An Integrated Evaluation Framework based on Generalized Likelihood Uncertainty Estimation for Quantifying Uncertainty in Flood Modeling

Tao Huang[1] and Venkatesh Merwade[1]

[1]Purdue University

December 8, 2022

Abstract

Evaluation of the performance of hydrologic and hydraulic models is a crucial step in the modeling process. Considering the limitations of single statistical metrics, such as the Nash Sutcliffe efficiency (NSE), the Kling Gupta efficiency (KGE), and the coefficient of determination (R2), which are widely used in the evaluation of model performance, an evaluation framework that incorporates multiple criteria and based on the generalized likelihood uncertainty estimation (GLUE) is proposed to demonstrate the uncertainty in the evaluation criteria and hence to quantify the overall uncertainty of flood models in a comprehensive way. This framework is applied to the one-dimensional HEC-RAS models of six reaches located in States of Indiana and Texas of the United States to quantify the uncertainty associated with the channel roughness and upstream flow input. Specifically, the effects of different prior distributions of the uncertainty sources, multiple high-flow scenarios, and various types of measurement errors (white noise, positive bias, and negative bias) in observations on the evaluation metrics are investigated by using the bootstrapping method and Monte Carlo simulations. The results show that the model performances based on the uniform and normal priors are comparable. The distributions of all the evaluation metrics in the framework are significantly different for the flood model under different high-flow scenarios, and it further indicates that the metrics are essentially random statistical variables. Additionally, the white-noise error in observations has the least impact on the metrics, while the positive and the negative biases would have opposite impacts, which depends on whether the model overestimated or underestimated the hydrologic variable.

An Integrated Evaluation Framework based on Generalized Likelihood Uncertainty Estimation for Quantifying Uncertainty in Flood Modeling

Tao Huang and Venkatesh Merwade

Lyles School of Civil Engineering, Purdue University

December 2022

AGU FALL MEETING

Introduction

➢Flood & Flood model

- Flooding is one of the most devastating natural disasters in the world
- Evaluation of reliability and accuracy of model predictions is a critical issue
- How to demonstrate and quantify the uncertainty in evaluation metrics?

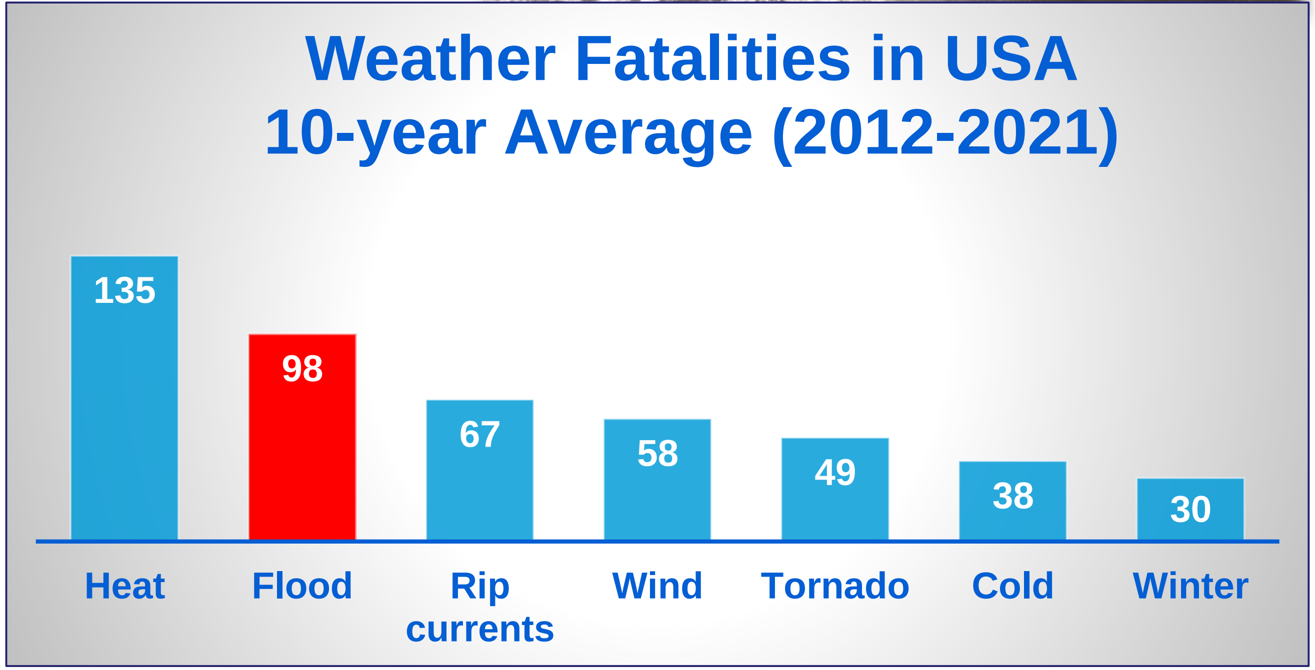
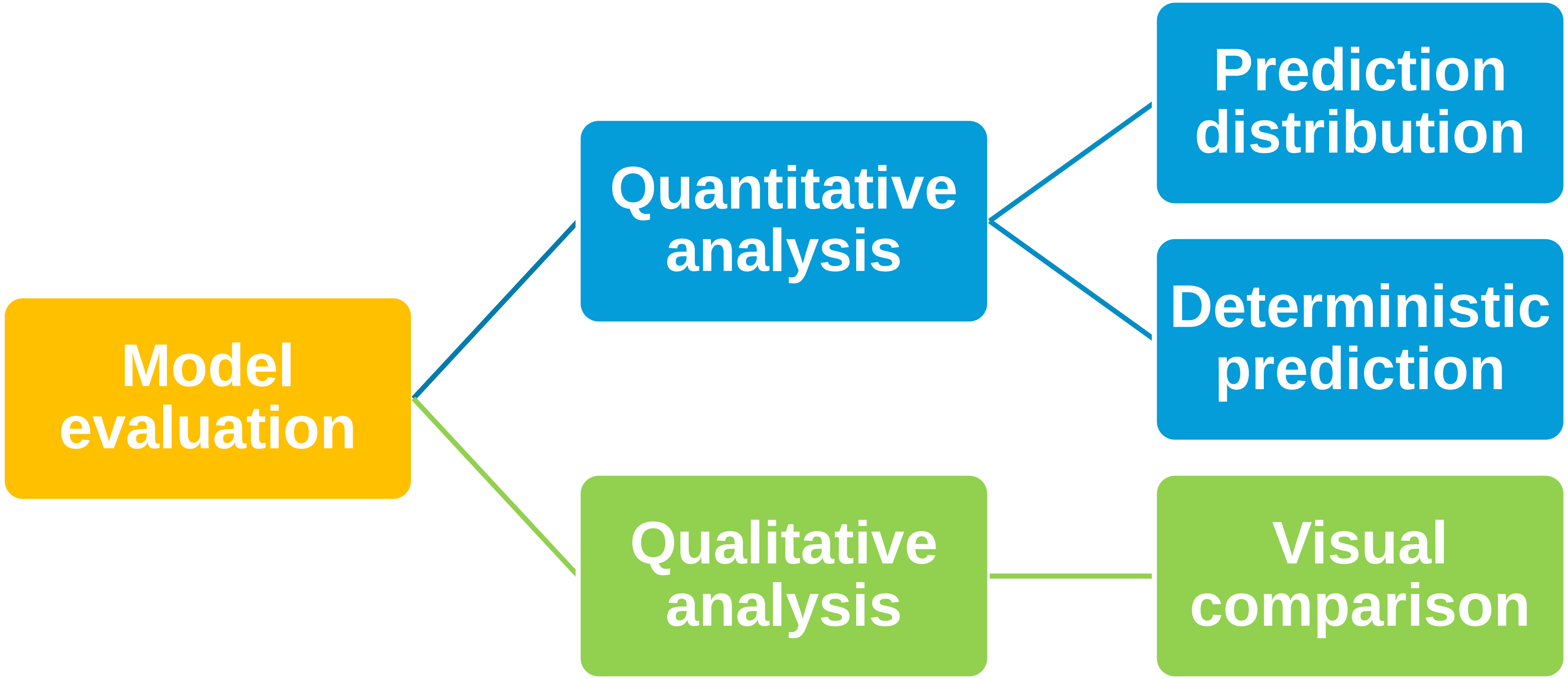


Image source: https://www.usatoday.com/picture-gallery/news/weather/2022/06/14/major-flooding-mudslides-yellowstone-national-park/7618177001/
Data source: https://www.weather.gov/hazstat/

Introduction

- **Uncertainty in evaluation criteria for flood models**



➢ **Statistical metrics**

❑ Reliability of prediction distribution

❑ *MSE/RMSE, NSE, KGE, R^2, etc.*

➢ **Limitations**

❑ No "perfect" single metric

❑ No broad consensus on uncertainty quantification for flood models

(Choi, 2022; Clark et al., 2021; Knoben et al., 2018; D. Liu, 2020; Rogelis et al., 2016; Siqueira et al., 2018; Towner et al., 2019)

Introduction

➢ **Objectives:** quantify uncertainty in evaluation metrics

✓ revisit the statistical meanings of existing metrics and propose an integrated evaluation framework

✓ investigate the effect of different prior distributions in GLUE analysis on the uncertainty metrics

✓ evaluate the effect of different high-flow scenarios on the uncertainty metrics

✓ explore the impact of different types of measurement errors on the uncertainty metrics

Integrated Evaluation Framework

➢ Uncertainty coefficients

$$UC1 = \frac{N_{obs-90\%}}{n} \cdot 100\%$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_{obs,i} - y_{sim,i})^2}{n}}$$

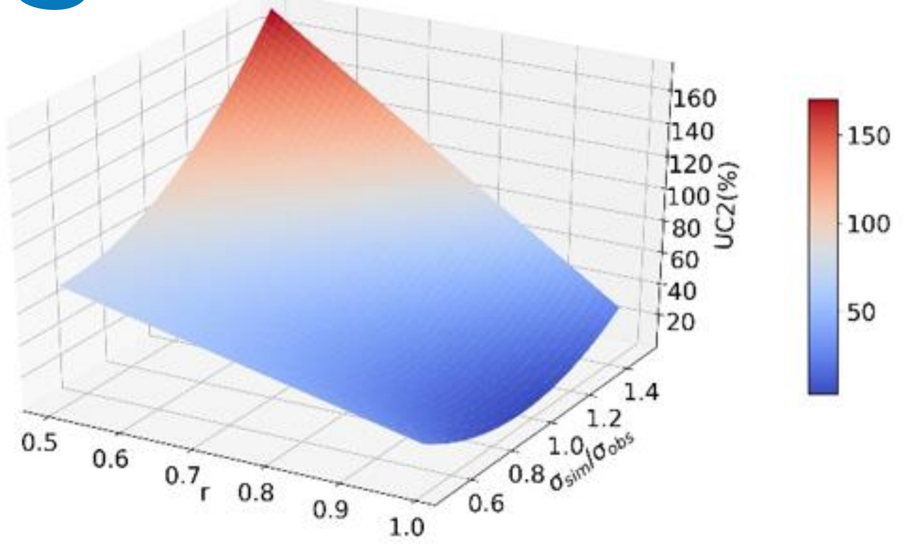
$$UC2 = 1 - NSE = \frac{\sum_{i=1}^{n}(y_{obs,i} - y_{sim,i})^2}{\sum_{i=1}^{n}(y_{obs,i} - \overline{y}_{obs,i})^2} = \frac{\sum_{i=1}^{n}(y_{obs,i} - y_{sim,i})^2 / n}{\sum_{i=1}^{n}(y_{obs,i} - \overline{y}_{obs,i})^2 / n} = \frac{RMSE^2}{\sigma_{obs}^{2}} \cdot 100\%$$

$$UC3 = 1 - KGE = \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\overline{y}_{sim}}{\overline{y}_{obs}} - 1)^2} \cdot 100\%$$

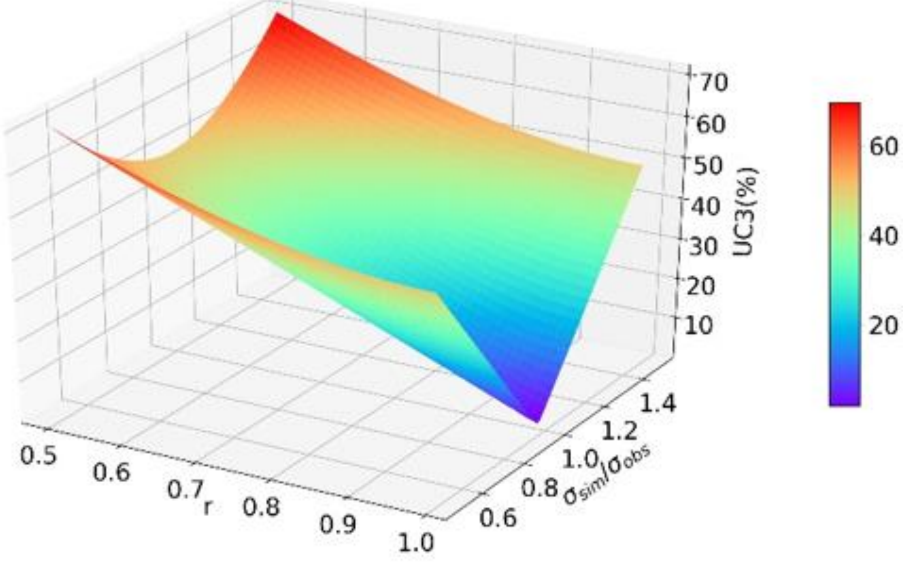
$$UC4 = 1 - R^2 = \frac{\sum_{i=1}^{n}(y_{obs,i} - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_{obs,i} - \overline{y}_{obs,i})^2} \cdot 100\% = \frac{\sum_{i=1}^{n}(y_{obs,i} - Slope \cdot y_{sim,i})^2}{\sum_{i=1}^{n}(y_{obs,i} - \overline{y}_{obs,i})^2} \cdot 100\%$$

(Anscombe, 1973; Nagelkerke, 1991; Nash and Sutcliffe, 1970; Gupta et al., 2009)

Integrated Evaluation Framework



(a) *UC2*.



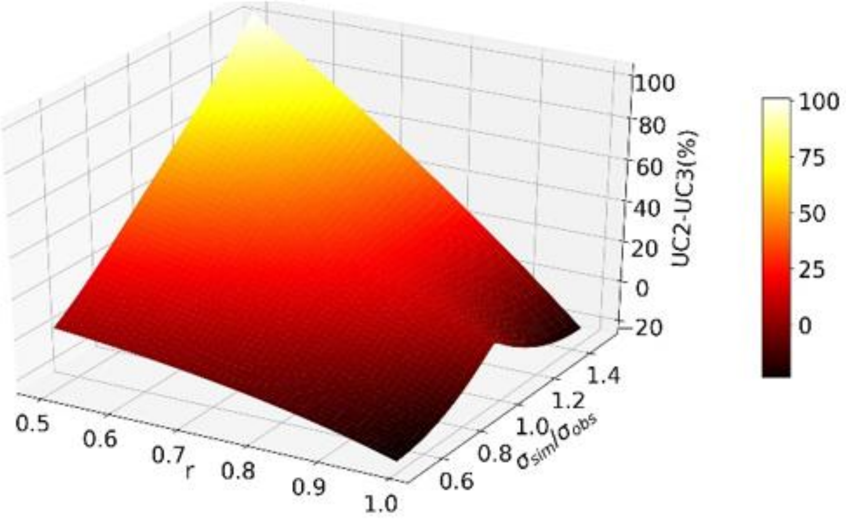
(b) *UC3*.

Assuming

$$\boxed{\overline{y}_{sim} = \overline{y}_{obs}}$$

$$UC2 = 1 - NSE = \left[1 - r^2 + \left(r - \frac{\sigma_{sim}}{\sigma_{obs}}\right)^2 + \left(\frac{\overline{y}_{sim} - \overline{y}_{obs}}{\sigma_{obs}}\right)^2\right] \cdot 100\%$$

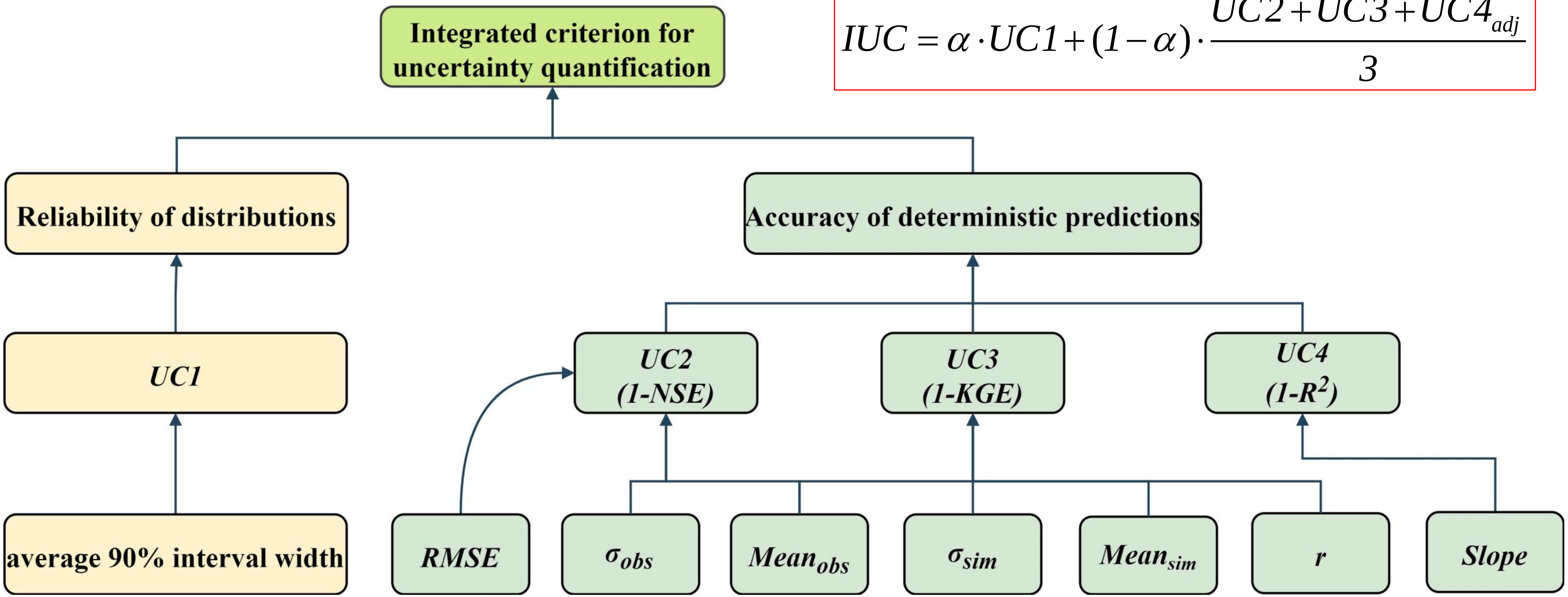
$$UC3 = 1 - KGE = \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\overline{y}_{sim}}{\overline{y}_{obs}} - 1\right)^2} \cdot 100\%$$



(c) *UC2-UC3*.

Integrated Evaluation Framework

$$IUC = \alpha \cdot UC1 + (1-\alpha) \cdot \frac{UC2 + UC3 + UC4_{adj}}{3}$$



Integrated evaluation framework for flood models

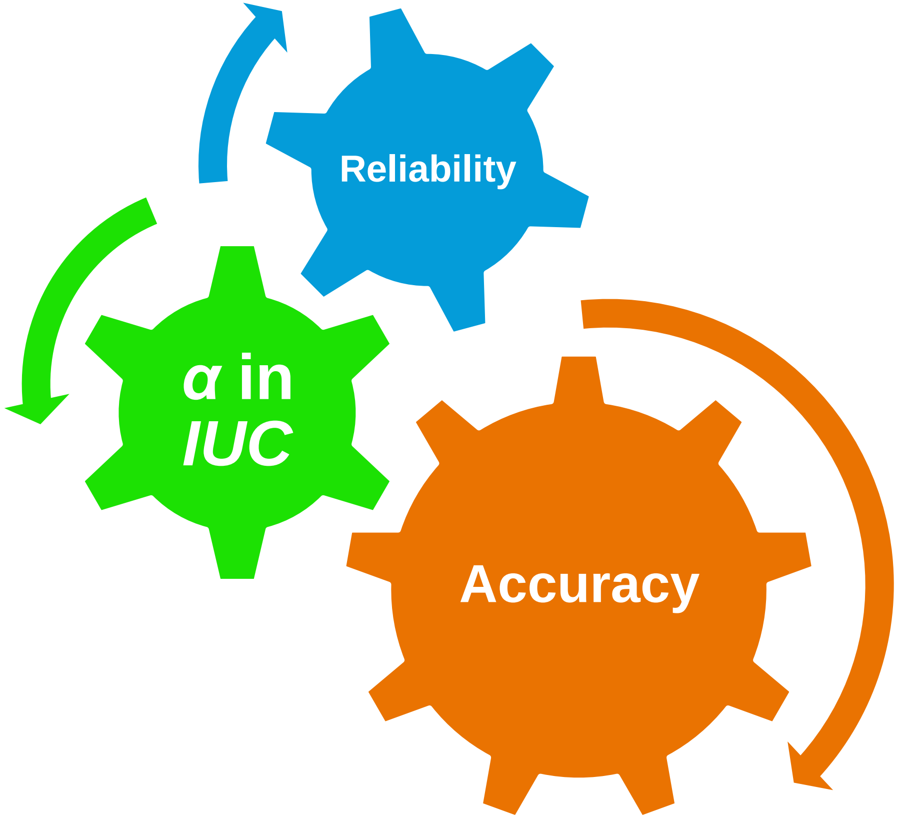
$$UC4_{adj} = 1 - R^2 + |1 - Slope|$$

Integrated Evaluation Framework

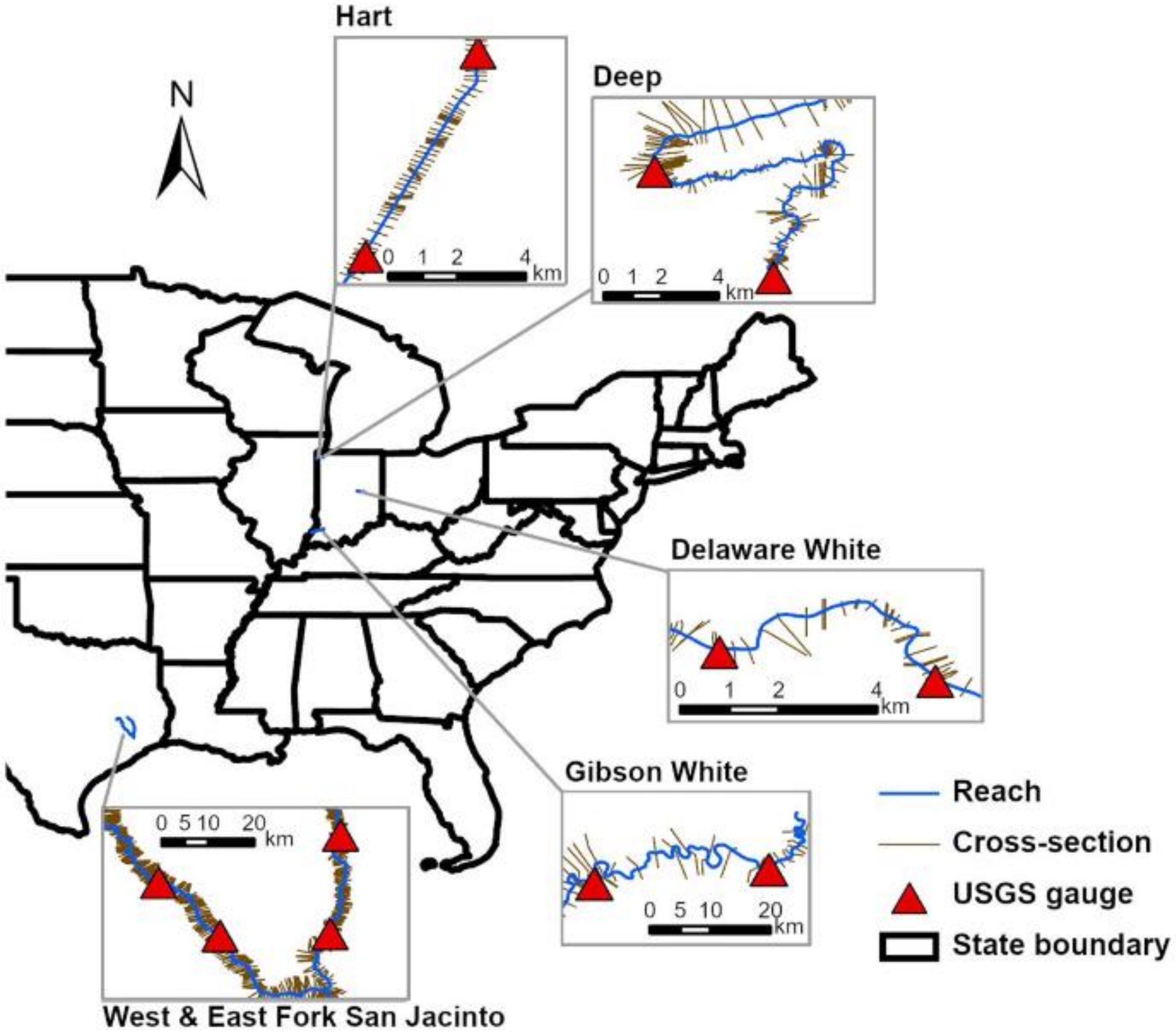
- Values of the empirical factor (*α*) in *IUC*

$$IUC = \alpha \cdot UC1 + (1-\alpha) \cdot \frac{UC2 + UC3 + UC4_{adj}}{3}$$

90% prediction interval of water stage distribution	α
average interval width ≤ 0.3 m	0.1
0.3 m < average interval width ≤ 0.9 m	0.25
0.9 m < average interval width ≤ 1.2 m	0.5
1.2 m < average interval width ≤ 1.8 m	0.25
average interval width > 1.8 m	0.1



Study Area and Data



Study Stream (State-No.)	Channel length (km)	Average channel width (m)	Channel slope (%)
Hart (IN-1)	8.45	16	0.1037
Deep (IN-2)	19.55	48	0.0095
Delaware White (IN-3)	6.76	64	0.0631
Gibson White (IN-4)	70.38	182	0.0087
West Fork San Jacinto (TX-1)	56.31	227	0.1624
East Fork San Jacinto (TX-2)	50.11	76	0.0438

FEMA model source: https://dnrmaps.dnr.in.gov/appsphp/model/index.php
https://webapps.usgs.gov/infrm/estbfe/

Simulation period: 200 days (summer & fall in 2021)

Methodology

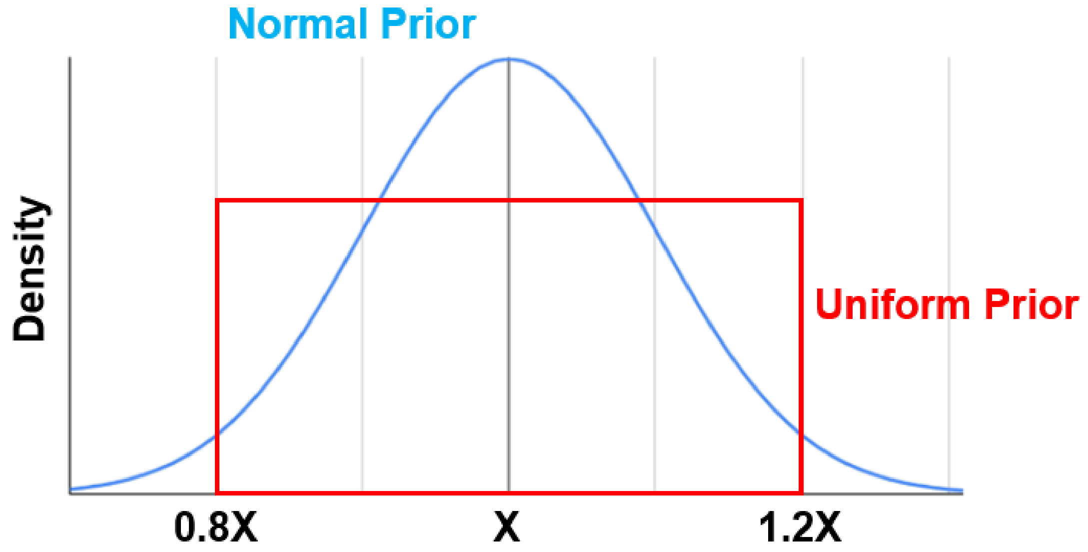
➢ Uncertainty quantification based on GLUE for FEMA models (1D HEC-RAS)

- **Generalized likelihood uncertainty estimation (GLUE)** incorporates both Monte Carlo sampling and the Bayesian analysis (Beven and Binley, 1992).

Sampling from prior distributions of uncertainty sources → **Model ensemble and multiple model runs** → **Multiple results and compare to observations based on Likelihood** → **Estimate uncertainty based on behavioral outputs**

Methodology

➢ **Uncertainty quantification based on GLUE for FEMA models (1D HEC-RAS)**



Uncertainty Type	Uncertainty Source	Prior Distribution-1 (Uniform)	Prior Distribution-2 (Normal)
Model Parameter	Channel roughness (n)	*U* (0.8n, 1.2n)	*N* (n, 0.1n)
Input Data	Upstream flow input (Q)	*U* (0.8Q, 1.2Q)	*N* (Q, 0.1Q)

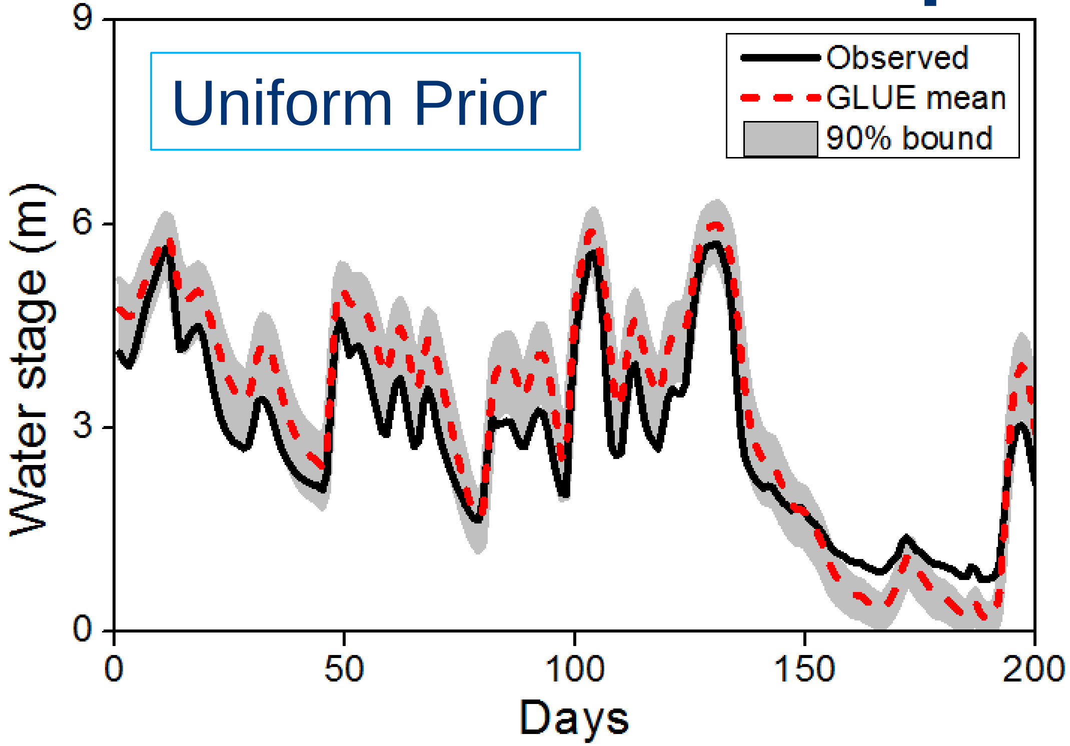
400 (=20×20) model configurations (plan files) in HEC-RAS

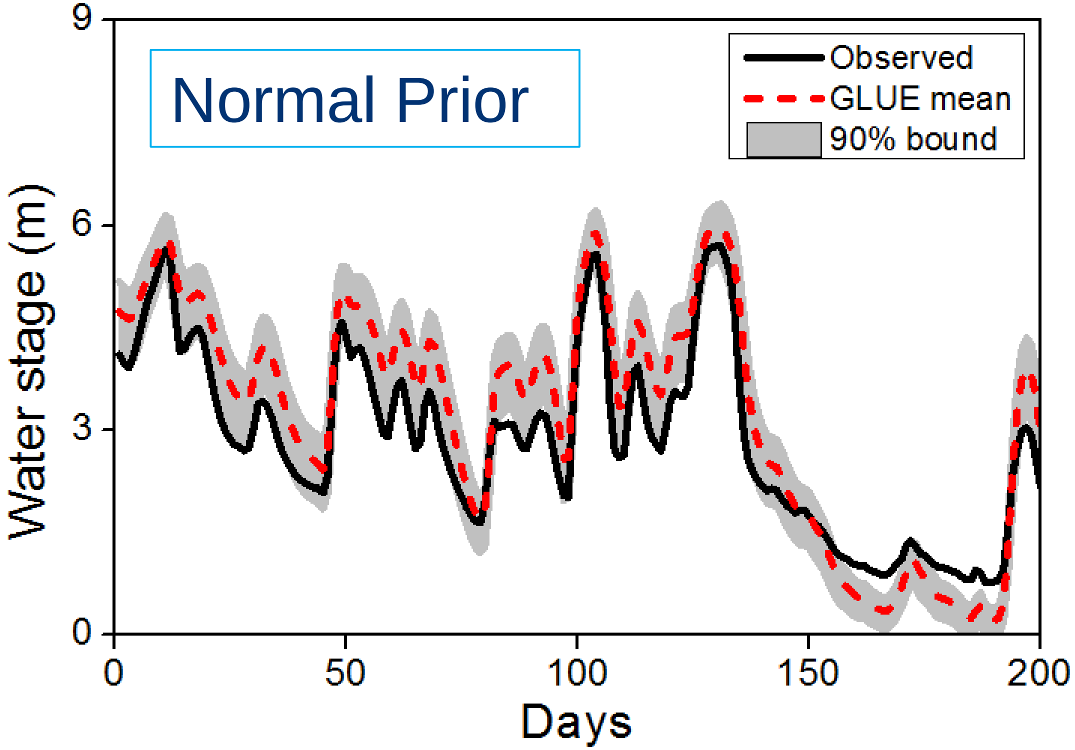
- **Likelihood function:** $$L(f_k \mid D) = \frac{1}{\sum_{t=1}^{T} (f_{k,t} - y_t^{obs})^2}$$

- **Cut-off threshold:** top 75%

Results and Discussion

➢ Effect of different prior distributions in GLUE

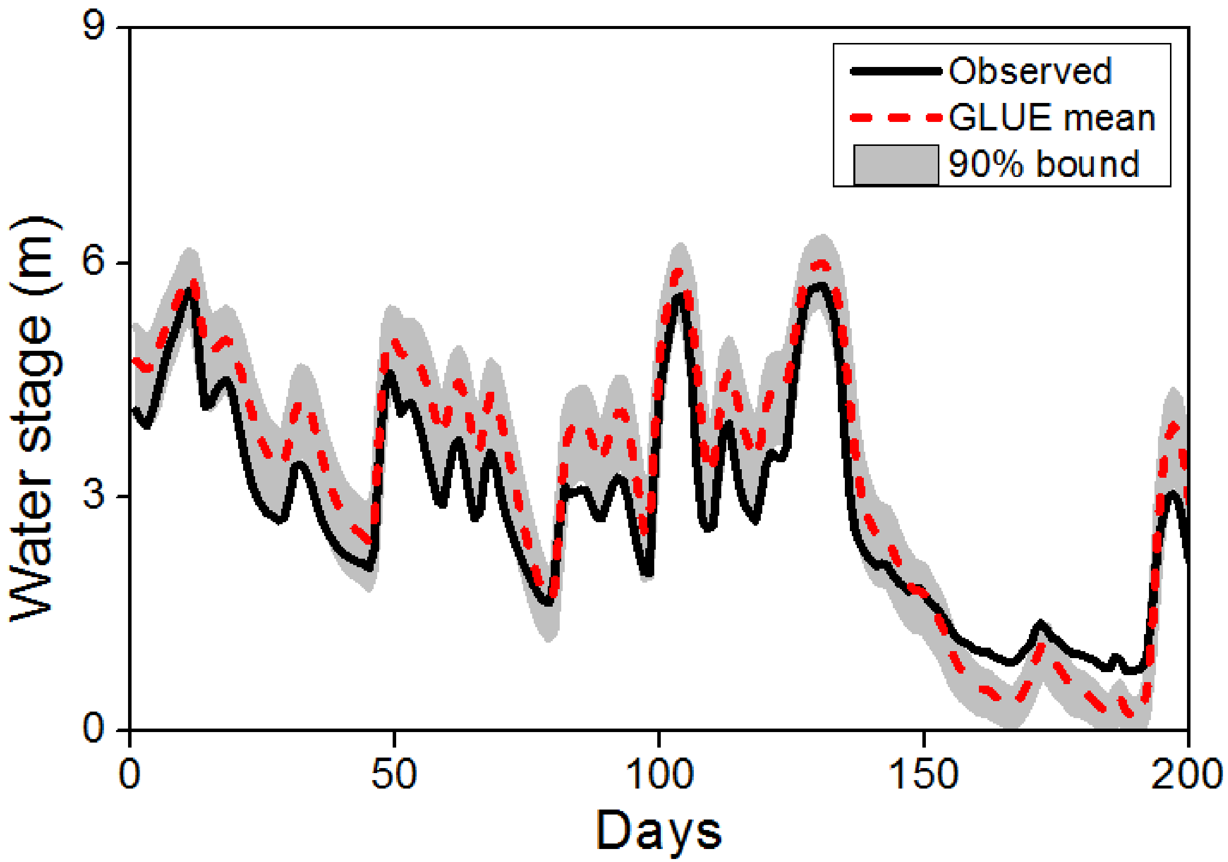




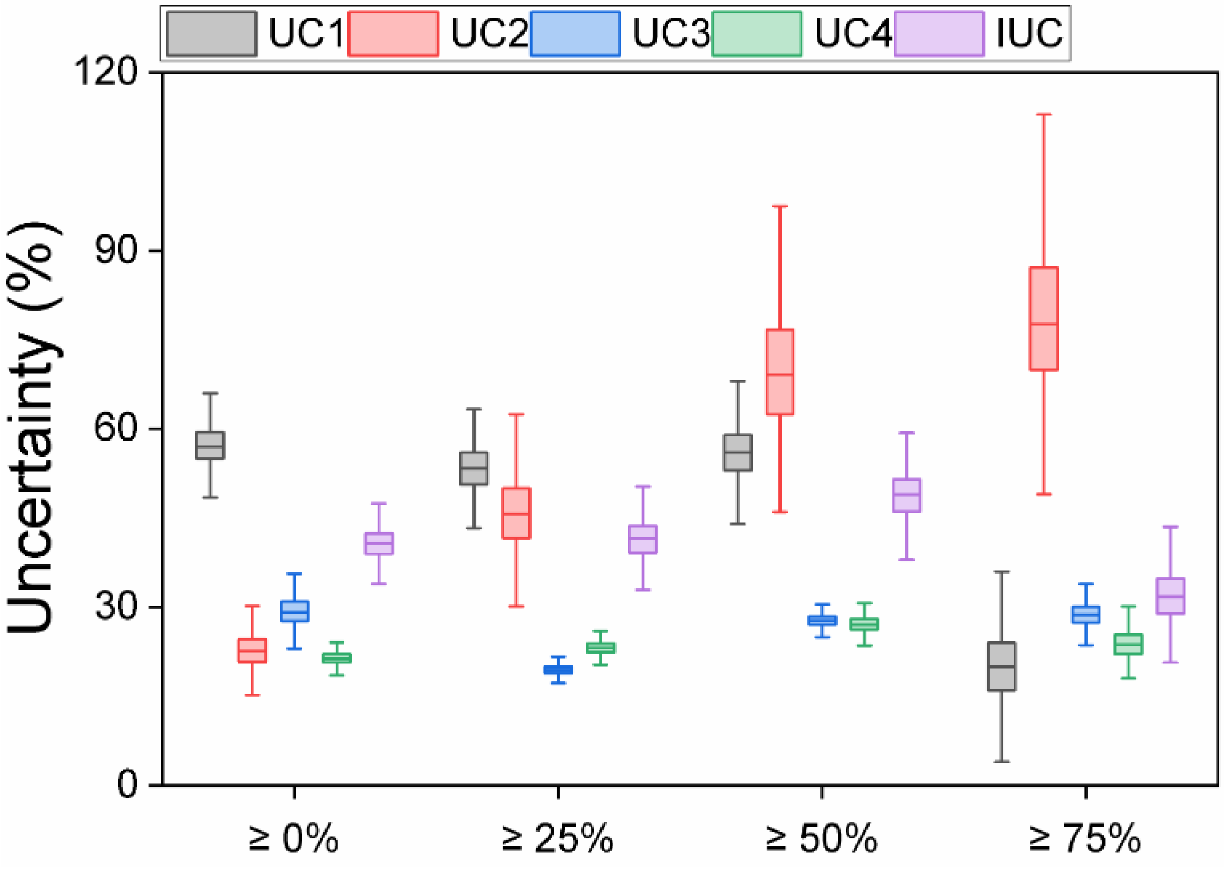
Model	Type	Mean (*U*)	90% CI (*U*)	Mean (*N*)	90% CI (*N*)
IN-4	*UC1 (%)*	57.36	[52.00, 63.50]	59.67	[54.00, 65.50]
	UC2 (%)	22.88	[18.72, 27.86]	23.06	[18.76, 28.06]
	UC3 (%)	29.36	[25.34, 33.41]	29.22	[25.56, 33.47]
	UC4 (%)	21.42	[19.71, 23.15]	21.58	[19.88, 23.35]
	IUC (%)	40.55	[35.20, 45.01]	39.62	[30.95, 45.91]

Results and Discussion

➢ Evaluation under different high-flow scenarios



Water stage predictions of IN-4 model

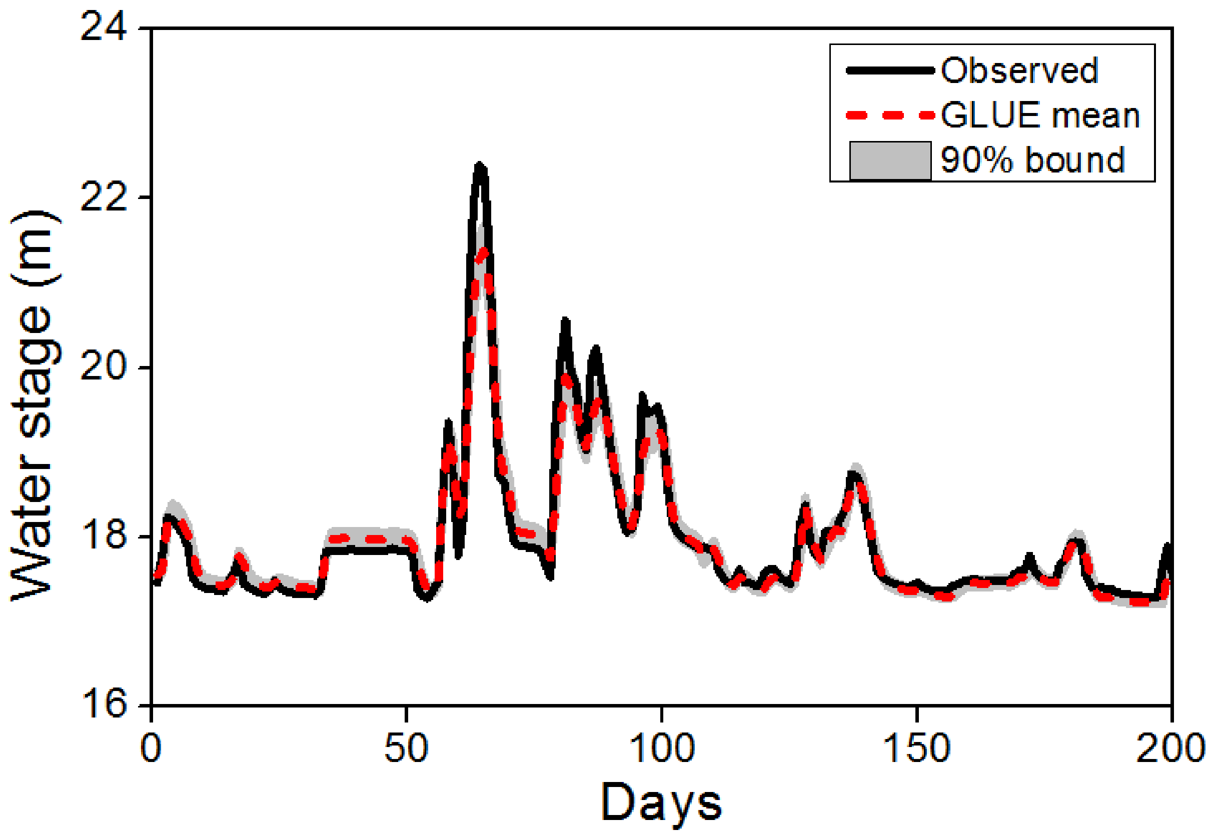


UC distributions of IN-4 model

Results and Discussion

➢ Integrated evaluation framework for flood models

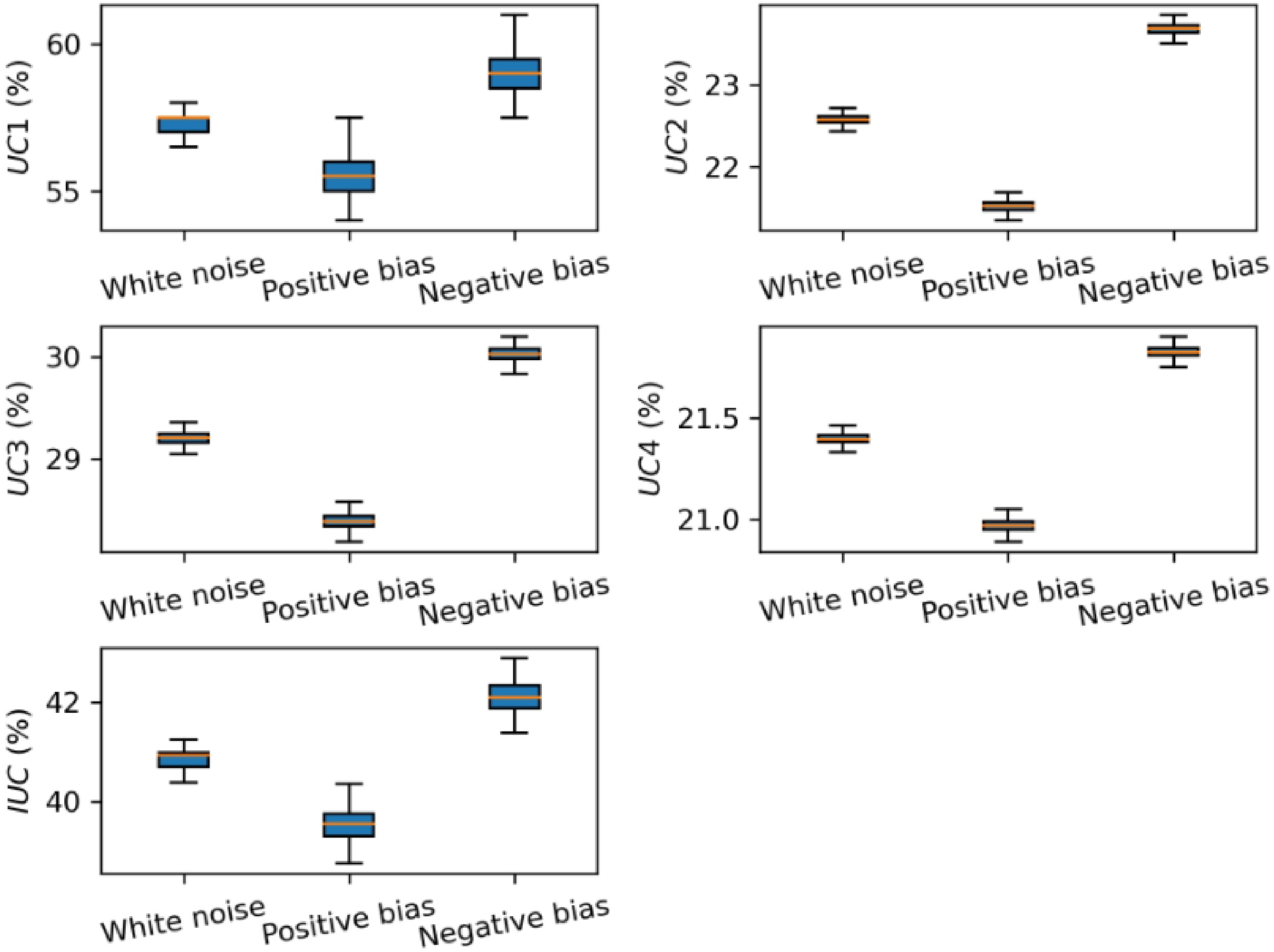
≥ Per (%)	*IUC* (%)	90% width (m)	*RMSE* (m)	Ratio of *sd*	Ratio of *mean*	*r*	*slope*
0	14.96	0.16	0.23	0.85	1.00	0.98	1.00
10	14.92	0.17	0.24	0.84	1.00	0.98	1.00
20	15.19	0.18	0.26	0.83	1.00	0.97	1.00
30	15.84	0.20	0.27	0.82	1.00	0.97	1.00
40	17.57	0.21	0.29	0.81	1.00	0.97	1.00
50	18.65	0.23	0.31	0.78	1.00	0.97	1.00
60	18.82	0.25	0.34	0.77	1.00	0.97	1.01
70	19.28	0.29	0.39	0.76	0.99	0.97	1.01
80	29.33	0.34	0.46	0.73	0.99	0.96	1.01
90	40.70	0.39	0.64	0.80	0.97	0.93	1.03



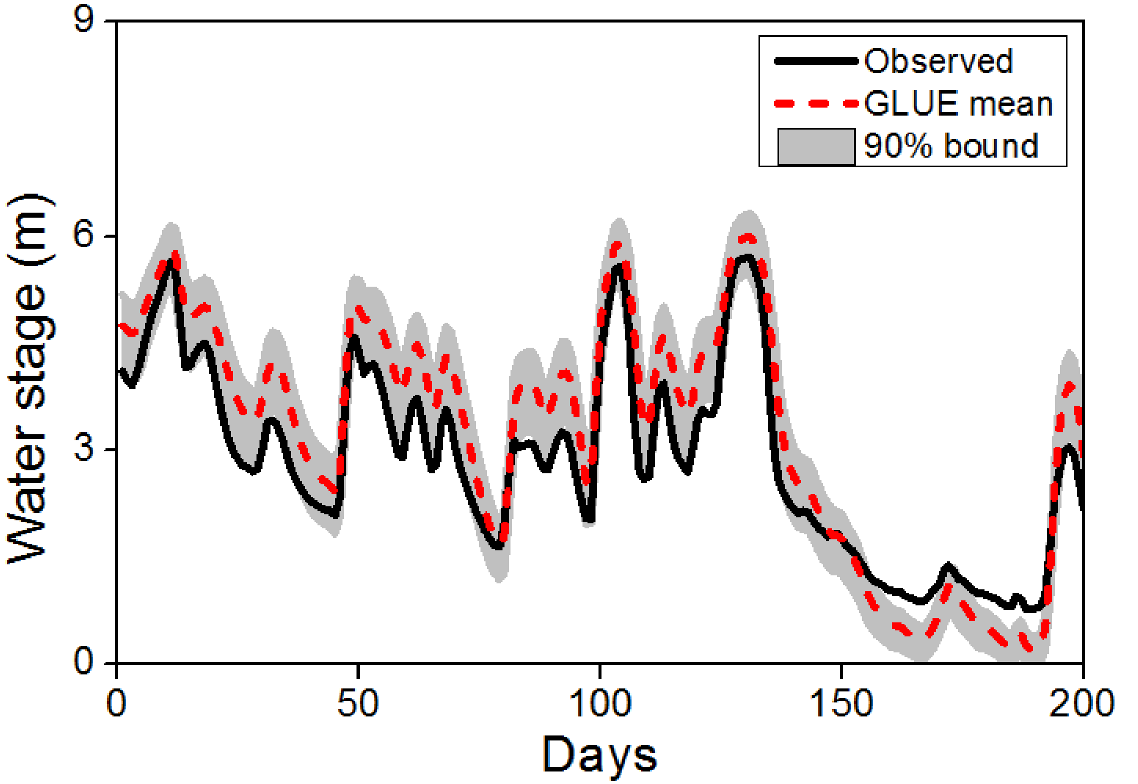
Water stage predictions of TX-1 model

Results and Discussion

➢ Impact of various measurement errors



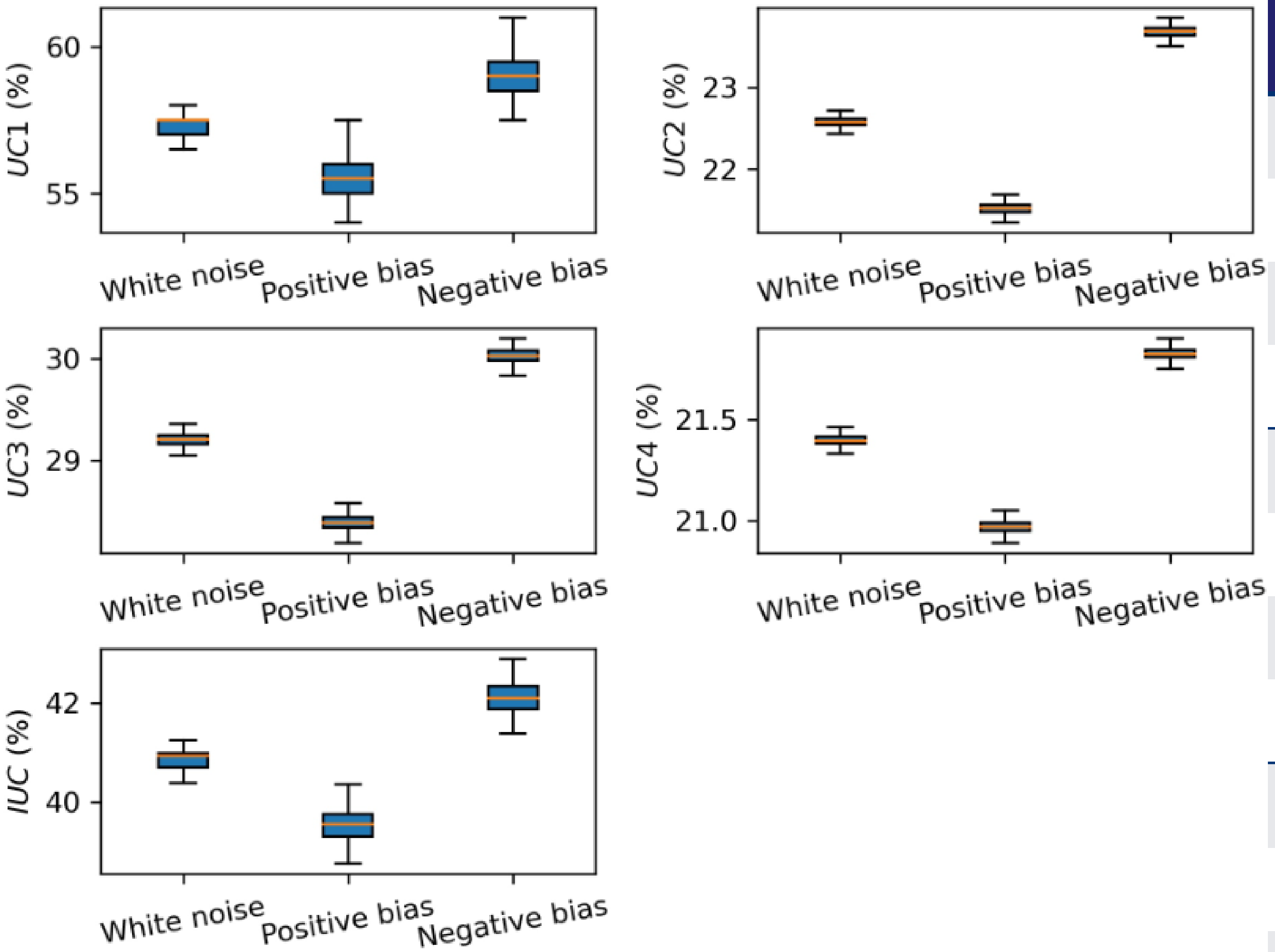
UC distributions of IN-4 model



Water stage predictions of IN-4 model

Results and Discussion

➢ Impact of various measurement errors



UC distributions of IN-4 model

Type of errors	Type of *UC*	Mean (%)	Relative change (%)	Type of *UC*	Mean (%)	Relative change (%)
No errors	*UC1*	57.50	/	*UC2*	22.57	/
WN		57.35	-0.26		22.58	0.04
PB		55.36	-3.72		21.52	-4.65
NB		59.11	2.80		23.69	4.96
No errors	*UC3*	29.20	/	*UC4*	21.40	/
WN		29.20	0.00		21.40	0.00
PB		28.39	-2.77		20.97	-2.01
NB		30.03	2.84		21.83	2.01
No errors	*IUC*	40.95	/			
WN		40.87	-0.20			
PB		39.49	-3.57			
NB		42.15	2.93			

Conclusions

$x;\ \theta$ → $f(x;\theta)$ → y ← Evaluation criteria

- A uniform prior in the GLUE analysis is adequate for the uncertainty quantification in the absence of solid prior knowledge.
- Evaluation metrics (*UC*s) are random variables: conditional on a specific flow scenario; present a statistical distribution.
- White-noise measurement errors have the least impact on *UC*s.
- The integrated evaluation framework based on GLUE can be applied to any other hydrologic variables.

THANK YOU

"No one trusts a model except the man who wrote it; everyone trusts an observation except the man who made it."

– Harlow Shapely

Email: huan1441@purdue.edu