AGU Fall Meeting 2018
Email: ksargsy@sandia.gov

Bayesian Framework for Structural Uncertainty Estimation of Land Models

Khachik Sargsyan[1], Daniel Ricciuto[2], Cosmin Safta[1]

[1] Sandia National Laboratories, Livermore, CA, USA, [2] Oak Ridge National Laboratory, Oak Ridge, TN, USA

Goal: Model Error Quantification

Develop statistical framework for model error representation, quantification and propagation for physical models.

- Represent and estimate the error associated with
  - Simplifying assumptions, parameterizations
  - Mathematical formulation, theoretical framework
  - Numerical discretization
- Inverse modeling context $\quad y_i = f(x_i;\lambda) + \epsilon_i$
- Given data, calibrate for λ, accounting for model error
- Model error is deviation from 'truth'

Truth $g(x) \neq f(x;\lambda)$ Model

Embedded Model Error Representation

x are operating conditions, design parameters, various QoIs
λ are model parameters to be inferred/calibrated

- Default: Ignore model errors:

$$g(x) = f(x;\lambda) + \epsilon$$

  - Biased or overconfident physical parameters
  - Wrong model predictions

- Conventional: Correct for model errors:

$$g(x) = f(x;\lambda) + \delta(x) + \epsilon$$

  - Physical parameters are ok
  - Wrong model predictions (data-specific corrections)
  - Model and data errors mixed up

- What we do: Correct *inside* the model:

$$g(x) = f(x;\lambda + \delta(x)) + \epsilon$$

  - Embedded model error
  - Preserves model structure and physical constraints
  - Disambiguates model and data errors
  - Allows meaningful extrapolation

Bayesian Estimation of Model Error

$$y_i = f(x_i;\lambda + \delta_\alpha) + \epsilon_i$$

- Given data y_i, perform *simultaneous* estimation of $\tilde{\alpha} = (\lambda, \alpha)$, i.e. model parameters λ and model-error parameters α.
- Bayes' theorem

$$\overbrace{p(\tilde{\alpha}|y)}^{\text{Posterior}} = \frac{\overbrace{p(y|\tilde{\alpha})}^{\text{Likelihood}}\;\overbrace{p(\tilde{\alpha})}^{\text{Prior}}}{\underbrace{p(y)}_{\text{Evidence}}}$$

- To estimate the likelihood $L_y(\tilde{\alpha}) = p(y|\tilde{\alpha}) = p(y|\lambda,\alpha)$, one needs uncertainty propagation through $f(x_i;\underbrace{\lambda+\delta_\alpha}_{\text{stochastic}})$

Forward Prediction with Polynomial Chaos

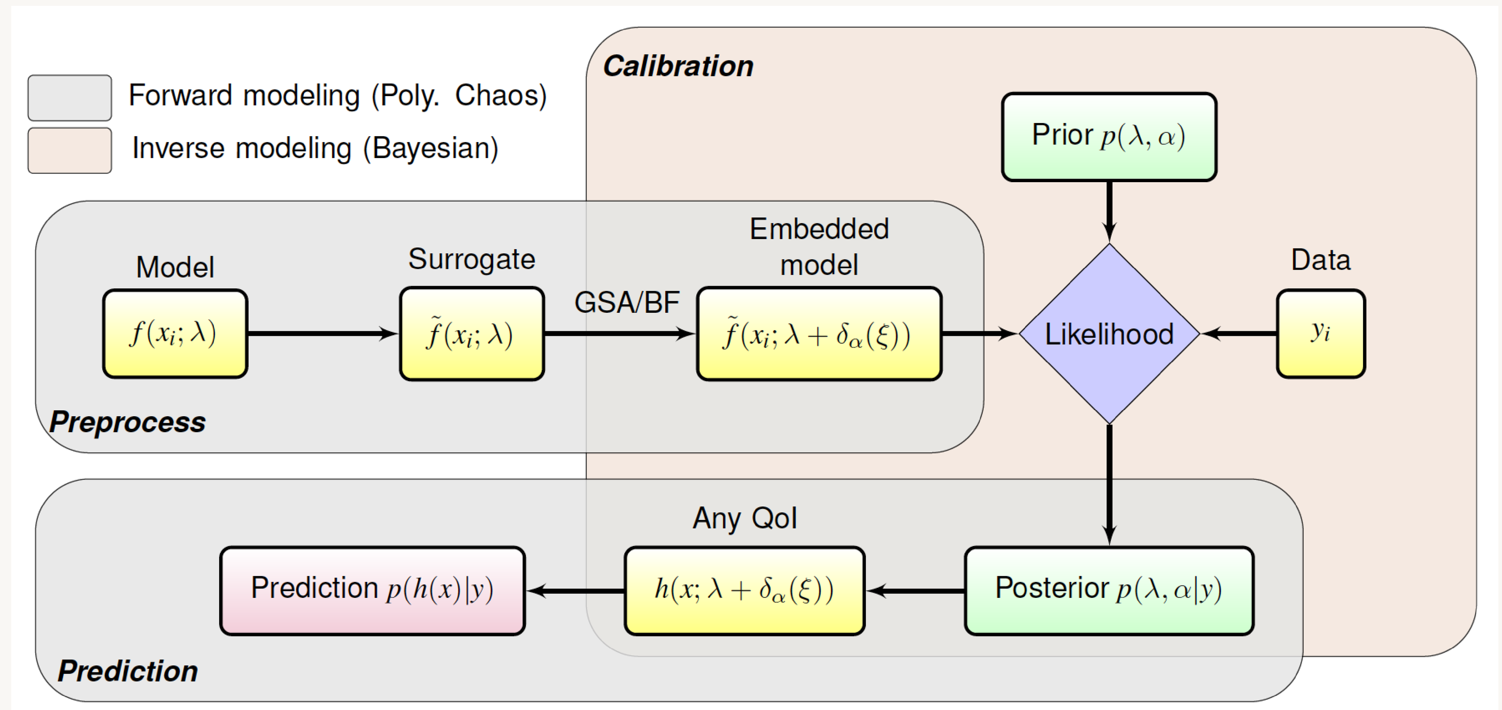
$$f(x;\lambda+\delta_\alpha) = f\left(x;\textstyle\sum_k \alpha_k \Psi_k(\xi)\right) \stackrel{NISP}{=} \textstyle\sum_k f_k(x;\alpha)\Psi_k(\xi)$$

- Non-intrusive spectral projection (NISP) employed for
  - Likelihood computation and posterior predictions
  - Easy access to first two moments:

$$\mu(x;\alpha) = f_0(x;\alpha), \qquad \sigma^2(x;\alpha) = \sum_{k>0} f_k^2(x;\alpha)||\Psi_k||^2$$

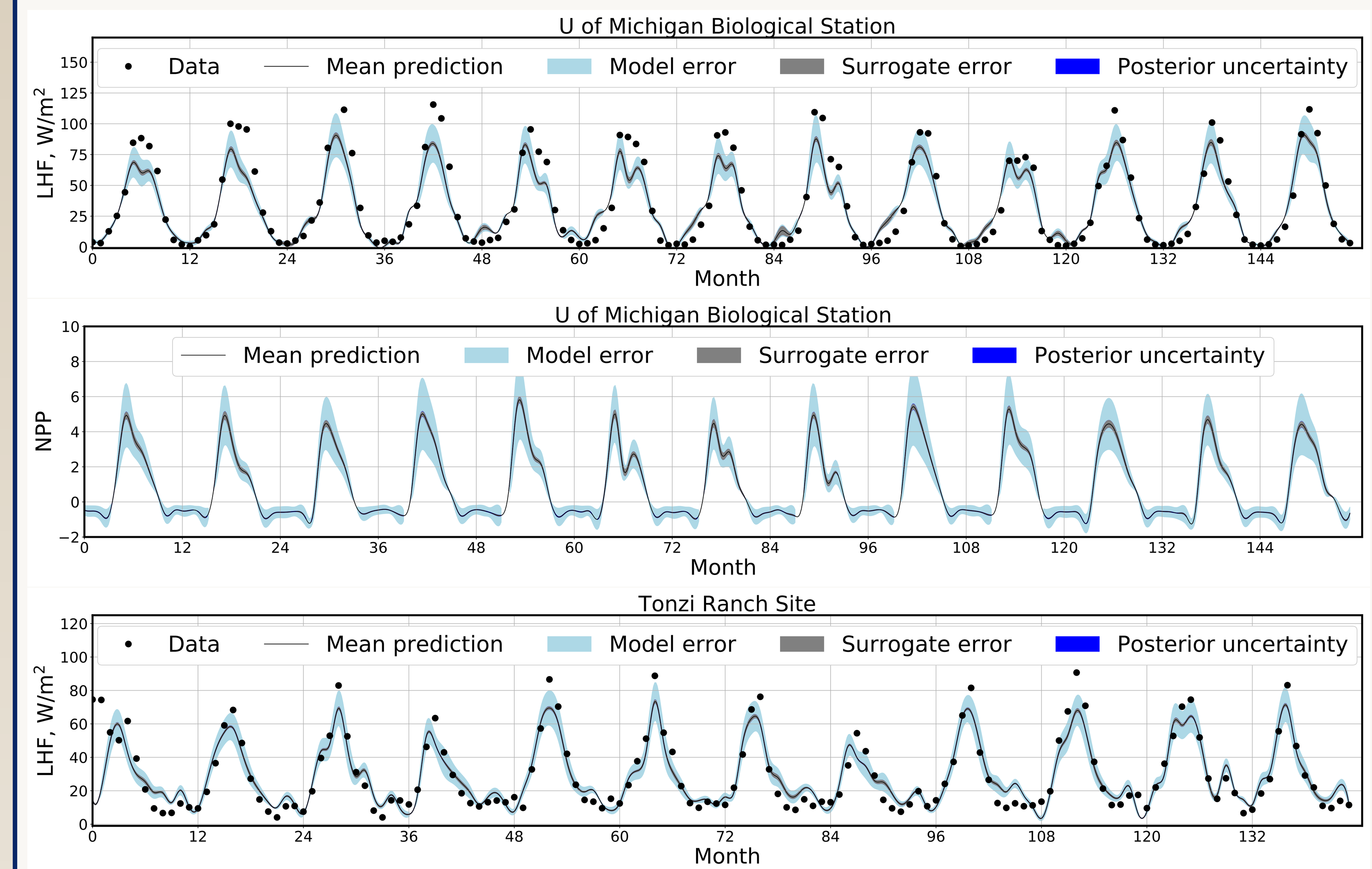
- Predictive mean $\quad \mathbb{E}[y(x)] = \mathbb{E}_\alpha[\mu(x;\alpha)]$
- Decomposition of predictive variance

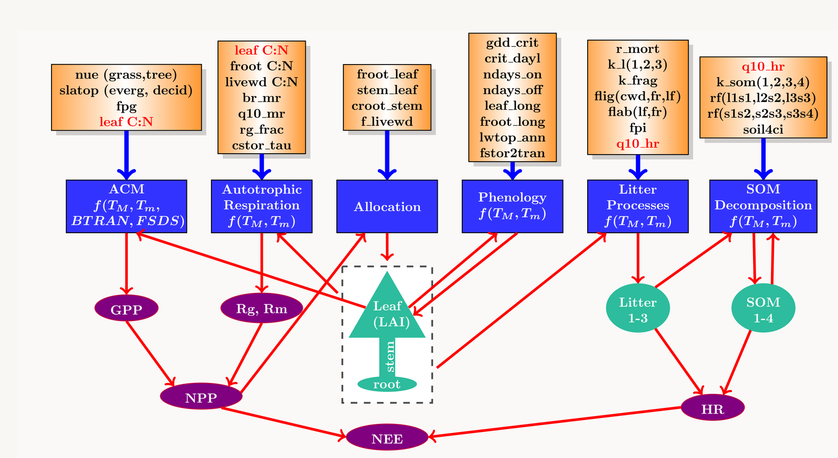
$$\mathbb{V}[y(x)] = \underbrace{\mathbb{E}_\alpha[\sigma^2(x;\alpha)]}_{\text{Model error}} + \underbrace{\mathbb{V}_\alpha[\mu(x;\alpha)] + \sigma_d^2}_{\text{Poserior/Data error}}$$



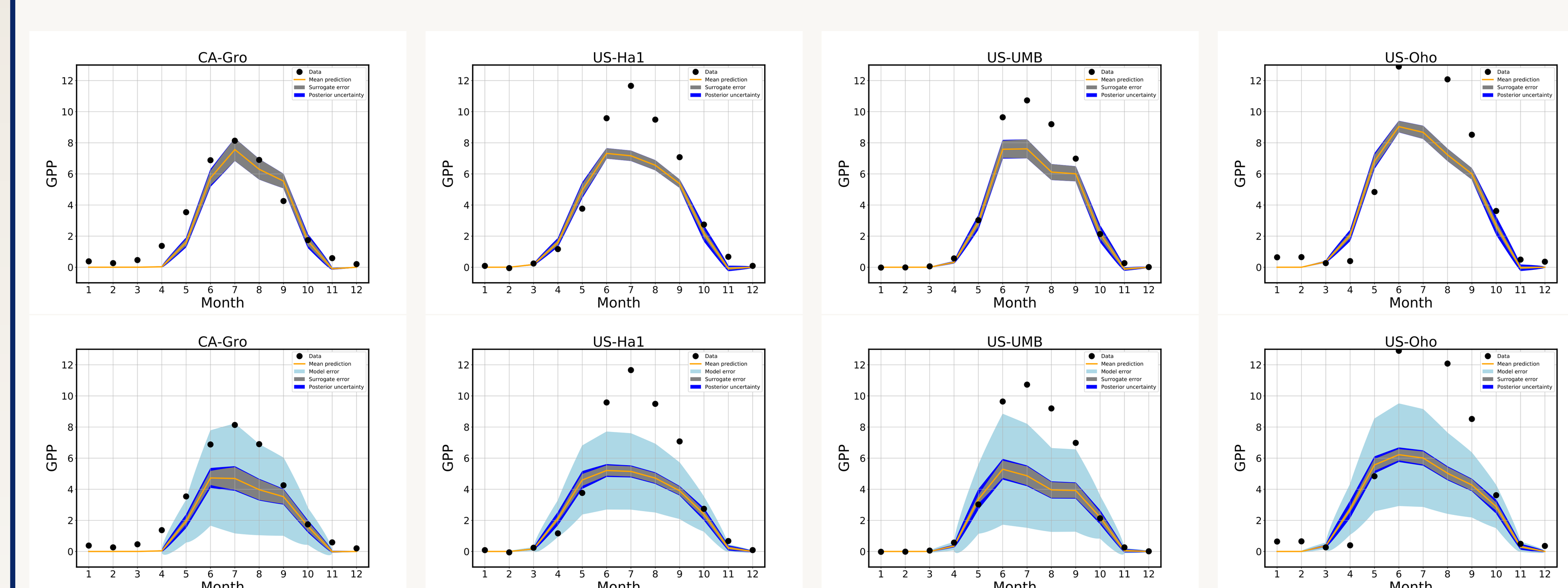
Application: E3SM Land Model

- US Dep-t of Energy (DOE) Earth system model
- Land, atmosphere, ocean, ice, human system components; high-resolution
- Employ DOE leadership computing facilities





- **ELM-LF** is a lower-fidelity, python version of ELM.
- Calibration with select FLUXNET sites data.
- Model error is the dominant uncertainty component: removes biases and overfitting.



Reference: K. Sargsyan, X. Huan, H. Najm, Embedded model error representation for Bayesian model calibration, arXiv preprint arXiv:1801.06768, 2018.

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (BER). the U.S. Department of Energy, Office of Science, Scientific Discovery through Advanced Computing (SciDAC) program through the FASTMath Institute.