AGU Fall Meeting 2016
Email: ksargsy@sandia.gov

Sandia National Laboratories   Oak Ridge National Laboratory

Weighted Iterative Bayesian Compressive Sensing (WIBCS) for High Dimensional Polynomial Surrogate Construction

Khachik Sargsyan[1], Daniel Ricciuto[2], Cosmin Safta[1], Bert Debusschere[1], Habib N. Najm[1], Peter Thornton[2]

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

Goal: Parametric Uncertainty Quantification

Quantify uncertainties due to parameter variability in high-dimensional, computationally expensive physical models.

Key: Build a Surrogate Model

$G(x) \approx g_c(x)$ **... otherwise called metamodel, proxy, response surface, emulator.**

- Surrogate approximation to input-output map
- Computationally inexpensive to evaluate
- Replaces the model in forward UQ (parametric UQ, global sensitivity analysis) and inverse UQ (calibration, optimization, tuning)

Polynomial chaos (PC) as a surrogate model.

- Interprets input parameters as random variables
- Propagates input uncertainties to outputs of interest, e.g.,

$$x_i \sim \text{Uniform}[a_i, b_i] \quad \Rightarrow \quad x_i = \frac{a_i + b_i}{2} + \frac{b_i - a_i}{2}\xi_i.$$

Output is represented with respect to Legendre polynomials of standard input $\xi_i \sim \text{Uniform}[-1, 1]$

$$G(x(\xi)) \approx g_c(\xi) \equiv \sum_{k=0}^{K} c_k \Psi_k(\xi).$$

Challenge: High Dimensionality

- Large number of input parameters, $x \in \mathbb{R}^d$ with $d \sim 30$ or above
- Model is expensive; $N \gg 1$ simulations of $G(x)$ are required
- Truncation of PC is challenging; too many basis terms $K > N$

- **Compressive sensing** allows learning sparse polynomial expansion and deals with high-dimensionality
- **Bayesian inference** works with any number of model simulations and constructs uncertain surrogate

Weighted Iterative Bayesian Compressive Sensing

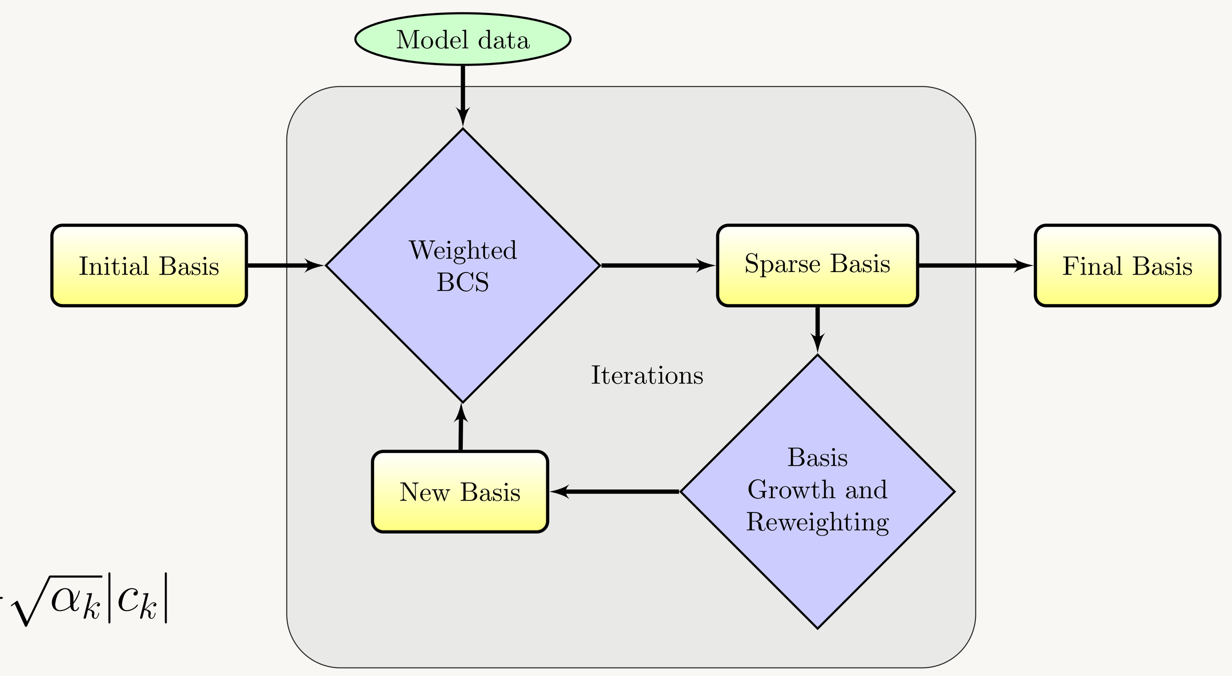
Bayesian inference of PC coefficients: $\underbrace{p(c|\mathcal{D})}_{\text{Posterior}} \propto \underbrace{L_{\mathcal{D}}(c)}_{\text{Likelihood}} \underbrace{p(c)}_{\text{Prior}}$

- Gaussian likelihood:

$$L_{\mathcal{D}}(c) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}s} \exp\left(-\frac{(G(x_i) - g_c(\xi_i))^2}{2s^2}\right)$$

- Hierarchical Laplace prior:

$$p_\alpha(c) = \int \prod_{k=0}^{K-1} \underbrace{p(c_k|\sigma_k^2)}_{\text{Gaussian}} \underbrace{p_\alpha(\sigma_k^2)}_{\text{Gamma}} d\sigma_k^2 = \prod_{k=0}^{K-1} \frac{\sqrt{\alpha_k}}{2} e^{-\sqrt{\alpha_k}|c_k|}$$



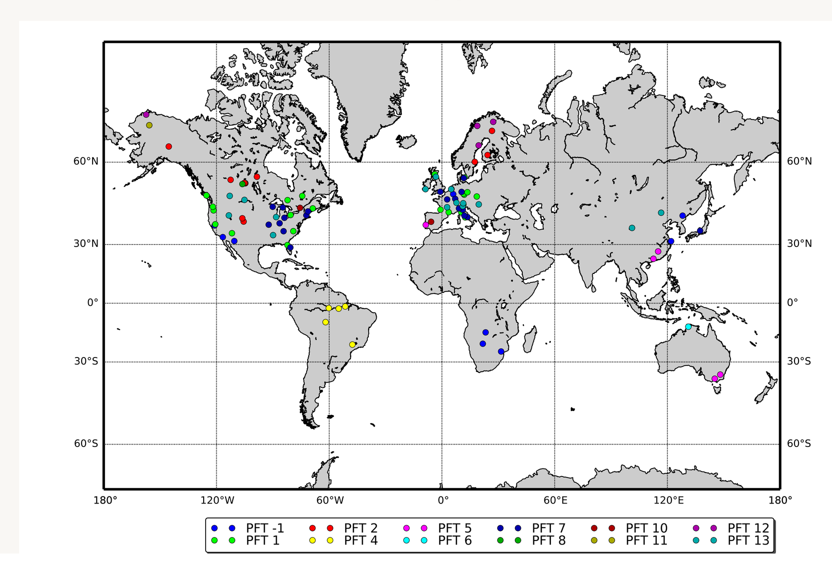
- **Weighted** Laplace prior, hyperparameter vector α_k
- **Iterative growth** of polynomial bases to capture high order terms
- Implemented in **UQTk** software library (www.sandia.gov/uqtoolkit)

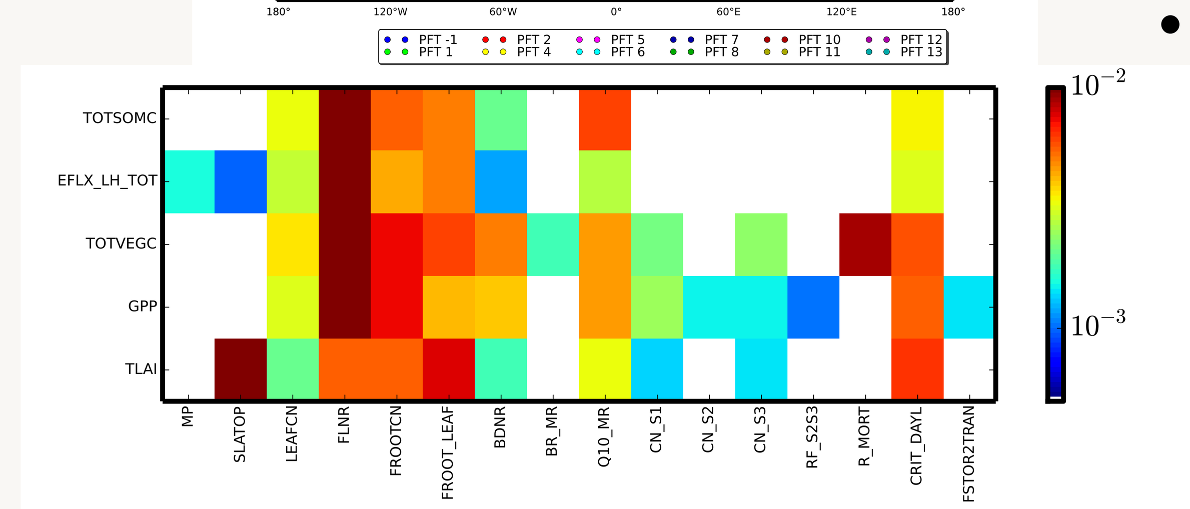
K. Sargsyan, C. Safta, H. Najm, B. Debusschere, D. Ricciuto, P. Thornton, "Dimensionality Reduction for Complex Models via Bayesian Compressive Sensing", *International Journal of Uncertainty Quantification*, 4(1), pp. 63-93, 2014.

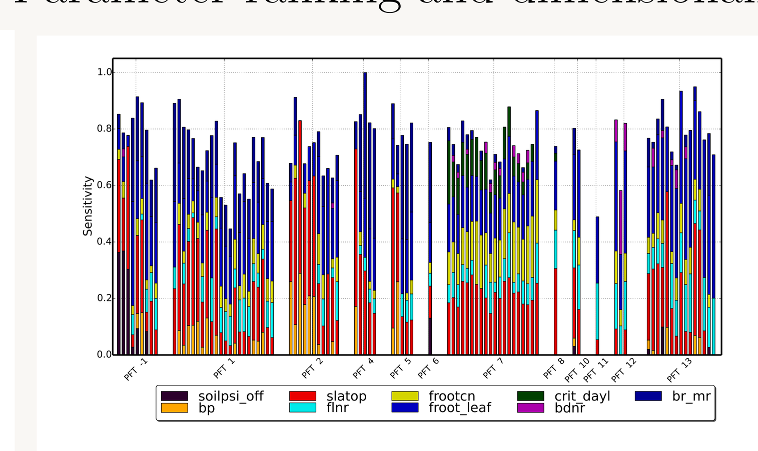
Application: ACME Land Model

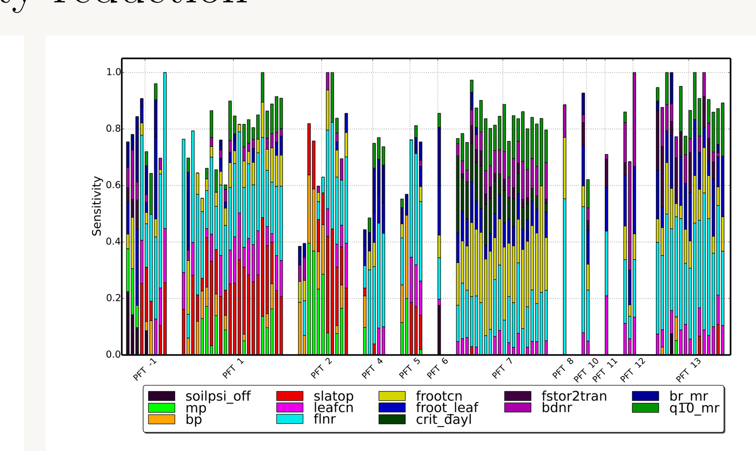
ACME Accelerated Climate Modeling for Energy

- Study 5 steady state outputs for 96 FLUXNET sites:
  - Gross primary productivity (GPP)
  - Total leaf area index (TLAI)
  - Total vegetation carbon (TOTVEGC)
  - Total soil organic matter carbon (TOTSOMC)
  - Latent heat flux (EFLX_LH_TOT)
- Vary 68 input parameters over selected ranges
- Built surrogate model approximations (per site, per output) w.r.t 68 inputs using 3000 ensemble simulations
- Performed variance-based decomposition
- Sensitivities grouped according to Plant Functional Types (PFTs)
- Only ∼15 (out of 68) parameters have significant impact to overall uncertainty
- Parameter ranking and dimensionality reduction









D. Ricciuto, K. Sargsyan, P. Thornton, "The Impact of Parametric Uncertainties on Biogeochemistry for ACME Land Model", *in prep.*, 2016.

U.S. DEPARTMENT OF ENERGY

This research was supported as part of the Accelerated Climate Modeling for Energy (ACME) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

National Nuclear Security Administration