SAND2015-3214R

Whitepaper prepared for DOE Workshop on Integrated Simulations for Magnetic Fusion Energy Sciences
Topic E: Beyond interpretive simulations No oral presentation
Prepared by Laura Swiler (lpswile@sandia.gov), Michael Eldred, and John Shadid, Sandia National Labs[1].

Motivation: Predictive simulations of magnetic confinement fusion and burning plasmas will require the quantification of all uncertainties and errors relating to the simulation capabilities. These include: discretization error (temporal and spatial); incomplete convergence error (nonlinear, linear, etc.); uncertainties in input data (initial conditions, boundary conditions, coefficients, thermo-physical properties, source terms, etc.); and uncertainties in the component models (the specific form and parameter values).[13] Typically, there is a focus on model parameter uncertainties but uncertainties about which model to use and about the bias or discrepancy of a particular model (sometimes called model form error) often dominate parameter uncertainty. This whitepaper addresses the challenges of performing uncertainty quantification (UQ) for expensive computational models.

Approach: Uncertainty can be classified as *aleatory uncertainty* and *epistemic uncertainty.*[7,14] Aleatory uncertainty describes inherent randomness or irreducible variability. Epistemic uncertainty is used to characterize lack of knowledge which could be reduced by additional data. Aleatory uncertainty is commonly estimated and propagated through computational models by probabilistic methods such as sampling methods[8] (Monte Carlo, quasi-Monte Carlo, Latin Hypercube techniques, etc.) and by direct propagation methods (stochastic Galerkin, stochastic collocation, Taylor expansion, etc.).

Sampling methods have general applicability, but can often converge very slowly. However, recent advances focus on ways to preferentially place samples on regions of interest or failure regions, often adaptively, to reduce the number of samples. These methods are designed to handle nonlinear, multimodal failure surfaces by creating global approximations (e.g. surrogate models) and adaptively refining the approximation in the vicinity of a particular response threshold. Some examples are the Efficient Global Reliability Analysis method (EGRA)[2], in which a Gaussian process model is adaptively refined to accurately resolve a particular contour of a response function, and GPAIS (Gaussian Process Adaptive Importance Sampling)[3]. The GPAIS method uses an iterative process to construct a sequence of improving component densities which approximate the importance density for importance sampling calculations. Another class of smarter sampling methods is called k-d darts and is based on ideas in computational geometry[4]. In this approach, samples are no longer “points”: they may be higher-dimensional structures such as lines or planes. Although much progress has been made to with these reliability methods and adaptive sampling, there still are many research challenges, the main one being how to scale these approaches to higher dimensions.

Direct propagation methods often converge much faster than sampling but are more intrusive. However, there are versions of these methods which are non-intrusive.[6] For example, polynomial chaos expansions (PCE) can be built by evaluating a black-box function at tensor product quadrature points, at cubature points, at sparse grid points and using a spectral projection approach to calculate the PCE coefficients.[1,16] Alternatively, one can use a variety of regression approaches to estimate PCE coefficients from a set of Monte Carlo samples, including least-squares, orthogonal matching pursuit, or least absolute shrinkage methods. The latter two methods are examples of compressive sensing techniques which are able to identify a reduced basis set from a large number of candidate bases. This is an important feature for non-intrusive stochastic expansion.[1] Additionally, there has been significant work on methods which further reduce the number of function evaluations required, through adaptive sparse grids. Adaptive schemes have been developed that utilize a variety of metrics to identify

1. 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.

important dimensions, such as the Sobol' sensitivity indices of the PCE[11], the spectral coefficient decay rates, or generalized sparse grids. Again, a major research challenge is to scale these methods to higher dimensions and to identify dominant basis sets from a limited number of simulation runs.

A major focus of the FES program is to understand how repeated disruptive events such as high current plasma contacting a tokamak wall will cause failure, and to develop robust designs to better handle such disruptive events. The adaptive sampling and PCE methods outlined above can be used to evaluate failure probabilities within an Optimization under Uncertainty (OUU) framework. The goal might be to determine optimal design parameters which maximize an expected output subject to minimizing the variance of that output. Such OUU formulations lead to designs which are more robust than deterministic optimization. Many research challenges remain, including methods to overcome the cost of performing a full UQ analysis as each point in design space. Stochastic expansion methods can be more efficient than sampling in OUU, since analytic forms for the statistics and their derivatives are available from the expansions, which facilitates gradient-based optimization.[5] OUU approaches have also been used to address mixed UQ problems which have both epistemic and aleatory uncertainty. [5]

A final area of interest is mixed-fidelity UQ. Often, one would like to use many function evaluations from a low-fidelity model and just a few high fidelity model evaluations to make a high-fidelity prediction with some level of accuracy. For example, in plasma physics there is a rich hierarchy of models, from relatively cheap magnetohydrodynamics (MHD) models to computationally expensive methods like particle-in-cell (PIC) techniques. One goal might to generate a PIC-scale prediction based on many MHD runs and a few PIC runs. The high-fidelity model is modeled as a sum of the low-fidelity model plus a discrepancy term. Various approaches have been used in multi-fidelity UQ, such as representing both the low fidelity model and the discrepancy term with stochastic expansions.[1,10] However, much work remains to be done. Another area of interest is relating a simulation model to experimental data. In this case, the model discrepancy is the difference between these terms. It has been modeled using Gaussian process models, for example.[9] However, recent work proposes an embedded modeling of the discrepancy.[12] Finally, different model forms for a single model (e.g. material model A, B, or C which are "peer models") are being modeled with categorical variables[15] and investigated using optimal model selection methods.

Research Challenges for UQ methods for fusion energy simulations:

1. Scaling UQ methods to handle higher input dimensions but minimize the samples required. Potential approaches include dimension reduction, compressive sensing, adaptive methods, etc.
2. Developing OUU formulations that handle the specifics of fusion energy problems such as robust design optimization. For example, one challenge is handling the cumulative effects of disruptive events over time, which is not currently treated in single-event failure calculations.
3. Multi-fidelity UQ approaches: the analytic form of the discrepancy term, the calculation of discrepancy in a non-intrusive vs. embedded manner, the identification of the number of low vs. high fidelity runs needed to maintain predictive accuracy, and selection of optimal model choices (hierarchical or peer models) are all significant research challenges.

Impact: Experimentally validated and integrated predictive capabilities will require validation of coupled multi-physics phenomena. These validation efforts rely on uncertainty characterization at all levels. The UQ methods outlined in this whitepaper are part of a larger effort which will involve sensitivity analysis, a posteriori error analysis, model calibration, adjoint-based methods, etc. The overall goal is to create a framework where one can identify and "trade-off" the numerical errors, parameter uncertainties, and model discrepancy errors to produce more accurate and efficient large-scale computational simulations of fusion device physics.

References:

1. Adams, B. M., Ebeida, M.S., Eldred, M.S., Jakeman, J.D., Swiler, L.P., Stephens, J.A., Vigil, D.M., Wildey, T.M., Bohnhoff, W.J., Dalbey, K.R., Eddy, J.P., Hu, T., Bauman, L.E., Hough, P.D., (2014), "Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.1 Theory Manual", Sandia National Laboratories, Report No. SAND2014-4253, Albuquerque, NM.

2. Bichon, B.J., M. S. Eldred, L. P. Swiler, Mahadevan, S., J. M. McFarland, (2008), "Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions," AIAA Journal, 46(10), pp. 2459-2468.

3. Dalbey, K. and L. P. Swiler. "Gaussian Process Adaptive Importance Sampling." *International Journal for Uncertainty Quantification*, Vol. 4(2). 2014. pp. 133-149.

4. Ebeida, M., Patney, A., Mitchell, S., Dalbey, K., Davidson, A. and J. Owens. k-d Darts: Sampling by k-dimensional flat searches. ACM Transactions on Graphics (TOG) 33 (1), 2013.

5. Eldred, M.S., Swiler, L.P., and Tang, G., "Mixed Aleatory-Epistemic Uncertainty Quantification with Stochastic Expansions and Optimization-Based Interval Estimation," *Reliability Engineering and System Safety (RESS)*, Vol. 96, No. 9, Sept. 2011, pp. 1092-1113.

6. Eldred, M.S., Webster, C.G., and Constantine, P.G., "Evaluation of Non-Intrusive Approaches for Wiener-Askey Generalized Polynomial Chaos," paper AIAA-2008-1892 in *Proceedings of the 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference (10th AIAA Non-Deterministic Approaches Conference)*, Schaumburg, IL, April 7-10, 2008.

7. Helton, J.C. Conceptual and Computational Basis for the Quantification of Margins and Uncertainty. Sandia Technical Report 2009-3005.

8. Helton, J.C. and F.J. Davis. Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. Reliability Engineering and System Safety 2003;81(1):23-69.

9. Kennedy, A., and O'Hagan, A., (2001), "Bayesian Calibration of Computer Models," Journal of Royal Society, 63(3), pp. 425-564.

10. Ng, L.W.T. and Eldred, M.S., "Multifidelity Uncertainty Quantification Using Nonintrusive Polynomial Chaos and Stochastic Collocation," paper AIAA-2012-1852 in *Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference (14th AIAA Non-Deterministic Approaches Conference)*, Honolulu, Hawaii, April 23-26, 2012.

11. Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models.* New York: Wiley; 2004.

12. Sargsyan, K., Najm, H. N. and Ghanem, R. (2015), On the Statistical Calibration of Physical Models. Int. J. Chem. Kinet., 47: 246–276. doi: 10.1002/kin.20906

13. Scientific Grand Challenges: 2009 Report on Fusion Energy Sciences and the Role of Computing at the Extreme Scale. Report prepared as an account of a U.S. Department of Energy workshop, Chair W. Tang, Co-chair D. Keyes. Report available at: http://science.energy.gov/fes/news-and-resources/workshop-reports/

14. Swiler, L.P. and A.A. Guinta, "Aleatory and Epistemic Uncertainty Quantification for Engineering Applications," Sandia Technical Report SAND2007-2670C.

15. Swiler, L.P, P. D. Hough, P. Qian, Xu Xu, C. Storlie, H. Lee, "Surrogate Models for Mixed Discrete-Continuous Variables" in Constraint Programming and Decision Making.   *Studies in Computational Intelligence Series*, Volume 539, 2014, pp. 181-202. Springer.

16. Xiu, D. *Numerical Methods for Stochastic Computations: A Spectral Method Approach.* Princeton University Press, 2010.