Parameter Estimation and Transferability Study across Major US Watersheds

Huiying Ren[1], Z. Jason Hou[1] , Maoyi Huang[1], Jie Bao[1], Jaideep Ray[3], Yu Sun[2], Teklu Tesfa[1], L. Ruby Leung[1], Laura P. Swiler[3]

[1]Pacific Northwest National Laboratory, Richland, WA, USA

[2]Department of Hydraulic Engineering, Tsinghua University, Beijing, China

[3]Sandia National Laboratory, Livermore, CA, USA

Sandia National Laboratories

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

Motivation

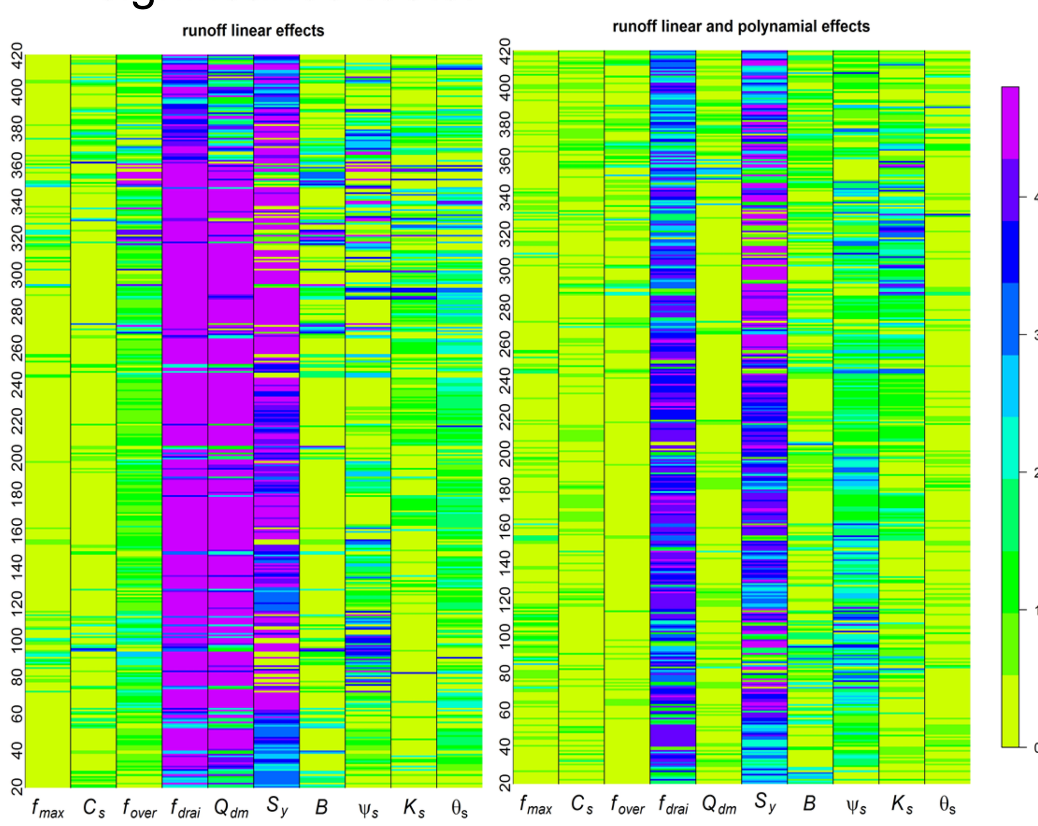
- Model calibration could improve estimates of model parameters and yields more accurate simulations
- Challenges: Curse of high-dimensionality; site-to-site variability in parameter identifiability; computational demand

Methodology

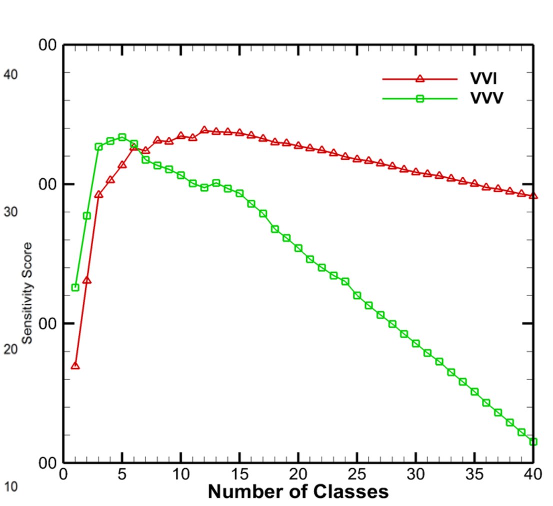
- Classification scheme to reduce the system complexity
- Principal Component Analysis and Expectation-Maximization based clustering approaches are use for classifying 431 US watersheds
- Classes are generated based on sensitivity patterns of simulated streamflow to hydrological parameters, as well as climate attributes
- Surrogate models are developed at representative sites, and validated as computationally efficient alternative to the numerical simulator (i.e., Community Land Model)
- Markov chain Monte Carlo(MCMC)-Bayesian approaches were conducted to evaluate the transferability of parameter values within and between the watershed classes

Study sites and parameterization

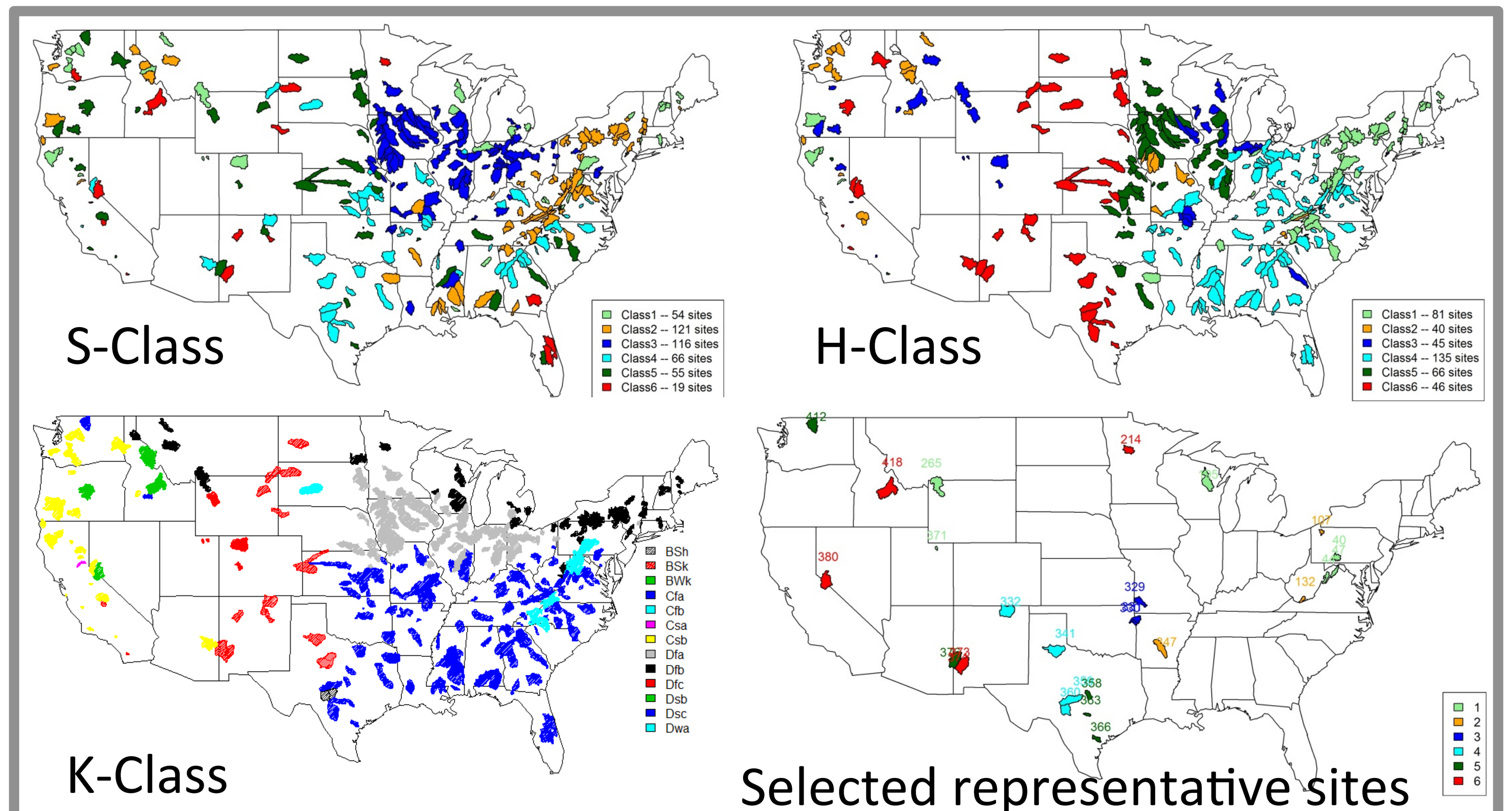
- 431 watershed across the US
- Ten hydrological parameters were selected given their significant impacts on surface and subsurface runoff, latent and sensible fluxes, and soil moisture.
- Parameter screening with an uncertainty quantification framework that integrates quasi-Monte Carlo sampling, minimum-relative-entropy theory for defining priors, and statistical parameter significance tests.



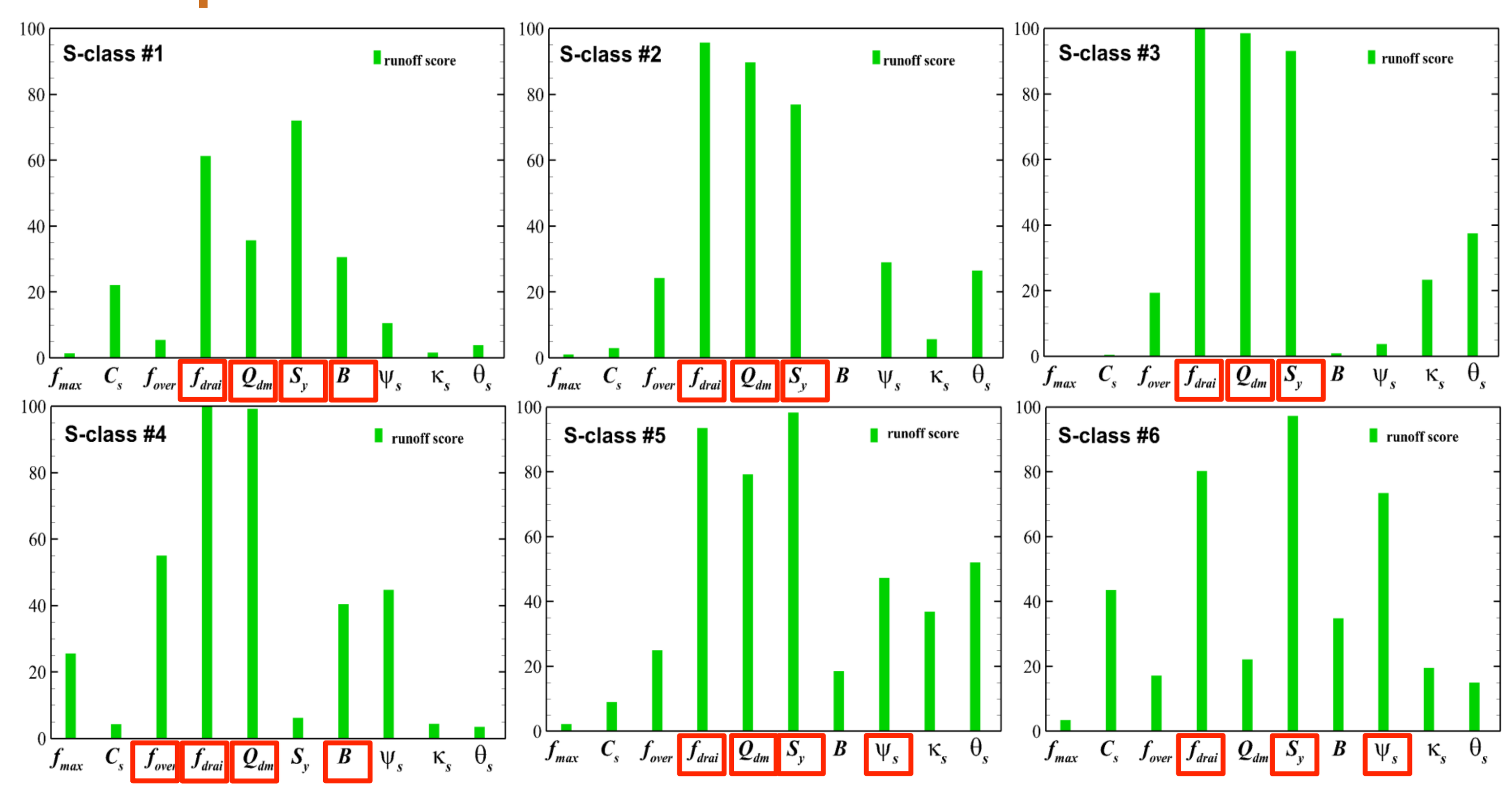
Summary sensitivity scores of the input parameters at the 431 MOPEX basins, with logarithmic runoff as the response variable. The scores are for parameters of a) linear main effects; b) linear and quadratic effects. A higher score means stronger sensitivity of model responses to the input parameter.



BIC for the clustering models against the number of classes for two model options "VVI" and "VVV". The sensitivity scores for clustering are based on GLM analysis with log runoff as the response variable.



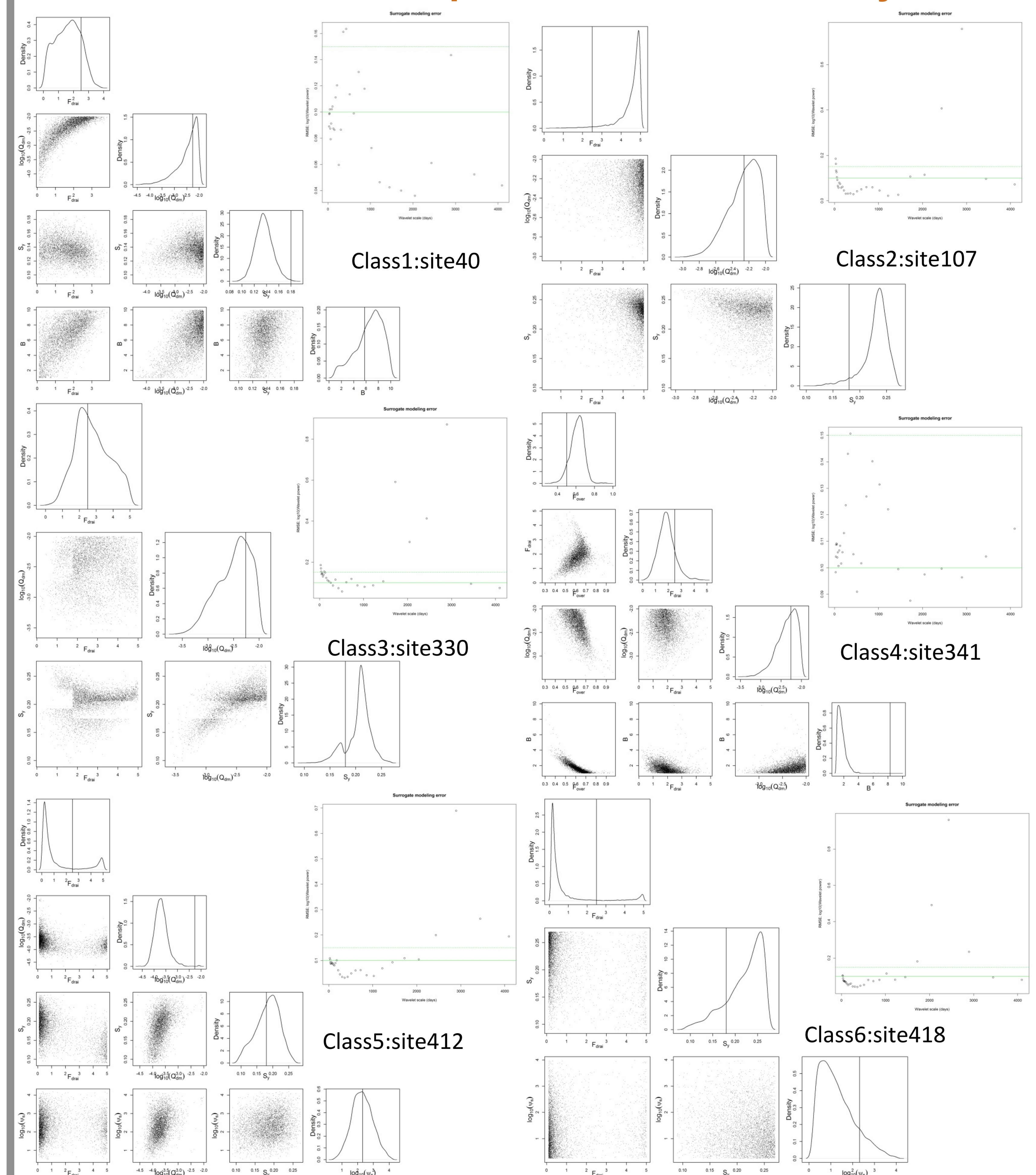
Dominant parameters for each S-Class



Inversion strategies

- MCMC-Bayesian approach for generating posterior samples
- Surrogate used as the forward model, and cross-validated (e.g., both training and testing errors < 15%)
- Modifications when surrogate development is difficult for certain sites:
  - Adopted a composite model, made by adding a kriging component to the fitting errors, that is, to construct quadratic + kriging surrogates, then set up the likelihood and the prior and uses adaptive MCMC as before
  - Made a surrogate (e.g., quadratic) model for a subset of the parameter space close to the 'true' parameter set. A classifier is added (e.g., using treed linear models) to define such a 'good' subspace when sampling the posterior
  - Modeled uncorrelated vs time-correlated errors

Inversion results and parameter transferability



Posterior probabilistic distributions and scatters

Conclusions

- The differences in hydrological parameter values between classes are related to the facts that both model parameters and forcing may drive the hydrologic behaviors at the watersheds to different regimes.
- By reducing the parameter dimensionality to a reasonably low number, the classification makes the inverse modeling possible and less ill-posed.
- Parameter transferability within the same class may not yield the best model performance if the soil and climate conditions vary substantially within the class.
- Further work will be done to evaluate the transferability of parameter posterior distributions within each class.

U.S. DEPARTMENT OF ENERGY

Office of Advanced Scientific Computing Research (OASCR), Applied Mathematics Program

Office of Biological and Environmental Research, Earth System Modeling Program

Contact: Zhangshuan.Hou@pnnl.gov

www.pnnl.gov

SAND2015-10842C