

Error Estimation Approaches for Progressive Response Surfaces –More Results*

V. J. Romero[†], R. Slepoy, L. P. Swiler, A. A. Giunta
Sandia National Laboratories[‡], Albuquerque, NM, USA

T. Krishnamurthy
NASA Langley Research Center, Hampton, VA, USA

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 built up incrementally as global information is added from new sample points added to the previous points in the parameter space. As the response surfaces are globally upgraded, indicators 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 or unstructured fashion. Whatever the approach, it is usually desirable to sample the entire parameter space uniformly (at least in early stages of sampling). At later stages of sampling, depending on the nature of the quantity being resolved, it may be desirable to continue sampling uniformly (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 previous stages of sampling ("adaptive" response surfaces). Here we consider progressive response surfaces where a balanced representation of global response over the parameter space is desired. We use Kriging and Moving-Least-Squares methods to fit Halton quasi-Monte-Carlo data samples and interpolate over the parameter space. On 2-D test problems we use the response surfaces to compute various response measures and assess the accuracy/applicability of heuristic error estimates based on convergence behavior of the computed response quantities. Where applicable we apply Richardson Extrapolation for estimates of error, and assess the accuracy of these estimates. We seek to develop a robust methodology for constructing progressive response surface approximations with reliable error estimates.

Keywords: Progressive Sampling and Response Surfaces, Error Estimation, Uncertainty Propagation

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/sensitivity analyses are cited in Giunta *et al.* (1994), Roux *et al.* (1996), Unal *et al.*, Venter *et al.* (1996), Romero (1998), and Simpson *et al.* (2004).

Some issues that arise when using response surface approximations (RSAs) are accuracy and the number and placement of the data samples on which the RSAs are built. With a sufficiently flexible global fitting/interpolating function over the parameter space, response surface accuracy ideally increases as the number of data points

*This paper is a work of the United States Government and is not subject to copyright protection in the U.S.

[†]Corresponding author, contact: vjromer@sandia.gov, 505-844-5890

[‡]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.

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.

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 economical replacement for the actual function. Monitoring convergence heuristics of progressive response surface approximations can help in this regard, and to get error estimates once the sampling is stopped. The latter is the principle investigative point of this paper.

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

Given these specifications, Romero et al. (2004a) examined several data fitting/interpolation methods (kriging, global regression polynomials, and finite-element piecewise polynomial interpolation) 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 (Romero and Bangston, 1998). Familiar structured designs such as Box-Behnken and Central Composite designs exist or are added at various stages or levels of the design.

PLS endeavors to preserve uniformity of sampling coverage over the parameter space in the various stages 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. The locations of previous samples are respected because it is desired to fully leverage them (with minimal marginalization of their information value) as new samples are added. This also means that the new added samples have maximal information value as well.

The advantage of the structured PLS approach is that it is thought to be close to an optimal incremental sampling method in that, as samples are added to attain each new Lattice level, the spacing of samples throughout the parameter space remains uniform. Thus, global uniformity of coverage is maintained at each level as the sampling levels progress.

A substantial *disadvantage* of PLS, however, is that its **structured** experimental-design nature allows only a quantized increment M of samples 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 at a time to maintain the uniform filling of the hypercube. Unfortunately, this quantized incremental cost M accelerates quickly as the PLS level and dimension of the parameter space increase.

Unstructured sampling methods do not have the prohibitive cost-scaling of PLS. However, unstructured sampling methods produce less-uniform sampling of the parameter space. Nevertheless, in various circumstances the uniformity can be quite good (see, e.g., Romero et al. (2003) for examples of Latin Hypercube Sampling (LHS) and Centroidal Voronoi Tessellation (CVT), and Romero et al. (2005a) for Latinized CVT and Hammersley “Quasi-Monte Carlo” (QMC) low-discrepancy sequence results).

However, the four methods just cited are **non-incremental** sampling methods; augmenting the number of samples implies a completely different sampling of the parameter space with all new point-locations. To go from N to $N+M$ samples in the space would therefore require $N+M$ new evaluations of the exact function (instead of just M

new samples). This is less efficient than an **incremental** sampling method like PLS, which at each new stage costs only the increment of M new samples.

Unstructured sampling methods that *are* incremental are the Halton QMC low-discrepancy sequence method that we examine in this paper, and Simple-Random Monte Carlo. These allow M as small as 1 without creating a prejudiced imbalance in the global coverage of the parameter space. Hence, regardless of the dimension of the space and the stage of sampling, we could incrementally add, say, 20% more samples at a time and the point sampling would remain as uniform as the method is capable of producing. The question is: "How uniform is the point spacing of these methods?"

The point placement of Halton sampling appears characteristically more uniform over the parameter space than that of Simple-Random Monte Carlo, but not as uniform as PLS or the four non-incremental sampling methods mentioned above (see Romero et al. 2003, 2005a). Among the incremental methods, Romero et al. (2004b) confirm that better sampling uniformity over the parameter space strongly correlates with better response surface accuracy (PLS most accurate, then Halton, then SRS). Because Halton does not suffer from the cost-scaling problems that PLS does, and it generally produces better point uniformity than Simple-Random Monte Carlo, we choose Halton sampling in this paper as a basis for our progressive response surfaces.

Given a set of sampling points over a parameter space, the quality of the response surface approximation also depends on the particular method used to fit and interpolate the data. We now turn to consideration of data fitting/interpolation methods.

Finite-element (FE) type interpolation of point data is very robust and accurate compared to other methods (Romero et al. 1998, 2004a,b; Krishnamurthy et al. 2002). However, FE interpolation is presently limited to structured point placement and low dimensional parameter spaces. For general unstructured point placement and arbitrary numbers of samples in arbitrary dimensions, it is not immediately obvious how anything but globally continuous C^0 (only piecewise smooth) interpolation functions from linear simplex elements could be built. Any higher-order convergence potential in the interpolation function would therefore be lost. Also, procedures for extrapolation to the hypercube boundaries are not immediately obvious, as spatial extrapolation is not normally encountered in the Finite Element Method (Strang & Fix, 1973).

Accordingly, four available data fitting and interpolation/extrapolation methods that are applicable to structured or unstructured progressive sampling have been evaluated by the authors. These are global polynomial regression (Krishnamurthy et al. 2002, Romero et al. 2004a); kriging (Krishnamurthy et al. 2002, Romero et al. 2004a, 2005b); Moving Least Squares (MLS) methods (Krishnamurthy et al. 2002, Romero et al. 2003, 2004b, 2005b); and Radial Basis Function (RBF) methods (Krishnamurthy 2003). The MLS, RBF, and kriging methods are capable of yielding much better results than global polynomial regression. However, the former methods are more complex and tricky to use. We present evidence of this in (Romero et al. 2005b). We used only kriging and MLS methods here because we are more familiar with these than with RBF methods, and global polynomial regression does not have the local fitting conformability we desire here (see Romero et al. 2004a).

In this paper we concentrate on the issue of error estimation in progressive RSAs. Upon upgrading a response surface by adding samples, the resulting change in values calculated from the response surface is a heuristic indicator of the magnitude of relevant error in the response surface approximation. When the incremental change goes to zero (for a sufficiently large number of added samples), this tentatively indicates that the relevant response-surface error has become negligible. We elaborate on this basis for convergence and error indication in this paper, and examine the accuracy of derived error estimates. We also examine whether Richardson Extrapolation, which is a familiar method for assessing the convergence of solution results from successively refined discrete approximations of partial differential equations (Roache 1998), can be successfully used to assess the convergence of results from successively refined response-surface approximations. In doing so, we summarize the RE technique and its applicability to progressive RSAs. As explained, the Richardson procedure sets constraints on the numbers of samples that can be added at each stage of progression.

In this paper we continue an investigation started in (Romero et al. 2005b). There we examined convergence and error estimates for global quantities that depend on the accuracy of the RSA over the entire parameter space. The assessment quantities were spatially averaged pointwise error, and mean and standard deviation of calculated response when the 2 input parameters to the performance function were assigned uncertainty prescribed by a 2-D joint-normal probability density function (PDF). Here we examine convergence and error estimates for calculated exceedence probabilities, given this same input PDF. As explained, these quantities test local RSA accuracy.

Application of Progressive Response Surface Approaches to 2-D Test Function

2-D Test Function

Figure 1 plots the 2-D performance function used to study the effect of sample point addition, sample placement, and interpolation method on the response surface approximations. The multimodal test function is defined as:

$$r(p_1, p_2) = \left[0.8R + 0.35 \sin\left(2.4\pi \frac{R}{\sqrt{2}}\right) \right] [1.5 \sin(1.3\theta)] \quad (1)$$

where $R = \sqrt{(p_1)^2 + (p_2)^2}$ and $\theta = \arctan\left(\frac{p_2}{p_1}\right)$ on the domain $0 \leq p_1 \leq 1$, $0 \leq p_2 \leq 1$.

Point Progression for Progressive Sampling of 2D Function

In this work, values (samples) of the target function are obtained at 3, 9, 27, and 81 points. This progression of points supports generalized Richardson Extrapolation (Roache 1998) for error estimation as pursued in a later section. This particular progression of points accompanies the choice of grid refinement factor $g = \sqrt{3}$ per space dimension. We chose $g = \sqrt{3}$ for reasons cited in the Richardson section later. As explained more fully in (Romero et al. 2005b), this factor is squared for two space dimensions (would be cubed for three space dimensions, etc.) to determine the growth in the number of sample points required to get to each new level of domain "discretization" for Richardson Extrapolation. Thus, for our 2-D problem the sampling total-growth-rate factor is $g_{\text{total}} = g^{D=2} = [\sqrt{3}]^2 = 3$. We start with 3 sample points in the domain because this is the minimum allowable for Moving Least Squares for reasons explained below. For the first refinement, we multiply the initial 3 sample points by the sampling growth rate of 3 to determine an increase to 9 samples. The next refinement implies $3 \times 9 = 27$ samples, and then $3 \times 27 = 81$ samples.

Halton Point Sets

For the reasons cited in the Introduction, we use Halton QMC sampling (see e.g. Owen 2003) to generate the 3, 9, 27, and 81 point sets. We use two different sets of prime-number bases (2-3 and 5-7) to generate two different sets of Halton point progressions. This allows us to examine any effect of prime-number base on the results. Figure 2 shows the specific Halton point sets used. The first 20 consecutive points in each progression are labeled so that the reader can see the uniformity (or lack thereof) in point spacing throughout the domain as sampling points are added one at a time.

Kriging and Moving-Least-Squares Data Fitting/Interpolation Approaches

The Halton point sets are fitted and interpolated with Moving Least Squares and kriging methods. The particular implementation of Moving Least Squares (MLS) we use is described in (Krishnamurthy et al. 2002). Linear basis functions are used for the 3-point Halton 2-3 and 5-7 sets. For the 9-point and larger sets a quadratic polynomial basis function is used. This requires at least $(D+1)(D+2)/2$ sample points (6 for $D=2$ dimensions) within a given interpolation point's local radius of influence. An optimal local radius of influence is found and used for each different point set so that this element of fitting error is minimized in this study. Optimal influence radii are determined by iterating the radius value to minimize the sum of the squared differences between the particular response surface and the exact function at a 21×21 grid of uniformly spaced points over the 2-D domain. A quintic weighting function was used to give C^2 smoothness to the MLS interpolation function over the 2-D domain.

The kriging method used in this study also produces a continuous C^2 interpolating function over the entire parameter space. The fundamentals of the method are described in (Giunta et al. 1998). The method is principally affected by the correlation length parameter Θ of the exponential spatial correlation model in the method. The correlation length can be specified or computed internally using a maximum-likelihood-estimation optimization procedure. We had some trouble with the latter approach, as shown in (Romero et al. 2005b). Hence, we manually varied the correlation length as described in (Romero et al. 2005b) until we achieved satisfactory results.

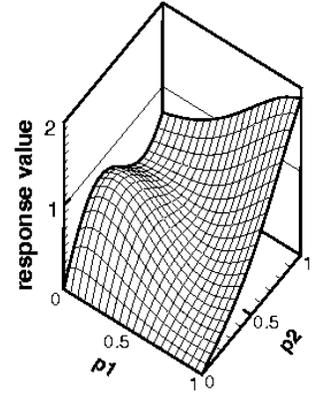


Figure 1. 2-D test function to be approximated (exact "target" function).

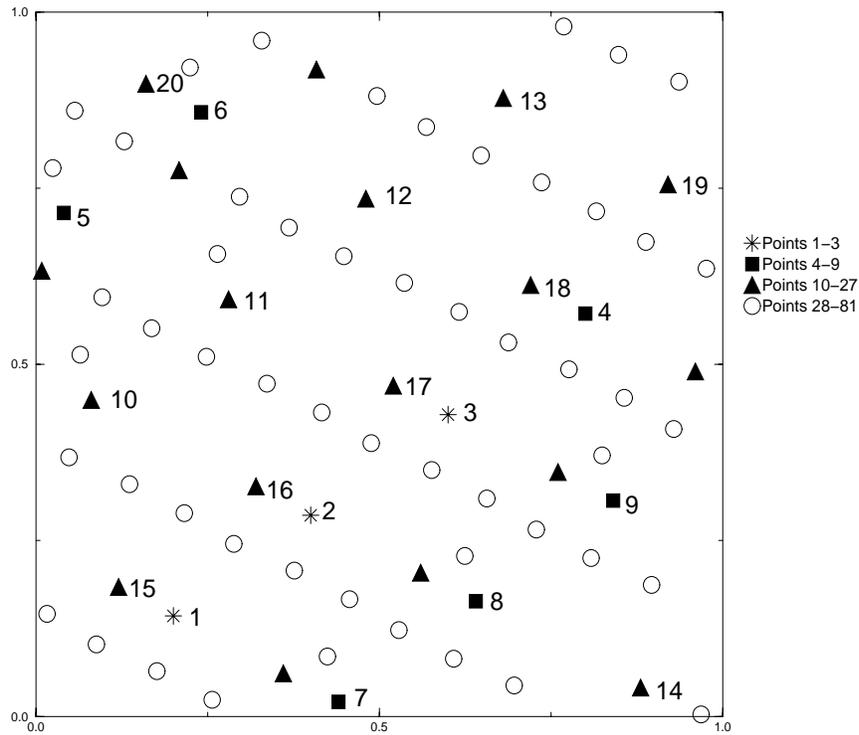
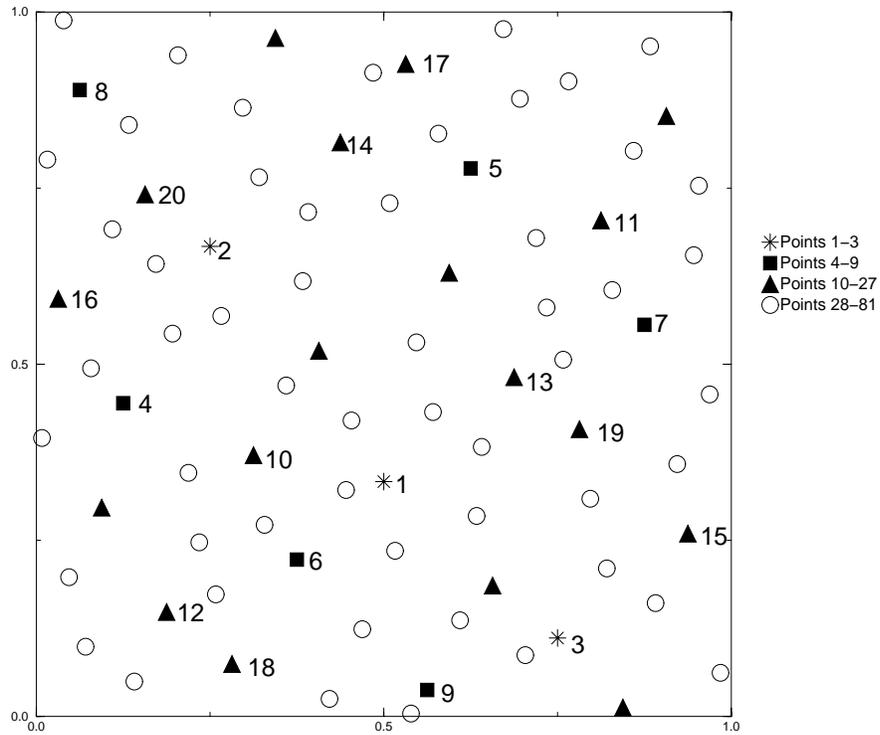


Figure 2. Halton Quasi-Monte-Carlo point sets (3, 9, 27, and 81 point sets) from two different prime-number Bases (2-3 [top chart] and 5-7 [bottom chart]). First 20 consecutive points in each progression are labeled to observe the uniformity (or lack thereof) in point spacing as points are added one at a time.

Again, as for the MLS method, a manual optimization was performed to minimize global pointwise difference from the exact function. Normally, such optimization could not be performed because we would only have values of the target function at the 3, 9, 27, or 81 sample points themselves, not everywhere over the parameter space as we employed for our parameter optimizations. We wanted to optimize the methods' free parameters to eliminate this source of RSA error in the following. We are therefore working with optimal or nearly optimal MLS and kriging results here. This indicates to us how our error estimators work under idealized MLS and kriging fitting performance. If they work under these conditions, then maybe the error estimators will still work effectively under more realistic conditions where non-optimal RSA fits exist. Conversely, if the error estimation methods do not work under the idealized conditions, then we'd have little reasons to think they would work under more realistic conditions.

Figure 3 shows the MLS and kriging RSAs. Generally, the manually optimized MLS and kriging RSAs improve toward the target function as sampling points are added. Beyond the noticeable differences at 3 points, the differences between kriging and MLS are relatively small at 9, 27, and 81 points. The effect of prime-number base is significant at lower numbers of samples (3 and 9 points) but not for the larger sample sets (27 and 81 points) when the parameter space becomes well populated.

The MLS results at 3 points are noticeably bad for both 2-3 and 5-7 bases. This is because a linear basis function is the only allowable MLS surface for 3 points in a 2-D space, and the first three points that define the plane are almost collinear for Base 2-3 and are collinear to within machine precision for Base 5-7 (see Figure 2). Thus, the 5-7 case is ill-conditioned so the MLS method yields noise. The Base 2-3 results are a steep plane (though shown in Figure 3 with flattened dog-ear extremes because of plotter limitations when the vertical scale was kept the same as the other plots on the page). We could have legitimately eliminated the MLS problems at 3 points by simply picking 3 best points out of the sets of 9 (assuming that in a real application we would plan to progress to at least a second level of sampling and would therefore know *a priori* nine locations for sample points and be able to pick the best three of these locations to sample first). However, to be consistent we would have had to use the same new Base 2-3 and 5-7 3-point sets for the kriging progressions and had already calculated and plotted kriging results with the original 3-point sets. Moreover, the unfortunate MLS results with the original 3-point sets don't adversely impact our error estimation findings later, so we stayed with the originals.

Accuracy and Convergence of Progressive Response Surface Methods

RSA accuracy as it impacts certain results of uncertainty propagation is investigated in the next two subsections. The system response function maps uncertain (random variable) inputs of the system into an output uncertainty distribution. Concretely, let the values of two uncertain inputs $p1$ and $p2$ come from independent normal distributions having means 0.5 and standard deviations $\sigma = 0.5/3$. The corresponding joint probability density function (JPDF) of these input random variables is represented in Figure 4 after truncation of the JPDF beyond the $p1$ - $p2$ unit-square domain. The JPDF is then renormalized to integrate to unity over this space. The relative height of the JPDF defines the relative propensity for attaining specific combinations of input values marked by any particular point (location) in the parameter space. This likelihood function for attaining given input combinations maps through the applicable input/output functional relationship of the system (such as our test function shown in Figure 1) into a corresponding likelihood function for attaining values of the system output.

Operationally, the JPDF of the input random variables can be mapped through the system input/output response function into an output response distribution via standard Monte Carlo (MC) sampling techniques. Toward this purpose, one million Latin Hypercube (Iman and Shortencarrier 1984) MC samples or $(p1, p2)$ parameter combinations (points in the parameter space) were generated according to the JPDF in Figure 4. The 2-D target function and the various RSAs in Figure 3 were then each evaluated at the 10^6 $(p1, p2)$ parameter sets, yielding for each function a population of 10^6 response values.

Response Probabilities Calculated from RSAs (indicators of local RSA accuracy)

Frequently in probabilistic analysis, it is desired to determine the probability that response lies above or below some threshold value r . The "exceedence probability" $EP(r)$ is estimated by the ratio of the number of calculated response values above the threshold value r , to the total number (10^6) of samples. For any threshold value r , we can then get estimates of $EP(r)$ from the 10^6 output values of the exact and approximate response functions.

We note that for a response threshold value of say $r=1.5$, the exact value of the exceedence probability is equal to the integral of the JPDF over the exceedence (shaded) region of the parameter space shown in Figure 5.

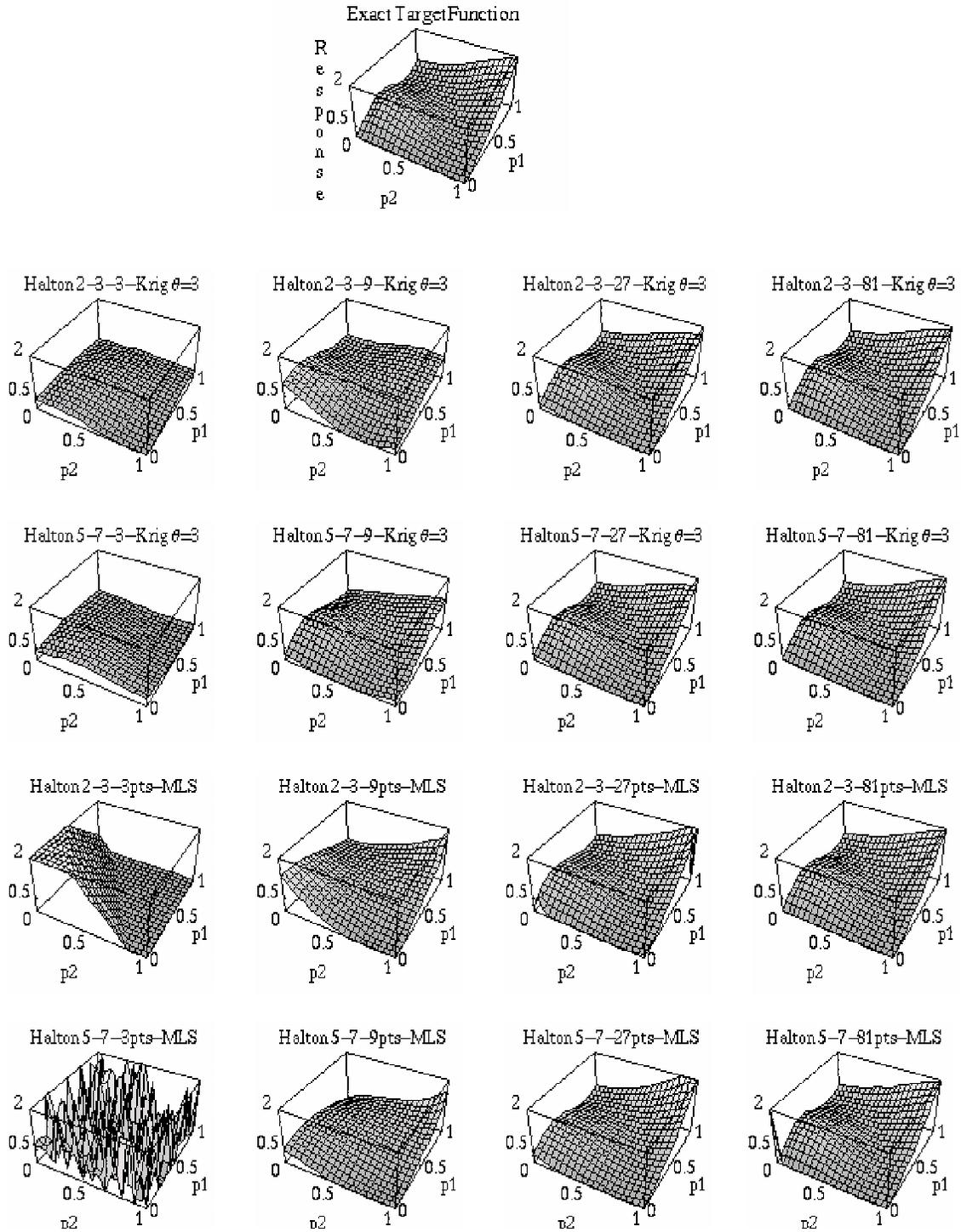


Figure 3. Convergence of progressive response surfaces to target function as samples are added to the parameter space. (Kriging and MLS spatial correlation length parameters optimized manually by comparison to exact target function.)

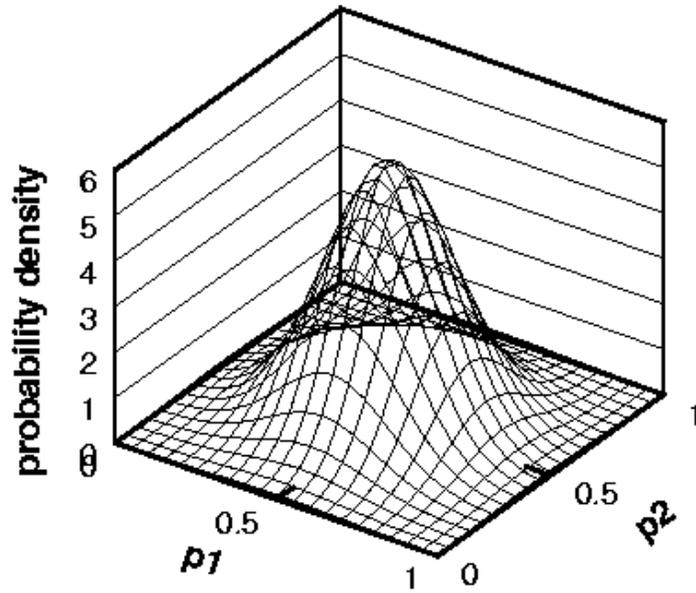


Figure 4. Joint Probability Density Function describing the random variables in the problem: normally distributed parameters p_1 and p_2 with means 0.5, std. deviations $\sigma=0.167$, and truncation of the unit square parameter space at 3σ above and below the mean values.

The figure shows a cutting plane through the exact function at a response level of $r=1.5$. The shaded region of the $r=1.5$ threshold plane marks the region of the (p_1, p_2) parameter space where the function exceeds the threshold value. Accordingly, a multidimensional integration of the JPDF over this shaded region yields $EP(r = 1.5)$.

Errors in RSA-derived MC probabilities are thus connected to RSA errors in the local vicinity of the exceedence region. These errors alter the perceived exceedence region and associated integration boundary, inside of which the MC samples that fall contribute to the numerator of the MC ratio, and hence to the EP estimate. Therefore, error of RSA-derived MC probability serves as a mechanism for sensing localized RSA error in the vicinity of the exceedence region for a particular response threshold level r .

We use the above mechanism to examine local RSA error for a diversity of exceedence regions and probability magnitudes corresponding to threshold levels $r = 1.5, 1.0, 0.5,$ and 0.2 . The associated exceedence regions are shown in Figures 5 and 6.

Results for all threshold levels are listed in Table 1 and plotted in Figures 7 to 10. Also shown for reference are the exact values of the statistics obtained from the exact 2-D function. The error bars in the plots and the 'S' and 'U' designators in some of the cells of the table will be explained later. Results are presented from 10^6 response values of MLS and kriging RSAs constructed using 3, 9, 27, and 81 points (for two Halton point progressions from prime-number generating bases 2-3 and 5-7).

The MLS results at 3 points are off the plotting scale because of the unfortunate oversight described earlier where the three points used were effectively collinear. However, this is viewed as an avoidable interaction with the Halton sampling method through the fix mentioned earlier.

At all threshold levels, as Halton sample points are added, the MLS and kriging RSAs produce non-monotonic convergence of the probability estimates for at least one if not both Halton point sets (Base 2-3 and Base 5-7). For both MLS and kriging, results are generally reasonably accurate at 27 points, and usually very accurate for 81

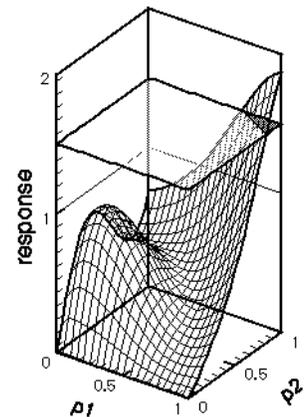


Figure 5. Cutting plane through exact function at a response value of 1.5.

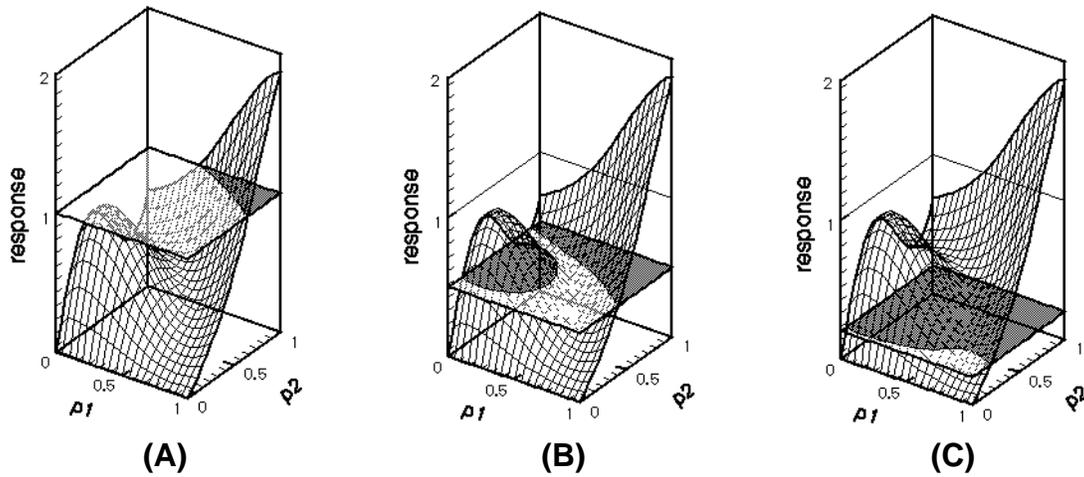


Figure 6. Cutting planes through exact function showing associated exceedence regions for response threshold values of 1.0, 0.5, and 0.2.

Table 1. Response Probabilities calculated from MLS and Kriging response surface approximations (indicators of local RSA accuracy)

Calculated quantity	# of Halton samples	kriging		Moving Least Squares	
		Base 2-3	Base 5-7	Base 2-3	Base 5-7
EP($r=1.5$) exact=0.000252	3	0.0	0.0	0.03458	0.11443
	9	0.0, U	0.0, U	0.00000, S	0.00000, S
	27	0.00023, S	0.00011, U	0.00029, S	0.00044, S
	81	0.00026, S	0.00023, S	0.00024, S	0.00025, S
EP($r=1.0$) exact=0.00756	3	0.0	0.0	0.11344	0.19141
	9	0.00232, U	0.00127, U	0.00363, S	0.00005, S
	27	0.00752, S	0.00773, S	0.00692, S	0.00749, S
	81	0.00777, S	0.00687, S	0.00742, S	0.00706, U
EP($r=0.5$) exact=0.44926	3	0.52822	0.28435	0.27175	0.43713
	9	0.76977, U	0.59636, S	0.76769, S	0.77031, S
	27	0.44815, S	0.44746, S	0.54894, S	0.59333, S
	81	0.46814, S	0.37717, U	0.46517, S	0.46168, S
EP($r=0.2$) exact=0.98439	3	0.98503	0.87991	0.40183	0.65150
	9	0.98907, U	0.98524, S	0.99079, S	0.98125, S
	27	0.98375, S	0.98437, S	0.98531, S	0.98317, S
	81	0.98519, S	0.97569, S	0.98454, S	0.98454, S

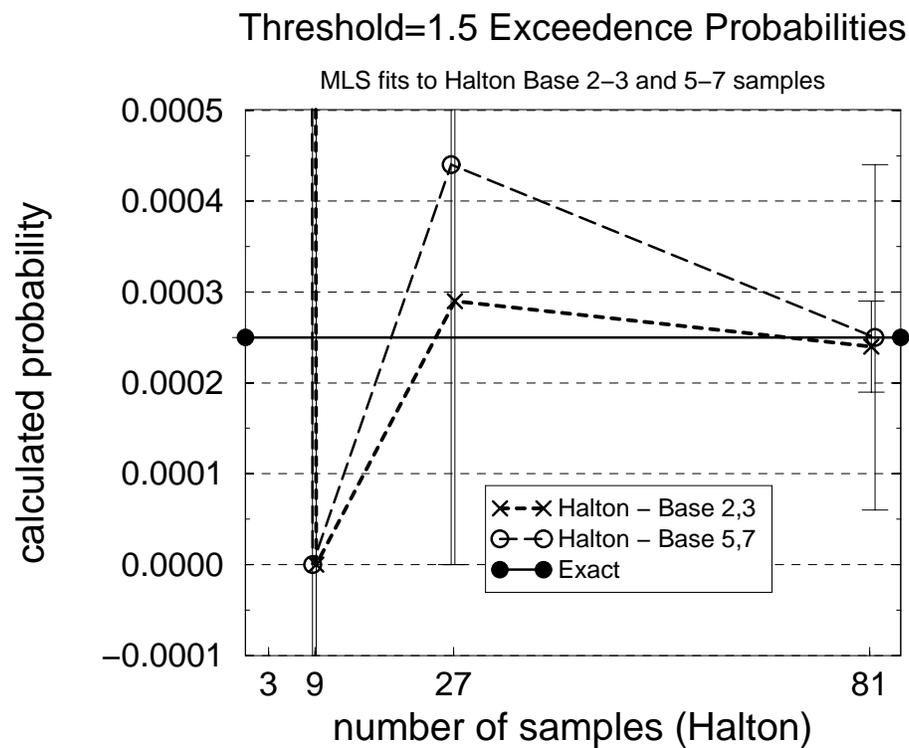
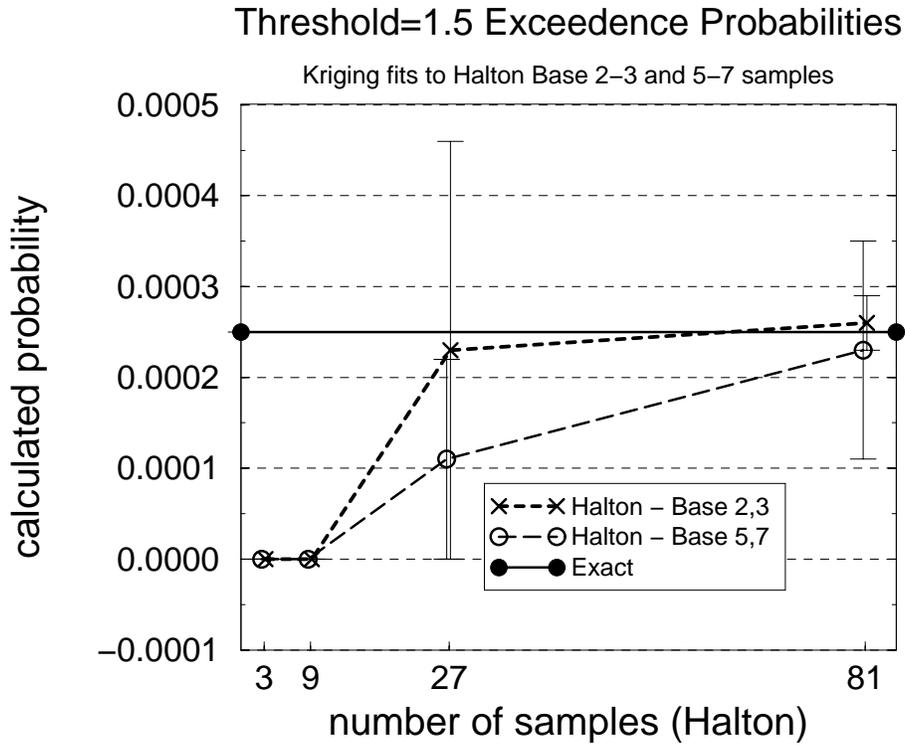


Figure 7. Convergence behavior of calculated exceedence probability (for $r=1.5$ threshold) as Halton sample points are added in kriging and MLS progressive RSAs.

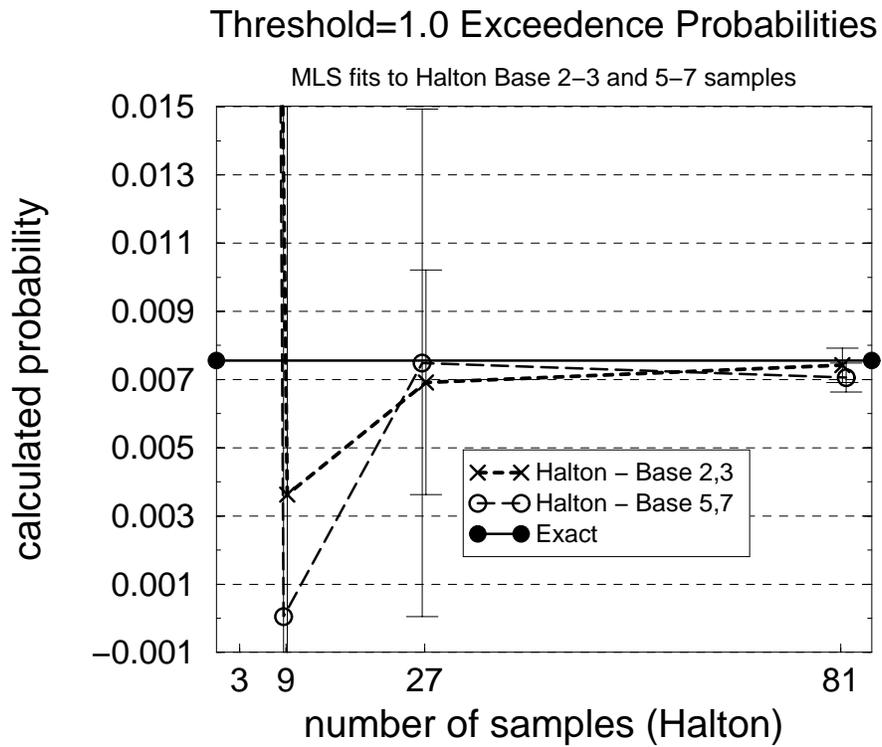
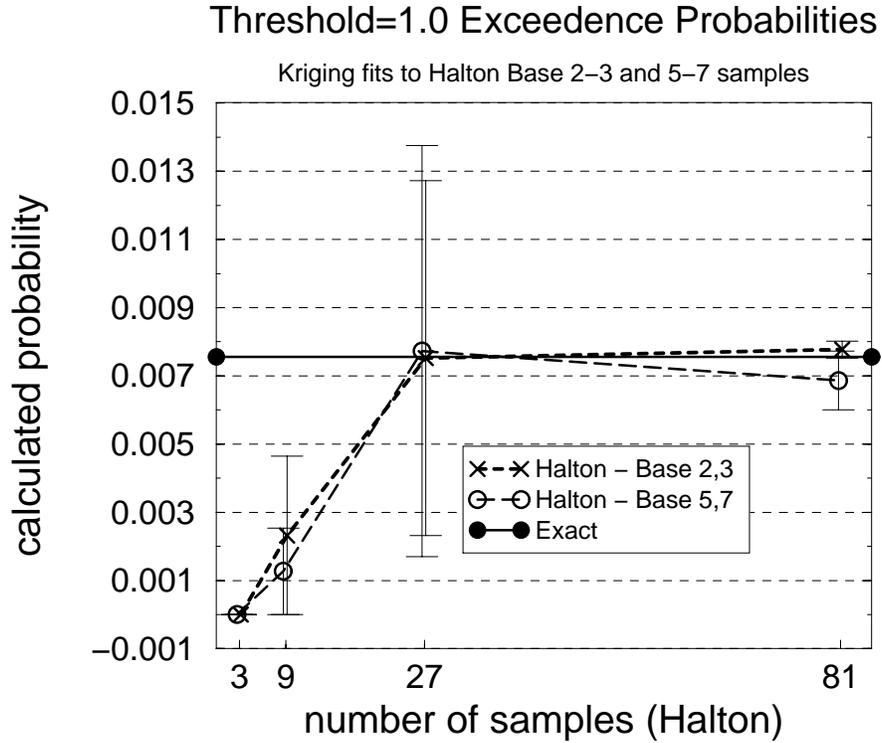
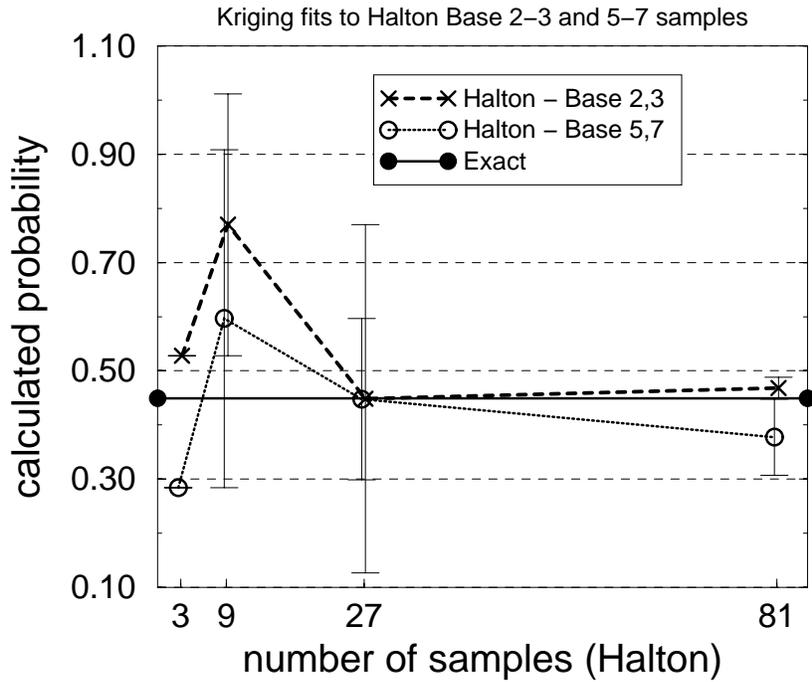


Figure 8. Convergence behavior of calculated exceedence probability (for $r=1.0$ threshold) as Halton sample points are added in kriging and MLS progressive RSAs.

Threshold=0.5 Exceedence Probabilities



Threshold=0.5 Exceedence Probabilities

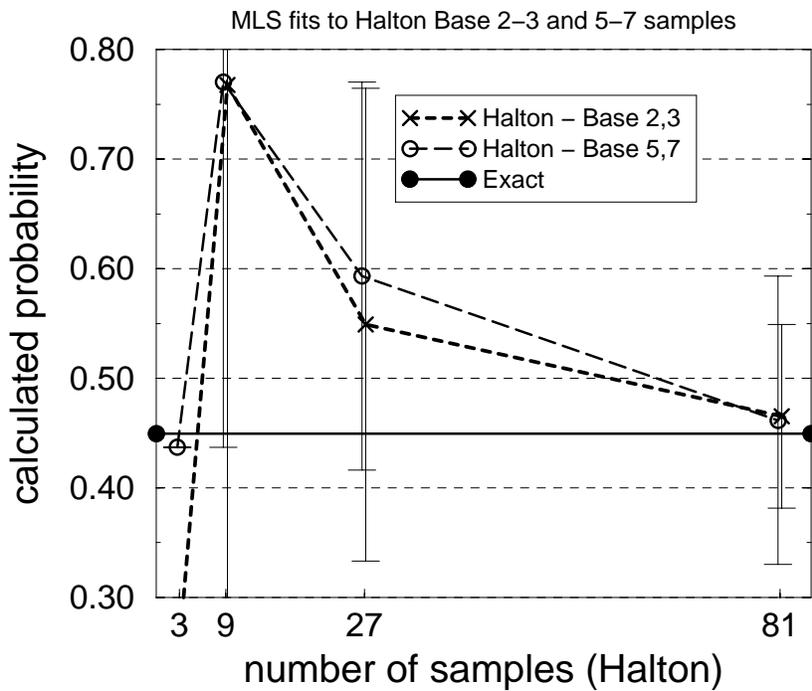


Figure 9. Convergence behavior of calculated exceedence probability (for $r=0.5$ threshold) as Halton sample points are added in kriging and MLS progressive RSAs.

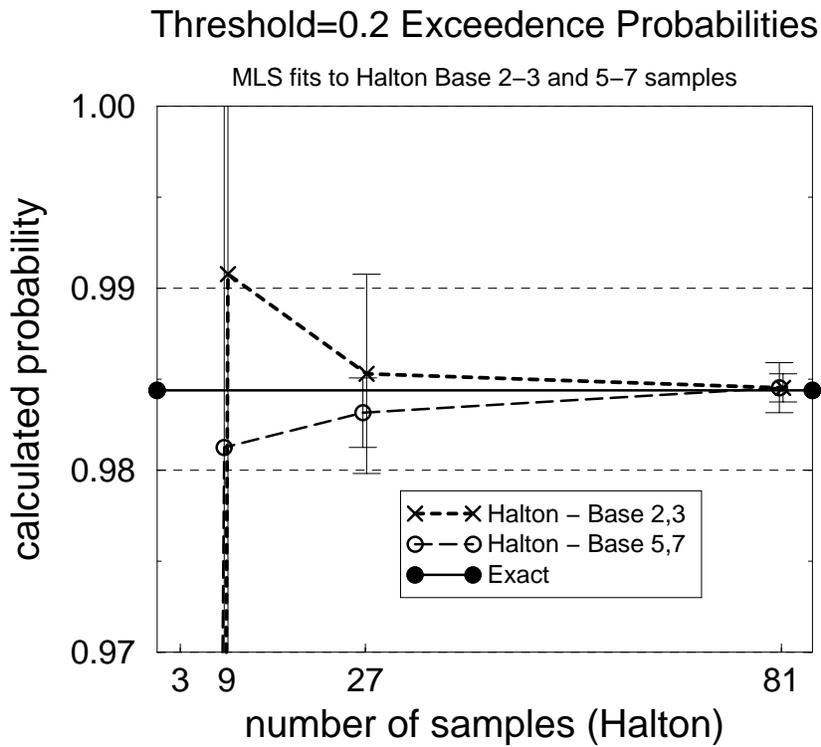
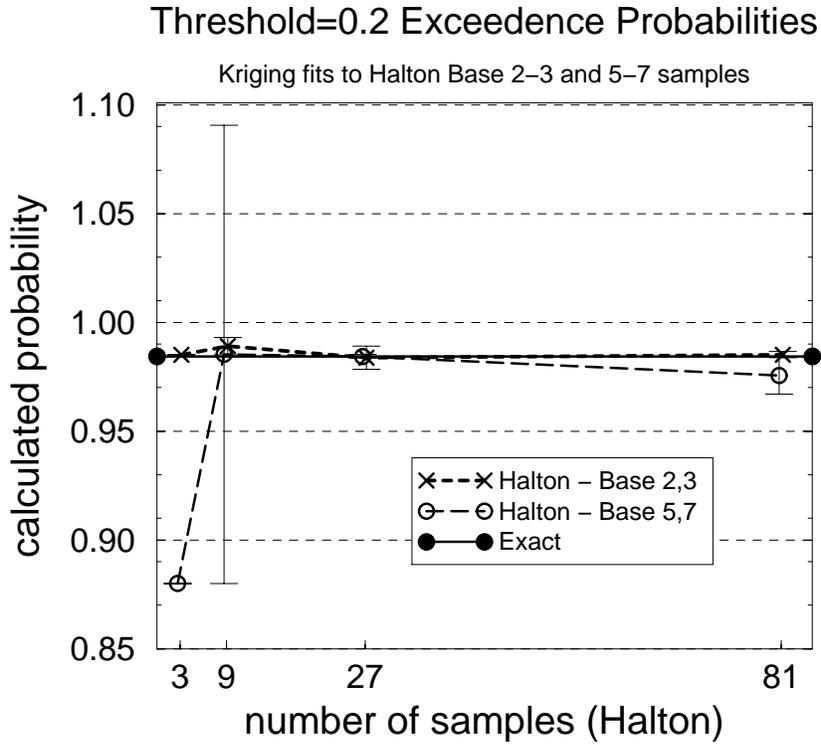


Figure 10. Convergence behavior of calculated exceedence probability (for $r=0.2$ threshold) as Halton sample points are added in kriging and MLS progressive RSAs.

points. The exception at 81 points is the kriging fit to the Halton Base 5-7 points. Although the value of the correlation length Θ was manually optimized over all the Halton point sets, kriging did not work particularly well (nearly ill-conditioned solution matrix) for the Halton Base 5-7 81-point set. Other than this, no strong advantage of one Halton prime-number generating Base over the other is apparent over the 64 cases.

Significant oscillation in convergence exists for both MLS and kriging methods, although more oscillation appears here for MLS than for kriging. This seems due to the bad MLS estimates at 3 sample points for both Halton Bases. Again, this could have been fixed with a different choice of the 3-point sets without loss of generality of the method, but was not done. Our other experience with MLS and kriging suggests that their performance is roughly comparable.

The main purpose of this paper is not to discern whether kriging or MLS works best with Halton sampling for progressive RSAs. The main purpose is to apply and assess error estimation approaches that work with both methods. We get to this shortly.

Global Response Quantities Calculated from RSAs (indicators of global RSA accuracy)

In (Romero et al. 2005b) we examined convergence and error estimates for global quantities that depend on the accuracy of the RSA over the entire parameter space. The assessment quantities were spatially averaged pointwise difference from the exact target function, and mean and standard deviation of the response distribution resulting from propagation of 10^6 LHS MC samples of the JPDF in Figure 4. RSA outputs for these global quantities are presented in Table 2 for completeness. Analysis in (Romero et al. 2005b) reveals similar trends and conclusions to those stated above for the local response quantities.

Table 2. Global Response Quantities calculated from MLS and Kriging response surface approximations (indicators of global RSA accuracy)

Calculated quantity	# of Halton samples	kriging		Moving Least Squares	
		Base 2-3	Base 5-7	Base 2-3	Base 5-7
spatially averaged pointwise absolute error (exact = 0.0)	3	0.25677	0.36857	1.39950	2.04166
	9	0.16766	0.08459	0.15890	0.09771
	27	0.00466	0.01434	0.03978	0.04558
	81	0.00450	0.02163	0.00588	0.01223
mean of response (exact=0.512108)	3	0.48112	0.39353	-0.00867	0.35116
	9	0.55544, S	0.53428, S	0.55234, S	0.55206, S
	27	0.51203, S	0.51209, S	0.51883, S	0.52303, S
	81	0.51772, S	0.48452, U	0.51382, S	0.51350, S
standard deviation of response (exact=0.163065)	3	0.09273	0.14604	0.83005	57.1842
	9	0.12401, U	0.15891, S	0.11631, S	0.11630, S
	27	0.16318, S	0.16389, S	0.14935, S	0.14674, S
	81	0.16293, S	0.16566, U	0.16029, S	0.16045, S

Error Estimation Approaches for Progressive Response Surface Methods

The convergence behavior of computed global and local quantities of response can be quite erratic with progressive response surfaces, as we have found before with other sampling and interpolation schemes (see Romero et al. 2004a). Nonetheless, we must assess RSA convergence and error. Here we look at “Simple Delta” and Richard Extrapolation approaches for convergence assessment and error estimation.

“Simple Delta” Method

A very simple error estimation approach exists that is applicable to progressive RSA results. The approach relies on simple change in computed response as samples are added. This “**Simple Delta**” (SD) method of error estimation is patterned after the control of convergence error in many fields such as local optimization, root-finding, and the iterative solution of sets of nonlinear simultaneous equations. (The latter is often associated with the solution of partial differential equations, e.g., P/Thermal Theory Manual, 1991).

In these applications it is necessary to determine when to stop iterating. Upon completing an iteration (or the addition of a sufficiently large number of samples to the RSA), if the incremental change in the quantity being computed is exactly zero, this generally indicates that the iteration has converged (although there are imaginable situations where this would not be the case). However, it is generally not cost effective to require machine-zero change in the computed quantity before terminating the iteration. Instead, an acceptably larger error tolerance is specified that terminates the iteration when the incremental change in the computed quantity becomes smaller than the tolerance. The tolerance is set according to how much error from the true result the user is willing to accept, in the context of the tradeoff that more accuracy implies more computing cost. The hope here is that when the incremental change becomes less than the prescribed tolerance, the answer is within that tolerance of the exact result.

This simple method of controlling and estimating convergence error appears to work reasonably well in practice, but there are certainly many circumstances where it will fail. In particular, if the tolerance condition is met in a slowly-converging regime of the iteration process, then termination can be premature and the exact result may not be within the specified error tolerance. Techniques are therefore sometimes employed to account for the actual rate of convergence (fast or slow, whatever it may be), but these are usually less-sophisticated variations of the RE procedure described later. We therefore assume that our empirical results with the RE method represent the best that can be practically attained when convergence-rate information is factored into the error estimates.

As an example, consider the kriging results at the top of Figure 7. Calculated exceedence probabilities are presented for the $r=1.5$ response threshold. For each curve we can derive error bars at 9, 27, and 81 points by subtracting the preceding coarser result. For instance, we subtract the result at 3 points from the result at 9 points. The absolute magnitude of this difference is taken as an estimate of how much more the exact result may differ from the current value at 9 points. In view of the often oscillatory nature of convergence of progressive RSAs, we form error half-bars of this magnitude above and below the result at 9 points.

It turns out that the formed error bars at 9 points do not contain the exact result. The error bars have zero magnitude because the RSA results at 3 points are the same (zero calculated probability) as at 9 points. This occurs for RSAs built on both the Base 2-3 and Base 5-7 Halton point sets. Figure 2 shows that neither point set has any of its first 9 samples near the 1.0-1.0 corner of the parameter space. Therefore, RSA accuracy near this corner is very bad. Yet accuracy at this corner is paramount to an accurate calculation of $r=1.5$ exceedence probability because, as Figure 5 shows, only that corner region contributes to the exceedence probability.

With the addition of more samples over the parameter space, this corner region of the parameter space becomes better sampled. Accordingly, the RSA over this region becomes better resolved and more accurate as Figure 3 shows. Corresponding probability estimates improve considerably as Figure 7 and Table 1 show. By 81 samples, the EP estimates are surprisingly accurate given the small magnitude of the exact probability (0.000252) and the fact that the 81 samples are distributed roughly evenly throughout the entire parameter space, and not (adaptively) concentrated about the exceedence region.

Applying the SD method at 27 points yields error bars in Figure 7 that do not contain the exact result for the Base 5-7 Halton sampling, but do contain the exact result for the Base 2-3 point set. At 81 samples, the error bars contain the exact result for the both the 5-7 and 2-3 Halton point sets. Hence, error bars calculated by the SD method contained the exact result in 3 of the 6 possible opportunities here.

We generally find a much higher success rate. For example, the MLS plot at the bottom of Figure 7 shows a 100% success rate, where all six Simple Delta error bars contained the exact result. More globally, we find a success rate of 81% (39 of 48 possible successes as denoted by the ‘S’ for ‘success’ designators in the eligible cells of Table 1). As indicated in Table 2, SD error bars were successful in 21 of the 24 possible cases (87.5% success rate) for computed global statistics of response (mean and standard deviation) analyzed in (Romero et al. 2005b). In total, a success rate of $(39 + 21)/(48 + 24)$, or 83%, was obtained. This rate is probably artificially high because the anomalous MLS results at 3 points were so bad that the SD error bars at 9 points were so large they’d almost be assured of containing the exact result if the estimates at 9 points were anything reasonable. When the MLS 9-point cases are discounted from eligibility because of this anomaly, the success rate for SD-derived error bars is still relatively high at 82% or 48 of the 60 eligible cases. This is quite satisfying for such a simple approach to error estimation. Success of the method appears to be independent of which Halton Base was used, or whether kriging or MLS was used.

Generalized Richardson Extrapolation Method

The more complicated generalized Richard Extrapolation method (RE, Roache 1998) did not perform nearly as well in our investigations as the Simple Delta method did. Therefore we are somewhat brief here with the presentation of the RE method and results.

In the solution of certain partial differential equations, certain discretizations of the equations possess an ideal theoretical rate p of convergence for computed results as the discretization of the PDE domain is appropriately refined. Then the exact grid-independent (grid-converged) result can be obtained from two results calculated on grids of “medium” and “fine” discretization spacings or cell size:

$$T_{exact} \approx T_f - \frac{T_m - T_f}{g^p - 1} \quad (2)$$

where g is the per-dimension refinement factor in going from medium to fine grid spacing, and T_m and T_f refer to results calculated on the medium and fine meshes. One caveat is that the medium and fine meshes must be fine enough that the calculated results are in the “asymptotic regime” of convergence of the discretized solution. This means that the results must be monotonically converging to the exact result at a slowing pace as the grid is refined.

In (Romero et al. 2005b) we present and argue an analogy between RSA sampling resolution in the domain, and discretization resolution of partial differential equations. There is some basis to argue that refined estimates from RSA progressions based on refined sampling in the domain are analogous to successive grid-refined solutions of partial differential equations. If the criteria for applicability are otherwise met, we can attempt an application of the RE idea to RSA-computed results at various sampling resolutions or “discretizations” of the domain, to estimate sampling-converged results. Strictly, the analogy requires that, at each stage of sampling ($N= 3, 9, 27,$ and 81 points), the N sample points are uniformly spaced throughout the 2-D domain. This associates each sample point with a region of volume $1/N$ of the domain volume. If the sampling method happens to be Centroidal Voronoi Tessellation, then this requirement is well met. However, CVT is not an incremental sampling method as progressive RSAs require, so Halton sampling appears to be the best choice of the incremental sampling methods. It gives the most uniform sample distributions of the incremental sampling methods that we presently know of. It approaches but does not strictly satisfy the uniform sampling requirement.

Another requirement for RE is that results converge at a constant rate p as the domain resolution is refined at a constant rate r per domain dimension. We have no theoretical basis to establish the constancy of convergence rates for results calculated from progressive RSAs. In fact, empirical evidence presented later indicates otherwise. Nonetheless, we apply RE to see if it actually works in practice, which is the most important indicator of its relevance.

Since we do not have a theoretically established rate p of convergence, we empirically solve for an apparent rate of convergence in an analogous manner as is done for PDE solutions. We must have results from three successive discretizations of the domain and the results monotonically converging at a slowing pace emblematic of the “asymptotic region” of convergence of the results. When these conditions are satisfied, the following is the standard formula (Roache 1998) for calculating the empirical rate of convergence p for appropriately discretized PDEs:

$$p = \frac{\ln\left(\frac{T_c - T_m}{T_m - T_f}\right)}{\ln(g)}. \quad (3)$$

In our case the per-dimension refinement factor is $g = \sqrt{3}$ and T_c , T_m , and T_f refer respectively to results calculated from RSA built on the “coarse”, “medium”, and “fine” samplings of the exact function. The “course”, “medium”, and “fine” sampling resolutions are the 3, 9, and 27 point sets for the 3-9-27 progression cases, and the 9, 27, and 81 point sets for the 9-27-81 progressions. The observed p order can be used in equation (2) to extrapolate from the two most accurate results (from the medium and fine samplings) to a “converged” result. We emphasize here that RE is applied to quantities calculated from the coarse, medium, and fine response surfaces, to improve these estimates or to get error estimates. RE is not applied to improve the response surfaces themselves.

Our choice of $g = \sqrt{3}$ as the per-dimension refinement factor in this work arises from several considerations. First, we want to be able to test the accuracy of Richardson Extrapolation and the consistency of the calculated rate of convergence p across two “spans” of response data. That is, we hope to apply the equations over the 3-9-27 point span and over the 9-27-81 point span to see if the calculated convergence rates remain constant, and whether extrapolated results from these two spans of data coincide closely. To get the necessary three refinements, the augmentation factor of $\sqrt{3}$ per space dimension or $[\sqrt{3}]^2$ for two dimensions yields a moderate number of top-end samples (81) in a progression that spans 4 levels. The next higher rate of refinement we could have chosen is $r = \sqrt{4}$, or a refinement doubling per space dimension. For two space dimensions this would have meant a factor of 4 growth in the total number of samples per refinement, which would have taken the sample count to a relatively high 192 for the third and last refinement.

As a demonstration example, we now consider the applicability of RE to the kriging results at the top of Figure 7. Neither of the 3-9-27 point spans of results based on Halton 2-3 and 5-7 sampling exhibit the necessary monotonic convergence, so are not eligible for RE. The 9-27-81 point span based on Halton 5-7 sampling does exhibit monotonic increase, but it is *divergent* rather than convergent. That is, a slowing pace of convergence does not occur as the grid is refined; the difference from the 27-point result to the 81-point result is greater than the difference from the 9-point to the 27-point result. Hence, the behavior is not yet in the required asymptotic regime of convergence. The 9-27-81 point span based on Halton 2-3 sampling is the only of the four candidate spans in the plot that RE is applicable to. That RE is applicable in this case is denoted by a ‘Y’ in Table 3. The N’s in Table 3 for the other three cases denote that RE was not applicable. Considering all the possible cases, RE was applicable in only 14 of the 48 possible cases (29%), as itemized in the table.

The results in the table show a strong dependence on point span. Only one of the 14 applicable instances of RE occurred over a 3-9-27 point span. The other 13 instances occurred over 9-27-81 point spans. These later spans in the sampling progression exhibited many more instances of an asymptotic regime of convergence. There appears to be no significant effect of Halton generating base, 2-3 versus 5-7. Seven of the applicable instances of RE occurred with Base 2-3 sampling, and seven occurred with Base 5-7 sampling. Within just kriging, two instances occurred with Base 2-3 and three with Base 5-7. Within just MLS, five instances occurred with Base 2-3 and four with Base 5-7. There does seem to be an impact of interpolation method; a total of five instances occurred with kriging, but nine occurred with MLS. Over the 9-27-81 point spans, MLS results entered the asymptotic regime in nine instances, but kriging achieved this regime only four times –less than half the MLS rate, based on the same Halton sample sets.

Beyond issues of applicability to only 29% of the eligible cases here, questions arise regarding RE effectiveness when it *is* applicable. Table 4 features the 14 cases that qualified for RE. The listed empirical rates of convergence were calculated from eqn. (3). The extrapolated values from use of eqn. (2) are listed under the column labeled ‘ T_{RE} ’ in the table. The next column, labeled ‘ $|T_f - T_{exact}| - |T_{RE} - T_{exact}| > 0$?’ asks whether the RE value is closer to the exact result than the fine-resolution RSA result (at 27 or 81 points, depending on whether the particular sampling progression was 3-9-27 or 9-27-81). A ‘Y’ indicates ‘Yes’ to this question, and a ‘N’ indicates ‘No’ to the question. The No’s dominate by a count of 10 to 5. Hence, in the 29% of eligible cases where RE was applicable, the RE estimate was actually worse in the large majority of cases than the last RSA result in the 3-9-27 or 9-27-81 progression.

Table 3. RSA progressions where Richardson Extrapolation was potentially applicable. ‘Y’ indicates all requirements were met and RE was applied.

Calculated quantity	Halton Point Progression (# samples)	kriging		Moving Least Squares	
		Base 2-3	Base 5-7	Base 2-3	Base 5-7
EP($r=1.5$)	3, 9, 27	N	N	N	N
	9, 27, 81	Y	N	N	N
EP($r=1.0$)	3, 9, 27	N	N	N	N
	9, 27, 81	Y	N	Y	N
EP($r=0.5$)	3, 9, 27	N	N	N	N
	9, 27, 81	N	Y	Y	Y
EP($r=0.2$)	3, 9, 27	N	N	N	N
	9, 27, 81	N	N	Y	Y
mean of response	3, 9, 27	N	N	N	N
	9, 27, 81	N	N	Y	Y
stndrd. dev. of response	3, 9, 27	N	Y	N	N
	9, 27, 81	N	Y	Y	Y

Not seeing reliable accuracy improvement with RE, we then ask whether we can at least put accurate error bars on the RE estimates, that reliably contain the exact results. The penultimate column in the table, labeled ‘ $|T_{RE-Tf} - |T_{RE-Textact}| > 0 ?$ ’, applies the Simple Delta method for error bar estimation. The absolute difference between the RE result and the fine-sampling result sets the error-bar extents above and below the RE estimate. The column asks whether the exact result is within this error bar. We get 8 results in the affirmative and 6 in the negative (a 57% success rate). So, even though the RE estimates were not a reliable improvement on the RSA results, at least the SD error bars on the RE estimates contained the exact result the majority of the time. Still, this is little consolation for RE estimates that are often worse than the RSA results. Furthermore, the success rate of SD error bars about the RSA estimates is much higher, above 80% in our investigation.

Finally, in the last column labeled ‘ $|T_{RE-Tf} - |Tf-Textact}| > 0 ?$ ’ we ask whether we can use the RE results to derive reliable error bars about the last (and usually best) RSA estimates in the 3-9-27 or 9-27-81 sampling progressions. Here we use a “backward” Simple Delta approach. For example, we assume the 27-point RSA estimate is the best from a 3-9-27 point RSA progression, and construct an error bar around the 27-point result by differencing from the RE estimate. In this context, the RE machinery is just a relatively complex way to get a backward SD error estimate on the most-refined RSA result in the progression. Error bars so constructed happened to contain the exact result 7 out of the 14 possible times (as itemized in the table). This is only a 50% success rate versus the >80% success rate of the forward SD method that did not employ the RE machinery.

Hence, it seems that the RE machinery is not beneficial for our purposes here. We make one last observation that RE worked notably better with the MLS interpolation than with kriging (for which, only one RE-affirming results exists in Table 4).

Table 4. When Richardson Extrapolation is applicable, is it effective?

Case where RE was applicable	empirical convergence rate, p	T_{RE}	$\frac{ T_f - T_{exact} }{ T_{RE} - T_{exact} } > 0 ?$	$\frac{ T_{RE} - T_f }{ T_{RE} - T_{exact} } > 0 ?$	$\frac{ T_{RE} - T_f }{ T_f - T_{exact} } > 0 ?$
krig, $r=1.5$ 9-27-81, Base2-3	3.708	0.000264	N	N	N
krig, $r=1.0$ 9-27-81, Base2-3	5.711	0.007781	N	N	N
krig, $r=0.5$ 9-27-81, Base5-7	1.367	0.314319	N	N	N
krig, std.dev. 3-9-27, Base5-7	1.728	0.167033	N	N	Y
krig, std.dev. 9-27-81, Base5-7	1.883	0.166636	N	N	N
MLS, $r=1.0$ 9-27-81, Base2-3	3.430	0.007510	Y	Y	N
MLS, $r=0.5$ 9-27-81, Base2-3	1.747	0.413181	N	Y	Y
MLS, $r=0.5$ 9-27-81, Base5-7	0.539	0.079334	N	Y	Y
MLS, $r=0.2$ 9-27-81, Base2-3	3.573	0.984414	Y	Y	N
MLS, $r=0.2$ 9-27-81, Base5-7	0.614	0.987953	N	N	Y
MLS, mean 9-27-81, Base2-3	3.460	0.512939	Y	Y	N
MLS, mean 9-27-81, Base5-7	2.028	0.508843	N	Y	Y
MLS, std.dev. 9-27-81, Base2-3	2.012	0.165706	Y	Y	Y
MLS, std.dev. 9-27-81, Base5-7	1.452	0.171685	N	Y	Y

Summary and Conclusion

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 replacement for the actual function. Monitoring convergence heuristics of progressive response surface approximations can help in this regard, and to get error estimates once the sampling is stopped. The latter is the principle investigative point of this paper.

A major difficulty here was that, even though sampling was incremented at a constant rate of refinement, computed results with the RSAs did not reliably converge at constant rates. In fact, the convergence behavior of com-

puted global and local quantities of response was typically quite erratic, often even oscillatory, creating difficulty in the convergence and error assessments.

Initial indications based on our observations are that the very straightforward Simple Delta method of error estimation was relatively successful. In this method, error bars on RSA estimates are obtained based on changes in RSA predictions that accompany additions of data samples to the approximation. We achieved over 80% success rates over the large number of diverse cases investigated here (60 or 72 in number, depending on whether certain “anomalous” cases were counted or not). However, success tendencies may be very problem dependent and certainly depend on the size of the sample increments. For example, would SD error estimates work as successfully under a doubling (say) of sample points, such that the progression would go from the lowest reasonable number (3 points in 2-D) to 6, 12, 24 points, etc? If so, then the slower rate of growth in this sampling progression could perhaps reduce the total number of samples required to decide when the RSA is accurate enough to stop sampling. Certainly, much investigation remains to be done to better characterize the general prospects of Simple Delta error estimation in general problems and under various sampling incrementation schemes.

Application of the Richardson Extrapolation methodology rested on a somewhat tenuous analogy with grid refinement of partial differential equations. Even if applicable in principle (still a very open question), a progression of at least three RSA results is required before RE is applicable, as opposed to just the two necessary for the SD method. Furthermore, of these fewer opportunities, the often erratic convergence behavior limited RE applicability to only 29% of the opportunities. Of these 29% of opportunities where RE was applicable, RE was found to improve upon RSA estimates and SD error bars in some cases, but only in a minority of the cases. Hence, we conclude that the added complexity of the RE machinery is not beneficial for our purposes here.

We noted some evidence that MLS generally worked better than kriging, but this may not be due to an innate advantage of the MLS method, but due to our limited experience in implementing the kriging method. No significant effect of Halton Sampling prime-number generator Base (2-3 versus 5-7) was evident in the results here.

We also note that we manually intervened in the kriging and MLS methods to set the methods’ free parameters. In doing so, we used knowledge of the global response function shape to optimize the spatial correlation length parameters input to these methods. We intervened in order to ensure favorable conditions for the applicability of the RE method where possible, in order to see whether the method might work (even if under the artificially favorable conditions). We found that even under well-controlled favorable conditions, RE was only mildly successful. In contrast, the Simple-Delta error estimation approach worked more successfully and in a greater variety of circumstances.

Although kriging and MLS have the potential for superior response surface accuracy (versus, e.g., global polynomial regression), at this point we are not confident that they can be used autonomously in general settings. It appears that much more work needs to be done to make them more robust and autonomous methods that could be used in automated procedures for surrogate-based optimization and uncertainty propagation.

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.
- Giunta, A.A., Dudley, J.M., Narducci, R., Grossman, B., Haftka, R.T., Mason, W.H., Watson, L.T., (1994) “Noisy Aerodynamic Response and Smooth Approximations in HSCT Design,” paper AIAA-94-4376, 5th AIAA Multidisciplinary Analysis and Optimization Symposium, Panama City, Fla., Sept. 7-9, 1994.
- Giunta, A.A., and Watson, L.T., (1998) “A Comparison of Approximation Modeling Techniques: Polynomial Versus Interpolating Models,” paper AIAA-1998-4758 presented at the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St. Louis, MO, Sept. 2-4, 1998.
- Iman, R.L., and Shortencarier, M.J. (1984), “A FORTRAN77 Program and User’s Guide for the Generation of Latin Hypercube and Random Samples to Use with Computer Models,” Sandia National Laboratories report SAND83-2365 (RG), printed March 1984.

- 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 Tutorial for SIGGRAPH 2003, Stanford University Statistics Dept., notes available at: <http://www-stat.stanford.edu/~owen/reports/siggraph03.pdf> Monte
- Patterson, T.N.L., (1968) "The Optimum Addition of Points to Quadrature Formulae," *Math. Comp.*, 22, pp. 847-856.
- P/Thermal Analysis Code Theory Manual (1991), Section 4.2.4 Convergence Criteria and Error Estimation, PDA Engineering, Costa Mesa, CA.
- Roache, P.J., (1998) *Verification and Validation in Computational Science and Engineering*, Hermosa Publishers, New Mexico.
- Romero, V.J. (1998) "Optimization and Non-Deterministic Analysis with Large Simulation Models: Issues and Directions," paper in ASME HTD-Vol. 357-2, 7th ASME/AIAA Joint Thermophysics and Heat Transfer Conference, Albuquerque, NM, June 15-18, 1998.
- 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., J.V. 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.
- Romero, V.J., L.P. Swiler, A.A. Giunta, (2004a) "Construction of Response Surfaces based on Progressive-Lattice-Sampling Experimental Designs with Application to Uncertainty Propagation," *Structural Safety*, 26, pp. 201-219.
- Romero, V.J., T. Krishnamurthy, L.P. Swiler, (2004b) "Progressive Response Surfaces," 9th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, Albuquerque, NM, July 26-28.
- Romero, V.J., J.V. Burkardt, M.D. Gunzburger, J.S. Peterson, (2005a) "Comparison of Pure and "Latinized" Centroidal Voronoi Tessellation Against Other Statistical Sampling Methods," accepted for special issue of *Reliability Engineering and System Safety* devoted to Sensitivity Analysis of Model Output (SAMO) Conference, Santa Fe, NM, March 8-11. Conference version of paper available at <http://library.lanl.gov/ccw/samo2004/>.
- Romero, V.J., R. Slepoy, L.P. Swiler, A.A. Giunta, T. Krishnamurthy, (2005b) "Error Estimation Approaches for Progressive Response Surfaces," paper AIAA-2005-1822, 46th Structures, Structural Dynamics, and Materials Conference, April 18-21, 2005, Austin, TX.
- Roux, W.J., Stander, N., and Haftka, R.T., (1996) "Response Surface Approximations for Structural Optimization," *6th Multi-Disciplinary Analysis and Optimization Symposium*, Bellevue, WA, Sept. 4-6 (paper AIAA 4042-CP).
- Simpson, T. W., Booker, A. J., Ghosh, D., Giunta, A. A., Koch, P. N. and Yang, R.-J., (2004) "Approximation Methods in Multidisciplinary Analysis and Optimization: A Panel Discussion," *Structural and Multidisciplinary Optimization*, Vol. 27, No. 5, pp. 302-313.
- Strang, G., and Fix, G.J. (1973), *An Analysis of the Finite Element Method*, Prentice Hall, Englewood Cliffs, NJ.
- 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*, Bellevue, WA, Sept. 4-6 (AIAA paper 96-4040-CP).