PROGRESSIVE RESPONSE SURFACES

V. J. Romero, T. Krishnamurthy, and L. P. Swiler
Sandia National Laboratories[*], *Albuquerque, NM 87185-0825*
vjromer@sandia.gov, t.krishnamurthy@nasa.gov, lpswile@sandia.gov

Abstract

Response surface functions are often used as simple and inexpensive replacements for computationally expensive computer models that simulate the behavior of a complex system over some parameter space. "Progressive" response surfaces are ones that are built up progressively as global information is added from new sample points in the parameter space. As the response surfaces are globally upgraded based on new information, heuristic indications of the convergence of the response surface approximation to the exact (fitted) function can be inferred. Sampling points can be incrementally added in a structured fashion, or in an unstructured fashion. Whatever the approach, at least in early stages of sampling it is usually desirable to sample the entire parameter space uniformly. At later stages of sampling, depending on the nature of the quantity being resolved, it may be desirable to continue sampling uniformly over the entire parameter space (Progressive response surfaces), or to switch to a focusing/economizing strategy of preferentially sampling certain regions of the parameter space based on information gained in early stages of sampling (Adaptive response surfaces). Here we consider Progressive response surfaces where a balanced indication of global response over the parameter space is desired. We use a variant of Moving Least Squares to fit and interpolate structured and unstructured point sets over the parameter space. On a 2-D test problem we compare response surface accuracy for three incremental sampling methods: Progressive Lattice Sampling; Simple-Random Monte Carlo; and Halton Quasi-Monte-Carlo sequences. We are ultimately after a system for constructing efficiently upgradable response surface approximations with reliable error estimates.

Introduction and Background

Large-scale optimization and uncertainty analyses are often made feasible through the use of response surfaces as surrogates for computational models that may not be directly employable because of prohibitive expense and/or noise properties and/or coupling difficulties in multidisciplinary analysis. Examples of response surface usage to facilitate large-scale optimization and uncertainty analyses are cited in Roux *et al.* (1996), Unal *et al*. (1996), and Venter *et al.* (1996).

Two issues that arise when using response surface approximations (RSA) are accuracy and the cost of procuring the data samples needed to create the RSA. With a sufficiently flexible global fitting/interpolating function over the parameter space, response surface accuracy generally increases as the number of data points increases (if the points are appropriately placed throughout the parameter space), until the essential character of the function is effectively mapped out. Thereafter, it is not cost effective to continue adding samples. Since a single high-fidelity physics simulation (*i.e*., one data sample) can take many hours to compute, it is desirable to minimize the number of simulations that are needed to construct an accurate response surface.

For our purposes here it is assumed that: 1) the computer model is relatively expensive to evaluate; 2) the parameter space is a unit hypercube or can be accurately and inexpensively mapped into one; 3) the sampled or "target" function is a continuous, deterministic function over the parameter space; 4) reasonably general, arbitrary target functions are to

[*]Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

be fitted; and 5) approximate response values are desired over the entire parameter space or subspace being considered –*i.e*., for global and local optimization or mapping inputs to outputs in uncertainty propagation.

Given these specifications, Romero et al. (2000) examined several formulations for constructing progressive response surfaces built on Progressive Lattice Sampling (PLS) incremental sampling designs. PLS is a paradigm for structured uniform sampling of a hypercube parameter space by placing and incrementally adding sets of samples such that all samples are efficiently leveraged as the design progresses from one level to the next. Figures 1 - 4 show successive PLS levels in 2-dimensions. (Also shown for comparison are point sets from classical simple-random Monte Carlo sampling -using three different seeds for the random-number generator, and from Halton “quasi- Monte Carlo” low-discrepancy sequences (see, e.g., Owen, 2003) where we have used two different sets of prime-number bases to create the two sets of Halton samples.)

PLS endeavors to preserve uniformity of sampling coverage over the parameter space in the various stages or levels of the incremental experimental design. Uniform coverage over the parameter space is desirable for general response surface construction because this reduces the redundancy or marginalization of new information from added samples. This is a basic concept of upgradable quadrature methods (Patterson-1968, and Genz & Malik - 1983).

PLS builds knowledge by reducing global knowledge deficit over the parameter space. It does not attempt to build specific or targeted knowledge by building on previous information in the manner of “adaptive” sampling, which efficiently maximizes knowledge over particular regions of the parameter space. Thus, PLS designs select sample locations strictly on geometric principles such that new samples are intended to be “maximally far” from each other and from all other existing samples at each level of sampling. Thus, global uniformity of coverage is maintained at each level as the scheme progresses.

The arrangement of samples in each PLS level allows the parameter space to be subdivided into a regular pattern of adjacent polygons, which for two parameter space dimensions results in triangular and quadrilateral 2-D finite elements (FEs) yielding linear to quadratic polynomial interpolation over each element (see Romero & Bankston, 1998). The collection of all the elements together creates a C^0-continuous global function over the parameter space. As such, the global RSA has considerable freedom to locally conform to the data values of the sample points (see Figures 6 - 8). A mathematical analysis (Strang & Fix, 1973) of finite element interpolation of this nature shows that for a continuous and infinitely differentiable function over the parameter space, the domain integral of the pointwise absolute error goes to zero as the spacing between samples goes to zero. This analysis can be applied to the FE/PLS method. Hence, in the limit of infinite sampling, FE/PLS response surfaces converge everywhere in the domain to target functions in this class (though the convergence rate is generally not uniform over the domain). Therefore the FE/PLS method provides a convergent reference against which the accuracy of other progressive response surface methods can be compared.

Upon upgrading from one level to the next of the PLS design, the resulting change in the response surface at any point in the domain is a heuristic indicator of the magnitude of the local error in the response surface approximation. When the incremental change goes to zero, this tentatively indicates that the local approximation error has become negligible.

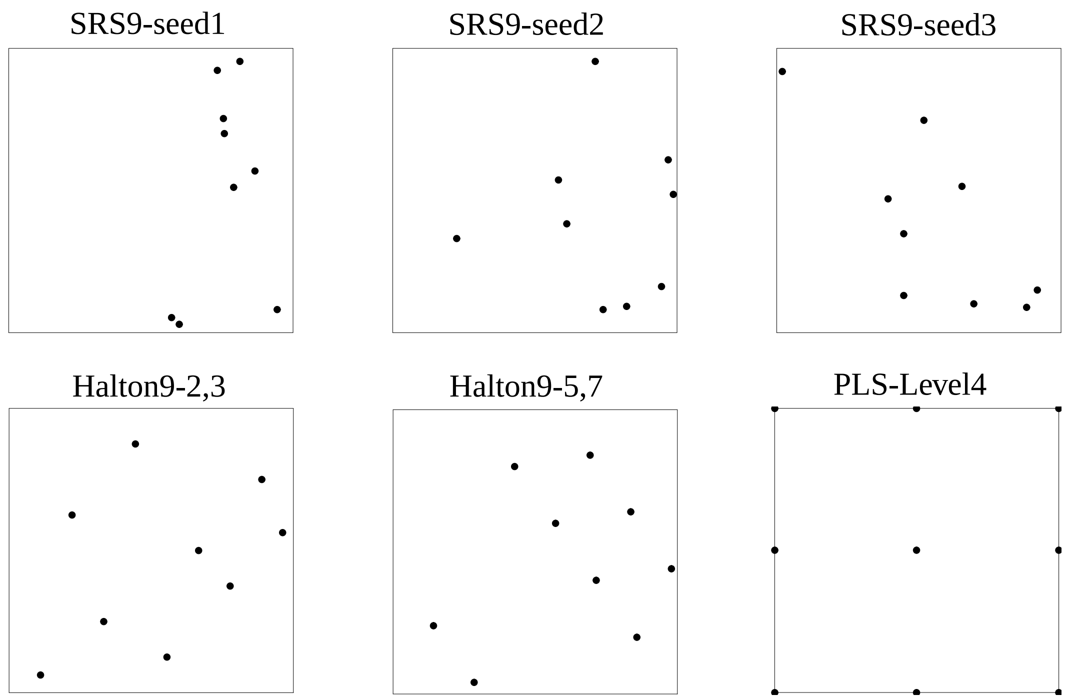

Figure 1. **9-point sample sets on a 2-D unit Hypercube from three types of incremental sampling methods: A) Top Row– SRS Monte Carlo with three different initial seeds; and B) Bottom Row– base 2,3 and 5,7 Halton sets, and PLS set.**

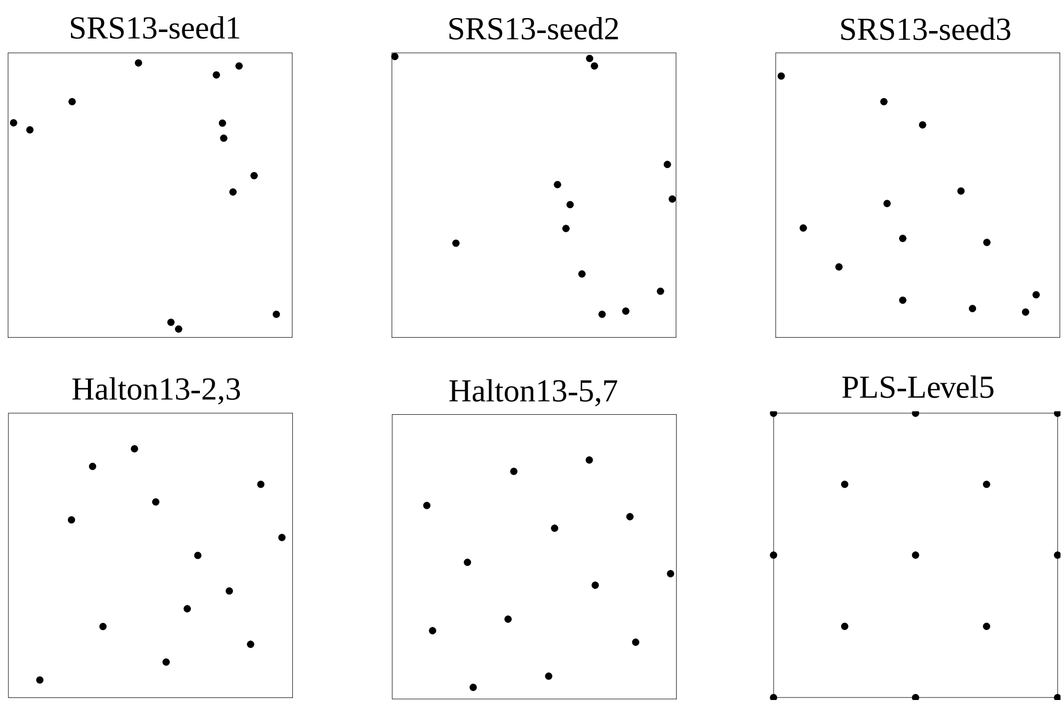

Figure 2. **13-point sample sets on a 2-D unit Hcube from three types of incremental sampling methods: A) Top Row– SRS Monte Carlo with three different initial seeds; and B) Bottom Row– base 2,3 and 5,7 Halton sets, and PLS set.**

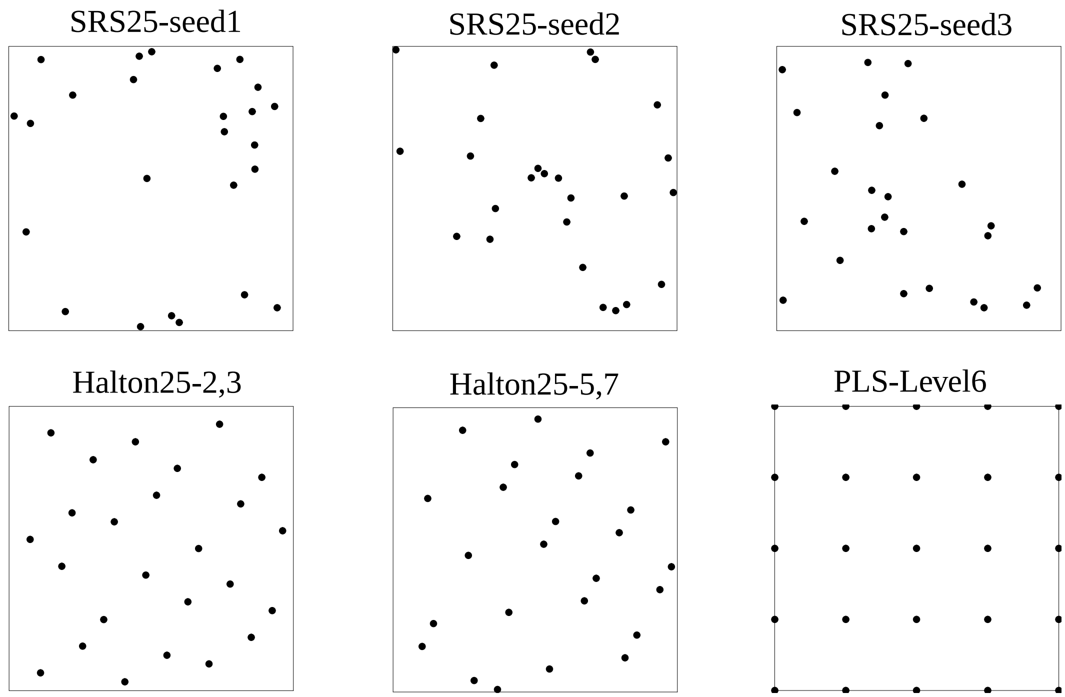

Figure 3. **25-point sample sets on a 2-D unit Hcube from three types of incremental sampling methods: A) Top Row– SRS Monte Carlo with three different initial seeds; and B) Bottom Row– base 2,3 and 5,7 Halton sets, and PLS set.**

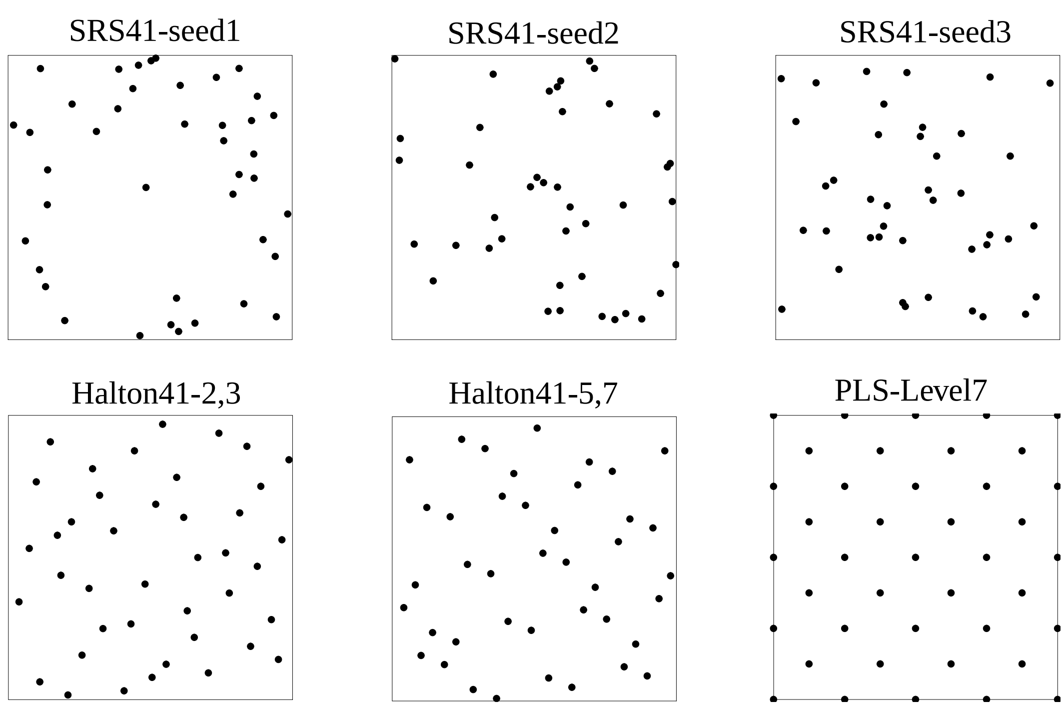

Figure 4. **41-point sample sets on a 2-D unit Hcube from three types of incremental sampling methods: A) Top Row– SRS Monte Carlo with three different initial seeds; and B) Bottom Row– base 2,3 and 5,7 Halton sets, and PLS set.**

The advantage of the structured PLS approach is that it is thought to be globally optimal or nearly optimal in that, as samples are added in attaining each new Lattice Sampling "Level", the spacing of samples throughout the parameter space remains uniform, or nearly so, given the constraint of previous sample locations. The locations of previous samples are respected because it is desired to fully leverage them (with minimal marginalization) as new samples are added.

This would appear to be the most efficient way (following the precedent of upgradable quadrature methods) to progressively build up a response surface. It may be that at any given level if complete freedom is allowed where the samples can be placed, then these may be arrangeable over the parameter space with better uniformity than PLS provides. For example, Romero et al. (2003) find that this appears to occur sometimes with Latin Hypercube Sampling and Centroidal Voronoi Tessellation –depending on the random number generator initial seed, number of samples, etc. However, these are **non-incremental sampling methods**; augmenting the number of samples would imply a completely different sampling of the parameter space at all new point locations. To go from X to X+Y samples in the space would therefore involve X+Y new evaluations of the generating function. This is more expensive than an **incremental sampling method** like PLS, which at each new stage costs only the increment of Y new samples (to be added to the original X for a total of X+Y uniformly dispersed samples).

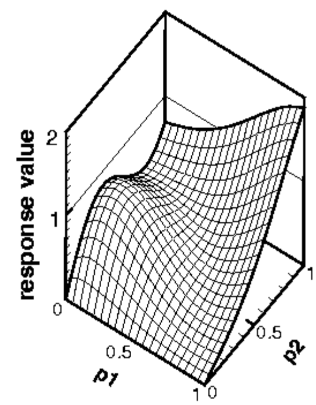

Figure 5. 2-D test function to be approximated ("generating" function).

A strong *dis*advantage of PLS, however, is that its **structured** experimental-design sampling nature allows only a quantized increment Y to be added to an existing PLS level (point set) to graduate to a new level. Hence, there is a constraint on the number of samples that can be added and still maintaining as uniform a filling of the hypercube as the

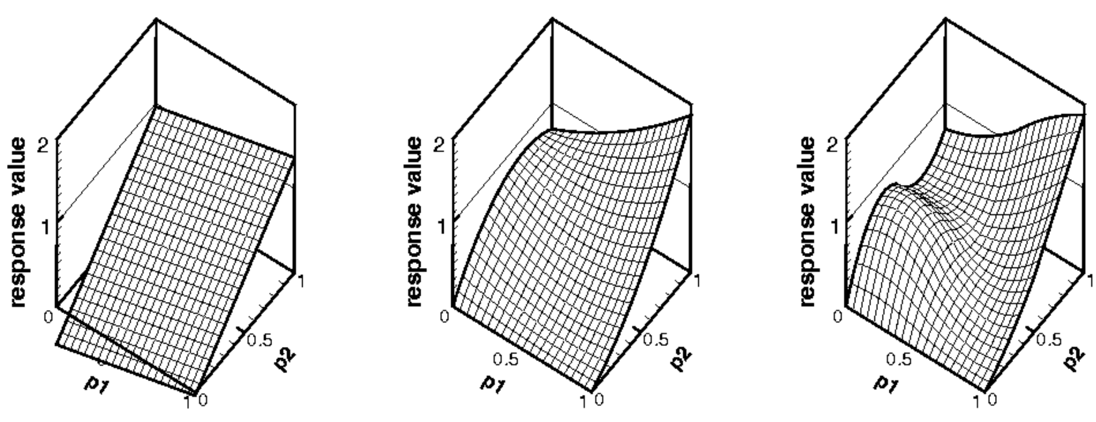

Figure 6. Successive response surface approximations to the exact generating function by finite-element interpolation of PLS designs: (A) - Level 1 having 3 samples; (B) - Level 3(4) having 9 samples; (C) - Level 6 having 25 samples.

scheme is capable of. In particular, this quantized incremental cost Y scales increasingly quickly as the PLS level and dimension of the parameter space increase.

Other, incremental sampling methods exist that are **unstructured** (non-quantized) that don't have the hard cost-scaling of PLS. Some popular unstructured incremental-sampling methods are the Halton "quasi- Monte Carlo" low-discrepancy sequence method (Owen, 2003), and Simple-Random Sampling (SRS) Monte Carlo. These allow Y as small as 1 without creating a prejudiced imbalance in the global coverage/filling of the parameter space. Hence, regardless of the dimension of the space and the stage of sampling, we could incrementally add, say, 50% more samples at a time in monitoring the convergence of the response surface results. If the sampling budget is reached before adequate convergence is established, then the final results will still be based on a point placement that is characteristic of the most uniform sample placement the method is capable of producing.

The point placement of Halton sampling appears characteristically more uniform over the parameter space than that of SRS, as can be seen in Figures 1 - 4. (We assert this based only on graphical appearance, not on formal measures of uniformity. Burkhardt et al. (2004) describe reliable metrics of point uniformity in hypercubes, but these have not been applied here.) In turn, the uniformity of PLS appears to be better than that of Halton. In the next section we will confirm that better sampling uniformity over the parameter space generally correlates with better response surface accuracy.

Given a set of sampling points over a parameter space, the quality of the response surface approximation (RSA) also depends on the particular method used to fit and interpolate the data. We now turn to consideration of data fitting methods to interpolate and extrapolate the sample data. Finite-element interpolation of sample data for general unstructured point placement and arbitrary numbers of samples in arbitrary dimensions is a difficult prospect. It is not immediately obvious how anything but linear tetrahedral (simplex) elements could readily be used, thereby sacrificing any higher-order convergence potential in the piecewise-linear interpolation scheme. Extrapolation procedures are also not immediately obvious, as this is not normally encountered in the Finite Element Method.

Four general data fitting and interpolation/extrapolation methods that can work with structured or unstructured progressive sampling schemes have been evaluated by the authors. These are global polynomial regression and kriging (Romero, et al., 2000, and Krishnamurthy & Romero, 2002); moving least squares (Krishnamurthy & Romero, 2002, and Romero et al., 2003); and Radial Basis Function methods (Krishnamurthy, 2003). Though our experience is very limited, of these, MLS appears to present a good balance of response surface accuracy, smoothness, robustness, and ease of use. Therefore, we use MLS in the next section to generate response surfaces from the 2-D sample sets in Figures 1 - 4.

2-D Example Problem for Examining Performance of Progressive Response Surfaces

Figure 5 plots a 2-D model function used to study the effect of sample point addition, sample placement scheme, and interpolation method on response surface accuracy. The value of the multimodal function in Figure 5 is defined as:

$$Y(\text{p1,p2}) = \left[0.8r + 0.35\sin\left(2.4\pi\frac{r}{\sqrt{2}}\right)\right][1.5\sin(1.3\theta)] \qquad \text{(EQ 1)}$$

where $r = \sqrt{(p1)^2 + (p2)^2}$ and $\theta = \arctan\left(\frac{p2}{p1}\right)$ on the domain $0 \le p1 \le 1$, $0 \le p2 \le 1$.

Exact values (samples) of this function are obtained at 9, 13, 25, and 41 points shown in Figures 1 - 4 for the PLS, SRS, and Halton incremental sampling methods. These numbers of samples correspond to PLS levels 4 - 7 in 2-D. (The high deterministic uniformity of the PLS sets provides a good standard to compare the other point sets against.)

The point data is then fitted and interpolated with Moving Least Squares (MLS). The particular implementation of MLS we use is described in Krishnamurthy & Romero (2002). The quintic weight function described in the reference is used here to give C2 smootheness to the MLS global interpolation function over the 2-D parameter space. A quadratic polynomial basis function is used for local interpolation, which requires at least (M+1)(M+2)/2 sample points (6 for M=2 dimensions) within a given evaluation point's local radius of influence. An optimal local radius of influence is calculated and used for each different point set, so that this element of fitting error is minimized in this study.

In order to assess the response surface error due to the fitting method (as opposed to the number and location of sample points), the finite element interpolation method of Romero & Bangston (1998) is also used to fit the PLS data sets. Figure 6 shows FE/PLS response surfaces for sets of 3, 9, and 25 samples. This illustrates the convergence of the progressive response surface to the target function as samples are added from the PLS design.

To examine the fitting performance of the progressive response surfaces, a simple measure of quality of global fit over the parameter space is used:

$$\text{approximate spatial average absolute error} = \frac{\sum_{i=1}^{441} \left|\text{exact}_i - \text{predicted}_i\right|}{441} \qquad \text{(EQ 2)}$$

where exact and predicted values in the summation come from respective evaluation of the exact function and the particular response surface approximation at 441 equally spaced points on a 21x21 square grid overlaid on the unit-square domain. This measure is an expedient approximation to the global average integrated absolute error over the domain, which would require a much more involved calculation.

Performance of Progressive Response Surface Methods on 2-D Problem

Figure 7 presents the response surface errors for the various sampling schemes, numbers of samples, and interpolation methods. All sampling methods yield a reduction of global fitting error as the number of samples increases in progressing to each new stage of the response surface.

At every population level (9, 13, 25, 41 samples), SRS-based sample placement performs worst (for all three of the random-number-generator initial seeds tried). This is a general reflection of the less uniformity with which the points are placed in the domain, as Figures 1 - 4 reveal. The more uniform Halton placement (for both the 2,3 and 5,7 prime bases for p1,p2 coordinates) performs significantly better than SRS in general, but still significantly less well than the deterministically uniform PLS point placement. The effect of point placement is substantial; at a population level the error difference between the best and worst point placements is roughly equal to the error difference between that and the next population level. That is, using a better point-placement scheme can substantially reduce the number of samples (and evaluation cost) needed to achieve a given level of response surface fitting error.

Among the incremental sampling methods (which are the most cost efficient type of samplers for progressive response surface construction), of the ones tried here Halton appears to be the best choice in general. Being one-at-a-time incremental sampling methods, Halton and SRS don't have the hard cost-scaling of PLS (where only quantized increments are allowed), and Halton point uniformity is much better than SRS uniformity in general.

With regard to interpolation methods, both FE and MLS are applied to the PLS sets for comparison. Which interpolator is used has relatively less effect here than the number and placement of sample points, but the effect is still non-negligible at 25 and 41 points. Though FE interpolation generally yields smaller error than MLS, this is not always true (e.g., the data for 13 samples). Futhermore, FE is not currently a viable choice for unstructured Halton sampling, which is the most viable of the incremental sampling methods considered here.

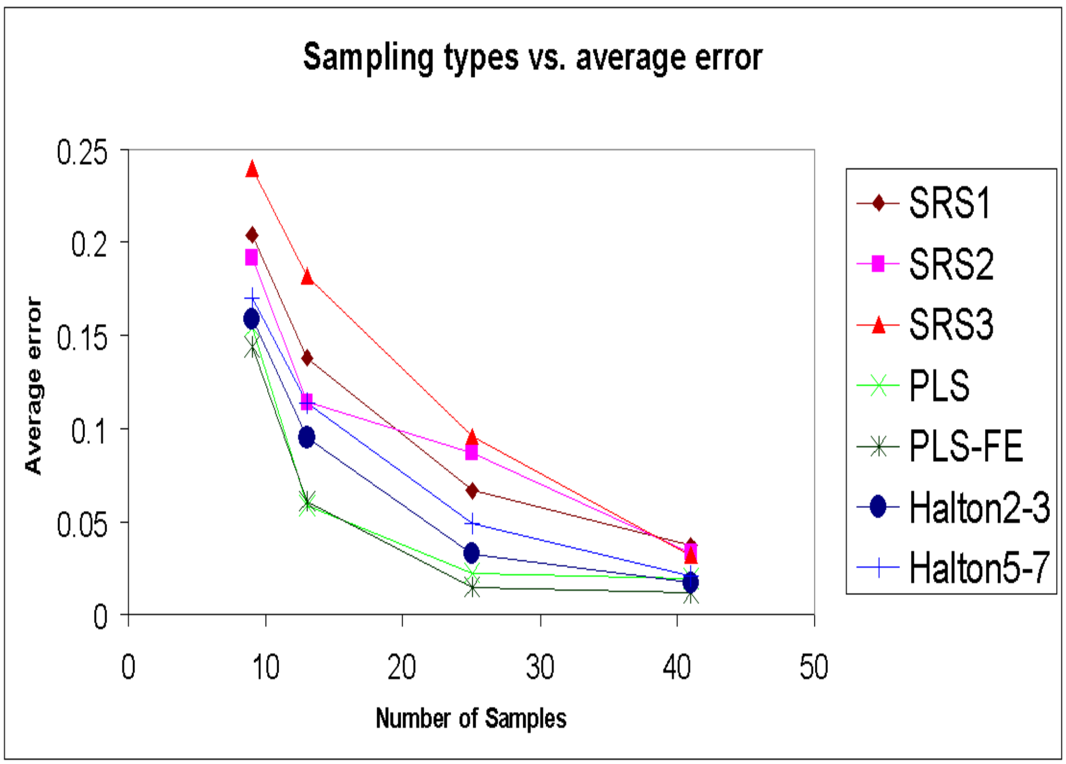

Figure 7. Convergence of progressive response surfaces to generating function as various sampling methods add samples. (Spatially averaged absolute error is plotted.)

Closing

It can be very difficult to determine when a particular sampling design and interpolation scheme sufficiently resolve a function, yet this must be done if the response surface is to be used as an effective inexpensive replacement for the actual function. Monitoring convergence heuristics of progressive convergent response surface approximations can help in this regard, and the Halton/MLS combination initially appears to be a generally viable approach for generating such response surfaces. However, there are still some subtle issues with both of these technologies that have to be characterized and addressed before they can be ready for general robust implementation. This will be the subject of future development, testing, and evaluation of the Halton/MLS method (and any other viable alternatives identified in the future).

References

Burkardt, J., M. Gunzburger, H. Nguyen, J. Peterson, and Y. Saka, (2004) “A new methodology for point sampling I: Uniform sampling in hypercubes” in preparation

Genz, A.C., and Malik, A.A., (1983) “An Imbedded Family of Fully Symmetric Numerical Integration Rules,” *SIAM J. Numer. Anal*., Vol. 20, No. 3, June.

Krishnamurthy, T., and Romero, V.J., (2002) "Construction of Response Surfaces with Higher Order Continuity and its Application to Reliability," paper AIAA-2002-1466, 43rd Structures, Structural Dynamics, and Materials Conference, April 22-25, 2002, Denver, CO.

Krishnamurthy, T., (2003) "Response Surface Approximation with Augmented and Compact Support Radial Basis Functions," paper AIAA-2003-7561, 44th Structures, Structural Dynamics, and Materials Conference, April 7-10, 2003, Norfolk, VA.

Owen, A. (2003), "Quasi-Monte Carlo Sampling," A Tutuorial for SIGGRAPH 2003, Stanford University Statistics Dept., notes available at: http://www-stat.stanford.edu/~owen/reports/siggraph03.pdfMonte

Patterson, T.N.L., (1968) “The Optimum Addition of Points to Quadrature Formulae,” *Math. Comp*., 22, pp. 847-856.

Romero, V.J., and S.D. Bankston (1998), “Finite-Element/Progressive-Lattice-Sampling Response Surface Methodology and Application to Benchmark Probability Quantification Problems,” Sandia National Laboratories report SAND98-0567, printed March, 1998.

Romero, V.J., L.P. Swiler, A.A. Giunta, (2000), "Application of Finite-Element, Global Polynomial, and Kriging Response Surfaces in Progressive Lattice Sampling Designs," paper PMC2000-176, 8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, July 24-26, 2000, Univ. of Notre Dame, IN. Extended version published as “Construction of response surfaces based on Progressive Lattice Sampling experimental designs with application to uncertainty propagation,” *Structural Safety*, Vol. 26, 2004.

Romero, V.J., J.S. Burkardt, M.D. Gunzburger, J.S. Peterson, T. Krishnamurthy, (2003), "Initial Application and Evaluation of a Promising New Sampling Method for Response Surface Generation: Centroidal Voronoi Tessellation," paper AIAA-2003-2008, 44th Structures, Structural Dynamics, and Materials Conference, Apr. 7-10, 2003, Norfolk, VA.

Roux, W.J., Stander, N., and Haftka, R.T. (1996), “Response Surface Approximations for Structural Optimization,” *6th Multi-Disciplinary Analysis and Optimization Symposium*, Belleview, WA, Sept. 4-6 (paper AIAA 4042-CP).

Unal, R., R.A. Lepsch, W. Engelund, and D.O. Stanley (1996), “Approximation Model Building and Multidisciplinary Design Optimization using Response Surface Methods,” *6th Multi-Disciplinary Analysis and Optimization Symposium*, Belleview, WA, Sept. 4-6 (paper AIAA 4044-CP).

Venter, G., Haftka, R.T., and Starnes, J.H. Jr. (1996), “Construction of Response Surfaces for Design Optimization Applications,” *6th Multi-Disciplinary Analysis and Optimization Symposium*, Belleview, WA, Sept. 4-6 (AIAA paper 96-4040-CP).