GOES-R and JPSS Proving Ground Demonstration at the Hazardous Weather Testbed 2019 Spring Experiment Final Evaluation

Project Title: GOES-R and JPSS Proving Ground Demonstration at the 2019 Spring 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), University of Oklahoma (OU), Cooperative Institute for Mesoscale Meteorological Studies (CIMMS)

Duration of Evaluation: 22 April 2019 – 7 June 2019

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

Submitted Date:

Table of Contents

1. Executive Summary.. 2
2. Introduction.. 3
3. Products Evaluated.. 5
3.1 Advanced Baseline Imager (ABI) Imagery, Baseline Derived Products RGB Composites, and Channel Differences.. 5
3.2 Probability of Severe (ProbSevere) Model.. 13
3.3 Geostationary Lightning Mapper (GLM) Lightning Detection................. 21
3.4 NOAA Unique Combined Atmospheric Processing System (NUCAPS) Temperature and Moisture Profiles.. 25
3.5 All-Sky LAP Stability Indices, Total Precipitable Water, and Layered Precipitable Water Products... 34
3.6 Advanced Blended Total Precipitable Water Product.............................. 39
4. Summary and Conclusions...42
5. References...44

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 2019 Spring Experiment, which took place at the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) in Norman, OK from 22 April to 7 June 2019. The Satellite Proving Ground activities were focused in the Experimental Warning Program (EWP). A total of 30 National Weather Service (NWS) forecasters representing five NWS regions, 2 Department of Defense (DOD) United States Air Force forecasters, and an additional three broadcast meteorologists participated in the EWP experiment. They evaluated six major baseline, future capability, and experimental GOES-R and JPSS products (Table 1) 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). Additionally, this year in the EWP the Satellite Proving Ground Experiment was combined with an experiment by the Radar Operations Center (ROC) and NSSL evaluating several new updates and experimental products utilizing radar data. This allowed both experiments to provide more forecasters and more weeks to evaluate their products and provide robust feedback on the development of these experimental products for use in NWS operations.

Some of the products demonstrated in 2019 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 2019 EWP Spring 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, the Probability of Severe statistical model (ProbSevere), the All-Sky Layer Atmospheric Profile (LAP) suites of Precipitable Water (PW) products and stability indices, and the Advanced Blended TPW product. 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 NOAA-20 satellite 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 (only applied to the NOAA-20 afternoon CONUS passes). Also, participants were able to view the NUCAPS derived parameters in a gridded plan or cross-section view, as well as evaluate a new NUCAPS forecast product. 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), The University of Alabama in Huntsville (UAH), Science and Technology Corporation (STC), Vaisala, NOAA, NESDIS, 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 Satellite

Proving Ground efforts in the HWT with support from 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 (http://hwt.nssl.noaa.gov/ewp/) . 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 were especially interested in evaluating GLM data and getting an early view of the data before others in operational offices.

2. Introduction

GOES-R Proving Ground (Goodman et al. 2012) demonstrations in the HWT have provided users with a glimpse into the capabilities, products and algorithms that are and will be available with the GOES-R satellite series, beginning with GOES-16 which launched in November 2016. The education and training received by participants in the HWT fosters interest and excitement for new satellite data and helps to promote readiness for the use of GOES-R data and products. Additional demonstration of JPSS products introduces and familiarizes users with advanced satellite data that are already available. The HWT provides a unique opportunity to enhance research-to-operations and operations-to-research (R2O2R) by enabling product developers to interact directly with operational forecasters, and to observe the satellite-based algorithms being used alongside standard observational and forecast products in a simulated operational forecast and warning environment. This interaction helps the developer to understand how forecasters use the product, and what improvements might increase the product utility in 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 2019, the GOES-R/JPSS Proving Ground activities were conducted during the weeks of April 22, April 29, May 6, May 13, May 20, and June 3 with six participants each week ranging from NWS Forecasters to forecasters from the U.S. Air Force to broadcast meteorologists from around the country. 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, three broadcast meteorologists sponsored by the GOES-R Program participated in the Spring Experiment, working alongside the other forecasters. Training modules in the form of Articulate Power Point presentations for each demonstrated product 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 survey and preparation and delivery of the “Tales from the Testbed” webinar for weeks 2 - 5.

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 Monday-Thursday 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 a blog (http://hwt.nssl.noaa.gov/ewp/). 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 mature, 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 weeks 2 - 5, forecasters participated in the “Tales from the Testbed” webinar, in association with the Warning Decision Training Division (WDTD). These presentations gave participants an opportunity to share their experience in the HWT with an average of 20 – 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 that particular week’s weather. Time was also 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 300 GOES-R and JPSS related blog posts were written during the lifetime of the Spring Experiment by forecasters, product developers, and the Satellite Liaison. At the end of each shift (Monday-Thursday), participants filled out a survey 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. Feedback from the GOES-R and JPSS demonstrations during the 2019 Spring Experiment is summarized in this report.

3. Products Evaluated

Table 1. List of GOES-R and JPSS products demonstrated within the HWT/EWP 2019 Spring Experiment

Demonstrated Product	Category
Advanced Baseline Imager (ABI) imagery, baseline derived products, RGBs, and channel differences	GOES-R Baseline & National Weather Service
ProbSevere version 2	GOES-R Risk Reduction
GLM Total Lightning Detection	GOES-R Baseline
NUCAPS Temperature and Moisture Profiles and associated products	JPSS Baseline
All-Sky LAP Stability Indices, Total Precipitable Water, and Layered Precipitable Water Products	GOES-R Risk Reduction
Advanced Blended TPW	GOES-R Risk Reduction
Category Definitions: **GOES-R Baseline Products** – GOES-R Level 1 Requirement products that are funded for operational implementation **GOES-R Risk Reduction** – New or enhanced GOES-R applications that explore possibilities for improving Algorithm Working Group (AWG) products. These products may use the individual GOES-R sensors alone, or combine with data from other in-situ and satellite observing systems or NWP models with GOES-R **National Weather Service** – Products created within AWIPS-II **JPSS Baseline** – Products funded through the JPSS program	

3.1 Advanced Baseline Imager (ABI) Imagery, Baseline Derived Products, RGB Composites, and Channel Differences

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

This was the third year that real time GOES-16 ABI imagery and associated derived products were evaluated in the HWT and the first year for GOES-17 imagery and associated products to be evaluated. The primary focus of evaluation for this year's experiment was on the derived products, particularly those associated with convection and severe weather forecasting. Since the imagery from GOES-16 has been declared operational, that imagery was used heavily, but there was not a lot of emphasis on evaluation, most of the emphasis in evaluation of the satellite imagery and products was for GOES-17 when it was advantageous to use the west satellite. The emphasis for the imagery primarily focused on the use and importance of the one-minute mesoscale sectors versus five-minute data in the forecasters' decision making capabilities. Feedback from this experiment on performance and display of the derived products and usefulness of the one-minute imagery is presented in this section. The RGB composite imagery and Channel differences from both satellites were also used and received feedback from the forecasters throughout the experiment.

Use of ABI Imagery and Derived Products in the HWT

GOES-R series data has been in the testbed for evaluation for three years. GOES-16 has been operational for approximately two years now in its east position and GOES-17 has been in the west position for a little under a year as the operational west satellite. The data from both satellites have been used widely in operations over the last two years. Therefore, the primary focus of the evaluation of ABI data was mainly tied to the importance of one-minute imagery and how that aided forecasters severe storm forecasting and warning decision making. This was the first year for GOES-17 data to be evaluated in the testbed, so another focus of the experiment was to find the usefulness of that data, particularly across the western CONUS. This year's experiment provided many different modes of severe storms and different forcing mechanisms that allowed forecasters throughout the weeks to examine the usefulness of imagery and products across a variety of forecast and warning situations around the country. At the end of each day, the forecasters were asked if the one-minute imagery provided significant information during warning operations not captured in the five minute data. Nearly 75% of the forecasters responded that the one minute data did provide more information and was a crucial data set during the daily operations. Overall, at the end of each week, the forecasters were asked to rate the impact that the one-minute imagery had on their operations, and 32 of the 35 total forecaster responded that it had a very positive or extremely positive impact. Some responses are shown below.

> "I was able to see how quickly storms will develop as opposed to the 5 minute imagery. It did help in my decision making during the warning process."
> *Forecaster, End-of-Day Survey*

> "The visible imagery was very helpful in identifying the change of characterization of clouds leading to convective initiation. The VIS/IR sandwich showed when the storms reached heights where glaciation began and was a precursor for lightning activity, as shown by the GLM"
> *Forecaster, End-of-Day Survey*

> "Getting data about the strength of updrafts in actual to the minute real-time will help improve a lot of our severe weather statistics...even more so than Dual Pol's introduction. On a marginal day, it will certainly help to pick out which storms have the best chance to become severe and which storms need more interrogation to issue a warning or SPS."
> *Forecaster, End-of-Week Survey*

> "Overall very helpful for marginal, pulse convection as it can help you gain lead time on when stronger cores are developing."
> *Forecaster, End-of-Week Survey*

For a case of a supercell over eastern Colorado, the one-minute imagery was able to show the forecaster some of the more in depth features and characteristics of the storm. The forecaster says "Many things visible here, including uplift and twisting of stratus deck in the inflow region, anvil plume texture, and updraft texture. The parallax error ends up helping the user get more info about the vertical structure and composition of the storm" (Figure 1). In this case and many other cases, the forecaster was able to determine some storm structure and low level features associated with stability when monitoring the one-minute imagery for convective initiation, storm maturity, or storm decay.

Figure 1: 2307 UTC 01 May 2019 GOES-16 0.64um “Red” Visible Imagery over Eastern Colorado.

Forecaster also evaluated some of the baseline derived products from both GOES-16 and GOES-17 in the HWT experiment this year. There was more emphasis put on evaluating some of the convection related products and RGBs this year compared to the imagery since the imagery has been in heavy use operationally for the last two years. The forecasters focused on products that were related to convective forecasting and warning operations. Therefore, not every derived product was evaluated during the experiment. The derived products in their current form continue to be sparingly used during convective operations. Some of the main products that got the most use during the experiment were the derived stability indices plus TPW, the cloud top products, and the derived motion winds. Each day during the surveys, forecasters were asked what products they used during their operations and what they were used for, and at the end of each week they were asked to explain which products would have the most impact on improving warning operations. Overwhelmingly, the products that got the most use were the derived stability indices and TPW. Though most forecasters wrote that they would much rather see the All-Sky version of these products in operations and feel that they would provide a lot more use than the current version. The All-Sky products will be talked about later in the report. At times these products were unusable because of the data gaps due to cloud cover in the area of interest, and provided no value. These products, most notably the CAPE, tended to underdo the values quantitatively when compared to RAP model derived values, 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 at times was the derived-motion winds. This product was used primarily in a forecast sense and looking for jet maxes and areas of deep-layer shear that would subsequently affect the storm initiation and growth. Some other products were also examined, such as the Cloud Top Temperature, Phase, and Pressure products. These products were used after storms had initiated while forecasters were monitoring for new updraft growth or decay within the storm. Feedback on these derived products is presented in this section.

> "CAPE and cold cloud tops helped me understand the environmental conditions the best."
> *Forecaster, End-of-Week Survey*

> "Yes. TPW, CAPE, and Cloud Top products were all utilized. Even tried out a few of the convective RGBs which proved useful in today's setup identifying outflow and other storm-scale boundaries."
> *Forecaster, End-of-Day Survey*

> "I initially used TPW and cloud top products to assess atmospheric moisture content and cloud glaciation of developing convection. Once convection had fully initiated and I went into warning mode, radar products proved most useful."
> *Forecaster, End-of-Day Survey*

> "Prefer to use the All Sky products to the baseline derived products."
> *Forecaster, End-of-Day Survey*

> "I used CAPE, LCL heights, and TPW. They all were helpful when I was nowcasting to see if storms were likely to intensify."
> *Forecaster, End-of-Day Survey*

Limitations and Recommendations for Derived Products

Some recommendations for improvement and use of these derived products in routine operations is given here. The most common suggestion from forecasters was that it is much more preferred to have the data gaps filled in on the stability indices and the TPW products like is done in the All-Sky version. There were numerous times throughout the experiment where these indices proved useless because they did not cover the area of interest due to cloud cover, even prior to convective initiation and primetime warning operations. This made it impossible to utilize these products to analyze the pre convective environment in many cases. Also, we recommend exploring improving the values of the stability indices, particularly CAPE, to get values more representative of other observations and data sets. CAPE values are consistently underdone, though the gradients are consistent. Values more in line with other data sets used by forecasters can provide greater trust in using the products in the future. One suggestion is using a different, higher resolution model as the first guess that could provide more accurate values than the GFS.

> "Did not use the baseline derived products today. I tended to use the All-Sky products instead."
> *Forecaster, End-of-Day Survey*

> “No, there was early morning cloud cover across ND, so the baseline derived products were not available in this area.”
> *Forecaster, End-of-Day Survey*

> “I don't trust derived products much due to the prior assumptions that must be made (Cloud Mask, etc.). I'm trying to use the base channels or RGB / Difference products exclusively”
> *Forecaster, End-of-Day Survey*

> “No, not as useful given the data holes”
> *Forecaster, End-of-Day Survey*

> “CAPE - did show area - but seemed to be low compared to SPC and RAP forecasts.”
> *Forecaster, End-of-Day Survey*

> “I looked the TPW and stability indices in the mesoanalysis phase. This was helpful, but the stability indices were relatively underdone.”
> *Forecaster, End-of-Day Survey*

Use of RGB Composites and Channel Differences in the HWT

The last of the ABI related products to be discussed that were evaluated during the experiment were the multiple RGB Composites and channel differences. These products are created on the fly within AWIPS-II and have also been in operations for two years now. The RGB composites have been slowly getting more use within local office operations as forecasters become more comfortable analyzing the imagery and as more training has been provided throughout the NWS. A number of the different convective and forecast related RGBs were used throughout the week. A large number of forecasters focused on three main RGBs throughout the experiment. Those were the Day Land Cloud Convection RGB (Fig. 3), the Day Cloud Phase Distinction RGB, and the VIS/IR Sandwich RGB. Nearly every forecaster that came through the experiment used these every day as part of their routine forecast and warning operations. These RGBs helped forecaster hone in on areas of developing convection by identifying clouds that were becoming glaciated and were growing deeper from those that were not growing and remained mostly water clouds. This allowed forecasters to identify areas of growing deep convection and made them aware of where to be focused for possible severe weather in the near future. The VIS/IR Sandwich RGB also was crucial in identifying cloud top features of mature storms such as overshooting tops, and above anvil cirrus plumes (AACP) which helped identify storms where updrafts were strongest and were maintaining their strength. A few other RGBs were used pretty extensively throughout the weeks, like the Simple Water Vapor RGB (Fig. 2), the Differential Water Vapor RGB, the Day Convection RGB, and the Air Mass RGB. These RGBs were primarily used for large scale analysis of synoptic patterns that could play a role in developing convection.

> “I mostly used the Day Cloud Convection RGB and the Day Convection RGB to monitor convective trends. I also used the Differential Water Vapor RGB to assess the synoptic environment.”
> *Forecaster, End-of-Day Survey*

> "Yes. I used Day land cloud convection, day convection, and day cloud phase distinction. This RBG along with radar gave me a wealth of information regarding the overall trend of storms. I used the day land cloud convection to look at storms downstream in order to see if they would intensify in southern Iowa."
> *Forecaster, End-of-Day Survey*

> "I used the ECONUS Sandwich to look at developing convection. The color scale made this imagery useful to identify updraft areas and really made the thunderstorm tops very evident."
> *Forecaster, End-of-Day Survey*

> "I used the Simple Water vapor and Vis/IR sandwich RGBs today. The Vis/IR sandwich helped to determine the depth of the convection, and the Simple Water Vapor was very useful in tracking the multi-vortex upper low that was moving into the CWA and was a potential trigger for convection. It also highlighted the return low level moisture that was partly responsible for the afternoon convection that developed."
> *Forecaster, End-of-Day Survey*

> "I looked at the day cloud convection and day cloud phase products. They were helpful in identifying glaciation and eventually electrification. This would be helpful for DSS events."
> *Forecaster, End-of-Day Survey*

In the case shown below from 06 June 2019 in the western United States, forecaster utilized several different RGBs to make their analysis throughout the day. One forecaster utilized the simple Water Vapor RGB along with RAOB observations to distinguish an upper level jet maximum that was moving into their area of interest and could play an important role in afternoon storm development (Fig. 2). Pairing this RGB with other model and analysis data sets gave him confidence that a shortwave was moving into the area with an area of upper level divergence moving into southern Idaho and western Wyoming that would provide the necessary lift for storm development during the afternoon hours. Later in the day, another forecaster utilized the Day Land Cloud Convection RGB to distinguish areas of growing, deepening cumulus clouds from upper level cirrus clouds (Fig. 3). "Looking downstream we're noticing a line of developing convection using the Day land cloud RGB. It's helpful to analyze just how high the tops of the storms are getting combined with cloud tops. Zooming in closer I can examine the few storms that have popped up more closely. The RBG helped differentiate between the ambient cirrus clouds from the growing storms. That might not have been as easy to see using a typical visible view." Here the RGBs enhanced the forecaster's ability to distinguish between different types of clouds and made it easier to analyze storms that were of interest over just using the visible imagery alone.

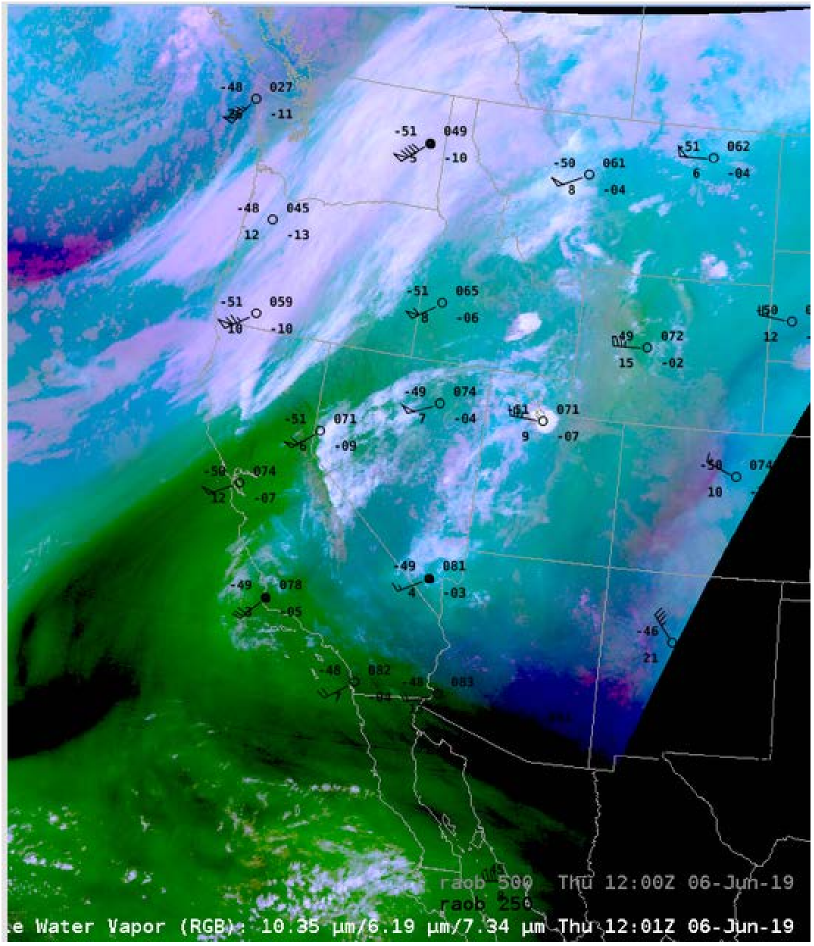

Figure 2: 1201 UTC 6 June 2019 Simple Water Vapor RGB with 250mb RAOB observations depicting a jet across the Pacific Northwest United States.

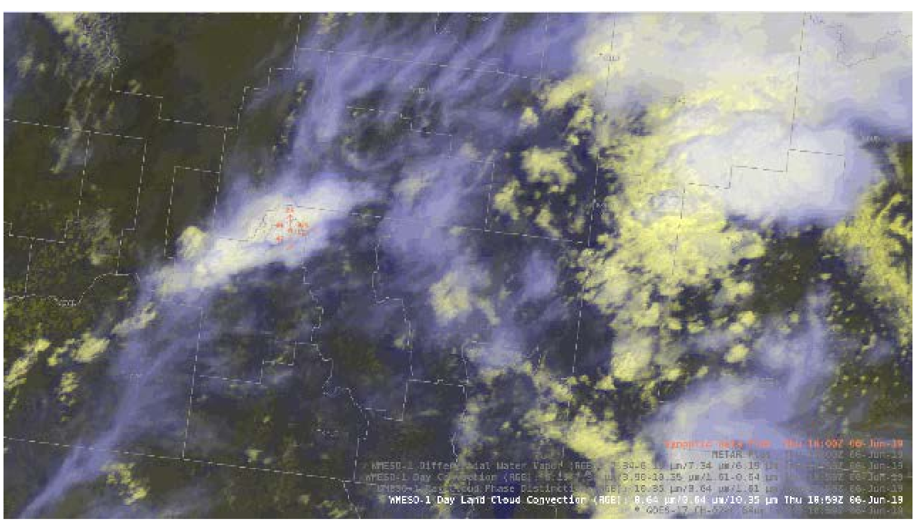

Figure 3: 1859 UTC 6 June 2019 Day Land Cloud Convection RGB depicting convective initiation occurring across the Pocatello, Idaho County Warning Area.

The ABI channel differences were also evaluated during the experiment this year. The main channel difference used by forecasters in convective operations was the split window difference, mainly for trying to identify areas of low level moisture pooling along boundaries in the absence of clouds. This channel difference largely proved to not be too useful throughout the experiment, and many forecasters were turned off to using it because the color table was very hard to look at and to interpret. Some forecasters did change to their own color table to try to use the product. The other channel differences went largely unused during the experiment as there were just too many products to evaluate to spend time looking at some of those products that weren't related to convection as closely.

> "Used split window difference to examine dryline, but color table made it very hard to look at for any length of time."
> *Forecaster, End-of-Day Survey*

Limitations and Recommendations for Improvement of RGBs and Channel Differences

The major recommendation is for the improvement of the color table for the Split Window Difference Product. Many forecasters want to use that product, but don't waste their time with it because of the color table. Another major recommendation is for the further training on the RGB suite of products, particularly more applications based training showing what the products can be used for different times of the year and for different types of forecast problems. The quick guides are great, but more can be done to enhance forecaster utilization of that imagery. Otherwise, feedback on all of the ABI products was overwhelmingly positive and has been a game changer in the forecaster utilize satellite data now in warning operations.

> "Want to leverage some of the split window products more like 12.3 - 10.3 diff product to better track low-level moisture discontinuities in the pre-convective environment, but color curves are not very good at all. Also some of the other WV difference products/RGBs are hard to use due to the color enhancements. I think these products would be used more operationally IF work could be done to enhance color enhancements. Bottom line is I would first and foremost like to see additional RGBs, difference products that would help forecasters better assess boundary layer changes and gradients. Would be especially helpful in more marginal pre-convective environments."
> *Forecaster, End-of-Week Survey*

> "Training on the best use of ABI imagery is important. There is so much more information in the GOES-R era, which makes it difficult to figure out which products are best to use, and for what particular scenario."
> *Forecaster, End-of-Week Survey*

> "I think some additional (reinforcing) training would be useful on how these different ABI products can be used."
> *Forecaster, End-of-Week Survey*

"All ABI products continue to offer great value in the warning and situational awareness process"
Forecaster, End-of-Week Survey

3.2 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 the coming year, was evaluated in the HWT for the fifth consecutive year, with updates made since last year's experiment. ProbSevere is currently undergoing tuning and assessment with the in-orbit 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 also 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 input values. This year, an outer contour specifically for ProbTor was added to the ProbSevere contour that could be set to appear whenever the ProbTor value reached a threshold set by the user. There was also a separate product the forecaster could load up to display a contour for each separate hazard if they so choose to. 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 tend to overlay ProbSevere guidance on either base radar data, or MRMS products (e.g., Composite Reflectivity, MESH, and Reflectivity at Lowest Altitude (RALA)) when storms begin to initiate. Early on in the lifecycle of the storms, ProbSevere alerted forecasters to the first storms of the day that were becoming potentially severe and which ones warranted closer monitoring and analysis. The consistent view among all of the forecasters was that ProbSevere continues to be very useful as a situational awareness tool, alerting forecasters of storms requiring further radar interrogation. This was found to be especially useful in busy warning environments, where there were many ongoing storms, to quickly rank storms in importance to interrogate based on storms with higher probabilities of severe. Forecaster also often loaded up a four panel layout on contours for the full ProbSevere along with the individual contours for each hazard type (ProbWind, ProbHail, ProbTor) to examine which hazard was the dominate threat for any given storm. ProbSevere also assisted in situations where warning issuance was more marginal or uncertain based on the environment or base radar data. ProbSevere would sometimes provide more confidence to issue or not issue the warning based on the probabilities and parameters available in ProbSevere. It is important to note that forecasters did not use ProbSevere alone to issue warnings, but instead used it as another valuable piece of information to increase confidence in the decision making of the warning forecaster. Both the uses and some limitations to the ProbSevere algorithm and display will be discussed in this section.

> "ProbSevere serves as a good triage tool and can help in the warning decision making process by being able to monitor trends within a storm."
> *Forecaster, End-of-Day Survey*

> "Similar to previous days, the ramp up in values increased my situational awareness and helps me with warning lead time. Coupling knowledge of the environmental conditions/setup with ProbSevere helps me increase warning lead time, especially with initial warnings) or the first warning on a new storm). ProbSevere is an effective tool for monitoring early storm intensification trends."
> *Forecaster, End-of-Day Survey*

> "Used ProbSevere to aid in monitoring strengthening/weakening of the scattered convection. The Time Series product was very beneficial as a Situational Awareness Tool. A particular storm in the CWA, which had previously been quite strong, had weakened and I diverted my attention elsewhere. However, I kept the ProbSevere Time Series displayed on another monitor. Later, I noticed that the Time Series indicated that the storm was potentially re-intensifying (top left image) and it caused me to direct my attention back to the thunderstorm."
> *Forecaster, End-of-Day Survey*

With the addition of some advancements in the algorithm this year, the ProbSevere version 2 "All Hazards" model allows for an outer contour to pop up, around the main ProbSevere contour, showing the ProbTor value for the storm. This contour shows up by default when the ProbTor value reached 3%, but this was able to be adjusted by the user depending on the day's environment or potential for tornadoes (Fig. 4). This allowed the forecaster to see both the All Hazards version but also be drawn to storms that were showing increased potential for tornadoes without having to have the ProbTor contour laded in a separate pane or window. Overall, feedback was positive on this addition and was appreciated by forecasters as a means to limit the screen real estate that loading up the separate hazards can take when multiple hazards are in play on any given forecast day. When asked at the end of the week if the outer contour for ProbTor was helpful, 94% of the forecasters responded that yes the outer contour was helpful. There were some thoughts on how to further improve this contour, as some thought it was too large and some would like for it to be easier to modify the threshold quickly in a warning setting. Otherwise, most all forecasters found it to be rather useful.

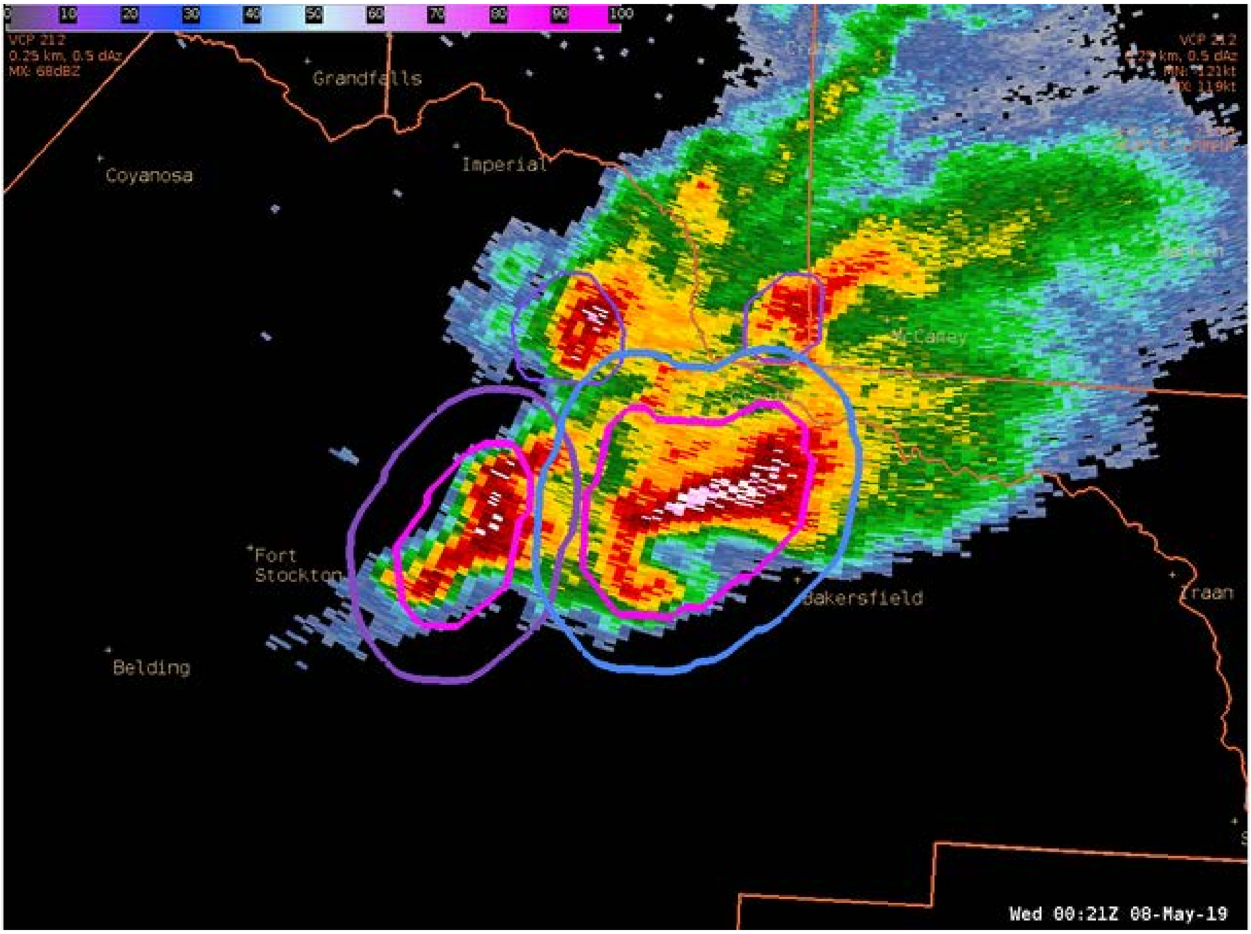

Figure 4: 0021 UTC 08 May 2019 0.5 degree Base Reflectivity and ProbSevere v. 2 with outer ProbTor contour in west Texas.

> "Seeing the ProbTor contour pop up lets you know quickly if you should be monitoring rotation."
> *Forecaster, End-of-Week Survey*

> "Yes, it's helpful to see trends in tor potential separate from wind/hail, but in the same easy to view and interrogate product"
> *Forecaster, End-of-Week Survey*

> "I found the additional contour for ProbTor to be very beneficial. Seeing the additional contour pop-up resulted in me investigating that particular cell with more scrutiny. The additional contour is also great as far as in increasing situational awareness."
> *Forecaster, End-of-Week Survey*

ProbSevere was also found to help enhance forecaster confidence when deciding to issue or not issue severe thunderstorm warnings, and to a lesser extent in tornado warnings. Forecasters were asked at the end of each day to assess the role ProbSevere had in their warning decisions as it related to confidence and lead time. Approximately 92% of forecasters responded that they felt ProbSevere increased their confidence in issuing severe thunderstorm warnings, while only 62%

responded that ProbTor helped increase their confidence in issuing tornado warnings. Similarly, when it came to lead time, generally 73% felt that they gained lead time on issuing severe thunderstorms warnings compared to using base data alone, while 52% felt they gained lead time while issuing tornado warnings. Forecasters noted that the most important value out of ProbSevere was the trend in the probabilities and related parameters. New this year, forecasters were able to pull up the ProbSevere website and obtain time series trend graphs of the probabilities and many of the parameters associated with ProbSevere to make it easy to identify growing or decreasing trends over the last hour or so (Fig. 6). This was a very welcome addition and was generally used by every forecaster and would be greatly anticipated for use within AWIPS-II.

> "There was only one storm where I considered issuing a warning. I held off on the warning and the fact the ProbSevere stayed fairly low gave me more confidence in that decision."
> *Forecaster, End-of-Day Survey*

> "Yes, being able to view the data in the time series, I could tell the storm was still strengthening based on visual and statistical trends, which gave me the confidence to issue sooner."
> *Forecaster, End-of-Day Survey*

> "It was more valuable in increasing confidence in warning decisions than in increasing lead time."
> *Forecaster, End-of-Day Survey*

> "In particular, the time-series plots of the ProbSevere data helped increase my lead time for severe thunderstorm warnings (primarily for hail) early in the event combined with GLM data, and the environmental conditions that were known (including analysis from NUCAPS modified soundings, All Sky CAPE/Lapse Rates)."
> *Forecaster, End-of-Day Survey*

> "More the reverse is true for today. It increased my opinion in NOT issuing. Low ProbTor values solidified my idea not to consider tornadoes a real threat other than including the 'possibility' of them in some warnings."
> *Forecaster, End-of-Day Survey*

The following example shows a case where ProbSevere helped give confidence in issuing a severe thunderstorm warning a little sooner than he would have otherwise without having ProbSevere. The time trend graphical readout information is also shown here to showcase the ease of pulling out information on trends in probabilities and parameters with this interface. The following is taken straight from the blog where the forecaster explains their decision making process for the storm in the image which posed mainly a wind threat (Fig. 5).

> "SPS's have been upgraded to a warning due to some interesting velocity couplets. At the time of the warning, the new ProbSevere model had (ProbHail: 5%, ProbWind: 56%, ProbTor: 2%). At the storm's severity peak, the new ProbSevere model had (ProbHail:

8%, ProbWind: 78%, ProbTor: 17%). While the warning decision was made based on the velocity signature showcasing 30+ kts inbound with some broad rotation, the added confidence of the ProbSevere model assisted in "pulling the trigger" and likely provided several minutes of added lead-time in this scenario. Also taken into account was the understanding that it was the northern-most cell in a line that is moving into a favorable environment characterized by ML CAPE values near/above 2000 J/KG based on the All-Sky LAPS analysis... There was also a considerable lightning jump that occurred in this cell as it underwent strengthening. Tree damage was reported within the warning shortly after it was issued."
Forecaster, 04 June 2019 Blog Post, "LSX Warning 1 (Mountain Bone)"

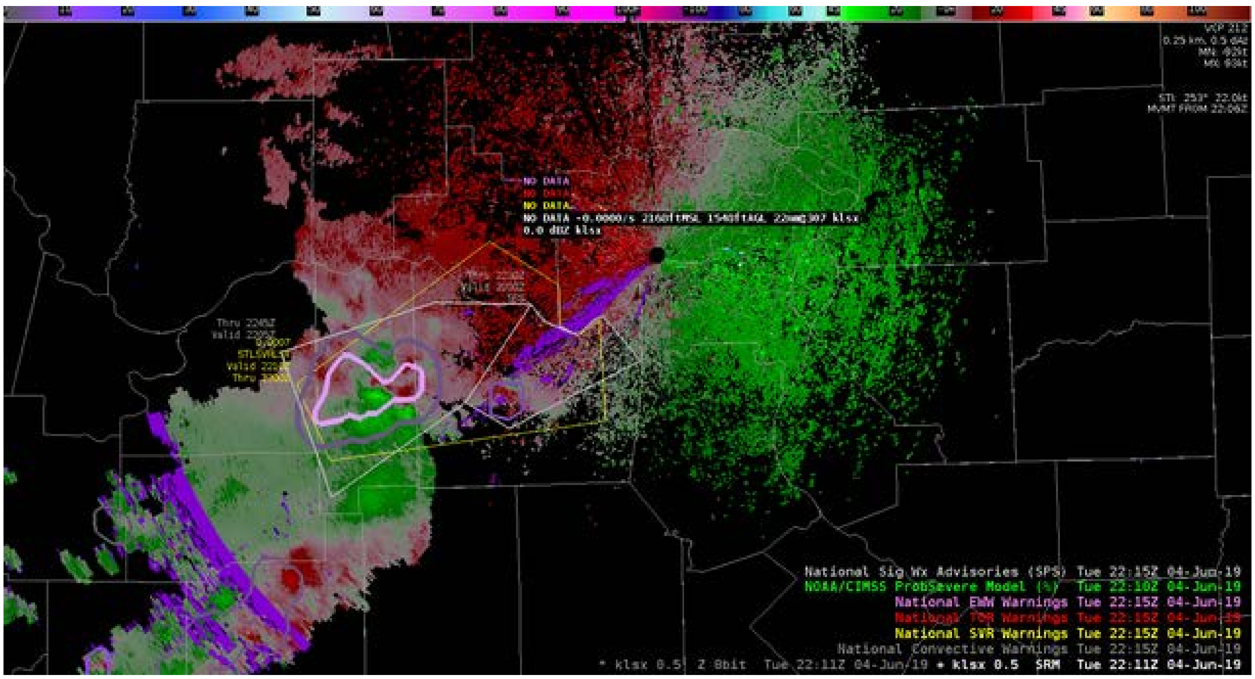

Figure 5: 2211 UTC 4 June 2019 0.5 degree Storm Relative Motion with ProbSevere v. 2 overlaid near St. Louis, Missouri

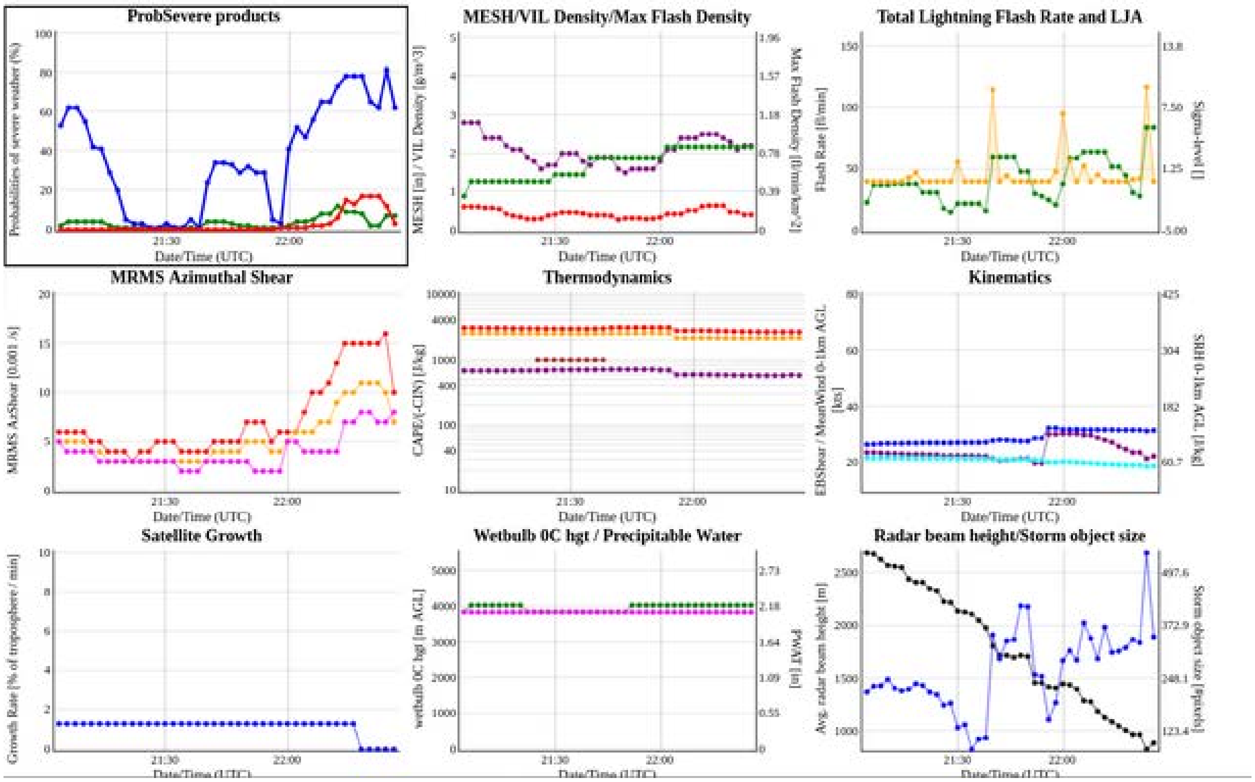

Figure 6: ProbSevere time trends for the storm shown in Figure 5 near St. Louis.

Limitations and Recommendations for Improvement of ProbSevere

There are still several limitations to ProbSevere that prevent it from being as effective in warning operations as it could be. Forecaster throughout the experiment found that latency with the product could be an issue in shaving valuable minutes off of warning issuance. There were several occasions where the increase in probabilities lagged well behind radar characteristics that led to warning issuance before there was a jump in probabilities. This can be attributed mostly to latency in ingesting the MRMS data into the algorithm causing delay in the processing of the ProbSevere probabilities. This issue seemed to make the biggest difference in environments conducive to rapidly developing severe deep convective storms. Also, as in previous years and iterations of ProbSevere, forecasters would like to see better tracking. This can be hard to accomplish without sacrificing the lead time, but it is recommended that steps be attempted to provide better tracking to enhance the use of the product in an operational warning environment. Problems with tracking during the experiment included clumping together of cells in close proximity to each other, but were still separate storms when looking at base radar data, Including lines of storms as one big ProbSevere contour making it difficult to discern which areas of the line might require more attention than other parts that might not be as strong, and not tracking important wind features such as outflow boundaries out ahead of the main reflectivity core which could be where the strongest winds are located for that storm. Another main limitation noticed for ProbSevere from the experiment was for the ProbTor algorithm. There were many wild fluctuations in the ProbTor values at times which seemed to be related to the MRMS values being incorporated into it. There were also some instances of ProbTor showing values in stratiform rain regions of trailing precipitation behind convective lines owing to bad values of azimuthal shear (Fig. 7). It is recommended to filter out these bad values to try to keep them from getting into the algorithm and causing decrease in confidence in the product. Forecasters would also like to see a single radar velocity product that could be incorporated into this for ProbTor to try to get away from some of these issues that inaccuracies in the MRMS data can provide.

> "There were major issues with ProbTor today - nearly every storm had high tornado probabilities due to strong signatures in the MRMS Rotation Tracks/AzShear. None of the storms had tornado potential via velocity data. The ProbTor model output was more of a hindrance."
> *Forecaster, End-of-Day Survey*

> "I think it's useful for relatively discrete convection, but with mixed modes and messy convection, AzShear is still too noisy for me to really rely on it. I'll use it more as a triage/SA tool, but still rely primarily on single-site base data for tornado warnings."
> *Forecaster, End-of-Day Survey*

> "Having a baseline product for forecasters to utilize that incorporates velocity is of extreme importance. I discount much of the ProbWind and ProbTor high-end readouts for this reason alone."
> *Forecaster, End-of-Week Survey*

> "While I didn't issue any warnings today we could have issued some Marine Warnings. ProbSevere helped in these situations. One item of note is that there were several times where the MCV/QLCS was one element in ProbSevere grouping the entire line of reflectivity together. While percentages remained high this didn't really help zero in on the strongest parts of the storm."
> *Forecaster, End-of-Day Survey*

> "Watching multicellular storms across northwest Oklahoma, ProbSevere had issues with how to identify storms. Initial cells were detected fairly well but as storms went upscale, mergers and dissipations resulted in tracking objects getting lost, expanded to include multiple cores, etc."
> *Forecaster, End-of-Day Survey*

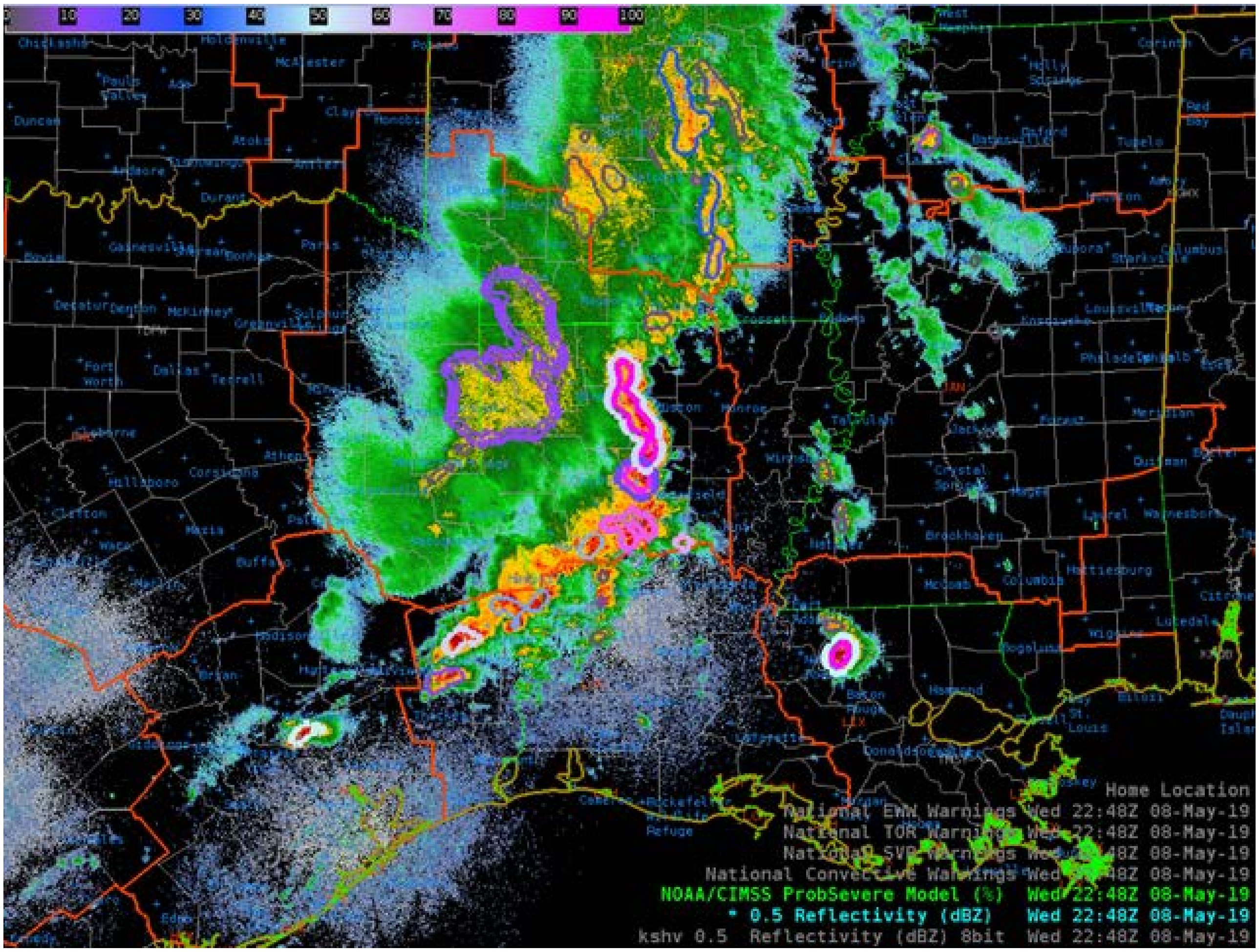

Figure 7: 2248 UTC 08 May 2019 ProbSevere Model overlaid on 0.5 degree reflectivity shows ProbTor values of 12% in the stratiform rain region behind the convective line over northwestern Louisiana.

Suggestions for Product Display

In general, forecasters were liked the ProbSevere contour display and the addition of the ProbTor outer contour, but there were some suggestions offered for improvement of the current display listed below.

- Provide a GUI or easier method to change the threshold for the ProbTor outer contour to be able to make a quick easy change without having to go into the localization perspective.
- Provide the time trends graphs within AWIPS-II to easily pull up when interrogating a storm.
- Have a separate color table for ProbTor that “pops” a little more at the lower end of the scale to bring out some of those changes in more marginal environments when storms will most likely not have high ProbTor values.

- Have all three hazard contours in the display like the outer ProbTor contour and have the thresholds at which they appear be user adjustable similar to the ProbTor now so that all the hazards can stand out apart from each other on one display.
- Focus the ProbTor contour on a smaller area more around the actual area of the storm than the entire reflectivity core, particularly in QLCS line events.

"I liked having the ProbTor contour in addition to probe severe. I think the best case would have been to have all 3 as separate contours where the user could modify when the contours started to appear (i.e. only show contours above 30% for wind, 30% for hail, and 20% for tor)."
Forecaster, End-of-Week Survey

"The ProbTor contour was much too large. It would be helpful to narrow that down to a smaller area of concern, particularly with QLCS type events where you might have a couple of areas of concern along a line. These were often just grouped into one large object."
Forecaster, End-of-Week Survey

"It was helpful...as we discussed, having an easier way to change the thresholds would be very nice given that on some days the environment is extremely conducive so having a low threshold is useless. Whereas on some days, the only storms that get a tor outer contour (even with a low value) were often those storms that required closer investigation. I know we can change the threshold in the localization perspective, but if it could be easier, it would greatly enhance its use."
Forecaster, End-of-Week Survey

3.3 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)

For the 2019 experiment, a variety of updated GLM products were created based on feedback from the initial review in the HWT and Operations Proving Ground in 2018. Initial products included (all at 1-min or 5-min with 1-min updates): Flash Extent Density (FED), Event Density, Average and Minimum Flash Size, Average Group Size, Total Optical Energy, Flash Centroid Density and Group Centroid Density. Prior to arrival at the HWT, some forecasters had access to the GLM Flash Extent Density at their local offices, a small number of other forecasters had access to the Average Flash Area and Flash Energy gridded products as well. However, most forecasters did not use the GLM products routinely in operations prior to the HWT experiment. All forecasters had previously taken the GOES GLM training, though some noted it was more than a year prior.

Overall, forecasters had a better understanding of the products in 2019 than in 2018. Average confidence in the products was much higher than in 2018 with most forecasters ranking most products high or very high in terms of understanding. This was likely due to a longer period of exposure and limited access at some NWS offices. The top two products (ranked from both in surveys in and from discussion) in terms of usefulness during severe weather operations with forecasters were the Flash Extent Density (FED) and Minimum Flash Area products. Event

Density and Flash Energy were also considered useful by the forecasters. Many forecasters also commented that having a subject matter expert available to answer questions made an impact on usefulness of the data. However, training was a consistent topic of weekly discussion with multiple suggestions for implementation discussed. First, forecasters believed strongly that locally-relevant training be provided, not a single example of a supercell storm in the plains. The questions and thoughts consistently brought up during live events and open-ended survey questions should also be considered within future product and training development. These main points are highlighted below:

Use of GLM in the HWT

As part of best practices for display of the GLM data, subject-matter experts encouraged the forecasters to overlay the ground-based systems (including both IC lightning from ENTLN and CG data from NLDN) over the GLM data. This was suggested to provide a holistic view of lightning activity - the spatial extent from GLM and IC / CG ratio and locations from the ground-based networks. However, this suggestion often led to discussion on why the products often showed different values and trends. Explanations on why the GLM and ground-based systems had differences often led to further discussion and forecasters finding evidence to support or provide caveats for the previously stated reasons. Throughout the experiment subject-matter experts provided context from ongoing research to better understand why detection efficiency of the GLM appeared lower in some storms than others. This included a discussion of optical depth (i.e., stronger storms with more and/or larger ice hydrometeors may make it difficult for the optical emission from many lightning flashes to reach cloud top. Additionally, flash size and rate were discussed as possible limitations on the GLM data, as GLM will not likely detect as many of the small flashes in and around the updraft since they could be grouped together due to the larger spatial resolution of GLM and may not emit much light. Since the GLM is new instrument, ongoing research will likely impact operational use of the products. We strongly suggest that training efforts actively involve subject matter experts for integration of the latest research results.

> "Earlier discussions with lightning detection experts suggested the low GLM FED count may be due to the location of lightning within the storm updraft region, which could impact how well GLM can sense it. That is difficult for the typical operational meteorologist to consider in real-time since it goes well beyond current training, and leads to decreased forecast confidence in the lightning data."
> 23 May 2019, Blog Post: *Lightning Jump in GLM FED but not Earth Networks*
> https://blog.nssl.noaa.gov/ewp/2019/05/22/lightning-jump-in-glm-fed-but-not-earth-networks-data/

Due to some of the inherent caveats with the GLM lightning data, forecasters frequently found it the most useful in situations where issues of optical depth and light blockage within intense thunderstorms were not a consideration. In particular, forecasters found the data useful in monitoring trends for marginally severe storms for signs of intensification or dissipation. Additionally, due to the coverage of the full spatial extent of lightning, forecasters also commonly highlighted the potential use for decision-support services.

"IDSS standpoint, the Minimum Flash Area and FED proved that it's necessary to look at both GLM products and ground based lightning products to see the "total" picture. The GLM products captured a larger flash that extended out into the stratiform area behind the main line that is not seen in the ENTLN and NLDN products. This information can be especially important for Airport Weather Warnings and/or outdoor venues."
16 May 2019, Blog Post: 'IDSS usage from GLM minimum flash area
https://blog.nssl.noaa.gov/ewp/2019/05/16/idss-usage-from-glm-minimum-flash-area/

When multiple observation platforms trend in the same manner, forecasters noted enhanced confidence in warning decisions. The gridded GLM products allowed forecasters to efficiently match the GLM data with satellite, lightning and radar trends. In cases where the total flash rate trends, as noted from FED, matched trends from either base radar or ProbSevere, forecasters commonly noted increased confidence and easier warning decisions.

"The GLM products increased confidence in a decision to issue two SVRs"
22 May 2019, Blog Post: *"GLM predicting strengthening updrafts"*
https://blog.nssl.noaa.gov/ewp/2019/05/22/glm-predicting-strengthening-updrafts/

During the 2018 experiment, forecasters consistently commented they liked the idea behind the flash area product, but the average aspect seemed to dampen the values. Since trends in the smallest flashes are most directly related to intense updrafts, active charging of hydrometeors, and regions of high turbulence, it was suggested to examine a minimum flash area product instead of average. Throughout the 2019 experiment, forecasters consistently used the minimum flash area product while the average flash area product was rarely used after the first day. The minimum flash area product was rated as the second-most useful product of the entire GLM product suite, behind only FED. For this reason, we suggest the Minimum Flash Area product be considered as one of the primary (or base) GLM gridded products for operations.

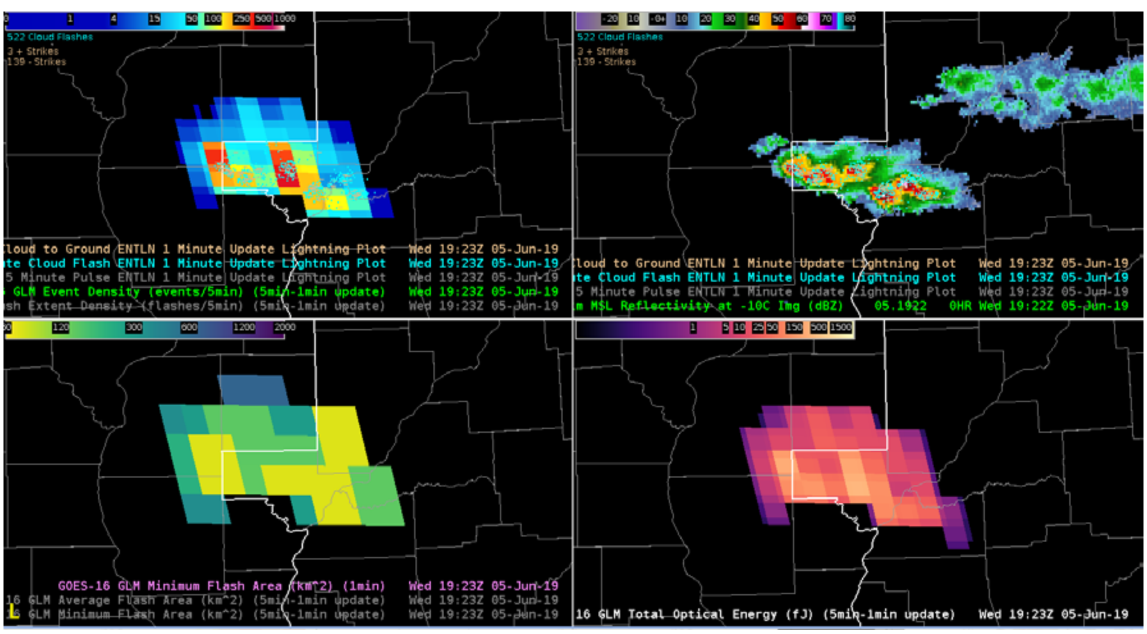

Figure 8: GLM Event Density w/ ENTLN Lightning Data (Top Left); MRMS -10C with ENTLN overlayed (Top Right); GLM Minimum Flash Area (Bottom Left); GLM Total Optical Energy (Bottom Right).

> "Had unique opportunity to watch a line of strong to eventually severe storms ignite and strengthen rapidly along an old outflow boundary this afternoon using GLM data. It was interesting using the above 4 panel display (Fig. X) to not only witness the ignition, but also the strengthening of each sequential cell along the line. By utilizing the GLM Minimum Flash Area (Bottom left) 1 minute imagery in a loop, we were able to sample the early-onset updraft core strengthening of each sequential cell along said line, and watch the event density jump up in accordance."
> 5 June 2019, Blog Post: *"ILX Case"*
> https://blog.nssl.noaa.gov/ewp/2019/06/05/ilx-glm-case/

A GLM-IR RGB product was created in collaboration with Drs. Christopher Schultz (NASA) and Eric Bruning (TTU) to help integrate the IR imagery with the GLM products due to 2018 HWT comments regarding the lack of screen space for integrating the GLM products. The goal of this product is to highlight the areas with the coldest cloud tops (higher blue colors using the IR temperatures from the 10.3 um band), highest FED (increased red values), and smallest flashes (increased green values from smaller MFA). While this was the initial product development, forecasters showed interest and further development will continue. Within this RGB, the strongest, mature updrafts will trend towards white. Newer convection, within warmer cloud tops, will trend towards red or yellow. Anvil flashes, with low flash rates and large size, but still within cold cloud tops will trend towards purple.

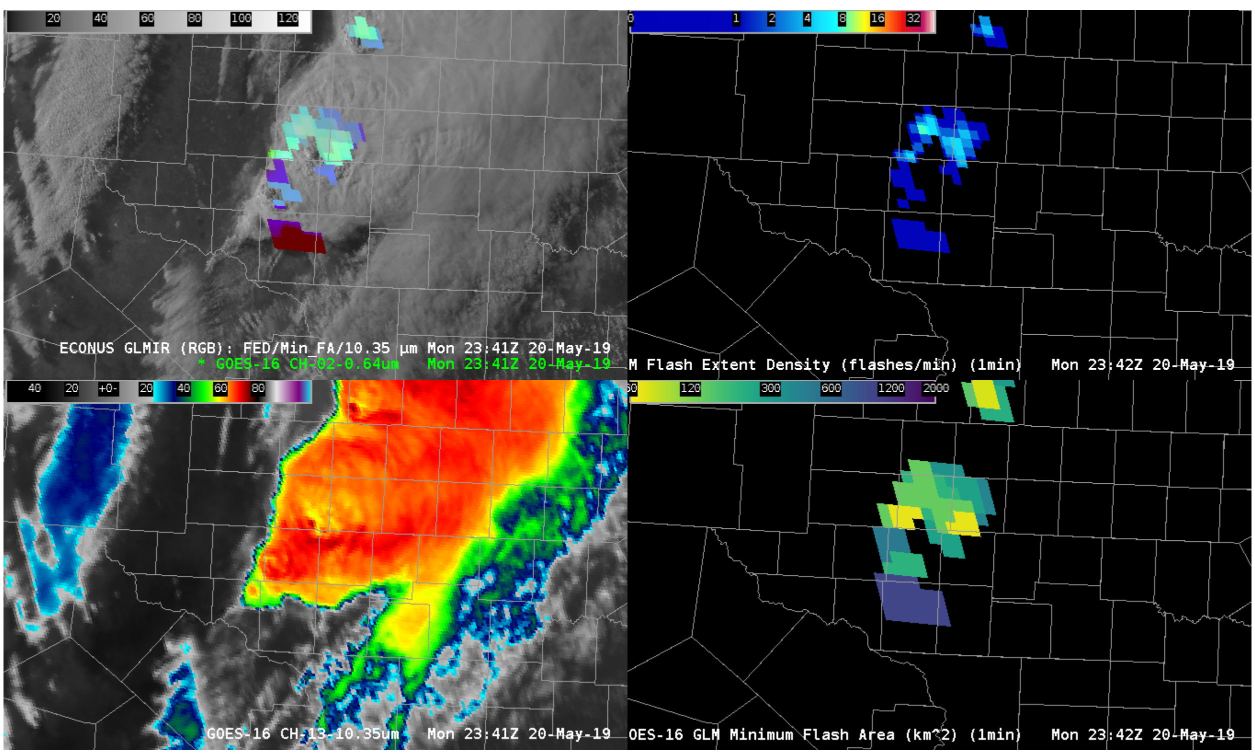

Figure 9: GLMIR RGB and Visible Satellite Imagery (Top Left); FED (Top Right); IR satellite (Bottom Left); GLM Minimum Flash Area (Bottom Right) at 2341 UTC on 20 May 2019.

Recommendations for Operational Implementation

Based on feedback from the 2019 HWT, we recommend the gridded product rollout of GLM continue and should become integrated into the baseline (Level2) products from GLM. In particular, FED, min flash size, event density, and flash energy should be made available. We also recommend the GOES-17 GLM CONUS projection move or expand to cover areas across the north central plains and central United States where it appears that GOES17 have the same or better detection efficiency as GOES-16. Additionally, we suggest the NWS take multiple steps to continue training efforts with local offices. Again in 2019, forecasters reported having a subject matter expert available to answer questions made the greatest impact on product understanding and use throughout the week. We would like to stress forecasters' use the GLM products in conjunction with the ground-based lightning products including Earth Networks Total Lightning Network, Vaisala's National Lightning Detection Network, and/or GLD360, as appropriate. It should also be emphasized in training that GLM may have a highest degree of impact in Decision Support Services as the network is the only one available to NWS forecasters that depicts the entire extent of lightning activity in a storm or region.

3.4 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 2019 for the fifth 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 newly launched NOAA-20 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 NOAA-20 overhead passes were created over the east coast around 1700 UTC, central US around 1830 UTC, and western US around 2000 UTC with a typical latency of one hour or so before the soundings were available for viewing by forecasters in AWIPS, which is manageable, but not ideal for operational use. 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 procedure just checks 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 an accurate representation of the atmosphere.

There were several other products presented for the NUCAPS evaluation in 2019 other than just the soundings. The effectiveness of these soundings and the impact the reduced latency had on forecaster use of the soundings was evaluated. An experimental version of the NUCAPS profiles was available again this year, with updates, for the NOAA-20 passes over CONUS 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 new boundary layer to replace the lowest levels of the sounding. 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 again this year for forecasters to view in AWIPS along with 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. Finally, a new NUCAPS Forecast product was tested this year that took the latest NUCAPS sounding and attempted to make a 6 hour forecast of stability parameters using the data from those soundings and advecting them forward in time via the HYSPLIT model.

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 quality and value of NUCAPS data for the severe weather forecast and warning process and to determine suggestions for improvement for readiness in operations.

Use of NUCAPS in the HWT

The most notable benefit of NUCAPS soundings and related products continues to be the ability of the soundings to fill in the gaps between regularly launched RAOB soundings both temporally and spatially. The timing of the soundings across the United States during the early afternoon is often times just an hour or two prior to expected convective initiation. These soundings allow the forecaster to get a sneak peek at the environment to look for areas where initiation might be likely to occur over the coming hours. In the HWT this year forecasters utilized the soundings from the newly launched NOAA-20 satellite which allows the latency issues of the soundings experienced in previous years to be cut down to less than an hour in most cases. This reduction in latency was generally received with positive feedback, making it much more useable in an operational setting than previous versions having longer latency. Given that the availability of the soundings is in the early afternoon, the soundings and associated data sets were primarily used by forecasters to assess the pre convective environment as well as the environment ahead of ongoing convection. Forecasters also noted the usefulness of NUCAPS soundings for data sparse areas and areas that don't have any RAOB sites nearby. The idea of an observational sounding is welcome to all forecasters who otherwise have to rely on model output for those areas. The NUCAPS soundings and the experimental boundary layer modified soundings were usually compared to model soundings and SPC Mesoanalysis parameters to assess their validity and usefulness for the day. Out of the products available with NUCAPS, the modified profiles were found to be the most useful by forecasters for making forecast decisions with approximately 70% of responses saying they found the modified profiles somewhat helpful to extremely helpful for their forecast, compared to about 57% for the unmodified profiles when they were asked at the end of each day.

> "The NUCAPS were most useful in assessing the region before convection fully developed."
> *Forecaster, End-of-Day Survey*

> "The modified profiles proved to be a good way to monitor the thermodynamic environment across the area and seemed representative when compared to other real time analysis data."
> *Forecaster, End-of-Day Survey*

> "I looked at the environment ahead of the Indiana storms and it gave me confidence that parameters were enhanced enough to sustain the strong convection (vs the high res models that killed the storms too early). Also, the soundings seemed to do okay in Montana, but overdid the CAPE and low level moisture."
> *Forecaster, End-of-Day Survey*

> "The modified profiles proved to be a good way to monitor the thermodynamic environment across the area and seemed representative when compared to other real time analysis data."
> *Forecaster, End-of-Day Survey*

One of the biggest evaluations for the NUCAPS soundings and products in terms of usefulness continues to be the product latency. This year, soundings from the NOAA-20 satellite, with an updated ground system, saw the sounding latency reduced by about half of what the previous latency was with the Suomi-NPP satellite. This was met with generally positive feedback from the forecasters as latency was typically less than an hour. Latency could still be an issue in some cases, like when convection starts before the sounding arrives in AWIPS or some instances when it took longer to get in front of the forecasters. The faster that the soundings can be delivered to the forecaster, the more utility they will have in operations, so that is an aspect that can continually be worked on. The modified sounding arrived shortly after the unmodified soundings which was also positive, as the forecasters tended to favor having the boundary layer modified soundings over the unmodified version when utilizing them for convective operations. Many forecasters stated that they would rather just have the boundary layer modified soundings in a convective environment because the boundary layer is so important for storm forecasting and the unmodified versions generally don’t represent the boundary layer well at all. The most value in severe storm operations tends to come from the modified soundings. Typically the latency associated with the gridded products and the NUCAPS-FCST proved them to be not as useful for this type of analysis.

> “The profiles and modified profiles arrived in time to sample an area that had not experienced any convection yet and provided information on the environment where storms were eventually expected to reach. For this case, the gridded and forecast profiles came in too late to examine the environment.”
> *Forecaster, End-of-Day Survey*

> “The modified NUCAPS soundings appeared to depict the boundary layer well.”
> *Forecaster, End-of-Day Survey*

> “Unmodified NUCAPS were not good for the boundary layer. Modified were definitely better and more usable in the boundary layer.”
> *Forecaster, End-of-Day Survey*

Below shows a case from May 23, 2019 in western Oklahoma where the forecaster compares the unmodified and modified NUCAPS soundings that displays the advantages that the boundary layer modification can have when trying to use these soundings in a severe storm forecasting environment. From the blog: “It uses a surface temperature of 63F, and a surface dew point of 50F, both more than 10 degrees below the nearby observations (Fig. 10). These inaccurately low values give, as one would expect, no sign of instability (CAPE=0). Looking at the modified NUCAPS gives a different picture, however. In the modified sounding, the surface temp is 69, and the dew point is 65, which is much closer to the observed surface obs (Fig. 11). This changes your surface-based CAPE to 2055 J/kg, vs. 0 from the unmodified. It also has a sharper low level inversion, which one would expect based on the stratus deck in place. To "verify", let's look at the SPC mesoanalysis. Sure enough, the mesoanalysis shows CAPE at around 2000 J/kg. Clearly, this is another case where the modified NUCAPS sounding is a noticeable improvement over the unmodified sounding.”

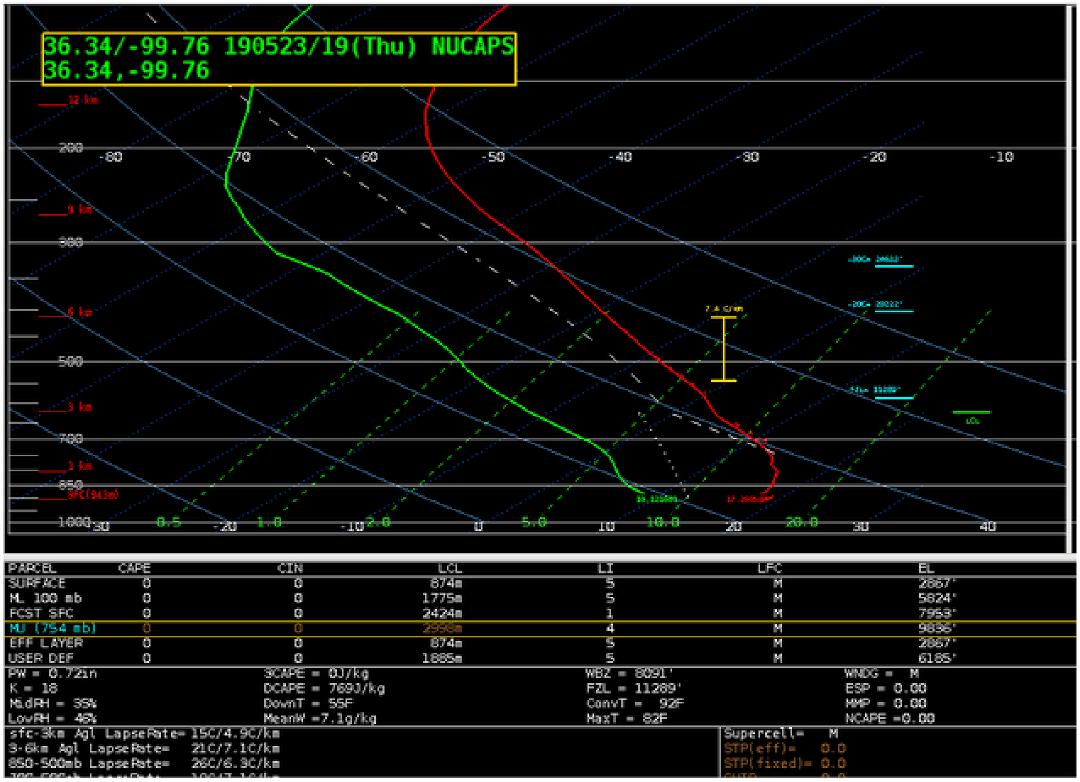

Figure 10: 23 May 2019 ~19 UTC NUCAPS sounding from western Oklahoma

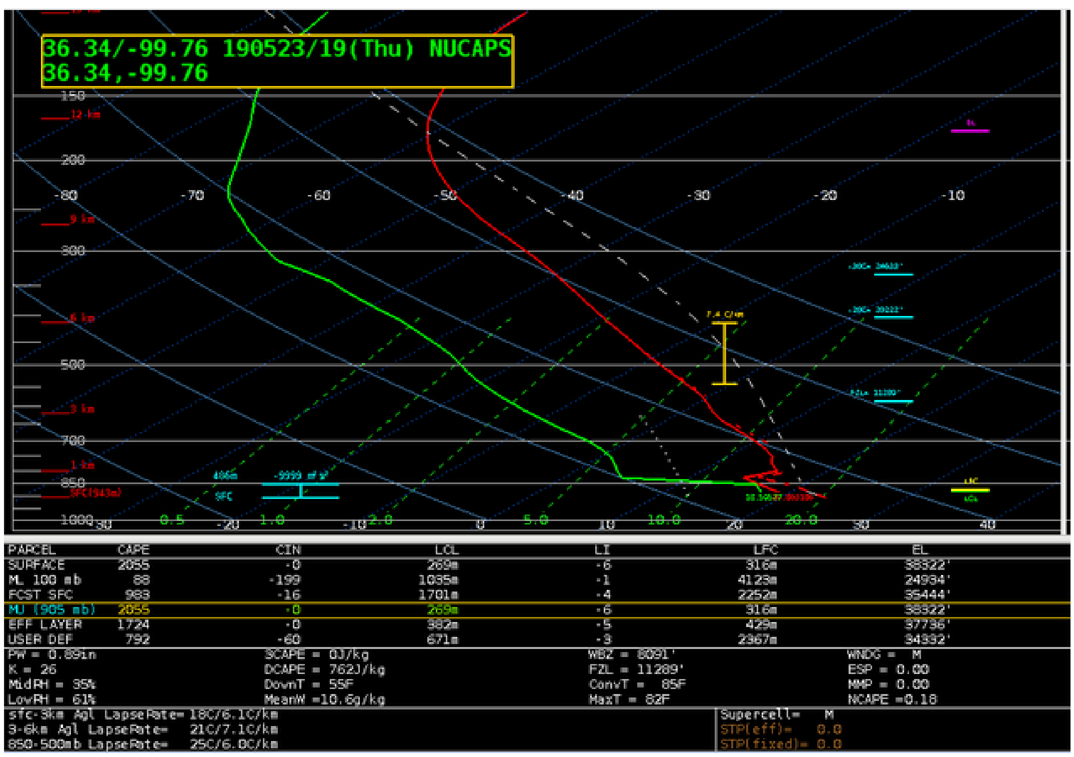

Figure 11: 23 May 2019 ~19 UTC Automatically Modified NUCAPS sounding from the same point as Figure 8 in western Oklahoma.

Along with the modified and unmodified soundings, there were other NUCAPS related products evaluated this year. The gridded plan views of NUCAPS profiles were once again evaluated within the testbed this year for the fourth time. These products allow the forecaster to view certain parameters that might be of some importance to the forecast in a plan view of data from all of the profiles in a swath instead of having to click through multiple soundings. Some of the plan views were useful, but there were still quite a few flaws with the data which will be talked about later in the limitations section. The parameter that forecasters most used the gridded data for was looking at mid-level lapse rates (Fig. 12). The lapse rates seemed to consistently match up fairly well to that from the SPC Mesoanalysis and when compared to high resolution model output. In addition to using the soundings and gridded plan views to look at lapse rates, forecasters also utilized the profiles and the gridded data to look for a few other parameters associated with severe weather. They were used to evaluate the height of the freezing level, the height of the -20 Celsius level for evaluating hail growth potential and severity, the height of the tropopause, as well as looking at 850 millibar and 700 millibar temperatures to try to identify any capping inversions or mid-level warming that could inhibit convection. Most of the use from the NUCAPS profiles and gridded data came at the mid to upper levels because of the issues noted in the boundary layer.

> "The pass happened around 1830Z and the CWA was on the edge of the swath, but NUCAPS was able to pick up on a weak low level cap that helped keep convection at bay most of the day."
> *Forecaster, End-of-Day Survey*

> "Used the information from the soundings to understand the freezing, -10, and -20 levels to understand where the hail-growth zones and environment. Primarily 21z - 23z in LSX."
> *Forecaster, End-of-Day Survey*

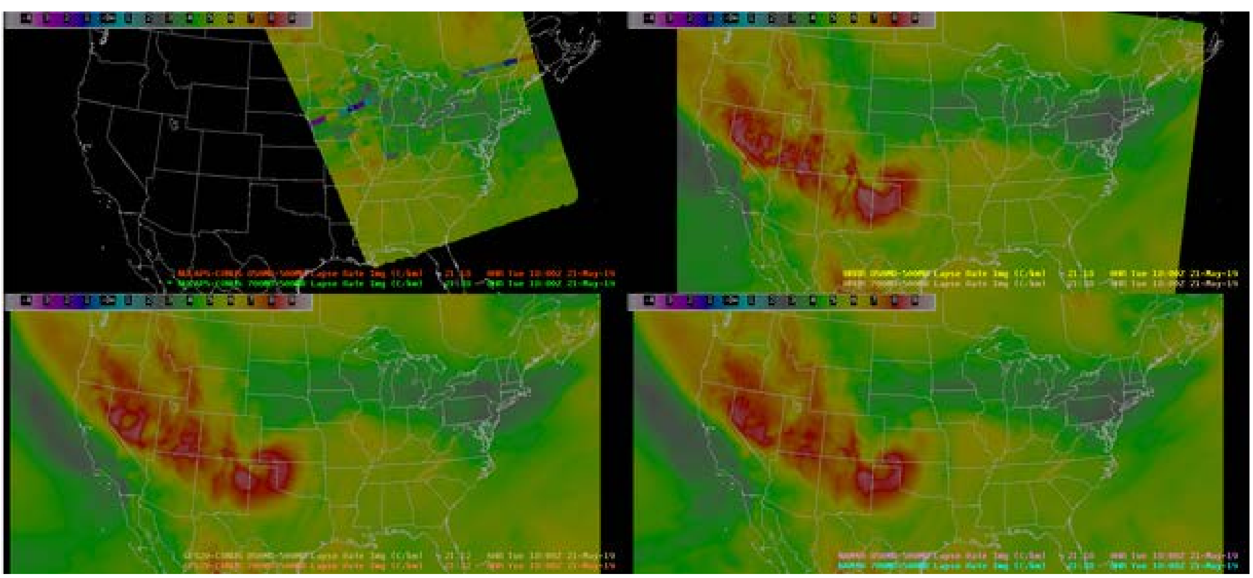

Figure 12: 21 May 2019 1800 UTC 700-500mb lapse rates for NUCAPS (top left), HRRR (top right), NAM (bottom right), and GFS (bottom left).

"NUCAPS 700-500mb lapse rates in northern Arkansas are close to model forecasts. The NUCAPS in this pass was indicating lapse rates around 6.2 C/km. Model data was generally a little bit lower around 5.8 C/km. The SPC mesoanalysis was on the high end near 6.5 C/km."
Forecaster, All Sky, TPW, and NUCAPS Moisture and Instability Profiles, *GOES-R HWT Blog*

New this year was the NUCAPS-FCST product (Fig. 13). This products takes the most recent swath of soundings and advects them forward in time using the HYSPLIT model to give plan view hourly forecasts out to six hours of CAPE, CIN, etc. These data were brand new and experienced many issues with data flow and usefulness in its first run in the testbed, mainly stemming from the numerous "holes" and discontinuities in the data that made it unusable in many instances. The concept was applauded by forecasters, but the product still needs a lot of work to become useful in an operational setting. When asked about the use of the NUCAPS-FCST product at the end of each day, a little over 70% of responses from the entire experiment indicated that they simply did not use the product for various reasons. The product was unavailable many times throughout the experiment, or it simply came in too late to be of much use to the forecaster who was already in warning mode. There were a few instances where the product showed some promise showing good instability and moisture values in an area when compared with model output. The forecast product was greatly affected by the quality of the soundings that it was initialized on.

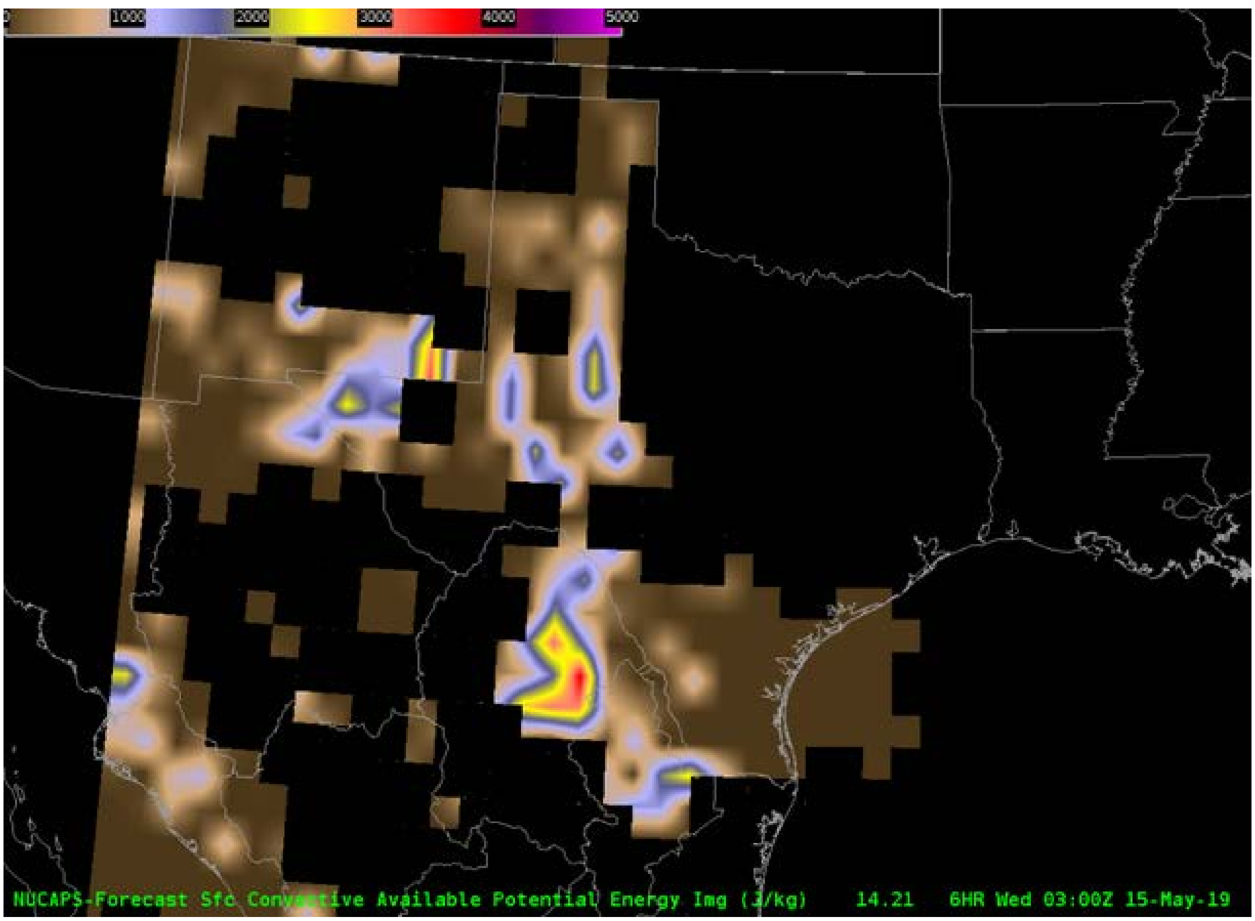

Figure 13: 15 May 2019 6 hour forecast valid at 0300 UTC of SBCAPE from the NUCAPS-FCST product across the south central U.S.

> "The FCST CAPE was too noisy to use."
> *Forecaster, End-of-Day Survey*

> "It arrived too late for use of the forecast data."
> *Forecaster, End-of-Day Survey*

> "It did along with rap analysis. It showed unstable air along with straight hodographs. The pwats also depicted a good amount of moisture for that area that was initialized well with sounding data."
> *Forecaster, End-of-Day Survey*

> "NUCAPS forecast did a good job at forecasting the maintenance of convection."
> *Forecaster, End-of-Day Survey*

Limitations and Recommendations for Improvement

As optimizing NUCAPs soundings continues to be worked on each year that the products are evaluated in the HWT, there are still a number of limitations with the products that keep them from being used regularly in operations. The number one limitation for the NUCAPS soundings for many years has been the latency of the product getting to AWIPS. With the introduction of the NOAA-20 satellite and updated ground system this year, that latency has been reduced greatly and forecasters in turn have appreciated that and find the soundings to arrive in a much more useable time frame than before. The latency can still be improved as many forecasters would like to still see the soundings earlier, but it is no longer the number issue relayed by forecasters in the experiment. The significant limitations to forecasters using these soundings for convective operations was the issues in resolving the boundary layer structure and that there were often times bad retrievals in the areas of interest or the swath of soundings would flat out skip over the area of interest and no sounding would be retrieved for that day. Many forecaster stated that they would rather just see the boundary layer modification soundings as the ones presented in AWIPS because they do a better job representing that part of the profile. While those soundings need to continue to be improved, forecasters overall found them more reliable than the unmodified soundings and it saves time for the forecaster not having to modify the soundings themselves on the fly. For the modified profiles, the latency for those can continue to be improved as well. Generally the modified soundings arrived fairly quickly after the unmodified, but there were times where they lagged behind by 30 minutes or more, which by that point the forecaster had moved on from sounding analysis. If the modified soundings are going to be used operationally, they need to arrive in a timely manner as well.

> "Main issue was dealing with the boundary layer, as mentioned above."
> Forecaster, End-of-Day Survey

> “Modified soundings were unavailable. The soundings that were available didn't look realistic as they did not depict any instability in the warm sector.”
> *Forecaster, End-of-Day Survey*

> “Unmodified NUCAPS were not good for the boundary layer. Modified were definitely better and more usable in the boundary layer.”
> *Forecaster, End-of-Day Survey*

> “The modified profiles seemed to be an improvement over the regular profiles, so I'm not sure there's much use for the non-modified profiles”
> *Forecaster, End-of-Week Survey*

> “Provide the modified profiles in AWIPS as an option since it is more efficient that having the forecaster modify the soundings themselves.”
> *Forecaster, End-of-Week Survey*

Some other limitations to using the NUCAPS profiles or modified profiles within operations is a problem with the display mechanism used within AWIPS. Many forecasters do not look at soundings in AWIPS-II/NSHARP regularly because of how the profiles can be distorted and it’s just not as functional as some other sounding programs. It is recommended that another option for displaying NUCAPS soundings be explored to increase use in forecast offices. Having NUCAPS available within SHarPy or BufKit would be ideal to gain a wider audience of forecasters looking at the profiles as part of their daily routine.

> “Mainly just better boundary layer representation, but the other battle you are facing is the use of NSHARP within AWIPS. A lot of forecasters don't like that. Having a way to view NUCAPS in SHARPpy or Bufkit then It would be more beneficial.”
> *Forecaster, End-of-Week Survey*

> “Using it in AWIPS is still a problem for me. I totally vote for SHARPpy or Bufkit ability. And again for today there was no pass over our forecast area for the day.”
> *Forecaster, End-of-Day Survey*

> “NSHARP is a poor platform to view these profiles.”
> *Forecaster, End-of-Day Survey*

Another main recommendation for improving the use of these profiles and other NUCAPS related products is to provide more satellite overpasses during the day to give more profiles than just the one. As mentioned earlier, there were several times throughout the experiment where a location would get missed between passes, or that the timing would be just slightly off to where the sounding wouldn’t arrive before convection started rendering them less useful in those situations. Forecasters continually stated that the more soundings that can be provided, particularly during the afternoon hours, the better. When asked at the end of each week to give some suggestions for improvement of the products some common themes stood out and are listed below.

- Provide gridded data for the modified profiles, especially instability related parameters

- Gridded data is too coarse of a resolution to be useful, look at ways to enhance the apparent resolution or interpolation between grid points.
- The NUCAPS-FCST data is too "splotchy" to use, find a way to interpolate between points better to fill in some of those gaps in the data
- Test out other model sources to use for advection and interpolation of the FCST product
- Adding identification to better know what sounding you are looking at and where it is located

"More satellite overpasses during the late afternoon."
Forecaster, End-of-Week Survey

"More satellite overpasses, Modified NUCAPS CAPE, Capping inversions, Cloud top height/temperature"
Forecaster, End-of-Week Survey

"MOAR DATA! More satellite overpasses from all available satellites would fill in gaps both spatially and temporally. It would also increase the chances of having a profile over any given area."
Forecaster, End-of-Week Survey

"Better navigation within NSHARP to know which dot you clicked on, and better comparison between points."
Forecaster, End-of-Week Survey

"Gridded for the modified soundings, especially CAPE"
Forecaster, End-of-Week Survey

"Not sure it was not very stable but when available it definitely had some blotchy data grids which made it unrealistic."
Forecaster, End-of-Week Survey, on NUCAPS-FCST Product

3.5 All-Sky LAP Stability Indices, Total Precipitable Water, and Layered Precipitable Water Products

University of Wisconsin Cooperative Institute for Meteorological Satellite Studies (CIMSS)

Blended all-sky moisture and stability fields were once again demonstrated in the HWT this year. These fields are derived via a fusion of GOES-16 radiance observations and numerical weather prediction (NWP) forecast data. The project has three components. The first component is the GOES-R Legacy Atmospheric Profile (LAP) retrieval algorithm, a baseline GOES-R product. The LAP algorithm generates retrievals in the "clear-sky" using information from GOES-16 ABI and the Global Forecast System (GFS) NWP model forecasts as a first guess. The second component computes retrievals in some cloudy regions, mainly thin/low clouds, also using information from ABI and a GFS first guess. Finally, the GFS NWP model "fills in" the areas where no retrievals are available from the previous two algorithms due to extensive sufficient cloud cover. Combining these components together provides one blended all-sky product. Fields that are derived from this all-sky algorithm and available to forecasters during the experiment included Total Precipitable Water (TPW), Layer Precipitable Water (LPW) in the SFC-0.9, 0.9-0.7, and 0.7-0.3 atmospheric layers in sigma coordinates, Convective Available Potential Energy (CAPE), Lifted Index (LI), K-Index (KI), Total Totals (TT), and Showalter Index (SI). There was also a data type product which showed what type of retrieval (clear, cloudy, or GFS) provided the value for each pixel. The all-sky products were available every 30 minutes at the top and bottom of the hours and provided full CONUS coverage from GOES-16 ABI at 6 kilometer resolution spatially. The purpose of this evaluation was to gather feedback for how the algorithm could be improved to better suit forecaster needs and to discover any technical issues with the product.

Use of LAP Products in the HWT

The All-Sky LAP fields were primarily utilized toward the beginning of the shift each day, prior to convective initiation to aid in pre-convective mesoscale analysis. The TPW and Layered PW fields were found to be the most beneficial in assessing moisture trends and moisture return into the region of interest. These products were met with great positivity throughout the experiment and is something that forecasters would like to see operationalized. When asked at the end of each week if they would like to see the three Layered PW products become part of the operational suite of products, 34/35 forecasters responded yes they would like to see them in operations. When asked which of the current All-Sky products they would like to see part of an operational version, the overwhelming answer was for the TPW and LPW products to be of highest importance, followed by CAPE and LI. The Total Totals, K-index, and Showalter Index got very limited use throughout the experiment and would not be considered a priority to be implemented operationally. The layered PW provided better definition of elevated mixed layers and low level moisture gradients, along which convection would eventually initiate. Forecasters also continuously stated that these PW products would have a wide range of applications aside from severe convection, like flash flood forecasting and winter storms. Regarding the CAPE and LI fields, forecasters found that looking at these values in a more relative sense and not getting stuck on the actual values was most useful when looking for areas of instability. Looking for trends and gradients in the instability often represented areas of future convective development or maintenance and helped enhance confidence in the forecast for a given area.

> "Being able to visualize and understand the distribution of the total PW values throughout different layers of the atmosphere is important. It can really have tremendous utility, especially in convective and flooding type environments."
> *Forecaster, End-of-Week Survey*

> "Yes, the low level PW product is extremely useful to see the PW gradients and where storms are most likely to develop. I saw this over and over again in the Testbed, and in the blog from previous weeks."
> *Forecaster, End-of-Week Survey*

> "The Layer PW is very useful for diagnosing moisture depth across the region which can tell us more about the quality of moisture. The different layers can also pinpoint areas of elevated moisture advection patterns for favorable elevated convection area."
> *Forecaster, End-of-Week Survey*

“I regularly use the operational GOES LAP products, but do have issue working with the blackout areas. Having these gaps filled up, the products we tested this week looked very smooth and realistic meteorologically. I would like to have access to them all, but if there is an issue with bandwidth, then the ones I'd want the most would be...in order of importance...CAPE, TPW, LPW, LI, TT, and KI.”
Forecaster, End-of-Week Survey

“The layered information would be very helpful for analyzing boundaries at various levels in the atmosphere for convective initiation, and would also be very helpful information for atmospheric river and monsoon events in the western CONUS”
Forecaster, End-of-Week Survey

The All-Sky LAP fields proved to be beneficial in a number of cases throughout the experiment for analyzing the pre-convective environment and identifying boundaries conducive to sparking convection. One such case is from West Texas on 17 May 2019. During the early afternoon mesoscale analysis, the forecaster noticed the low-level PW field showing a sharp dryline across the Texas New Mexico border with a pronounced dryline bulge forming in the west Texas panhandle (Fig. 14). This became the primary focal point for future convective initiation that the forecaster would monitor the rest of the afternoon. Later on, another forecaster checked the All-Sky CAPE to notice an area of increasing CAPE across southwest Texas in the Midland and Lubbock CWAs (Fig. 15). This dryline and associated sharp gradient in CAPE eventually spawned many supercells that went on to produce severe weather across the region (Fig 12 & 13).

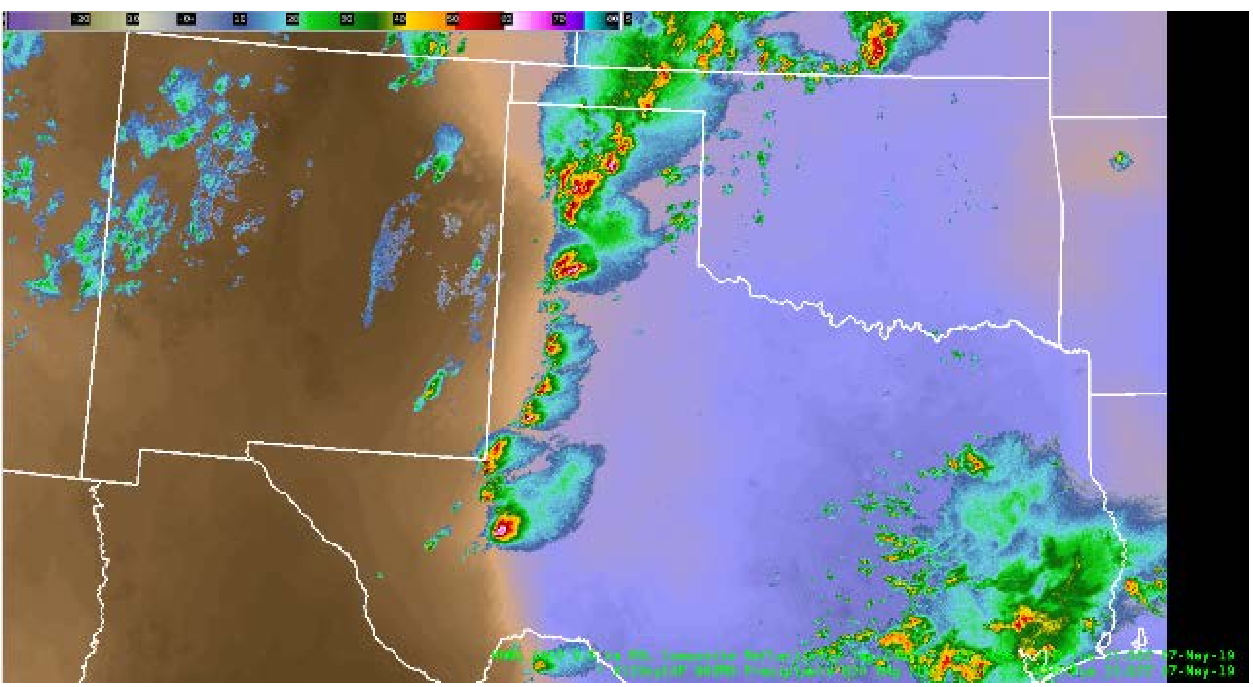

Figure 14: 2157 UTC 17 May 2019 All-Sky LAP sfc-0.9 sigma level “low-level” precipitable water with MRMS Composite Reflectivity across the Southern Plains

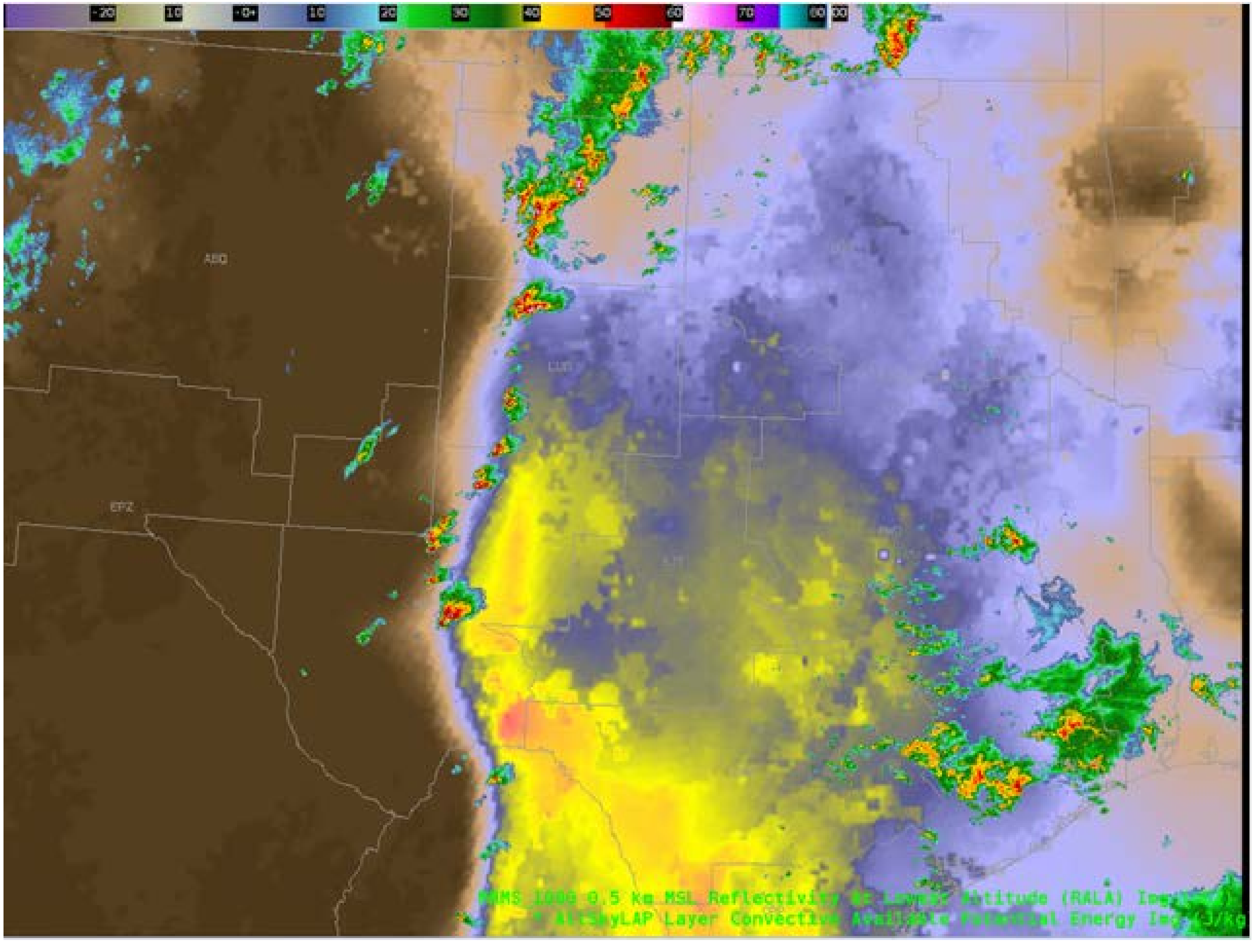

Figure 15: 2157 UTC 17 May 2019 All-Sky LAP CAPE with MRMS Reflectivity at Lowest Altitude (RALA) across the Southern Plains

The forecasters were also asked to evaluate the resolution of the All-Sky data, both spatial and temporal, for operational use. The version of the All-Sky evaluated in the HWT this year was provided at a 30 minute temporal and 6 kilometer spatial resolution. When asked if this would suffice for operational use, the majority of forecasters (72%) responded that the spatial resolution was pretty good for these types of products, but that 30 minutes temporal was a little long when looking on the fast time scales of severe convection. There were also times where the product was a little slow getting into AWIPS and caused the data that was being looked at to be 30 minutes to an hour old which was another operational concern when thinking about its daily use.

> "6km is good, but I think finer resolution data, if possible, would be better!"
> *Forecaster, End-of-Day Survey*

> "Yes and now...6km is quite good, but I do find myself wondering if there would be more mesoscale utility if at 2-3km"
> *Forecaster, End-of-Day Survey*

> "Would prefer to see at least every 15 minutes."
> *Forecaster, End-of-Day Survey*

"30 min ok - 15 minutes would be better to help diagnose trends given latency"
Forecaster, End-of-Day Survey

Limitations and Recommendations for Improvement

The LAP fields were often used in concert with and compared to other datasets typically utilized by forecasters during convective warning operations such as radiosonde data, model data (e.g. RAP, HRRR, NAM), SPC Mesoanalysis, and NUCAPS soundings. Some of the advantages of using the All-Sky LAP fields have already been discussed, but there were some limitations noticed throughout the experiment and recommendations to address for future improvements. This suggestions are listed below:

- Number one is that forecasters would like to see the LAP fields used with a different higher resolution model, particularly over CONUS, such as the RAP or HRRR. This might help with some of the discontinuities seen and provide better more accurate values as opposed to a coarser GFS forecast.
- As mentioned above, there were often large discontinuities between the clear sky retrieval areas and the GFS only areas which made the forecaster less confident in using the data in those situations and can hinder confidence in the product down the line in operations. Correcting those large discontinuities is of high importance, either by better interpolation techniques or trying different models to fill in those areas.
- Higher temporal resolution is desired by a fair number of forecasters to better assess trends in the convective environment, while they would welcome higher spatial resolution, the 6 kilometer spatial resolution seems to work well with most forecasters.
- Continue to improve on the latency of the product. Data would often be 30 minutes to an hour old by the time forecasters could see the updates in their AWIPS, which in a warning environment can mean a lot.
- Provide some more dynamic range on the low end of the color table, especially with the LPW products, to better pull out some of those more subtle boundaries and gradients easier on the low end of the spectrum.
- When the new GFS model runs were incorporated, the data would drastically change values over the course of a 30 minute update which threw the forecasters off when the data became jumpy and more difficult to assess which values have confidence in.

"More temporal resolution would be preferred. Try using NAM, RAP, or HRRR to incorporate more rapidly updating datasets?"
Forecaster, End-of-Day Survey

"Again, the latency and low temporal resolution renders this product to only just a synoptic overview and of little use during actual severe weather operations."
Forecaster, End-of-Day Survey

"Noticed today that the latency was up to an hour, which meant I got half hour data an hour too late. That's not operationally-feasibly usable."
Forecaster, End-of-Day Survey

"Would be curious to see the 15-minute resolution with RAP or other high-res model data filling in instead of the GFS."
Forecaster, End-of-Day Survey

"Noticed areas where the GFS kicked in and there was a noticeable drop where everywhere around it was higher"
Forecaster, Daily Debrief May 9, 2019

"Enhancing the low end of the color scale would be useful in lower CAPE environments"
Forecaster, Daily Debrief June 4, 2019

"There was a noticeable drop in the CAPE fields when new model runs come in which caused the data to be jumpy"
Forecaster, Daily Debrief May 9, 2019

Overall, the reception of these products was positive among the forecasters that participated in the experiment this year. A majority would like to see the All-Sky version replace the current LAP indices as the operational baseline product currently in operations, while still working to improve the product to provide more uses during warning operations.

"Yes, although, I'd recommend a further testing utilizing other data rather than GFS at least at a CONUS scale. Filling in cloud gaps with GFS data can get outdated rather quickly, especially in a convective type environment where cold pool propagation can render CAPE, TPW, LPW values almost useless rather quickly. Otherwise, would love the ability to analyze TPW, LPW, LI, and CAPE in a pre-convective type environment."
Forecaster, End-of-Week Survey

3.6 Advanced Blended Total Precipitable Water

Colorado State University (CSU) and the Cooperative Institute for Research in the Atmosphere (CIRA)

The advanced Blended Total Precipitable Water (ATPW) product was evaluated for the first time in the HWT during this experiment. This product uses a blend of polar orbiting microwave retrievals advected to the top of each hour using the methodology of Wimmers and Velden (2011) and GOES-16 TPW in the clear sky areas over the CONUS. The ATPW are produced every hour at ~:20 past the hour and at a spatial resolution of 16 kilometers. Feedback about the performance of the product compared to other operational and experimental TPW fields was desired for the experiment. It was of particular interest to evaluate if this blended TPW version was more useful than the current operational Blended TPW product that uses a blend of polar orbiting microwave retrievals and surface based measurements.

Use of ATPW in the HWT

A lot like the All-Sky products, the ATPW was viewed mainly toward the beginning of each shift when making the forecast for the day and looking for areas of moisture return or any semblance of boundaries that might be a focus for convection later in the day. The ATPW was used mainly to evaluate its effectiveness compared to the All-Sky TPW product and to the current operational blended TPW product. When asked at the end of each day if the ATPW performed better than the operational product, 68% of the responses were that yes, the ATPW performed better that day. So in general the ATPW was superior to the operational blended TPW but some major work still needs to be done on the product for many of the forecasters to begin using this product regularly in operations. Some of the limitations and caveats will be discussed later on in this section.

> "The ATPW product showed higher values than the operational product, which seemed consistent with the environment."
> *Forecaster, End-of-Day Survey*

> "The ATPW values were more realistic. The blended TPW seemed to be too high at times, especially across central and eastern OK, southward into Texas."
> *Forecaster, End-of-Day Survey*

> "ATPW did better somewhat but both struggled in depicting the dry air in the wake of the MCS and behind the dryline. In reality the gradient of moisture (at least in low levels) was much greater than depicted in the TPW products. The ATPW had more reasonable values in the warm sector."
> *Forecaster, End-of-Day Survey*

Throughout the experiment, forecasters were also asked to evaluate if they feel it is important to have a blended product that is independent of model data and possible biases. The current operational version of the product uses no model data in its calculations, whereas the ATPW uses GFS winds to advect the polar data and the GOES-16 product that goes into the ATPW also utilizes some GFS data. Overall, forecasters were pretty indifferent to whether or not they feel like the products needs to be independent of model data. When asked at the end of each week, 50% responded that it was moderately important, while only 8% felt that it was extremely important. Many responses from forecasters felt that if the data was incorporated correctly, that it isn't a problem to have it as part of the product if it presents a better result. The idea of a totally observational product is a good one in theory, but is hard to accomplish in practice to make it a good product. Forecasters were also asked about the hourly resolution, and 60% responded at the end of the week that with the data that goes into it, that one hour is sufficient while others would like to see a little higher resolution data on the convective scale to analyze trends and boundaries better.

> "It is sufficient, but the more temporal resolution the better for analyzing trends."
> *Forecaster, End-of-Week Survey*

> "We need to rely on model data to fill in gaps in certain areas. As long as model data is properly incorporated and weighted to observed data, then I have no issue with having model data in a blended PW product. I do not want to see model data override good satellite data. However, I do not like to see discontinuities, so properly incorporating model data to have clean product that still shows gradient is useful."
> *Forecaster, End-of-Week Survey*

"Although it would be nice to have a blended TPW product that is completely independent of model fields, this blended TPW product does not demonstrate much utility in its current state. I really like the idea of an all-observation product so as to avoid model bias/error, but the polar-orbiting satellites do not cut it in the cloudy areas. Model guidance is much closer and small errors in PW based on models is still superior to the current blended TPW."
Forecaster, End-of-Week Survey

"While it would be ideal (the less models the better in my opinion), I understand that is not always possible with satellite derived products. I prefer my satellite products to try to limit the use of model data, but if the sensor is unable to derive things because of clouds or what-not, then it's best to blend model data into it rather than have expansive no data areas. I like that the All-Sky products provide a data type graphic that you can underlay beneath the products so that you know if that data is coming from the instrument or is model data."
Forecaster, End-of-Week Survey

Limitations and Recommendations for Improvement

There were a couple of significant limitations associated with the ATPW product noticed throughout the experiment. A lot of these caused the product to be not very useful for severe storm forecasting and warning operations during the weeks. A list of the noticed problems is below:

- There were some very noticeable gradients between the clear sky areas using the GOES retrievals and the cloudy sky areas using the advected polar orbiting data (Fig. 16). This caused the data to become often times unusable if there were clouds in the area of interest. Though it did seem to have less of these discontinuities and weird gradients than the operational blended product.
- The 16 kilometer resolution made the ATPW appear very "blocky" and missed out on some of the smaller scale boundaries that could be picked up on by the All-Sky LAP TPW and other datasets utilized by forecasters (Fig. 16).

"The ATPW continues to struggle over cloudy areas with erroneously low values."
Forecaster, End-of-Day Survey

"The ATPW product seems to have problems resolving cloudy areas and generates artificial gradients. The AllSky PW imagery seems to perform better."
Forecaster, End-of-Day Survey

"Blended CIRA product not useful - blocky and low update times."
Forecaster, End-of-Day Survey

"Blended CIRA product very blocky - with artificial "odd" gradients. Forecasters would generally dismiss it."
Forecaster, End-of-Day Survey

> "Yes, there was a sharp gradient in the ATPW between cloudy and clear areas. PW values were reasonable in the clear areas but too low in the cloudy areas."
> *Forecaster, End-of-Day Survey*

Below is an example of those artificial gradients across Oklahoma in an area of cloud cover. This was a common site throughout the experiment. In this example there was an 1800 UTC RAOB launched from Norman, Oklahoma that showed TPW values around 1.52 inches while the ATPW only shower about an inch. Meanwhile, the forecaster also compared the RAOB values to the All-Sky LAP TPW which showed values near 1.70 inches with a much more uniform area of higher values across Oklahoma as opposed to the ATPW.

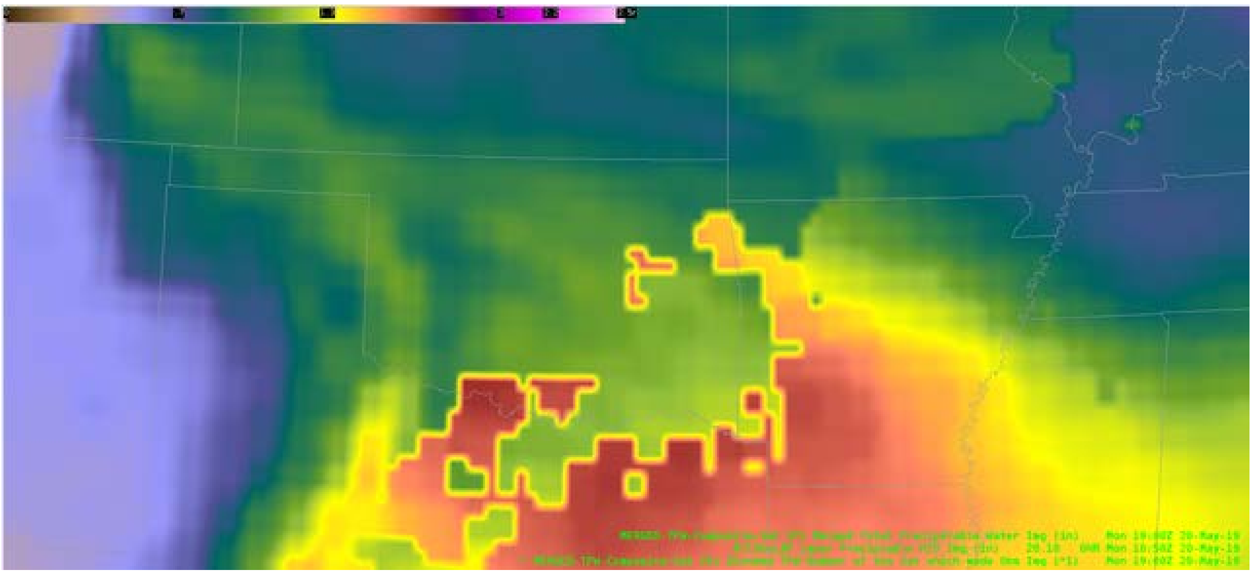

Figure 16: 1900 UTC 20 May 2019 ATPW product over Oklahoma and Arkansas showing the discontinuities between the clear and cloudy areas as well as the poor spatial resolution showing the dry line in the Texas panhandle not as sharp as compared to other datasets.

4. Summary and Conclusions

The GOES-R and JPSS Proving Ground conducted six weeks of satellite product evaluations alongside evaluations by the Radar Operations Center and NSSL and various experimental radar products during the 2019 Spring Experiment in the Hazardous Weather Testbed. Thirty NWS forecasters, three broadcast meteorologists, and two U.S. Air Force meteorologists evaluated many GOES-R and JPSS products and capabilities, and interacted directly with algorithm developers during the experiment. Participants had ample opportunity to subjectively evaluate, identify strengths and weaknesses, and suggest potential improvements for all of the products presented in this experiment. 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. Based on past feedback, more time was spent at the start of each week as a group going through each of the products in detail within AWIPS. This included a brief refresher about each product, a tutorial on where to load the products in AWIPS, recommendations for pre-built procedures, uses 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. Participants also received quick guides for most of the products to help with refresher training throughout the experiment when using the product in operations. The pre-built procedures were also well appreciated (especially by the broadcast meteorologists) as they facilitated a quick start to operations.

For the fifth 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 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 continued use of the vast 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.

More detailed feedback and case examples from the HWT 2019 GOES-R/JPSS Spring Experiment can be found on the GOES-R Proving Ground HWT blog at: www.goesrhwt.blogspot.com

Archived weekly “Tales from the Testbed” webinars can be found at: http://hwt.nssl.noaa.gov/tales/

Acknowledgements

Multiple NWS and broadcast meteorologists participated in this experiment and provided detailed feedback that has gone into the recommendations in this report, it would not be possible to complete research-to-operations without their willingness to openly play with experimental data. Tiffany Meyer also played a crucial role in the experiment as the one to install AWIPS 2

builds to handle the experimental data and applied many changes and modifications to the products before and during the experiment. Eric Bruning developed the code base that created the GLM products viewed in the HWT. The following people also participated as focal points and subject matter experts throughout the week for the products demonstrated in the HWT: Eric Bruning (Texas Tech University), Chris Schultz (NASA-Marshall Space Flight Center), Justin Sieglaff (University of Wisconsin - CIMSS), John Cintineo (University of Wisconsin - CIMSS), Jack Dostalek (Colorado State University – CIRA), Kris White (NWS WFO Huntsville, AL), Rebekah Esmaili (Science and Technology Corporation), Brian Kahn (Jet Propulsion Laboratory), and Bill Sjoberg (NUCAPS)

5. 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

Mecikalski, J. R., J. K. Williams, C. P. Jewett, D. Ahijevych, A. LeRoy, and J. R. Walker, 2015: Probabilistic 0–1-h convective initiation nowcasts that combine geostationary satellite observations and numerical weather prediction model data, *J. Appl. Meteor. Climatol.*, **54**, 1039–1059. doi:10.1175/JAMC-D-14-0129.1

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.

Wimmers A. J., and C. S. Velden, 2011: Seamless advective blending of total precipitable water retrievals from polar-orbiting satellites. *J. Appl. Meteor. Climatol.*, **50**, 1024-1036, https://doi.org/10.1175/2010JAMC2589.1.