IM3
INTEGRATED
MULTISECTOR
MULTISCALE
MODELING

Large Ensemble Diagnostic Evaluation of Hydrologic Parameter Uncertainty in the Community Land Model Version 5 (CLM5)

Hongxiang Yan[1], Ning Sun[1], Hisham Eldardiry[1], Travis Thurber[1], Patrick Reed[2], Keyvan Malek[2], Rohini Gupta[2], Daniel Kennedy[3], Sean Swenson[3], Zhangshuan Hou[1], Yanyan Cheng[1], and Jennie Rice[1]

1 Pacific Northwest National Laboratory, Richland, WA, USA
2 Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA
3 National Center for Atmospheric Research, Boulder, CO, USA

This research is supported by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program

Pacific Northwest National Laboratory | Baylor University | Berkeley Lab | Boston University | UC Davis University of California | Cornell University | NC State University | The University of North Carolina at Chapel Hill | Oak Ridge National Laboratory | Sandia National Laboratories

Objectives and Storyline

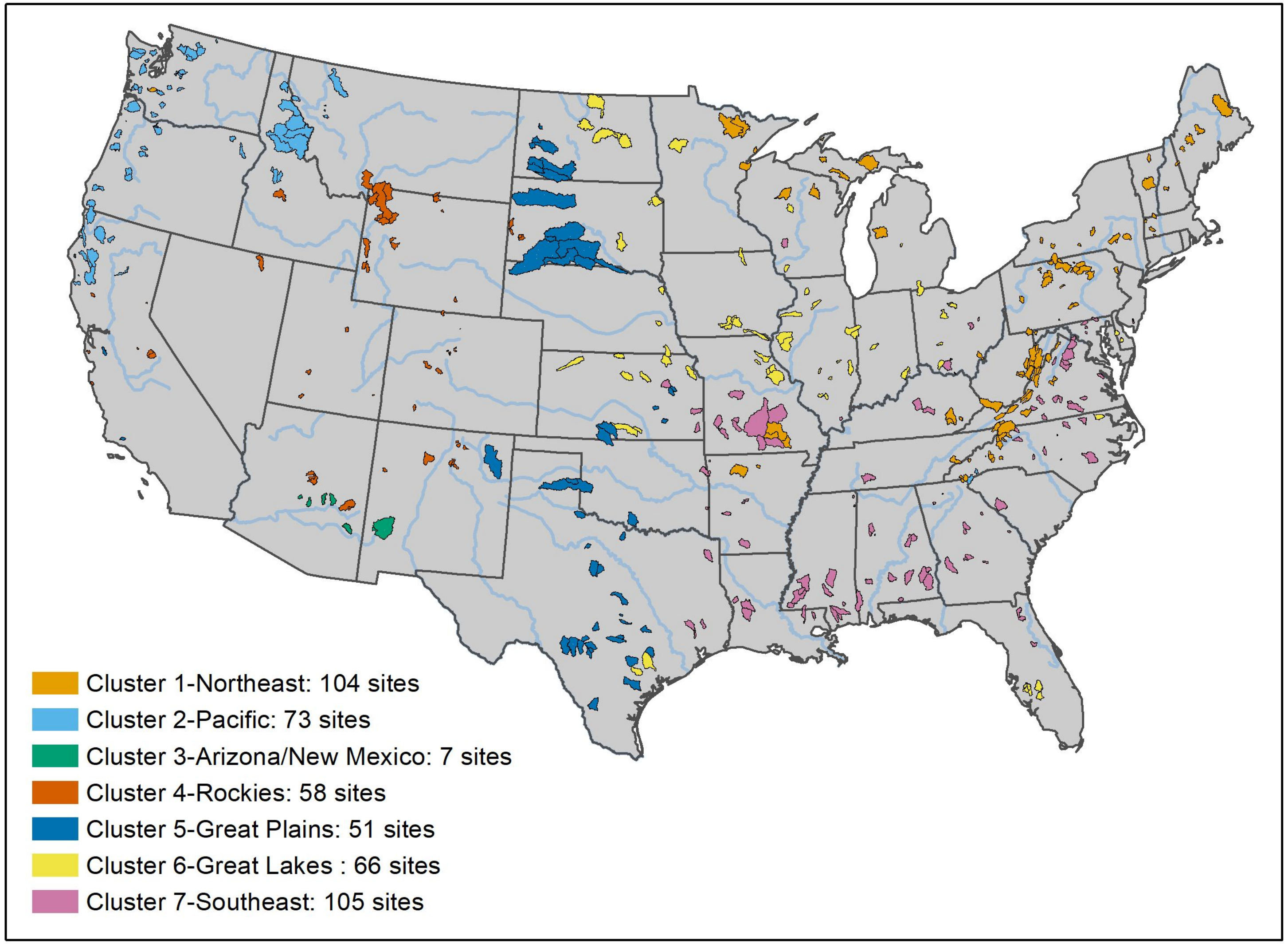
1. Benchmark the performance of CLM5 based on the default hydrologic parameters (**Now**).
2. Characterize the effects of hydrologic parameter uncertainty on CLM5 streamflow predictions (**How**).
3. Perform a global sensitivity analysis to evaluate which CLM5 parameters most influence a suite of diagnostic hydrologic metrics (**Why**).

Are the baseline deterministic parameters and streamflow predictions acceptable?

How is hydrologic parameter uncertainty shaping CLM5's streamflow predictions?

Diagnostically assess which parameterized processes control key features of streamflow prediction errors.

Study Area: Configure CLM5 from CONUS 1/8° gridcell to Basin scale



- Selected 464 CAMELS basins (natural flow basins) over CONUS to represent different hydroclimate regimes.

- The basins are classified into seven clusters for regional analysis using *k*-means++ clustering.

27 Diagnostic Error Metrics

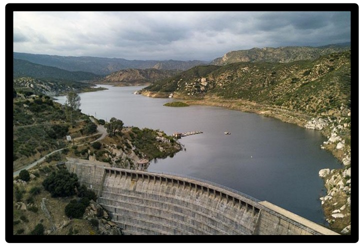
Num.	Metric and Flow Temporal Scale	Unit	Relevance
1	Daily Kling Gupta Efficiency (D-KGE)	-	Multiobjective metric
2	Daily Mean Absolute Error (D-MAE)	m^3/s	Overall daily flow
3	Daily Nash Sutcliffe Efficiency (D-NSE)	-	High daily flow
4	Daily Root Mean Square Error (D-RMSE)	m^3/s	High daily flow
5	Daily Transformed Root Mean Square Error (D-TRMSE)	m^3/s	Low daily flow
6	Daily Flow Duration Curve Slope Bias	-	Daily flow flashiness
7	Daily Variance Bias	-	Daily flow variability
8	Num. of Days Contributing 50% of D-RMSE (D-E50)	day	Days of large error
9	Monthly Kling Gupta Efficiency (M-KGE)	-	Multiobjective metric
10	Monthly Mean Absolute Error (M-MAE)	m^3/s	Overall monthly flow
11	Monthly Nash Sutcliffe Efficiency (M-NSE)	-	High monthly flow
12	Monthly Root Mean Square Error (M-RMSE)	m^3/s	High monthly flow
13	Monthly Transformed Root Mean Square Error (M-TRMSE)	m^3/s	Low monthly flow
14	Monthly Flow Duration Curve Slope Bias	-	Monthly flow flashiness
15	Monthly Variance Bias	-	Monthly flow variability
16	Num. of Months Contributing 50% of M-RMSE (M-E50)	month	Months of large error
17	Annual Volume Bias	-	Total water balance
18	Flow Regime Quantile 0-10% (Q0-10) Volume Bias	-	Low flow water balance
19	Flow Regime Quantile 10-25% (Q10-25) Volume Bias	-	Low flow water balance
20	Flow Regime Quantile 25-50% (Q25-50) Volume Bias	-	Moderate flow water balance
21	Flow Regime Quantile 50-75% (Q50-75) Volume Bias	-	Moderate flow water balance
22	Flow Regime Quantile 75-90% (Q75-90) Volume Bias	-	High flow water balance
23	Flow Regime Quantile 90-100% (Q90-100) Volume Bias	-	High flow water balance
24	Winter (DJF) Volume Bias	-	Seasonal water balance
25	Spring (MAM) Volume Bias	-	Seasonal water balance
26	Summer (JJA) Volume Bias	-	Seasonal water balance
27	Fall (SON) Volume Bias	-	Seasonal water balance

Most of literature focused only on 1 or 2 error metrics.

Here:

- A total of 27 diagnostic error metrics
- Different metrics ask different questions, e.g.

Hydrological Drought



Reservoir Operation

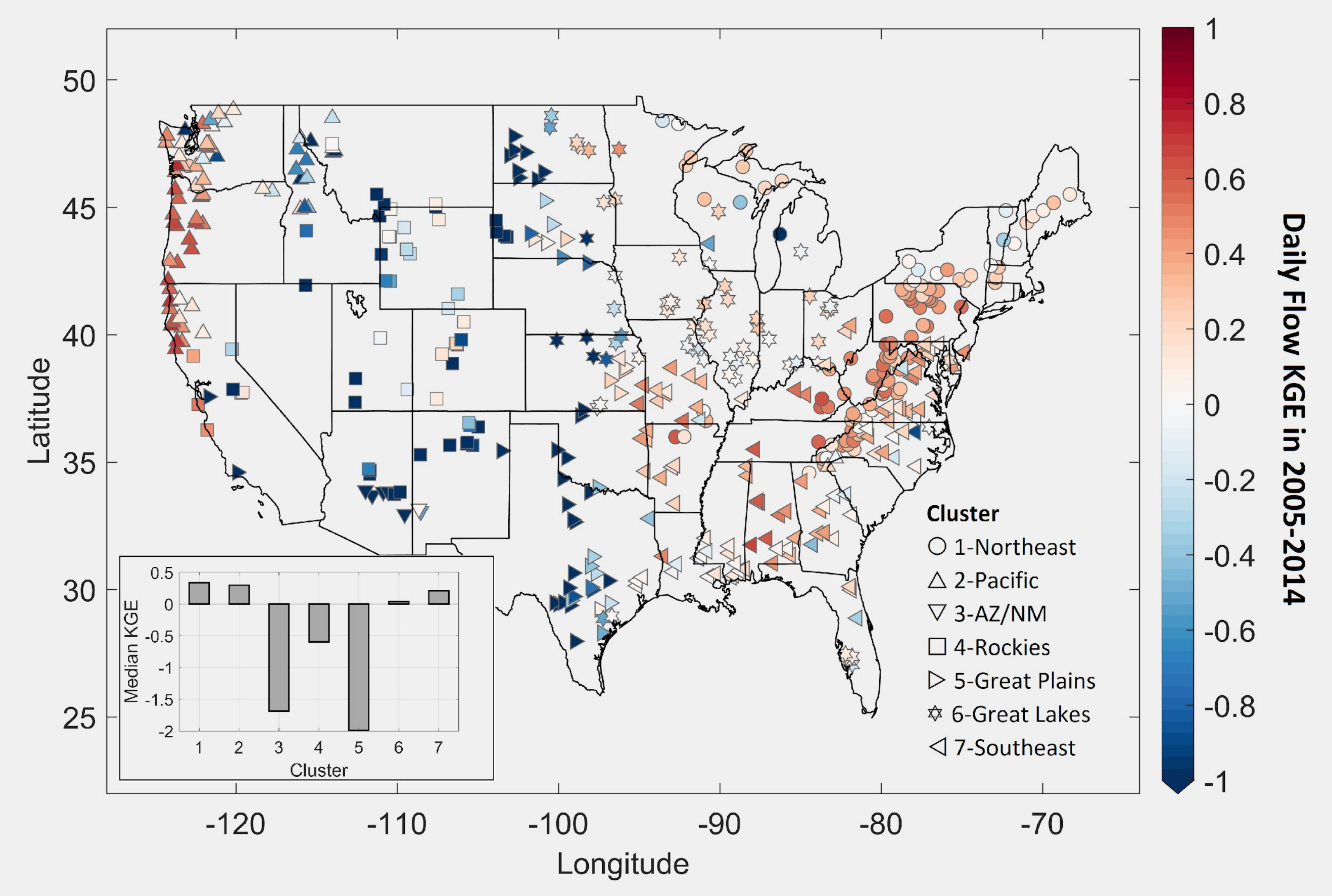
Flood Management

Objectives and Storyline

1. Benchmark the performance of CLM5 based on the default hydrologic parameters (**Now**).
2. Characterize the effects of hydrologic parameter uncertainty on CLM5 streamflow predictions (**How**).
3. Perform a global sensitivity analysis to evaluate which CLM5 parameters most influence a suite of diagnostic hydrologic metrics (**Why**).

Are the baseline deterministic parameters and streamflow predictions acceptable?

Default Parameters Yield Strong Regional Differences in KGE Metric

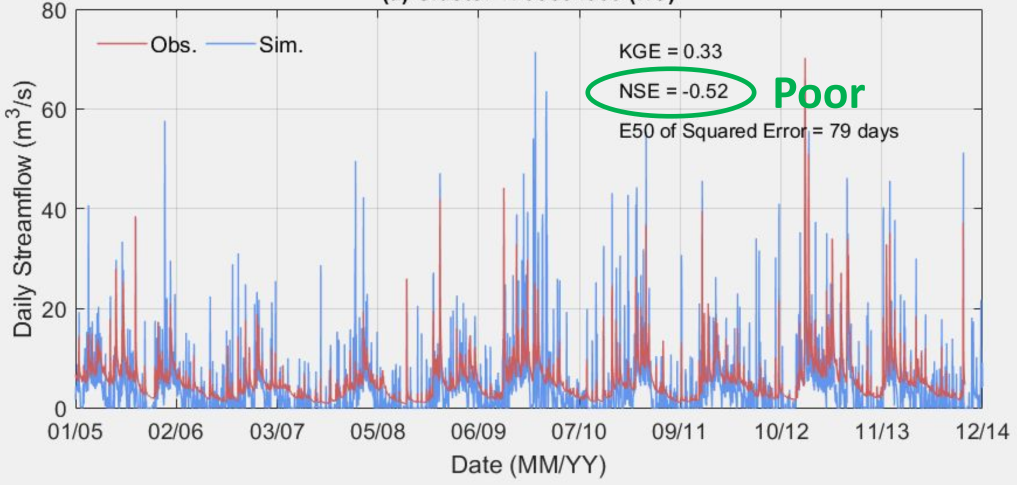


- KGE ranges in (-∞, 1]
- KGE=1: perfect model prediction
- KGE>0.7: acceptable model performance
- KGE<0.5: some model skill
- KGE<0: very poor

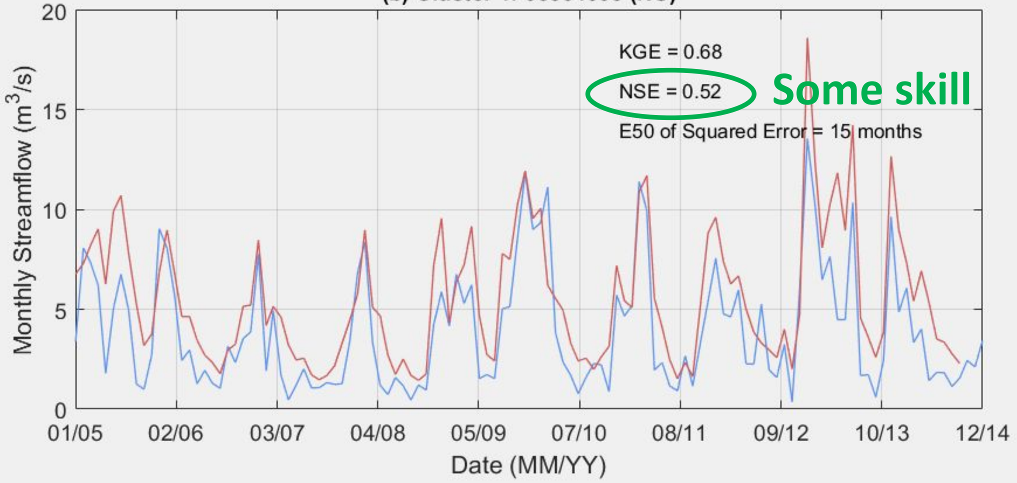
IM3

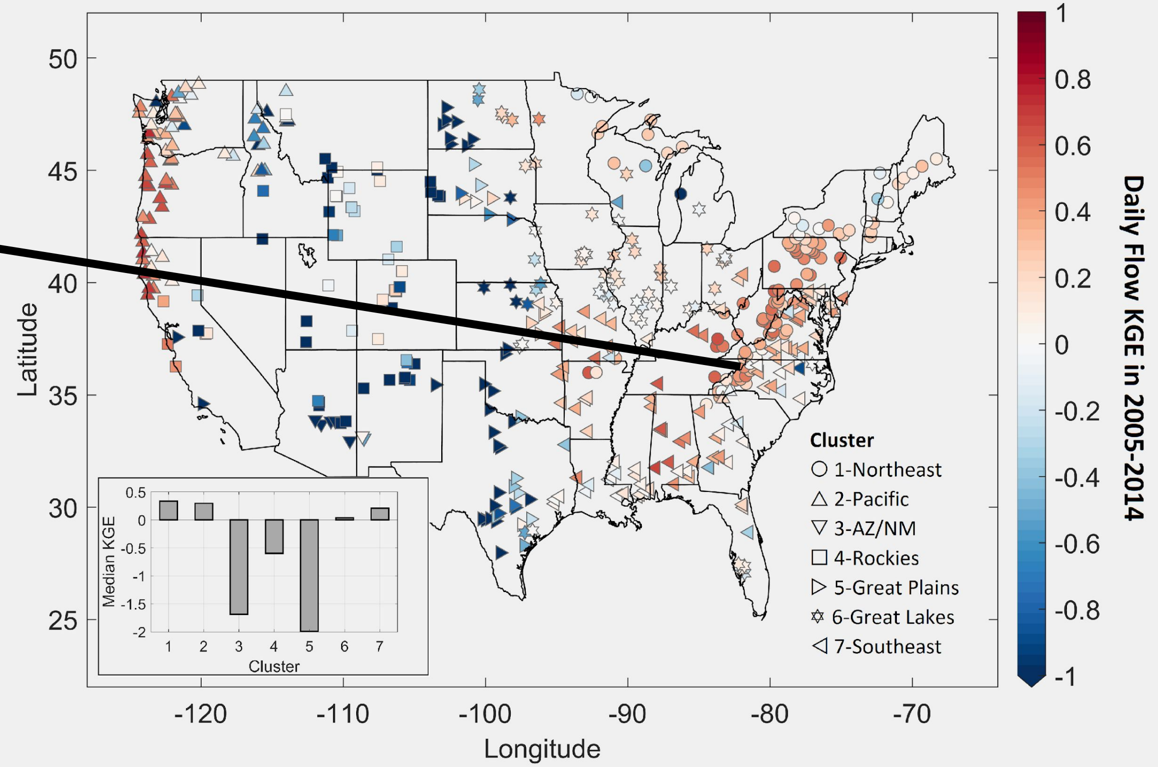
Performances Vary in Different Time Scales (Daily Flow vs. Monthly Flow)

(a) Cluster 1: 03504000 (NC)



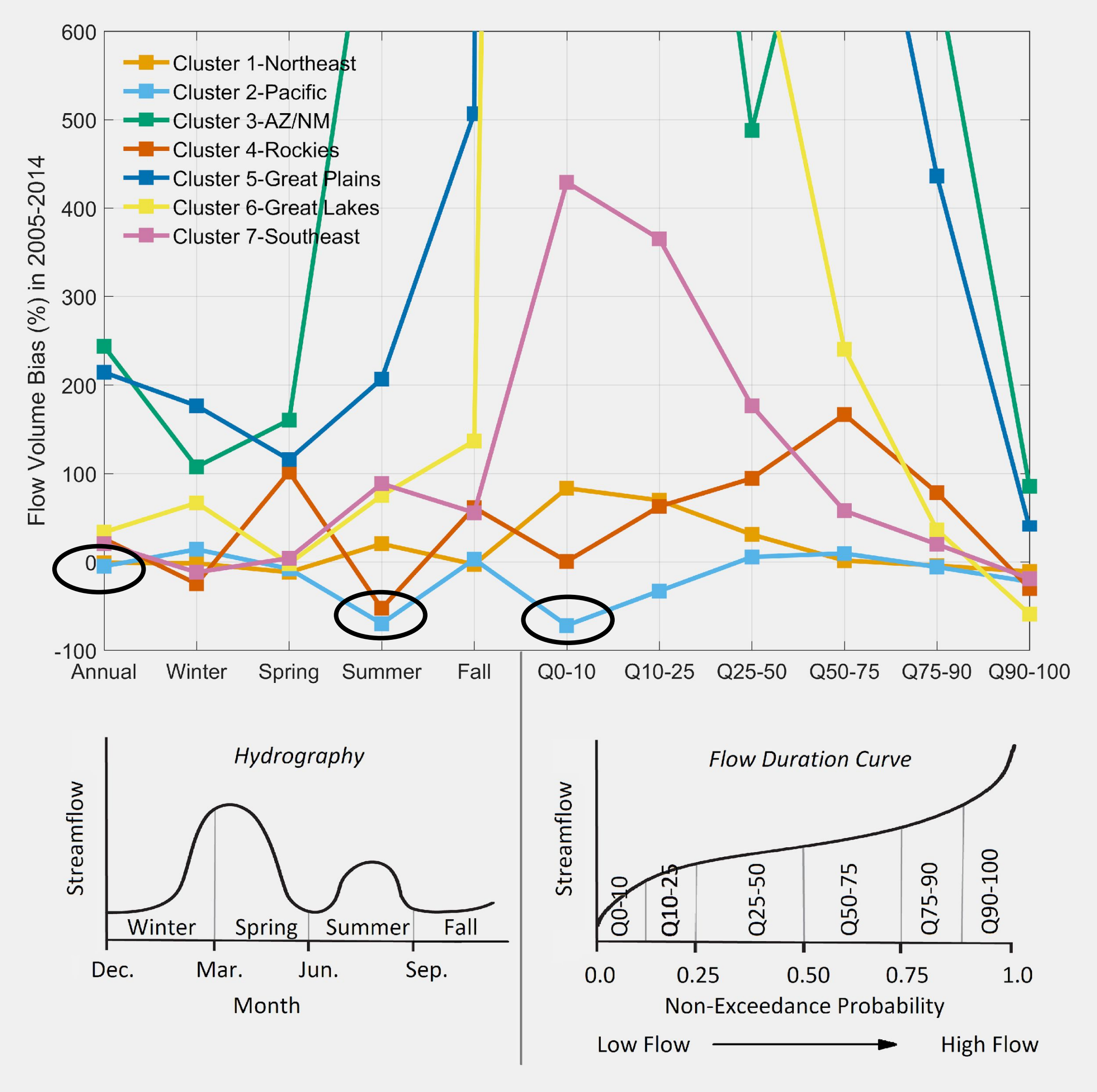
(b) Cluster 1: 03504000 (NC)





Why you need multiple diagnostic error metrics to get the full picture

Default Parameters Yield Strong Regional Differences in Flow Volume Bias



Even within the same cluster, default parameters showed large difference in hydrologic predictions across different flow regimes.

- Good annual water balance ≠ Good low flow
- Good annual water balance ≠ Good high flow

Objectives and Storyline

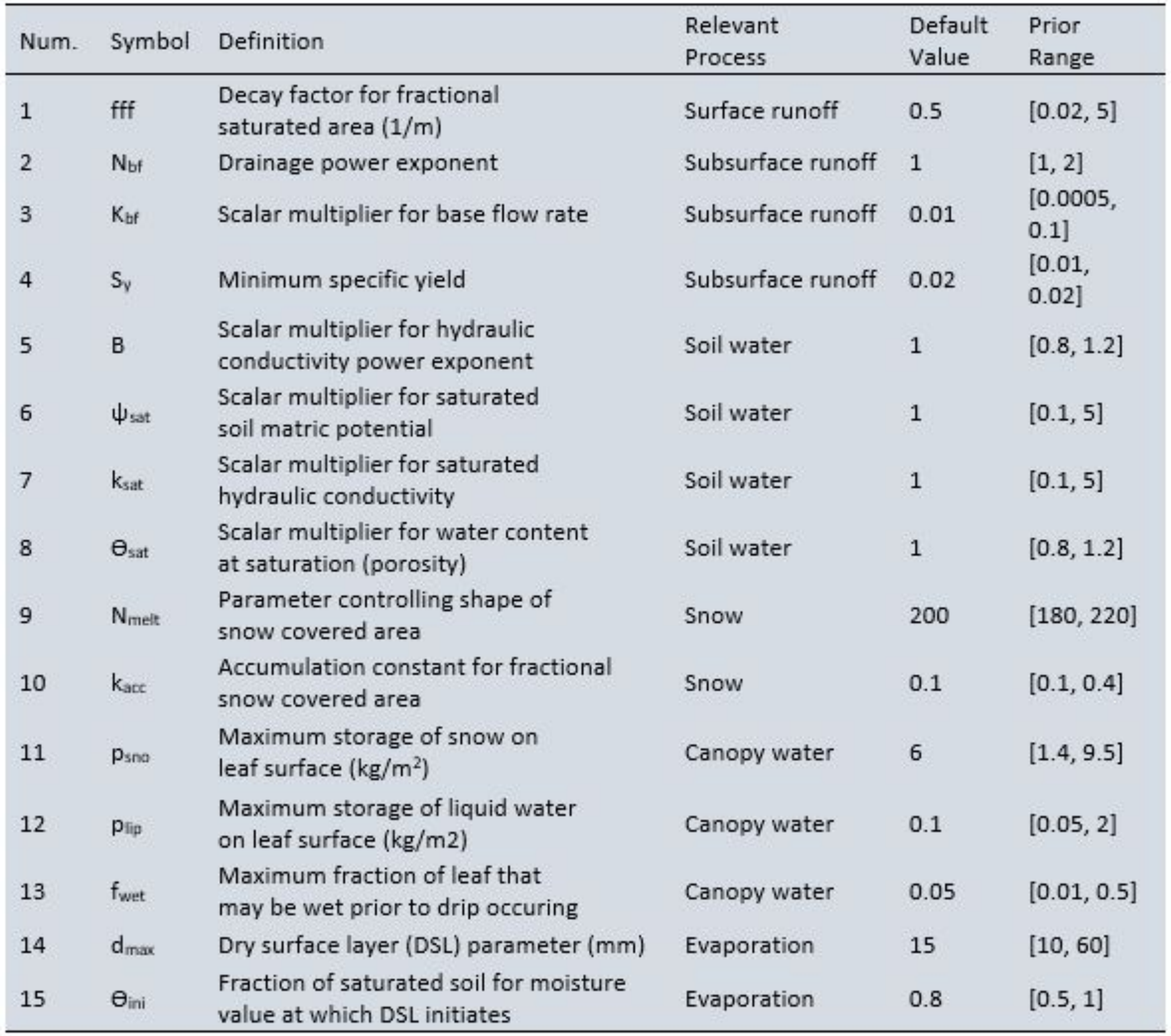
1. Benchmark the performance of CLM5 based on the default hydrologic parameters (Now).

2. Characterize the effects of hydrologic parameter uncertainty on CLM5 streamflow predictions (**How**).

3. Perform a global sensitivity analysis to evaluate which CLM5 parameters most influence a suite of diagnostic hydrologic metrics (**Why**).

How is hydrologic parameter uncertainty shaping CLM5's streamflow predictions?

Parameter Uncertainty Characterization (UC)

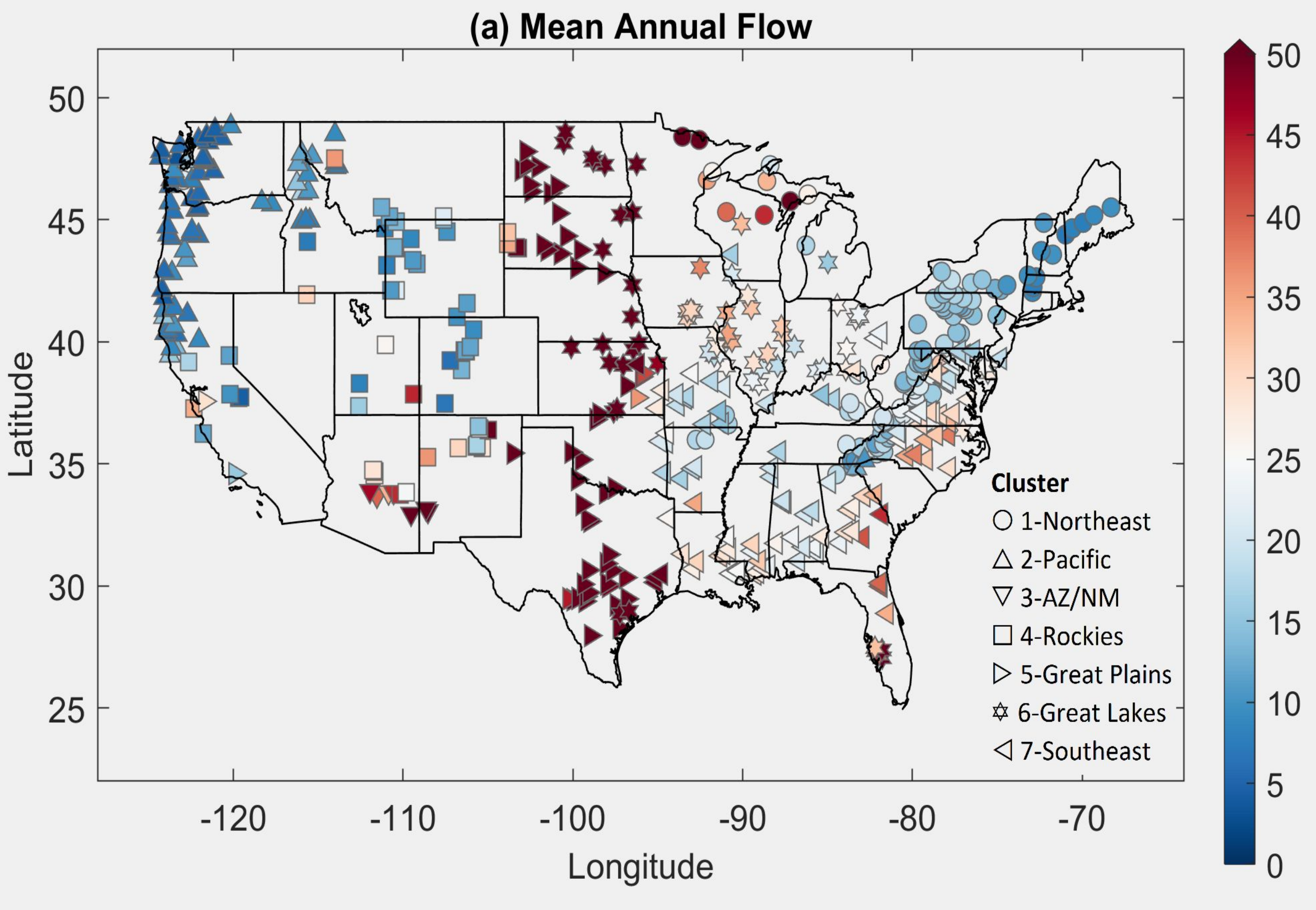


Num.	Symbol	Definition	Relevant Process	Default Value	Prior Range
1	fff	Decay factor for fractional saturated area (1/m)	Surface runoff	0.5	[0.02, 5]
2	N_{bf}	Drainage power exponent	Subsurface runoff	1	[1, 2]
3	K_{bf}	Scalar multiplier for base flow rate	Subsurface runoff	0.01	[0.0005, 0.1]
4	S_y	Minimum specific yield	Subsurface runoff	0.02	[0.01, 0.02]
5	B	Scalar multiplier for hydraulic conductivity power exponent	Soil water	1	[0.8, 1.2]
6	ψ_{sat}	Scalar multiplier for saturated soil matric potential	Soil water	1	[0.1, 5]
7	k_{sat}	Scalar multiplier for saturated hydraulic conductivity	Soil water	1	[0.1, 5]
8	θ_{sat}	Scalar multiplier for water content at saturation (porosity)	Soil water	1	[0.8, 1.2]
9	N_{melt}	Parameter controlling shape of snow covered area	Snow	200	[180, 220]
10	k_{acc}	Accumulation constant for fractional snow covered area	Snow	0.1	[0.1, 0.4]
11	p_{sno}	Maximum storage of snow on leaf surface (kg/m^2)	Canopy water	6	[1.4, 9.5]
12	p_{lip}	Maximum storage of liquid water on leaf surface (kg/m2)	Canopy water	0.1	[0.05, 2]
13	f_{wet}	Maximum fraction of leaf that may be wet prior to drip occuring	Canopy water	0.05	[0.01, 0.5]
14	d_{max}	Dry surface layer (DSL) parameter (mm)	Evaporation	15	[10, 60]
15	θ_{ini}	Fraction of saturated soil for moisture value at which DSL initiates	Evaporation	0.8	[0.5, 1]

- 15 hydrologic parameters that include hydrologic processes of surface, subsurface, soil, snow, canopy, and evaporation.
- Draw 1,500 ensembles using Latin Hypercube Sampling (LHS); each ensemble = [15 parameter values]
- Total runs: 1,500 x 464 = 696,000 for the period 2005–2014

IM3

Parameter Uncertainty Strongly Affects Mean Annual Flow Prediction

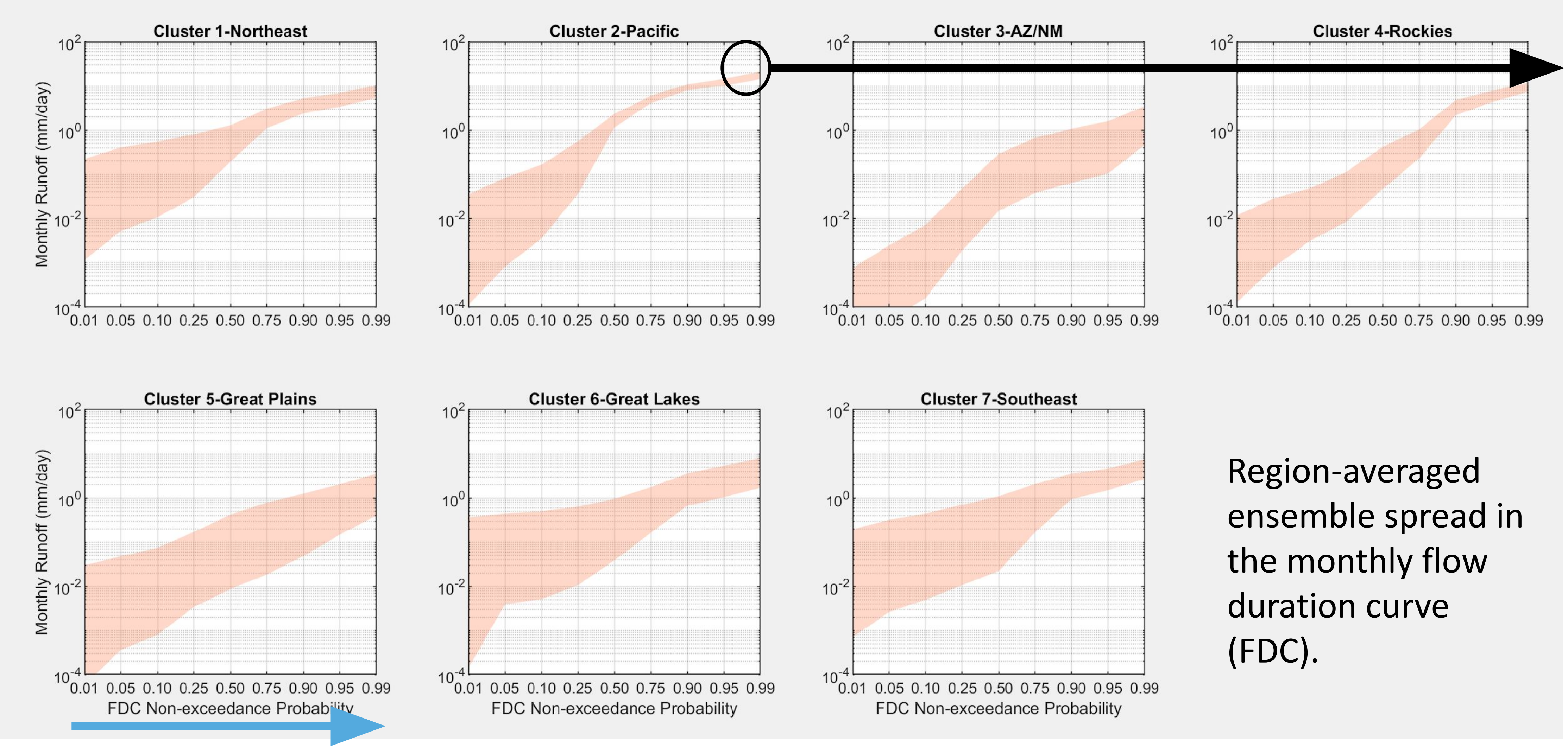


Coefficient of Variation (%) of Ensemble Predictions during 2005-2014

- Coefficient of Variation (CV) = standard deviation / mean
- CV measures the dispersion of ensemble predictions.
- If CV=20% and mean=1, it suggests the standard deviation of ensembles is 0.2.

IM3

Parameter Uncertainty Yield Strongly Regional Differences in Flow Regime

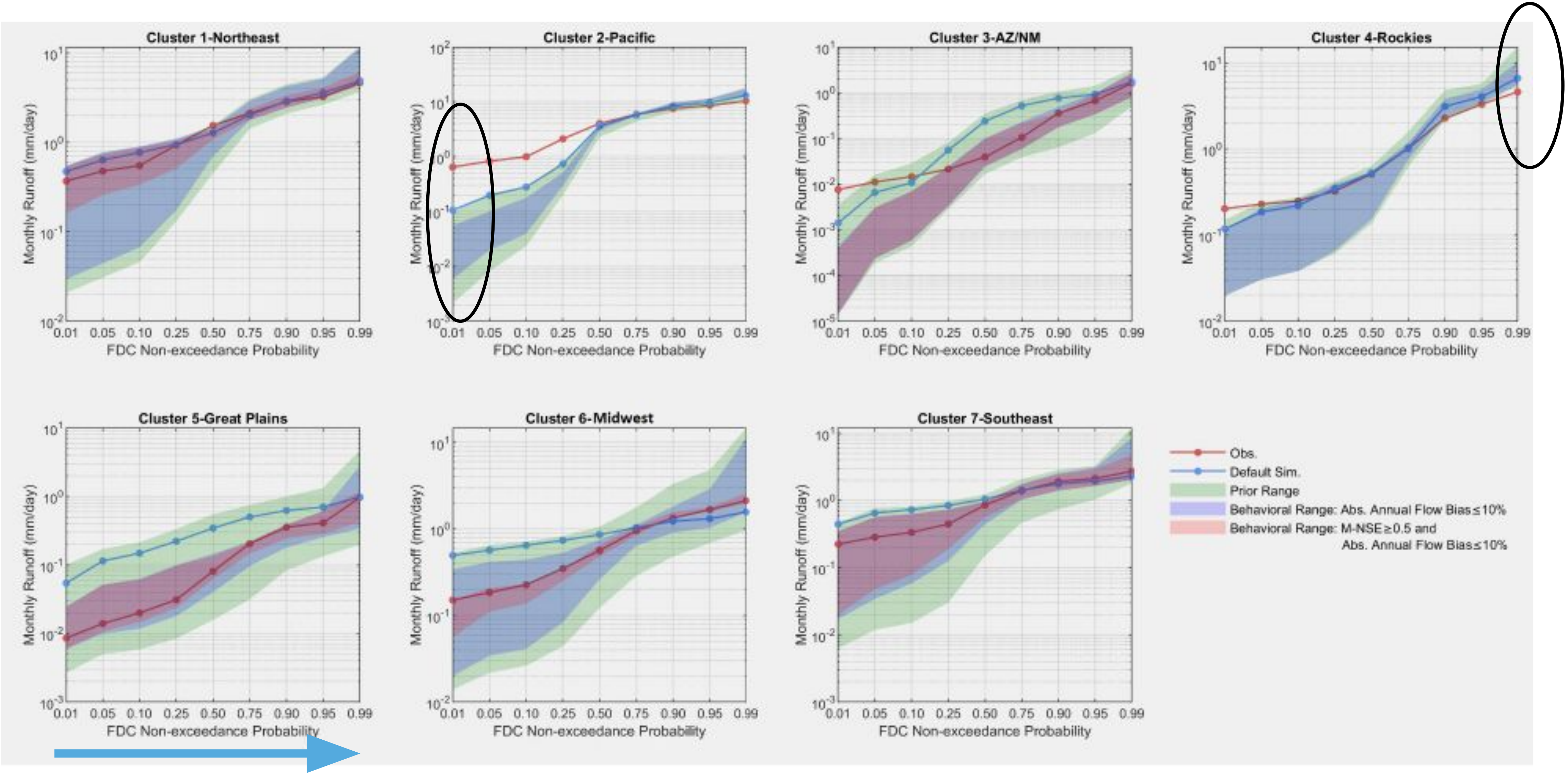


99% quantile [14.5, 21.2]; 46% relative diff.

Region-averaged ensemble spread in the monthly flow duration curve (FDC).

Low Flow High Flow

Behavioral Parameters with >1 Constrains



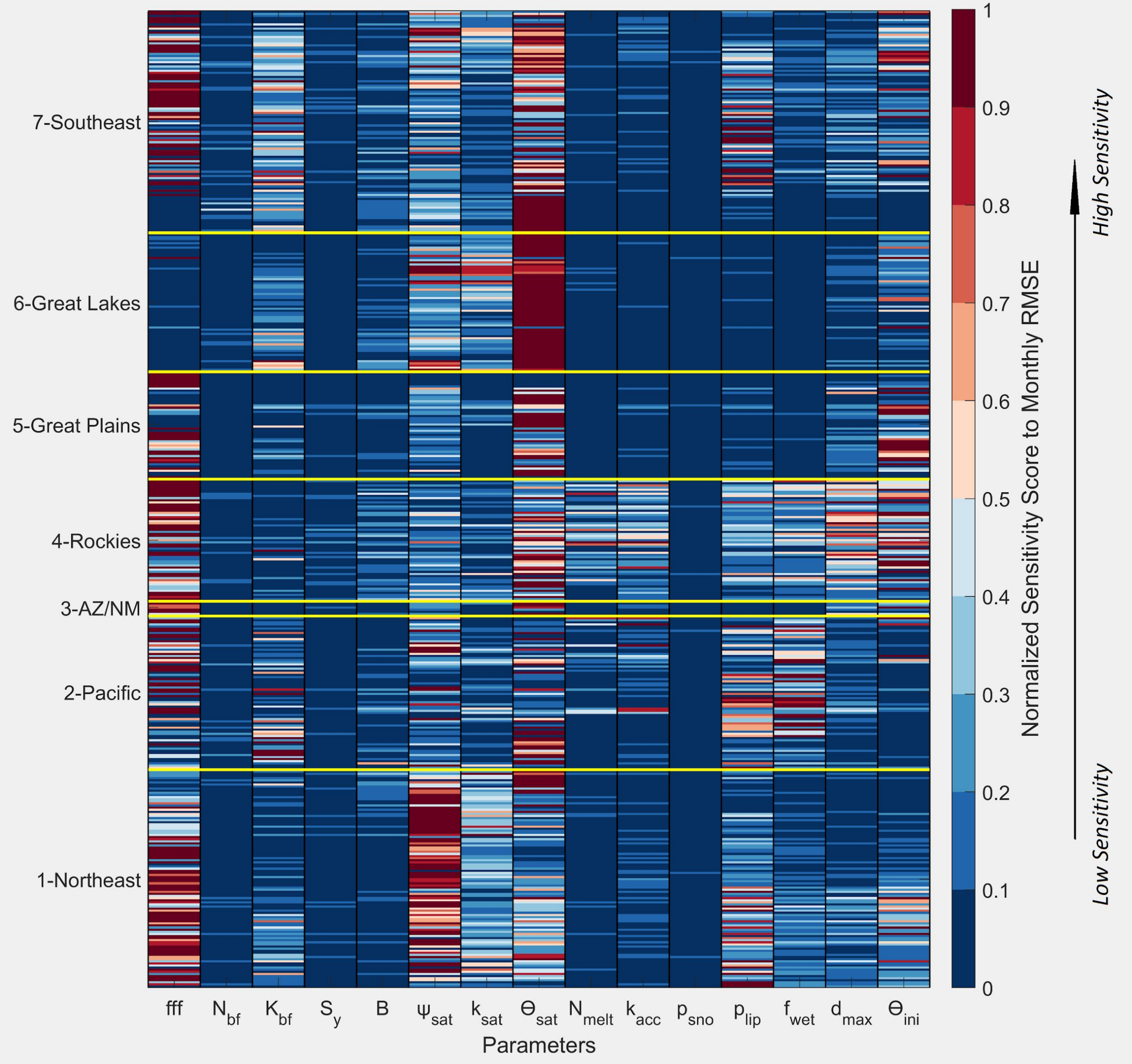
Low Flow High Flow

Objectives and Storyline

1. Benchmark the performance of CLM5 based on the default hydrologic parameters (Now).
2. Characterize the effects of hydrologic parameter uncertainty on CLM5 streamflow predictions (**How**).
3. Perform a global sensitivity analysis to evaluate which CLM5 parameters most influence a suite of diagnostic hydrologic metrics (**Why**).

Diagnostically assess which parameterized processes control key features of streamflow prediction errors.

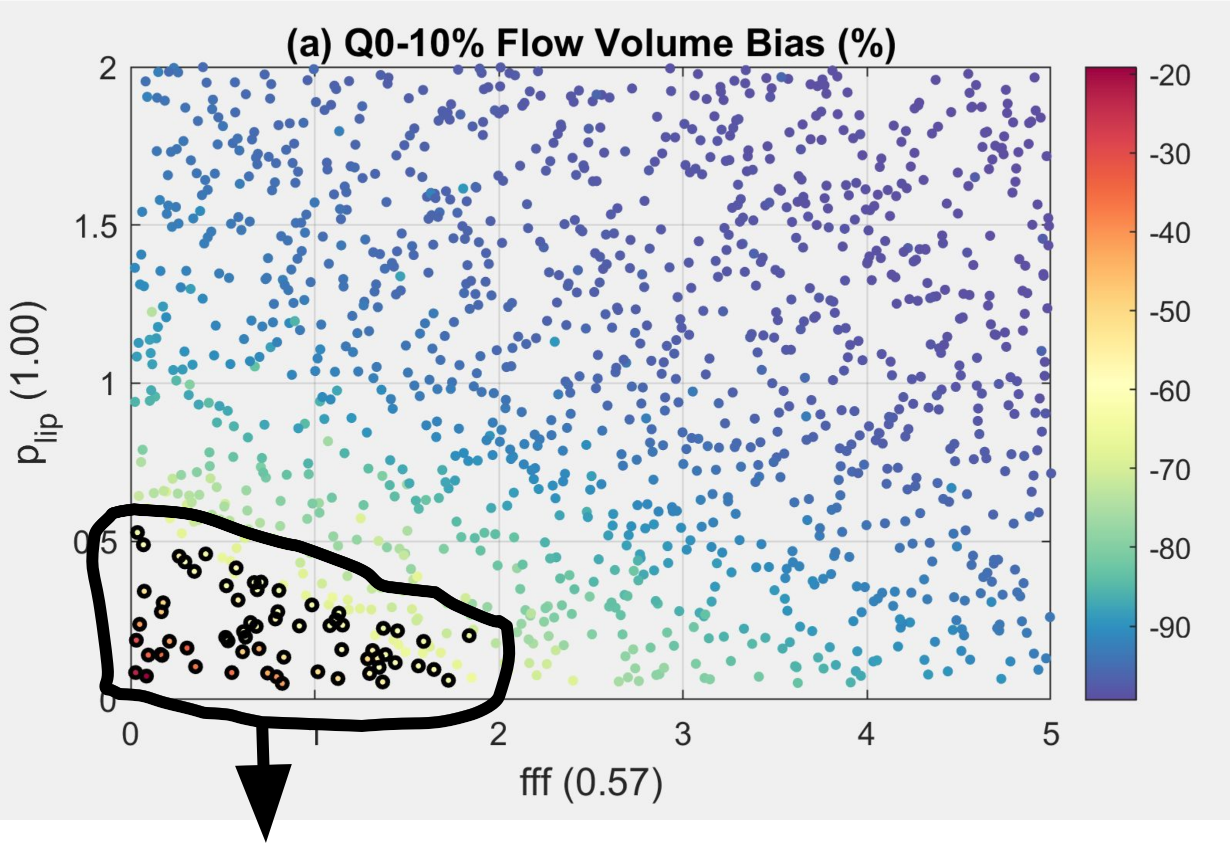
464 Basin Sensitivity Score to Monthly Root-Mean-Square-Error (RMSE) Metric (focusing high flow errors)



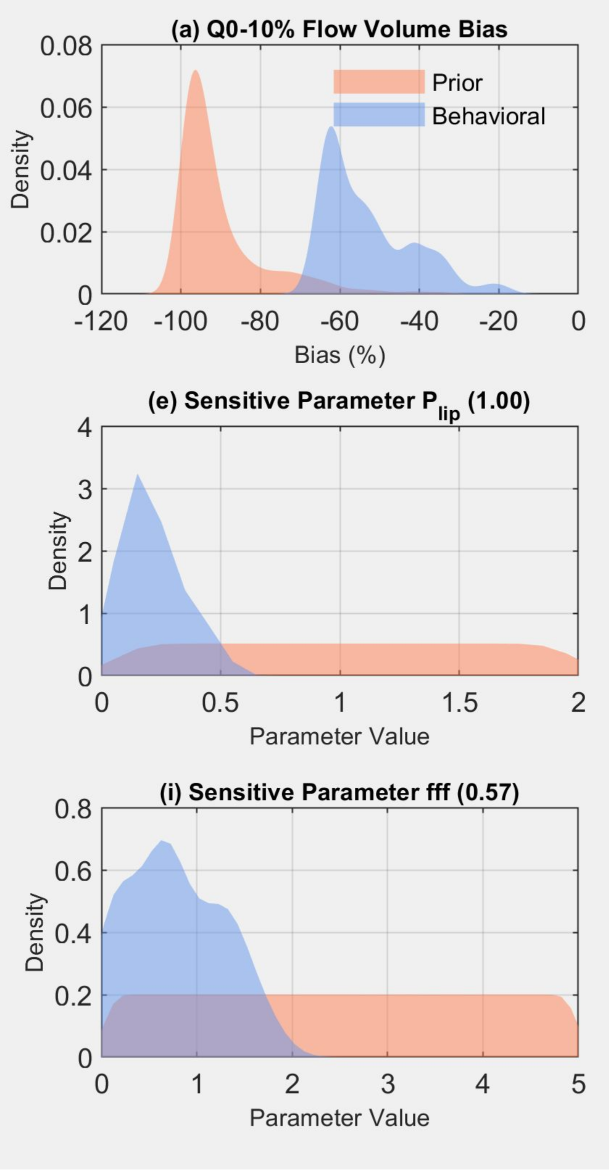
- Parameter sensitivity score showed regional pattern and related to CLM5 model structure (e.g., runoff generation structure, soil depth)
- Surface runoff parameter (*fff*) and soil parameters are the most sensitive parameters in most of the basin.
- These results can be guided in model calibration and provide model diagnosis.

Two Parameters Dominate Low Flow and Top 5% Parameters Partition (Behavioral Parameters)

A Basin in Cluster-1 as an example

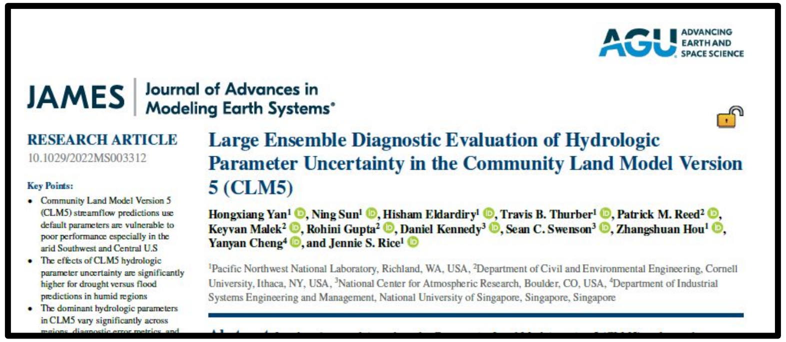


Behavioral Parameters



- Linear separation
- Those behavior parameters will produce enhanced ensemble streamflow predictions for drought studies.

Publications, Datasets, and Relevant Presentation



AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2022MS003312

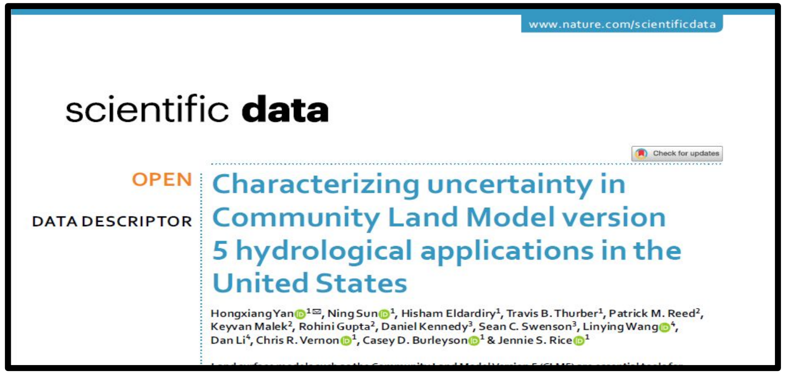
Large Ensemble Diagnostic Evaluation of Hydrologic Parameter Uncertainty in the Community Land Model Version 5 (CLM5)

Hongxiang Yan[1], Ning Sun[1], Hisham Eldardiry[1], Travis B. Thurber[1], Patrick M. Reed[2], Keyvan Malek[2], Rohini Gupta[2], Daniel Kennedy[3], Sean C. Swenson[3], Zhangshuan Hou[1], Yanyan Cheng[4], and Jennie S. Rice[1]

[1]Pacific Northwest National Laboratory, Richland, WA, USA, [2]Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA, [3]National Center for Atmospheric Research, Boulder, CO, USA, [4]Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore, Singapore

Key Points:

- Community Land Model Version 5 (CLM5) streamflow predictions use default parameters are vulnerable to poor performance especially in the arid Southwest and Central U.S
- The effects of CLM5 hydrologic parameter uncertainty are significantly higher for drought versus flood predictions in humid regions
- The dominant hydrologic parameters in CLM5 vary significantly across



www.nature.com/scientificdata

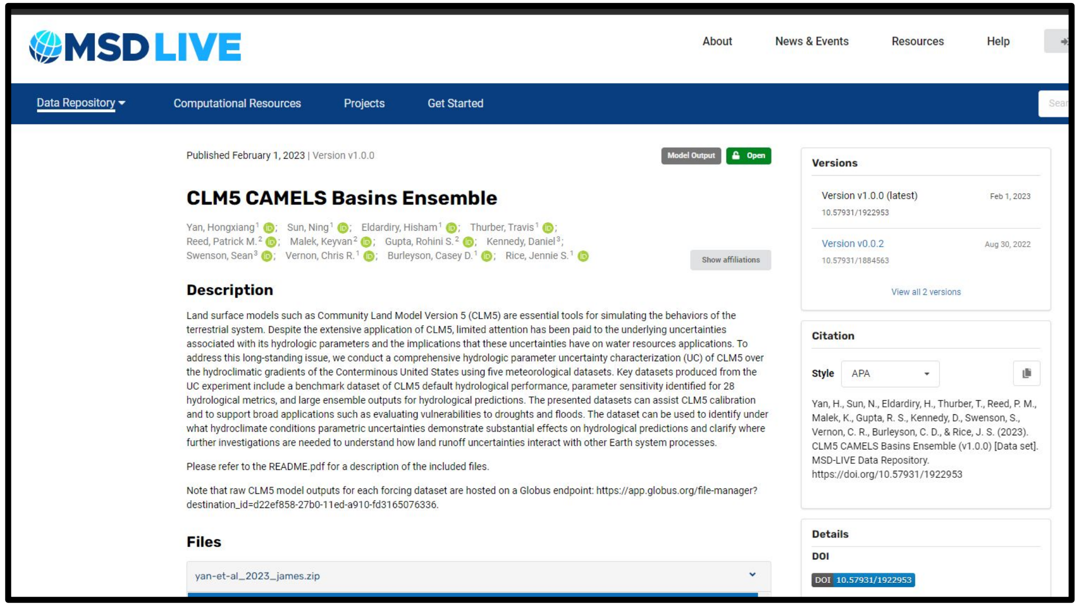
scientific **data**

OPEN

DATA DESCRIPTOR

Characterizing uncertainty in Community Land Model version 5 hydrological applications in the United States

Hongxiang Yan[1]✉, Ning Sun[1], Hisham Eldardiry[1], Travis B. Thurber[1], Patrick M. Reed[2], Keyvan Malek[2], Rohini Gupta[2], Daniel Kennedy[3], Sean C. Swenson[3], Linying Wang[4], Dan Li[4], Chris R. Vernon[1], Casey D. Burleyson[1] & Jennie S. Rice[1]



Published February 1, 2023 | Version v1.0.0

CLM5 CAMELS Basins Ensemble

Yan, Hongxiang[1]; Sun, Ning[1]; Eldardiry, Hisham[1]; Thurber, Travis[1]; Reed, Patrick M.[2]; Malek, Keyvan[2]; Gupta, Rohini S.[2]; Kennedy, Daniel[3]; Swenson, Sean[3]; Vernon, Chris R.[1]; Burleyson, Casey D.[1]; Rice, Jennie S.[1]

Description

Land surface models such as Community Land Model Version 5 (CLM5) are essential tools for simulating the behaviors of the terrestrial system. Despite the extensive application of CLM5, limited attention has been paid to the underlying uncertainties associated with its hydrologic parameters and the implications that these uncertainties have on water resources applications. To address this long-standing issue, we conduct a comprehensive hydrologic parameter uncertainty characterization (UC) of CLM5 over the hydroclimatic gradients of the Conterminous United States using five meteorological datasets. Key datasets produced from the UC experiment include a benchmark dataset of CLM5 default hydrological performance, parameter sensitivity identified for 28 hydrological metrics, and large ensemble outputs for hydrological predictions. The presented datasets can assist CLM5 calibration and to support broad applications such as evaluating vulnerabilities to droughts and floods. The dataset can be used to identify under what hydroclimate conditions parametric uncertainties demonstrate substantial effects on hydrological predictions and clarify where further investigations are needed to understand how land runoff uncertainties interact with other Earth system processes.

Please refer to the README.pdf for a description of the included files.

Note that raw CLM5 model outputs for each forcing dataset are hosted on a Globus endpoint: https://app.globus.org/file-manager?destination_id=d22ef858-27b0-11ed-a910-fd3165076336.

Files

yan-et-al_2023_james.zip

Versions

Version v1.0.0 (latest) Feb 1, 2023
10.57931/1922953

Version v0.0.2 Aug 30, 2022
10.57931/1884563

View all 2 versions

Citation

Style: APA

Yan, H., Sun, N., Eldardiry, H., Thurber, T., Reed, P. M., Malek, K., Gupta, R. S., Kennedy, D., Swenson, S., Vernon, C. R., Burleyson, C. D., & Rice, J. S. (2023). CLM5 CAMELS Basins Ensemble (v1.0.0) [Data set]. MSD-LIVE Data Repository. https://doi.org/10.57931/1922953

Details

DOI

DOI 10.57931/1922953

- More detailed analysis
- Archived datasets
- Five meteorological forcing
- Default parameter simulation
- 1,300+ ensemble simulation
- Parameter sensitivity score to 20+ metrics

Upcoming presentation. June 13 (Tuesday): 11:00 AM - 11:15 AM

Forcing-dependent parametric uncertainty in community land model simulations: evaluation of hydrologic signatures over the conterminous United States by Dr. Hisham Eldardiry