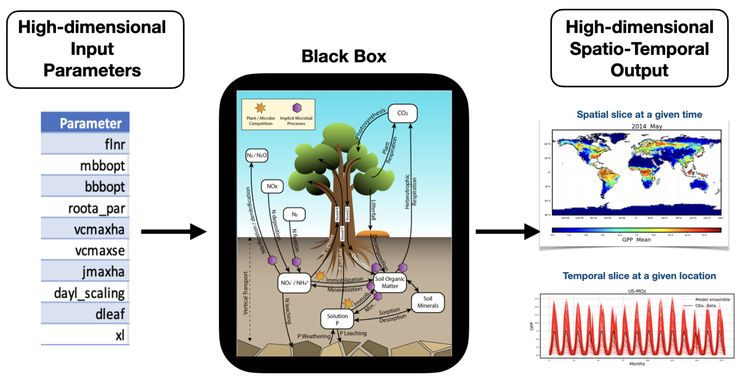
Hit twice by the curse of dimensionality: spatio-temporal land model calibration using Karhunen-Loève expansions

AGU Fall Meeting 2021
Email: ksargsy@sandia.gov

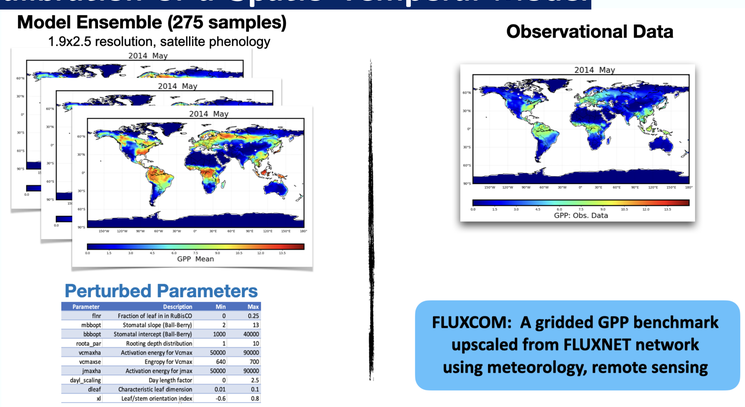
Khachik Sargsyan[1], Daniel Ricciuto[2], Cosmin Safta[1]

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

E3SM Land Model



Calibration of a Spatio-Temporal Model



Goal:
Calibrate model given observational data

$$f(\lambda; x, t) \approx g(x, t)$$

Need for a Surrogate Model

- Forward modeling: Global sensitivity analysis requires prohibitively many model evaluations
- Inverse modeling: Likelihood requires online evaluation of the model, e.g.

$$p(D|\lambda) \propto \exp\left(-\frac{1}{2}||f(\lambda; x, t) - g(x, t)||^2\right)$$

- Construct a surrogate, inexpensive approximation

$$f(\lambda; x, t) \approx f_s(\lambda; x, t)$$

using the perturbed ensemble of model simulations.

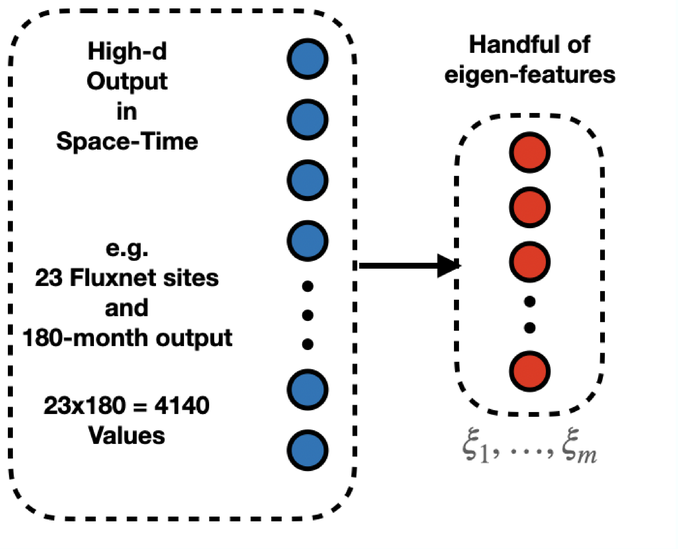
High-Dimensional Output

Karhunen-Loève (KL) Expansion:

- Optimal linear expansion in terms of eigenvalues of the covariance kernel.

$$f(\lambda; x, t) = \bar{f}(x, t) + \sum_{j=1}^{\infty} \xi_j(\lambda)\sqrt{\mu_j}\phi_j(x, t)$$

- Similar to principal components analysis, but in the continuous, (x, t) space.
- The analysis of spatio-temporal field $f(\lambda; x, t)$ reduces to a handful of eigen-features $\xi_j(\lambda)$
- KL expansion is an optimal linear auto-encoder



High-Dimensional Input

Option 1: Sparse Polynomial Chaos (PC) Surrogates for each $\xi_j(\lambda)$

$$\xi_i(\lambda) \approx \xi_i^{\mathrm{PC}}(\lambda) = \sum_{k=0}^{K} c_{ik}\Psi_k(\lambda)$$

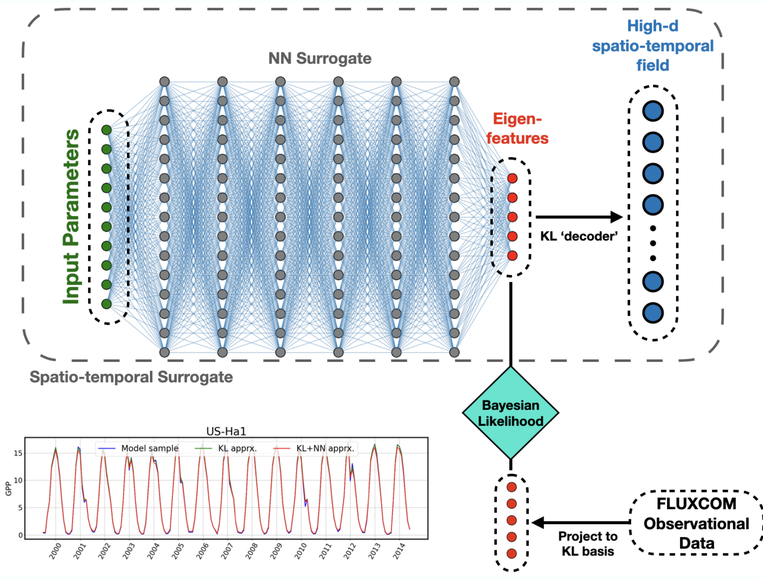
Option 2: Joint Neural Network (NN) Surrogate for all $\xi_j(\lambda)$'s

$$\boldsymbol{\xi}(\lambda) \approx \boldsymbol{\xi}^{\mathrm{NN}}(\lambda)$$

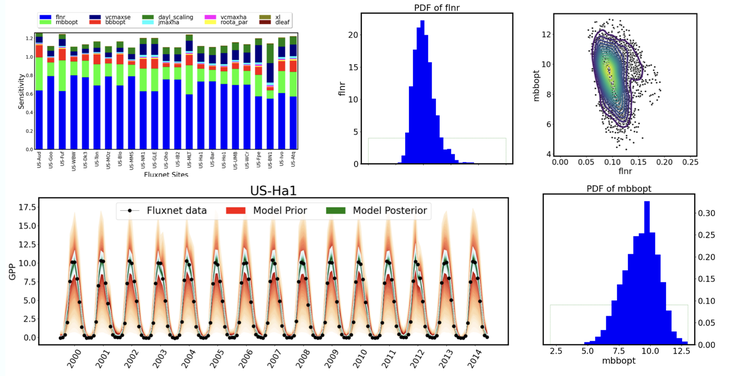
Surrogate-enabled Bayesian Calibration

Joint spatio-temporal surrogate

$$f(\lambda; x, t) = \bar{f}(x, t) + \sum_{j=1}^{m} \xi_j^{\mathrm{NN}}(\lambda)\sqrt{\mu_j}\phi_j(x, t)$$



Results and Summary



- High-dimensional spatio-temporal output well-represented with a few KL eigen-features with a fraction of ELM cost
- 4140 spatio-temporal cells down to 5 eigen-features with 1% loss of accuracy only
- Global sensitivity analysis using the KL+NN surrogate
- Bayesian calibration performed in the eigen-space

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).