JAERI-M
89-190

CURRENT STATUS OF UNCERTAINTY ANALYSIS METHODS FOR COMPUTER MODELS

November 1989

Tsutomu ISHIGAMI

日 本 原 子 力 研 究 所
Japan Atomic Energy Research Institute

JAERI-M レポートは，日本原子力研究所が不定期に公刊している研究報告書です。
入手の問合わせは，日本原子力研究所技術情報部情報資料課（〒319-11茨城県那珂郡東海村）あて，お申しこしください。なお，このほかに財団法人原子力弘済会資料センター（〒319-11 茨城県那珂郡東海村日本原子力研究所内）で複写による実費頒布をおこなっております。

JAERI-M reports are issued irregularly.
Inquiries about availability of the reports should be addressed to Information Division Department of Technical Information, Japan Atomic Energy Research Institute, Tokai-mura, Naka-gun, Ibaraki-ken 319-11, Japan.

編集兼発行　日本原子力研究所
印　　刷　いばらき印刷㈱

Current Status of Uncertainty Analysis Methods for Computer Models

Tsutomu ISHIGAMI

Department of Reactor Safety Research
Tokai Research Establishment
Japan Atomic Energy Research Institute
Tokai-mura, Naka-gun, Ibaraki-ken

(Received October 20, 1989)

This report surveys several existing uncertainty analysis methods for estimating computer output uncertainty caused by input uncertainties, illustrating application examples of those methods to three computer models, MARCH/CORRAL II, TERFOC and SPARC. Merits and limitations of the methods are assessed in the application, and recommendation for selecting uncertainty analysis methods is provided.

Keywords: Uncertainty Analysis, Sensitivity Analysis, Importance Analysis, Response Surface, Regression Analysis, Monte Carlo Simulation, Latin Hypercube Sampling, Computer Code

計算モデルに対する不確実さ解析手法の現状

日本原子力研究所東海研究所原子炉安全工学部
石神　　努

（1989年10月20日受理）

本報は，計算モデルの入力データに含まれる不確実さに起因する計算結果の不確実さを解析するためにこれまでに開発されてきた幾つかの手法と，それらをMARCH／CORRAL II，TER－FOCおよびSPARCの3つの計算モデルに応用した解析例についてまとめたものである。解析例を通じて手法の適用限界を論じるとともに，応用に際しての解析手法の選択に関する考察を行う。

東海研究所：〒319－11　茨城県那珂郡東海村白方字白根2－4

目　次

Contents

1. Introduction

In order to assess the safety of complex technological systems, computer models are widely used in many fields such as probabilistic safety assessment (PSA) of nuclear power plants,[(1),(2)] dose assessment for disposal of radioactive waste[(3)] and so on. However, the results (outputs) calculated include uncertainties associated with both the complex computer models themselves and values of input variables or parameters. An adequate quantification of uncertainties associated with such assessments as PSA is crucial to their use as a decision-making tool.[(2)]

The general problem with which uncertainty analysis is concerned can be expressed as a functional relationship whose output depends in a deterministic manner on various input variables.[(4)] Thus,

$$Y = h(X_1,\ldots,X_K) \ , \tag{1.1}$$

where, Y denotes system output, and $X_1,\ldots,X_K$ are system variables. As is shown in Figure 1.1, in practice, the values of the input variables are not precisely known and consequently, some imprecision attaches to the estimate of output Y. The objectives of uncertainty analysis are to quantify the uncertainty in Y, and, often to partition that uncertainty among the contributing input variables.

When the functional relationship, Eq.(1.1), is complex and thus not given in analytic expression, quantification of the uncertainty in Y needs a set of input/output data which will be obtained from a large number of numerical calculations based on the computer model of concern. As is larger the size of the data set, the uncertainty in Y will be, in general, estimated more precisely because of decreased errors in statistics. The Monte Carlo simulation method is known to be the direct and proper method for the uncertainty quantification, where a large number of repeated calculations with the computer model are needed. However, the Monte Carlo method loses its effectiveness for very long-running computer codes, which is often the case for assessment of the safety of complex technological systems using computer models. Hence some uncertainty analysis methods have been proposed and studied, which aim to assess uncertainty using a less number of input/output data.

The objectives of this report are to provide a current status of uncertainty analysis methods and to illustrate application examples of the methods to three computer models, MARCH/CORRAL II, TERFOC and SPARC.

Chapter 2 describes a procedure of uncertainty analysis. Chapter 3 describes several methods for uncertainty analysis. Application of the methods to uncertainty analysis for computer models is illustrated in chapter 4. Summary and recommendation on use of uncertainty analysis methods considered in this report are provided in chapter 5.

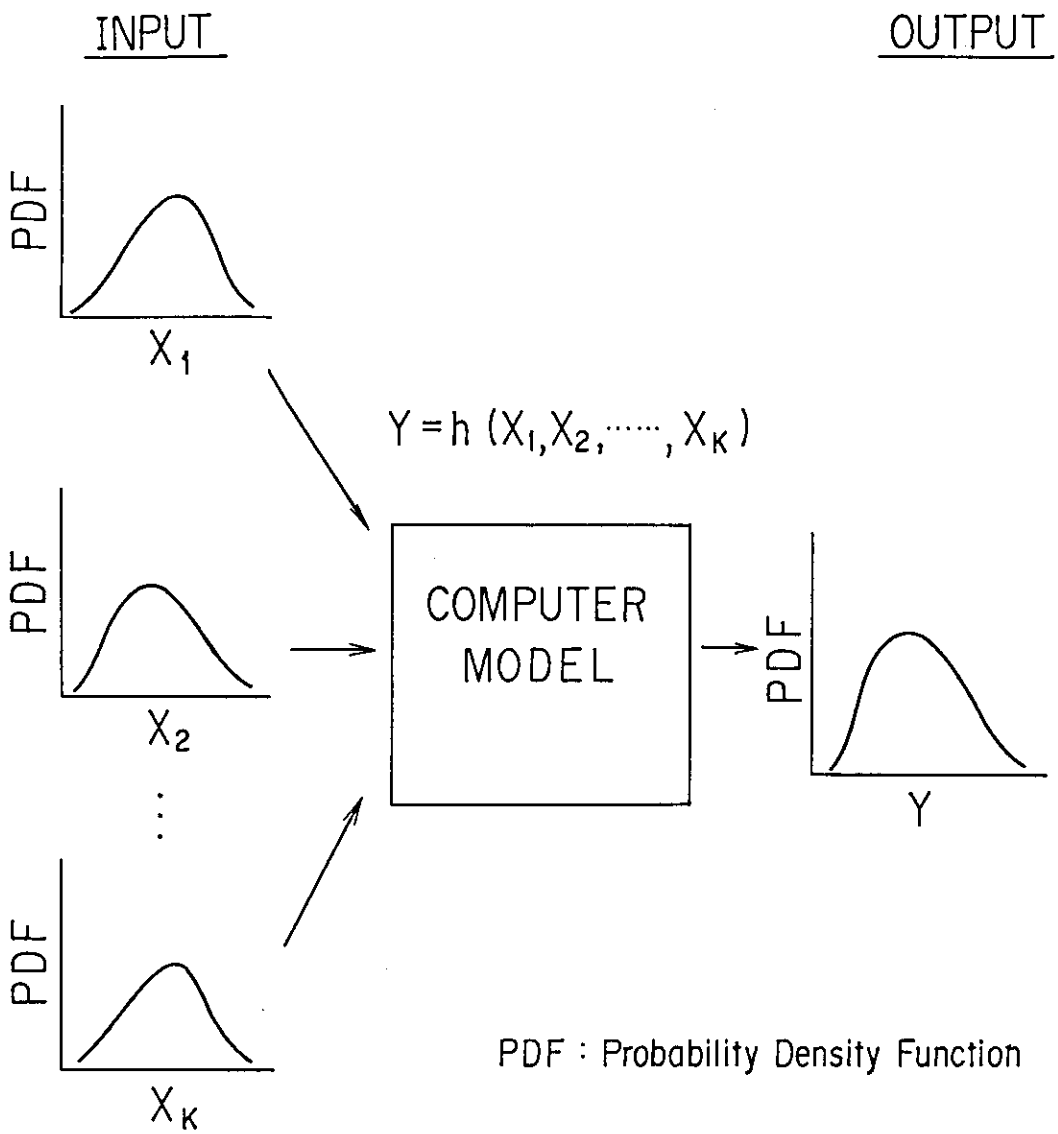

Fig. 1.1 Concept of uncertainty propagation through a computer model.

2. Procedure of Uncertainty Analysis

In order to estimate the uncertainties included in the output predicted by the computer model of concern, the following steps are to be followed:[5]

(1) Screening Analysis: This stage is necessary to reduce the number of input variables to a manageable size. This is accomplished by parametric sensitivity studies on the computer code, where calculations by varying input variable values one by one are made and change of the output is estimated.

(2) Uncertainty Propagation Analysis: This stage consists of (a) identification and classification, (b) quantification, and (c) propagation. Identification and classification of uncertainties entails a detailed examination of the model included in the computer code. The quantification process will entail using the available experimental data base to establish reasonable upper and lower bound estimates together with probability density functions (PDFs) for the sensitive input variables to the computer code. Then the propagation of input uncertainties through the code will be accomplished giving upper and lower bound values together with the PDF for the output.

(3) Uncertainty Importance Analysis: This stage will establish importance ranking of input variables with respect to the output uncertainty for the computer code.

(4) Output Distribution Sensitivity Analysis: Although uncertainty about the correct values of input variables can be quantified by treating the variables as random variables with appropriate PDFs, or, cumulative distribution functions (CDFs), these distributions are assigned on the basis of available data, combined with the judgment of experts. Because experimental data is often scarce, and expert opinions can vary widely, the input uncertainty distributions are not definitive; therefore, the impact of the input PDFs on the output PDF must be assessed. In this analysis stage, the sensitivity of the output PDF to the input PDFs, that is the relationships between alternative input PDFs and the corresponding output PDF, will be established.

This report describes the current status of techniques for the second, third and fourth steps. In the following discussion, we suppose that the first step, and the items (a) and (b) in the second steps have been established. In other words, we start with the assumption that the number of input variables included in the uncertainty analysis has been reduced to a manageable size, and that the input PDFs have been given in the following uncertainty analyses.

3. Description of Uncertainty Analysis Methods

3.1 Introduction

Quantification of uncertainty in the output Y, in general, needs a set of input/output data which will be obtained from numerical calculations with the original computer model. A larger size of the data set will be useful in estimating the uncertainty in Y more precisely because of decreased errors in statistics. If the computer model of concern is very simple and thus the computation time is short enough to make a large number of numerical calculations with the computer model possible, the uncertainty in Y will be estimated directly using the large number of input/output data based on the original computer model. This approach is known to be the Monte Carlo method. However, if it takes a long time to make computer runs with the computer model, the Monte Carlo method loses its effectiveness.

The current uncertainty analysis methods may include two kinds of approaches: One approach is based on replacing the original complex computer model by a simplified surrogate model such as a response surface model(6) (hereafter called response surface method), and the other does not rely on such a simplified surrogate model but relies on a limited number of calculations based on the original computer models (hereafter called direct method). The response surface method involves the simplified simulation of physical or logical models. In other words, the original output function, Eq.(1.1), is approximated by a simplified function:

$$Y = R(X_1, X_2, \ldots, X_K) \tag{3.1}$$

often termed a response surface. In building a replacement for the computer model, various sampling techniques are used such as random sampling, Latin hypercube sampling,(7),(8) experimental design,(9) and so on, in conjunction with a regression analysis.(10)-(12) After building the simplified surrogate model for the original computer model, a series of uncertainty analyses (the uncertainty propagation analysis, the uncertainty importance analysis, and the output distribution sensitivity analysis) will be performed with respect to the simplified surrogate model not to the original model. The results obtained, therefore, are only as valid as the approximations attendant in fits of the simplified surrogate model to the original model. The response surface methods which will be described in the following sections include the classical regression method, and the modified regression method.(13),(14)

The direct method refers to the approach that utilizes the input/output relationships provided by the original computer model calculations. Although the Monte Carlo method is the direct and proper tool for this approach if computation time is allowable, it is not practical for very long-running computer codes. In obtaining reasonable results from a limited number of input/output relationships, the key in the direct method exists in a sampling technique. Today the Latin hypercube sampling (LHS) technique is known to be useful and has been extensively used in uncertainty analysis for many fields. Also several methods for uncertainty analysis based on the LHS technique have been developed. The direct methods described in this report include the Monte Carlo method, the LHS method,(7),(8) the uncertainty reduction method,(15),(16) the method of closest distance,(17) the weighting method,(18) and the rejection method.(19)

Applicability of those uncertainty analysis methods mentioned above to each analysis step is shown in Table 3.1. The following sections will be devoted for describing these uncertainty analysis methods.

3.2 Response Surface Methods

3.2.1 Classical Regression Method

(1) Regression Equation

In a linear regression method, the function R in Eq.(3.1) is replaced by a linear polynomial function of input variables:

$$Y = a_o + \sum_{i=1}^{K} a_i X_i \ , \tag{3.2}$$

where a_o and a_i are constants fitted to the computer model input/output relationships based upon the LHS or an experimental design of the inputs.

As measures of input importance, the coefficients a_i become meaningful only in the case that the parameter inputs are dimensionally comparable. The problem of different units of measurement in the input variables can be eliminated by standardizing all variables:

$$X_i \longrightarrow X_i^* = (X_i - \langle X_i \rangle) / \sigma(X_i) \ , \tag{3.3a}$$

$$Y \longrightarrow Y^* = (Y - \langle Y \rangle) / \sigma(Y) \ , \tag{3.3b}$$

where $\langle X_i \rangle$ and $\langle Y \rangle$ are the means of X_i and Y for some samples, and $\sigma(X_i)$ and $\sigma(Y)$ the standard deviations of X_i and Y for the sample, respectively. Eq.(3.2) can now be rewritten in the following standardized form,

$$Y^* = \sum_{i=1}^{K} a_i^* X_i^* \ . \tag{3.4}$$

Here the coefficients a_i^* are called the standardized regression coefficients (SRCs).

To better account for nonlinearity in the original model, it is often more sensible to formulate rank regression equations using the variable ranks instead of the original variables.(10),(11) Specifically, the smallest value of each variable across sample members is assigned the rank 1, the next smallest value is assigned the rank 2, and so on up to the largest value which is assigned the rank N, where N denotes the total number of observations in the sample. Therefore, the rank regression form of Eq.(3.2) is given by

$$r_Y = \hat{a}_o + \sum_{i=1}^{K} \hat{a}_i r_i \ . \tag{3.5}$$

Here r_i and r_Y are ranks of X_i and Y, respectively. The r_i and r_Y are standardized according to the following relations:

$$r_i \longrightarrow r_i^* = (r_i - \langle r_i \rangle) / \sigma(r_i) \ , \tag{3.6a}$$

$$r_Y \longrightarrow r_Y^* = (r_Y - \langle r_Y \rangle) / \sigma(r_Y) \ , \tag{3.6b}$$

where $\langle r_i \rangle$ are $\langle r_Y \rangle$ are the mean values of r_i and r_Y for the sample, and $\sigma(r_i)$ and $\sigma(r_Y)$ are the standard deviations of r_i and r_Y for the sample, respectively. Therefore, using Eqs.(3.6a) and (3.6b), the standardized form of Eq.(3.5) is as follows:

$$r_Y{}^* = \sum_{i=1}^{K} \hat{a}_i{}^* \, r_i{}^* , \tag{3.7}$$

where $\hat{a}_i{}^*$ is called the standardized rank regression coefficient (SRRC). A value of Y is easily obtained from the rank r_Y by using an interpolation method. It is known that this linear rank regression equation can well approximate Y, when Y is a monotonic function of the X_i's.(10)

The goodness-of-fit of the regression model is measured by the quantity,(11)

$$R^2 = \sum_{j=1}^{N} (\hat{Y}_j - \langle Y \rangle)^2 \Big/ \{ \sum_{j=1}^{N} (\hat{Y}_j - \langle Y \rangle)^2 + \sum_{j=1}^{N} (Y_j - \hat{Y}_j)^2 \}, \tag{3.8}$$

called the coefficient of determination, where $\hat{Y}_j$ and Y_j are the raw values of Y given by the regression equation and the original computer model, respectively.

The regression equation, Eq.(3.2), can be expanded to include not only linear terms but also non-linear terms such as quadratic terms of X_i and X_j, by which the goodness of fit will be expected to be better.

(2) Standardized Regression and Partial Correlation Coefficients(12)

In the standardized regression equation, Eq.(3.4), the uncertainty width of Y^* or the variance of Y^*, V_{Y^*}, is given by

$$V_{Y^*} = \sum_{i=1}^{K} a_i{}^{*2} \, V(X_i{}^*) , \tag{3.9}$$

where $V(X_i{}^*)$ is the variance of the variable $X_i{}^*$, and independency among variables $X_i{}^*$'s has been assumed. Since the variables $X_i{}^*$ and Y^* are standardized, their variance should be one. Thus the Eq.(3.9) is rewritten as

$$V_{Y^*} = 1 = \sum_{i=1}^{K} a_i{}^{*2} . \tag{3.10}$$

The above equation, Eq.(3.10), shows that the uncertainty in Y^* is the sum of the square of SRC for X_i. The SRC, therefore, provides an importance measure with which to identify the variables which should be accounted for in a regression model. In other words, as is larger the absolute value of the SRC for X_i, the variable X_i makes larger contribution to the uncertainty in Y.

The partial correlation coefficient (PCC) is a measure of the unique linear relationship between two variables that cannot be explained in terms of the relationships of these two variables with any other variables. Thus, it provides an importance measure with which to identify the variables which should be accounted for in a regression model.

As an example, consider a linear model having only one input variable:

$$\hat{Y} = a_0 + a_1 X_1 . \tag{3.11}$$

The residuals from this model are denoted by $Y_i - \hat{Y}_i$ where Y_i = i-th observation value by original computer model, $\hat{Y}_i$ = i-th prediction using Eq.(3.11). The partial correlation for any remaining variable not in the model is found by computing the sample correlation coefficient between the residuals and that variable. Thus, a measure of linearity between any remaining variable and Y is obtained, given that an adjustment has been made for the variable(s) already in the model.

When nonlinear relationships are involved, it is often more appropriate to calculate PCCs on variable ranks rather than on the actual values for the variables: such coefficients are known as partial rank correlation coefficients (PRCCs).

3.2.2 Modified Regression Method(13),(14)

In this subsection, a modified regression model based upon the classical approach is described. Given a functional relationship between the computer model inputs and outputs of the form

$$Y = h(X_1, X_2, \ldots, X_K) , \tag{3.12}$$

a Taylor series expansion of the function h around a sample vector $\mathbf{X}_s = (X_{1s}, X_{2s}, \ldots, X_{Ks})$ can be effected to obtain:

$$Y \simeq h(\mathbf{X}_s) + \sum_{i=1}^{K} (X_i - X_{is}) \left. \frac{\partial h}{\partial X_i} \right|_{X_i = X_{is}} . \tag{3.13}$$

Here, $\mathbf{X}_s$ is one of the original sample members generated by the LHS method relative to the original input distributions. Now, in the light of alternative input distributions, a new sample of input vectors $\mathbf{X}$, where $\mathbf{X} = (X_1, X_2, \ldots, X_K)$ is generated. Eq.(3.13) is implemented relative to a new vector $\mathbf{X}$ by effecting the Taylor expansion about the original vector $\mathbf{X}_s$ (generated from the original distributions) that is the closest to the new vector $\mathbf{X}$, i.e., the original vector $\mathbf{X}_s$ that minimizes the quantity

$$\sum_{i=1}^{K} a_i^{*2} \, (X_i^* - X_{is}^*)^2 \quad . \tag{3.14}$$

The constant a_i^* is a weight that reflects the importance of the i-th input variable as measured by the SRC, and X_i^* and X_{is}^* are the standardized values of X_i and X_{is}, respectively. Here, Eq.(3.14) may be viewed as a modified Euclidean distance measure that accounts for the importance of an individual dimension (i.e., variable). It should be noted that the analytic form of the function h is unknown; however, the value of $h(\mathbf{X}_s)$ which is the computer model output corresponding to the sample input vector $\mathbf{X}_s$ is known. This is the case since the computer model utilized the original samples as input for the purpose of formulating the regression fit. The gradient of h at X_i is approximated as:

$$\left. \frac{\partial h}{\partial X_i} \right|_{X_i = X_{is}} \approx a_i \; , \tag{3.15}$$

that is, as a derivative with respect to the regression model of Eq.(3.2). Substituting Eq.(3.15) into Eq.(3.13) yields,

$$Y = h(\mathbf{X}_s) + \sum_{i=1}^{K} a_i (X_i - X_{is}). \tag{3.16}$$

Equation (3.16) can be recasted into the rank form by replacing X_i, Y, and a_i with r_i, r_Y, and $\hat{a}_i$, respectively. Therefore,

$$r_Y = \hat{h}(\mathbf{r}_s) + \sum_{i=1}^{K} \hat{a}_i (r_i - r_{is}). \tag{3.17}$$

Here, $\hat{h}(\mathbf{r}_s)$ indicates the rank of output Y corresponding to an input rank vector $\mathbf{r}_s$ of the original sample where the vector $\mathbf{r}_s$ is chosen to minimize the quantity

$$\sum_{i=1}^{K} \hat{a}_i^{*2} (r_i^* - r_{is}^*)^2 \tag{3.18}$$

and r_{is}^* is the standardized version of r_{is}. The constant $\hat{a}_i^*$ is a weight that reflects the importance of the i-th input rank as measured by the SRRC. Eq.(3.17) is now used as a surrogate for the original computer model.

3.3 Direct Methods

3.3.1 Monte Carlo Method

The Monte Carlo Method directly simulates the distribution of the output Y. After a joint probability distribution to the input variables $X_1, X_2, \ldots, X_K$ is assigned, a large number of independent samples $(X_{1i}, X_{2i}, \ldots, X_{Ki})$, i=1,...,M, from the assigned joint distribution are taken and the corresponding outputs Y_i, i=1,...,M, are calculated with the original computer model. The laws of probability ensure that this spectrum of outputs provides a good representation of the true output distribution. From the output data, summary statistics such as the mean, variance and percentile values of Y may be estimated.

Although the Monte Carlo method has the principal advantage of its very general applicability that there is no restriction on the form of the joint input distribution or on the nature of the relationship between input and output, the major disadvantage to the Monte Carlo method is the relatively large number, M, of runs often needed to obtain reliable information. Sometimes, M can be on the order of several thousands, which may rules out the use of the Monte Carlo method for very long-running computer codes. Another disadvantage is that Monte Carlo results do not tell us which variables are the most important contributors to output uncertainty.

In order to overcome the disadvantage of the relatively large number of computer runs needed in the Monte Carlo method, a type of stratified Monte Carlo sampling known as the LHS was proposed by McKay, Conover and

Beckman.[7] As for the disadvantage that the Monte Carlo method is powerless with respect to the uncertainty importance analysis, Ishigami and Homma[16] proposed the calculation technique called the uncertainty reduction method which is based on the importance measure proposed by Hora and Iman[15] and needs repeated runs according to the several different Monte Carlo or LH samples. The LHS method and the uncertainty reduction method will be described in the following subsections.

3.3.2 Latin Hypercube Sampling Method[7],[8]

The LHS technique provides an alternative to the conventional Monte Carlo approach. The LHS method can provide a precise estimate of the model's response to input variability with a smaller total number of calculations. This is accomplished by a constrained sampling scheme. LHS selects N different values from each of K variables $X_1, \ldots, X_K$ in the following manner. The range of each variable is divided into N intervals on the basis of equal probability. One value from each interval is selected at random with respect to the probability density in the interval. The N values thus obtained for X_1 are paired in a random manner (equally likely combinations) with the N values of X_2. These N pairs are combined in a random manner with the N values of X_3 to form N triplets, and so on, until N K-tuplets are formed. This is the Latin hypercube sample. Thus, for given values of N and K, there exist $(N!)^{K-1}$ possible interval combinations for a LHS. It is convenient to think of the LHS as forming an NxK matrix of input where the i-th row contains specific values of each of the K input variables to be used on the i-th run of the computer model. A particular row is often referred to as a LHS input vector.

3.3.3 Uncertainty Reduction Method[15],[16]

The uncertainty reduction method aims to quantify importance of input variables including uncertainties to the output uncertainty. The method is based on the importance measure proposed by Hora and Iman,[15] and the calculation technique to practically estimate the importance measure was developed by Ishigami and Homma.[16] The uncertainty reduction method does not rely on a simplified surrogate model of the original computer model but relies on a limited number of calculations based on the original model using the Monte Carlo or Latin hypercube sampling.

The importance measure I_j for the input variable X_j proposed by Hora and Iman is given by

$$I_j = \sqrt{V_Y - V_Y{}^j} \ , \qquad (3.19)$$

where V_Y is the variance of the output variable Y with uncertainties in all input variables $X_1, X_2, \ldots, X_K$ of concern being considered, and $V_Y{}^j$ is variance of Y with uncertainties in the other input variables than the variable X_j being considered by ascertaining the specified value of X_j. The expression of V_Y and $V_Y{}^j$ is given by, with an assumption that input variables are independent,

$$V_Y = \int .. \int \{ h(X_1, X_2, \ldots, X_K) - \langle Y \rangle \}^2 \prod_{i=1}^{K} f_i(X_i) dX_i$$

$$= \int .. \int h(X_1, X_2, \ldots, X_K)^2 \prod_{i=1}^{K} f_i(X_i) dX_i - \langle Y \rangle^2, \qquad (3.20)$$

$$V_Y{}^j = \int V_Y (\underline{x}_j) f_j(\underline{x}_j) d\underline{x}_j, \qquad (3.21)$$

where,

$$V_Y(\underline{x}_j) = \int .. \int \{ h(X_1, \ldots, X_{j-1}, \underline{x}_j, X_{j+1}, \ldots, X_K) - \langle h(\underline{x}_j) \rangle \}^2 \prod_{\substack{i=1 \\ (i \neq j)}}^{K} f_i(X_i) dX_i, \qquad (3.22)$$

$$\langle h(\underline{x}_j) \rangle = \int .. \int h(X_1, \ldots, X_{j-1}, \underline{x}_j, X_{j+1}, \ldots, X_K) \prod_{\substack{i=1 \\ (i \neq j)}}^{K} f_i(X_i) dX_i, \qquad (3.23)$$

$f_i(X_i)$ = PDF for the input variable X_i.

Here $\langle h(\underline{x}_j) \rangle$ and $V_Y(\underline{x}_j)$ are a mean and a variance of Y, respectively, with the input variable X_j ascertained by the specified value $\underline{x}_j$. Combining

Eqs.(3.20) through (3.23) leads to the relation:

$$V_Y - V_Y{}^j = U_j - \langle Y\rangle^2, \tag{3.24}$$

$$U_j = \int \langle h(\underline{x}_j)\rangle^2 f_j(\underline{x}_j)d\underline{x}_j. \tag{3.25}$$

Since the quantity $\langle Y\rangle^2$ is constant, the importance rankings are determined by the value of U_j. In other words, the input variable X_j is more important than X_i, if $U_j>U_i$. Hence the importance analysis to determine the importance rankings of input variables is reduced how to practically estimate the quantity U_j.

A calculation technique to estimate the quantity U_j of Eq.(3.25) is illustrated by rewriting it as

$$U_j = \int \langle h(\underline{x}_j)\rangle^2 f_j(\underline{x}_j)d\underline{x}_j$$

$$= \int\!..\!\int h(X_1,\dots,X_K)\ h(X'_1,\dots,X'_{j-1},X_j,\ X'_{j+1},\dots,X'_K)$$

$$\Big(\prod_{i=1}^{K} f_i(X_i)dX_i\Big)\Big(\prod_{\substack{i=1\\(i\neq j)}}^{K} f_i(X'_i)dX'_i\Big). \tag{3.26}$$

The above equation, Eq.(3.26), shows that the quantity U_j is nothing more than an expectation value of the function defined by

$$H(X_1,\dots,X_K,\ X'_1,\dots,X'_{j-1},\ X'_{j+1},\dots,X'_K)$$

$$= h(X_1,\dots,X_K)\ h(X'_1,\dots,X'_{j-1},X_j,X'_{j+1},\dots,X'_K), \tag{3.27}$$

with (2K-1) independent random variables $(X_1,\dots,X_K,\ X'_1,\dots,X'_{j-1},\ X'_{j+1},\dots,X'_K)$ whose PDF is given by $\Big(\prod_{i=1}^{K} f_i(X_i)\Big)\Big(\prod_{\substack{i=1\\(i\neq j)}}^{K} f_i(X'_i)\Big)$. Thus the quantity U_j would be estimated numerically with allowable Monte Carlo calculations. It is noted that to estimate U_j the required computer run number for the original model function $h(X_1,\dots,X_K)$ is 2xM, where M is a number presenting Monte Carlo sample size, because the function H is expressed by twofold product of the original model function h.

If Monte Carlo calculations for the original model $h(X_1,X_2,\ldots,X_K)$ is costly and impractical, the LHS method would be applicable which requires much smaller number of samples than that of the Monte Carlo method. Now the computer run number to estimate U_j is reduced to 2xN and thus the computer run number needed to determine importance rankings of input variables is reduced to 2xNxK, where N is a number presenting LH sample size. It should be emphasized that this technique is expected to give more accurate results as the size of the Monte Carlo or the LH samples becomes larger, owing to reduced error in statistics. In addition, it should be noted that the uncertainty reduction method is applicable in the case where input variables of concern include such ones as selection of models embedded in the computer models.

3.3.4 Method of Closest Distance(17)

The method of closest distance can be used for the output distribution sensitivity analysis to determine the impact of the probability distribution functions characterizing the input variable uncertainties on the output distributions.

The method is based upon the modified regression method. It is a method that utilizes the input/output relationships provided by the original computer model calculations based upon a single LH sample.

Elimination of the second term of Eqs.(3.16) and (3.17) yields,

$$Y \approx h(\mathbf{X}_s), \tag{3.28}$$

and

$$r_Y \approx \hat{h}(\mathbf{r}_s). \tag{3.29}$$

The method comprises the following steps:

(1) A Latin hypercube sample is generated:

$$\mathbf{X}_i = (X_{1i},X_{2i},\ldots,X_{Ki}), \quad i=1,2,\ldots,N\ . \tag{3.30}$$

Here, the N sample members correspond to N combinations of values for the K parameter inputs. The input vector (i.e., sample member) $\mathbf{X}_i$ yields the output Y_i from the computer code where, for simplicity, just one output variable is considered.

(2) In order to ascertain the effects of the input PDFs on the output distributions, another Latin hypercube input sample,

$$\mathbf{X}_j' = (X_{1j}', X_{2j}', \ldots, X_{Kj}'), \quad j=1,2,\ldots,N' \tag{3.31}$$

is generated. This sample is obtained with respect to the different input PDFs from those employed in step (1).

(3) The output value Y_j' corresponding to the randomly sampled input vector $\mathbf{X}_j'$ is approximated by the LHS output value Y_s whose corresponding LHS rank vector $\mathbf{r}_s$ is "closest" to the rank vector $\mathbf{r}_j'$ corresponding to $\mathbf{X}_j'$, where Eq.(3.18) provides the definition of "closeness". Then the corresponding original output Y_s is used as an approximate replacement for the output Y_j'. Hence N' random output values are approximated by the nearest of the N' LHS output values obtained from the original computer code calculations.

(4) An approximation of the output distributions resulting from the second set of input distributions is thereby compiled in the light of the original computer model calculations. These new output distributions may be compared to the original output distributions in order to ascertain their sensitivity to the input PDFs. This approach is identical to the modified regression method described in subsection 3.2.2 except that the regression based terms are excluded from the surrogate model of the original computer model.

3.3.5 Weighting Method(18)

A small sample sensitivity analysis technique which directly utilizes the computer model results generated based upon LHS of input distributions has been proposed by Iman, et al.(18) This method can be used to determine the impact of the probability distribution functions characterizing the input variable uncertainties on the output distributions without the need for computer calculations in addition to those used to determine the original output PDFs and without relying on a response surface representation of the physical model.

Iman, et al. show that if the probability density function of a single input variable X_i is changed from $f_i(X_i)$ to $q_i(X_i)$, then the mean $\langle Y \rangle$, standard deviation $\sigma(Y)$, and cumulative probability function $C(Y)$ of the output variable Y may be approximated by

$$\langle Y \rangle_W = \sum_{j=1}^{N} W_j \, Y(j), \tag{3.32}$$

$$\sigma_W(Y) = \sum_{j=1}^{N} W_j \, (Y(j) - \langle Y \rangle_W)^2, \tag{3.33}$$

$$C_W(Y) = \sum_{j=1}^{N} W_j \, u(Y - Y(j)), \tag{3.34}$$

where u is the unitary step function defined by

$$u(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \, , \end{cases} \tag{3.35}$$

$Y(j)$ is the computer model output corresponding to the LHS input vector in which the rank of X_i is j, N is the LH sample size, and the weighting factor, W_j, is given by the probability with respect to $q_i(X_i)$ that the reference variable X_i takes a value in the j-th interval of the original stratification of the parameter space with respect to $f_i(X_i)$:

$$W_j = \int_{X_i(j-1)}^{X_i(j)} q_i(X_i) dX_i . \tag{3.36}$$

Here $X_i(j)$ is the upper bound of X_i in the j-th interval and, for given j, is determined by

$$j/N = \int_{X_i(0)}^{X_i(j)} f_i(X_i) dX_i , \tag{3.37}$$

since in the LHS approach, the range of each variable is divided into N nonoverlapping equiprobable intervals.

3.3.6 Rejection Method(19)

In this method proposed by Beckman and McKay,(19) a subset of the original computer model outputs corresponding to a Monte Carlo sample from the original input distributions is selected to provide the appropriate statistical outputs corresponding to a new set of input distributions.

Consider the input variable vector $\mathbf{X} = (X_1, X_2, \ldots, X_K)$ and the sample vectors

$$\mathbf{X}_j = (X_{1j}, X_{2j}, \ldots, X_{Kj}), \quad j=1 \text{ to } N \tag{3.38}$$

which are generated with respect to the PDF $f(\mathbf{X})$. Let the output of the computer model corresponding to the input $\mathbf{X}_j$ be Y_j. The rejection method relies on selection of some subset of the sets of variables $(\mathbf{X}_j, Y_j)$. Let the new input PDF be $q(\mathbf{X})$.

It is necessary that there exists a uniform bound M such that,

$$\frac{q(\mathbf{X})}{f(\mathbf{X})} < M \tag{3.39}$$

for all $\mathbf{X}$ and that the domain of $q(\mathbf{X})$ be contained within the domain of $f(\mathbf{X})$. Let the random variable V, given as sample vector $\mathbf{X}_j$, be assigned a uniform probability distribution between 0 and $M \cdot f(\mathbf{X}_j)$. The data set $(\mathbf{X}_j, Y_j)$ is then retained as a sample from $q(\mathbf{X})$ if a random realization of V is less than $q(\mathbf{X}_j)$. The theoretical basis for this approach is expounded in Ref. 19.

Table 3.1 Applicability of various methods to each step of the uncertainty analysis.

Method	Uncertainty propagation analysis	Uncertainty importance analysis	Output distribution sensitivity analysis
Response surface method			
Classical regression method	A	A	A
Modified regression method	A	A	A
Direct method			
Monte Carlo method	A	NA	A
LHS method	A	NA	A
Uncertainty reduction method	NA	A	NA
Method of closest distance	NA	NA	A
Weighting method	NA	NA	A
Rejection method	NA	NA	A

A : Applicable
NA : Not applicable
LHS : Latin hypercube sampling

4. Application to Deterministic Models

The uncertainly analysis methods described in the previous chapter have been applied to three computer models, the MARCH/CORRAL II Model,(20),(21) the TERFOC model(22) and the SPARC model.(23) Table 4.1 summarizes methods utilized in each analysis step for three computer models.

4.1 Application to the MARCH/CORRAL II Model(24)

This section illustrates results of the uncertainty analysis for the MARCH/CORRAL II model using the regression method to estimate uncertainties in the fission product release to the environment during core meltdown accidents at a BWR Mark-I plant. Investigated are identification of important input variables and uncertainty propagation. A sampling technique of an orthogonal factorial design is used to determine a response surface, and analysis of variance and regression analysis are performed for the MARCH/CORRAL II input/output relationships. Uncertainty propagation is analyzed using the Monte Carlo sampling technique based on the response surface obtained.

In the analysis, the computer code system VARS(25) developed by Japan Atomic Energy Research Institute (JAERI) has been utilized.

4.1.1 Description of the MARCH/CORRAL II Model

The purpose of the MARCH/CORRAL II code is to estimate the radiological release following a severe nuclear reactor accident. The MARCH code is a computer code to analyze the thermal hydraulic response of the reactor core, the primary coolant system, and the containment system in light water reactors during a course of core meltdown accidents. The CORRAL II code is a computer code to analyze the fission product transport and deposition in the containment system of water-cooled reactors and to evaluate environmental fission product release fractions of the core inventory. Here the MARCH output supplies thermal hydraulic conditions in the containment to the CORRAL II input.

4.1.2 Accident Sequences Analyzed

The accident sequences analyzed are a station blackout accident and a loss of decay heat removal accident at a BWR plant with Mark-I type containment.

Station Blackout Accident

In the accident, AC electric power of both the off-site power and the diesel generator is lost completely and only the DC electric power in the station is available. In the analysis, DC electric power was assumed to be exhausted in four hours after the initiation of the accident. Therefore high pressure coolant injection (HPCI) and reactor core isolation cooling (RCIC) systems using the turbine driven pumps were assumed to be operable for the first four hours. After water injection capability from HPCI/RCIC pumps is lost, the station blackout would develop a core meltdown accident, involving subsequent containment failure which would bring a large amount of fission products released to the environment.

Loss of Decay Heat Removal Accident

In the sequence, a transient occurs that renders the power conversion system unavailable as a heat sink for the reactor, and then a reactor scram occurs. As decay heat of the core continues to be transferred to the coolant, pressure in the reactor coolant system (RCS) increases until the safety relief valves (SRVs) open. After the SRVs open, steam from the RCS is discharged into the suppression pool. Then the HPCI and RCIC systems automatically start to recover the water level in the reactor pressure vessel (RPV). Then SRVs are manually opened (assumption) to depressurize the RPV and the low pressure coolant injection (LPCI) is used to retain the water level. Steam in the RPV discharges into the suppression pool through the SRVs. Thus the decay heat of the core is transferred to the torus water. However, failure of the residual heat removal (RHR) system causes the torus water temperature and the containment pressure to increase until the containment fails. In the study, it was assumed that the drywell would fail when the pressure or temperature in the containment reaches 174.7 psia or 500 ^{o}F, respectively. Upon containment failure the core cooling injection was assumed to fail. Loss of all injection results in decreasing the water level in RPV, core uncovery, core melt and subsequent vessel failure.

4.1.3 Variables Considered in the Analysis and Sampling of Input Variable Values

(1) Variables considered in the analysis

The output from the CORRAL II code includes environmental release fractions of the core inventory of eight groups of fission products; noble gas, organic iodine, elemental iodine, and particulates (Cs-Rb, Te, Ba-Sr, Ru, La). Two output variables were selected for consideration: the environmental release fraction of the core inventory of elemental iodine and Cesium-Rubidium (Cs-Rb).

The input variables selected in the analysis of the station blackout accident and the loss of decay heat removal accident are shown in Tables 4.2 and 4.3, respectively, with the ranges and distribution patterns of the input variables. These ranges of the variable values were determined based on experimental data and/or a subjective judgment.

(2) Selection of input variable values

The input variable values were selected using an orthogonal factorial design, one of the experimental designs, where input variables were treated as ones with three levels. The orthogonal factorial designs used are $L_{243}(3^{121})$ and $L_{81}(3^{40})$ in the analysis of the station blackout accident and the loss of decay heat removal accident, respectively.

4.1.4 Analysis of Variance and Regression Analysis

Based on the MARCH/CORRAL II input/output relationships using the factorial orthogonal design, analysis of variance was performed to identify important input variables for the amount of fission products released to the environment. The analysis of variance partitions variance of the output into sum of squares of a variable (main effect) and/or an interactive variable set (interaction).

In regression analysis, input variables identified to be important by analysis of variance were considered as candidates for terms in a regression equation. The regression equation was determined in a step wise manner; At each step one term was added to the regression equation so as to make the greatest reduction in the error sum squares. The resulting regression equation took a simple form composed of polynomials up to second order of the input variables.

The uncertainty propagation analysis was made based on the regression equation using a Monte Carlo sampling technique with joint probability distributions for the input variables being assigned.

4.1.5 Results for the Station Blackout Accident

(1) Elemental iodine

Table 4.4 shows the results obtained by analysis of variance. It can be seen that the important input variables affecting the calculated elemental iodine release fraction to the environment are FDROP (fractional holdup of melted core at core slumping), DCF (suppression pool decontamination factor) and DFI (natural deposition factor for elemental iodine) whose values of mean squares are relatively large. The interaction between TMELT (fuel melting temperature) and FDROP is recognized to be important. In the uncertainly propagation analysis, two sets of distribution functions were assumed; in one case (case 1) the PDFs of input variables were assumed to be uniform or loguniform distributions, and the other case (case 2) they were assumed to be normal or lognormal distributions.

Figures 4.1 and 4.2 show the calculated probability distributions of elemental iodine release fraction to the environment in the case 1 and in the case 2, respectively. The mean and median values calculated are 0.022 and 0.021, respectively in the case 1, being 0.019 and 0.018 in the case 2. The upper (the 95th percentile) and the lower (the 5th percentile) bounds calculated are 0.043 and 5.6×10^{-3}, respectively in the case 1, being 0.039 and 4.6×10^{-3} in the case 2. Thus the difference between the upper and the lower bounds is about one order of magnitude.

(2) Cesium-Rubidium (Cs-Rb)

As for Cs-Rb, an analysis similar to the one for elemental iodine was performed.

Table 4.5 shows the results obtained by analysis of variance. We can see that the important input variables affecting the calculated Cs-Rb release fraction to the environment are DFP (natural deposition factor for particulate), DCF (suppression pool decontamination factor) and DCF*DFP whose values of mean squares are relatively large. It should be noted that the interaction between DCF and DFP is recognized to be important.

Figures 4.3 and 4.4 show the calculated probability distributions of Cs-Rb release fraction to the environment in the case 1 and in the case 2, respectively. The mean and median values calculated are 0.080 and 0.085, respectively in the case 1, being 0.079 and 0.083 in the case 2. The calculated upper (the 95th percentile) and the lower (the 5th percentile) bounds calculated are 0.15 and 9.3×10^{-3}, respectively in the case 1, being 0.14 and 0.011 in the case 2. Thus the difference between the upper and the lower bounds is about one order of magnitude.

4.1.6 Results for the Loss of Decay Heat Removal Accident

(1) Elemental iodine

Table 4.6 shows the results of analysis of variance. It can be seen that the important input variables affecting the calculated elemental iodine release fraction to the environment are MWORNL (zirconium-water reaction model), FDROP and DFI2 (natural deposition factor for iodine) whose values of mean squares are relatively large.

In the uncertainly propagation analysis, the value of the variable MWORNL was assumed to be 1 (Cathcart) or 0 (Baker-Just) with the same probability (fifty percent). For the other variables, two sets of distribution functions were assumed; in one case (case 1) the PDFs of input were assumed to be uniform or loguniform distributions, and the other case (case 2) they were assumed to be normal or lognormal distributions.

Figures 4.5 and 4.6 show the calculated probability distributions of elemental iodine release fraction to the environment in the case 1 and in the case 2, respectively. The mean and median values calculated are 0.032 and 0.031, respectively in the case 1, being 0.028 and 0.028 in the case 2. The upper (the 95th percentile) and the lower (the 5th percentile) bounds calculated are 0.064 and 0.0053, respectively in the case 1, being 0.055 and 0.0042 in the case 2. Thus the difference between the upper and the lower bounds is about one order of magnitude.

(2) Cesium-Rubidium (Cs-Rb)

Table 4.7 shows the results obtained by analysis of variance. We can see that the important input variables affecting the calculated Cs-Rb release fraction to the environment are PDIA (particle diameter), DFP, MWORNL and FDROP whose values of mean squares are relatively large. It is noted that the interaction between variables is small.

Figures 4.7 and 4.8 show the calculated probability distributions of Cs-Rb release fraction to the environment in the case 1 and in the case 2, respectively. The mean and median values calculated are 0.10 and 0.095, respectively in the case 1, being 0.089 and 0.085 in the case 2. The upper (the 95th percentile) and the lower (the 5th percentile) bounds calculated are 0.20 and 0.015, respectively in the case 1, being 0.17 and 0.016 in the case 2. Thus the difference between the upper and the lower bounds is about one order of magnitude.

4.1.7 Summary

The uncertainty analysis for environmental fission product release included two main features: One was identification of important parameters affecting the calculated results and the other was a quantitative estimate of uncertainty for fission product release fraction to the environment. The former was obtained by analysis of variance and the latter by Monte Carlo simulation based on the response surface.

For the station blackout accident, important parameters identified are suppression pool scrubbing factor, natural deposition factor and a fractional holdup of melted core at core slumping as main effects and the interaction between core melting temperature and a fractional holdup of melted core at core slumping. For the loss of decay heat removal accident, important parameters identified are a fractional holdup of melted core at core slumping, natural deposition factor and zirconium-water reaction model, while the interaction between input variables was small. The reason why suppression pool decontamination factor was not significant is conjectured as follows. In the accident, the containment had failed before fission product was released. Thus the released fission products were directly carried to the atmosphere with a small effect of suppression pool scrubbing.

The results obtained by Monte Carlo simulation were similar for both the accident sequences. It was shown that in both accident sequences, the difference between the upper and lower bounds of environmental fission product release fractions was about one order of magnitude.

4.2 Application to the TERFOC Model(16)

This section illustrates results of the uncertainty importance analysis for the TERFOC model using the regression method and the uncertainty reduction method to identify important input variables. A sampling technique of the LHS was utilized in both the methods.

4.2.1 Description of the TERFOC Model

The purpose of the TERFOC (Terrestrial Food Chain) computer code developed by JAERI is to assess the potential radiological impact of routine releases of radioactive effluents from nuclear facilities.(22) The TERFOC model calculates the concentration of radionuclides in vegetation and animal products as the result of deposition of radionuclides on agricultural land.

To simplify the problem for the analysis an average long-term concentration of ^{131}I [Bq/m^3] in the above ground atmosphere was assumed and the physico-chemical form of iodine was assumed to be 10% elemental, 50% organic and 40% particulate. The equation used for predicting the i-th radionuclide concentration, C_m [Bq/ℓ], in milk is given by

$$C_m = F_m \, (Q_{FF}C_{VF} + Q_{FS}C_{VS}) \exp(-\lambda_i t_f), \qquad (4.1)$$

F_m = feed to milk transfer factor [d/ℓ]
Q_{FF} = daily dry intake of fresh forage by dairy cows [kg/d]
Q_{FS} = daily dry intake of stored forage by dairy cows [kg/d]
t_f = time delay from production to consumption of milk [d]
C_{VF} = radionuclide concentration in fresh forage [Bq/kg]
C_{VS} = radionuclide concentration in stored forage [Bq/kg]
λ_i = physical decay constant of the i-th radionuclide [1/d]

4.2.2 Variables Considered in the Analysis

Of many output variables provided by the TERFOC model, one output variable was selected in the importance analysis for simplicity. The output variable selected is the ^{131}I concentration in milk, C_m given by Eq.(4.1), at the time of 30 years. The TERFOC input variables selected are given in Table 4.8 together with their assigned ranges and distribution patterns as used for the importance analysis. These twelve input variables were sampled independently using the LHS technique in accordance with the ranges and distribution patterns of the input variables.

4.2.3 Results of the Importance Analysis

Based on the TERFOC-run results for two hundreds LHS input vectors, the evaluation of PCCs, PRCCs, SRCs, and SRRCs was performed using the computer program in Ref. 12. Table 4.9 shows these coefficients and the coefficient of determination, R^2, for the ^{131}I concentration in milk, C_m, at the time of 30 years. The results show that the fit of the regression model to the TERFOC model is almost satisfactory ($R^2 > 0.8$). It is found that the three input variables, F_m, RP and ω_P, as revealed by their high correlation coefficients, predominantly govern the uncertainty in the output C_m.

In calculating the quantity U_j relating to the importance measure I_j for the j-th input variable, the TERFOC computer runs for four hundreds LHS input vectors were performed to obtain the expectation value of the function H, Eq.(3.27), corresponding to the output variable C_m. Here the number of "input" variables in the function H is not twelve but twenty three (=2x12-1), and this is why the LH sample size is larger by two times than the LH sample size used in the regression analysis. The calculated values of U_j for twelve input variables are shown in Table 4.10. The results indicate that the input variables, F_m, ω_P and RP, make dominant contribution to the uncertainty in the output variable C_m.

The rankings of input variable obtained from the uncertainty reduction method are also shown in Table 4.10 together with the results by the regression analysis of PCCs, PRCCs, SRCs and SRRCs. The results show the following;

(a) Both the uncertainty reduction method and the regression method identify the three dominant input variables, F_m, ω_P and RP, that contribute to the output uncertainty.

(b) There can be seen some differences between the results by the uncertainty reduction method and those by the regression method. This may reveal that the fit of the regression model to the TERFOC model is not completely adequate in spite of high value of R^2.

4.3 Application to the SPARC Model(13),(14)

This section illustrates results of the output distribution sensitivity analysis for the SPARC model using several methods described in the previous chapter. Merits and limitations of the methods are examined by detailed application to SPARC.

4.3.1 Description of the SPARC Model

The suppression pool aerosol removal code (SPARC) calculates the scrubbing of fission products released from the reactor coolant system (RCS) into the pressure suppression pool of boiling water reactors (BWRs) during postulated severe reactor accidents. This code is part of the source term code package (STCP), and is particularly suited for the purpose of demonstrating merits of the various sensitivity analysis methods for the following reasons:

(1) Relatively small computational requirements.

(2) Due to a limited number of input variables, a small of LH samples will suffice for the analysis.

(3) Owing to the above, the SPARC code can be readily exercised for several different LHS inputs. This enables comparison of the sensitivity analysis methods with direct SPARC sensitivity calculations.

4.3.2 Variables Considered in the Analysis

For the purpose of the sensitivity analysis with the SPARC model, the following calculational outputs were tracked.

(a) The integral decontamination factor (DF) for CsI defined by:

$$DF = \frac{\int_{t_i}^{t_f} M_{in}\, dt}{\int_{t_i}^{t_f} M_{out}\, dt} . \qquad (4.2)$$

(b) The total leakage amount of all radionuclides into the wetwell airspace defined by;

$$L = \int_{t_i}^{t_f} \sum_j M^j_{out}\, dt , \qquad (4.3)$$

where M_{in} is the mass of CsI aerosols entering the pool from the RCS, M_{out} is the mass of CsI aerosols leaving the pool and entering the suppression pool's wetwell airspace region, t_i is the initial time, t_f is the final time (t_f-t_i is the scrubbing duration), and the superscript j corresponds to the j-th radionuclide species entering the wetwell airspace region.

The SPARC input variables selected together with their assigned ranges and probability distribution patterns as used for the reference analysis are given in Table 4.11. These six input variables were sampled using the LHS technique in accordance with the ranges and distribution patterns of the input variables. A sample size of fifty was considered corresponding to fifty combinations of values for the six input variables.

4.3.3 Reference Analysis

Based on the SPARC-run results for fifty initial LHS input vectors (LHS-1), the evaluation of PRCCs and SRRCs was performed using the computer program in Ref. 12. Table 4.12 shows the PRCCs, SRRCs and the coefficient of determination, R^2, for the integral DF for CsI and the total leakage of all radionuclides into the wetwell atmosphere. These results indicate that the fit of the regression model to SPARC for the reference output variables is satisfactory ($R^2 > 0.9$). It is found that two input variables, X_1 (RATIO) and X_3 (VSWARM), as revealed by their high correlation coefficients, predominantly govern the magnitude of uncertainties in the outputs DF and L.

Table 4.13 shows some properties of the output distributions, resulting from propagation of the 50 input LH vectors through SPARC. These include the mean, standard deviation, the 5th, 50th and 95th percentile values for DF and L, respectively.

Of course, numerical simulation techniques such as LHS provide only an estimate of the output distributions that would in principle be generated by the exact analytic propagation of the input distributions. In order to provide an appreciation for the impact of the LHS approach on the calculated results, additional SPARC calculations were performed using a different set of fifty LHS input vectors (LHS-2), although sampled from the same input distributions. These comparisons are given in Tables 4.12 and 4.13. Even though relatively large differences in the calculated PRCCs and SRRCs exist for the unimportant input variables X_4, X_5 and X_6, the impact of LHS on the important input variables is shown to be insignificant.

4.3.4 Output Distribution Sensitivity Analysis

The reference analysis of the previous subsection showed that the two SPARC variables X_1 (RATIO) and X_3 (VSWARM) were the most significant contributors to both the integral DF for CsI (DF), and the total leakage of all radionuclides into the wetwell atmosphere (L). The sensitivity analysis, therefore, focused attention on the effect of varying the PDFs of the most important input variables, X_1 and X_3, on the PDFs of the output variables DF and L.

Table 4.14 lists the assumed distributions for X_1 and X_3 in the sensitivity cases as compared with the reference analysis of subsection 4.3.3. The mean and the range of each input variable in the sensitivity analysis were assumed to be the same as those given in Table 4.11 for the reference analysis.

The sensitivity of the output variables will be determined by changing the distributions for X_1 (case S-1) and X_3 (case S-2) from uniform to normal using the sensitivity methods described in Chapter 3. Hence, the approach to be adopted is one in which the original LHS-1 results provide the basis for applying the response surface methods and direct methods described in Chapter 3. By comparing the results thereof with the output distributions based upon LHS of the new input distributions (LHS-3 and LHS-4 of Table 4.14) and runs of the actual computer model SPARC, the success of the methods may be assessed. Here, three LH sample sizes of fifty are considered. Note that since simple regression models are being utilized to propagate the samples LHS-3 and LHS-4, we could equally well have used larger samples acquired from the same distributions with respect to which LHS-3 and LHS-4 were generated. However, for more direct comparison with the original computer code predictions, we use the regression models to propagate the actual LHS-3 and LHS-4 samples.

The normal distribution is of the form:

$$q(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \tag{4.4}$$

where the distribution parameters, μ and σ are determined by requirement[8] of the same mean and range between the normal distribution q(x) and uniform distribution f(x) = 1/(b-a):

$$\int_a^b x\, f(x)dx = \int_{-\infty}^{\infty} x\, q(x)dx \tag{4.5}$$

and

$$\int_{-\infty}^{a} q(x)dx = \int_b^{\infty} q(x)dx = 0.001. \tag{4.6}$$

Eqs. (4.5) and (4.6) give the values of μ and σ as:

$$\mu = (a+b)/2 \tag{4.7}$$

$$\sigma = (b-a)/6.182 \tag{4.8}$$

4.3.4.1 Response Surface Methods

(1) Classical Regression Model

The adopted standardized rank regression equation is:

$$r_Y^* = \sum_{i=1}^{K} \hat{a}_i^* \, r_i^* \, . \tag{4.9}$$

Table 4.15 lists the SRRCs, $\hat{a}_i^*$, and the coefficient of determination, R^2, for the integral DF and the total leakage as calculated by the computer program in Ref. 12. These results indicate that the fit of Eq.(4.9) to SPARC for the reference output variables is satisfactory ($R^2>0.9$).

(2) Modified Regression Model

Based on the rank regression analysis, the adopted modified rank regression equation is

$$r_Y = \hat{h}(\mathbf{r}_s) + \sum_{i=1}^{K} \hat{a}_i \, (r_i - r_{is}) \, . \tag{4.10}$$

Here the coefficient $\hat{a}_i$ is identical to the SRRC, $\hat{a}_i^*$, in Table 4.15. This is ensured since both the ranks of X_i and Y cover the same range, 1 to 50. The nearest LHS-1 input rank vector, $\mathbf{r}_s$, to a given LHS-3 or LHS-4 vector was determined using Eq.(3.18) with the values of $\hat{a}_i^*$ given in Table 4.15.

4.3.4.2 Direct Methods

(1) Weighting Method

In the weighting method, statistical parameters such as mean, the standard deviation and the CDF for the reference output variable Y were calculated by Eqs.(3.32) through (3.34). The weighting factors W_j corresponding to the change in distributions given in Table 4.14 were calculated using Eqs.(3.36) and (3.37), and the results are shown in Table 4.16.

It is noted that the weighting method requires neither the LHS-3 nor the LHS-4 samples. This is the case since this method requires only a knowledge of the new input distributions, but does not require a sampling of these distributions to be effected.

(2) Rejection Method

This method requires a uniform bound M defined in Eq.(3.39) as

$$M = \mathrm{MAX}\,\{\,q(\mathbf{X})/f(\mathbf{X})\,\}. \tag{4.11}$$

Such a bound indeed exists and in this application is given by

$$\begin{aligned} M &= \frac{b-a}{\sqrt{2\pi}\ \sigma} \\ &= \frac{6.182}{\sqrt{2\pi}} \\ &= 2.47\ , \end{aligned} \tag{4.12}$$

where Eqs.(4.4) through (4.8) have been used. Again this method requires a knowledge only of the new input distributions and does not require the samples LHS-3 and LHS-4.

(3) Method of Closest Distance

In this method, Eq.(3.29) based upon the rank data was used. The nearest LHS-1 input rank vector, $\mathbf{r}_s$, to a vector of LHS-3 or LHS-4 is determined using Eq.(3.18) with the values of $\hat{a}_i^*$ given in Table 4.15.

4.3.4.3 Results and Discussions

Tables 4.17 through 4.20 summarize the calculated statistical parameters for the output variables DF and L based directly upon the SPARC results as well as upon the five sensitivity analysis methods. A comparison of the calculated cumulative distribution functions (CDFs) based on the five methods is given in Figures 4.9 through 4.12.

Comparison of the sensitivity methods against the SPARC results shows the following:

(1) Generally good agreement is achieved between the direct SPARC results and the results of the classical rank regression method, the modified rank regression method, the rejection method and the method of closest distance.

(2) The weighting method shows a good agreement for the calculated mean

and median, however, the calculated standard deviation, the 5th and 95th percentiles show large deviations compared with other methods.

(3) Of the original 50 LHS sample members, application of the rejection method dictated the retention of 22 sample members for both sensitivity cases S-1 and S-2. These retained sample members then provided a basis for making inferences relative to the new input PDFs. Note that this number 22 is broadly consistent with the theoretical frequency of retention(19) of 1/M, where M (=2.47 in this application) is the uniform bound defined in subsection 3.3.6. While the rejection method has performed well in the current investigation, it should be borne in mind that this performance would be expected to degrade in circumstances where the initial (pre-rejected) number of sample members is small or where the bound M is large. In either case, the number of retained sample members would be small.

Table 4.1 Methods utilized in each analysis step.

Analysis step	MARCH/CORRAL II	TERFOC	SPARC
Uncertainty propagation analysis	. regression method (response surface method)	---	. LHS method
Importance analysis	. method of analysis of variance	. regression method . uncertainty reduction method	. regression method
Output distribution analysis	---	---	. regression method . modified regression method . method of closest distance . weighting method . rejection method

Table 4.2 MARCH/CORRAL II input variables (parameters), their ranges and distributions considered in the uncertainty analysis for the station blackout accident.

Variable	Description	Range		Distribution	
		Lower bound	Upper bound	Case 1	Case 2
DCF	Suppression pool decontamination factor [-]	1.2	10^6	loguniform	lognormal
WDED	Mass of water left in primary at the end of meltdown [1b]	1.75×10^4	7.0×10^4	loguniform	lognormal
TMELT(a)	Fuel melting temperature [F]	3362	5072	uniform	normal
FDROP(b)	Fractional holdup of melted core before slumping into bottom head starts [-]	0.4	0.9	uniform	normal
FZOCR(c)	Fraction of ZrO_2 in the center of debris particle [-]	0.0	0.5	loguniform	lognormal
COND	Thermal conductivity of debris [Btu/(hr·ft·K)]	4.31	25.1	loguniform	lognormal
DP	Debris particle diameter [inch]	0.0025	2.5	loguniform	lognormal
FLRMC	Mass times heat capacity of structural metal below the vessel [Btu/F]	1680	6720	loguniform	lognormal
HIM	Interface heat transfer coefficient between metallic layer of debris and concrete [$W/(cm^2 \cdot K)$]	0.04	1.0	loguniform	lognormal
HIO	Interface heat transfer coefficient between oxide layer of debris and concrete [$W/(cm^2 \cdot K)$]	0.018	0.45	loguniform	lognormal
ABRK	Containment break area [ft^2]	0.2097	41.94	loguniform	lognormal
FPSM	Fission product core release fraction by meltdown release [-]	Low RSS(d)	High RSS	uniform	normal
FPSV	Fission product core release fraction by varoprization release [-]	Low RSS	High RSS	loguniform	lognormal
DFI	Natural deposition factor for elemental iodine [1/hr]	$\lambda/2$	2λ(e)	loguniform	lognormal
DFP	Natural deposition factor for particulate [1/hr]	$\lambda/5$	5λ	loguniform	lognormal

(a) includes TFUS (melting temperature of fuel plus the temperature equivalent of the heat of fusion) and TMLT (melting temperature of debris).
(b) includes FCOL (minimum fraction of core melted for total core collapse).
(c) includes FZOS1 (fraction of ZrO_2 in the first shell of particles, remaining ZrO_2 in outside layer).
(d) RSS means the data used in WASH-1400.
(e) : calculated within the CORRAL II code based on input data.

Table 4.3 MARCH/CORRAL II input variables (parameters), their ranges and distributions considered in the uncertainty analysis for the loss of decay heat removal accident.

Variable	Description	Range		Distribution	
		Lower bound	Upper bound	Case 1	Case 2
SAREA	Slab heat transfer area [ft^2]	566.7	5100	loguniform	lognormal
HIF	Interface heat transfer coefficient for 2-material slab [$Btu/(hr \cdot ft^2 \cdot F)$]	0.5	4.5	loguniform	lognormal
FDROP(a)	Fractional holdup of melted core before slumping into bottom head starts [-]	0.4	0.9	uniform	normal
MWORNL	Zirconium-water reaction model [-] 1: Cathcart, 0: Baker-Just	1	0	discrete	discrete
TCOOLS	Temperature for drywell cooler failure [F]	150	250	uniform	normal
FPSM	Fission product core release fraction by meltdown release [-]	Low RSS	High RSS	loguniform	lognormal
FPSV	Fission product core release fraction by vaporization release [-]	Low RSS	High RSS	loguniform	lognormal
DCF	Pool decontamination factor [-]	1.2	10^6	loguniform	lognormal
PDIA	Particle diameter [m]	$\lambda/2$	2λ(c)	loguniform	lognormal
DFI2	Natural deposition factor for I_2 [1/hr]	$\lambda/2$	2λ	loguniform	lognormal
DFP	Natural deposition factor for particulate [1/hr]	$\lambda/5$	5λ	loguniform	lognormal
C1	Pressure for containment failure [psia]	140	210	uniform	normal

(a) Includes FCOL (minimum fraction of core melted for total core collapse).
(b) RSS means the data used in WASH-1400.
(c) Calculated within the CORRAL II code based on input data.

Table 4.4 Results of analysis of variance for elemental iodine release fraction to the environment in the station blackout accident.

Variable			Degree of freedom	Sum of squares	Mean squares
DCF			2	7.0226D-03*	3.5113D-03
WDED			2	2.7755D-03	1.3878D-03
TMELT			2	2.5017D-03	1.2508D-03
FDROP			2	1.0634D-02	5.3171D-03
FZOCR			2	6.3982D-04	3.1991D-04
COND			2	1.3066D-03	6.5329D-04
FLRMC			2	9.0058D-04	4.5029D-04
HIM			2	2.6348D-04	1.3174D-04
HIO			2	4.3098D-04	2.1549D-04
FPSM			2	9.4865D-04	4.7432D-04
FPSV			2	6.3483D-04	3.1741D-04
ABRK			2	9.4290D-04	4.7145D-04
DFI			2	4.4618D-03	2.2309D-03
DCF	*	WDED	4	7.8048D-04	1.9512D-04
DCF	*	TMELT	4	7.8402D-04	1.9601D-04
DCF	*	FDROP	4	1.5460D-03	3.8650D-04
DCF	*	COND	4	1.2956D-03	3.2391D-04
DCF	*	DP	4	5.1624D-04	1.2906D-04
DCF	*	FPSV	4	5.7865D-04	1.4466D-04
DCF	*	FPSM	4	1.0398D-03	2.5996D-04
WDED	*	FDROP	4	4.2473D-04	1.0618D-04
WDED	*	COND	4	1.1353D-03	2.8382D-04
WDED	*	DP	4	8.4231D-04	2.1058D-04
WDED	*	FLRMC	4	6.1262D-04	1.5316D-04
WDED	*	FPSV	4	5.9601D-04	1.4900D-04
TMELT	*	FDROP	4	5.8029D-03	1.4507D-03
TMELT	*	COND	4	7.2848D-04	1.8212D-04
TMELT	*	DFI	4	4.5028D-04	1.1257D-04
FDROP	*	DP	4	8.1977D-04	2.0494D-04
FDROP	*	FPSV	4	5.0637D-04	1.2659D-04
FZOCR	*	DFI	4	4.6717D-04	1.1679D-04
COND	*	FPSV	4	5.9265D-04	1.4816D-04
COND	*	FPSM	4	9.5872D-04	2.3968D-04
DP	*	FLRMC	4	7.2141D-04	1.8035D-04
DP	*	DFI	4	6.0843D-04	1.5211D-04
FLRMC	*	FPSM	4	6.5177D-04	1.6294D-04
FPSV	*	DFI	4	5.1281D-04	1.2820D-04
DFI	*	FPSM	4	6.5116D-04	1.6279D-04
ERROR			116	5.8273D-03	5.0235D-05

Cut off value of variance: 1.0×10^{-4}

* 7.0226D-03 reads 7.0226×10^{-3}

Table 4.5 Results of analysis of variance for Cs-Rb release fraction to the environment in the station blackout accident.

Variable			Degree of freedom	Sum of squares	Mean squares
DCF			2	2.0914D-01*	1.0457D-01
WDED			2	1.9427D-02	9.7135D-03
TMELT			2	4.5993D-03	2.2996D-03
FDROP			2	2.4175D-02	1.2087D-02
FZOCR			2	6.0361D-03	3.0181D-03
COND			2	1.2604D-02	6.3021D-03
DP			2	5.9054D-03	2.9527D-03
FLRMC			2	5.9609D-03	2.9804D-03
FPSV			2	6.5581D-03	3.2790D-03
ABRK			2	1.3225D-02	6.6124D-03
DFP			2	3.8099D-01	1.9049D-01
DCF	*	COND	4	1.4201D-02	3.5502D-03
DCF	*	FPSV	4	8.1044D-03	2.0261D-03
DCF	*	DFP	4	1.6437D-01	4.1093D-02
DCF	*	FPSM	4	2.1089D-02	5.2723D-03
WDED	*	COND	4	1.0009D-02	2.5024D-03
TMELT	*	FDROP	4	1.6494D-02	4.1236D-03
TMELT	*	DFP	4	1.2599D-02	3.1498D-03
FZOCR	*	FPSV	4	9.0066D-03	2.2516D-03
COND	*	FPSM	4	1.1480D-02	2.8699D-03
DP	*	FLRMC	4	1.0342D-02	2.5856D-03
ERROR			180	1.5762D-01	8.7566D-04

Cut off value of variance: 2.0×10^{-3}

* 2.0914D-01 reads 2.0914×10^{-1}

Table 4.6 Results of analysis of variance for elemental iodine release fraction to the environment in the loss of decay heat removal accident.

Variable	Degree of freedom	Sum of squares	Mean squares
MWORNL	2	9.8030D-03*	4.9015D-03
FDROP	2	5.1129D-03	2.5565D-03
DFI2	2	3.0740D-03	1.5370D-03
TCOOLS	2	2.0444D-03	1.0222D-03
SAREA	2	1.7260D-03	8.6301D-04
PDIA	2	1.5636D-03	7.8182D-04
SAREA * C1	4	2.7212D-03	6.8031D-04
C1	2	1.2302D-03	6.1510D-04
HIF	2	9.5716D-04	4.7858D-04
FPSM * DCF	4	1.8377D-03	4.5942D-04
FPSM	2	7.2343D-04	3.6171D-04
DCF	2	5.2073D-04	2.6037D-04
MWORNL * DCF	4	1.0406D-04	2.6014D-04
DCF * DFI2	4	9.4114D-04	2.3529D-04
MWORNL * DFI2	4	7.4512D-04	1.8628D-04
FDROP * DFI2	4	6.2548D-04	1.5637D-04
FDROP * DCF	4	5.3906D-04	1.3476D-04
FDROP * MWORNL	4	2.3913D-04	5.9783D-05
HIF * C1	4	8.2724D-05	2.0681D-05
ERROR	24	6.0293D-03	2.5122D-04
TOTAL	80	4.1557D-02	

* 9.8030D-03 reads 9.8030 x 10^{-3}

Table 4.7 Results of analysis of variance for Cs-Rb release fraction to the environment in the loss of decay heat removal accident.

Variable	Degree of freedom	Sum of squares	Mean squares
PDIA	2	1.2600D-01*	6.3001D-02
DFP	2	6.5246D-02	3.2623D-02
MWORNL	2	2.7200D-02	1.3600D-02
FDROP	2	2.2685D-02	1.1343D-02
SAREA	2	8.8959D-03	4.4480D-03
DCF	2	8.7633D-03	4.3817D-03
FPSM	2	8.4184D-03	4.2092D-03
FPSM * DCF	4	1.5381D-02	3.8452D-03
TCOOLS	2	6.0860D-03	3.0430D-03
SAREA * C1	4	9.8593D-03	2.4648D-03
DCF * DFP	4	8.2093D-03	2.0523D-03
HIF	2	2.3716D-03	1.1858D-03
FDROP * DCF	4	3.3419D-03	8.3548D-04
MWORNL * DFP	4	2.2570D-03	5.6424D-04
FDROP * DFP	4	2.2433D-03	5.6083D-04
MWORNL * DCF	4	1.8604D-03	4.6509D-04
C1	2	7.4097D-04	3.7049D-04
HIF * C1	4	1.4768D-03	3.6919D-04
FDROP * MWORNL	4	1.0760D-03	2.6901D-04
ERROR	24	5.0385D-02	2.0994D-03
TOTAL	80	3.7250D-01	

* 1.2600D-01 reads 1.2600×10^{-1}

Table 4.8 TERFOC input variables, their ranges and distributions considered in the importance analysis.

Variable	Description	Range			Distribution
		Lower bound		Upper bound	
ω_E	washout ratio for elemental iodine [-]	$2x10^4$		$2x10^6$	lognormal
ω_P	washout ratio for particulate iodine [-]	$3x10^4$		$3x10^6$	lognormal
V_{gE}	deposition velocity for elemental iodine [cm/s]	0.02		26	lognormal
V_{gP}	deposition velocity for particulate iodine [cm/s]	0.03		3	lognormal
RP	Mass interception factor for pasture forage (r/Y_V) $[m^2/kg(wet)]$	0.47		8.5	lognormal
λ_W	weathering removal rate constant [1/d]	$5.9x10^{-3}$		0.42	lognormal
F_m	feed to milk transfer factor [d/ℓ]	$1.9x10^{-3}$		$5.2x10^{-2}$	lognormal
Q_{FF}	daily dry intake of fresh forage by dairy cows [kg/d]	3	4^*	5	triangular
Q_{FS}	daily dry intake of stored forage by dairy cows [kg/d]	9	13^*	18	triangular
t_e	time period that vegetation is exposed to contamination during the growing season [d]	15	30^*	60	triangular
t_h	time delay between harvest and consumption of stored forage [d]	15	90^*	180	triangular
t_f	time delay from production to consumption of milk [d]	0	2^*	7	triangular

* denotes the most probable value in the triangular distribution.

Table 4.9 PCCs, PRCCs, SRCs, SRRCs and R^2 values for the TERFOC output variable C_m.

Variable	PCC	PRCC	SRC	SRRC
ω_E	0.25	0.32	0.11	0.10
ω_P	0.64	0.82	0.36	0.43
V_{gE}	0.36	0.51	0.17	0.18
V_{gP}	0.30	0.55	0.14	0.19
RP	0.71	0.86	0.44	0.49
λ_W	-0.52	-0.70	-0.27	-0.29
F_m	0.80	0.88	0.59	0.54
Q_{FF}	0.27	0.41	0.12	0.13
Q_{FS}	0.06	0.02	0.02	0.01
t_e	0.07	0.10	0.03	0.03
t_h	-0.02	-0.07	-0.01	-0.02
t_f	-0.16	-0.40	-0.07	-0.13
R^2	0.81	0.91	0.81	0.91

Table 4.10 Values of quantity U_j with the uncertainty reduction method (URM) and rankings of input variables for the TERFOC output variable C_m.

Variable	U_j	Ranking				
		URM	PCC	PRCC	SRC	SRRC
ω_E	2.00×10^5	5	8	9	8	9
ω_P	2.38×10^5	2	3	3	3	3
V_{gE}	2.04×10^5	4	5	6	5	6
V_{gP}	1.97×10^5	6	6	5	6	5
RP	2.29×10^5	3	2	2	2	2
λ_W	1.90×10^5	10	4	4	4	4
F_m	2.56×10^5	1	1	1	1	1
Q_{FF}	1.91×10^5	9	7	7	7	7
Q_{FS}	1.88×10^5	12	11	12	11	12
t_e	1.88×10^5	11	10	10	10	10
t_h	1.93×10^5	8	12	11	12	11
t_f	1.93×10^5	7	9	8	9	8

Table 4.11 SPARC input variables, their ranges and distributions considered in the reference uncertainty analysis.

Variable	Description	Range		Distribution
		Lower bound	Upper bound	
X_1: RATIO	Bubble aspect ratio [-]	1	4	Uniform
X_2: DIAM	Mean bubble diameter [mm]	3	20	Uniform
X_3: VSWARM	Bubble swarm rise velocity [cm/s]	20	120	Uniform
X_4: VIMPT	Inlet impact velocity [cm/s]	0	30,000	Uniform
X_5: NRISE	Number of time steps for the calculation of decontamination factors during bubble rise [-]	100	1,000	Uniform
X_6: CDIF	A constant imbedded in the diffusional removal model [-]	1	4	Uniform

Table 4.12 PRCCs, SRRCs and R^2 values for the SPARC output variables, DF and L.

	Integral DF				Total Leakage (L)			
	LHS-1		LHS-2		LHS-1		LHS-2	
Variable	PRCC	SRRC	PRCC	SRRC	PRCC	SRRC	PRCC	SRRC
X_1	0.95	0.77	0.94 (-1.1%)	0.80 (3.9%)	-0.95	-0.77	-0.95 (0%)	-0.80 (3.9%)
X_2	-0.76	-0.29	-0.72 (-5.3%)	-0.29 (0%)	0.76	0.29	0.75 (1.3%)	0.30 (3.4%)
X_3	-0.90	-0.51	-0.85 (-5.6%)	-0.46 (-9.8%)	0.90	0.52	0.86 (-4.4%)	0.46 (-12%)
X_4	0.31	0.08	0.25 (-19%)	0.07 (-13%)	-0.32	-0.08	-0.27 (-16%)	-0.07 (-13%)
X_5	0.35	0.09	0.00 (-100%)	0.00 (-100%)	-0.35	-0.09	0.00 (-100%)	0.00 (-100%)
X_6	0.19	0.05	0.15 (-21%)	0.04 (-20%)	-0.18	-0.05	-0.17 (-5.6%)	-0.05 (0%)
R^2	0.94	0.94	0.92	0.92	0.94	0.94	0.93	0.93

NOTE: Percentage departure from LHS-1 results is given in parentheses.

Table 4.13 Summary of the reference analysis results for the SPARC output variables, DF and L.

	Integral DF		Total Leakage (L)	
Statistical parameter	LHS-1	LHS-2	LHS-1	LHS-2
Mean	240.3	243.7 (1.4%)	2503	25067 (0.1%)
Standard deviation	129.9	138.8 (6.9%)	904.9	907.1 (0.2%)
5th	131.7	128.6 (-2.4%)	879.7	923.2 (4.9%)
50th	199.7	196.5 (-1.6%)	2488	2535 (1.9%)
95th	574.6	547.8 (-4.7%)	3766	3481 (2.0%)

NOTE: Percentage departure from LHS-1 results is given in parentheses.

Table 4.14 Assumed probability distributions of SPARC input variables in the output distribution sensitivity analysis.

Variable	Reference analysis	Sensitivity analysis S-1	Sensitivity analysis S-2
X_1	Uniform	Normal	Uniform
X_3	Uniform	Uniform	Normal
All others	Uniform	Uniform	Uniform
Sample	LHS-1	LHS-3	LHS-4

Table 4.15 SRRCs and R^2 values for the SPARC output variables, DF and L.

Variable	Integral DF $\hat{a}_i{}^*$	Total leakage (L) $\hat{a}_i{}^*$
X_1	0.776	-0.775
X_2	-0.304	0.294
X_3	-0.502	0.512
X_4	--	-0.079
X_5	0.091	-0.092
X_6	--	--
R^2	0.92	0.95

Table 4.16 Weighting factor W_j in the weighting method.

j	1	2	3	4	5
W_j	5 (-4)*	7.3 (-4)	1.04 (-3)	1.45 (-3)	1.99 (-3)

j	6	7	8	9	10
W_j	2.71 (-3)	3.60 (-3)	4.77 (-3)	6.19 (-3)	7.84 (-3)

j	11	12	13	14	15
W_j	9.93 (-3)	1.23 (-2)	1.49 (-2)	1.80 (-2)	2.13 (-2)

j	16	17	18	19	20
W_j	2.14 (-2)	3.16 (-2)	3.21 (-2)	3.57 (-2)	3.91 (-2)

j	21	22	23	24	25
W_j	4.22 (-2)	4.49 (-2)	4.70 (-2)	4.85 (-2)	4.92 (-2)

* 5 (-4) $\equiv 5 \times 10^{-4}$

NOTE: Owing to the symmetric property of the normal distribution, $W_{25+j} = W_{26-j}$. $(j=1,2,\ldots,25)$

Table 4.17 Statistical parameters for the SPARC output variable DF when the PDF for X_1 (RATIO) is changed from uniform to normal distribution.

Statistical parameter	SPARC	Classical rank regression	Modified rank regression	Weighting method	Rejection method	Method of closest distance
Mean	227.7	218.0 (-4.3%)	221.3 (-2.8%)	246.2 (8.1%)	210.4 (-7.6%)	242.1 (6.3%)
Standard deviation	91.3	85.9 (-5.9%)	91.5 (0.2%)	143.1 (57%)	65.6 (- 28%)	130.0 (42%)
5th	138.3	141.0 (2.0%)	135.3 (-2.2%)	116.5 (-16%)	126.2 (-8.7%)	131.7 (-4.8%)
50th	186.1	199.5 (7.2%)	201.9 (8.5%)	206.2 (9.7%)	199.8 (7.4%)	199.9 (7.4%)
95th	434.2	353.0 (-19%)	402.2 (-7.4%)	571.4 (32%)	349.6 (-19%)	574.6 (32%)

NOTE: Percentage departure from SPARC results is in parentheses.

Table 4.18 Statistical parameters for the SPARC output variable DF when the PDF for X_3 (VSWARM) is changed from uniform to normal distribution.

Statistical parameter	SPARC	Classical rank regression	Modified rank regression	Weighting method	Rejection method	Method of closest distance
Mean	219.0	228.6 (4.4%)	221.3 (1.1%)	205.4 (-6.2%)	222.2 (1.5%)	227.0 (3.7%)
Standard deviation	91.1	120.2 (32%)	91.5 (0.4%)	45.8 (-50%)	125.1 (37%)	113.6 (25%)
5th	128.8	135.8 (5.4%)	135.3 (5.0%)	147.4 (14%)	118.0 (-8.4%)	131.7 (2.3%)
50th	202.3	201.1 (-0.6%)	201.9 (-0.2%)	199.7 (-1.3%)	190.7 (-5.7%)	199.7 (-1.3%)
95th	396.7	552.5 (39%)	402.2 (1.4%)	287.2 (-28%)	577.7 (46%)	379.6 (-4.3%)

NOTE: Percentage departure from SPARC results is in parentheses.

Table 4.19 Statistical parameters for SPARC output variable L when the PDF for X_1 (RATIO) is changed from uniform to normal distribution.

Statistical parameter	SPARC	Classical rank regression	Modified rank regression	Weighting method	Rejection method	Method of closest distance
Mean	2470	2507 (1.5%)	2533 (2.6%)	2532 (2.5%)	2579 (4.4%)	2497 (1.1%)
Standard deviation	749.3	663.2 (-11%)	790.5 (5.4%)	927.6 (24%)	731.5 (-2.4%)	906.7 (21%)
5th	1163	1553 (34%)	1304 (12%)	693.1 (-40%)	1451 (25%)	880.0 (-24%)
50th	2641	1493 (-5.6%)	2521 (-4.5%)	2499 (-5.4%)	2492 (-5.6%)	2487 (-5.8%)
95th	3564	3552 (0.3%)	3567 (0.1%)	3764 (5.6%)	3945 (11%)	3766 (5.7%)

NOTE: Percentage departure from SPARC results is in parentheses.

Table 4.20 Statistical parameters for the SPARC output variable L when the PDF for X_3 (VSWARM) is changed from uniform to normal distribution.

Statistical parameter	SPARC	Classical rank regression	Modified rank regression	Weighting method	Rejection method	Method of closest distance
Mean	2589	2511 (-3.0%)	2559 (-1.2%)	2504 (-3.3%)	2710 (4.7%)	2564 (1.0%)
Standard deviation	831.4	784.9 (-5.6%)	825.6 (-0.7%)	452.3 (-46%)	941.5 (13%)	867.3 (4.3%)
5th	1270	963.2 (-24%)	1265 (0.4%)	1728 (36%)	875.8 (-31%)	1327 (4.5%)
50th	2446	2491 (1.8%)	2456 (-1.4%)	2487 (1.7%)	2615 (6.9%)	2487 (1.7%)
95th	3849	3648 (-5.2%)	3648 (-5.2%)	3325 (-14%)	4203 (9.2%)	3766 (-2.2%)

NOTE: Percentage departure from SPARC results is in parentheses.

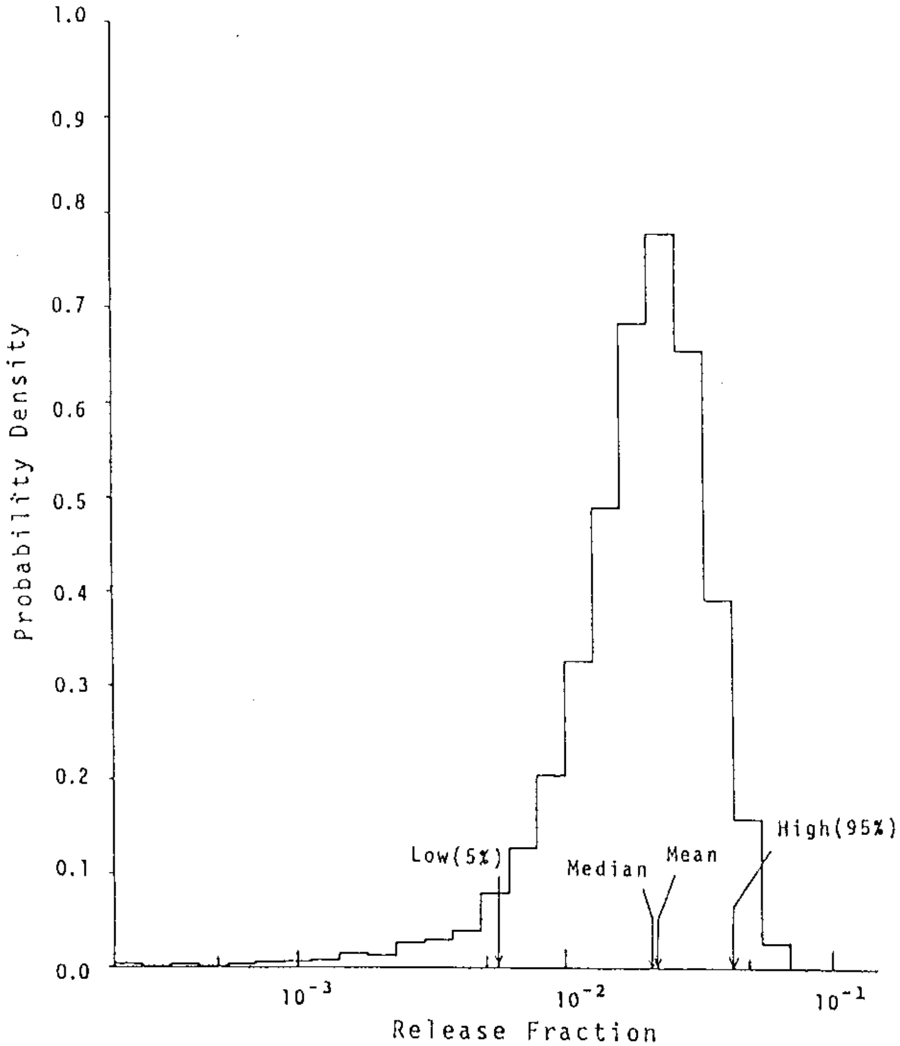

Fig. 4.1 Probability distribution of elemental iodine release fraction to the environment in the station blackout accident (Case 1).

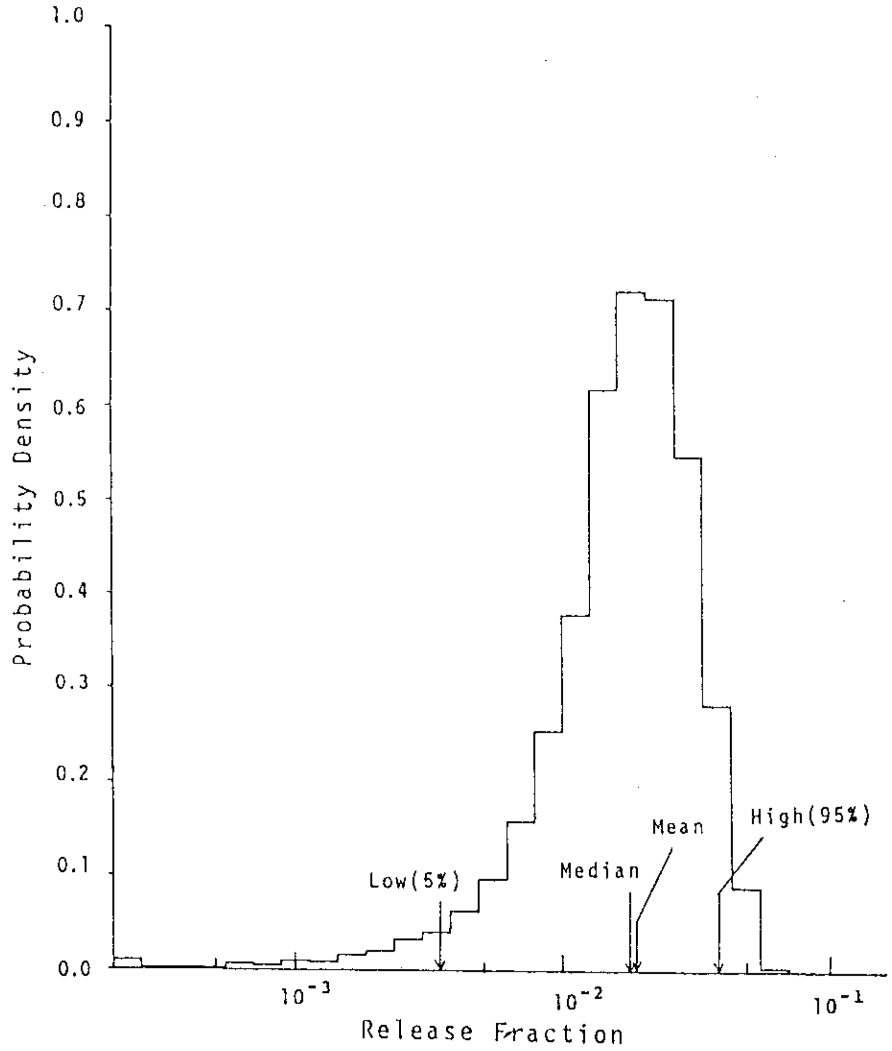

Fig. 4.2 Probability distribution of elemental iodine release fraction to the environment in the station blackout accident (Case 2).

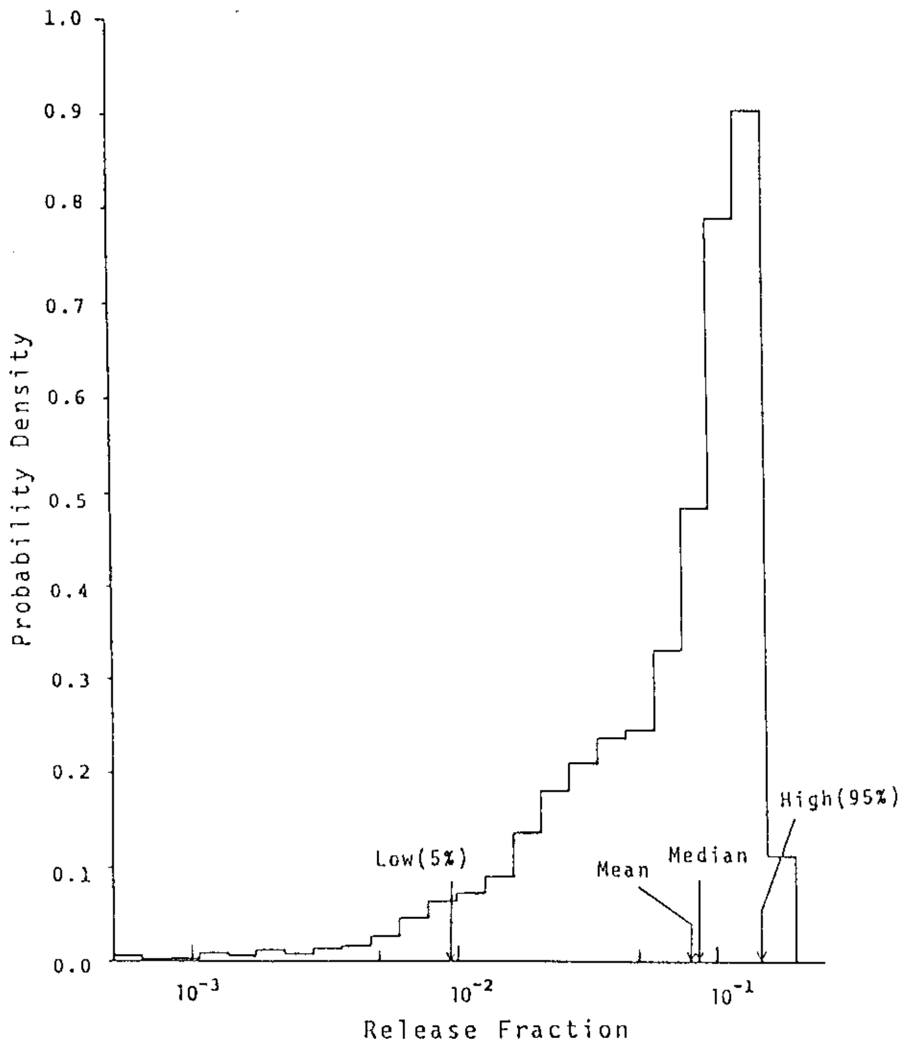

Fig. 4.3 Probability distribution of Cs-Rb release fraction to the environment in the station blackout accident (Case 1).

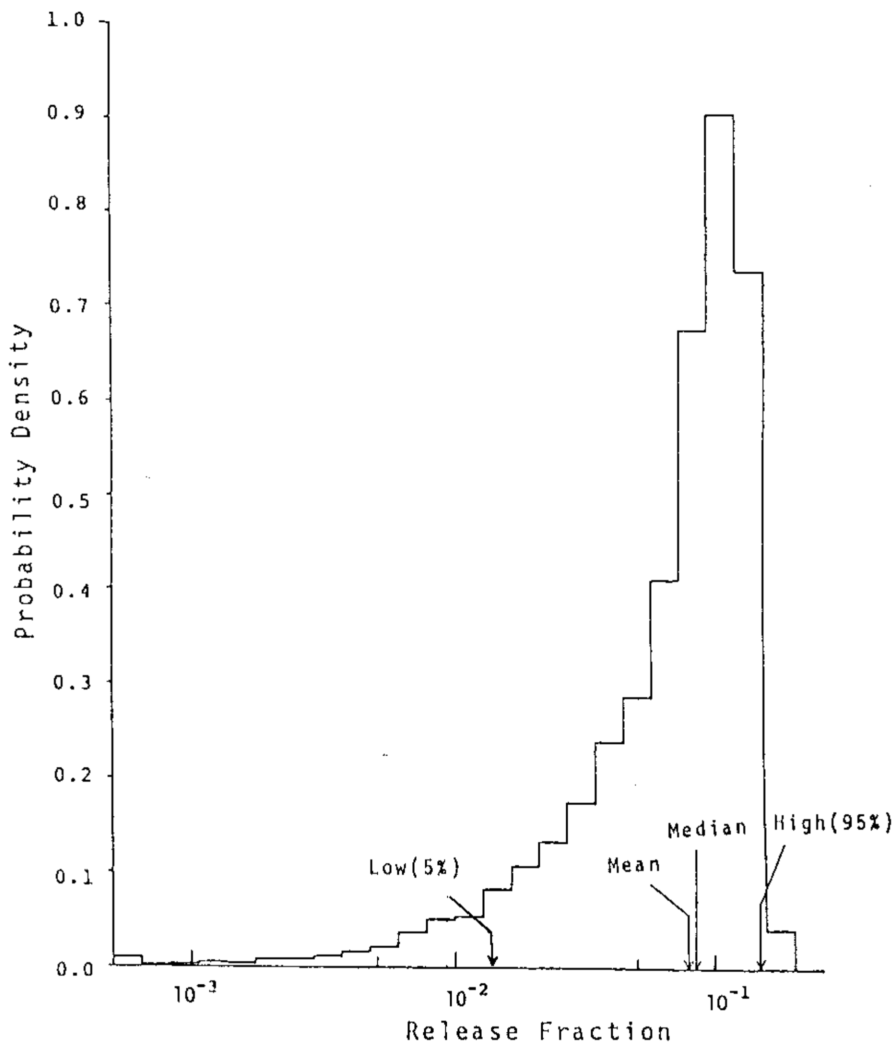

Fig. 4.4 Probability distribution of Cs-Rb release fraction to the environment in the station blackout accident (Case 2).

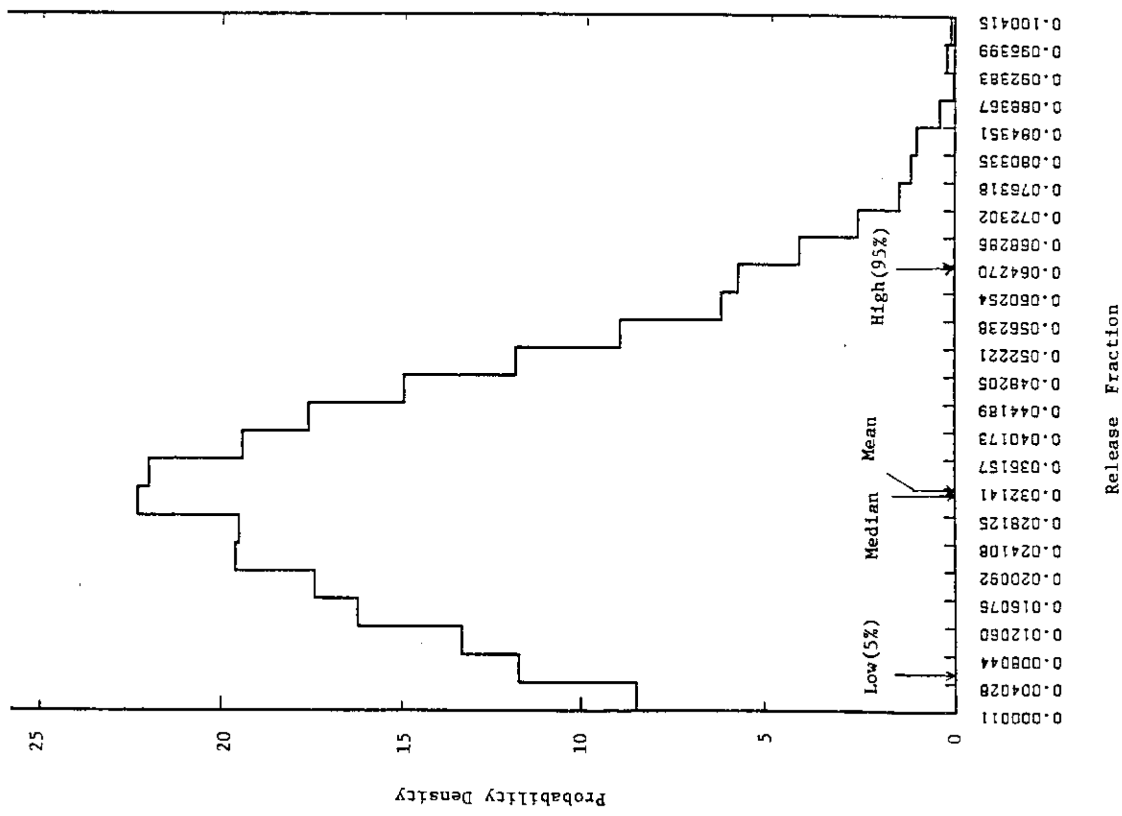

Fig. 4.5 Probability distribution of elemental iodine release fraction to the environment in the loss of decay heat removal accident (Case 1).

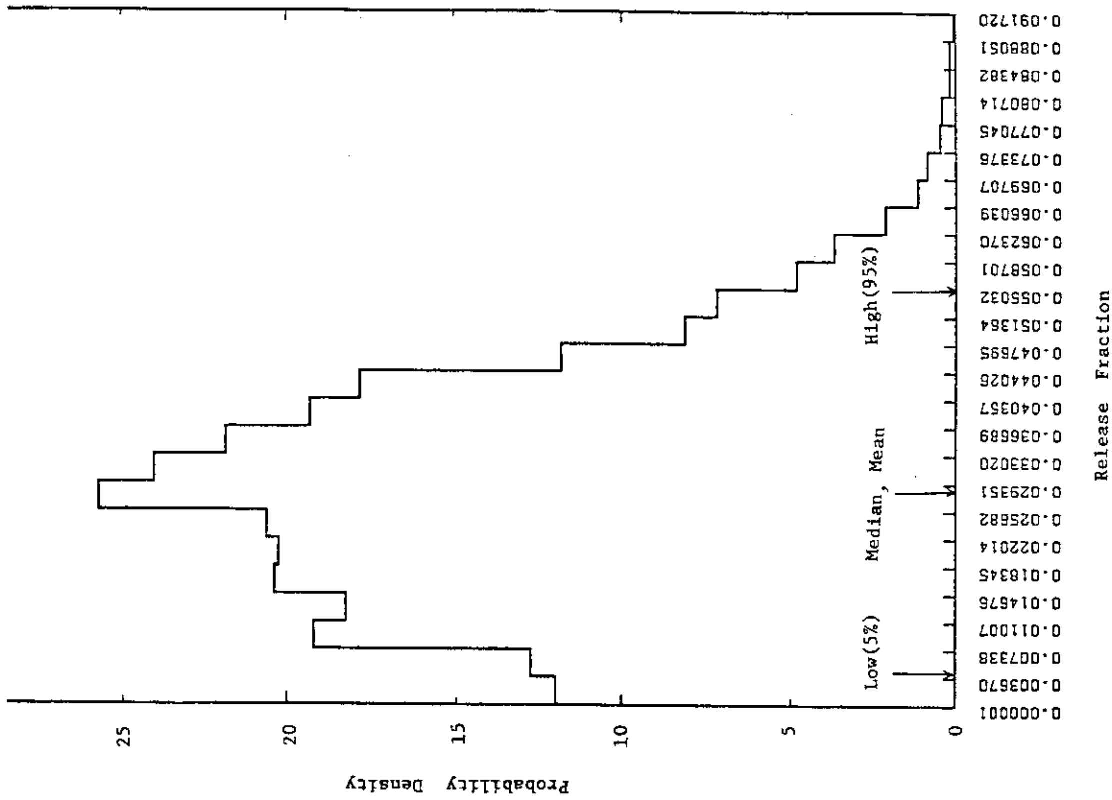

Fig. 4.6 Probability distribution of elemental iodine release fraction to the environment in the loss of decay heat removal accident (Case 2).

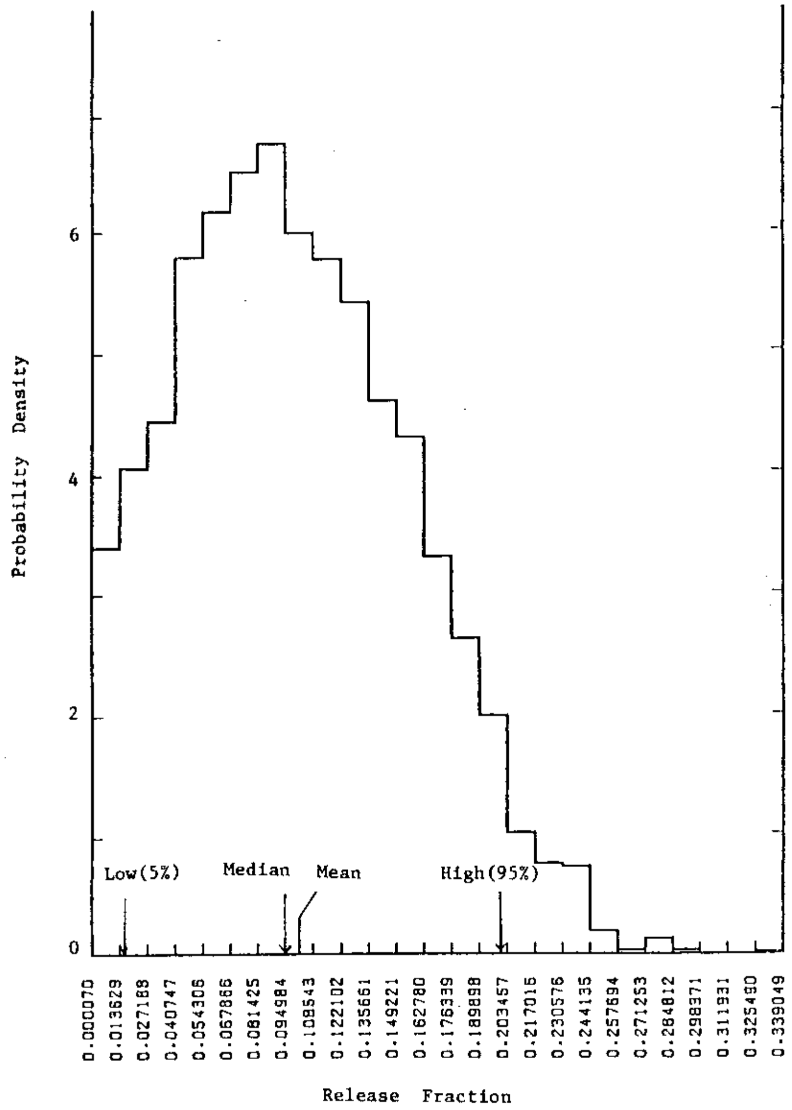

Fig. 4.7 Probability distribution of Cs-Rb release fraction to the environment in the loss of decay heat removal accident (Case 1).

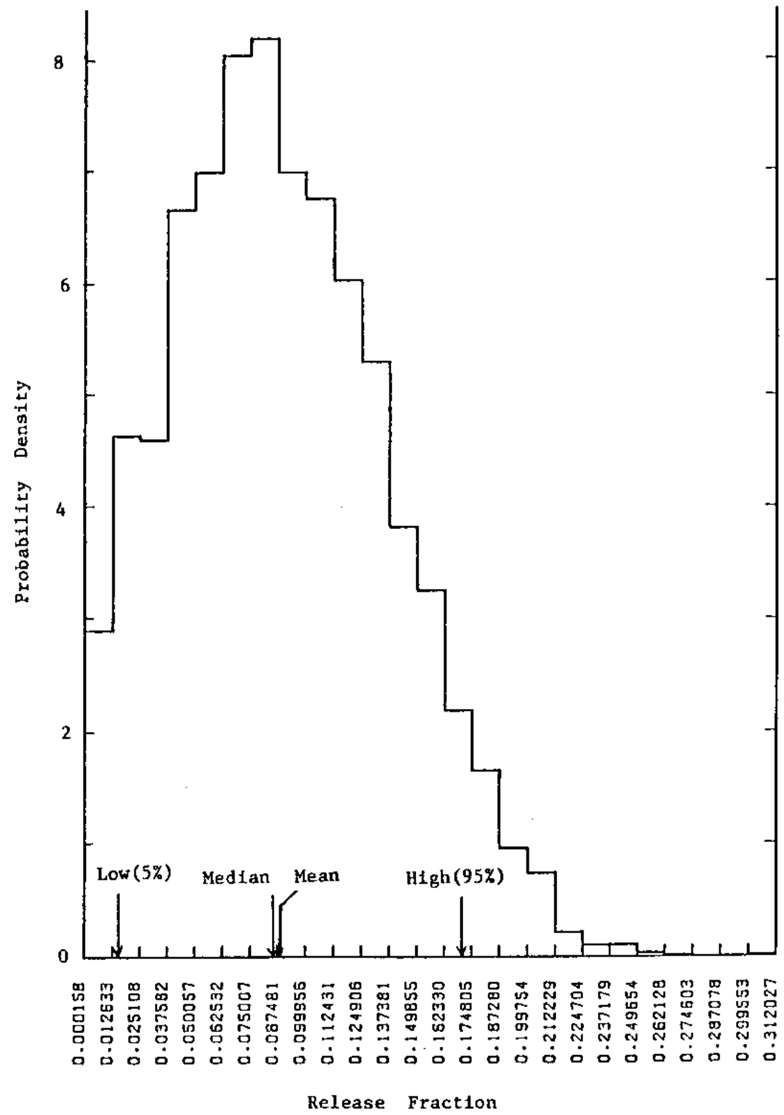

Fig. 4.8 Probability distribution of Cs-Rb release fraction to the environment in the loss of decay heat removal accident (Case 2).

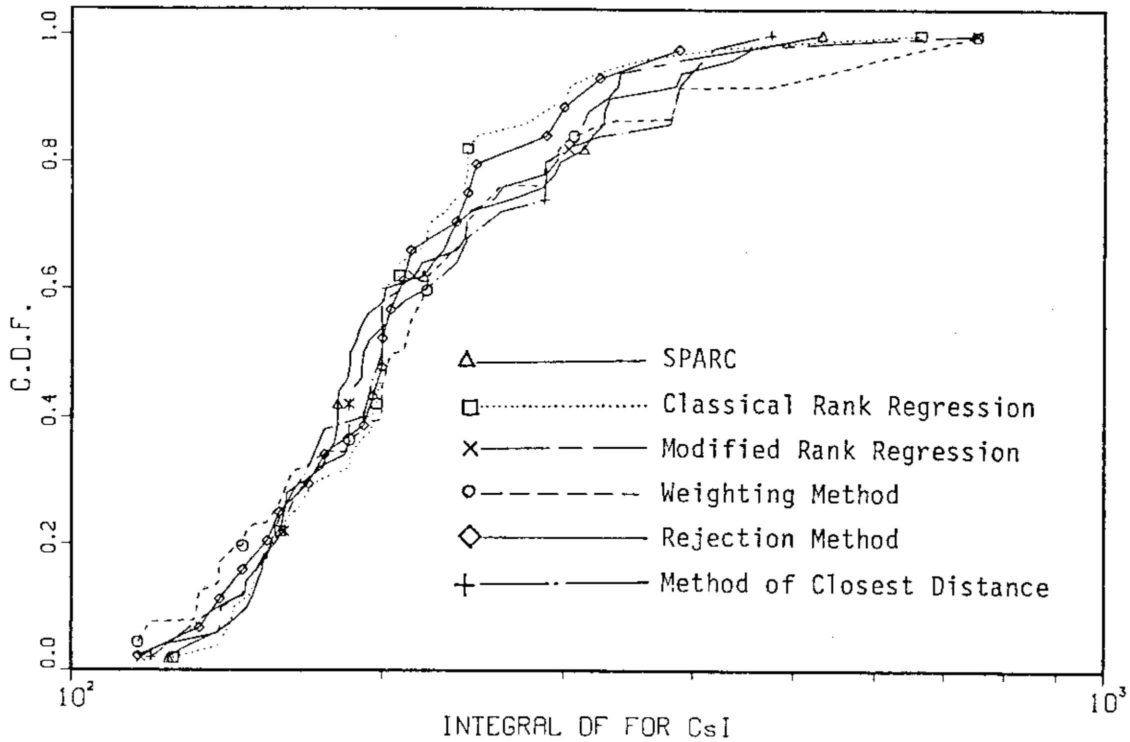

Fig. 4.9 Comparison of cumulative distribution functions (CDFs) for the SPARC output variable DF when the PDF for X_1 (RATIO) is changed from uniform to normal distribution.

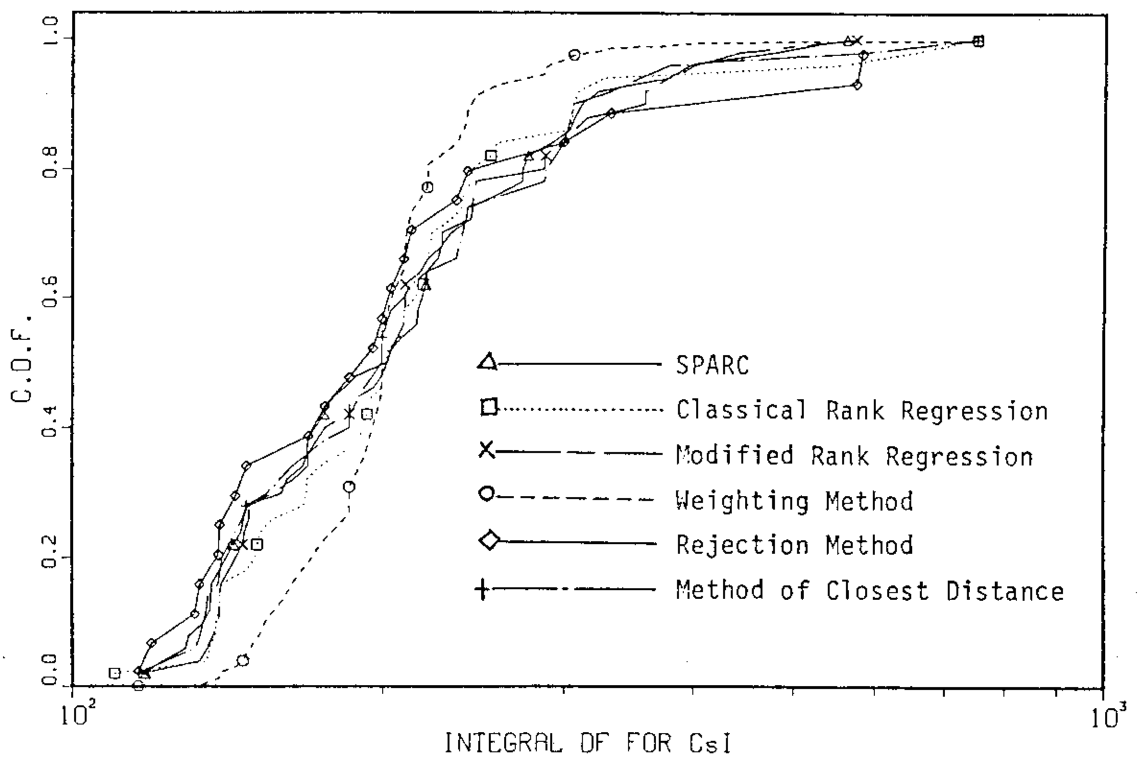

Fig. 4.10 Comparison of cumulative distribution functions (CDFs) for the SPARC output variable DF when the PDF for X_3 (VSWARM) is changed from uniform to normal distribution.

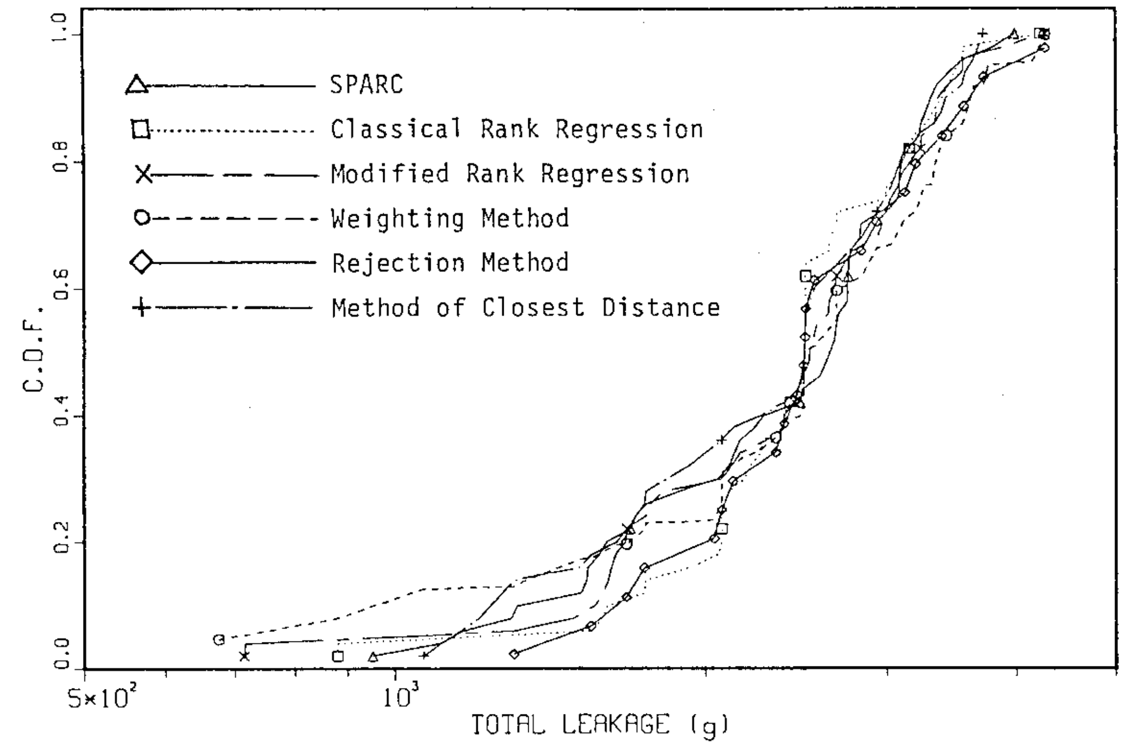

Fig. 4.11 Comparison of cumulative distribution functions (CDFs) for the SPARC output variable L when the PDF for X_1 (RATIO) is changed from uniform to normal distribution.

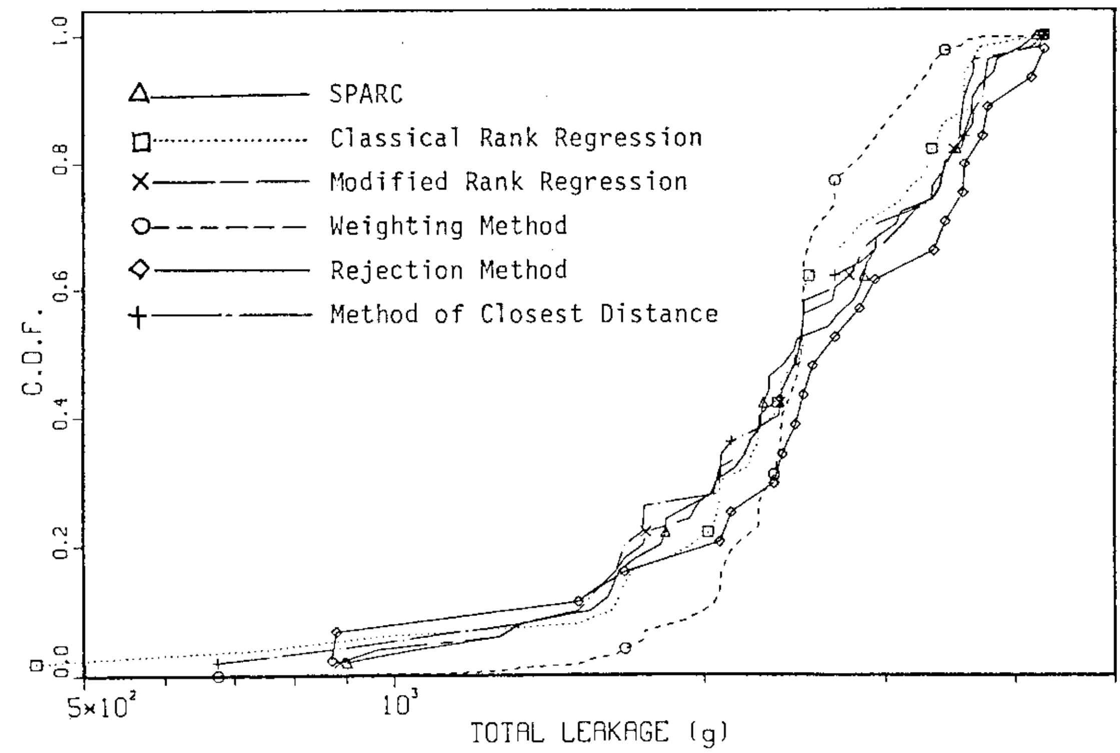

Fig. 4.12 Comparison of cumulative distribution functions (CDFs) for the SPARC output variable L when the PDF for X_3 (VSWARM) is changed from uniform to normal distribution.

5. Summary

The objectives of the uncertainty analysis for deterministic computer models are to quantify the output uncertainty and to partition the uncertainty among the contributing variables. The uncertainty analysis considered here consists of four analysis steps of (1) screening analysis, (2) uncertainty propagation analysis, (3) uncertainty importance analysis and (4) output distribution sensitivity analysis. Several existing methods for the analysis steps (2) through (4) have been surveyed. The approach to the uncertainty analysis includes the response surface method and the direct method.

The response surface methods described are (a) the classical regression method and (b) the modified regression method, and the direct methods described are (c) the Monte Carlo method, (d) the LHS method, (e) the uncertainty reduction method, (f) the method of closest distance, (g) the weighting method, and (h) the rejection method.

Application of these methods to three computer models, MARCH/CORRAL II, TERFOC and SPARC, have been illustrated. In application to the MARCH/CORRAL II model, the regression method was used to estimate uncertainties in the fission product release to the environment during core meltdown accidents at the BWR Mark-I plant. Investigated are identification of important input variables and uncertainty propagation. Calculated results showed that the difference between upper and lower bounds of the amounts of fission products (elemental iodine and Cs-Rb) was about one order of magnitude.

In application to the TERFOC model, the uncertainty importance analysis was made using the regression method and the uncertainty reduction method, where the sampling technique of the LHS was utilized. The results showed that although both the method identified the same input variables contributing to the output uncertainty, there could be seen some differences between the results by the two methods.

In application to the SPARC model, several output sensitivity distribution methods were assessed including the regression method, the modified regression method, the method of closest distance, the weighting method, and the rejection method. The analysis showed that although the modified rank regression method performed the best of the five methods, remarkable differences among the results stemming from the five methods was not marked. Thus, judicious application of each of the five methods, weighed by a knowledge of the goodness of fit to the regression model formulated, will provide a basis for the output distribution sensitivity analysis.

Finally recommendation is provided for selecting several uncertainty analysis methods in each analysis step of uncertainty propagation, importance rankings of input variables, and output distribution sensitivity.

When the computer code of concern is such a fast-running code to make a large number of repeated calculations with the computer code possible, the Monte Carlo method is maximally used for the uncertainty propagation and the output distribution sensitivity analyses. As for the importance analysis, the uncertainty reduction method can be used by combining with the Monte Carlo sampling.

When the computer code of concern is a long-running code, which is often the case, utilization of the methods to be recommended is described below:

(1) In the uncertainty propagation analysis step, the LHS method would be useful to provide the output distribution. Then the computer model input/output relationships obtained based on the LHS are used as a data base in the following regression analysis to determine P(R)CCs and the S(R)RCs.

(2) When the coefficient of determination, R^2, in the regression analysis is close to one, say $R^2 > 0.9$, the response surface obtained may be satisfactory. Therefore P(R)CCs and S(R)RCs would give importance brankings of input variables, and the output distribution sensitivity analysis can be performed with respect to the response surface using a Monte Carlo sampling.

(3) However, if the regression analysis gives poor results, namely the R^2 value obtained is largely apart from one, the direct method will be applicable in the following importance and output distribution analyses. In the importance analysis, utilization of the uncertainty reduction method together with a sampling technique of the LHS is recommended. As for the output distribution sensitivity analysis, it is not possible to make general recommendation among several direct methods. Judicious application of each of the direct methods will provide a basis for the output distribution sensitivity analysis.

The above is summarized in Table 5.1 which shows the methods to be recommended in utilizing several methods in each uncertainty analysis step.

Table 5.1 Methods to be recommended in each uncertainty analysis step.

Computer code	Uncertainty propagation analysis	Importance analysis	Output distribution sensitivity analysis
Short-running code	OM + MCM	URM: OM + MCM	OM + MCM
Long-running code	OM + LHS method ↓ Regression analysis		
For $R^2>0.9$:	RS + MCM	Regression method	RS + MCM
For $R^2<0.9$:	OM + LHS method	URM: OM + LHS method	Direct methods

OM : Original computer model
MCM: Monte Carlo method
LHS: Latin hypercube sampling
RS : Response surface
URM: Uncertainty reduction method
R^2 : Coefficient of determination in a regression analysis

Acknowledgment

The author is grateful to K. Namatame and K. Kobayashi of Japan Atomic Energy Research Institute for reviewing and making many helpful remarks on the manuscript.

References

1. U.S. Nuclear Regulatory Commission, "Reactor Safety Study, an Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants," WASH-1400 (NUREG-75/014) (1975).

2. U.S. Nuclear Regulatory Commission, "Reactor Risk Reference Document," NUREG-1150 (Draft) (1987).

3. D.M. Wuschke, P.A. Gillespie, K.K. Mehta, W.F. Heinrich, D.M. Leneveu, V.M. Guvanasen, G.R. Sherman, D.C. Donahue, B.W. Goodin, T.H. Andres and R.B. Lyon, "Second Interim Assessment of the Canadian Concept for Nuclear Fuel Waste Disposal-Volume 4; Post Closure Assessment," Atomic Energy of Canada Ltd Report AECL-8373-4 (1985).

4. D.C. Cox and P. Baybutt, "Methods for Uncertainty Analysis: A Comparative Survey," Risk Analysis, **1**, 251 (1981).

5. C. Park and M. Khatib-Rahbar, "Quantification and Uncertainty Analysis of Source Terms for Severe Accidents in Light Water Reactors (QUASAR), Part 1 - Methodology and Program Plan," NUREG/CR-4688, BNL/NUREG-52008, Volume 1 (1986).

6. R.H. Myers, "Response Surface Methodology," Allyn and Bacon, Boston, MA (1975).

7. M.D. McKay, W.J. Conover and R.J. Beckman, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, **21**, 239 (1979).

8. R.L. Iman and M.J. Shortencarier, "A FORTRAN 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use with Computer Models," NUREG/CR-3624, SAND83-2365 (1984).

9. G.E.P. Box and J.S. Hunter, "The 2^{k-p} Fractional Factorial Designs, Part I," Technometrics, **3**, 311 (1961).

10. R.L. Iman and W.J. Conover, Sensitivity Analysis Techniques: Self-Teaching Curriculum," NUREG/CR-2350, SAND81-1978 (1982).

11. R.L. Iman and M.J. Shortencarier, "Stepwise Regression with PRESS and Rank Regression (Program User's Guide)," SAND79-1472 (1980).

12. R.L. Iman M.J. Shortencarier and J.D. Johnson, "A FORTRAN 77 Program and User's Guide for Calculation of Partial Correlation and Standardized Regression Coefficients," NUREG/CR-4122, SAND85-0044 (1985).

13. T. Ishigami, E. Cazzoli, M. Khatib-Rahbar and S.D. Unwin, "Techniques to Quantify the Sensitivity of Deterministic Model Uncertainties," Nucl. Sci. Eng., **101**, 371 (1989).

14. T. Ishigami, E. Cazzoli, M. Khatib-Rahbar and S.D. Unwin, "Quantification and Uncertainty Analysis of Source Terms for Severe Accidents in Light Water Reactors (QUASAR), Part II - Sensitivity Analysis Techniques," NUREG/CR-4688, BNL-NUREG-52008, Volume 2 (1987).

15. S.C. Hora and R.L. Iman, "A Comparison of Maximus/Bounding and Bayes/Monte Carlo for Fault Tree Uncertainty Analysis," SAND85-2839 (1986).

16. T. Ishigami and T. Homma, "An Importance Quantification Technique in Uncertainty Analysis for Computer Models," JAERI-M 89-111 (1989).

17. T. Ishigami, E. Cazzoli, M. Khatib-Rahbar and S.D. Unwin, "A Sensitivity Analysis Technique for Application to Deterministic Models," Trans. Am. Nucl. Soc., **55**, 313 (1987)

18. R.L. Iman, "Risk Methodology for Geologic Disposal of Radioactive Waste: Small Sample Sensitivity Analysis Techniques for Computer Models, With an Application to Risk Assessment," NUREG/CR-1397, SAND80-0020 (1980).

19. R.J. Beckman and M.D. McKay, "Monte Carlo Estimation under Different Distributions Using the Same Simulation," Technometrics, **29**, 153 (1987).

20. R.O. Wooton and H.I. Avci, "MARCH (Meltdown Accident Response CHaracteristics) Code Description and User's Manual," NUREG/CR-1711 (1980).

21. R.J. Burian and P. Cybulskis, "CORRAL II Users Manual," Battelle Columbus Laboratories (1977).

22. O. Togawa, "A Computer Code TERFOC-N to Calculate Dose to the Public due to Atmospheric Releases to Radionuclides in Normal Operations of Nuclear Facilities," to be published.

23. P.C. Owczarski et al., "Technical Bases and User's Manual for Prototype of a Suppression Pool AerosolRemoval Code (SPARC)," NUREG/CR-3317, PNL-4742 (1985).

24. K. Kobayashi, T. Ishigami, H. Asaka and M. Akimoto, "Current Status of BWR Severe Accident Analysis," Nihon-Genshiryoku-Gakkai Shi (J. At. Energy Soc. Japan), **27**, 1093 (1985) [in Japanese].

25. T. Ishigami, K. Kobayashi, H. Horii, T. Miyabe and H. Take, "User's Manual of VARS: A Computer Code System for Uncertainty Analysis," JAERI-M 88-082 (1988) [in Japanese].