NATIONAL WEATHER SERVICE

Building a Weather-Ready Nation

Validation of NWS Hydrologic Ensemble Forecast Service (HEFS) Real-time Products at the Middle Atlantic River Forecast Center

Seann Reed, Alaina MacFarlane

1/13/2020 AMS Annual Meeting Boston

Overview

- What is HEFS?
- Validation methods for short-term products (0-10 days)
  - Compare to MMEFS
  - Compare to deterministic forecasts
- Results
- Next steps

Hydrologic Ensemble Products at MARFC

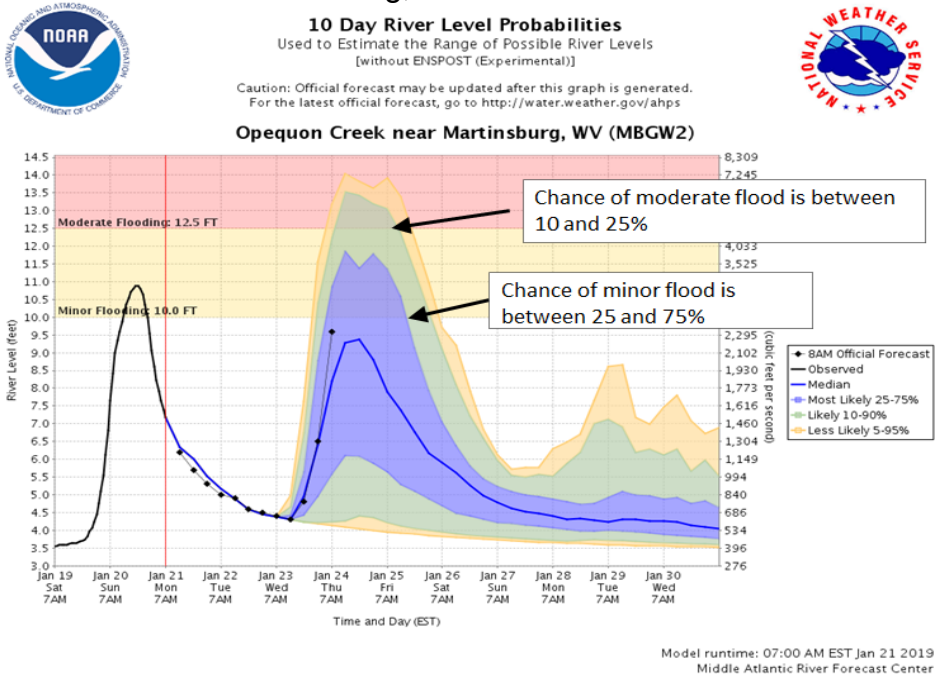
ESP	MMEFS	HEFS*
1990's	2012	2017
Accounts for soil moisture and snow states	Accounts for soil moisture and snow states	Accounts for soil moisture and snow states
Runs historical precipitation and temperature through hydrologic models	Runs raw meteorological model output through hydrologic models	Runs bias corrected and downscaled met model output through hydrologic models
30 to 90 day outlooks	7 day outlooks	Seamlessly blends forecasts for different lead times up to 365 days
No met models	GEFS v.11, NAEFS, SREF	GEFS v.10 (current implementation)
Once per day	Multiple times per day	Once per day

* Software, science, and support from the NWS Office of Water Prediction (OWP)

HEFS Graphics Currently on water.weather.gov/ahps

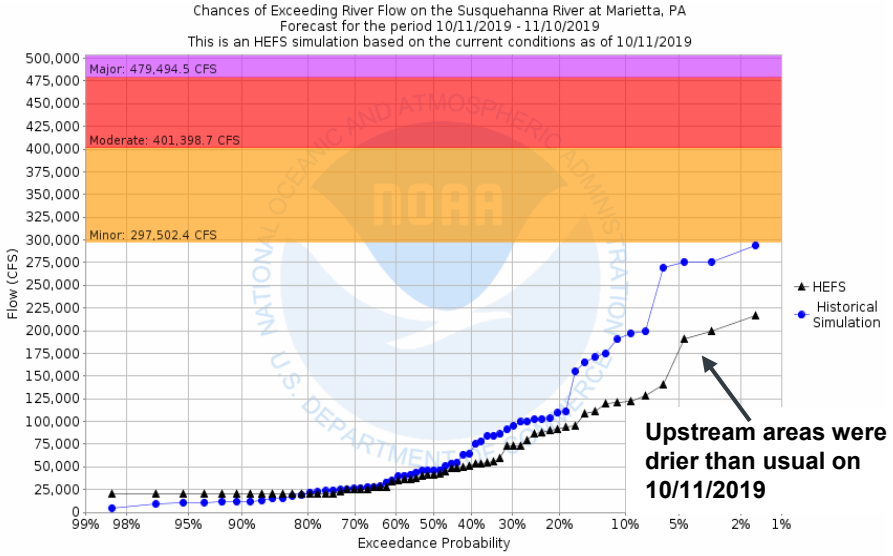
10-day Short-term Hydrograph Product

Martinsburg, WV 273 mi^2



30-day Exceedance Probability Product

Marietta, PA 26,000 mi^2



HEFS graphics are available at 117 of our forecast points

Who Uses Short-term Hydrologic Ensemble Forecasts?

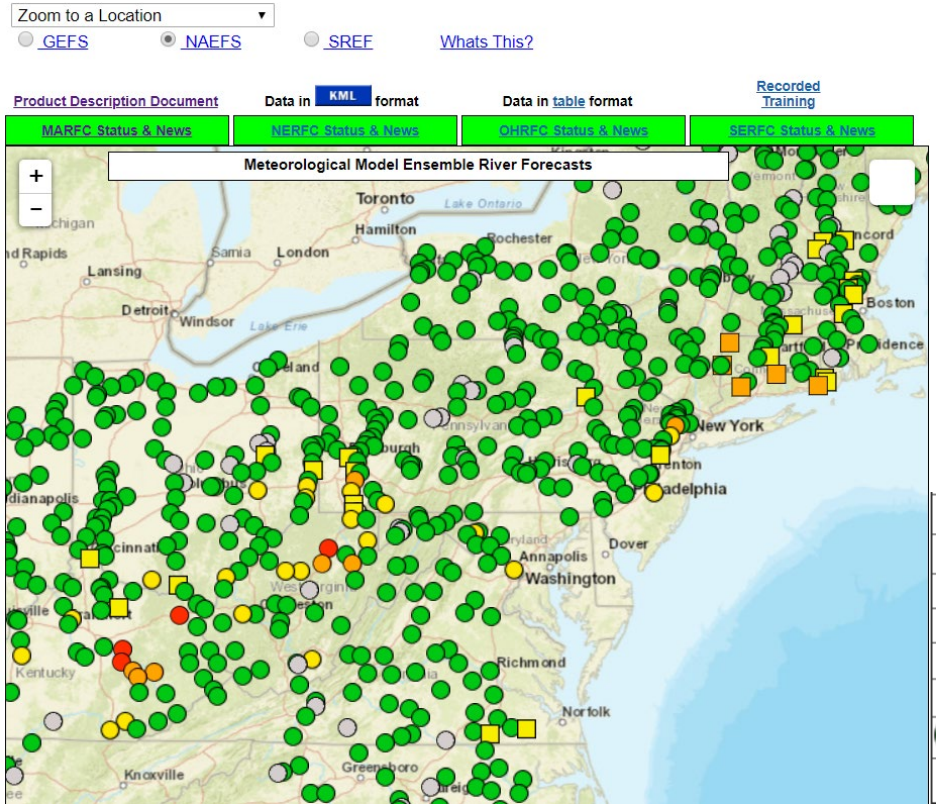
- RFC and WFO Forecasters
- Emergency managers
- Water Resources Managers
- Hydroelectric Power Plant Operators

Transition from MMEFS to HEFS

- MMEFS 7-day outlook products popular since 2012
- We are now producing 10-day HEFS products
- Do we continue to produce both products?
- How accurate are they?

MMEFS Map-based Web Interface

https://www.weather.gov/erh/mmefs?Lat=40&Lon=-77&Zoom=7



- Four RFCs
- GEFS, NAEFS, and SREF
- **Will add HEFS to this interface**

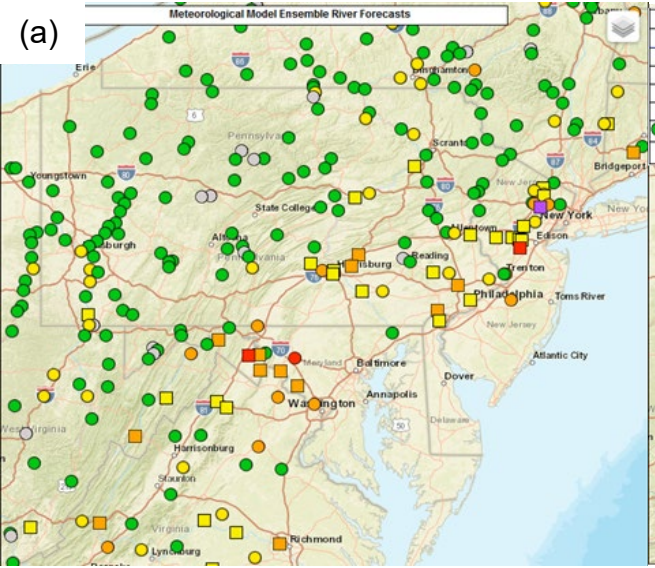
Chance of Exceedance		
30%	Level	70%
	Action	
	Minor Flood	
	Moderate Flood	
	Major Flood	
= less than 30% chance of reaching Action level		
= no critical levels defined for this point		

Map-based Event Validation Example

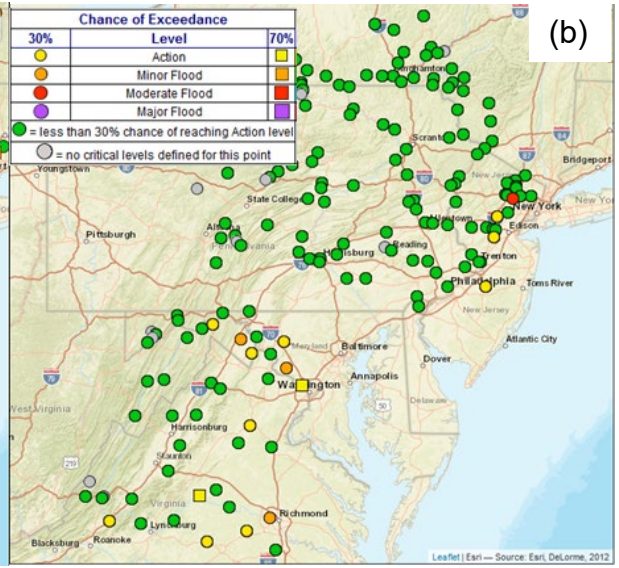
MMEFS-NAEFS Jan 22, 2019 0z

HEFS Jan 22, 2019 0z*

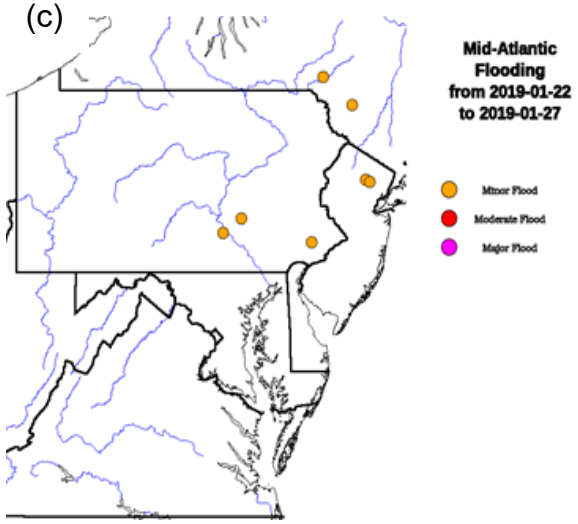
(a)



(b)



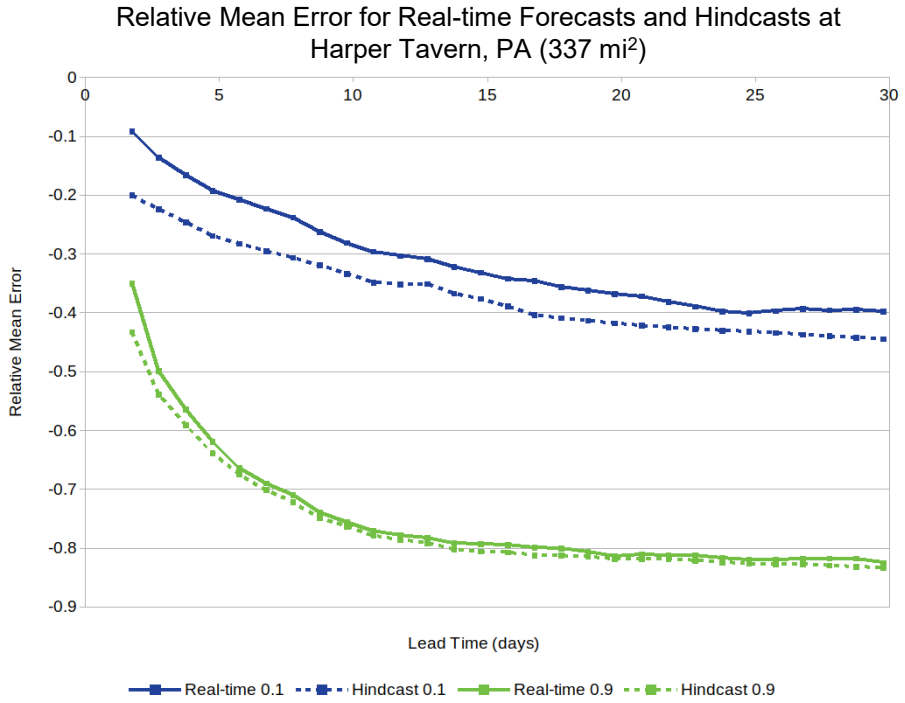
(c)



*Offline HEFS maps

MMEFS-NAEFS: 5 correct, 22 false alarms	HEFS: 1 correct, 3 false alarms	2.5 Day Lead Time

Which forecasts to validate?



- Real-time products – shorter record
- Hindcasts – no state updating
- We validated real-time products in this study

Validation Data

- Archived HEFS and MMEFS-NAEFS forecasts from Jan 2017 to Sep 2019 (**wet years!**)
- Observed stage data from 103 points
  - ice affected stages set to missing
- For contingency statistics:
  - **pooled** data from many points given short analysis period
  - exceedances in a **2 – 6 day forecast window** are tallied, effectively pooling lead times

Contingency Statistics

Contingency	Obs Y	Obs N
Fcst Y	A	B
Fcst N	C	D

POD = A/(A+C)

Probability of Detection=Fraction of obs floods predicted correctly

FAR = B/(A+B)

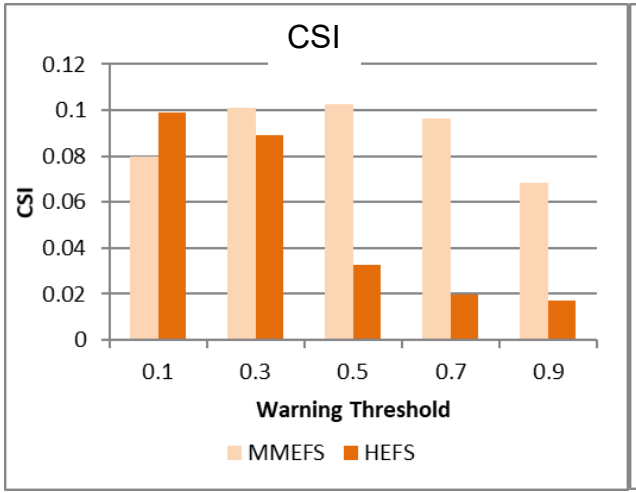
False alarm ratio=fraction of fcst floods which do not occur

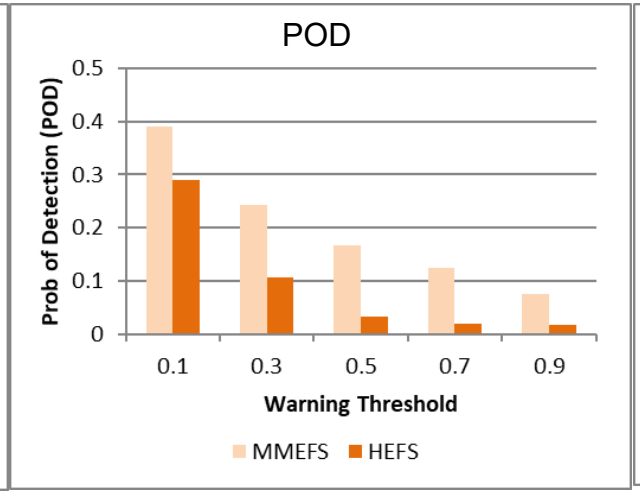
CSI = A/(A+B+C)

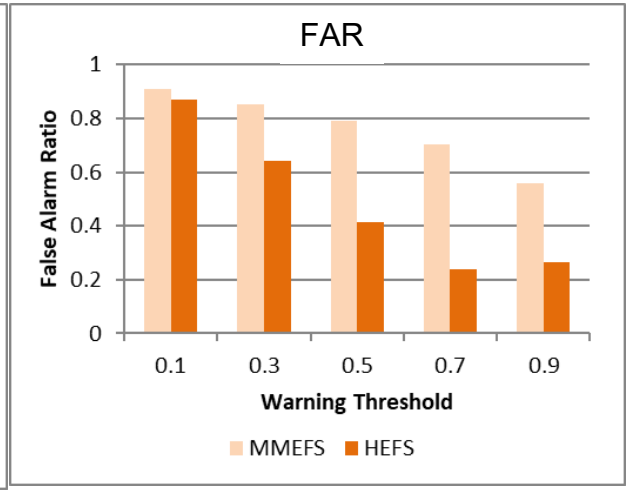
Critical success index=fraction of either obs or fcst floods predicted correctly

CSI, POD, FAR for Predictions Exceeding Minor Flood

(results pooled for 103 locations, 2.75 years of daily forecasts)



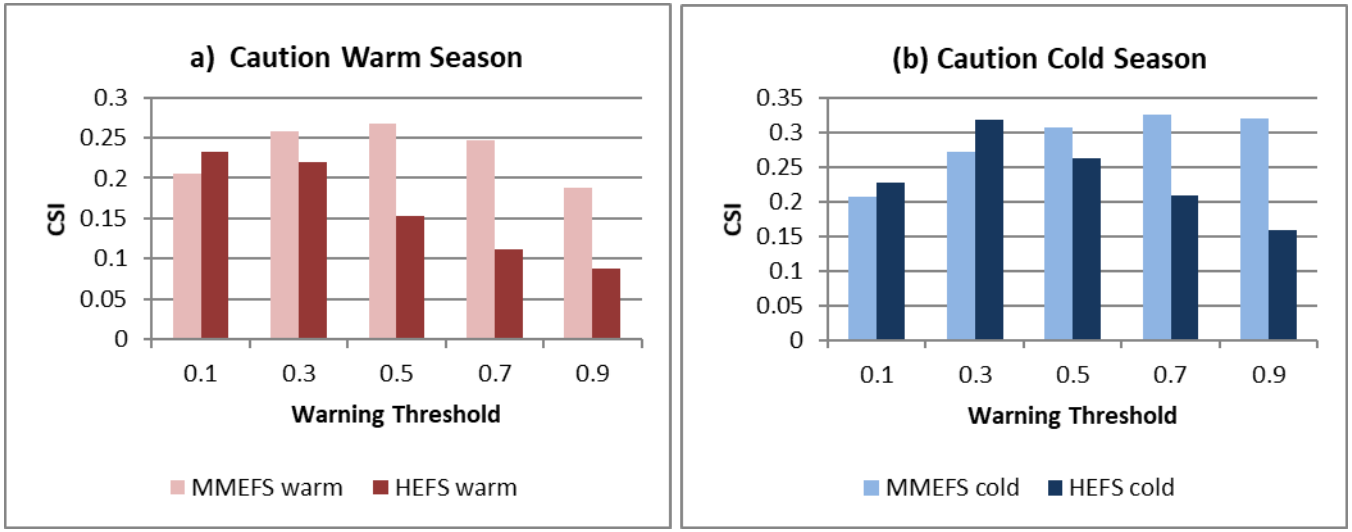




A 0.1 warning threshold means the ensemble forecasts predict action stage if 10% of the ensemble members exceed action stage. 0.1 is a much more conservative threshold than 0.9 when 90% of members would have to exceed action stage.

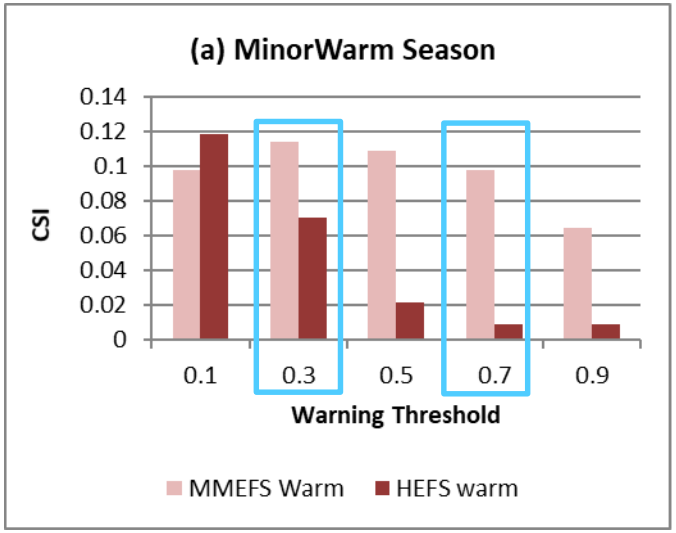
CSI for Predictions Exceeding Caution

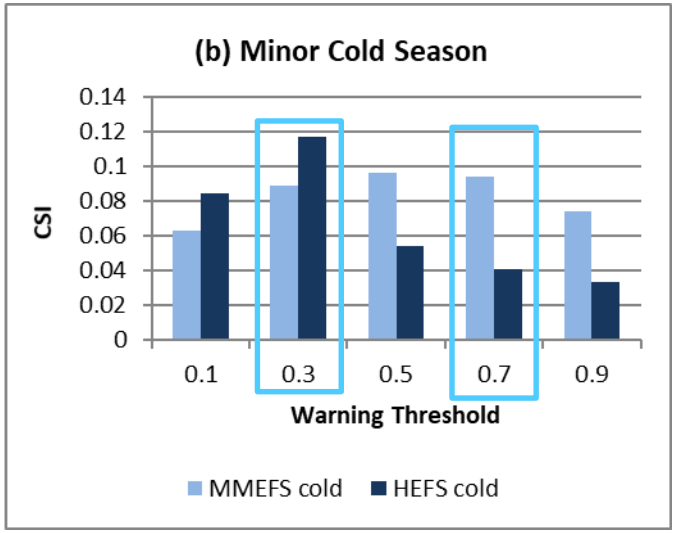
(results pooled for 103 locations, 2.75 years of daily forecasts, and two seasons)



CSI for Predictions Exceeding Minor Flooding

(results pooled for 103 locations, 2.75 years of daily forecasts, and two seasons)

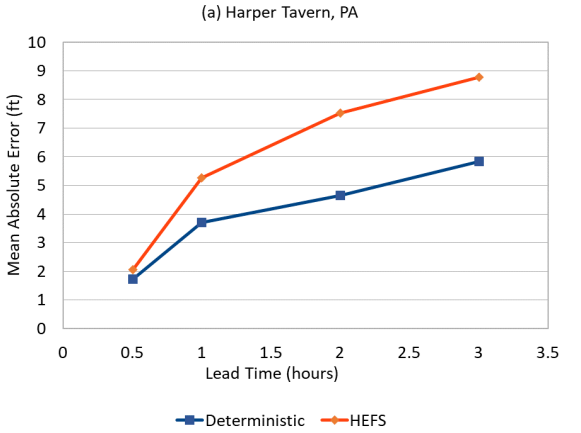


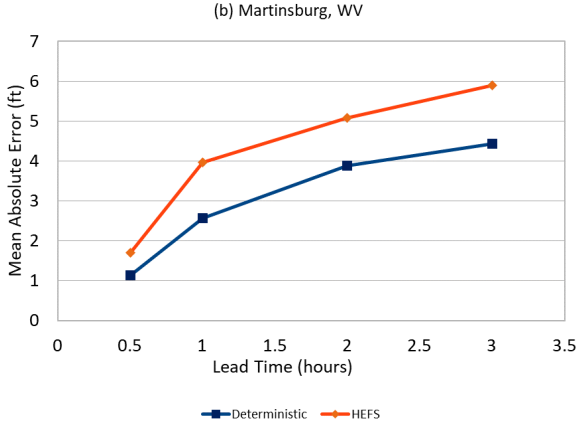


Mapped thresholds

Deterministic vs. HEFS Ensemble Mean Forecast

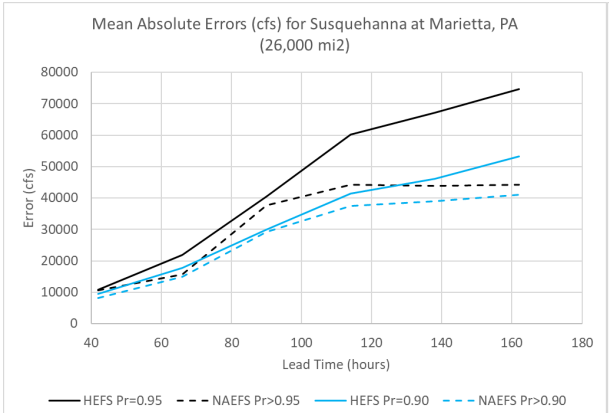
Small, flashy basins

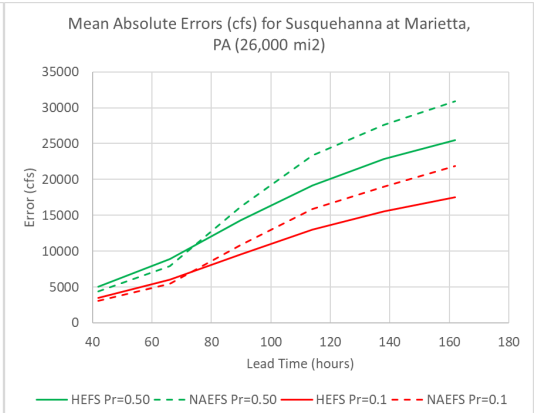


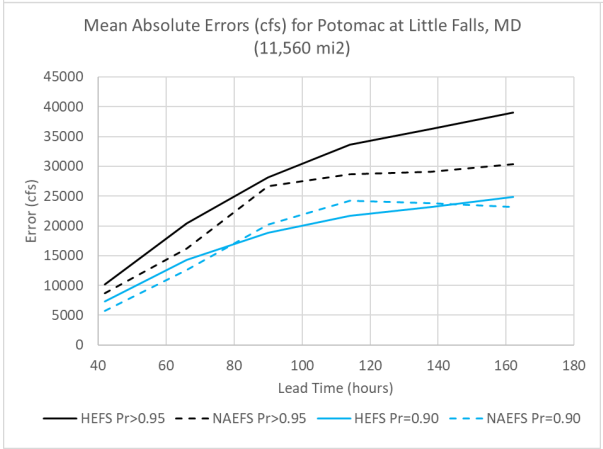


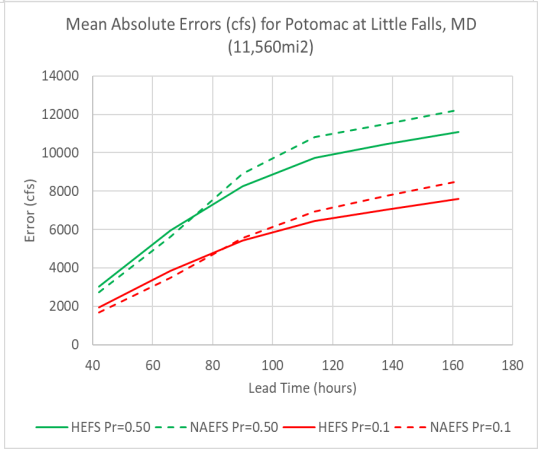
- Mean absolute errors (MAE) for only forecast-observed pairs **above flood stage**
- Includes 14 flood events for Harper Tavern and 20 for Martinsburg
- For floods, we emphasize deterministic forecasts in days 1-3, then use ensemble forecasts for 3 days and beyond; however, ensemble information may be used to inform our deterministic QPF.

MAE for High and Low Flows in Large Basins









- Left panels: MAE for only forecast-observed pairs above the 90th and 95th percentiles
- Right panels: MAE for only forecast-observed pairs above the 10th and 50th percentiles
- NAEFS has lower errors for flows>95th percentile; HEFS has lower errors for flows >10th and flows > 50th percentile

Conclusions

- MMEFS has higher PODs but also higher FARs compared to HEFS
- HEFS and MMEFS CSIs are comparable at low warning thresholds but MMEFS CSIs are better at higher warning thresholds
- For minor flooding, HEFS does better in the cold season at the 0.3 warning threshold, which is the lower of two thresholds we display on our maps; HEFS does poorly at the other map display threshold of 0.7
- Both models predict Caution stage more accurately than Minor flooding
- HEFS does better than MMEFS in terms of MAE for low and medium flows in large basins – water supply applications
- For flood events in small basins, deterministic model MAE is substantially lower than HEFS ensemble mean MAE in days 1-3
- **Continue to run and validate both HEFS and MMEFS; work to improve HEFS. . .**

Next Steps

- GEFS v. 12 into HEFS
- HEFS multiple runs per day – get closer to deterministic forcings/Mods
- HEFS improved bias correction for high flow events (NWS Office of Water Prediction is working on this.)