Meeting Report
Hydrometeorology Testbed and Extreme Precipitation Forecasting Improvement

October 15, 2019

National Center for Weather and Climate Prediction, College Park, MD

Final Report 27 May 2020

The NOAA Hydrometeorological Testbed (HMT) is a joint OAR-NWS testbed motivated to make communities that are more resilient to the impacts of extreme precipitation on lives, property, water supply and ecosystems. HMT is co-managed by the NWS Weather Prediction Center (WPC) and the OAR Physical Sciences Laboratory (PSL) in partnership with the NWS Office of Water Prediction (OWP).

The mission of HMT is "Improving forecasts of extreme precipitation and forcings for hydrologic prediction."

Hydromet Testbed Executive Oversight Council:

David Novak, Director, NWS Weather Prediction Center (WPC)
Robert S. Webb, Director, OAR/ESRL Physical Sciences Laboratory (PSL)
Ed Clark, Director, NWS National Water Center (NWC)

Report writing team:

Andrea J. Ray, PSL HMT Coordinator
James Correia, WPC HMT Coordinator
James Nelson, Development and Training Branch Chief, WPC

Acknowledgements: Barbara DeLuisi, PSL, report cover design, and Lisa Darby for comments on the report.

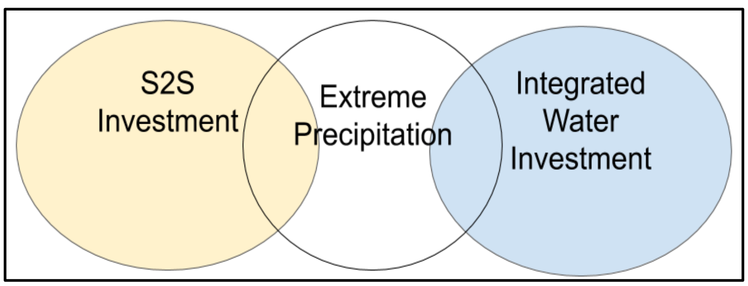
Cover photo credits:
Snow plow: USAF
Flooded street: USGS
Flood in Denham Springs, LA: DOD
People pushing car: DLA

Executive Summary

The Nation has experienced increasing devastation from heavy precipitation events recently. In just the past 3 years, 13 precipitation-related billion-dollar disasters in the Nation have resulted in over 200 deaths. This trend has dramatically increased the demand and expectations from core decision makers for accurate, consistent, and understandable rainfall forecasts. Heavy precipitation and resulting flash flooding occur across the year with seasonal and geographic variations. The predictability of these events varies with event type, region, and season.

Several ongoing NOAA efforts might aid in improving forecasts of extreme precipitation, however, precipitation forecasting from minutes to 10 days is not a focus among these efforts. These include the Earth Prediction Innovation Center (EPIC), the NOAA Water Initiative, and recent investments to improve prediction on subseasonal to seasonal (S2S) time scales. Therefore, there is a need for significant investment for precipitation in the weather time scale.

Motivated by the need to improve forecast skill and threat communication, the NOAA Hydrometeorological Testbed (HMT) organized a workshop with diverse participation of NOAA operational and research partners. Presentations illustrated key operational challenges and emerging new science and technology. Workshop participants discussed next steps to meet these goals, high priority challenges, and metrics for tracking improvements. Participants recommended initiating a planning process for an extreme precipitation forecasting improvement project, with the investments in this area benefitting S2S and water prediction as well.



Investments in extreme precipitation prediction will have benefits for integrated water and subseasonal to seasonal (S2S) prediction.

High-priority hydrometeorology challenges identified:

- Improve numerical model quantitative precipitation forecast (QPF) and streamflow forecast skill, with reliability (the ability to give unbiased probability estimates) and sharpness (the ability to forecast extreme values)
- Create continuum of decision support services (DSS) for extreme rainfall from weather to S2S for more effective risk communication to core partners, including water resource management
- Improve risk communication through the cross disciplinary engagement between physical and social sciences
- Establish the ability to disentangle errors from coupled meteorological and hydrological models

Candidate metrics identified:

- Double the rate of improvement of numerical weather prediction (NWP) precipitation skill, from the current rate of 1% a year to 2% a year using a baseline of the Global Forecast System (GFS) precipitation skill.

- Extend extreme rainfall services to Day 10.
- Improve the reliability and sharpness of probabilistic information for extreme events.
- Narrow the 5-day landfalling position of atmospheric rivers on the west coast and Alaska from 500 km.
- Improve the skill of atmospheric forcings provided to the National Water Model (NWM) and other hydrological models.

Two overarching approaches were identified to address these challenges:

- *Improve NWP forecasts of rainfall.* The skill of rainfall events has improved at a slower rate of improvement than even hurricane intensity skill. The pace of improvement in NWP rainfall forecasts needs to accelerate, including improved calibration of probabilistic forecasts (reliable and sharp distributions).
- *Narrow the gap between what we know (full model data/post processing/scenarios) and what we tell (quality and quantity of DSS)*. A first step is to make use of good information not yet being used by NWS forecasters, and thus not available to partners and the public. Addressing this involves both the physical and social sciences, including mining critical information from ensembles for forecasters to use, visualizing such information, as well as product design, and communication best practices.

The group recognized the need to continue discussions and called for an effort to develop an HFIP-like improvement plan **for a program to improve predictions of extreme precipitation, and provide better forcings for flood predictions, and more effective watches and warnings.** This program would encompass extreme precipitation and its impacts with timescales ranging from minutes to 10 days, as well as the forcings necessary to meet the challenges of hydrologic modeling. This program would be a NOAA-focussed interagency effort within the US and western North America that advances the science and makes improvements in predictions in these time scales (sub-daily to 10 days). Achieving these goals will require leveraging efforts of the Next Generation Global Prediction System (NGGPS), Unified Forecast System (UFS), and the Earth Prediction Innovation Center (EPIC). **Such an effort will directly contribute to advancements needed to address the Precipitation Grand Challenge, including providing better set up for S2S precipitation outlooks**. The group supported continuing current HMT experiments including the Winter Weather and Flash Flood and Intense Rainfall Experiments (WPC-HMT 2019a,b; Erickson et al. 2019; Barthold et al. 2015). Investment in extreme precipitation will have benefits across Climate Weather, and Water.

Near term activities towards such a program include:

- Contribute to the Precipitation Grand Challenge effort.
- Develop research questions to address both near term and long-term forecast challenges.
- Contribute to the UFS effort, in particular, the forecast goals and science priorities related to UFS applications for medium-range weather (atmospheric behavior out to about two weeks) and S2S (atmospheric and ocean behavior from about two weeks to about one year)
- Conduct near real-time investigations of high impact extreme hydrologic events.
- Begin a more comprehensive process to more fully lay out science and implementation plans for an extreme precipitation improvement program.

1. Introduction

The Nation has experienced increasing devastation from heavy rainfall events recently. In just the past 3 years, there have been 13 separate billion-dollar disasters in the nation, resulting in over 200 deaths. This trend has dramatically increased the demand and expectations from core decision makers for accurate, consistent, and understandable rainfall forecasts. Heavy precipitation and resulting flash flooding occur throughout the year with seasonal and geographic variations (Fig. 1). The predictability of these events varies with event type, region, and season (Fig. 2).

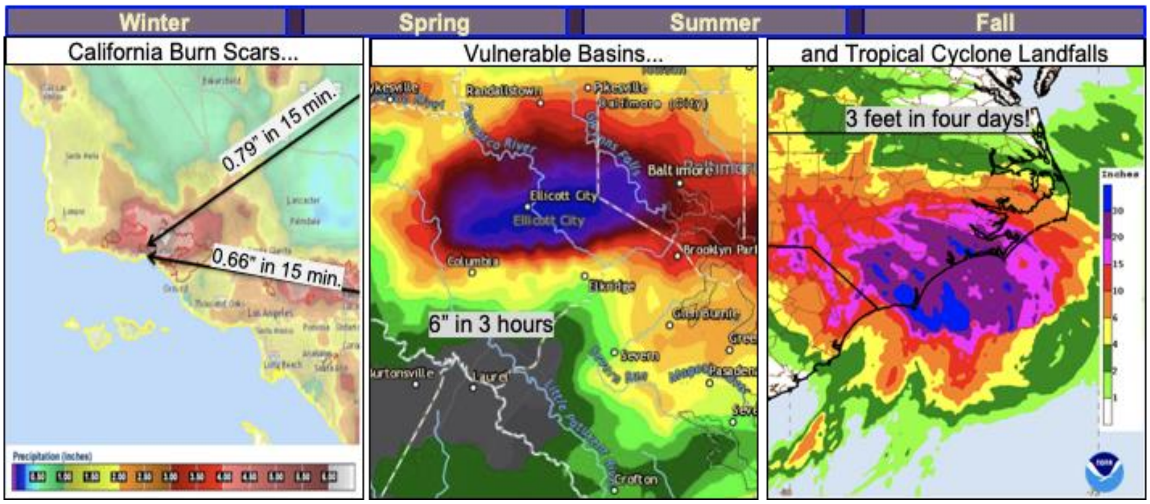


Figure 1. Examples of heavy precipitation events spanning diverse geographical areas and across seasons. (Novak and Carbin presentations)

Extreme precipitation spans the Weather - Water - Climate interface, and is key to the Precipitation Prediction Grand Challenge that has recently been established by the NOAA Climate, Weather, and Water Board. The set-up for extreme precipitation risk spans seasons to minutes (Fig. 3), providing a connection to other aspects of climate and water. Furthermore, systematic model biases start in the weather time-frame (first hour of the forecast). Therefore, improving near-term errors and biases in the first 10 days will cascade to improvements in the S2S timeframe (weather out to seasonal timescale).

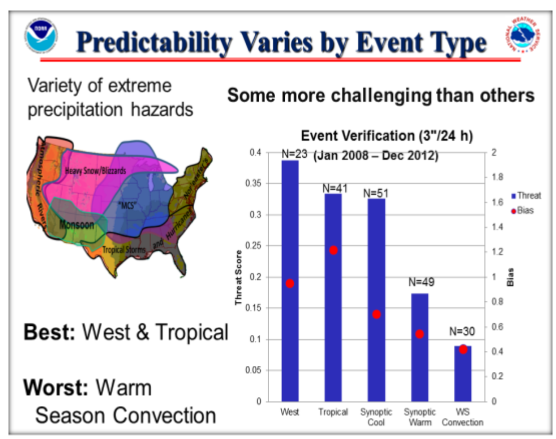


Figure 2. Predictability by event type.

Furthermore, with the advent of Impact-Based Decision Support Services (IDSS, https://www.weather.gov/about/idss) to core partners, there is a need for clear and effective communication in heavy rainfall events. Finally, to translate the precipitation into flood impact information, there is a need for both improved hydrology and atmospheric forcings as well as social sciences to inform product needs, design, and communication. Advances in hydrologic information will plateau if they are not matched by improvements in precipitation forcings.

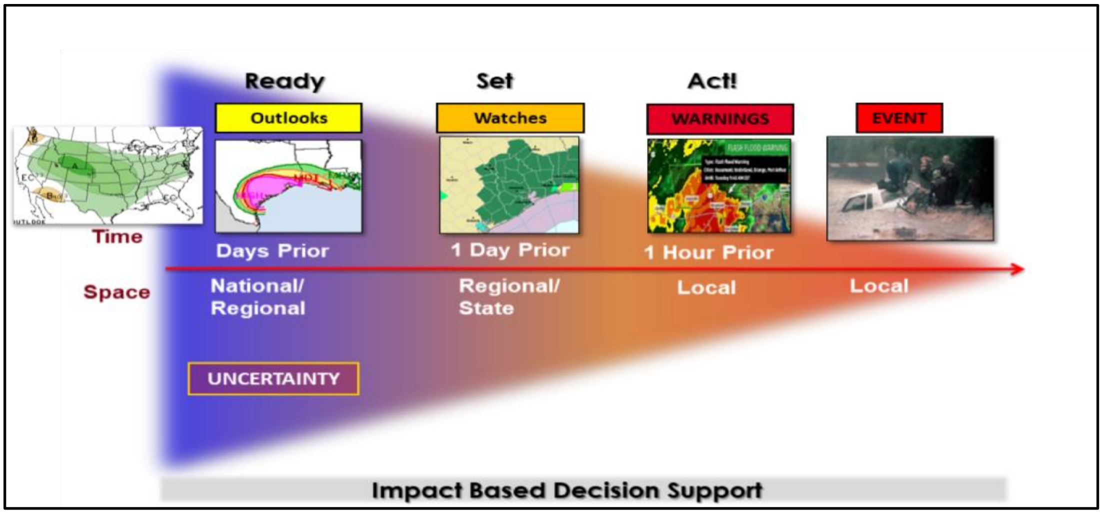


Figure 3, Risk of extreme precipitation sets up days to seasons prior to the event. From Introduction presentation by Novak.

Given the need for improvements in forecast skill and threat communication, the NOAA Hydrometeorological Testbed (HMT) organized a workshop with diverse participation of NOAA operational and research partners. The workshop was designed to identify operational forecast challenges in the Day 0 - 10 timeframe, with specific goals to:

- Review science goals for the HMT
- Discuss the state of the science and operational needs for decision support
- Foster ongoing integration among operations & research partners
- Explore the concept for a focused improvement program on extreme precipitation forecasting and modernizing flood forecasting, and towards that goal, identify a small set of high-priority hydrometeorology challenges and candidate goals and metrics

To foster discussion, the workshop was organized around brief talks illustrating key operational challenges and emerging new science and technology. This report uses key figures from these presentations to illustrate the current problems found in forecasting precipitation. Organizers intentionally invited a diverse group of participants from NWS operational and regional offices, NWS headquarters and policy offices, and several NOAA research labs. Invited presentations and panel discussions focused on types of forecast challenges, operational forecasting perspectives, and emerging science. Much of the afternoon was devoted to discussion on themes and foci to advance science and forecasting of extreme precipitation. A list of attendees is provided in Appendix A, along with links to the agenda and presentations, discussion notes (Apdx. B), background on the HMT (Apdx. C), HMT-relevant metrics tracked by NOAA (Apdx. D), and a list of acronyms (Apdx. E).

At the end of this document we propose a planning process for an extreme precipitation forecasting improvement project, with the investments in this area benefitting S2S and water prediction as well.

2. Key Challenges For Forecasting Precipitation

NWS forecasters have been leveraging NWP to provide rainfall forecasts for many decades. As weather modeling improved, so have the forecasts, but recently this pace has slowed (Fig 4). The Day 1 Equitable Threat Scores (ETS) for 1" QPF have remained around 0.34 and the 2" QPF has remained around 0.25 for the past 10 years. The trends for 2" QPF are on pace to match the skill for Day 1 from a decade ago. Similarly, for current Day 3, the ETS nearly matches the skill of Day 2 from a decade ago. These metrics are tracked by NOAA, see Appendix D for HMT-relevant metrics. In his presentation, WPC Director David Novak noted that fine-scale precipitation extremes remain a challenge, and while forecasters still add value over model guidance, the gap between forecasters and models (specifically ensembles) is closing.

The NWS Global Forecast System (GFS) skill of 1" rainfall events has improved ~15% over the past 16 years (Fig. 4), a slower rate of improvement than even hurricane intensity skill. Skill also varies by event type and by region of the country (Fig. 5, and Sukovich et al. 2014). For a primer on verification metrics for probabilistic forecasts, including reliability, discrimination, and sharpness, see Hudson (2017) and the NWS Glossary of Verification Metrics (undated).

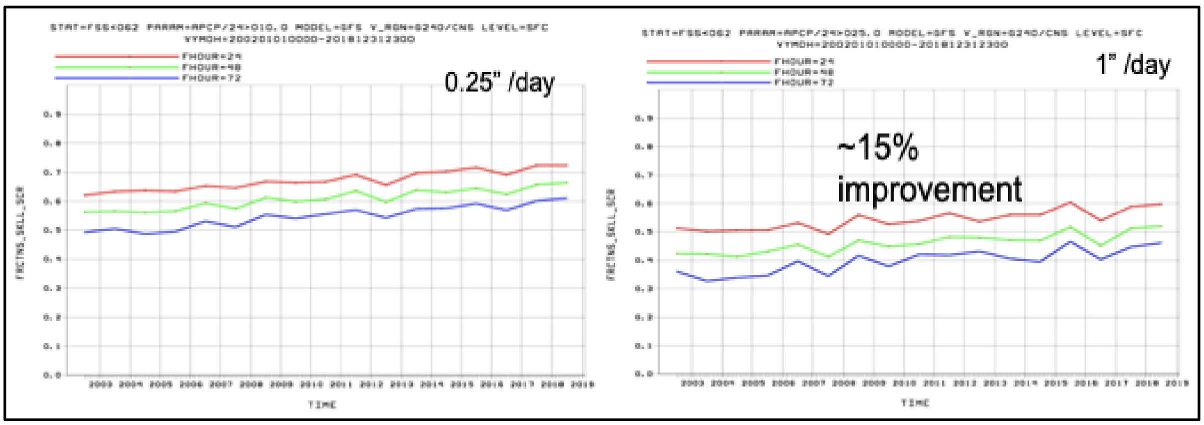


Figure 4. Global Forecast System (GFS) 24/48/72h annual FSS, 2003-2019. The rate of improvement of GFS precipitation skill over this 16 year period is ~15%. This lags improvements in other phenomena, including hurricane track (30-50% improvement) and 500mb Anomaly Correlation (~20%), and is on par with Hurricane Intensity (~15%

Mesoscale convective systems in the Great Plains in the summer pose a particular challenge, as documented by Fritsch and Carbone (2004). Then - and still - the prediction of summer rainfall is the Achilles' heel of weather prediction – particularly how to reliably forecast deep, moist convection. Finer resolution models may handle these systems better, as they can more explicitly represent the convective features, but the spatio-temporal accuracy is much lower than other precipitation features - even at short lead times. Participants discussed the variations in predictability by event type, region, and season, citing the work of Sukovich et al. (2014), in particular regarding the Missouri River basin (Fig. 5).

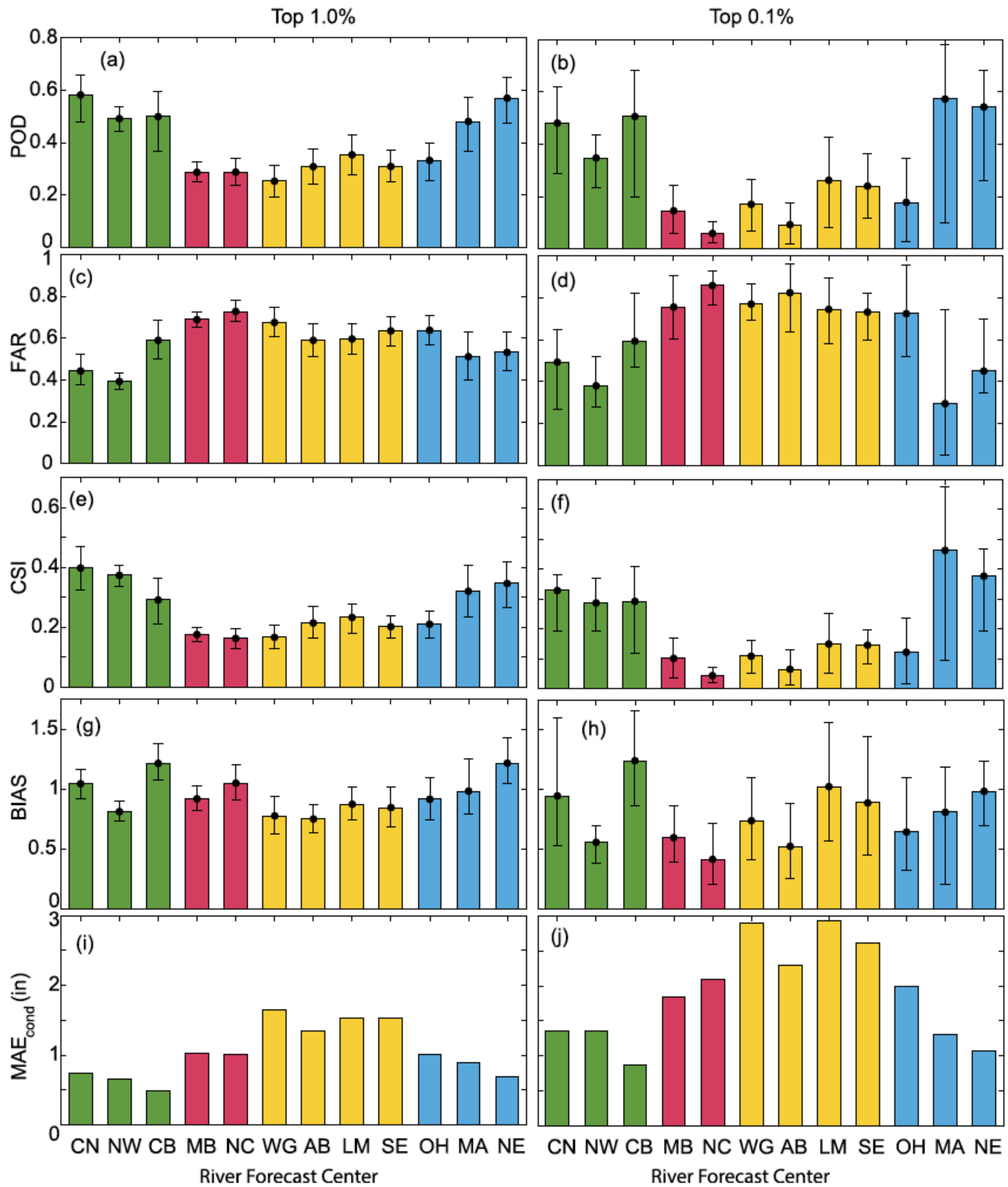


Figure 5. Regional (a) probability of detection (POD), (c) false alarm ratio (FAR), (e) CSI, (g) bias, and (i) MAE_{cond} values by River Forecast Center (RFC) aggregated over 2007–11 for the top 1.0% of precipitation events using the regional precipitation thresholds in Table 2 of Sukovich et al. (2014). Regional (b) POD, (d) FAR, (f) CSI, (h) bias, and (j) MAE_{cond} values by RFC aggregated over 2007–11 for the top 0.1% of precipitation events using the some regional precipitation thresholds. Bar graphs are color coded into four broad U.S. geo-graphical regions: West (green), upper Midwest (red), South/Southeast (yellow), and East/Northeast (blue). Brackets indicate 95% confidence intervals for the skill scores and bias. Table 9 in Sukovich et al. (2014).

New ways for forecasters to contextualize the information from QPF and outlooks may ultimately lead to improvements to the end hydrologic impact forecast. In the theme of contextualizing information, continuing to extract more useful and relevant information from our models is needed. For example, the precipitation rate over a 15-minute period or smaller, not currently output by our fine scale modeling systems, could help decision makers**.**

Novak discussed predictability challenges associated with winter storms near significant metropolitan areas. The sudden change in precipitation type (from rain to snow) can have a large effect on the public and core partners. In some cases, forecasts verified well and in others small errors in timing, location, duration, and intensity resulted in large snow amount errors. This highlights the sensitivity that comes with forecasting for large population areas in particular when there may be a more severe consequence for a low Probability of Detection (POD) than a high False Alarm Ratio (FAR).

Curtis Alexander of the Global Systems Lab described efforts funded in part by the USWRP/HMT and tested in the HMT Winter Weather Experiment. These efforts are making enhancements to the RAP-HRRR model suite and are intended to expand on the current RAP/HRRR winter forecast fields to include Variable-Density Snow Accumulation and Surface Precipitation Type Estimation (Figure 6).

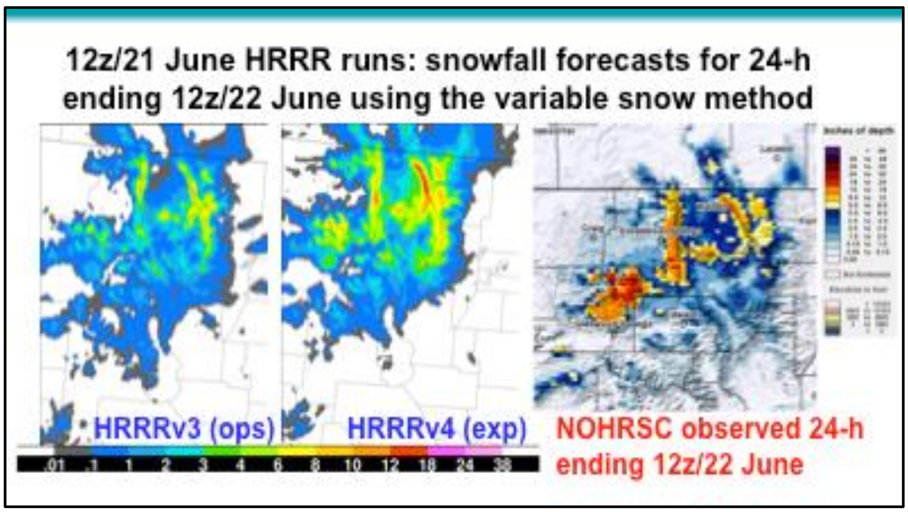


Figure 6. HRRR comparison of snowfall amount between v3 (left) and v4 (middle) compared to NOHRSC snowfall analysis (right), where v4 uses a variable density snow-to-liquid ratio which varies by lowest model level temperature (Alexander presentation).

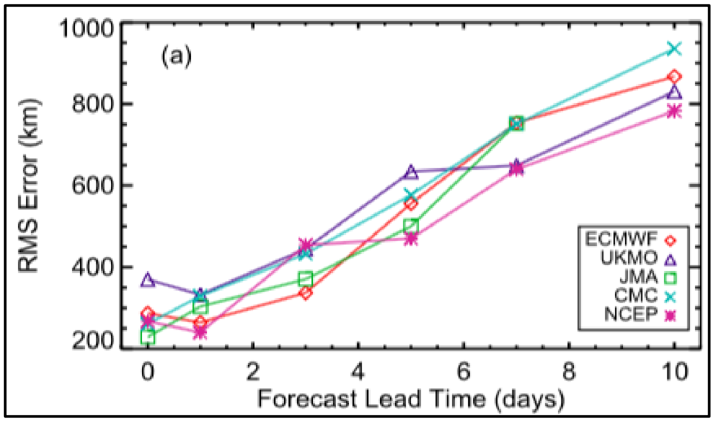


Figure 7. AR landfalling RMS (root mean square in kilometers) by forecast lead time in days. From Wick et al. (2013), fig 11.

Atmospheric rivers are an important source of precipitation on the west coast of the US and Alaska. They often result in flooding causing an average of $1.1 Billion in damages annually on the west coast (San Francisco Chronicle, 2019). However, the average GEFS AR landfall error in the 5-day forecast is ~500 km (Fig 7, and Wick et al. 2013), which is not sufficient for managing reservoirs to mitigate flooding and maximize water storage (Haynes presentation). Better forecasts from seasonal to individual atmospheric rivers would help local water management be better prepared.

Rowden echoed NWS deep core partners' need for improved precipitation and reservoir inflow forecasts in much of the west, where reservoir managers are often navigating a fine line between drought and flood operations.

Tropical systems provide a large amount of precipitation to the southern US coast along with thunderstorms. Lance Wood, the Science and Operations Officer at the WFO Houston/Galveston, described examples from several hurricanes, including Hurricane Harvey in Houston. The examples highlighted how forecast confidence with extreme accumulations over a large area contrasted with more recent Hurricane Imelda (Fig. 8) with lower forecast confidence and extreme accumulations over a smaller area. Workshop participants discussed whether products depicting specificity, like the High Resolution Ensemble Forecast (HREF) maximum rainfall, could lead to communication challenges for users trying to understand if their counties were at risk. Products depicting flooding risk like the Excessive Rainfall Outlook (ERO) may help communicate the risk more effectively. Key to either product approach is the recognition that extreme rainfall is best expressed probabilistically.

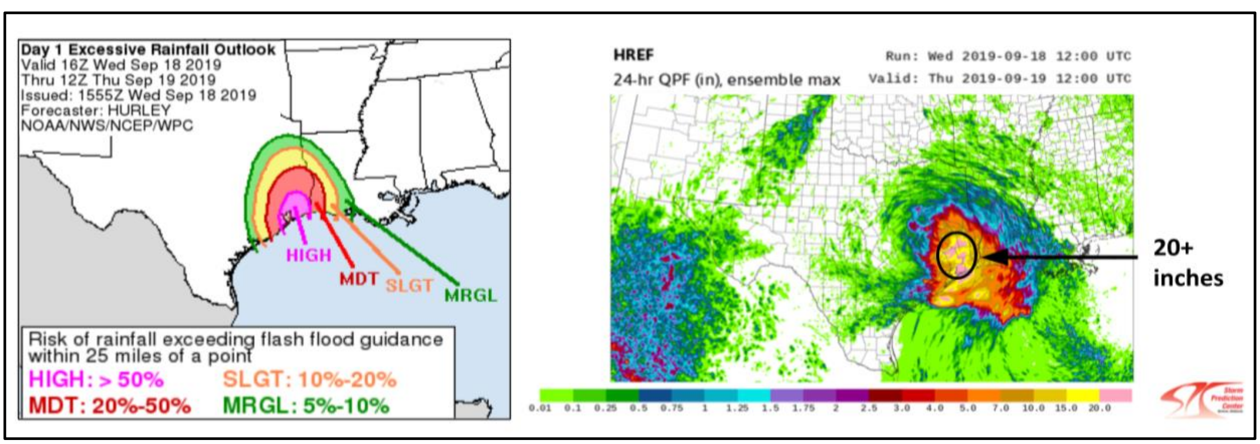


Figure 8. Successful collaboration with the WPC resulted in WFO Houston issuing an accurate Day 1 High risk area for Imelda(Wood presentation).

3. Key Challenges for Hydrological Forecasting

There is a need to improve ability to anticipate the most extreme flooding days, and yet forecasting these events remains a science challenge as well as a forecast communication challenge. Capturing risk of these events in national scale products has had mixed success as flash flooding is a small-scale phenomenon. Among the challenges is that heavy precipitation and resulting flash flooding occurs across the year with seasonal and geographic variations. Depending on the place, time of year, and other factors, heavy rainfall may or may not result in flooding. Forecasting flood events requires an appreciation of underlying hydrology (e.g. soil moisture, soil saturation, drainage characteristics), and understanding which basins are more vulnerable than others.

Improvements to flood and hydrologic impact forecasts as well as related IDSS will require understanding both the error characteristics from hydrology *and* meteorology and improving those modeling systems. In other words, while hydrology modelers may point to the "forcing" (e.g., QPF) for forecast error, and meteorologists may point to complex or black-box-like hydro models for forecast error, investments in both aspects of flood forecasting are needed. Forecasting flood events also requires an appreciation of underlying hydrology (e.g. soil moisture, soil saturation, drainage characteristics), because depending on the place, time of year, and other factors, heavy rainfall may result in different flooding characteristics. Thus, it is necessary to identify more vulnerable basins. Further, it is necessary to provide ways for forecasters to synthesize this information in the context of the QPF and outlooks, to enable the issuance of effective watches and warnings.

The Ellicott City, MD flooding events came up numerous times in the workshop. These events illustrate the challenge of predicting heavy precipitation events at sufficient temporal and spatial detail. The localized and extreme flood event of July 30, 2016 ensued from 6.6" of rain in three hours. WPC Forecast Operations Branch Chief Greg Carbin described how the risk was difficult to represent in an outlook product (a 'Slight' risk was in effect in the ERO product) because of the challenge of forecasting precipitation at fine spatial scales. Although a warning was issued 40 min in advance, the catastrophic nature of the event was not anticipated (Carbin presentation).

Kelly Mahoney of the OAR Physical Sciences Laboratory described a case study from Ellicott City flooding in 2018 (Viterbo et al. 2020). This study began as rapid-response, near-real-time investigation intended to complement ongoing work with OWP/NWC to disentangle forecast errors from NWM possibly resulting from precipitation forcing. In this case, the HRRR spatial error resulted in an underforecast of flooding for Ellicott City (Fig. 9). However, when precipitation objects were moved to account for the spatial error, flooding was over-forecast for this small basin. Additionally, using observations to drive the water model, they found that small basins were straightforward with predictions from rainfall to streamflow. However, in large basins, teasing out the root causes of errors was more difficult. In particular, Mahoney discussed how errors amplify: spatial displacement or temporal error in the precipitation forecast can result in a forecast bust from the hydrologic model.

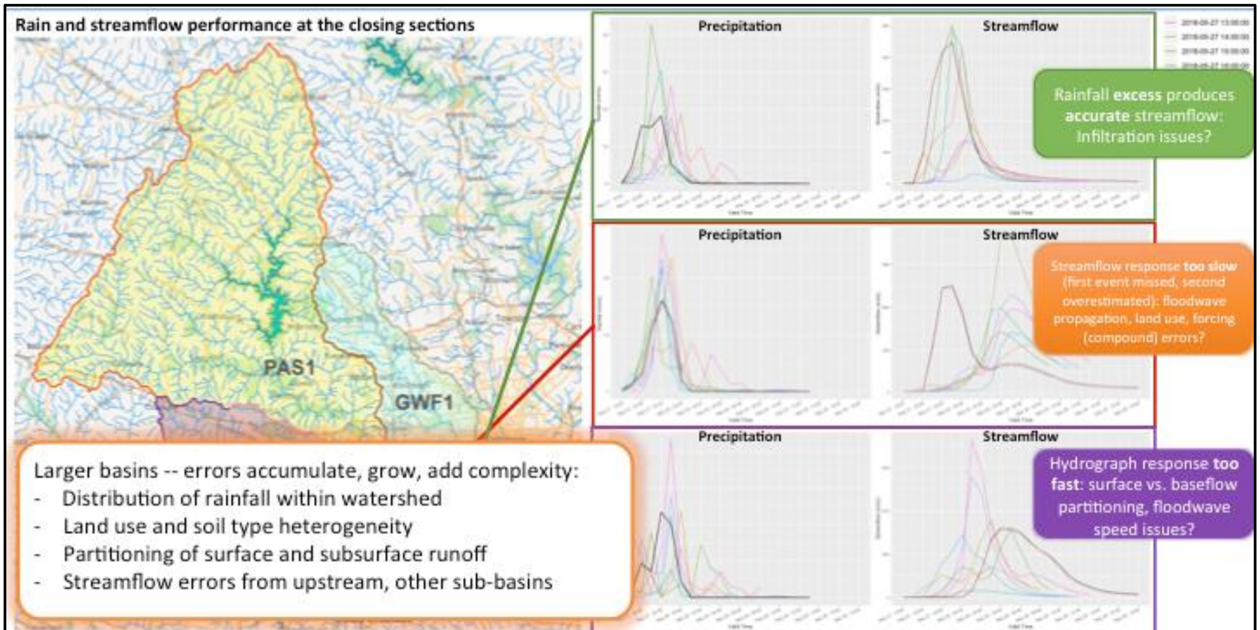


Figure 9. Watershed evaluation of the Ellicott City Flood. While errors accumulate in larger basins, Small basins show promise for NWM in capturing flooding signals. (Mahoney presentation)

The cross-timescale nature of some hydrologic events was illustrated by the various flooding events experienced in the Missouri River Basin in 2019. Wendy Pearson, the Hydrologist-in-Charge of the Missouri Basin River Forecast Center, described the set-up of high seasonal snowpack in the upper basin, then heavy rain on snow covered and frozen ground. As high flows from spring melting occurred, there were ice jams in many areas, and even excess sediment and river debris transport. The Army Corps Cold Regions Research and Engineering Laboratory (CRREL) has an ice break/jam model that provides approaches for predicting these jam events, but it needs to be simplified to be useful in real time river forecast operations. To better forecast the streamflow of the these S2S to storm scale events, RFCs need better forecasting of QPF and precipitation types, ways of handling differences between snow and rain, and better ways to deal with these in RFC hydrologic models.

The significant hazard of post-fire flooding and debris flows, especially in the west, was described by Katherine Rowden, the NWS Western Region Hydrology Program Manager. In debris flow cases, the instantaneous rainfall intensity, not accumulation, is what triggers initiation of debris flows. Rowden described an experimental product from NSSL which overlays estimated observed rainfall with debris flow thresholds used by WFOs for flash flood warnings, which are derived by the US Geological Survey (USGS) (Fig. 10), and is being tested in California and Colorado. Products like this may be useful in WFO's for

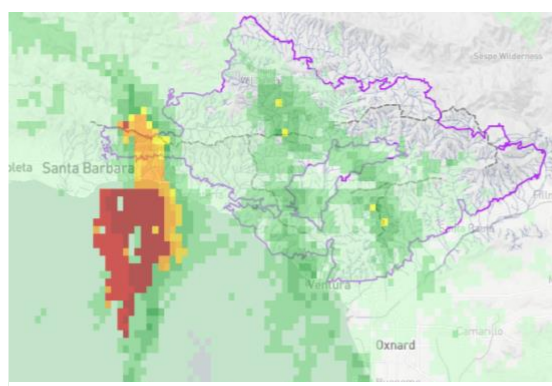


Figure 10. An experimental product from NSSL overlays mapped estimated rainfall over debris flow thresholds. (Rowden presentation)

IDSS, as discussed in a recent workshop on post-wildfire hydrology (NWS, 2019).

As the complexity of the water modeling suite increases so does the pressure to ensure that upstream modeling components are accurate and downstream models to be properly initialized, updated, and maintained. Model coupling will challenge our abilities to describe and discriminate between errors among the various sources of input data, as was demonstrated by Mahoney earlier.

To more fully deal with accuracy errors of NWP, ensemble hydrological modeling has been performed using the FLASH system, forced with precipitation forecasts from the Warn-on-Forecast ensemble system, described by J.J. Gourley of the OAR National Severe Storms Laboratory. In the discussion, probabilistic flooding information was seen as beneficial to the issuance of warnings with greater lead times since uncertainty was explicitly accounted for. Being able to quantify the uncertainty and make better short-range forecasts can be a game changer for short term IDSS, participants said.

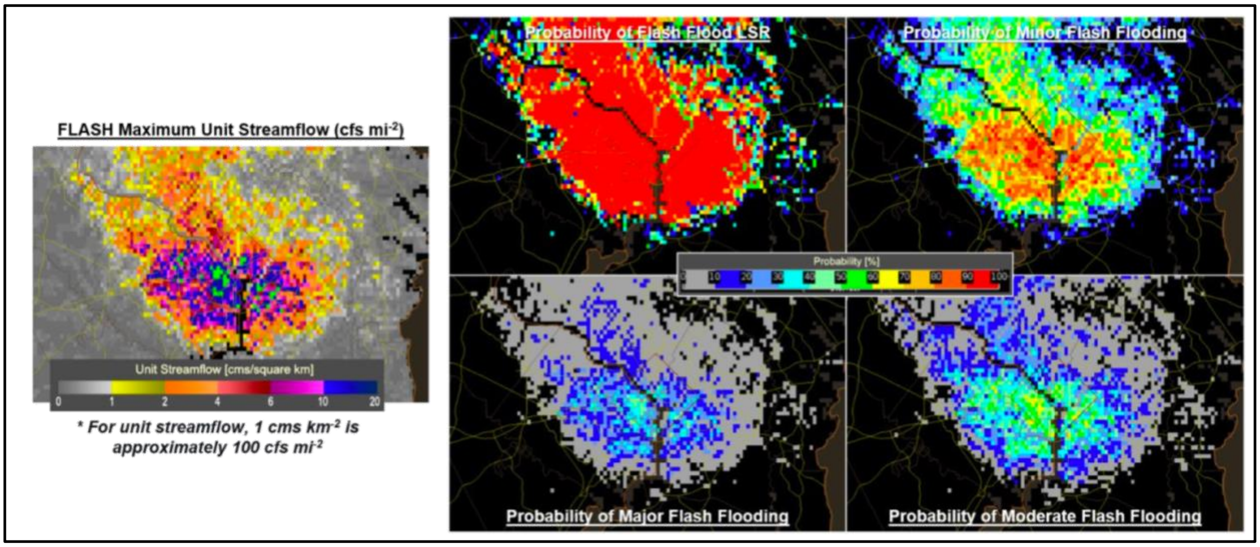


Figure 11. FLASH CREST maximum unit streamflow product showing an example of real-time outputs during Washington DC flash flood emergency (08 July 2019). Deterministic (left) and probabilistic (right) data at 1400 UTC 8 July 2019. Images taken from flash.ou.edu webpage, which was used during the project evaluations (Gourley presentation).

Atmospheric rivers present a hydrologic forecasting challenge as well as a precipitation challenge. Alan Haynes, the Hydrologist-in-Charge of the California-Nevada River Forecast Center, described the challenges including better hydrologic modeling of small, flashy basins, especially those with high vulnerability, including high population. Thus there is a need is for better pinpointing which basins will be hit by ARs at longer lead times for decision makers. Situations that result in full reservoirs with large cumulative releases, and saturated soils, while also being hit with multiple AR events, can overwhelm the flood control system and threaten levees protecting urban populations, such as those described in White et al. (2019). Reservoir managers need to coordinate and manage water releases from multiple reservoirs and for multiple purposes, and thus need both short and S2S range forecasts to make informed decisions, including those for fisheries and ecosystems. NOAA has been involved in an effort with

California Department of Water Resources to leverage improvements in weather and water forecasts to better manage reservoirs (Forecast Informed Reservoir Operations, FIRO). Based on a preliminary viability assessment of forecast use (Jasperse et al 2017), in particular forecasts of intense precipitation and atmospheric rivers, the Army Corps of Engineers approved a deviation in operations to further test the use of these forecasts in the 2019 rainy season (Coleman 2019). FIRO efforts provide a pathway for use of improved forecast skill in reservoir operations.

4. Key Challenges for Impact based Decision Support Services

A theme that arose often is that making good forecasts doesn't always translate to meeting user needs. User needs are better met by incorporating social and behavioral science assessments into the NWS IDSS. Equally necessary is helping forecasters acquire the skills and demands of the IDSS era. Impact-based Decision Support Services (IDSS) are being rolled out to core partners by the NWS. Partner demands and expectations are ever increasing. If we are to meet the communication challenges of being useful, usable and understandable with our products and services then so must our forecasts be timely and accurate for them to have value.

During a panel discussion, participants from NWS operations described some of the communications challenges, from the perspectives of the NWS Flash Flood Program (Kate Abshire), the Eastern Region Science Services Division (Jeff Waldstreicher), the Western Region Hydro Program (Katherine Rowden), and an WFO Science Operation Officer (Lance Wood). These include unprecedented heavy rains from tropical systems and snowfall amount and timing.

Participants discussed how using probabilistic language and probability better communicates our scientific conception of uncertainty but leaves much to be desired in communicating in a context for various publics to understand. When crafting IDSS messages, forecasters must distill ensemble forecast information into bite sized pieces of information. Lance Wood, the Science and Operations Officer of the WFO/Houston/Galveston, said that they find that the methods and tools NWS uses to construct summary information may be taken too literally or seem excessive depending on the kind of data presented. This point was echoed by Jeff Waldstreicher of the Eastern Region Science and Services Division: while forecasters have a sense of the uncertainty, that uncertainty may be poorly codified in a map or chart. In particular, it is a challenge to find the best ways to communicate the potential of an extreme rainfall event in the Day 4-7 forecast period, such as the period in which a tropical storm or hurricane threatens the coast, and in which

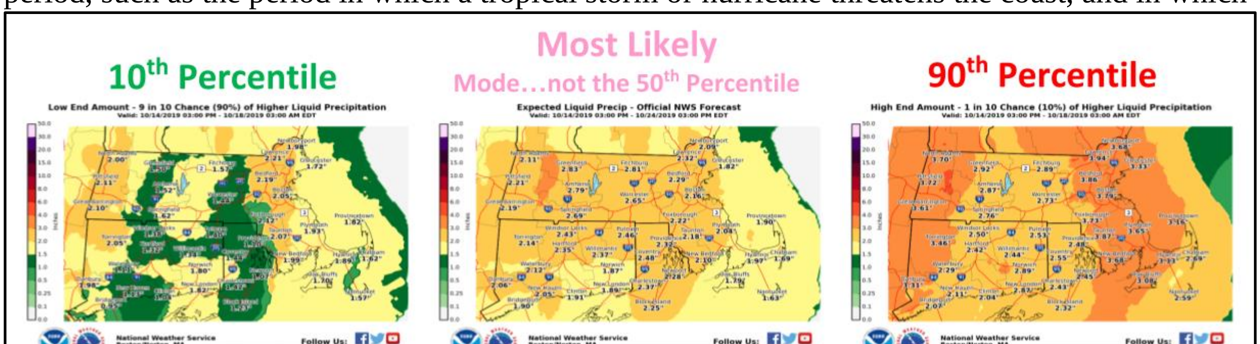


Figure 12. Communicating what is possible. Low end-Expected-High end maps being tested by the Eastern Region and elsewhere define the "goalposts" for an event (Waldstreicher presentation).

actions might be taken to mitigate its impacts. They are testing maps that provide a range of expected conditions (Fig. 12).

User understanding is increased when probabilistic information was accompanied by forecaster remarks (Karstens et al. 2018; ERG 2017, Carr et al. 2018, and Waldstreicher presentation). The role of the forecaster for decision makers includes being the Expert Interpreter, providing Contextual Information, relaying forecaster confidence and alternative scenarios from the perspective of the forecaster's expert knowledge. Contextual information may also improve decision making abilities for core partners.

Core partners need more nuanced information for particular areas like burn scars (e.g. 2014 Carleton Complex fire & debris flows, Rowden presentation), or highly developed urban areas now made vulnerable to flooding (e.g. Ellicott City; Carbin and Mahoney presentations). Understanding the people goes a long way to designing the messages, but we need well developed information and well-trained communicators to design and deliver that message.

Meeting partners needs and addressing their challenges could mean adding longer range products that detail risks/threats due to rainfall and flooding. But what information should be contained in such products, and how might it be displayed with an appropriate level of specificity and detail accounting for uncertainty? What information do water managers/planners need to enhance their decision making out to seasonal time scales? And does that information capability currently exist within the National Water Modeling portfolio?

5. Findings

Participants agreed that the HMT charter mission, vision, and activities are still valid and general enough to continue guiding the HMT program. The workshop identified several high priority challenges and approaches for addressing the challenges, and candidate metrics for tracking progress. These findings are the result of the diverse and informed perspectives at the workshop. It is recognized that these are not yet comprehensive, and further discussion and input are needed from a broader representation of science, operations, and partner communities. Presentations and discussions covered a range of emerging science, including that funded by the U.S. Weather Research Program testbeds, the Joint Technology Transfer Initiative (JTTI), and base funding at OAR labs.

A broad theme of the meeting was that impacts based decision support is a growing need. To improve the value of the forecast for core partners in particular requires that messages contain actionable and understandable forecast information. However, messaging must also be precise for many individual users. On the meteorology side, the precision placement (or in a probabilistic framework, the confidence) of high rainfall-rate and/or large accumulations has not reached a skill commensurate with users' needs. For hydrology, improved forecasts of precipitation out to 10 days can help the NWC provide better decision support, but work is needed to establish the capabilities to simulate debris flows (sediment, ice, or debris transport).

The WPC Hydromet Testbed, and its Flash Flood and Extreme Rainfall, Winter Weather, and other experiments, were recognized as important efforts for testing advances in a quasi-operational environment. Th OAR/Weather Program Office (WPO) support for testbeds and transition projects has provided dedicated funding for these and other testbeds, which have provided a successful pathway for R2O.

As Fritsch and Carbone (2004) emphasize, probabilistic prediction is key to the extreme rainfall challenge. Meteorological ensembles have reached the point where they can be beneficial but how to make the best use of them in IDSS is an ongoing challenge. The challenges range from messaging to the publics to training forecasters how best to use the ensemble guidance to communicate the forecast.

Discussions should continue on ways to build more coupled, integrated modeling systems - either linking atmospheric and hydrologic models or linking hydrologic models with debris flow models - that could assist more directly with user needs. But these modeling systems all propagate errors and teasing out the contributions to those errors is difficult. Insight, either through verification or evaluation, is needed. Reforecasting (such as the new GEFSv12) could be a foundation from which insight with current models could be constructed. However, all of these approaches require considerable computational investment.

The theme of providing probabilistic information to express certainty/uncertainty (i.e. confidence) in the forecast will help start the conversation to prepare a long-term improvement plan. Through evaluation and verification of Numerical Weather and Hydrological Prediction (NWHP) models, we need to frame where the best allocation of resources can lead to improvements in probabilistic information. This improved information will better meet the needs of partners.

High-priority hydrometeorology challenges:

- Improve numerical model quantitative precipitation forecast (QPF) and streamflow forecast skill, with reliability (the ability to give unbiased probability estimates) and sharpness (the ability to forecast extreme values)
- Create continuum of decision support services (DSS) for extreme rainfall from weather to S2S for more effective risk communication to core partners, including water resource management
- Improve risk communication through the cross disciplinary engagement between physical and social sciences
- Create continuum of extreme rainfall services from weather to S2S for more effective risk communication to core partners, including water resource management
- Establish the ability to disentangle errors from coupled meteorological and hydrological models

Two overarching approaches were identified to address these challenges:

- *Improve NWP forecasts of rainfall.* The skill of rainfall events has improved at a slower rate of improvement than even hurricane intensity skill. The pace of improvement in NWP rainfall forecasts needs to accelerate, including improved calibration of probabilistic forecasts (reliable and sharp distributions). Discussion highlighted that there is likely no silver bullet, and the opportunity to leverage NGGPS, UFS, EPIC, and other efforts.

- *Narrow the gap between what we know (full model data/post processing/scenarios) and what we tell (quality and quantity of DSS)*. A first step is to make use of what we already know. Several participants thought that there was good information that's not yet being used in the NWS, and thus not available to partners and the public. Addressing this involves both the physical and social sciences, including mining critical information from ensembles for forecasters to use (perhaps using machine learning), visualizing such information, as well as product design, and communication best practices.

Candidate metrics. The following candidate metrics were identified:

- Double the rate of improvement of NWP precipitation skill, from the current rate of improvement of 1% a year to 2% a year using GFS precipitation skill as baseline
- Extend heavy rainfall services to Day 10 (Improve DSS and connect to S2S)
- Improve the reliability and sharpness of probabilistic information for extreme events
- Narrow the 5 day landfalling position of atmospheric rivers on the west coast and Alaska from 500 km to 400 km.
- Improve the skill of atmospheric forcings provided to the NWM and other hydrological models.

These metrics will need further discussion to refine and develop baselines and numerical goals. See current HMT-related agency metrics in Appendix D.

Discussion also highlighted opportunities within the multidisciplinary fields and communities working on extreme precipitation that may benefit the understanding, forecasting, and communication of extreme precipitation risks. To name a few, technological advances include the development of UFS; emergence of gap filling radars to cover observation gaps; new AWIPS functionalities to create variables/analysis; and machine learning and artificial intelligence capacities. The program on Forecasting a Continuum of Threats (FACETS, Rothfusz et al. 2018) and the Warn-on-Forecast program (Stensrud et al. 2009) are working to improve communication of forecasts, and leverage an emerging Social and Behavioral Science (SBS) capacity in NOAA. Along with NOAA's partners such as cooperative institutes this SBS capacity has grown in the last decade, and now includes a social science program within the OAR WPO. Geospatial analysis is becoming easier and more common, and in combination with social science may allow identifying where more vulnerable basins and places exist, given any particular level of physical/meteorological events.

Finally, participants endorsed the concept for a focused improvement program on extreme precipitation forecasting. The group recognized the need to continue discussions and called for an effort to develop a HFIP-like improvement plan (Gall et al. 2013) for extreme precipitation and its impacts with timescales ranging from minutes to a season.

6. Conclusions and Next Steps

The group recognized the need to continue discussions and called for an effort to develop an HFIP-like improvement plan **for a program to improve predictions of extreme precipitation, and provide better forcings for flood predictions, including more effective watches and warnings.** This program would encompass extreme precipitation and its impacts with timescales ranging from minutes to 10 days, as well as the forcings necessary to meet the challenges of hydrologic modeling. This program would be a NOAA-focussed interagency effort within the US and western North America that advances the science and makes improvements in predictions in these time scales (sub-daily to 10 days). Such an effort will directly contribute to advancements needed to address the Precipitation Grand Challenge, including providing better "set up" for subseasonal to seasonal precipitation outlooks. Next steps include identifying a planning team and time line for such an effort to develop a science and implementation plan to address the challenges identified above. These include the science and research needed in the three areas (Meteorology, Hydrology, Social and Behavioral Science) to support forecast and operational challenges, as well as to estimate the resource needs and opportunities. This science and implementation plan document would establish some specific and aspirational goals, informed by recent trends in WPC and NWHP verification, and consistent with the NWS Strategic Plan. These three main topical areas each require improvements, and as such, it is necessary to understand the crosscut linkages that will help place resources in a configuration to maximize our chances of realizing these improvements.

Specifically, near term activities towards such a program include:

a) Contribute to the Precipitation Grand Challenge effort.
b) Develop research questions to address both near term and long-term forecast challenges.
c) Contribute to the UFS effort, in particular, the forecast goals and science priorities related to UFS applications for medium-range weather (atmospheric behavior out to about two weeks) and S2S (atmospheric and ocean behavior from about two weeks to about one year)
d) Conduct near real-time investigations of high impact extreme hydrologic events.
e) Begin a more comprehensive process to more fully lay out plans for an extreme precipitation improvement program.

Several ongoing NOAA efforts might aid in improving extreme precipitation, however, precipitation forecasting from minutes to 10 days is not a focus among these efforts. These include the EPIC, the NOAA Water Initiative, and recent investments to improve prediction on S2S time scales. Therefore, participants concluded that there is a ‘gap’ in investments for precipitation on this time scale, as depicted in Fig 13. To address this gap, the group recommended significant investment on this time scale, and supported continuing current HMT experiments and broadening the discussion to address the NOAA Precipitation Grand Challenge. Investment in extreme precipitation will have benefits across Climate Weather, and Water.

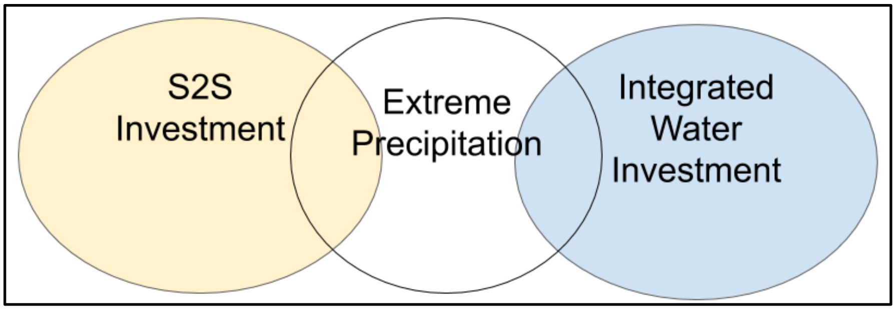


Figure 12. Investments in extreme precipitation prediction will have benefits for integrated water prediction and S2S prediction.

Overall, the workshop served to galvanize a community around the heavy precipitation and flooding challenge. We are not meeting the ever-increasing needs of partners, as partners are looking for specific information on rainfall **amount**s (including intensity), **where** it will occur, **when and the possible impacts it will have**. Participants are motivated to address this challenge.

REFERENCES

Barthold, F. E., T. E. Workoff, B. A. Cosgrove, J. J. Gourley, D. R. Novak, and K. M. Mahoney, 2015: Improving Flash Flood Forecasts: The HMT-WPC Flash Flood and Intense Rainfall Experiment. *Bull. Amer. Meteor. Soc.*, **96**, 1859–1866, https://doi.org/10.1175/BAMS-D-14-00201.1.

Blake, B.T., J.R. Carley, T.I. Alcott, I. Jankov, M.E. Pyle, S.E. Perfater, and B. Albright, 2018: An Adaptive Approach for the Calculation of Ensemble Gridpoint Probabilities. Wea. Forecasting, 33, 1063–1080, https://doi.org/10.1175/WAF-D-18-0035.1)

Carr, R. H., Montz, B., Semmens, K., Maxfield, K., Connolly, S., Ahnert, P., . . . Elliott, J., 2018: Major Risks, Uncertain Outcomes: Making Ensemble Forecasts Work for Multiple Audiences. Weather and Forecasting, 33(5), 1359-1373. doi:10.1175/Waf-D-18-0018.1

Coleman, C. 2019:The Corps approves major deviation for Forecast Informed Reservoir Operations effort. January 26, 2019. https://www.army.mil/article/216622/the_corps_approves_major_deviation_for_forecast_informed_reservoir_operations_effort

Gall, R., Franklin, J., Marks, F., Rappaport, E. N., & Toepfer, F., 2013:The Hurricane Forecast Improvement Project. Bulletin of the American Meteorological Society, 94(3), 329-343. doi:10.1175/Bams-D-12-00071.1

Eastern Research Group (ERG), 2017: Communicating Probabilistic Information for Decision-Makers: A Case Study Using Experimental Snow Forecast Products: Summary of Decision-Maker Focus Groups and Simulations. Written under contract for the NOAA's National Weather Service. August 2017. 42 pp.

Erickson, M. J., J. S. Kastman, B. Albright, S. Perfater, J. A. Nelson, R. S. Schumacher, and G. R. Herman (2019): Verification Results from the 2017 HMT–WPC Flash Flood and Intense Rainfall Experiment. *J. Appl. Meteor. Climatol.*, **58**, 2591–2604, https://doi.org/10.1175/JAMC-D-19-0097.1.

Fritsch, M. J., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season: A USWRP research and development strategy. Bulletin of the American Meteorological Society, 85, 955-965. https://doi.org/10.1175/BAMS-85-7-955

HMT Charter, 2016: https://drive.google.com/drive/folders/1Njp8APZYs-5OYeeHXW2Wu94GBI6-vhtg

Hudson, D.: 2017. Ensemble Verification Metrics. Presentation at the ECMWF Annual Seminar. https://www.ecmwf.int/en/elibrary/17626-ensemble-verification-metrics

Kain, J.S., S.R. Dembek, S.J. Weiss, J.L. Case, J.J. Levit, and R.A. Sobash, 2010: Extracting Unique Information from High-Resolution Forecast Models: Monitoring Selected Fields and

Phenomena Every Time Step. Wea. Forecasting, 25, 1536–1542, https://doi.org/10.1175/2010WAF2222430.1

Karstens, C.D., J. Correia, D.S. LaDue, J. Wolfe, T.C. Meyer, D.R. Harrison, J.L. Cintineo, K.M. Calhoun, T.M. Smith, A.E. Gerard, and L.P. Rothfusz, 2018: Development of a Human–Machine Mix for Forecasting Severe Convective Events. Wea. Forecasting, 33, 715–737, https://doi.org/10.1175/WAF-D-17-0188.1

NWS Office of Hydrology: Undated. Glossary of Forecast Verification Metrics. https://www.nws.noaa.gov/oh/rfcdev/docs/Glossary_Verification_Metrics.pdf

NWS Post Wildfire Hydrology Working Group, 2019. Notes and Findings from the Post-Wildfire Hydrology WorkshopColorado Springs, Colorado, September 17 - 21, 2018. https://drive.google.com/drive/folders/1Njp8APZYs-5OYeeHXW2Wu94GBI6-vhtg

Jasperse, J., Ralph, M., Anderson, M., Brekke, L.D., Dillabough, M., Dettinger, M.D., Haynes, A., Hartman, R., Jones, C., Forbis, J., Rutten, P., Talbot, C., and Webb, R. 2017. Preliminary viability assessment of Lake Mendocino forecast informed reservoir operations. Report of the Center For Western Weather and Water Extremes, https://pubs.er.usgs.gov/publication/70192184

Rothfusz, L. P., Schneider, R., Novak, D., Klockow-McClain, K., Gerard, A. E., Karstens, C., ...Smith, T. M. (2018). FACETs A Proposed Next-Generation Paradigm for High-Impact Weather Forecasting. Bulletin of the American Meteorological Society, 99(10), 2025-2043. doi:10.1175/Bams-D-16-0100.1

San Francisco Chronicle. Big atmospheric rivers do a lot of damage — especially in Northern California. Dec. 4, 2019. https://www.sfchronicle.com/environment/article/Big-atmospheric-rivers-do-a-lot-of-damage-14881960.php

Stensrud, D. J., Xue, M., Wicker, L. J., Kelleher, K. E., Foster, M. P., Schaefer, J. T., . . . Tuell, J. P. (2009). Convective-scale Warn-on-Forecast System A Vision for 2020. Bulletin of the American Meteorological Society, 90(10), 1487-+. doi:10.1175/2009bams2795.1

Sukovich, E. M., F. M. Ralph, F. E. Barthold, D. W. Reynolds, and D. R. Novak, 2014: Extreme Quantitative Precipitation Forecast Performance at the Weather Prediction Center from 2001 to 2011. Weather Forecast, 29, 894-911.

Viterbo, F., K. Mahoney, L. Read, F. Salas, B. Bates, J. Elliott, B. Cosgrove, A. Dugger, D. Gochis, and R. Cifelli, 2020: A Multiscale, Hydrometeorological Forecast Evaluation of National Water Model Forecasts of the May 2018 Ellicott City, Maryland, Flood. *J. Hydrometeor.*, **21**, 475–499, https://doi.org/10.1175/JHM-D-19-0125.1

White, A.B.; Moore, B.J.; Gottas, D.J.; Neiman, P.J. Winter Storm Conditions Leading to Excessive Runoff above California's Oroville Dam during January and February 2017. *B Am. Meteorol. Soc.* **2019**, *100*, 55–69, doi:10.1175/Bams-D-18-0091.1.

Wick, G.A., P.J. Neiman, F.M. Ralph, and T.M. Hamill, 2013: Evaluation of Forecasts of the Water Vapor Signature of Atmospheric Rivers in Operational Numerical Weather Prediction Models. *Wea. Forecasting*, **28**, 1337–1352, https://doi.org/10.1175/WAF-D-13-00025.1

WPC-HMT, 2019a: The 2018-19 HMT-WPC winter weather experiment: Final report. https://www.wpc.ncep.noaa.gov/hmt/experimentsummaries.shtml

WPC-HMT, 2019b: 2019 Flash Flood and Intense Rainfall Experiment (FFaIR): Final report. https://www.wpc.ncep.noaa.gov/hmt/experimentsummaries.shtml

APPENDIX A: Workshop Materials and Attendees

Workshop background materials are available in here:
https://drive.google.com/drive/folders/1iQwehw0KGTRJ-_kDnkbA7p7gwmlx3BsH

Link to Agenda:
https://docs.google.com/document/d/1wLtBvk3TCA9qIVV3UB0r8KZitmAeYihIRGRQBmloF8o/edit

Link to Presentations:
https://drive.google.com/drive/folders/1DR1qSIUWoW_6TAnMMupeMkVhBrcmoQ6x

HMT Charter, 2016. https://drive.google.com/drive/folders/1Njp8APZYs-5OYeeHXW2Wu94GBI6-vhtg

Workshop Attendees

Kate Abshire – NWS National Flash Flood Services Lead
Curtis Alexander – NOAA Office of Oceanic and Atmospheric Research (OAR)/Global Systems Division (GSD) Acting Assimilation Development Branch Chief
Tamara Battle – NOAA Office of Weather and Air Quality (OWAQ) Weather Act Policy Coordinator
Mike Bodner – NWS Weather Prediction Center (WPC) Meteorologist/Developer
Kandis Boyd **-** NOAA OAR OWAQ Acting Director
Gregory Carbin – NWS WPC Forecast Operations Branch Chief
Rob Cifelli – NOAA Earth System Research Laboratory (ESRL) Physical Sciences Division (PSD) Hydrometeorology Modeling and Applications Team Lead
Edward Clark – NWS National Water Center (NWC) Director, HMT Executive Committee
Grant Cooper – NWS National Centers for Environmental Prediction (NCEP) Acting Director
James Correia Jr – NWS WPC Hydrometeorology Testbed (HMT) Coordinator Contractor
David DeWitt – NWS Climate Prediction Center (CPC) Director
Gina Eosco – NOAA OAR OWAQ Social Science Program Manager
Michael Erickson – NWS WPC Research Scientist Contractor
Kathryn Gilbert – NWS WPC and Ocean Prediction Center (OPC) Deputy Director
JJ Gourley – OAR National Severe Storms Laboratory (NSSL) Research Hydrometeorologist
Brian Gross – NWS Environmental Modeling Center (EMC) Director
Kirstin Harnos – NWS WPC Winter Weather Experiment (WWE) Lead Contractor
Alan Haynes – NWS California-Nevada River Forecast Center Hydrologist in Charge
Joshua Kastman – NWS WPC Winter Storm Severity Index (WSSI) Research Scientist Contractor
Mark Klein – NWS WPC Science and Operations Officer
Bill Lamberson – NWS WPC Research Scientist Contractor
Christopher Lauer – Economist, NOAA Social Science, Performance & Strategy Division
Kelly Mahoney – ESRL/PSD Hydrometeorology Modeling and Applications Team
Mary Mullusky – NWS Analyze, Forecast, and Support (AFS) Office Water Resource Services Chief

David Myrick – NWS Office of Science & Technology Integration (STI) Field-Driven R2O Team Lead and National Science and Operations Officer
James Nelson – NWS WPC HMT Coordinator and Development and Training Branch (DTB) Chief
David Novak – NWS WPC Director, HMT Executive Committee
Michele Olson – NOAA OAR OWAQ Social Science Program Coordinator
Wendy Pearson – NWS Missouri Basin River Forecast Center, Hydrologist in Charge
Matthew Pyle – NWS EMC Engineering & Implementation Branch
Andrea Ray – OAR ESRL Physical Sciences Division, PSD HMT Coordinator and Physical Scientist
Katherine Rowden – NWS Western Region Hydrology Program Manager and NWS Post-Fire Hydrology Working Group Lead
Beverly Sobel – NWS WPC Project Manager Contractor
Sarah Trojniak – NWS WPC Flash Flood and Intense Rainfall Experiment (FFaIR) Coordinator Contractor
Bruce Veenhuis – NWS WPC Meteorologist
Jeff Waldstreicher – NWS Eastern Region HQ, Scientific Services Division
Robert Webb – ESRL/PSD Director, HMT Executive Committee
Lance Wood – NWS Houston/Galveston Science and Operations Officer
Nusrat Yussouf – Cooperative Institute for Mesoscale Meteorological Studies (CIMMS) & NSSL Research Scientist
Jian Zhang – NSSL Warning Research & Development Division

APPENDIX B: Brainstorming Notes from Group Discussion

1. **Key Challenges**
   a. Need 5-15 min QPE and (P) QPF to capture the rainfall rates that are impactful. Post-fire flood and debris flows, especially in the western U.S. -- these require 5-15 min QPE, (P) QPFs, and rain rates to capture the rainfall rates that are high impact for urban flooding as well. The MRMS v12 15 min QPF will help forward this.
   b. Need extension of precipitation information to the Day 4-7 timeframe. This is beyond existing CAMs, but general threat information can be provided.
   c. Central and southeastern US QPF -- skill in these regions is lower than other parts of the country, in particular for warm season convection
   d. Winter precipitation type challenges (rain-freezing ppt-snow and timing), in particular in the Northeast; rain-snow line and timing; melt level to determine p-type (rain or snow)
   e. Distinguishing between 'typical' and 'catastrophic' impacts in advance
   f. Flood Outlook Potential (FOP) - categorical: [Flood] severity index (like winter storm severity index)? Method to communicate between outlooks, watches, and warnings
   g. Spatial (dis)placement/object-based verification
   h. Disentangling model errors (e.g., what is causing spatial placement or intensity "bust"
   i. Communication of range of possibilities and baseline uncertainty from model; conditional climatology
   j. Reforecast of guidance is necessary but, hard; computationally expensive
   k. Understand forecasters as a user (e.g., use of ensembles), including sufficient training
   l. Forecasts of the most extreme flooding days, such as Ellicott City or tropical storm-induced flood days and to be able to designate these in risk/threat products. Operations and core partners would like to be able to distinguish between 'typical' and 'catastrophic' impacts in advance.
   m. Communicating threats and uncertainty for extreme precipitation events, including hurricanes
   n. Heavy rain events inland and the S2S and medium-range set-up for flooding
   o. Precipitation outlook information out to the Day 4-7 How do we best communicate the potential of an extreme rainfall event in the Day 4-7 forecast period?
   p. Predicting mesoscale frontal wave formation which can significantly increase precipitation duration over a given basin

2. **What is Needed to Address Challenges**
   a. Computing (address span of requirements from first hour to Day 10)
   b. Machine learning to tease out information on extreme rainfall threats
      i. Data and people to work on this (PSD 20-year reanalysis)
   c. New kinds of verification - e.g., object-based, accounting for displacement
      i. "Cone of uncertainty," "error" 3-4 days out?

d. Training - particularly of forecasters
e. What is already available but not yet used in WFOs?

3. Opportunities

a. Gap filling radars - for ppt rate at short-time scales
b. MRMS v12 15 min QPF and HRRR v4 hourly update to 18 hour - what outputs needed to learn from these models?
c. AWIPS functionalities (JJ) to create variables/analysis without using bandwidth
d. Machine learning capacities, AI
    i. Need to elevate NOAA corporate capabilities
e. Leverage Social Behavioral Science (SBS) capacity in NOAA
f. NOAA big data initiative for HRRR - data mining, model climatology
g. Testbed participation by WFO forecasters as opportunity to understand use

4. Other

a. Timing - amount and where? Are the key precipitation issues
    i. Complement to Mesoscale Precipitation discussion
    ii. Vulnerability of some places
    iii. Machine Learning and geospatial => Where are more vulnerable places? Given a level of physical/meteorological events
b. Integrate expected meteorology and expected hydrology - engineering/built environment
    i. Social Behavioral Science (SBS) - improving communication, understanding needs
c. Uncertainty - different types - QPF and response uncertainty; static and dynamic hydrological uncertainty and how to communicate 3 dimensions

Appendix C: Background on the WPC and NWS-OAR Hydrometeorological Testbeds

The Hydromet Testbed refers to a suite of efforts with related goals and missions: 1) the NWS/WPC testbed, one of twelve testbeds and proving grounds that facilitate the orderly transition of research capabilities to operational implementation through development testing, and pre-deployment testing and operational readiness/suitability evaluation. A competitive grant program provides funding for projects to be developed and tested, under the U.S. Weather Research Program operated out of the OAR Office of Weather and Air Quality (OWAQ). OWAQ also funds projects on precipitation forecasting under the parallel Hazardous Weather Testbed operated by the NWS Storm Prediction Center. These projects complement projects in the HMT, and include some efforts discussed at this workshop.

In 2016, a Charter was developed to encourage collaborative effort and a common mission and vision among the NWS WPC and National Water Center, and OAR to improve forecasts of extreme precipitation and forcings for hydrologic prediction (see https://drive.google.com/drive/folders/1Njp8APZYs-5OYeeHXW2Wu94GBI6-vhtg). HMT supports both OAR and NWS strategic objectives (HMT Charter, 2016). The Charter lays out a business model in which the HMT functions:

- Fosters the adoption of new techniques, models, observing systems, and other advances with potential for improving forecast guidance and hydrologic forcings.
- Incorporates well-established practices from the Joint Hurricane Testbed to expedite the transfer of research advances into an operational setting.
- Is guided by operational challenges and requirements identified through the Capabilities and Requirements Decision Support (CaRDS) process.
- Identifies high-priority hydrometeorology challenges in the forecast of extreme precipitation and the investigation of new techniques, models, observing and forcings for hydrologic prediction as candidate for demonstration and testing.

Every year HMT meteorologists at WPC organize the Winter Weather Experiment (WWE), Medium Range Forecast, Flash Flood and Intense Rainfall (FFaIR) Experiment and other efforts to test potential new forecasting products and methods with researchers, forecasters, partners, etc. This interaction enhances the R2O process.

Participants at the Flash Flood and Intense Rainfall Experiment (FFaIR)

FFaIR participants display forecast maps.

Appendix D. HMT-relevant Metrics tracked by NOAA

The Following are Government Performance and Results Act (GPRA) measures tracked by NOAA, except as noted. GPRA descriptions are taken from the NWS Performance Management Web Portal (https://verification.nws.noaa.gov/services/public/index.aspx) and Department of Commerce Performance dataset (https://performance.commerce.gov/KPI-NOAA/NOAA-Accuracy-threat-score-of-Day-1-precipitation-/mvk5-p955). Current and recent goals and scores for each measure can be found at these websites.

Precipitation Forecast Day 1 Threat Score

This Threat Score (TS) measures how well WPC forecasts the occurrence of precipitation (rain or the water equivalent of melted snow or ice pellets) of 1 inch or greater in the 24-hour period 6 to 30 hours into the future across the contiguous U.S. The TS is a verification score that takes into account two types of precipitation forecasting errors: the errors associated with not forecasting rainfall where it does occur and those associated with forecasting rainfall where it does not occur. The Threat Score is defined as the number of correct forecasts divided by the total of number of correct forecasts plus the number of wrong forecasts. The TS varies from 0, no correct forecasts, to 100, when the forecast area exactly matches the observed area of 1 inch rainfall over the entire U.S. The FY20 annual score goal is 34. The scores vary seasonally during the year with higher values generally occurring during the fall and winter when weather systems are larger and more well-defined and lower values occurring in the spring and summer when precipitation is scattered and on a smaller scale.

WPC also tracks the Day 3, two-inch threat score as a DOC Agency Priority Goal (https://www.performance.gov/commerce/APG_commerce_3.html). This is the same score as the Day 1 GPRA, except with a higher threshold (2" in 24 hours) and longer lead time (3 days in advance). WPC has a goal to maintain a 3-year running average of this TS of 0.14. According to Novak, sustaining a 0.14 3-day Threat Score gives confidence to provide 'high risk' excessive rainfall outlook category on Day 3

Flash Flood Warnings - Lead Time and Accuracy

Flash flood warning statistics are based on product issuance information and confirmation of actual flash floods by the local Weather Forecast Offices. The metric includes all warned events with zero lead times and all unwarned events. NWS forecasters issue on average approximately 4,000 flash flood warnings per year.

Flash Flood Warning Lead Time is the difference between the time the warning was issued and the time flash flooding was first reported (based on certified storm reports). The 2020 Lead Time goal is 65 minutes.

Flash Flood Accuracy (POD) reflects both the spatial and temporal accuracy of the flash flood warning(s). This metric represents the percentage of the flash flood event (area and time) that was warned. The 2020 Accuracy goal is 76%.

Winter Storms - Lead Time and Accuracy

Similar to Flash Flood Warnings, the Winter Storm GPRA measures include lead time and accuracy. Any winter storm event that meets warning criteria verifies any type of winter storm warning. Winter storm warning lead times have more than doubled since FY99, to an average of just over 20 hours, with some recent leveling since FY13. The accuracy, or Probability of Detection (POD), for winter storm warnings has held steady (0.85 to 0.90) since the NWS started tracking it as a nationwide measure in FY99. However, some years have shown and will continue to show lower hit rates. This is because a drop in the event count for a given year tends to produce a lower score, as forecasters have fewer opportunities. For example, drops in the annual counts of winter storms from 8748 (FY14) to 6169 or less (FY15 thru FY18) may have contributed to the recent decline in hit rates. The FY2020 goals are 20 hours lead time and 90% accuracy.

Appendix E. Acronyms

APG: Agency priority goal
CaRDS: Capabilities and Requirements Decision Support
CIMMS: Cooperative Institute for Mesoscale Meteorological Studies
CNRFC: NWS California-Nevada River Forecast Center
CSI: Critical success index (i.e., threat score)
DOC: Department of Commerce
EMC: NWS NCEP Environmental Modeling Center
EPIC: Earth Prediction Innovation Center
ERO: WPC Excessive Rainfall Outlook
ESRL: OAR Earth System Research Laboratory
ETS: Equitable Threat Score
FAR: False Alarm Ratio
FFaIR: HMT Flash Flood and Intense Rainfall Experiment
GFS: Global Forecast System
GPRA: Government Performance and Results Act
GSD: OAR/ESRL Global Systems Division
HFIP: Hurricane Forecast Improvement Project
HREF: High Resolution Ensemble Forecast
HRRR: High Resolution Rapid Refresh model
HWT: OAR-NWS Hazardous Weather Testbed
HMT: OAR-NWS Hydrometeorological Testbed
IDSS: Impact-Based Decision Support Services
JHT: OAR-NWS Joint Hurricane Testbed
JTTI: OAR-NWS Joint Technology Transfer Initiative
MAE_{cond}: conditional mean absolute error
MBRFC: NWS Missouri Basin River Forecast Center
MRMS: NSSL Multi-radar/Multi-sensor system
NCEP: NWS National Centers for Environmental Prediction, including WPC, EM
NGGPS: Next Generation Global Prediction System

NOHRSC: NWS National Operational Hydrologic Remote Sensing Center
NSSL: National Severe Storms Laboratory
NWC: NWS National Water Center
NWM: National Water Model
NWS: NOAA National Weather Service
OAR: NOAA Office of Oceanic and Atmospheric Research
OPC: NCEP Ocean Prediction Center
OWAQ: OAR Office of Weather and Air Quality, now WPO*
OWP: NWS Office of Water Prediction
POD: Probability of Detection
PQPF: Probabilistic quantitative precipitation forecast
PSL: OAR ESRL Physical Sciences Laboratory (formerly Division)*
QPE: Quantitative precipitation estimation
QPF: Quantitative precipitation forecast
S2S: Subseasonal to seasonal time scale
SSD: NWS Scientific Services Division (one for each region)
STI: NWS Office of Science & Technology Integration (STI)
TS: Threat Score
UFS: Unified Forecast System
USGS: U.S. Geological Survey
USWRP: U.S. Weather Research Program, part of WPO
WFO: NWS Weather Forecast Office
WPC: NWS NCEP Weather Prediction Center
WPO: OAR Weather Program Office, formerly OWAQ*
WWE: HMT Winter Weather Experiment
WSSI: WPC Winter Storm Severity Index

*The names of these organizations were changed between the workshop and the final report.