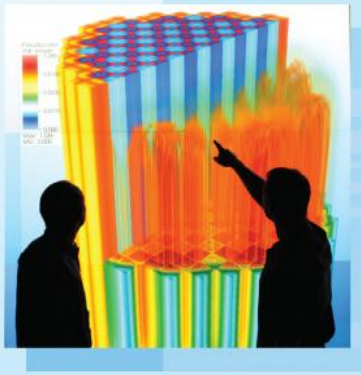
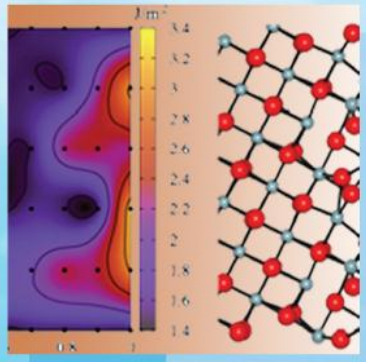
Power uprates
and plant life extension



Engineering design
and analysis

Science-enabling
high performance
computing



Fundamental science

Plant operational data

Consortium for Advanced Simulation of LWRs

L3:VUQ.SAUQ.P5.02

Jim Stewart
SNL

Completed: 6/30/12

U.S. DEPARTMENT OF ENERGY | Nuclear Energy

SANDIA REPORT

SAND2012-xxxx
Unlimited Release
Printed June 2012

Propagation of Model Form Uncertainty for Thermal Hydraulics using RANS Turbulence Models in Drekar

Michael S. Eldred and Timothy M. Wildey

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

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.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

CASL-U-2012-0080-000

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2012-xxxx
Unlimited Release
Printed June 2012

Propagation of Model Form Uncertainty for Thermal Hydraulics using RANS Turbulence Models in Drekar

Michael S. Eldred and Timothy M. Wildey
Optimization and Uncertainty Quantification Department
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1318
mseldre@sandia.gov, tmwilde@sandia.gov

Abstract

This document summarizes the results from a level 3 milestone study within the CASL VUQ effort. It demonstrates the propagation of model form uncertainty that arises from the presence of multiple turbulence models within the context of thermal hydraulics analyses. The lack of knowledge associated with an inability to *a priori* identify an appropriate turbulence model is modeled as discrete epistemic uncertainty. This approach provides an alternative to model selection processes, for use when data is unavailable or inadequate for reducing the model form uncertainty. In this case, the alternative is to propagate the model form uncertainty and report UQ results that include this epistemic uncertainty source alongside other parametric sources. The study calculates epistemic intervals on aleatory statistics for several quantities of interest, where the epistemic intervals are computed using mixed continuous-discrete optimization methods and the aleatory statistics are computed using polynomial chaos expansions. We first investigate two simple algebraic test problems with multiple model forms and then deploy the methods to the Drekar application. The Drekar study employs a set of Reynolds-averaged Navier-Stokes (RANS) turbulence models, including Spalart-Allmaras and k-ε. Results highlight the importance of efficient mixed continuous-discrete optimizers and the challenges in employing surrogate emulators within mixed domains.

CASL-U-2012-0080-000

Acknowledgment

The authors thank the Drekar team, especially Tom Smith, John Shadid, and Eric Cyr, for delivering the thermal hydraulics examples used in this study. We also thank Patty Hough for assisting with the mixed-integer evolutionary algorithms in COLINY and Keith Dalbey for assisting with Gaussian process models in Surfpack.

Contents

1 Background . . . 7
2 Interval Estimation with Discrete Variables . . . 9
3 Computational Experiments . . . 11
3.1 Rosenbrock example . . . 11
3.1.1 UQ Results: LHS . . . 11
3.1.2 UQ Results: EA . . . 11
3.1.3 UQ Results: SBGO . . . 12
3.2 Short column example . . . 13
3.2.1 UQ Results: LHS . . . 14
3.2.2 UQ Results: SBGO . . . 15
3.2.3 UQ Results: EA . . . 15
3.3 Thermal Hydraulics with Drekar . . . 16
3.3.1 UQ Results . . . 17
4 Observations . . . 19
References . . . 21

Figures

1 Example CDF ensemble. Commonly referred to as a “horsetail” plot. . . . 8
2 Convergence of MINLP approaches for Rosenbrock example. . . . 13
3 The steady-state x-velocity for typical realization computed using a RANS model in Drekar. . . . 16
4 The steady-state x-velocity for typical realization computed using a RANS model in Drekar. . . . 17
5 The profile of of the steady-state x-velocity along the outflow boundary for a typical realization computed using a RANS model in Drekar. . . . 17

Tables

1 UQ results using LHS for Rosenbrock example. . . . 12
2 UQ results using EA for Rosenbrock example. . . . 12
3 UQ results using SBGO for Rosenbrock example. . . . 12
4 UQ results using LHS for short column example. . . . 14
5 UQ results using SBGO for short column example. . . . 15
6 UQ results using EA for short column example. . . . 15
7 UQ results for Drekar channel flow problem . . . 18

1 Background

Uncertainty quantification (UQ) is the process of determining the effect of input uncertainties on response metrics of interest. These input uncertainties may be characterized as either aleatory uncertainties, which are irreducible variabilities inherent in nature, or epistemic uncertainties, which are reducible uncertainties resulting from a lack of knowledge. Since sufficient data is available for characterizing aleatory uncertainties, probabilistic methods are commonly used for computing response distribution statistics based on input probability distribution specifications. Conversely, for epistemic uncertainties, data is generally too sparse to support objective probabilistic input descriptions, leading either to subjective probabilistic descriptions (e.g., assumed priors in Bayesian analysis) or nonprobabilistic methods based on interval specifications (e.g., pure intervals or Dempster-Shafer basic probability assignments).

In this milestone, we are interested in the propagation of a mixture of aleatory and epistemic uncertainties, where the epistemic uncertainties include discrete parameterizations of model form. A common approach to quantifying the effects of mixed aleatory and epistemic uncertainties is to separate the aleatory and epistemic variables and perform nested iteration. From an algorithmic standpoint, this separation allows the use of strong probabilistic inferences where possible, while employing alternative inferences only where necessary. From a conceptual standpoint, the epistemic portion should capture uncertainty that is fully reducible (to constants) given perfect information, and the aleatory portion should capture all of the irreducible variability. To perform this separation rigorously, it may be necessary to separate aleatory and epistemic portions within a single random source, e.g. by modeling an aleatory random variable that is parameterized by an epistemic mean. Given this rigorous separation, the interpretation of the results of the nested iteration becomes straightforward: any particular epistemic realization represents a possible state of the full variability of the uncertain process. If the separation is not performed in this manner and irreducible uncertainty is hidden within the epistemic parameterization, then the results are more opaque: an epistemic realization is instead a partial expectation involving a subset of the actual random variability. Statistics derived from these partial expectations can be misleading and are easily misinterpreted; moreover, this mistake apears to be widespread due to the additional sophistication required to perform a rigorous separation.

It can be argued that model form uncertainty readily conforms to a rigorous epistemic separation in that we do not typically view it as having underlying stochastic variability; rather, in a perfect state of information and modeling, we would reduce to a single correct model form. Since, in reality, we do not have access to this perfect model form, we instead try to capture a representative range (i.e., an epistemic interval) of the possible model outcomes using the set of imperfect models that we have available. This is of course not guaranteed to bound the perfect model case, but rather provides a quantitative assessment of the uncertainty reflected within the current model form ensemble when we lack the data necessary to perform an informed model down-selection.

Traditionally, the analysis of mixed aleatory and epistemic uncertainty has involved a nested sampling approach, in which each sample drawn from the epistemic variables on the outer loop results in a sampling over the aleatory variables on the inner loop. In this fashion, we generate families or ensembles of response distributions, where each distribution represents the irreducible

aleatory uncertainty. Plotting an entire ensemble of cumulative distribution functions (CDFs) in a "horsetail" plot allows one to visualize the upper and lower bounds on the family of distributions (see Figure 1). However, nested iteration can be computationally expensive when it is implemented

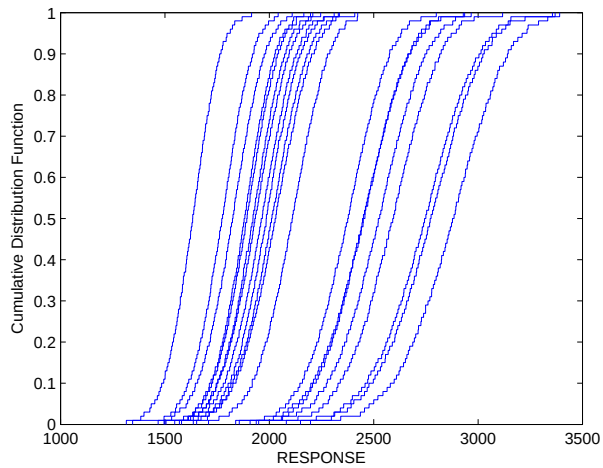


Figure 1. Example CDF ensemble. Commonly referred to as a "horsetail" plot.

using two random sampling loops. Consequently, when employing simulation-based models, the nested sampling must often be under-resolved, particularly at the epistemic outer loop, resulting in an under-prediction of credible output ranges.

In [3, 4], a central goal was to preserve the advantages of uncertainty separation while addressing issues with accuracy and efficiency by closely tailoring the algorithmic approaches to the propagation needs at each level. In particular, fast-converging stochastic expansion approaches (nonintrusive polynomial chaos and stochastic collocation) are employed for aleatory propagation, and optimization-based interval estimation is performed for the epistemic propagation. A significant amount of background information is provided in these references and is not repeated here; please refer to [3, 4] for additional information on polynomial chaos and stochastic collocation expansion methods for aleatory propagation and on interval-valued probability (IVP), Dempster-Shafer, and second-order probability (SOP) approaches for mixed UQ. In this report, we focus on the IVP formulation and build on the previous work to extend the optimization-based interval estimation approaches to include discrete epistemic parameterizations of model form.

2 Interval Estimation with Discrete Variables

In the IVP formulation, we seek the minima and maxima of aleatory statistical quantities of interest (QoI) over the epistemic parameters, where we extend the formulation in [3, 4] to include a mixture of continuous and discrete parameters:

$$\begin{aligned} \text{minimize} \quad & Q_i(\mathbf{c}, \mathbf{d}_r, \mathbf{d}_s) \\ \text{subject to} \quad & \mathbf{c}_L \leq \mathbf{c} \leq \mathbf{c}_U \\ & \mathbf{d}_L \leq \mathbf{d}_r \leq \mathbf{d}_U \\ & \mathbf{d}_s \in \mathbb{D} \end{aligned} \tag{1}$$

$$\begin{aligned} \text{maximize} \quad & Q_i(\mathbf{c}, \mathbf{d}_r, \mathbf{d}_s) \\ \text{subject to} \quad & \mathbf{c}_L \leq \mathbf{c} \leq \mathbf{c}_U \\ & \mathbf{d}_L \leq \mathbf{d}_r \leq \mathbf{d}_U \\ & \mathbf{d}_s \in \mathbb{D} \end{aligned} \tag{2}$$

where $\mathbf{Q}$ denotes the vector of QoI, $\mathbf{c}$ denotes the continuous epistemic parameters which are limited by real-valued bounds $[\mathbf{c}_L, \mathbf{c}_U]$, $\mathbf{d}_r$ denote discrete range parameters that are defined by a sequence of integers and are therefore limited by integer bounds $[\mathbf{d}_L, \mathbf{d}_U]$ on the sequence, and $\mathbf{d}_s$ denote discrete set parameters that are defined by finite sets of admissible real or integer values, one distinct admissible set per discrete parameter.

Since the QoI are nonlinear functions of the epistemic parameters in general, we require mixed-integer nonlinear programming (MINLP) solvers to compute solutions to these types of optimization problems. MINLP solvers are generally distinguished based on their support for categorial versus noncategorical discrete parameters. In the noncategorical case, the restriction to discrete values can be relaxed allowing the discrete parameters to take on continuous values during the solution process. An example of this type of approach is the branch and bound algorithm, which solves a series of continuous relaxation subproblems to compute bounds, prune branches, and ultimately arrive at a final solution which satisfies the restrictions to discrete values. When applicable, this type of approach is preferred due to computational efficiency, but it requires the ability to evaluate the model at non-discrete intermediate values. Thus, the more challenging case algorithmically is the case of categorical discrete variables, for which relaxed values cannot be simulated (e.g., 1.5 satellites in a constellation). The enumeration of model form choices generally falls into the categorical discrete variable case, such that relaxation-based algorithms are not directly applicable. MINLP for categorical variables must rely on combinatorial techniques and tends to be much more computationally expensive.

To enable MINLP with categorical variables to be applied for computationally expensive thermal hydraulics models, we leverage our capability for surrogate modeling, in particular Gaussian process (GP) modeling. In [3, 4], we utilize the prediction variance estimates of GPs to formulate approaches based on expected improvement. Continuous optimization based on maximizing

expected improvement functions from GPs is commonly known as efficient global optimization (EGO) [6], which can be a highly effective technique for balancing the competing desires to *explore* regions where little is known and *exploit* regions where promising solutions have already been found. Longer term, it is our intent to extend the EGO approach to support discrete parameters through a generalization to the expected improvement formulation.

For initial demonstration of IVP with categorical discrete parameters, we have initially focused on three algorithmic approaches for the interval estimation at the epistemic outer loop. First, as a benchmark and sanity check, we have the simple approach of employing Latin hypercube sampling [9] over the parameter set $\mathbf{c}, \mathbf{d}_r, \mathbf{d}_s$ and then reporting the observed minimum and maximum sample as the interval bounds. In this case, there is no adaptation or refinement, only a single randomly generated sample set.

Second, we can start from an initial random sample, and then evolve this population of candidate optima using a mixed-integer evolutionary algorithm. Here we apply the EAminlp solver from COLINY [5] directly to the simulation model, with no surrogate emulation.

Third, we combine the LHS and EAminlp solvers with adaptive surrogate emulation. We have implemented an adaptive surrogate-based global optimization (SBGO) approach that takes an initial LHS sample, forms a set of GP models for the epistemic QoI, and then iteratively adapts each GP while computing approximate minima and maxima using EAminlp. In particular, the SBGO approach involves the following steps:

1. Perform LHS over the range of the parameter set $\mathbf{c}, \mathbf{d}_r, \mathbf{d}_s$ and form initial GP models for the quantities of interest $\mathbf{Q}$ at the epistemic outer loop. Each epistemic realization involves an aleatory UQ propagation.

2. Apply a categorical MINLP solver (in this case, EAminlp) to minimize the GP prediction to compute an approximate minimum. This solution is tested for soft convergence in terms of change in solution $\mathbf{c}^*, \mathbf{d}_r^*, \mathbf{d}_s^*$ and change in approximate QoI minimum. If converged, skip to step 4; else, continue to step 3.

3. the approximate minimum is validated with an aleatory analysis and the GP is updated with these truth model results. Return to step 2.

4. The MINLP solver is now applied to maximize the prediction of the adapted GP. Note that the previously-adapted GP can be fully reused, only the sense of the optimization is changed. This MINLP solution is tested for soft convergence; if converged, advance to the next QoI and return to step 2 (or stop if no more QoI); else, continue to step 5.

5. the approximate maximum is validated with an aleatory analysis and the GP is updated with these truth model results. Return to step 4.

Relative to the EGO approach for successive adaptation of GP models, SBGO strictly exploits regions with good solutions, and does not support global exploration beyond the initial GP construction in step 1.

3 Computational Experiments

Two simple algebraic tests are included to demonstrate algorithm behavior and relative performance in greater detail. For these problems, distinct model forms are drawn from previously published multifidelity examples; although all fidelities are treated as model form peers in this context. We conclude with the targeted thermal hydraulics application using Drekar, demonstrating the approach for CASL-relevant physics.

3.1 Rosenbrock example

We start with a simple example of two model forms for the Rosenbrock polynomial, originally published in [2]:

$$\text{Form 1}: \quad f_1 = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \tag{3}$$

$$\text{Form 2}: \quad f_2 = 100(x_2 - x_1^2 + .2)^2 + (0.8 - x_1)^2 \tag{4}$$

where form 1 is the traditional polynomial and the small offsets in form 2 are sufficient for it to differ from form 1 in function, gradient, and Hessian values. The two variables are independent standard normals, and aleatory metrics include the mean and standard deviation of f.

If only the model form were epistemic, then the interval optimization could be performed by simple enumeration of two aleatory analyses, one for each model form. In this case, the epistemic intervals for aleatory mean and standard deviation of f using an aleatory sparse grid level of two are $[365.64, 402.00]$ and $[997.52, 1050.0]$, respectively. To make the example more interesting, we add continuous epistemic variables that define the means of the aleatory random variables x_1 and x_2. In particular, $x_1 \sim N(\mu_1, 1), x_2 \sim N(\mu_2, 1)$ with $\mu_1 \in [-1, 1], \mu_2 \in [-1, 1]$. These results are reported in the following two sections.

3.1.1 UQ Results: LHS

Table 1 shows Latin hypercube sampling results for the Rosenbrock example, again using an aleatory sparse grid level of two. It is evident that the intervals are converging (lower bounds from above and upper bounds from below), but only appear to be accurate to a few digits after extensive sampling (10^5 outer loop samples with greater than two million total evaluations).

3.1.2 UQ Results: EA

Table 2 shows results for mixed-integer evolutionary optimization with the population size set to 100 and the maximum number of population cycles for each QoI bound set to 5, 10, and 25

Table 1. UQ results using LHS for Rosenbrock example.

Outer Evals	Total Evals	μ_f	σ_f
10	210	[346.152, 1153.31]	[1005.22, 2552.71]
100	2100	[303.160, 1385.31]	[796.943, 2861.09]
1000	21000	[302.098, 1524.23]	[789.075, 3005.77]
10000	210000	[301.680, 1579.54]	[777.004, 3097.92]
100000	2100000	[301.656, 1602.87]	[776.399, 3122.45]

Table 2. UQ results using EA for Rosenbrock example.

Outer Evals	Total Evals	μ_f	σ_f
2380	44457	[302.003, 1605.00]	[778.958, 3124.36]
4360	69283	[301.998, 1605.00]	[776.052, 3124.36]
10300	99869	[301.638, 1605.00]	[776.109, 3124.36]

(corresponding to 2380, 4360, and 10300 total outer loop evaluations, respectively). It is evident that the intervals are converging more rapidly than for LHS, as would be expected when using directed optimization techniques. The EA intervals after approximately 10^5 total evaluations are more converged than those after 2.1 million total LHS evaluations. However, even after a reduction of a factor of 20, the EA function evaluation counts are still unacceptably high for use in expensive CASL applications.

3.1.3 UQ Results: SBGO

Table 3 shows results for surrogate-based global optimization for two different initial sample sizes: 10 and 20 random initial LHS points (corresponding to 26 and 47 outer loop evaluations after refinement). The first result (26 outer loop evaluations) starts from the same data as the 10 outer

Table 3. UQ results using SBGO for Rosenbrock example.

Outer Evals	Total Evals	μ_{f_1}	σ_{f_1}
26	525	[301.642, 1485.04]	[776.044, 3027.57]
47	945	[301.998, 1605.00]	[813.394, 3121.79]

loop sample case in Table 1, and it is evident that the optimization-driven refinement significantly improves the accuracy of the bounds with relatively few additional evaluations. With the exception of the lower bound for σ_f, the second result (47 outer loop evaluations) has comparable accuracy to 10^5 outer loop samples, after a reduction of a factor of 2700 in evaluation cost. These evaluation counts become much more practical for use with Drekar.

Taking a 50 iteration EA result as a reference solution, we plot the convergence rates for each of the three techniques in Figure 2. Errors are averaged over the four QoI and the two SBGO results are shown relative to their LHS initial sample starting points (10 and 20 outer loop evaluations).

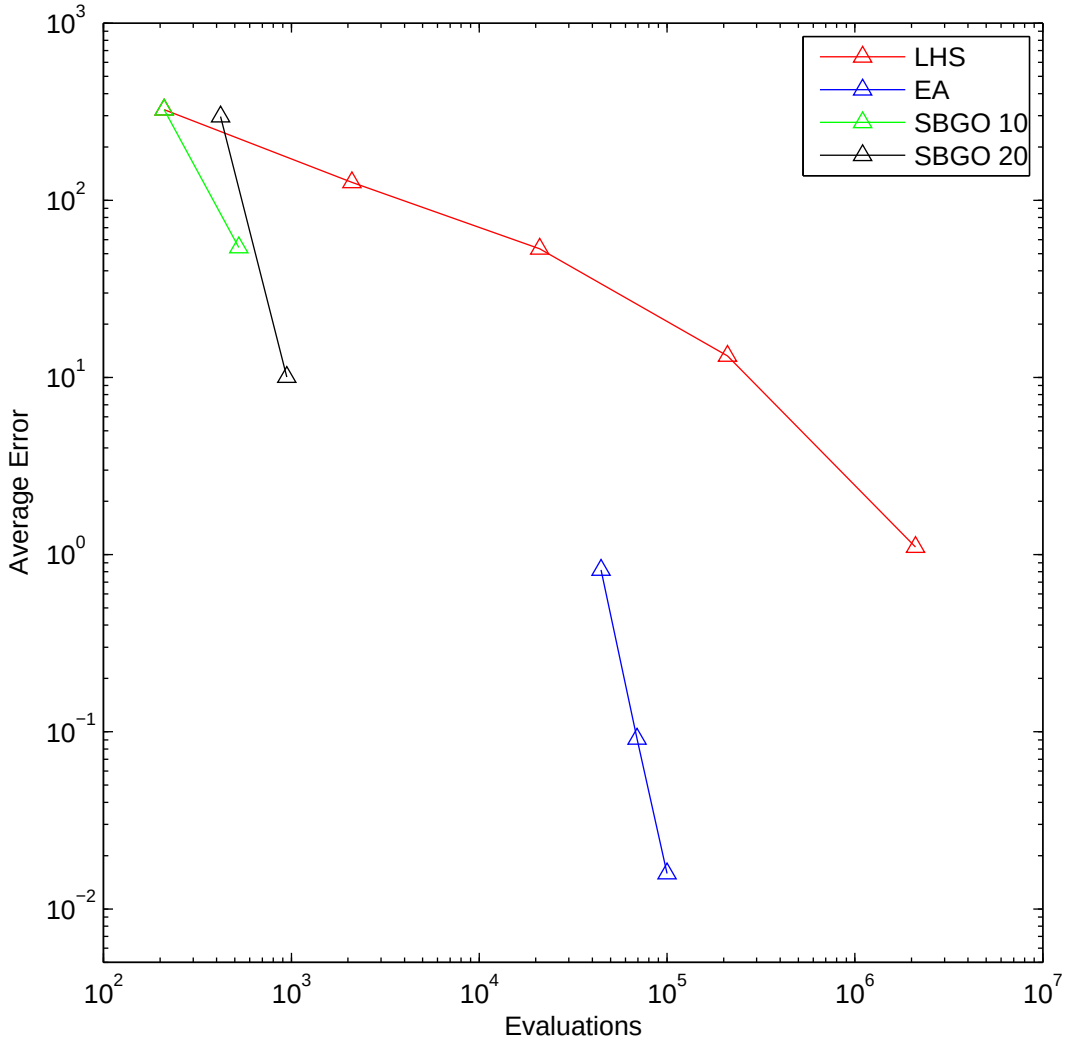


Figure 2. Convergence of MINLP approaches for Rosenbrock example.

3.2 Short column example

Relative to the previous example, this test problem increases the aleatory dimension from two to five, increases the number of model form alternatives from two to four, and replaces polynomials with rational functions. It involves the plastic analysis and design of a short column with rectangular cross section (width b and depth h) having uncertain material properties (yield stress Y) and

subject to uncertain loads (bending moment M and axial force P) [7]. The model forms for the limit state are:

$$\text{Form 1}: \quad f_1 = 1 - \frac{4M}{bh^2Y} - \left(\frac{P}{bhY}\right)^2 \tag{5}$$

$$\text{Form 2}: \quad f_2 = 1 - \frac{4P}{bh^2Y} - \left(\frac{P}{bhY}\right)^2 \tag{6}$$

$$\text{Form 3}: \quad f_3 = 1 - \frac{4M}{bh^2Y} - \left(\frac{M}{bhY}\right)^2 \tag{7}$$

$$\text{Form 4}: \quad f_4 = 1 - \frac{4M}{bh^2Y} - \left(\frac{P}{bhY}\right)^2 - \frac{4(P-M)}{bhY} \tag{8}$$

where Form 1 is the traditional form and Forms 2–4 were published in [8]. The distributions of the random variables are uniform for b and h ($[5,15]$ and $[15,25]$, respectively), normal for P and M (nominally $N(500,100)$ and $N(2000,400)$, respectively), and lognormal for Y (nominally $(\mu,\sigma) = (5,.5)$). P and M are correlated with a correlation coefficient of 0.5 (uncorrelated otherwise).

We again augment the discrete epistemic model form variable with continuous epistemic variables that define the means of the aleatory random variables P, M, and Y. In particular, $P \sim N(\mu_P,100), M \sim N(\mu_M,400), Y \sim logN(\mu_Y,0.5)$ with $\mu_P \in [400,600], \mu_M \in [1750,2250], \mu_Y \in [4,6]$. The aleatory metrics are again the mean and standard deviation of f, where the aleatory moments are evaluated using a polynomial chaos expansion with a sparse grid level of two (requiring 85 evaluations for five variables).

3.2.1 UQ Results: LHS

Table 4 shows Latin hypercube sampling results for the short column example. Three of the four

Table 4. UQ results using LHS for short column example.

Outer Evals	Total Evals	μ_f	σ_f
10	850	[-4.53457, 8.40151]	[0.306464, 4.90532]
100	8500	[-6.91096, 8.25776]	[0.292929, 6.75695]
1000	85000	[-7.16286, 8.42517]	[0.286231, 6.95294]
10000	850000	[-7.17142, 8.56947]	[0.285428, 6.95959]
100000	8500000	[-7.17298, 8.59563]	[0.285359, 6.96081]

results appear to be converging monotonically, with approximately two or three digits of accuracy after 10^5 outer loop samples. The convergence of the upper bound for μ_f is non-monotonic and its accuracy is less certain.

3.2.2 UQ Results: SBGO

Table 5 shows results for surrogate-based global optimization with initial samples set at 10 and 100 and the maximum number of refinement iterations for each QoI bound set to 10.

Table 5. UQ results using SBGO for short column example.

Outer Evals	Total Evals	μ_f	σ_f
50	3825	[-4.53457, 8.40151]	[0.285353, 4.90532]
140	11815	[-6.91096, 8.49273]	[0.292929, 6.75695]

Results for SBGO are disappointing for this problem, as the refined intervals only differ from the corresponding LHS intervals in two cases: the σ_f lower bound started from 10 LHS samples and the μ_f upper bound started from 100 LHS samples. Upon closer examination of the iteration history (not shown), it is evident that significant inaccuracy is present in the extrema prediction from the GPs, as the validations of the predicted extrema often differed in order and sign. Modifications that provided some benefit included restricting the number of model form alternatives and reordering the model forms for each QoI (aleatory mean and standard deviation, in this case) to allow monotonicity in their epistemic variation. However, none of the options investigated were fully satisfactory in taming the accuracy issues with the GP models. This difficulty provided the original motivation for adding the direct EA option to this study, in order to provide an alternative with greater robustness.

3.2.3 UQ Results: EA

Table 6 shows results for mixed-integer evolutionary optimization with the population size set to 100 and the maximum number of population cycles for each QoI bound set to 5, 10, and 25. It is

Table 6. UQ results using EA for short column example.

Outer Evals	Total Evals	μ_f	σ_f
2380	197350	[-7.17299, 8.22448]	[0.285353, 6.96081]
4360	329610	[-7.17299, 8.46808]	[0.285353, 6.96081]
10300	527825	[-7.17299, 8.60743]	[0.285353, 6.96081]

evident that direct application of the mixed-integer EA without the GP surrogate indirection results in much more reliable interval estimation. However, the total number of evaluations is prohibitive for use with thermal-hydraulics codes such as Drekar. Compared to LHS, the EA results with 10^4 outer loop samples can be seen to be more converged than those for 10^5 outer loop LHS samples. Thus, the optimization approach still demonstrates savings, albeit not nearly as significant as with effective surrogate emulation.

3.3 Thermal Hydraulics with Drekar

Fuel rods within light-water nuclear reactors are cooled from water flowing through the core. Turbulence within these fluid flows has a significant effect on cooling effectivness, yet the modeling of turbulence is not a mature science; significant uncertainty exists in turbulence model formulations and a relatively large family of candidates exist.

In this study, RANS turbulence models are explored within the framework of the Drekar simulation code. At this time, two models were viewed as being sufficiently mature and robust with respect to parameter variations for this study: the Spalart-Allmares (SARANS) model and the k-ε (KERANS) model with Neumann boundary conditions for the turbulent kinetic energy (TKE).

In this study, we consider flow in a 3D channel as shown in Figure 3. The mesh is much finer along the top and bottom boundaries to efficiently resolve the near-wall region. We enforce a pressure drop in the x-direction via a source term in the Navier Stokes equations. No-slip boundary conditions ($\mathbf{u} = \mathbf{0}$) are enforced along the top and bottom boundaries, while the remaining boundaries are chosen to be stress-free. In Figure 4 we plot the x-velocity for a typical realization computed using a SARANS model, and in Figure 5 we plot the profile of the x-velocity along the outflow boundary.

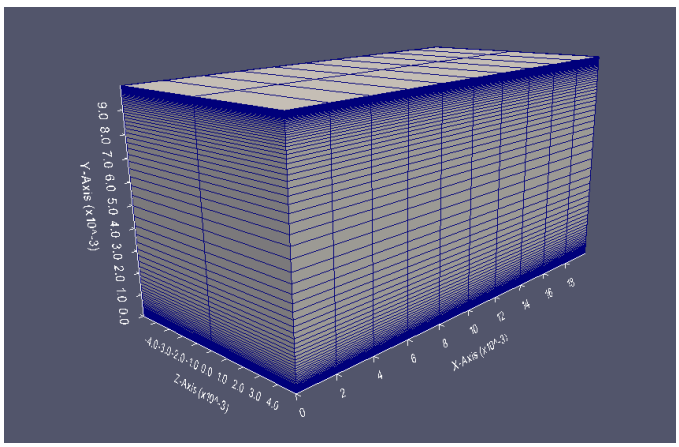


Figure 3. The steady-state x-velocity for typical realization computed using a RANS model in Drekar.

We consider the molecular viscosity (denoted by ν) and the source term for the x-momentum equation (denoted by f) to be aleatoric uniform random variables with epistemic upper bounds: $\nu \sim U(1E-5, \eta_\nu)$ and $f \sim U(10, \eta_f)$ with $\eta_\nu \in [5e-5, 5e-4]$ and $\eta_f \in [20, 40]$. Our simulation quantities of interest are the spatially averaged x-velocity and the spatially averaged pressure. The aleatory statistics of interest are the stochastic means of these simulation QoI, and we seek epistemic intervals on these aleatory statistics computed over model form, η_ν and η_f.

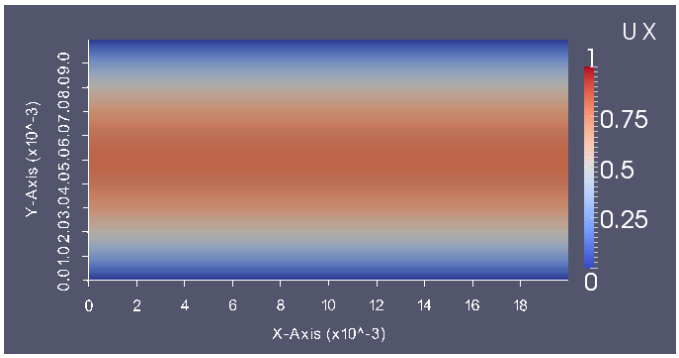


Figure 4. The steady-state x-velocity for typical realization computed using a RANS model in Drekar.

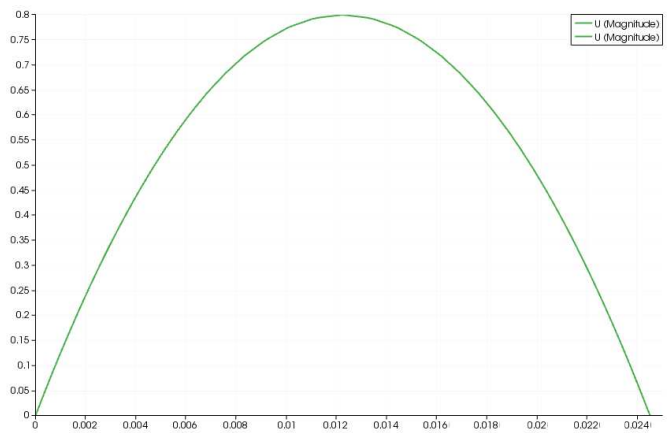


Figure 5. The profile of of the steady-state x-velocity along the outflow boundary for a typical realization computed using a RANS model in Drekar.

3.3.1 UQ Results

We start with 10 Latin hypercube samples on the epistemic outer loop, and employ 5th-order tensor quadrature (4th-order tensor polynomial chaos expansions) within each inner loop aleatory propagation. SBGO adapts this initial sample based on sequential application of the EA to a GP emulator. Table 7 summarizes the results for these two methods. As for the short column test problem, SBGO performance is disappointing: it shows improvement in two of the bounds, but fails to identify a better upper bound for μ_{ux} or a better lower bound for $\mu_{pressure}$. As for the short column test problem, accuracy issues with the predicted GP optima were again the issue. Direct application of the EA without GP emulation would be expected to provide much more accurate

Table 7. UQ results for Drekar channel flow problem

Method	Outer Evals	Total Evals	μ_{ux}	$\mu_{pressure}$
LHS	10	250	[0.727604, 2.78150]	[32.6109, 282.237]
SBGO	22	550	[0.723086, 2.78150]	[32.6109, 292.450]

intervals, but was cost prohibitive without deployment to more substantial computer resources.

4 Observations

Deployment of discrete epistemic parameterizations for propagating model form uncertainty has led to the following primary observations:

- In the algebraic examples, clear benefit was shown in utilizing optimization approaches relative to LHS sampling. In the Rosenbrock example, a cost reduction of a factor of 20 was observed for the EA, which rose to a factor of 2700 for SBGO.
- It is evident that the application of standard emulation techniques, such as Gaussian process modeling, to discrete domains is challenging. In the short column and Drekar examples, SBGO results were disappointing, leaving us the expensive option of MINLP optimization without emulation. Fortunately, internal LDRD projects and external academic collaborations have been focusing on the discrete emulation challenge, such that a number of directions are possible for improving discrete GP performance within our interval estimation techniques.
- Given an effective discrete variable GP modeling capability, extensions to the adaptive refinement machinery are also needed. Unlike the EGO-based approach for continuous epistemic parameters, the SBGO algorithm for mixed epistemic parameters is purely exploitative and lacks any adaptive refinement pressure for exploring regions that may contain good solutions. The SBGO approach must rely on the initial LHS sample for exploration of the space. For this reason, SBGO was not as reliable in identifying global optima as EGO is in [3, 4], presumably due to cases in which the region of the global extrema are not adequately explored. Extension of the EGO-based approach to support discrete epistemic parameters is of interest in future work.
- In this study, we have enumerated different model form options from a single modeling source. However, it is for more complex multi-component simulations (e.g., multiphysics simulation of reactor cores) for which manual enumeration becomes impractical and the automation and greater efficiency within these approaches can provide the greatest benefit. To fully realize this potential, additional improvements such as those discussed above are needed to tackle this combinatorial growth in complexity.

In addition, the following algorithmic details are noted:

- The order of GP trend functions was an issue when the number of model alternatives was small. In particular, use of a quadratic trend was problematic when only two model alternatives were present.
- EA performance was not carefully optimized for direct application without emulation. For example, initial population sizes were large and could be better optimized for efficiency.

Given an effective emulator-based IVP capability, the following additional directions could be pursued:

- The gradient-enhanced kriging (GEK) capability could be employed at the outer epistemic loop to incorporate sensitivies of the aleatory statistics with respect to the continuous epistemic parameters, as enabled by the underlying stochastic expansions [1, 4].

- Extension from IVP to Dempster-Shafer is straight-forward, but has not been explored in this work. Greater efficiency per cell is needed before supporting combinatorial growth in cell counts.

- The Drekar::CFD team expressed interest in exploring the intervals of the statistic for the case where the parameters of each turbulence model are uncertain. This can be explored in the future using the tools developed in this study.

References

[1] M. S. Eldred. Design under uncertainty employing stochastic expansion methods. *International Journal for Uncertainty Quantification*, 1(2):119–146, 2011.

[2] M. S. Eldred, A. A. Giunta, and S. S. Collis. Second-order corrections for surrogate-based optimization with model hierarchies. In *Proceedings of the 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Albany, NY,, Aug. 30–Sept. 1, 2004. AIAA Paper 2004-4457.

[3] M. S Eldred and L. P. Swiler. Efficient algorithms for mixed aleatory-epistemic uncertainty quantification with application to radiation-hardened electronics. part 1: Algorithms and benchmark results. Technical Report SAND2009-5805, Sandia National Laboratories, Albuquerque, NM, 2009.

[4] M. S. Eldred, L. P. Swiler, and G. Tang. Mixed aleatory-epistemic uncertainty quantification with stochastic expansions and optimization-based interval estimation. *Reliability Engineering and System Safety*, 96(9):1092–1113, 2011.

[5] W. E. Hart. The coliny project. Web site, 2007. `http://software.sandia.gov/Acro/Coliny/`.

[6] D. Jones, M. Schonlau, and W. Welch. Efficient global optimization of expensive black-box functions. *J. Global Optimization*, 13:455–492, 1998.

[7] N. Kuschel and R. Rackwitz. Two basic problems in reliability-based structural optimization. *Math. Method Oper. Res.*, 46:309–333, 1997.

[8] L. W. T. Ng and M. S. Eldred. Multifidelity uncertainty quantification using nonintrusive polynomial chaos and stochastic collocation. In *Proceedings of the 14th AIAA Non-Deterministic Approaches Conference*, number AIAA-2012-1852, Honolulu, HI, April 23-26, 2012.

[9] L. P. Swiler and G. D. Wyss. A user's guide to Sandia's latin hypercube sampling software: LHS UNIX library and standalone version. Technical Report SAND04-2439, Sandia National Laboratories, Albuquerque, NM, July 2004.

DISTRIBUTION:

1	MS 1318	Brian M. Adams, 1441
1	MS 1318	Keith Dalbey, 1441
3	MS 1318	Michael S. Eldred, 1441
1	MS 1318	James R. Stewart, 1441
3	MS 1318	Timothy M. Wildey, 1441
1	MS 1320	Eric C. Cyr, 1426
1	MS 1320	Thomas M. Smith, 1442
1	MS 1321	John N. Shadid, 1444
1	MS 9159	Patricia D. Hough, 8954
1	MS 0899	Technical Library, 9536 (electronic copy)

Sandia National Laboratories

CASL-U-2012-0080-000