SAND2018-XXXX

SAND2018-8614C

Exceptional service in the national interest

Sandia National Laboratories

Dakota within the NEAMS Workbench

NEAMS IPL Workshop
Aug. 8-10, 2018

Laura Swiler

U.S. DEPARTMENT OF ENERGY

National Nuclear Security Administration

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

What is Dakota?

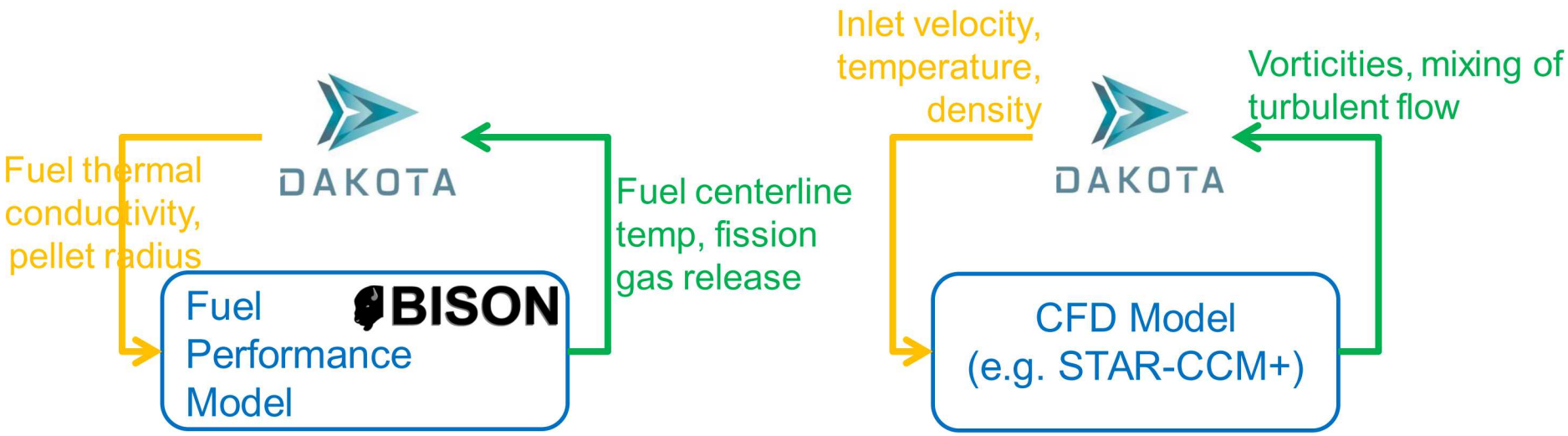
Sandia National Laboratories

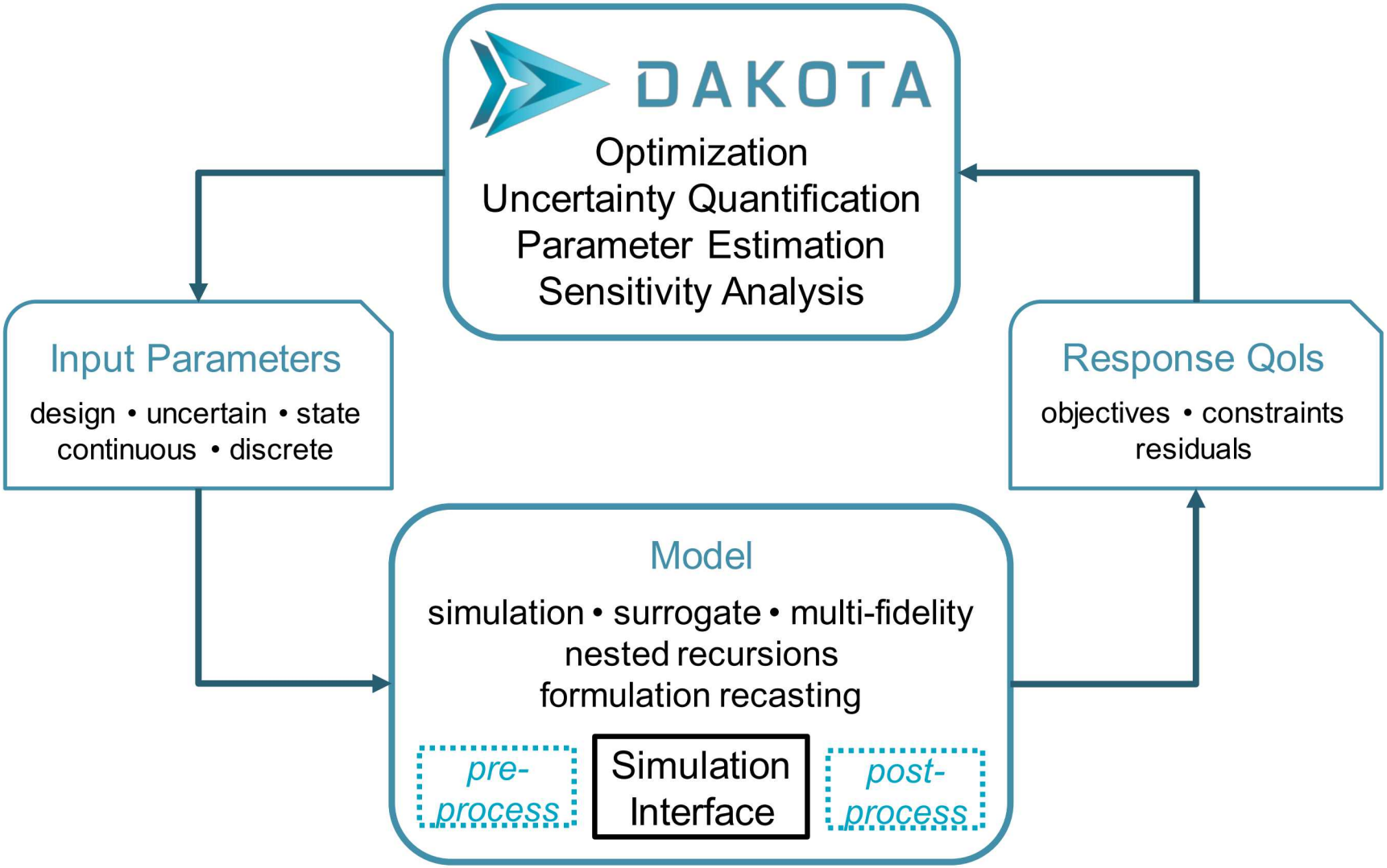
- A software framework developed at Sandia National Laboratories since 1995, primarily funded by the Advanced Simulation and Computing (ASC) program of NNSA.
- A set of algorithms that allows a user to "iterate" over their simulation for the purposes of uncertainty analysis, optimization, and calibration.
- Provides scientists and engineers (analysts, designers, decision makers) greater perspective on model predictions:
  - *Enhances understanding of risk by quantifying margins/uncertainties*
  - *Improves products through simulation-based design, calibration*
  - *Assesses simulation credibility through verification and validation*

…by analyzing ensembles

- Strategically selects model parameters
- Manages concurrent simulations
- Analyzes responses (model outputs)
- ➢ Automates one-pass parameter variation/ analysis to advanced goal-oriented studies

Run	Input	Output
1	0.814	91.3
2	0.906	63.24
…		
N	1.270	9.75





Available at https://dakota.sandia.gov

Sandia National Laboratories

WHY USE DAKOTA?

Dakota: Distinguishing Strengths

Sandia National Laboratories

- Makes sensitivity analysis, optimization, and uncertainty quantification practical for costly computational models
- Flexible interface to simulation codes: one interface; many methods
- Combined deterministic/probabilistic analysis
- Continual advanced algorithm R&D to tackle computational challenges (particularly in SNL's national security mission)
  - Treats non-smooth, discontinuous, multi-modal responses
  - Surrogate-based, multi-fidelity, and hybrid methods
  - Risk-informed decision-making: epistemic and mixed UQ, rare events, Bayesian
- Scalable parallel computing from desktop to HPC

Many Methods in One Tool

Sandia National Laboratories

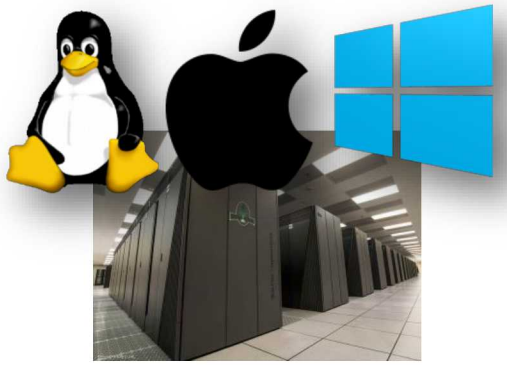
Sensitivity Analysis	Uncertainty Quantification
• Designs: MC/LHS, DACE, sparse grid, one-at-a-time • Analysis: correlations, scatter, Morris effects, Sobol indices	• MC/LHS/Adaptive Sampling • Reliability • Stochastic expansions • Epistemic methods
Optimization	**Calibration**
• Gradient-based local • Derivative-free local • Global/heuristics • Surrogate-based	• Tailored gradient-based • Use any optimizer • Bayesian inference

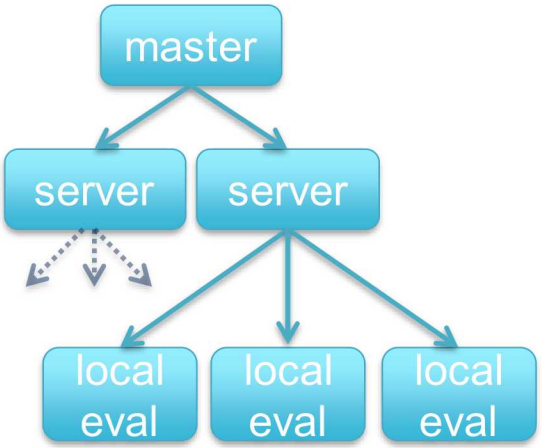
Interface Dakota to your simulation once, then apply various algorithms depending on your goal…

Computing and Parallelism

Sandia National Laboratories

- Runs in various computing environments
  - Desktop: Mac, Linux, Windows
  - HPC: Linux clusters, IBM Blue Gene/P and /Q, IBM AIX, including many DOE machines
  - Distributed workstation computing
- Exploits concurrency at multiple levels
  - Multiprocessor simulations
  - Multiple simulations per response
  - Samples in a parameter study
  - Optimizations from multiple starting points
- File management features, including
  - Work directories to partition analysis files
  - Template directories share files common among analyses



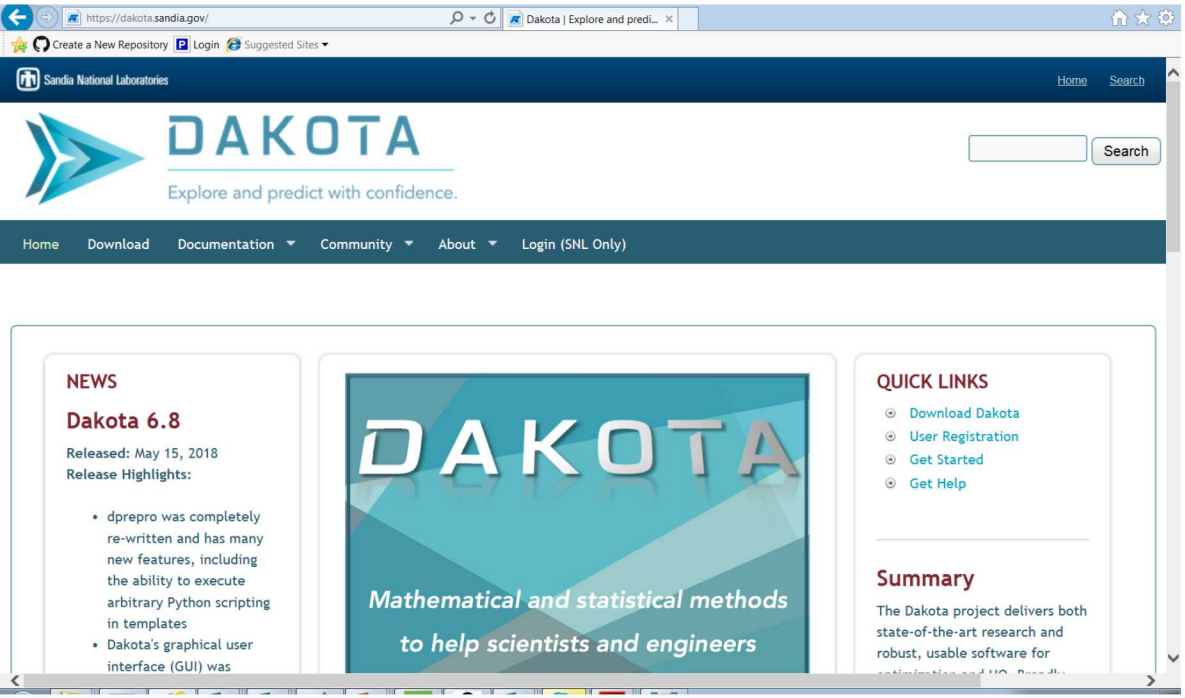


Dakota History and Resources

- Genesis: 1994 optimization LDRD
- Modern software quality and development practices
- Released every May 15 and Nov 15
- Established support process for SNL, partners, and beyond

http://dakota.sandia.gov

Mike Eldred, Founder



Lab mission-driven algorithm R&D deployed in production software

- Extensive website: documentation, training materials, downloads
- Open source facilitates external collaboration; widely downloaded

Getting Started and Getting Help

Sandia National Laboratories

Tour http://dakota.sandia.gov at a high level

- Getting Started
  - Download (LGPL license, freely available worldwide): http://dakota.sandia.gov/download.html
  - Getting Started: http://dakota.sandia.gov/quickstart.html
  - User's Manual, Chapter 2: Tutorial with example input files https://dakota.sandia.gov/sites/default/files/docs/6.3/Users-6.3.0.pdf
- Getting Help
  - Extensive documentation (user, reference, developer): http://dakota.sandia.gov/content/manuals
  - Support mailing list (reaches Dakota team and user community): dakota-users@software.sandia.gov

SNL Center for Computing Research and its partners have broad and deep expertise in embedded and black-box optimization and uncertainty quantification

Sandia National Laboratories

Uncertainty Quantification

Uncertainty Quantification (UQ)

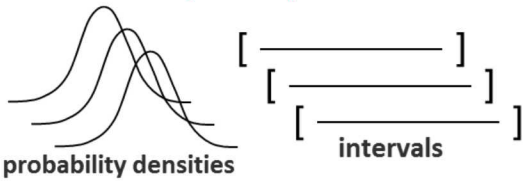
Sandia National Laboratories

Goal: Account for natural variability and state-of-knowledge uncertainties affecting computational models.

Why? Uncertainty-aware predictions support risk-informed decision making, model validation, and robust design.

Forward UQ: propagate uncertainties to put error bars on simulation predictions

Uncertainty in input variables *u*



physical parameters / properties, initial/boundary conditions, numerical accuracy, geometry, data

Simulation Model *s(u)*

Statistics or intervals on output *s(u)*

Observations: experimental or upstream model

Inference / Inverse UQ: characterize parameter uncertainty from observations

Uncertainty Quantification Methods

Sandia National Laboratories

- Monte Carlo sampling:
  - LHS, Quasi MC, Design of Experiments, Adaptive Sampling
- Local and global reliability
  - Employ optimization to efficiently determine failure probabilities
- Polynomial chaos expansions / stochastic collocation
  - Large research investment, adaptive sparse grids, compressive sensing
- Mixed aleatory-epistemic approaches
  - Regulatory Requirements, "uncertainty on uncertainties"
- Surrogate Models: Gaussian processes, PCE, splines, etc.

We focus on UQ methods that are as efficient and accurate as possible assuming simulations are very costly.

Dakota: State-of-the-Art UQ R&D

Sandia National Laboratories

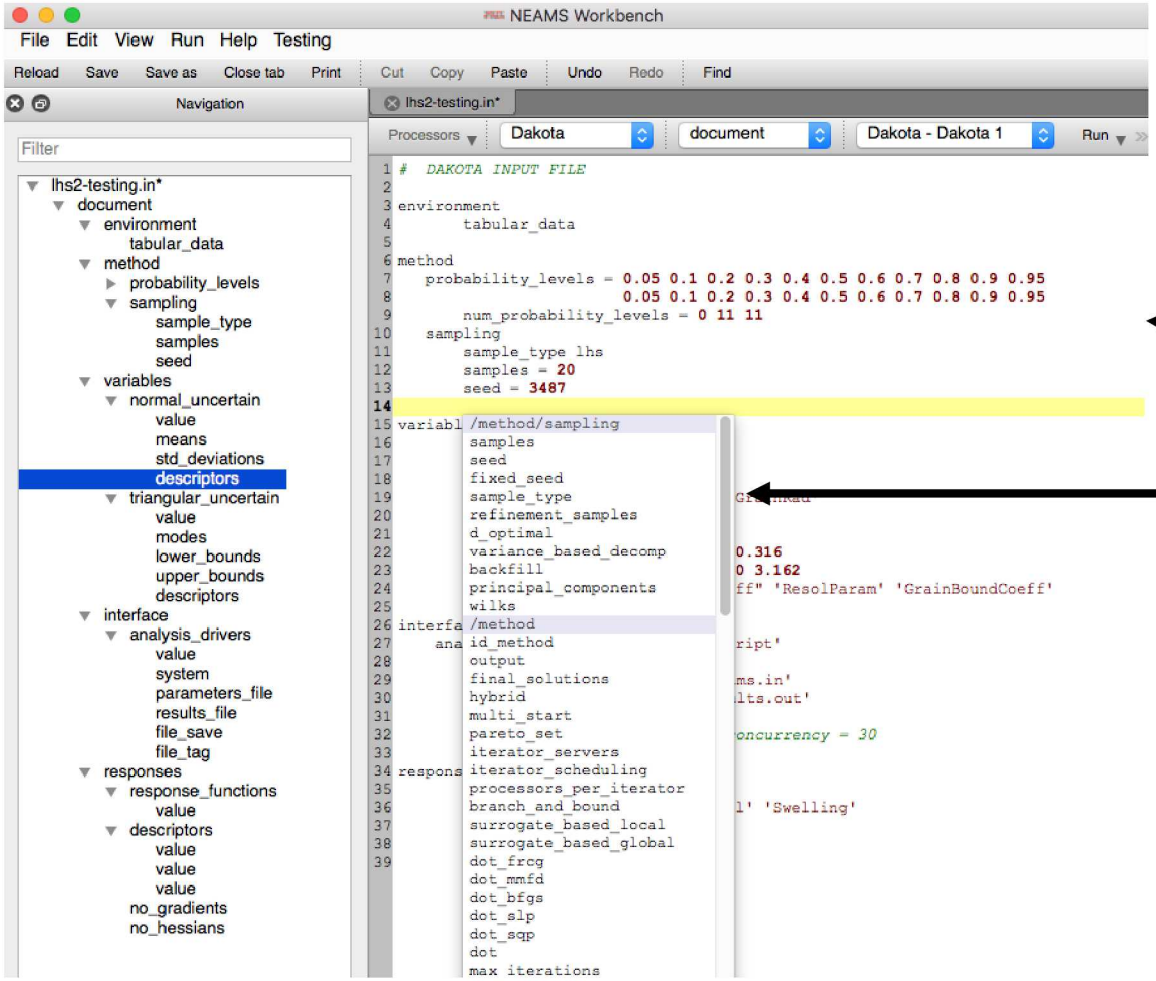
- Continual advanced algorithm R&D to tackle computational challenges (particularly in SNL's national security mission):
  - Severely constrained simulation budgets
  - High-dimensional parameter spaces
  - Nonsmooth or unreliable quantities of interest

- Notable active areas:
  - Multi-level, multi-fidelity UQ (and optimization)
  - Bayesian inference
  - Dimension reduction and surrogate modeling
  - Random field modeling
  - Discrepancy modeling

Sandia National Laboratories

Running Dakota from within NEAMS Workbench AND Dakota driving another code (BISON)

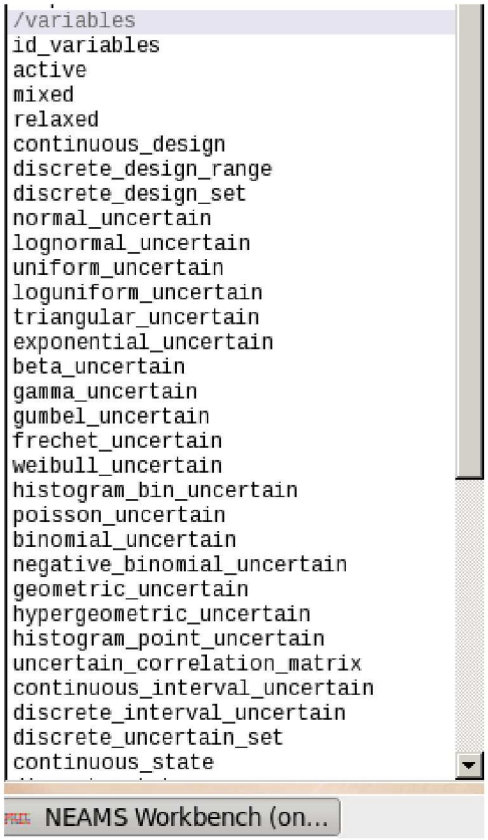
Dakota within the NEAMS Workbench



Dakota input file

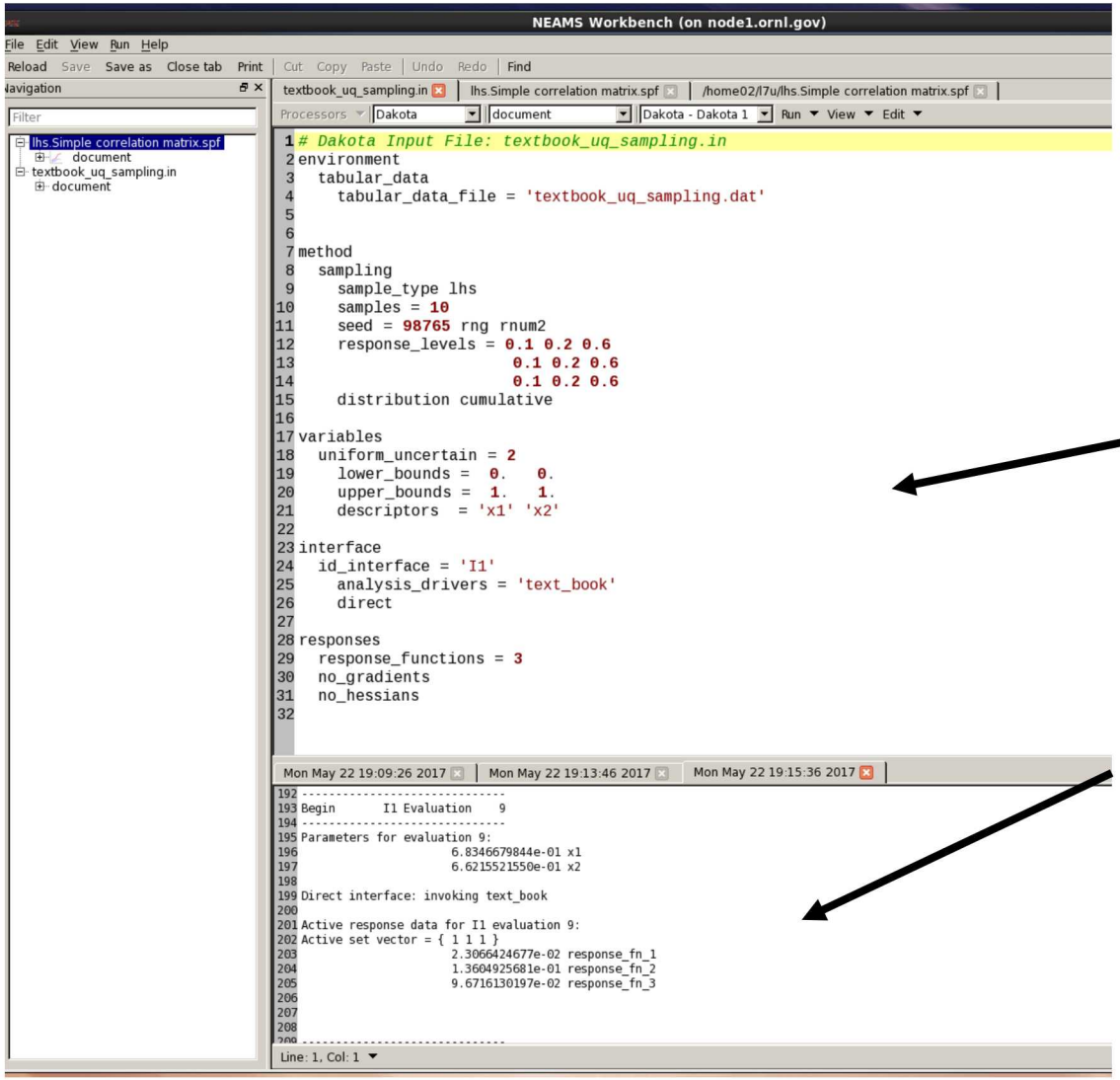
Autocompletion showing options

Dakota within the NEAMS Workbench



Pull down showing options for specifying variable types

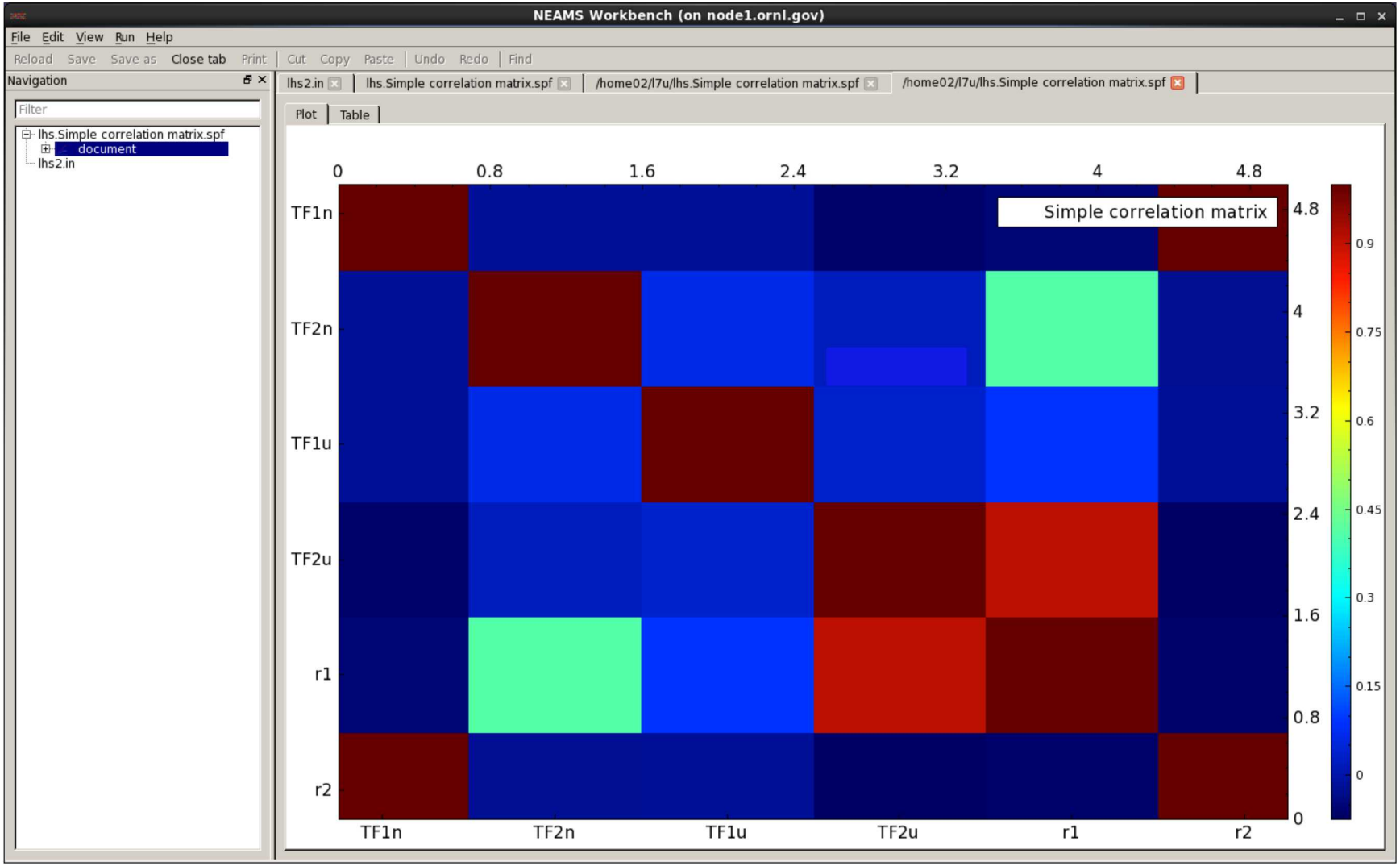
Dakota within the NEAMS Workbench

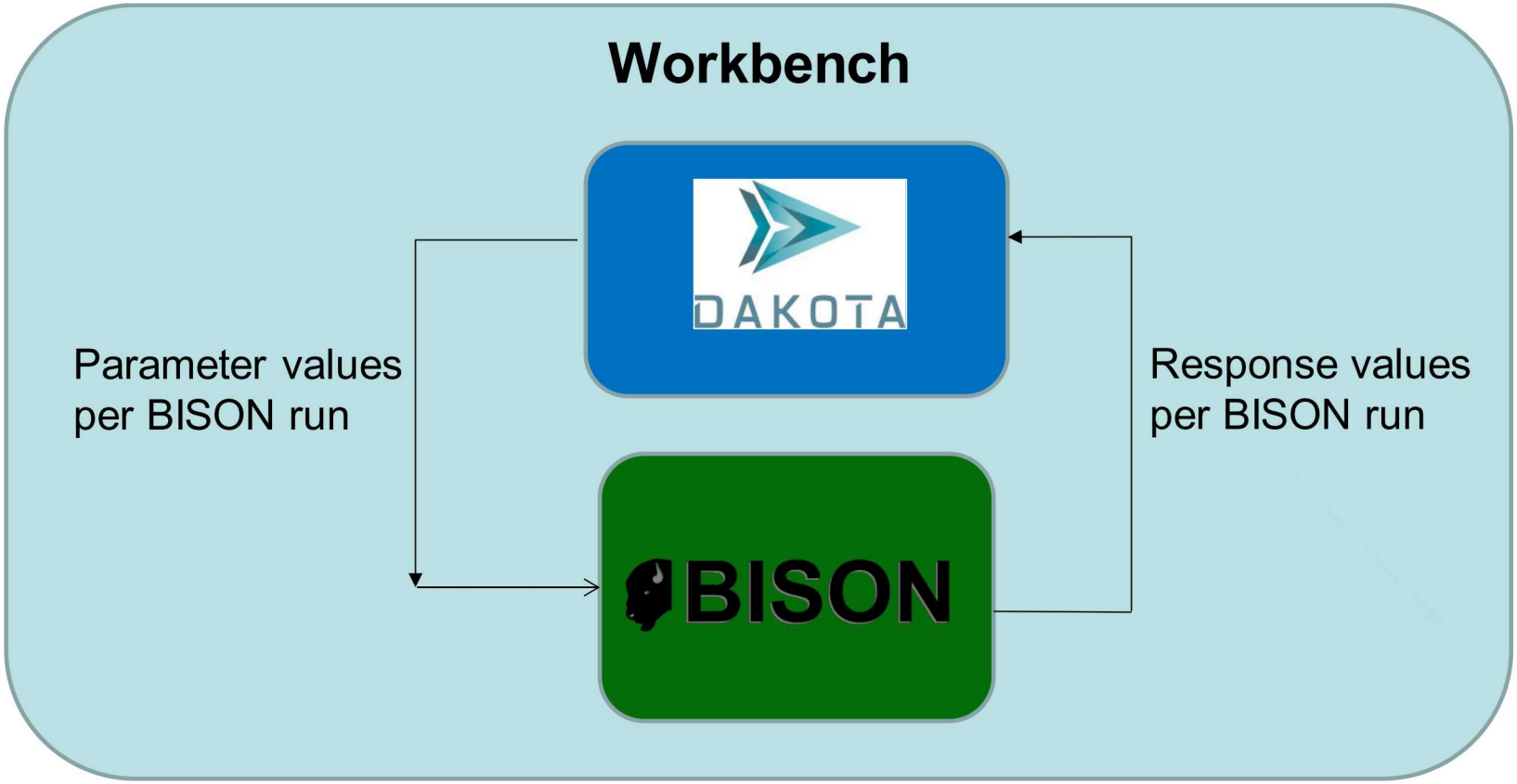


Dakota input file

Results panel showing function evaluations

Dakota within the NEAMS Workbench

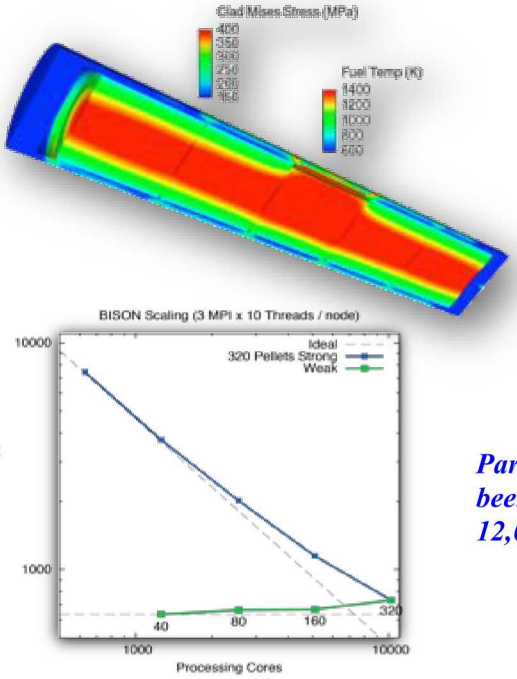



Workbench
DAKOTA
Parameter values per BISON run
Response values per BISON run
BISON

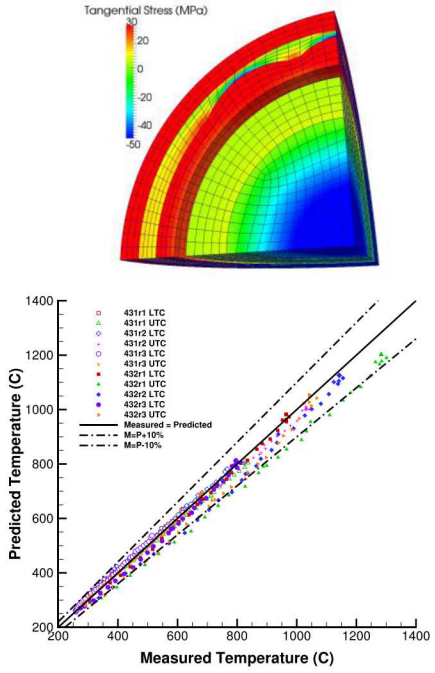
BISON Fuel Performance Code

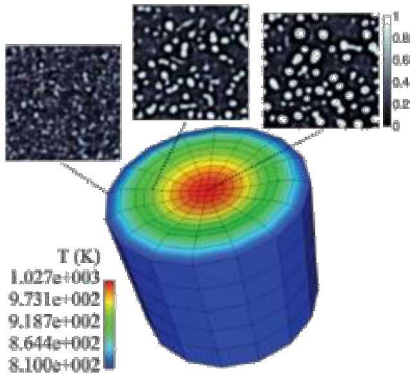
Solution method: Implicit finite element solution of the coupled thermomechanics and species diffusion equations using the MOOSE framework; 1D spherical, 2D axisymmetric or 3D

Multiphysics constitutive models: large deformation mechanics (plasticity and creep), cracking, thermal expansion, densification, fission gas release, radiation effects (swelling, thermal conductivity degradation, etc.)



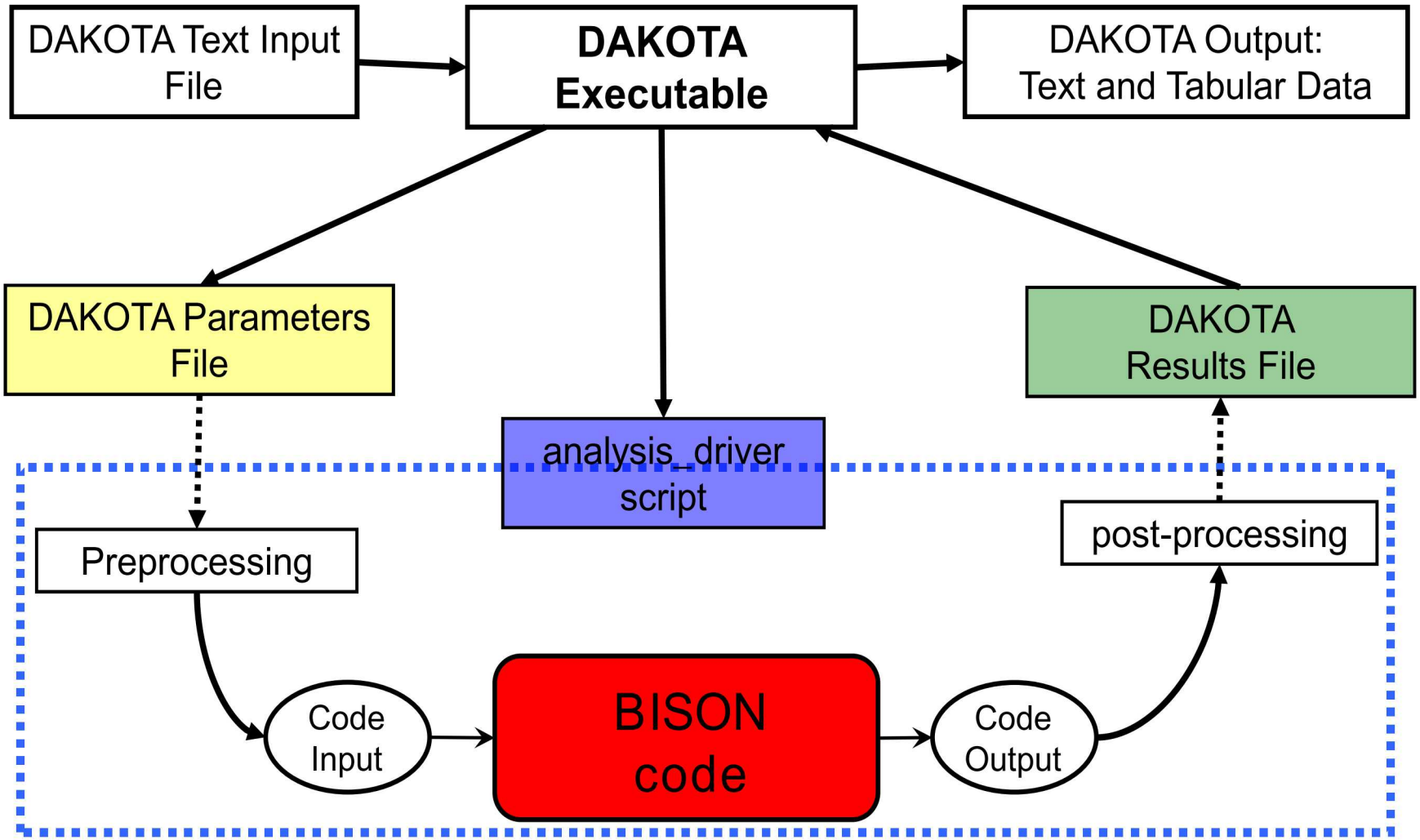
Parallel: has been run on 1 - 12,000 cpus.



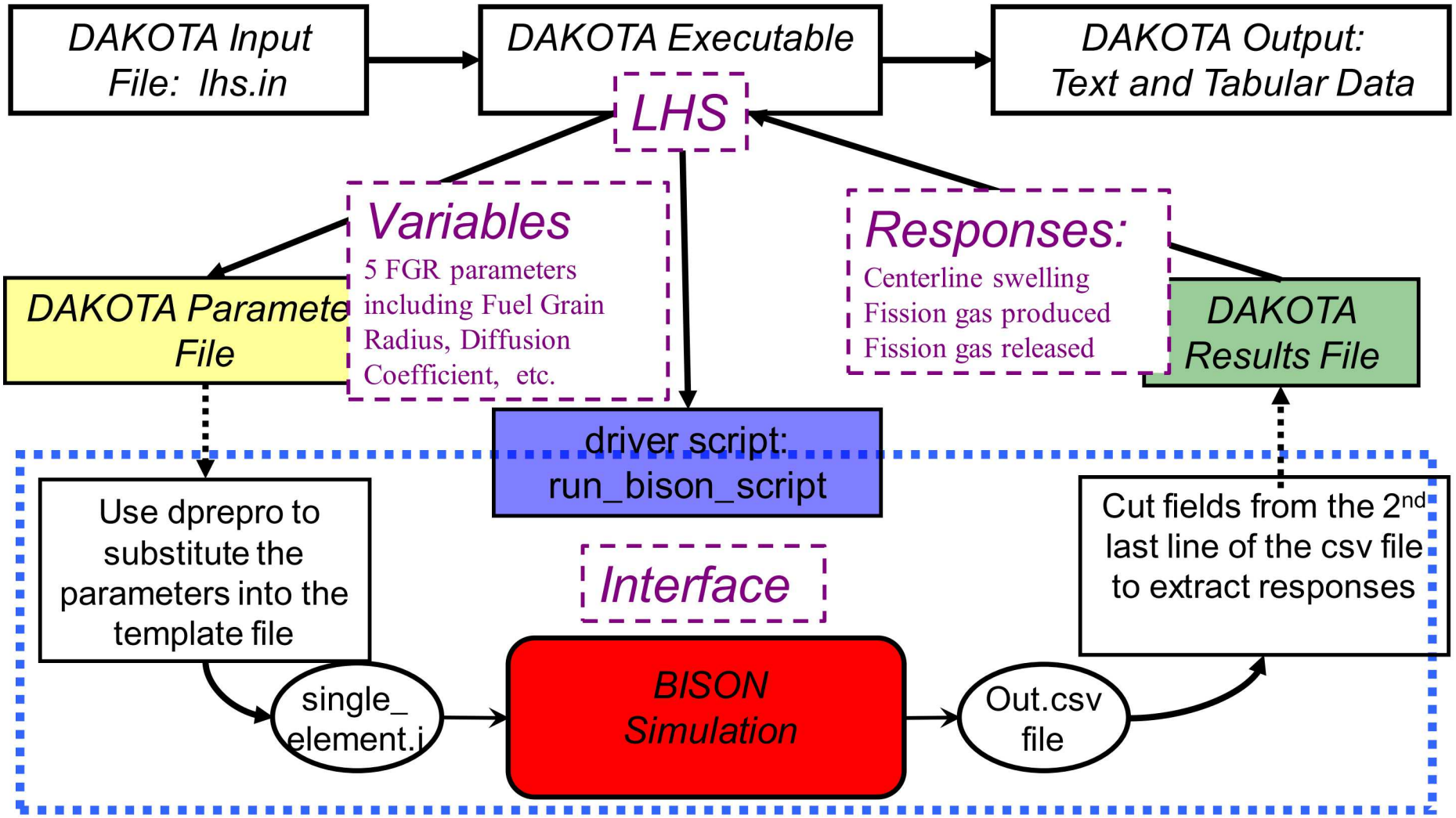


Couples readily to lower length scale models (MARMOT)

Dakota-Bison Coupling



Dakota Execution Workflow



Dakota Input File: lhs.in

```
environment
      tabular_graphics_data

method,
      sampling
      sample_type lhs
      samples = 20
      seed = 3487

variables,
       histogram_point_uncertain  = 5
       num_pairs  = 3 3 3 3 3
       abscissas = 0.95 1.0 1.05 0.4 1.0 1.6 0.1 1.0 10.0 0.1 1.0 10.0 0.316 1.0 3.162
       counts  = 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
       descriptor 'DakTemp'  'FuelGrainRad'  'DiffusCoeff'  'ResolParam' 'GrainBoundCoeff'
interface,
      system
       analysis_driver = 'run_bison_script'
       parameters_file = 'params.in'
       results_file    = 'results.out'
       file_tag file_save

responses,
      num_response_functions  = 3
      descriptors = 'FG_gen'  'FG_Rel' 'Swelling'
      no_gradients
      no_hessians
```

We produce tabular output

We are doing LHS sampling with 20 samples

We have five uncertain variables that are Histogram variables at the specified locations

Each time Dakota determines another set of parameters, it will execute the run_bison_script. Parameters and results will be stored in the named files.

There are three responses. Only response functions are returned, not the gradients or Hessians.

DAKOTA Parameters File

- The file is typically named “params.in” and is generated by DAKOTA for each function evaluation (for “system” or “fork” interface)
- The file lists the number of variables and the variable values in order, one value per line, followed by additional information:

```
{ DAKOTA_VARS     =                          5 }
{ DakTemp         =  9.500000000000000e-01 }
{ FuelGrainRad    =  4.000000000000000e-01 }
{ DiffusCoeff     =  1.000000000000000e+00 }
{ ResolParam      =  1.000000000000000e+00 }
{ GrainBoundCoeff =  1.000000000000000e+00 }
{ DAKOTA_FNS      =                          3 }
{ ASV_1:FG gen    =                          1 }
{ ASV_2:FG Rel    =                          1 }
{ ASV_3:Swelling  =                          1 }
{ DAKOTA_DER_VARS =                          0 }
{ DAKOTA_AN_COMPS =                          0 }
```

- Your script will extract the variables from this file and insert them into your code input deck.

Text substitution process

- "dprepro" is an executable perl script distributed with DAKOTA under the GNU LGPL
- Use dprepro in run_bison_script (explained next) to substitute the Dakota parameters (found in the params.in file) into a template file to create the BISON input file
- The syntax is:

```
>>  dprepro params.in bison.template bison.i
```

- For our example, we use the following command:

```
>>  dprepro --left-delimiter=% --right-delimiter=%
    $argv[1] single_element.template single_element.i
```

Driver file: run_bison_script

```
# --------------
# PRE-PROCESSING
# --------------
# Incorporate the parameters from DAKOTA into the template
set num = `echo $argv[1] | cut -c 11-`
mkdir workdir.$num
cp $argv[1] workdir.$num/params.in
cp single_element.template workdir.$num/.
cp cube_111.e workdir.$num/.
cp dprepro workdir.$num/.
cd workdir.$num

./dprepro --left-delimiter=% --right-delimiter=% params.in single_element.template
single_element.i

# --------
# ANALYSIS
# --------
/class/bison-opt -i single_element.i > temp.out
# ---------------
# POST-PROCESSING
# ---------------
# extract function value from the simulation output
tail -2 out_single_element.csv|cut -f2 -d',' >> results.tmp
tail -2 out_single_element.csv|cut -f3 -d',' >> results.tmp
tail -2 out_single_element.csv|cut -f4 -d',' >> results.tmp
cp results.tmp ../$argv[2]
cd ..
```

NOTE:

$argv[1] is “params.in.n” (from Dakota)

$argv[2] is “results.out.n” (-> to Dakota)

where n is the sample number

Run Dakota

- We now have set up the flow: the Dakota input file is lhs.in. Every time Dakota determines a new sample, it will invoke run_bison_script. The script will substitute the parameters into the BISON input file, run BISON, and post-process the results to return responses to Dakota.

```
>> dakota -i lhs.in
```

Optionally:

```
>> dakota -i lhs.in -c check
>> dakota -i lhs.in -o lhs.out
```

Run Dakota

```
Writing new restart file dakota.rst
methodName = nond_sampling
gradientType = none
hessianType = none

>>>>> Running nond_sampling iterator.

NonD lhs Samples = 20 Seed (user-specified) = 3487
------------------------------
Begin Function Evaluation    1
------------------------------
Parameters for function evaluation 1:
                     9.5000000000e-01 DakTemp
                     4.0000000000e-01 FuelGrainRad
                     1.0000000000e+00 DiffusCoeff
                     1.0000000000e+00 ResolParam
                     1.0000000000e+00 GrainBoundCoeff

run_bison_script params.in.1 results.out.1

Active response data for function evaluation 1:
Active set vector = { 1 1 1 }
                     5.0099634673e+02 FG_gen
                     5.5082361489e+01 FG_rel
                     7.4027484507e-02 Swelling
```

Run Dakota

Sandia National Laboratories

```
<<<<< Function evaluation summary: 20 total (20 new, 0 duplicate)

Statistics based on 20 samples:

Moment-based statistics for each response function:
                            Mean           Std Dev          Skewness          Kurtosis
          FG_gen  5.0099634673e+02  1.1664023442e-13  1.0829771494e+00 -2.2352941176e+00
          FG_rel  4.3127004142e+01  4.0357474122e+01  1.6414892498e+00  2.6450662130e+00
        Swelling  5.1849404115e-02  3.7028882698e-02  1.4493006479e+00  1.1234191683e+00

95% confidence intervals for each response function:
                      LowerCI_Mean      UpperCI_Mean    LowerCI_StdDev    UpperCI_StdDev
          FG_gen  5.0099634673e+02  5.0099634673e+02  8.8703777868e-14  1.7036141802e-13
          FG_rel  2.4239124846e+01  6.2014883438e+01  3.0691471409e+01  5.8944982008e+01
        Swelling  3.4519353559e-02  6.9179454671e-02  2.8160109604e-02  5.4083335785e-02
```

Run Dakota

Sandia National Laboratories

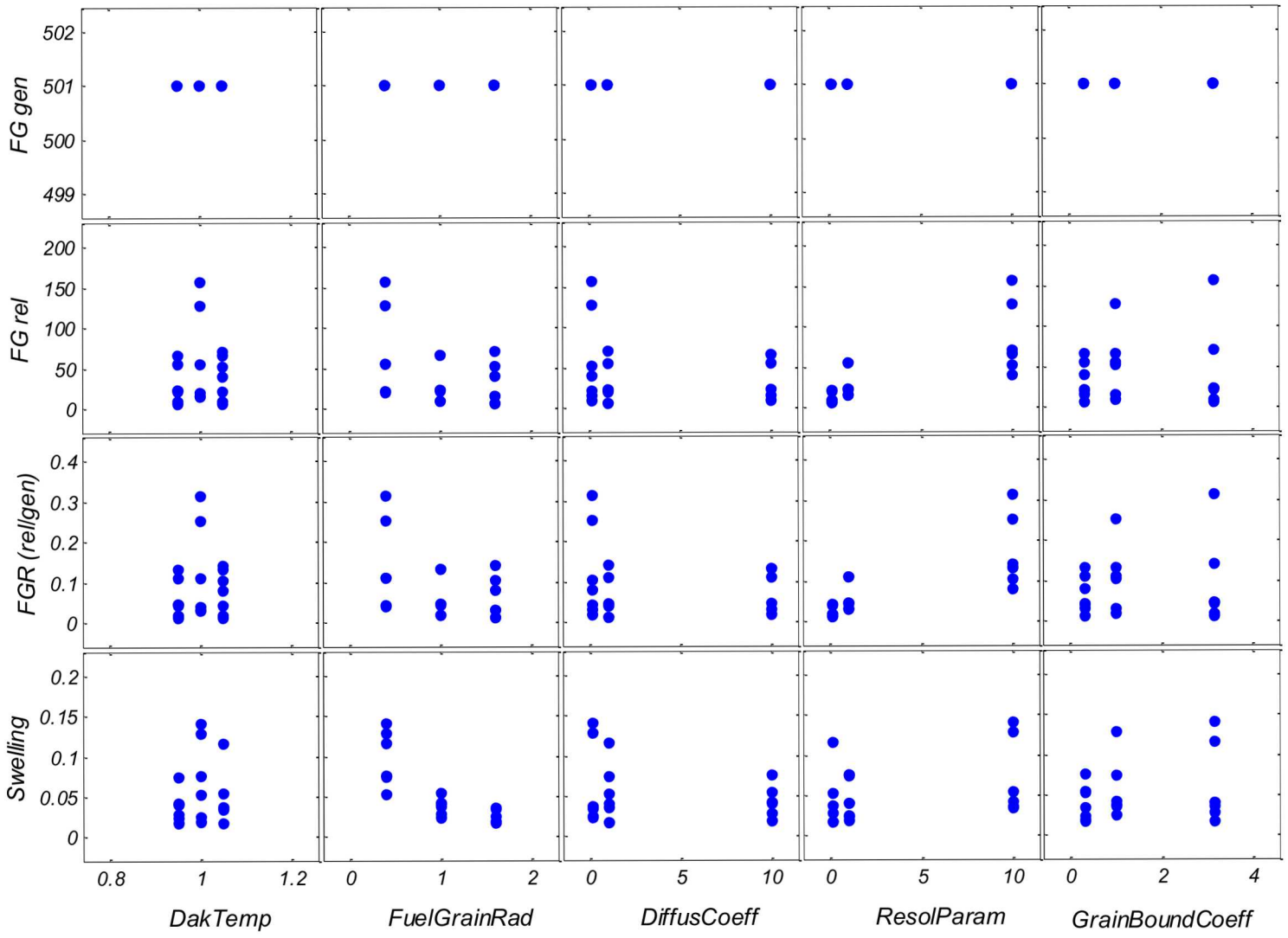
```
Simple Correlation Matrix among all inputs and outputs:
                  DakTemp FuelGrainRad  DiffusCoeff   ResolParam GrainBoundCoeff   FG_gen
 DakTemp      1.00000e+00
FuelGrainRad  2.22804e-01  1.00000e+00
 DiffusCoeff -2.72213e-01 -4.69579e-02  1.00000e+00
  ResolParam  3.46201e-01  8.44051e-02 -3.84643e-02  1.00000e+00
GrainBoundCoeff -2.74813e-01 -6.03050e-02 -4.59231e-02 -6.99213e-02  1.00000e+00
      FG_gen  0.00000e+00  0.00000e+00  0.00000e+00  0.00000e+00  0.00000e+00  1.00000e+00
      FG_rel  9.24460e-02 -4.34988e-01 -9.31770e-02  7.51561e-01  7.69795e-02  0.00000e+00
    Swelling  1.12108e-01 -7.84487e-01 -1.65934e-01  3.09223e-01  1.98031e-01  0.00000e+00


Partial Correlation Matrix between input and output:
                   FG_gen       FG_rel     Swelling
     DakTemp -0.00000e+00 -1.94114e-01  4.41138e-01
FuelGrainRad -0.00000e+00 -7.61412e-01 -9.09650e-01
 DiffusCoeff -0.00000e+00 -2.44164e-01 -2.95894e-01
  ResolParam -0.00000e+00  8.87268e-01  6.14260e-01
GrainBoundCoeff -0.00000e+00  1.74623e-01  4.89549e-01
```

dakota_tabular.dat file

Sandia National Laboratories

%eval_id	DakTemp	FuelGrainRad	DiffusCoeff	ResolParam	GrainBoundCoeff	FG_gen	FG_rel	Swelling
1	0.95	0.4	1	1	1	500.9963467	55.08236149	0.074027485
2	0.95	1.6	1	0.1	3.162	500.9963467	5.90327614	0.017369679
3	1	1.6	10	1	0.316	500.9963467	14.94445219	0.019417587
4	1	0.4	1	0.1	0.316	500.9963467	20.5047579	0.052740915
5	1	0.4	0.1	10	1	500.9963467	127.1503408	0.128708461
6	1.05	1.6	0.1	10	0.316	500.9963467	40.19855927	0.034059363
7	1	0.4	0.1	10	3.162	500.9963467	156.8925373	0.1408676
8	1.05	1	10	10	0.316	500.9963467	66.14807271	0.054998274
9	0.95	1	10	1	3.162	500.9963467	23.83181335	0.03984616
10	1.05	1.6	1	0.1	0.316	500.9963467	5.816189239	0.017687502
11	1.05	1	0.1	0.1	1	500.9963467	9.586800228	0.037444359
12	1	1.6	0.1	1	1	500.9963467	14.8602684	0.025451241
13	1	0.4	10	1	0.316	500.9963467	55.17689751	0.075503977
14	1.05	0.4	1	0.1	3.162	500.9963467	22.16840092	0.11637827
15	1.05	1.6	1	10	3.162	500.9963467	70.20286002	0.0352169
16	1.05	1.6	0.1	10	1	500.9963467	52.80899606	0.0352169
17	0.95	1	10	0.1	3.162	500.9963467	9.485112967	0.028143603
18	0.95	1	0.1	1	0.316	500.9963467	22.11395758	0.02264172
19	0.95	1	10	10	1	500.9963467	65.94341485	0.0414814
20	0.95	1	1	1	3.162	500.9963467	23.72101398	0.03978669

dakota_tabular.dat file
Minitab plotmatrix(x(:,1:5),x(:,6:9))



DAKOTA Results File

- The file is typically named "results.out" and is generated by the simulation code (or script) for return to DAKOTA.
- You need to extract the relevant response(s) from your code output and write to the results.out file with the function values in order, one value per line. Here is the results.out.1 file listing created by our run_bison_script:

f1
f2
…
fM

```
500.99634672866

55.082361489447

0.074027484506627
```

- You can add a text label after the function value (on the same line) to help you keep track of the f-values / constraints.
- If your code generates gradients of the function-values and/or Hessian values (matrix of 2nd derivatives), also report in this file.

RECAP

Sandia National Laboratories

- We have set up the framework:
  - The Dakota input file is lhs.in.
  - Every time Dakota determines a new sample, it will invoke run_bison_script.
  - This script will substitute the parameters into the BISON input file, run BISON, and post-process the results to return responses to Dakota.
  - We looked at the Dakota screen output and the dakota_tabular.dat file.
  - We looked at the format that Dakota generates with params.in files and results.out files.

Description of BISON Case Study

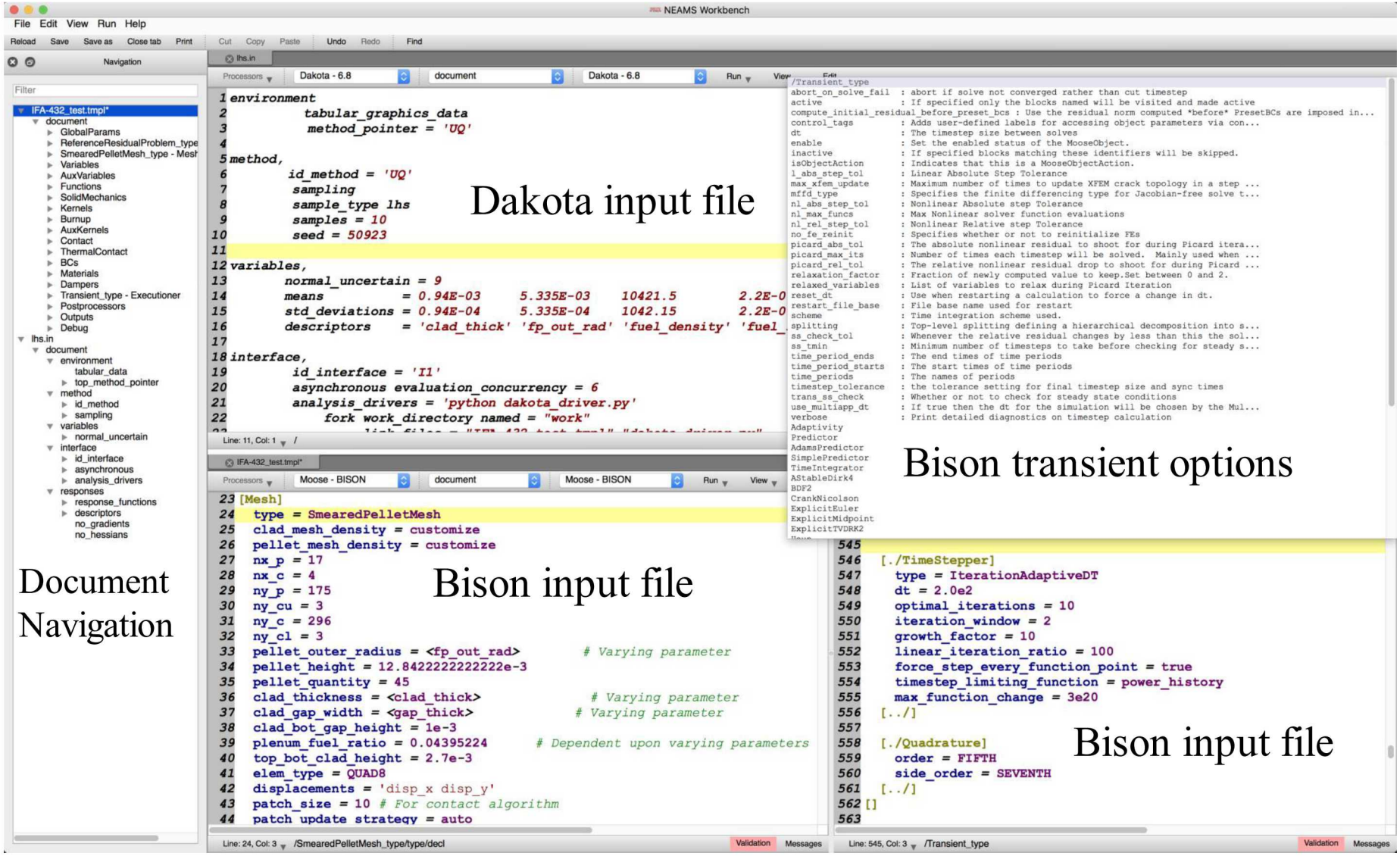
- The Bison fuel performance code was coupled to Dakota to analyze rod 1 from the IFA-432 **experimental assembly irradiated at the Halden reactor in Norway.** [1]
- Kyle Gamble (Idaho National Laboratory) provided the Bison input and example power history, axial peaking factors, etc.
- Input uncertainties were assumed in manufacturing parameters (e.g., cladding thickness), and model parameters (e.g., fuel thermal conductivity).
- Outputs of interest include fuel centerline temperature and fission gas release.
- Related work in reference 2 and 3.

1. Hann, C.R., et al. "Data Report for the NRC/PNL Halden Assembly IFA-432," NUREG/CR-0560, PNL-2673, 1978.
2. Swiler, L., Gamble, K., Schmidt, R. and R. Williamson. SAND Report 2015-8088 "Sensitivity Analysis of OECD Benchmark Tests in BISON."
3. Gamble, K. and L. Swiler. "Uncertainty Quantification and Sensitivity Analysis Applications to Fuel Performance Modeling." *TOPFUEL American Nuclear Society Conference*, Sept. 2016.

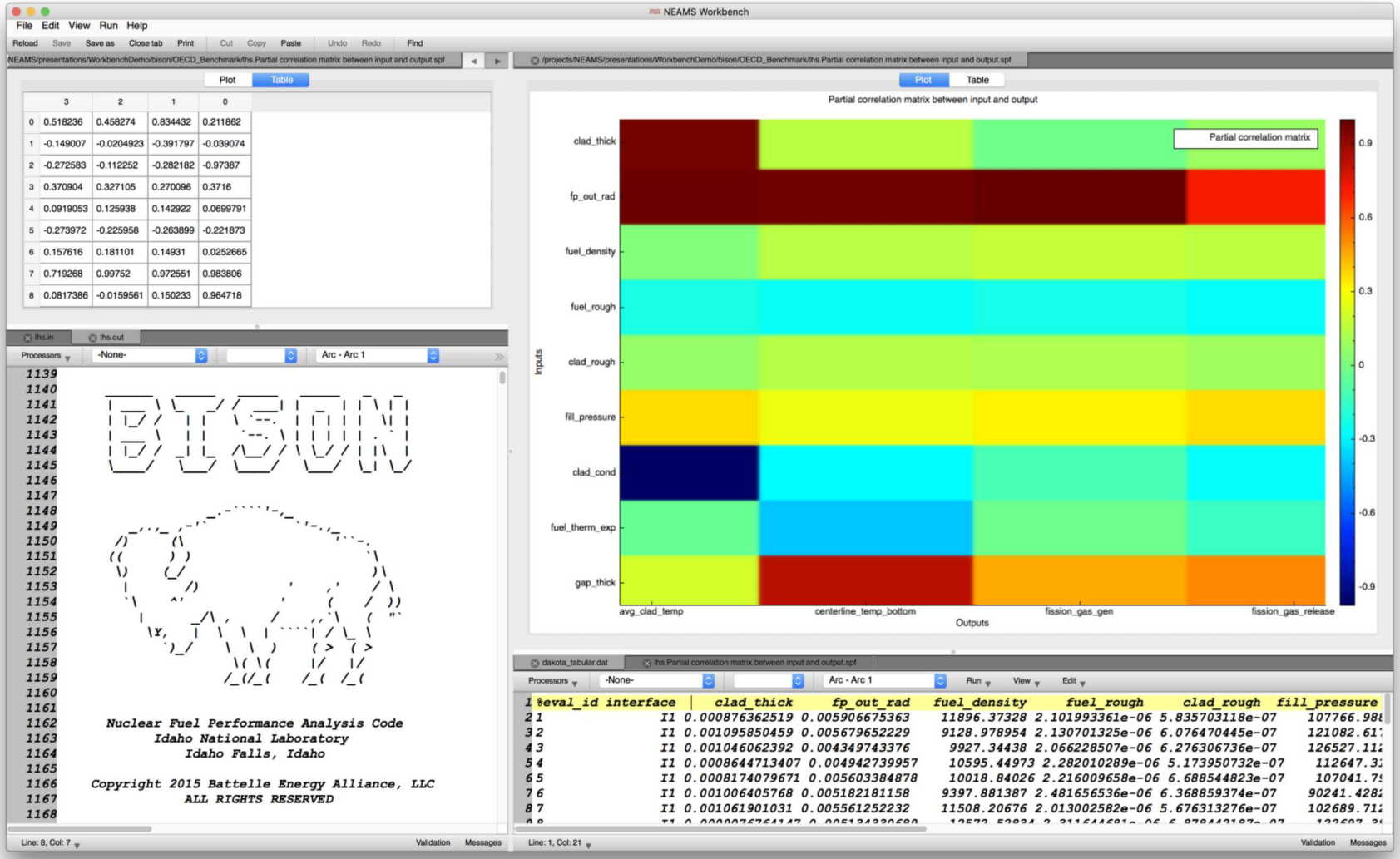
Input / Outputs for Sampling Study

- Inputs
  - Fill Pressure
  - Fuel conductivity
  - Clad conductivity
  - Fuel thermal expansion
  - Gas conductivity factor
  - Gap thickness
- Outputs
  - Clad Temperature
  - Fuel Centerline Temperature
  - Fission Gas Release
  - Clad Diameter

Workbench – Dakota - BISON



Workbench – Dakota - BISON



	3	2	1	0
0	0.518236	0.458274	0.834432	0.211862
1	-0.149007	-0.0204923	-0.391797	-0.039074
2	-0.272583	-0.112252	-0.282182	-0.97387
3	0.370904	0.327105	0.270096	0.3716
4	0.0919053	0.125938	0.142922	0.0699791
5	-0.273972	-0.225958	-0.263899	-0.221873
6	0.157616	0.181101	0.14931	0.0252665
7	0.719268	0.99752	0.972551	0.983806
8	0.0817386	-0.0159561	0.150233	0.964718

	%eval_id	interface	clad_thick	fp_out_rad	fuel_density	fuel_rough	clad_rough	fill_pressure
2	1	I1	0.000876362519	0.005906675363	11896.37328	2.101993361e-06	5.835703118e-07	107766.98
3	2	I1	0.001095850459	0.005679652229	9128.978954	2.130701325e-06	6.076470445e-07	121082.61
4	3	I1	0.001046062392	0.004349743376	9927.34438	2.066228507e-06	6.276306736e-07	126527.11
5	4	I1	0.0008644713407	0.004942739957	10595.44973	2.282010289e-06	5.173950732e-07	112647.3
6	5	I1	0.0008174079671	0.005603384878	10018.84026	2.216009658e-06	6.688544823e-07	107041.7
7	6	I1	0.001006405768	0.005182181158	9397.881387	2.481656536e-06	6.368859374e-07	90241.428
8	7	I1	0.001061901031	0.005561252232	11508.20676	2.013002582e-06	5.676313276e-07	102689.71

Summary

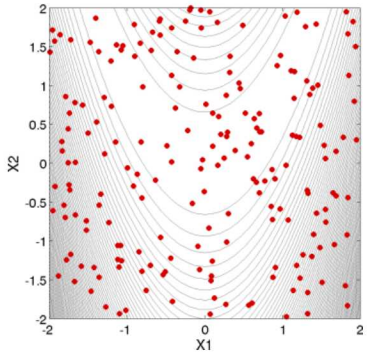
- Dakota is a SNL tool that performs uncertainty quantification, optimization, and sensitivity analysis of simulation models.
- NEAMS Workbench is an ORNL tool to allow the user to create, validate, parse, and process input files for various codes including SCALE, ARC, and Bison.
- Bison is an INL tool to model performance of nuclear fuels.
- We demonstrated the integration of all three: both Dakota and Bison are integrated within Workbench. In one study, we can define the Dakota input and the corresponding Bison input.
- Demonstrated post-processing of sample results with visual display of correlation matrices.
- Interfacing between Dakota and BISON (or any other code) can be done with a user-provided script or by the “Dakota driver” utility provided by the Workbench team

Sandia National Laboratories

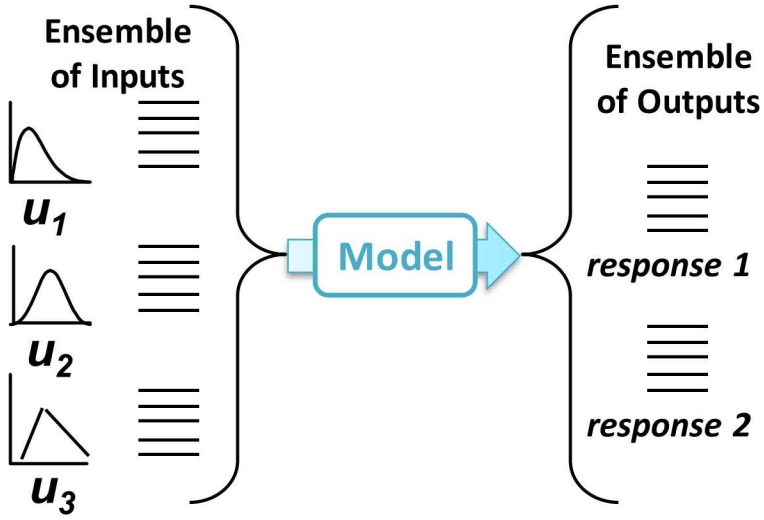
Backup slides

Workhorse UQ Method: Monte Carlo Sampling

- Sampling methods draw (pseudo-random) realizations from the specified input distributions, run the simulation, and calculate sample statistics:
  - Sample moments, min/max, empirical PDF/CDF, based on ensemble of calculations



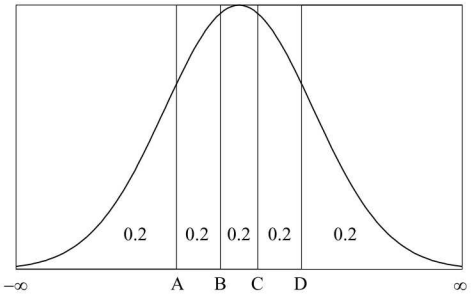
Monte Carlo sample of two input variables



- Robust even for complex, poorly-behaved simulations
- Slow, though reliable convergence: $O(N^{-1/2})$, (in theory) independent of dimension
- Parallelism: all samples are known at onset and can be evaluated concurrently

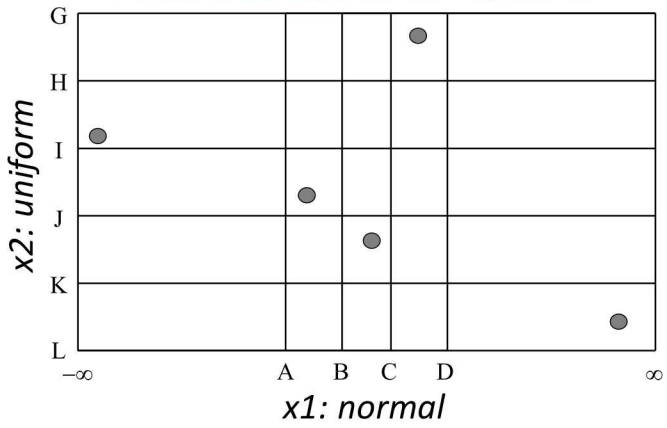
Latin Hypercube Sampling (LHS)

- Dakota has `sample_type` options `random` and `lhs`
- LHS is recommended when possible
  - Better convergence rate and stability across replicates
  - Any follow-on studies must double the sample size
- LHS (McKay and Conover): stratified random sampling among equal probability bins for all 1-D projections of an n-dimensional set of samples

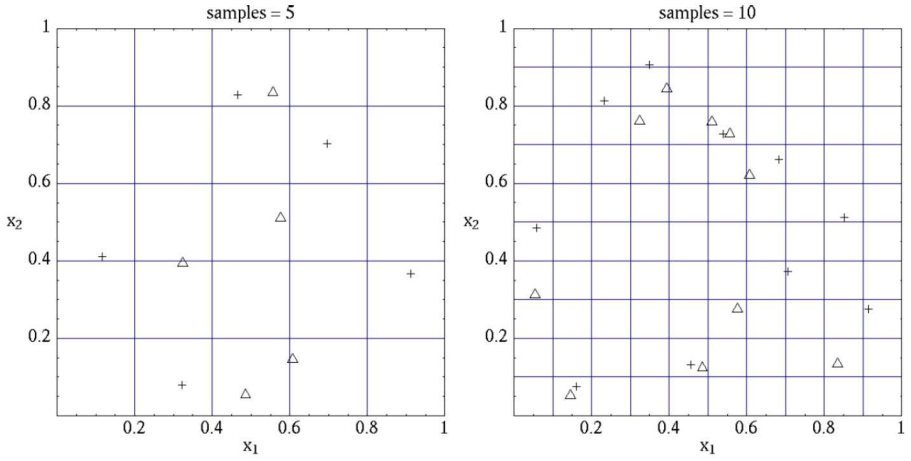


example equi-probable intervals for an LHS of size 5 on a normal random variable

A two-dimensional LHS of size 5



Uniform LHS designs of sizes 5 and 10



Selecting a UQ Method

Sandia National Laboratories

Consider variable characterizations, model properties, ultimate UQ goal to choose a method

Sampling (Monte Carlo, LHS)

- Robust, understandable, and applicable to most any model
- Slow to converge
- Moments, PDF/CDF, correlations, min/max

Stochastic Expansions

- Surrogate models tailored to UQ for continuous variables
- Highly efficient for smooth model responses
- Moments, PDF/CDF, Sobol indices

Reliability

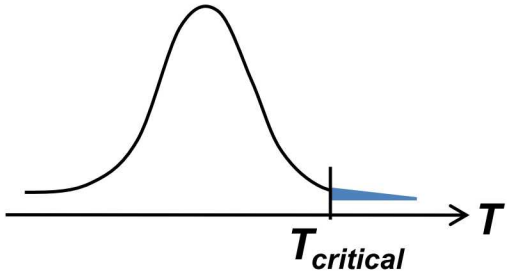
- Goal-oriented; target particular response or probability levels
- Efficient local (require derivatives) / global variants
- Moments, PDF/CDF, importance factors

Epistemic

- Non-probabilistic methods
- Generally applicable, can be costly when no surrogate
- Belief/plausibility, intervals, probability of frequency

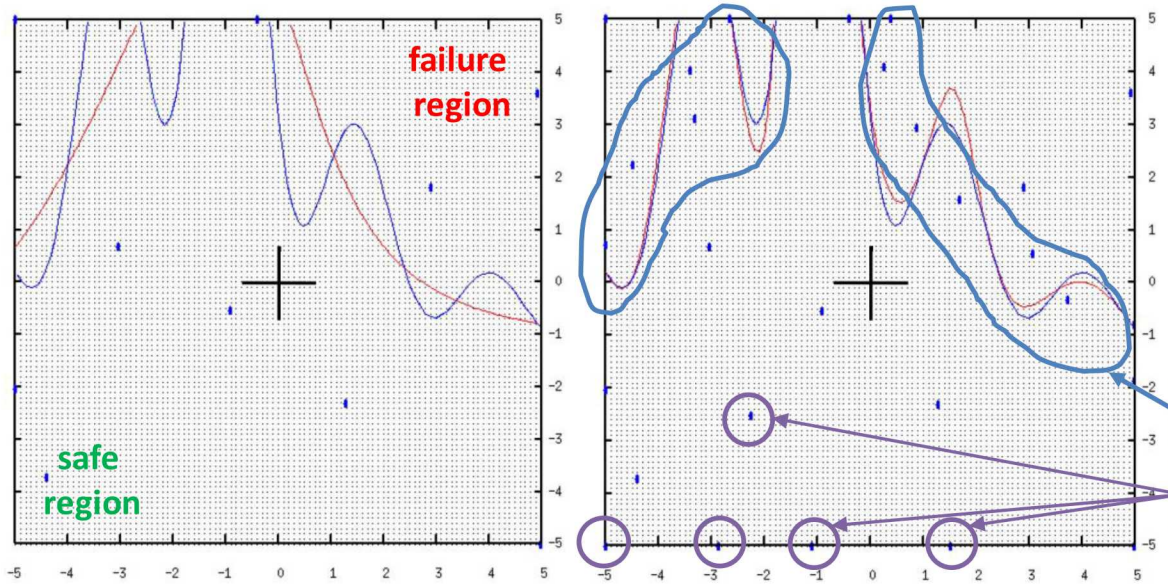
Reliability Methods: What Are They?

- Goal-oriented methods that focus on regions of probability or response space of interest, for example:
  - What temperature is achieved with 99% probability?
  - What is the probability of exceeding $T_{critical}$?
- Naïve sampling can be ineffective / under-resolved
  - Run 10,000 samples, only 5 are in relevant region
- Need to specify to Dakota
  - Probability or response threshold(s) of interest using `probability_levels, response_levels`
  - Method choice
    - Mean-value: best for linear problems, normally distributed parameters, efficient derivatives; specify `local_reliability` (with no `mpp_search`)
    - MPP: computes most probable point of failure when failure boundary is near linear or quadratic; specify `local_reliability` (with an `mpp_search` option)
    - Adaptive: computes probability of failure for complicated failure boundaries; specify `global_reliability`



Reliability Methods: How Do They Work?

- Reliability methods try to directly calculate statistics of interest:
  - Make simplifying approximations and/or
  - Recast the UQ as an iterative procedure, such as iteratively refined sampling or as a nonlinearly constrained optimization problem



Mean-value: uses derivatives; make a linearity (and possibly normality) assumption and project

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

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

MPP: solve an optimization problem to directly determine input values giving rise to most probable point of failure

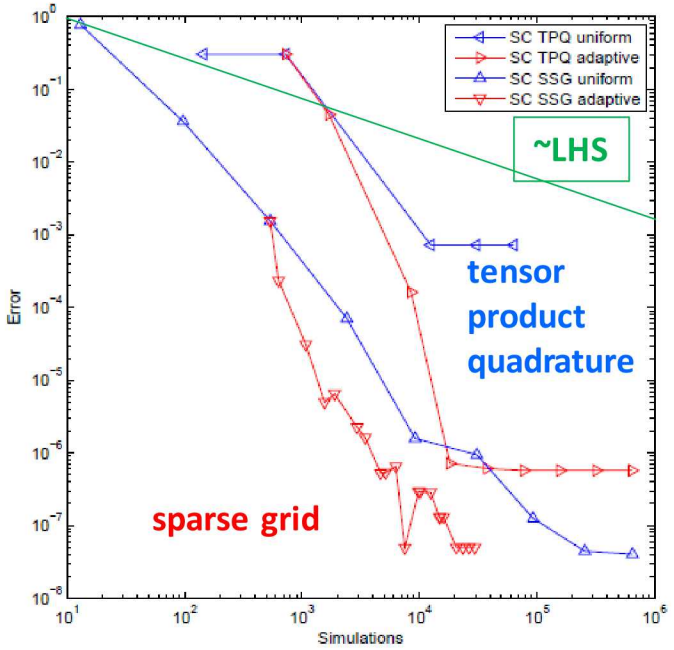
$$\text{minimize} \quad u^T u$$

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

Adaptive: iteratively refine understanding of failure region

Stochastic Expansions: What Are They?

- General-purpose UQ methods that build UQ-tailored polynomial approximations of the output responses
- Perform particularly well for smooth model responses
- Resulting convergence of statistics can be considerably faster than sampling methods
- Two categories
  - Polynomial Chaos: specify the type of orthogonal polynomials and coefficient estimation scheme
  - Stochastic Collocation: specify the type of polynomial basis and the points at which the response will be interpolated

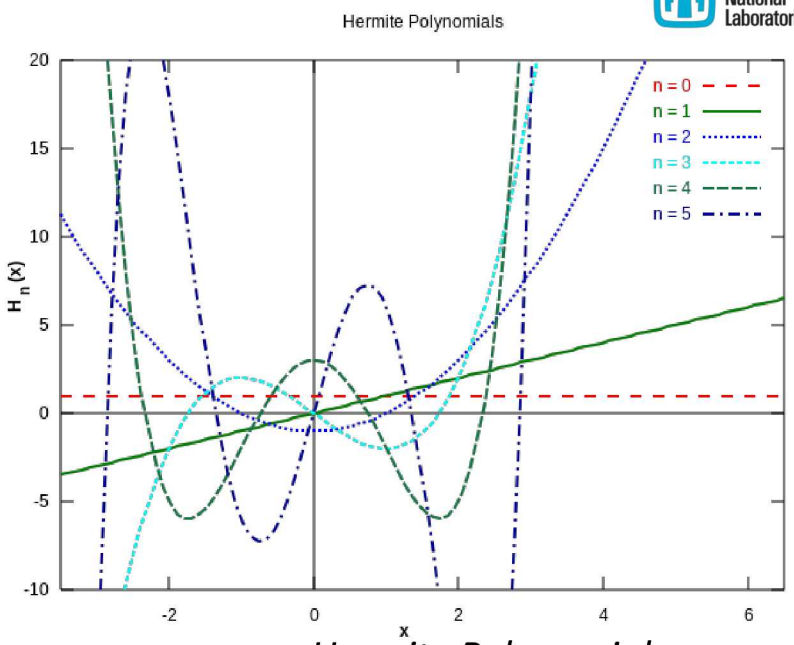


Polynomial Chaos: How Does It Work?

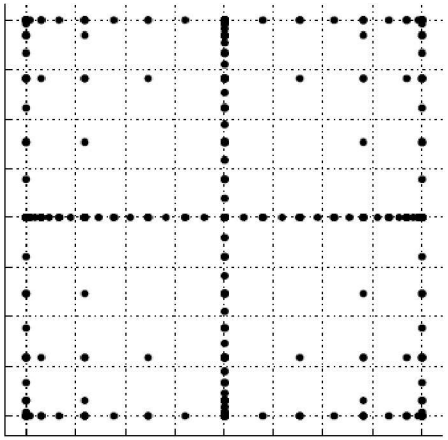
- Uses an orthogonal polynomial basis $\Psi_j(\xi)$, e.g., Hermite polynomials orthogonal w.r.t. normal density
- Evaluates the model in a strategic way (sampling, quadrature, sparse grids, cubature)…
- …to efficiently approximate the coefficients of an orthogonal polynomial approximation of the response

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

$$\alpha_j = \frac{\langle R, \Psi_j \rangle}{\langle \Psi_j^2 \rangle} = \frac{1}{\langle \Psi_j^2 \rangle} \int_{\Omega} R \, \Psi_j \, \varrho(\boldsymbol{\xi}) \, d\boldsymbol{\xi}$$



Hermite Polynomials



Sparse Grid

Adaptive Basis Selection: Compressed Sensing in High Dimensions

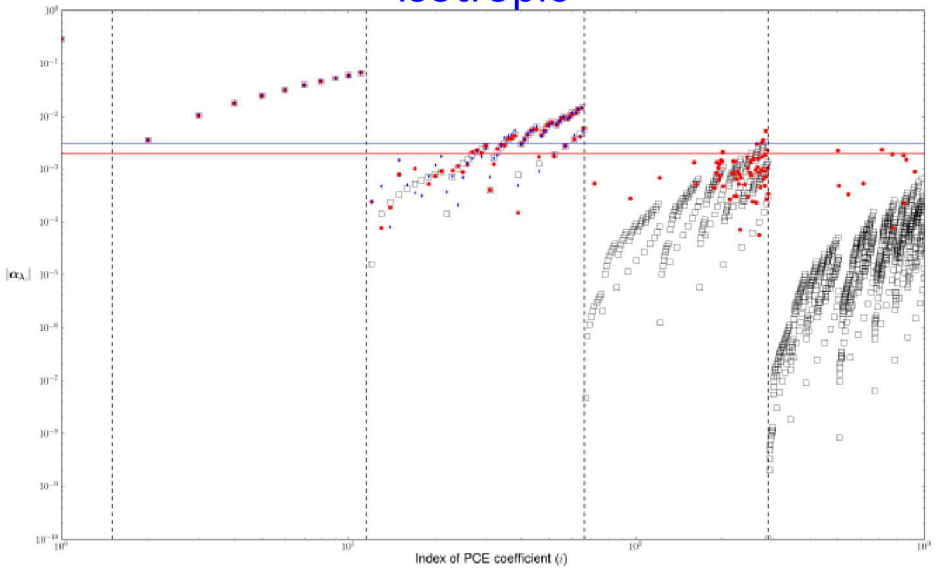
In high dimensions, we may only be able to consider a 2nd or 3rd degree total-order basis

p	3	4	5
$\mathcal{A}_{3,1}^{40}$	$12,341$	$135,751$	$1,221,759$

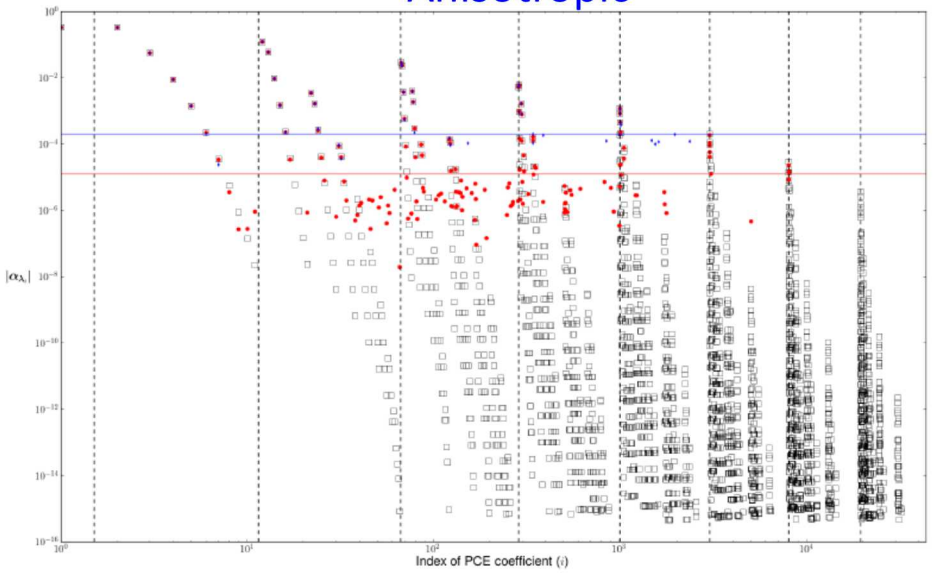
What if the function is anisotropic and important coefficients correspond to p > 3?

We seek algorithms that can adaptively determine an effective basis → *expanding front*

Isotropic



Anisotropic



Bayesian Calibration

- **Goal of Bayesian methods: Obtain statistical characterization of parameters consistent with data and its associated uncertainty**
- Generate posterior distributions on model parameters, given
  - Experimental data
  - A prior distribution on model parameters
  - A presumed probabilistic relationship between experimental data and model output that can be defined by a likelihood function

$$\pi(\theta \mid d) \propto \pi(\theta) L(d \mid \theta)$$

Posterior distribution

Model parameters

Observed Data

Prior parameter distribution

Likelihood function which Incorporates the model

Bayesian Calibration

- Experimental data = Model output + error

$$d_i(\boldsymbol{x}_i) = M(\boldsymbol{\theta}, \boldsymbol{x}_i) + \varepsilon_i$$

- If we assume error terms are independent, zero mean Gaussian random variables with variance σ^2, the likelihood is:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(d_i(\boldsymbol{x}_i) - M(\boldsymbol{\theta}, \boldsymbol{x}_i))^2}{2\sigma^2} \right]$$

- How do we obtain the posterior?
  - We use a technique called Markov Chain Monte Carlo (MCMC)
  - In MCMC, the idea is to *generate a sampling density that is approximately equal to the posterior.*
  - MCMC algorithms are computationally expensive: Typically requires $O(10^4) - O(10^5)$ samples → prohibitive for high fidelity models

Bayesian Calibration

- Efficient methods for MCMC
  - Surrogate models, PCE. Adapt the emulator in regions of high likelihood
  - DRAM algorithms: Adapt the proposal covariance based on the accepted samples. Precondition the proposal covariance by the inverse of the Hessian of the posterior.
  - Hybrid MCMC: introduce an auxiliary momentum vector and implement Hamiltonian dynamics, so the potential energy function is the target density.
- Problems:
  - Bayesian analysis is conceptually attractive
  - **But it is difficult to obtain good posteriors**
  - MCMC algorithms are very expensive, need a good "mixing" of the chains, easy to get garbage posteriors and not realize it, results are highly dependent on assumptions made about process variance

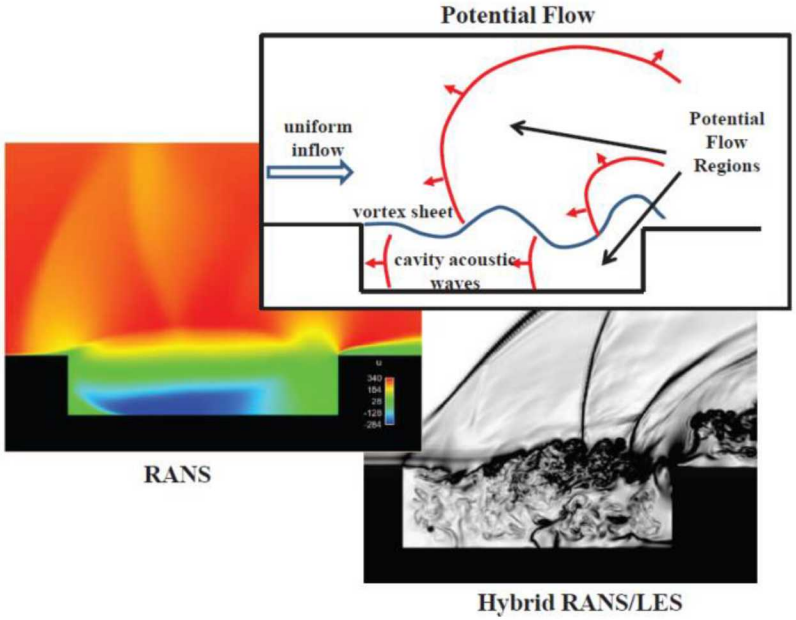
Multi-fidelity

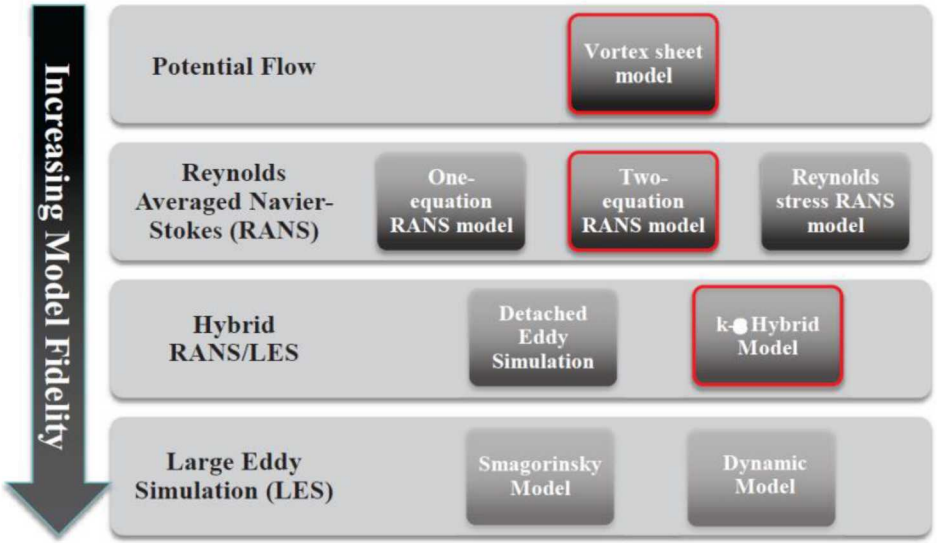
Discrete model choices for same physics

- A clear hierarchy of fidelity (low to high)
- The model inputs and outputs are the same

Horizontal Axis Wind Turbine

Vertical Axis Wind Turbine





Multifidelity UQ using Stochastic Expansions

- High-fidelity simulations (e.g., RANS, LES) can be prohibitive for use in UQ
- Low fidelity "design" codes often exist that are predictive of basic trends
- Can we leverage LF codes w/i HF UQ in a rigorous manner? → global approxs. of model discrepancy

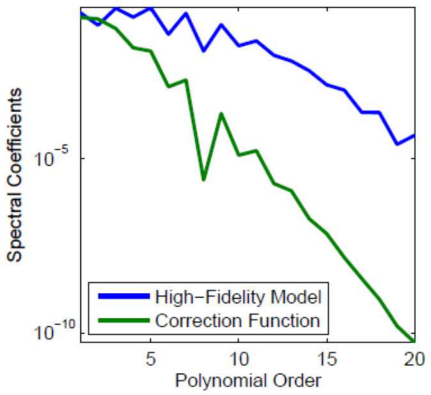
$$\hat{f}_{hi}(\xi) = \sum_{j=1}^{N_{lo}} f_{lo}(\xi_j) L_j(\xi) + \sum_{j=1}^{N_{hi}} \Delta f(\xi_j) L_j(\xi)$$

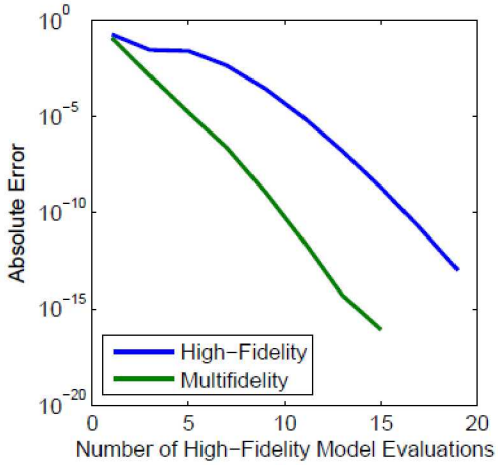
$N_{lo} >> N_{hi}$

$$R_{\text{high}}(\xi) = e^{-0.05\xi^2} \cos 0.5\xi - 0.5 e^{-0.02(\xi-5)^2}$$

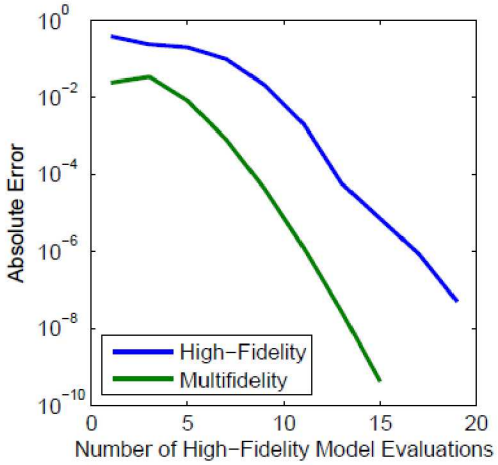
$$R_{\text{low}}(\xi) = e^{-0.05\xi^2} \cos 0.5\xi,$$

discrepancy





(a) Error in mean



(b) Error in standard deviation

Low fidelity

CACTUS: Code for Axial and Crossflow TUrbine Simulation

High fidelity: DG formulation for LES Full Computational Fluid Dynamics/ Fluid-Structure Interaction

Multi-level Monte Carlo

Monte Carlo estimator: $\hat{Q} = \frac{1}{N} \sum_{i=1}^{N} Q_i$ → analytic variance $Var[\hat{Q}] = \frac{s_Q^2}{N}$

Geometrical MLMC – targeting discretization levels

Multilevel Monte Carlo estimator

$$\hat{Q}_M^{\mathrm{ML}} \overset{\mathrm{def}}{=} \sum_{\ell=0}^{L} \hat{Y}_{\ell,N_\ell}^{\mathrm{MC}} = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \left(Q_{M_\ell}^{(i)} - Q_{M_{\ell-1}}^{(i)} \right)$$

$$\left.\begin{aligned} \mathcal{C}(\hat{Q}_M^{ML}) &= \sum_{\ell=0}^{L} N_\ell C_\ell \\ \sum_{\ell=0}^{L} N_\ell^{-1} \mathbb{V}ar\,(Y_\ell) &= \varepsilon^2/2 \end{aligned}\right\} \xrightarrow{\textit{Lagrange multipliers}} N_\ell = \frac{2}{\varepsilon^2} \left[\sum_{k=0}^{L} \left(\mathbb{V}ar\,(Y_k)\, C_k \right)^{1/2} \right] \sqrt{\frac{\mathbb{V}ar\,(Y_\ell)}{C_\ell}}$$

[Giles, 2008]

Dakota UQ Methods Summary

Sandia National Laboratories

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

Also see Usage Guidelines in User's Manual

Examples

Sandia National Laboratories

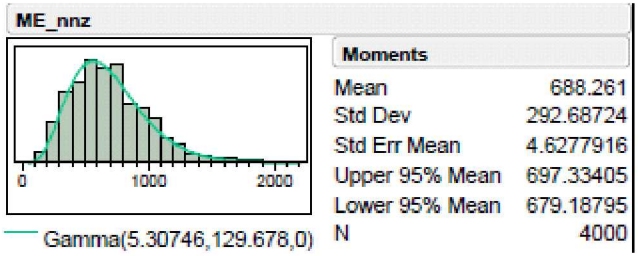
Example: Uncertainty in Boiling Rate in Nuclear Reactor Core (DOE CASL)

Sandia National Laboratories

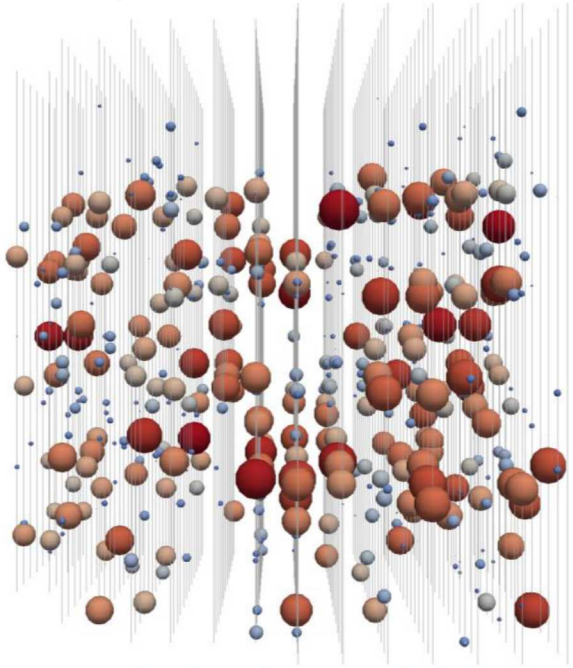
- Use nuclear reactor thermal-hydraulics model to assess uncertainty in localized boiling due to variable operating conditions
- Compare Dakota UQ approaches and modelling assumptions

Method	ME_nnz		ME_meannz		ME_max	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
LHS (40)	651.225	297.039	127.836	27.723	361.204	55.862
LHS (400)	647.33	286.146	127.796	25.779	361.581	51.874
LHS (4000)	688.261	292.687	129.175	25.450	364.317	50.884
PCE (Θ(2))	687.875	288.140	129.151	25.7015	364.366	50.315
PCE (Θ (3))	688.083	292.974	129.231	25.3989	364.310	50.869
PCE (Θ (4))	688.099	292.808	129.213	25.4491	364.313	50.872

mean and standard deviation of key metrics



normally distributed inputs need not give rise to normally distributed outputs…



anisotropic uncertainty distribution in boiling rate throughout quarter core model (side view)

Studies with Dakota-Bison

Sandia National Laboratories

- BISON is a nuclear fuels performance code developed at INL
- This work is sponsored by DOE's Nuclear Engineering Advanced Modeling and Simulation Program (NEAMS)

- Several years of collaboration, extensive set of case studies performed
  - Swiler, L.P., Williamson, R. L. and D. M. Perez. "Calibration of a Fuel Relocation Model in BISON", American Nuclear Society Mathematics and Computational Methods (M&C) meeting, May 2013.
  - "Uncertainty and sensitivity analysis of fission gas behavior in engineering-scale fuel modeling." Giovanni Pastore, L.P. Swiler, J.D. Hales, S.R. Novascone, D.M. Perez, B.W. Spencer, L. Luzzi, P. Van Uffelen, and R.L. Williamson. *Journal of Nuclear Materials* 456 (2015) 398–408.
  - Swiler, L. P., Pastore, G., Williamson, R. L. and D. M. Perez. "Sensitivity Analysis of the Fission Gas Behavior Model in BISON." SAND Report 2013-3906.
  - Swiler, L. P., D. Mandelli, C. Rabiti, and A. Alfonsi. "DAKOTA Reliability Methods applied to RAVEN-RELAP7". SAND Report 2013-8439.
  - SAND2014-1839. Dakota Uncertainty Quantification Methods Applied to the NEK-5000 SAHEX Model, V.G. Weirs.
  - SAND2014-18550. Sensitivity Analysis of the Gap Heat Transfer Model in BISON. R. C. Schmidt, L. P. Swiler, D.M. Perez, R. L. Williamson
  - SAND Report 2015-8088 "Sensitivity Analysis of OECD Benchmark Tests in BISON."
  - L. Swiler and K. Gamble. "Initial Calibration of the Solid Swelling Factor in BISON using Dakota." NEAMS Milestone report, Sept. 2016.
  - Gamble, K. and L. Swiler. "Uncertainty Quantification and Sensitivity Analysis Applications to Fuel Performance Modeling." *TOPFUEL American Nuclear Society Conference*, Sept. 2016.
  - Delchini, M.O., Popov, E., Pointer, D., and L. Swiler. "Assessment of SFR Wire wrap simulation uncertainties." ORNL/TM-2016/540. Sept. 2016.