GOES-R and JPSS Proving Ground Demonstration at the Hazardous Weather Testbed 2017 Summer Experiment Final Evaluation

Project Title: GOES-R and JPSS Proving Ground Demonstration at the 2017 Summer Experiment - Experimental Warning Program (EWP)

Organization: NOAA Hazardous Weather Testbed (HWT)

Evaluator(s): National Weather Service (NWS) Forecasters, Broadcast Meteorologists, Storm Prediction Center (SPC), National Severe Storms Laboratory (NSSL)

Duration of Evaluation: 19 June 2017 – 21 July 2017

Prepared By: Michael Bowlan (OU/CIMMS and NOAA/SPC) and Kristin Calhoun (OU/CIMMS and NSSL)

Submitted Date:

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 2017 Summer Experiment, which took place at the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) in Norman, OK from 19 June to 21 July 2017. The Satellite Proving Ground activities were focused in the Experimental Warning Program (EWP). A total of 12 National Weather Service (NWS) forecasters representing four NWS regions and an additional four broadcast meteorologists participated in the EWP experiment. They evaluated four major (Table 1) baseline, future capability, and experimental GOES-R and JPSS products in the real-time simulated short-term forecast and warning environment of the EWP using the second generation Advanced Weather Interactive Processing System (AWIPS-II).

Some of the products demonstrated in 2017 were involved in previous HWT experiments and have received updates based on participant feedback from the HWT and other demonstrations. GOES-R products demonstrated in the 2017 EWP Summer Experiment included: GOES-16 Advanced Baseline Imager (ABI) Cloud and Moisture Imagery, baseline derived products and numerous multispectral Red Green Blue (RGB) products, the Geostationary Lightning Mapper (GLM) Lightning Detection, and the Probability of Severe statistical model (ProbSevere). Additionally, GOES-16 provides 1-minute imagery via two 1000-km x 1000-km mesoscale sectors, and its value was also assessed in monitoring convective storm life cycles. As a JPSS Proving Ground activity, the NOAA Unique Combined Atmospheric Processing System (NUCAPS) temperature and moisture profiles were displayed using the AWIPS-II sounding analysis program. These soundings were created using data from three different polar orbiting satellites: the Suomi National Polar-orbiting Partnership (Suomi-NPP) and Europe's MetOp-A and MetOp-B. Additionally, a modified version of NUCAPS was also examined in which an automated correction incorporating surface observations was applied to the boundary layer to improve the accuracy of the sounding. Also, participants were able to view the NUCAPS derived parameters in a plan or cross-section view. Several visiting scientists attended the EWP over the four weeks to provide additional product expertise and interact directly with operational forecasters. Organizations represented by those individuals included: The University of Wisconsin Cooperative Institute for Meteorological Satellite Studies (UW/CIMSS), The University of Oklahoma Cooperative Institute for Mesoscale Meteorological Studies (OU/CIMMS), the National Severe Storms Laboratory (NSSL), the NASA Short-term Prediction Research and Transition Center (SPoRT), Science and Technology Corporation (STC) and NWS. The Storm Prediction Center (SPC) and HWT Satellite Liaison, Michael Bowlan (OU/CIMMS and NOAA/SPC), provided overall project management and subject matter expertise for the GOES-R Proving Ground efforts in the HWT with support from William Line (NWS) and Kristin Calhoun (OU/CIMMS and NOAA/NSSL).

Forecaster feedback during the evaluation was collected using several different methods, including daily surveys, weekly surveys, daily debriefs, weekly debriefs, blog posts, informal conversations in the HWT and a weekly "Tales from the Testbed" webinar. Typical feedback included: suggestions for improving the algorithms, ideas for making the displays more effective for information transfer to forecasters, best practices for product use, suggestions for training, and situations in which the tools worked well and not so well. Forecasters' favorite aspect of this year's experiment was being able to evaluate real time GOES-16 imagery and products, and seeing the benefits of certain products or RGBs to their operations. The ProbSevere model continues to provide useful guidance, especially when applied to discrete storms, though improvements to performance are needed for multicellular/linear convective modes, and particularly when wind and tornadoes are the main hazard. Forecasters were also especially interested in evaluating GLM data for the first time and getting an early view of the data before others in operational offices. However, they indicated many improvements are needed before it is ready for efficient use in operations. Finally, participants found the NUCAPS information to be helpful in filling spatial and temporal gaps that exist in atmospheric sounding information, and liked that the plan view displays provided a quick look at certain parameters and levels in a NUCAPS swath.

2. Introduction

GOES-R Proving Ground (Goodman et al. 2012) demonstrations in the HWT provide users with a glimpse into the capabilities, products and algorithms that will be available with the future GOES-R geostationary 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 an operational environment. Feedback received from participants in the HWT has proven invaluable to the continued development and refinement of GOES-R and JPSS algorithms. Furthermore, the EWP facilitates the testing of satellite-based products in the AWIPS-II data processing and visualization system currently used at NWS Weather Forecast Offices (WFOs).

In 2017, the GOES-R/JPSS Proving Ground activities were conducted during the weeks of June 19, June 26, July 10, and July 17 with three NWS forecasters and one broadcast meteorologist participating each week. In an effort to extend the satellite knowledge and participation to the broader meteorological community, and to recognize the critical role played by the private sector in communicating warnings to the public, broadcast meteorologists sponsored by the GOES-R Program participated in the Summer Experiment for the fourth year in a row, working alongside NWS forecasters. Training modules in the form of an Articulate Power Point presentation for each product demonstrated were sent to and completed by participants prior to their arrival in Norman. Each week, participants arrived in Norman on Sunday, worked eight hour experimental warning shifts Monday-Thursday and a half-day on Friday before traveling home Friday afternoon.

Much of Monday was a spin-up day that included a one hour orientation, familiarization with the AWIPS-II system, and one-on-one hands-on training between participants, product developers, and the Satellite Liaison. The shifts on Tuesday, Wednesday and Thursday were "flex shifts", meaning the start time was anywhere between 9 am and 3 pm, depending on when the most active convective weather across the Contiguous United States (CONUS) was expected to occur. The next day start time was determined the previous evening by the Weekly Coordinator. The Friday half-day involved a weekly debrief and preparation and delivery of the "Tales from the Testbed" webinar.

Shifts typically began a couple of hours before convective initiation was expected to occur as many of the products demonstrated this year have their greatest utility in the pre-convective environment. At the start of each Mon-Thurs experimental warning shift, the Satellite Liaison and forecasters interrogated the large scale weather pattern across the CONUS and determined where to operate for the day. Forecasters, working in pairs, provided experimental short-term forecasts for their assigned County Warning Area (CWA) via the Hazardous Weather Testbed blog. Early in the shift, these were primarily mesoscale forecasts discussing the environment, where convection was expected to occur, and what the applicable demonstration products were showing. Once convection began to grow upscale, one forecaster in the pair would switch to issuing experimental warnings for their CWA while the other forecaster would continue to monitor the mesoscale environment and compose blog posts. Blog posts regarding the use of demonstration products in the warning decision-making process were written during this period along with continued updates on the mesoscale environment. If severe convective activity in a CWA ceased or was no longer expected to occur, the Satellite Liaison would transition the pair of forecasters to focus on a more convectively active CWA.

At the end of each week, the three NWS forecasters and one broadcast meteorologist participated in the "Tales from the Testbed" webinar, prepared by the Satellite Liaison, via GoToMeeting. These 22 minute presentations gave participants an opportunity to share their experience in the HWT with an average of greater than 30 remote locations each week, including NWS Headquarters, NWS WFOs and research scientists at satellite cooperative institutes nationwide, providing widespread exposure for the GOES-R and JPSS Proving Ground products. Topics for each of the four webinars were chosen based on the particular week's weather. Sixteen minutes were allowed afterward for questions and comments from viewers on the webinar.

Feedback from participants came in several forms. During the short-term experimental forecast and warning shifts, participants were encouraged to blog their decisions along with any thoughts and feedback they had regarding the products under evaluation. Over 400 GOES-R and JPSS related blog posts were written during the four weeks of the Summer Experiment by forecasters, product developers, and the Satellite Liaison. At the end of each shift (Monday-Thursday), participants filled out a survey of questions for each product under evaluation. The Tuesday-Thursday shifts began with a "daily debrief" during which participants discussed their use of the demonstration products during the previous day's activities. Friday morning, a "weekly debrief" allowed product developers an opportunity to ask the participants any final questions, and for the participants to share their final thoughts and suggestions for product improvement. Additionally on Friday morning, participants completed one last "end of the week" survey of questions. Feedback from the GOES-R and JPSS demonstrations during the 2017 Summer Experiment is summarized in this report.

3. Products Evaluated

Table 1. List of GOES-R and JPSS products demonstrated within the HWT/EWP 2017 Summer Experiment

Demonstrated Product	Category
Advanced Baseline Imager (ABI) imagery, baseline derived products	GOES-R Baseline
RGB Composites and Channel Differences	National Weather Service
ProbSevere Model	GOES-R Risk Reduction
GLM Lightning Detection	GOES-R Baseline
NUCAPS Temperature and Moisture Profiles	JPSS
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** – Products funded through the JPSS program	

3.1 Advanced Baseline Imager (ABI) Imagery, Baseline Derived Products

National Oceanic and Atmospheric Administration (NOAA), National Weather Service (NWS), National Environmental Satellite, Data, and Information Service (NESDIS), and GOES-R Program

For the first time, real time ABI data were able to be evaluated in the HWT. All available imagery and baseline derived products were up for evaluation in this year's experiment. The imagery was heavily utilized throughout the experiment with a focus on using some of the new channels and how they can apply to convective situations. Many convectively applicable baseline products, such as Total Precipitable Water (TPW), derived stability indices, and derived motion winds, were also heavily evaluated in this real time experimental warning environment. In addition, many forecasters also evaluated the RGB composites and channel differences. Different temporal resolution imagery from 1 minute to 5 minutes to 15 minutes was evaluated and compared to see how the advancements from GOES-16 can improve forecaster decision-making in warning situations from using legacy GOES-13/15 data. Feedback from this experiment, including product usefulness and display, is presented in this report.

Use of ABI imagery in the HWT

This was the first time that forecasters used actual (not proxy) in-orbit ABI data in the testbed, and the first time for many to see and use the derived products and RGB composites in a simulated operational environment. The ABI was one of the most anticipated capabilities to be evaluated in the testbed by the participants. At the end of each day, forecasters were asked a number of questions regarding the ABI data. Perhaps the most beneficial of all the ABI data was having the very high temporal resolution imagery coupled with the increased spatial resolution, especially for assessing convective initiation, tracking boundaries and storm/boundary interactions, and storm trend monitoring. Forecasters were asked if the 1-minute imagery provided significant information not captured in the routine 5-minute imagery, and 65 % of forecasters responded that the 1-minute imagery did provide value over the 5-minute data. At the end of each week, they were asked to rate the overall impact of the 1- minute imagery and 18 of the 19 responses indicated it had a very positive to extremely positive impact, with the lone other response rating it as a moderately positive impact. Overall the 1-minute imagery was found to be extremely valuable to experimental convective warning operations.

> "The 5 minute Vis sat and IR didn't detect the small perturbations in the convective cloud tops compared to the 1 minute imagery. I could really get a sense of the explosive updraft development using the 1 min imagery."
> *Forecaster, End-of-Day Survey*

> "Early on in the day, you could see towers building and dissipating in the 1-minute data that could not be seen in the 5-minute data."
> *Forecaster, End-of-Day Survey*

> "I was able to see changes in cloud top temperature and updraft trends quicker than with the 5- minute data which helped in warning decision."
> *Forecaster, End-of-Day Survey*

> "…during marginal less obvious days I think having the 1-minute data can help focus in on potential changes in convective strength more quickly than you'd otherwise be able to do."
> *Forecaster, End-of-Week Survey*

> "Monitoring convective updrafts and tracking outflow boundaries/interactions. The 1 minute data made detecting these features much easier. As far as a nowcasting environment, you can really focus your thunderstorm chances along and ahead of boundaries, helping to improve the forecast, and eventually help aid in warning operations as we transition into them."
> *Forecaster, End-of-Week Survey*

Not only did the 1-minute imagery aid in the initiation and updraft monitoring phase of convection, but it also aided forecasters in null cases and in identifying areas that were more stable and might not result in storm development. The stable character of a cumulus cloud field was telling to a forecaster that storms were not imminent in that area and allowed the forecaster to focus on other parts of the forecast area. For a developed storm, the lack of overshooting tops or texture at the storm top indicated that updrafts were not particularly robust. Downward trends in overshooting top abundance and strength or warming of the cloud tops were indicators that storm updraft intensity and/or coverage may be waning. Similarly, the slowing spread and decreasing sharpness of storm anvil edge were also indicators of convective decay that were captured nicely in the 1-minute imagery.

> "Being able to tell in nearly real time that disturbed cu field was not growing much in the vertical. Operationally allowed me to determine that the threat for new storms was low. Not sure I would have seen that in 5 minute data."
> *Forecaster, End-of-Day Survey*

> "I was able to see how quickly the cumulus clouds were becoming shredded or dying off since there was an abundance of dry air in the middle levels capping off vertical growth."
> *Forecaster, End-of-Day Survey*

> "Served best in targeting storms of interest. Cold IR tops and cumulus texture were key due to the high temporal and spatial resolution."
> *Forecaster, End-of-Day Survey*

> "It will allow you to see the subtle changes in the storms or intensifying updrafts. When we were watching storms in SD with the 1-min data, you could see towers developing and dissipating that weren't captured by the 5-min data."
> *Forecaster, End-of-Week Survey*

Furthermore, the forecasters were asked about what ABI imagery channels they used and which were found to be most useful. A majority of forecasters found the "red" visible and "clean" Infrared (IR) channels to be most useful in convective operations. Of the newer channels available from GOES-16, the three water vapor channels were used extensively and found to be moderately to very useful by a large majority of the forecasters. Forecasters found the lower level water vapor channel useful for tracking some more subtle features otherwise missed by current water vapor imagery. These channels were used to identify short wave troughs, jets, and moisture streams used as a forecast and Nowcast tool for storm initiation and development. Forecasters also made use of combining the visible channel with the clean IR semi-transparent overlay, dubbed the "sandwich" product by the Europeans (Setvák et al. 2012) (Fig. 1), to monitor cloud top temperature information of the IR imagery while maintaining the texture and resolution of the visible imagery within the updraft/overshooting top. Additionally, many found that the increased resolution of the IR imagery made it much more useful at night compared to current IR imagery, with some even calling it as good as having visible imagery at night. Many of the other channels were used more sparingly throughout the experiment.

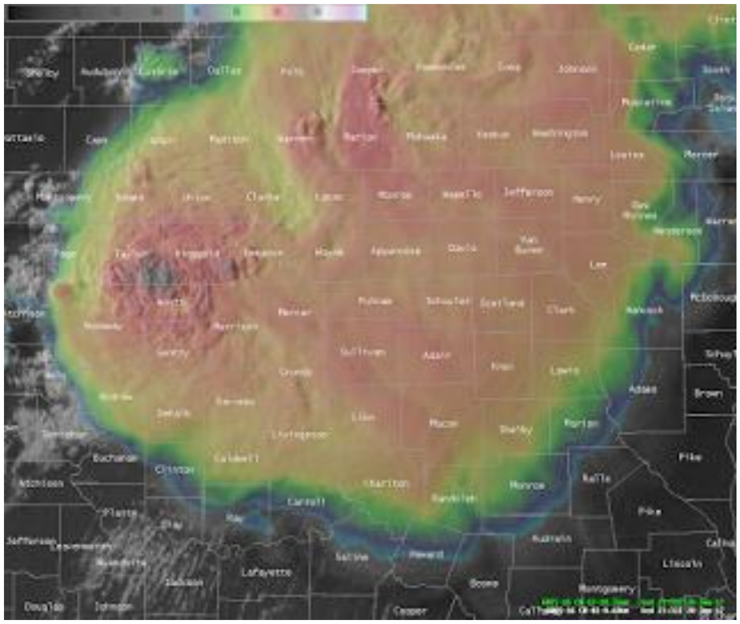

Figure 1: 2332 UTC 28 June 2017 GOES-16 VIS (underlay) and IR (transparent overlay) imagery for a storm on the Missouri/Iowa border.

“During the experiment, I became I big fan of the sandwiched Red Visible and Clean IR windows. This allowed for easy identification of stronger updraft/vertical development within environments already impacted by convective debris, such as anvil cirrus or events developing within a region of stratus.”
Forecaster, End-of-Week Survey

“I certainly think the red visible and clean IR provide a lot of useful information regarding convective trends and seem like the most important. The use of the three water vapor channels also provides a lot of useful information in terms of identifying subtle features which were possibly not evident in previous GOES imagery.”
Forecaster, End-of-Week Survey

“All of the water vapor channels were a huge help in identifying sort waves, jet streams and moisture. The visible channels by day were obviously a huge asset but the IR higher resolution and temporal resolution make night-time use as good as using daytime visible.”
Forecaster, End-of-Week Survey

“The "Sandwich" combo is very useful for situational awareness and developing convection. When the first developing cell intensified, colder cloud tops were clearly shown in the IR, with the cu field with less vertical depth clearly identifiable with weaker cells.”
*Forecaster, “*VIS/IR Sandwich Procedure (20 June 2017)”, *GOES-R HWT Blog*

“I used the Water Vapor channels to find short wave features, knowing that a mid-level (500hPa) short wave was lifting into western Montana.”
Forecaster, End-of-Day Survey

In one case of land falling Tropical Storm Cindy on 22 June 2017, the forecaster wanted to investigate why storms were not intensifying in eastern Mississippi as opposed to farther west. There was a tornado watch in effect throughout the area. The forecaster looked at a number of products including the three water vapor channels. He noticed that while there was ample amounts of moisture on the mid- and upper-level water vapor channels, the low-level water vapor showed a large dry slot extending into eastern Mississippi (Fig. 2). This made the forecaster confident that storms were not likely to intensify in the dry slot and he should focus his attention on the storms to the west within the focus area.

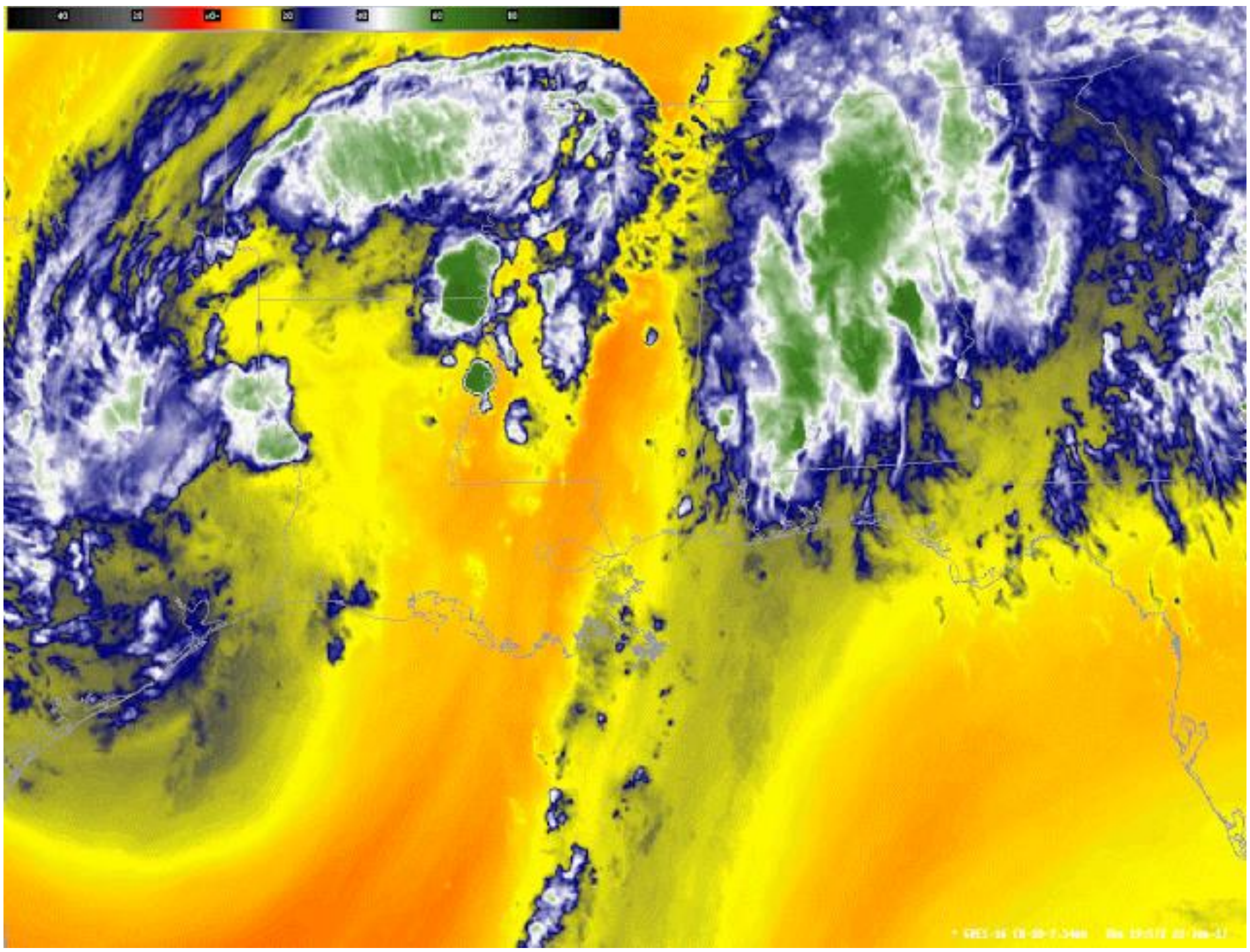

Figure 2: 2057 UTC 22 June 2017 GOES-16 7.3um “Low-Level” Water Vapor imagery over the central Gulf Coast.

Use of ABI Baseline Derived Products in the HWT

Forecasters were also able to evaluate the baseline derived products within the HWT experiment this year. The forecasters focused mainly on products of relative importance to convective initiation and warning operations, so some products were not evaluated for this experiment. They were asked to use these products each day to make mesoscale forecasts and updates during their experimental warning operations. Each day forecasters were asked which products they used and what they were used for, and at the end of the week, they were asked to explain which products would have the most impact on improving warning operations in the field. Total Precipitable Water was overwhelmingly seen as the product that provided the most impact and had the most consistent use in the testbed. The derived stability indices were also examined carefully, especially the Convective Available Potential Energy (CAPE) and to some extent the Lifted Index (LI) and K-index products. These products had issues at times quantitatively identifying the correct values when compared with SPC Mesoanalysis, but in a qualitative sense were seen as helpful for forecasting convective development and in some cases decay by identifying regions of greater relative instability and air mass gradients. Another product used were the derived motion winds. These provided value at times, but had a number of issues including data display to accuracy which will be discussed in more detail in a later section. Some other products such as fire "hot spot" identification, aerosol detection, and some cloud top products were looked at by a few forecasters, but no formal evaluation was conducted for these products. Forecaster feedback on these derived products are presented in this section.

> "I used the Total Precipitable Water Product which was valuable in discerning where the deeper moisture was located. The CAPE product underestimated the amount of instability. The GOES derived winds helped to verify the placement of the upper jet and a mid-level speed max."
> *Forecaster, End-of-Day Survey*

> "The CAPE/LI products were really helpful in highlighting the areas of focus on a Situational Awareness basis. While the values for CAPE were generally lower than expected, the areal extent and gradients were generally in line with what I would have expected."
> *Forecaster, End-of-Week Survey*

> "For derived, I found the total PW to be very helpful in representing the environmental conditions, which did have an impact on how convection would evolve. LI was not bad either. The winds were really useful, mainly for the same reason as the PW in analyzing the environment.
> *Forecaster, End-of-Week Survey*

> "Yes, CAPE, LI, and PW. Actually all were useful in showing a gradient in the same location. More specifically, higher CAPE, LI, and PW were see across eastern MT and western ND, with then a transition to lower values towards south central SD. This ended up being where convection began to weaken."
> *Forecaster, End-of-Day Survey*

> "I looked at the stability products as well as the cloud phase product. There was a lot of missing data in the stability products, but the cloud phase was useful in helping identify glaciation."
> *Forecaster, End-of-Day Survey*

The derived CAPE and TPW played an important role in decision making for an event on 21 June 2017 in Eastern North Dakota and Western Minnesota. A weak front or wind shift boundary was moving through the area and the forecaster in this case noticed that the CAPE and TPW products highlighted this boundary very well, with more unstable CAPE values out ahead of the boundary and a more stable air mass behind the boundary. The forecaster also noticed a tongue of higher CAPE moving north into the forecast area as the day was progressing. He predicted storms to develop quickly along gradient of higher CAPE values along the boundary and that they should intensify as they move southeast into the zone of higher CAPE. As the scenario progressed, his forecast verified well as storms developed all along the CAPE gradient, then intensified and became severe as they moved southeast into the more unstable air mass noted in the GOES stability indices (Fig. 3).

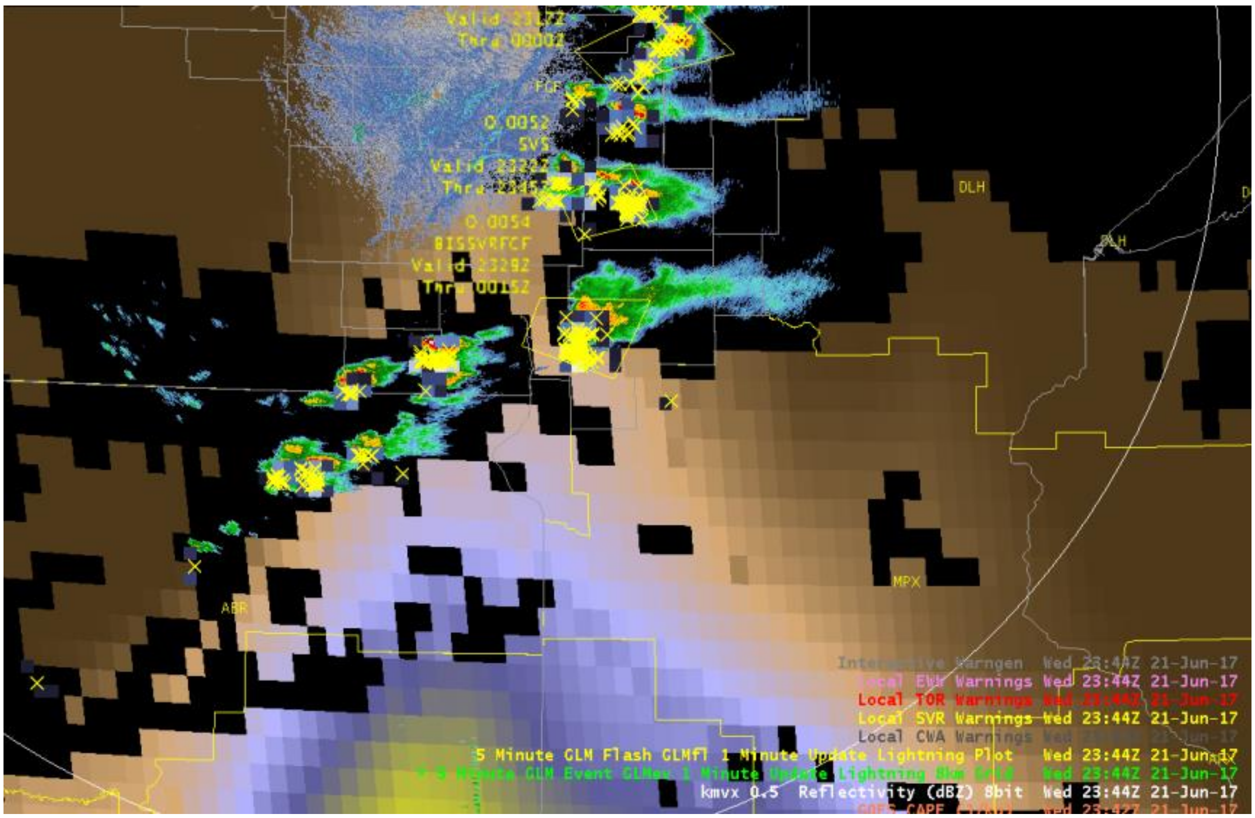

Figure 3: 2344 UTC 21 June 2017 GOES-16 derived CAPE (color fill in J/kg) with radar reflectivity, GLM lightning (denoted by yellow x), and experimental severe thunderstorm warnings (yellow polygons)

Suggestions for Improvement of Derived Products

There were a number of suggestions that forecasters provided to improve some of the derived products for routine use in operations. The derived CAPE values were consistently 50% or more lower when compared to other data such as SPC Mesoanalysis and the Rapid Refresh (RAP) model CAPE. Generally the GOES-16 derived CAPE values were not very indicative of the magnitude of CAPE present, although the qualitative regions of larger/smaller CAPE and gradients often lined up very well with the gradients seen in the SPC Mesoanalysis. There were also numerous issues that the forecasters identified with the derived winds. One idea for improvement was to have all levels time matched to each other. When all of the levels were loaded into AWIPS-II, different levels show up at different times as the images moved forward in time causing the wind barbs to appear "jumpy" and sporadic in nature and difficult to assess any type of trends. Many would like to see the wind barbs stay on the screen until the next update at that particular level and to be time matched to update at the same times at all levels. Another idea for the displaying the wind barbs is to color code by height instead of color coding every height based on speed. Since the wind barb already has the speed on the barb, color coding by height so that the different levels are easily discernable when multiple levels are loaded was considered ideal by some forecasters. In addition, the accuracy of the derived winds seemed to be lacking somewhat. Forecasters noted that there very few winds below about 500 hPa in most cases, even in cases where there was ample cloud cover below the 500 hPa level. There also seemed to be many spurious wind barbs, usually around cumulus fields, that showed speeds and direction that did not appear to be physically realistic and did not correspond well with the imagery or other observations such as soundings and radar Velocity Azimuth Display (VAD) profiles from the area. Some quotes and ideas from forecasters are posted below.

> "I looked at baseline derived winds. Some levels (pressure) seem to have a lot of plots, others I didn't really get anything in the area of interest. Seemed real hit and miss. I expect more consistency in the plot. I did have one level at 300 hPa which had a good plot of winds I could compare to cloud and model trends."
> *Forecaster, End-of-Day Survey*

> "I also viewed the derived motion wind for the tropical storm at the New Orleans CWA. Unfortunately, the performance of the derived motion wind was poor. There were a lot of spurious wind vectors across southern Texas and Mexico, which had values between 50-80 knots. These are likely due to the algorithm tracking small areas of cumulus. In addition, the wind vectors across the southern part of the tropical storm were in the opposite direction. These winds should be from the west instead of the east. The overall poor performance of the derived motion wind gives me zero confidence in operational utility of the data."
> *Forecaster, End-of-Day Survey*

> "I view derived CAPE. Though the values of CAPE appear too low (by about 50% when compared to SPC Mesoanalysis), it is a great situational awareness tool to find the location of the greatest instability."
> *Forecaster, End-of-Day Survey*

> "Total precipitable water was the most useful of the derived products. CAPE was probably the least representative of the derived products. Also used derived winds to look at an MCV. The 200-400hPa winds were probably the most useful. Winds from 500hPa and down were really not detected."
> *Forecaster, End-of-Day Survey*

> "While the values for CAPE were generally lower than expected, the areal extent and gradients were generally in line with what I would have expected."
> *Forecaster, End-of-Week Survey*

3.2 GOES-16 RGB Composites and Channel Differences

National Weather Service (NWS)

The last GOES-16 related products to be evaluated during the experiment were the numerous RGB composites and channel differences. These products are created on the fly within AWIPS-II and combine multiple channels together to highlight certain features related to a forecast problem. A RGB is a way to combine information from multiple channels into a single image to highlight many different features associated with certain phenomena. Numerous convective and wildfire RGBs were examined in the HWT to assist with convective forecasting and development. A couple of channel differences were also evaluated for their use in convective environments. Channel differences are also created within AWIPS-II by subtracting the values of one channel from another to pull out information that can be important and otherwise not noticed when looking at a single channel.

Use of RGB Composites in the HWT

First, there was an overall sentiment that better training was needed for forecasters at WFOs on the RGBs and their recommended use in operations. Many forecasters had to "start from scratch" learning about the RGBs and would like some sort of reference sheet or "quick guide" before being able to use these. However, after forecasters became more familiar with the RGBs and learned what certain ones are useful for including appropriate interpretation of the enhancements, a number of them that were used throughout the remainder of each week during the experiment. A large number of forecasters found that the day cloud phase RGB (Fig. 4) was most useful when monitoring cumulus fields for convective initiation. That RGB allowed forecasters to see when clouds glaciated and matured into deep convective clouds. There was also some utility in the advanced Day Convection RGB, although it is only available in Full Disk mode right now in AWIPS-II and at a degraded resolution of 6 kilometers. At the full resolution of 2 kilometers, forecasters believe that this RGB will also assist in monitoring towering cumulus and mature convection. Many of the other convective RGBs were also examined, but were found to have little to no utility over visible and IR satellite imagery. The simple water vapor and Air mass RGBs were also found to have some utility in monitoring for subtle features like shortwave troughs, jets, and mid-level fronts that were a little less discernable in the single channel water vapor imagery.

> "The day cloud product was by far the most useful for monitoring convective trends. The simple water vapor product did a good job in identifying shortwave features as well, and at times features stood out better in this product than in any of the 3 ABI water vapor channels."
> *Forecaster, End-of-Week Survey*

> "Day Cloud Phase- Was able to monitor storm glaciation giving information on potential storm development. Very useful.
> Simple Water Vapor - Used to assimilate 3 different channels highlighting more subtle features."
> *Forecaster, End-of-Week Survey*

> "I honestly did not find much of an improvement in the RGBs over the baseline channels. I could see some use in the Day Cloud Phase, but many times, the RGBs didn't offer any more information than I already gathered from the baseline channels."
> *Forecaster, End-of-Week Survey*

> "There seemed to be way too many RGB's and it would be nice to have some kind of guidance on what each RGB can be used for. Without having much background going in, it's really hard to evaluate them when you don't know what practical applications each one may be good for."
> *Forecaster, End-of-Week Survey*

> "The day cloud phase provided a lot of valuable info on whether clouds were water based, or once they had glaciated. Using the correct color curve also added a lot of value. The simple water vapor product also helped identify the location of a mid-level speed max, and also helped identify a region of ascent associated with this feature."
> *Forecaster, End-of-Day Survey*

In a case from 20 July 2017 in northwestern North Dakota, shown below, a forecaster was able to use the Day Cloud Phase RGB at the beginning of the forecast shift to monitor for convective initiation across the area. He noticed that there were numerous towering cumulus going up along a boundary ahead of an associated shortwave trough entering the forecast area from the west. In the areas where the cyan-like color changes to a greener and eventually green-yellow color indicate that those towers are beginning to glaciate and become more ice clouds than water clouds. The forecaster took note of this and knew that initiation was occurring which alerted him to anticipate experimental warning operations to begin very soon considering the favorable environment. This RGB was used many times throughout the week to keep forecasters aware of when initiation was beginning and when to prepare for warning operations.

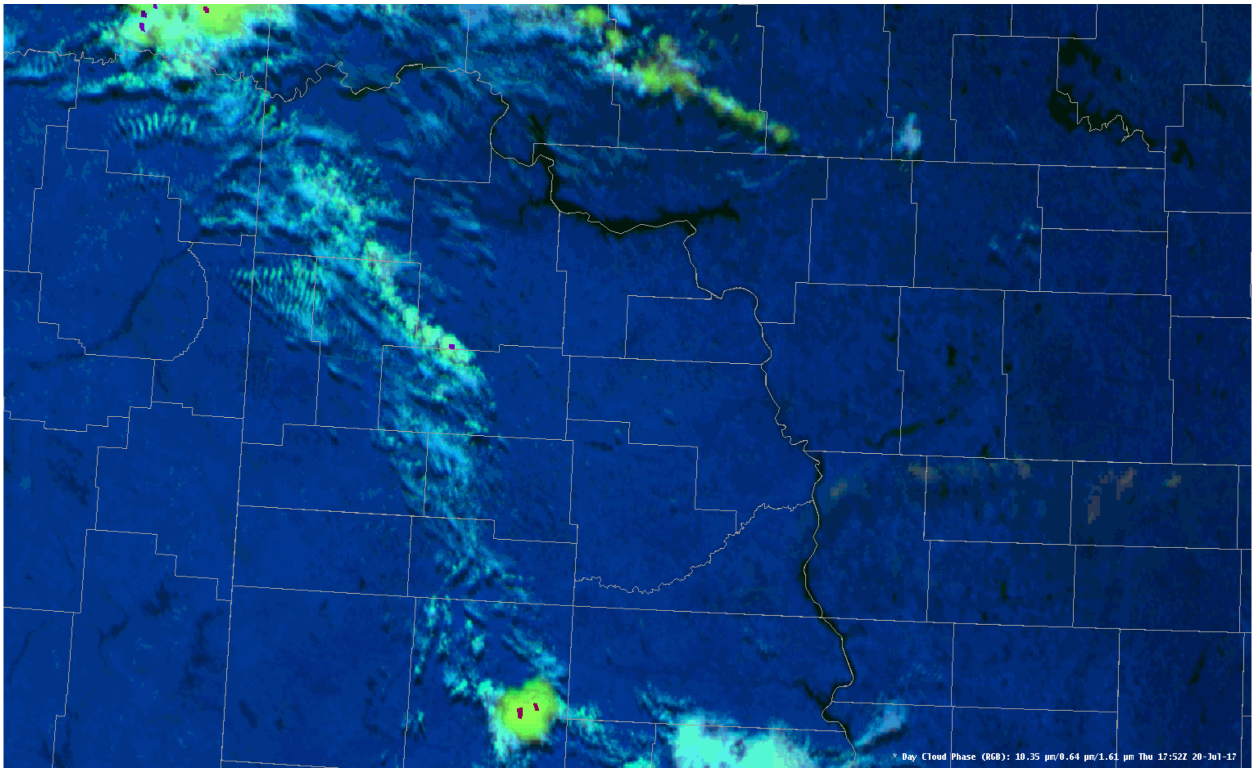

Figure 4: 1752 UTC 20 July 2017 Day Cloud Phase RGB over western North Dakota. Cyan colors denote low-level water clouds while the greener colors denote clouds that have become mixed phase ice particles and water droplets.

In addition to these convective RGBs, forecasters also had opportunities to evaluate a few of the other RGBs that are available within AWIPS-II. Far and away one of the most useful was the Fire Temperature RGB (Fig. 5). Most forecasters thought this RGB added value over the 3.90 micron channel to easily detect wildfires. The fire stands out clearly on this RGB as a bright red pixel or pixels and the contrast of colors helped make the fire easily apparent, even on a zoomed out CONUS view. Many of the other RGBs were briefly looked at as well, but didn’t provide much added information during this experiment.

> “I used RGBs at the end and I was really impressed with the Fire Temperature. It was outside of CWA, but I was amazed how it picked up the hot spots, even with a small fire in the TX Panhandle.”
> *Forecaster, End-of-Day Survey*
>
> “Fire Temperature, it was very useful in finding hot spots across the western United States.”
> *Forecaster, End-of-Day Survey*
>
> “Day Land Convection, Fire Temperature and Air mass were the most use this week. I really liked the fire RGB's a lot they were very good at identifying fire even on wide map views.”
> *Forecaster, End-of-Week Survey*

"The FIRE temperature was very useful in showing the smallest of fires that lasted even less than 30 min. I know that's not convective warning operations, but it helped with fire weather concerns. The cloud top phase was also key in showing rapidly increasing updraft strength."
Forecaster, End-of-Week Survey

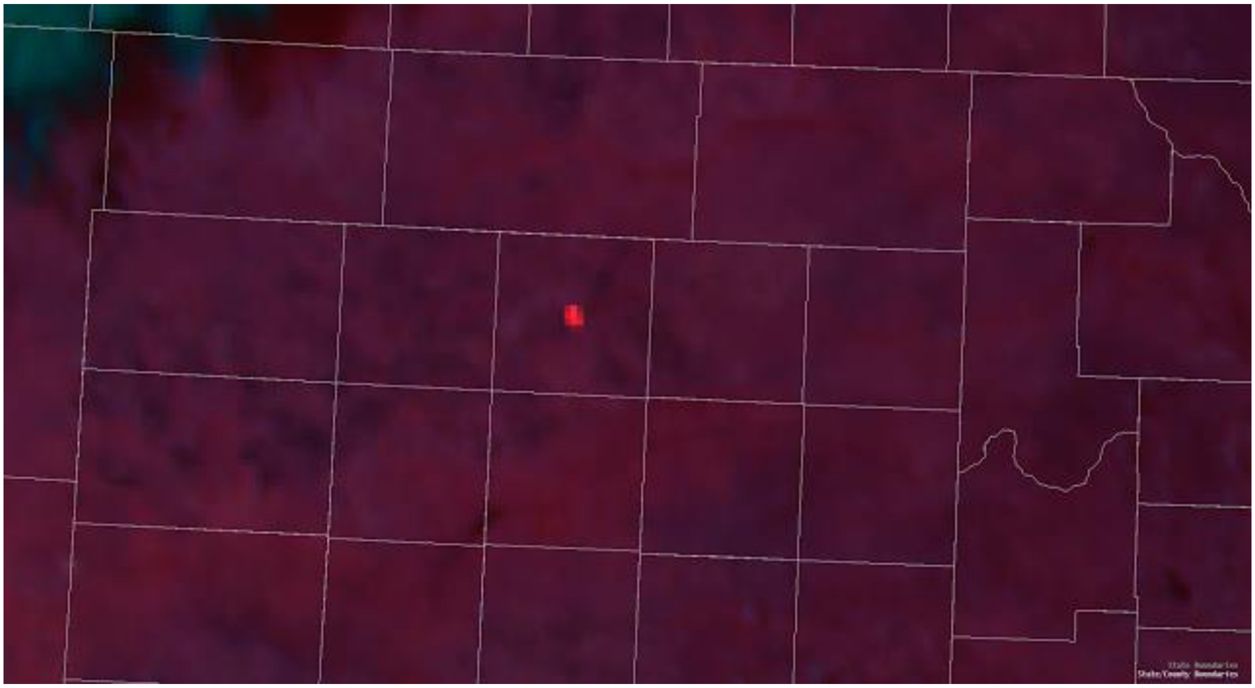

Figure 5: 12 July 2017 Fire Temperature RGB identifies a hot spot associated with a wildfire in the Texas Panhandle

Use of Channel Differences in the HWT

The ABI split window channel difference, also known as the low level moisture channel difference appeared to provide limited utility in certain cases of convective development. This channel difference was able to highlight some areas of low level moisture pooling along boundaries prior to convective initiation and cloud contamination. This product showed the most use in specialized circumstances along dry lines or fronts before there were large scale cumulus fields present. The other channel difference products were found to have little to no use in convective operations, although many forecasters did think that the Fog difference was valuable to have in operations.

"Had some limited use of the Moisture difference product. Was able to highlight moisture pooling along warm front with it."
Forecaster, End-of-Week Survey

"SWD_IR is still one of my favorite band differences, really allows me to see multiple features all on one loop."
Forecaster, End-of-Day Survey

> "Used the moisture difference product. Took a bit to get adjusted to data, but was able to highlight area of moisture pooling along front."
> *Forecaster, End-of-Day Survey*

Suggestions for Improvement of RGBs and Channel Differences

The major downfall for the RGBs and channel differences for the forecasters this year was the insufficient training in how to apply the RGBs within their operational workflow. By the end of their week, many forecasters understood better how to use and apply many of the RGBs, but the HWT results strongly suggest field forecasters need much more training on the subject to be able to expertly apply these complex images in their day to day operations. It was suggested to spend more time on the first day of each week of the experiment to go over these in more detail before starting simulated operations. It was also suggested there be "quick guides" developed for all RGBs showing basic use and application to specific forecast problems. Many forecasters also commented that the list of RGBs should be reduced as there are too many to choose from, especially since some RGBs are very similar and have very similar applications.

A suggestion for the "split window" difference product was to change the color scale, as the higher difference values are a very bright pink color that is harder to detect visually. Also, in the summertime when the land is warmer, there are a lot more of these bright pink values in the background and the moisture convergence field does not show up well. Many suggested a grey scale type color map that is visually apparent, with coloring for the max values that show more clearly the zones of low level moisture pooling.

> "Color curve on moisture difference could be improved. Bright pink colors were hard to interpret at times. Otherwise data was very useful."
> *Forecaster, End-of-Week Survey*

> "Consider permanently changing the default color curve for the moisture difference product to the grey scale that was available in the procedures this week."
> *Forecaster, End-of-Week Survey*

> "There are numerous "Convective" RGB's that are very similar to the normal Visible imagery and don't really provide much additional impact to the forecaster. Thus, I don't really see myself using them in the future."
> *Forecaster, End-of-Week Survey*

> "In addition, it would be nice to have quick guides on the RGB products to help highlight the reason they should be used and what you should focus on when looking at them. I spend too much time trying to remember the color cube and mentally deriving out why certain colors were appearing that it doesn't make it operationally useful to use in real-time."
> *Forecaster, End-of-Week Survey*

Other ABI Comments

> "There are many benefits with these new products/tools to improve on a day by day weather scenario as well as severe weather days. The reality is that forecasters may not get to use them all and probably set several procedures with 1/2 of the products available. I will suggest to continue working with forecasters (NWS and media) and get their feedback on these new tools. Collaboration is critical when creating and implementing new products for us to use and for the public to understand."
> *Forecaster, End-of-Week Survey*

> "Overall I think the ABI products are really useful and a huge improvement from previous GOES imagery. I'd imagine with more experience I'll find additional uses for many of the products I'm not currently using."
> *Forecaster, End-of-Week Survey*

3.3 Probability of Severe (ProbSevere) Model

Cooperative Institute for Meteorological Satellite Studies (CIMSS)

The NOAA/CIMSS ProbSevere statistical model, planned for operational implementation by NCO as an update to MRMS in 2018, was evaluated in the HWT for the third consecutive year, with updates made since last year's experiment. ProbSevere is currently undergoing tuning and assessment with the in-orbit ABI and GLM data for future demonstrations. The statistical model produces a probability that a storm will first produce any severe weather in the next 60 minutes (Cintineo et al. 2014). The data fusion product merges RAP model-based instability and shear parameters, satellite vertical growth and glaciation rates, radar derived maximum expected size of hail (MESH), and Earth Networks (ENI) total lightning information. Additional RAP and Multi-Radar Multi-Sensor (MRMS) fields such as azimuthal shear were used in the model this year to provide guidance on specific severe hazards of tornado, wind, and hail. ProbSevere tracks a developing storm incorporating data from both satellite and radar imagery using an object-oriented approach. As the storm matures, the Numerical Weather Prediction (NWP) information, lightning data, and satellite growth trends are applied to the overlapping radar objects. The product updates approximately every two minutes and is displayed as contours with different colors and thicknesses corresponding to different probability value bins that are overlaid on radar imagery. Data readout is available by moussing over the probability contour, revealing the probability of severe for each hazard (hail, wind, and tornado), along with the model predictor values. The product was evaluated on its ability to increase forecaster confidence and skillfully extend lead time to severe hazards for NWS warnings during potential severe weather situations. Additionally, feedback regarding the product display and readout was solicited.

Use of ProbSevere in the HWT

Forecasters loaded the ProbSevere guidance as an overlay on either the base radar data, or MRMS products (e.g., Composite Reflectivity, MESH, and Reflectivity at Lowest Altitude) at the beginning of each shift. Early in the shift, ProbSevere alerted forecasters to the first storms of the day that were becoming potentially significant and warranted closer monitoring. Forecasters consistently commented ProbSevere was very useful for the situational awareness it provides to alert forecasters to storms that need further interrogation. This was especially important in busy warning situations where many storms were present. ProbSevere was often used as a guide to quickly rank storms in terms of importance to interrogate based on the higher probabilities (Fig. 6). In other situations where warning issuance was marginal or more uncertain based on the base radar data, ProbSevere would sometimes provide more confidence to issue the warning. It is important to recognize that forecasters did not use ProbSevere alone to issue warnings, but instead based their decisions on what ProbSevere was showing in context with other observational datasets. Forecasters would also often take note of the parameters within the readout of the ProbSevere contour in AWIPS-II to monitor changing input parameters to help better interpret the trends in the probability values. This gave confidence to the forecaster in how the storm attributes were evolving and how the local environment was changing, and provided much more insights into the ProbSevere algorithm performance. With all of the great uses for ProbSevere, there were also several limitations and suggestions from the forecasters to improve both the model and product display. Both uses and limitations will be discussed in this section.

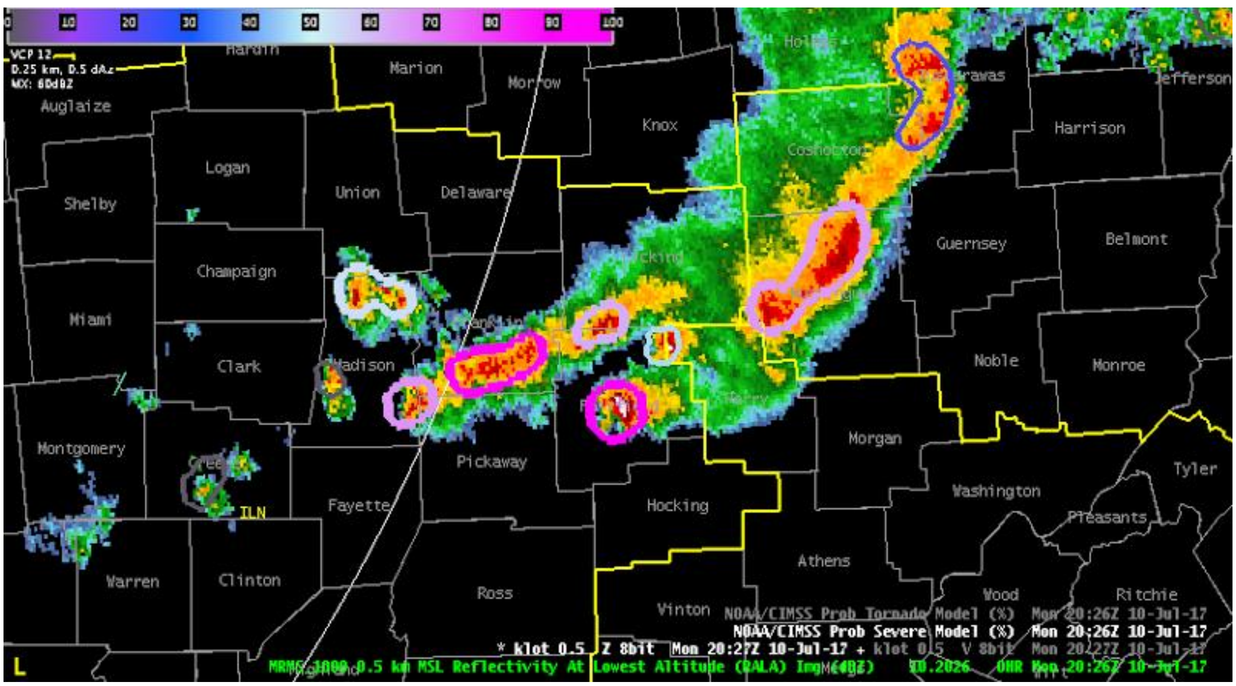

Figure 6: 2026 UTC 10 July 2017 MRMS Reflectivity at Lowest Altitude (RALA), and ProbSevere Model contours in western Ohio and far eastern Indiana. The ProbSevere Model helped forecasters rank storms during this busy warning environment.

> "It is a great tool to use for monitoring storms/situational awareness. It will definitely help the forecaster confidence when the tool provides elevated/high probability numbers."
> *Forecaster, End-of-Week Survey*

> "I do not think the ProbSevere model should be used solely for warning decisions, but it can be a very useful tool in situational awareness and monitoring trends. There is a lot of MRMS data available and ProbSevere pulls all of this information together in one tool that can be combined with radar during storm interrogation. I did not have any tornadic

storms so I do not have any comments on the ProbTor model. I do however like how the new version separates the probabilities into the different categories."
Forecaster, End-of-Week Survey

"Useful for SA with lots of storms"
Forecaster, "Day 2 Wrap Up and Feedback (19 July 2017)", *GOES-R HWT Blog*

"Good as an SA tool, especially in a busy environment and in the junky convection days. Helps bring attention to the storms that are different."
Forecaster, "Day 4 wrap up and weekly feedback (14 July 2017)", *GOES-R HWT Blog*

With the addition of many more inputs into the algorithm, such as low-level azimuthal shear, mid-level azimuthal shear, 0-1km Storm Relative Helicity (SRH), and Wet bulb zero height, for the first time ProbSevere was tested in the HWT to provide probability guidance for three different hazard types (tornado, wind, and hail) (Fig. 7). Also available to the forecasters was a full readout of the input parameters for each probability contour and a "light" version of the readout with fewer parameters so that the readout would not take up the entire screen when sampling the probability contour. Lastly, a separate ProbTor contour was available to examine the probability of tornadoes based only on the inputs for tornado guidance. Overall feedback was positive on these additions, with a few negatives which will be discussed below, and forecasters liked being able to see what parameters were utilized by the model and how changes in the parameter values impacted the resultant probability guidance. This helped the forecaster better identify the type(s) of severe risk and improve the wording of hazards within the warning. Many forecasters liked having both the full readout and the light readout depending on how their display was set up at the time. In fact, 72% of forecasters answered that they liked having both readouts available to look at (41/57 answers). The full readout was good for full screen analysis, but when forecasters loaded ProbSevere on a 4-panel display, many wanted to use the light version as the full readout would overlap into other panels.

"Like having the full text and light options... preferring to start events will all the information and then decreasing to light to declutter as event progresses."
Forecaster, End-of-Week Survey

"I really like the two being separate and watching for trends in both is a good situational awareness tool. Like we have mentioned this week, its best to focus on trends and not the actual value."
Forecaster, End-of-Week Survey

"Very helpful! I usually do not like to use products like this, mainly because I never know what is going into it. Having the option to display each parameter in the algorithm as well as showing the value for each one gave me more confidence in the accuracy of the display it was showing. I still used it mostly for situational awareness, but it agreed very well with the strongest convection and where I would've issued a warning without its aid."
Forecaster, End-of-Week Survey

At the very least, ProbSevere enhanced forecaster confidence when issuing severe thunderstorm warnings, but played a lessor role for tornado warnings. On days when ProbSevere was used in warning decisions, forecasters felt that it increased their confidence in issuing severe thunderstorm warnings 88% of the time (37/42 answers), whereas they only felt that ProbTor increased their confidence in issuing tornado warnings 43% of the time (6/14 answers) when tornado warnings were issued. The small number of tornado cases during this experiment could have prevented a more robust evaluation of the ProbTor model. Forecasters noted that the trends in ProbSevere were the most important in focusing their attention on storms and adding confidence to the warnings. A quick jump in probabilities over a couple of radar volume scans was a key indicator to forecasters that the storm was intensifying rapidly and would most likely soon become severe. In these situations, forecasters recommended it was best to wait for one or two scans of rapidly increasing probabilities (e.g. from say 10% to 50% to 80%) and to interrogate the base data further before making a warning decision. 60% of the time, forecasters felt that ProbSevere increased their lead time for severe thunderstorm warning issuance, and felt that ProbTor increased lead time in tornado warning issuance 43% of the time. When asked after each shift if they would use ProbSevere in operations, 100% answered yes.

> "I feel there is utility to the ProbSevere model and could supplement our normal radar interrogation. It gave me more confidence in warning issuance and helped with lead time and wording in my warnings."
> *Forecaster, End-of-Day Survey*

> "ProbSevere really increased warning lead time with storms in FSD area."
> *Forecaster,* "Day 2 Wrap Up and Feedback (12 July 2017)", *GOES-R HWT Blog*

> "The ProbSevere model was impressive today. It showed increased in probabilities as the storms were maturing and allowed me to easily focus on the storms that were the strongest."
> *Forecaster, End-of-Day Survey*

> "I think the trends on ProbTor and all the Probs are helpful. The ProbTor seemed to follow SRM trends and that sort of helps on keep an eye on things with those trends. Interesting to watch percentages as well."
> *Forecaster, End-of-Day Survey*

Many forecasters saw great utility in having the separate hazards and corresponding severe probabilities listed in the readout. An example from 22 June 2017 in eastern Colorado, shown below, shows how a forecaster found utility in using the different hazard percentages and how it affected their warning decision. ProbWind steadily increased from 1% to 73% over 20 minutes (Fig.7), which indicated an increasing damaging wind threat with the thunderstorms (in addition to hail). There is evidence in the radial velocity of accelerating outbound winds coincident with the time period that ProbWind increased. "Based on other environmental conditions", he decided that the wind threat had increased from 60 mph to 70 mph and incorporated this into his warning. Later on, wind damage was reported with this storm, although the wind speed was unknown. This example shows the utility the specific ProbSevere hazards can have in a warning environment.

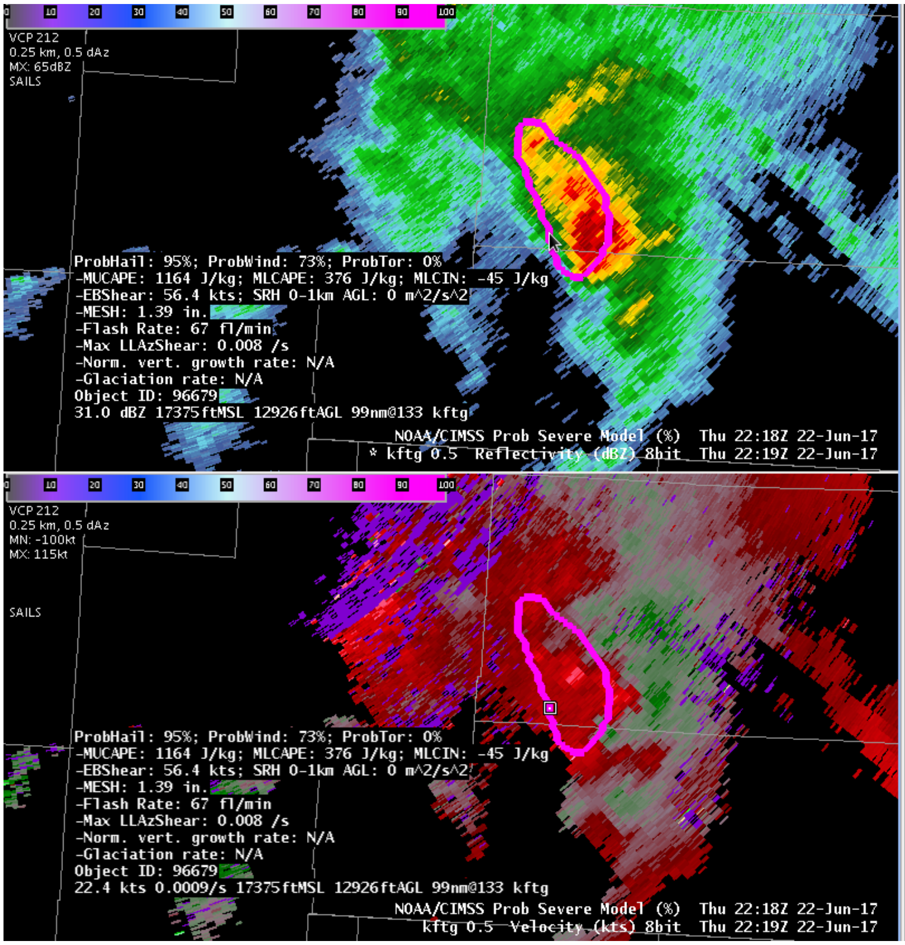

Figure 7: 2218 UTC 22 June 2017 ProbSevere probability contour (overlay), ProbSevere readout (text), radar reflectivity (top), and radar radial velocity (bottom) for a storm in eastern Colorado.

Limitations and suggestions for improvement of ProbSevere

There were still some instances commonly pointed out by forecasters where ProbSevere was not effective. Forecasters often found that the ProbSevere probabilities often lagged slightly behind the strengthening in base radar data both spatially and temporally due to the latency of the MRMS data processing and subsequent ingesting into the model. In rapidly strengthening scenarios this seemed to hamper the lead time of ProbSevere compared to that of the base radar data. Also, when storms were within close proximity of each other, the ProbSevere model would typically group the storms together into one contour, often times reducing the severe probability value drastically and making ProbSevere much less useful and trustworthy. Some ideas offered from forecasters included an option to have a user defined box drawn around a storm of interest and use that to track storms, or to possibly have a user-defined reflectivity threshold to follow the storm core more effectively.

> "Found that as storms morphed, got bigger, the ProbSevere became one big contour. That was frustrating when looking at ProbTor because one storm area had 3 circulations (few

> counties apart) and only one ProbTor. Maybe a user defined (drawn the box you want) option for ProbSevere/Tor would help."
> *Forecaster, End-of-Week Survey*

> "ProbSevere and ProbTor are useful for situational awareness; however, there is some lag from base data because of the use of MRMS. Though it probably isn't feasible from a bandwidth standpoint, I could see a significant advantage of incorporating data from individual radars."
> *Forecaster, End-of-Week Survey*

As has been the case in previous years of the HWT experiment, ProbSevere was most useful in cases where severe hail was the primary threat and in discrete storm modes. This is to be expected as there is no other input for wind or tornado that performs as well as MESH does for hail guidance. Many forecasters did comment that this iteration of ProbSevere seemed to do better with wind and tornado events than previous versions, but could still use some further improvements before being ready for use in primetime in operations. Many of the issues with the ProbWind seemed to stem from storm tracking issues and the significant wind being displaced from the reflectivity core, such as in cases of outflows racing out ahead of the storm. Forecasters also commented that the model could be better trained for different regions of the country with different environments and threats than others.

> "ProbSevere is great. Best with hail, not so much due to displacement issues in wind."
> *Forecaster, End-of-Week Survey*

> "In one case, we observed a gust front that we issued a SVR for a storm. A subsequent LSR had measured 67 mph winds, the gust front was outside of the circled "body" of the storm, and hence had very low ProbWind. We had another case where there was a large area of severe outbound winds in a storm. ProbWind was low. I think the explanation was that there was not enough gradient (not sure about that). Either way, it was in the circle and was easily seen on Base V. There were many instances when we were in a marginal or low severe environment that it was off. I should mention that none of my cases were in a good tornado environment, so I never really had a chance to evaluate ProbTor."
> *Forecaster, End-of-Week Survey*

> "ProbWind seem to have trouble with regards to severe outflow winds. The outflow separates from the monitored convective cell so it is no longer tracked but it would be useful for the algorithm to have an assessment of the environment to where it can recognize the conditions in which severe outflows are possible. This is a problem we often face out west where we can get severe outflow winds even with 40 dbz showers."
> *Forecaster, End-of-Week Survey*

Suggestions for Product Display

In general, forecasters liked the ProbSevere contour display and readout, but there were circumstances that forecasters offered suggestions for improvement to the current display. There were times when forecasters preferred an alternative color scheme that would make the jumps in probabilities stand out more. One such suggestion was to create a color scale showing the probability contours in ten percent increments with more contrasting colors, so that the magnitude increases are more readily apparent compared to the gradient color scale available now. Also, given the importance forecasters placed in monitoring the trends of ProbSevere (ProbHail and ProbWind) and ProbTor, an easy method to view trends (e.g., time series plot) was suggested to more quickly evaluate how a storm's probabilities were changing. Forecasters also noted that they would like to see qualitative wording in the readout alongside the azimuthal shear values to give them a better sense of what values are strong or weak, similar to what is done with the satellite vertical growth rates and glaciation rates. There were also comments suggesting that the ProbTor contour be a little thicker so that it can be overlaid and seen with the ProbSevere contour at the same time. Finally, many forecasters wanted to manually configure what parameters to display in the readout, so that each user would have their own configurable ProbSevere readout with parameters they feel are most important.

> "The color map (color curve) would be better in 10% increments. I personally didn't like the blue to white (default) color map. I know this is a personal configuration, but I feel a color table similar to radar (green to red to white) doesn't blend in with the contours. I would also make the ProbSevere or the ProbTor contours a little wider or at least make them different. If you overlap them... the ProbSevere contour is a little larger than the ProbTor contour. I wished it would focus on the actual circulation on the radar instead of contouring the same area as the ProbSevere contour. When a cluster of storms had 2 or more circulations, the contour would group the storm and the % would go down. In actuality, one of the circulations increased dramatically, but the % dropped because of the grouping."
> *Forecaster, End-of-Week Survey*

> "I would like to see either the ProbTor or ProbSevere outline be a little wider. That way it will be easier to view the data when they are overlaid. Currently there is a slight difference in size, but it made it too difficult to have both on at the same time and be able to see what the values were."
> *Forecaster, End-of-Week Survey*

> "Make the list for ProbSevere light user configurable. Forecaster could list only those parameters that are the focus for that event. Similar to ProbTor have a ProbWind and a ProbHail available as well."
> *Forecaster, End-of-Week Survey*

> "I'd like to see trends in each product displayed visually somehow. I found that the trends as important as the actually % numbers."
> *Forecaster, End-of-Week Survey*

> "Need some frame of reference for the parameters that come up when sampling, especially the AZ shear. I was not sure what values of those were significant."
> *Forecaster, End-of-Week Survey*

3.4 Geostationary Lightning Mapper (GLM) Lightning Detection

University of Oklahoma (OU) /Cooperative Institute for Mesoscale Meteorological Studies (CIMMS), NOAA/National Severe Storms Laboratory (NSSL) and
NASA-Short-term Prediction Research and Transition Center (SPoRT)

The demonstration of the GLM Level 2 products consisting of Events, Groups, and Flashes in the HWT provided a unique opportunity for forecasters to view data from this new instrument in a simulated operational setting for the first time. Unfortunately, a number of modifications were required early in the post launch testing of GLM prior to forecasters viewing the data in the HWT. Additionally, following this evaluation, the recommendation is that the data should not be provided to NWS operational offices until both initial visualization difficulties and erroneous geo-location problems (yet to be fixed in the satellite ground system for dissemination to users) can be addressed.

Since the GLM data were not available via the satellite broadcast network (SBN) at the time of the experiment, NOAA/NSSL and NASA/SPoRT utilized a local data manager (LDM) connection to provide data in real-time to the HWT following the methodology for previous data sharing (e.g., Lightning Mapping Array) already in place between the institutions.

Upon the arrival of the data at the HWT and viewing within the AWIPS-II software platform, the local PIs deemed it necessary to revise the initial visualization implementation prior to forecaster use and analysis. This initial default visualization merely forced the GLM data through the AWIPS "Lightning Data Ingest" that was created for ground-based point data of the National Lightning Detection Network (NLDN). As such, this initial visualization incorrectly displayed the groups, events, and flashes from GLM data simply as negative Cloud-to-Ground (-CG) points including within legends and data descriptions (Fig. 8). Locally, updates were made to the ingest code to change formatting and read out of the data on the screen, this included but was not limited to, moving away from the –CG nomenclature and plotting and updating the color scales and tables (Fig. 9). Additionally, a "GLM combination" product was created that utilized the gridded Event Density and flash points for the forecasters (Fig. 9b). This combination product provided the forecasters with the count of total flashes for trends and comparison with other networks while also providing an extent density that is unique to GLM. The code base to modify the default visualizations was provided to the NWS Meteorological Development Laboratory (MDL) as a temporary fix until a different ingest and display process can be developed.

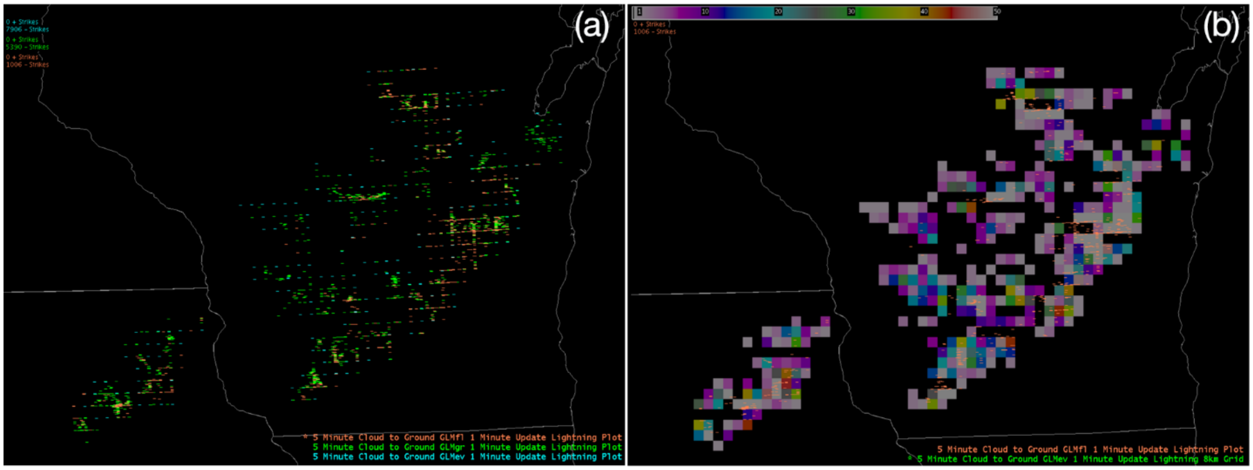

Figure 8: Default GLM visualization. (a) Points are displayed for groups, events, and flashes. All are displayed with (- symbol, reserved for –CG lightning). (b) Gridded event density with default color map. Legends refer to data as Cloud to Ground lightning in both images.

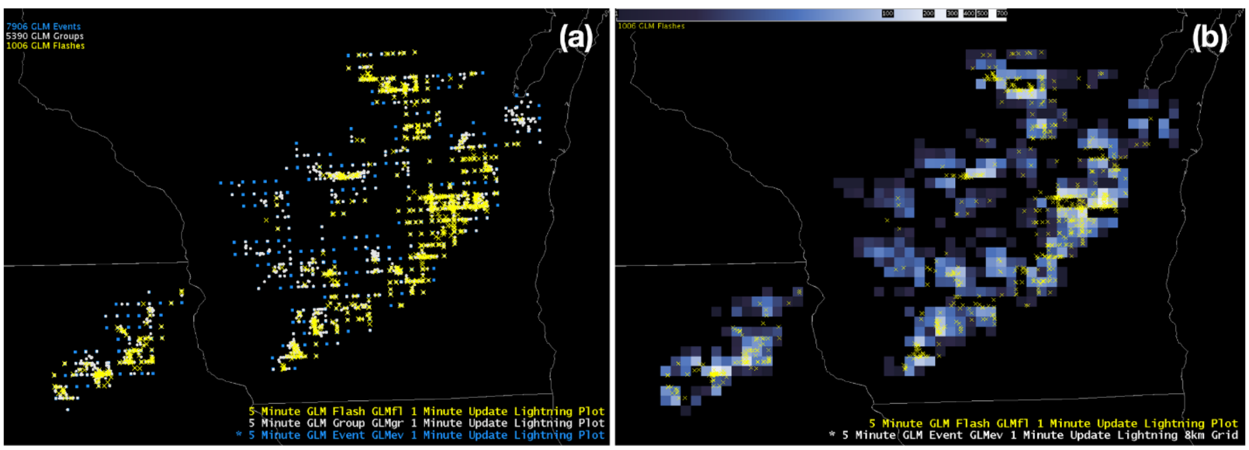

Figure 9: HWT GLM visualization. (a) Points are displayed for groups, events, and flashes as blue square, white circles, and yellow x, respectively. (b) Gridded event density with updated color scale and table (multiple color tables were produced for the evaluation).

GLM Use in the HWT and Ideas for Improvement

One of the primary recommendations following this evaluation is that the data are provided to the NWS at the native resolution of the instrument (8 km near NADIR and stretching to 10-12 at the edges of the field of view), to reduce visualization errors that occur when the data are viewed on a fixed grid (Fig. 10). Within the evaluation, multiple grid resolutions were utilized by forecasters, while forecasters preferred different grid resolutions for multiple reasons, 9 km spacing appeared to be the best option across multiple forecast offices across the CONUS when forced to fixed grid. As shown in detail in Fig. 10, and commented by forecasters in blog posts, the 8 km gridding often produces false gaps in the data when operating in regions away from the center of the field of view, but the 10 km gridding often has the appearance of a false increase in density when 2-4 pixel centroids fell within a single window.

> "Ultimately... the 8km grid of the events doesn't line up with the satellite's actual solution. If the events grid spacing is set at 8km... there are times no events are recorded,

since the points don't fall within the 8 km grid. If the grid is set to 10 km... there are double or even triple counts of the events. The best grid from what I worked with was 9 km. That gave the best representation of event counts without duplicating counts or losing counts where they turned to 0 events. There is promise with the GLM."
Forecaster, End-of-Day Survey

"They need to do away with the uniform grid sizing and go with the grid projections of GOES-16. For the most part the 9 km grid size is preferred due to its similarity to the GOES-16 projection. However, as you move farther from the nadir, you begin to see data issues where GLM events begin to get counted in two grid boxes due to the larger GOES-16 grid size the farther you move from the nadir. The rest of the GOES-16 data adjusts to the satellite's projection, why doesn't the GLM?"
Forecaster, End-of-Day Survey

"Would focus mainly on the 9km data being the default in AWIPS."
Forecaster, End-of-Week Survey

In regards to color tables, forecasters preferred a rainbow color curve that had an "increased contrast", highlighted peaks clearly between storms, and was the "best for showing dramatic lightning jumps" (Fig. 11d). Additionally, there was a strong preference by a majority of forecasters to turn on interpolation of the gridded data (Fig. 11). While this did reduce peak values in both the grids and mouse-over read out, forecasters found the interpolated product easier to understand quickly when comparing to other fields such as radar reflectivity since it "caught my eye better" and "seems a little more intuitive."

"I liked the color table with interpolation... it seemed to "look" more realistic, and to get some idea of lightning trends connected with a given cell. I would combine that with 1 minute strike data operationally... although, again, the need is there to give better strike location or else it's back to nldn."
Forecaster, End-of-Week Survey

"I assigned a warm color curve for use on top of vis satellite data which seem to stick out better in my mind. However, the assignment I made for the colors was arbitrary and developing color curves that may have better physical meaning derived from research would be better."
Forecaster, End-of-Week Survey

"I would recommend the rainbow color curve as baseline as it shows the most contrast from low to high values."
Forecaster, End-of-Day Survey

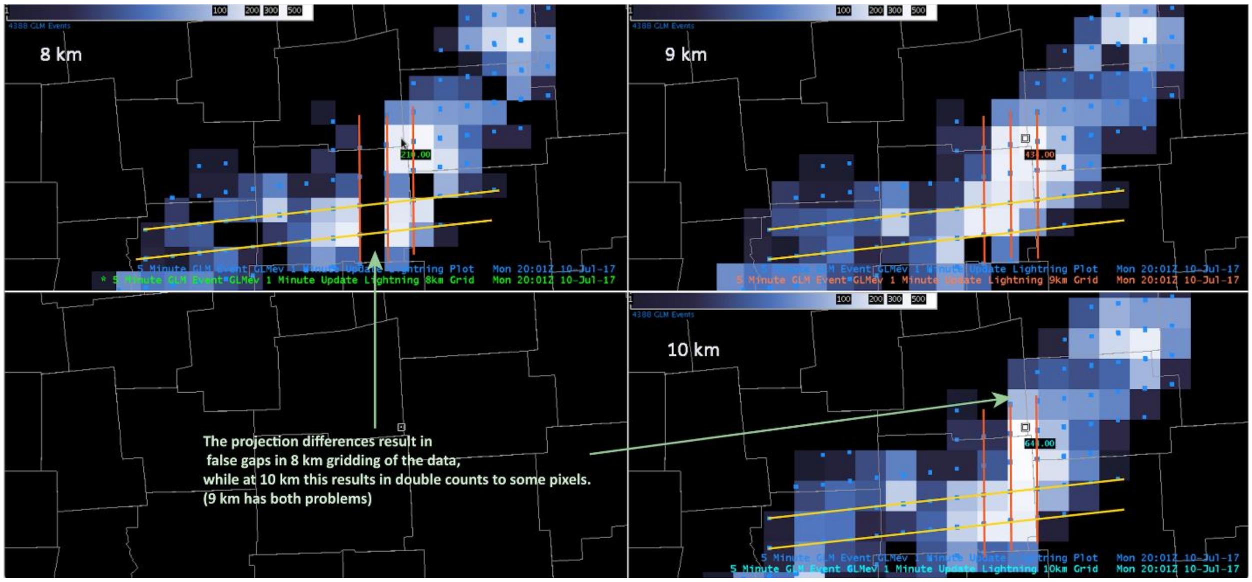

Figure 10: GLM fixed grid resolutions evaluated within the AWIPS visualizations. Blue squares are located at the Level 2 latitude and longitude locations of events. Gridding at 8 km (top left), 9 km (top right), and 10 km (bottom right) was evaluated by forecasters.

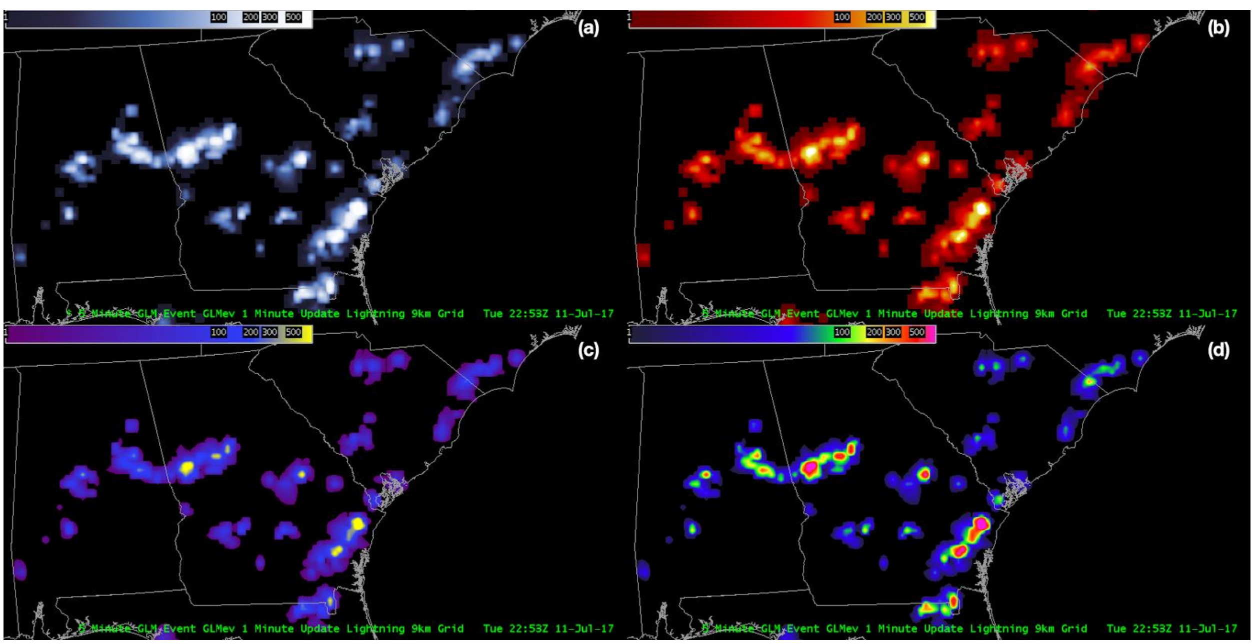

Figure 11: Interpolated GLM color tables tested by forecasters for gridded density products. (a) Initial HWT color table. (b) Red-yellow color table designed to highlight regions of increased activity when overlaid on visible satellite. (c) Color table designed to highlight lowest and highest rates to pull eye to extent and peak areas. (d) Rainbow color table designed to highlight contrast between storms. The rainbow color table was the preferred by a majority of forecasters.

Forecasters commonly noted both the increased spatial extent of lightning when examining the event density grids and the lead time and increased data coverage relative to the ground based lightning detection and other systems such as radar ("The faster updating GLM data also helped show trends in the storms quicker than just viewing radar data alone"). For lightning safety and decision support services this was seen as an additional advantage of the GLM over the ground-based lightning detection networks.

> "Interesting to see a cloud to ground lightning strike occur well away from main updraft cores in area with little if any meaningful reflectivity."
> *Forecaster,* "Anvil Lightning Near Yankton", *GOES-R HWT Blog*

> "First CG strike was detected at 1957 UTC with the GLM giving 6 minute lead time before the first strike."
> *Forecaster,* "First Lightning Along in south central SD", *GOES-R HWT Blog*

> "A splitting storm over Hamilton, Ontario Canada on to Lake Ontario. The GLM coverage for events was much greater than that of ENTLN for both Canada and the lake. This is a limitation of ground based ENTLN sensors, especially over the marine areas. This is an example where GLM could be very useful when interrogating storms and watching for convective development for the issuance of Special Marine Warnings."
> *Forecaster,* "GLM Usage Over Canada/Marine Zone", *GOES-R HWT Blog*

> "Allowing us to see not only the information with the various flashes but the spatial extent of the events was very helpful. Will be useful in aiding customers of lightning threat."
> *Forecaster, End-of-Day Survey*

Suggestions for Improvement

A number of suggestions were made by forecasters to better understand and utilize GLM data for both storm interrogation and for comparison with other lightning data such as from NLDN. The forecasters repetitively stressed that they needed to be able to monitor storm trends from multiple networks simultaneously as they commonly overlaid point data from multiple networks. Each week of the experiment, forecasters suggested they would prefer to also view the data through time series. Forecasters often commented on rapid increases in the flash rate as viewed via the gridded density as the determining point for a severe warning or providing them with additional confidence in doing so (Fig. 12). While the gridded data were useful for this purpose, in addition to the time series information above they also wanted a formal storm-based jump defined via an algorithm. It was unclear to the forecasters, however, if the jump should be defined via the total flash rate or event rate over an area and would like research conducted to determine why one method is or may be better than another.

> "Ability to create time series information for storms or areas would be very useful."
> *Forecaster, End-of-Day Survey*

> "It would be useful to see trends in lightning (i.e., time series)."
> *Forecaster, End-of-Day Survey*

Even though the forecasters found the training "adequate" most forecasters were still confused regarding the event, group, and flash definitions and when to use each one and how this related to other lightning data and information they are more commonly accustomed. With "a variety of networks and terminology from each network (flashes, cloud flashes, pulses, events, groups, strikes, etc.) efficient use and physical meaning of lightning data is a constant point of confusion for forecasters." Following group discussions with forecasters, we suggest the development of locally-focused training highlighting regionally important use of the data (e.g., fire weather in California, tropical cyclones near the gulf coast, and severe storms in the central and southern regions). Additionally, forecasters frequently commented that menu options for choosing lightning data within AWIPS are overwhelming and that GLM additions would increase confusion further. Regarding this point, **we suggest only providing the gridded data at the native resolution and only at 1 min, 2 min and 5 min intervals; all with one min updates**.

Finally, it is strongly suggested based on forecaster feedback the parallax algorithm using an average cloud-height be removed from the Level 0/Level 2 data and added in only following testing of the non-parallax corrected data with the standard lookup table against one utilizing cloud-top height from the ABI. Forecasters noted that they found the current GLM parallax correction confusing when relating to other radar, ground-based lightning detections, and also satellite data. Forecasters noted they found it more confusing to relate lightning within storms of different heights to the other data due to this offset implementation. Specifically, in blogs forecasters noted: "I found myself more drawn to watching the ENTLN trend because the GLM was displaced out ahead of the convection and seemed a little distracting", "At first I thought that the GLM lightning was lining up with the gust front instead of the core behind it", and "it was a bit difficult to use GLM to assess trends in convection…lightning [was] observed well away from the actual updrafts…complex trends were difficult to assess."

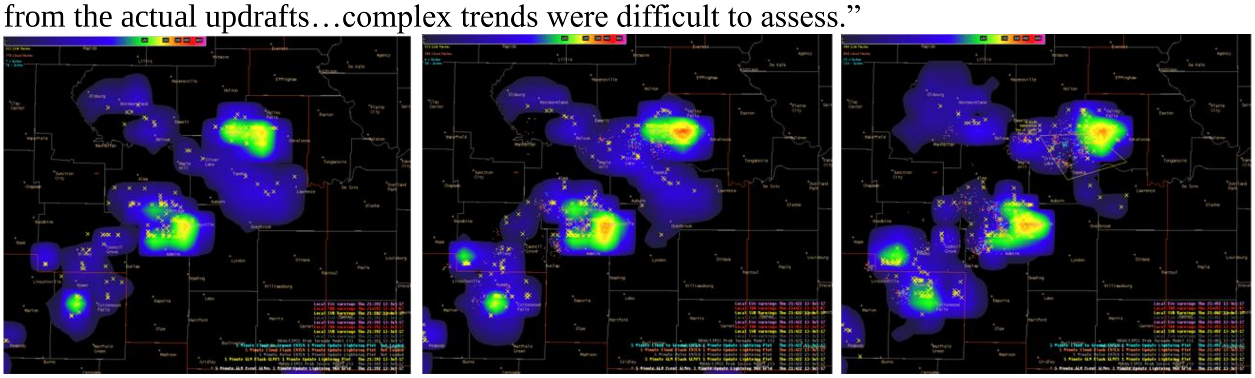

Figure 12: GLM gridded event density with GLM flash centroids (yellow X) and Earth Networks (cloud and CG) points overlaid in Topeka, Kansas area on 13 July 2017. Forecaster noted: we "issued our first warning, mainly due to a notable jump in lightning data from the GLM."

Overall, even with the caveats mentioned above, forecasters routinely found use for the GLM data in severe weather operations. The "GLM combination" product created at the HWT was the primary utilized option for the data with most forecasters utilizing the 5 min product and 1 min updates. However, in addition to local training development, it is recommended that the geo-location errors present during the experiment are fixed and a native resolution gridded product is provided to the NWS before fielding the GLM data.

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

Joint Polar Satellite System (JPSS)

The NOAA Unique Combined Atmospheric Processing System (NUCAPS) was demonstrated in the HWT in 2017 for the third year in a row. The atmospheric temperature and moisture profiles are generated using an algorithm that combines both statistical and physical retrieval methods. NUCAPS combines information from both the Cross-track Infrared Sounder (CrIS) and the Advanced Technology Microwave Sounder (ATMS) instruments aboard the Suomi-NPP polar orbiting satellite to provide soundings as close to the surface as possible. These profiles are produced at NESDIS/NDE and delivered over the AWIPS Satellite Broadcast Network (SBN) for display in the National Skew-T and Hodograph Analysis and Research Program (NSHARP) application in AWIPS-II. During the experiment, swaths of NUCAPS profiles from Suomi-NPP overhead passes were created over the east coast around 1800 UTC, central US around 1930 UTC, and western US around 2100 UTC with a typical latency of one and a half to two hours before the soundings were available for viewing by forecasters in AWIPS. Quality control (QC) flags associated with the NUCAPS profiles were also evaluated in AWIPS. These flags allow forecasters to quickly and easily identify which profiles within a swath passed (green) or failed (red/yellow) automated QC checks. The QC checks just check that a clean retrieval was obtained from both the infrared and microwave imager (green), just the microwave imager and not the infrared (yellow), or neither provided a clean retrieval (red). These QC flags do not directly determine the accuracy of the sounding and whether the sounding is am accurate representation of the atmosphere.

There were several other additions to the NUCAPS evaluation for 2017. An experimental version of the NUCAPS profiles was available for the Suomi-NPP passes during the afternoon. This version provides a correction in the boundary layer to surface temperature and dew point using nearby surface data. The correction inputs the Real-time Mesoscale Analysis (RTMA) surface observations for the new surface temperature and dew point of the sounding and then creates a mixed layer to the top of the boundary layer. The boundary layer height is determined by using Equation (1) below and then creates a new boundary layer for the existing NUCAPS profile based on these data. Plan view displays and vertical cross-sections of NUCAPS-derived thermodynamic fields were also available for forecasters to view in AWIPS. Finally, NUCAPS temperature and moisture profiles generated using data from instruments aboard the European MetOp-A and MetOp-B satellites were also made available in AWIPS. Swaths of NUCAPS profiles from MetOp-B were created from passes over the east coast around 1500 UTC, central US around 1630 UTC, and western US around 1800 UTC with MetOp-A soundings created approximately one hour later. This allowed for more sampling of the atmosphere between the typical 1200 UTC and 0000 UTC Universal Rawinsonde Observation Program (RAOB) soundings. The latency of the MetOp soundings was similar to that from Suomi-NPP.

Equation (1): $$z_{i+1} = [{z_i}^2 + \frac{2}{\gamma} C_H |V| (\theta_{Skin} - \theta_{Air}) \Delta t]^{\frac{1}{2}}$$

z – Height of mixed layer

θ_{Skin} – Potential Temperature of surface skin (GOES-16 11/12 um)

θ_{Air} – Potential temperature of surface air (RTMA)
|V| – Wind Speed (RTMA)
γ – Lapse rate of free atmosphere (NUCAPS)
C_H – Exchange coefficient (constant)

The purpose of the NUCAPS demonstration was to assess the value added of NUCAPS data to the severe weather Nowcast and warning process and to determine suggestions for improvement for readiness in operations.

Use of NUCAPS in the HWT

The key benefit of NUCAPS noted by forecasters was that it can help fill the gap between the morning 1200 UTC and the evening 0000 UTC RAOB soundings. The timing of the soundings is often just prior to convective initiation and during developing convection, although latency concerns at times cause the soundings to not be viewable until after initiation has occurred. Given that its availability is in the early afternoon, NUCAPS was primarily used by forecasters to assess the pre-convective environment and environments in the vicinity of developing storms. After the storms were ongoing, NUCAPS would be used to analyze environment ahead of the convection for anticipating intensity fluctuations in the storms. Forecasters also noted the utility of NUCAPS in data sparse regions between radiosonde data. Although there were no observational RAOB soundings to compare to in these regions, forecasters enjoyed having some type of data to compare to models and provide some input on the state of the atmosphere in an otherwise data void area. NUCAPS soundings were also used to compare to model soundings to determine if the environment was evolving as expected from the model forecasts and analyses. Forecasters were also able to see how the atmosphere had evolved from the early morning sounding, to the late morning overpass of the Met-Op profiles, and eventually to the afternoon profiles from Suomi-NPP. This was important at the beginning of each shift to see how things had changed since that 1200 UTC sounding. In addition to the profiles, fields derived from NUCAPS that forecasters found particularly useful included lapse rates, freezing level, -20C level, mid-level temperatures for capping inversion monitoring, and precipitable water and other moisture trends. Many of the surface based parameters were not too helpful from the NUCAPS soundings because of the satellite's inadequate sampling of the boundary layer.

> "Pre convection and developing they were most useful. I like the gridded during convection but found the soundings prior to development useful as well. Seeing how the boundary layer is changing from the morning sounding was nice to see."
> *Forecaster, End-of-Day Survey*

> "Pre-convective & developing: In our case today both NUCAPS and GOES16 Derived indices both showed a lack of available instability. This held true with the lack of deep convection."
> *Forecaster, End-of-Day Survey*

> "Temperature fields were what I used most, to try and assess the strength of the cap."
> *Forecaster, End-of-Day Survey*

"700 hPa, 850 lapse rates... looking for cap strength, and whether existing convection was moving into a more favorable environment from the west."
Forecaster, End-of-Day Survey

"For today, the mid-level lapse rates, because when comparing it to some of the GOES derived products such as CAPE and PW, the same gradient from western to south-central ND showed up. This was helpful in determining when convection might weaken."
Forecaster, End-of-Day Survey

"I think the NUCAPS does provide useful information between raob soundings, particularly for monitoring convective trends, but also perhaps things like inversion layers and such."
Forecaster, End-of-Week Survey

NUCAPS also proved to be valuable for monitoring sub-severe convection along with a fire on 10 July in the Billings, Montana CWA. In this case, the NUCAPS sounding pass was around 1930z (Fig. 13). Several soundings were plotted near Billings and locations to the east, ahead of convection. A noticeable inversion was detected near/just above 700hPa. Compared to the HRRR (Fig. 14), RAP, and NAM soundings (not shown) valid at a similar time, the guidance failed to resolve this feature. Additionally, the special 18z observed sounding from Glasgow, Montana (much farther north and east than Billings) was compared to the nearest NUCAPS profile to that sounding location. The forecaster then stated that "no noticeable differences that would have a sizable impact on mesoscale forecasting were noticed."

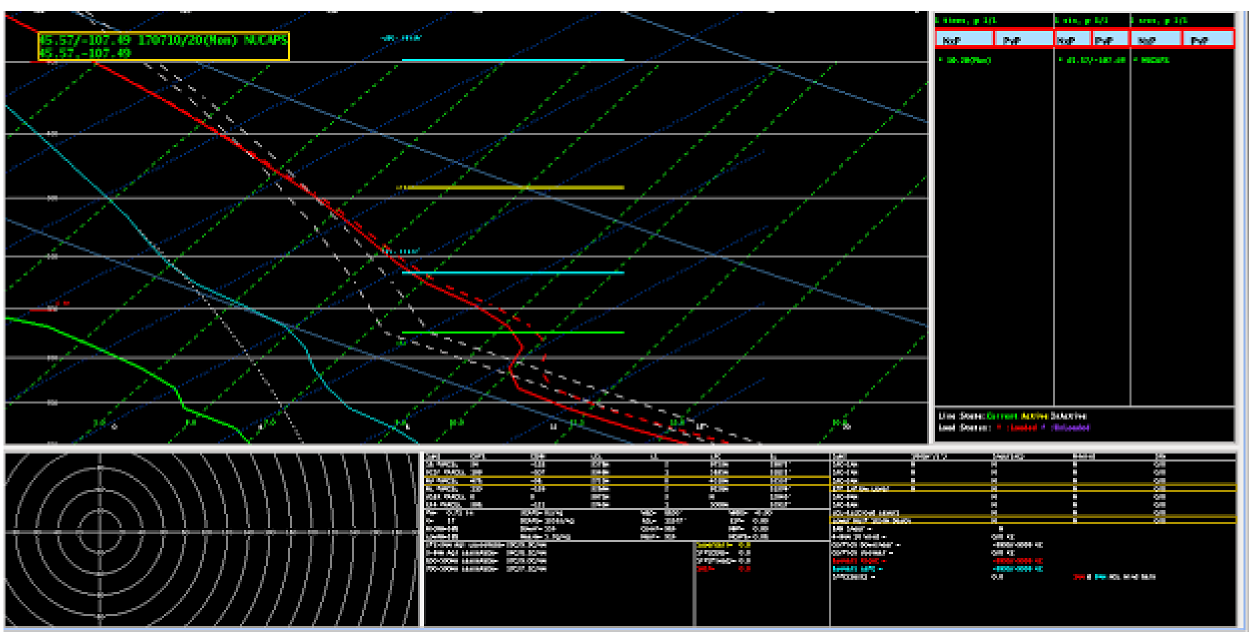

Figure 13: 1930 UTC 10 July 2017 NUCAPS temperature and moisture profile plotted in a skew-t diagram near Billings, Montana.

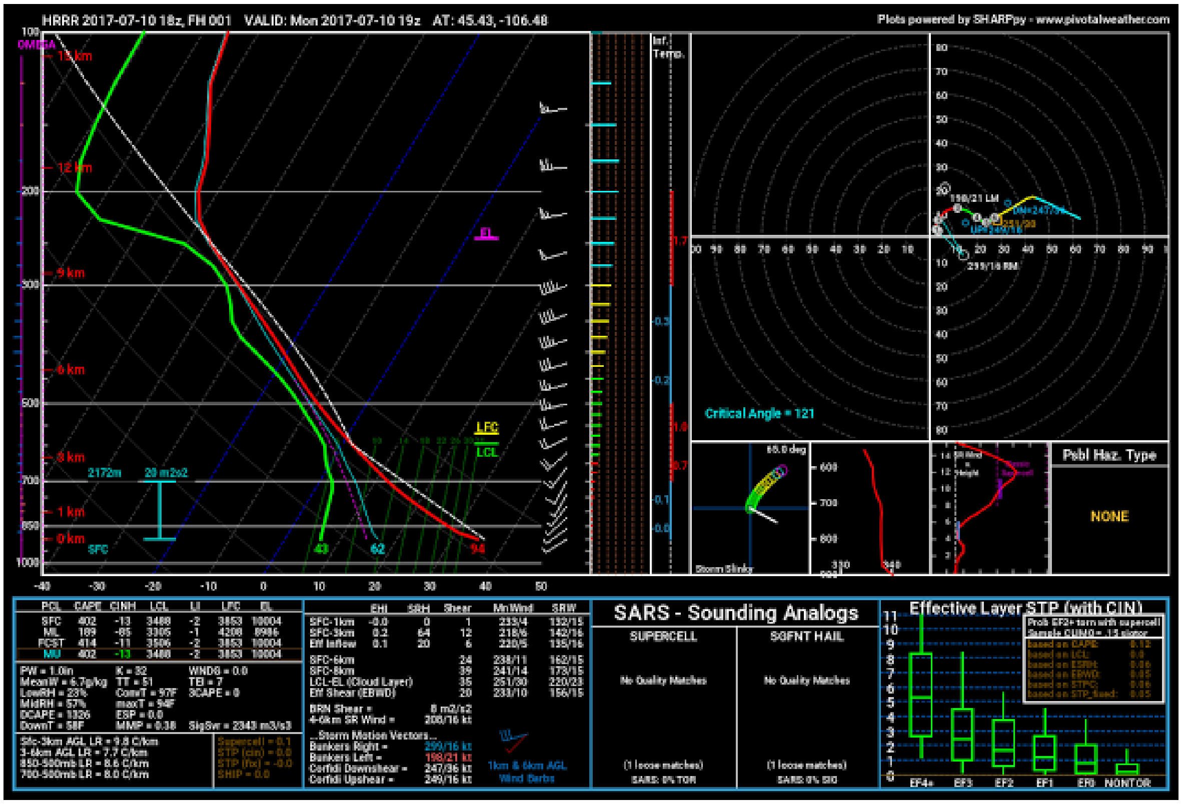

Figure 14: 1800 UTC HRRR 1 hour forecast sounding valid at 1900 UTC 10 July 2017 plotted in a skew-t diagram.

To further investigate the accuracy of the NUCAPS profiles taken east of Billings, forecasters used a smoke plume visible at 0.5 and 0.9 degrees on the Billings (KBLX) radar, near Birney, Montana (Fig. 15). This fire is approximately 90 miles from the Radar site, with the 0.5 degree scan intersecting at ~9800 feet above ground level (AGL). From the nearest NUCAPS sounding to the fire, the strongest point of the inversion was ~8300 ft. AGL. The placement of the fire and smoke plume suggests some accuracy of the NUCAPS capture of the inversion, which is missing from model guidance. They chose to investigate the smoke plume, because, given the distance from the Billings Radar, this plume would not likely be visible without an inversion causing the smoke to level off and begin to spread out horizontally. Therefore, the conclusion was that there is some confidence that the NUCAPS soundings/profiles were a fair representation of the atmosphere. Additionally, it was noticed that as convection moved eastward in the afternoon, its intensity decreased, which showed a potential impact of the inversion on the convective environment. Although there can be other factors that contribute to convective decay, the forecasters felt that the profiles in this situation were trustworthy and helped contribute to their forecast for the evolution of the convection throughout the day.

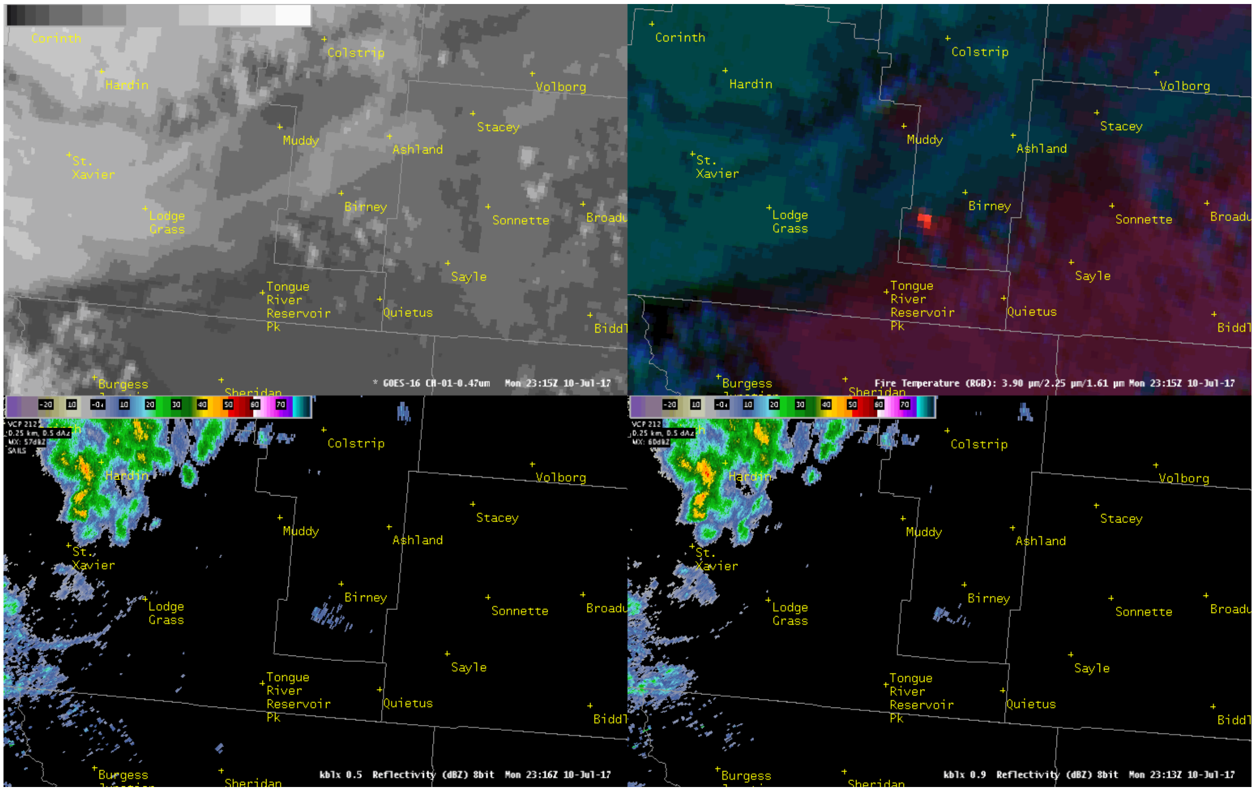

Figure 15: 2313 UTC 10 July 2017 AWIPS 4 panel; GOES-16 0.47um visible in the upper left, GOES-16 Fire Temperature RGB in the upper right, KBLX 0.5 degree elevation reflectivity in the lower left, and KBLX 0.9 degree elevation reflectivity in the lower right. Notice the smoke plume near Birney on the KBLX radar and associated hotspot on Fire Temperature RGB.

Throughout the experiment, forecasters consistently compared the NUCAPS profiles with those from other datasets such as model analysis or forecast soundings. Such comparisons allowed them to learn the strengths and weaknesses of the NUCAPS data, identify inaccuracies, and visualize environmental trends. In situations where a special afternoon radiosonde sounding was available, forecasters took advantage of the opportunity to learn key differences between RAOB or model point soundings and satellite-based sounding profiles. Most obvious was the smoother appearance of the satellite-based soundings, lacking the vertical detail provided by a radiosonde, and thus unable to resolve some smaller scale features visible in radiosondes. The occasional availability of temporally and spatially collocated radiosondes also allowed forecasters to gauge the accuracy of the NUCAPS data with observed soundings (Fig. 16). When the corrections were made by the forecasters to the surface temperature and dewpoint of the NUCAPS profiles, values of surface based fields like Surface Based Convective Available Potential Energy (SBCAPE) and lapse rates typically matched up well with those from a radiosonde. What was lacking when comparing these soundings was the accuracy of how the profile looked in the boundary layer. Often, the low levels were not very indicative of the boundary layer in the observed soundings. The upper levels (typically above ~700hPa) seemed to line up fairly well with observed soundings in terms of profile shape. Forecasters also compared the new experimental version of the NUCAPS soundings which modifies the boundary layer to these special soundings and NWP model soundings (Fig. 17 and 18). This allowed the forecaster to compare various stability parameters from the modified sounding with the radiosonde and model soundings, and assess the validity of this experimental sounding. Often times, the experimental sounding showed much more realistic values of SBCAPE and other surface based features than the non-experimental sounding when compared with model soundings and the SPC mesoanalysis. However as is mentioned previously, this correction applied a well-mixed boundary layer to the NUCAPS sounding, which was often times not very well representative of the sounding shape.

Figure 16: Experimental 19Z NUCAPS (left) and 20Z observed RAOB for North Platte, Nebraska on 27 June 2017.

The SBCAPE in, Figure 16, and other surface based parameters match up pretty well with the observed sounding. However, while the NUCAPS sounding has the big mid-level inversion it is quite a bit higher in the atmosphere than what was observed and therefore Mixed Layer Convective Available Potential Energy (MLCAPE) was drastically different. Above the inversion the dew point is too moist on the NUCAPS profile in the mid-levels while the temperature profile is pretty close. This is just one example of how the NUCAPS soundings were compared with other soundings and data points.

> "The experimental NUCAPS that was modified did very well today and really seemed to capture the pre-convective environment. I was really impressed by it overall compared with the raw NUCAPS soundings."
> *Forecaster, End-of-Day Survey*

> "For today, the Experimental NUCAPS performed much better than the SNPP NUCAPS. The primary reason was the experimental soundings I observed, better sampled or incorporated the surface data. The operational or SNPP soundings had poor surface/boundary layer data, which severely impacted its data, especially with derived parameters from the soundings, such as instability. Even still, data manipulation/adjustments had to be made to the boundary layer (not the surface data points) to bring the sounding more in line with the more thought of conditions, based on the initialized RAP and HRRR."
> *Forecaster, End-of-Day Survey*

> "NUCPAS was available to sample the pre-convective environment. Compared the 19Z experimental NUCAPs to the 20Z special RAP sounding. Saw some differences in the lower levels. The cap was gone on both the 19Z NUCAPS and 20Z RAP sounding but

the surface temperature at RAP was near 84/57 at 19z and 90/52 at 20Z. Mainly pre and developing. Was interesting to compare special sounding data, operational NUCAPS and Experimental. Strong CAP was evident early ion but NUCAPS data showed the CAP weakening."
Forecaster, End-of-Day Survey

The newer additions for this year's NUCAPS evaluation were met with overall positivity from the participants. Forecasters took advantage of the late morning NUCAPS profiles from the Met-Op satellite, and allowed forecasters to monitor the evolution of the atmosphere from the early morning radiosonde up through the time of convective initiation. Also the Met-Op NUCAPS allowed forecasters to examine the pre-convective environment on those days where there was early initiation or out west when Suomi-NPP NUCAPS was post initiation. Overall, forecasters would welcome more NUCAPS soundings from other satellites and getting more timely data to monitor the evolution of the environment. Forecasters also found that the experimental NUCAPS soundings were somewhat beneficial for examining the convective environment. They were consistently more in line with observations when it came to values of SBCAPE and Surface Based Convective Inhibition (SB CINH), while also retaining the mid-level and upper-level accuracy of the traditional NUCAPS soundings. Many forecasters preferred the experimental over the non-experimental due to these reasons, although further advancements in the experimental soundings are required to make them more accurate and to match up more realistically with observed atmospheric profiles.

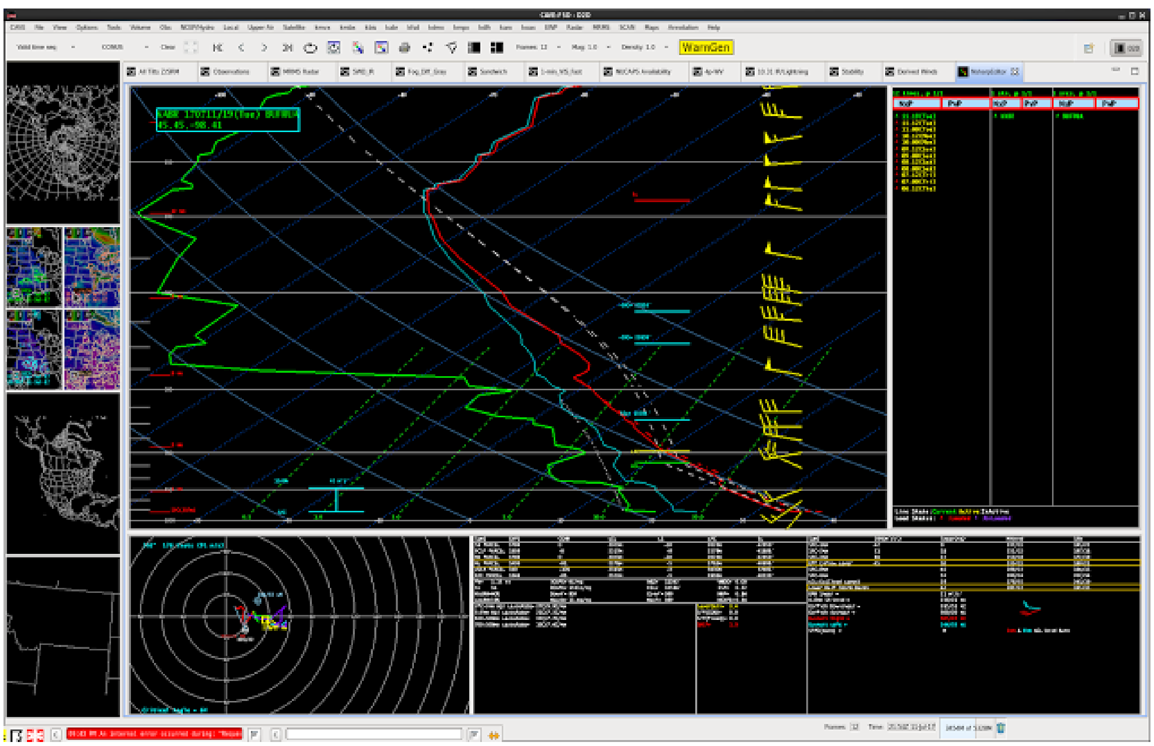

Figure 17: 1900 UTC 11 July 2017 observed sounding from Aberdeen, South Dakota

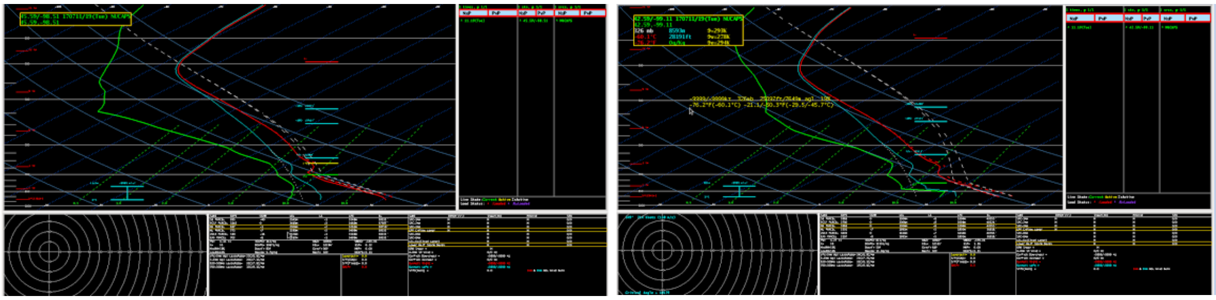

Figure 18: 1920 UTC 11 July 2017 NUCAPS sounding (left) and experimental NUCAPS sounding (right). Notice the modified dewpoint and temperature in the boundary layer on the experimental sounding, and overall the sounding shape seems much more accurate in the upper levels than in the lower levels. Also notice the inversion is picked up in the NUCAPs profile, but is displaced vertically much lower in the atmosphere than what was observed.

> "It would very useful in the field to be able to get multiple passes during the day to be able to monitor evolution of fields of interest. Use of MetOpA to supplement temporal resolution is useful and would be good to have these soundings on future polar orbiters. Ideally an hourly availability would be quite useful and even game changing for potential research items."
> *Forecaster,* "18Z NUCAPS (Op vs Experimental), viewing, and latency", *GOES-R HWT Blog*

> "I enjoyed using the soundings and the Experimental modified soundings when surveying the pre-storm environment. The Experimental modified soundings were far closer to reality when it came to instability. However, one of my fellow forecasts used the unmodified sounding to see 0C and -20C heights to help in hail forecasts, as that data seemed accurate."
> *Forecaster, End-of-Week Survey*

The gridded NUCAPS data were found to be somewhat useful at times as well, such as looking at the mid-level lapse rates. A lot of forecasters looked at these over the weeks and found having the satellite-derived lapse rates can be helpful in assessing the convective environment. In the image below over the northern Plains region (Fig. 19) the correspondence of the lower resolution gridded NUCAPS data (left) with the RAP model forecast (right) is shown for 850-700hPa and 700-500hPa lapse rates. This not only helped give the forecaster confidence in the observations and the convective potential on this day, but also helped give a little confidence in the model forecast for the situation as well.

Another case looks at mixing ratios to see how they compared with the latest RAP model run (Fig. 20). At 700 hPa (top two panels with NUCAPS on left and RAP on right), both were generally showing a dry tongue stretching from Tennessee into Arkansas and Missouri. They were also both in good agreement on the mixing ratio magnitudes over the primary forecast area for the day of Sioux Falls, South Dakota. Looking lower at 850hPa (lower panels) however, mixing ratio values overall were considerably less than what the RAP was indicating.

"Thus, confidence may be a bit better at levels at or above 700hPa, but not so good for 850hPa or lower. Overall, as you get closer to the surface, it looks like there is a tendency for NUCAPS to trend towards a drier solution than the models."
Forecaster, "NUCAPS Mixing Ratio", *GOES-R HWT Blog*

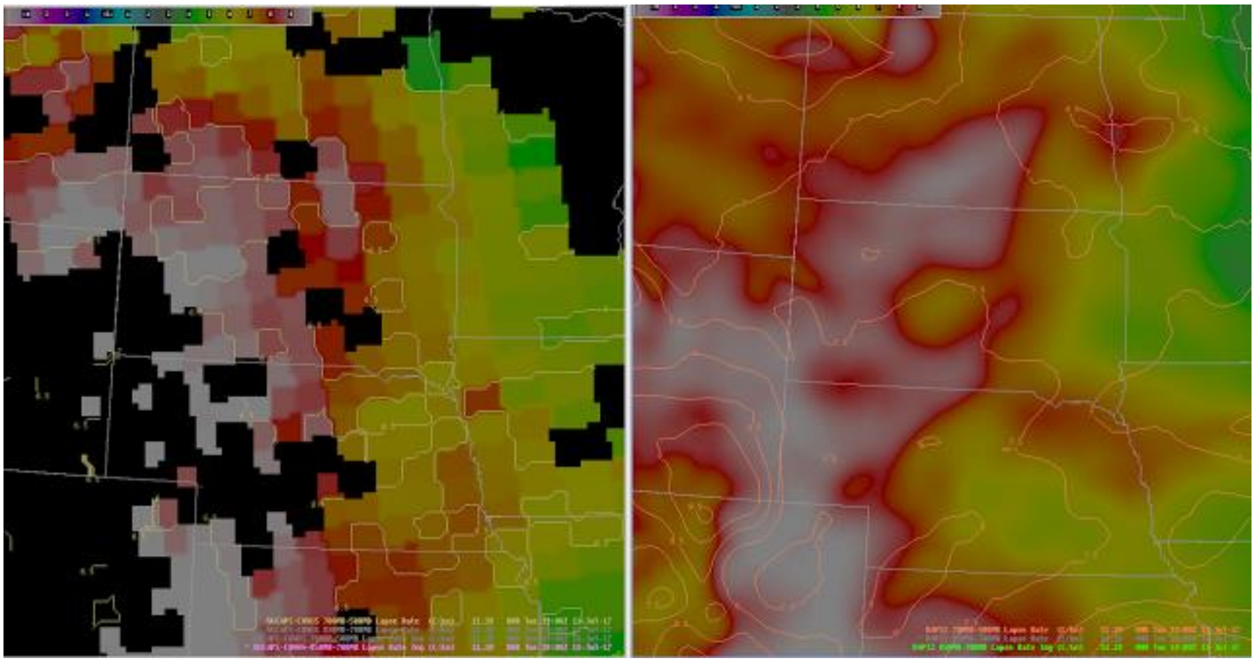

Figure 19: 1900 UTC 11 July 2017 gridded 850 – 700 hPa Lapse Rates (C/km) with 700 – 500hPa Lapse Rates (C/km) contoured for NUCAPS (left) and RAP analysis (right). Black pixels in NUCAPS image denotes missing data because of cloud cover or "bad" sounding retrievals.

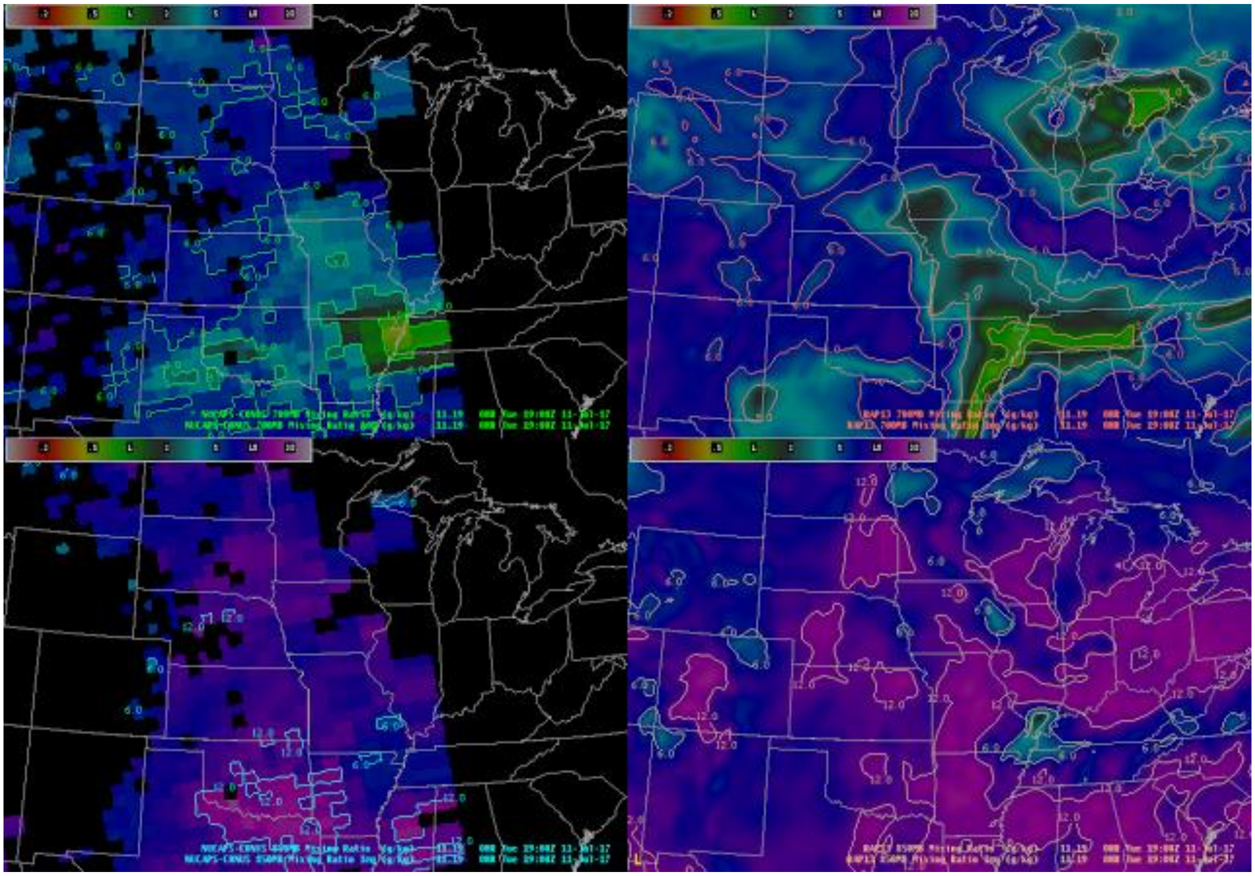

Figure 20: 1900 UTC 11 July 2017 Mixing Ratio (g/kg) for NUCAPS (left) and RAP analysis (right) at 700hPa on the top row and 850hPa on the bottom row.

Limitations and Suggestions for Improvement

As has been noted in past experiments, the latency of the NUCAPS profiles from timestamp to when they appear in AWIPS II is among the biggest drawbacks to using NUCAPS operationally in a convective environment. Many forecasters thought that the NUCAPS profiles provided value, especially in data void regions and having the experimental surface data modifications was applauded by all as being helpful for not having to manually modify the surface conditions in the soundings. But many times, owing to the long processing latency, the soundings were not available in AWIPS until after convective initiation had occurred, and were mainly used to confirm what had already been noted through other analyses. The data processing and transmission latency into AWIPS was consistently 1.5 to 2 hours, which made the data much less useful to most forecasters. Many forecasters commented that they would like to see the latency less than one hour for the data to be useful in a real-time rapidly evolving convective environment. Primarily because of the latency issue, 65% of forecasters said that the gridded NUCAPS provided small to very small operational impact on their forecast problem of the day.

> "I think they would be great if it didn't come in so late."
> *Forecaster, End-of-Day Survey*

“Latency was an issue for gridded data, would be useful to have during the pre-convective stage and is too slow to arrive after the pass. Would rely on other analysis tools.”
Forecaster, End-of-Day Survey

“The MOST important issue that needs to be addressed with NUCAPS is the time latency. In order to make it useful, it needs to be timelier. In a convective environment... the storms will probably be ongoing by 21 to 22Z, so I won't care about what the sounding looked like 2 hours ago.”
Forecaster, End-of-Week Survey

“It was discussed through the week, but the main issue I saw this week was the latency of the data. With development occurring during the afternoon, we sometimes saw the data after convection developed and it made it unusable in those cases.”
Forecaster, End-of-Week Survey

Forecasters frequently commented on the gridded plan-view data and some of the improvements they would like to see to make it more useful. Most stemmed from having better, more baseline parameters to evaluate. While the gridded data were considered useful for mid-level lapse rates and temperatures, they were not as helpful in analyzing values of CAPE and other stability indices. This was mainly due to having to use the volume browser and compute CAPE from a certain pressure level instead of from the surface. Forecasters would like to see additional parameters such as SBCAPE, MLCAPE, and Most Unstable Convective Available Potential Energy (MUCAPE) that are available in NSHARP sounding analyses. They would also like to have some pre-made menu options in AWIPS to display these parameters quickly without using the less efficient volume browser. Finally, many forecasters commented on the blockiness of the gridded data display, and all of the missing data in clouds (Fig. 19). They would like to see either better retrievals in cloudy regions, or fill in some of the gaps by another method to make the data field more continuous. On a positive note, 12/16 forecasters answered that they would be likely to use the gridded data back at their home office if some of these issues were fixed, particularly the latency issues.

“Stability parameters need to be calculated on surface level instead of the pressure levels for the plan view. There also needs to be selections for the different kinds of CAPE (e.g. MUCAPE, MLCAPE and SBCAPE). The plan view NUCAPS wasn't helpful all IMO. The data was a mess and didn't represent the environment at all.”
Forecaster, End-of-Week Survey

“Yes, getting CAPE values to be in the same format that meteorologists are used to (Surface based, Mixed Layer, and Most Unstable) would lead to more use. The current method of pressure levels is pretty useless.”
Forecaster, End-of-Week Survey

“If the data is of good quality, I could see this data being helpful in the winter time when the thermal profile is very important. BUT, with a slight change in temperatures around

> 0-3C making a large difference, the data must be very good to gain support by NWS forecasters."
> *Forecaster,* "Testing out the NUCAPS Cross Section", *GOES-R HWT Blog*

Forecasters also noted a few instances of NUCAPS sounding points passing the QC check (appearing green in the AWIPS display) but they were not considered to be representative soundings of the environment when compared with other observations and model soundings. Forecasters would like to see the QC process refined to improve the accuracy of soundings that pass the QC check. Seeing soundings that are designated to be reasonably accurate but appear to be problematic will cause many forecasters to doubt the validity of NUCAPS soundings and may discourage use of NUCAPS in the future.

> "When loading NUCAPS soundings, make sure you overlay satellite data with sounding points. Even though several dots were green and passes quality control, the sounding wasn't representative of the expected environment. There was also quite a bit of cloud cover (GOES16 Vis sat 0.64 (red channel) over that point, so even though it passed quality control... it probably wasn't a good sounding point."
> *Forecaster,* "NUCAPS Sounding Data", *GOES-R HWT Blog*

There were also some additional suggestions to make NUCAPS more desirable for use in WFO operations. Many forecasters would like to see a mouse-over readout, similar to ProbSevere, when sampling the sounding points to provide important parameters in order to get a quick look at the sounding-derived environment. There were also numerous comments about needing some type of station ID or better way of labeling the dots to make it easier to remember which point was chosen to look at previously. Some suggested having a better pop out Skew-t solution to be able to mouse-over the points and an NSHARP like sounding is displayed with a readout of parameters as opposed to the pop-up Skew-t now that just shows the sounding with no parameters. Forecasters would also like to be able to overlay multiple soundings at once as well as overlay observed and model soundings with NUCAPS soundings to better be able to compare between the environments. Lastly, many forecasters would like to see winds integrated into the sounding, whether that be from satellite derived winds or another source of wind data.

> "I would like the derived winds to show up in the soundings if possible. I would also like to see a better location ID when clicking on point soundings."
> *Forecaster, End-of-Week Survey*

> "Just started mulling this over... I think some sort of processing algorithms could be implemented when data comes in, to "alert" the user when certain thresholds are met.... something like "here's where the highest lapse rates are right now", or "here is where the 700 hPa cap is strongest/weakest." I think automation could be employed to great effect as soon as soundings arrive to highlight areas of concern."
> *Forecaster, End-of-Week Survey*

> "There should be a way to load the NUCAPS sounding window in an additional pop-up window. That way you would be able to know what point you selected."

Forecaster, “NUCAPS Sounding Data”, *GOES-R HWT Blog*

“Need some location ID on these sounding. Hard to remember which 'dot' one actually chose. Only one sounding at a time now but someday it may be more.”
Forecaster, “NuCaps/MET OP sounding Locations thought”, *GOES-R HWT Blog*

4. Summary and Conclusions

The GOES-R and JPSS Proving Ground conducted four weeks of satellite product evaluations during the 2017 Summer Experiment in the Hazardous Weather Testbed. Twelve NWS forecasters and four broadcast meteorologists evaluated many GOES-R and JPSS products and capabilities, and interacted directly with algorithm developers during the experiment. With GOES-R and JPSS being the sole focus of the demonstration, participants agreed that they had ample opportunity to subjectively evaluate, identify strengths and weaknesses, and suggest potential improvements for all of the products. An abundance of feedback was captured from participants via multiple methods, including daily and weekly surveys, daily and weekly debriefs, real-time blog posts, informal conversations in the HWT, and the “Tales from the Testbed” webinars. This feedback included suggestions for improving the algorithms, ideas for making the displays more effective, best practices for product use, and highlighting specific forecast situations in which the tools worked well and not so well.

Training, in the form of Articulate PowerPoint presentations for each product, was generally well received by participants. They were able to complete the training before arriving in Norman, and felt that it provided them with a basic understanding of each of the products. The only product(s) that forecasters felt less prepared to use were the RGBs, and more training of RGB applications was requested by every forecaster. Based on past feedback, more time was spent at the start of each week as a group going through each of the products in AWIPS. This included a brief refresher about each product, a tutorial on where to load the products in AWIPS, recommendations for pre-built procedures, and caveats. Starting the week with this walkthrough was applauded by participants, and contributed to a smooth start to experimental operations. Similar to last year, an information sheet listing each product under evaluation, its location in AWIPS-II, and contents of notable procedures was created for reference during experimental operations. The pre-built procedures were well appreciated (especially by the broadcast meteorologists) as they facilitated a quick start to operations.

For the fourth year, broadcast meteorologists participated in the Proving Ground Experiment equally with the NWS forecasters. Once again, the inclusion of broadcast meteorologists in the HWT activities went smoothly and proved to be fruitful for all participants. The broadcasters received a unique glimpse into the life of a NWS forecaster during simulated severe weather operations, noting the massive amount of data a forecaster must sift through and the substantial responsibility and stress one feels in such situations. Similarly, the interaction allowed NWS forecasters to gain insight from the broadcast meteorologists on some of their responsibilities, helping to unify the two groups. Broadcasters found at least some utility in all of the products demonstrated, and especially look forward to the high temporal resolution satellite imagery and using the GLM data to communicate lightning threats to the public. AWIPS familiarization at a nearby WFO prior to their arrival in Norman was vital for their successful participation in HWT activities during the week.

Overall, participants enjoyed their experience in the HWT, and felt that the experiment was very well organized. With the emphasis being on baseline satellite products and future capabilities, this activity helps to reinvigorate the use of satellite data in severe warning operations, fostering excitement and increased preparedness for the use of the GOES-R series satellite technology. Participants found at least some utility in all of the satellite products demonstrated, and look forward to using the data more in operations.

Summary of Recommendations for Improvement

There were numerous recommendations and suggestions for improvement of the accuracy and display of the products evaluated during the four week experiment. Starting with the ABI baseline derived products. It is recommended that the display for the derived motion winds be changed significantly to be useful to the forecasters. It is recommended that the wind barbs be color coded by height and not by speed to de clutter the display and to be able to easily determine what height a wind measurement is taken at. It is also recommended that the wind barbs stay loaded on the screen until the next update and that all of the observations are synched and time matched to avoid the “jumpiness” of the observations.

It is recommended to the satellite training team to spend sufficient time ASAP on training material for the RGBs and channel differences. Specifically it was desired that better training be provided on how and when to apply the RGBs and channel differences to a particular forecast problem. It is also recommended that a change to the color scale be made for the split window difference to use the more grey scale like what was used in the testbed instead of the bright pink and yellow default AWIPS color map.

For ProbSevere many of the recommendations are tied to improvements in the product display within AWIPS. Forecasters enjoyed having the separate hazard probabilities and recommend continuing to hone those probabilities to see continued improvement. It is also recommended that a better storm tracking feature be developed, instead of just reflectivity, most importantly for ProbWind because many times the ProbWind contours would not be tracking the area of strongest winds and therefore provide low probabilities. Many forecasters suggested some type of velocity tracking to try and reduce this problem. It is also recommended to “train” the model for different geographic regions which have different environments and severe threats. In terms of the product display, it is recommended that some qualitative wording be added to the product readout for the azimuthal shear, similar to the satellite growth rates, that indicate if the shear values are considered strong or weak. Finally it is also recommended that the ProbTor contour be slightly thicker so that it shows up when overlaid with the ProbSevere contour.

For the GLM, again, the display was the main focus for improvement from the forecasters aside from the obvious data quality issues that were present during the testbed experiment and GLM testing period. It is recommended from this experiment that all of the geolocation errors are fixed and that a native resolution gridded product is provided before fielding the GLM data to the forecasters. It is also recommended to the training group that locally focused training is developed that highlight the use of GLM in the region of interest. It is also recommended that more training is needed to provide the forecasters with the physical meaning of "events", "groups", and "flashes" and when they can be applied to certain forecast situations.

For NUCAPS, the main recommendation is still tied to the latency of the product into AWIPS. It is recommended to cut the latency down to one hour or less from product time to when it is displayed in AWIPS. It is also suggested that a mouse-over readout be developed to get a readout of various convective parameters when sampling the sounding points. Also, along the same lines with the display, it is recommended that a "station ID" be assigned to the points to more easily keep track of which soundings have been examined and which have not. It is also recommended to incorporate some type of wind information into the soundings to get a view of the vertical wind shear when looking at the sounding. Further suggestions are that forecasters would like to see more "relevant" parameters from the gridded NUCAPS (e.g. SBCAPE, MLCAPE, and SBCINH) instead of parameters tied to a certain pressure level. Finally it is also recommended that work continue on the automated boundary layer correction of NUCAPS to hone it to provide a more realistic boundary layer in all cases instead of creating solely a well-mixed boundary layer.

More detailed feedback and case examples from the HWT 2017 GOES-R/JPSS Summer Experiment can be found on the GOES-R Proving Ground HWT blog.

Archived weekly "Tales from the Testbed" webinars.

References

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. *Wea. Forecasting*, **29**, 639–653, doi:10.1175/WAF-D-13-00113.1.

Goodman, S. J., and Coauthors, 2012: The GOES-R Proving Ground: Accelerating user readiness for the next-generation geostationary environmental satellite system. *Bull. Amer. Meteor*. Soc., **93**, 1029–1040, doi:10.1175/BAMS-D-11-00175.1.

Goodman, S. J., R. J. Blakeslee, W. J. Koshak, D. Mach, J. Bailey, D. Buechler, L. Carey, C. Schultz, M. Bateman, E. McCaul Jr., and G. Stano, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmos. Res.*, **125–126**, 34-49. doi:10.1016/j.atmosres.2013.01.006

Setvák, M., K. Bedka, D. T. Lindsey, A. Sokol, Z. Charvát, J. Šťástka, and P. K. Wang, 2013: A-Train observations of deep convective storm tops. *Atmos. Res.*, **123,** 229–248, doi:10.1016/j.atmosres.2012.06.020.