THE EXPERIMENTAL WARNING PROGRAM

2015 Experiment Summary

NOAA Hazardous Weather Testbed, Norman, OK

23 February 2016

Gabriel Garfield[1, 2]

K. Bowden[1,2], R. Clark[1,3], S. Cobb[1,3], J. Correia[1,4], K. Calhoun[1,3], Z. Flamig[1,3], J. Gourley[3], P. Heinselman[3], C. Karstens[1,3], D. Kingfield[1,3], C. Ling[5], D. LaDue[6], B. Line[1,4], S. Martinaitis[1,3], G. Stumpf[1,3,7]

[1] OU/Cooperative Institute for Mesoscale Meteorological Studies, Norman, OK
[2] NOAA/NWS/Weather Forecast Office Norman, OK
[3] NOAA/NWS/NCEP Storm Prediction Center, Norman, OK
[4] NOAA/NWS/NCEP Storm Prediction Center, Norman, OK
[5] University of Akron, Akron, OH
[6] OU/Center for Analysis and Prediction of Storms, Norman, OK
7 NOAA/NWS/Meteorological Development Laboratory Norman, OK

TABLE OF CONTENTS

1. INTRODUCTION 3
2. PROJECT OVERVIEWS 3
    i. Spring Warning Project 5
    ii. Probabilistic Hazards Information Experiment 6
    iii. Hydrology Experiment 7
    iv. Phased Array Radar Innovative Sensing Experiment 8
3. PROJECT DETAILS AND RESULTS 9
    i. Spring Warning Project 9
    ii. Probabilistic Hazards Information Experiment 16
    iii. Hydrology Experiment 20
    iv. Phased Array Radar Innovative Sensing Experiment 22
4. PUBLICATIONS 24
5. PERSONNEL 25
6. ACKNOWLEDGEMENTS 26

1. INTRODUCTION

The HWT is a joint project of the National Weather Service (NWS) and the National Severe Storms Laboratory (NSSL). The HWT provides a conceptual framework and a physical space to foster collaboration between research and operations to test and evaluate emerging technologies and science for NWS operations. The HWT was borne from the "Spring Program" which, for the last decade, has been used to test and evaluate new forecast models, techniques, and products to support NWS Storm Prediction Center (SPC) forecast operations. Now, the HWT consists of two primary programs. The original NSSL/SPC "Spring Program" is now known as the Experimental Forecast Program (EFP)[1].

The other activity in the HWT, and the subject of this summary, is the **Experimental Warning Program (EWP),** which is designed to test and evaluate new applications, techniques, and products to support Weather Forecast Office (WFO) severe convective weather warning operations. This was the ninth year for warning activities in the testbed. We gathered feedback from NWS operational meteorologists. User comments were collected during shifts, forecasters participated in live blogging, electronic surveys were given at the end of shifts, and discussions occurred during post-mortem de-briefings. Input from NWS operational meteorologists is vital to the improvement of the NWS warning process, which ultimately saves public lives and property. The NWS feedback on this test is most important for future development for the NWS and eventual implementation of new application, display, and product concepts into AWIPS2 and other operational systems.

[1] Note that the EFP Spring Program is not the subject of this Operations Plan. For more information on the EFP Spring Program, please contact Israel Jirak (SPC).

2. PROJECT OVERVIEWS

The National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) Experimental Warning Program (EWP) at the National Weather Center (NWC) in Norman, Oklahoma hosted the 2015 EWP Spring Program (EWP2015). Several experiments to improve National Weather Service severe weather warnings were conducted this spring in the NOAA Hazardous Weather Testbed (HWT) as part of the annual Experimental Warning Program, a joint project of the National Weather Service and NSSL/CIMMS to support NOAA's goal to evolve the National Weather Service and build a Weather-Ready Nation. This year, the 2015 EWP Spring Program featured 4 projects, which operated for 14 calendar weeks.

EWP Project	Operation Dates	Operational Weeks	Number of Forecasters
Spring Warning Project	4 May – 12 June	5 weeks	30
Probabilistic Hazards Information Experiment	4 May – 5 June	3 weeks	6*
Hydrology Experiment	6 July – 24 July	3 weeks	18
Phased Array Innovative Sensing Experiment	4 August – 21 September	6 weeks	30

Table 1: Details for the 2015 Experimental Warning Program.

* Plus 12 partners (emergency managers and broadcasters)

1. Spring Warning Project

The Spring Warning Project (SWP) was a joint project conducted by scientists representing the Geostationary Operational Environmental Satellite – R Series (GOES-R), Joint Polar Satellite System (JPSS), Lightning Jump Algorithm (LJA), and Earth Networks Total Lightning Network (ENTLN) groups. The experiment spanned five weeks in May and June 2015. During the experiment, 25 National Weather Service forecasters and 5 broadcast meteorologists participated as evaluators. Feedback included live blogging, experimental warnings, daily debriefs and surveys, weekly debriefs and surveys, conversations, and the "Tales from the Testbed" webinar. The following sub-projects were a part of the SWP.

A. GOES-R and JPSS Convective Applications

During the SWP, the GOES-R Proving Ground conducted a pre-operational demonstration of recently developed products and capabilities. These products are associated with the next generation GOES-R series of geostationary satellites, subject to the constraints of existing data sources to emulate the satellite sensors. This early exposure was designed to increase forecaster familiarity with future GOES-R capabilities. In this way, SWP forecasters were readied for receipt and use of the GOES-R data prior to the launch. Additionally, feedback received from participants is being utilized in the continued development of GOES-R algorithms. The first of the GOES-R series of satellites is scheduled to launch in October 2016. Additional demonstration of JPSS products introduces and familiarizes users with advanced satellite data that are already available.

B. Lightning Jump Algorithm

The lightning jump algorithm was also evaluated during the SWP. In severe storms, rapid increases in lightning flash rate, or "lightning jumps," are coincident with pulses in the storm updraft and typically precede severe weather, such as tornadoes, hail, and straight-line winds, at the surface by tens of minutes. The GOES-R Geostationary Lightning Mapper (GLM) provided a general path to operations for the use of continuous total lightning observations and the lightning jump concept over a hemispheric domain. SWP forecasters evaluated a gridded sigma-level based lightning jump product on a CONUS scale in real time.

C. Earth Networks Total Lightning Network

Earth Networks Incorporated's (ENI) total lightning and total lightning derived products were evaluated in real-time as part of the SWP. This experiment built upon the initial evaluation in 2014, including enhancements following forecaster feedback of product use and incorporation into the warning-decision process. ENI-derived products that were evaluated included storm-based flash rates tracks, and time-series as well as three levels of thunderstorm alerts. The 2015 evaluation tested the feasibility of use and performance under the stress of real-time warning operations. Forecasters evaluated an updated Lightning Jump Algorithm (LJA), based on the GOES-R Geostationary Lightning Mapper, which was enhanced based on feedback from forecasters participating in the 2014 program. These evaluations will help prepare for possible operational implementation in 2016 following the launch of GOES-R. Earth Networks' total lightning and total lightning derived products, including storm-based flash rates tracks, time-series, and three levels of thunderstorm alerts were evaluated in real time, building upon the initial evaluation in 2014. The 2015 evaluation tested the feasibility of use and performance under the stress of real-time warning operations.

Web Presence:

GOES-R HWT Blog	http://goesrhwt.blogspot.com/
EWP Blog	http://hwt.nssl.noaa.gov/ewp/internal/blog/
Forecaster Training	http://hwt.nssl.noaa.gov/ewp/internal/2015/ *

*(LDAP user name / password required)

Project Contacts:

Kristin Calhoun	kristin.kuhlman@noaa.gov	(Lightning)
Bill Line	bill.line@noaa.gov	(Satellite)
Tiffany Meyer	tiffany.meyer@noaa.gov	(Lightning)

2. Probabilistic Hazard Information Experiment

During the three weeks of the experiment, forecasters assessed a new tool using rapidly-updating high-resolution gridded Probabilistic Hazard Information (PHI) as the basis for next-generation severe weather warnings. This experiment was part of a broad effort to revitalize the NWS watch/warning paradigm known as Forecasting a Continuum of Environmental Threats (FACETs). The major emphasis of the HWT PHI experiment was on initial testing of concepts related to human-computer interaction while generating short-fused high-impact Probabilistic Hazard Information for severe weather. The long-term goal of this effort was to move the refined concepts and methodologies that result from this experiment into Hazard Services, the next generation warning tool for the NWS, for further testing and evaluation in the HWT prior to operational deployment.

This year marked the inaugural HWT Experiment with emergency managers (EMs). The EMs provided feedback on their interpretation of experimental probabilistic forecasts generated in the HWT from the PHI experiment and the Experimental Forecast Program (EFP). This feedback was used in conjunction with feedback from forecasters to refine how the uncertainty information is generated and disseminated.

Web Presence:

FACETS Program	http://www.nssl.noaa.gov/projects/facets/

Project Contacts:

Chris Karstens	chris.karstens@noaa.gov	(PHI Tool developer)
Daphne LaDue	dzaras@ou.edu	
Chen Ling	cl99@uakron.edu	
Greg Stumpf	gregs.stumpf@noaa.gov	

3. Hydrology Experiment

The Multi-Radar / Multi-Sensor (MRMS) Hydro Experiment (hereafter, "Hydro"), was a part of the 2015 United States Weather Research Program (USWRP) Hydrometeorological Testbed (HMT). The HMT-Hydro experiment was conducted in conjunction with the Flash Flood and Intense Rainfall (FFaIR) Experiment at the Weather Prediction Center (WPC) from 6 July to 24 July. During the experiment, National Weather Service and River Forecast Center forecasters worked with research scientists to assess emerging hydrometeorological concepts and products to improve the accuracy, timing, and specificity of flash flood watches and warnings. In particular, forecasters evaluated short-term predictive tools derived from MRMS quantitative precipitation estimates (QPE) and Flooded Locations and Simulated Hydrographs (FLASH) hydrologic modeling framework. The Hydro Experiment also explored the utility of experimental watch and warning products conveying uncertainty and magnitude issued through the Hazard Services software. This allowed research scientists to investigate human factors to determine operationally relevant best practices for the warning decision making process and the system usability of the Hazard Services platform.

Web Presence:

FLASH Website http://flash.ou.edu/

Project Contacts:

J.J. Gourley jj.gourley@noaa.gov
Steve Martinaitis steven.martinaitis@noaa.gov

4. Phased Array Radar Innovative Sensing Experiment (PARISE)

The 2015 Phased Array Innovative Sensing Experiment (PARISE) ran for six weeks during August and September. During the experiment, National Weather Service forecasters from the Great Plains utilized phased array radar data in displaced real-time warning operations. The primary goal of the experiment was to assess the impacts of higher-temporal resolution radar data on the warning decision processes and performance of the forecasters. Similar studies - conducted in 2010, 2012, and 2013 - showed encouraging results, but the sample size was too small for generalization. Thus, to improve the reliability of previous findings, this study increased the sample by increasing the number of forecasters to 30 and the number of cases worked to 9.

The 2015 PARISE featured three parts: 1) the traditional experiment, 2) an eye-tracking experiment, and 3) a focus group. While the traditional experiment built on knowledge obtained from previous experiments, the eye-tracking experiment brought a new and exciting avenue to the work of PARISE. Forecasters' eye gaze data was collected as they worked a case in simulated real-time. This data provided new insight into impacts of higher-temporal resolution on the forecaster warning decision process and allowed PARISE scientists to analyze and compare forecasters' cognitive processes objectively. Finally, the focus group session drew on participants' experiences throughout the whole week, and generated insightful feedback and ideas important to the development of a future PAR network. The traditional experiment took place in the Hazardous Weather Testbed, and the eye-tracking experiment took place in RRDD space on the 4th floor of the National Weather Center.

Web Presence:

PARISE Webpage	https://www.nssl.noaa.gov/projects/parise/

Project Contacts:

Katie Bowden	katie.bowden@noaa.gov
Pam Heinselman	pam.heinselman@noaa.gov

3. PROJECT DETAILS AND RESULTS

1) Spring Warning Project

During the Spring Warning Project, experimental products from the Geostationary Operational Environmental Satellite R-Series (GOES-R), Joint Polar Satellite System (JPSS), Lightning Jump Algorithm (LJA), and the Earth Networks Total Lightning Network (ENTLN) were evaluated in the Hazardous Weather Testbed. Product descriptions and some early results will be discussed in the following.

A) GOES-R / JPSS

GOES-R and JPSS baseline and experimental products were evaluated during the SWP. Among the key goals of this experiment were to a) generate user feedback, b) prepare users for new and future satellite systems, and c) foster interactions between algorithm developers and end-users.

i) NUCAPS Point Soundings (JPSS)

The NOAA Unique CrIS ATMS Processing System (NUCAPS) soundings are real-time temperature and moisture profiles developed by JPSS using information from the CrIS and ATMS instruments aboard the Suomi NPP satellite. These profiles are derived using an algorithm that combines both statistical and physical retrieval methods. NUCAPS resolves ~10 vertical layers of water vapor and ~20 vertical layers of temperature in the atmosphere. Forecasters select a profile from a swath of points in AWIPS-II, which can then be viewed and edited in NSHARP. The purpose of this demonstration was to capture the value added by NUCAPS to convective forecasting, learn what adjustments could be made to enhance its operational usefulness, and enlighten participants to the existence of NUCAPS in AWIPS-II. During the SWP, forecasters found these soundings to be useful in filling spatiotemporal gaps that exist in observed vertical sounding information. The general shape of the profiles and stability/moisture parameters appeared to be realistic when compared to observed soundings, increasing forecaster confidence when using these data. However, forecasters would like to see the process of editing the lowest layers of the profile (which often contain errors) be automated.

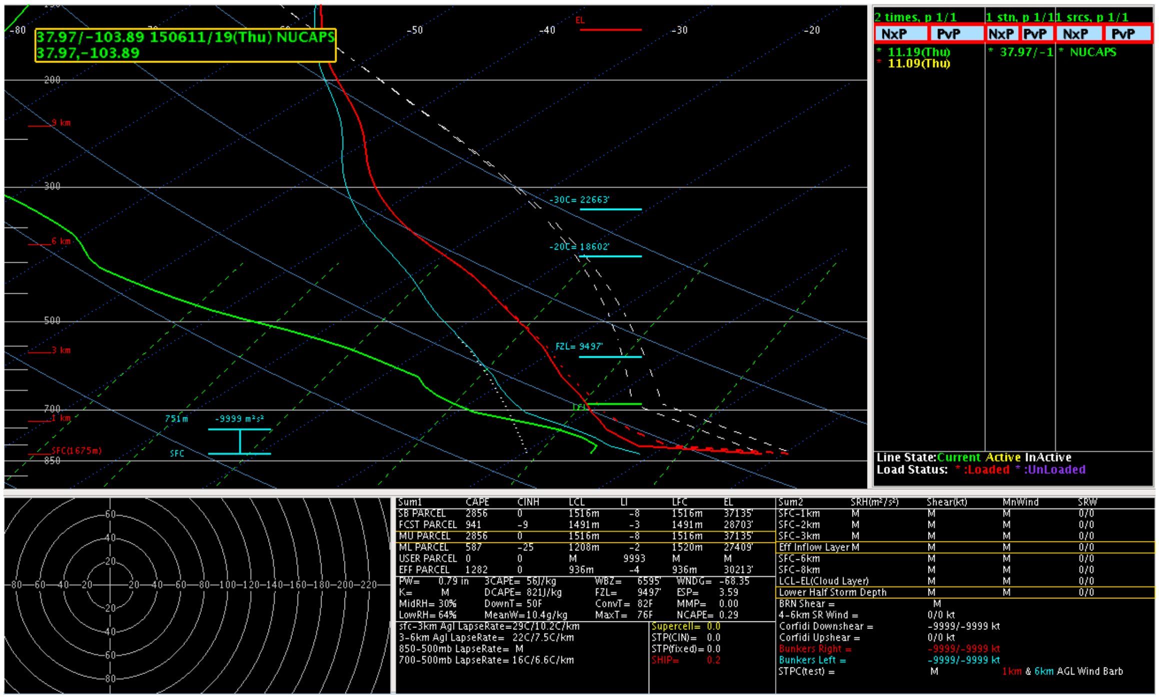

Figure 1: NUCAPS sounding used during the Spring Warning Project.

i) LAP Stability and Moisture Indices (UW/CIMSS)

The Legacy Atmospheric Profile (LAP) all-sky project combines a clear-sky retrieval algorithm, cloudy-sky retrieval algorithm, and NWP to provide an all-sky, plan view display of atmospheric stability and moisture indices. The retrieval algorithms use GOES Sounder data as a proxy for the GOES-R Advanced Baseline Imager (ABI). Among the evaluated fields were Convective Available Potential Energy, K-Index, Lifted Index, Showalter Index, Total Totals, Total Precipitable Water and Layer Precipitable Water (3 layers). Participants found the LAP products to be most valuable in depicting the gradients in stability and moisture along which convection would later focus development, as well as significant trends in the fields (increasing moisture, increasing instability) over an area. Forecasters primarily utilized the LAP fields to assess the recent evolution and current state of the environment at the beginning of the shift prior to CI, helping in their initial forecasts. Forecasters would like to have a method of knowing which of the three algorithms the data is from at a given point, along with improved training on layer PW.

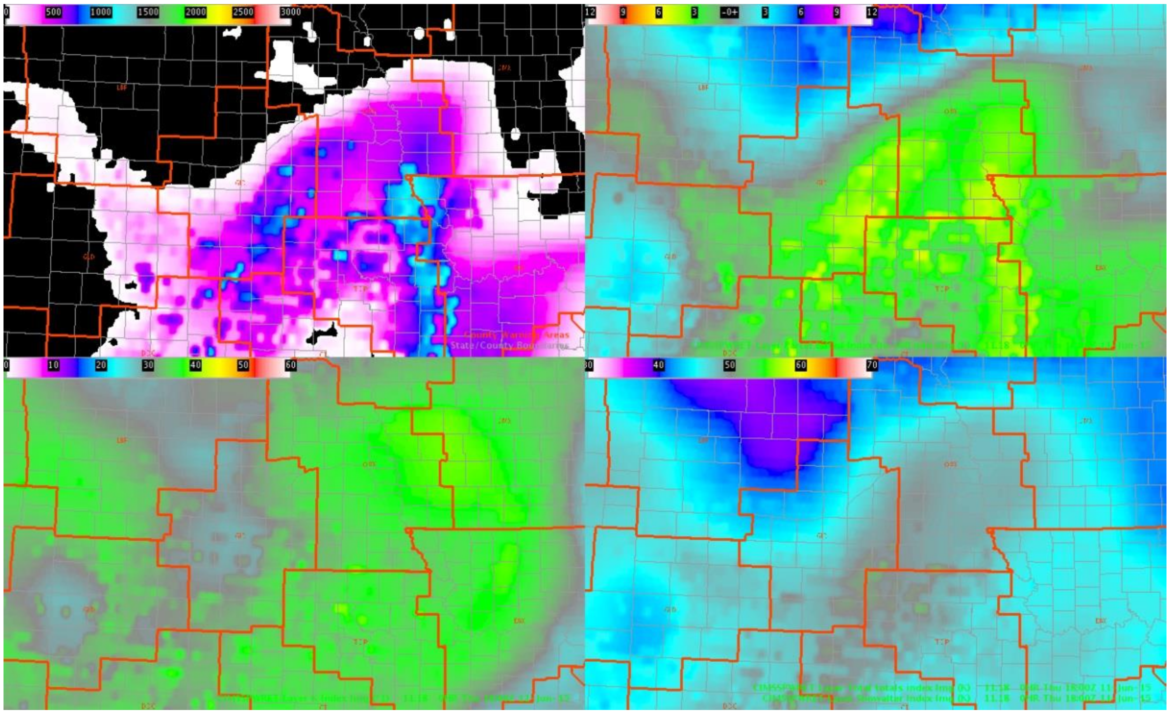

Figure 2: LAP stability indices used during the Spring Warning Project.

i) GOES-R Convective Initiation (UAH)

The GOES-R Convective Initiation algorithm fuses GOES cloud products and RAP-derived environmental fields and uses a logistic regression framework to produce a probability of future convective initiation for a given cloud object. Convective initiation here is defined as a 35 dBz reflectivity echo at the -10C level. Using objective validation techniques, a training database of over 500,000 objects has been developed, representing convective regimes much better when compared to earlier iterations of the algorithm. Participants thought the algorithm was effective in drawing user focus to areas where convective initiation was about to occur and away from where it was less probable to occur. They would like to see improved performance under cirrus clouds and in areas of congested cu, and would like an algorithm that provides CI probabilities to severe convection.

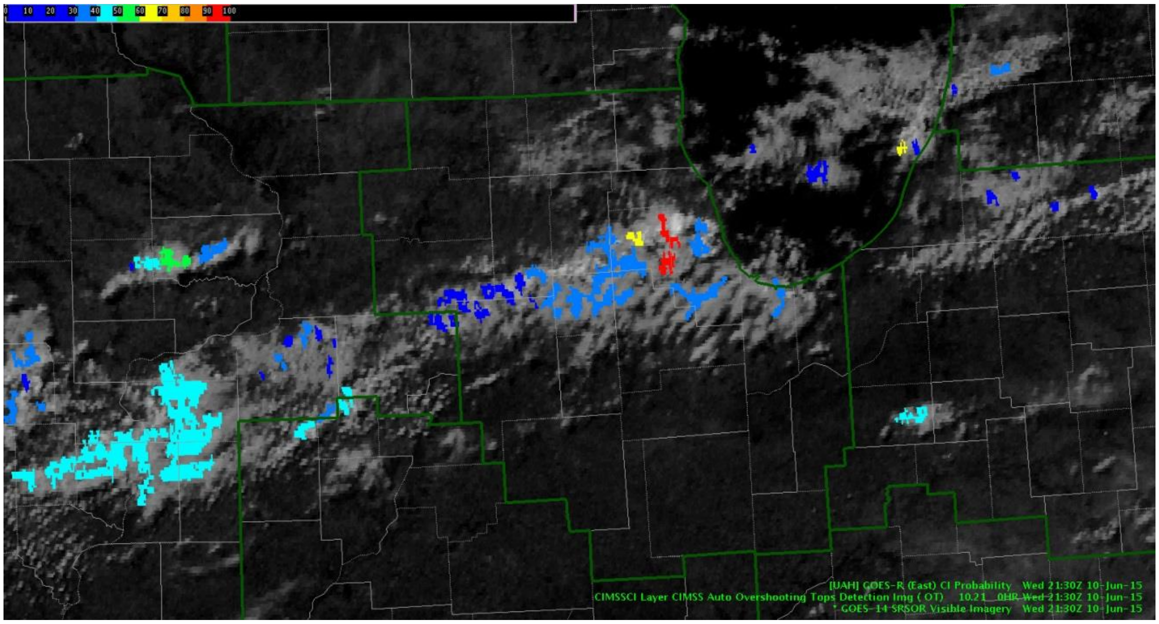

Figure 3: GOES-R Convective Initiation output used during the Spring Warning Project.

i) Probability of Severe Model (UW/CIMSS)

The Probability of Severe (ProbSevere) Model displays the probability that a developing storm will produce its first incidence of severe weather (i.e., hail, wind, or tornado) within the next hour. The data fusion product merges NWP-based instability and shear parameters, satellite vertical growth and glaciation rates, and radar derived maximum expected size of hail (MESH). A developing storm is tracked in both satellite and radar imagery using an object-oriented approach. The probabilities are displayed as contours around storm objects that change color and thickness with changing probability. Forecasters may interrogate the individual model input through cursor read-out, and the product updates every 2-minutes. ProbSevere was very popular with forecasters, as it helped them to maintain situational awareness to where the most significant storms were developing. Forecasters indicated that it often influenced their warning decisions, especially during busy severe weather situations. Participants would like to see probabilities by specific threat, a time series of recent probabilities, and improved performance when severe wind was the main threat.

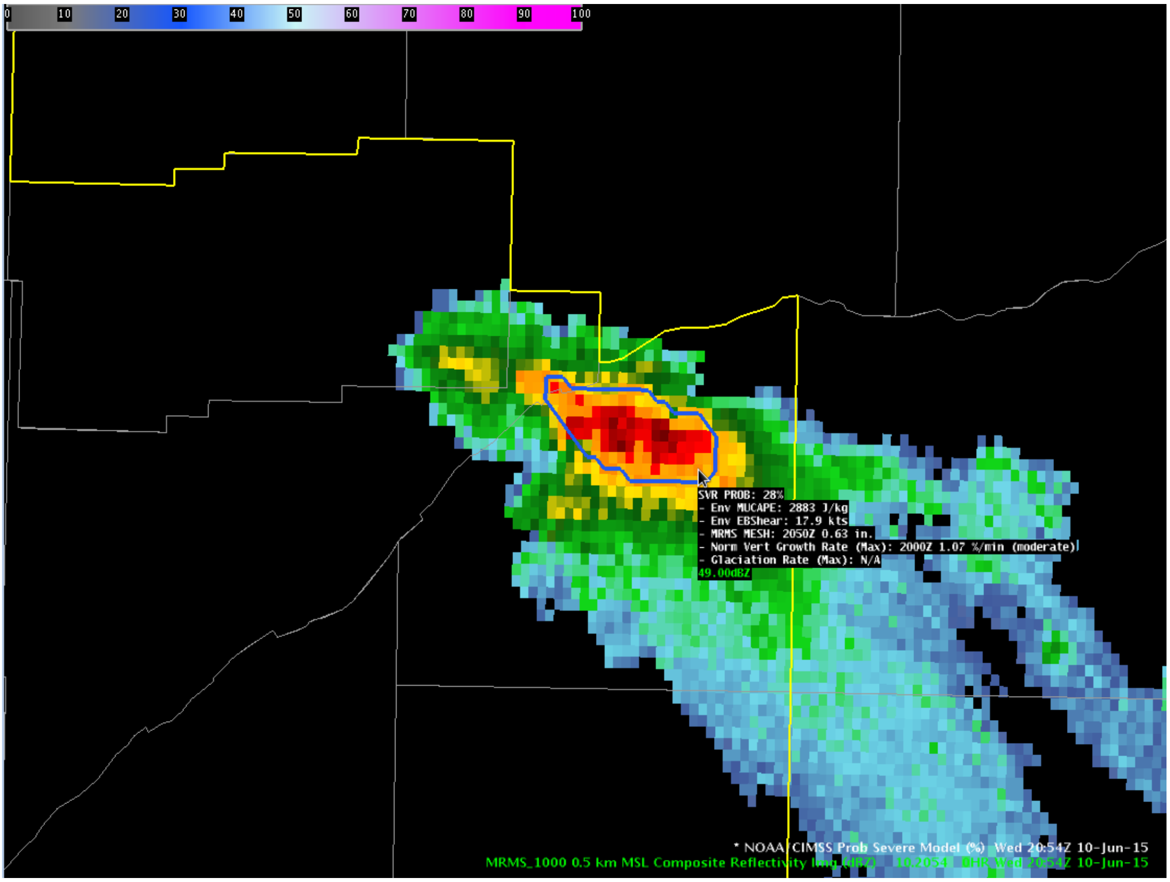

Figure 4: ProbSevere model output used during the Spring Warning Project.

i) GOES-14 Super Rapid Scan Operations for GOES-R (SRSOR)

The GOES-14 Imager was operated in an experimental 1-min mode known as "Super Rapid Scan Operations for GOES-R (SRSOR)" from 18 May to 11 Jun 2015 during the SWP. The location of the approximately 1500 km x 2000 km sector of 1-min satellite imagery was adjusted daily based on the expected region of most active hazardous weather. GOES-14 SRSOR demonstrates a capability of the GOES-R ABI when in Mode 3 "flex mode", which will include 30-sec imagery over one 1000 km x 1000 km sector, or two 1000 km x 1000 km sectors of 1-min imagery. In addition to the imagery, an automated Overshooting Top Detection algorithm (OTD) and Atmospheric Motion Vectors (AMVs) were derived from the 1-min data. Participants commented that they were able to diagnose subtleties in environmental changes that were not apparent in current operational satellite data (5-30 minutes). They mentioned that the fluid animations of 1-min imagery kept them ahead of the game, revealing features and processes much early than is otherwise possible. By combining the satellite imagery with other very-high resolution datasets such as lightning and radar, forecasters were able to verify their conceptual model of storm structures such as supercells.

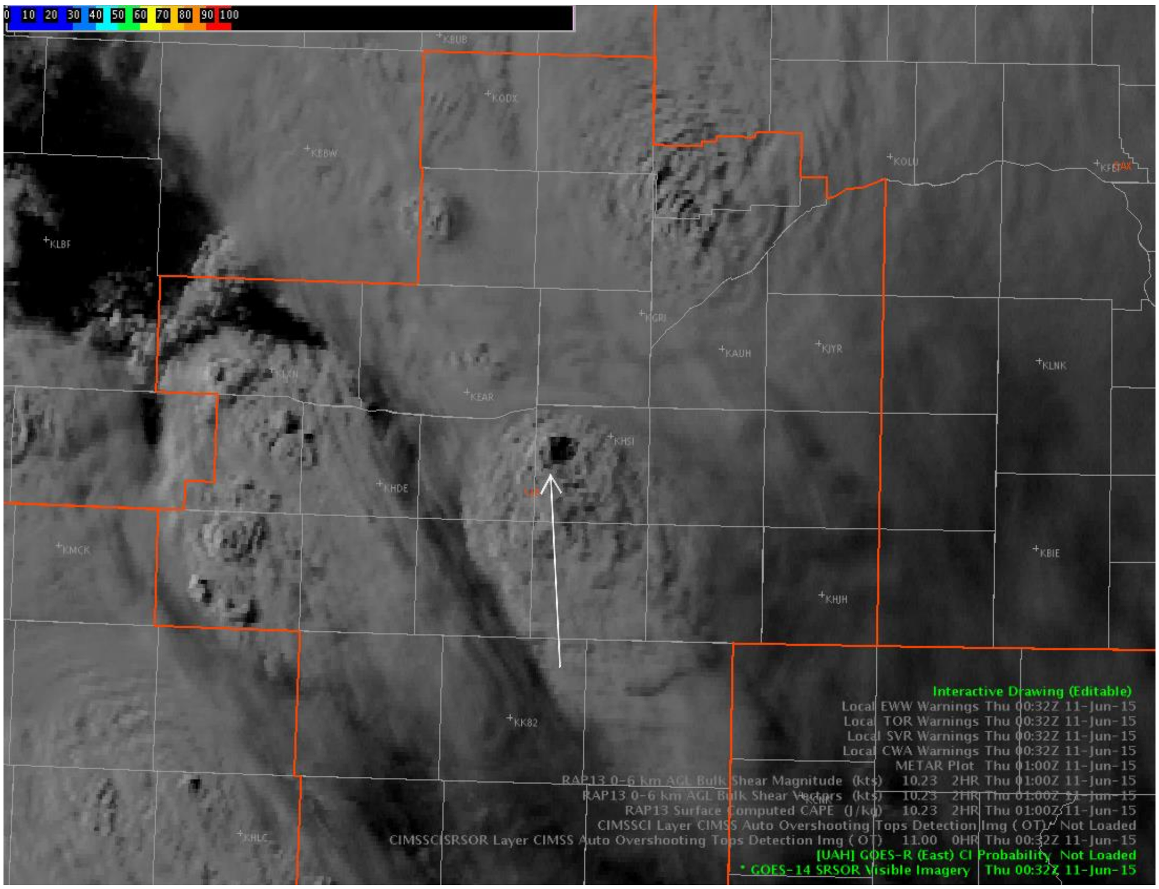

Figure 5: Overshooting tops seen using Super Rapid Scan during the SWP.

PGLM Total Lightning

ii) PGLM Total Lightning (NASA/SPORT)

The Psuedo Geostationary Lightning Mapper (PGLM) total lightning products were created using Lightning Mapping Array (LMA) total lightning data from regional networks around the CONUS. The LMA networks detect very-high frequency (VHF) radiation which provide the areal extent of total lightning activity. These data are sorted into flashes, which are then gridded and remapped to the spatial resolution of GLM (8 km x 8 km), and displayed as flash extent density. This demonstration allows forecasters to become familiar with lightning displays that may be avialable with the GOES-R GLM. Since total lightning can be used as a proxy for updraft intensity, forecasters used the data to monitor storm intesity changes in near real-time. Forecasters would like to see continued improvement to training for incorporating total lightning information into the warning-decision process.

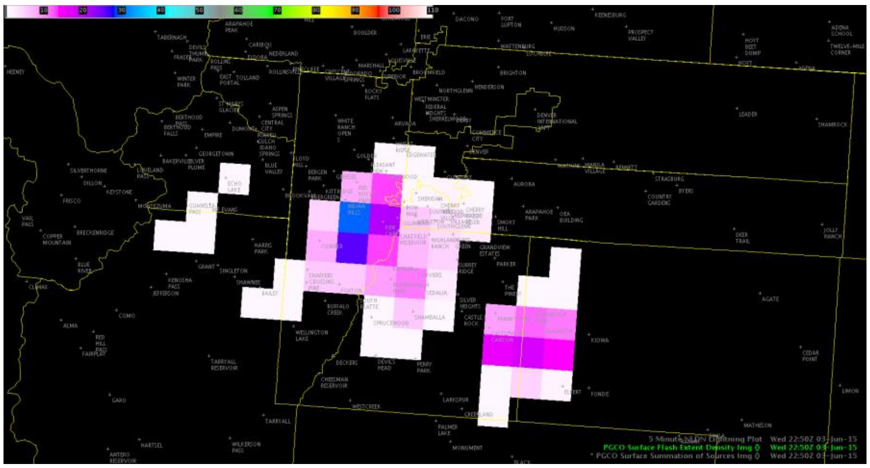

Figure 6: Total flash extent density seen during the Spring Warning Project.

iii) Lightning Jump

A lightning jump is defined as a rapid increase in a storm's flash rate. The Lightning Jump algorithm (LJA, Figure 7) leverages the physical relationships that produce lightning jumps to deduce storm intensity in real-time. How can it do this? To start, storm electrification is dominated by the non-inductive charging mechanism. In non-inductive charging, ice collisions that occur within the presence of super-cooled water cause charge separations to occur. This can be either positive or negative, depending on storm type, liquid water content, and temperature. Thus, flash rate is positively correlated with mixed-phase ice mass and cloud liquid water content. As these quantities increase, so does the flash rate. Lightning jumps are correlated with a higher probability of severe weather (Schultz et al. 2009).

During the Spring Warning Project, warning forecasters evaluated 1-minute LJA data in a real-time operational environment. Forecasters said the LJA increased their confidence in issuing severe thunderstorms warnings. Additionally, many forecasters commented that the LJA drew their attention to the most important storms in their area of responsibility. Finally, one forecaster observed that a series of lightning jumps seemed to be well-associated with a storm's intensity.

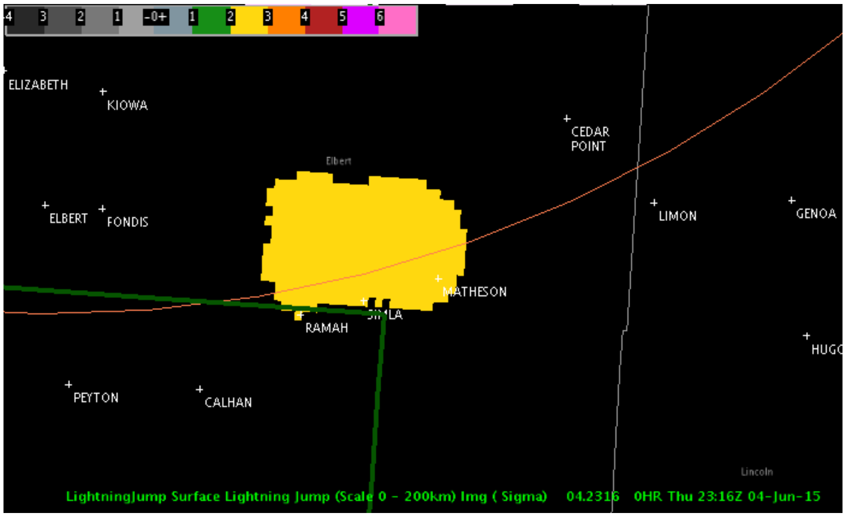

Figure 7: Lightning jump algorithm, as seen during the Spring Warning Project.

i) Earth Networks Total Lightning Network Experiment

Earth Networks Incorporated (ENI) has developed several algorithms for real-time severe weather operations based on data from their total lightning network. In order to evaluate these products, ENI algorithms were tested in the Hazardous Weather Testbed. In Year 1 (2014), the ENI product suite was tested in displaced real-time scenarios. Data was denied to forecasters to test the value of the products. In Year 2, the ENI products were introduced into real-time operations at the HWT (2015).

Among the products tested in Year 2 were configurable thresholds for alerts, time series of lightning parameters, and flash rate color tables. Forecaster feedback was generally positive for the flash rate products: one forecaster wrote, “It focused my attention to storms that were developing quickly with strong updrafts” and another wrote that the flash rates led to “higher confidence in issuing warnings.” Forecasters found the time series products useful, as they were an “easy way to visualize trends in total lightning” (Figure 8). Overall, forecasters found total lightning helpful in identifying storms that are intensifying or re-intensifying. However, the thunderstorms alerts were found to handle convective lines poorly. In general, one forecaster commented that “more thorough training is needed for NWS forecasters.

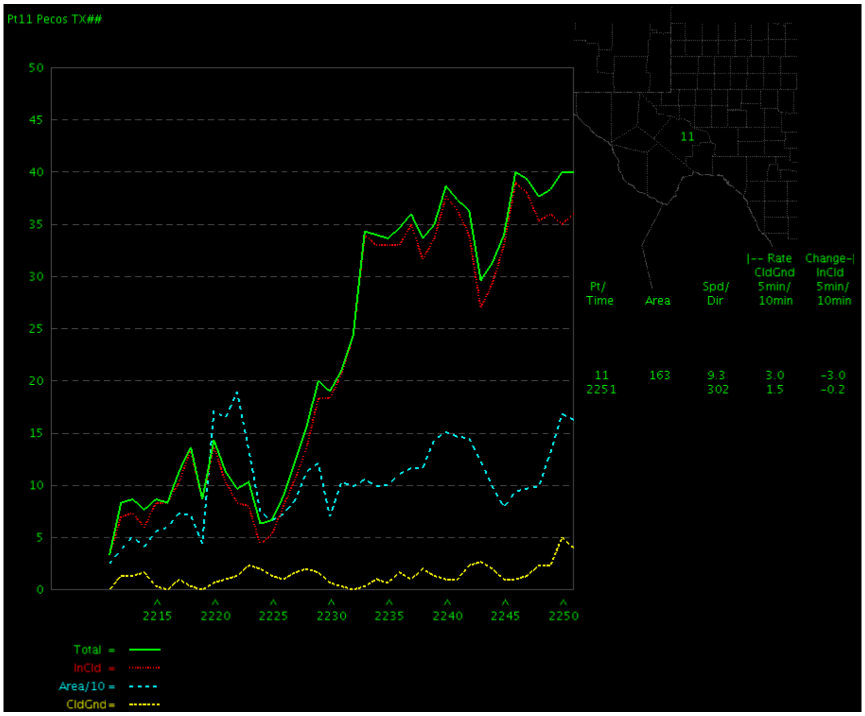

Figure 8: Time series of ENTLN total lightning product.

2) Probabilistic Hazards Information Experiment

For the first time, this year's Hazardous Weather Testbed (HWT) featured the simultaneous and synergistic creation and evaluation of probabilistic tornado, wind, and hail forecasts with National Weather Service (NWS) forecasters and emergency managers. These forecasts belong to the concept of probabilistic hazard information, a key feature of FACETs, aiming to give forecasters the ability to convey forecast uncertainty in a continuous manner while providing key decision makers with location-specific information about the timing and likelihood of impending severe weather events. To help forecasters, rapidly-updating automated probabilistic guidance, such as NOAA's ProbSevere model, was made available within a prototype tool for creating forecasts of probabilistic hazard information on short time scales (0-2 hour lead times). Likewise, emergency managers used a prototype interface for viewing real-time forecasts of probabilistic hazard information to assist in making simulated decisions (Figure 9). With these features in mind, the following list of motivating questions helped guide the experiment design:

- In what ways can forecasters utilize automated guidance in generating forecasts of probabilistic hazard information?

- What are the strengths and weaknesses of each approach to utilizing the automated guidance?
- What tools are necessary to effectively combine the best human forecaster and machine abilities?
- What are the strengths and weaknesses of probabilistic hazard information forecasts for decision-making by emergency managers?

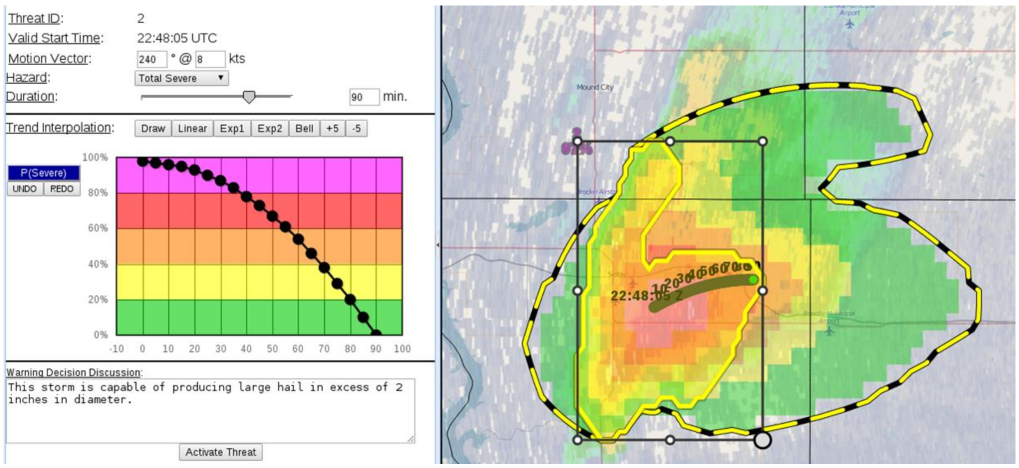

Figure 9: Probabilistic hazards information (PHI) tool.

The simultaneous experiments occurred for 3 weeks, every other week, during May and June of 2015. Each week, 2 forecasters and 3-4 emergency managers participated, while a diverse group of atmospheric scientists, human factors engineers, and social scientists observed and collected information. Debriefing discussions after each case or operational period enabled the rapid generation of ideas, grounded in the storms just experienced. Each week included a range of severe weather, from marginally severe storms to a complicated set of waves of short line segments of storms training over a metro area and a high-end tornado case. Although each week of the project had unique participants, each week iterated toward the same types of information to include in the discussion box associated with each forecast of probabilistic hazard information: storm history and forecaster thinking.

From the experiment observations, four levels of forecaster utilization of automated guidance were identified:

- Level 1: Forecasters generate all forecasts of probabilistic hazard information; access to automated guidance is disabled
- Level 2: Forecasters optionally use automated guidance; automated guidance is running but can be overridden.

- Level 3: Forecasters partially override automation; automated guidance is running but mechanical aspects of guidance cannot be overridden
- Level 4: Forecasters observe automatic generation of probabilistic hazard information; automated guidance is running unabated

The level 1 approach yielded results similar to those published from the 2014 experiment (with forecasters only). The concept of probabilistic hazard information removes the notion that a storm must achieve some level of intensity before issuing information (i.e., warning product). Thus, any storm with the potential for producing severe weather is a candidate for forecaster generation. However, monitoring and updating multiple (approximately 4-5+) hazard areas simultaneously becomes problematic from a workload perspective. Forecasters were observed to fall back on WarnGEN (current software using for generating warnings by the NWS) approaches (e.g., simplifying and stacking object geometries as parallelograms) when workload become problematic. Emergency managers noticed when objects became large, or when storm objects were combined, strongly preferring the more refined, precise location information that guidance to forecasters provided; they understood there was forecast uncertainty and that storms could abruptly shift direction. In a case of a larger-than-needed tornado forecast passing close to the town an EM was assigned to, he reported that he relied on radar instead to make decisions related to spotter placement and emergency response. It is important to note that for tornado hazards, the level 1 approach usually worked well, given the simple and consistent shape (circle or ellipse) typical of these hazard areas and the rareness of having to monitor and update 4-5+ tornado hazard areas simultaneously for a given NWS County Warning Area (CWA).

On the other end of the spectrum, the level 4 approach frustrated forecasters and emergency managers alike. The forecasters wanted to be involved the process of issuing forecasts of probabilistic hazard information and were concerned about the performance of a completely automated system. Without human forecaster intervention, the emergency managers had a difficult time interpreting the automated forecasts, and were relying on their training and intuition to interpret radar signatures themselves. Emergency managers articulated that without the additional information such as storm history and radar interpretation that forecasters typically provide, they were left trying to figure out on their own what the storm was capable of. Emergency managers also shared forecasters' wariness of automated information, strongly preferring to know that a forecaster had either changed the information to make it more accurate, or that the forecaster agreed with the guidance.

Initially, the level 2 option was the only option presented to forecasters for interacting with the automated guidance. Tools such as a slider bar were created to give forecasters the ability to mask automated guidance below or above a set probability threshold. Additionally, forecasters could click on the map to allow, block, or modify the automated guidance (Figure 10). Temporally-varying aspects of the automated guidance that could be modified included the position and shape of the automated object, as well as the speed, direction, duration, probability of occurrence, and discussion. In this level 2 approach, forecasters trended toward turning off the guidance as events unfolded. A forecaster stated, "We are taught in the Distance Learning Operations Course (DLOC) to be leery of algorithms," so perhaps there is an inherent distrust of such information. However, it was clear that the system design was also a factor. Forecasters would commonly create and issue a forecast of probabilistic hazard information (i.e., level 1 approach) while the automated system was running and generating forecasts for the same hazard. This dynamic complication resulted in concurrent forecasts overlapping, therefore resulting in confusion by the emergency managers.

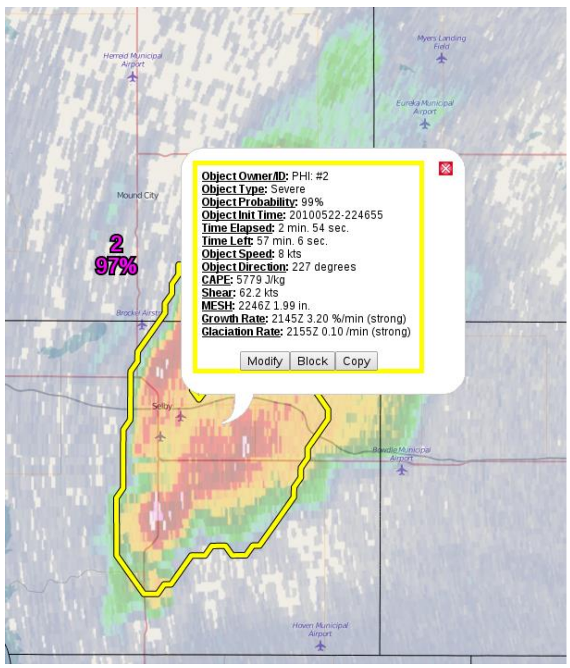

Figure 10 Object attributes table in the PHI tool.

Although the system design was identified as a limiting factor in the level 2 approach, this observation led to an important insight into how the system could be leveraged in a way to maximize the abilities of the human forecaster and the automated guidance. It was apparent that the rapidly-updating object positions and shapes from the automated system (i.e., ProbSevere) were fairly robust and often resembled those generated manually. Thus, a new approach (level 3) was created whereby forecasters could partially override specific aspects the automated guidance. More specifically, forecasters could not modify the object position or shape from the automated guidance, but could optionally override the object speed, direction, duration, probability of occurrence, and discussion.

The development of the level 3 approach gave us new and interesting results. Perhaps an obvious benefit is the significant amount of time that is saved in creating a forecast, due to leveraging the rapid generation and updating of objects encompassing the hazard areas. This time savings allowed forecasters to focus more intently on interrogating the storms and developing effective and ongoing threat communication, which emergency managers greatly appreciated. Through debriefings each week it became clear that emergency managers cannot devote consistent attention to weather, and thus rely on clear, meteorological assessments from the NWS to help them and their constituents make decisions. Thus, in level 3, forecasters were developing a better understanding the hazard potential and serving key stakeholder needs.

Finally, it was observed that forecasters commonly raised and lowered the diagnostic probability values recommended to the forecaster from the automated guidance. More specifically, areas receiving severe weather reports often contained high diagnostic probability values (near 100%), and the opposite true in areas devoid of severe weather reports. Questions were raised from both the forecasters and emergency managers about why non-binary values of probability were provided for events when clearly nothing or something was happening. For example, a live stream video showed a tornado occurring, while the automated guidance probability value was at 70%. Such discussion helped us refine the difference in perspectives between forecasters and users, which serves as a challenge for guiding future research efforts:

- <u>Forecaster Perspective</u>: Given **a storm** of x,y,z characteristics, what is the **probability** that it will produce severe weather?
- <u>User Perspective</u>: Given **my location**, will severe weather **affect me or not**?

Both groups found the project to have great potential for operations. Forecasters appreciated our researcher team's careful approach to co-creating a useful tool for them and their users. Both groups favored the greater and more useful information stream to the structure of traditional warnings. Many issues were discussed that would need to be addressed before implementation, such as how to derive legacy warnings. By providing only probabilities, emergency managers quickly realized they would choose to act on much lower probabilities on days showing potential for high-impact severe weather as opposed to a marginally severe event. Both groups greatly valued the ability to connect the warning desk to the key decision maker, gaining great insights into the others' roles and considerations.

3) Hydrology Experiment

During the experiment, National Weather Service and River Forecast Center forecasters worked with research scientists to assess emerging hydrometeorological concepts and products to improve the accuracy, timing, and specificity of flash flood watches and warnings. In particular, forecasters evaluated short-term predictive tools derived from MRMS quantitative precipitation estimates (QPE) and Flooded Locations and Simulated Hydrographs (FLASH, Figure 11) hydrologic modeling framework. The Hydro Experiment also explored the utility of experimental watch and warning products conveying uncertainty and magnitude issued through the Hazard Services software. This allowed research scientists to investigate human factors to determine operationally relevant best practices for the warning decision making process and the system usability of the Hazard Services platform.

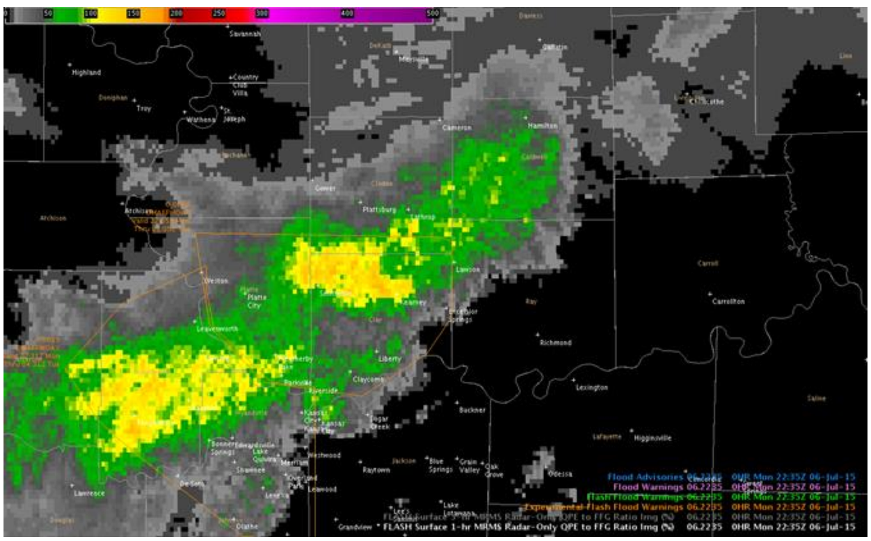

Figure 11: FLASH model output in AWIPS-2.

The principal investigators (PIs) of the Hydrology Experiment collaborated with the PIs of the Flash Flood and Intense Rainfall Experiment. FFaIR simulated a national center with quantitative probability forecast / flash flood probabilities and discussions while the HMT-Hydro forecasters simulated a WFO in watch / warning operations. The previous day's forecast was discussed between the two testbeds, including a comparison of HMT-Hydro operations and local storm reports. Additionally, the Hydrology Experiment PIs leveraged in-house collaboration efforts within the National Weather Center including Severe Hazards Analysis and Verification Experiment (SHAVE) and Meteorological Phenomenon Identification Near the Ground (mPING). Daily activities included evaluating FFaIR briefings, subjective evaluation of products and experimental watches and warnings, and real-time forecast activities.

Goals:

- Evaluate skills of experimental flash flood monitoring and short-term predictive tools from FLASH suite of products
- Evaluate FLASH products for development of flash flood recommenders
- Determine the benefit of extrapolated quantitative precipitation estimates and High Resolution Rapid Refresh 0-6 hour forecasts in FLASH suite for increased lead time
- Determine the utility in assigning probabilities (uncertainties) and magnitudes to experimental watches and warnings
- Evaluate potential use of Hazards Services using human factors research methodologies
- Enhance cross-testbed collaboration on short-term flash flood flash flood forecasting challenges

Results:

- It is challenging to estimate western US precipitation with MRMS and FLASH products

- Quantitative precipitation forecast – forecast CREST model can help generate lead time, but be aware of inconsistencies between model runs.

- First evaluation of Hazards Services shows that it has unique functionalities as well as some limitations.

- FLASH products with multi-radar / multi-sensor (MRMS) quantitative precipitation estimation (QPE) can quickly highlight areas with greater potential for flash flooding

- The FLASH products can also provide information that might lead to a decision on not to provide a flash flood warning

- Take the time to develop procedures using FLASH product to help with quick situational awareness and warning decision making

4) Phased Array Radar Innovative Sensing Experiment

During the experiment, National Weather Service forecasters from the Great Plains utilized phased array radar data in displaced real-time warning operations. The 2015 PARISE featured three parts: 1) the traditional experiment, 2) an eye-tracking experiment, and 3) a focus group. While the traditional experiment built on knowledge obtained from previous experiments, the eye-tracking experiment brought a new and exciting avenue to the work of PARISE. Forecasters' eye gaze data was collected as they work a case in simulated real-time (Figure 12). This data provided new insight into impacts of higher-temporal resolution on the forecaster warning decision process and allowed us to analyze and compare forecasters' cognitive processes objectively. Finally, the focus group session drew on participants' experiences throughout the whole week, and I generating insightful feedback and ideas important to the development of a future PAR network.

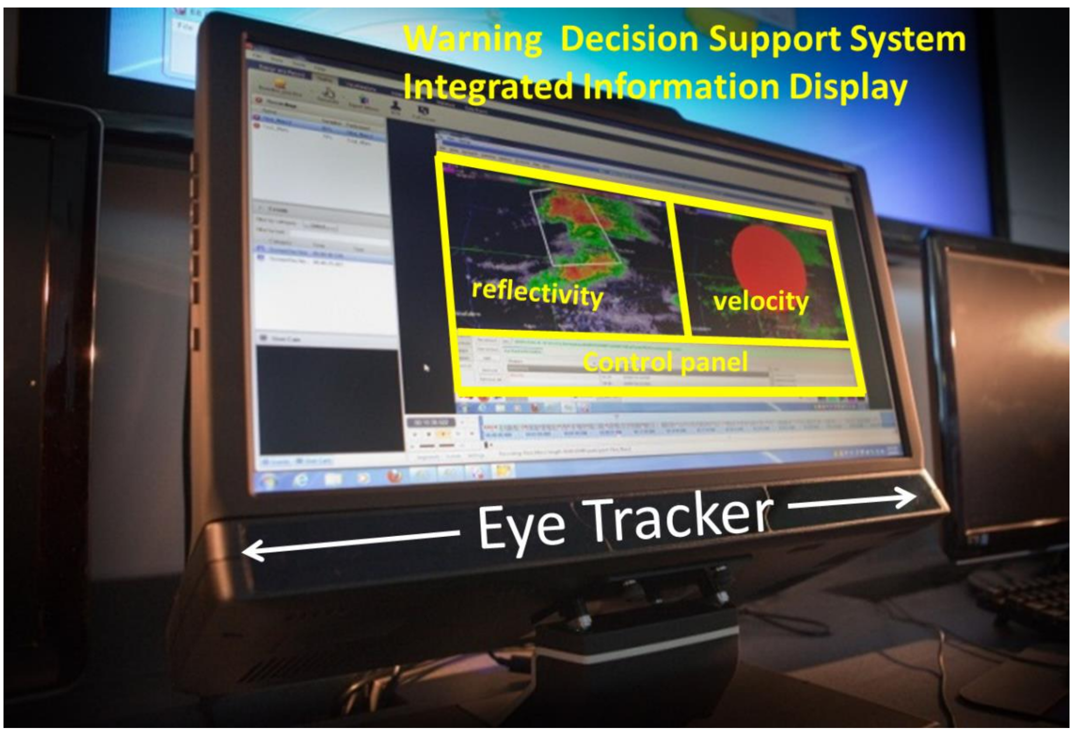

Figure 12: Eye-tracking technology used in PARISE 2015.

Goals:

- Investigate whether the benefits of higher-temporal resolution radar data are evident with an increased sample size of participants and number of cases work

- Obtain a deeper understanding of forecaster's cognitive processes as they interact with different temporal resolution of radar data
- Understand how temporal resolution impacts forecasters' cognitive workload and if and when data overload exists
- Incorporate eye-tracking technology to further develop our understanding of forecasters' cognitive processes and workload issues.

Results:

- Eye-tracking methods found to be a viable method for developing our understanding of forecaster's cognitive processes

- Used focus group to work through a set of predetermined questions. Helped to generate new ideas and give feedback on experiment design.

4. PUBLICATIONS

Bowden, K. A., P. L. Heinselman, D. M. Kingfield, and R. P. Thomas, 2015: Impacts of phased-array radar data on forecaster performance during severe hail and wind events. *Wea. Forecasting*, **30**, 389 – 404.

Heinselman, P., D. LaDue, D. Kingfield, and R. Hoffman, 2015: Tornado warning decisions using phased-array radar data. *Wea. Forecasting*, **30**, 57 – 78.

Karstens, C. D., G. Stumpf, C. Ling, L. Hua, D. Kingfield, T. M. Smith, J. Correia, Jr., K. M. Calhoun, K. L. Ortega, C. J. Melick, and L. P. Rothfusz, 2015: Evaluation of a probabilistic forecasting methodology for severe convective weather in the 2014 Hazardous Weather Testbed. *Wea. Forecasting*, **30**, 1551-1570.

5. PERSONNEL

EWP Officers

Gabe Garfield	
Operations Coordinator	gabriel.garfield@noaa.gov
Darrel Kingfield	
Information Technology Coordinator	darrel.kingfield@noaa.gov
Tiffany Meyer	
Information Technology Coordinator	tiffany.meyer@noaa.gov
Alan Gerard	
NSSL HWT Liaison	lans.rothfusz@noaa.gov
Travis Smith	
EWP Co-Manager	travis.smith@noaa.gov
David Andra	
EWP Co-Manager	david.andra@noaa.gov
Lans Rothfusz	
EWP Co-Manager	lans.rothfusz@noaa.gov

6. ACKNOWLEDGMENTS

EWP2015 wouldn't have been possible without contributions from a number of individuals and organizations. Those organizations include the Cooperative Institute for Mesoscale Meteorological Studies, the National Severe Storms Laboratory, the GOES-R Program Office, the Meteorological Development Laboratory, and the National Weather Service Forecast Office in Norman, Oklahoma.

We would like to acknowledge the contributions of the following individuals: Katie Bowden, Kristin Calhoun, John Cintineo, Race Clark, Karen Cooper, Jimmy Correia, Vicki Farmer, Zac Flamig, Aimee Franklin, J.J. Gourley, Pam Heinselman, Chris Jewett, Darrel Kingfield, Chris Karstens, Bill Line, Chen Ling, Daphne LaDue, Steve Martinaitis, John Mecicalksi, Tiffany Meyer, Brian Motta, Lans Rothfusz, Elise Schultz, Justin Sieglaff, Travis Smith, Geoffrey Stano, Sarah Stough, Bret Williams, and others. And we'd also like to give special thanks to James Murnan and Keli Pirtle, who managed the media associated with the experiment.

This work has been primarily funded via sources from the National Severe Storms Laboratory, the National Weather Service Meteorological Development Laboratory, Earth Networks Incorporated, the U.S. Weather Research Program, Multi-Function Phased Array Radar, and via the NOAA-University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce. Also, we would like to thank the National Weather Service regional directors, who sponsored an additional 5 forecasters.