H31L-0768

Quantifying and Communicating the Uncertainty of National Water Model Evaluation Metrics

Jason A. Regina[1]

1 Office of Water Prediction, National Weather Service, NOAA, Tuscaloosa, AL, USA.

AGU24 Annual Meeting
Washington, D.C. | 9–13 December 2024
H029 - Advancing the State-of-the-Science of Water Resources Modeling - What's Next for Community Development at the Intersection of Domain, Computer, Data, and Decision Sciences

OWP OFFICE OF WATER PREDICTION

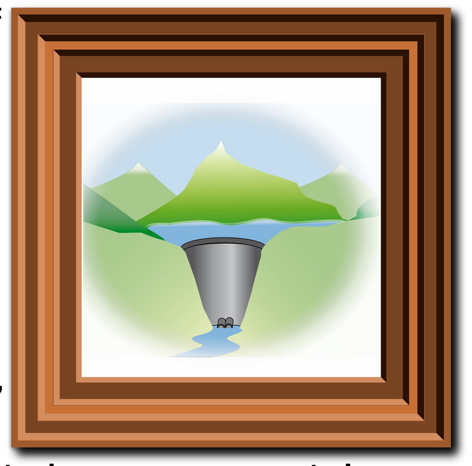
Choose your own Adventure!

Continental scale hydrological modeling takes all the challenges of catchment scale modeling and makes them even more complex. This includes quantifying and communicating uncertainty. Appropriate *framing* of uncertainty becomes especially important when supporting decision makers.

Choose one or more uncertainty frames[1] to learn more!

1 — Maturity and Utility
2 — Expert Knowledge
3 — Scope
4 — Robustness
5 — Confidence

1 — Maturity & Utility



Framing uncertainty in terms of maturity and utility involves fairly characterizing model features in relative terms. For example, the National Water Model v3.0 includes a reservoir module that simulates releases Downstream of dams. However, we expect model performance downstream of these locations to be more uncertain than at unregulated locations due to the inherent complexities of reservoir management.

2 — Expert Knowledge

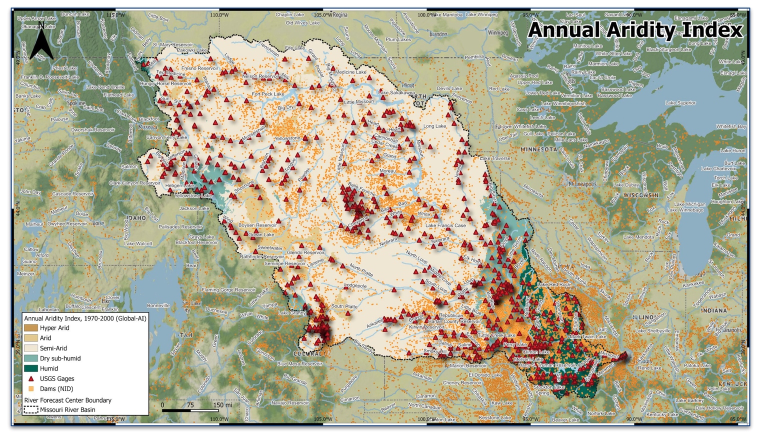
Framing uncertainty with expert knowledge lends context to evaluation metrics given what we know about hydrology, forecasting, or observations. For example, observed discharge error tends to be higher during very high and very low flow[2].

This increased observational uncertainty has implications on the uncertainty of evaluations conditioned on high or low flow. Image by Sachin Shah. Public Domain. "Gage in Flood." https://www.usgs.gov/media/images/gage-flood.

The more you learn, the less you know!

3 — Scope

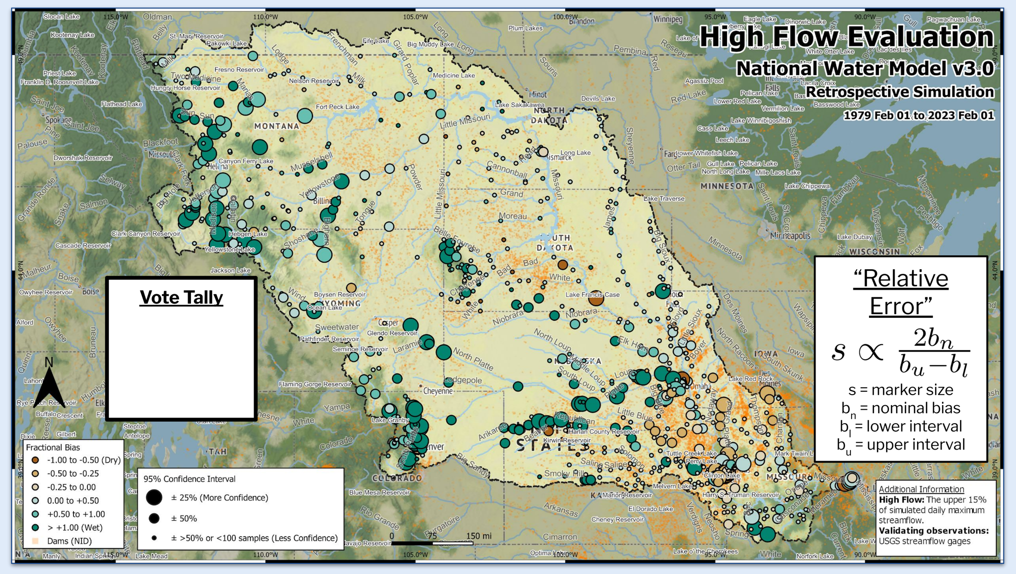
Framing uncertainty in terms of scope makes it clear to users what model output does and does not take into consideration. For example, National Water Model v3.0 based flood inundation maps are limited to riverine flooding. Inferences about the potential for pluvial flooding will require alternative models, accurate forcing, and expert opinion.

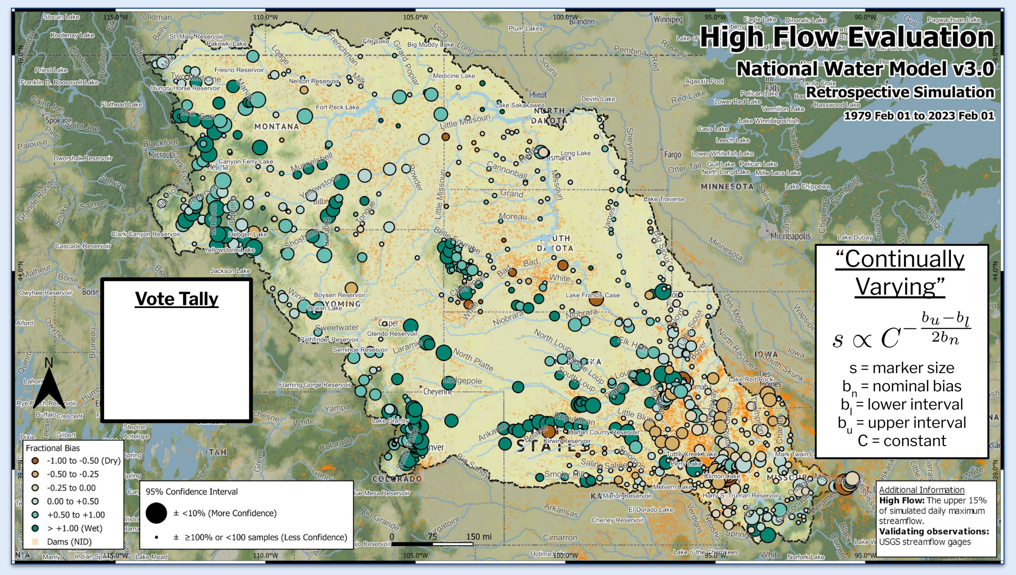


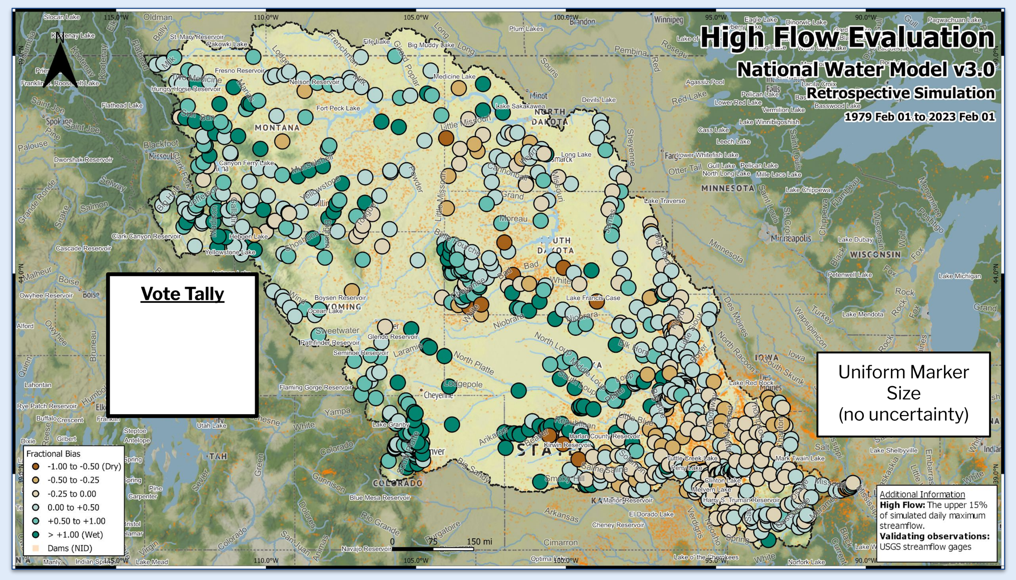
Establishing scope can also involve limiting applicable metrics to catchments with similar attributes like land cover, top soils, or annual aridity index. Model performance in the humid lower Missouri basin does not necessarily extend to the arid upper basin, for example.

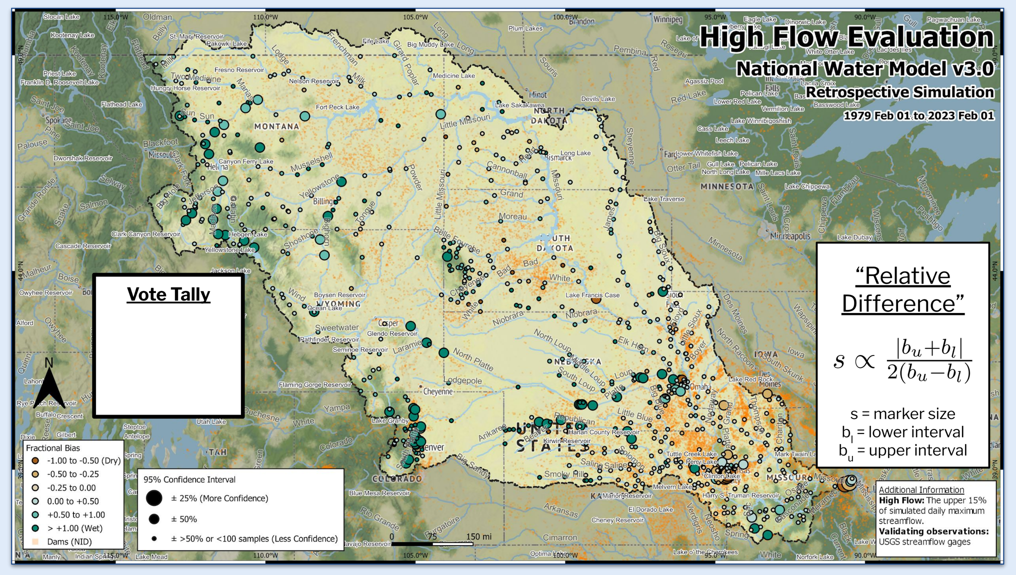
5 — Confidence

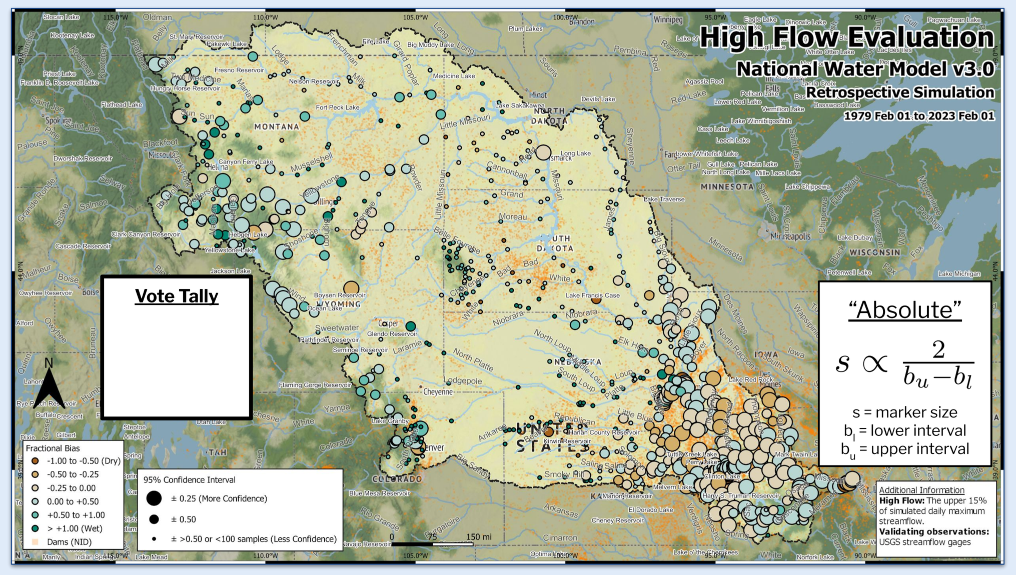
Statistical "confidence," in particular *confidence intervals*, often come to mind when considering how to frame uncertainty. Confidence intervals for evaluation metrics offer ranges of estimates for model performance. There are many ways to construct and communicate confidence intervals. These maps demonstrate five different ways to symbolize 95% confidence intervals by varying marker size. Use the stickers to vote for your favorite!





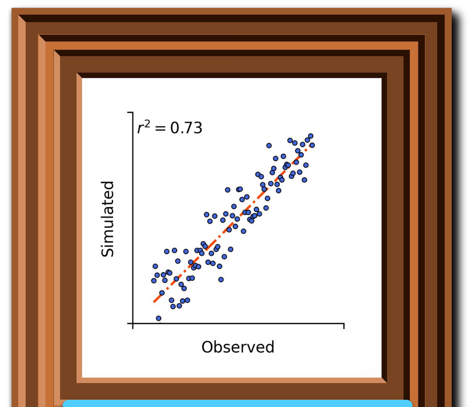




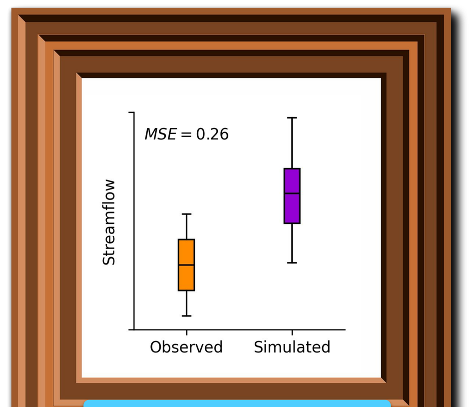


Use the stickers to vote for your favorite map of uncertainty!

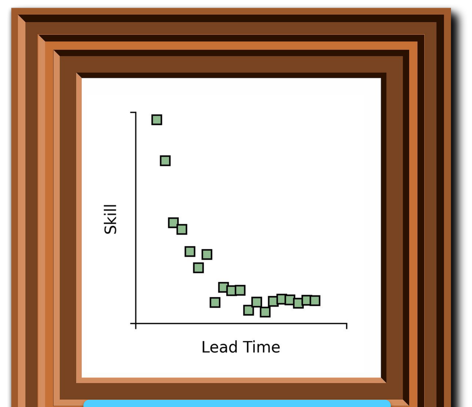
4 — Robustness



Association



Bias

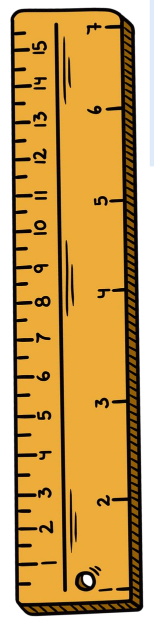


Skill

Framing uncertainty in terms of robustness involves applying multiple independent methods of analysis to the same problem to quantify the reliability of your conclusions. In the context of evaluations, adopting a set of uncorrelated or weakly correlated metrics offers a way to test model reliability across different measures of performance.

For more on evaluation challenges, see our sister poster!
Challenges in Establishing Evaluation Standards for R2O Decision-Making in Hydrology
IN33B: Breaking Down Silos: Advancing Interoperability in Computational Hydrosciences Through Standardization
Wednesday, 11 December 2024, 13:40 – 17:30 EST

Evaluation Metrics



Model evaluation metrics offer a summary of a model's accuracy compared to a benchmark (often observations). Popular metrics include Nash-Sutcliffe Efficiency, Peak Streamflow Error, or False Alarm Rate. The example evaluation maps shown above visualize Fractional Bias.

Fractional Bias

Tendency to over or under simulate streamflow magnitude (above "high water")
Range: -1 to infinity
Values < 0: under-simulation ("dry" tendency)
Values > 0: over-simulation ("wet" tendency)

REFERENCES

1. Guillaume, J. H. A., Helgeson, C., Elsawah, S., Jakeman, A. J., & Kummu, M. (2017). Toward best practice framing of uncertainty in scientific publications: A review of Water Resources Research abstracts. Water Resources Research, 53(8), 6744–6762. https://doi.org/10.1002/2017WR020609
2. McMillan, H., Krueger, T., & Freer, J. (2012). Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. Hydrological Processes, 26(26), 4078–4111. https://doi.org/10.1002/hyp.9384

ACKNOWLEDGEMENTS:

The author would like to thank James D. Brown, Katie Van Werkhoven, Jeff Zimmerman, Trey Flowers, Nels Frazier, and Austin Raney for their review and critical feedback.

DATASETS:

1. NOAA National Water Model CONUS Retrospective Dataset—Registry of Open Data on AWS, accessed 28 Nov 2023, https://registry.opendata.aws/nwm-archive/.
2. U.S. Geological Survey, 2016, National Water Information System (USGS Water Data for the Nation), accessed 28 Nov 2023, http://waterdata.usgs.gov/nwis/.

CONTACT

Website: *https://water.noaa.gov*
Email: *nws.nwc@noaa.gov*

View my poster and other conference materials