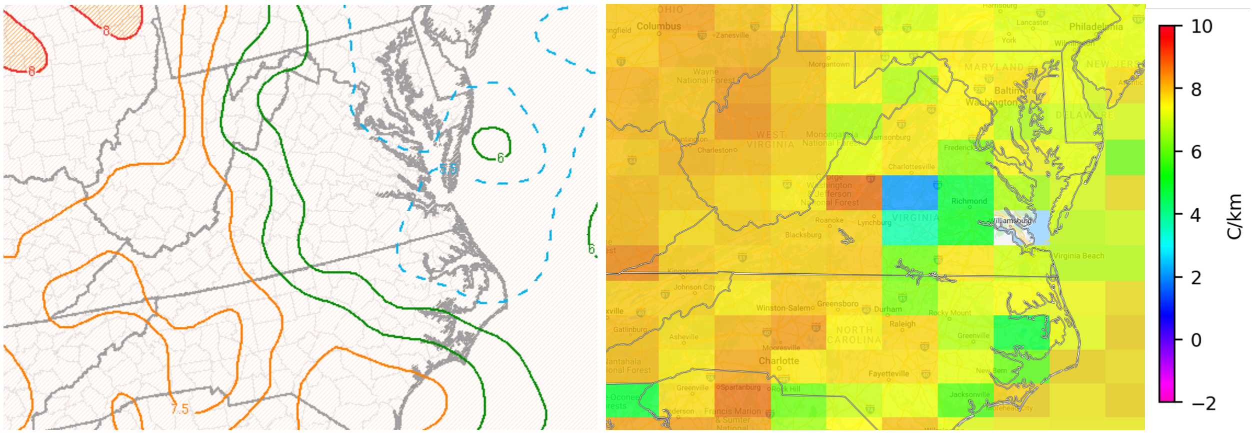

Figure 7: (Left) SPC Meso-A 700-500 mb lapse rates at 19z 6/14. (Right) NUCAPS 700-500 mb lapse rates at 1819z 6/14.

‘we had three overpasses right before the convection initiated…They all combine to increase confidence on what conditions COULD be in areas where the models are saying one thing and an observational system is indicating potential reality…In our case, the bulls-eye of 3000 J/kg in east central Minnesota is likely real and needs to be an area to watch’ [Figure 8]

14 June 2022, Blog Post: *LightningCast/NUCAPS and Isolated-to-Scattered Convection in FGF.*
https://inside.nssl.noaa.gov/ewp/2022/06/24/lightningcast-nucaps-and-isolated-to-scattered-convection-in-fgf/

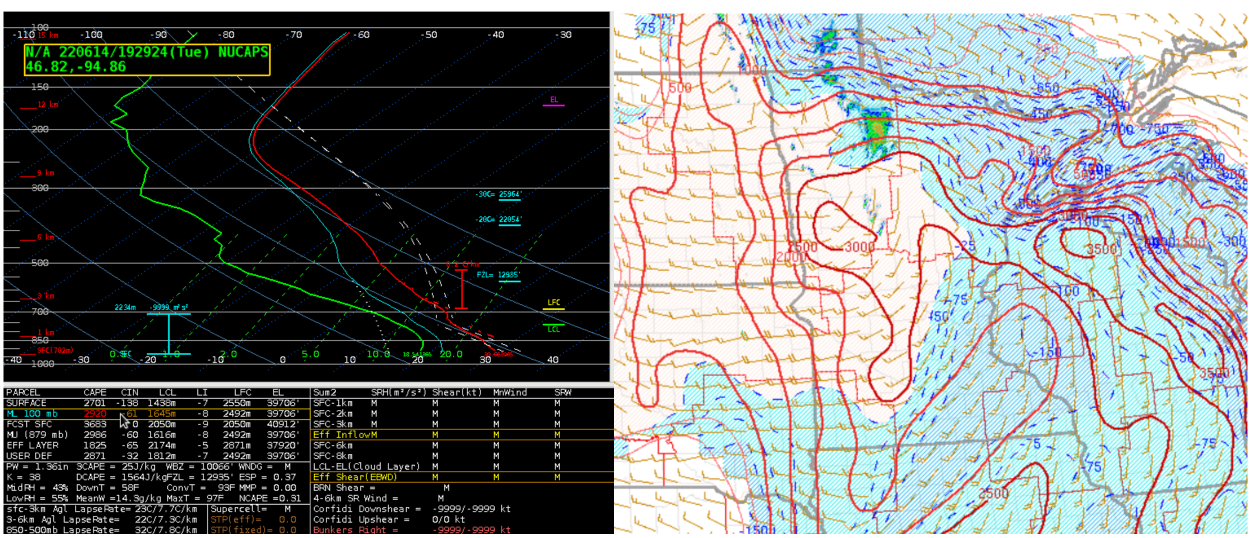

Figure 8: (Left) NUCAPS Profile from AQUA in north central Minnesota around 19Z on 14 June 2022. (Right) ML-CAPE and ML-CIN from the 21Z SPC Mesoanalysis data on 14 June 2022.

When both satellites were available, forecasters gravitated toward using information from consecutive overpasses from NOAA-20 and Aqua to validate mesoanalysis and RAOB data. The approximately one-hour difference in overpass times gave forecasters the opportunity to assess temporal changes in the thermodynamic profiles, often by monitoring for increasing CAPE, decreasing CIN, and changes in lapse rates (low-level and mid-level). When forecasters had these data available, they were asked how it applied to their usage. Across all three product suites, the leading response was ‘The increased temporal coverage allowed me to assess trends in the environmental or convective development’. The second most used response for NUCAPS Profiles and Modified NUCAPS stated ‘One satellite overpass was insufficient’, signaling that the use of thermodynamic profiles with updates on the order of 1 to 2 hours were useful. However, for Gridded NUCAPS the second most used response was ‘The temporal coverage was insufficient (there was too much time between overpasses)’, meaning that for gridded datasets, more rapid updates over larger domains appear necessary.

> ‘had a NOAA-20, and AQUA pass for the polar orbiting satellites [Figure 9], that we could then compare to the special observed sounding from GRB [Figure 10]…between Aqua and NOAA-20 you can see that the environment becomes much more moist over time (AQUA came around 19Z and NOAA-20 came around 18Z). The increase in temperatures and dew points in the low levels between the two NUCAPS soundings show that there was increasing low level lapse rates and increasing CAPE through time…you can see AQUAs vast improvement in the low level over NOAA-20. The one caveat seems to be the smoothing of the values in the mid levels.’
> 15 June 2022, Blog Post: *Pre-Convective Environment Across GRB*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/pre-convective-environment-across-grb/

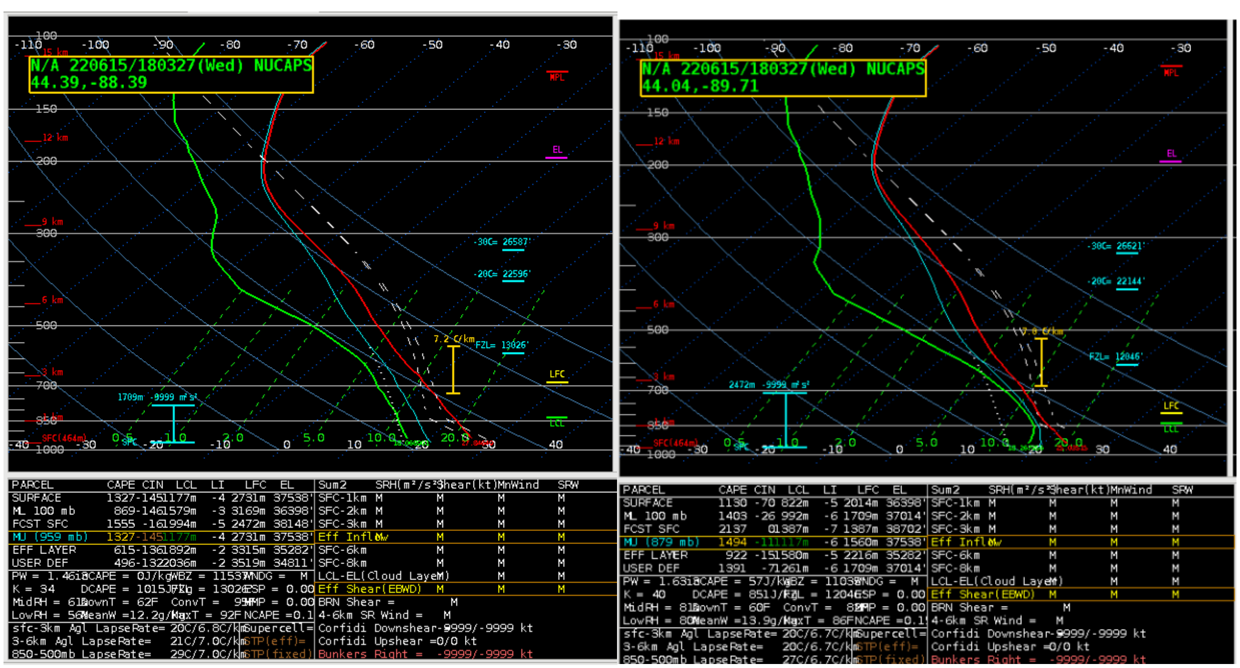

Figure 9: NUCAPS Profiles from NOAA-20 around 18Z (left) and Aqua around 19Z (right) in the Green Bay, Wisconsin CWA on 15 June 2022.

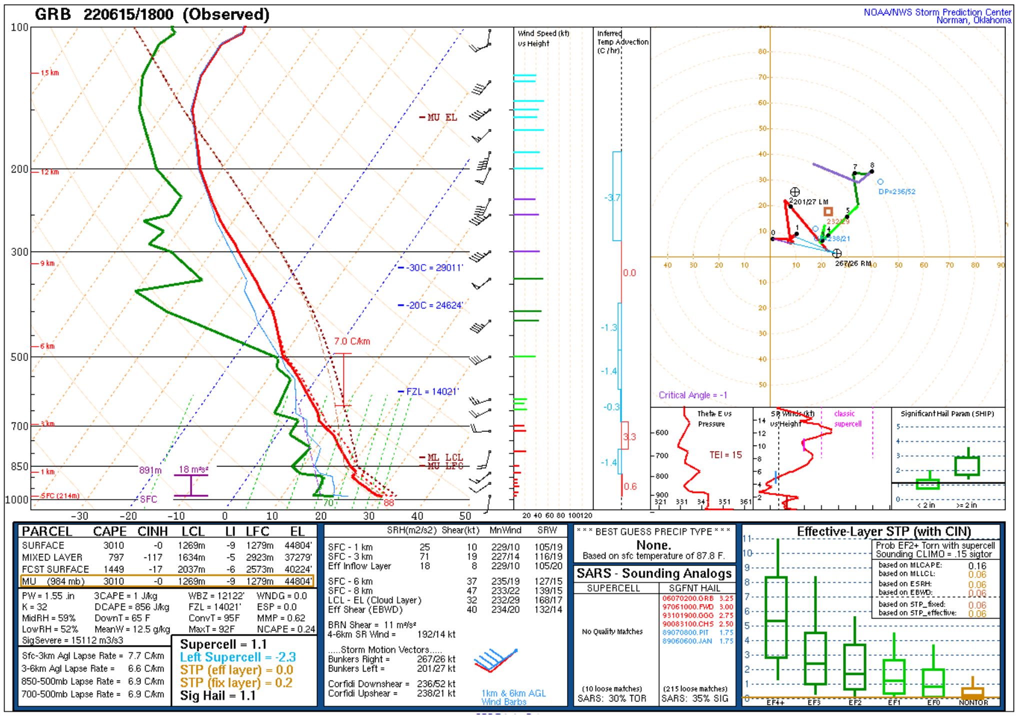

Figure 10: GRB sounding around 18Z on 15 June 2022.

> ‘Between the passes of NOAA-20, and Aqua there was enough of a time difference between them that you [can] see changes in the near term environment.’
> *Forecaster – End of Day Survey*

> ‘I wish that there was another overpass across BGM to see exactly how the environment had changed after the initial sounding/pass was made.’
> *Forecaster – End of Day Survey*

The adjusted temperature and moisture values in the Modified NUCAPS product provided another area of investigation for the forecasters, who often compared them to the original NUCAPS profiles. When asked in the weekly survey to describe the NUCAPS products, participants readily recalled key elements of the modified NUCAPS algorithm. This included modifying only values in the boundary layer and the use of GOES and RTMA data to make the adjustments. The following quotes from forecasters highlight the applications of Modified NUCAPS in the testbed.

> ‘I noticed a much better match between the gridded NUCAPs and PHS today in the instability fields. I also noticed a weaker CAP in the modified NUCAPS soundings in my limited time with those. I think the modified NUCAPS was much better with convection developing much more rapidly with the weaker cap.’
> *Forecaster – End of Day Survey*

> ‘NUCAPS and NUCAPS Modified soundings displayed [Figure 11] were along a warm frontal passage lifting south to north across southern Indiana. The modified NUCAPS sounding provided a more realistic depiction of the environment with higher levels of instability (CAPE and DCAPE). Furthermore, the LCL values represented in the

modified sounding were lower; depicting an environment more favorable for tornadoes along the warm front.'
8 June 2022, Blog Post: *Comparing NUCAPS Soundings Along Warm Front Passage*
https://inside.nssl.noaa.gov/ewp/2022/06/24/comparing-nucaps-soundings-along-warm-front-passage/

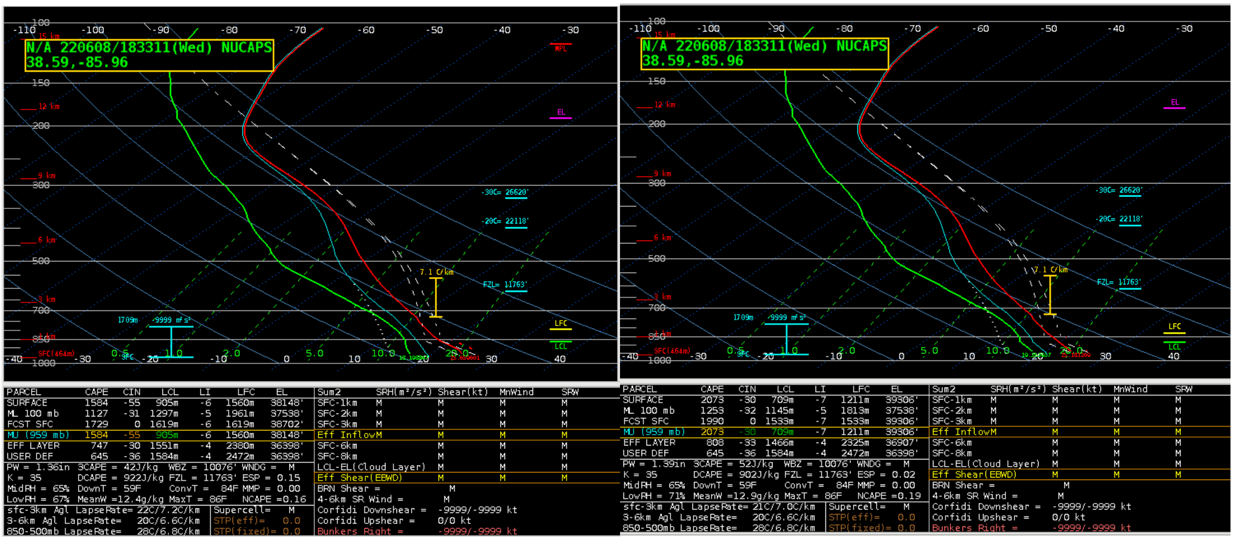

Figure 11: (Left) NUCAPS sounding from 8 June 2022. (Right) Modified NUCAPS sounding from 8 June 2022.

'Modified NUCAPS profiles did a decent job of representing the boundary layer. However, there were a few times where the modified NUCAPS profiles introduced too much moisture to the sounding and made the atmosphere more unstable that what it may have actually been.'
Forecaster – End of Week Survey

Data availability, driven by the timing and location of each NUCAPS overpass along with their latency, had the most impact on the utility of all three products across all convective scenarios. Among a list of common issues provided in the daily survey, for all three products, 'Data were not available when or where I needed them' was the most often selected choice. When asked about using data from multiple overpasses to assess temporal changes, 30 of 60 forecasters responded 'No, the data were not available'. In the weekly surveys, recommendations for NUCAPS profiles mentioned a desire for more satellite overpasses in 10 of 15 responses. While unfortunate timing or locations of the satellite overpasses with respect to initiating or approaching convection were prevalent, other issues included poor retrievals in dense cloud cover and product latency also relative to timing. Overall, the feedback from participants points to a desire for more consistent coverage and availability of satellite soundings.

'Definitely more satellite overpasses would be extremely useful. At this stage, we are limited by the time of the overpass, which can be either during convective minimum or maximum, depending on the time of day.'
Forecaster – End of Week Survey

> 'Dense cloud cover and a mis-aligned satellite pass led to the data being unavailable for my area. Soundings just to the west of my CWA were available, but no discernable features were seen.'
> *Forecaster – End of Day Survey*

> 'while the one pass we had was useful, no other passes came over our CWA in the times when we could have benefited from them.'
> *Forecaster – End of Week Survey*

Forecasters also provided feedback on the NUCAPS display, and additional ways to enhance the product. Periodically in daily discussions, forecasters listed the lack of vertical wind data in NUCAPS soundings as a future product improvement. While only temperature and moisture profiles are retrieved from NUCAPS, forecasters showed a desire to integrate wind data for a more complete view of the atmosphere. In terms of display enhancements, participants stated that the process of comparing soundings from different data sources (NUCAPS, RAOBs, models, etc.) was cumbersome; more so in NSHARP than SHARPpy. Regarding Gridded NUCAPS, this year's forecasters consistently disliked that information from all soundings were included in the display. This included profiles that did not pass the NUCAPS quality control algorithm. Some participants remarked that the contours and values in these areas appeared unrealistic and unhelpful. This feedback stands in direct contrast to responses from the earlier SPGs. Previously, only values that passed quality control were included in Gridded NUCAPS, but all quality values were displayed in more recent SPGs after previous participants wanted fewer gaps in the data. Examples of these recommendations can be found in the following quotes.

> 'AWIPS display is klunky for multiple sounding displays, whether temporal or spatial, making it hard to analyze raw NUCAPS soundings that way'
> *Forecaster – End of Week Survey*

> 'didn't entirely prevent use, but just struggled with workflow of making direct comparisons between NUCAPS and other sounding types in one easy panel'
> *Forecaster – End of Day Survey*

> 'The biggest draw back to me is the lack of wind data. Screen space is limited in an operational setting and incorporating some type of wind data like the RAP would be very beneficial to get a quick snapshot of the atmosphere all at one time. I also think the soundings are much easier to look at and interrogate in SHARPpy.'
> *Forecaster – End of Week Survey*

Recommendations for Operational Implementation

Based upon the evaluation of NUCAPS products in the 2022 HWT Satellite Proving Ground, the following items have been recommended:

- **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. These data may include the Aqua satellite, or the NOAA-21 satellite after its launch.

- **It is strongly recommended that improvements continue to be made in the display of NUCAPS data.** Suggestions for improvements include the ability to compare soundings between multiple data sources, and the use of Gridded NUCAPS in situations with poorly resolved profiles.
- **It is recommended that comparisons between NUCAPS and other sources of thermodynamic data continue to be emphasized in training to build forecaster confidence.** Training may consider the mesoanalyst position as its primary audience.
- **It is recommended that the integration of wind information with NUCAPS profiles and gridded products be considered as a product enhancement.** Potential avenues of development may include the use of wind data from models (e.g. RAP, HRRR, etc.) or adjacent RAOBs to create additional convective indices.

3.3 Optical Flow Winds

The Optical Flow Winds product from NOAA-NSSL and UW-CIMSS (Rabin 2021) was evaluated in the testbed for a second consecutive year. Wind data are created using 1-minute ABI imagery from the GOES-16/17 mesoscale scenes using an optical flow technique (Brox et al. 2004). The product provides wind vectors at a higher spatial (2 km) and temporal (5 minutes) resolution when compared to the GOES Atmospheric Motion Vectors (AMVs), with the intended purposes of monitoring divergence signatures near the overshooting tops of thunderstorms, vertical shear between cloud layers, and mesoscale boundaries. For the 2022 satellite testbed, the five most recent Optical Flow Wind outputs were displayed in a web-based tool for each mesoscale scene. Users could select color-coated wind barbs with a resolution of 10 km from five product levels (1000-800, 800-600, 600-400, 400-200, and 200-100 mb) along with an 'All' product which provided a color-enhanced image of wind speed across all pressure levels at a resolution of 2 km. For the layered wind vectors, enhanced imagery from the ABI Clean-Infrared band (10.3µm) was used as a background image for the products to depict fine changes in brightness temperature at a resolution of 0.1 K. Based upon participant feedback and recommendations from the 2021 HWT Satellite Proving Ground experiment, a divergence product was created and made available to forecasters in the web display. This new development was the primary motivation for demonstrating the Optical Flow Winds product again in the testbed.

Prior to the testbed, forecasters were given a product guide outlining how the product was made, features of the web-based tool, and case studies showing how the product responds in cases of severe convection. When deciding locations each day for warning and DSS operations, regions of interest coinciding with GOES-16/17 mesoscale scenes were given an elevated priority, so forecasters had more opportunities to use the Optical Flow Winds product. Overall forecasters were able to apply the Optical Flow Winds product to convection as outlined in the training, along with finding additional applications in pre-convective environments. Display modifications were frequently suggested by the forecasters, along with discussion of the suitability of the Optical Flow Winds product for NWS operations. These themes are described in the following section.

Use of Optical Flow Winds in the HWT

Throughout the testbed, forecasters used the Optical Flow Winds product to monitor updraft strengths through divergence signatures at cloud top, and to analyze features in the mesoscale environment such as upper-level jets and wind shear. These general applications were met with mixed reviews from the forecasters. In the daily survey responses, a minority of 39% (20/51) of responses indicated that the Optical Flow Winds product provided useful information not available from the currently operational GOES Atmospheric Motion Vector (AMV) product. This was a similar response rate to the same question asked in 2021 (33%, 14/43). Those who found these data useful noted the ability to monitor cloud top divergence signatures that supported trends from other products such as a GLM, GOES RGB products, and radar. This was especially true for discrete storms with stronger updraft speeds, and therefore robust cloud-top divergence signatures. Frequently the participants gravitated towards the new divergence product for these applications. In addition to cloud-top divergence signatures, forecasters also used the layered wind barbs to assess to identify divergence within the ambient environmental flow pattern and interpret low-level winds from the motion of cumulus clouds. Examples of these applications can be found below.

> ‘Optical flow winds showed the increasing storm divergence in the intensifying convection over the ARX CWA and also well sampled the strong mid and upper tropospheric winds indicative of strong deep layer bulk shear.’
> *Forecaster – End of Day Survey*

> ‘While this was helpful in a warning environment to look at storm top divergence and speed of the winds at the tops of clouds, I was able to find another great use for it. In the pre convective environment I had pulled up the Optical Flow Winds and noticed that it was tracking winds and speeds of clouds over Lake Michigan [Figure 12]. In an area where any wind information and observation data can be very sparse to near non-existent. The optical flow winds could be very helpful for open waters forecasting.’
> 15 June 2022: Blog Post, *Pre-Convective Environment Across GRB*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/pre-convective-environment-across-grb/

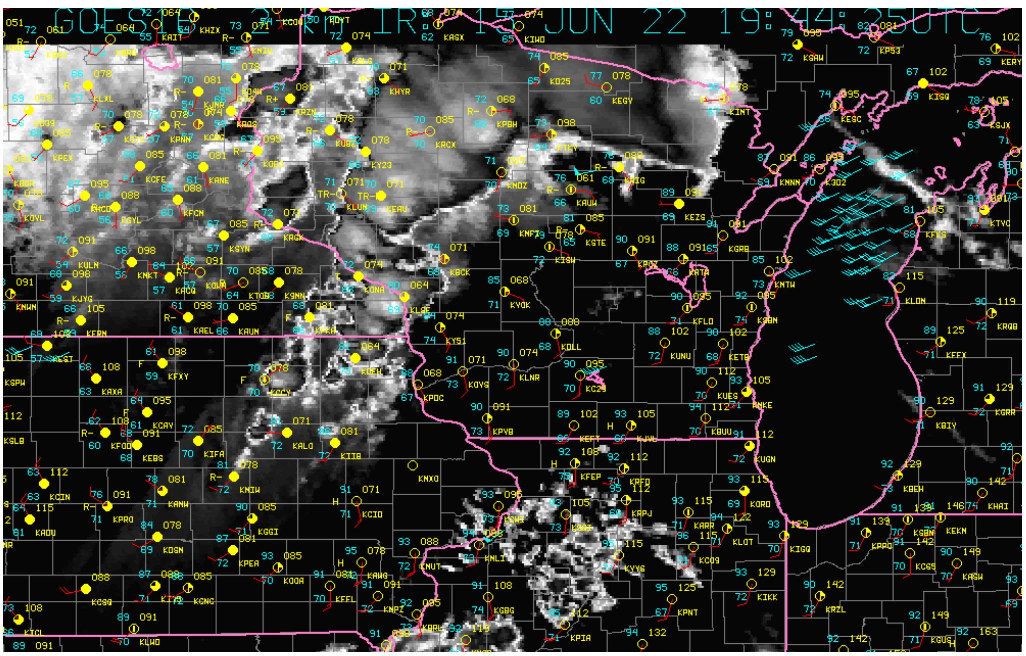

Figure 12: Optical Flow Winds output for 15 June 2022 at 1944 Z. 1000-800 mb winds (cyan) have been plotted on top of surface observations and enhanced-IR imagery.

> ‘Strong storm top divergence signals an intense or intensifying thunderstorm. Matched with upper level radar scans, satellite interrogation (clean IR), and ProbSevere, optical flow wind products may be another tool to aid the warning forecaster and/or storm scale mesoanalysis, Below are the corresponding optical flow wind storm top divergence images approaching 4 PM EDT (3 PM CDT) when the sig severe hail was reported [Figure 13]. The thunderstorm of interest is centered just south of eastern Lake Ontario.’
> 16 June 2022, Blog Post: *Observations from BUG on Thursday June 16th*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/observations-from-buf-on-thursday-june-16th/

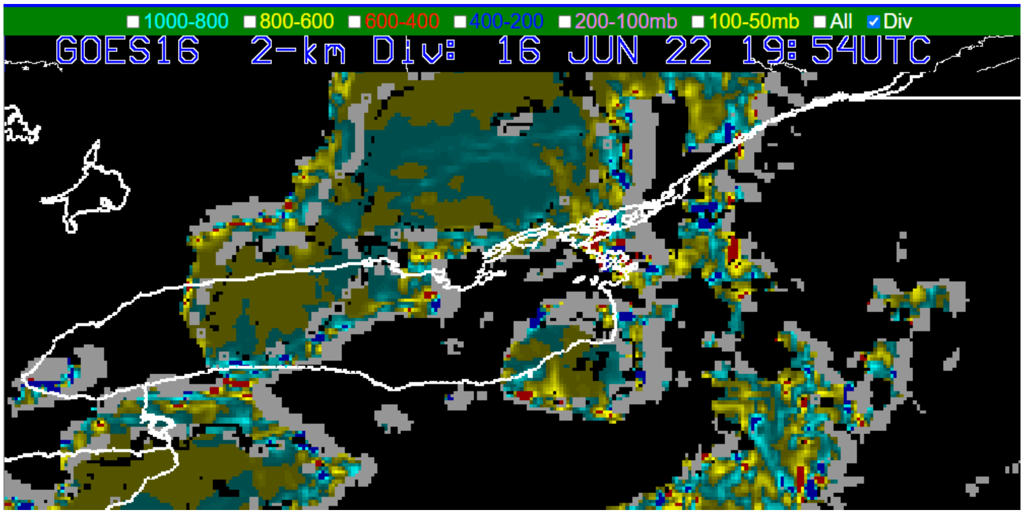

Figure 13: Display of the divergence field from the Optical Flow Winds product at 1954 Z on 16 June 2022.

> 'Once storms developed a bit more, you can pick out a slight directional divergence signal in the flow. Looking at SPC mesoscale analysis there is also directional divergence is present in the mesoscale analysis [Figure 14]…Overall looking at these two examples, using these tools together can be a way to verify information and give the meteorologist more confidence (or less) in a specific product to help them with forecasts, DSS, and warning operations.'
> 8 June 2022: Blog Post, *PHS, NUCAPS, Optical Flow, Prob Severe Fun*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/phs-nucaps-optical-flow-prob-severe-fun/

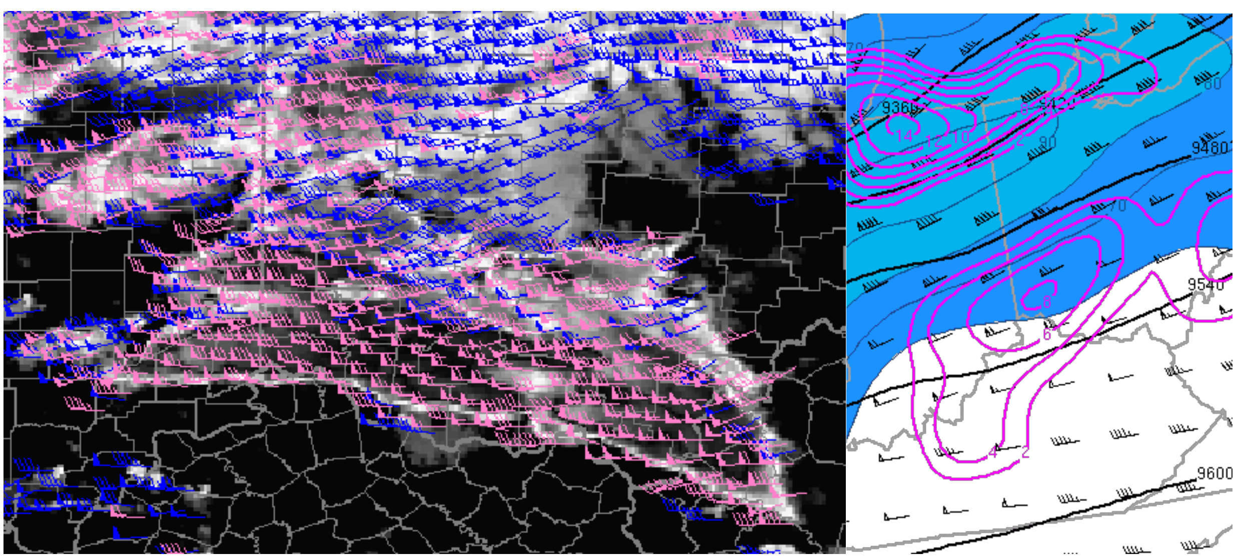

Figure 14: (Left) 400-200 mb (blue) and 200-100 mb (pink) winds from the Optical Flow Winds product on 8 June 2022 near 2100 Z. (Right) 300 mb geopotential heights (black), winds, and divergence (pink) fields from the 2100 Z SPC Mesoanalysis on 8 June 2022.

In blog posts, group discussions, and surveys, forecasters constantly mentioned the display of the Optical Flow Winds product as its primary limitation. Similar to feedback received in the 2021 HWT Satellite Proving Ground experiment, forecasters found the web interface of the product difficult to interpret. Geospatial referencing of cloud-top features with respect to convection displayed in AWIPS-II was mentioned as a limitation, as the mapping features were paired with specific wind layers in the product. Oftentimes, the forecasters also found the wind barbs too closely packed to identify features of interest in the full view of the mesoscale scene. Due to these limitations with the current web interface, forecasters often suggested that putting the Optical Flow Winds product into AWIPS may help alleviate these display issues. This was supported by 82% (14/17) forecasters recommending that Optical Flow Winds be displayed in AWIPS-II. Additionally, feedback regarding the new divergence product noted that the divergence signatures were noisy between each output, making it hard to identify trends in the strength of cloud-top divergence. Changes to the divergence color scale were recommended as well, with forecasters commenting that only coloring values above a certain value may help reduce the noise observed in the product. Further discussion of this feedback can be found in the following quotes from forecasters.

> ‘Found the display of Optical Flow Winds to be difficult to see, due to low resolution of underlaying satellite data. The wind barbs were also clustered too close together to really get a good sense of smaller scale features like storm-top divergence if it existed. Would prefer to be able to change the satellite imagery beneath to another channel or RGB, Day-Cloud Phase Distinction RGB for example.’
> *Forecaster – End of Day Survey*

> ‘The display online was difficult to read with many overlapping wind barbs. Some layers like the dark blue color were very hard to see over the darker satellite background. I imagine this would be much easier to read/interpret in AWIPS.’
> *Forecaster – End of Week Survey*

> ‘The divergence field is really hard to gather meaningful intel from and the existing platform outside of AWIPS limits its overall usage. A suggestion in our group today was that divergence could be useful if the noise is limited. Perhaps remove values above and below a certain threshold. Instead of widespread values, draw attention to the important outliers.’
> 24 May 2022, Blog Post: *MAF Testbed Observations*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/maf-testbed-observations/

> ‘There is a lot of variability from scan to scan on the strength of the divergence field but there is enough of a signal to figure out where the strongest couplets could be and which storm tops they could be associated with. We couldn’t overlay radar data or the 3.9 micron “Red/visible” channel with a divergence product to make a 1:1 comparison; something to consider would be a grid that could be overlaid on a different ABI image to do a visual comparison to this product.’[Figure 15]
> 15 June 2022, Blog Post: Supercells and Optical Flow Winds
> https://inside.nssl.noaa.gov/ewp/2022/06/24/supercells-and-optical-flow-winds/

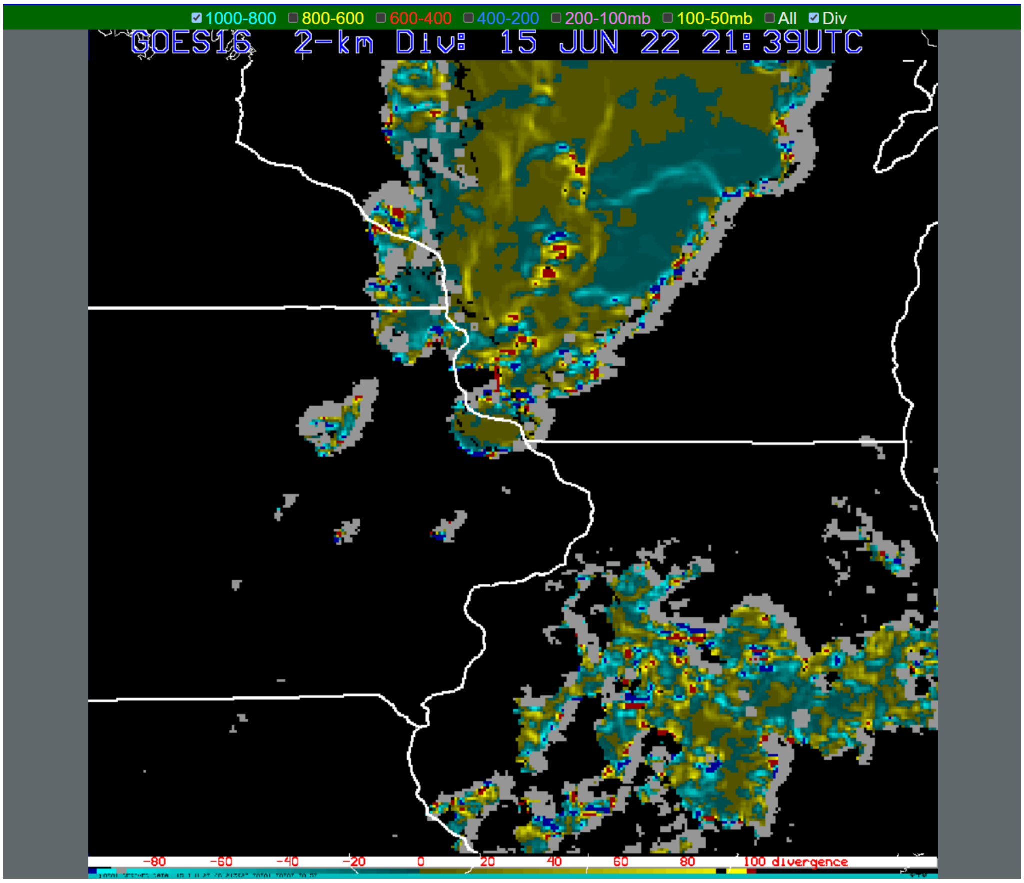

Figure 15:Optical Flow Winds divergence fields from 15 June 2022 at 2139 Z.

Recommendations for Operational Implementation

Based upon the evaluation of the Optical Flow Winds product in the 2022 HWT Satellite Proving Ground, the following items have been recommended:

- **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.** Improvements may include additional flexibility in geolocation layers, removing divergence values below a set threshold, modification of the divergence color bar, or a transition of the product suite to AWIPS-II.
- **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.** Training efforts may include intercomparisons between the Optical Flow Winds product and GOES Atmospheric Motion Vectors, case studies of divergence signatures prior to severe weather occurrence which coincide with similar signatures from radar or satellite imagery, and recorded presentations.

3.4 PHSnMWnABI Model

New to the HWT Satellite Proving Ground in 2022 was the PHSnMWnABI Model. Information from hyperspectral sounders and microwave imagers on the NOAA-20, Suomi-NPP, and MetOp-B/-C satellites in low-Earth orbit are integrated with the GOES-R ABI, and then ingested into a numerical weather model similar to the Rapid Refresh (RAP) and the High-Resolution Rapid Refresh (HRRR) models (Smith et al. 2020). Temperature and moisture data from the low-Earth orbiting satellites are associated with information provided by the GOES-R ABI, stepped forward in time, and then used to initialize the numerical model for each forecast hour. The PHSnMWnABI model is initialized every two hours, with hourly output available out to nine hours, a horizontal resolution of 3 km, and covering nearly all the central and eastern continental United States. The expectation of this numerical model was to leverage the high spectral resolution of the hyperspectral sounders and microwave imagers in low-Earth orbit with the high spatiotemporal resolution of the ABI, improving the initial conditions for the model and providing more accurate forecasts of the mesoscale environment coincident with deep convection.

Prior to the testbed, participants were provided with a training video, a 'one-pager' overview, example applications, and a user's guide. Within the testbed, forecasters were able to access three PHSnMWnABI products at the 3 km resolution in AWIPS-II, with a larger suite of model output available online in a web display. The three products available in AWIPS-II were the lifted index (LI), most-unstable convective available potential energy (MU-CAPE), and the significant tornado parameter (STP). Procedures made using the PHSnMWnABI data to compare the PHSnMWnABI values of LI and MU-CAPE against the LI and MU-CAPE values available from the GOES-R L2 products. Forecasters most often used PHSnMWnABI model outputs from 18 Z and 20 Z each day, and compared model fields to the SPC Mesoanalysis and NCUAPS data to interrogate the ambient mesoscale environment within their forecast area. The forecasters frequently referred to the PHSnMWnABI model as 'PHS', and may be used in the following sections when quoting blog posts.

Use of PHSnMWnABI in the HWT

Throughout the testbed, forecasters consistently used the PHSnMWnABI model output to help identify the location of convection initiation, and the characteristics of ongoing thunderstorms. Gradients in MU-CAPE were used by forecasters to locate low-level boundaries, their ability to initiate convection, and their movement over subsequent forecast hours. Initially this was referenced by the PHSnMWnABI developers at the beginning of each week, and in group discussions was well received by the forecasters. Comparisons were frequently made between the model and similar data sources like NUCAPS and the SPC Mesoanalysis. Similar to the applications of NUCAPS, the PHSnMWnABI model provided context for the certainty, or uncertainty, of the ambient mesoscale environment. In group discussions each week, forecasters noted that the PHSnMWnABI was suited for the mesoscale analyst role within WFOs during warning operations. Once convection initiated or entered their forecast area, the participants continued to use PHSnMWnABI model output to investigate how thunderstorms would evolve over the coming hours. The following quotes from blog posts highlight the previously described applications.

‘Down in Tallahassee, there are two boundary layers where storms are initiating or ongoing. There is a lingering MCS that moved down from the Midwest overnight and a Sea Breeze. You can see the CAPE gradients along the both boundaries and how that progresses forward in time with both the boundaries interacting with one another… This tracks very well when you overlay the visible satellite imagery with the PHS images as you can see the cumulus field along the CAPE gradient [Figure 16]. This gives a good visualization of where storms are initiating along the Sea Breeze and the strongest storm movement along the MCS.’
14 June 2022, Blog Post: *Storm Movement and Severity at TAE*
https://inside.nssl.noaa.gov/ewp/2022/06/24/storm-movement-and-severity-at-tae/

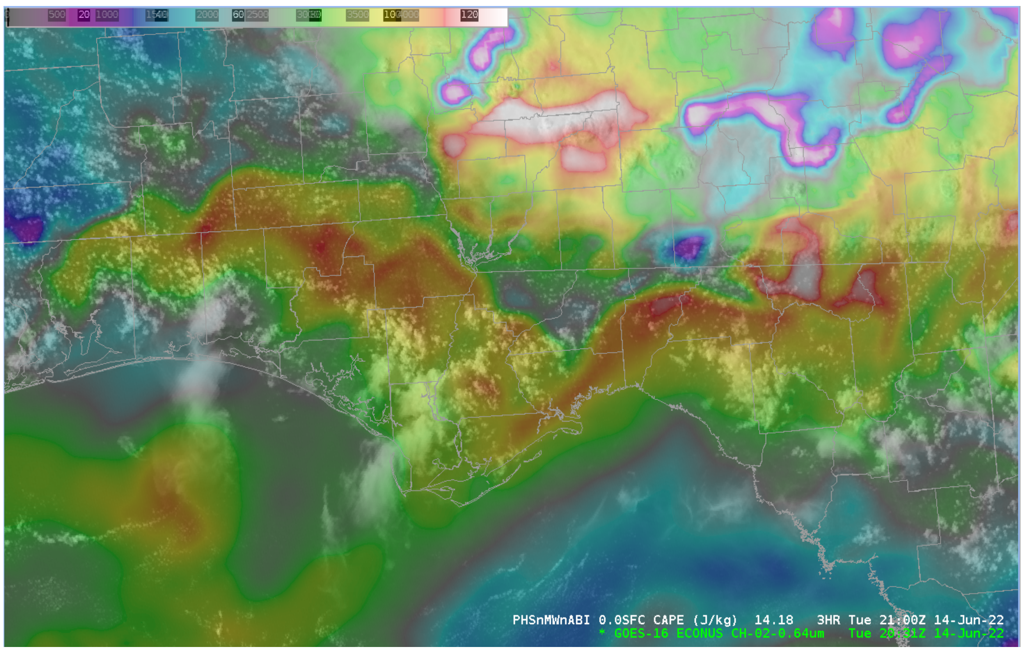

Figure 16: An overlay of the PHSnMWnABI MU-CAPE on top of the ABI visible band (Ch 2).

‘We decided to compare the output between PHS, NUCAPS, and SPC mesoscale analysis. The variable that we chose was surface CAPE… They all seemed to match up well highlighting the higher instability to the south that would gradually push north this afternoon. It was definitely a confidence builder in each product to see the agreement between them.’
6 June 2022, Blog Post: *PHS, NUCAPS, Optical Flow, Prob Severe Fun*
https://inside.nssl.noaa.gov/ewp/2022/06/24/phs-nucaps-optical-flow-prob-severe-fun/

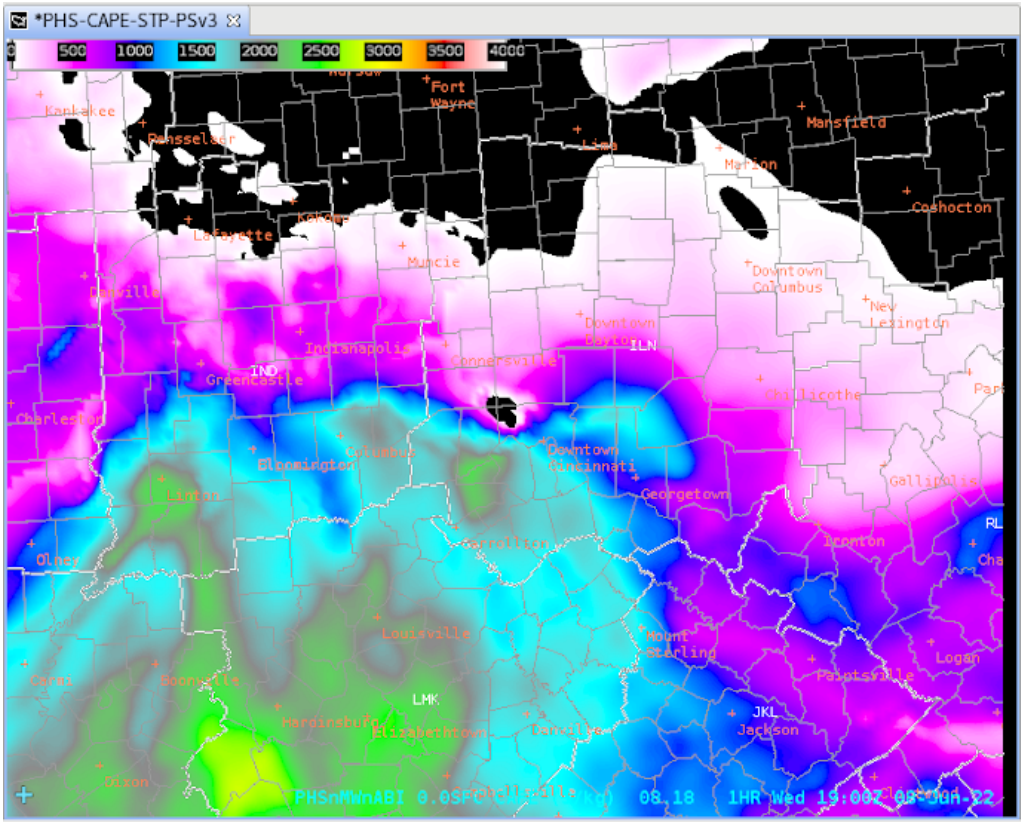

Figure 17: MU-CAPE from the PHSnMWnABI model across southern Ohio, southern Indiana, and Kentucky on 8 June 2022.

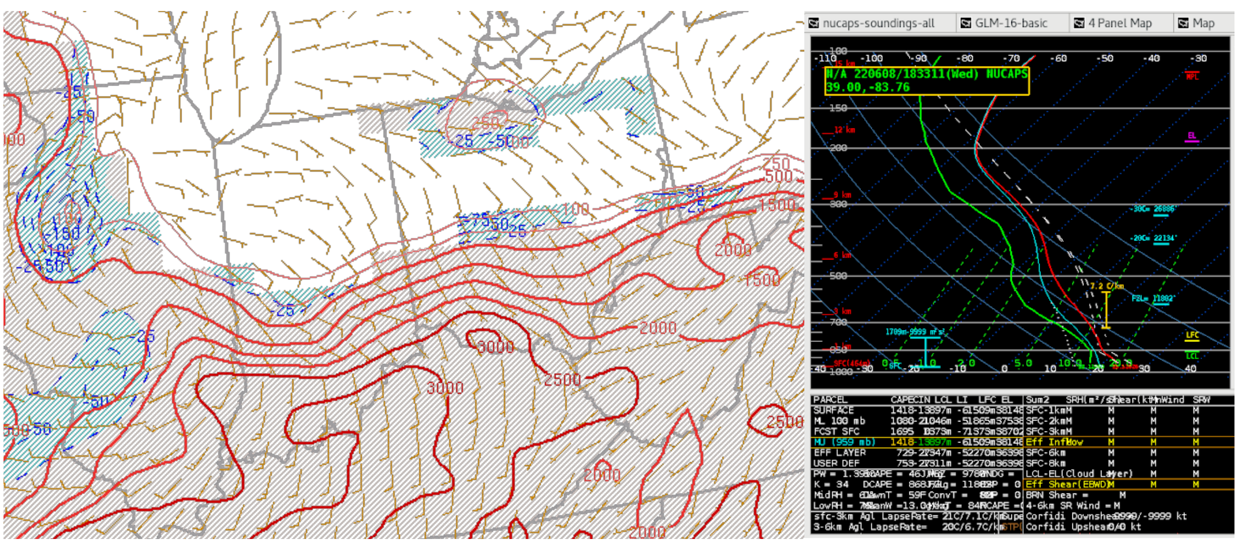

Figure 18: (Left) SPC Mesoanalysis of CAPE (red contours) at 18 Z on 8 June 2022. (Right) NUCAPS Profile from far southwest Ohio at 1833 Z on 8 June 2022.

'The PHS Sfc CAPE procedure was helpful in diagnosing the mesoscale environment [Figure 19]. In particular, the depiction of CAPE gradients matched well with where the Day-Cloud-Phase Distinction showed where these boundaries lay as could be construed from the cumulus field [Figure 20]…The PHS Sfc CAPE depicted this surface boundary

migrating southward through the central NE through the 21Z-00Z time frame. Observed cells moving left (east) off the boundary into a more stable environment as resolved by the PHS Sfc CAPE field all decreased in intensity and saw their convective updrafts weaken…Seeing this after the first hour raised forecast confidence in the forecast thinking of today's severe weather potential and was shared in a graphicast for this test case scenario.'
6 June 2022, Blog Post: *PHS CAPE Gradient Use in Mesoanalysis*
https://inside.nssl.noaa.gov/ewp/2022/06/24/phs-cape-gradient-use-in-mesoanalysis/

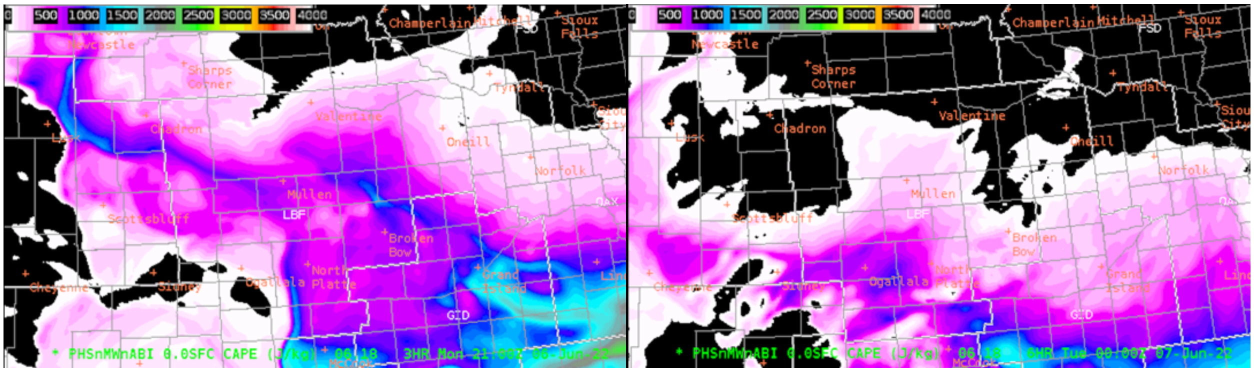

Figure 19: PHSnMWnABI output of MU-CAPE at 21Z (left) and 00Z (right) from the 6 June 2022, 18 Z run.

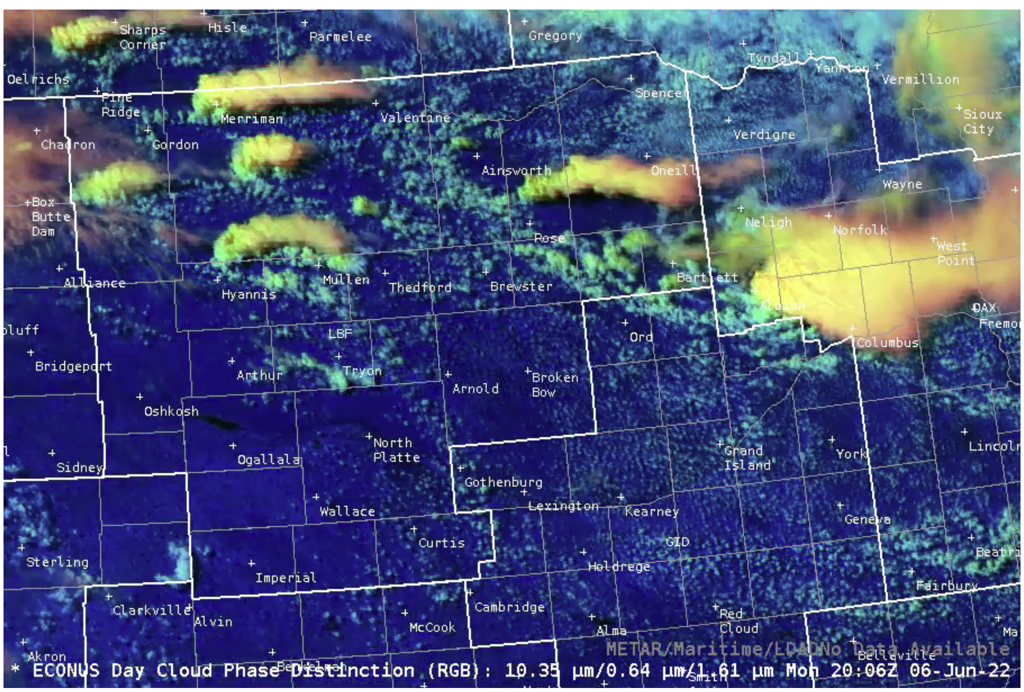

Figure 20: Day cloud phase distinction RGB from 2006 Z on 6 June 2022.

Due to its use in interrogating the mesoscale environment, the PHSnMWnABI model was often not directly referenced by forecasters when making warning decisions and DSS messaging. However, participants who did cite the PHSnMWnABI model in the warning/DSS form reinforce the previously mentioned applications. From the total of 52 DSS communications reported in the testbed, the model was mentioned as a contributing factor six times. These often used the MU-CAPE and LI fields available in AWIPS-II, and their trends over the forecast period, to better inform their DSS messaging. Of the 36 total graphical forecasts created, three included output from the PHSnMWnABI model with two directly concerning mock DSS events. While a few forecasters used the model output in these graphics, they did note sharing the fields in AWIPS-II would be reserved for decision makers with a stronger meteorological background, or additional commentary would be needed for clarification. Amongst the 82 warnings issued during the experiment, seven directly mentioned the PHSnMWnABI model as an influence in their warning decisions, with a majority referring to forecasted STP values as one factor when coupled with current observations like radar, GLM, and ProbSevere v3.

> DSS Message: 'UPDATE: Moderate showers are approaching the area within the next 30-40 minutes. No lightning is expected at this time.'
> Forecaster Thoughts: 'Continued downward trend in the lightning cast product and the GLM showing no significant increases with the "showers" or residual convection on the extreme southern edge of this line of convection. PHS also indicates very little available instability at this time with a minimum indicated on the SPC mesoanalysis.'
> *Forecaster – DSS Messaging Form*
> *Kids Triathlon – Tuscaloosa, AL, 25 May 2022*

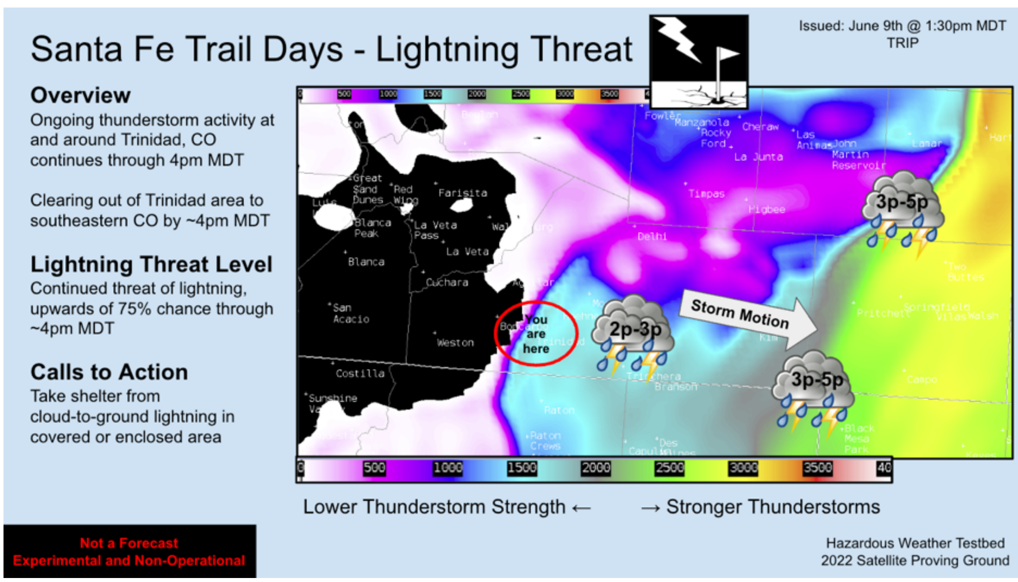

Figure 21: A public graphic featuring the PHSnMWnABI MU-CAPE field from a mock DSS event on 9 June 2022 in Colorado.

> 'Based on the PHS forecast combined with satellite analysis, we were able to focus the convective threat for the Madison area toward 6PM and onward, tied to the stronger forcing and better moisture arriving from the west where the ongoing convection resided closer to the cold front. It appears that the PHS sampling of moisture in the column applied to the near-term forecast strongly outperformed the SPC/RAP Mesoanalysis model background and OA algorithm.'
> 15 June 2022, Blog Post: *Can PHS Improve Mesoanalysis and Near-Term Convective Forecasts?*

https://inside.nssl.noaa.gov/ewp/2022/06/24/can-phs-improve-mesoanalysis-and-near-term-convective-forecasts/

Feedback from the daily and weekly surveys echoed the utility of the PHSnMWnABI model as described previously and provided some additional quantitative insights. Forecasters overwhelmingly felt that the PHSnMWnABI model gave a more accurate view of where strong convection was more likely, with up to 92% (49/53) of responses supporting this statement. In terms of the forecast hours most frequently used by participants, 45% (25/56) of forecasters stated that they used the 0-2 hour forecasts, 50% (28/56) stated they used the 2-4 hour forecasts, and 5% (3/56) stated they used the 4-6 hour forecasts. While no forecasters stated that they used the forecasts beyond six hours, the approximately four hours spent in the simulated operations each day during the testbed may have disincentivized participants from viewing data beyond four hours.

Additionally, participants stated that they wished there were more model fields from PHSnMWnABI available in AWIPS-II, including moisture, shear, and convective inhibition. While these parameters were available online, forecasters frequently commented that the online display of additional model fields were difficult to integrate into their analysis in real time. Factors such as the static display of images and the inability to quickly change between forecast hours, model runs, and parameters were brought up each week in group discussions. Visual improvements to the model fields in AWIPS-II were also discussed, including changing the default color tables to match those used for the corresponding GOES-R fields.

Recommendations for Operational Implementation

Based upon the evaluation of the PHSnMWnABI model in the 2022 HWT Satellite Proving Ground, the following items have been recommended:

- **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.** Verification of the PHSnMWnABI model forecasts against other mesoscale sources of environmental data such as the HRRR, RAP, and the SPC Mesoanalysis is also encouraged to show the unique value of this satellite data integration technique into numerical weather prediction.
- **It is strongly recommended that the web display of PHSnMWnABI model output improve to aid in future integration and testing for NWS forecasters.** The ability to display model fields quickly and intuitively by forecast hour, model run, and region should be considered.
- **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.** The fields most requested from forecasters include moisture data (dew points and relative humidity), shear parameters, convective inhibition, and convective indices like those in the SPC Mesoanalysis, RAP, or HRRR.

3.5 Probability of Severe (ProbSevere) LightningCast Model

The NOAA/CIMSS Probability of Severe (ProbSevere) LightningCast model (Cintineo et al. 2022) was evaluated for the first time in the 2022 HWT Satellite Proving Ground Experiment. A set of four images from the GOES-R ABI's spectral bands (Band 2: 0.64 µm, Band 5: 1.6µm, Band 13: 10.3µm, and Band 15: 12.3µm) are used in a machine-learning model to predict the probability that the GLM will observe lightning in the next 60 minutes. The model learns spatial and multi-spectral features from the training data that are important for short-term lightning forecasting. LightningCast was trained using the four predefined ABI bands as input, and 60-minute accumulations of the GLM Flash Extent Density product as truth (i.e., maximum flash-extent density at every point over the ensuing 60 minutes). Output probabilities are displayed as contours with the option for parallax-corrected data, available for the GOES-16/-17 CONUS and mesoscale scenes, and accessible both day and night.

Prior to the testbed, forecasters were provided with a training video describing how the ProbSevere LightningCast model generates its output, the definition of the lightning related probabilities, and example applications including forecasting the initiation of lightning and decision support service. Additionally, AWIPS-II procedures overlaying the LightningCast model on the Day Cloud Convection RGB, GLM Flash Extent Density, and the MRMS Isothermal Reflectivity were provided to forecasters. The default probability contours were set to 10%, 25%, 50%, and 75%. Overall, forecasters heavily used the LightningCast product for mock DSS events, along with maintaining situational awareness of areas where deep convection may be initiating. At times, the forecasters referred to the ProbSevere LightningCast model as 'PS LC' or 'LC'. Applications and recommendations of the LightningCast product from forecasters are provided in the following section.

Use of ProbSevere LightningCast in the HWT

Amongst all six product suites demonstrated in the 2022 HWT Satellite Proving Ground experiment, the LightningCast model was by far most often used to support forecasters in the mock DSS events. Forecasters mentioned the LightningCast model as a contributing factor in 41 of the 52 DSS messages (79%) submitted throughout the experiment. Trends in LightningCast probabilities, associated with approaching or departing storms, at the DSS event location were consistently mentioned throughout the experiment. Times when the default probability contours crossed the mock DSS event location were also noted for creating 'decision points' for forecasters to communicate thunderstorm hazards to their partners. Based on daily survey results, forecasters most often felt that LightningCast provided them 10-19 minutes (36%, 20/56) and 20-29 minutes (32%, 18/56) of actionable lead time for the advection of storms. The following blog posts from testbed participants highlight these applications.

> 'Initially the LightningCast for our DSS event surged to near 50% or slightly above. This was an initial concern for the DSS area. As these storms weakened, the probabilities of lightning also fell to under 25%. I liked that these probability decreases were not rapid, but a gradual fall after the initial peak. GLM and LightningCast both had a consistent drop in probability and lightning activity as the "storms" weakened…While an initial picture of the probabilities looks concerning, pairing this with other satellite products for

context and seeing the overall trend of this data led to an easy decision to wait for additional data.'
25 May 2022, Blog Post: *DSS in the Birmingham CWA*
https://inside.nssl.noaa.gov/ewp/2022/06/24/dss-in-the-birmingham-cwa/

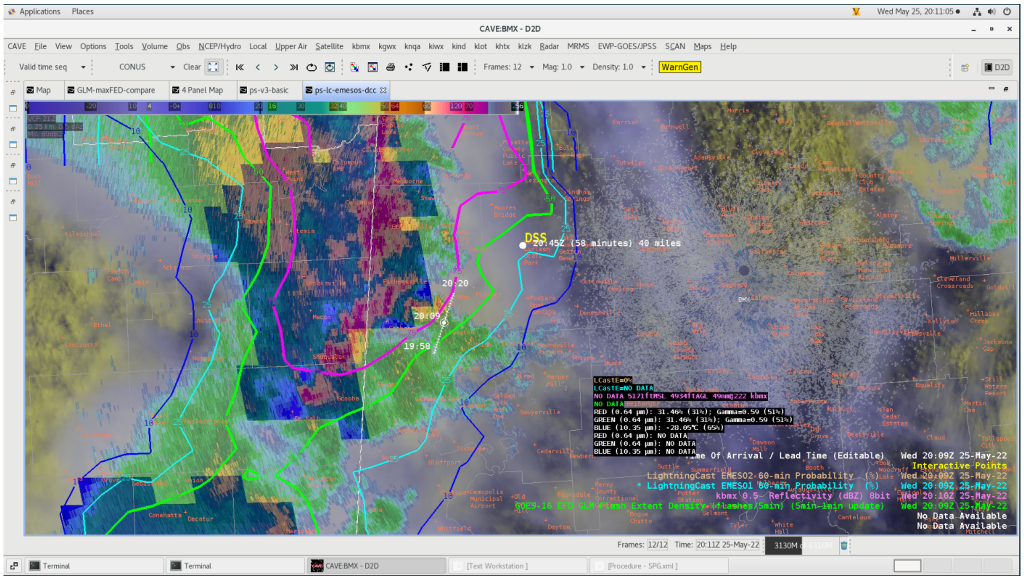

Figure 22: ProbSevere LightningCast overlaid with radar reflectivity.

'I utilized the Lightning Cast to provide a probable end time of the lightning threat for the Riverfest event in La Crosse, WI. This was a valuable tool as it provided some added confidence when the storms would exit the event area. I did my best to line up the TOA tool with the 25 percentile contour. Once I got my estimated time that the end of the lightning threat would reach the event, then I added about 30 minutes to ensure it was well east of the event circle.'
15 June 2022, Blog Post: *GLM Parallax and LightningCast Fun*
https://inside.nssl.noaa.gov/ewp/2022/06/24/glm-parallax-and-lightning-cast-fun/

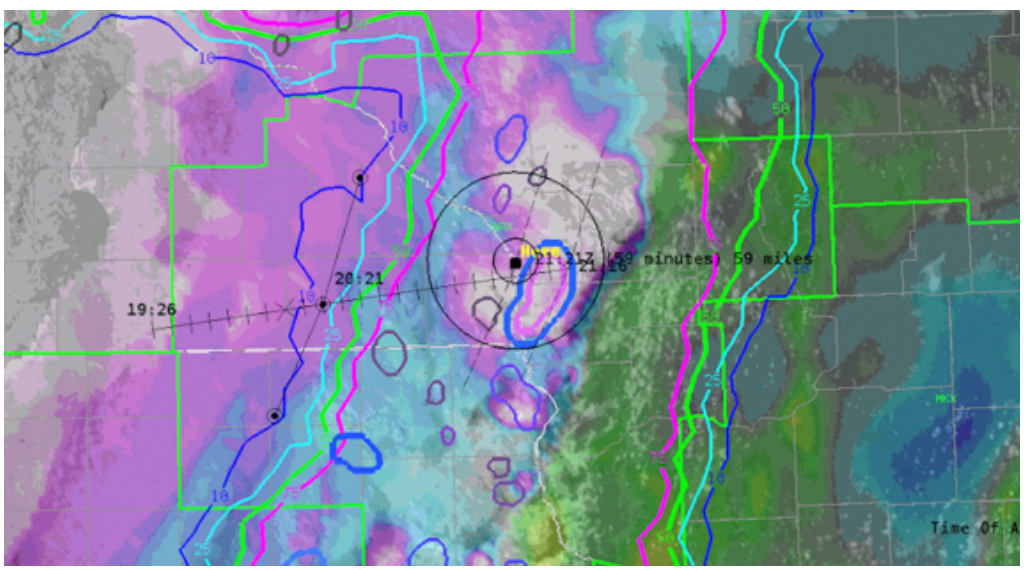

Figure 23: ProbSevere LightningCast model output overlaid on ProbSevere v3, PHSnMWnABI, and ABI data. 'Home' represents the DSS location and the AWIPS time of arrival tool was used.

On top of its frequent use in mock DSS event messaging, the LightningCast model was featured in 18 of the 36 forecast graphics forecasters created for the public and DSS partners. Messaging of these graphics focused on the movement of probability contours toward and away the participant's forecast area or simulated DSS event. Annotations on the forecast graphics often provided timing guidance for the incoming threat of lightning. Probability contours depicted the current lightning hazard to an area when thunderstorms were ongoing and provided forecasters a way to communicate the initiation and movement of lightning risk while in operations. Forecast graphics from two mock DSS events are provided below as examples.

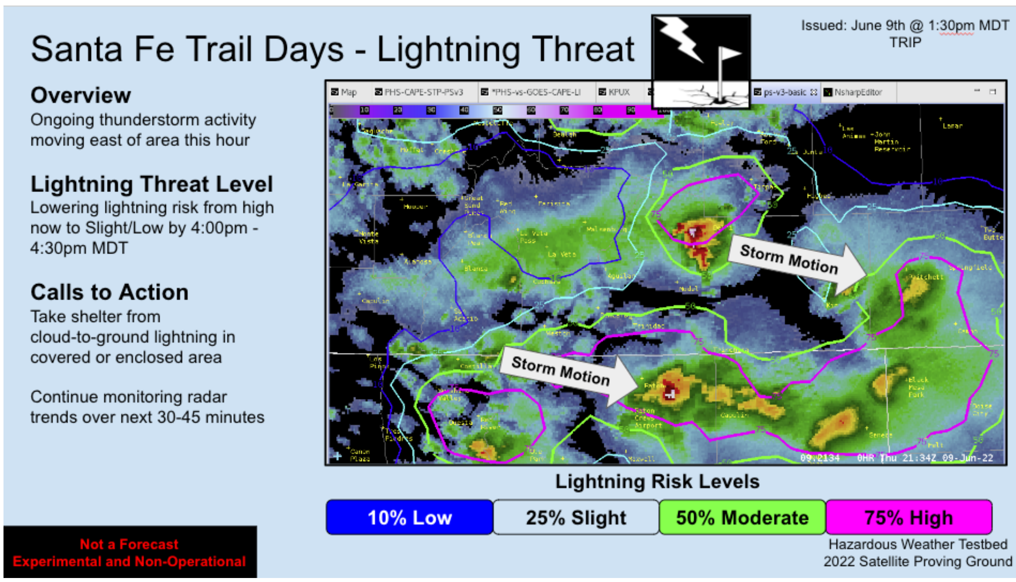

Figure 24: A public forecast graphic featuring the LightningCast model from 9 June 2022 for a simulated DSS event in Colorado.

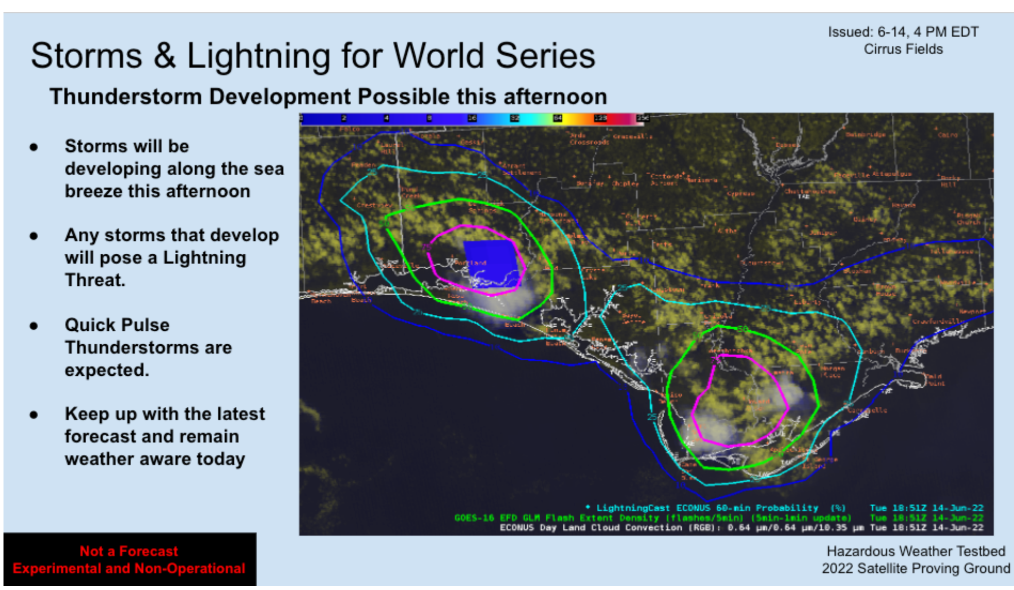

Figure 25: A public forecast graphic featuring the LightningCast model from 14 June 2022 for a simulated DSS event in Florida.

ProbSevere LightningCast's predictive ability for initiating thunderstorms was frequently scrutinized by participants in the testbed, both for mock DSS events and general situational awareness. Similar to the forecaster's responses regarding LightningCast's average actionable lead time for the advection of storms, participants felt that the LightningCast model most frequently provided 10 to 19 minutes (36%, 20/56) and 20 to 29 minutes (32%, 18/56) of actionable lead time for lightning initiation. Increasing LightningCast probabilities in areas showing no lightning activity helped forecasters confirm, or in some cases adjust, areas of focus for initiating thunderstorms.

Forecasters noticed that the LightningCast model struggled in areas of initiating thunderstorms under mid-level and upper-level clouds. Probabilities in these scenarios would remain low (below 25%) before the first GLM flash, or not achieve the 10% contour before the first GLM flash. In response to this, forecasters suggested in group discussions that the LightningCast algorithm look to incorporate information from radar. In the weekly survey, forecasters rated proposed research efforts from the ProbSevere LightningCast developers, and 'Incorporate radar data as model predictors into LightningCast' received the highest average rating, with 43% of forecasters listing this as their highest priority. The following blog posts from forecasters highlight applications of the ProbSevere LightningCast model for initiating convection.

> 'Initially we were thinking storms would form first over the southern portion of the CWA based off of modeled convective parameters but at 2101Z, a 25% chance of lightning popped up via LC for the north-central portion of the forecast area. This 25% contour appeared a few minutes before radar reflectivity started showing up for the same area. Seeing even the 10% contour show up earlier on, clued me into the fact that we needed to shift our focus further north than we originally thought.'
> 9 June 2022, Blog Post: *LBF HWT Blog Day 4*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/lbf-hwt-blog-day-4/

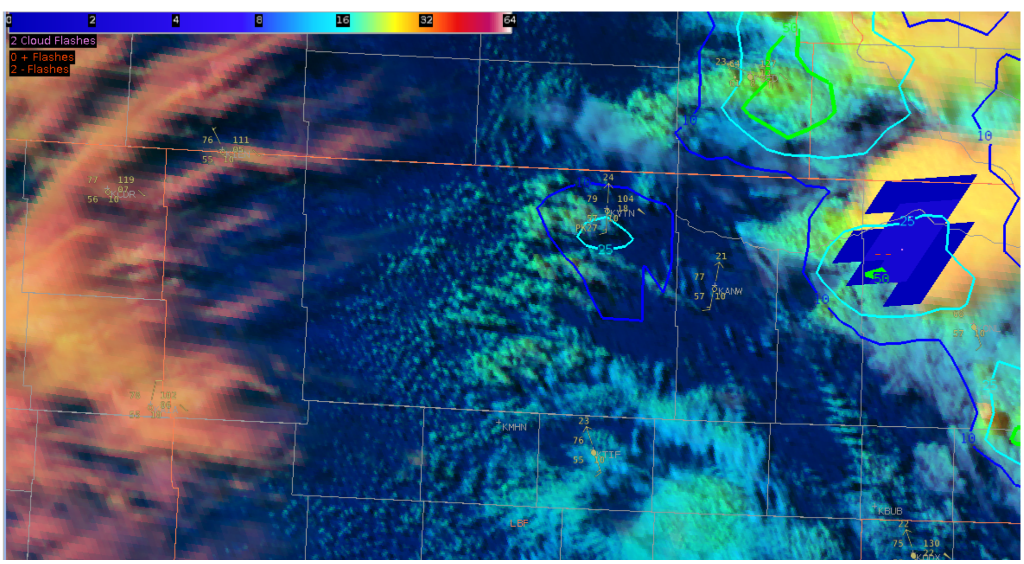

Figure 26: ProbSevere LightningCast overlaid on GLM Flash Extent Density and the Day Cloud Phase Distinction in northern Nebraska on 9 June 2022.

> 'a few weaker cells had developed under a mix of sun and high clouds…we noticed that although the -10C MRMS data would suggest lightning, the LightningCast data was actually much lower than anticipated (below 10 percent)… it was surmised that this may be a case where convective debris or high clouds could mask the signal of weaker convection…Eventually LightningCast did highlight this area with a lightning risk, but this case is relevant because if we are using this tool to aid in DSS to our partners, we must also identify some potential limitations.'
> 6 June 2022, Blog Post: *Potential LightningCast Limitations Under Convective Debris*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/potential-lightningcast-limitations-under-convective-debris/

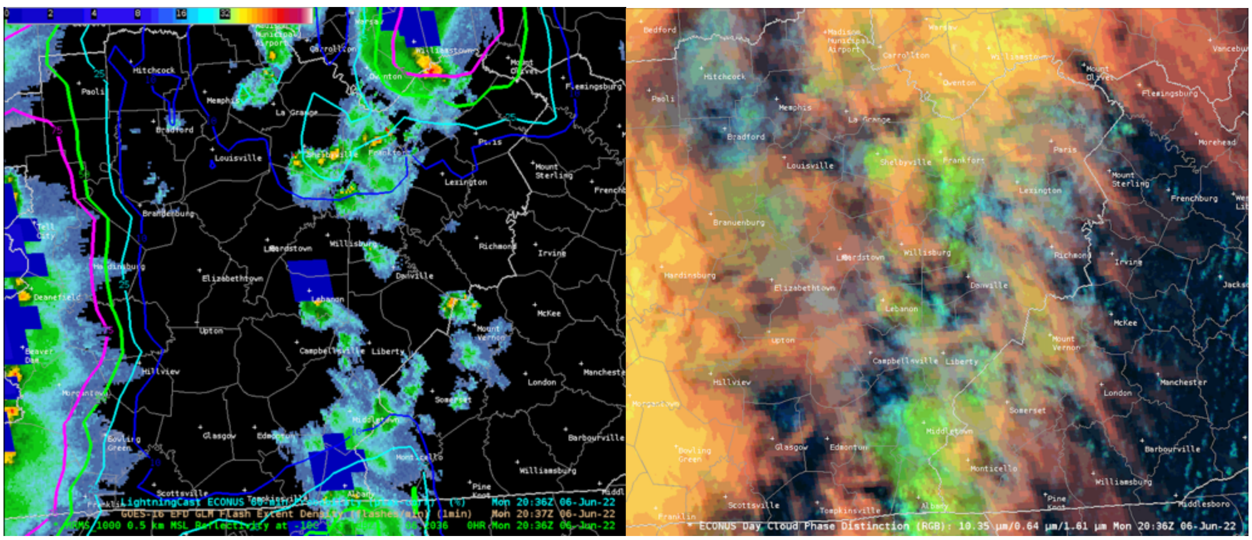

Figure 27: (Left) ProbSevere LightningCast overlaid on GLM Flash Extent Density and MRMS Isothermal Reflectivity at -10˚C. (Right) ABI Day Cloud Phase Distinction RGB.

Overall, forecasters showed confidence in using the ProbSevere LightningCast model in convective monitoring for situational awareness and in mock DSS events. When asked about the calibration of LightningCast probabilities in the daily surveys, approximately 88% of responses (58/66) found the model's output well calibrated, with only about 11% finding it too high (7/66). In the weekly surveys, forecasters were asked how often they would use the LightningCast model in AWIPS, along with the tasks they would use it for. Of the 19 participants, 7 responded they would 'Always or Almost Always' use LightningCast, and 12 responded they would 'Regularly' use the model. The three less frequent options (Occasionally, Seldom, and Never) were not selected by any forecaster. In terms of applicable tasks for the model, all forecasters responded that they would use the LightningCast model for 'DSS support' and 'Convection initiation or convective maintenance situational awareness'. Forecasters also frequently responded that they would use the LightningCast model for terminal aerodrome forecast amendments (17 of 19) and airport weather warnings (11 of 19). While beyond the scope of the 2022 HWT Satellite Proving Ground experiment, this presents another potential avenue for applying the ProbSevere LightningCast model in NWS operations.

Forecasters were also asked in the weekly surveys and end-of-week discussions for ways to improve the LightningCast model based on their experiences in the testbed. One topic that was brought up each week in the testbed, along with the surveys, was the potential applications of a point-based time series tool in AWIPS. This was often compared to the time series tool available in ProbSevere v3, described in the following section. In AWIPS forecasters used the time-of-arrival tool to predict when thunderstorms, and their associated LightningCast probabilities, would approach or leave areas of interest like DSS locations. The LightningCast model was also available to view in the University of Wisconsin's Space Science and Engineering Center's RealEarth website, which contained a point-based time series tool. While not frequently used during the experiment, those forecasters who did view LightningCast data in RealEarth remarked that this tool was useful for following trends in the modeled probabilities. Additional comments by forecasters often included the display of probability contours, or best practices when integrating the LightningCast model with other products. Some forecasters showed a preference to overlaying the probability contours on the Day Cloud Phase Distinction RGB over the Day Cloud Convection RGB used in the pre-built AWIPS procedures. Additionally, frequent use of the LightningCast product in graphical forecasts for the public and DSS partners placed an increased emphasis on which contours were 'significant' for decision points and their ability to be displayed in format readily understood by the general public.

Recommendations for Operational Implementation

Based upon the evaluation of the ProbSevere LightningCast model in the 2022 HWT Satellite Proving Ground, the following items have been recommended:

- **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.** Applications for aviation-related forecasts and products may warrant future investigation.
- **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.** Additional themes may include limitations in areas of dense cloud cover, best practices for display in AWIPS-II, and how to communicate these data to the public.
- **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.**

3.6 Probability of Severe (ProbSevere) Model – Version 3

The NOAA/CIMSS Probability of Severe (ProbSevere) Model (Cintineo et al. 2014) was evaluated again in the HWT, and was the second to include Version 3 (v3) of the model. ProbSevere v3 is a statistical model that provides probabilistic guidance regarding the occurrence of severe weather within the next 60 minutes, updating every two minutes, along with its associated hazard models for severe hail (ProbHail), severe wind (ProbWind), and tornadoes (ProbTor). Several fundamental changes have been made to v3 in an attempt to improve model calibration (Cintineo et al. 2020). ProbSevere v3 leverages specific ENTLN, GOES-16 GLM, GOES-16 ABI, Multi-Radar/Multi-Sensor (MRMS), and SPC Mesoanalysis data into new machine-learning models for each hazard type, producing output in the form of polygons centered around each storm.

Forecasters were able to display ProbSevere and individual hazard data in AWIPS in the form of storm-relative contours. Contours could then be sampled to provide additional probability information, along with the most common meteorological parameters used to produce the probabilities. A new addition to ProbSevere v3 was the ability to display probabilities of each object as a time series by double-clicking the contour. Prior to the testbed, forecasters were provided a training video discussing how v3 differs from v2 in terms of input variables, calibrated probabilities from each model, and how to access the time series tool. AWIPS-II display procedures were also made to overlay the ProbSevere v3 storm objects on MRMS variables which frequently are the leading predictands for ProbHail, ProbWind, and ProbTor. This frequently involved MRMS Composite Reflectivity, Maximum Expected Size of Hail, Vertically Integrated Liquid (VIL), and Low-Level (0-2 km) Azimuthal Shear.

Use of ProbSevere v3 in the HWT

Along with liberally using the pre-made display procedures, forecasters frequently overlaid the ProbSevere products on data from individual radar or satellite imagery to aid in the real-time interrogation of convective hazards. Per the weekly survey, all forecasters used the ProbSevere (all-hazards-in-one) product to display in AWIPS-II during the week, while approximately half used the individual hazards. Throughout the entirety of the experiment, the ProbSevere v3 model was liberally used in warning operations to quickly identify and monitor the severity potential of thunderstorms. Amongst the 82 reported warnings from forecasters in the experiment, ProbSevere v3 was mentioned as a contributing factor to their warning decision 91% (75/82) of the time; the most of any product demonstrated in the experiment. Trends in hazard probabilities were cited directly in 37% (28/75) of all ProbSevere mentions, with the next highest factor being probabilities exceeding specific values in approximately 7% (5/75) of responses.

In the daily surveys, 94% (51/54) and 91% (39/43) of forecasters agreed that ProbSevere v3 helped increase their confidence and lead time for issuing warnings, respectively. For tornado warnings, 73% (16/22) of forecasters felt that ProbSevere v3 helped increase their confidence issuing tornado warnings, while 61% (11/18) felt it improved their lead time. Group discussions revealed similar results from the surveys and warning report form, with forecasters consistently using trends from ProbSevere output to bring attention to or support timely warning decisions. The blog posts from forecasters help to highlight these applications.

‘As convection developed, we also practiced relying on probSevere and probTor for lead time in anticipating warnings. The following shows an example where the probTor trends corresponded well with ARX’s actual decision to issue a tornado warning (Figure 28)… Similarly, intensification of the convective line appeared to be well detected. In fact, depending on what threshold of the probSevere parameters is relied on (probably depends on environment and other factors), the escalating value could have given useful lead time for a severe issuance decision.’
15 June 2022, Blog Post: *Situational Awareness and Lead Time with LightningCast and ProbSevere Tor*
https://inside.nssl.noaa.gov/ewp/2022/06/24/situational-awareness-and-lead-time-with-lightningcast-and-probsevere-tor/

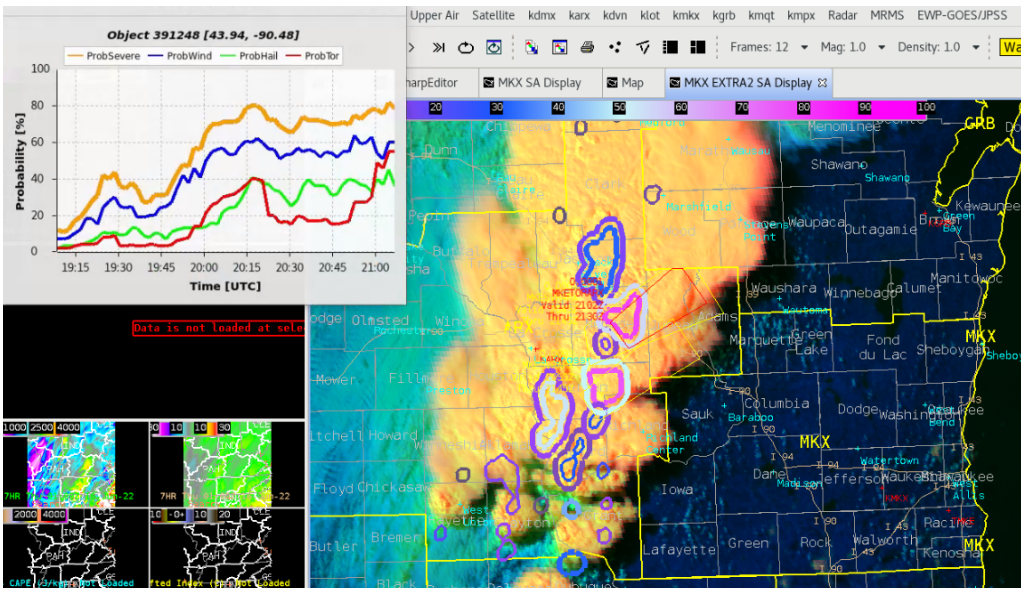

Figure 28: A display of ProbSevere storm objects and a time series plot of a storm object in western Wisconsin on 15 June 2022.

‘The ProbSevere v3 drew attention to a cell near Bloomington, IN on the afternoon of June 8th. Specifically it was drawing attention to the wind threat with a high ProbWind percentage reaching 50% to 60% (Figure 29). However, radar interrogation was only initially yielding velocity of 30-40kts on the inbound side of the cell, sub-severe but certainly SPS worthy. There were a few scans yielding up to 45kts as well. So initially there was a little wondering why there was higher ProbWind percentages. Quick realization of the storm track being perpendicular to the radar beam could explain the lower velocity signatures here. Knowing this, ProbSevere and its ProbWind portion was drawing attention to a scenario that otherwise may have been missed if only looking at velocity data and not realizing the storm track relative to the radar beam.’
8 June 2022, Blog Post: *ProbWind and ProbTor Performance*
https://inside.nssl.noaa.gov/ewp/2022/06/24/probwind-and-probtor-performance/

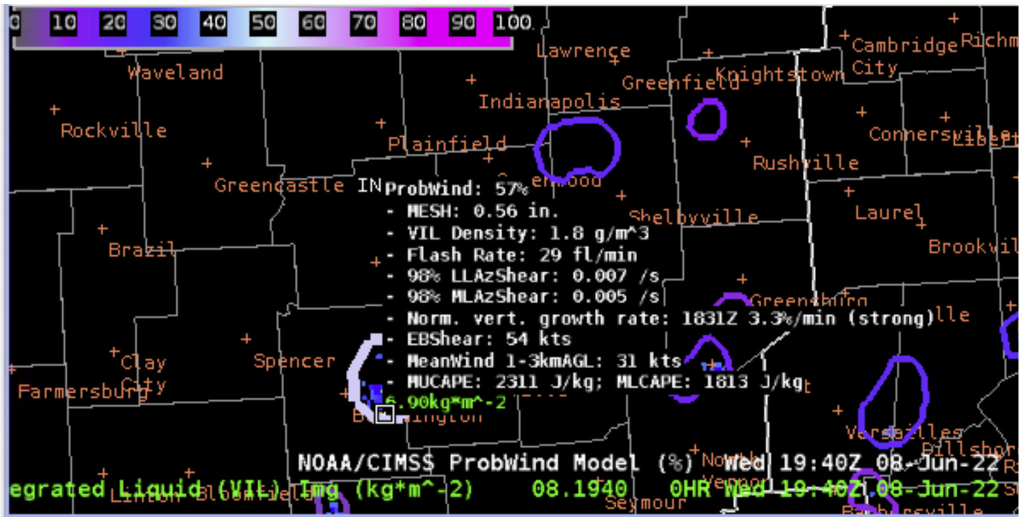

Figure 29: ProbWind sampled output for a storm in southern Indiana on 8 June 2022.

Considerable changes to the ProbSevere v3 model provided forecasters the opportunity to compare it with the currently operational ProbSevere v2 model. When asked to compare the performance of ProbSevere v2 and v3 in the daily surveys, approximately 54% (37/69) of forecasters responded that ProbSevere v3 was either ‘much better’ or ‘slightly better’ than ProbSevere v2. Nearly 10% responded that their performances were similar, and only 4% (3/69) responded that v2 was ‘slightly better’ or ‘much better’ than v3. Additionally, almost 32% (22/69) participants responded that they did not compare the two models that day, meaning that they primarily used ProbSevere v3 during the day without needing to use the probabilities from v2 for comparison. Responses from the weekly surveys when comparing v2 and v3 were similar, with 17 of 19 forecaster preferring v3 over v2, and 2 of 19 stating that they were unsure. In discussions with forecasters, it was frequently stated that the relative probabilities between each individual hazard ProbSevere v3 (ProbHail, ProbTor, and ProbWind) appeared to better match the hazards expected from mesoanalysis and current observations than v2. Additionally, participants mentioned each week that sequential outputs in v3 were more consistent than v2, making it easier to quickly identify trends in the probabilities. The following blog posts from forecasters highlight these comparisons.

> ‘ProbSevere V3 has been showing much improvement over version two, with this particular day featuring somewhat pusley mixed-mode storms across upstate New York…The ProbSevere V3 time series reflected the marginally severe nature of these storms very well, with values peaking at 60 to 65%. The very encouraging sign was the peaks and valleys in the ProbSevere V3 time series that showed this group of cells peaking at over 50% severe probabilities, dipping below 50% as the new updraft takes over, and then once again peaking above 50% once the new updraft strengthened. I’d definitely recommend forecasters to take a look at the timeseries to build confidence in cell trends during warning applications.’
> 16 June 2022, Blog Post: *ProbSevere v3 Adjusting During New Cell Formation*
> https://inside.nssl.noaa.gov/ewp/2022/06/24/probsevere-v3-adjusting-during-new-cell-formation/

> ‘A great example of the differences between PSv3 and PSv2 is with one of the strongest storms of the day for the Tallahassee CWA (Figure 30). Version 2 seems to try to

highlight a hail threat at 48% while Version 3 has prob severe hail at 6%. It seems that Version 2 is overestimating the Hail threat for this area. Especially given the subpar mid level lapse rates at 5 C/km or less, storm motion of around 5 kts or less and the upper level subsidence. The storm environment just is not conducive to produce quarter sized or larger hail.'
14 June 2022, Blog Post: *Storm Movement and Severity at TAE*
https://inside.nssl.noaa.gov/ewp/2022/06/24/storm-movement-and-severity-at-tae/

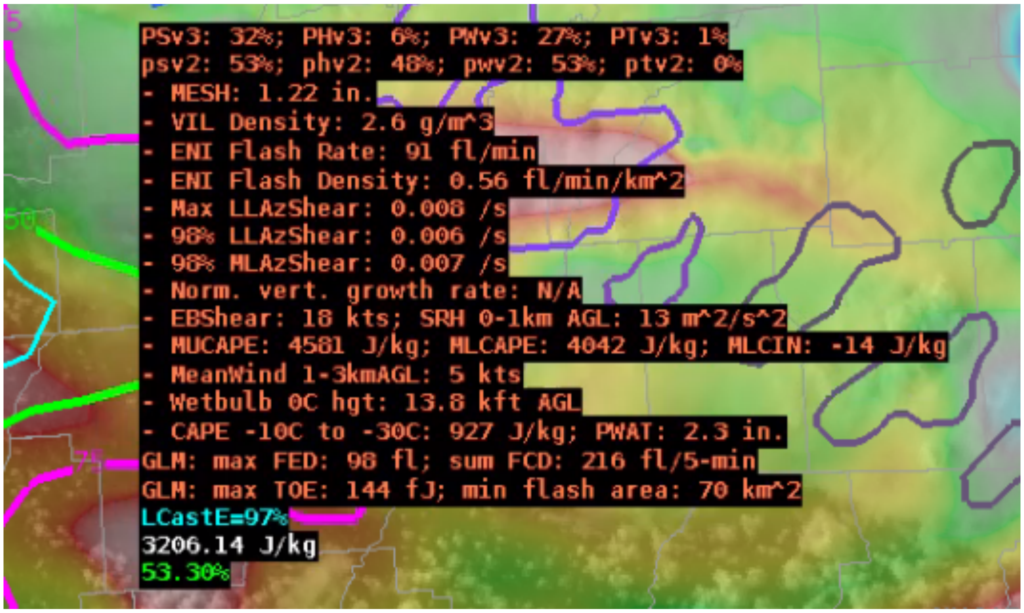

Figure 30: Sampled output from a ProbSevere v3 storm object in Florida, highlighting the differences between the ProbSevere v3 and ProbSevere v2 probabilities.

Another new feature in ProbSevere v3 was the addition of a time series tool, which displayed probabilities from each selected storm object in a time series graph. In discussions with forecasters, use of the ProbSevere time series tool was often mentioned to maintain situational awareness and observe trends in the model probabilities. In daily surveys, all respondents (45) indicated that the ProbSevere time series was useful in their warning decision process. When asked how often forecasters would use the ProbSevere time series tool in CAVE if available to them, all but two responded with 'All of the time' or 'Most of the time', the highest two categories available. The lowest two categories, 'Seldom' and 'Never', received zero responses. Forecasters were also able to interrogate ProbSevere v2 and v3 model output online in a web display, which also included time series plots for several key variables the model uses. In daily discussions a few forecasters did express a desire to have these input variables included in the AWIPS-II time series display. The following quotes from forecasters highlight the utility of the ProbSevere time series display in AWIPS-II.

> 'Getting the component time-series that are available in RealEarth into AWIPS so we can see how different parameters may be influencing ProbSevere output. It isn't a big deal to move over to the PC and look at them that way but having the data in AWIPS would make it more usable in a severe weather event by saving a bit of time from having to move from AWIPS to a computer.'

Forecaster – End of Day Survey

'When analyzing a thunderstorm developing over western South Dakota, a noticeable jump occurs near 20:25 – 20:30 UTC as seen on the ProbSevere Time Series (Figure 31). At this same time, there was a distinct uptick in lightning activity seen in the GLM 4 panel. This would correlate with a strengthening of the thunderstorm at this time… This thunderstorm was beginning to exhibit severe hail potential.
6 June 2022, Blog Post: *ProbSevere v3 and NUCAPS*
https://inside.nssl.noaa.gov/ewp/2022/06/24/probsevere-v3-and-nucaps/

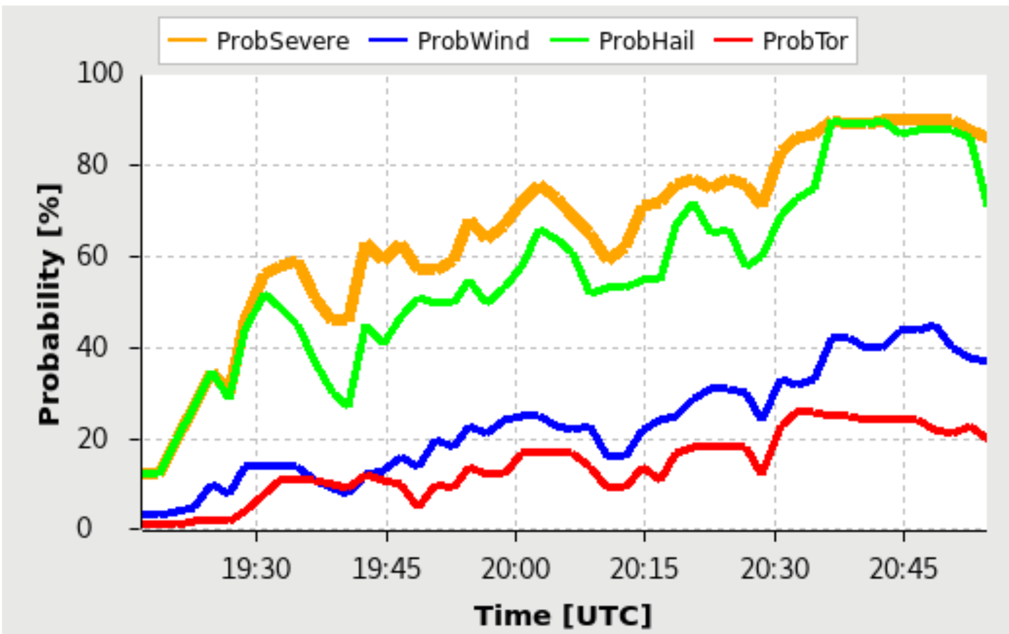

Figure 31: ProbSevere v3 time series output for a storm in South Dakota on 6 June 2022.

Along with recommendations regarding the new time series tool, forecasts also provided recommendations to improve ProbSevere v3 performance and visualization in AWIPS-II. One topic brought up this year during these conversations was the fact that data from the SPC Mesoanalysis are fed into ProbSevere v3, and these fields change hourly. This may cause the probabilities to suddenly change at the time of ingest in rapidly changing mesoscale environments coincident with convection. Similarly, a few forecasters noted errors in the storm object identification algorithm which then led to probabilities that were less calibrated to the storm feature of interest. If storm objects became too large, certain signals became more muted and decreased the sensitivity of the model. Lastly, forecasters mentioned that sampling storm objects to view the probabilities and model input fields could take up large amounts of screen space, especially when in multi-panel displays.

'At 2005Z the KIND radar showed a solid and persistent tornado signature on a cell near Shelbyville, IN. To go along with this in the overall outlook for the day was an MD highlighting the tornado risk and a Tornado Watch that included the southern half of IN. However, the ProbTor parameter did not tick up in response to the tornado signature on radar. The observed uptick at 2000Z, coincided with the intake of the new SPC mesoanalysis data that ProbSevere v3 ingests as described by the product's providers. Even still, this uptick was small only amounting to a ~5% increase.'

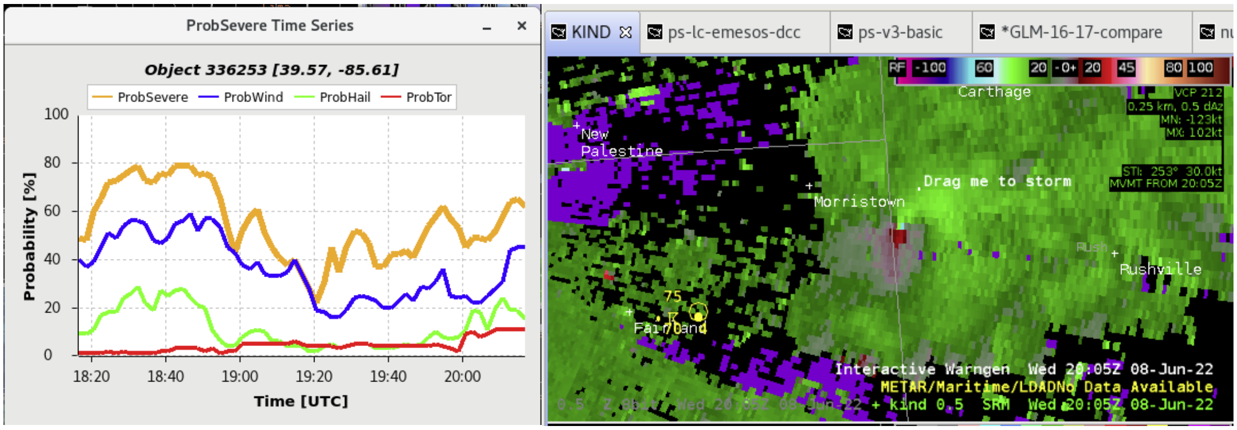

Figure 32: (Left) ProbSevere v3 time series output for a storm object in Indiana on 8 June 2022. (Right) Radar radial velocity data from the KIND WSR-88D, which coincides with the selected storm object.

'ProbSevere seemed to join objects at times leading to rather large objects. It would be nice to keep these separate. The timeseries was excellent and really enjoyed using that for trends.'
Forecaster – End of Week Survey

'In multiple panels, specifically using at least 4, the sample readout would be challenging to view at times and you had to scroll down to view the entire readout. Not sure the workaround from this, but found it difficult to navigate when in a warning situation.'
Forecaster – End of Week Survey

Recommendations for Operational Implementation

Based upon the evaluation of the ProbSevere v3 model in the 2022 HWT Satellite Proving Ground, the following items have been recommended:

- **It is highly recommended that the ProbSevere v3 model be integrated into NWS operations and replace ProbSevere v2.** This is based upon improved model calibration to increase forecaster confidence in warning decisions, and the new time series tool to quickly identify trends in thunderstorm severity potential.
- **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.** Training methods may include workshops, future testbed and proving ground efforts, and case studies highlighting how to apply these data in the warning decision process.
- **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.**

4. Summary and Conclusions

The GOES-R and JPSS Satellite Proving Ground conducted three weeks of remote satellite product evaluations during the 2022 Spring Experiment in a virtual format of the Hazardous Weather Testbed. Nineteen NWS forecasters evaluated six GOES-R and JPSS products and interacted with multiple algorithm developers and subject matter experts during the experiment. Quantitative feedback was collected primarily through surveys administered at the end of each day and week. Qualitative feedback came from daily discussions with forecasters and their blog posts. Expanding upon a recommendation from the 2021 experiment, mock decision support service events were created each day to coincide with NWS warning operations. New to the experiment, forecasters submitted information regarding their warning decisions and decision support communications through a form. Additionally, forecasters could create graphics using the demonstrated products to communicate hazards to the public or DSS partners. Participants received training materials prior to the testbed for each product through a combination of user guides, PowerPoint presentations, and online learning modules. Products were also summarized at the beginning of each week by their developers, which included product applications, limitations, locations in AWIPS, and recommended display practices.

During the end of week surveys forecasters were presented with questions regarding their experience in the virtual testbed, providing feedback that may be helpful for future experiments. First, participants were asked how they would describe the testbed to someone in their local office. Nearly all forecasters found the experience enjoyable and provided them with new product knowledge, hazardous weather scenarios, and forecasting techniques they could apply in their local offices. Several participants commented in the surveys, along with during the end of week discussions, that participating virtually in the experiment did at times cause fatigue; however, most stated that the organization of virtual elements in the experiment was effective for evaluation, collaboration, and discussion. A few forecasters also stated that they would not be able to participate in an in-person experiment due to personal commitments that may not allow them to travel away from their home areas for a full week. Product developers also supported these statements, often suggesting that traveling to Norman, Oklahoma for multiple weeks of evaluations would be cost intensive. With this said however, nearly all forecasters and developers agreed there would still be a considerable benefit to holding at least a portion of the experiment in-person. Comments from forecasters summarizing their week in the testbed are provided below.

> ‘Participating in my very first testbed as a General Forecaster was very informative and engaging. Acquiring a breadth of experiential knowledge and integrated decision support services in these testbed activities was indeed a great asset to my professional career growth. I hope to utilize the practical applications someday when performing warning operations.’
> *Forecaster – End of Week Survey*

> ‘This is a unique experience not only to get hands-on experience with future, potentially operational products, including having easy access to product experts in the (virtual) "room" to allow for stimulating conversation and a direct channel to provide feedback, but also to gain perspective on how experienced forecasters around the country handle forecasting problems both similar and different from our own. The mix of both

operational accuracy with the live data AWIPS environment and yet the flexibility to take a pause and dig into data or discussion is a really great chance to hone skills and reflect on your forecasting philosophies, all in one package.'
Forecaster – End of Week Survey

'New satellite products were made available to us, both on various websites and through the cloud AWIPS, that we could test out during warning operations. People involved in the research and development of these were there with us to observe how we used them in our work, answer questions, and share opinions. Though it was virtual, I didn't feel that I was missing out on too much, as we got a lot of use out of video calls and online chatrooms.'
Forecaster – End of Week Survey

Additional comments from forecasters and developers featured suggestions for improving the organization of future experiments, and feedback on changes from the 2021 HWT Satellite Proving Ground Experiment. This year, forecasters provided decision support throughout the experiment for 21 simulated events during operations. Increased emphasis on DSS events were driven by evolving forecaster needs within the NWS and preliminary testing of two DSS events in the 2021 experiment. Forecasters and developers found the added DSS events were a new challenge and provided new avenues for evaluating the demonstrated products for hazardous weather. Another change, driven by feedback in previous testbeds, was pre-built display procedures in AWIPS-II. These were widely embraced by the participants throughout each week. Forecasters expressed that the procedures provided them with an opportunity to quickly get started on their tasks in operations and leverage best display practices for each product.

Recommendations for Future GOES-R/JPSS HWT Satellite Proving Ground Experiments

Based upon the 2022 GOES-R/JPSS HWT Satellite Proving Ground, the following recommendations for future testbeds are included below:

- **It is strongly recommended that future HWT Satellite Proving Ground Experiments consider in-person and virtual setups for participants, developers, and visiting scientists.** This may maximize the resources available for developers and testbed organizers and create a more diverse pool of forecasters. Consideration should be given to the feasibility of virtual, in-person, and hybrid methods for the experiment.
- **It is strongly recommended that simulated DSS events continue as a portion of the testbed.** The applications shown by forecasters and subsequent feedback increased the scope and impact of the product beyond warning operations.

Acknowledgements

Multiple NWS meteorologists participated in this experiment and provided detailed feedback that went into the recommendations in this report. It would not be possible to complete research-to-operations without their willingness to work with experimental data, ask questions, and participate in crucial discussions with product developers. The developer teams themselves have continued to leverage the HWT Satellite Proving Ground as a collaborative environment to test emerging technologies and transfer impactful knowledge to NWS operations. Johnathan Madden and Justin

Monroe (OU CIWRO) provided exceptional technical support during all phases of the experiment, which allowed forecasters to effectively use the cloud-based AWIPS-II instances for the evaluation. The following people participated as product focal points, subject matter experts, and observers for the products demonstrated in the experiment. The Satellite Proving Ground at the HWT thanks them for their contributions.

GLM

Kristin Calhoun	NOAA – National Severe Storms Laboratory
Kevin Thiel	University of Oklahoma – CIWRO

NUCAPS

Rebekah Esmaili	Science and Technology Corporation
Emily Brendt	NASA – Marshall Space Flight Center
Jack Dostalek	Colorado State University – CIRA
Paul Casey	Science and Technology Corporation
Brian Kahn	NASA – Jet Propulsion Laboratory
Peter Kalmus	NASA – Jet Propulsion Laboratory
Nadia Smith	Science and Technology Corporation
Bill Sjoberg	NOAA – National Environmental Satellite, Data, and Information Service
Kristopher White	NWS – Huntsville, Alabama WFO

Optical Flow Winds

Bob Rabin	NOAA – National Severe Storms Laboratory

PHSnMWnABI

Scott Lindstrom	University of Wisconsin – CIMSS
William Smith	University of Wisconsin – CIMSS
Qi Zhang	Hampton University

ProbSevere (v3 and LightningCast)

John Cintineo	University of Wisconsin – CIMSS
Justin Sieglaff	University of Wisconsin – CIMSS
Lee Cronce	NOAA – National Environmental Satellite, Data, and Information Service

Observers and Subject Matter Experts

Dan Bikos	Colorado State University – CIRA
Dan Lindsey	NOAA – National Environmental Satellite, Data, and Information Service
Connor Nelson	NWS – Operations Proving Ground
Joseph Patton	University of Maryland – CISESS
Jordan Rabinowitz	NWS – Eastern Region Headquarters
Alexa Ross	University of Wisconsin – CIMSS
Michael Stock	NOAA – National Severe Storms Laboratory
Jorel Torres	Colorado State University – CIRA
Kyle Ziolkowski	Environment Canada – Storm Prediction Centre

5. References

Barnet, C. D., and Coauthors, 2021: NUCAPS Algorithm Theoretical Basis Document.

Berndt, E., and Coauthors, 2020: Gridded Satellite Sounding Retrievals in Operational Weather Forecasting: Product Description and Emerging Applications. *Remote Sensing*, **12**, https://doi.org/10.3390/rs12203311.

Blumberg, W. G., K. T. Halbert, T. A. Supinie, P. T. Marsh, R. L. Thompson, and J. A. Hart, 2017: SHARPpy: An Open-Source Sounding Analysis Toolkit for the Atmospheric Sciences. *Bulletin of the American Meteorological Society*, **98**, 1625–1636, https://doi.org/10.1175/BAMS-D-15-00309.1.

Brox, T., A. Bruhn, N. Papenberg, and J. Weickert, 2004: High Accuracy Optical Flow Estimation Based on a Theory for Warping. *Computer Vision - ECCV 2004*, T. Pajdla and J. Matas, Eds., Berlin, Heidelberg, Springer Berlin Heidelberg, 25–36.

Bruning, E. C., and Coauthors, 2019: Meteorological Imagery for the Geostationary Lightning Mapper. *J. Geophys. Res. Atmos.*, **124**, 14285–14309, https://doi.org/10.1029/2019JD030874.

Calhoun, K. M., K. L. Berry, D. M. Kingfield, T. Meyer, M. J. Krocak, T. M. Smith, G. Stumpf, and A. Gerard, 2021: The Experimental Warning Program of NOAA's Hazardous Weather Testbed. *Bulletin of the American Meteorological Society*, 1–51, https://doi.org/10.1175/BAMS-D-21-0017.1.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and D. T. Lindsey, 2014: An Empirical Model for Assessing the Severe Weather Potential of Developing Convection. *Weather and Forecasting*, **29**, 639–653, https://doi.org/10.1175/WAF-D-13-00113.1.

——, ——, ——, A. Wimmers, J. Brunner, and W. Bellon, 2020: A Deep-Learning Model for Automated Detection of Intense Midlatitude Convection Using Geostationary Satellite Images. *Weather and Forecasting*, **35**, 2567–2588, https://doi.org/10.1175/WAF-D-20-0028.1.

——, ——, and ——, 2022: ProbSevere LightningCast: A Deep-Learning Model for Satellite-Based Lightning Nowcasting. *Weather and Forecasting*, **37**, 1239–1257, https://doi.org/10.1175/WAF-D-22-0019.1.

Esmaili, R. B., and Coauthors, 2020: Adapting Satellite Soundings for Operational Forecasting within the Hazardous Weather Testbed. *Remote Sensing*, **12**, https://doi.org/10.3390/rs12050886.

Goodman, S. J., and Coauthors, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmospheric Research*, **125–126**, 34–49, https://doi.org/10.1016/j.atmosres.2013.01.006.

Rabin, B., 2021: Wind Estimates in Storms from an Optical Flow Technique Using GOES Imagery.

Smith, W. L., Q. Zhang, M. Shao, and E. Weisz, 2020: Improved Severe Weather Forecasts Using LEO and GEO Satellite Soundings. *Journal of Atmospheric and Oceanic Technology*, **37**, 1203–1218, https://doi.org/10.1175/JTECH-D-19-0158.1.

Thiel, K. C., K. M. Calhoun, A. E. Reinhart, and D. R. MacGorman, 2020: GLM and ABI Characteristics of Severe and Convective Storms. *J. Geophys. Res. Atmos.*, **125**, https://doi.org/10.1029/2020JD032858.