Large-sample, high-resolution climate model data sets and flood modeling (rainfall runoff) of warm season floods in mountainous terrain

NASEM Workshop: Extreme Rainfall in Mountainous Terrain: Modeling and observational challenges for warm-season precipitation

Andrew Newman
Senior Scientist, NSF NCAR

4 November 2025

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Introduction

- Warm season extreme precipitation and subsequent floods are high impact. Globally, annual average losses (1979-2021) associated with flooding total <u>$388 billion</u>. WMO, 2023
- Increasing in frequency and severity
- Need to understand future change and our risk profiles

From Spectrum News 1, Louisville KY

Jack Morgan, Texas Public Radio

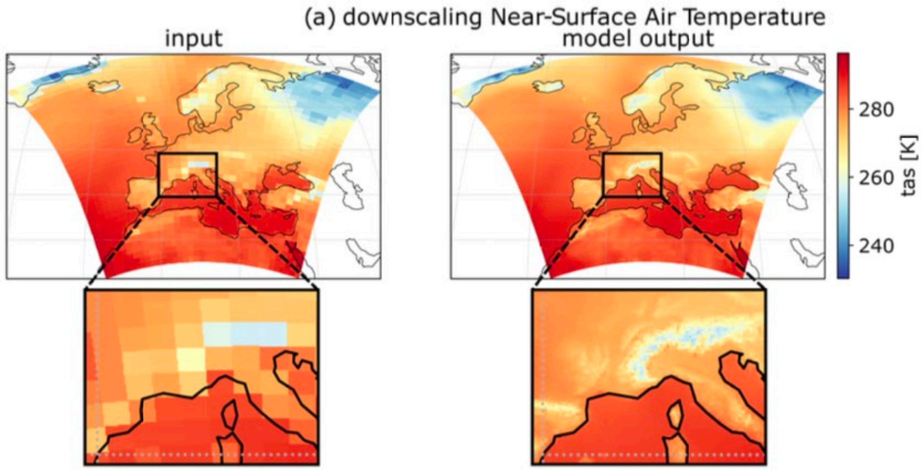
Definitions of Large-Sample, High-Resolution Hydroclimate Modeling and Data

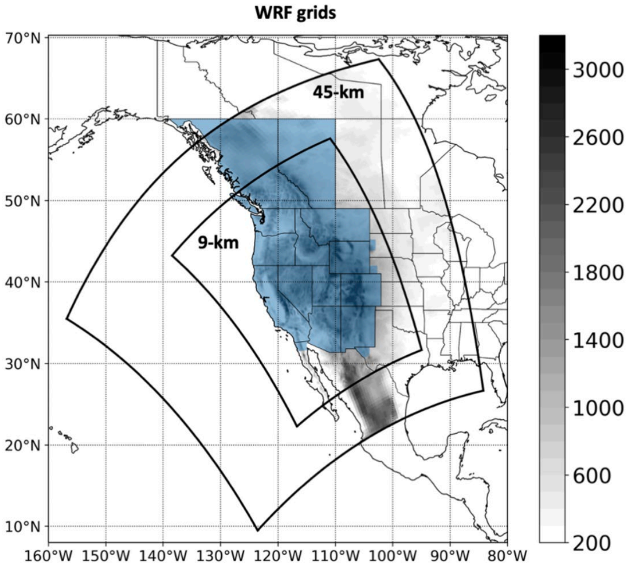
- High-resolution for this discussion:
  - Atmosphere: Orography, storm, meso-beta resolving (< 5-12 km grid spacing)
  - Hydrology: Mesoscale hydrologic modeling on the small catchment (HUC12 ~100 km^2) or similar grid spacing (5-10 km) scale (e.g. elevation gradients, headwaters)
  - Sub daily to daily output
- Large-sample definitions for this discussion:
  - Atmosphere: regional to continental scale; decades to centuries of simulation years
  - Hydrology: Hundreds to thousands of catchments or regional to continental scale; hundreds to thousands of events/hundreds to thousands of simulation years
- Why are these datasets important?
  - Diversity of events
  - Capture non-stationarities such as forced trends and changes in variability
  - Robust statistical sampling and evaluation

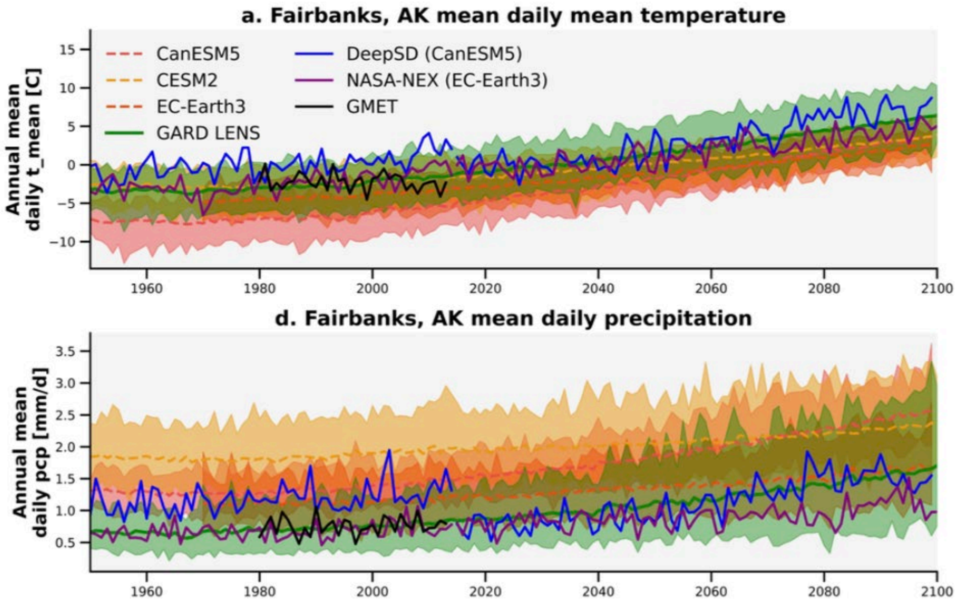
Current Large-Sample Climate Modeling and Datasets

- Rapid development of high-resolution regional and global datasets
  - Continuum of methods – simpler statistical, dynamic, AI/ML (downscaling and emulators)
- Downscaling of global Earth System Models and their large ensembles (e.g. SMILEs)
  - LOCA2, GARD-LENS, WUS-D3, d4PDF, CONUS-404

Downscaling Method Class	Class Definition	Method Examples
Simple Statistical	Statistical methods that primarily correct or adjust the mismatch between large- and local-scale (or model and observation) climatologies.	Bias Corrected Spatially Disaggregated (BCSD), Quantile Delta Method (QDM), Quantile-Quantile Delta Change Method (QQDC)
Complex Statistical	Statistical methods that use a form of linear or complex non-linear regression (machine learning) or analog methods to filter the information transfer between large- and local-scale (or model and observation) climatologies.	Statistical DownScaling Model (SDSM), LOcally Constructed Analogs (LOCA), Multivariate Bias Correction (MBCn), EnGARD, Convolutional Neural Networks (CNNs), Random Forests
Quasi-dynamical	Hybrid methods that combine statistical and dynamical tools to perform downscaling	Intermediate Complexity Atmospheric Research (ICAR) model
Fully Dynamical	The application of mesoscale atmospheric models using GCM or ESM boundary conditions to generate outcomes at local scales.	Weather Research and Forecasting (WRF) model



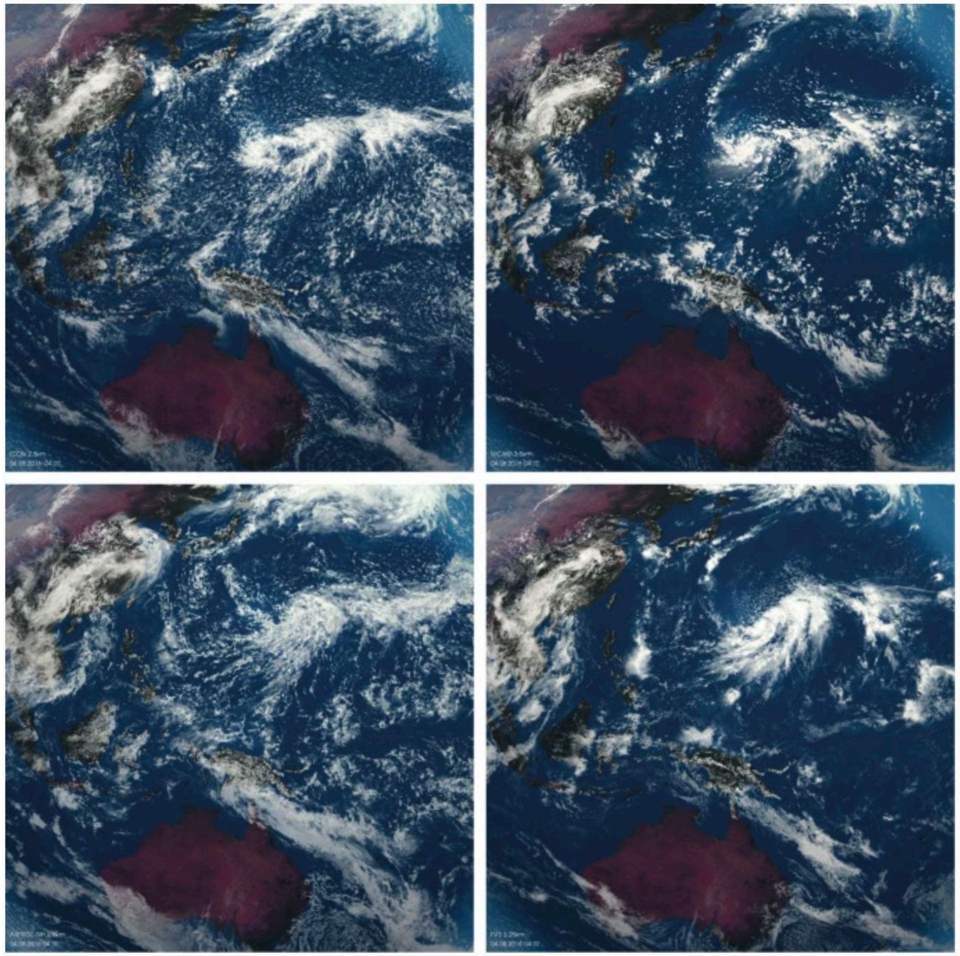




Schmude et al. (2024) Rahimi et al. (2024) Hartke et al. (2024)

Current Large-Sample Climate Modeling and Datasets

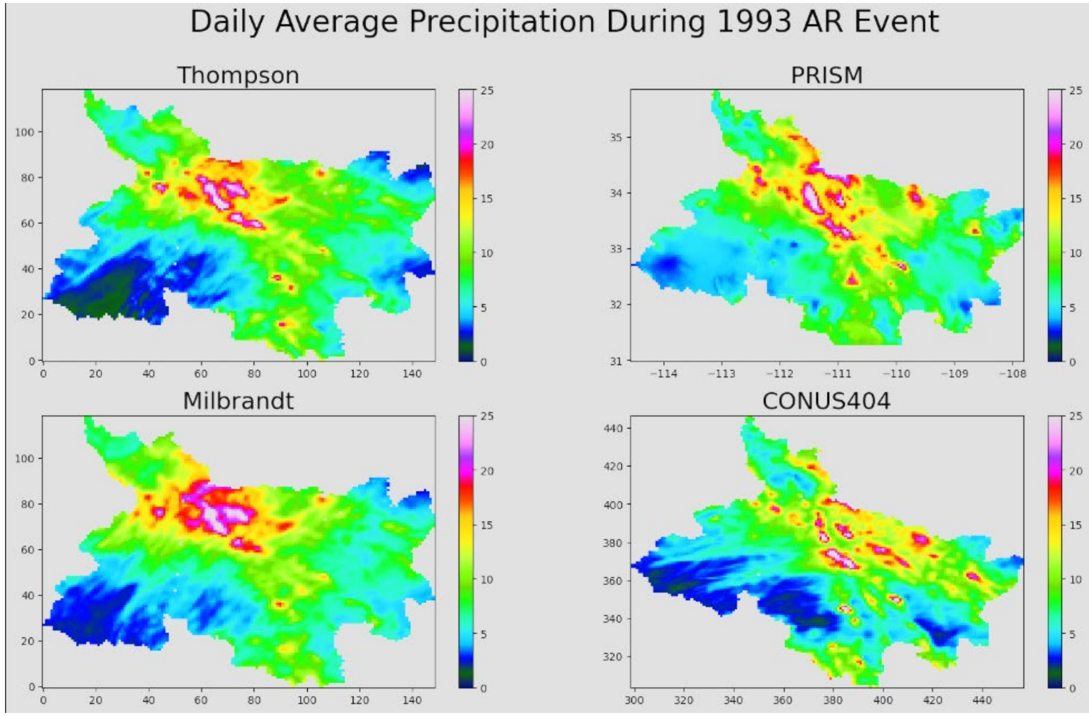
- Global storm resolving model (GSRM) simulations
  - DYAMOND simulations, other single model simulations (e.g. NICAM, SCREAM, MPAS-A)
  - Very limited temporal length, atmosphere only



Stevens et al. (2019)

Current Large-Sample Climate Modeling and Datasets

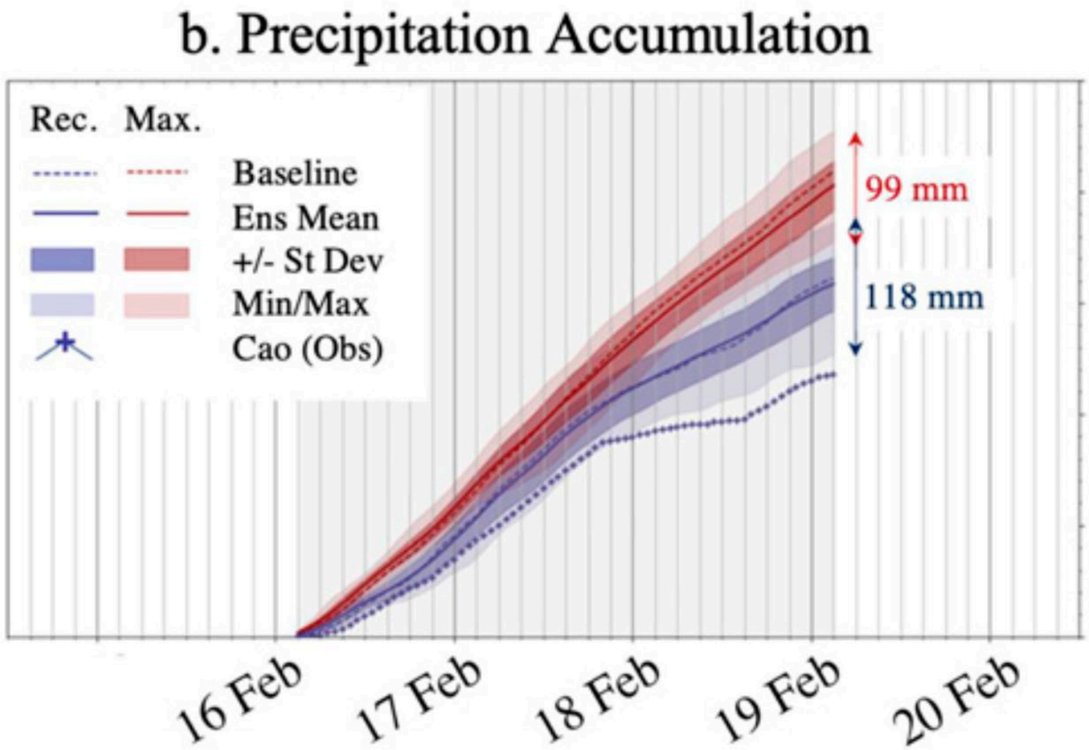
- Global storm resolving model (GSRM) simulations
  - DYAMOND simulations, other single model simulations (e.g. NICAM, SCREAM, MPAS-A)
  - Very limited temporal length, atmosphere only
- Event based datasets
  - Regional simulations of only events of interest
    - Boost event sample sizes
    - Process studies



Lybarger et al. (2025), in prep

Current Large-Sample Climate Modeling and Datasets

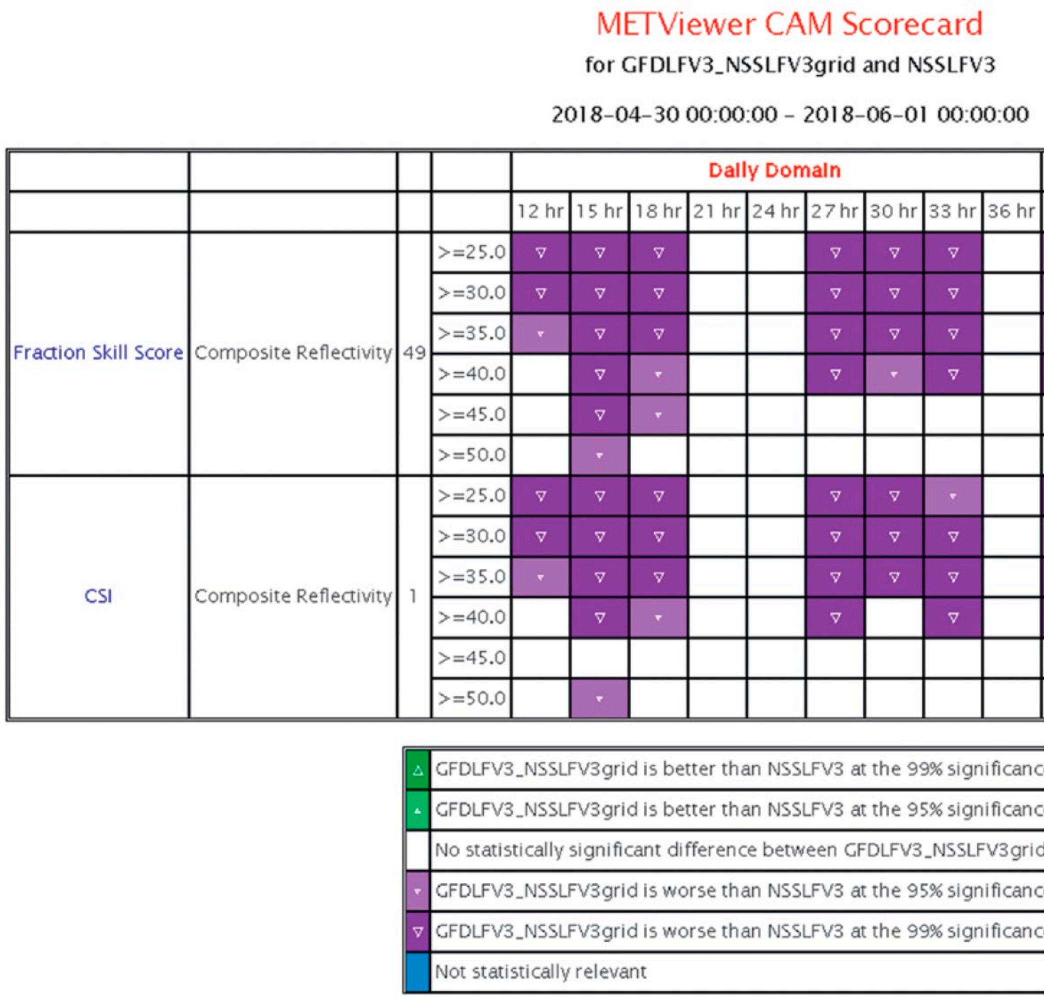
- Global storm resolving model (GSRM) simulations
  - DYAMOND simulations, other single model simulations (e.g. NICAM, SCREAM, MPAS-A)
  - Very limited temporal length, atmosphere only
- Event based datasets
  - Regional simulations of only events of interest
    - Boost event sample sizes
    - Process studies
  - PMP method development and evaluation
    - Uncertainty quantification
    - Model evaluation
    - Non-stationarity



Tarouilly et al. (2023)

Large-Sample Climate Modeling and Datasets - Gaps

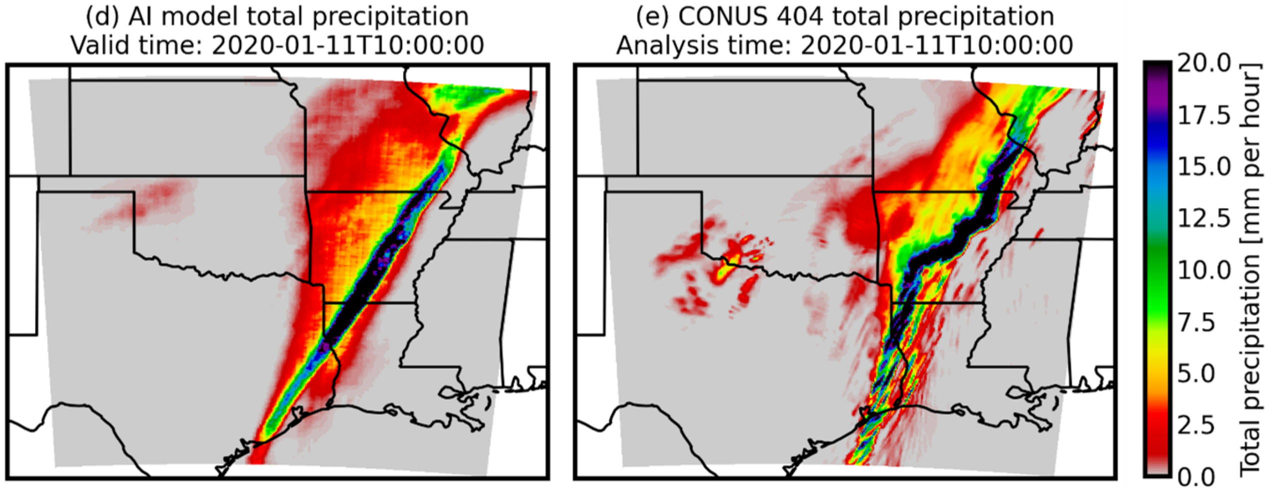
- What methods are fit (or adequate) for purpose?
  - Comprehensive evaluation and benchmarking is hard, but essential
  - Relative and absolute performance criteria needed
    - Better than previous model?
    - Good enough for decision making?



Gallo et al. (2019)

Large-Sample Climate Modeling and Datasets - Gaps

- What methods are fit (or adequate) for purpose?
  - Comprehensive evaluation and benchmarking is hard, but essential
  - Relative and absolute performance criteria needed
    - Better than previous model?
    - Good enough for decision making?
- Managing the AI/ML frontier
  - AI/ML is poised to exponentially increase methods and data
  - Need robust evaluation
  - Is explainable AI of value?
    - Increase partner trust?



Sha et al., in prep

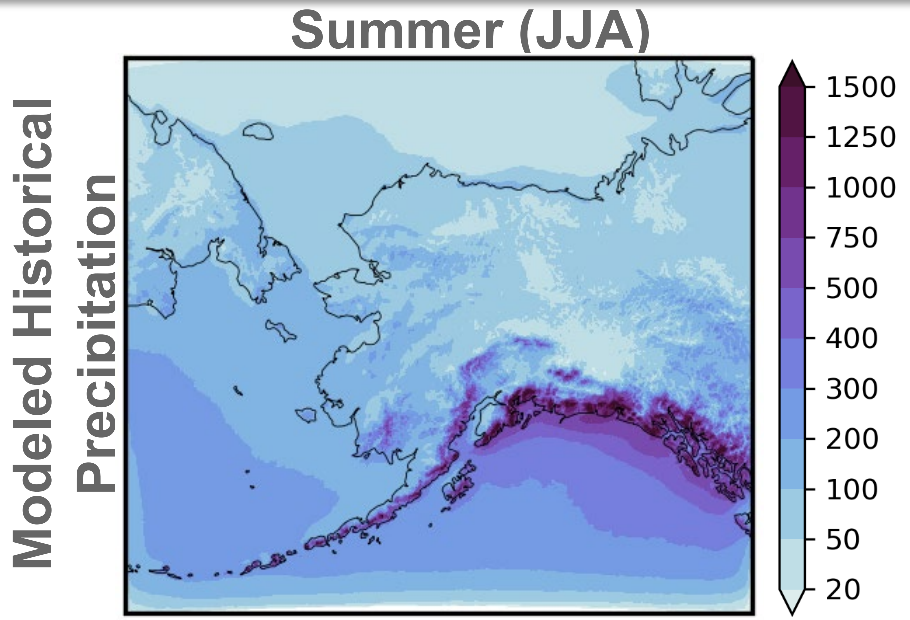
Large-Sample Climate Modeling and Datasets - Gaps

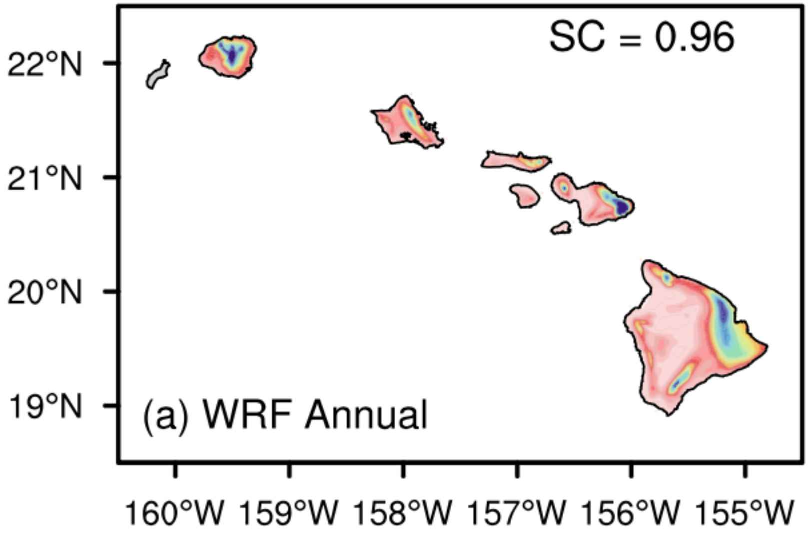
- Sample sizes and uncertainty quantification
  - Breaking computational bounds and advanced methods (e.g., Dr. Wright's talk next)
  - Continue GPU transition, how to integrate AI/ML into models, or just emulate everything?

OLCF at ORNL

Large-Sample Climate Modeling and Datasets - Gaps

- Sample sizes and uncertainty quantification
  - Breaking computational bounds and advanced methods (e.g., Dr. Wright's talk next)
  - Continue GPU transition, how to integrate AI/ML into models, or just emulate everything?
- Outside CONUS (OCONUS)
  - Few datasets for Hawai'i and Alaska
  - Fewer for Puerto Rico, other territories and possessions
  - Limited observational training/validation datasets

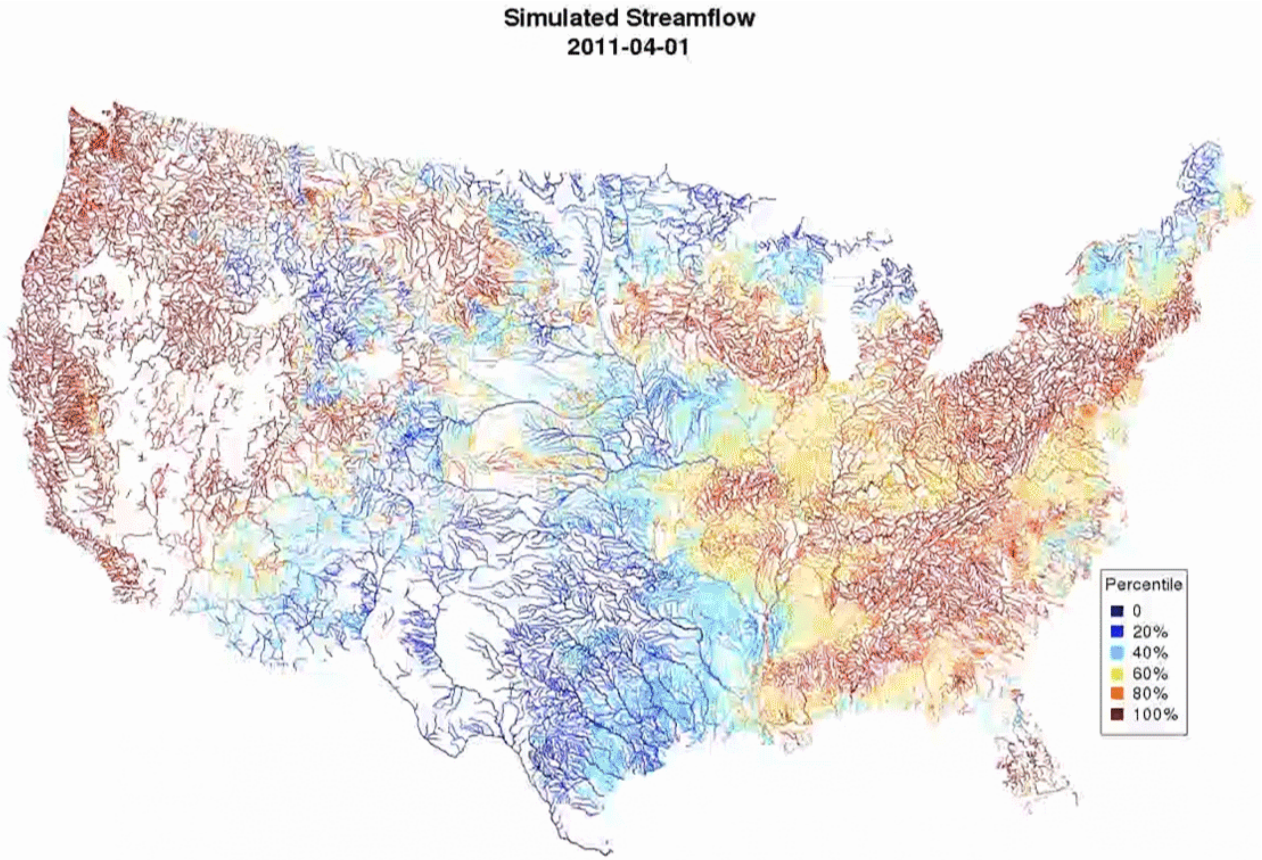




Cheng et al. (2025); Xue et al. (2020)

Large-Sample Surface Hydrologic Modeling and Datasets

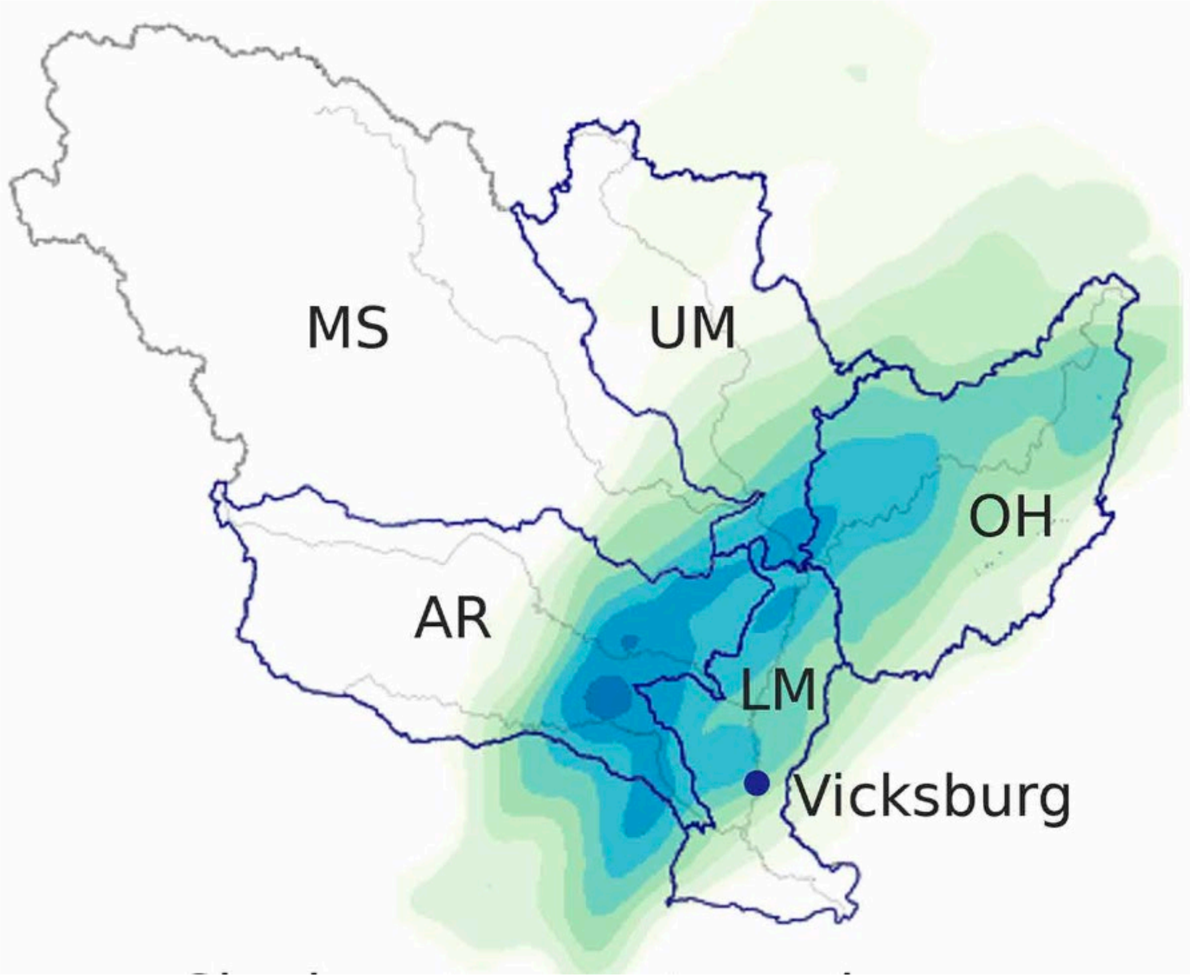
- Many national scale models and historical datasets
  - National Water Model (NWM), National Hydrologic Model (NHM), other models and groups produce retrospectives



USGS

Large-Sample Surface Hydrologic Modeling and Datasets

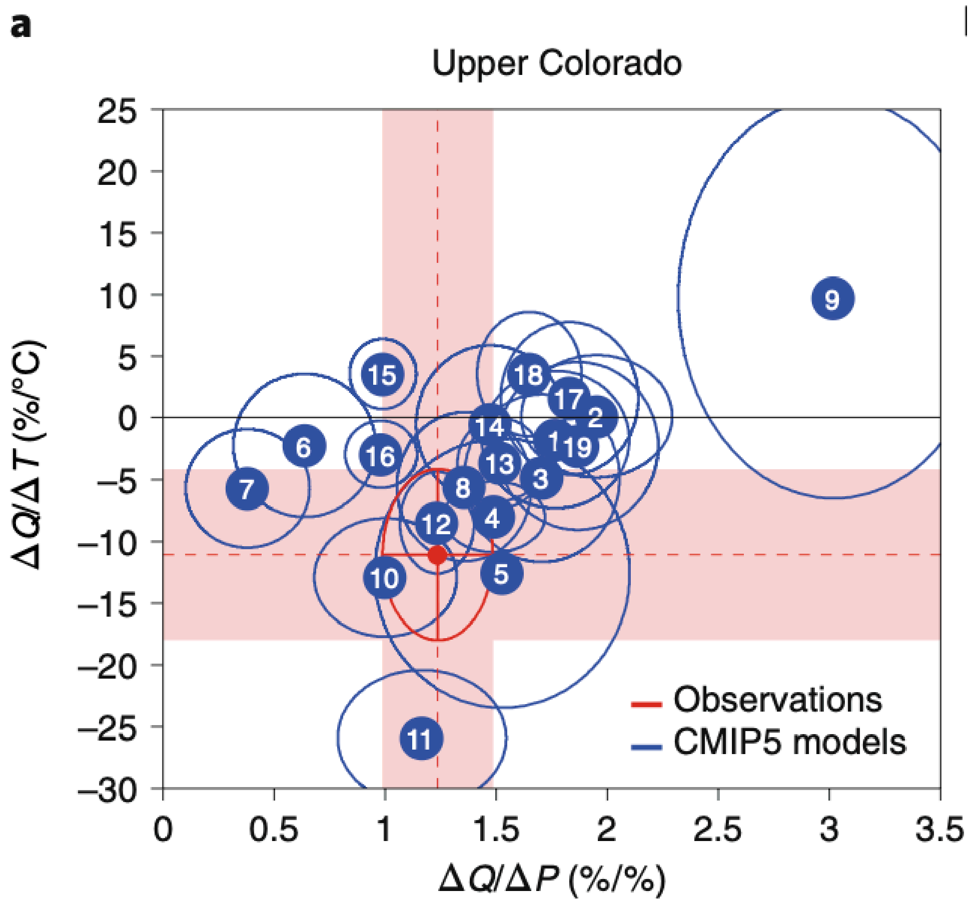
- Many national scale models and historical datasets
  - National Water Model (NWM), National Hydrologic Model (NHM), other models and groups produce retrospectives
- Fewer fit-for-purpose regional to continental scale datasets available for future periods
  - Lack of inputs
  - Model chain complexity
    - Many decision points
  - High computational demands



Liu et al. (2025)

Large-Sample Surface Hydrologic Modeling and Datasets - Gaps

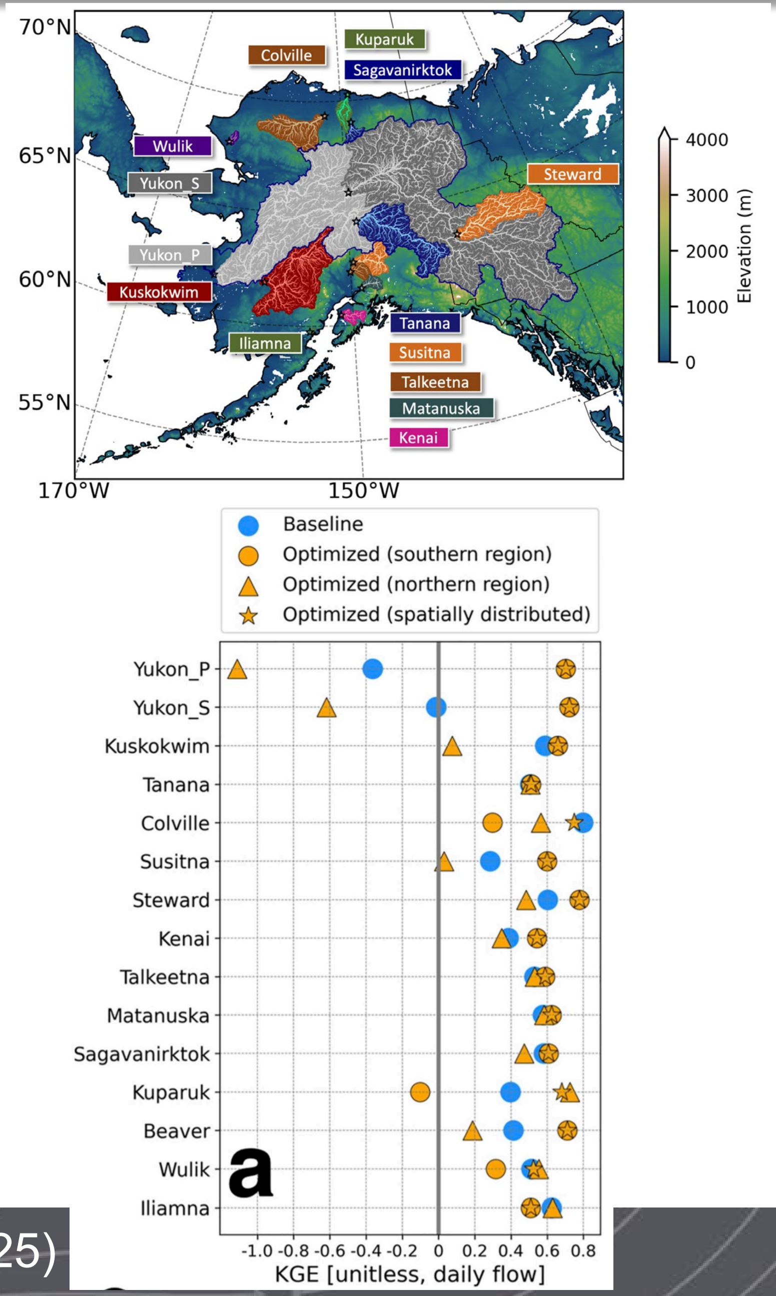
- Caution using most current land model output directly from high-resolution (or other) models
  - Often little effort is made for hydrologic performance
  - E.g. ESM land models may have unrealistic climate sensitivities (Lehner et al. 2019)



Lehner et al. (2019)

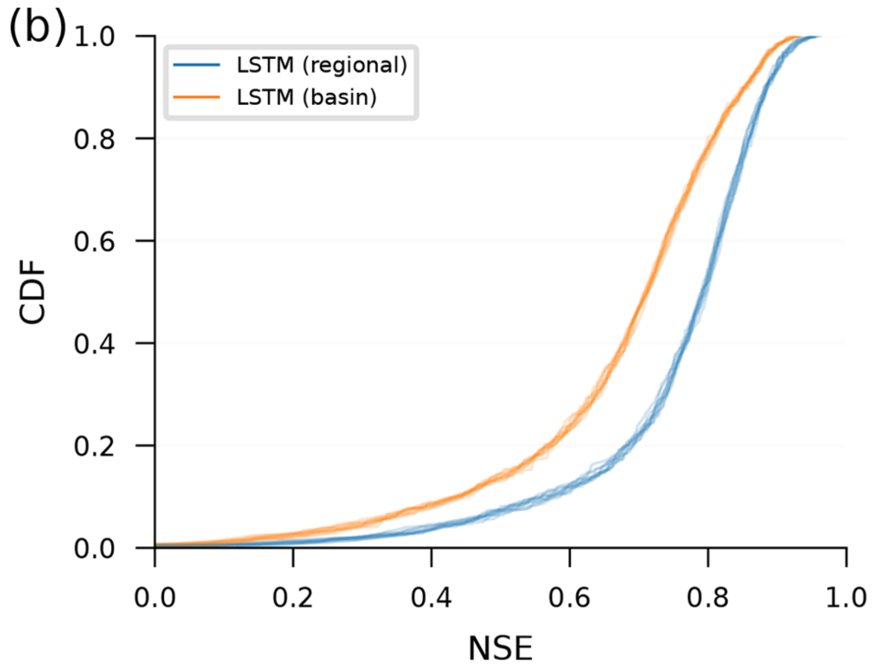
Large-Sample Surface Hydrologic Modeling and Datasets - Gaps

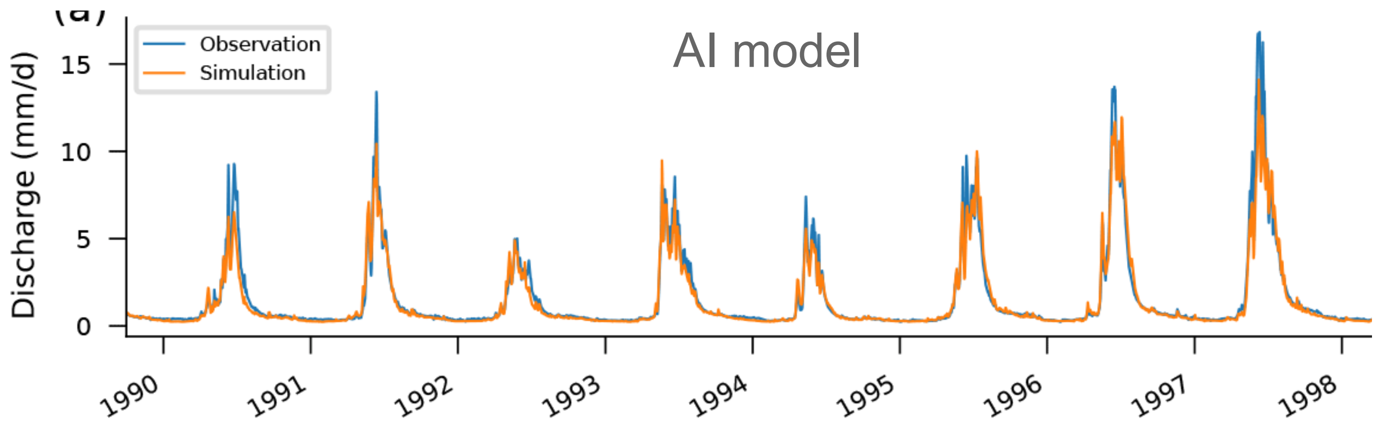
- Caution using most current land model output directly from high-resolution (or other) models
  - Often little effort is made for hydrologic performance
  - E.g. ESM land models may have unrealistic climate sensitivities (Lehner et al. 2019)
- There are efforts towards improving coupled model hydrology
  - Land-model configuration and optimization for hydrology
  - Fit-for-purpose inspired evaluation



Large-Sample Surface Hydrologic Modeling and Datasets - Gaps

- Benchmarking
  - Need to continue to push for robust, *a priori*, partner co-produced benchmarks
  - Open-source, community software for hydrology à la ILAMB or METplus
- AI/ML modeling
  - Continue integrating partners as soon as possible in development process
  - Explainability to build trust
  - Robust benchmarking and model intercomparisons

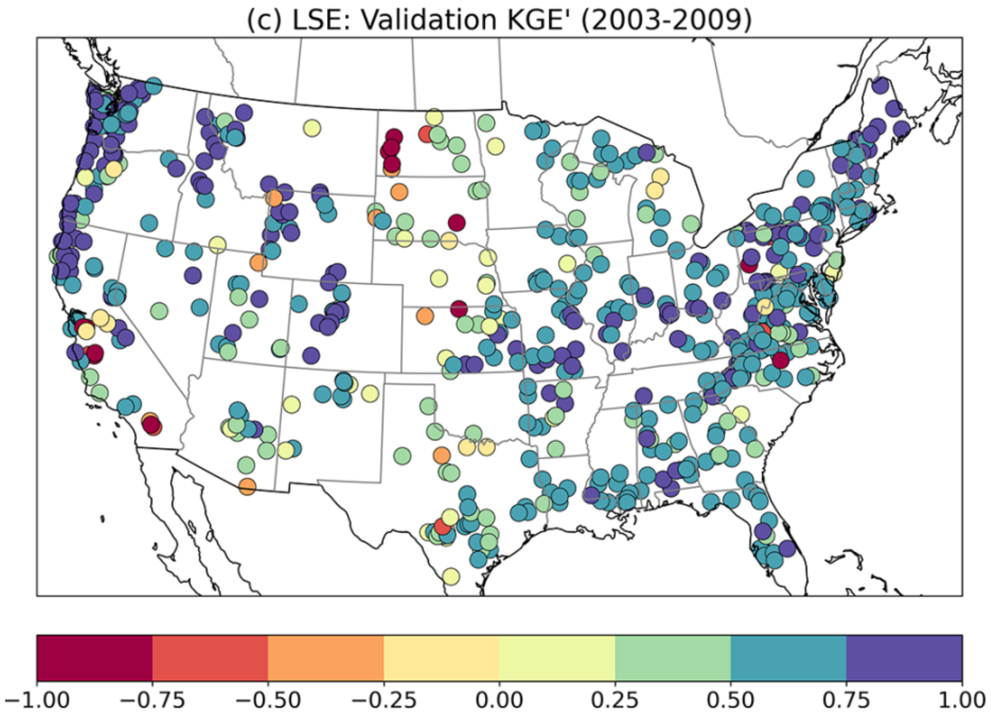




Kratzert et al. (2024)

Large-Sample Surface Hydrologic Modeling and Datasets - Gaps

- Regionalization and optimization
  - Model structure, process parameterization, and parameter optimization are unconstrained problems
  - AI/ML based optimization and regionalization is promising

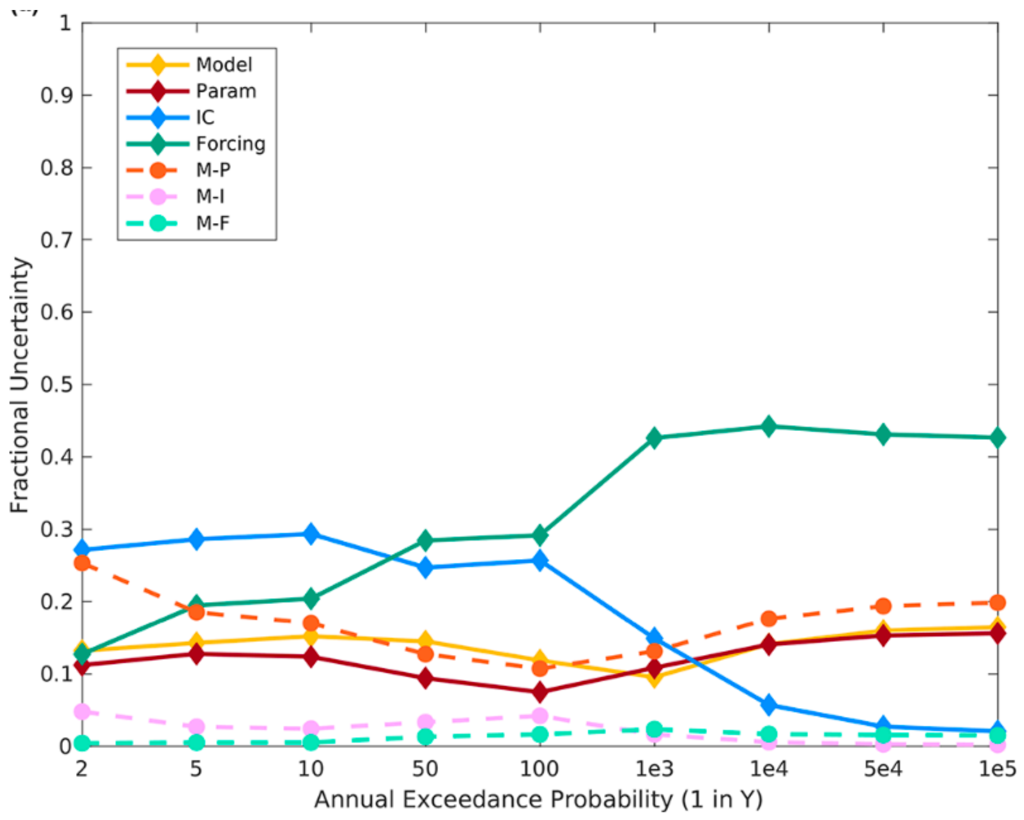


Farahani et al. (2025)

Large-Sample Surface Hydrologic Modeling and Datasets - Gaps

- Regionalization and optimization
  - Model structure, process parameterization, and parameter optimization are unconstrained problems
  - AI/ML based optimization and regionalization is promising
- Ensemble simulations and uncertainty quantification
  - Need much larger regional continuous and event based ensembles
  - Formal uncertainty analysis across all dimensions of flood simulation

Altus, OK



Newman et al. (2021)

Opportunities to Connect to PMP Modernization

- Model Evaluation Project (MEP)
  - Comprehensive benchmarking of both atmospheric and hydrologic models across a range of metrics (5-2, 5-12)
    - **Co-designed *a priori* performance criteria and benchmark models (5-2)**
      - What are meaningful metrics?
      - What are meaningful improvements (e.g. practical significance)?
    - Climate sensitivities
    - Right answer for the right reasons - process representation (5-8, 5-12)
- Uncertainty quantification
  - Large ensembles (5-7, 5-10)
  - Multi-method evaluation and assessment for near-term change factors (5-9)
- Emulators and other computational infrastructure advances (e.g. GPU based physical models) (5-10)

Key Take-Home Messages

- Method and data explosion is an opportunity for analysis
  - Need to quantify adequate/fit-for-purpose (MEP is critical)
  - Move to multi-century record lengths; increasing model process representation using AI/ML and computational hardware advances
- Improving hydrology, both in offline and coupled models
  - Need optimization and regionalization of model structure and parameters
  - Need large-scale benchmarking studies – Do our models respond realistically?
- Improving uncertainty quantification
  - Need more formal uncertainty analysis across scenario, model structure and parameters, meteorology, and hydrologic states
- Partner input and participation is strong in places
  - Need deeper partner co-design of models, experiments, and data