To appear in the Proceedings of the Winter Simulation Conference,
Arlington, Virginia, December 13-16, 1992

LAUR 92-2338

LATIN HYPERCUBE SAMPLING AS A TOOL IN UNCERTAINTY ANALYSIS OF COMPUTER MODELS

Michael D. McKay

Statistics Group, Los Alamos National Laboratory
Los Alamos, New Mexico 87545, U.S.A.

ABSTRACT

This paper addresses several aspects of the analysis of uncertainty in the output of computer models arising from uncertainty in inputs (parameters). Uncertainty of this type, which is separate and distinct from the randomness of a stochastic model, most often arises when input values are guesstimates, or when they are estimated from data, or when the input parameters do not actually correspond to observable quantities, e.g., in lumped-parameter models. Uncertainty in the output is quantified in its probability distribution, which results from treating the inputs as random variables. The assessment of which inputs are important with respect to uncertainty is done relative to the probability distribution of the output.

1 INTRODUCTION

Uncertainty in a model output—how big it is and what it is attributable to—is not a new issue. The accuracy of models has always been a concern, and things are no different today with computationally intensive models on computers. Three things make the analysis of computer models usually more difficult than that of other models. First, the nature of the relationship between the output and the inputs is often very complex. Second, there can be very many inputs for which the cost of data collection is high. Finally, when the output is something like a forecast, comparison between predicted and calculated values is essentially impossible. Despite these difficulties, questions like "What is the uncertainty in the calculation?" continue to be heard both in scientific circles and in political ones, where the cost of decisions resting on model calculations can be high.

Although there are several useful questions one might ask about computer models, there seems to be a tendency to use the same kind of method to answer many of them. It is unlikely, however, that any single method of analysis exists that answers all questions in model evaluation. Moreover, an examination of typical questions would likely suggest that different kinds of methods are not only desirable but necessary. To find appropriate methods, one needs precise statements of objectives. Unfortunately for many investigators, what begins as an intuitive concept, like uncertainty or sensitivity, can easily end up as an imprecisely stated and misleading question that suggests an inappropriate method of answering it. This paper tries to address that problem by providing a framework within which questions of uncertainty and importance are posed.

This discussion will be limited to particular considerations from the diversity of issues comprising model analysis. First of all, this paper is not an empirical comparison or evaluation of methods currently used in the analysis of computer models. Examples of such studies are Saltelli and Homma (1992), Saltelli and Marivoet (1990), Iman and Helton (1988) and Downing, Gardner and Hoffman (1985). Secondly, we will consider questions dealing with the values of the output and the input, and not, specifically, with the form or structure of their relationship. In particular, we will focus on the issue of uncertainty in the calculated value due to uncertainty in input values. To provide a setting, we will introduce current methods from two different perspectives related to model analysis. With that background, we will develop a new and useful paradigm for analysis and methods development for issues related to uncertainty in the output value.

2 TWO PERSPECTIVES

Model analysis can be thought of as a collection of questions asked about output and input values. Although a simplification, it is useful to distinguish the questions as arising from one of two perspectives. This idea, previously discussed by McKay (1978, 1988), allows the introduction of a new way of viewing importance of inputs with regards to uncertainty in the output.

The first perspective is from the space of input values, and tends to focus at one point, like a nominal input value. Because of this, quantities of interest, like a derivative, can seem to be treated as constant over the input space, so that the focal point really does not matter.

This perspective is termed a local perspective relative to the input space.

The second perspective is from the space of output values. As such, its focus is not constrained a priori in the input space, so that it is termed a global perspective relative to the input space.

The reason for making the distinction between local and global perspectives has to do with the problem of identifying important inputs. Although "importance" has not yet been defined, it seems reasonable to suppose that qualities that make an input important locally are not necessarily those that make it important globally, and vice versa. Therefore, it is necessary to realize the perspective of interest.

2.1 A Local Perspective

Let us suppose that there is some value of the inputs, x_0, worth attention and that we are interested in changes in the output Y for small perturbations in inputs X about x_0. A common question in this situation concerns propagation of error, characterized by the derivatives of Y with respect to the components of X. Similarly, one might be interested in the direction, not necessarily parallel to a coordinate axis, in which Y changes most rapidly, or in the change in Y in an arbitrary direction. Issues like these lead one to the concept of "critical" or "important" variable (or direction) as being one(s) that most accounts for change in Y. For propagation of error, it seems to make sense to talk about each component of X as being important or not important. When the direction becomes arbitrary, it seems natural to talk about subsets of the components, rather than about individual ones.

2.2 A Global Perspective

Suppose that interest lies in the event that Y, the output or prediction, exceeds some specified value. Questions that could arise in this case might be concerned with associating particular inputs (components of X) or segments of ranges with that event. Objectives of study for this question might be related to controlling the event, or with reducing its probability of occurrence in the real world by adjusting the values of some of the inputs. If costs are associated with the inputs, minimum cost solutions might be sought.

Clearly, both perspectives have a place in model analysis. In the local perspective, interest in X is restricted to a (small) neighborhood of a single point, and the derivative seems to come into play. In the global perspective, interest is in values of Y, which might translate into a subset of, or possibly just a boundary in, the input space. In this case, the role of the derivative is less clear. What tends to blur the distinction between the two perspectives is the use of the derivative to answer questions of a global nature. The practice is defensible if the model is essentially linear, meaning that the derivative does not change substantially with x_0; or that, to first order approximation, an "average" derivative is sufficient to characterize the model, again meaning that the model is essentially linear. In what follows, a global approach is taken and the role of the derivative is not paramount.

3 UNCERTAINTY ANALYSIS

We are interested in the type of uncertainty in the output of a model that can be characterized as being due to the values used for the inputs. A related uncertainty, due to the structure or form of the model itself, will not be addressed explicitly. Neither will we be concerned with uncertainty due to errors in implementation of the model on a computer. On the other hand, it is certainly acceptable that the output might have the randomness of a stochastic process. In that case, we will think of the output of the model as being the cumulative distribution function of the observable output value. With this in mind, the purpose of uncertainty analysis is to quantify the variability in the output of a computer model due to variability in the values of the inputs.

We proceed by first describing the variability in the inputs with probability functions. Commonly, input values are uncertain because they are guesstimates, or when they are estimated from data, or when the input parameters do not actually correspond to observable quantities, e.g., in lumped-parameter models. Treating the inputs as random variables introduces another layer of complication, namely, that of assigning to them probability distributions. Everything that follows will depend on the distributions used for the inputs, which means uncertainty in the input probability distributions leads to corresponding uncertainty in the analysis. As an alternative to quantifying that uncertainty, some kind of variational study used to measure the effect of the distributional assumptions is a possibility.

When the inputs are treated as random variables, the output becomes a random variable because it is a transformation of the inputs. Uncertainty in the output, then, is characterized by its probability distribution. Therefore, when we consider questions related to uncertainty in the output, Y, we will look to the probability distribution of Y for answers.

We assume that interest in uncertainty in Y can be summed up by in these two questions: "How big is it?" and "Can it be attributed to particular inputs?" An obvious motivation for these questions is a desire to minimize uncertainty in the model output, which might be achieved by reducing the variance of some of the inputs. Thus, an

important problem might be how to find a minimum-cost reduction in the variance of Y by reducing the variances of components of X. This problem presupposes, of course, that reduction in the variance of the inputs makes sense.

4 MEASURING UNCERTAINTY

We look to the probability distribution function of the output Y for information about uncertainty. Questions like "What is the uncertainty in Y?" might be answered using a probability interval constructed with quantiles of the distribution of Y. For example, the 0.05 and the 0.95 quantiles define an interval covering 90% of the probability content of the distribution of Y. Alternatively, the difference in the two quantiles provides a range of 90% coverage. The use of probability intervals as a measure of uncertainty has an advantage over the use of the variance of the distribution in that the variance may not directly relate to coverage. This is not the case, though, in the familiar normal distribution where quantiles depend in a straightforward manner on the mean and variance of the distribution.

Ideally, the probability distribution of Y would be known once the distribution of X is specified. Realistically, the distribution will have to be estimated, most likely, with a sample of runs using the model which relates Y to X. Simple random sampling (SRS) could be used for this purpose, as well as could other sampling schemes. Latin hypercube sampling (LHS), introduced by McKay, Conover and Beckman (1979), is a preferred alternative to SRS when the output is a monotone function of the inputs. Additionally, Stein (1987) shows that LHS yields an asymptotic variance smaller than that for SRS. Besides being used to estimate the probability distribution, sample values could be used to construct a tolerance interval, which covers at least a specified portion of the probability distribution of Y with a specified confidence level. (For a short discussion of probability interval and tolerance interval, see Tietjen (1986, p. 36).) Generally, a tolerance interval, which corresponds to a probability interval when the probability distribution is estimated, is based on a random sample. If usual methods for constructing tolerance intervals based on nonparametric techniques or on the normal distribution—e.g., in Bowker and Lieberman (1972, pp. 309-316) where they are called tolerance limits—are applied to the LHS, however, the results are only approximate.

In the remainder of this paper, we will be concerned only with probability distributions and their moments. Furthermore, we will assume that sample sizes are sufficiently large to rule out concern about sampling error in all regions of interest in estimated distribution functions.

Although it has been suggested that a probability interval is a more appropriate measure of uncertainty than is variance, the use of variance to "partition" or allocate uncertainty to components of X cannot be overlooked. In fact, three categories of techniques for the determination of "important" inputs, relative to uncertainty in the output, look primarily at the variance of Y. We will review these techniques using variance as the measure of uncertainty before introducing the new paradigm for uncertainty analysis.

5 PARTITIONING UNCERTAINTY

Statements like "20% of the uncertainty in Y is due to X_1" have a nice sound, but may be very misleading without explanation. If we suppose that uncertainty in Y is measured by its variance, then a reasonable interpretation of the statement is that the variance of Y can be written, approximately, as the sum of two functions, one depending on the distribution of X_1 alone and the other independent of the distribution of X_1. This picture captures a motivation for, but does not limit, the classes of techniques to be discussed.

5.1 Linear Propagation of Error

When we are using variance to measure uncertainty, the problem of partitioning uncertainty reduces to that of finding suitable decompositions for the variance of Y. The simplest of these is the usual propagation of error method in which Y is expressed as a Taylor series in the inputs X about some point x_0. To first order approximation, the variance of Y is expressed as a linear combination of the variances of the components of X by choosing x_0 to be μ, the mean value of X.

$$Y(X) = Y(x_0) + \sum_i \frac{\partial Y(x_0)}{\partial X_i}(X_i - x_{0i}) + \cdots$$

$$V[Y] \simeq \sum_i \left(\frac{\partial Y(\mu)}{\partial X_i}\right)^2 V[X_i]$$

When the derivatives of Y are not determined numerically, but estimated by the coefficients from a linear regression of Y on X, one seems to be making a stronger assumption about the linear dependence of Y on X. However, it is generally unknown whether the value of the actual derivative of Y or the value of an average slope is preferred in the variance approximation. In a technique that could be related to linear propagation of error, Wong and Rabitz (1991) look at the principal components of the partial derivative matrix.

Although not precisely a variance decomposition, correlation coefficients have been used to indicate relative importance of the inputs. They are mentioned here

because they are closely related to linear regression coefficients. In a similar way, rank transformed values of Y and X have been used for rank correlation and rank regression by McKay, Conover and Whiteman (1976), and Iman, Helton and Campbell (1981a, 1981b). Also, Oblow (1978) and Oblow, Pin and Wright (1986) use a technique whereby the capability of calculating derivatives into the model is added using a precompiler called GRESS.

5.2 General Analytical Approximation

The natural extension of linear propagation of error, to add more terms in the Taylor series, makes it difficult to interpret variance decomposition component-wise for X. That is, the introduction of cross-product terms brings cross-moments into the variance approximation, which makes the approximation no longer separable with respect to the inputs. Nevertheless, one may feel it necessary to use higher order terms in variance approximation. The adequacy of the approximation to Y might be used as a guide to the adequacy of the variance approximation. However, there is no particular reason to think that one implies the other.

Similarly, the linear approximation of Y used in the regression can be generalized to an arbitrary analytical approximation from which, in theory, the variance of Y can be derived either mathematically or through simulation. Alternatively, one can use a method proposed by Sacks, Welch, Mitchell and Wynn (1989), which looks at the model as a realization of a stochastic process. The difficulties in interpretation and assessing adequacy just mentioned for the higher order Taylor series expansion apply here, too.

5.3 Sampling Methods

This final category of partitioning techniques relies on a sample (usually, some type of random sample) of values of Y whose variability can be partitioned according to the inputs without an apparent assumed functional relation between Y and X. In this category is a Fourier method of Cukier, Levine and Shuler (1978). The procedure says that values of each component of X are to be sampled in a periodic fashion, with different periods for each component. The variability (sum of squares) of the resulting values of Y can be written as a sum of terms corresponding to the different periods, and thus associated with the different components. It is unclear how this relates to linear propagation of error, but it may be just another way to estimate the same quantities. The original Fourier method applies to continuous inputs. It is extended to binary variables by Pierce and Cukier (1981). Again, the relation to linear propagation of error is unclear. Another procedure suggested by Morris (1991) examines a probability distribution of the partial derivatives of the output arising from particular sampling designs.

Finally, I mention a partition of variance described by Cox (1982). Though not actually a sampling method, the elements of the decomposition are likely to estimated from sampled data, in practice. The identity used involves the variances of conditional expectations of the output given subsets of the inputs. As with general analytical approximation, it is not possible to isolate terms for all the individual components of X.

6 MATHEMATICAL FRAMEWORK

The uncertainty in the output that we focus on is that attributable to the inputs. Specifically, we are ignoring the uncertainty in calculations due to the possibility that the structure of the model might be deficient. We let Y denote the calculated output, which depends on the input vector, X, of length p through the computer model, h(•). Because proper values of the components of X may be unknown or imprecisely known, or because they can only be described stochastically, it seems reasonable to treat X as a random variable and to describe uncertainty about X with a probability function. Uncertainty in the calculation Y is captured by its own probability function, which is what we will study. In summary, then,

$$\begin{aligned} Y &= h(X) \\ X &\sim f_x(x)\,,\ x \epsilon R^p \\ Y &\sim f_y(y)\,. \end{aligned}$$

For now, we will think of f_x as known, although in practice, knowledge about it is at best incomplete.

We look to the probability distribution, f_y, for answers to the question "What is the uncertainty in Y?" That is to say, we can use the quantiles of the distribution of Y to construct probability intervals. Alternatively, one might use the variance of Y to quantify uncertainty. In either case, under the assumption that f_y can be adequately estimated, questions answerable with quantiles or moments are covered. However, as has already been mentioned, the issue of how well f_x is known will surely have to be addressed in practice.

We relate questions of importance of inputs to the probability distribution of Y. That is, we will consider questions like "Which variables really contribute to (or affect) the probability distribution of the output?" What it means to be important is defined in somewhat of a backwards way as being the complement of unimportant. We say that a subset of inputs is unimportant if the conditional distribution of the output given the subset is

essentially independent of the values of the inputs in the subset. We now examine these ideas in more detail.

Suppose that the vector X of inputs is partitioned into X_1 and X_2. Corresponding to the partition, we write

$$\begin{aligned} Y &= h(X) \\ &= h(X_1, X_2) \,. \end{aligned}$$

Furthermore, we assume that X_1 and X_2 are stochastically independent, meaning that

$$\begin{aligned} X_i &\sim f_i(x_i) \,, \;\; i = 1, 2 \\ f_x(x) &= f_1(x_1) f_2(x_2) \,. \end{aligned}$$

We address the question of the unimportance of X_2 by looking at

$$f_{y|x_2} = \text{ distribution } Y \text{ given } X_2 = x_2 \,.$$

as compared to f_y. We say that X_2 is unimportant if f_y and $f_{y|x_2}$ are not substantially different for all values of X_2 of interest. Similarly, we say that X_1 contains all the important inputs if X_2 is unimportant. Of course, the actual way to compare f_y and $f_{y|x_2}$ must be determined.

We use the term "screening" to mean an initial process of separating inputs into X_1, potentially important ones, and X_2, potentially unimportant ones. In the next section, a simple method of partitioning the inputs, following McKay, Beckman, Moore and Pickard (1992), will be discussed.

7 A SIMPLE SCREENING HEURISTIC

We now describe a simple, two-step screening process. The first step is to partition X into a set of "important" components, X_1, and a set of "unimportant" components, X_2. The second step is a partial validation to estimate how the components in X_2 actually change $f_{y|x_2}$, to be used to decide if X_2 is really unimportant.

7.1 Partitioning the Input Set

We say that X_2, a subset of X, is (completely) unimportant when the marginal distribution of Y, equals the conditional distribution of Y given X_2.

$$f_y = f_{y|x_2} \text{ for all values of } X_2 \qquad (1)$$

A way to get an idea of how closely the equality in (1) holds is through the variance expression (2) which expresses the marginal variance of f_y in terms of the conditional mean and variance of $f_{y|x_2}$. The variance of Y can be written as

$$V[Y] = E[V[Y \mid X_2]] + V[E[Y \mid X_2]] \,. \qquad (2)$$

Equality of the marginal and conditional distributions in (1) implies that the conditional mean and variance are equal to their marginal counterparts for all values of X_2. Specifically, the variance (over X_2) of the conditional expectation in (2) is zero. It is unlikely, of course, that any (realistic set) of the inputs is completely unimportant. Therefore, the equality between marginal and conditional quantities will be true only in approximation, with the degree of approximation linked to the level of acceptance of the difference between the marginal and conditional distributions of the output, Y.

By inference, if X_1, the complement to X_2, is (completely, singly) important, the conditional variance of Y given X_1 is zero, and the variance of the conditional expectation of Y given X_1 is the marginal variance. As before, these relations usually hold only in approximation. Nevertheless, a comparison of terms in (2) will offer a way to look at the degree of importance.

The variance decomposition in (2) suggests a related identity from a one-way analysis of variance, in which the total sum of squares is written as the sum of two components, a "between level" component and a "within level" component. It will be the analysis of variance approach we will use to suggest which components of X belong in X_1 and which in X_2. What we will do is to use r "replicate" Latin hypercube samples of size k. The same k values of each component of X will appear in each replicate but the matching within each one will be done independently. The k values will correspond to the k levels in the sum of squares decomposition.

In an LHS as introduced by McKay, Conover and Beckman (1979), when the inputs are continuous and stochastically independent, the range of each component of X is divided into k intervals of equal probability content. Simple modifications can be made to handle discrete inputs (McKay 1988) and dependence (McKay 1988, Stein 1987, Iman and Conover 1982). For a true LHS, a value is selected from each interval according to the conditional distribution of the component on the interval. For this application, it will be sufficient to use the probability midpoint of the interval as the value. The k values for each input are matched (paired) at random to form k input vectors. For the replicates needed in this screening heuristic, r independent matchings of the same values are used to produce the $n = k \times r$ input vectors in total.

A design matrix, M, for an LHS is given in (3). Each column contains a random permutation of the k values for an input. Each row of the matrix corresponds

to a random matching of values for the p inputs used in a computer run.

$$M = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1p} \\ v_{21} & v_{22} & \cdots & v_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ v_{k1} & v_{k2} & \cdots & v_{kp} \end{bmatrix} \tag{3}$$

A design matrix for any of the r replicates in this application is obtained by randomly and independently permuting the values in every column of M.

After making the necessary n computer runs using replicated LHS, we begin by looking at the components of X one at a time. Let U denote the component of interest in X, and denote the k values of U by $u_1, u_2, \ldots, u_k$. We label n values of the output as y_{ij} to correspond to the ith value u_i, in the jth replicate (sample). The sum of squares partition corresponding to the input U takes the form

$$\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\overline{y})^2 = r\sum_{i=1}^{k}(\overline{y}_i-\overline{y})^2 + \sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\overline{y}_i)^2$$

$$\text{SST} = \text{SSB} + \text{SSW}$$

where

$$\overline{y}_i = \frac{1}{r}\sum_{j=1}^{r} y_{ij} \text{ and } \overline{y} = \frac{1}{k}\sum_{i=1}^{k}\overline{y}_i \ .$$

The statistic we have chosen to use to assess the importance of U is $R^2 = \text{SSB}/\text{SST}$. Although R^2 is bounded between 0 and 1, the attainment of the bounds is not necessarily a symmetric process. The upper bound is reached if Y depends only on U. In that case, for any fixed value of U, say u_i, the value of Y will also be fixed, making SSW equal to 0. As a result, R^2 will be 1. On the other hand, if Y is completely independent of U, we do not expect SSB (and, therefore, R^2) to be 0. We now examine this last point in more detail.

In general, the probability distribution of R^2 will be unknown. To gain a little insight, however, suppose that we arbitrarily partition a random sample of size n from a normal distribution to form R^2. (An arbitrary partition would correspond to Y independent of U.) The expected value of R^2 is $(k-1)/(n-1)$, which goes to zero with k/n as n increases. Thus, one might consider $(k-1)/(n-1)$ as a working lower bound associated with a completely unimportant input.

Issues that still need to be addressed include the apportionment of n between r and k, the extension of the design and decomposition to more that one component at a time, and the interpretation of values of R^2.

Whether or not one uses R^2 or additional methods to develop the sets X_1 and X_2, there remains the issue of evaluating the partition to see how effective it is in satisfying (1). In fact, iterating between a partition and validation is what one would do in practice. The next section discussion validation.

7.2 Validation of the Partition

Very simply stated, in the validation step we look at X_1 and X_2 and try to assess how well the partition meets the objective of isolating the important inputs to X_1. We propose using a very elementary sequence of steps that begins with a sample design resembling Taguchi's (1986) inner array/outer.

1. Select a sample, S2, of the X_2s and a sample, S1, of the X_1s.

2. For each sample element $x_2 \in$ S2, obtain the sample of Y corresponding to $\{x_2 \otimes \text{S1}\}$.

3. Calculate appropriate statistics for each sample in Step 2, e.g., $\overline{Y}(x_2)$, $s_y^2(x_2)$ and $\widehat{F}_{y|x_2}$.

4. Compare the statistics and decide if the difference x_2 makes is acceptable.

The differences seen in the statistics in Step 4 are due only to the different values of x_2 because the sample values for X_1 are the same in each. Hence, the comparisons are reasonable.

The reliability of any validation procedure needs to be evaluated. In this case, S2 may not adequately cover the domain of X_2, particularly as the dimension of X_2 increases. Merely increasing the size of S2 may not be an acceptable solution if the increase in the number of runs to generate the sample of Ys becomes impossible to accommodate. Inadequate coverage can be due to two reasons. First, regions where the conditional distribution of Y really changes with X_2 alone may be missed. Second, there may be regions where the interaction between X_2 and X_1 in the model has a significant impact on the conditional distribution of Y. Although it has obvious deficiencies, LHS is an appropriate sampling method for generating S2 because it provides marginal stratification for each input in X_2, meaning that the individual ranges within the components likely have been sampled adequately. Whether or not interaction between X_1 and X_2 will be detected is unknown. As an alternative to LHS, one might use an orthogonal array as described by Owen (1991), which provides marginal stratification for all pairs of input variables.

8 APPLICATION

For an application of these methods, the reader is referred to McKay, Beckman, Moore and Pickard (1992).

ACKNOWLEDGMENTS

This work was supported by the United States Nuclear Regulatory Commission, Office of Nuclear Regulatory Research. The author expresses his thanks for many valuable conversations with and help from Richard Beckman, Richard Pickard, Lisa Moore, Thomas Bement, Paul Kvam and other members of the Statistics Group at the Los Alamos National Laboratory.

REFERENCES

Bowker, A. H. and Lieberman, G. K. (1972). *Engineering Statistics*. Prentice-Hall, Englewood Cliffs, NJ.

Cox, D. C. (1982). An analytical method for uncertainty analysis of nonlinear output functions, with application to fault-tree analysis. *IEEE Transactions on Reliability*, R-31(5):465–468.

Cukier, R. I., Levine, H. B., and Shuler, K. E. (1978). Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26:1–42.

Downing, D. J., Gardner, R. H., and Hoffman, F. O. (1985). An examination of response-surface methodologies for uncertainty analysis in assessment of models. *Technometrics*, 27(2):151–163.

Iman, R. L. and Conover, W. J. (1982). A distribution free approach to inducing rank correlation among input variables. *Communications in Statistics—Simulation and Computation*, B11:311–334.

Iman, R. L. and Helton, J. C. (1988). An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis*, 8(1):71–90.

Iman, R. L., Helton, J. C., and Campbell, J. E. (1981a). An approach to sensitivity analysis of computer models: Part I—introduction, input variable selection and preliminary variable assessment. *Journal of Quality Technology*, 13(3):174–183.

Iman, R. L., Helton, J. C., and Campbell, J. E. (1981b). An approach to sensitivity analysis of computer models: Part II—ranking of input variables, response surface validation, distribution effect and technique synopsis. *Journal of Quality Technology*, 13(4):232–240.

McKay, M. D. (1978). A comparison of some sensitivity analysis techniques. Presented at the ORSA/TIMS annual meeting, New York.

McKay, M. D. (1988). Sensitivity and uncertainty analysis using a statistical sample of input values. In Ronen, Y., editor, *Uncertainty Analysis*, chapter 4, pages 145–186. CRC Press, Boca Raton, FL.

McKay, M. D., Beckman, R. J., Moore, L. M., and Picard, R. R. (1992). An alternative view of sensitivity in the analysis of computer codes. In *Proceedings of the*

American Statistical Association Section on Physical and Engineering Sciences, Boston, MA.

McKay, M. D., Conover, W. J., and Beckman, R. J. (1979). 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.

McKay, M. D., Conover, W. J., and Whiteman, D. E. (1976). Report on the application of statistical techniques to the analysis of computer codes. Technical Report LA-NUREG-6526-MS, Los Alamos National Laboratory, Los Alamos, NM.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174.

Oblow, E. M. (1978). Sensitivity theory for reactor thermal-hydraulics problems. *Nuclear Science and Engineering*, 68:322–337.

Oblow, E. M., Pin, F. G., and Wright, R. Q. (1986). Sensitivity analysis using computer calculus: A nuclear waste isolation application. *Nuclear Science and Engineering*, 94:46–65.

Owen, A. B. (1992). Orthogonal arrays for computer integration and visualization. *Statistica Sinica*, 2(2):439–452.

Pierce, T. H. and Cukier, R. I. (1981). Global nonlinear sensitivity analysis using Walsh functions. *Journal of Computational Physics*, 41:427–443.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical Science*, 4(4):409–435.

Saltelli, A. and Homma, T. (1992). Sensitivity analysis for a model output: Performance of black box techniques on three international benchmark exercises. *Computational Statistics & Data Analysis*, 13:73–94.

Saltelli, A. and Marivoet, J. (1990). Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques. *Reliability Engineering and System Safety*, 28:229–53.

Stein, M. (1987). Large sample properties of simulations using Latin hypercube sampling. *Technometrics*, 29(2):143–151.

Taguchi, G. (1986). *Introduction to Quality Engineering*. Kraus International Publications, White Plains, NY.

Tietjen, G. L. (1986). *A Topical Dictionary of Statistics*. Chapman and Hall, New York.

Wong, C. F. and Rabitz, H. (1991). Sensitivity analysis and principal component analysis in free energy calculations. *Journal of Physics and Chemistry*, 95:9628–9630.