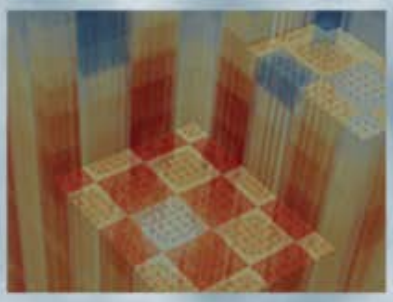
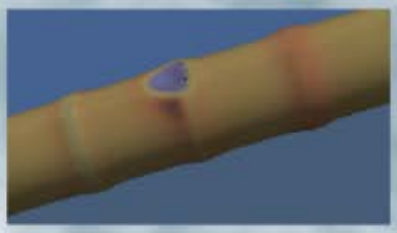
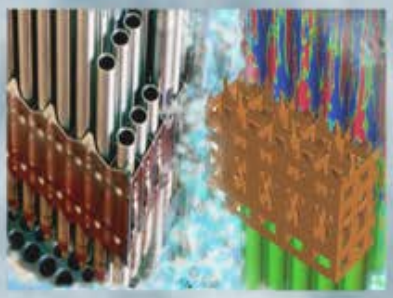
CASL-U-2015-0101-000

Sensitivity and Uncertainty Methods
(SAND2013-5431C)







Brian Adams
Sandia National Laboratory

July 8-10, 2013

U.S. DEPARTMENT OF ENERGY | Nuclear Energy

CASL-U-2015-0101-000

SAND2013-5431C
based primarily on
SAND2012-7387P, SAND2012-7388P

Sensitivity and Uncertainty Methods

Brian M. Adams
Optimization and Uncertainty Quantification, SNL
Validation and Uncertainty Quantification, CASL

Consortium for Advanced Simulation of Light-Water Reactors (CASL)
Education Council
Summer Student Workshop 2013

July 8—10, 2013

Oakridge, TN

CASL-U-2015-0101-000

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.

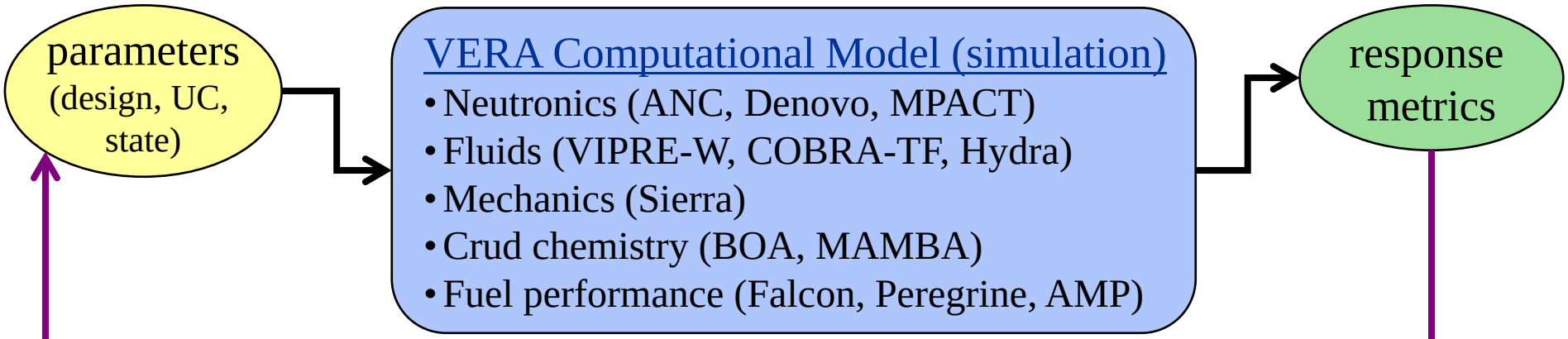
Learning Goals:
Sensitivity and Uncertainty Methods

- **Identify representative parametric inputs and response metrics for computational models in your area of expertise**
- **Understand purpose of sensitivity analysis (SA) and uncertainty quantification (UQ)**
- **Differentiate a few key classes of algorithms to perform parametric SA/UQ**
- **Know how to get started with CASL's VERA/Dakota software for SA/UQ**

CASL-U-2015-0101-000

Brainstorm: Nuclear Engineering Computational Models

- **What physical phenomena do you simulate or observe?**
- **On what input data or parameters does your simulation or experiment depend? Are they crisply defined?**
- **What output response metrics do you typically examine?**



- **Our focus: parametric studies that systematically vary code inputs to assess effects on code outputs**

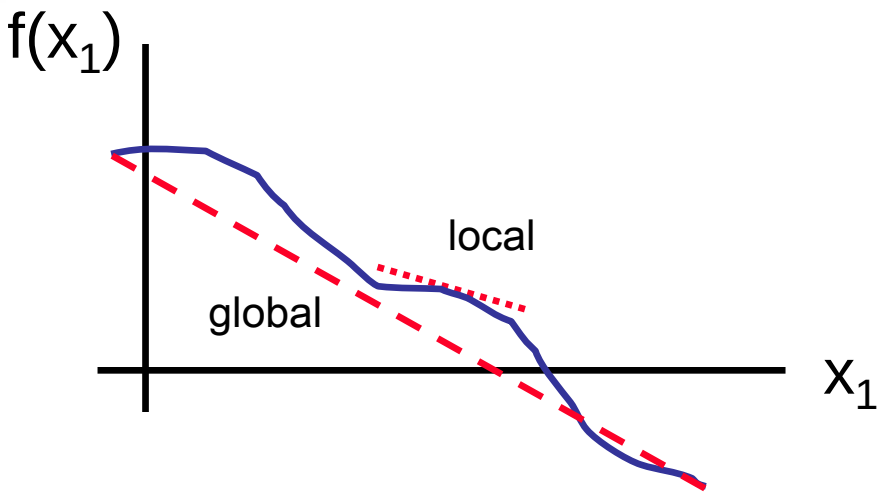
CASL-U-2015-0101-000
Sandia National Laboratories

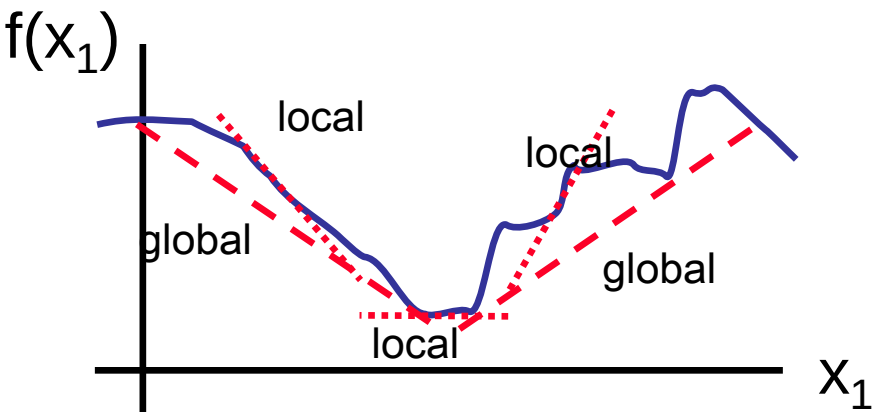
Why Perform Sensitivity Analysis?

- What? Understand code output variations as input factors vary
- Why? Identify most important variables and their interactions
  - Ranking/screening: Identity the most important variables, down-select for further UQ or optimization analysis
  - Provide a focus for resources
    - Data gathering and model development
    - Code development
    - Uncertainty characterization
  - Identify key model characteristics: smoothness, nonlinear trends, robustness
- Secondarily:
  - Identify code and model issues
  - Reuse designs to construct surrogate models

CASL-U-2015-0101-000
Sandia National Laboratories

Sensitivity Analysis: Influence of Inputs on Outputs





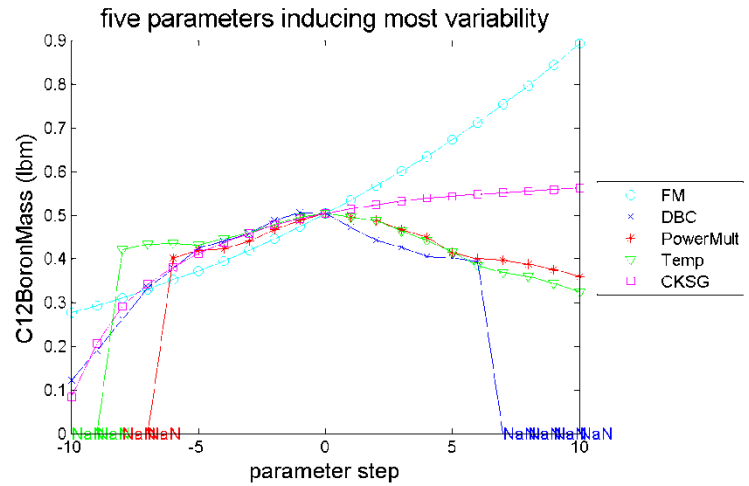
Assess variations in f(x1) due to (small or large) perturbations in x1.

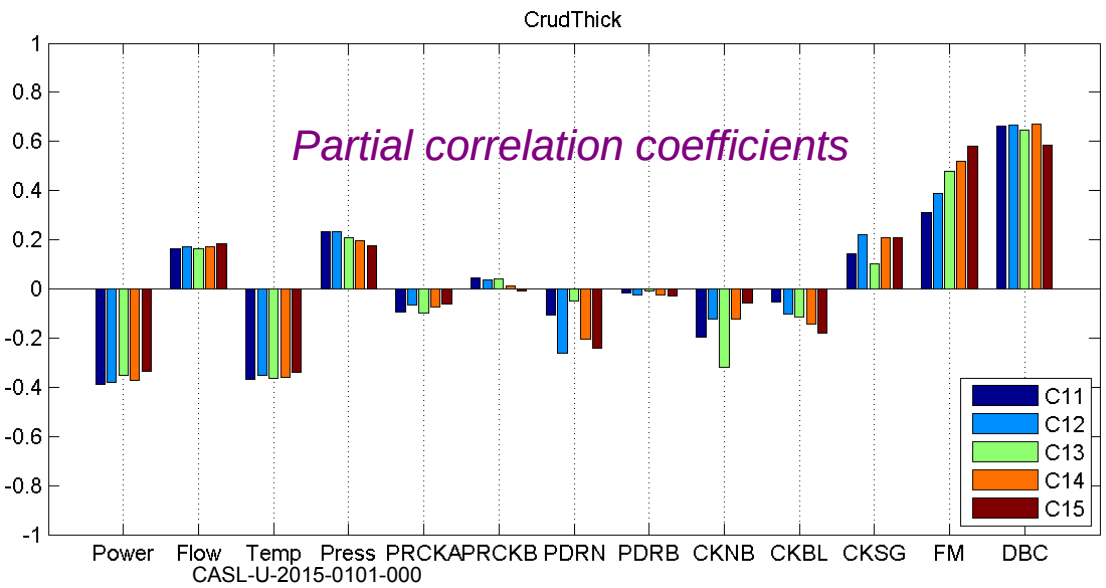
- **Local sensitivities**
  - Partial derivatives (first- or higher-order) at a specific point in input space.
  - Typically, given a specific x1, what is the slope at that point?
  - Determine via calculus, adjoints, automatic differentiation, finite differences
- **Global sensitivities**
  - What is the general trend of the function over all values of x1?
  - Found via sampling and regression.
  - Typically consider inputs uniformly over their whole range

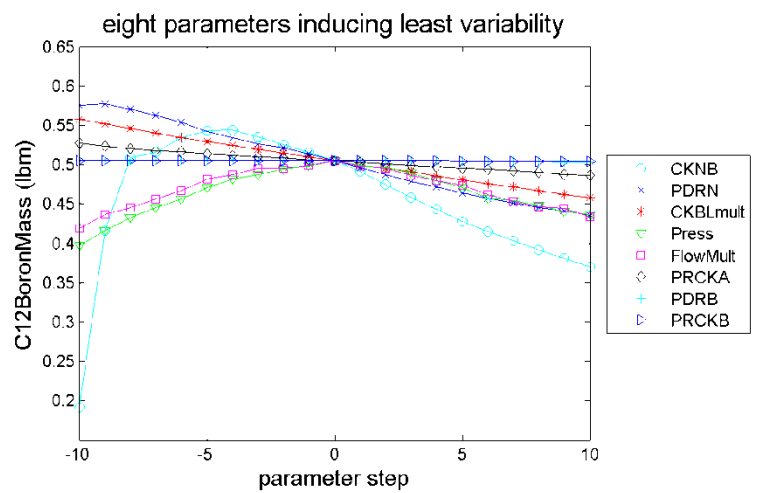
Example: Sensitivity Analysis for Crud Thickness

- CASL coupled VIPRE-W/BOA study
- Model predicts crud thickness, boron mass as a function of reactor operating conditions, heating, fluid flow and crud chemistry reaction rates
- Exposed code issues, parameter relations
- Found relative sensitivity of crud thickness to various parameters over five cycles

Univariate sensitivities

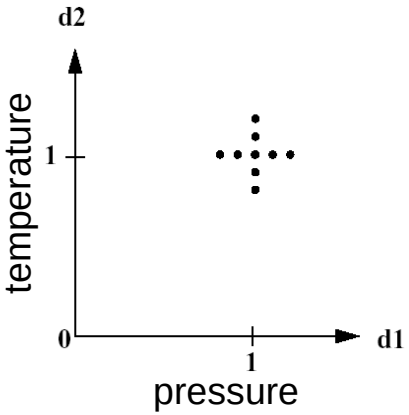


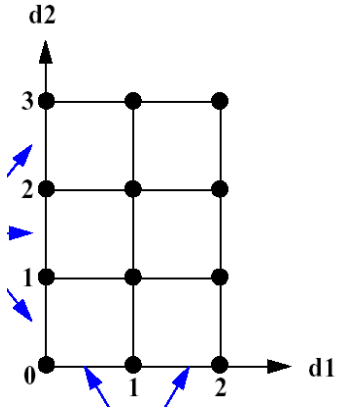


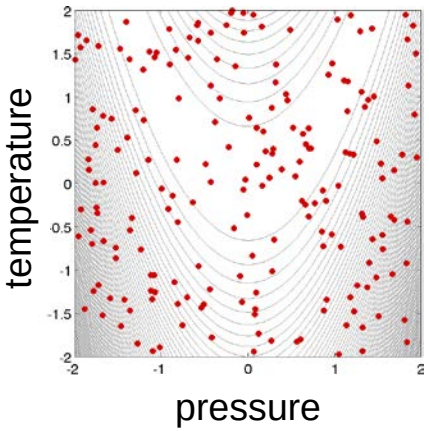


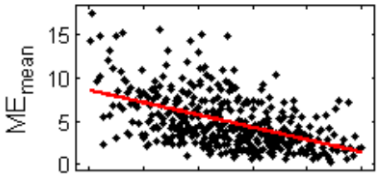
Selected Approaches to Sensitivity Analysis

- Local derivatives or sensitivities
- Univariate or joint parameter studies
- Global methods: sample designs spanning input space
  - Sampling: Monte Carlo, Latin hypercube, Quasi-MC, CVT
  - DOE/DACE: factorial, orthogonal arrays, Box-Behnken, CCD
  - Morris one-at-a-time
- Typical analysis results for responses
  - Simple and partial (including rank) correlation coefficients
  - Regression and resulting coefficients
  - Variance-based decomposition
  - Importance factors
  - Scatter plots

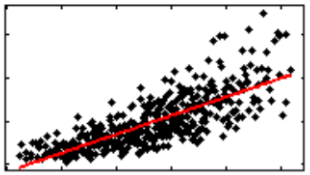




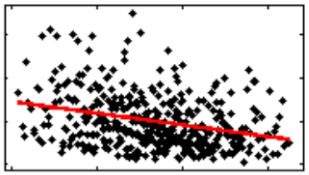




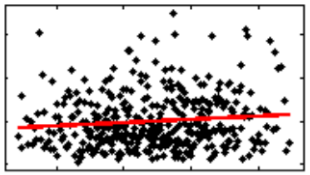
pressure



temperature



flow



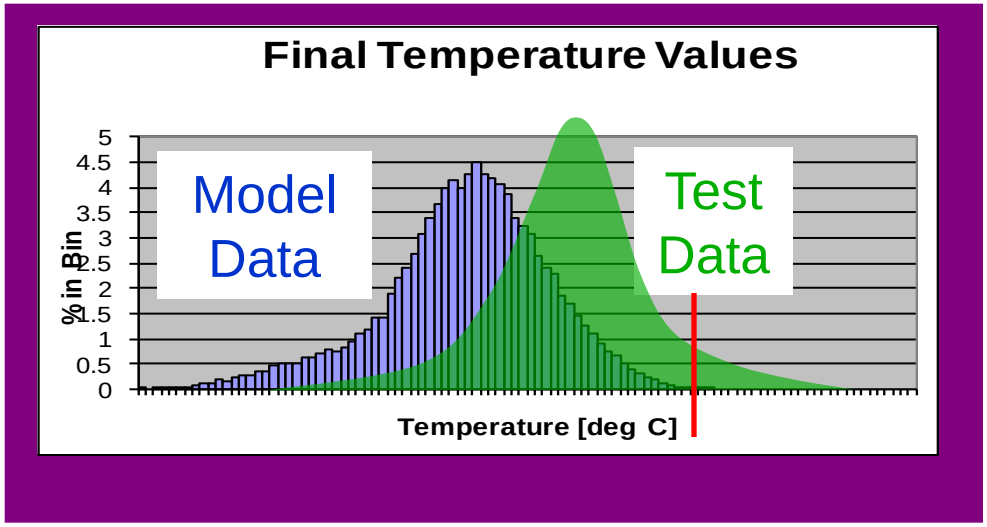
power

CASL-U-2015-0101-000
Sandia National Laboratories

Uncertainties in Simulation and Validation

A *few* uncertainties affecting computational model output/results:

- **physics/science parameters**
- **statistical variation, inherent randomness**
- **model form / accuracy**
- **material properties**
- **manufacturing quality**
- **operating environment, interference**
- **initial, boundary conditions; forcing**
- **geometry / structure / connectivity**
- **experimental error (measurement error, measurement bias)**
- **numerical accuracy (mesh, solvers); approximation error**
- **human reliability, subjective judgment, linguistic imprecision**



Parameterized…

The effect of these on model outputs should be integral to an analyst's deliverable: *best estimate PLUS uncertainty!*

CASL-U-2015-0101-000
Sandia National Laboratories

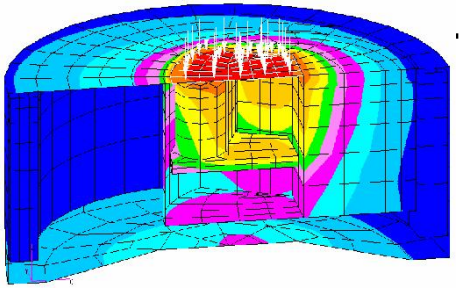
Why Perform Uncertainty Quantification?

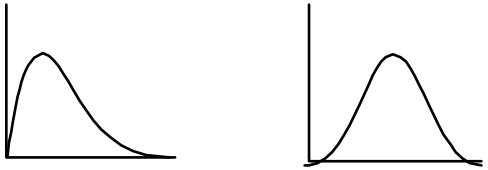
- **What? Determine variability, distributions, statistics of code outputs, given uncertainty in input factors**
- **Why? Assess likelihood of typical or extreme outcomes. Given input uncertainty…**
  - **Determine mean or median performance of a system**
  - **Assess variability in model response**
  - **Find probability of reaching failure/success criteria (reliability metrics)**
  - **Assess range/intervals of possible outcomes**
- **V&V, QMU: assess how close uncertainty-endowed code predictions are to**
  - **Experimental data (validation, is model sufficient *for the intended application?)***
  - **Performance expectations or limits (quantification of margins and uncertainties; QMU)**

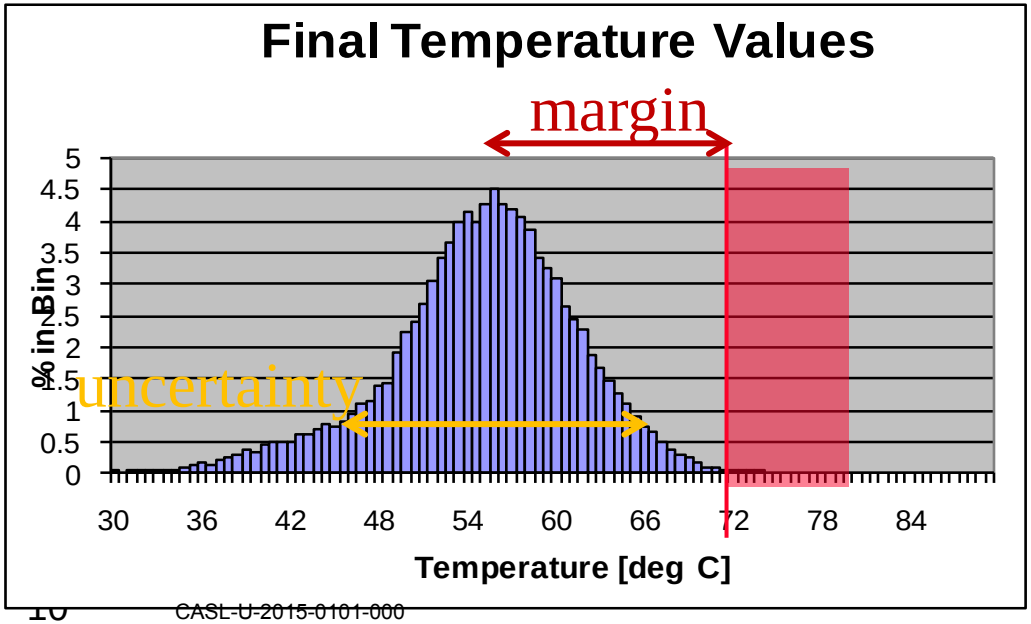
CASL-U-2015-0101-000
Sandia National Laboratories

Example: Thermal Uncertainty Quantification

- **Device subject to heating (experiment or computational simulation)**
- **Uncertainty in composition/ environment (thermal conductivity, density, boundary), parameterized by $u_1, \ldots, u_N$**
- **Response temperature $f(u)=T(u_1, \ldots, u_N)$ calculated by heat transfer code**







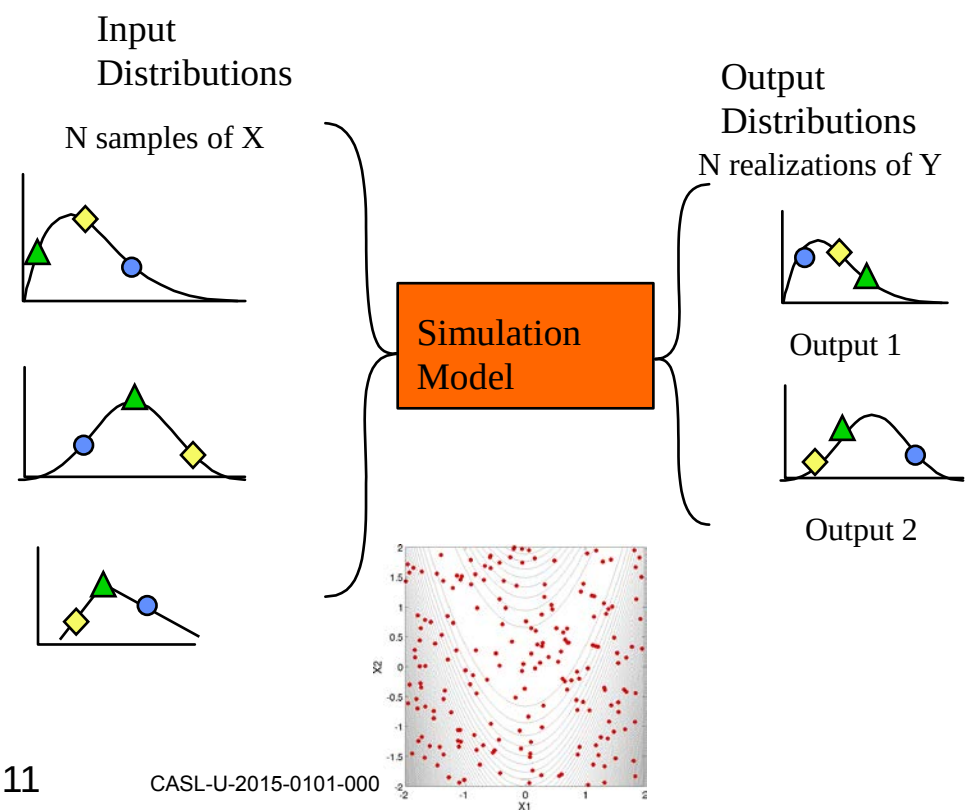
Given distributions of $u_1,\ldots,u_N$, UQ methods calculate statistical info on outputs:

- **Mean(T), StdDev(T), Probability($T \geq T_{critical}$)**
- **Probability distribution of temperatures**
- ***Correlations (trends) and sensitivity of temperature***

CASL-U-2015-0101-000
Sandia National Laboratories

Prevalent UQ Method: Random Sampling

- **Assume distributions on each of the *n* uncertain input variables**
- **Sample from each distribution and pair into *N* samples**
- **Run the simulation model for each of the *N* samples**
- **Use results ensemble to build up a distribution for each of the *m* outputs**

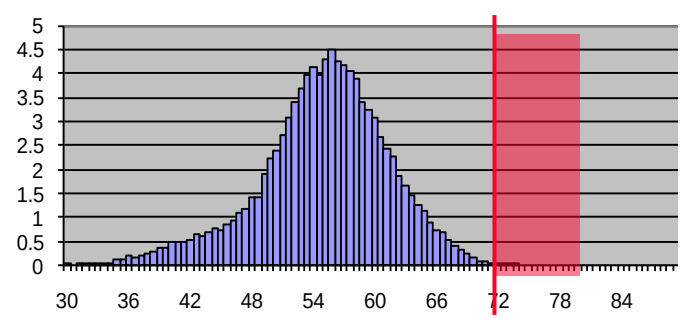


- **sample mean**

$$\bar{T} = \frac{1}{N}\sum_{i=1}^{N} T(u^i)$$

- **sample variance**

$$T_{\sigma^2} = \frac{1}{N}\sum_{i=1}^{N}\left[T(u^i) - \bar{T}\right]^2$$



CASL-U-2015-0101-000

Selected UQ Approaches

- **Sampling:** **robust, understandable, can require many samples (Monte Carlo, LHS, importance)**
- **Reliability methods:** **efficient optimization search to find behaviors or failure modes (mean value, MPP, FORM, SORM)**
- **Stochastic expansions:** **highly efficient on smooth responses (tailored surrogates: polynomial chaos, stochastic collocation)**
- **Other:** **interval estimation, evidence theory, mixed aleatory/ epistemic**

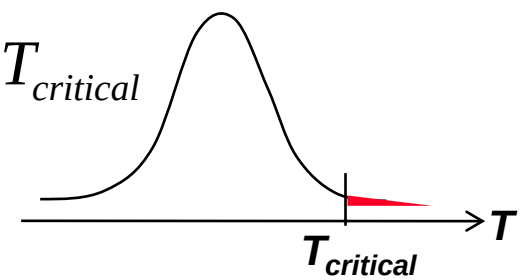
- ***All can be adaptive, surrogate-enhanced to be efficient on costly simulation models***

$$\mu_T = T(\mu_u)$$

$$\sigma_T = \sum_i \sum_j Cov_u(i,j) \frac{dg}{du_i}(\mu_u) \frac{dg}{du_j}(\mu_u)$$

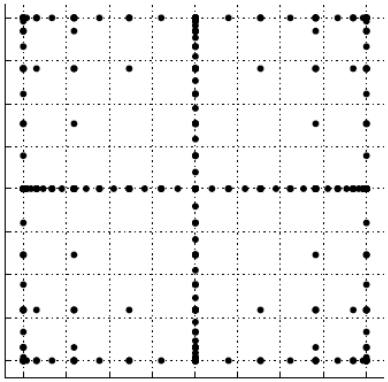
$$\text{minimize} \quad u^T u$$

$$\text{subject to} \quad T(u) = T_{critical}$$



$$R = \sum_{j=0}^{P} \alpha_j \Psi_j(\xi)$$

$R(\xi) \approx f(u)$

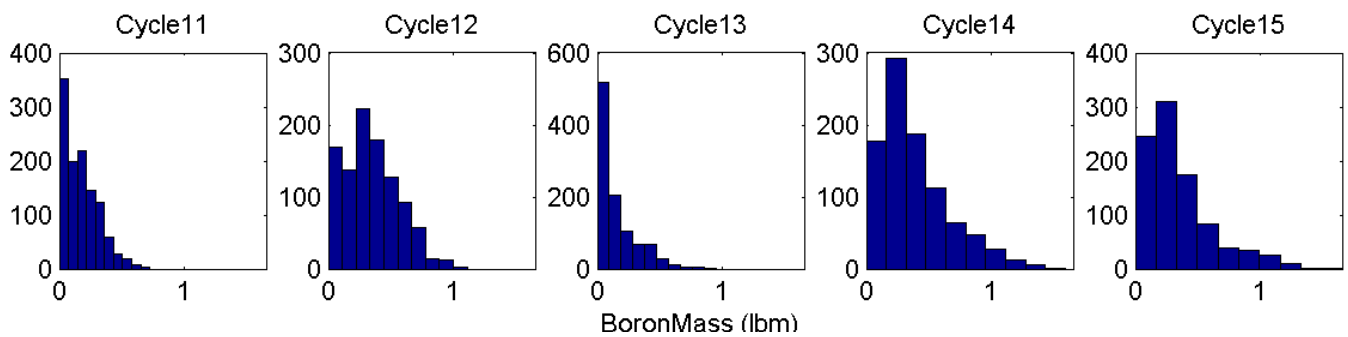


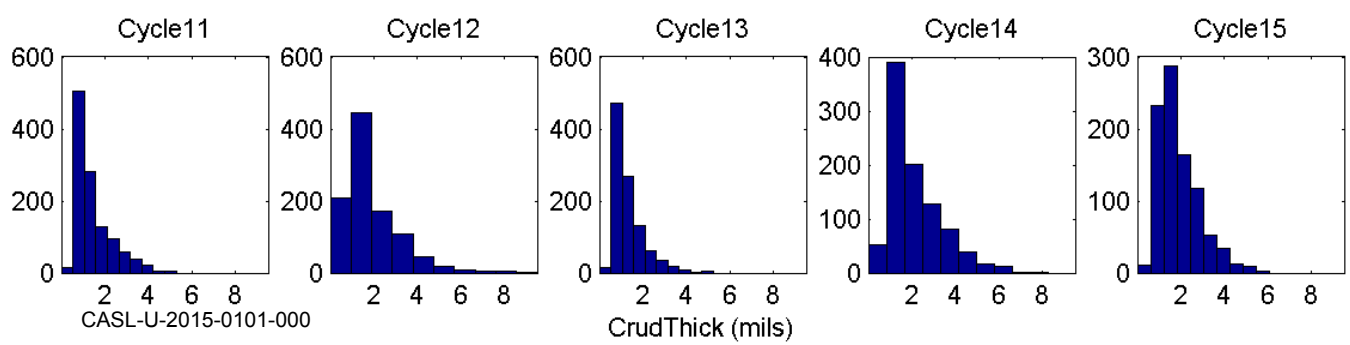
CASL-U-2015-0101-000
Sandia National Laboratories

Crud Metric Uncertainties Based on 1300 Samples

cycle 11 response	stats from samples		
	mean	std. dev.	coefficient var.
BoronMass (lbm)	0.170	0.141	0.83
CrudThick (mils)	**1.453**	**0.866**	**0.60**
NiFeRatio (1)	**1.524**	**0.298**	**0.20**
NiMetalMass (lbm)	3.798	0.908	0.24
MedianNi (ppb)	8.738	8.950	1.02
MedianFe (ppb)	6.810	5.033	0.74
CrudFraction (1)	0.087	0.148	1.71

- Note relative vs. absolute variability
- Discuss with industry and compare to field data (not really validation)
- Bold are roughly **compatible with BOA calibration data**

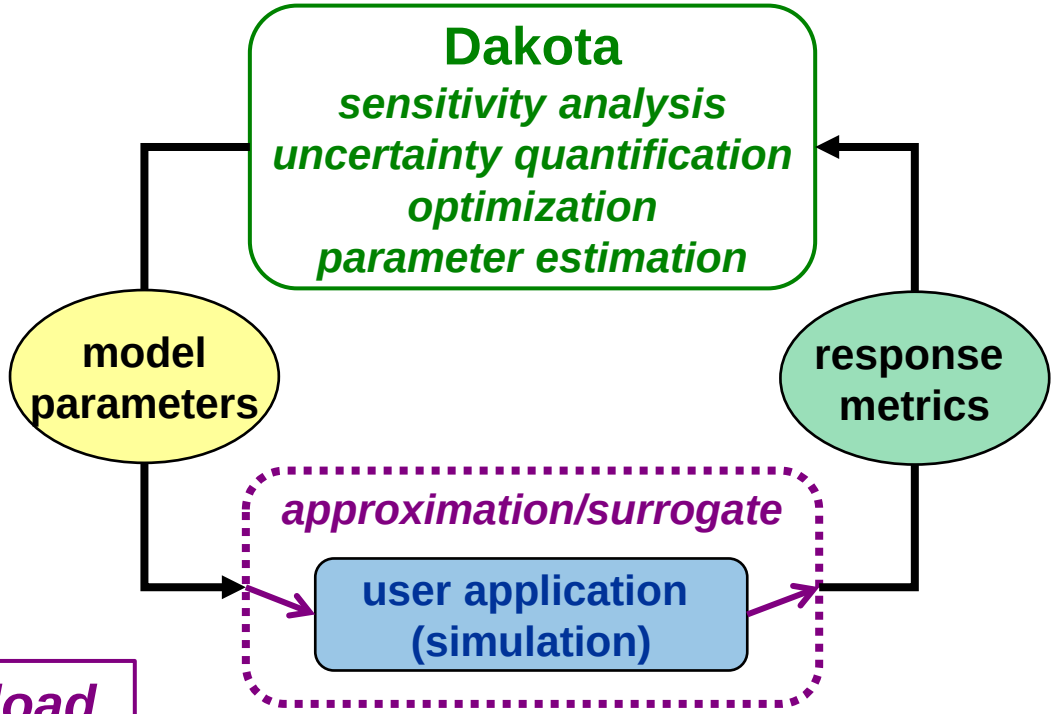




Dakota: Software Delivery Vehicle for CASL VUQ Algorithms

- Apply sensitivity analysis, UQ, optimization, and calibration to computational models
- Facilitate design trade-offs, V&V, QMU, risk-based decision making

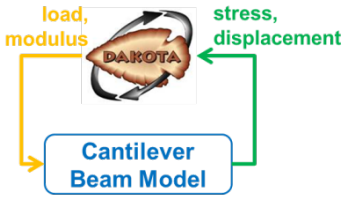
- Extensive algorithm toolbox
- Non-intrusive iterative systems analysis
- Parallel computing: desktop to HPC
- Simulation management



Included with VERA, or download from http://dakota.sandia.gov/

Steps to Using DAKOTA

1. Define analysis goals; understand how DAKOTA helps, learn about and select from possible methods
2. Access DAKOTA and understand help resources
3. **Workflow:** create an automated workflow so DAKOTA can communicate with your simulation
   - Parameters to model, responses from model to DAKOTA
   - Typically requires scripting (Python, Perl, Shell, Matlab) or programming (C, C++, Java, Fortran)
   - Workflow usually crosscuts DAKOTA analysis types



built-in cantilever beam analysis

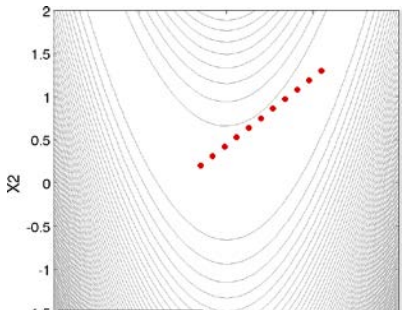
4. **DAKOTA input file:** Jaguar GUI or text editor to configure DAKOTA to exercise the workflow to meet your goals
   - Tailor variables, methods, responses to analysis goals
   - Syntax documented in Reference Manual
5. Run DAKOTA: command-line; text input / output

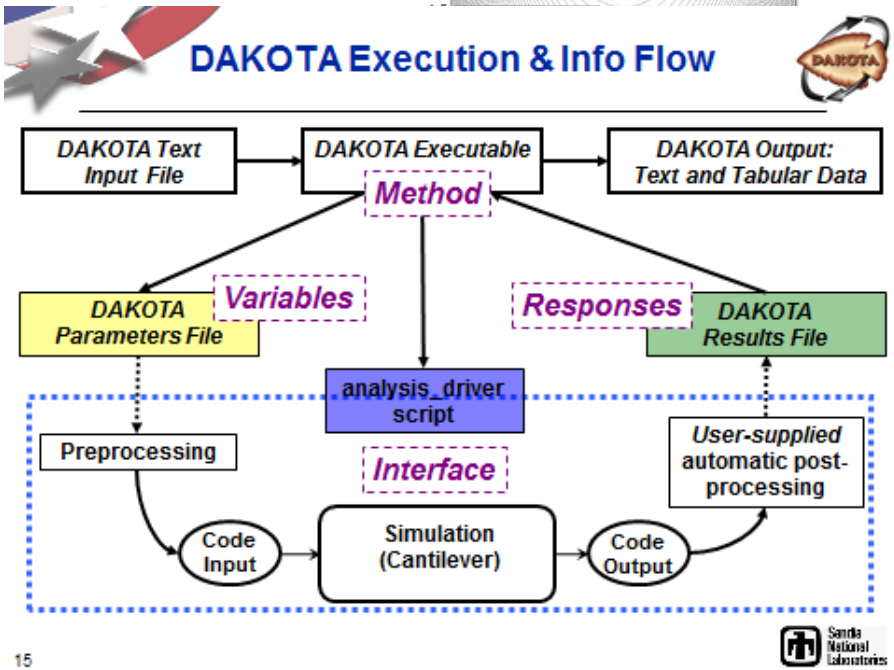
Configure DAKOTA Input File for a Vector Parameter Study

```
strategy
  single_method
  graphics, tabular_graphics_data

method
  vector_parameter_study
  num_steps = 10
  final_point  4.0 4.0 40000. 29.E+6 500. 1000.

variables
  continuous_design = 2
    initial_point  1.0          1.0
    descriptors    'beam_width' 'beam_thickness'
  continuous_state = 4
    initial_state  40000. 29.E+6 500. 1000.
    descriptors    'R'    'E'    'X'  'Y'

interface
  direct
  analysis_driver = 'mod_cantilever'

responses
  num_objective_functions = 3
  descriptors = 'area' 'stress' 'displacement'
  no_gradients
  no_hessians
```

(cantilever_vector.in)

Define Flow / Algorithm





Define Problem / Mapping

CASL-U-2015-0101-000

Mechanics of Running Dakota on a VERA Problem

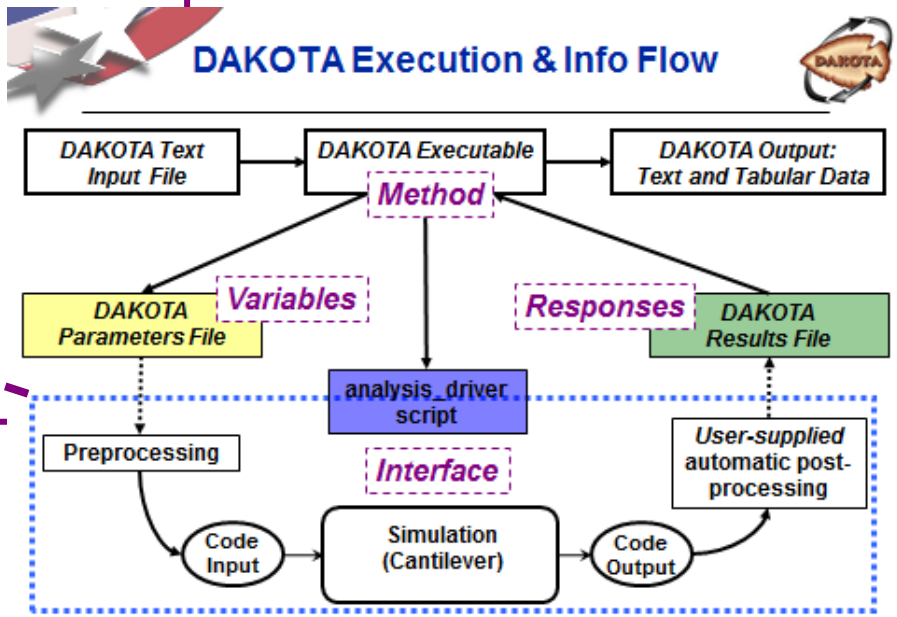
- **Dakota Input file**
- **Workflow components for a VERA simulation**
  - **parameter insertion to simulation input file**
  - **automated simulation execution**
  - **scripted post-processing (Python)**
- **Dakota output**

Extensive help resources at
http://dakota.sandia.gov

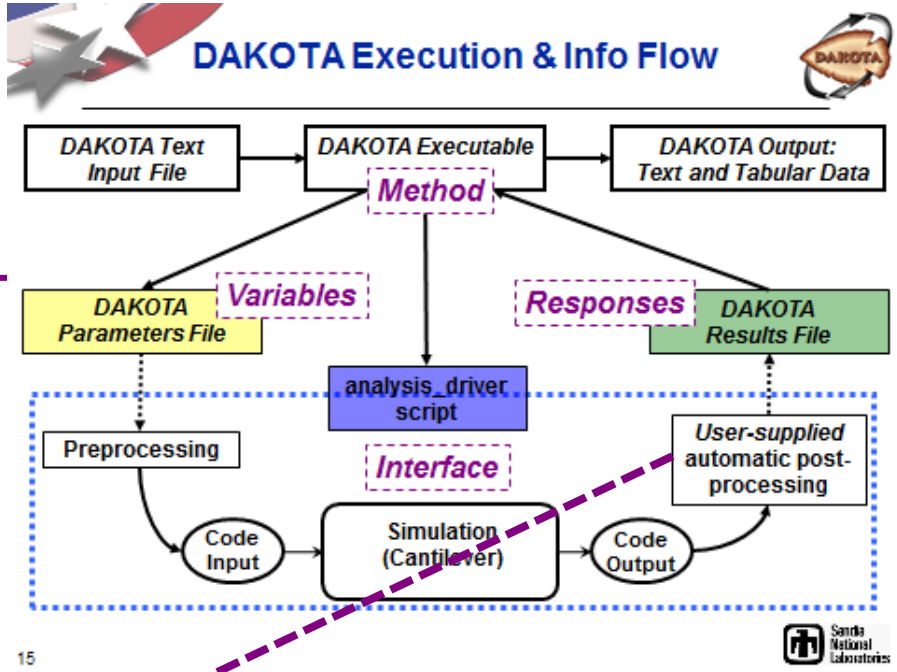
CASL-U-2015-0101-000
Sandia National Laboratories

Dakota/Cobra-TF Workflow: Example Driver Script

```
#!/bin/sh

#echo "Args are $@"

# $1 will be params.in from Dakota
# $2 will be results.out to Dakota

# Iterate over experimental configurations
#   61121 61451 62441 64561 66561
# Ultimately creating 3 responses for each (upper, middle, low)
# For a total of 15 responses
# Give option to turn them off

for d in 61121 61451 62441 64561 66561;
do
  echo "INFO: Running in $d"
  cd $d
  dprepro ../$1 deck.inp.template deck.inp
  ctf.x 1>ctf.log 2>&1
  ../../data_extract_oneexp.py >> ../$2
  cd ..
done
```

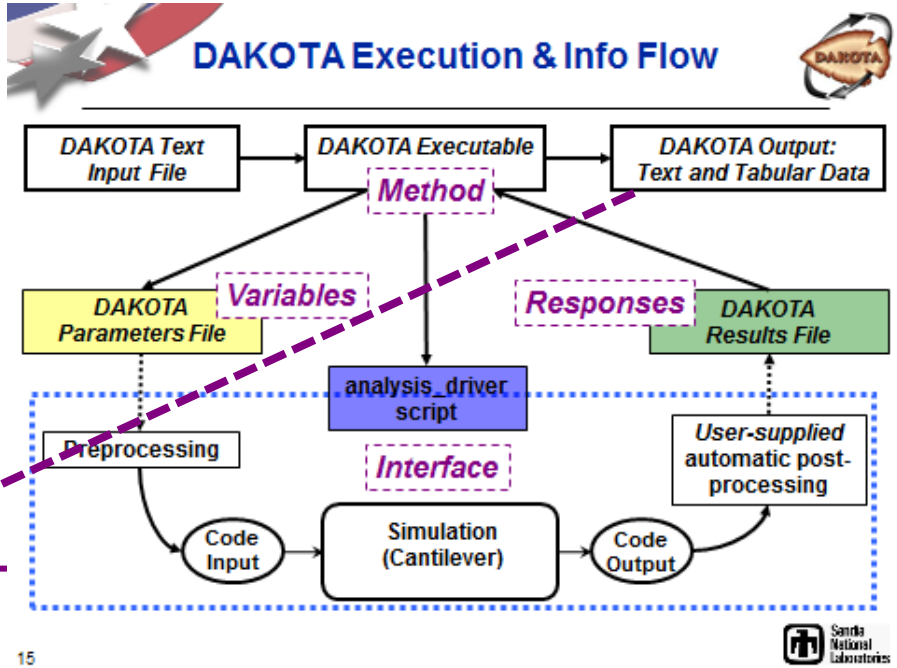


Dakota/Cobra-TF Workflow: Example Driver Script




```
#!/usr/bin/env python

# Data extraction for one experiment (taken from Bob's
data_extract.py)

# Open the results_channels.out file
#   output_file = open(path+'/results_channels.out','r')
output_file = open('results_channels.out','r')

...

# Upper Level (3.177 m)
splitline = data[6229].split()
ch22void.append(float(splitline[5]))

# Calculate the average void fractions
lowvoid = (ch15void[0]+ch16void[0]+ch21void[0]+ch22void[0])/4.0
midvoid = (ch15void[1]+ch16void[1]+ch21void[1]+ch22void[1])/4.0
uppvoid = (ch15void[2]+ch16void[2]+ch21void[2]+ch22void[2])/4.0

...
```

Dakota/Cobra-TF Workflow: Example Dakota Output




```
Statistics based on 59 samples:

Moment-based statistics for each response function:
                            Mean             Std Dev          Skewness          Kurtosis
        61121L  6.0000000000e-03  2.5120399733e-03  8.9165060021e-01 -6.3493455488e-01
        61121M  7.6864406780e-02  8.4413955929e-03  1.0856962276e-01 -1.1906324408e+00
        61121U  1.4667796610e-01  9.8110435277e-03  2.2461556484e-02 -1.1726090950e+00
        61451L  8.0000000000e-03  5.2488422381e-18 -1.0262783019e+00 -2.0714285714e+00
        61451M  9.9084745763e-02  5.9777983942e-03  1.2030932736e-01 -9.9176117733e-01
        61451U  2.2735593220e-01  8.3392770543e-03  7.4271101465e-03 -1.1978798374e+00
...
Partial Correlation Matrix between input and output:
                   61121L       61121M       61121U       61451L       61451M       61451U
DeltaTemp    -9.06421e-01 -9.98804e-01 -9.99591e-01  0.00000e+00 -9.96488e-01 -9.99426e-01
```

Learning Goals Revisited:
Sensitivity and Uncertainty Methods

- **Can you identify representative parametric inputs, uncertainties, and response metrics for computational models in your area of expertise?**
- **Do you understand the principal use of sensitivity analysis (SA) versus uncertainty quantification (UQ)? How would they influence your decisions?**
- **Can you name a few algorithms for each of SA and UQ?**
- **Where can you acquire Dakota for SA and UQ?**
- **What are the steps to getting started with it?**

Thank you for your participation!

briadam@sandia.gov

http://dakota.sandia.gov

Sandia National Laboratories

Dakota Reference Materials

CASL-U-2015-0101-000

DAKOTA Supports

- **Efficient simulation-based engineering design: optimize virtual (computational) prototypes**
- **Risk analysis and quantification of margins and uncertainty (QMU): assess the effect of parametric uncertainty on the probability of achieving desired system performance**
- **Verification and validation: automate mesh convergence or solver tolerance studies, generate ensembles of possible simulations or statistics to compare to experimental data**

CASL-U-2015-0101-000

DAKOTA in a Nutshell

DAKOTA includes a wide array of algorithm capabilities to support engineering transformation through advanced modeling and simulation.

Adds value to simulation-based analysis by answering fundamental science and engineering questions:

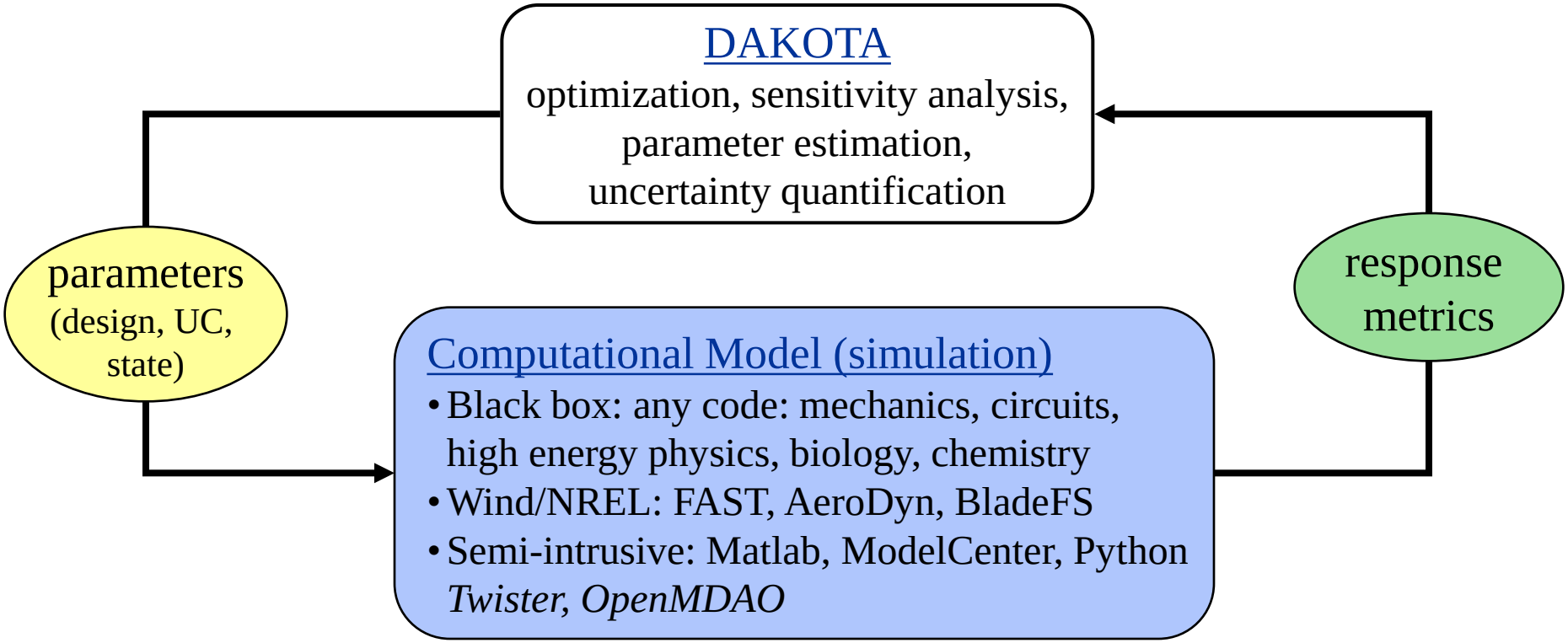
- **What are the crucial factors/parameters and how do they affect key metrics? *(sensitivity)***
- **How safe, reliable, robust, or variable is my system? *(quantification of margins and uncertainty: QMU, UQ)***
- **What is the best performing design or control? *(optimization)***
- **What models and parameters best match experimental data? *(calibration)***

- ***All rely on iterative analysis with a computational model for the phenomenon of interest***

CASL-U-2015-0101-000
Sandia National Laboratories

Automated Iterative Analysis of Computational Models

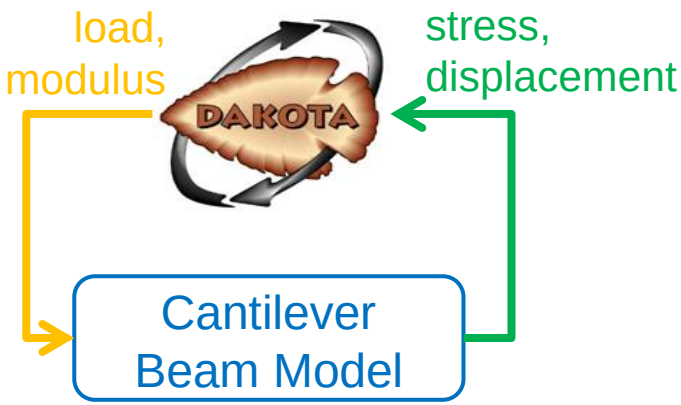
Automate typical "parameter variation" studies with various advanced methods and a generic interface to your simulation

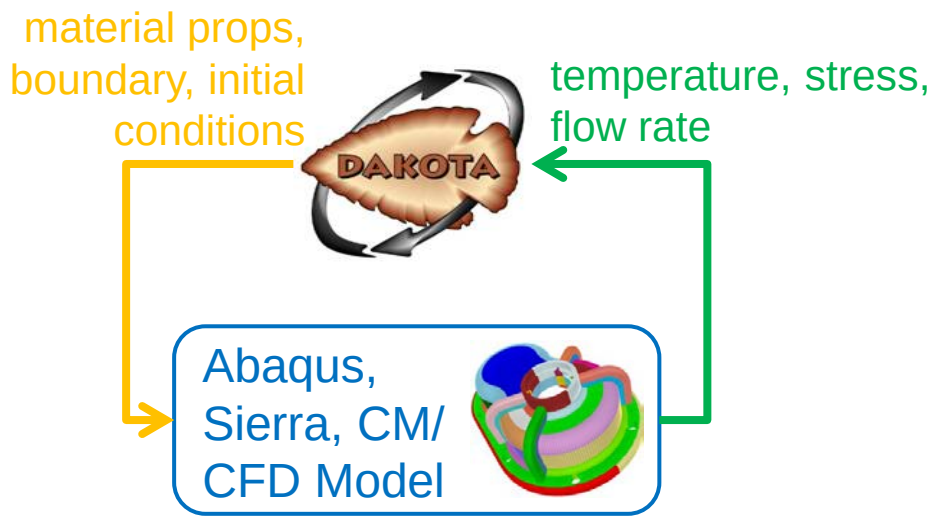


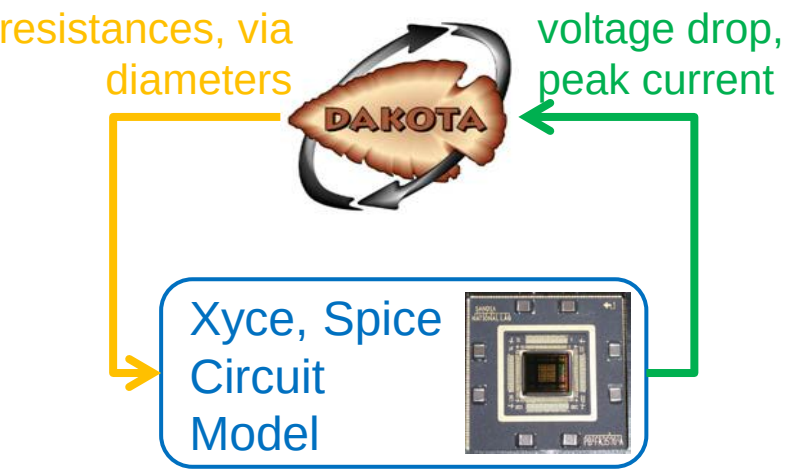
- Can support experimental testing: examine many accident conditions with computer models, then physically test a few worst-case conditions.

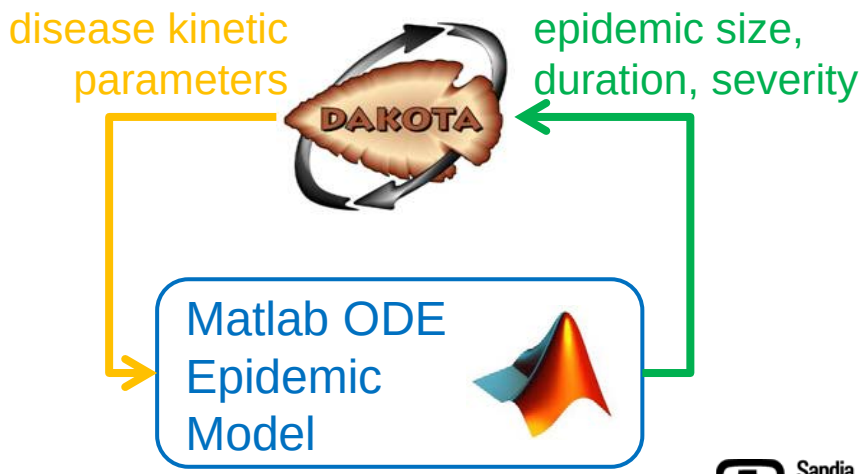
CASL-U-2015-0101-000
Sandia National Laboratories

DAKOTA Analysis: Iterating over Parameters of Computational Models









CASL-U-2015-0101-000
Sandia National Laboratories

Simulation Challenges DAKOTA Addresses

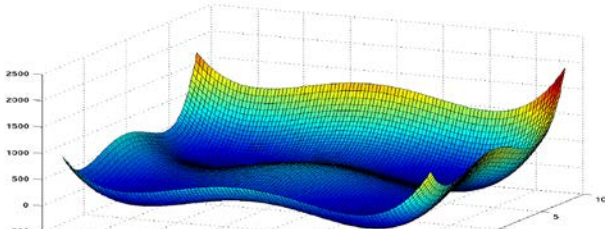
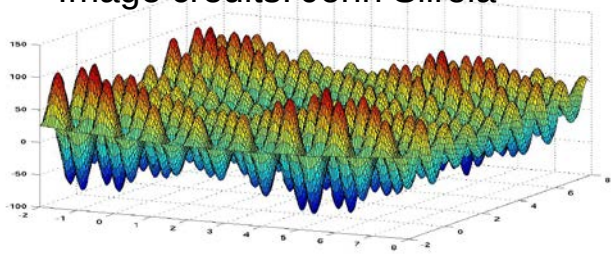
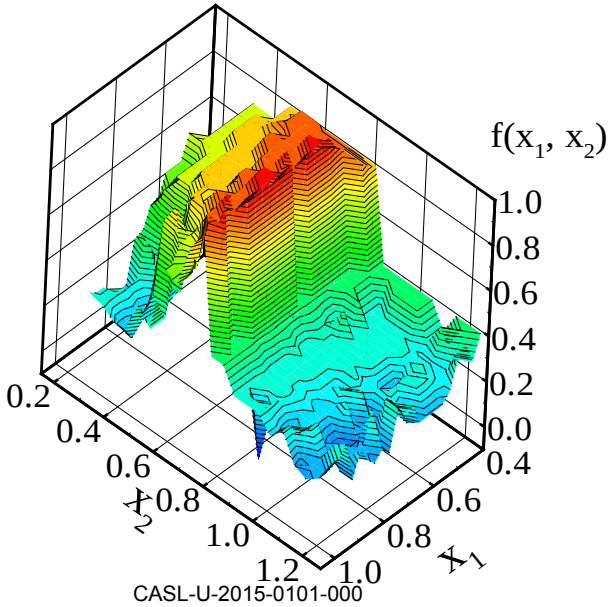


Image credits: John Siirola





In science and engineering problems of interest, we typically have:

- **no explicit function for $f(x_1,x_2)$**
  - **can't leverage algebraic structure**
- **limited number of evaluations/samples**
  - **expensive to evaluate $f(x_1,x_2)$ (long runtime even on many processors)**
  - **simulation may fail (hidden constraints)**
- **noisy / non-smooth**
  - **can't reliably estimate derivatives**
- **local extrema, non-convex**
  - **globally optimal solutions challenging**

Considerable research has been done to mitigate these issues.

Sandia National Laboratories

Key DAKOTA Capabilities
Largely Addressed in Advanced Topics

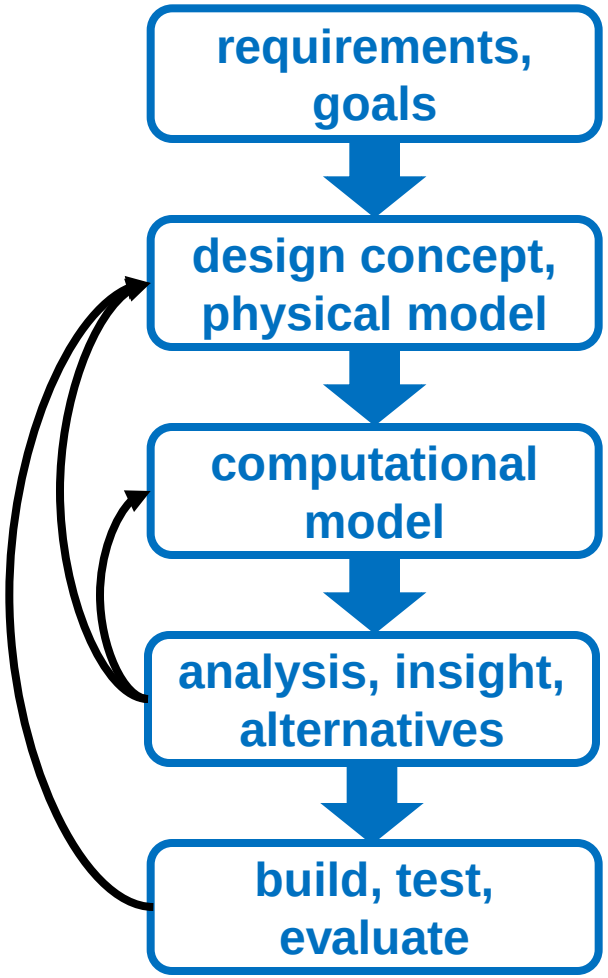
- **Generic interface to simulations: parameters in, responses out**
- **Time-tested and advanced research algorithms to address challenging science and engineering simulations**
- **Strategies to combine methods for advanced studies or improve efficiency with surrogates**
- **Mixed deterministic / probabilistic analysis**
- **Scalable parallel computations from desktop to clusters**
- **Primarily command-line and text-based interface; JAGUAR DAKOTA GUI and DART Workbench ease creating studies**
- **Object-oriented code; modern software quality practices**

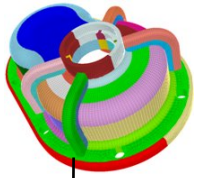
- **Additional details: http://dakota.sandia.gov/**
  - **Extensive documentation, including a tutorial**
  - **Support resources: http://dakota.sandia.gov/resources.html**
  - **Software downloads: stable releases and nightly builds (freely available worldwide via GNU LGPL)**

Sandia National Laboratories

DAKOTA in (Simplified) Simulation-based Design Process

Time check: 0:20





DAKOTA embedded in design refinement process to

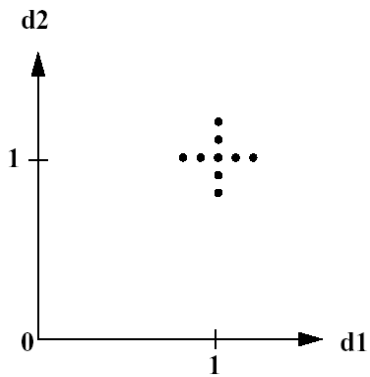
- **Tune model parameters to get desired behavior or match data**
- **Do simulation/design space exploration; explore alternative possibilities**
- **Validate designs/alternatives with their uncertainty**

Parameterized by mesh/geometry, materials, equations

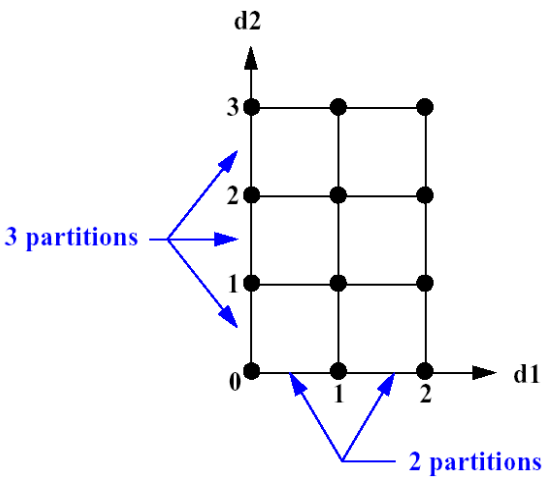
CASL-U-2015-0101-000

Sandia National Laboratories

Basic SA in DAKOTA: Parameter Studies

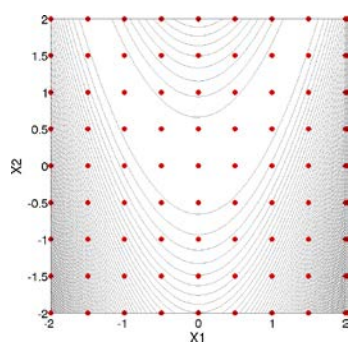


- Start at nominal values, perturb up and down
- Together: perform a DAKOTA centered parameter study on cantilever beam problem
- Convey to DAKOTA the parameter variations and which responses to study
- Example DAKOTA screen output and tabular file



- Group exercise: what would you change in the DAKOTA input to instead perform the grid parameter study at left?
- Use DAKOTA Reference Manual and/or JAGUAR
- What do you see as benefits/drawbacks of these methods?

Example: uniform grid over [-2.0, 2.0]



Time check: 0:20

CASL-U-2015-0101-000

DAKOTA Sensitivity Analysis Summary

- **What? Understand code output variations as input factors vary; main effects and key parameter interactions.**
- **Why? Identify most important variables and their interactions**
- **How? What DAKOTA methods are relevant? What results?**

Category	DAKOTA method names	univariate trends	correlations	modified mean, s.d.	main effects Sobol inds.	importance factors / local sensis
Parameter studies	`centered, vector, list`	P				
	`grid`		D		P	
Sampling	**`sampling`**, `dace lhs, dace random, fsu_quasi_mc, fsu_cvt with variance_based_decomp...`	P	D		D	
DACE (DOE-like)	`dace {oas, oa_lhs, box_behnken, central_composite}`		D		D	
MOAT	`psuade_moat`			D		
PCE, SC	`polynomial_chaos, stoch_collocation`				D	D
Mean value	`local_reliability`					D

multi-purpose!

D: DAKOTA
P: Post-processing (3rd party tools)

- **Also see DAKOTA Usage Guidelines in User's Manual**

CASL-U-2015-0101-000
Sandia National Laboratories

SA References

- Saltelli A., Ratto M., Andres T., Campolongo, F., et al., Global Sensitivity Analysis: The Primer, Wiley, 2008.
- J. C. Helton and F. J. Davis. Sampling-based methods for uncertainty and sensitivity analysis. Technical Report SAND99-2240, Sandia National Laboratories, Albuquerque, NM, 2000.
- Sacks, J., Welch, W.J., Mitchell, T.J., and Wynn, H.P. Design and analysis of computer experiments. Statistical Science 1989; 4:409–435.
- Oakley, J. and O'Hagan, A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. J Royal Stat Soc B 2004; 66:751–769.

- **DAKOTA User's Manual**
  - Parameter Study Capabilities
  - Design of Experiments Capabilities/Sensitivity Analysis
  - Uncertainty Quantification Capabilities (for MC/LHS sampling)
- Corresponding Reference Manual sections

DAKOTA UQ Summary and Relevant Methods

- **What? Understand code output uncertainty / variability**
- **Why? Risk-informed decisions with variability, possible outcomes**
- **How? What DAKOTA methods are relevant?**

character	method class	problem character	variants
aleatory	probabilistic sampling	nonsmooth, multimodal, modest cost, # variables	Monte Carlo, LHS, importance
	local reliability	smooth, unimodal, more variables, failure modes	mean value and MPP, FORM/SORM,
	global reliability	nonsmooth, multimodal, low dimensional	EGRA
	stochastic expansions	nonsmooth, multimodal, low dimension	polynomial chaos, stochastic collocation
epistemic	interval estimation	simple intervals	global/local optim, sampling
	evidence theory	belief structures	global/local evidence
both	nested UQ	mixed aleatory / epistemic	nested

- **See DAKOTA Usage Guidelines in User's Manual**
- **Analyze tabular output with third-party statistics package**

CASL-U-2015-0072-000

UQ References

- SAND report 2009-3055. “Conceptual and Computational Basis for the Quantification of Margins and Uncertainty” J. Helton.
  - Helton, JC, JD Johnson, CJ Sallaberry, and CB Storlie. “Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis”, *Reliability Engineering and System Safety* 91 (2006) pp. 1175-1209
  - Helton JC, Davis FJ. Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. *Reliability Engineering and System Safety* 2003;81(1):23-69.
  - Haldar, A. and S. Mahadevan. *Probability, Reliability, and Statistical Methods in Engineering Design* (Chapters 7-8). Wiley, 2000.
- Eldred, M.S., "Recent Advances in Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Analysis and Design," paper AIAA-2009-2274 in Proceedings of the 11th AIAA Non-Deterministic Approaches Conference, Palm Springs, CA, May 4-7, 2009.

  - **DAKOTA User’s Manual:** Uncertainty Quantification Capabilities
  - DAKOTA Theory Manual
  - Corresponding Reference Manual sections

Sandia National Laboratories

Advanced Capabilities Topics

General features

- **Restart**
- **Evaluation cache**
- **Utilities in dakota_restart_util**
- **Tabular graphics data**
- **Failure capturing: abort, retry, recover, ignore**
- **Constraint specification: linear, nonlinear; equality, inequality**
- **Input/output scaling**
- **Matlab interface**

Approximation methods

- **Global data fit surrogate methods (polynomials, MARS, Kriging, etc.)**
- **Local surrogate methods (Taylor series, multipoint)**
- **Hierarchical: high/low fidelity models**
- **Corrections**

Strategies/Advanced approaches

- Nested models: OUU
- Multi-objective (Pareto) optimization
- Multistart; multi-level hybrid
- Surrogate-based optimization (variety of constraint handling approaches): trust region; EGO/EGRA
- Reliability-based design optimization
- Advanced UQ topics: polynomial chaos, second-order probability, Dempster-Shafer, surrogate-based UQ
- AMPL: for analytic problems / algebraic mappings

Parallel capabilities: message passing, asynchronous local, hybrid

- Asynchronous evaluations
- Dakota parallel, application serial
- Dakota serial, application parallel
- Multi-level parallel: concurrent iteration, concurrent function evaluations, concurrent analyses,
- multiprocessor simulations

Sandia National Laboratories