



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

A Global Sensitivity Analysis Methodology for Multi-physics Applications

C. H. Tong, F. R. Graziani

February 6, 2007

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.



UCRL-NUMBER

A GLOBAL SENSITIVITY ANALYSIS METHODOLOGY FOR MULTI-PHYSICS APPLICATIONS

CHARLES TONG
CENTER FOR APPLIED SCIENTIFIC COMPUTING

FRANK GRAZIANI
DNT B DIVISION



A Global Sensitivity Analysis Methodology for Multi-physics Applications

Charles Tong

Frank Graziani

Lawrence Livermore National Laboratory

Livermore, CA 94550-0808

Abstract

Experiments are conducted to draw inferences about an entire ensemble based on a selected number of observations. This applies to both physical experiments as well as computer experiments, the latter of which are performed by running the simulation models at different input configurations and analyzing the output responses. Computer experiments are instrumental in enabling model analyses such as uncertainty quantification and sensitivity analysis. This report focuses on a global sensitivity analysis methodology that relies on a divide-and-conquer strategy and uses intelligent computer experiments. The objective is to assess qualitatively and/or quantitatively how the variabilities of simulation output responses can be accounted for by input variabilities. We address global sensitivity analysis in three aspects: methodology, sampling/analysis strategies, and an implementation framework. The methodology consists of three major steps: (1) construct credible input ranges; (2) perform a parameter screening study; and (3) perform a quantitative sensitivity analysis on a reduced set of parameters. Once identified, research effort should be directed to the most sensitive parameters to reduce their uncertainty bounds. This process is repeated with tightened uncertainty bounds for the sensitive parameters until the output uncertainties become acceptable. To accommodate the needs of multi-physics application, this methodology should be recursively applied to individual physics modules. The methodology is also distinguished by an efficient technique for computing parameter interactions. Details for each step will be given using simple examples. Numerical results on large scale multi-physics applications will be available in another report. Computational techniques targeted for this methodology have been implemented in a software package called PSUADE.

1 Introduction

A mathematical model is characterized by a set of equations (for example, partial differential equations), a number of parameters (for example, Prandtl number), and other variables (for example, initial and boundary conditions, forcing functions) which together enable a well-posed solution to the model. Since most real-life applications are too complex for the existence of analytical solutions, these mathematical models are often discretized and solved on computers. The corresponding simulation model is then validated against data, from, ideally, well-orchestrated physical experiments. Many sources of uncertainties, however, are at play to hinder the effort of reconciling the simulation with the experimental data. For example, insufficient understanding of the physical system may result in model uncertainties, that is, the formulation of an incorrect set of equations. Parameter uncertainties may arise as a result of measurement errors or lack of data. Furthermore, the discrepancies between the mathematical model and the corresponding simulation model may arise due to choices of numerical algorithms and code correctness. Therefore, building a credible simulation capability requires that model assumptions are clearly and meticulously stated and (model and parameter) uncertainties are clearly defined and assessed.

In this report we consider mathematical models arising from multi-physics applications typified by some or all of the following features:

- the phenomena (output responses) may be highly nonlinear (with the inputs),
- the simulation may be expensive to evaluate,
- there are many uncertain variables,
- the uncertain variables may be correlated,
- there may be many output responses that are of interest,
- the output responses may have complex constraints, and
- the output responses may not be deterministic.

In this context, it may be useful to explore how the output responses behave when the inputs are subject to variations, and specifically, which input variables contribute most significantly to output variations. This question can be answered by sensitivity analysis (SA) studies.

Sensitivity analysis has found widespread applications in many disciplines. In the following we quote a few definitions of sensitivity analysis from different applications.

European Commission Joint Research Centre: Sensitivity analysis is the study of how the variation in the output of a model (numerical or otherwise) can be apportioned, qualitatively or quantitatively, to different sources of variation.

Definition in financial world: Investigation into how projected performance varies along with changes in the key assumptions on which the projections are based.

EPA definition: Sensitivity analysis is a formalized procedure to identify the impact of changes in various model components on model output. It is an integral part of simulation experimentation and may influence model formulations. It is commonly used to examine model behavior. The general procedure is to define a model output variable that represents an important aspect of model behavior. The values of various inputs of the model are then varied and the resultant change in the output variable is monitored. Large changes in the output variable imply that the particular input varied is important in controlling model behavior.

Uncertainty analysis (UA) concerns with the quantitative characterization of the relationship between input and output uncertainties. Sensitivity analysis, on the other hand, concerns with assessing the individual contribution of the uncertain parameters to the output variabilities. Therefore, there is a close tie between UA and SA, in the sense that SA is a natural next step to UA.

SA methods are generally grouped into two classes: local SA and global SA methods. Local SA computes or approximates partial derivatives of the outputs with respect to individual input variables at some nominal coordinate. The partial derivatives can be obtained by solving, together with the original set of equations for the model, a set of sensitivity equations. Alternately, the simulation source code can be preprocessed with automatic differentiation tools resulting in a modified code that yields additional derivative information. Still a third alternative is to compute a finite difference approximations to the derivatives using multiple simulation runs. Global SA analysis, on the other hand, studies the effects of the variations of input parameters on the outputs in the entire allowable ranges of the input space. Saltelli *et al.* [20, 22] have defined global methods by two properties:

1. The inclusion of influence of scales and shapes of the probability density functions for all inputs.
2. The sensitivity estimates of individual inputs are evaluated while varying all other inputs (multi-dimensional averaging).

Many argue that global SA is more suitable in revealing the true sensitivity information especially for applications with wide variability ranges, large nonlinearities and complex interactions exemplified by many multi-physics models. However, it should be noted that local SA and global SA are not mutually exclusive, since local SA techniques can play important roles in global sensitivity studies (which is one of our current research topics). This report is dedicated exclusively to global sensitivity analysis techniques that are non-intrusive (that is, the simulation model is treated as a “black box” rather than having to be modified.) In Section 2 we describe the basic steps of the methodology. Section 3 and 4 provide the background for the second and third step of the methodology. Section 5

discusses ways to validate this global sensitivity analysis methodology. Section 6 describes an implementation framework to facilitate the use of this methodology. Finally, a brief summary will be given in Section 7.

2 A Global Sensitivity Analysis Methodology

The sensitivity analysis methodology described in this report is sampling-based (“non-intrusive”) in which the model is executed repeatedly for combinations of values sampled from the distribution (assumed known) of the input variables. The essence of this methodology is to use a two-stage divide-and-conquer approach to deal with high dimensionality (large number of uncertain inputs). This two-stage approach has been advocated in several other applications [20, 26]. Our methodology, in addition to incorporating the two-stage analysis, also includes a recursive mechanism to handle individual physics sensitivities, an enhanced screening method, and an efficient computational technique for two-way (two-input) interaction studies.

Prior to performing sensitivity analysis, much diligence is required to compile a detailed specification of the model to be studied, since the results concluded from the analysis are valid only with respect to the given specification. The specification includes the simulation model, uncertain parameters that are to be varied (for example, material strength parameters); uncertain parameters which are fixed (for example, the grid resolution or algorithmic parameters such as convergence tolerances) in the current study but which may affect the outcome of the analysis if they are allowed to vary; the simulation output responses; and any other assumptions about the model. Subsequent steps in our global sensitivity analysis methodology consists of the following:

1. Construct a complete description of the uncertain input parameters (that is, the ranges and forms of the distributions) from
 - (a) experimental data (from resident physicists),
 - (b) literature search (data from published papers elsewhere),
 - (c) expert judgment (resident physicists and code developers), and/or
 - (d) results of applying this methodology recursively to individual physics modules.
2. If the number of uncertain input parameters is relatively small (say, less than 10), skip to Step (3). Otherwise, perform a down-select screening analysis on all uncertain parameters. There are several alternatives as to which screening method is best. In the absence of sufficient knowledge about the functional relationship between the input and output variables, we recommend the model-independent Morris-one-at-a-time (MOAT) design [14]. So, given the ranges and forms of the input distributions.
 - (a) Generate an MOAT design matrix.

- (b) Perform the simulation experiments (which consists of preprocessing, running multiple simulations, and post-processing).
 - (c) Compute the MOAT measures and create screening and scatter plots.
 - (d) Analyze the result and identify a subset of input parameters for Step (3).
3. Perform a quantitative sensitivity analysis via variance decomposition techniques.
- (a) If the simulation is computationally intensive and the output response is a “relatively” smooth function of the uncertain inputs, we should consider using response surfaces (other names are: surrogate functions and emulators) to represent the simplified models. For the response surface approach to be effective, a space-filling sampling design is desired. Direct variance decomposition, on the other hand, can use sampling design such as the replicated Latin hypercube design for first order sensitivity indices, the replicated orthogonal array for second order sensitivity indices, and the extended Fourier Amplitude Sampling Test (FAST) or direct integration methods for total sensitivity indices.
 - (b) Perform the simulation experiments.
 - (c) Analyze the results using statistical and graphical analysis techniques such as sensitivity indices, and scatter and main effect plots.
4. Expend effort to reduce uncertainties of the most important inputs. If needed, repeat this whole process with new set of input ranges.

The most important yet challenging task in this methodology is the first step. It is important because proper prescription of the ranges and shapes of the input distributions can dramatically alter the outcome of the analysis. It is challenging because oftentimes these ranges are obtained by carefully analyzing data from physical experiments. For large scale multi-physics applications, physical experiments for the entire multi-physics processes may not be feasible, but smaller physical experiments on individual physics may be viable (For example, by observing the actual bubble growth rates of the Rayleigh-Taylor experiments, one can prescribe realistic uncertainty bounds for some hydrodynamics parameters). For this reason, we emphasize the benefit of applying this methodology recursively for multi-physics applications (meaning applying this same methodology to individual physics for the purpose of determining credible ranges for some of the uncertain parameters). Validating simulation data against experimental data for individual physics may provide an indispensable tool for making the result of Step (1) more defensible.

For multi-physics applications which are plagued by complex interactions and nonlinearities, it is important that the techniques in Step (2) and (3) above are model-independent, that is, they do not make too many assumptions about the model behaviors. In addition, the techniques should include rigorous analysis of parameter interactions and correlations.

Finally, we would like to emphasize on the iterative nature of this methodology. After important parameters are identified, their effects can be reduced by tightening the parameter ranges through better understanding of the process. Once accomplished, other input parameters may emerge as important in the next iteration. Also, the iteration will safeguard against the Type II errors (when important parameters are treated as unimportant) in the previous iteration. Furthermore, for very large number of uncertain parameters (say several hundreds), the number of parameters identified as important in the screening analysis may still be too large for the subsequent quantitative step. In this case, additional divide-and-conquer steps coupled with the iterative loop will make the task more manageable.

3 Parameter Screening

There are a number of approaches to perform sensitivity analysis when the number of uncertain parameters is large. One approach is to perform quantitative sensitivity analysis on all input parameters simultaneously using Bayesian techniques (this may be very expensive or many assumptions about the probabilistic structure of the application have to be made). Another approach is to put all the uncertain parameters into a small number of groups and then perform group sensitivity analysis. Still another approach, which is the one explored in this report, uses a two-phase divide-and-conquer strategy [20, 26]. The first phase, to be discussed in this section, consists of employing a low-cost ($O(M)$ simulation runs where M is the number of uncertain input parameters) coarse sampling and the corresponding analysis method to identify a subset of parameters which are candidates for more detailed analysis in Step (3). Screening methods generally perform well when the number of important inputs is small compared to the total number of inputs under investigation.

For the choice of screening methods, model independence is an important consideration. If it is known that the output variable behaves approximately linearly or monotonically with all the inputs and there is little parameter interactions, we can use a relatively inexpensive screening design such as the Plackett-Burman design [18], which requires only $M + 1$ simulation runs (again, M is the number of inputs). However, if no such knowledge is available, the Morris one-at-a-time (MOAT) method [14] is recommended. The MOAT method has been demonstrated to be effective in capturing parameters that have direct impact on the output of interest throughout the parameter space (that is, when local features are not dominant). Screening methods such as MOAT can also be used to provide an independent verification of prior knowledge about the model. For example, if prior experience tells us that there is no interaction between parameter X and Y, screening provides another venue to confirm this.

3.1 The MOAT Designs

The Morris method is called one-at-a-time (OAT) method because the design varies one input at a time. To generate the Morris design, a base sample point $X^{(1)}$ is created such that each of the M (M is the number of input parameters) components of $X^{(1)}$ is randomly selected from

the set $\{0, 1/(p-1), 2/(p-1), \dots, 1\}$, where p is a pre-selected integer (for example, $p = 4$). The second sample point $X^{(2)}$ is created from $X^{(1)}$ by perturbing one of its inputs by Δ which is a pre-selected multiple of $1/(p-1)$ (an even p and $\Delta = p/(2p-2)$ are recommended to preserve equal probabilities). Subsequent sample points $X^{(i)}$, $i = 3, \dots, M+1$ are created in a similar way, namely from $X^{(i-1)}$ by perturbing one of its inputs which has not been perturbed before. After perturbing all inputs, we have $M+1$ sample points. A necessary condition to satisfy in generating the sample points is that all $X^{(i)}$ still lie in the design space (for example, $[0, 1]^M$).

This process is captured in the following mathematical form:

$$B^* = (J_{M+1,1}X^{(1)} + (\Delta/2)[(2B - J_{M+1,M})D + J_{M+1,M}])P \quad (1)$$

where B^* is an $(M+1) \times M$ normalized design matrix (each row is one normalized sample point), $J_{m,n}$ is an $m \times n$ matrix of all 1's, D is an $M \times M$ diagonal matrix in which each diagonal element is either $+1$ or -1 with equal probability, $X^{(1)}$ is a row vector having the base sample, P is an $M \times M$ permutation matrix, and

$$B = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

To insure this design creation process is a random selection from the distribution of elementary effect (to be defined later), the process uses three randomizations: (1) the base sample point is randomly selected, (2) the direction of the perturbation is random (that is, the creation of D), and (3) the choice of which input to perturb next is a randomized process (reflected by P). After B^* has been generated, this normalized sample has to be mapped onto the actual parameter ranges and distributions before running them through the simulations. This procedure is repeated r times to ensure enough regions in the design space has been explored.

3.2 The MOAT Analysis

After these sample points have been evaluated, the elementary effects for the inputs can then be calculated by:

$$d_{c(k)} = \frac{f(X^k) - f(X^{k-1})}{\Delta}, k = 2, \dots, M+1$$

where $d_{c(k)}$ is defined as the elementary effect for input $c(k)$ (c maps k to its true input number). Note that since we randomly select which input to perturb next, the elementary effect computed using the k -th and $(k+1)$ -th points in general is not the elementary effect for the k -th input which explains the $c(k)$ mapping.

Now that we have r elementary effects for each input (let's label them d_i^j for input i and replication j .) Morris proposed two sensitivity measures to analyze the data: μ which estimates the overall effect of each input on the output, and σ which estimates the higher order effects such as nonlinearity and interactions between inputs. The formulas for them are:

$$\mu_i = \frac{1}{r} \sum_{j=1}^r d_i^j, \quad \text{and} \quad \sigma_i = \sqrt{\frac{\sum_{j=1}^r (d_i^j - \mu_i)^2}{r}},$$

respectively. Campolongo *el al.* [20] proposed an improved measure μ^* in place of μ where

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^r |d_i^j|.$$

If μ_i^* is substantially different from zero, it indicates that input i has an important “overall” influence on the output. A large σ_i implies that input i has a nonlinear effect on the output, or there are interactions between input i and the other inputs.

The MOAT measures can also be shown graphically by screening plots, which have the x- and y- axes the modified means and standard deviations, respectively. Each input is represented on the screening plot by a point with coordinate (μ_i^*, σ_i) . The plot can then be divided into four quadrants with the two quadrants on the right containing sensitive inputs. Expert judgment should be used when defining the four quadrants through intense discussion between physicists and analysts.

An example screening plot is given in Figure 1. This example is taken from Morris' original paper:

$$y = \beta_0 + \sum_{i=1}^{20} \beta_i w_i + \sum_{i<j}^{20} \beta_{ij} w_i w_j + \sum_{i<j<k}^{20} \beta_{ijk} w_i w_j w_k + \sum_{i<j<k<l}^{20} \beta_{ijkl} w_i w_j w_k w_l$$

where $w_i = 2(X_i - 0.5)$ except for $i = 3, 5, 7$, where $w_i = 2(1.1X_i/(X_i + 0.1) - 0.5)$. The other coefficients are assigned as:

$$\begin{aligned} \beta_i &= 20, \quad i = 1, \dots, 10, \\ \beta_{ij} &= -15, \quad i, j = 1, \dots, 6, \\ \beta_{ijk} &= -10, \quad i, j, k = 1, \dots, 5, \\ \beta_{ijkl} &= 5, \quad i, j, k, l = 1, \dots, 4, \end{aligned}$$

and the remainder of first and second order coefficients were generated independently from a normal distribution with zero mean and unit standard deviation, and the remainder of third and fourth order coefficients were set to 0. Here we observe that parameter 1 through 10 are on the right half plane, demonstrating their importance. Parameter 1 through 7, in addition, are located on the upper right quadrant, demonstrating strong nonlinearities and/or interactions.

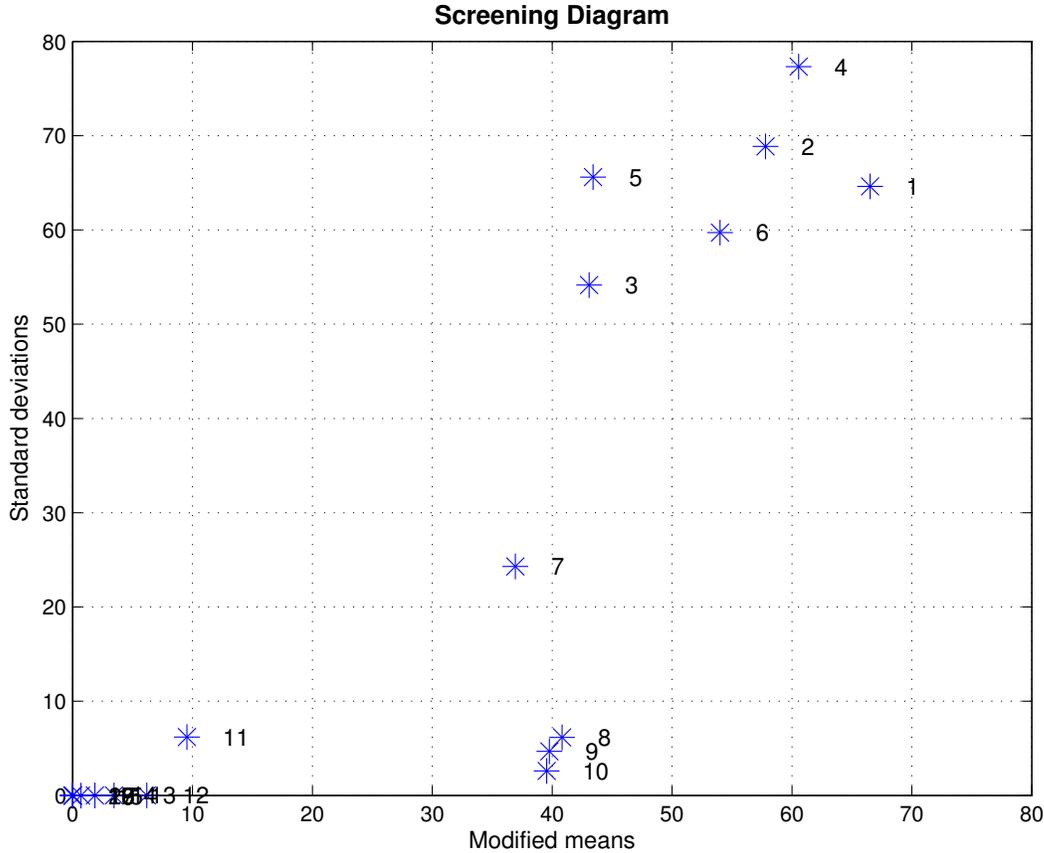


Figure 1: An example screening plot

3.3 Useful Enhancement to Screening Studies

Through our experience with working with a large-scale multi-physics code, we have accumulated a few suggestions for making this screening process more informative and productive. We discuss some of them in the following.

3.3.1 More Graphical Analyses

In addition to the screening plots, scatter plots are also very helpful. The scatter plot for the Morris example is given in Figure 2. Each triangular symbol represents, for the corresponding input, an elementary effect. The orientation and color of the triangle indicates at which level the input (value) was at when the elementary effect was calculated. The scatter plot, in addition to giving detailed information on the standard deviations of the elementary effects, also help identify anomalies as well as separate nonlinearities from interactions.

In addition to scatter plots, bar graphs showing the modified means and standard deviations can also be revealing.

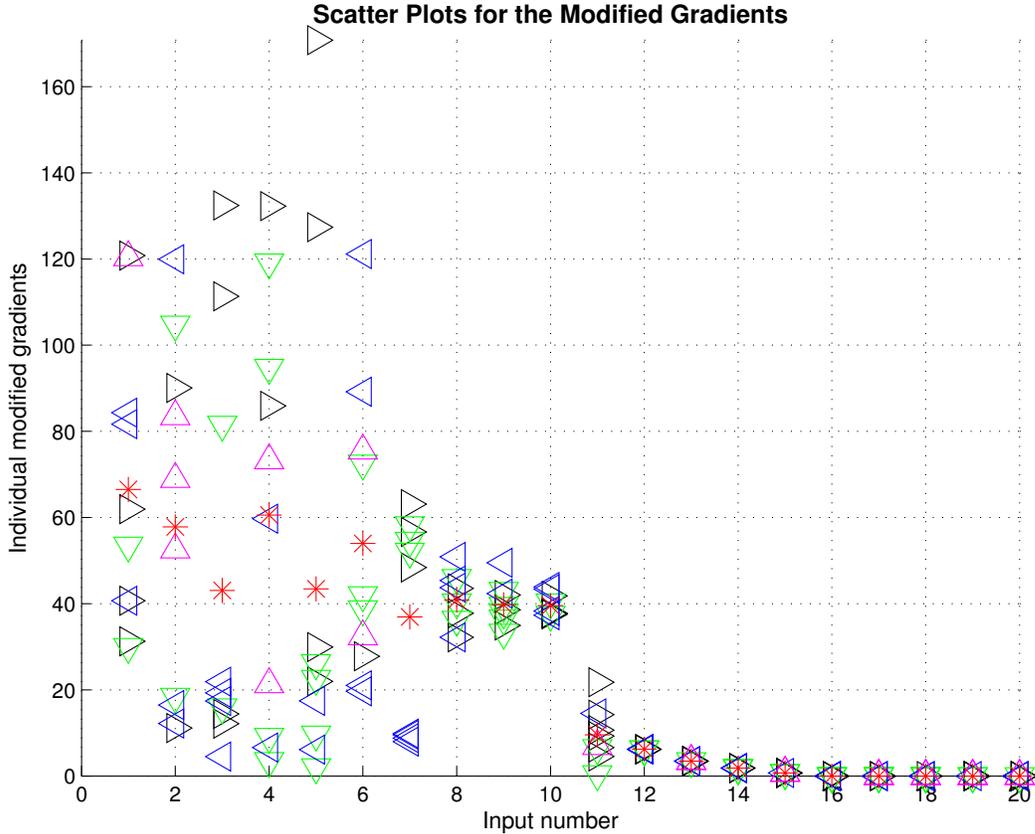


Figure 2: An example scatter plot

3.3.2 Variable and Multiple Resolutions

The Morris method uses coarse sampling to explore input sensitivities. For example, when $p = 4$, each input is assigned 4 levels. The total number of possible elementary effects for each input is $p^{M-1}[p - \Delta(p-1)]$. A large measure of central tendency for the distribution indicates an input with an important “overall” influence on the output on this particular p scale. If, however, we have the knowledge that the output varies more quickly with a particular input, then it will be more sensible to use a large p for this input. Hence, additional knowledge on the model behavior may be helpful in gauging good values of p for each input. This modification can be easily modified from the original Morris method.

Furthermore, if the value of p for a given input is not known and additional resources are available for parameter screening, then a second pass with finer resolution (larger p) can be used to check the earlier assumptions.

3.3.3 Selective Replication

Multi-physics applications often encounter outliers in computing the elementary effects of some inputs. After checking that these outliers are not due to code or other errors, additional

runs may be needed to further study whether these outliers are indeed isolated ones. The idea of selective replication is to collect the subset of inputs which we want to study more and create a reduced Morris design for these inputs. The number of replications used for this second step can be determined adaptively when coupled with hypothesis testing of the elementary effects. The selective replication has been implemented in our software package.

3.3.4 Local Refinement

A simple yet powerful idea to study anomalies is the use of local refinement. Suppose for input i , one of its r elementary effect is an outlier. After careful examination, the large gradient computed for the i -th input at $X_i^{(1)}$ and $X_i^{(2)}$ is due to one of its end points (say, at $X_i^{(2)}$) appearing to be an anomaly. Focus should then be on investigating whether the offending sample point ($X_i^{(2)}$) is set up and run correctly. In the absence of any evidence of gross errors, one may want to refine the region between $X_i^{(1)}$ and $X_i^{(2)}$ (while keeping all other inputs the same) and study what happens when X_i approaches from $X_i^{(1)}$ to $X_i^{(2)}$.

3.3.5 Screening for Two-parameter Interactions

The computational requirement of the MOAT method is relatively inexpensive ($r(M + 1)$ where $r \approx 10$ in general). However, one shortcoming of this method is that nonlinearities and parameter interactions are confounded. It is highly desirable to separate nonlinearities from interactions, since negligible interaction between a given parameter with the remaining ones implies that this parameter can be studied independently (that is, decoupled from the rest) in Step (3) due to the additive nature of variance decomposition in this case. To qualitatively study two-parameter interactions, Campolongo and Braddock [1] proposed the extended Morris method. However, the number of simulation runs for this new method is now $O(M^2)$.

In our method implementation, we also provide a qualitative assessment of parameter interaction without additional simulation cost, taking advantage of MOAT's multiple visits to individual input levels. Specifically, let $X_i^{(j)}, j = 1, \dots, p$ (say, $p = 4$) be the p levels used for input i . If the number of replications r is 10, then it is likely that each of the p levels is used more than once. By comparing the elementary effects on the same levels (again computing standard deviations) a rough assessment of interactions can be made. If a more accurate assessment is needed, one can use more replications or employ the extended Morris method.

3.4 Dealing with Correlated Inputs

The original MOAT method can only handle models with non-correlated or orthogonal inputs. Correlations or non-orthogonalities, however, are common in multi-physics applications exhibiting themselves in the form of:

1. joint probability density functions,

2. inequality constraints in some analytical forms, or
3. constraints imposed by results of related physics simulation.

If most or all uncertain inputs are interacting with each other through correlations, the screening process may be greatly complicated. The complication arises because input correlations often disguise as parameter interactions. In many scientific applications, however, input correlations usually appear in groups pertaining to individual physics. In this case, a natural modification is to put all inputs into uncorrelated groups and then analyze qualitative importance based on group contributions. This is reasonable since some inputs become important only because they are correlated with other inputs.

In our global sensitivity analysis methodology, we also accommodate correlations by providing a modified screening technique. To illustrate the idea, let the function be $Y = F(X_1, X_2) = X_2$ with X_1 and $X_2 \in [0, 1]$ which are correlated such that $X_2 < \alpha X_1$ where α is some measure of correlations. We consider the following four scenarios:

1. X_1 has small influence on Y , but it has strong influence on X_2 .
2. X_1 has small influence on both Y and X_2 .
3. X_1 has strong influence on Y , but it has small influence on X_2 .
4. X_1 has strong influence on both Y and X_2 .

The screening procedure should identify X_1 as important for all but the second scenario.

The modified Morris procedure is summarized as follows.

1. Construct an modified MOAT design so that all sample points lie inside the feasible region (that is, the correlation criteria are satisfied);
2. Propagate the MOAT design matrix via simulation runs;
3. Analyze the MOAT measures without considering the constraints. This should identify parameters that are important themselves without the constraints. For example, for the function above, this step will identify X_2 alone (and not X_1) as important; and
4. Examine how the likely unimportant inputs (X_1 here) constrain the important ones. For example, if $\alpha = 0.1$, then we determine that X_1 has strong influence on X_2 (scenario (1)) and hence X_1 should be included in the group of important inputs. If, however, $\alpha = 10$, then X_1 has small influence on X_2 (scenario (2)) and thus can be ignored.

In Step (4) above, if the constraint is a complex function requiring running another physics code, this analysis may require approximating the constraint in some functional form (parametric or nonparametric) using response surface methods.

3.5 An Enhanced MOAT Method

Using some of the enhancements described above, we have implemented an enhanced MOAT design and analysis tool with the following features:

Variable resolution: In order to support this feature, the original Morris equation (1) has been modified to become:

$$B^* = J_{M+1,1}X^{(1)} + 1/2[(2BP - J_{M+1,M})D + J_{M+1,M}]G. \quad (2)$$

Here, in addition to the notations given before, we have a new diagonal matrix G which has diagonal elements $\{\delta_i = 1/(p_i - 1)\}$ where p_i is the number of levels for parameter i . In addition, the initial seed points are stratified to allow a more even exploration of the parameter space.

Selective replications: This feature creates reduced Morris designs for a small subset of input parameters that have been identified for further screening.

Constrained sampling: This feature creates Morris sample points in the parameter space constrained by input correlation in the form of inequalities. The corresponding Morris analysis will take into consideration of the constraints in computing the parameter importance.

Interaction screening: A simple analysis is added to perform a crude analysis of parameter screening. The results can be used to confirm previous knowledge, or to decouple the uncertain parameters for independent studies.

Visual analysis: This feature creates scatter plots and bar graphs for examining individual sample points and individual gradients.

3.6 Additional Notes

The MOAT method has been successfully used across many applications. There are, however, concerns about its efficacy. Specifically, because of the curse of dimensionality (first used by Bellman to describe the exponential increase in volume associated with adding extra dimensions to a mathematical space), any inexpensive attempts to thoroughly explore this high dimensional uncertain parameter space inevitably incur error of judgment. There are two types of possible errors in the screening exercise. The first type (Type I error), in which unimportant parameters are mistaken as being important, is benign in nature, causing only minor nuisance in Step (3) of the aforementioned methodology. The second type (Type II error), in which important parameters are mistaken as being unimportant, is more serious.

To address this concern, we first quote Campolongo *et al.* [2] “Previous exercises, ..., seem to indicate that Morris does not make Type II errors. This is reasonable, since the influence of factors in models follows, according to our experience, a Pareto-like distribution, with few

factors accounting for most of variance, and most of the inputs taking up the remaining bit. One has to build a model artificially, in order to get a case where the influence of factors is more uniformly balanced.” We also quote Morris: “If an input has a direct impact on the outputs throughout the experimental region, even the nature of that impact changes with location, it takes only a few replications to see it. The danger would come with an input that is important only in a small subregion. Then the probability of stumbling across it would be related to the volume of the subregion in which it is active.”

In this section we advocate a probabilistic understanding of the Type II error by using hypothesis testing on the elementary effects. The idea is that for each input variable, the Morris design gives r elementary effects. We can combine the means and the standard errors of means together with bootstrapping in testing more rigorously the likelihood of an Type II error for each input. The hypothesis testing requires physicists and analysts working together to formulate the hypothesis. This test will be described in more details in a related report by Jason Lenderman.

4 Variance Decomposition

This section focuses on using variance as an indicator of input importance. The analysis techniques based on regression coefficients, correlation coefficients and their variants have limited applicability since they assume that the model is nearly linear or monotonic. Our attention is mainly on variance-based measures such as correlation ratio and Sobol’/Saltelli sensitivity indices which are less restrictive on the models. For applications with uncorrelated inputs, it is based on the full decomposition of the model variance $V(Y)$ (another notation for σ^2) into

$$V(Y) = \sum_i V_i + \sum_{i < j} V_{ij} + \sum_{i < j < k} V_{ijk} + \cdots + V_{12\dots M} \quad (3)$$

where $V_i = V(E(Y|X_i))$, $V_{ij} = V(E(Y|X_i, X_j)) - V_i - V_j$, and so on. The total number of terms for M inputs is thus $2^M - 1$. In the present context, we are interested primarily in the main effects (V_i) and two-way interactions (V_{ij}). In the event that the inputs are correlated, the above relationship no longer holds. However, variance-based measures are still useful sensitivity indicators. In this section we describe the computational techniques for main effect and two-way interaction calculations.

4.1 Using Response Surface Models

Response surface methodology (RSM) is a collection of statistical and mathematical techniques useful for developing, improving, and optimizing processes [16]. Provided that the model response Y is smooth enough, RSM can be very helpful in reducing the computational cost for uncertainty quantification. The idea is to create a set of sample points that are space-filling and use them to train a response surface model (or a function approximator). The choice of response surface methods for a given simulation model depends on the knowledge

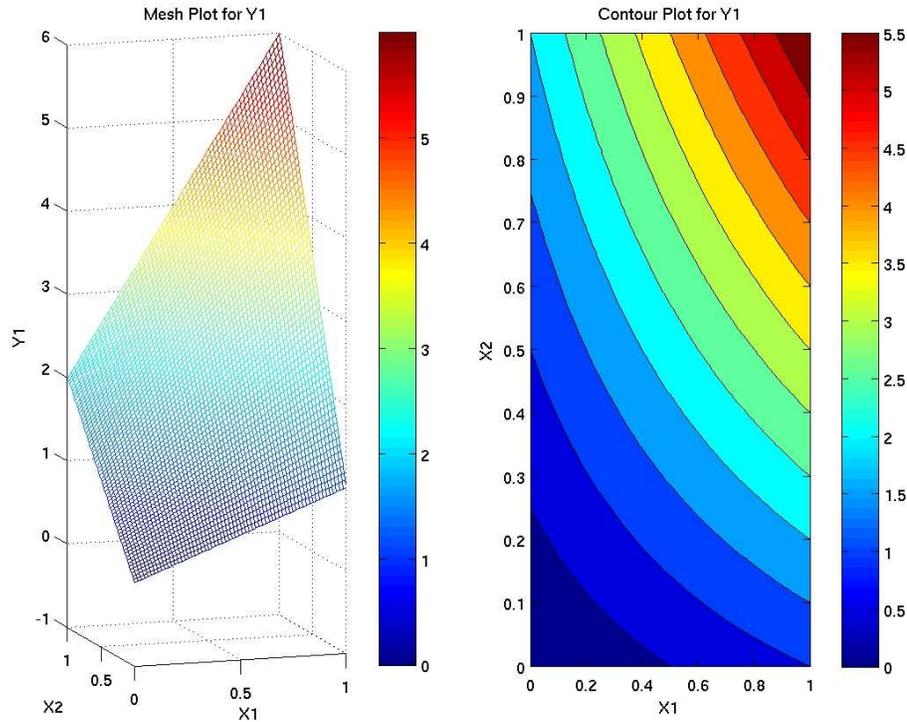


Figure 3: An example response surface plot

about the simulation model itself. If the simulation model is known to be a linear mapping of its uncertain inputs, then a first-order regression model suffices. If no such knowledge is available about the mapping, more general nonparametric models such as splines, neural network, or Gaussian process models may be more appropriate. An example of a response surface is given in Figure 3. Many tools to create and validate response surfaces have been incorporated in our software package.

Once a good response surface model has been constructed with sufficient accuracy, subsequent analysis can rely on this response surface model which is inexpensive to evaluate. This will facilitate the efficacy of the more quantitative analyses that require a large number of evaluations.

4.2 Main Effect Study

The main effect analysis described in this section is due to McKay [11]. The essence of this analysis is the statistical measure called variance of condition expectation. Again, let $E(Y)$ and $V(Y)$ be the prediction mean and variance of an output variable Y , statisticians can tell us that

$$V(Y) = V(E(Y|X_i)) + E(V(Y|X_i)) \quad (4)$$

where X_i is the i -th input. Here the first term on the right hand side is the variance of the conditional expectation of Y , conditioned on X_i ; and the second term is an error or residual term. A simple interpretation of $V(Y|X_i = X_i^*)$ (the conditional variance) is that it is the reduced variance obtained by setting X_i to its true value X_i^* . In reality, X_i is uncertain so X_i^* is not known exactly. Therefore, it makes sense to average the conditional variances sampled at different values of X_i in its variability range, giving rise to $E(V(Y|X_i))$. It should be obvious now that the difference $V(Y) - E(V(Y|X_i)) = V(E(Y|X_i))$ is a sensible indicator of the importance of input i . Another interpretation of the above relationship is that $V(E(Y|X_i))$ measures the variability in the conditional expected value of Y as the input X_i takes on different values. The residual term represents the variability in Y not accounted for by the input subset X_i .

A related measure is the correlation ratio η^2 , which is defined as

$$\eta^2(X_i) = V(E(Y|X_i))/V(Y).$$

A high correlation ratio shows that X_i is important in the variability of the output. If all uncertain inputs are uncorrelated and there are no multi-way interactions, then the sum of the correlation ratios is 1.

There are many ways to compute the correlation ratios. A popular method is to use replicated Latin hypercube sampling together with the corresponding main effect equations. Using Latin hypercube designs, X_i takes on values $X_{ij}, j = 1, \dots, S$ where S is the number of symbols (or levels). This is to be replicated R times so that the total number of simulations (or emulations) required is $N = SR$. The estimator for the expectation conditioned on $X_i = X_{ij}$ is

$$\bar{Y}(X_i = X_{ij}) = \frac{1}{R} \sum_{r=1}^R Y^{(r)}(X_i = X_{ij})$$

where $Y^{(r)}(*)$ is the output corresponding to the r -th replication. The estimator for the total variance is approximated by (for any $i = 1, \dots, M$)

$$V(Y) = \frac{1}{SR} \sum_{j=1}^S \sum_{r=1}^R [Y^{(r)}(X_i = X_{ij}) - \bar{Y}]^2.$$

The estimator for variance of conditional expectation on X_i is approximated by

$$V(Y|X_i = X_{ij}) = \frac{1}{R} \sum_{r=1}^R [Y^{(r)}(X_i = X_{ij}) - \bar{Y}(X_i = X_{ij})]^2,$$

and hence

$$E(V(Y|X_i)) = \frac{1}{SR} \sum_{j=1}^S \sum_{r=1}^R [Y^{(r)}(X_i = X_{ij}) - \bar{Y}(X_i = X_{ij})]^2.$$

Finally, we have

$$V_i = V(E(Y|X_i)) = V(Y) - E(V(Y|X_i)),$$

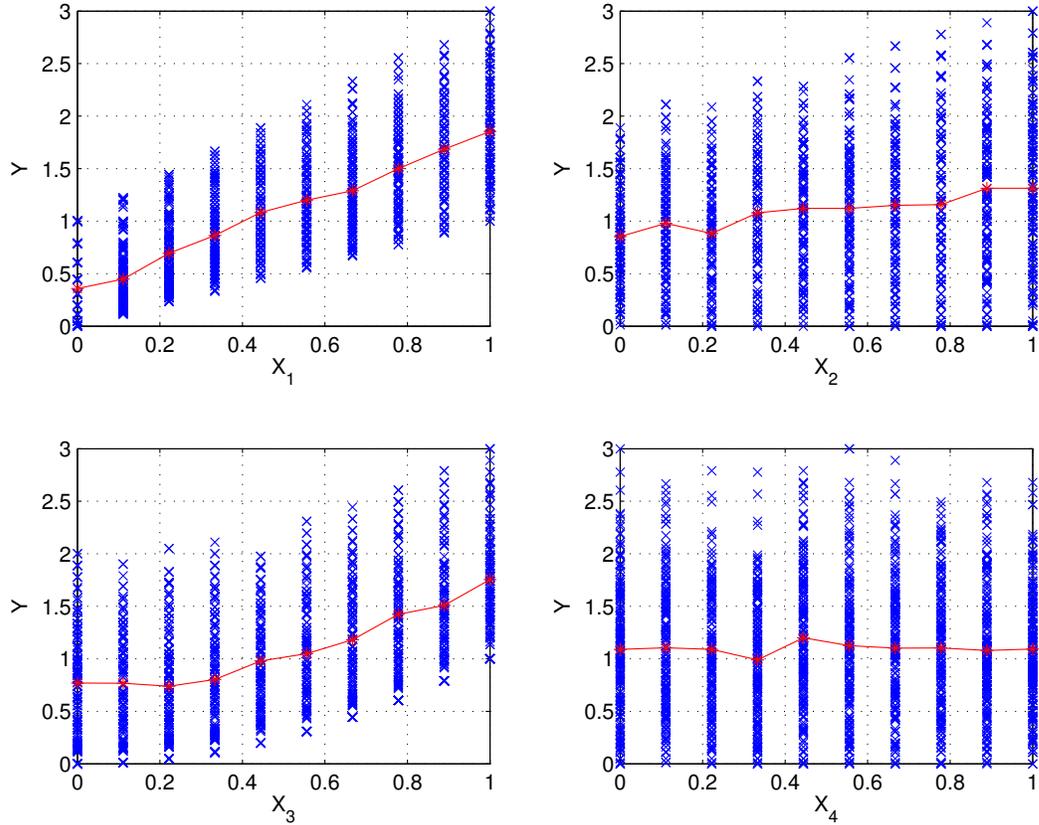


Figure 4: An example scatter plot from replicated Latin hypercube

and the correlation ratio is computed by dividing $V(E(Y|X_i))$ by the total variance. This correlation ratio is a quantitative measure for the input importance. In addition, the scatter plot from the replicated Latin hypercube also gives the trend as X_i takes on different values X_{ij} (by inspecting the line joining the means of all groups, namely X_{ij} 's.) In addition, when R is large enough (say, > 50), the parameter space is sufficiently explored such that the difference between the group maxima and minima also indicates interaction effect (that X_i is interacting with some other inputs) although it will take more studies to identify which other inputs it interacts with. An example is given in Figure 4 where $Y = X_1 + X_1 * X_2 + X_3^2$ for four inputs. Here a close examination of the scatter plots reveals that X_1 and X_2 are involved in some interactions, X_3 is nonlinear, and X_4 has negligible effect.

4.3 Interaction Study

We have extended the idea for main effect analysis to two-way interaction studies for uncorrelated inputs. In this case, we employ the following relationship

$$V(Y) = V(E(Y|X_i, X_k)) + E(V(Y|X_i, X_k))$$

where X_i and X_k are two distinct inputs under consideration. The first term on the right hand side is the variance of the conditional expectation of Y , conditioned on X_i and X_k . Again, the second term is the error or residual term measuring the estimated variance of Y by fixing X_i and X_k . Also, the correlation ratio for the input pair (X_i, X_k) is

$$\eta^2(X_i, X_k) = V(E(Y|X_i, X_k))/V(Y).$$

A high correlation ratio shows that X_i and X_k taken together are important contributors to the output variability. The variance due to the interaction term alone is defined as

$$V(X_i, X_k) = V(E(Y|X_i, X_k)) - V(E(Y|X_i)) - V(E(Y|X_k)).$$

$V(X_i, X_k)$ can be computed using many different techniques, for example, by directly evaluating the corresponding integral. Here we illustrate its evaluation with the use of replicated orthogonal array sampling. Using orthogonal array design with a strength of 2, X_i and X_k take on values $X_{ij}, j = 1, \dots, S$ and $X_{kl}, l = 1, \dots, S$ where S is the number of symbols (or levels). The variance estimator for the expectation conditioned on $X_i = X_{ij}$ and $X_k = X_{kl}$ is

$$\bar{Y}(X_i = X_{ij}, X_k = X_{kl}) = \frac{1}{R} \sum_{r=1}^R Y^{(r)}(X_i = X_{ij}, X_k = X_{kl})$$

where $Y^{(r)}(X_i = X_{ij}, X_k = X_{kl})$ is the output corresponding to $X_i = X_{ij}$ and $X_k = X_{kl}$ in the r -th replication (that is, keeping the two inputs at some fixed values and varying all others). Similarly, the conditional variance is approximated by

$$V(Y|X_i = X_{ij}, X_k = X_{kl}) = \frac{1}{R} \sum_{r=1}^R \left[Y^{(r)}(X_i = X_{ij}, X_k = X_{kl}) - \bar{Y}(X_i = X_{ij}, X_k = X_{kl}) \right]^2.$$

Averaging the conditional variance over the ranges of X_i and X_k , we have

$$\begin{aligned} E(V(Y(X_i = X_{ij}, X_k = X_{kl}))) &= \\ \frac{1}{S^2 R} \sum_{j=1}^S \sum_{l=1}^S \sum_{r=1}^R &\left[Y^{(r)}(X_i = X_{ij}, X_k = X_{kl}) - \bar{Y}(X_i = X_{ij}, X_k = X_{kl}) \right]^2. \end{aligned}$$

The estimator for the total variance is

$$V(Y) = \frac{1}{S^2 R} \sum_{j=1}^S \sum_{l=1}^S \sum_{r=1}^R \left[Y^{(r)}(X_i = X_{ij}, X_k = X_{kl}) - \bar{Y} \right]^2,$$

and

$$V(E(Y|X_i, X_k)) = V(Y) - E(V(Y|X_i, X_k)).$$

Finally, we have

$$V_{ik} = V(E(Y|X_i, X_k)) - V(E(Y|X_i)) - V(E(Y|X_k)), \quad (5)$$

where $V(E(Y|X_i))$ and $V(E(Y|X_k))$ can be estimated in the main effect analysis or in the current study using the replicated orthogonal array sample.

This same idea can be applied to the analysis of higher order interaction. For example, to analyze 3-way interaction, a replicated orthogonal array design of strength 3 can be used together with the corresponding formulas for computing the variance of conditional expectations.

4.4 Dealing with Correlated Inputs

If the inputs are correlated, the variance decomposition (3) is no longer valid. The main effect and interaction terms $V(X_i)$ and $V(X_i, X_k)$ can still be computed. However, pure interaction terms can no longer be written as in (5). There are several approaches for handling the case with correlated inputs.

Main effect analysis: The main effect analysis presented previously can be used in this case. However, the first order sensitivity indices will be confounded with interactions due to parameter correlations. If the correlation is in the form of joint probability density functions, the replicated Latin hypercube methods due to Iman and Conover [6] or Stein [23] are useful to produce suitable sample designs for main effect analysis. Direct integration can be used for correlations due to inequality constraints. Ratto, Tarantola, and Saltelli compared several different sampling techniques for estimating the main effect [19].

Group analysis: Jacques and Lavergne [7] proposed a modified approach by grouping the correlated input variables into the same group. Sensitivity indices are then defined for each group (instead of individual inputs) and these multi-dimensional indices are computed based on the Sobol' estimation method using Monte Carlo simulations. If the goal is to reduce the output uncertainty by a certain amount, Tarantola's [26] modified approach can be used to identify a minimal subset of inputs such that fixing these inputs will reduce the variance of Y to a targeted variance V_T .

Total sensitivity indices: Another measure that may be useful in the case of correlated inputs is the total sensitivity index S_{T_i} for input i which is defined as

$$S_{T_i} = V_i + \sum_{i \neq j} V_{ij} + \sum_{i \neq j, i \neq k} V_{ijk} + \dots + V_{12\dots M}.$$

To compute this quantity, we first recall the identity from statistics

$$V(Y) = V(E(Y|X_{\sim i})) + E(V(Y|X_{\sim i})) \quad (6)$$

where $\sim i$ means all inputs except input i . For uncorrelated inputs, the total sensitivity index for input i is defined as:

$$S_{T_i} = E(V(Y|X_{\sim i}))/V(Y) \quad (7)$$

which can be computed by the Sobol' method or by using the extended Fourier Amplitude Sampling Test (FAST) [20].

For correlated inputs, the Equation (6) does not hold. However, Equation (7) is still a valid measure for total sensitivity. In this case, Equation (7) can be computed by evaluating the integrals corresponding to the partial variances of all inputs. Since this process is very expensive, it only makes sense to perform this analysis on the response surface models. Both the FAST and Sobol' method for total sensitivity analysis has been implemented in our software package.

To illustrate the case of correlated inputs, we use the following example:

$$Y = F(X_1, X_2, X_3) = X_2 + X_3 \quad (8)$$

where $X_1, X_2, X_3 \in [0, 1]$ and $X_1 + X_2 \leq 1$. The sample mean and variance for this example are $5/6$ and $5/36$, respectively. The unnormalized main effects of all inputs are $V_1 = 1/48, V_2 = 1/12$ and $V_3 = 1/12$ respectively. However, we also note that

$$V_1 + V_2 + V_3 = 1/48 + 1/12 + 1/12 = 3/16 > V(Y) = 5/36.$$

which is typical in models with correlated inputs. Even though X_1 has no contribution to the output Y , the correlation due to the inequality constraint has induced a small variance of conditional expectation for X_1 . In more details, we have

$$V(E(Y|X_1)) + E(V(Y|X_1)) = 1/48 + 1/9 = 19/144,$$

$$V(E(Y|X_2)) + E(V(Y|X_2)) = 1/12 + 1/12 = 1/6,$$

$$V(E(Y|X_3)) + E(V(Y|X_3)) = 1/12 + 1/18 = 5/36,$$

which do not obey the relationship in Equation (4).

Similarly, we can evaluate the variance integrals to compute the 2-input interaction effects as $V_{12} = 1/18, V_{13} = 5/48$, and $V_{23} = 1/6$, and

$$V(E(Y|X_1, X_2)) + E(V(Y|X_1, X_2)) = 1/18 + 1/12 = 5/36,$$

$$V(E(Y|X_1, X_3)) + E(V(Y|X_1, X_3)) = 5/48 + 1/36 = 19/144,$$

$$V(E(Y|X_2, X_3)) + E(V(Y|X_2, X_3)) = 1/6 + 0 = 1/6.$$

Finally, the total sensitivity indices (unnormalized) for all inputs are $0.0, 1/36$ and $1/12$, respectively. More specifically,

$$V(E(Y|X_{\sim 1})) + E(V(Y|X_{\sim 1})) = 1/6 + 0 = 1/6,$$

$$V(E(Y|X_{\sim 2})) + E(V(Y|X_{\sim 2})) = 5/48 + 1/36 = 19/144,$$

$$V(E(Y|X_{\sim 3})) + E(V(Y|X_{\sim 3})) = 1/18 + 1/12 = 5/36.$$

For this example, based on the total sensitivity indices we conclude that X_1 is not sensitive at all, which agrees with the the original function (Equation (8)).

5 Validating our Sensitivity Analysis Methodology

So far we have presented a global sensitivity analysis methodology suitable for large scale multi-physics simulations with large number of uncertain parameters and complex constraints. A natural question to consider is: how reliable are the results from this methodology? In this section we take a critical look at this methodology by first pointing out its assumptions and then describing a few built-in mechanisms to self-validate the appropriateness of the assumptions for a given application. We approach this critical analysis by closely examining the following three major steps of the methodology:

1. the screening analysis,
2. the response surface analysis, and
3. the quantitative sensitivity analysis.

5.1 Self-Validation in the Screening Analysis

First of all, as discussed in Section 3.6, we would emphasize on the importance of understanding when the screening analysis can be most effective. The two desirable conditions are: (1) screening methods generally perform well when the number of important inputs is small compared to the total number of inputs under investigation (also called Pareto effect, which states that roughly 20% of the inputs account for 80% of the output variability); and (2) the MOAT method has been demonstrated to be effective in capturing parameters that have direct impact on the output of interest throughout the parameter space. Due to the coarse sampling in the screening procedure, Type II errors (when important parameters are treated as unimportant) may occur. Although many practitioners reported that Type II errors are unlikely for real applications, we have incorporated three mechanisms in our methodology to safeguard against Type II errors: (1) append to the screening procedure with a multi-resolution step which adds replications using finer resolution (that is, larger p_i as described in Section 3.3.2); (2) add a rigorous hypothesis testing on the Morris elementary effects to give probabilistic bounds on the believability of the results (briefly described in Section 3.6); and (3) provide an iterative feedback loop to capture the errors missed in the last iteration (Step 4 in the methodology presented in Section 2). Mechanism 1 above helps to verify that the coarseness of the screening sample is adequate, at the expense of additional simulation runs. For mechanism 2 to be effective, some assumptions such as small asymmetry of the distribution of the elementary effects may be needed. Having additional information about the output response such as smoothness and upper/lower bounds will help to detect the degree of asymmetry and improve the robustness of the hypothesis testing. Mechanism 2 can also be part of an adaptive procedure to satisfy the hypothesis by launching more simulation runs (details will be given in another report).

5.2 Cross-Validation in the Response Surface Analysis

As stated previously, if the simulation runs are very expensive, it may make more sense to create a response surface model first before performing quantitative sensitivity analyses. To construct a reliable and accurate response surface model for general applications, we need a suitable sampling method, a good response surface method, and effective validation tools. A good sampling method should be space-filling and should provide adequate boundary coverage (to avoid extrapolation in the analysis stage). Furthermore, adequate sample size is an important concern without any smoothness assumptions about the output responses. Adaptive response surface methodologies together with good error estimation tools will greatly reduce the computational costs. A few response surface methods have been mentioned in Section 4.1. Cross validation tools such as bootstrapping and jackknifing are useful to validate the response surface models. Many of these features are available from our software package.

5.3 Self-Validation in Variance Decomposition

Variance decomposition has a very strong mathematical foundation. The only concern we have is the accuracy of the estimates which depends largely on the sample size. In our variance decomposition we provide a multi-resolution approach to systematically study the accuracy of the estimates by observing the convergence of the estimates using different combinations for sample sizes.

6 PSUADE

To be able to perform designed simulation experiments like the ones used in our global sensitivity analysis methodology, many detailed tasks such as setting up the test problem, negotiating the computational resources, extracting and processing the appropriate outputs for examination, packaging the outputs for visualization, applying the analysis tools, optimizing the responses, interpreting the results, etc. are involved. To simplify these tedious bookkeeping chores, we need an integrated environment for such an analysis. There are many important issues in designing the architecture of such an integrated environment:

- a user-friendly interface for setting up the problem and for analyzing the results;
- an automated job execution environment suitable for different computer platforms;
- design and analysis tools for this methodology;
- graphics tools for visualizing the results; and
- a convenient scheme for storing and retrieving the results.

Many of the above considerations have been captured in a framework called PSUADE, which is a software package for sensitivity and uncertainty analysis. PSUADE supports a user-friendly interface via input and output filters. Once PSUADE has created a particular sampling design, it systematically feeds the design points into parameter files and calls a user handler which absorbs the sample data and inserts them into the model input files. The user handler then requests computational resource for the evaluation, waits for its completion, and extracts the output data from the model output files into PSUADE’s output files. This protocol allows all information exchange to be done via a user-written handler. There is no need to change the simulation source code to accommodate the analysis, and it is sometimes called a “non-intrusive” interface.

To allow multiple evaluations of the model, PSUADE allows function evaluations to be launched asynchronously. PSUADE then monitors the existence of the corresponding output files to determine that an evaluation has been completed. One can prescribe the degree of parallelism so that PSUADE can adequately request resources. PSUADE also has a rich collection of sampling methods such as Monte Carlo, full and fractional factorial, Latin hypercube, orthogonal arrays, and one-at-a-time methods. In addition, some of these methods have been equipped with refinement strategies [27] so that users can start with a small sample size and add more sample points as needed using refinements. Furthermore, uniform, normal and lognormal distributions are available for the inputs.

PSUADE provides a rich set of statistical tools such as the different moments, Pearson correlation coefficients, variance decomposition using main effect analysis, and others. It also provides a few optimization methods via public domain optimization tools.

PSUADE provides graphics support primarily through Matlab scripts. Screening plots, scatter plots, and response surface plots are created in Matlab format so that users can launch Matlab to view them.

PSUADE was written in C++ and runs on many different Unix-based operating systems.

7 Summary

In this report we describe a global sensitivity analysis methodology suitable for large scale multi-physics applications. Various computational and analysis aspects of this methodology from screening to main effect and interaction studies have been discussed in detail. Our methodology is distinguished by an enhanced screening phase, an efficient analysis for two-way interactions, and a recursive mechanism suitable for multi-physics applications. We have applied this methodology to several applications with satisfactory results. Specifically, there are a few lessons learned from our experience with the screening exercises:

1. The screening process is manageable (it is easy to understand, yet effective);
2. The process of applying the screening is as important as getting the results. The process promotes useful communications between scientists and analysts to verify that the observed phenomena make sense, and

3. Since the screening design exhibits a pattern, anomalies can easily be spotted and investigated.

The variance decomposition exercises likewise are revealing, especially with the help of graphical analysis tools. While PSUADE has been able to provide most of the functionalities of this methodology, a better user interface and more graphical tools will be more beneficial.

Acknowledgment:

The author would like to thank Dana Goto and Henry Hsieh for providing helpful feedbacks to enhance this methodology.

References

- [1] F. Campolongo and R. Braddock, *The use of graph theory in the sensitivity analysis of the model output: a new screening method*, Reliability Engineering and System Safety 64 (1997), pp. 1-12.
- [2] F. Campolongo, S. Tarantola, and A. Saltelli, *Tackling quantitatively large dimensionality problems*, Computer Physics Communications 117 (1999) pp. 75-85.
- [3] J. H. Friedman, *Multivariate adaptive regression splines*, *Annals of Statistics* 19.1, 1-141, 1991.
- [4] A. S. Hedayat, N. J. a. Sloane, and John Stufken, *Orthogonal Arrays: Theory and Applications*, Springer Series in Statistics, 1999.
- [5] J. C. Helton and F.J. Davis, *Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems*, Reliability Engineering and System Safety, Vol. 81, No. 1, pp. 23-69, 2003.
- [6] R. L. Iman and W. J. Conover, *A Distribution-free Approach to Inducing Rank Correlation Among Input Variables*, *Commun. Statist. Simula. Computa.*, Vol 11, pp. 311-334, 1982.
- [7] J. Jacques and C. Lavergne, *Sensitivity Analysis for Model with Correlated Inputs*, Reliability Engineering and System Safety 91 (2006), Issue 10-11, pp 1126-1134.
- [8] J. R. Koehler and A. B. Owen, *Computer Experiments*.
- [9] T. Kolda, *Revisiting Asynchronous Parallel Search for Nonlinear Optimization*, Technical Report SAND2004-8055, Sandia National Laboratories, Livermore, California, February 2004.

- [10] T. J. Lorenzen and V. L. Anderson, *Design of Experiments : A No-Name Approach*, Marcel Dekker, Inc. 1993.
- [11] M. D. McKay, *Evaluating Prediction Uncertainty*, Los Alamos National Laboratory Technical Report NUREG/CR-6311, LA-12915-MS.
- [12] M. McKay, R. Beckman, W. Conover, *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code*, *Technometrics*, 21(2):239-245, 1979.
- [13] M. McKay, M. A. Fitzgerald, and R. J. Beckman, *Sample Size Effects when Using R^2 to Measure Model Input Importance*, LANL technical report.
- [14] M. D. Morris, *Factorial Sampling Plans for Preliminary Computational Experiments*, *Technometrics*, 21(2), pp. 239-245, 1991.
- [15] M. D. Morris, *Private communication via email*.
- [16] R. H. Myers and D. C. Montgomery, *Response Surface Methodology*, Second Edition, Wiley Series in Probability and Statistics, 2002.
- [17] A. B. Owen, *Orthogonal Arrays for Computer Experiments, integration, and visualization*, *Statist. Sinica* 2, pp. 439-452, 1992.
- [18] R. L. Plackett and J. P. Burman, *The Design of Optimum Multifactorial Experiments*, *Biometrika*, 33, pp 305-325.
- [19] M. Ratto, S. Tarantola and A. Saltelli, *Estimation of Importance Indicators for Correlated Inputs*, Proceedings of the European Conference on Safety and Reliability, ESREL 2001, E. Zio, M. Demichela, N. Piccinini (eds.), torino, 16-20 September 2001, vol. 1, pp. 157-164.
- [20] A. Saltelli, K. Chan, E. M. Scott (editors), *Sensitivity Analysis*, Wiley Series in Probability and Statistics, 2000.
- [21] A. Saltelli, S. Tarantola, and K. Chan, *A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output*, *Technometrics*, Vol. 41, No. 1, pp. 39-55, 1999.
- [22] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity Analysis in Practice*, Wiley, 2004.
- [23] M. Stein, *large sample properties of simulations using Latin hypercube sampling*, *Technometrics*, 29(2), pp. 143-151.
- [24] B. Tang, *Latin Hypercubes and Supersaturated Designs*, Dissertation, Dept. of Statistics and Actuarial Science, University of Waterloo, 1992.

- [25] B. Tang, *Orthogonal Array-based Latin Hypercubes*, J. Amer. Statist. Assn. 88, 1392-1397, 1993.
- [26] S. Tarantola, *Quantifying uncertainty importance when Inputs are correlated*, Foresight and Precaution, edited by Cottam, Harvey, Pape and Tait (eds), pp. 1115-1120, 2000.
- [27] C. Tong, *Refinement Strategies for Stratified Sampling Methods*, Reliability Engineering and System Safety 91 (2006), Issue 10-11, pp 1257-1265.