NUREG/CR-4688
BNL-NUREG-52008
VOLUME 2

QUANTIFICATION AND UNCERTAINTY ANALYSIS OF SOURCE TERMS FOR SEVERE ACCIDENTS IN LIGHT WATER REACTORS (QUASAR)

PART II — SENSITIVITY ANALYSIS TECHNIQUES

Mohsen Khatib-Rahbar
Principal Investigator

October 1987

ACCIDENT ANALYSIS GROUP
DEPARTMENT OF NUCLEAR ENERGY, BROOKHAVEN NATIONAL LABORATORY
UPTON, LONG ISLAND, NEW YORK 11973

Prepared for Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
Under Contract No. DE-AC02-76CH00016

8802240354 871031
PDR NUREG
CR-4688 R PDR

NUREG/CR-4688
BNL-NUREG-52008
VOLUME 2
AN — R-4

QUANTIFICATION AND UNCERTAINTY ANALYSIS OF SOURCE TERMS FOR SEVERE ACCIDENTS IN LIGHT WATER REACTORS (QUASAR)

PART II — SENSITIVITY ANALYSIS TECHNIQUES

Mohsen Khatib-Rahbar
Principal Investigator

Authors
T. Ishigami*, E. Cazzoli, M. Khatib-Rahbar, and S.D. Unwin

* Visiting Scientist from Japan Atomic Energy Research Institute, Japan

Date Completed — September 1987
Date Published — October 1987

ACCIDENT ANALYSIS GROUP
DEPARTMENT OF NUCLEAR ENERGY
BROOKHAVEN NATIONAL LABORATORY
UPTON, NEW YORK 11973

UNITED STATES NUCLEAR REGULATORY COMMISSION
WASHINGTON, D.C. 20555
UNDER CONTRACT NO. DE-AC02-76CH00016
FIN A-3286

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

The views expressed in this report are not necessarily those of the U.S. Nuclear Regulatory Commission.

Available from
Superintendent of Documents
U.S. Government Printing Office
P.O. Box 37082
Washington, DC 20013-7982
and
National Technical Information Service
Springfield, Virginia 22161

ABSTRACT

Existing methods for sensitivity analysis are described and new techniques are proposed. These techniques are evaluated through consideration relative to the QUASAR program. Merits and limitations of the various approaches are examined by a detailed application to the Suppression Pool Aerosol Removal Code (SPARC).

BLANK PAGE

CONTENTS

CONTENTS
(Continued)

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The authors are grateful to W.T. Pratt and R.A. Bari of BNL for reviewing and making many helpful remarks on the manuscript. Thanks are due to Ms. Deborah Faivre for the excellent job in typing the manuscript.

The work described in this report is supported by the Division of Accident Evaluation, United States Nuclear Regulatory Commission. The authors acknowledge the continued support and guidance of P. Wood, J. Mitchell and M. Silberberg of the USNRC.

The visit of the first author to Brookhaven National Laboratory was made possible through funds provided by the Science and Technology Agency (STA) of Japan. The first author is thankful to both BNL and STA for this opportunity.

NOMENCLATURE

Symbol	Definition
a_i	regression coefficient
a_i^*	standardized regression coefficient
$\hat{a}_i$	rank regression coefficient
$\hat{a}_i^*$	standardized rank regression coefficient
$C_W(y)$	cumulative probability at y calculated by weighting method
DF	integral decontamination factor for CsI
$f(x)$	uniform distribution function for x
$f(\bar{x})$	original PDF for $\bar{x}$
$F(x_1,x_2,....x_k)$	functional relationship between inputs and outputs for the original computer model
$\hat{F}(\bar{r}_s)$	the rank of y corresponding to the sample rank vector $\bar{r}_s$
$f_i(x_i)$	original PDF for x_i
L	total leakage of all radionuclides into the wetwell atmosphere
M	uniform bound
M_{in}	mass of CsI aerosol entering the suppression pool from the reactor coolant system
M_{out}	mass of CsI aerosol leaving the suppression pool and entering the wetwell atmosphere
M_{out}^j	mass of j-th radionuclide species leaving the suppression pool and entering the wetwell atmosphere
$q(x)$	normal distribution function for x

NOMENCLATURE
(Continued)

Symbol	Definition
$q(\bar{x})$	new PDF for $\bar{x}$
$q_i(x_i)$	new PDF for x_i
r_i	rank of x_i
r_i^*	standardized form of r_i
r_y	rank of y
r_y^*	standardized form of r_y
$r_{y(j)}$	rank value of y given by the classical rank regression model for rank vector $\bar{r}_j$
$\bar{r}_s$	rank vector for original sample
$\bar{r}_j$	j-th rank vector for the new sample with the new input PDFs
r_{ij}	rank of x_i corresponding to X_{ij}
R^2	coefficient of determination in the regression model
t_i	time of scrubbing initiation
t_f	time of scrubbing termination
$u(x)$	unitary step function
W_j	weighting factor in weighting method
x_i	input variable
x_i^*	standardized form of x_i
$\bar{x}$	vector of input variables
$\bar{x}_i(\bar{x}_s)$	i-th (s-th) input vector for the original sample

NOMENCLATURE
(Continued)

Symbol	Definition
$\bar{X}_j$	j-th input vector for the new sample with the new input PDFs
X_{ij}	j-th input value of x_i for the new sample with the new input PDFs
$x_i(j)$	upper bound of x_i in the j-th interval for the original sample
y	output variable
y^*	standardized form of y
y_j	value of y given by the original computer model for the input vector $\bar{x}_j$
$\hat{y}_j$	value of y given by the regression model for the input vector $\bar{x}_j$
$y(j)$	value of y given by the original computer model for original sample input vector in which the rank of x_i is j
Y_j	value of y given by the original computer model for the input vector $\bar{X}_j$
$\langle y\rangle_W$	mean of y calculated by weighting method
μ	mean of x given the PDF q(x)
σ	standard deviation of x given the PDF q(x)
$\sigma_W(y)$	standard deviation of y calculated by weighting method
$\sigma(\cdot)$	standard deviation of the variable within the parentheses
$\langle\cdot\rangle$	mean of the variable within the parentheses

EXECUTIVE SUMMARY

The complex physical processes governing the phenomena that determine source terms are not completely understood. Thus, characterization of the magnitudes of uncertainties associated with predictions of computer models is necessary. The radiological releases following a severe nuclear reactor accident can be estimated using semi-mechanistic computer codes such as the Source Term Code Package (STCP).

In order to better establish the estimates of source terms using these computer codes, quantification of the uncertainties in the resulting source term is essential.

The Quantification and Uncertainty Analysis of Source Terms for Severe Accidents in Light Water Reactors (QUASAR) program is aimed at addressing uncertainties associated with input parameters as well as phenomenological models. In order to achieve the former objective, it is necessary to perform: (1) uncertainty analysis which yields the Probability Density Functions (PDFs) of the model outputs, and (2) sensitivity analysis which determine the sensitivity of the output PDFs to the input PDFs.

The objectives of this report are:

1. to describe existing methods and to propose new approaches by which to assess the sensitivity of the output probability distributions characterizing source term uncertainties to the input distribution assumptions,

2. to apply the existing and the proposed methods to a member code of the Source Term Code Package in order to assess the feasibility of their implementation. The Suppression Pool Aerosol Removal Code (SPARC) is adopted for this purpose, and

3. to assess and to compare the success of the methods under consideration.

Existing methods for assessment of the sensitivity output distributions to the form of the distributions attached to input parameters to computer models have been described. These are referred to as the classical regression method, the weighting method and the rejection method. Further, modifications to and variations upon these established methods have been proposed. These are referred to as the modified regression method and the method of closest distance. Through application to SPARC, the possibility of implementing these techniques within the context of the QUASAR program has been assessed.

The study shows that:

a) The contrast amongst the results stemming from the five methods is not marked. In general, the modified rank regression technique performed better than the classical rank regression model while the method of closest distance outperformed the weighting method.

b) The performance of the two regression methods reviewed clearly would in general be dependent upon the degree to which the regression model adopted can provide a good fit to the underlying computer model. Hence, in circumstances where the regression model provides a poor surrogate for the original computer model, then the weighting method and the rejection method might be expected to produce better approximations for the output distributions than would the regression methods. The method of closest distance would also be anticipated to display some reliance upon the goodness of the regression fit since the distance measure utilized therein incorporates regression-based information. However, one might expect the method of closest distance to provide a better surrogate for the computer model than the regression model in circumstances where the regression model would provide a poor fit since the method of closest distance does not explicitly resort to response surface techniques but instead relies directly upon outputs obtained by the original computer model.

c) It is not possible to draw general conclusions about the relative performance of the methods assessed regarding application to a range of computer models. Judicious application of each of the five approaches (i.e., the classical and modified regression approachs, the weighting method, the rejection method and the method of closest distance), weighed by a knowledge of the goodness of fit of the regression models formulated will provide a basis for output distribution sensitivity analysis in the QUASAR program.

1. INTRODUCTION

1.1 Background

The complex physical processes governing the phenomena that determine source terms are not completely understood. Thus, characterization of the magnitudes of uncertainties associated with predictions of computer models is necessary. The radiological releases following a severe nuclear reactor accident can be estimated using semi-mechanistic computer codes such as the Source Term Code Package (STCP).[1]

In order to better establish the estimates of source terms using these computer codes, quantification of the uncertainties in the resulting source term is essential.

The sources of uncertainty include both the models themselves as well as model input parameters required to characterize the physio-chemical processes.

The Quantification and Uncertainty Analysis of Source Terms for Severe Accidents in Light Water Reactors (QUASAR)[2] program is aimed at addressing uncertainties associated with input parameters as well as phenomenological models. In order to achieve the former objective, it is necessary to perform: (1) uncertainty analysis which yields the Probability Density Functions (PDFs) of the model outputs, and (2) sensitivity analysis which determine the sensitivity of the output PDFs to the input PDFs. The current report is designed largely to address the second goal. The first goal is discussed elsewhere.[3]

1.2 Review of Sensitivity Analysis Methods

A list of current sensitivity analysis methods include:

1. Differential Analysis Methods[4,5,6],
2. Expansion Methods[7,8]
3. Response Surface Methods[9]
4. Direct Methods.[10,11]

Most of the differential and expansion methods are based on the calculation of the first order partial derivative of output variable with respect to each input variable about their reference values. The sensitivity measures based on these methods depend generally, therefore, on the reference values. Since higher order derivatives are not included, this first order derivative measure is valid only for the case of functional relationships that deviate minimally from linearity. The complexities involved in source term uncertainty analysis make it very difficult to use analytical sensitivity analysis methods (i.e., both differential and expansion methods).

The response surface method (RSM) is a direct simplified simulation of physical or logical models. The RSM method is based on building a replacement for the computer model under consideration with the use of various sampling techniques (Latin Hypercube Sampling,[12,13] Random Sampling, Experimental Design, etc.) and regression analysis.[14,15,16] The resulting regression model is then used as a replacement for the original model in the sensitivity analyses. It should be noted that the sensitivity analysis is then performed with regard to the regressed models and not on the original models; thus, the results obtained are only as valid as the approximations attendant in regressed models themselves.

Direct methods utilize a limited number of calculations based upon the original model. Iman et al.,[10] have proposed a method which determines the impact of the input PDFs on the output PDFs without the need for computer calculations and response surface technique. In this method, the statistical parameters for the output variable are calculated based upon a weighting factor which reflects the change of the input PDFs. Henceforth, this approach will be referred to as the weighting method. Beckman and McKay[11] have proposed the so-called rejection method which permits the effects of the new PDFs to be ascertained through inspection of a subset of the original computer model outputs.

In the QUASAR sensitivity analysis where the sensitivity of the output PDFs to the input PDFs is to be evaluated, both response surface methods and direct methods seem to be applicable. As an example of response surface methods, the classical regression model is briefly described in Section 3.1.1 and a modified regression model is proposed in Section 3.1.2. Regarding direct methods, the weighting method and the rejection method are described in Sections 3.2.1 and 3.2.2, respectively. Further, a method related to the weighting approach but based upon the modified regression model is proposed in Section 3.2.3.

1.3 Objectives and Organization of the Report

The objectives of this report are:

1. to describe existing methods and to propose new approaches by which to assess the sensitivity of the output probability distributions characterizing source term uncertainties to the input distribution assumptions,

2. to apply the existing and the proposed methods to a member code of the Source Term Code Package in order to assess the feasibility of their implementation. The Suppression Pool Aerosol Removal Code (SPARC)[17] is adopted for this purpose, and

3. to assess and to compare the success of the methods under consideration.

Chapter 2 briefly describes the QUASAR methodology. The existing and the proposed sensitivity analysis methods are described in Chapter 3. Chapter 4 demonstrates the merits and limitations of the methods through application to SPARC. The conclusions are summarized in Chapter 5.

2. QUASAR METHODOLOGY

In order to estimate the uncertainties associated with the severe accident source terms predicted by the STCP, the following steps are to be followed as part of the QUASAR (Quantification and Uncertainty Analysis of Source Terms for Severe Accidents in Reactors) program at Brookhaven National Laboratory (BNL):[2]

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 various codes in the STCP.

2. Uncertainty 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 various models and their associated computer codes in the STCP. The quantification process in QUASAR 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 parameters/options to the STCP. The propagation of input uncertainties through the STCP will be accomplished through a stratified Monte Carlo simulation using the Latin Hypercube Sampling approach.[12,13]

3. Sensitivity Analysis: Following the completion of the uncertainty analysis stage, the sensitivity of the output PDFs will be established.

This report provides techniques for the last of these steps: sensitivity analysis following a detailed uncertainty analysis. This sensitivity analysis will address the impact of the assumptions regarding the subjective input PDFs.

In the following discussion, we suppose that the results of screening analysis and uncertainty analysis are available. That is, the number of input variables included in the uncertainty analysis has been reduced to a manageable size, and the values of the output statistical parameters such as mean, standard deviation, cumulative distribution function have been obtained through stratified Monte Carlo simulation using the Latin Hypercube Sampling approach.

Sensitivity analysis techniques most suitable for the present application will be discussed and developed in Chapter 3.

3. METHODS FOR OUTPUT DISTRIBUTION SENSITIVITY ANALYSIS

In order to perform post-uncertainty sensitivity analyses, extensive mathematical relationships between model/code inputs and outputs are required. These relationships can be obtained through (a) regression and response surface analysis in which the original computer model is replaced by a simplified surrogate model, or (b) direct utilization of a limited number of calculations based upon the original model as proposed by Iman et al.,[10] and Beckman and McKay.[11]

Examples of both regression and direct methods are reviewed and possible alternatives are developed and outlined in the following sections.

3.1 Regression Methods

3.1.1 Classical Regression Approach

In a linear regression approach, the output variable y is approximated by a linear polynomial function of input variables $x_1, x_2, \ldots, x_K$, that is:

$$y = a_0 + \sum_{i=1}^{K} a_i x_i \tag{3.1}$$

where a_0 and a_i are constants fitted to the computer model results based upon Latin Hypercube Sampling or 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 as:

$$x_i \rightarrow x_i^* = (x_i - \langle x_i \rangle)/\sigma(x_i), \tag{3.2}$$

$$y \rightarrow y^* = (y - \langle y \rangle)/\sigma(y), \tag{3.3}$$

where $\langle x_i \rangle$ and $\langle y \rangle$ are the means and $\sigma(x_i)$ and $\sigma(y)$ the standard deviations of the variables x_i and y, respectively. Eq. (3.1) can now be rewritten in the following standardized form,

$$y^* = \sum_{i=1}^{K} a_i^* x_i^*. \tag{3.4}$$

Here the 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.[14,15] Specifically, the smallest value of each variable 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 number of observations. Therefore, the rank regression form of

Eq. (3.1) is given by

$$r_y = \hat{a}_0 + \sum_{i=1}^{K} \hat{a}_i r_i. \tag{3.5}$$

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:

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

and

$$r_y \rightarrow r_y^* = (r_y - \langle r_y \rangle)/\sigma(r_y) \tag{3.7}$$

where $\langle r_i \rangle$ is the mean value of some sample set r_i, $\langle r_y \rangle$ is the mean value of the resultant output set r_y, $\sigma(r_i)$ and $\sigma(r_y)$ are the standard deviations for the sample sets r_i and r_y, respectively. Therefore, using Eqs. (3.6) and (3.7), the standardized form of Eq. (3.5) is as follows:

$$r_y^* = \sum_{i=1}^{K} \hat{a}_i^* \, r_i^*. \tag{3.8}$$

A value of y is easily obtained from the rank r_y by using an interpolation method. It is known that the rank regression equation can well approximate y, even if it does not only include linear terms, when y is a monotonic function of the x_i's.[14]

The goodness-of-fit of the regression model is measured by the quantity,[15]

$$R^2 = \sum_{j=1}^{N} (\hat{y}_j - \langle y \rangle)^2 / \{ \sum_{j=1}^{N} (\hat{y}_j - \langle y \rangle)^2 + \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 \}, \tag{3.9}$$

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.

With regard to output distribution sensitivity analysis, the sensitivity calculations are now performed with regard to the regression model rather than with regard to the original computer code. That is, alternative distributions on the inputs x_i are considered and their effects upon the output distribution as predicted by the regression model are ascertained.

3.1.2 Modified Regression Approach

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

$$y = F(x_1, x_2, \ldots, x_K), \tag{3.10}$$

a Taylor series expansion of F around a sample vector $\bar{x}_s = (x_{1s}, x_{2s}, \ldots, x_{Ks})$ can be performed to obtain:

$$y \approx F(\bar{x}_s) + \sum_{i=1}^{K} (x_i - x_{is}) \left. \frac{\partial F}{\partial x_i} \right|_{x_i = x_{is}} . \tag{3.11}$$

Here, $\bar{x}_s$ is one of the original samples generated by LHS methods relative to the original input distributions. Now, in the light of alternative input distributions, a new set of input vectors $\bar{x}$, where $\bar{x} = (x_1, x_2, \ldots, x_K)$ are generated. Eq. (3.11) is implemented relative to a new vector $\bar{x}$ by effecting the Taylor expansion about the original vector $\bar{x}_s$ (generated from the original distributions) that is the closest to the new vector $\bar{x}$, i.e., the original vector $\bar{x}_s$ that minimizes the quantity:

$$\sum_{i=1}^{K} a_i^{\star 2} (x_i^\star - x_{is}{}^\star)^2 . \tag{3.12}$$

The constant $a_i^\star$ is a weight that reflects the importance of the i-th input variable (e.g., as measured by the Partial Correlation Coefficients (PCCs) and/or Standardized Regression Coefficients (SRCs) to be discussed in Chapter 4), and $x_i{}^\star$ and $x_{is}{}^\star$ are the standardized values of x_i and x_{is}, respectively. Here, Eq. (3.12) may be viewed as a modified Euclidean distance measure that accounts for the importance of the individual dimension (i.e., variables). It should be noted that the analytic form of function F is not known, however, the value of $F(\bar{x}_s)$ which is the computer model output corresponding to the sample input vector $\bar{x}_s$ 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 F at x_i is approximated as

$$\left. \frac{\partial F}{\partial x_i} \right|_{x_i = x_{is}} \approx a_i , \tag{3.13}$$

that is, as the derivative given by the regression model of Eq. (3.1).

Substituting Eq. (3.13) into Eq. (3.11) yields,

$$y = F(\bar{x}_s) + \sum_{i=1}^{K} a_i (x_i - x_{is}). \tag{3.14}$$

Equation (3.14) 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,

$$r_y = \hat{F}(\bar{r}_s) + \sum_{i=1}^{K} \hat{a}_i (r_i - r_{is}) . \tag{3.15}$$

Here, $\hat{F}(\bar{r}_s)$ indicates the rank of output y corresponding to the rank vector $\bar{r}_s$ of the original sample and the vector $\bar{r}_s$ is chosen to minimize the quantity

$$\sum_{i=1}^{K} \hat{a}_i^{*2} (r_i^* - r_{is}^*)^2 \qquad (3.16)$$

where $\bar{r}_s^*$ is the standardized version of $\bar{r}_s$. The constant $\hat{a}_i^*$ is a weight that reflects the importance of the i-th input rank as measured by the Partial Rank Correlation Coefficients (PRCCs) and/or Standardized Rank Regression Coefficients (SRRCs) to be discussed in Chapter 4. Eq. (3.15) is now used as a surrogate for the original computer model in considering alternative input distributions and their effects upon the output distributions.

3.2 Direct Methods

Direct methods refer to those approaches that utilize the input/output relationships provided by the original computer model calculations (with respect to a sample of inputs) without relying upon regression fits to those relationships.

3.2.1 Weighting Method

A small sample sensitivity analysis technique which directly utilizes the computer model results generated based upon Latin Hypercube Sampling of input distributions has been proposed by Iman, et al.[10] This method can be used to determine the impact of the probability distribution functions characterizing the input variables on the outputs without the need for additional computer calculations 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 assignment of the output variable y, c(y) may be approximated by

$$\langle y \rangle_W = \sum_{j=1}^{N} W_j \, y(j), \qquad (3.17)$$

$$\sigma_W (y) = \sqrt{\sum_{j=1}^{N} W_j \, (y(j) - \langle y \rangle_W)^2}, \qquad (3.18)$$

$$C_W (y) = \sum_{j=1}^{N} W_j \, u(y - y(j)), \qquad (3.19)$$

where u is the unitary step function defined by

$$u(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 , \end{cases} \tag{3.20}$$

$y(j)$ 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 $f_i(x_i)$:

$$W_j = \int_{x_i(j-1)}^{x_i(j)} q_i(x_i)\, dx_i \ . \tag{3.21}$$

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

$$j/N = \int_{x_i(0)}^{x_i(j)} f_i(x_i)\, dx_i, \tag{3.22}$$

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

3.2.2 Rejection Method

In this method proposed by Beckman and McKay,[11] a subset of the original computer model outputs associated with samples 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 $\bar{x} = (x_1, x_2, \ldots, x_K)$ and the independent sample vectors

$$\bar{x}_j = (x_{1j}, x_{2j}, \ldots, x_{Kj}) \tag{3.23}$$

which are generated with respect to the probability density function $f(\bar{x})$. Let the output of the computer model corresponding to the input $\bar{x}_j$ be y_j.

The rejection method relies on a random selection of the existing sets of variables $(\bar{x}_j, y_j)$ to determine the sensitivity of the output distribution to the input joint distribution. Let the new input PDF be $q(\bar{x})$.

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

$$\frac{q(\bar{x})}{f(\bar{x})} < M \tag{3.24}$$

for all $\bar{x}$, and that the domain of $q(\bar{x})$ be contained within the domain of $f(\bar{x})$.

Let the random variable V, given as sample vector $\bar{x}_j$, be uniform between 0 and $M \cdot f(\bar{x}_j)$. The data set $(x_j, \bar{y}_j)$ are retained as a sample from $q(\bar{x})$ if a random realization of V is less than $q(\bar{x}_j)$. The theoretical basis for this approach is expounded in Reference 11.

3.2.3 Method of Closest Distance

In this section, a method based upon the modified rank egression approach developed in Section 3.1.2 is proposed. It is a method that utilizes the input/output relationships provided by the original computer model calculations based upon the LH samples.

In order to require only the original computer model outputs, eliminating the second term of Eqs. (3.14) and (3.15) yields,

$$y \approx F(\bar{x}_s) \tag{3.25}$$

and

$$r_y \approx \hat{F}(\bar{r}_s) \ . \tag{3.26}$$

The currently proposed technique comprises the following steps:

1. A set of Latin Hypercube input sample vectors is generated:

$$\bar{x}_i = (x_{1i}, x_{2i}, \ldots, x_{Ki}), \quad i=1, 2, \ldots, N \tag{3.27}$$

   Here, the N samples correspond to N combinations of values for the K parameter inputs. The input $\bar{x}_i$ yields the output y_i from the computer code where, for simplicity, just one output is considered.

2. In order to ascertain the effects of the input probability density functions (PDFs) on the output distributions, another set of Latin Hypercube input samples,

$$\bar{X}_j = (X_{1j}, X_{2j}, \ldots, X_{Kj}), \quad j=1, 2, \ldots, M \tag{3.28}$$

   is generated. These samples are 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 $\bar{X}_j$ is approximated by the LHS output value y_s whose corresponding LHS input vector $\bar{x}_s$ or rank vector $\bar{r}_s$ is "closest" to the vector $\bar{X}_j$ or rank vector $\bar{r}_j$ corresponding to $\bar{X}_j$, respectively, where Eq. (3.12) or Eq. (3.16) provides the definition of "closeness". Then the corresponding original output y_s is used as an approximate replacement for the output Y_j. Hence M random output values are approximated by the near-

est of the N Latin Hypercube Sample output values obtained from the original computer code calculations.

4. Hence, an approximation of the output distributions resulting from the second set of input distributions is 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 Section 3.1.2 except that the regression based terms are excluded from the surrogate model of the original computer model.

4. APPLICATION TO THE SUPPRESSION POOL AEROSOL REMOVAL CODE

The Suppression Pool Aerosol Removal Code (SPARC)[17] 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 current purpose of demonstrating merits of the various sensitivity analysis methods for the following reasons:

(1) Relatively small computational requirements (2 minutes per sequence on an IBM-3090 machine).

(2) Due to a limited number of input variables, a small number 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 techniques with direct SPARC sensitivity calculations.

For the purpose of the present sensitivity analysis with the SPARC model, the following calculational outputs will be tracked.

(a) The integral Decontamination Factor (DF) for CsI defined by:

$$DF = \frac{\int_{t_i}^{t_f} M_{in} \, dt}{\int_{t_i}^{t_f} M_{out} \, dt} . \tag{4.1}$$

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

$$L = \int_{t_i}^{t_f} \sum_j M_{out}^j \, dt \tag{4.2}$$

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.

4.1 Reference Analysis

The selected SPARC input variables together with their assigned ranges and probability distributions as used for the present reference analysis are given in Table 4.1.

Table 4.1 Input Variables, Initial Ranges and Distributions

Variable	Range		Distribution
	Lower Bound	Upper Bound	
x_1: RATIO	1	4	Uniform
x_2: DIAM (mm)	3	20	Uniform
x_3: VSWARM (cm/sec)	20	120	Uniform
x_4: VIMPT (cm/sec)	0	30,000	Uniform
x_5: NRISE	100	1,000	Uniform
x_6: CDIF	1	4	Uniform

RATIO: Bubble aspect ratio

DIAM: Mean bubble diameter

VSWARM: Bubble swarm rise velocity

VIMPT: Inlet impact velocity

NRISE: Number of time steps for the calculation of decontamination factors during bubble rise

CDIF: A constant imbedded in the diffusional removal model

These six input variables will be sampled using the Latin Hypercube Sampling techniques in accordance with their distributions and variable ranges. A sample size of fifty has been considered corresponding to fifty combinations of values for the six input variables.

4.1.1 Partial Correlation and Standardized Regression Coefficients[16]

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:

$$\hat{y} = a_0 + a_1 x_1. \tag{4.3}$$

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. (4.3). 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 standardized regression coefficients and partial correlation coefficients on variable ranks rather than on the actual values for the variables: such coefficients are known as standardized rank regression coefficients (SRRCs) and partial rank correlation coefficients (PRCCs). Specifically, the smallest value of each variable 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 number of observations. The standardized regression coefficients and/or partial correlation coefficients are then calculated on these ranks rather than upon the underlying raw variables. The rank transformation permits a better fit of the regression model to the actual model since then the weaker assumption of montonicity between raw outputs and inputs replaces the linearity requirement.

Based on the SPARC-run results for fifty initial LHS input vectors (LHS-1), the evaluation of partial rank correlation coefficients (PRCCs) and standardized rank regression coefficients (SRRCs) was performed using the computer program in Ref. 16. Tables 4.2 and 4.3 show the PRCCs, SRRCs and coefficient of determination, R^2, for the integral DF for CsI and the total leakage of all radionuclides into the wetwell atmosphere, respectively. These results also 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 the outputs DF and L.

Table 4.2 PRCCs, SRRCs and R^2 Values for the CsI Integral DF

Input Variable	LHS-1		LHS-2	
	PRCC	SRRC	PRCC	SRRC
x_1	0.95	0.77	0.94 (-1.1%)	0.80 (3.9%)
x_2	-0.76	-0.29	-0.72 (-5.3%)	-0.29 (0%)
x_3	-0.90	-0.51	-0.85 (-5.6%)	-0.46 (-9.8%)
x_4	0.31	0.08	0.25 (-19%)	0.07 (-13%)
x_5	0.35	0.09	0.00 (-100%)	0.00 (-100%)
x_6	0.19	0.05	0.15 (-21%)	0.04 (-20%)
R^2	0.94	0.94	0.92	0.92

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

Table 4.3 PRCCs, SRRCs and R^2 Values for the Total Leakage (L) of All Radionuclides Into the Wetwell Atmosphere

Input Variable	LHS-1		LHS-2	
	PRCC	SRRC	PRCC	SRRC
x_1	-0.95	-0.77	-0.95 (0%)	-0.80 (3.9%)
x_2	0.76	0.29	0.75 (1.3%)	0.30 (3.4%)
x_3	0.90	0.52	0.86 (-4.4%)	0.46 (-12%)
x_4	-0.32	-0.08	-0.27 (-16%)	-0.07 (-13%)
x_5	-0.35	-0.09	0.00 (-100%)	0.00 (-100%)
x_6	-0.18	-0.05	-0.17 (-5.6%)	-0.05 (0%)
R^2	0.94	0.94	0.93	0.93

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

4.1.2 Statistical Parameters for Output Variables

Tables 4.4 and 4.5 show some properties of the output distributions, resulting from propagation of the 50 input LH vectors through SPARC. These include the mean, standard deviation, 5th, 50th and 95th percentile values for DF and L, respectively. It is found that the quantity, characterizing distribution width, defined by

$$\sigma_{DF}/\langle DF\rangle \text{ or } \sigma_L/\langle L\rangle \qquad (4.4)$$

is the same order of magnitude as the distribution width of the input variables.

The calculated cumulative distribution functions for the outputs DF and L are shown in Figs. 4.1 and 4.2, respectively.

Of course, numerical simulation techniques such as LH sampling 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 Latin Hypercube Sampling 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 4.2 through 4.5. 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.

4.2 Sensitivity Analysis

The reference analysis of the previous section has shown that the two SPARC variables x_1 (RATIO) and x_3 (VSWARM) are the most significant contributors to both the integral DF for CsI (DF), and the total leakage of all radionuclides into the wetwell atmosphere (L). This sensitivity analysis, therefore, focuses 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 4.6 lists the assumed distributions for x_1 and x_3 in the current sensitivity cases as compared with the reference analysis of Section 4.1. The mean and the range of each input variable in the sensitivity analysis are assumed to be the same as those given in Table 4.1 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 regression and direct sensitivity analysis methods described in Chapter 3. By comparing the results thereof with the output distributions based upon LH sampling of the new input distributions (LHS-3 and LHS-4 of Table 4.6) and runs of the actual computer model SPARC, the success of the methods may be assessed.

Table 4.4 Summary of Results for the CsI Integral DF

Statistical Parameter	LHS-1	LHS-2
Mean, <DF>	240.3	243.7 (1.4%)
Standard Deviation, σ_{DF}	129.9	138.8 (6.9%)
$\sigma_{DF}/\langle DF\rangle$	0.54	0.57 (5.6%)
5th	131.7	128.6 (-2.4%)
50th	199.7	196.5 (-1.6%)
95th	574.6	547.8 (-4.7%)

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

Table 4.5 Summary of Results for the Total Leakage (L) of All Radionuclides Into the Wetwell Atmosphere

Statistical Parameter	LHS-1	LHS-2
Mean, $\langle L \rangle$	2503	2506 (0.1%)
Standard Deviation, σ_L	904.9	907.1 (0.2%)
$\sigma_L / \langle L \rangle$	0.36	0.36 (0%)
5th	879.7	923.2 (4.9%)
50th	24877	2535 (1.9%)
95th	3766	3481 (2.0%)

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

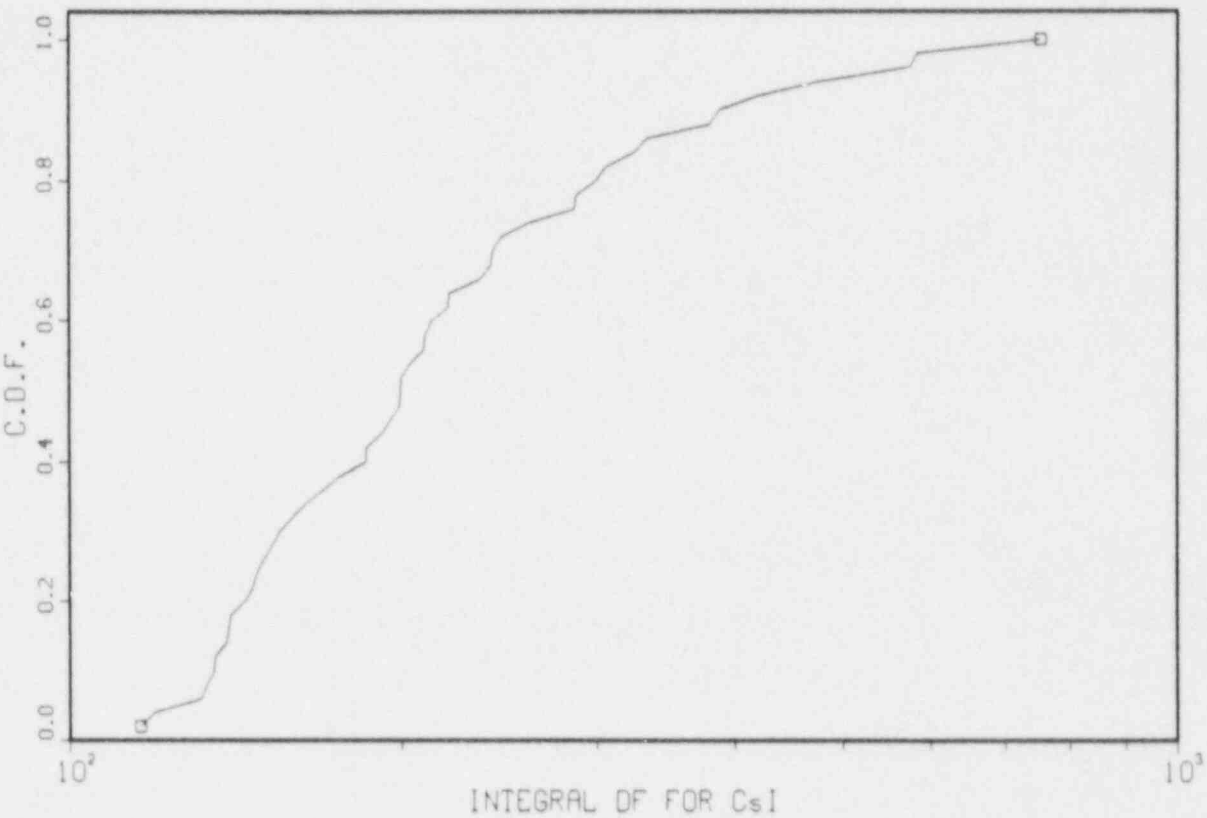

Figure 4.1 Cumulative Distribution Function for the CsI Integral DF (LHS-1)

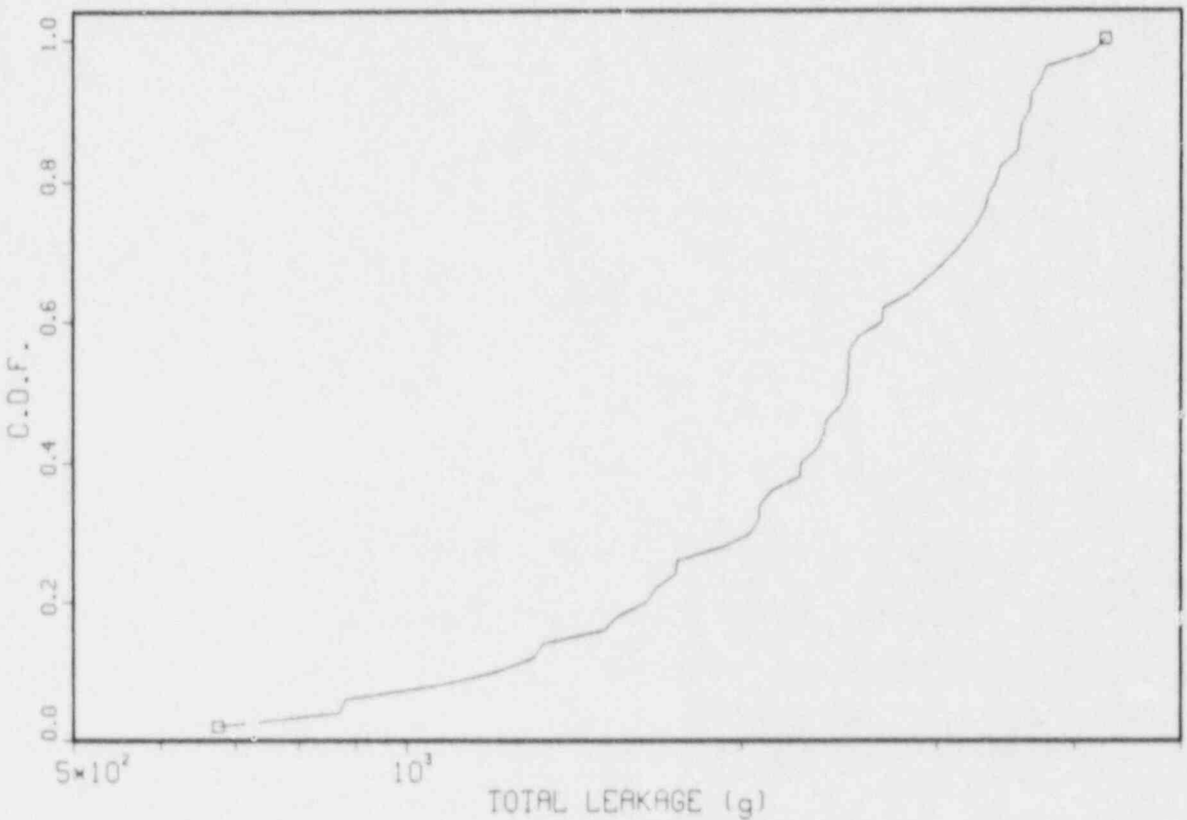

Figure 4.2 Cumulative Distribution Function for the Total Leakage of All Radionuclides into the Wetwell At   here (LHS-1)

Table 4.6 Assumed Probability Density Functions

Input Variable	Reference Analysis	Sensitivity Analysis	
		S-1	S-2
x_1	Uniform	Normal	Uniform
x_3	Uniform	Uniform	Normal
All Others	Uniform	Uniform	Uniform
Sample Set	LHS-1	LHS-3	LHS-4

The normal distribution is of the form (Fig. 4.3):

$$q(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \tag{4.5}$$

where the distribution parameters, μ and σ are determined as follows. Requirement of the same mean and range between the normal distribution $q(x)$ and uniform distribution

$$f(x) = 1/(b-a), \tag{4.6}$$

provides the following relationships[13]:

$$\int_a^b x\, f(x)\, dx = \int_{-\infty}^{\infty} x\, q(x)\, dx, \tag{4.7}$$

and

$$\int_{-\infty}^{a} q(x)\, dx = \int_b^{\infty} q(x)\, dx = 0.001 \tag{4.8}$$

That is, a negligible probability is attached to those parts of the parameter space that fall outside the original range. Equations (4.7) and (4.8) can be used to determine the values of μ and σ as:

$$\mu = (a+b)/2 \tag{4.9}$$

$$\sigma = (b-a)/6.182 \tag{4.10}$$

4.2.1 Regression Methods

4.2.1.1 Classical Regression Model

The adopted standardized rank regression equation is:

$$r_y^* = \sum_{i=1}^{K} \hat{a}_i^*\, r_i^*. \tag{4.11}$$

Table 4.7 lists the standardized rank regression coefficients, $\hat{a}_i^*$, and coefficient of determination, R^2, for the integral DF and the total leakage (L) as calculated by the "Stepwise Regression with PRESS and Rank Regression" Program.[15] These results indicate that the fit of Eq. (4.11) to SPARC for the reference output variables is satisfactory (R^2>0.9).

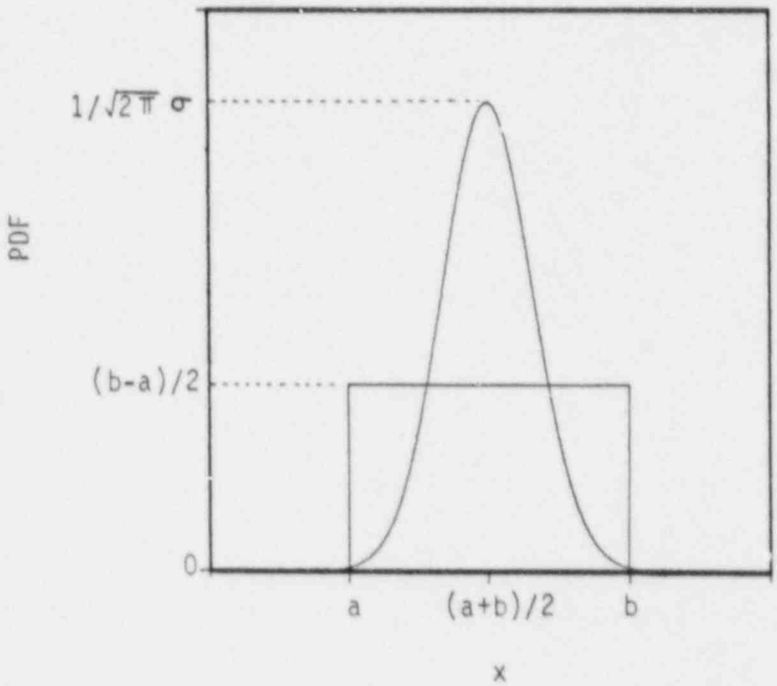

Figure 4.3 Illustration of Uniform and Normal Distributions

Table 4.7 Standardized Rank Regression Coefficients and R^2 Values

Input Variable	Integral DF for CsI $\hat{a}_i^*$	Total Leakage Into Wetwell Atmosphere $\hat{a}_i^*$
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

Eq. (4.11) can be used to obtain the output values (DF and L) for the LHS-3 or LHS-4 input vectors in the following manner:

(1) For an LHS-3 or LHS-4 input vector,

$$\bar{X}_j = (X_{1j}, X_{2j}, \ldots, X_{Kj}), \tag{4.12}$$

the rank of each variable x_i, r_{ij}, corresponding to its value X_{ij} is assigned using the relation

$$x_i(r_{ij} - 1) < X_{ij} < x_i\ (r_{ij}), \tag{4.13}$$

where $x_i(r_{ij})$ is the upper bound of the r_{ij}-th stratum of the original (LHS-1) input distributions given by

$$x_i(r_{ij}) = x_i(0) + r_{ij} \cdot (x_i(N) - x_i(0))/N. \tag{4.14}$$

This expression relies upon the uniformity of the original input distributions in LHS-1.

(2) Those sets of rank data denoted

$$\bar{r}_j = (r_{1j}, r_{2j}, \ldots, r_{Kj}),\ j = 1, 2, \ldots, N \tag{4.15}$$

are substituted into Eq. (4.11) in a standardized form. We thus obtain the corresponding rank data $r_y(j)$ for the output variable y.

(3) The raw value corresponding to the rank value, $r_y(j)$, is obtained by an interpolation method using the relationships between the raw data y_i and the corresponding rank data r_y which were determined relative to the LHS-1 data.

4.2.1.2 Modified Regression Model

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

$$r_y = \hat{F}(\bar{r}_s) + \sum_{i=1}^{K} \hat{a}_i\ (r_i - r_{is})\ . \tag{4.16}$$

Here the coefficient $\hat{a}_i$ is identical to the SRRC, $\hat{a}^*$, in Table 4.7. 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, $\bar{r}_s$, to a given LHS-3 or LHS-4 vector was determined using Eq. (3.16) with the values of $\hat{a}_i^*$ given in Table 4.7.

The calculational procedure using Eq. (4.16) to obtain the output value y for the LHS-3 or LHS-4 input vectors is that described in subsection 4.2.1.1. Note that since simple regression models are being utilized to propagate the sample sets LHS-3 and LHS-4, we could equally well have used larger sample sets 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 sample sets.

4.2.2 Direct Methods

4.2.2.1 Weighting Method

In the weighting method, statistical parameters such as mean, the standard deviation and the CDF for the reference output variable y are calculated by Eqs. (3.17) through (3.19). The weighting factors W_j corresponding to the change in distributions given in Table 4.6 were calculated using Eqs. (3.21) and (3.22) and the results are shown in Table 4.8.

It is noted that the weighting method requires neither the LHS-3 nor the LHS-4 inputs. 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.

4.2.2.2 Rejection Method

This method requires the uniform bound M given by Eq. (3.24) as:

$$\begin{aligned} M &= \mathrm{MAX}\ \{q(\bar{x})/f(\bar{x})\} \\ &= \frac{b - a}{\sqrt{2\pi}\ \sigma} \\ &= \frac{6.182}{\sqrt{2\pi}} \\ &= 2.47, \end{aligned} \tag{4.17}$$

where Eqs. (4.5) through (4.10) 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.

4.2.2.3 Method of Closest Distance

In this method, Eq. (3.26) based upon the rank data has been used. The nearest LHS-1 input rank vector, $\bar{r}_s$, to a vector of LHS-3 or LHS-4 is determined using Eq. (3.16) with the values of $\hat{a}_i^*$ given in Table 4.7.

Table 4.8 Weighting Factor 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, ..., 25)

4.2.3 Results and Discussions

Tables 4.9 through 4.12 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 approaches is given in Figures 4.4 through 4.7.

Comparison of the SPARC results with the reference analysis results of Tables 4.4 and 4.5 as well as Figures 4.1 and 4.2 indicate that the standard deviation for both output variables (DF and L) are reduced due to a reduction in the standard deviation for x_1 or x_3 when a normal distribution is used.

Comparison of the sensitivity methods against the SPARC results show that:

(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, 5th and 95th percentiles show large differences as compared with other methods.

(3) Of the original 50 LHS samples, application of the rejection method dictated the retention of 22 samples for both sensitivity cases S-1 and S-2. These retained samples 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[11] of 1/M, where M (= 2.47 in this application) is the uniform bound defined in Subsection 3.2.2. 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 samples is small or where the bound M is large. In either case, the number of retained samples would be small.

Table 4.9 Statistical Parameters for the CsI Integral 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, ⟨DF⟩	[illegible]	218.0 (-4.3%)	221.3 (-2.8%)	246.2 (8.1%)	210.4 (-7.6%)	242.1 (6.3%)
Standard Deviation, σ_{DF}	91.3	85.9 (-5.9%)	91.5 (0.2%)	143.1 (57%)	65.6 (-28%)	130.0 (42%)
σ_{DF}/⟨DF⟩	0.401	0.394 (-1.7%)	0.413 (3.0%)	0.581 (45%)	0.312 (-22%)	0.537 (34%)
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	[illegible] (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 4.10 Statistical Parameters for the CsI Integral 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, <DF>	219.0	228.6 (4.4%)	221.3 (1.1%)	205.4 (-6.2%)	222.2 (1.5%)	227.0 (3.7%)
Standard Deviation, σ_{DF}	91.1	120.2 (32%)	91.5 (0.4%)	45.8 (-50%)	125.1 (37%)	113.6 (25%)
σ_{DF}/<DF>	0.416	0.526 (26%)	0.413 (-0.7%)	0.223 (-46%)	0.563 (35%)	0.500 (20%)
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 4.11 Statistical Parameters for the Total Leakage of All Radionuclides Into the Wetwell Atmosphere 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, <L>	2470	2507 (1.5%)	2533 (2.6%)	2532 (2.5%)	2579 (4.4%)	2497 (1.1%)
Standard Deviation, σ_L	749.3	663.2 (-11%)	790.5 (5.4%)	927.6 (24%)	731.5 (-2.4%)	906.7 (21%)
$\sigma_L/\langle L\rangle$	0.303	0.265 (-13%)	0.312 (3.0%)	0.366 (21%)	0.284 (-6.3%)	0.363 (20%)
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 4.12 Statistical Parameters for the Total Leakage of All Radionuclides Into the Wetwell Atmosphere 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, <L>	2589	2511 (-3.0%)	2559 (-1.2%)	2504 (-3.3%)	2710 (4.7%)	2564 (1.0%)
Standard Deviation, σ_L	831.4	784.9 (-5.6%)	825.6 (-0.7%)	452.3 (-46%)	941.5 (13%)	867.3 (4.3%)
σ_L/<L>	0.321	0.313 (-2.5%)	0.323 (0.6%)	0.181 (-44%)	0.347 (8.1%)	0.338 (5.3%)
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.

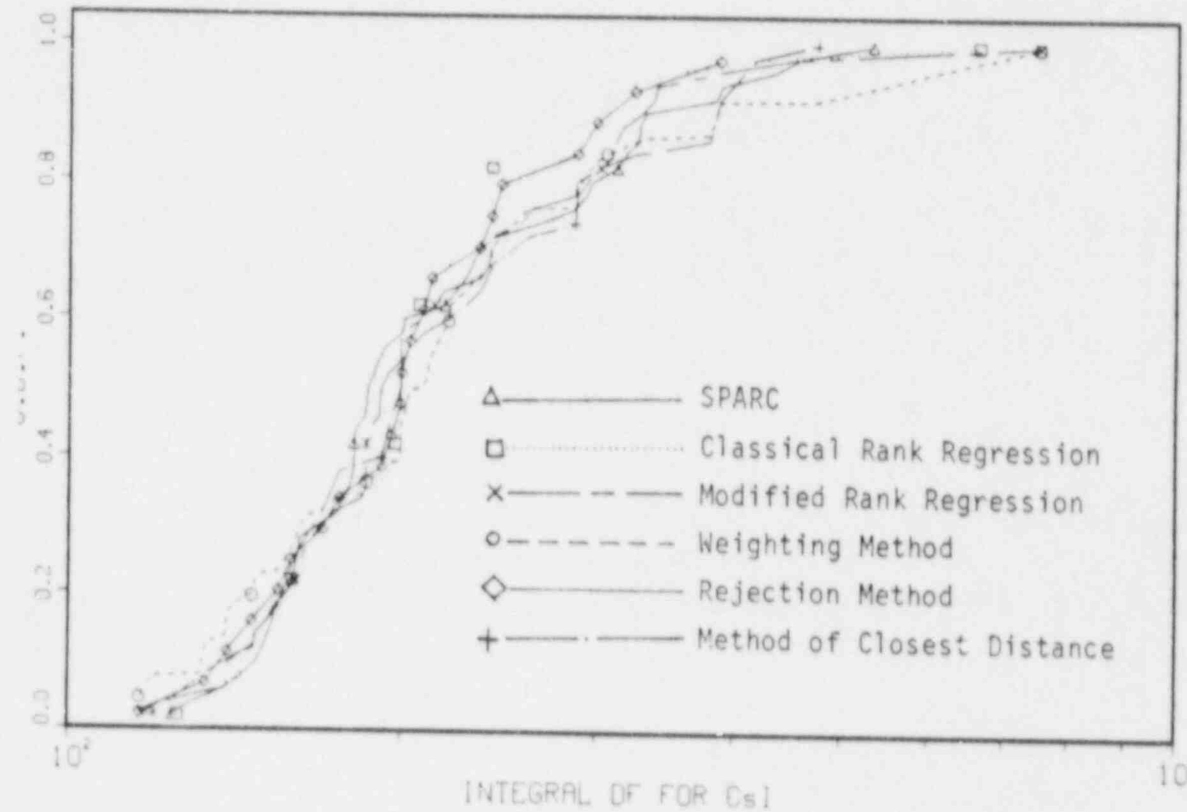

Figure 4.4 Comparison of Cumulative Distribution Functions (CDFs) for the Integral DF for CsI When the PDF for x_1 (RATIO) is Changed from Uniform to Normal Distribution

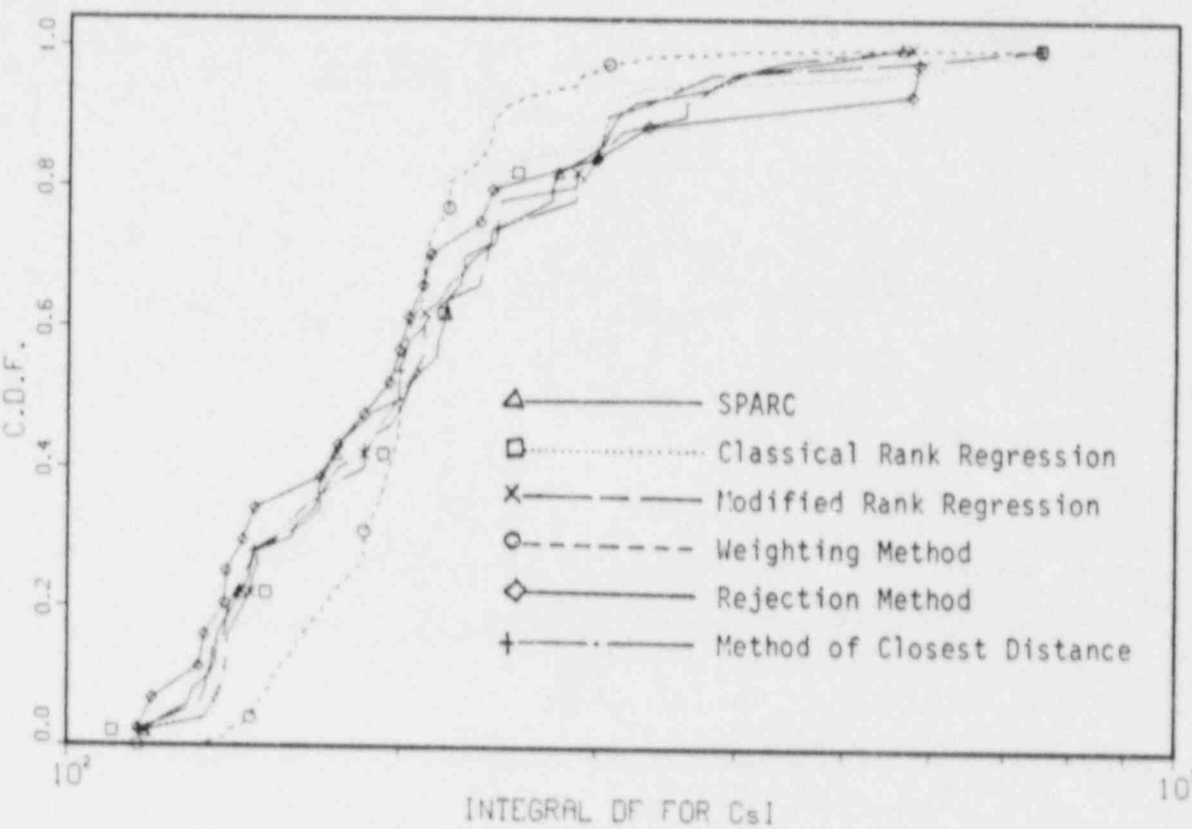

Figure 4.5 Comparison of Cumulative Distribution Functions (CDFs) for the Integral DF for CsI When the PDF for x_3 (VSWARM) is Changed from Uniform to Normal Distribution

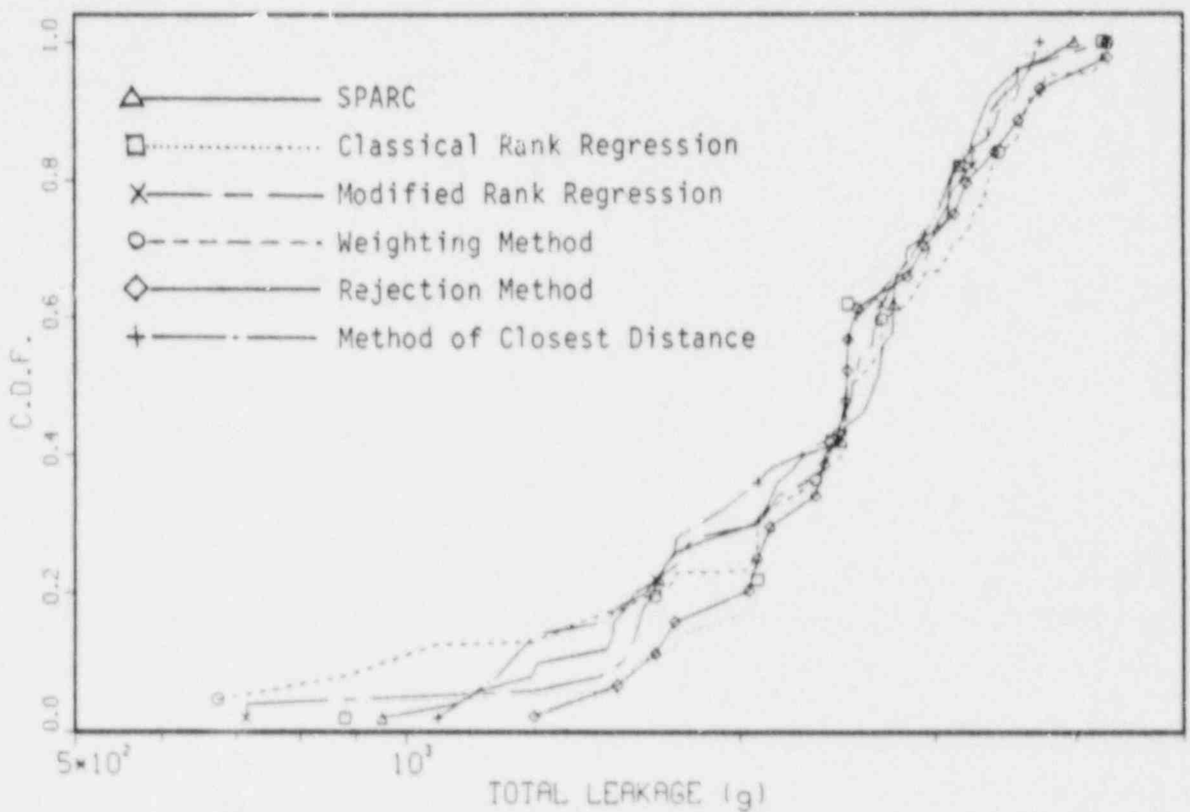

Figure 4.6 Comparison of Cumulative Distribution Functions (CDFs) for the Total Leakage of All Radionuclides Into Wetweil Atmosphere When the PDF for x_1 (RATIO) is Changed from Uniform to Normal Distribution

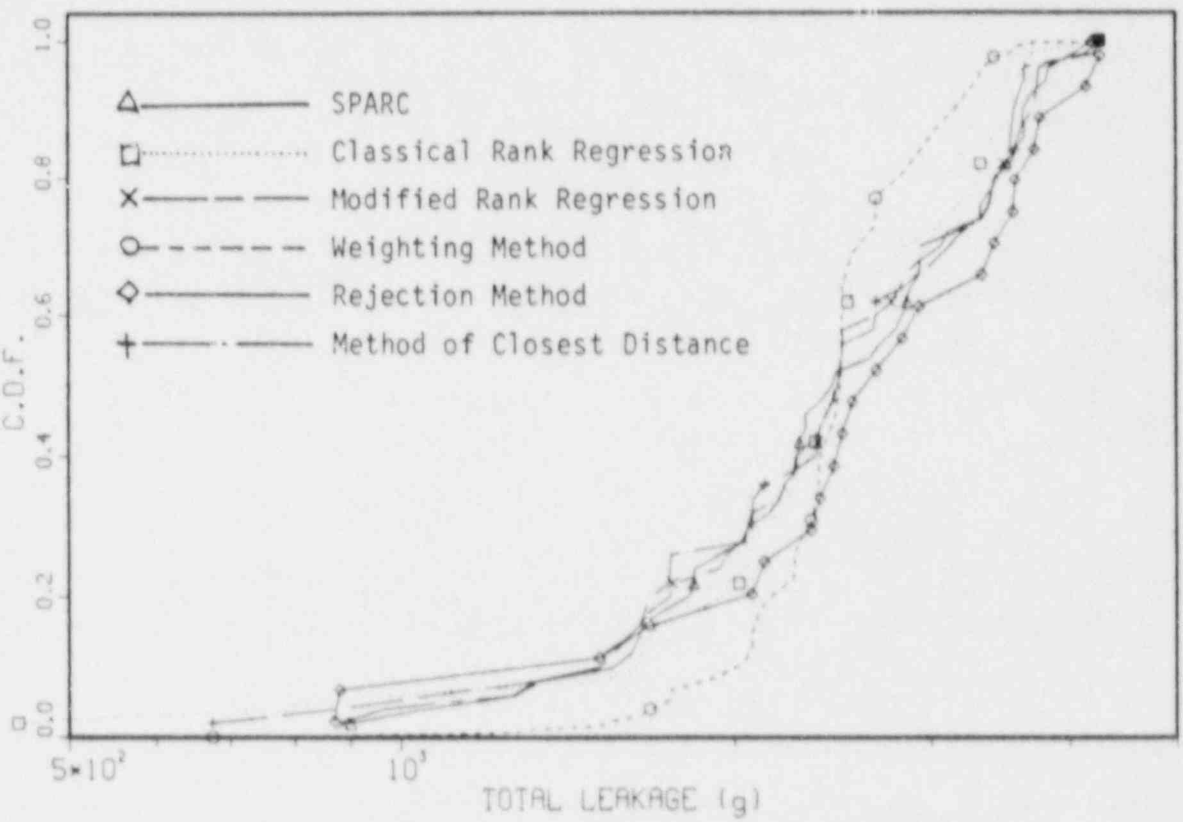

Figure 4.7 Comparison of Cumulative Distribution Functions (CDFs) for the Total Leakage of All Radionuclides Into Wetwell Atmosphere When the PDF for x_3 (VSWARM) is Changed from Uniform to Normal Distribution

5. SUMMARY AND CONCLUSIONS

Methods for assessment of the sensitivity of output probability distributions generated by computer models to the distributions assigned to the input parameters have been reviewed and modifications have been proposed. Through application to SPARC, a member code of the Source Term Code Package, the possibility of implementing these techniques within the context of the QUASAR program has been assessed. The methods considered include: (1) classical regression method, (2) modified regression method, (3) weighting method, (4) rejection method and (5) method of closest distance.

The conclusions can be summarized as follows:

a) The contrast amongst the results stemming from the five methods is not marked. In general, the modified rank regression technique performed better than the classical rank regression model while the method of closest distance outperformed the weighting method.

b) The performance of the two regression methods reviewed clearly would in general be dependent upon the degree to which the regression model adopted provides a good fit to the underlying computer model. Hence, in circumstances where the regression model provides a poor surrogate for the original computer model, then the weighting method and the rejection method might be expected to produce better approximations for the output distributions than would the regression methods. The method of closest distance would also be anticipated to display some reliance upon the goodness of the regression fit since the distance measure utilized therein incorporates regression-based information. However, one might expect the method of closest distance to provide a better surrogate for the computer model than the regression model in circumstances where the regression model would provide a poor fit since the method of closest distance does not explicitly resort to response surface techniques but instead relies upon outputs obtained by the original computer model.

c) It is not possible to draw general conclusions about the relative performance of the methods assessed regarding application to a range of computer models. Judicious application of each of the five approaches (i.e., the classical and modified regression approaches, the weighting method, the rejection method and the method of closest distance), weighed by a knowledge of the goodness of fit of the regression models formulated, will provide a basis for output distribution sensitivity analysis in the QUASAR program.

6. REFERENCES

1. J.A. Gieseke et al., "Source Term Code Package: User's Guide (MOD1)," NUREG/CR-4587, BMI-2138, July 1986.

2. 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, June 1986.

3. S.D. Unwin et al., "The Formulation of Probability Distributions for the QUASAR Program," BNL Technical Report A-3286, August 28, 1987.

4. R.P. Dickinson and R.J. Gelinas, "Sensitivity Analysis of Ordinary Differential Equation Systems - A Direct Method," J. of Computational Physics, Vol. 21, 123-143, 1976.

5. E.M. Oblow, "Sensitivity Theory for Reactor Thermal-Hydraulics Problems," ORNL/TM-6303, Oak Ridge National Laboratory, July 1978.

6. J. Hwang et al., "The Green's Function Method of Sensitivity Analysis in Chemical Kinetics," J. Chemical Physics, Vol. 69(11), 5180-5191, 1978.

7. R.I. Cukier et al., "Study of Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficient. I. Theory," J. of Chemical Physics, Vol. 59(8), 3873-3878.

8. T.H. Pierce and R.I. Cukier, "Global Nonlinear Sensitivity Analysis Using Walsh Functions," J. of Computational Physics, Vol. 41, 427-443, 1981.

9. R.H. Myers, "Response Surface Methodology," Allyn and Bacon, Boston, MA, 1975.

10. R.L. Iman et al., "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, March 1980.

11. R.J. Beckman and M.D. McKay, "Monte Carlo Estimation Under Different Distributions Using the Same Simulation," Technometrics, Vol. 29, 153-160, 1987.

12. M.D. McKay et al., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code," Technometrics, Vol. 21, 239-245, 1979.

13. 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, March 1984.

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

15. R.L. Iman et al., "Stepwise Regression With PRESS and Rank Regression (Program User's Guide)," SAND79-1472, January 1980.

16. R.L. Iman et al., "A FORTRAN 77 Program and User's Guide for the Calculation of Partial Correlation and Standardized Regression Coefficients," NUREG/CR-4122, SAND85-0044, June 1985.

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

NRC FORM 335 (2-84) NRCM 1102, 3201, 3202

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

SEE INSTRUCTIONS ON THE REVERSE

1. REPORT NUMBER (Assigned by TIDC, add Vol. No., if any)
NUREG/CR-4688
BNL-NUREG-52008
Volume 2

2. TITLE AND SUBTITLE
Quantification and Uncertainty Analysis of Source Terms for Severe Accidents in Light ter Reactors (QUASAR) Part II - Sensitivity Analysis .echniques

3. LEAVE BLANK

4. DATE REPORT COMPLETED
MONTH: September YEAR: 1987

5. AUTHOR(S)
T. Ishigami, E. Cazzoli, M. Khatib-Rahbar and S.D. Unwin

6. DATE REPORT ISSUED
MONTH: October YEAR: 1987

7. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Brookhaven National Laboratory
Upton, NY 11973

8. PROJECT/TASK/WORK UNIT NUMBER

9. FIN OR GRANT NUMBER
A-3286

10. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
USNRC
Office of Research
Washington, DC 20555

11a. TYPE OF REPORT

b. PERIOD COVERED (Inclusive dates)

12. SUPPLEMENTARY NOTES
M. Khatib-Rahbar, Principal Investigator, BNL; P. Wood, Project Manager, NRC

13. ABSTRACT (200 words or less)

Existing methods for sensitivity analysis are described and new techniques are proposed. These techniques are evaluated through consideration relative to the QUASAR program. Merits and limitations of the various approaches are examined by a detailed application to the Suppression Pool Aerosol Removal Code (SPARC).

14. DOCUMENT ANALYSIS - a. KEYWORDS/DESCRIPTORS
QUASAR
Source Term Uncertainty Analysis
Sensitivity and Importance Analysis
Latin Hypercube Sampling

b. IDENTIFIERS/OPEN-ENDED TERMS

15. AVAILABILITY STATEMENT

16. SECURITY CLASSIFICATION
(This page)
(This report)

17. NUMBER OF PAGES

18. PRICE

120555078877 1 1AN1R4
US NRC-OARM-ADM
DIV OF PUB SVCS
POLICY & PUB MGT BR-PDR NUREG
W-537
WASHINGTON DC 20555