ACME

Accelerated Climate Modeling for Energy

Parametric Uncertainty Analysis for ACME Land Model

Khachik Sargsyan (SNL-CA), Daniel Ricciuto (ORNL)

Nov 5, 2015

Forward UQ Analysis for Multiple Sites

- 96 FLUXNET sites covering major biomes and plant functional types
- Varying 68 parameters over given ranges
- Ensemble of 3000 runs on Titan
- 5 steady state QoIs extracted

Create Surrogates for Input-Output Maps

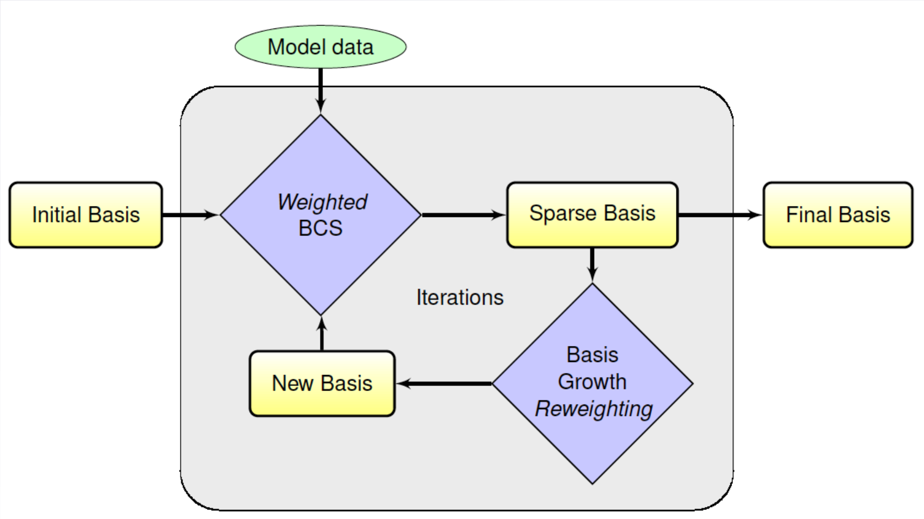
Surrogate models are needed for computationally intensive tasks:

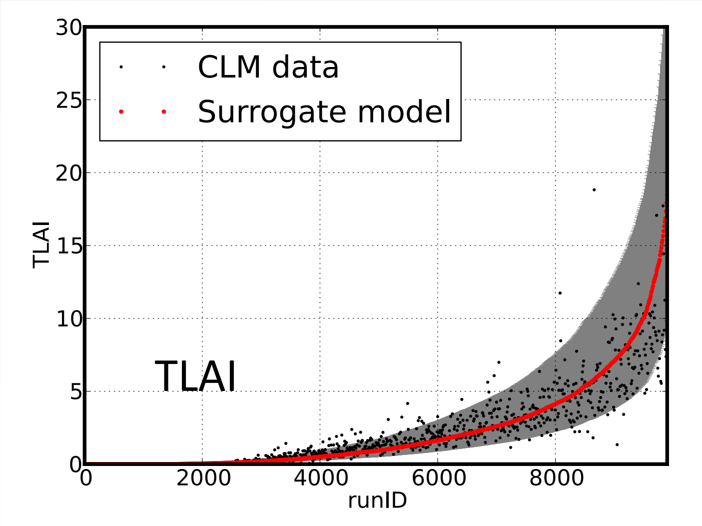
- Parameter estimation
- Optimization
- Experimental/computational design
- Forward uncertainty propagation

Major UQ challenges

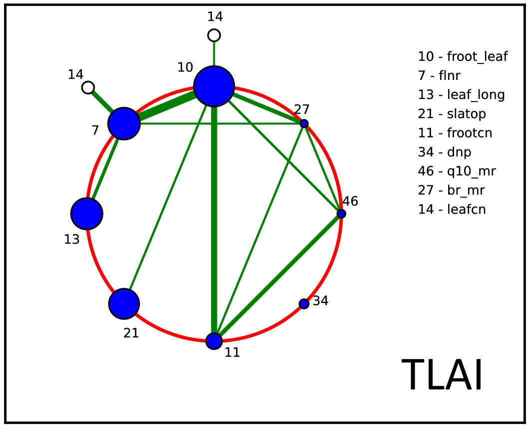
- Scarce information (3000 samples in 68-dimensional space)
- High-dimensionality (How to represent such a high-d function?)

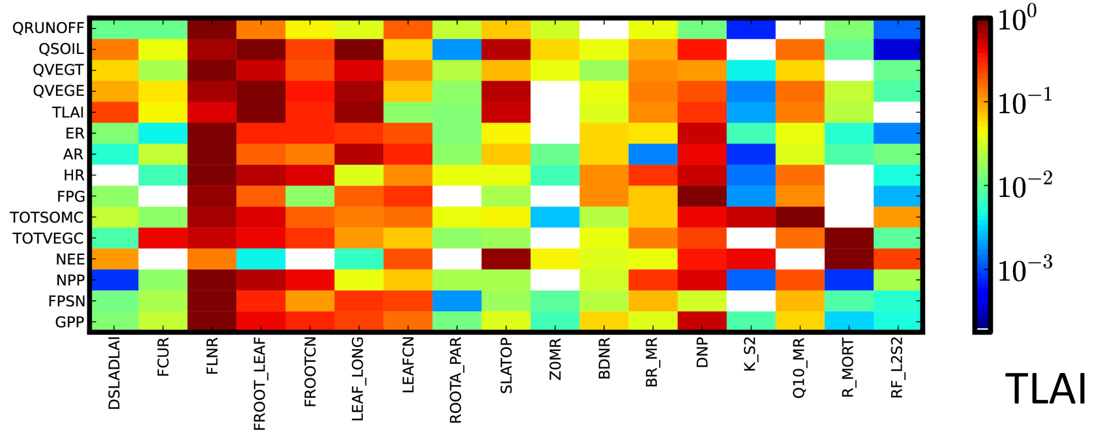
- Tools from machine learning for polynomial basis selection
- Iterative Bayesian Compressive Sensing -> Uncertain Surrogate

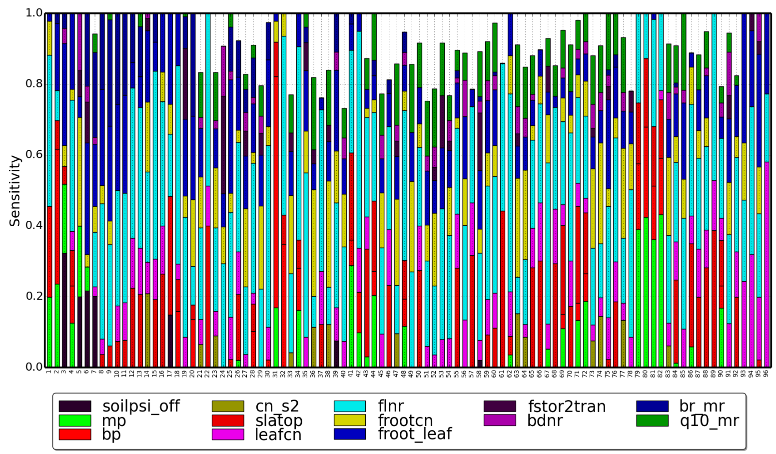


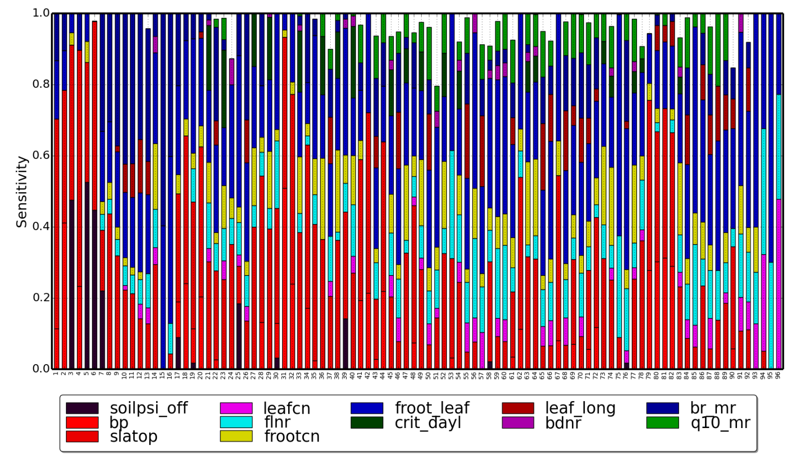


Uncertainty Decomposition / Global Sensitivity Analysis









Current / Future

- Forward UQ workflow for automatic parameter ranking
- A set of Python scripts as an interface to UQTk v2.2
- Expand toward other users

- Lower-dimensional, more detailed ensemble
- Create automatic workflow for Inverse UQ
- Calibration / Parameter Tuning with surrogates

Major goal: create a *surrogate model*

Surrogate model is a "good-enough" approximation of the full model over a range of parameter variability.

… otherwise called

- Metamodels
- Response surfaces
- Emulators
- Low-fidelity model

Black Box

$$Y = f(X)$$

Surrogate models are needed for computationally intensive tasks:

- Parameter estimation
- Optimization
- Experimental/computational design
- Forward uncertainty propagation

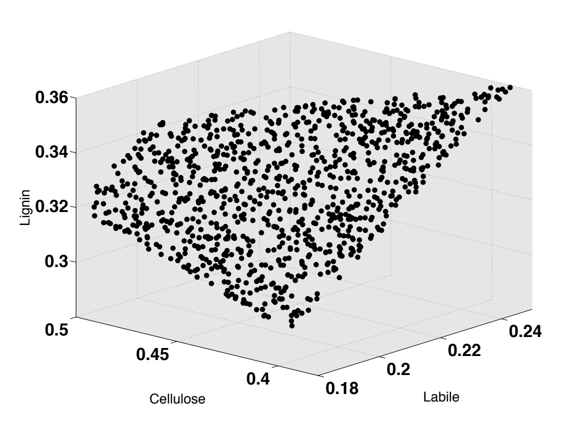
$$f(X) \approx f_{surr}(X)$$

Polynomial Chaos for the input

$$X = \sum_{k=0}^{P-1} a_k \Psi_k(\xi)$$

The input PC can be obtained from

- Preliminary calibration
- ✓ Expert opinion
- ✓ Accounting for constraints



Often it is a simple matter of scaling, i.e. linear PC:

- Gauss-Hermite $$X = \mu + \sigma\xi$$
- Legendre-Uniform $$X = \frac{a+b}{2} + \frac{b-a}{2}\xi$$

In such cases, the PC surrogate is simply a polynomial fit/regression.

Polynomial Chaos is the main workhorse

PC provides convenient means of representing model inputs and outputs
In a probabilistic way.

$$X = \sum_{k=0}^{P-1} a_k \Psi_k(\xi) \qquad \boxed{Y = f(X)} \qquad Y = \sum_{k=0}^{P-1} c_k \Psi_k(\xi)$$

ξ are standard variables (uniform, normal)

$\Psi_k(\cdot)$ are standard orthogonal polynomials (Legendre, Hermite)

- Think of Fourier-type expansions, only w.r.t. polynomials.
- Uncertain inputs *X* and outputs *Y* are represented via vectors of PC modes a_k and c_k

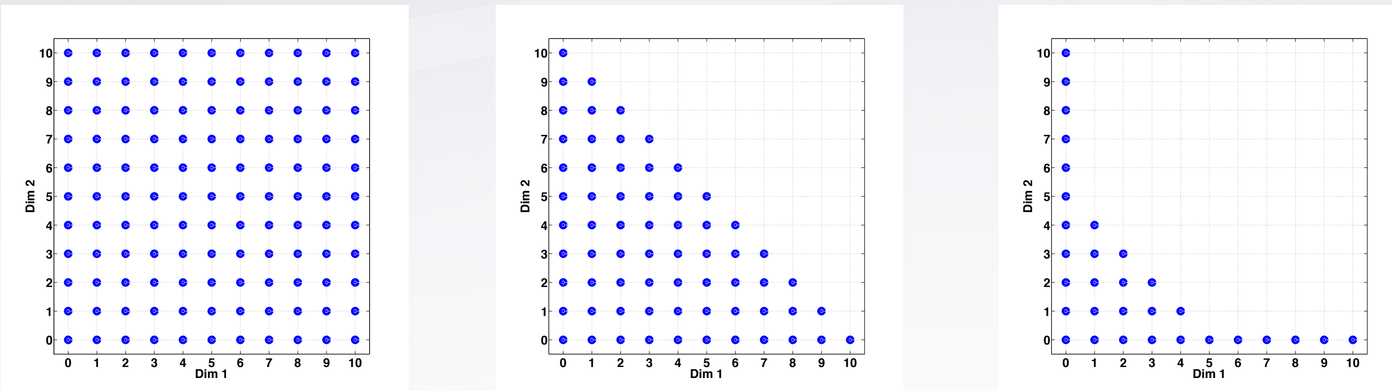
Non-intrusive setting: black-box ALM

Run the model at selected parameter settings: $y_i \approx CLM(x_{i,})$ for *i=1, …, N.*

Find PC surrogate coefficients s.t. $\vec{y} \approx A\vec{c}$ where $A_{ij} = \Psi_j(x_i)$

- Classical least-squares: $\min_{\vec{c}} \| \vec{y} - A\vec{c} \|_2$
- Bayesian inference: more flexible, provides errorbars on coefficients.

Key challenge: how to truncate polynomial expansion? Often *N<P*.



$$Y = \sum_{k=0}^{P-1} c_k \Psi_k(X)$$

(Bayesian) Compressed Sensing helps find the sparsest signal,
i.e. selects as few polynomial terms as possible.