UNCERTAINTY QUANTIFICATION USING RESPONSE SURFACE APPROXIMATIONS

A. A. Giunta, M. S. Eldred, and J. P. Castro
Sandia National Laboratoires, Albuquerque, NM 87185*
aagiunt@sandia.gov, mseldre@sandia.gov, jpcastr@sandia.gov

Abstract

This report describes an initial investigation into the error convergence trends in sampling-based uncertainty quantification (UQ) studies performed both with and without response surface approximations. The data provided by this limited study indicate that RS-based UQ methods exhibit error trends that are as good or better (converging faster to zero) when compared to conventional sampling-based UQ methods.

Introduction

An uncertainty quantification (UQ) study using traditional sampling-based approaches (e.g., Monte Carlo sampling and its variants) can be prohibitively expensive when applied to a high-fidelity simulation code that requires hours of supercomputer time for a single code execution. One approach to this problem is to employ response surface (RS) approximation methods to create a mathematical model of the high fidelity simulation code output data, and then perform UQ sampling on the computationally inexpensive mathematical model. While this approach is useful, it is important to note that it introduces additional error in the estimation of the UQ statistical metrics beyond the error created by traditional sampling. There are numerous choices for the sampling method that is used to generate the data for the RS approximation, as well as numerous choices for the mathematical form of the RS approximation. There are no clear guidelines in the statistical literature on how to best choose the sampling method and the RS approximation type. This study is an initial attempt to gain insight into when the RS-based UQ approach is more accurate than direct Monte Carlo sampling, and which combination of sampling method and RS approximation method yields the most accurate UQ results.

This study investigates three types of sampling methods to generate data to build the RS approximations: Monte Carlo (MC) sampling, Latin hypercube (LH) sampling, and orthogonal array (OA) sampling, along with two types of RS approximation methods: kriging interpolation and multivariate adaptive regression splines (MARS). A single test function is used in this study to provide some insight into the utility of the RS-based UQ approach. More extensive testing will follow to provide more definitive results.

---

* Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

It is expected that for a low number of data samples the RS-based UQ sampling methods will provide more accurate estimates of the test function mean value than direct sampling of the test function. As the number of samples increases, it is expected that there will be less of an advantage of the RS-based UQ approach over direct sampling.

Technical Approach

Test Function

Figure 1 shows a plot of the Rosenbrock function (Gill et al, 1981) that is used in this study since it provides an algebraic, nonlinear response function that exhibits some of the nonlinear trends often found in data from computationally expensive engineering simulation codes. The Rosenbrock function is:

$$f(x_1,x_2) = 100(x_2-x_1^2)^2 + (1-x_1)^2, \tag{1}$$

where, for the purposes of this UQ study, x_1 and x_2 are independent and uniformly distributed on the interval [-2, 2]. The mean value of $f(x_1,x_2)$ over the region $[-2, 2]^2$ is 455.667, and can be obtained analytically. When an explicit response function is not available, the mean value of the function can be estimated using data values via:

$$\tilde{\mu} = \frac{1}{N}\sum_{i=1}^{N} f(x_1, x_2), \tag{2}$$

where N is the number of samples. When the mean value is estimated using samples taken from the RS approximation, the mean value is denoted as:

$$\hat{\mu} = \frac{1}{M}\sum_{i=1}^{M} \hat{f}(x_1, x_2), \tag{3}$$

where $\hat{f}$ is the surface approximation function computed from N data points, and the mean value is estimated by sampling the approximation function at M points ($M>>N$).

Sampling Methods

Monte Carlo, Latin hypercube (McKay et al, 1979), and orthogonal array (Koehler and Owen, 1996) sampling methods involve an element of random sampling. However, the LH and OA sampling methods partition the parameter space into bins of equal probability, with the goal of attaining a more even distribution of sample points in the parameter space than typically occurs with MC sampling. Given N sample points to locate in an n-dimensional parameter space, LH sampling partitions the parameter space into an N^n grid of bins, with N bins along each axis. Then, sample points are distributed such that all one-dimensional projections of the samples yield one sample per bin. The OA sampling method is similar to the LH sampling approach, but in OA sampling the parameter space is partitioned into a $\left(\sqrt[t]{N}\right)^n$ grid of bins, and the samples are distributed such that all t-dimensional projections ($t<n$) of the samples yield at least one sample per bin. For the test function used in this study, the “strength 2” (i.e., t=2) OA method places one sample in each bin in the parameter space. However, this would not be the case for a

general *n*-dimensional test function. See the paper by Giunta et al. (2003) for additional detail on LH and OA sampling methods.

Surface Approximation Methods

Kriging interpolation (Cressie, 1991) and MARS (Friedman, 1991) are approximation methods that are intended to model arbitrary surfaces. Thus, they are well suited for modeling the nonlinear trends in the Rosenbrock function. The kriging method employs Gaussian basis functions, with all correlation parameters set to a value of unity. The kriging surface approximation model exactly interpolates all data points. The MARS method employs a combination of regression functions and cubic splines. It is not guaranteed to identically interpolate the data points, but in practice, the MARS surface approximation is very close to the data. The kriging and MARS surface approximation methods used in this study are those available in the DAKOTA version 3.1 software toolkit (Eldred et al, 2001).

Results and Discussion

Figure 2 illustrates the convergence of the estimated mean value of the function, as the number of samples increases. Note that the MC and LH sample sizes were 10, 25, 50, 75, 100, and 121, while the OA sample sizes were 9, 25, 49, and 121. The special bin structure required by the OA limits the sample sizes that can be generated with this method. Equation 2 was used to compute the mean values in this plot. In Figure 2, there are 10 replicates for each mean value estimate. The data points for the MC, LH, and OA methods have been slightly shifted along the horizontal axis to facilitate viewing of the data. Overall, the trend shown in Figure 2 is exactly as expected, i.e., that the LH and OA methods converge more quickly to the true mean value of the test function than MC sampling. Qualitatively, it appears that the OA data have as good or better (i.e., lower) dispersion than the LH data, but this trend is not pronounced until reaching 121 samples.

Figure 3 compares the convergence trends of the mean value computed using LH sampling on the test function (i.e., Equation 2) versus sampling on the surface approximation (i.e., Equation 3). In the case of the kriging and MARS data, the surface approximation was constructed using the number of samples listed on the horizontal axis, and then the mean value was computed from 10,000 LH samples of the approximation function. Clearly, using a kriging surface or a MARS surface provides a more accurate mean value estimate than can be obtained from the LH samples alone. The convergence rate of the kriging and MARS data points appears to be roughly the same.

Figure 4 shows the convergence trend of the OA sampling method versus the trends of the kriging and MARS surface approximations. As was seen for the LH data in Figure 3, in Figure 4 there is equivalent or lower dispersion in the mean value computed from the kriging and MARS surface approximations than from the OA samples alone. [Note that MARS-OA data is not available for N=9 samples due to a software problem.]

Figures 5 and 6 examine the interaction between the sampling methods and the surface approximation methods. In the MARS data shown in Figure 5, the OA sampling method appears to have slightly lower dispersion than the LH sampling method. However, the differences may not be statistically significant. The kriging data shown in Figure 6 are even more ambiguous in that the OA data points and LH data points have little or no discernable variation in convergence trend.

Summary

This study provides some initial insight into the tradeoffs among the choices for sampling methods and surface approximation methods in RS-based uncertainty quantification. While it would be premature to draw conclusions on the basis of a single low-dimensional test problem, it is worth noting the trends that were observed in these tests.

1. Using test function sample data to first build a response surface approximation (either kriging or MARS) and then sampling the surface, produced as good or lower errors in mean value estimates versus using the original sample data alone.

2. Not surprisingly, the OA sampling method generated slightly lower errors than did LH sampling. This trend was more pronounced when the original MC, LH, and OA mean value data were compared, while the trend was less pronounced when response surface approximations were constructed from the data prior to sampling. The advantage of OA sampling over MC and LH sampling will increase as the dimensionality of the test problem grows. However, the OA sample size rules place limits on the utility of the OA method.

Future work will expand on this initial study to include test problems having dimension size $n>2$. In addition, other sampling techniques, such as quasi-Monte Carlo methods, will be investigated, along with other response surface approximation methods, such as radial basis functions and artificial neural networks.

References

Cressie, N. (1991), *Statistics of Spatial Data*, John Wiley and Sons, New York, NY.

Eldred, M.S., Giunta, A.A., van Bloemen Waanders, B.G., Wojtkiewicz, S.F., Jr., Hart, W.E. and Alleva, M.P. (2001), *DAKOTA Users Manual: Version 3.1*, Sandia Technical Report SAND2001-3796, Sandia National Laboratories, Albuquerque, NM. (http://endo.sandia.gov/DAKOTA/software.html)

Friedman, J.H. (1991), ‘‘Multivariate Adaptive Regression Splines,’’ *Annals of Statistics*, Vol. 19, No. 1, pp. 1-141.

Gill, P.E., Murray, W., and Wright, M.H. (1981), *Practical Optimization*, Academic Press, San Diego, CA.

Giunta, A.A., Wojtkiewicz, S.F., Jr., and Eldred, M.S. (2003), "Overview of Modern Design of Experiments Methods for Computational Simulations," paper AIAA-2003-0649 in Proceedings of the 41st AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV.

Koehler, J.R., and Owen, A.B. (1996) “Computer Experiments,” in *Handbook of Statistics* (eds. S. Ghosh and C.R. Rao), Vol. 13, Elsevier-Science, pp. 261-308.

McKay, M.D., Beckman, R.J., and Conover, W.J. (1979), “A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code,” *Technometrics*, Vol. 21, No. 2, pp. 239-245.

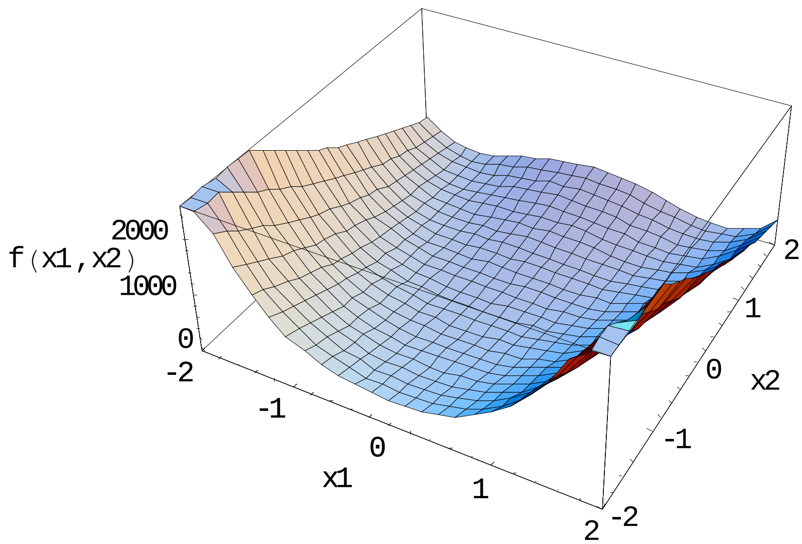

Figure 2. Rosenbrock's function.

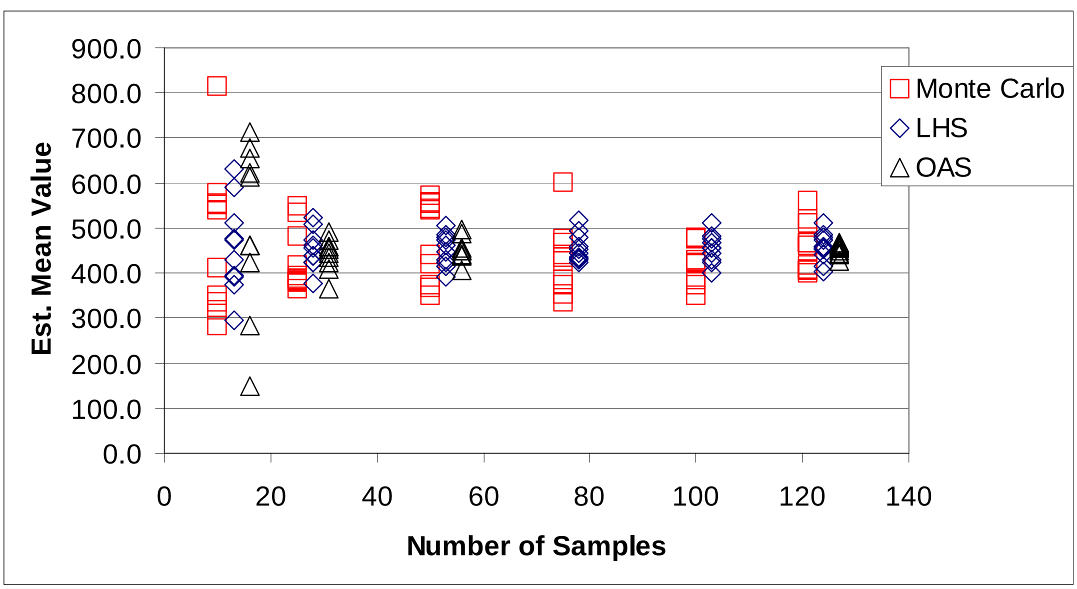

Figure 1. Convergence rate comparison in mean value estimation of Rosenbrock’s function using Monte Carlo, Latin Hypercube, and Orthogonal Array sampling.

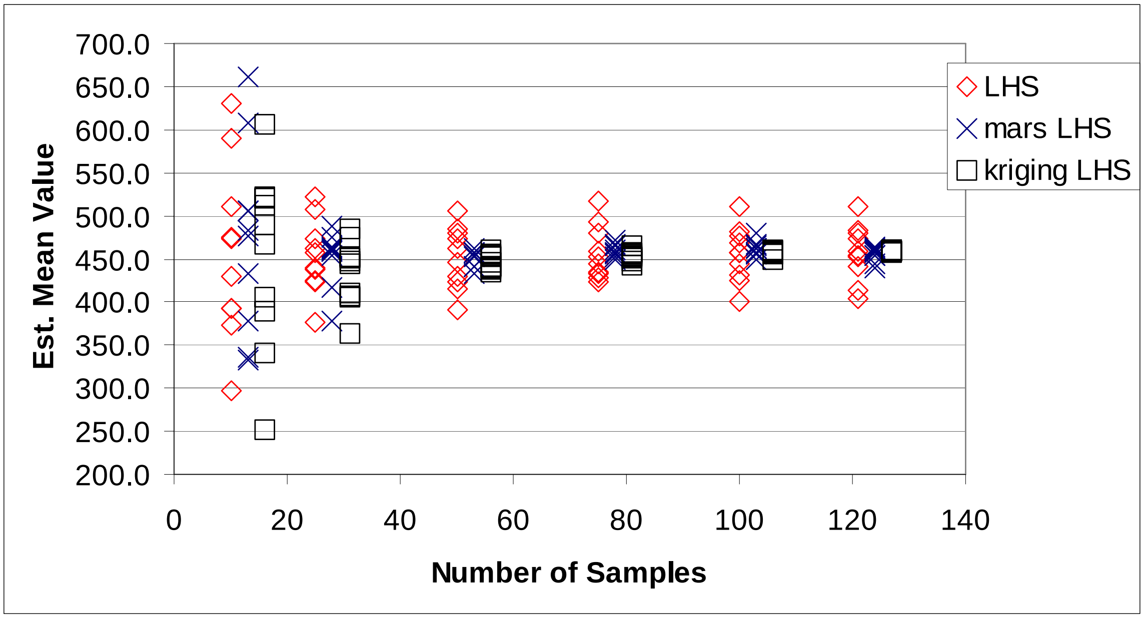

Figure 3. Convergence rate comparison of LH sampling of the Rosenbrock function versus LH sampling of response surface approximations to the Rosenbrock function.

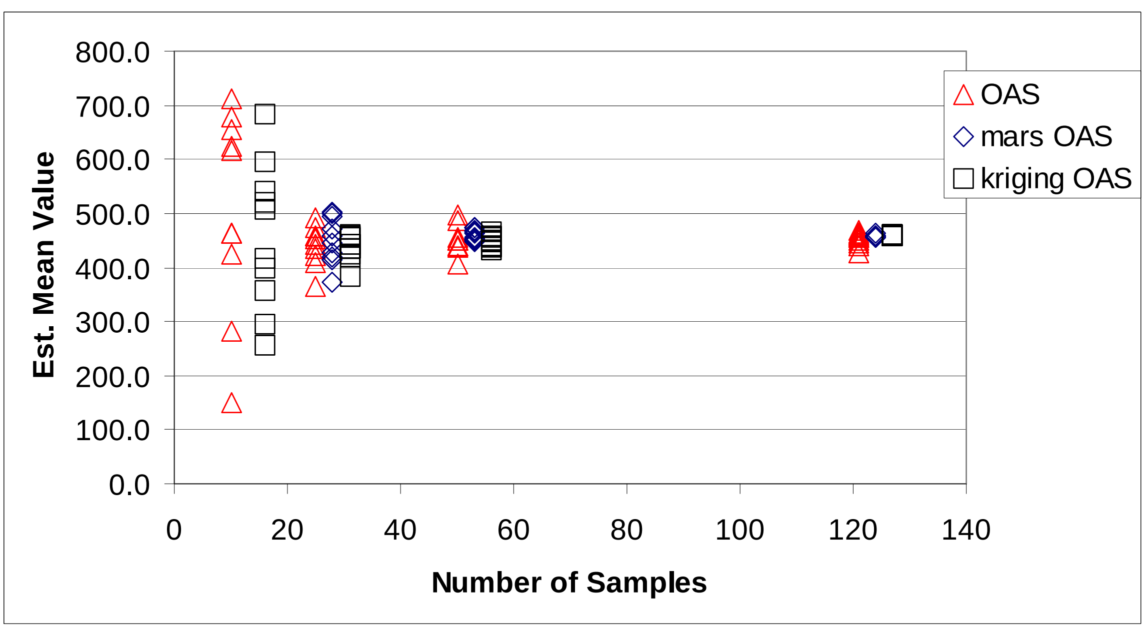

Figure 4. Convergence rate comparison of OA sampling of the Rosenbrock function versus OA sampling of response surface approximations to the Rosenbrock function.

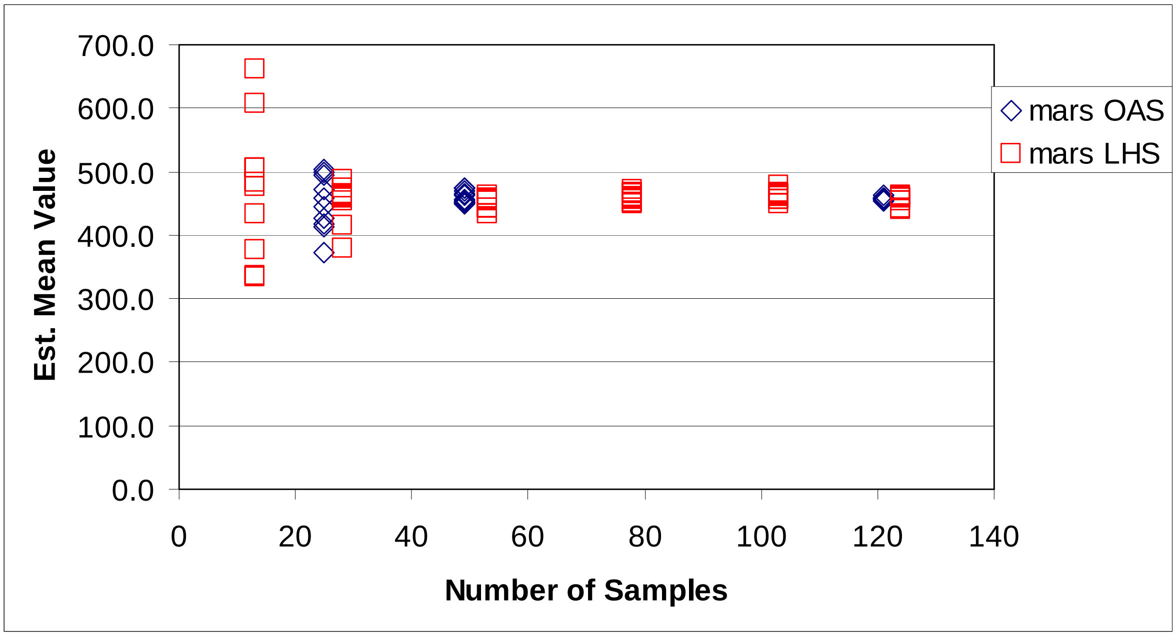

Figure 5. Convergence rate comparison using the MARS surface approximation method, with the MARS surface constructed using either OA or LH samples.

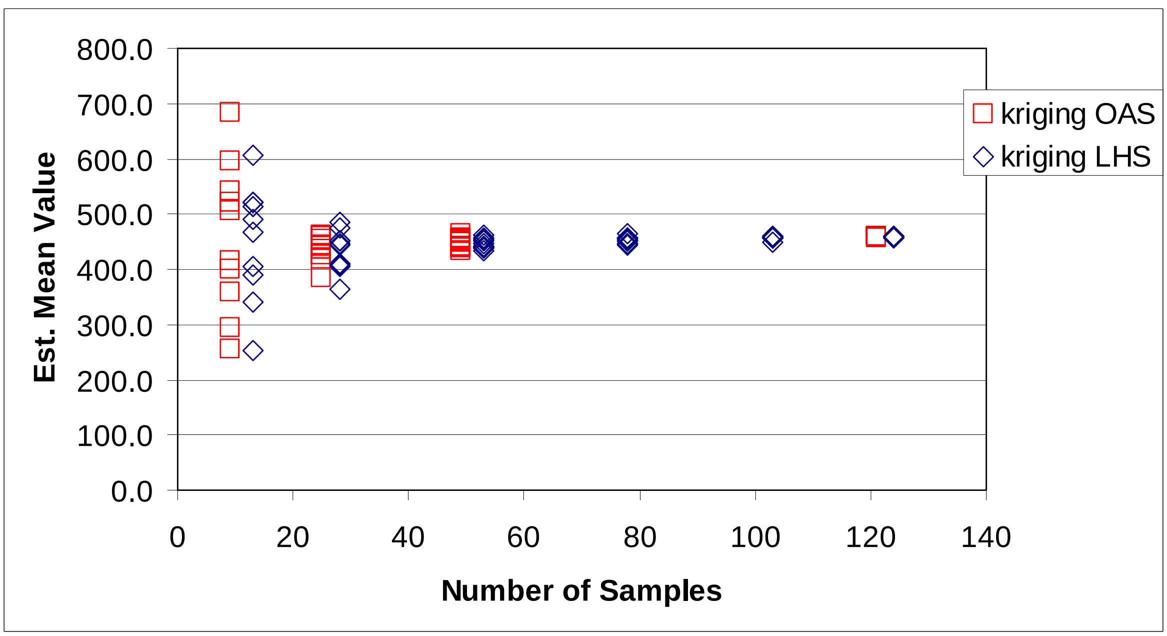

Figure 6. Convergence rate comparison using the kriging surface approximation method, with the kriging surface constructed using either OA or LH samples.