Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Entropy-Bayesian Inversion of Hydrological Parameters in the Community Land Model Using Heat Flux and Runoff Data

Z HOU, M HUANG, PNNL; J RAY, L SWILER, SNL

Pacific Northwest National Laboratory; Sandia National Laboratory

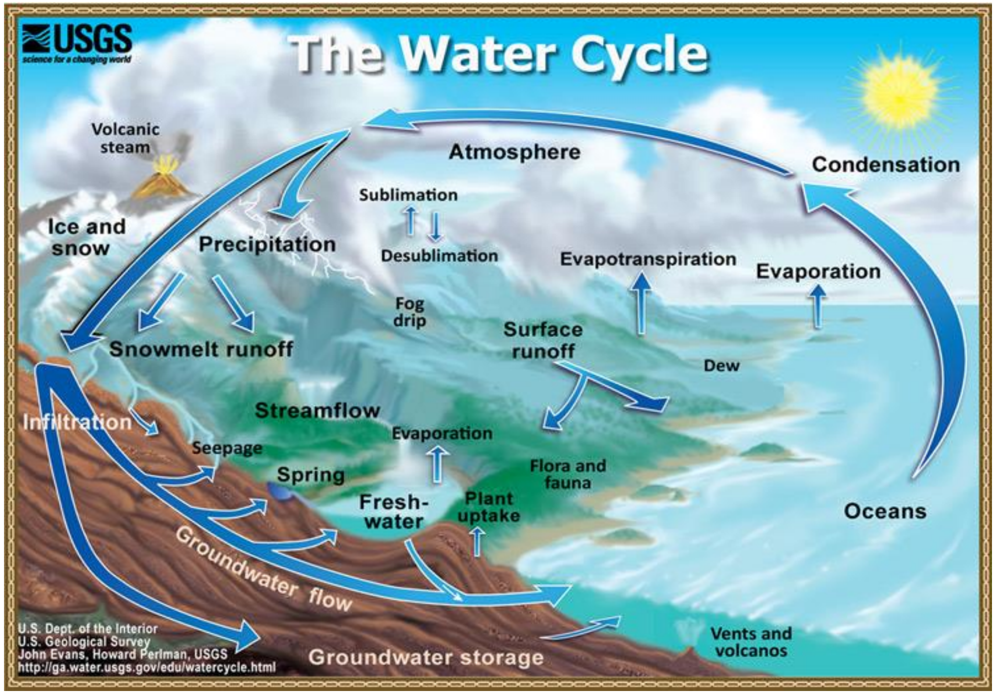
Mar 31 2014

Outline

- Motivation
  - To reduce uncertainty in the Earth system models and improve prediction capability
- Background
  - Sources of uncertainty
  - Uncertainty identification and representation
  - Parameter ranking/screening
  - Surrogate / ROM development
  - System classification
  - Stochastic calibration/inversion
    - MCMC-Bayesian using numerical forward models and/or surrogates
    - Multi-chain MCMC
- Entropy (minimum-relative-entropy, MRE) approach for uncertainty quantification (UQ) and Bayesian inversion
  - UQ based on MRE-derived prior probability density functions
  - Bayesian inversion using MRE-priors

Motivations

- Complex physical phenomena are of multi-phase, multi-component, and involve multiple biogeophysical/biogeochemical processes
- Integrated models introduce numerous model and coupling parameters and therefore formidable high-dimensional parameter spaces
- Many model parameter values were assigned without calibration which results in significant modeling errors (e.g., mismatches between observed and simulated flow and energy fluxes in land surface models)



Sources of Uncertainty

- Model uncertainty
  - Simplifications, structural model formulations/structures, extrapolations, resolution, model initial/boundary conditions
- Data uncertainty
  - Instrumental errors, consistency, gaps, resolution, scaling
- Natural uncertainty/variability/heterogeneity
  - Intrinsic quantities vary over time, over space, or across individuals in a population
  - Physical processes/mechanisms/features vary over space, time, and individuals
- Parameter uncertainty
  - Non-measurable, measurement errors, non-uniqueness, inaccurate calibration, mis-classification due to under-sampling…

Uncertainty Reduction

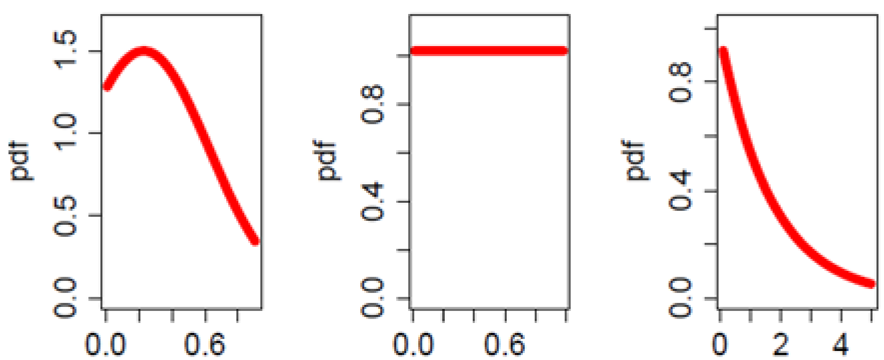
- Sensitivity analyses
- Direct survey data – hard information
- Indirect observations for inference – soft information
- Classification of systems
- Filtering, gap filling, decomposition, co-kriging
- Surrogates (ROMs) for system predictions and risk analyses
- Parameter significances → guidance on conceptual model development and improvement
- Parameter calibration/inversion → improve model accuracy and confidence of predictions

Uncertainty Representation – MRE priors

- In practice, measurements are not available, given prior knowledge including databases and experiences, close-form pdfs can be derived using minimum-relative-entropy (MRE) concept (Hou and Rubin 2005):

$$f(x) = \frac{\sqrt{\frac{\gamma}{\pi}} \exp\left[-\gamma \left(x + \frac{\beta}{2\gamma} \right)^2 \right]}{\Phi\left[\sqrt{2\gamma} \left(U + \frac{\beta}{2\gamma} \right) \right] - \Phi\left[\sqrt{2\gamma} \left(L + \frac{\beta}{2\gamma} \right) \right]} \qquad L \le x \le U$$

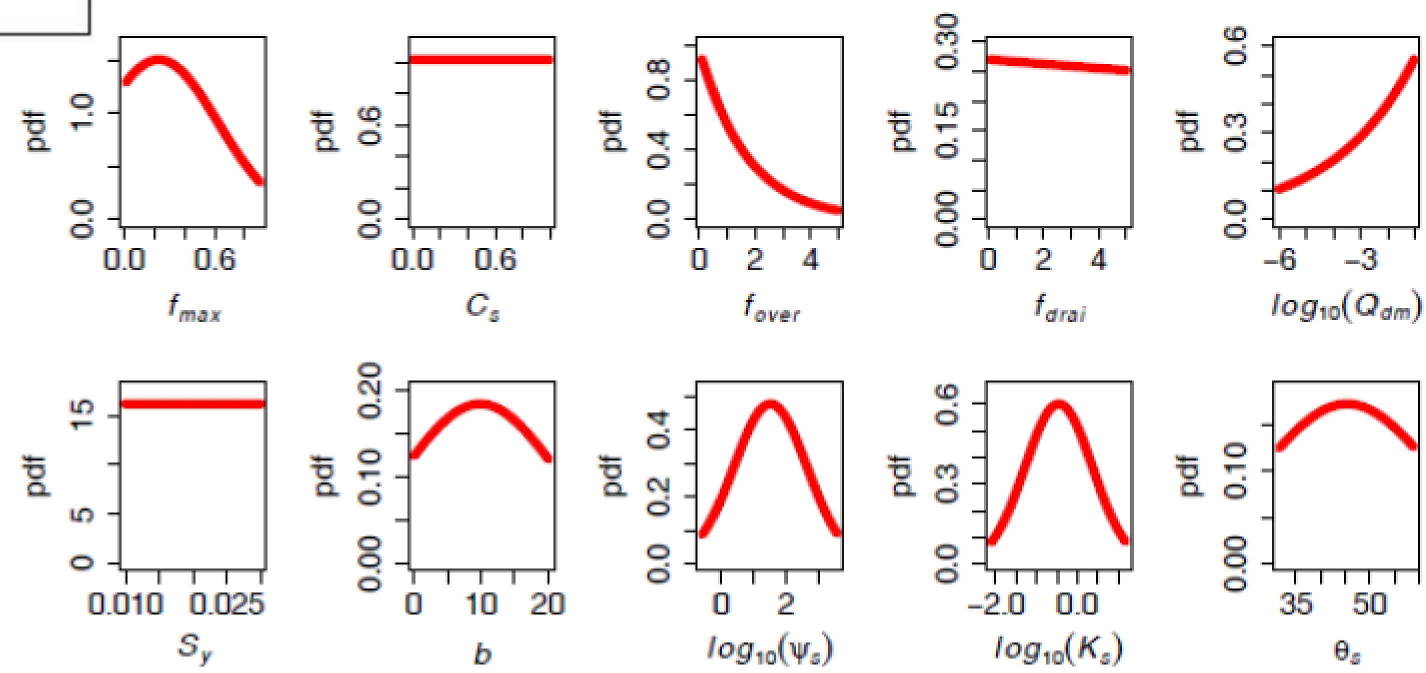
$$f(x) = \frac{\beta e^{-\beta x}}{e^{-\beta L} - e^{-\beta U}} \qquad L \le x \le U$$



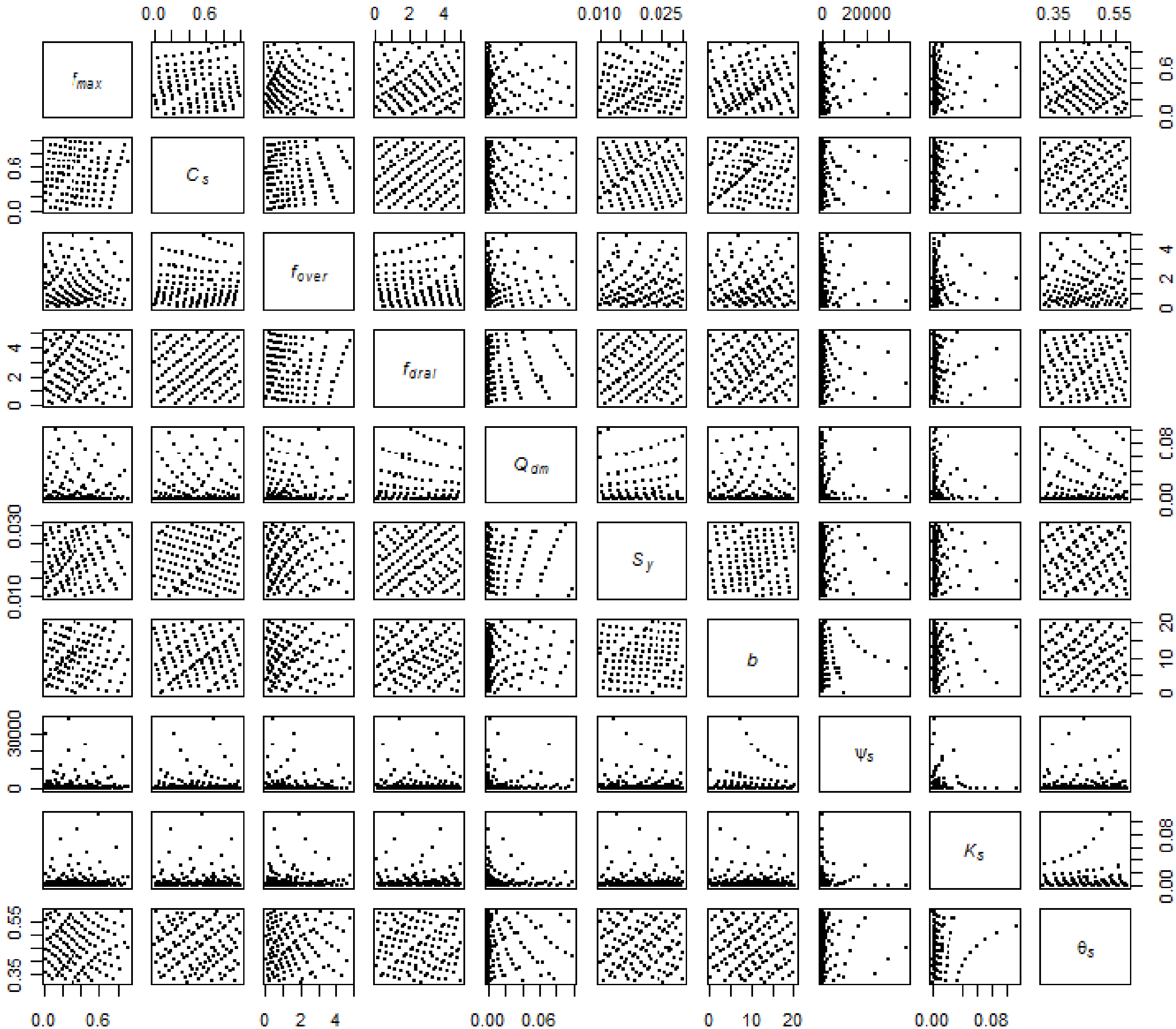
MRE priors for major hydrological parameters in Land Modeling

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

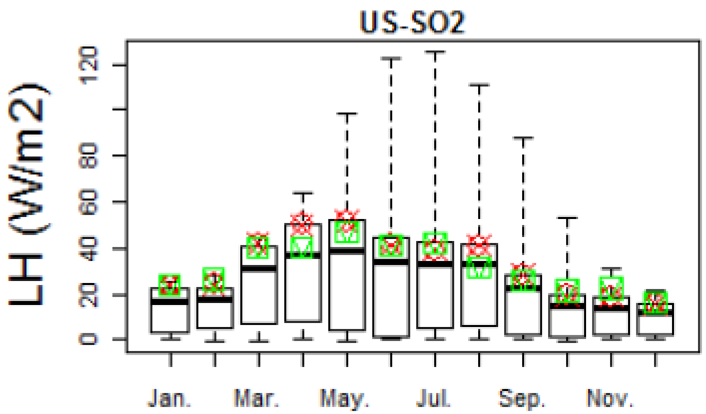
Symbol	Definition
f_{max}	Max fractional saturated area, from DEM
C_s	Shape parameter of the topographic index distribution
f_{over}	Decay factor (m^{-1}) for fsat
b	Clapp and Hornberger exponent
K_s	Hydraulic conductivity (mm s^{-1})
θ_s	porosity
Ψ_s	Saturated soil matrix potential (mm)
f_{drai}	Decay factor (m^{-1}) drainage
$q_{drai,max}$	Max drainage (kg m^{-2} s^{-1})
S_y	Average specific yield

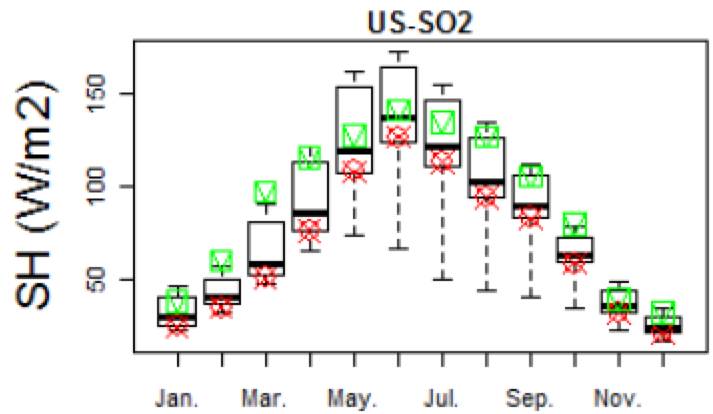


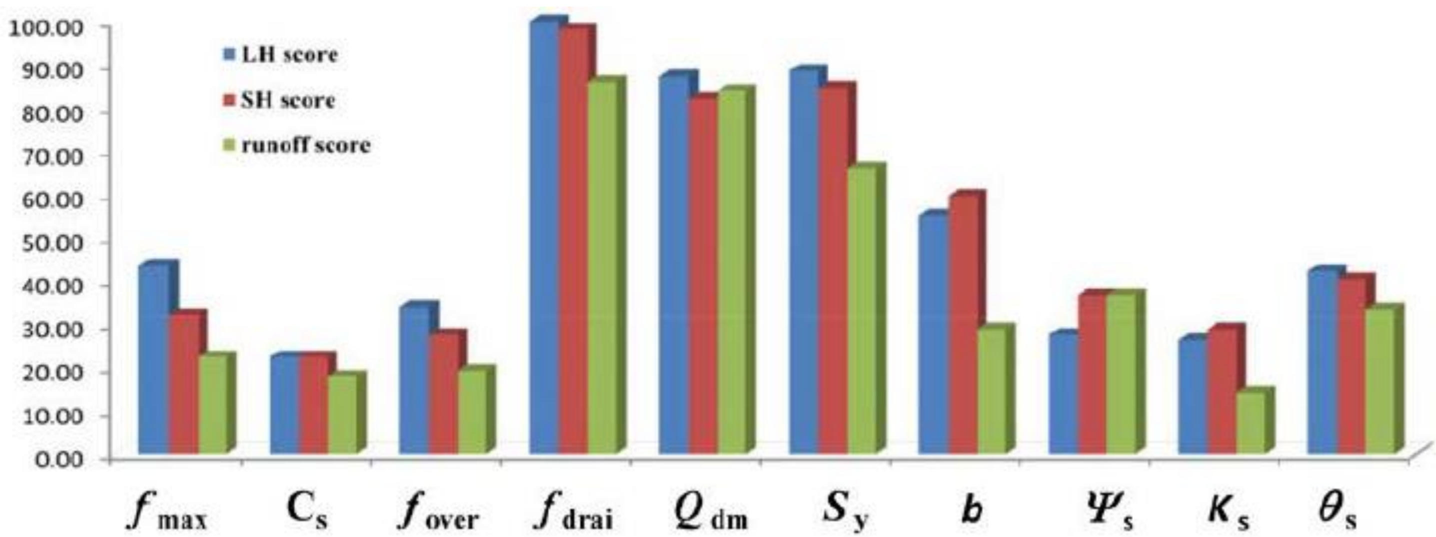
MRE-priors for UQ and SA: sampling



MRE priors for UQ and SA: response surface and parameter ranking







Bayesian Inversion

- Bayesian updating: $f(m/d,I) \propto f(d/m,I)^{*}\, f(m/I)$
- Prior pdf

$$f_{\boldsymbol{M}|\,\boldsymbol{I}(\boldsymbol{m}|\boldsymbol{I})} = \prod_{j=1}^{p} \frac{\sqrt{\frac{\gamma_j}{\pi}}\exp\left[-\gamma_j(m_j+\frac{\beta_j}{2\gamma_j})^2\right]}{\Phi\left[\sqrt{2\gamma_j}\left(U_{j+}\frac{\beta_j}{2\gamma_j}\right)\right]-\Phi\left[\sqrt{2\gamma_j}\left(L_{j+}\frac{\beta_j}{2\gamma_j}\right)\right]}$$

- Likelihood function

$$f_{D|\boldsymbol{M},\boldsymbol{\Sigma},\boldsymbol{I}(\boldsymbol{d}^*|\boldsymbol{m},\boldsymbol{\Sigma},\boldsymbol{I})} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}\boldsymbol{\pi}|\boldsymbol{\Sigma}|}}\left[\exp\{-\frac{1}{2}\left(d^*-G(m)\right)\Sigma^{-1}\left(d^*-G(m)\right)^T\}\right]$$

Bayesian Inversion

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

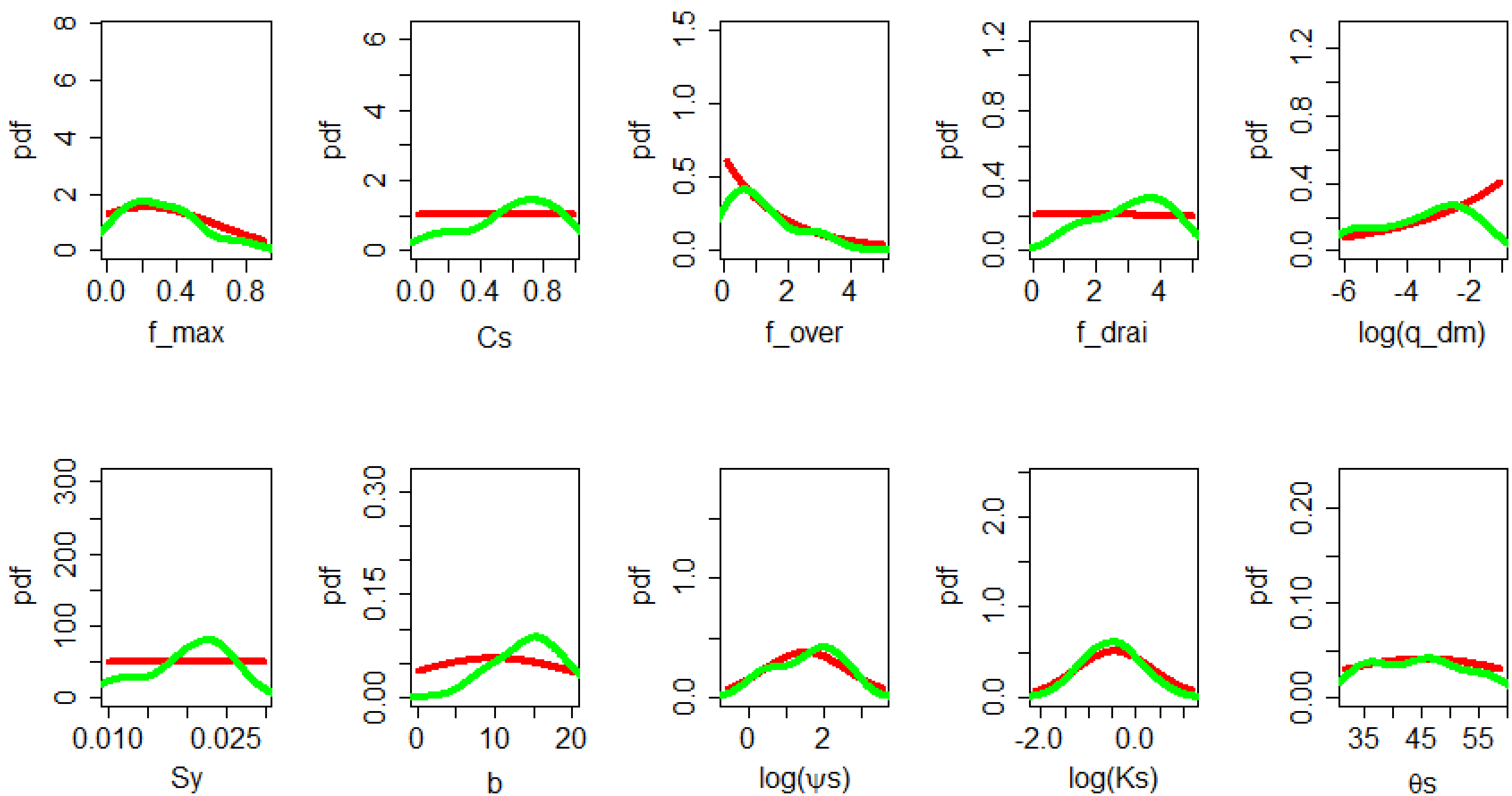
- Metropolis-Hasting sampling with community land model simulator as the forward model (Sun et al 2013)
- MCMC inversion using computationally-efficient surrogates (Ray et al 2014, presentation MS29 Apr 1 2014, 9:30-9:55)
- Entropy-Bayesian inversion has the following features:
  - Honor all the prior knowledge in the form of MRE prior pdfs
  - Enables simultaneous ensemble sampling and therefore parallel computing
  - Can estimate dependence structure (i.e., Σ) of residuals ε=d-G(m) (differences between observed and simulated responses such as heat fluxes and streamflow rate), and help reduce data redundancy

Entropy-Bayesian Procedure

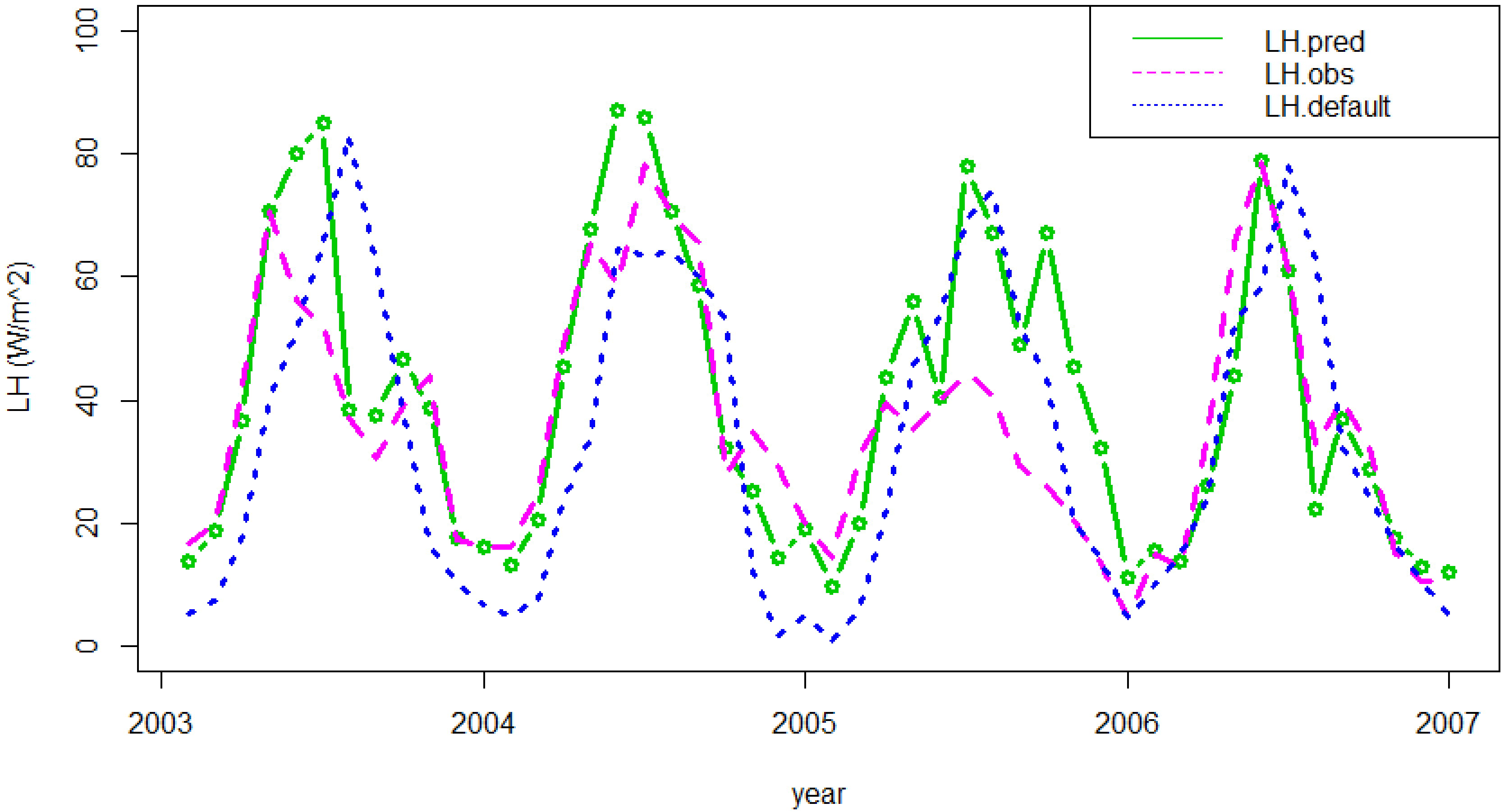
1. Derive closed-form prior pdfs using MRE theory, given bounds and statistical moments information about unknown parameters to be calibrated
2. Generate the first ensemble of parameter sets and simulate the land models in a task-parallel manner
3. Calculate the residuals at each data points for each sample set and then the error covariance matrix and likelihood functions
4. Generate a new ensemble of parameter sets from the current intermediate posterior pdfs
5. Repeat steps 3 and 4 until convergence of posterior pdfs

Posterior pdfs of hydrological parameters at US-ARM flux tower site

- Entropy-Bayesian procedure has been applied to the latent heat flux observation data for calibrating the major hydrological parameters



Posterior pdfs of hydrological parameters at US-ARM flux tower site

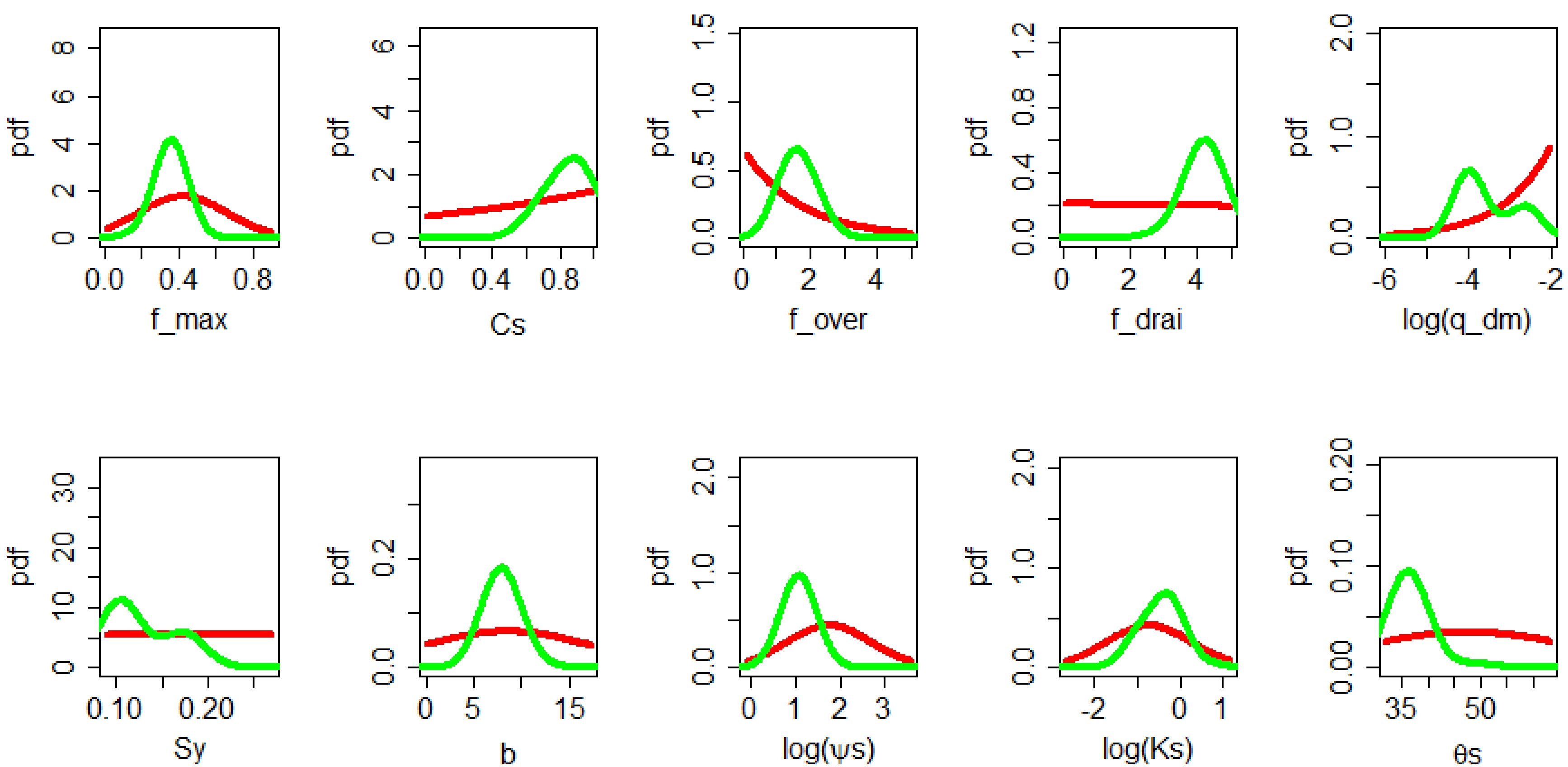


Posterior pdfs of hydrological parameters at a MOPEX basin (id 07147800)

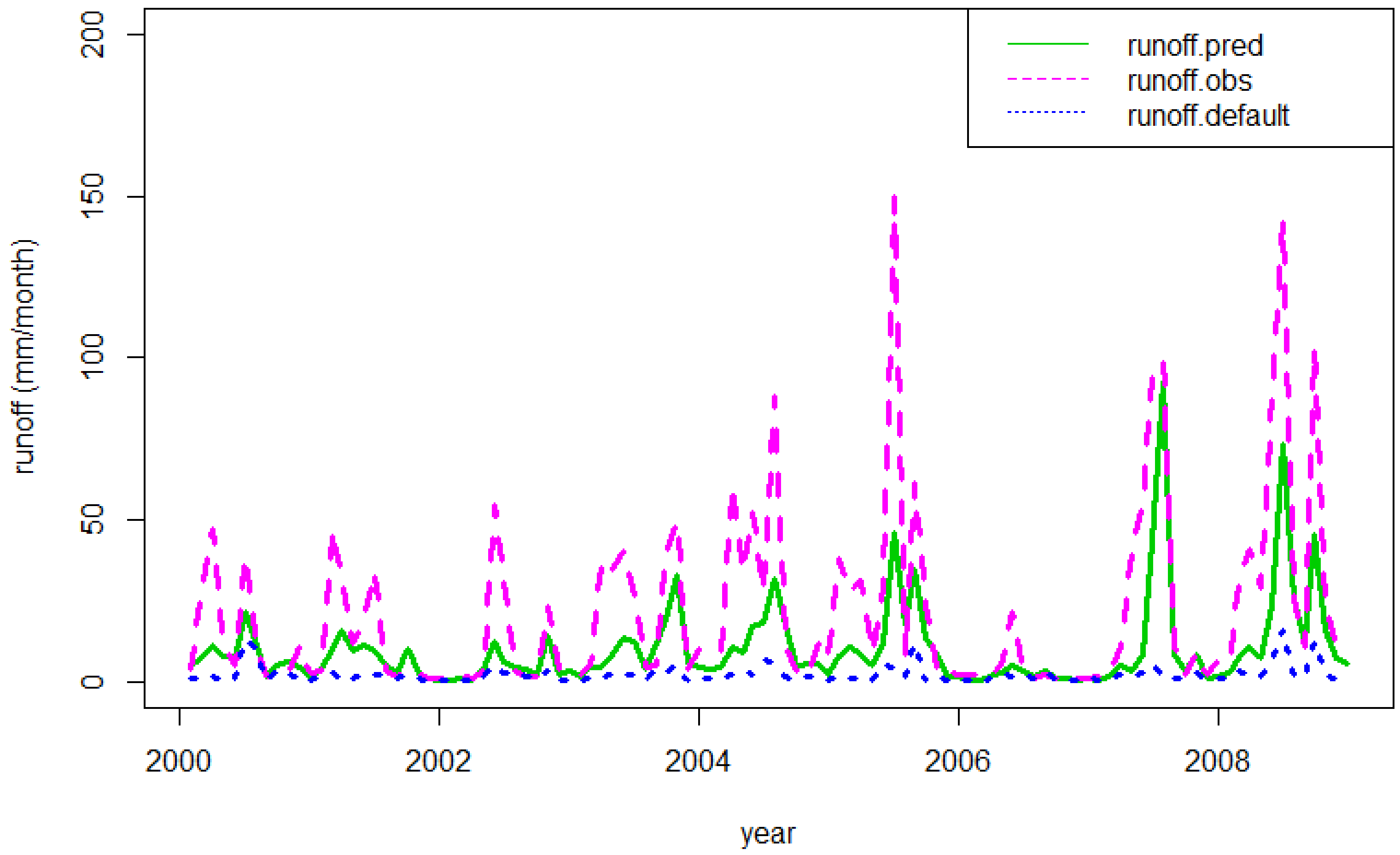
Pacific Northwest NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

- Entropy-Bayesian procedure has been applied to the runoff (streamflow) data for calibrating the major hydrological parameters



Posterior pdfs of hydrological parameters at US-ARM flux tower site



Summary

- Entropy (MRE) concept can be used to derive prior probability density functions that serves as the basis for further sensitivity analyses, parameter screening, uncertainty propagation, and surrogate development
- Integrated Entropy-Bayesian inversion enables ensemble sampling and task parallel computing, which helps reduce the computational time to achieve posterior estimates of unknown parameters
- Entropy-based UQ and Bayesian calibration have been applied to various flux tower sites and US MOPEX basins, for improving the Community Land Model (CLM) –
  - significant improvement in the modeling accuracy can be achieved using the calibrated parameters than using the existing hard-coded values