Sandia Software Enabling Extreme-Scale Uncertainty Quantification

Michael Eldred, Bert Debusschere, Kenny Chowdhary, John Jakeman, Habib Najm, Prashant Rai, Cosmin Safta, Khachik Sargsyan
Sandia National Laboratories

www.quest-scidac.org www.sandia.gov

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.

U.S. DEPARTMENT OF ENERGY Office of Science

Support for this work was provided through the Scientific Discovery through Advanced Computing (SciDAC) project funded by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research.

Process of quantifying the effect of uncertainties typically includes:

- *(Global) sensitivity analysis:* identification of input set with greatest influence on output QoIs
- *Uncertainty characterization:* model or infer from observable data; parametric/non-parametric/KDE
- *Uncertainty propagation:* input distributions → output QoI distributions
- *Decision making:* model validation, prediction, design under uncertainty

SNL software tools within QUEST support a range of:

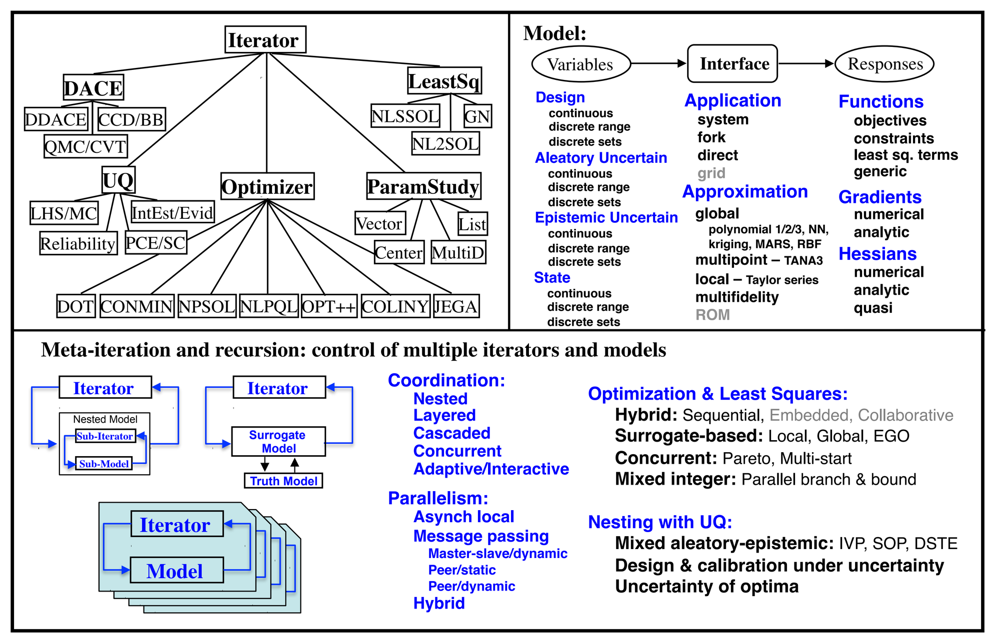
- *UQ studies:* sensitivity analysis, uncertainty propagation, statistical inference
- *Environments:* rapid prototyping in interpreted languages ↔ production computing in compiled languages on parallel platforms
- *Intrusion:* embedded ↔ linked ↔ black box

An interoperable set of tools that can be tailored:

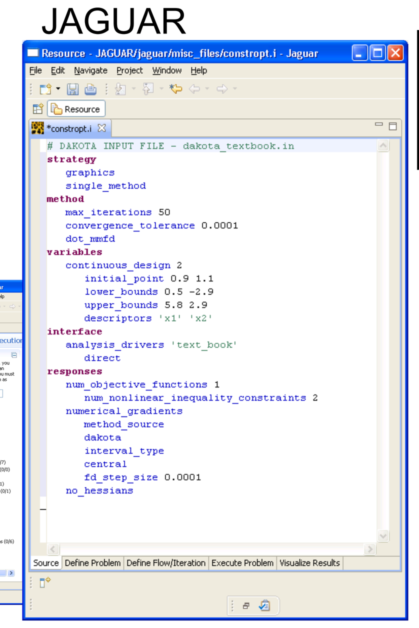
- DAKOTA + QUESO/GPMSA with PCE/SC/GP emulators
- Production deployment of stable capabilities in frameworks
- Close collaboration of SAPs with library developers for custom capabilities

DAKOTA (dakota.sandia.gov) is a C++ application that provides a variety of non-intrusive algorithms for design optimization, model calibration, uncertainty quantification, global sensitivity analysis, parameter studies, and solution verification. It can be used as either a stand-alone application or as a set of library services, and supports multiple levels of parallelism for scalability on both capability and capacity HPC resources.

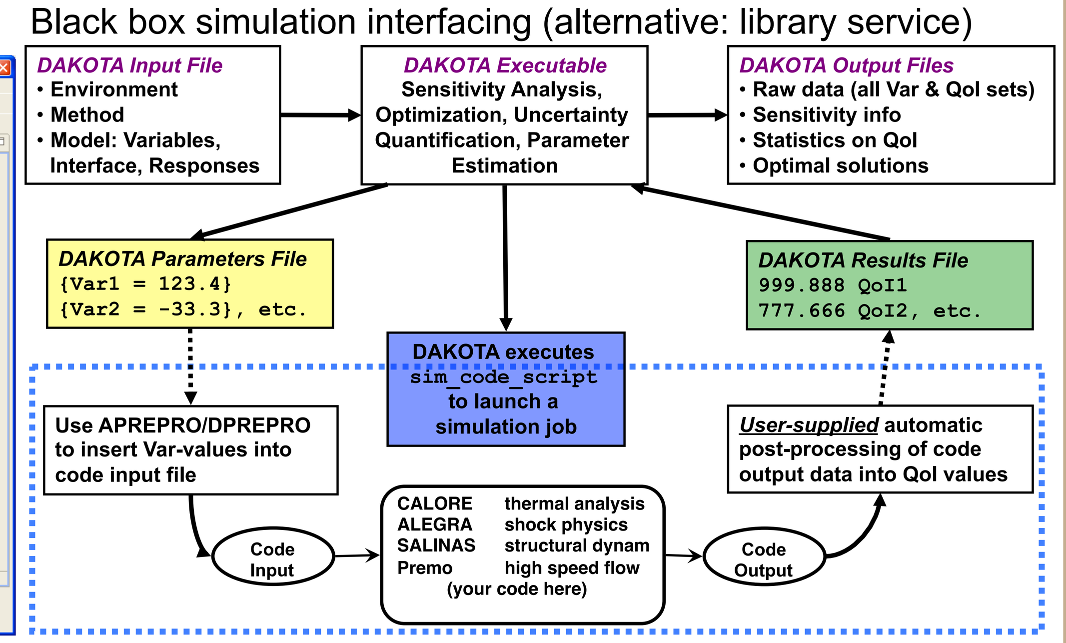
- Contact: dakota-developers@development.sandia.gov



JAGUAR



Black box simulation interfacing (alternative: library service)

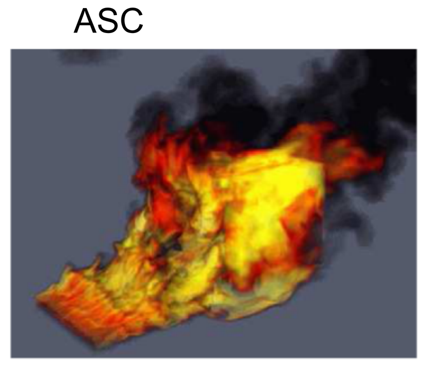


UQ Capabilities in v6.2 (released 5/15/15):

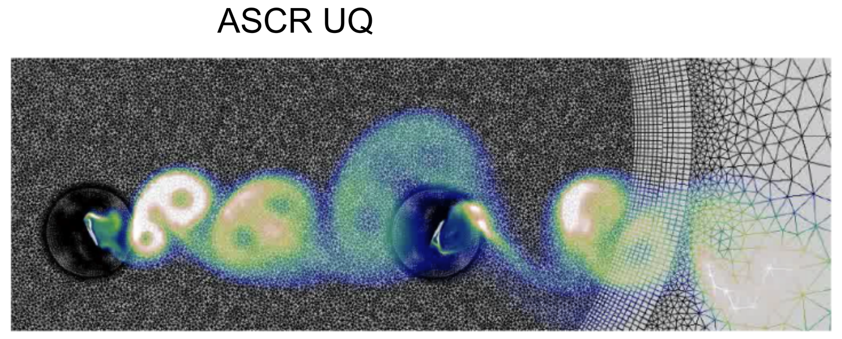
- **Sampling methods**
  - **Random: LHS, MC, Incremental**
  - **Importance: IS, AIS, MMAIS**
  - **Adaptive: Morse-Smale et al.**
- **Reliability methods**
  - **Local: MV, AMV, AMV+, AMV^2+, TANA-3, FORM, SORM**
  - **Global: EGRA, GPAIS, POF Darts, RKD**
- **Stochastic expansion methods**
  - **Polynomial chaos: projection, regression (see Algs poster)**
  - **Stochastic collocation: tensor & sparse; nodal & hierarchical**
- **Epistemic methods**
  - **Interval estimation: local, global, mixed-integer**
  - **Dempster-Shafer**
- **Bayesian methods**
  - **QUESO, GPMSA, DREAM**
  - **Emulator-based MCMC: PCE, SC, GP (Algs poster)**
  - **Random field inference (PISCEES at bottom)**
- **Meta-iteration and recursion**
  - **Mixed aleatory-epistemic UQ**
  - **Design / calibration under uncertainty**

Defense, Science, and Energy Applications

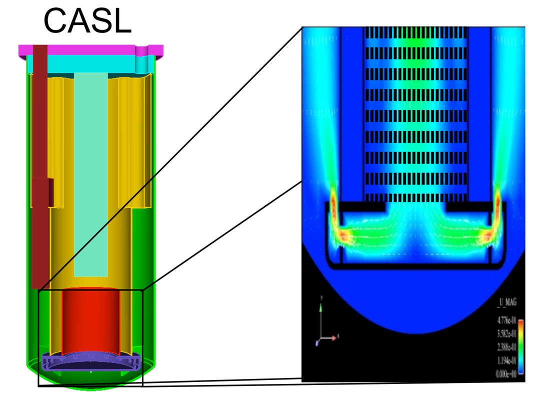
ASC



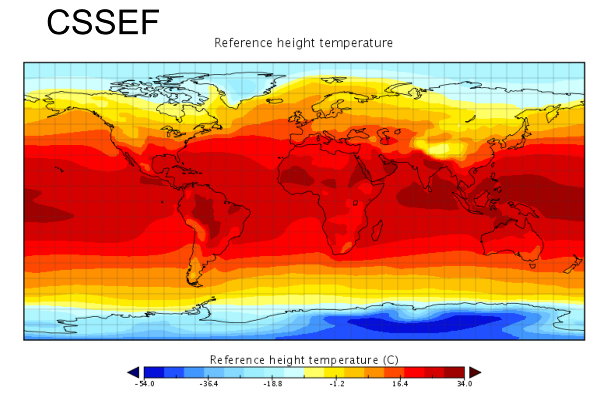
ASCR UQ



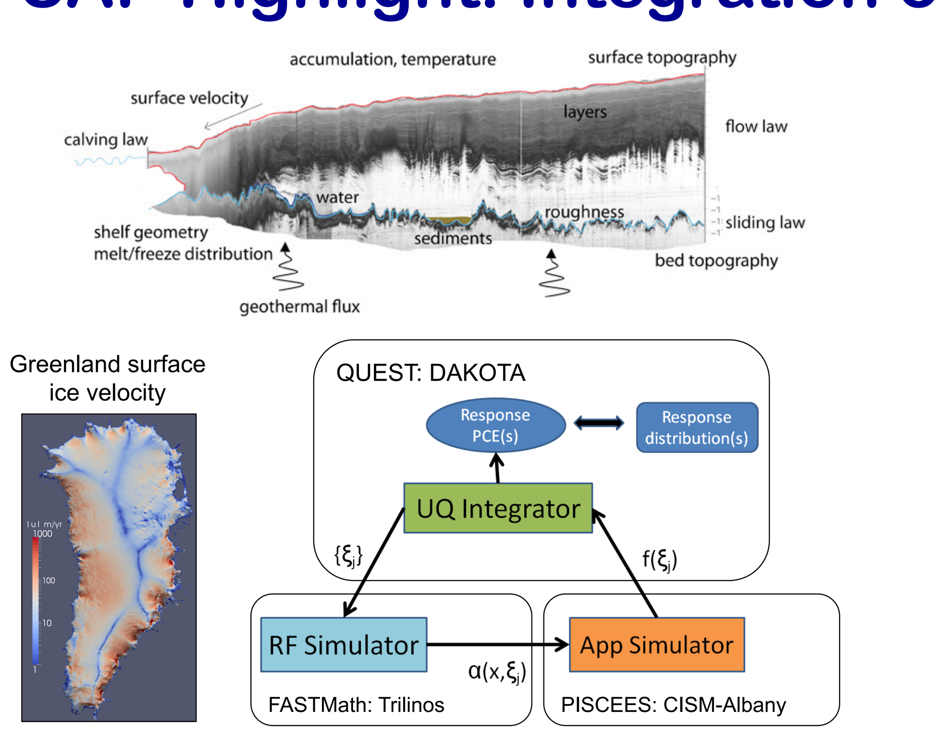
CASL



CSSEF



SAP Highlight: Integration of Albany/Dakota/Trilinos for PISCEES



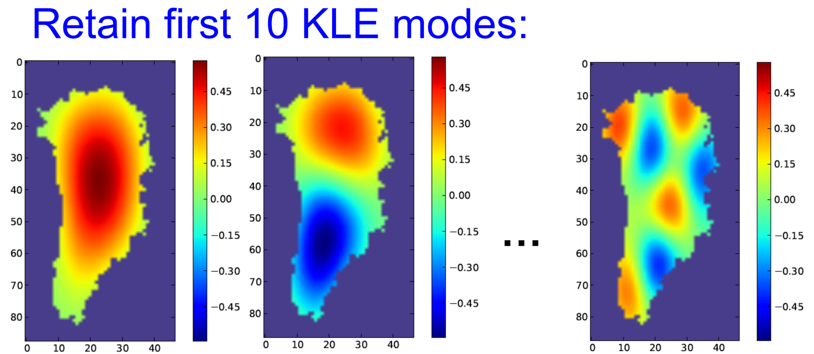
Karhunen-Loève expansion (KLE):

Assume analytic spatial covariance kernel (squared exponential) for random field:

$$C(r_1,r_2)=e^{-(r_1-r_2)^2/L^2}$$

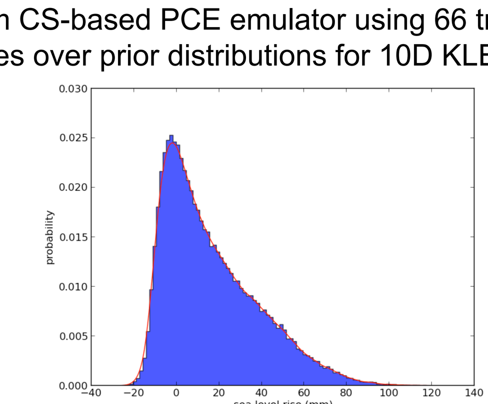
and integrate over domain for modes. Length scale (L) balances feature resolution vs. # KLE modes.

Retain first 10 KLE modes:



Dimension reduced inference and propagation for prediction of SLR:

- Form CS-based PCE emulator using 66 transient solves over prior distributions for 10D KLE parameters



- SLR statistics based on unconstrained priors
- Next steps: emulator-based inference (see Algs poster)
  → SLR statistics for posteriors constrained by surface velocity data

UQTk (www.sandia.gov/UQToolkit) is an LGPL open source library of functions for characterization and propagation of uncertainty in computational models.

- Mainly relies on spectral Polynomial Chaos Expansions (PCEs) for representing random variables and stochastic processes
- Complementary to production tools, UQTk targets:
  - Rapid prototyping
  - Algorithmic research
  - Outreach: Tutorials / Educational
- Contact: Bert Debusschere: bjdebus@sandia.gov

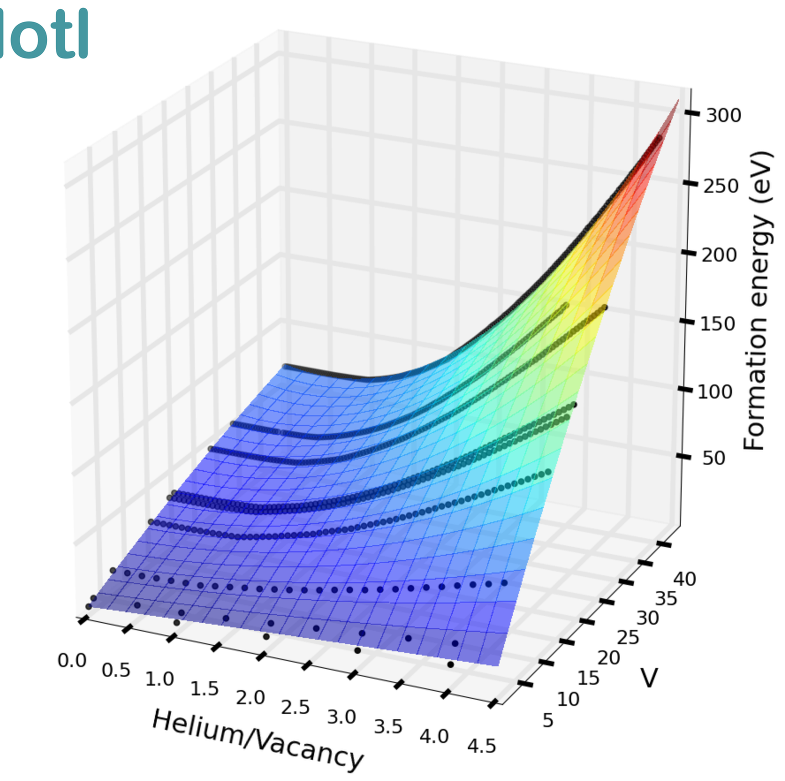
Capabilities:

- Intrusive and non-intrusive (quadrature) approaches for PCE stochastic Galerkin projection
  - Full and sparse quadrature approaches
- Markov Chain Monte Carlo library for Bayesian inference
- Bayesian Compressive Sensing
- Karhunen-Loève expansions
- Sensitivity analysis
- Core libraries in C++
- Examples and postprocessing tools in Python
- Python interface will be released as part of UQTk v3.0, Fall 2015

Applications:

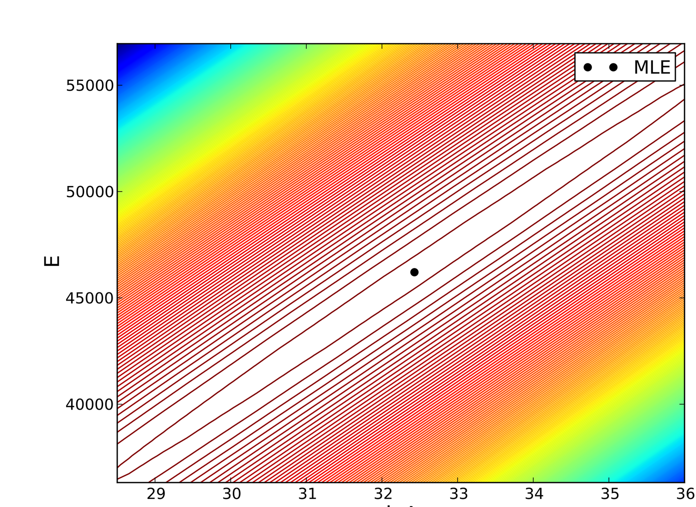
- UQTk components can be combined as needed into an end-to-end UQ workflow:
  - Surrogate construction → sensitivity analysis → parameter inference → PCE construction → forward propagation
  - Bayesian compressive sensing used in climate modeling for surrogate construction and dimensionality reduction of land, atmosphere and cloud models (CSSEF, ACME, Multiscale Earth Models, ACES4GCM)
  - UQ workflows set up in multiple SciDAC partnership projects: e.g. UQ in Xolotl (PSI)
- Development of lecture material and hands-on exercises for UQ tutorials
  - Nationally and Internationally

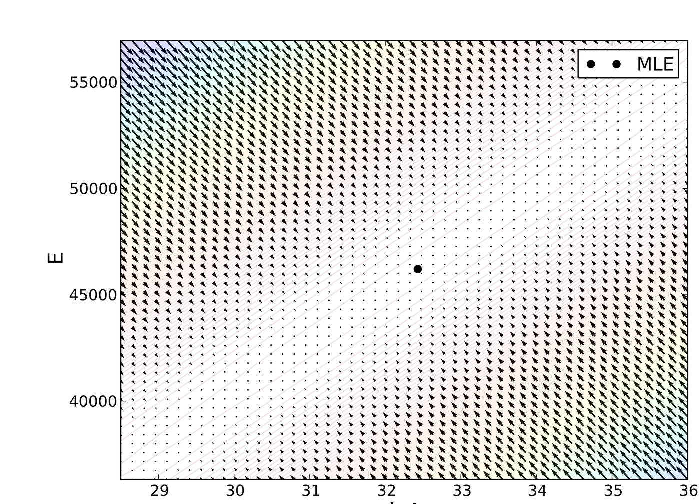
Uncertainty Quantification in Xolotl



- 2[nd] order Legendre-Uniform PC surrogate obtained with Bayesian regression from formation energies computed with MD
- Input to Xolotl, which computes cluster dissociation rates in plasma-surface interactions

Inference of Combustion Model Parameters





- 5[th] order Legendre-Uniform PC surrogate for an ignition time model, as a function of activation energy and pre-exponential (top left)
- Derivative of the surrogate (bottom left)
- Both the surrogate and its derivative obtained with UQTk
- Used in optimization to get better initial guess for MCMC
- Used to accelerate likelihood computation in MCMC