Reliability Engineering and System Safety **42** (1993) 327–367

Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal

Jon C. Helton*

Department of Mathematics, Arizona State University, Tempe, Arizona 85287, USA

Uncertainty and sensitivity analysis techniques for use in performance assessments for radioactive waste disposal are reviewed. Summaries are given for the following techniques: differential analysis, Monte Carlo analysis, response surface methodology, and Fourier amplitude sensitivity test. Of these techniques, Monte Carlo analysis is felt to be the most widely applicable for use in performance assessment. Monte Carlo analysis involves five steps: (1) selection of a range and distribution for each input variable; (2) generation of a sample from the input variables; (3) propagation of the sample through the model under consideration; (4) performance of uncertainty analysis; and (5) performance of sensitivity analysis. These steps are discussed and illustrated with an analysis performed as part of a preliminary performance assessment for the Waste Isolation Pilot Plant (WIPP).

1 INTRODUCTION

Performance assessment for radioactive waste disposal is a complex process and involves significant uncertainties.[1–4] The need to assess the impact of these uncertainties is now widely recognized.[5–10] As discussed by many authors, two basic types of uncertainty are present in performance assessments: stochastic uncertainty and subjective uncertainty.[11–16] Stochastic uncertainty results from the fact that the system under consideration can behave in many different ways (e.g. a variety of disruptive events might be possible at a waste disposal site). Subjective uncertainty results from the existing state of knowledge with respect to the correctness of the assumptions used in an analysis (e.g. what is the appropriate value to use for a spatially-averaged distribution coefficient over a specified region, or should radionuclide transport by flowing groundwater be represented with a single-porosity or a dual-porosity model at a given waste disposal site). Thus, stochastic uncertainty is a property of the system under study, while subjective uncertainty is a property of the analysts performing the study. However, these two types of uncertainty cannot be entirely separated. For example, there can be subjective uncertainty in a quantity used to characterize stochastic uncertainty (e.g. a Poisson process may be used to characterize the occurrence of drilling at a waste disposal site, which is a stochastic uncertainty, and the appropriate value to use for the rate constant λ associated with this process may be poorly known, which is a subjective uncertainty).

The primary focus of this paper is the analysis of subjective uncertainty, although the combined analysis of stochastic and subjective uncertainty will be discussed at the end of the paper. Thus, unless indicated otherwise, the discussions in this paper will involve assessments of subjective uncertainty. Such assessments are often divided into the closely related areas of uncertainty analysis and sensitivity analysis. Uncertainty analysis involves determining the uncertainty in model predictions that results from imprecisely known input variables, and sensitivity analysis involves determining the contribution of individual input variables to the uncertainty in model predictions. The input variables considered in most uncertainty/sensitivity analyses are real-valued quantities (e.g. solubilities, retardations, . . .), although input variables can also be used to represent more complex entities such as flow fields or alternative models.

One informal and four formal approaches have

* Department 6342, WIPP, Sandia National Laboratories, Albuquerque, NM 87185, USA

Reliability Engineering and System Safety 0951-8320/93/$06.00

been developed for performing uncertainty and sensitivity studies. The informal approach involves varying one parameter or set of assumptions at a time and observing the deviation in the resultant model prediction from a base-case prediction. This is sometimes designated the *ceteris paribus* approach. The formal approaches are differential analysis, Monte Carlo analysis, response surface methodology, and Fourier amplitude sensitivity test (FAST). These approaches are based on Taylor series, random sampling, response surface construction, and Fourier series, respectively. Although the implementation of these approaches is more complex than a *ceteris paribus* analysis, they often yield more information with less computational effort.

This review provides a description of the four formal approaches to uncertainty and sensitivity analysis. Of these approaches, Monte Carlo analysis is felt to be the most widely applicable to problems that arise in performance assessment for radioactive waste disposal. Therefore, after high-level descriptions are given for each of the approaches, a more detailed description will be given for the performance of a Monte Carlo analysis. This description will be illustrated with results obtained in uncertainty/sensitivity studies performed as part of a preliminary performance assessment for the Waste Isolation Pilot Plant (WIPP).[1,5]

2 REVIEW OF TECHNIQUES

For convenience in the following discussion, it is assumed that the model under consideration can be represented by a function of the form

$$y = f(x_1, x_2, \ldots, x_n) = f(\mathbf{x}) \tag{1}$$

where y is the model prediction (i.e. the output variable) and $x_1, x_2, \ldots, x_n$ are the model inputs (i.e. the input variables). The output variable y is assumed to be single-valued to keep the notation from becoming unwieldy although, in practice, there is usually more than one predicted variable of interest.

The expression in (1) is, in essence, a representation for a model at the level of the computer program that implements it. The mathematical model evaluated within the program often has a more complex structure. For example, many models used in radioactive waste management are based on an equation of the form

$$\mathbf{g}(t, x, y, z, \mathbf{y}, \mathbf{x}) = \mathbf{h}(t, x, y, z, \mathbf{x}) \tag{2}$$

where

t = temporal variable,
x, y, z = spatial variables,
$\mathbf{y}$ = $[y_1, y_2, \ldots, y_{nOV}]$,
y_i = output variable i,
nOV = number of output variables,
$\mathbf{x}$ = $[x_1, x_2, \ldots, x_{nIV}]$,
x_j = input variable j,
nIV = number of input variables,

and the functions $\mathbf{g}$ and $\mathbf{h}$ are vector-valued. The equation in (2) is often an ordinary or partial differential equation, in which case $\mathbf{g}$ is a differential operator and there are also initial and boundary value conditions that depend on t, x, y, z and $\mathbf{x}$. However, when viewed as a 'black box', a program that evaluates such an equation has the form shown in (1). In practice, the expression in (1) is often a sequence of linked models in which each model receives input generated by one or more preceding models and produces output which then becomes input to one or more following models.

The implementation of most uncertainty and sensitivity analysis procedures is independent of the form of the model shown in (2), in which case the model representation in (1) is entirely consistent with the use of these procedures. However, the implementation of some specialized techniques that can be used in conjunction with differential analysis is closely tied to the specific form of the equation in (2). When these techniques are used, it is necessary to know and work with this equation. Further, specialized strategies may be needed for analyses that involve complex sequences of linked models.

2.1 Differential analysis

Differential analysis is based on using a Taylor series to approximate the model under consideration. Once constructed, this series is used as a surrogate for the original model in uncertainty and sensitivity studies. A differential analysis involves four steps.

In the first step, base values, ranges and distributions are selected for the input variables x_j, $j = 1, \ldots, n$, appearing in (1). The base values can be represented by the vector

$$\mathbf{x}_0 = [x_{10}, x_{20}, \ldots, x_{n0}] \tag{3}$$

and form the point about which the Taylor series will be developed. The ranges and distributions will be used later in the analysis process when the effects on the output variable y of perturbations away from $\mathbf{x}_0$ are investigated.

In the second step, a Taylor series approximation to y is developed. If the series is restricted to first-order terms, the approximation has the form

$$y(\mathbf{x}) \doteq y(\mathbf{x}_0) + \sum_{j=1}^{n} \left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right](x_j - x_{j0}) \tag{4}$$

If first- and second-order terms are used, the

approximation becomes

$$y(\mathbf{x}) \doteq y(\mathbf{x}_0) + \sum_{j=1}^{n} \left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right](x_j - x_{j0}) + \tfrac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\left[\frac{\partial^2 f(\mathbf{x}_0)}{\partial x_j\,\partial x_k}\right](x_j - x_{j0})(x_k - x_{k0}) \quad (5)$$

Higher-order expansions are also possible if f has the necessary partial derivatives. In concept, the order of the approximation should be determined by the curvature of the surface $y = f(\mathbf{x})$. In practice, determination of an appropriate order for the approximation may be difficult. Evaluation of the partial derivatives in (4) and (5) is usually the most demanding part of a differential analysis.

In the third step, variance propagation techniques are used to estimate the uncertainty in y. If the approximation in (4) is used, the expected value and variance of y can be estimated by

$$E(y) \doteq y(\mathbf{x}_0) + \sum_{j=1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]E(x_j - x_{j0}) = y(\mathbf{x}_0) \quad (6)$$

and

$$V(y) \doteq \sum_{j=1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]^2 V(x_j) + 2\sum_{j=1}^{n}\sum_{k=j+1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_k}\right]Cov(x_j, x_k) \quad (7)$$

respectively, where E, V and Cov denote expected value, variance and covariance, respectively, and x_{j0} is assumed to equal $E(x_j)$. If the x_j are uncorrelated, the variance approximation in (7) reduces to

$$V(y) \doteq \sum_{j=1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]^2 V(x_j) \quad (8)$$

If the approximation in (5) is used, the x_j are uncorrelated, and fourth-order and higher terms are ignored in the derivation of $V(y)$, the estimates for expected value and variance are

$$E(y) \doteq y(\mathbf{x}_0) + \tfrac{1}{2}\sum_{j=1}^{n}\left[\frac{\partial^2 f(\mathbf{x}_0)}{\partial x_j^2}\right]V(x_j) \quad (9)$$

and

$$V(y) \doteq \sum_{j=1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]^2 V(x_j) + \sum_{j=1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]\left[\frac{\partial^2 f(\mathbf{x}_0)}{\partial x_j^2}\right]\mu_3(x_j) \quad (10)$$

where $\mu_3(x_j)$ denotes the third central moment for x_j. As correlations between the x_j and higher-order terms are included, approximations to the expected value and variance for y rapidly become very complicated.[17–20] Together, $E(y)$ and $V(y)$ provide a representation for the uncertainty in y that results from the uncertainty in the x_j. It is also possible to obtain a representation for the uncertainty in y by using the approximation in (4) or (5) and letting the x_j vary over their ranges.

In the fourth and final step, the Taylor series approximations are used to estimate the importance of the individual x_j. Fractional contribution to variance can be estimated from the expressions in (8) and (10) and provides one means of measuring variable importance. If the representation for y in (4) is used, then the contribution of x_j to the variance of y can be estimated with the ratio

$$\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]^2 V(x_j)/V(y) \quad (11)$$

where $V(y)$ is defined in (8). Similarly, if the representation for y in (5) is used, then the contribution of x_j to the variance of y can be estimated with the ratio

$$\left\{\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]^2 V(x_j) + \left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\right]\left[\frac{\partial^2 f(\mathbf{x}_0)}{\partial x_j^2}\right]\mu_3(x_j)\right\}/V(y) \quad (12)$$

where $V(y)$ is defined in (10). The fractional contributions to variance shown in (11) and (12) can be used to order the x_j with respect to their contribution to the uncertainty in y. This ordering involves both the absolute effect of the x_j, as measured with their partial derivatives, and the effect of distributions assigned to the x_j, as measured by $V(x_j)$ and $\mu_3(x_j)$.

Another way of measuring variable importance is by normalizing the partial derivatives in the first order Taylor series approximation to y appearing in (4). Specifically, this approximation can be rewritten in the two following forms:

$$\frac{y(\mathbf{x}) - y(\mathbf{x}_0)}{y(\mathbf{x}_0)} \doteq \sum_{j=1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\frac{x_{j0}}{y(\mathbf{x}_0)}\right]\left[\frac{x_j - x_{j0}}{x_{j0}}\right] \quad (13)$$

and

$$\frac{y(\mathbf{x}) - y(\mathbf{x}_0)}{SD(y)} \doteq \sum_{j=1}^{n}\left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\frac{SD(x_j)}{SD(y)}\right]\left[\frac{x_j - x_{j0}}{SD(x_j)}\right] \quad (14)$$

where SD is used to denote standard deviation and it is assumed that there is no problem with division by zero. The relation in (8) is used to provide an estimate for $SD(y)$. The normalized coefficients

$$\frac{\partial f(\mathbf{x}_0)}{\partial x_j}\frac{x_{j0}}{y(\mathbf{x}_0)} \quad \text{and} \quad \frac{\partial f(\mathbf{x}_0)}{\partial x_j}\frac{SD(x_j)}{SD(y)} \quad (15)$$

provide related, though different, measures of the impact of the individual x_j on y. The coefficient from (13) measures the effect on y that results from perturbing x_j by a fixed fraction of its base value x_{j0}. The coefficient from (14) measures the effect on y that results from perturbing x_j by a fixed fraction of its

standard deviation. Further, the sign of the coefficient indicates whether x_j and y tend to move up and down together or in opposite directions. The absolute values of these coefficients can be used to rank the relative importance of the individual x_j. A ranking based on the coefficients from (14) includes the effects of distribution assumptions while a ranking based on the coefficients from (13) does not.

As noted earlier, the largest effort in a differential analysis is usually the determination of the partial derivatives appearing in (4) and (5). Because of this, a number of specialized techniques have been developed to facilitate the calculation of these derivatives. Included in these are adjoint techniques,[21-23] Green's function techniques,[24-26] and the GRESS/ADGEN compiler.[27-30] Adjoint and Green's function techniques can provide significant computational savings in the calculation of partial derivatives but can be very tedious and time-consuming to implement because they require detailed manipulations involving the model equations shown in (2). The GRESS/ADGEN compiler was developed to avoid the human cost associated with these manipulations by performing the necessary manipulations directly from the FORTRAN code that implements the model. However, the basic ideas in differential analysis outlined in this section remain the same regardless of the procedure used to calculate the necessary partial derivatives.

A differential analysis can also use multiple Taylor series. In this approach, series of the form appearing in (4) and (5) are constructed at a number of different base-case values $\mathbf{x}_0$. For example, initial selections for the $\mathbf{x}_0$ and their associated Taylor series can be used to provide guidance on the selection of additional points at which to develop expansions. Then the individual series can be pieced together to provide a better representation for y than any of the individual series by themselves could.

Additional information on differential analysis is available in a number of references.[31-36]

2.2 Monte Carlo analysis

A Monte Carlo analysis is based on performing multiple model evaluations with probabilistically selected model input, and then using the results of these evaluations to determine both the uncertainty in model predictions and the input variables that give rise to this uncertainty. For Monte Carlo analyses, the model can be assumed to be a function of the form shown in (1). It is usually not necessary to consider the details of the model, although this helps sometimes in planning strategies for the analysis of very complex models. In general, a Monte Carlo analysis involves five steps.

In the first step, a range and distribution are selected for each x_j. These selections will be used in the next step in the generation of a sample from the x_j. If the analysis is primarily of an exploratory nature, then rather crude (e.g. uniform, loguniform, triangular) distribution assumptions may be adequate. However, if precise uncertainty results are desired for y, then corresponding care must be used in specifying the distributions for the x_j.

In the second step, a sample is generated from the ranges and distributions specified in the first step. The result of this step is a sequence of sample elements of the form

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}], \quad i = 1, 2, \ldots, m \tag{16}$$

where n is the number of input (i.e. sampled) variables and m is the sample size. The most widely used sampling techniques are random sampling, importance sampling and Latin hypercube sampling.[37]

In the third step, the model is evaluated for each sample element shown in (16). This creates a sequence of results of the form

$$y_i = f(x_{i1}, x_{i2}, \ldots, x_{in}) = f(\mathbf{x}_i), \quad i = 1, 2, \ldots, m \tag{17}$$

In essence, these model evaluations create a mapping from the analysis inputs (i.e. the $\mathbf{x}_i$) to the analysis results (i.e. the y_i) that can be studied in subsequent uncertainty and sensitivity analyses.

In the fourth step, the results shown in (17) are used as the basis for an uncertainty analysis. One way to represent the uncertainty in y is with a mean value and a variance. When random sampling or Latin hypercube sampling is used to generate the sample shown in (16), the expected value and variance for y can be estimated by

$$E(y) \doteq \sum_{i=1}^{m} y_i/m \tag{18}$$

and

$$V(y) \doteq \sum_{i=1}^{m} [y_i - E(y)]^2/(m-1) \tag{19}$$

respectively. Use of expected value and variance to characterize uncertainty reduces all of the information in (17) about the variability in y to two numbers. Clearly, information is lost in this process. Another way to summarize the variability in y is through the use of an estimated distribution function. In particular, this function is given by the step function F defined by

$$F(y) = \begin{cases} 0 & \text{if } y < y_1 \\ i/m & \text{if } y_i \le y < y_{i+1}, \quad i = 1, 2, \ldots, m-1 \\ 1 & \text{if } y_m \le y \end{cases} \tag{20}$$

where it is assumed that the y_i have been ordered so

that $y_i \leq y_{i+1}$. This creates a plot that displays all the information contained in (17) about the uncertainty in y. A very important aspect of the uncertainty studies that can be performed as part of a Monte Carlo analysis is that it is not necessary to introduce a surrogate or intermediate model to obtain the results in (18), (19) and (20). In contrast, both differential analysis and response surface methodology require the construction of an intermediate model before uncertainty analysis results can be obtained.

The fifth and final step is sensitivity analysis, which is based on an exploration of the mapping from analysis input to analysis results defined by the relationship in (17). Many techniques are available for this exploration. One of the simplest but also most useful is the generation of scatterplots. A scatterplot for input variable x_j and the output variable y is a plot of the points

$$(x_{ij}, y_i), \quad i = 1, 2, \ldots, m \tag{21}$$

Such plots often reveal thresholds or nonlinearities in the relationship between x_j and y. Another useful procedure is stepwise regression analysis. In this procedure, a regression model relating the x_j to y is constructed by bringing in one variable at a time. Variable importance is indicated by the order in which variables enter the model, the size and sign of the standardized regression coefficients, and the changes in R^2 values as additional variables enter the model. The predictor variables in the regression model do not have to be the original x_j; analysis results are often improved by using squares (i.e. x_j^2), cross products (i.e. $x_j x_k$), logarithms (i.e. $\ln x_j$) or other transformations as predictor variables. In many analyses, model predictions are not single-valued as shown in (1); rather, multiple values are produced due to temporal or spatial variation. When this is the case, plots of standardized regression coefficients or partial correlation coefficients as functions of time or location may be revealing.

Additional information on Monte Carlo analysis is available elsewhere.[38–47]

2.3 Response surface methodology

This procedure is based on developing a response surface approximation to the model under consideration. This approximation is then used as a surrogate for the original model in subsequent uncertainty and sensitivity analyses. For convenience, it is again assumed that the model under consideration is of the form shown in (1). A response surface methodology analysis for this model involves six steps.

In the first step, a range and distribution must be selected for each x_j. At the minimum, the ranges for the x_j must be decided on since this will impact the selection of design points in the next step. However, in a later step, uncertainty results will be obtained by a Monte Carlo simulation of a response surface approximation to the original model. At this point, distributions for the x_j are required.

In the second step, an experimental design is developed that defines the combinations of variable values for which model evaluations will be performed. Ultimately, these model evaluations will be used in the construction of a response surface replacement for the original model. Many different types of experimental designs are available, such as factorial, fractional factorial, Plackett–Burman, and central composite.[48–50] Designs developed specifically for computer experimentation are discussed by Sacks *et al.*[51,52] The design ultimately selected will depend on many factors, including the number of independent variables under consideration, the possible presence of quadratic or higher order effects, the possible importance of variable interactions, and the computational expense of evaluating the model. The final result of this step is a sequence of design points of the form

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}], \quad i = 1, 2, \ldots, m \tag{22}$$

where n is the number of input variables and m is the number of points in the design. In our notation, the design points in (22) look exactly like the sample elements shown in (16) for a Monte Carlo study. However, there is an important distinction. The design points in (22) are selected with a classical experimental design, which assures that a specified structure exists between the values of the individual x_j but allows no probabilistic weight to be assigned to the $\mathbf{x}_i$. In contrast, the sample elements in (16) are selected with a probabilistic procedure, and it is possible to assign a weight (e.g. $1/m$ in the case of random and Latin hypercube sampling) to each element for use in the construction of estimated means, variances and distribution functions. Some of the desirable properties of classical experimental designs and random or Latin hypercube sampling can be combined through the use of the restricted pairing technique developed by Iman & Conover for inducing a specified correlation structure within a sample.[53]

In the third step, the model is evaluated for each design point shown in (22). This creates a sequence of results of the form

$$y_i = f(x_{i1}, x_{i2}, \ldots, x_{in}) = f(\mathbf{x}_i), \quad i = 1, 2, \ldots, m \tag{23}$$

In the fourth step, the results shown in (23) are used in the construction of a response surface approximation to the original model. In most cases, this construction is based on least squares techniques.[48–50] If only first order terms are considered, the resulting response surface approximation

is of the form

$$y \doteq b_0 + \sum_{j=1}^{n} b_j x_j \tag{24}$$

although response surfaces involving quadratic terms, cross-products and other terms are possible. The exact types of response surfaces that are possible depend on the properties of the experimental design used to select the points shown in (22). The response surface shown in (24) plays the same role in a response surface methodology analysis that the Taylor series shown in (4) and (5) play in a differential analysis. In addition to procedures based on least squares, response surfaces can also be constructed with polynomial fits[54,55] and splines.[56]

In the fifth step, the approximation to y shown in (24) is used to estimate the uncertainty in y that results from the ranges and distributions assigned to the x_j in the first step. Two different approaches can be used. In the first approach, the expected value and variance of y are estimated from

$$E(y) \doteq b_0 + \sum_{j=1}^{n} b_j E(x_j) \tag{25}$$

and

$$V(y) \doteq \sum_{j=1}^{n} b_j^2 V(x_j) + 2 \sum_{j=1}^{n} \sum_{k=j+1}^{n} b_j b_k \, Cov(x_j, x_k) \tag{26}$$

respectively. If the x_j are uncorrelated, the variance approximation in (26) reduces to

$$V(y) \doteq \sum_{j=1}^{n} b_j^2 V(x_j) \tag{27}$$

When a more complex response surface approximation to y is used, then correspondingly more complex expressions for the expected value and variance of y result. Conceptually, the relations in (25), (26) and (27) are the same as the expected value and variance formulas shown in (6), (7) and (8) for use with a Taylor series as part of a differential analysis. In the second approach, the distribution function for y is estimated by performing a Monte Carlo simulation with the expression in (24). Since the cost of evaluating the approximation to y in (24) is small, a large sample size can be used. The result is a very good estimate for the distribution function associated with the response surface approximation to y in (24). However, whether or not this also provides a good approximation to the true distribution function for y depends on how well the response surface in (24) approximates y.

In the sixth and final step, the approximation to y in (24) is used to assess the sensitivity of y to the individual x_j. This assessment is obtained by normalizing the coefficients in (24) so that the effects of changes in the x_j are apparent. The two most useful normalizations are

$$\frac{y(\mathbf{x}) - E(y)}{E(y)} \doteq \sum_{j=1}^{n} \left[\frac{b_j E(x_j)}{E(y)}\right]\left[\frac{x_j - E(x_j)}{E(x_j)}\right] \tag{28}$$

and

$$\frac{y(\mathbf{x}) - E(y)}{SD(y)} \doteq \sum_{j=1}^{n} \left[\frac{b_j SD(x_j)}{SD(y)}\right]\left[\frac{x_j - E(x_j)}{SD(x_j)}\right] \tag{29}$$

where E and SD denote expected value and standard deviation, respectively. The standard deviation for y can be estimated from the relationship in (27). The normalized coefficients

$$\frac{b_j E(x_j)}{E(y)} \quad \text{and} \quad \frac{b_j SD(x_j)}{SD(y)} \tag{30}$$

appearing in (28) and (29), respectively, are analogous to the normalized coefficients appearing in (13) and (14) for differential analysis. The coefficient $b_j E(x_j)/E(y)$ indicates the importance of the individual x_j with respect to equal-sized perturbations from their expected values. The coefficient $b_j SD(x_j)/SD(y)$ indicates the importance of the individual x_j with respect to perturbations from their expected values that are equal to fixed fractions of their standard deviations. This coefficient is the same as a standardized regression coefficient.

Additional background information on response surface methodology is available elsewhere.[48–50,57–60]

2.4 Fourier Amplitude Sensitivity Test (FAST)

The FAST approach is based on performing a numerical calculation to obtain the expected value and variance of a model prediction. The basis of this calculation is a transformation that converts a multidimensional integral over all the uncertain model inputs to a one-dimensional integral. Further, a decomposition of the Fourier series representation of the model is used to obtain the fractional contribution of the individual input variables to the variance of the model prediction. As before, assume that the model under consideration can be represented in the form shown in (1). An analysis of this model based on the FAST approach involves four steps.

In the first step, ranges and distributions are constructed for the x_j. The final outcome of this step is a density function p for the variable $\mathbf{x}$ on a domain Ω, where

$$\mathbf{x} = [x_1, x_2, \ldots, x_n] \tag{31}$$

As a reminder, Ω is simply the set of all possible values that $\mathbf{x}$ can take on, and p is the characterization of the distribution of $\mathbf{x}$ over these values.

Once p and Ω are known, it is possible to state the expressions for the expected value and variance of y

formally. Specifically,

$$E(y) = \int_{\Omega} f(\mathbf{x})p(\mathbf{x})\, d\mathbf{x} \tag{32}$$

and

$$V(y) = \int_{\Omega} [f(\mathbf{x}) - E(y)]^2 p(\mathbf{x})\, d\mathbf{x} \tag{33}$$

where f is the defining function for y shown in (1). The preceding integrals are multidimensional and, in practice, are very difficult to evaluate. This leads to the second step in the approach.

In the second step, the multidimensional integral appearing in (32) is converted to a one-dimensional integral by the definition of a suitable space-filling curve in Ω. Construction of this curve is based on the density function p, which characterizes the distribution of $\mathbf{x}$, and requires the solution of a nonlinear differential equation for each independent variable x_j. The ultimate outcome of this construction process is a sequence of functions G_j, $j = 1, 2, \ldots, n$, and a corresponding sequence of integers ω_j, $j = 1, 2, \ldots, n$, such that

$$E(y) = \int_{\Omega} f(\mathbf{x})p(\mathbf{x})\, d\mathbf{x} \doteq \frac{1}{2\pi}\int_{-\pi}^{\pi} f[G_1(\sin \omega_1 s), G_2(\sin \omega_2 s), \ldots, G_n(\sin \omega_n s)]\, ds \tag{34}$$

The second integral, which is one-dimensional, is often easier to evaluate than the first integral, which is multidimensional. In essence, construction of the G_j and ω_j has converted f from a function of n variables (i.e. the elements x_j of $\mathbf{x}$) to a function of one variable (i.e. s).

In the third step, the one-dimensional representation for f is used to estimate the expected value and variance of y. This provides a representation of the uncertainty in y that results from the specified distributions for the x_j. As already indicated in (34), the expected value for y can be approximated by

$$E(y) \doteq \frac{1}{2\pi}\int_{-\pi}^{\pi} f[G_1(\sin \omega_1 s), G_2(\sin \omega_2 s), \ldots, G_n(\sin \omega_n s)]\, ds \tag{35}$$

Further, the variance of y can be approximated by

$$V(y) \doteq \frac{1}{2\pi}\int_{-\pi}^{\pi} \{f[G_1(\sin \omega_1 s), G_2(\sin \omega_2 s), \ldots, G_n(\sin \omega_n s)]\}^2\, ds - E^2(y) \tag{36}$$

In practice, a numerical procedure must be written to evaluate the integrals in the approximations to $E(y)$ and $V(y)$. Further, this procedure will require values for f, which must come from evaluating the model that predicts y.

In the fourth and final step, the sensitivity of y to the individual x_j is determined on the basis of fractional contribution to variance. By using properties of the Fourier series representation for f, it can be shown that

$$V(y) \doteq \sum_{k=1}^{\infty} (A_k^2 + B_k^2) \tag{37}$$

where

$$A_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f[G_1(\sin \omega_1 s), G_2(\sin \omega_2 s), \ldots, G_n(\sin \omega_n s)] \cos(ks)\, ds \tag{38}$$

and

$$B_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f[G_1(\sin \omega_1 s), G_2(\sin \omega_2 s), \ldots, G_n(\sin \omega_n s)] \sin(ks)\, ds \tag{39}$$

Further, it can also be shown that the part of the variance of y that derives from x_j can be approximated by

$$V_j(y) \doteq \sum_{k=1}^{\infty} (A_{k\omega_j}^2 + B_{k\omega_j}^2) \tag{40}$$

where ω_j is the integer associated with G_j in the conversion from a multidimensional integral to a one-dimensional integral shown in (34). Then, the ratio

$$V_j(y)/V(y) \tag{41}$$

gives the fractional contribution of x_j to the variance of y. This ratio can be used to rank the importance of the individual x_j with respect to their impact on the uncertainty in y. In practice, the integrals that define the A_k and B_k must be approximated numerically. Further, the series in (40) must be truncated.

Additional background and information on the FAST approach is available elsewhere.[61–65]

2.5 Relative merits of individual techniques

Differential analysis is based on developing a Taylor series approximation to the model under consideration. Ultimately, the quality of the analysis results will depend on how well this series approximates the original model. Desirable properties of differential analysis include the following: (1) the effects of small perturbations away from the base-case value about which the Taylor series was developed are revealed; (2) uncertainty and sensitivity analyses are straightforward once the Taylor series is developed; (3)

specialized techniques (e.g. adjoint, Green's function, GRESS/ADGEN) exist to facilitate the calculation of derivatives; and (4) the approach has been widely studied and applied.

However, there are two important drawbacks to differential analysis that should always be considered when selecting the procedure to be used in an uncertainty/sensitivity study. First, differential analysis is inherently local. The farther a perturbation moves from the base-case value about which the Taylor series was constructed, the less reliable the analysis results become. In particular, differential analysis is a poor choice for use in estimating distribution functions and provides no information on the possible existence of thresholds or discontinuities in the relationships between input and output variables. Overall, the more nonlinear the relationships between the input and output variables, the more difficult it is to employ a differential analysis effectively. A partial solution to these problems is to construct Taylor series at a number of points in the input space and then assemble the individual series into a single model. Second, differential analyses can be very difficult to implement and often require large amounts of human and/or computer time. This difficulty arises from the need to calculate the partial derivatives required in the Taylor series. The possible use of sophisticated techniques such as the GRESS/ADGEN procedures offers some encouragement in this area. Even so, the need to calculate the required derivatives should not be taken lightly.

Monte Carlo analysis is based on the use of a probabilistic procedure to select model input. Then, uncertainty analysis results are obtained directly from model predictions without the use of an intermediate surrogate model, and sensitivity analysis results are obtained by exploring the mapping from model input to model predictions that formed the basis for the uncertainty analysis. Desirable properties of Monte Carlo analysis include the following: (1) the full range of each input variable is sampled and subsequently used as model input; (2) uncertainty results are obtained without the use of a surrogate model; (3) extensive modifications to the original model are not necessary (such modifications are often required when adjoint or Green's function techniques are used as part of a differential analysis); (4) the full stratification over the range of each input variable facilitates the identification of nonlinearities, thresholds and discontinuities; (5) a variety of regression-based sensitivity analysis techniques are available; and (6) the approach is conceptually simple, widely used, and easy to explain.

Two particularly appealing features of Monte Carlo analysis are the full coverage of the range of each input variable and the ease with which an analysis can be implemented. Full stratification or coverage results from the fact that sampling procedures such as Latin hypercube sampling, which is discussed in Section 3.2, divide the range of each input variable into a number of intervals and then force the selection of a value from each interval, thus providing a full coverage of the range of each input variable. This feature is particularly important when the input variables have large ranges and the existence of nonlinear relationships between the input and output variables is a possibility. With respect to the second feature, essentially any variable that can be supplied as an input or generated as an output can be included in a Monte Carlo analysis without any modification to the original model.

The major drawback to Monte Carlo procedures is the fact that multiple model evaluations are required. If the model is computationally expensive to evaluate or many model evaluations are required, then the cost of the required calculations may be large. Computational cost should always be considered when selecting a technique, but it is rarely the dominant cost in performing an analysis. Special techniques such as Latin hypercube sampling and importance sampling can often be used to reduce the number of required model evaluations without compromising the overall quality of an analysis. Further, it is important to recognize that, in practice, the other analysis techniques discussed in this section can require as much computational time as Monte Carlo analysis.

Response surface methodology is based on constructing a response surface approximation to the original model. This approximation is then used as a surrogate for the original model in subsequent uncertainty and sensitivity studies. Desirable properties of response surface methodology include the following: (1) complete control over the structure of model input through the experimental design selected for use; (2) near optimum choice for a model whose predictions are known to be a linear or quadratic function of the input variables; and (3) uncertainty and sensitivity analyses that are inexpensive and straightforward once the necessary response surface approximation has been constructed. Further, the development of experimental designs has been widely studied, although typically for situations that are considerably less involved than those encountered in performing an uncertainty/sensitivity study for a complex model.

There are also several drawbacks to response surface methodology that should be considered when an approach to uncertainty/sensitivity analysis is being selected. These include the following: (1) difficulty in development of an appropriate experimental design because of many input variables, many output variables, unknown form for the model, or

spatial/temporal variability; (2) use of few values for each input variable; (3) possible requirement of many design points; (4) difficulties in detecting thresholds, discontinuities and nonlinearities; (5) difficulties in including correlations and restrictions between input variables; and (6) difficulty in construction of an appropriate response surface approximation to the original model, which may require a considerable amount of statistical sophistication and/or artistry. Ultimately, the final uncertainty/sensitivity results are no better than the response surface approximation to the original model. Response surface methodology will work when there are only a few (typically, less than 10) input variables, a limited number of distinct output variables (because a design that is appropriate for one output variable may not be appropriate for a different output variable), and the relationships between the input and output variables are basically linear or quadratic or involve a few cross-products. Otherwise, the structure of the input–output relationships is too complicated to be captured by a classical experimental design (or a sequence of designs if a sequential approach is being used) in an efficient manner.

The FAST approach is based on performing a numerical calculation to estimate expected value and variance. Further, sensitivity results are obtained by decomposing the variance estimate into the variances due to the individual input variables. Desirable properties of the FAST approach include the following: (1) full range of each input variable is covered; (2) estimation of expected value and variance is by a direct calculation rather than by use of a surrogate model; and (3) modifications to the original model are not required.

There are also several drawbacks to using the FAST approach. These include the following: (1) the underlying mathematics is complicated and difficult to explain; (2) the approach is not widely known or used; (3) developing the necessary space-filling curve and performing the numerical integration over this curve to obtain expected value and variance is complicated; (4) many model evaluations may be required; (5) an estimate for the cumulative distribution function of the output variable is not provided; and (6) it is not possible to specify correlations or other types of restrictions between variables. Fortunately, software has been developed to facilitate the implementation of an uncertainty/sensitivity study based on the FAST approach.[50] As analyses are currently performed with the FAST approach, no information on discontinuities, thresholds or nonlinearities is obtained. However, it is probably possible to investigate this type of behavior with the model evaluations that must be performed in the numerical integrations to obtain expected value and variance.

2.6 Monte Carlo as a preferred approach

Each approach to uncertainty and sensitivity analysis has its advantages and disadvantages, and all approaches have been successfully applied. It would be a mistake to state categorically that one approach will always be superior to the others regardless of the model under consideration. For a given analysis problem, the available approaches should be considered, and the approach that seems most appropriate for the problem should be selected. This selection should take into account the nature of the model, the type of uncertainty and sensitivity analysis results desired, the cost of modifying and/or evaluating the model, the human cost associated with mastering and implementing a technique, the time period over which an analysis must be performed, and the programmatic risk associated with unanticipated complications in the implementation of a technique.

The comments of the preceding paragraph notwithstanding, it is felt that Monte Carlo techniques provide the best overall approach for studying problems related to performance assessment for radioactive waste disposal. This statement is made for several reasons.

First, there are often large uncertainties in such problems. Due to full stratification over the range of each variable, Monte Carlo techniques are particularly appropriate for analysis problems in which large uncertainties are associated with the input variables. In particular, differential analysis and response surface methodology are likely to perform poorly when the relationships between the input and output variables are nonlinear and the input variables have large uncertainties.

Second, it is often necessary to estimate distribution functions (e.g. for comparison with the standard proposed by the U.S. Environmental Protection Agency).[66,67] Monte Carlo techniques provide direct estimates for distribution functions. Neither differential analysis nor the FAST approach is intended for the estimation of distribution functions. The estimates obtained with response surface methodology are no better than the response surface approximation to the original model. It should be possible to estimate distribution functions with results generated as part of the FAST approach, but this possibility apparently has not been investigated and applied.

Third, Monte Carlo techniques do not require a large amount of sophistication that goes beyond the analysis problem of interest. In contrast, differential analysis, response surface methodology and the FAST approach require a large amount of specialized knowledge to make them work. Developing this knowledge and making these techniques work can be very costly in terms of analyst time. Conceptually,

Monte Carlo techniques are simpler and do not require modifications to the original model or additional numerical procedures. For example, both differential analysis and the FAST approach can require sophisticated numerical calculations. The application of response surface methodology can require specialized knowledge in experimental design and response surface construction.

Fourth, Monte Carlo techniques can be used to propagate uncertainties through a sequence of separate models. Examples of this type of analysis can be found in performance assessments for radioactive waste disposal sites[2,4] and probabilistic risk assessments for nuclear power plants.[68–70] Due to the use of a number of independent computer programs and the necessity to handle information at model interfaces appropriately, the other methods do not seem to be applicable to this type of analysis.

Fifth, Monte Carlo techniques create a mapping from analysis input to analysis results. This mapping is rich in information because of the full stratification over the range of each input variable and the wide variety of output variables that can be generated and saved. Once produced and stored, this mapping can be explored in many ways. Differential analysis is inherently local. Response surface methodology employs a very sparse stratification. For example, a random or Latin hypercube sample of size 100 will have 100 distinct values for each input variable; a classical experimental design using 100 design points will have far fewer distinct values for each input variable (probably less than five, although the exact number will depend on the specific design selected for use). The exact nature of the mapping produced by the FAST approach has not been investigated.

This review of the four formal approaches to uncertainty and sensitivity analysis techniques has been relatively brief. Additional information can be found in the cited references for each approach and also in other reviews.[5–8,71–82] There also exist a number of studies that compare the application of two or more approaches to the same analysis problem.[9,83–102]

3 IMPLEMENTATION OF A MONTE CARLO ANALYSIS

The preceding section provides an overview of the four main approaches to uncertainty and sensitivity analysis for computer models. As discussed there, the approach with the widest applicability for use in performance assessments for radioactive waste disposal is Monte Carlo analysis. For this reason, the implementation of a Monte Carlo analysis will be discussed in more detail in the present section.

In the larger literature, the term Monte Carlo analysis is used in two different, but related, contexts. In the first context, the emphasis is on sampling from one family of distributions in order to estimate another distribution. This approach has a long history of use in statistics[103–106] and is sometimes referred to as 'distribution sampling'. In the second context, emphasis is on transforming a deterministic problem into a stochastic problem and then using sampling-based techniques to solve the stochastic problem. This approach arose from work performed at Los Alamos National Laboratory in the late 1940s[107] and has been successfully applied to a variety of problems, including evaluation of integrals, solution of systems of equations, and radiation transport/shielding.[108–112] Distribution sampling invariably arises in this context but often plays a small part in the overall formulation and solution of a problem. In this paper, the term Monte Carlo analysis is being used in the first context (i.e. that of distribution sampling).

The use of Monte Carlo techniques in performance assessment for radioactive waste disposal was initiated by a program supported by the U.S. Nuclear Regulatory Commission (NRC) at Sandia National Laboratories in the late 1970s.[113] There, an approach to uncertainty and sensitivity analysis based on Latin hypercube sampling and regression analysis was proposed[114] and applied to several of the models under development as part of the NRC's performance assessment methodology for the geologic disposal of high-level radioactive waste.[115–117] Further, the ideas associated with this approach were developed for use in performance assessment[44,118,119] and applied in several example performance assessments.[2,4] In addition, these ideas also contributed to the development of the U.S. Environmental Protection Agency's proposed standard (40 CFR 191)[66] for the disposal of high-level radioactive waste.[67]

As already indicated, a Monte Carlo analysis involves five steps. The performance of each of these steps will be discussed and illustrated with results from an uncertainty/sensitivity study performed as part of a preliminary performance assessment for the Waste Isolation Pilot Plant (WIPP).[1] The analysis problem being used for illustration is shown in Fig. 1 and involves two boreholes through the waste repository. One borehole penetrates a pressurized brine pocket beneath the repository and the other does not. Flow is up the borehole from the brine pocket, through the repository, up the other borehole to the Culebra formation, and then in the Culebra to the accessible environment, which is defined to be 3 km from the repository. There is also direct transport to the surface due to cuttings removal. This is referred to as scenario E1E2 in the analyses performed for the WIPP. A diagrammatic representation for the analysis is given in Fig. 2, and a detailed description of the analysis is given elsewhere.[127]

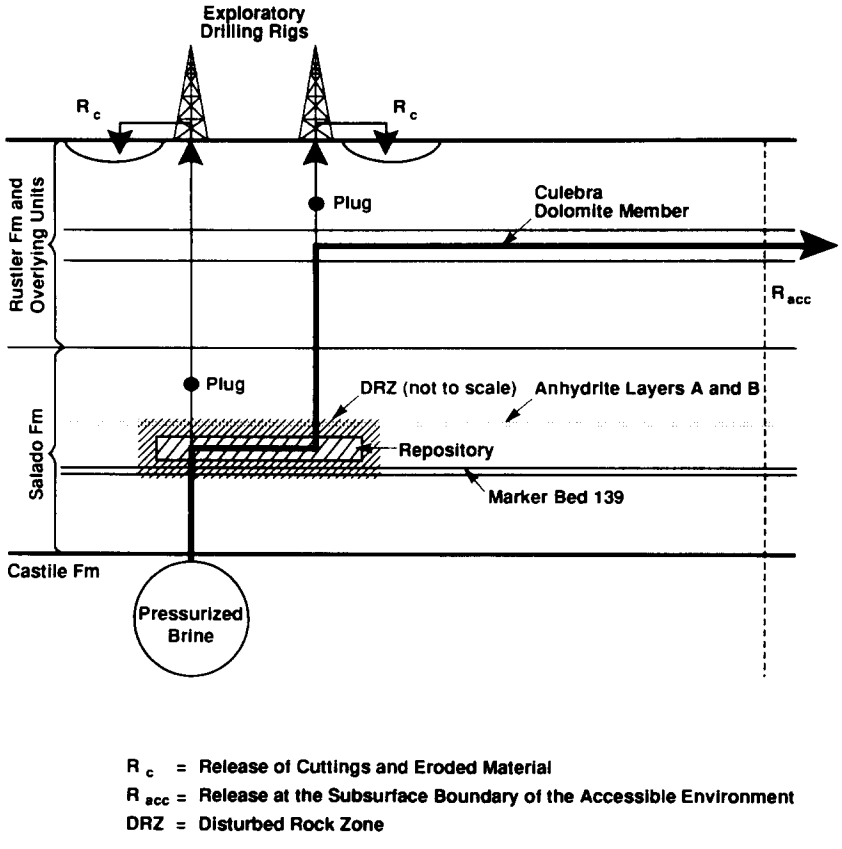

Fig. 1. Conceptual model for scenario E1E2 in the 1990 Preliminary Performance Assessment for the Waste Isolation Pilot Plant (WIPP) (Ref. 1). Arrows indicate assumed direction of flow.

3.1 Selection of variable ranges and distributions

Monte Carlo analyses use a probabilistic procedure for the selection of model input. Therefore, the first step in a Monte Carlo analysis is the selection of ranges and distributions for the input variables under consideration. When performed carefully, this can be the largest and most expensive part of a Monte Carlo analysis. However, the amount of effort expended here depends strongly on the purpose of the analysis.

If the analysis is primarily exploratory, then rather crude characterizations of the ranges and distributions for the input variables may be adequate. For example, physical plausibility arguments might be used to establish ranges, and uniform or loguniform distributions could be assumed within these ranges. These assumptions are often adequate to bound the ranges for output variables of interest and also to determine which input variables have the greatest influence on the output variables. The estimated range for an output variable and associated sensitivity results are primarily determined by the ranges assigned to the input variables. Thus, even for exploratory studies, care should be taken to avoid assigning unreasonably large ranges to variables. Sensitivity results are generally less dependent on the actual distributions assigned to the input variables than they are on the ranges chosen for the variables. However, distributional assumptions can have a large impact on the distributions estimated for output variables. Thus, when distributions for output variables must be estimated accurately, care must be used in developing distributions for the input variables.

Resources can often be used most effectively by performing a Monte Carlo analysis in an iterative manner. In a first iteration, rather crude range and distribution assumptions can be used to determine which input variables dominate the behavior of output variables of interest. Often, most of the variation in an output variable will be caused by a relatively small subset of the input variables. Once the most important input variables are identified, resources can be concentrated on characterizing their uncertainty. This avoids spending a large effort to characterize carefully the uncertainty in variables that have little impact on the ultimate outcome of an analysis.

As discussed earlier, computer models can be viewed as functions of the form shown in (1). In performance assessment, the expression in (1) may correspond to a complex sequence of linked computer programs rather than a single program. The individual variables x_j, $j = 1, \ldots, n$, can represent any parameter used in an analysis, including hydraulic conductivities, retardations, solubility limits, scenario

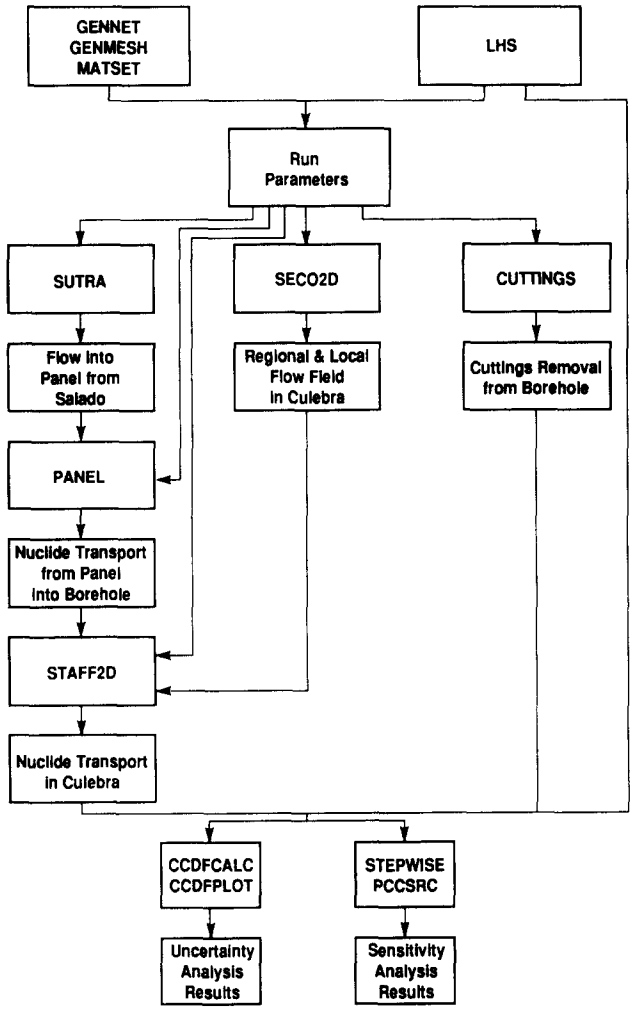

Fig. 2. Diagrammatic representation of example analysis. Program names are in capital letters, and program output is indicated in boxes following the program names. A detailed description of the analysis is given in Ref. 1. Descriptions of the individual programs are available elsewhere: GENNET, GENMESH and MATSET (Ref. 120). LHS (Ref. 121), SUTRA (Ref. 122), SECO2D (Ref. 123), CUTTINGS (Ref. 120), PANEL (Ref. 120), STAFF2D (Ref. 124), CCDFCALC and CCDFPLOT (Ref. 120), STEPWISE (Ref. 125), and PCCSRC (Ref. 126).

probabilities, parameters in distributions, probabilistic cutoffs used to eliminate low probability scenarios, and parameters that characterize numerical calculations such as mesh sizes and error bounds. The defining characteristic of these variables is that the analysis as represented in (1) requires a single value for each variable but it is uncertain as to what the value should be. Thus, the range assigned to each variable represents the set of possible values for that variable, and the corresponding distribution characterizes the likelihood that the appropriate value to use for this variable falls in various subsets of this range. As indicated in the introduction, these distributions are characterizing subjective uncertainty.

It is very important that the range assigned to a variable be consistent with its usage in the computer program that implements the underlying model. In particular, the range assigned to a variable should be consistent with the scale on which the variable is used in the specific implementation of the model under consideration. A common mistake is to estimate a variable on a local scale and then to infer uncritically that the observed local variability is the same as the uncertainty in this variable on a much larger scale. This can lead to serious misestimates of the range for the 'effective' variable value that is actually used in an analysis.

For example, a computer program might take a single value for the solubility limit of a radionuclide as input, with this single value being used throughout a room in a waste repository or perhaps even throughout the entire repository. Further, theoretical calculations or experimental results might be available for solubility limits under conditions that could occur in subregions of a room but which would be very unlikely to occur uniformly over the entire room. In this case, it would be a mistake to use the range of local results to characterize the range of solubility limits for a room or the repository since this range was developed for isolated sets of conditions that would not exist over large areas. The available information should be used in the construction of a range of 'effective' solubility limits that is consistent with the use of this parameter in the particular analysis being performed. Similar situations can occur in the characterizations of hydraulic conductivities, retardations and other variables where the scale on which data are measured is very different from the scale on which estimated variables are actually used.

The preceding discussion quite naturally leads to the following question: How should the ranges and distributions for input variables be determined for use in a Monte Carlo analysis? This is a reasonable question to ask, and a hard question to answer. Clearly, the answer must depend on the goals of the analysis, the time and resources available, and the type of information that exists for use in estimating ranges and distributions.

The simplest and most desirable situation would be to have a sequence

$$e_{1j}, e_{2j}, \ldots, e_{nE,j} \tag{42}$$

of independent, unbiased, normally and identically distributed estimates for a variable x_j exactly as it is used by a model in a particular analysis and by the computer program that implements this model. In this case, each e_{ij} is an estimate for the corresponding model input x_j, and the single best estimate for x_j is given by

$$\bar{x}_j = \sum_{i=1}^{nE} e_{ij}/nE \tag{43}$$

Further, the standard deviation, or standard error as it is sometimes called when population parameters are being considered, for $\bar{x}_j$ is given by

$$SD(\bar{x}_j) = \left[\sum_{i=1}^{nE} (e_{ij} - \bar{x}_j)^2\right]^{1/2} \Big/ \sqrt{nE(nE-1)} \quad (44)$$

The quantity

$$t = (\bar{x}_j - x_j)/SD(\bar{x}_j) \quad (45)$$

is distributed as a t-distribution with $nE - 1$ degrees of freedom, where x_j is the appropriate but unknown variable value for use in the analysis.[128] The preceding expression can be rearranged algebraically to obtain

$$x_j = \bar{x}_j - tSD(\bar{x}_j) \quad (46)$$

Thus, the t-distribution can be used to define a distribution for x_j. Further, a confidence interval (e.g. 95%, 99%) for x_j can also be obtained from the t-distribution and used to define the range of x_j. This is equivalent to excluding specified regions in the tails of the t-distribution when generating x_j from the expression in (46). The justification for using the t-distribution as a probability distribution for an uncertain variable comes from Bayes' theorem. This justification follows from using a diffuse prior distribution for the mean and standard deviation of the sampling process.[129]

As just illustrated, it may be possible to estimate the range and distribution for some variables with formal statistical procedures. Such procedures should always be used when data have been collected in an appropriate manner. Appropriate data collection usually requires prior knowledge of the precise variable to be estimated and use of a carefully planned experimental design. The exact statistical procedures selected for use would depend on the experimental design and the assumed relationships between the variable to be estimated and the data from the design.

Unfortunately, most parameters used in a performance assessment are not amenable to direct statistical estimation for various subsets for the following reasons: (1) The time scales over which parameters can be estimated are often much shorter than the time scales over which they will actually be used. (2) The physical scale on which parameters can be observed is often much smaller than the physical scale on which they will be used. As a result, heterogeneities in the system prevent individual observations from being used as estimates for system parameters. (3) Estimation of some parameters (e.g. distribution coefficients) requires the removal of material from the system. This removal can alter the properties of the material and thus lead to incorrect parameter estimates. (4) The exact conditions that will exist within the system (e.g. in a waste disposal room) are not known. Thus, it is not possible to design experiments to match the exact conditions for which parameter values are needed. (5) Collection of some types of data involves a degradation of the site (e.g. the sinking of boreholes). As a result, the collection of such data is necessarily limited. (6) Some data involve the occurrence of rare events (e.g. scenario probabilities). Although the geological and historical records can be searched for more information, designed experiments are not possible. (7) The time scales associated with future human activities make it impossible to design experiments to estimate parameters (e.g. drilling rates) associated with such activities.

Due to reasons of the type outlined in the preceding paragraph, ranges and distributions for most parameters used in a performance assessment cannot be obtained by formal statistical procedures. Nonetheless, there is still a large body of relevant information that can be used in estimating ranges and distributions. Much of this information is field data collected at the site. Other sources of information include theoretical calculations, mechanistic code calculations, physical data from other sites, and knowledge of the differences between the conditions under which data were collected and the conditions under which estimated parameters are to be used.

The challenge in developing ranges and distributions for use in a Monte Carlo study is to incorporate this diverse body of information meaningfully. Indeed, the importance of such ranges and distributions is that they provide a mathematical structure that summarizes the available information in a form that can be incorporated into further analyses. In most situations, the only practical way to develop these summary ranges and distributions is through an expert review process.

The ultimate outcome of this review process would be a distribution function $F(x)$ of the form shown in Fig. 3 for each input variable of interest. For a particular variable x_j, the function F is defined such that

$$prob(x < x_j \leq x + \Delta x) = F(x + \Delta x) - F(x) \quad (47)$$

That is, $F(x + \Delta x) - F(x)$ is equal to the probability that the appropriate value to use for x_j in the particular analysis under consideration falls between x and $x + \Delta x$. The probabilities involved in this representation will be subjective in the sense that they represent a degree of belief as to where the appropriate value for x_j falls conditional on all the information available to the reviewer or reviewers. However, when formal statistical procedures can be used as is indicated in conjunction with (46), the final result will again be a distribution of the form shown in Fig. 3. In both cases, the data summary process will have arrived at the same place: a distribution based on available information that characterizes where the appropriate value for x_j is likely to be located.

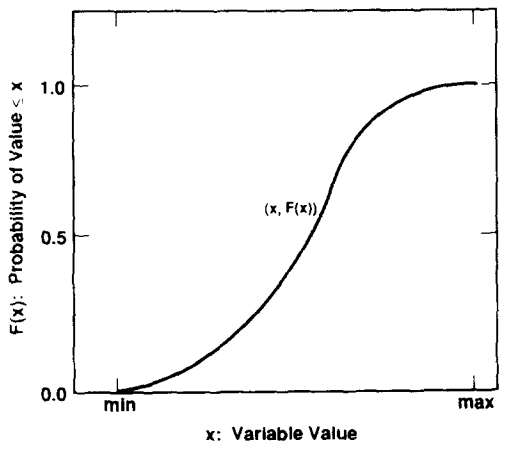

Fig. 3. Distribution function for an imprecisely known analysis variable. For each value x on the abscissa, the corresponding value $F(x)$ on the ordinate is the probability that a value less than or equal to x will occur.

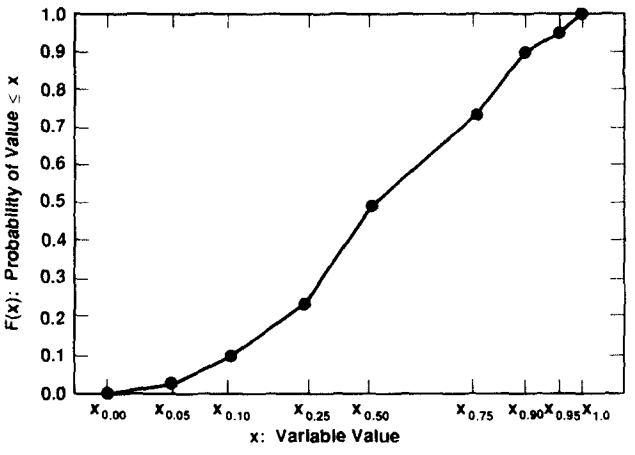

Fig. 4. Estimated distribution function for an imprecisely known analysis variable. This distribution function was built up from estimates for the following quantiles: 0·00, 0·05, 0·10, 0·25, 0·50, 0·75, 0·90, 0·95 and 1·00.

In many situations, the most appropriate way to construct a subjective distribution of the form shown in Fig. 3 is through the estimation of quantiles. For example, the process might start by determining minimum and maximum values for x_j, which defines the 0·00 and 1·00 quantiles. This provides estimates for the points

$$(x_{0\cdot00}, 0\cdot00) \text{ and } (x_{1\cdot00}, 1\cdot00) \tag{48}$$

on the distribution function in Fig. 3. The next point to estimate might be the median, which divides the range of x_j into two intervals of equal probability, followed by estimates for the 0·25 and 0·75 quantiles. This produces the following additional points on the distribution function:

$$(x_{0\cdot25}, 0\cdot25), (x_{0\cdot50}, 0\cdot50), (x_{0\cdot75}, 0\cdot75) \tag{49}$$

This process would continue by estimating additional points (e.g. the 0·05, 0·10, 0·90 and 0·95 quantiles) until the shape of the distribution is reasonably characterized. The rest of the distribution could then be filled in by assuming that the distribution function is linear between the specified quantiles, which is equivalent to fitting a maximum entropy distribution[130] as is sometimes proposed.[131,132] Figure 4 illustrates what the outcome of this process might look like.

Distribution functions for imprecisely known analysis variables can also be obtained by selecting parameter values such as the mean and standard deviation for established distributions (e.g. normal, lognormal, beta). However, it is generally best to avoid this approach for several reasons. First, there is usually no conceptual basis to pick a particular distribution. Second, it is hard to justify why a particular set of distribution parameters was selected (e.g. why a particular mean and standard deviation was selected for use with a lognormal distribution). In contrast, it is often much easier to relate the assignment of quantiles to specific information available to the reviewer. Third, most reviewers are not trained statisticians and often do not have an intuitive feeling for the relationship between the shape of a highly skewed distribution and the parameters that define it. Thus, selected parameters may not produce a distribution of the shape anticipated by the reviewer. For the preceding reasons, the use of formal distributions is undesirable because it puts an unnecessary transformation between the information possessed by the reviewer and the form in which this information is used in the analysis. In contrast, distributions constructed from quantiles are based on information that corresponds more closely to that available to the reviewer.

The scale of an expert review process can vary widely. At one extreme, a single individual might be involved in reviewing the available information on a particular variable and constructing the distribution shown in Fig. 4. The actual construction of this distribution could range from being entirely subjective to using sophisticated computational procedures to relate variability in data collected at one scale to uncertainty in a parameter for use on a different scale. At the other extreme, several teams of experts could be used to estimate a distribution independently, and then the final distribution used in the analysis would be calculated by aggregating (e.g. by averaging) the distributions obtained by the individual teams. An intermediate approach would be to have several knowledgeable individuals independently estimate a distribution and then aggregate these estimates. Bonano *et al.*[133] provide a detailed discussion on the elicitation and use of expert judgment in performance assessment for radioactive waste disposal. Discussions

of expert review processes are also provided by Berger[134] and Meyer & Booker.[135]

The U.S. Nuclear Regulatory Commission's reassessment of the risk from commercial nuclear power plants (NUREG-1150) provides an excellent example of the application of a formal expert review process to develop variable ranges and distributions for use in a Monte Carlo analysis.[69] This study involves probably the most extensive use of a formal expert review process performed to date. The general approach used and the experiences gained in its implementation are summarized in several articles.[136,137] Further, the actual performance of the expert review process is summarized in a sequence of technical reports.[138–142] This analysis used several experts to assess independently the range and distribution for each input variable of interest; then, the distributions supplied by the individual experts were averaged, with equal weight being given to each expert. A recent study of seismic hazard curves provides an example of the use of the team approach to estimating distributions.[143]

Table 1 presents the variables and their associated distributions used to illustrate Monte Carlo analysis. The indicated distributions were assigned by the principal investigators for the variables in an iterative performance assessment (i.e. the assessment is updated each year) for the WIPP being performed at Sandia.[131,144]

3.2 Generation of sample

The generation of a sample from the distributions developed in the first step of a Monte Carlo analysis is now discussed. For this discussion, suppose that the multidimensional variable $\mathbf{x}$ is under consideration and that the corresponding model can be represented by a function of the form shown in (1). Many sampling procedures have been proposed for use in Monte Carlo studies to generate samples from the joint probability distribution developed for $\mathbf{x}$.[149] The following often-used techniques are discussed below: random sampling, stratified sampling and Latin hypercube sampling.[37]

In random sampling, the observations

$$\mathbf{x}_i = [x_{i1}, \ldots, x_{in}], \quad i = 1, \ldots, m, \tag{50}$$

where m is the sample size, are selected according to the joint probability distibution for $\mathbf{x}$. Points from different regions of the sample space (i.e. range) of $\mathbf{x}$ occur in direct relationship to the probability of occurrence of these regions. Further, each sample element is selected independently of the preceding sample elements. Thus, a large sample size may be required to ensure adequate coverage of regions believed to be important but having low probabilities of occurrence.

A systematic coverage of the sample space of $\mathbf{x}$ is forced in stratified sampling. Specifically, the sample space S of $\mathbf{x}$ is partitioned into nS distinct strata S_j, $j = 1, \ldots, nS$. In general, each stratum has a different probability p_j of occurring that can be represented by

$$p_j = prob(\mathbf{x} \in S_j) \tag{51}$$

A random sample of size m_j is then obtained from each stratum S_j. That is, the points $\mathbf{x}_{jk}$, $k = 1, \ldots, m_j$, are selected at random from S_j. When all the $\mathbf{x}_{jk}$ are brought together, the result is the sequence of observations

$$\mathbf{x}_i = [x_{i1}, \ldots, x_{in}], \quad i = 1, \ldots, m = \sum_{j=1}^{nS} m_j \tag{52}$$

With stratified sampling, it is possible to force the selection of points from regions believed to be important even if these regions have a low probability of occurrence. This sampling technique is sometimes called importance sampling. When only one stratum is used, stratified sampling is the same as random sampling.

Stratified sampling operates to ensure the full coverage of specified regions in the sample space. This idea is carried further in Latin hypercube sampling to ensure the full coverage of the range of each variable. Specifically, the range of each variable (i.e. the x_j) is divided into m intervals of equal probability and one value is selected at random from each interval. The m values thus obtained for x_1 are paired at random with the m values obtained for x_2. These m pairs are combined in a random manner with the m values of x_3 to form m triples. This process is continued until a set of m n-tuples is formed. These n-tuples are of the form

$$\mathbf{x}_i = [x_{i1}, \ldots, x_{in}], \quad i = 1, \ldots, m, \tag{53}$$

and constitute the Latin hypercube sample. The individual x_j must be independent for the preceding construction procedure to work; a method for generating Latin hypercube and random samples from correlated variables has been developed by Iman & Conover[53] and will be discussed briefly. Latin hypercube sampling is an extension of quota sampling[150] and can be viewed as an n-dimensional randomized generalization of Latin square sampling.[151]

Random sampling, stratified sampling and Latin hypercube sampling are illustrated in Fig. 5 for a sample of size 10 from two uniformly distributed variables. The elements in the random sample are obtained by independent selection and subsequent pairing of values for the two variables. The individual sample elements are generated independently of the other sample elements. When a large sample size is used, the coverage of the sample space associated with x_1 and x_2 will closely approximate their joint distribution. However, when a small sample size is used, some sample elements may be close together,

Table 1. Ranges and distributions for variables in example Monte Carlo aanalysis (Refs. 131, 144)

1. SALCAP	Salado capacitance (Pa^{-1}). Distribution: Lognormal. Range: 1×10^{-11} to 1×10^{-10}. Source: Assigned by principal investigator.
2. SALPERM	Salado permeability (m^2). Distribution: Piecewise linear. Range: 1×10^{-22} to 3×10^{-20}. Source: Percentiles from data provided by principal investigator.
3. SALPRES	Salado pressure (MPa). Distribution: Uniform. Range: 7–15. Source: Range provided by principal investigators.
4. SOLALL	Waste solubility (per element, kg/kg). Distribution: Loguniform. Range: $2{\cdot}4 \times 10^{-10}$ to $2{\cdot}4 \times 10^{-4}$. Source: Assigned by principal investigator.
5. TIMINTR	Time of intrusion (sec). Distribution: Modified exponential. Range: $3{\cdot}16 \times 10^{9}$ to $3{\cdot}6 \times 10^{11}$. Source: Appendix C, Ref. 145.
6. BPPRESS	Brine pocket initial pressure (MPa). Distribution: Piecewise linear. Range: 7–17·4. Source: Range and median provided by principal investigator.
7. BHHYCND	Borehole hydraulic conductivity (m/s). Distribution: Lognormal. Range: 1×10^{-5} to 1×10^{-2}. Source: Ref. 146.
8. BHPOR	Borehole porosity (dimensionless). Distribution: Normal. Range: 0·25–0·5. Source: Ref. 146.
9. BPVOL	Brine pocket bulk volume (m^3). Distribution: Uniform. Range: $4{\cdot}8 \times 10^{3}$ to $1{\cdot}4 \times 10^{7}$. Source: Range provided by principal investigator.
10. CULTORT	Culebra tortuosity (dimensionless). Distribution: Piecewise linear. Range: 0·03–0·33. Source: Percentiles from data in Table E-9, Ref. 147.
11. CULDCOEF	Culebra diffusion coefficient (all radionuclide species, m^2/s). Distribution: Uniform. Range: $4{\cdot}8 \times 10^{-11}$ to $4{\cdot}3 \times 10^{-10}$. Source: Range from maximum and minimum values given in Table A-8, Ref. 144.
12. CULFRCSP	Culebra fracture spacing (m). Distribution: Piecewise linear. Range: 0·25–7. Source: Range and median provided by principal investigator.
13. CULRECH	Culebra recharge factor (dimensionless). Distribution: Uniform. Range: 1–2. Source: Ref. 148.
14. CULPREC	Culebra precipitation factor (dimensionless). Distribution: Uniform. Range: 1–2. Source: Ref. 148.
15. BHCSAREA	Borehole cross-sectional area (m^2). Distribution: Empirical. Range: $1{\cdot}1 \times 10^{-2}$ to $1{\cdot}6 \times 10^{-1}$. Source: Data provided by principal investigator.
16. RTRDMPU	Culebra matrix retardation factor (dimensionless) for plutonium. Distribution: Piecewise linear. Range: $1–1{\cdot}6 \times 10^{4}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
17. RTRDMAM	Culebra matrix retardation factor (dimensionless) for americium. Distribution: Piecewise linear. Range: $1–5{\cdot}6 \times 10^{3}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
18. RTRDMNP	Culebra matrix retardation factor (dimensionless) for neptunium. Distribution: Piecewise linear. Range: $1–1{\cdot}5 \times 10^{2}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
19. RTRDMU	Culebra matrix retardation factor (dimensionless) for uranium. Distribution: Piecewise linear. Range: $1–1{\cdot}1 \times 10^{2}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
20. RTRDFPU	Culebra fracture retardation factor (dimensionless) for plutonium. Distribution: Piecewise linear. Range: $1–5 \times 10^{4}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
21. RTRDFAM	Culebra fracture retardation factor (dimensionless) for americium. Distribution: Piecewise linear. Range: $1–5{\cdot}1 \times 10^{3}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
22. RTRDFNP	Culebra fracture retardation factor (dimensionless) for neptunium. Distribution: Piecewise linear. Range: $1–6{\cdot}4 \times 10^{1}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
23. RTRDFU	Culebra fracture retardation factor (dimensionless) for uranium. Distribution: Piecewise linear. Range: $1–6{\cdot}4 \times 10^{1}$. Source: Percentiles (0, 25, 50, 75, 100) provided by principal investigator.
24. CULHYCN1	Culebra hydraulic conductivity (m/s) for Zone 1. Distribution: Piecewise linear. Range: $2{\cdot}7 \times 10^{-6}$ to $5{\cdot}5 \times 10^{-5}$. Source: Percentiles from data provided by principal investigator.
25. CULHYCN2	Culebra hydraulic conductivity (m/s) for Zone 2. Distribution: Piecewise linear. Range: $9{\cdot}9 \times 10^{-9}$ to $4{\cdot}3 \times 10^{-8}$. Source: Percentiles from data provided by principal investigator.
26. CULHYCN3	Culebra hydraulic conductivity (m/s) for Zone 3. Distribution: Piecewise linear. Range: $1{\cdot}3 \times 10^{-7}$ to $3{\cdot}2 \times 10^{-7}$. Source: Percentiles from data provided by principal investigator.
27. CULHYCN4	Culebra hydraulic conductivity (m/s) for Zone 4. Distribution: Piecewise linear. Range: $3{\cdot}5 \times 10^{-8}$ to $1{\cdot}2 \times 10^{-7}$. Source: Percentiles from data provided by principal investigator.
28. CULHYCN5	Culebra hydraulic conductivity (m/s) for Zone 5. Distribution: Piecewise linear. Range: $4{\cdot}0 \times 10^{-6}$ to $4{\cdot}8 \times 10^{-6}$. Source: Percentiles from data provided by principal investigator.
29. CULHYCN7	Culebra hydraulic conductivity (m/s) for Zone 7. Distribution: Piecewise linear. Range: $1{\cdot}6 \times 10^{-5}$ to $2{\cdot}0 \times 10^{-4}$. Source: Percentiles from data provided by principal investigator.

with the result that coverage is reduced in other areas of the sample space. In particular, a very large sample size will be required to assure that specific low probability regions in the sample space have sample elements selected from them.

Stratified sampling divides the sample space into strata (i.e. the S_j, $j = 1, \ldots, 10$, appearing in Fig. 5) before the generation of the sample. Then, sample elements are randomly selected from the individual stratum. This provides a way to force sample elements to be taken from low-probability regions in the sample space that are believed to be important with respect to the analysis. However, a tradeoff must be made because, for a fixed sample size, increased coverage of

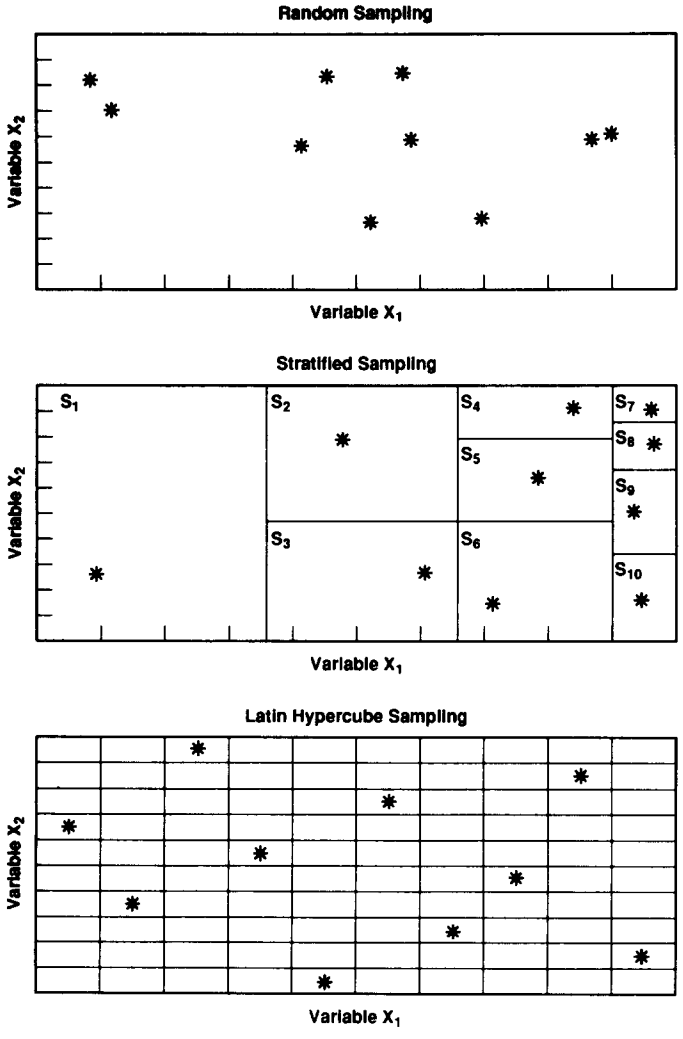

Fig. 5. Illustration of random sampling, stratified sampling and Latin hypercube sampling for a sample of size 10 from two uniformly distributed variables.

selected low probability regions will be accompanied by reduced coverage of other higher probability regions.

Latin hypercube sampling operates to ensure full coverage (i.e. stratification) over the range of each variable. Thus, as shown in Fig. 5 for a sample of size 10, the range of each variable is divided into 10 intervals of equal probability and one value is selected from each interval. In this example, the entire sample space is stratified into 100 strata of equal probability but only 10 of these strata are used in the generation of the sample. Thus, although the individual ranges of x_1 and x_2 are well-covered, there is no assurance that sample elements will be taken from specific regions in the sample space.

At the end of their comparison of sampling techniques, McKay *et al.*[37] conclude that Latin hypercube sampling has a number of desirable properties and recommend its consideration for use in Monte Carlo studies. These properties include (1) full stratification across the range of each variable, (2) relatively small sample sizes, (3) direct estimation of means, variances and distribution functions, and (4) the availability of a variety of techniques for sensitivity analysis. Another desirable property of Latin hypercube sampling is that it is possible to determine the effect of different distributions for the input variables on the estimated distribution for an output variable without rerunning the model.[44] As a result of these properties, Latin hypercube sampling has become a widely used sampling technique.[2,4,38,69,88,152-159]

Control of correlation within a sample used in a Monte Carlo analysis can be very important. If two or more variables are correlated, then it is necessary that the appropriate correlation structure be incorporated into the sample if meaningful results are to be obtained in subsequent uncertainty/sensitivity studies. On the other hand, it is equally important that variables do not appear to be correlated when they are really independent.

It is often difficult to induce a desired correlation structure on a sample. Indeed, most multivariate distributions are incompatible with the majority of correlation patterns that might be proposed for them. Thus, it is fairly common to encounter analysis situations where the proposed variable distributions and the suggested correlations between the variables are inconsistent; that is, it is not possible to have both the desired variable distributions and the requested correlations between the variables.

In response to this situation, Iman & Conover[53] have proposed a method of controlling the correlation structure in random and Latin hypercube samples that is based on rank correlation (i.e. on rank-transformed variables) rather than sample correlation (i.e. on the original data). With their technique, it is possible to induce any desired rank-correlation structure onto the sample. This technique has a number of desirable properties: (1) It is distribution free. That is, it may be used with equal facility on all types of distribution functions. (2) It is simple. No unusual mathematical techniques are required to implement the method. (3) It can be applied to any sampling scheme for which correlated input variables can logically be considered, while preserving the intent of the sampling scheme. That is, the same numbers originally selected as input values are retained; only their pairing is affected to achieve the desired rank correlations. This means that in Latin hypercube sampling the integrity of the intervals is maintained. If some other structure is used for selection of values, that same structure is retained. (4) The marginal distributions remain intact.

For many, if not most, uncertainty/sensitivity analysis problems, rank-correlation is probably a more natural measure of congruent variable behavior than is the more traditional sample correlation. What is

known in most situations is some idea of the extent to which variables tend to move up or down together; more detailed assessments of variable linkage are usually not available. It is precisely this level of knowledge that rank correlation captures.

The following discussion provides an overview of the Iman/Conover procedure for inducing a desired rank correlation structure on either a random or a Latin hypercube sample. A more detailed discussion of the procedure is given in the original article.[53] The procedure begins with a sample of size m from the n input variables under consideration. This sample can be represented by the $m \times n$ matrix

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{54}$$

where x_{ij} is the value for variable j in sample element i. Thus, the rows of $\mathbf{X}$ correspond to sample elements, and the columns of $\mathbf{X}$ contain the sampled values for individual variables.

The procedure is based on rearranging the values in the individual columns of $\mathbf{X}$ so that a desired rank correlation structure results between the individual variables. For convenience, let the desired correlation structure be represented by the $n \times n$ matrix

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \tag{55}$$

where c_{kl} is the desired rank correlation between variables x_k and x_l.

Although the procedure is based on rearranging the values in the individual columns of $\mathbf{X}$ to obtain a new matrix $\mathbf{X}^*$ that has a rank correlation structure close to that described by $\mathbf{C}$, it is not possible to work directly with $\mathbf{X}$. Rather, it is necessary to define a new matrix

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{m1} & s_{m2} & \cdots & s_{mn} \end{bmatrix} \tag{56}$$

that has the same dimensions as $\mathbf{X}$, but is otherwise independent of $\mathbf{X}$. Each column of $\mathbf{S}$ contains a random permutation of the m van der Waerden scores[160] $\Phi^{-1}(i/m + 1)$, $i = 1, 2, \ldots, m$, where Φ^{-1} is the inverse of the standard normal distribution. The matrix $\mathbf{S}$ is then rearranged to obtain the correlation structure defined by $\mathbf{C}$. This rearrangement is based on the Cholesky factorization[161] of $\mathbf{C}$. That is, a lower triangular matrix $\mathbf{P}$ is constructed such that

$$\mathbf{C} = \mathbf{P}\mathbf{P}^T \tag{57}$$

This construction is possible because $\mathbf{C}$ is a symmetric, positive-definite matrix (Ref. 161, p. 88).

If the correlation matrix associated with $\mathbf{S}$ is the $n \times n$ identity matrix (i.e. if the correlations between the values in different columns of $\mathbf{S}$ are zero), then the correlation matrix for

$$\mathbf{S}^* = \mathbf{S}\mathbf{P}^T \tag{58}$$

is $\mathbf{C}$ (Ref. 162, p. 25). At this point, the success of the procedure depends on the following two conditions: (1) that the correlation matrix associated with $\mathbf{S}$ be close to the $n \times n$ identity matrix; and (2) that the correlation matrix for $\mathbf{S}^*$ be approximately equal to the rank correlation matrix for $\mathbf{S}^*$. If these two conditions hold, then the desired matrix $\mathbf{X}^*$ can be obtained by simply rearranging the values in the individual columns of $\mathbf{X}$ in the same rank order as the values in the individual columns of $\mathbf{S}^*$. This is the first time that the variable values contained in $\mathbf{X}$ enter into the correlation process. When $\mathbf{X}^*$ is constructed in this manner, it will have the same rank correlation matrix as $\mathbf{S}^*$. Thus, the rank correlation matrix for $\mathbf{X}^*$ will approximate $\mathbf{C}$ to the same extent that the rank correlation matrix for $\mathbf{S}^*$ does.

The condition that the correlation matrix associated with $\mathbf{S}$ be close to the identity matrix is now considered. For convenience, the correlation matrix for $\mathbf{S}$ will be represented by $\mathbf{E}$. Unfortunately, $\mathbf{E}$ will not always be the identity matrix. However, it is possible to make a correction for this. The starting point for this correction is the Cholesky factorization for $\mathbf{E}$:

$$\mathbf{E} = \mathbf{Q}\mathbf{Q}^T \tag{59}$$

This factorization exists since $\mathbf{E}$ is a symmetric, positive-definite matrix. The matrix $\mathbf{S}^*$ defined by

$$\mathbf{S}^* = \mathbf{S}(\mathbf{Q}^{-1})^T\mathbf{P}^T \tag{60}$$

has $\mathbf{C}$ as its correlation matrix. In essence, multiplication of $\mathbf{S}$ by $(\mathbf{Q}^{-1})^T$ transforms $\mathbf{S}$ into a matrix whose associated correlation matrix is the $n \times n$ identity matrix; then, multiplication by $\mathbf{P}^T$ produces a matrix whose associated correlation matrix is $\mathbf{C}$. As it is not possible to be sure that $\mathbf{E}$ will be an identity matrix, the matrix $\mathbf{S}^*$ used in the procedure to produce correlated input should be defined in the corrected form shown in (60) rather than in the uncorrected form shown in (58).

The condition that the correlation matrix for $\mathbf{S}^*$ be approximately equal to the rank correlation matrix for $\mathbf{S}^*$ depends on the choice of the scores used in the definition of $\mathbf{S}$. On the basis of empirical investigations, Iman & Conover[53] found that van der Waerden scores provided an effective means of defining $\mathbf{S}$, and these scores are incorporated into the rank correlation procedure in the widely used LHS program.[121] Other possibilities for defining these

scores exist but have not been extensively investigated. The user should always examine the rank correlation matrix associated with $\mathbf{S}^*$ to ensure that it is close to the target correlation matrix $\mathbf{C}$. If this is not the case, the construction procedure used to obtain $\mathbf{S}^*$ can be repeated until a suitable approximation to $\mathbf{C}$ is obtained. Results given in Iman & Conover[53] indicate that the use of van der Waerden scores leads to rank correlation matrices for $\mathbf{S}^*$ that are close to the target matrix $\mathbf{C}$.

Additional information on the Iman/Conover (i.e. restricted pairing) technique to induce a desired rank-correlation structure is given in the original article.[53] The result of various rank-correlation assumptions is illustrated in Iman & Davenport.[163] The LHS (Latin Hypercube Sample) program[121] generates both random and Latin hypercube samples with user-specified rank correlations between variables. This is a very useful technique and its use should be considered in any sampling-based analysis.

For the example analysis, a Latin hypercube sample of size 40 was generated from the 29 variables listed in Table 1. Further, the restricted pairing technique was used to force the rank correlations between the variables to be close to zero.

3.3 Propagation of sample through analysis

The next step is the propagation of the sample through the analysis. Conceptually, this step is quite simple. Each element of the sample is supplied to the model as input, and the corresponding model predictions are saved for use in later uncertainty and sensitivity studies. This creates a sequence of results of the form

$$y_i = f(x_{i1}, x_{i2}, \ldots, x_{in}) = f(\mathbf{x}_i), \quad i = 1, 2, \ldots, m, \quad (61)$$

where n is the number of input (i.e. sampled) variables and m is the sample size. Typically, there are many model predictions of interest, in which case y_i would be a vector rather than a single number.

In its simplest form, this step involves little more than putting a 'DO loop' around the model within which (1) each sample element is read and supplied to the model as input, (2) the model is evaluated, and (3) the results of each model evaluation are written to a file that is saved after all model evaluations have been completed. In practice, this step can be somewhat more complicated than this. For example, a sampled variable may not be in exactly the form the model takes as input or model predictions may not be in the form desired for subsequent uncertainty and sensitivity analysis. In such cases, a preprocessor and a postprocessor can be added to the loop immediately before and immediately after model evaluation to perform the necessary transformations. A more complex situation sometimes arises when the model under consideration is actually a sequence of individual models, each of which supplies input to the next model in the sequence. When each model produces many distinct cases for analysis by the next model, it is sometimes necessary to use a clustering procedure at the interfaces to control the total number of cases that are propagated through the entire analysis. Otherwise, the number of individual cases can increase until the overall analysis becomes intractable due to computational cost. As an example, the NUREG-1150 analyses found it necessary to group results at model interfaces to make the Monte Carlo calculations being used to propagate uncertainties practical on a computational basis.[68,70]

The performance of sampling-based uncertainty/sensitivity studies is sometimes facilitated by the use of a special code package to control the overall analysis. As an example, the Compliance Assessment Methodology Controller (CAMCON)[120,164,165] has been developed to facilitate the performance and archival storage of the many complex calculations that will be required in the WIPP performance assessment. This methodology incorporates databases, sampling procedures, model evaluations, data storage, uncertainty and sensitivity analysis procedures, and plotting capabilities into a unified structure. CAMCON was used to control the analysis shown in Fig. 2. Other controllers have also been developed for use in performance assessment.[166,167]

3.4 Uncertainty analysis

Once a sample has been generated and propagated through the model, uncertainty analysis is straightforward. If random or Latin hypercube sampling is being used, then the expected value and variance for the output variable y can be estimated by the relations in (18) and (19) respectively. Several studies have shown that under various conditions Latin hypercube sampling results in more stable estimates than random sampling.[37,44,168] Both estimates are unbiased for random sampling. The estimated expected value is also unbiased for Latin hypercube sampling but the estimated variance is known to contain a bias. Empirical studies suggest that this bias is small.[37,53,92] When stratified sampling is used, the factors $1/m$ and $1/(m-1)$ in (18) and (19) must be replaced by weights w_i, $i = 1, \ldots, m$, that reflect the probability and number of observations associated with the individual stratum (Ref. 37, Section 8.2).

In practice, the distributions for the output variables considered in performance assessment are often highly skewed. Due to the disproportionate impact of large but unlikely values, the estimates for the means and variances associated with such distributions tend to be unstable. Here, unstable means that there is a large amount of variation among

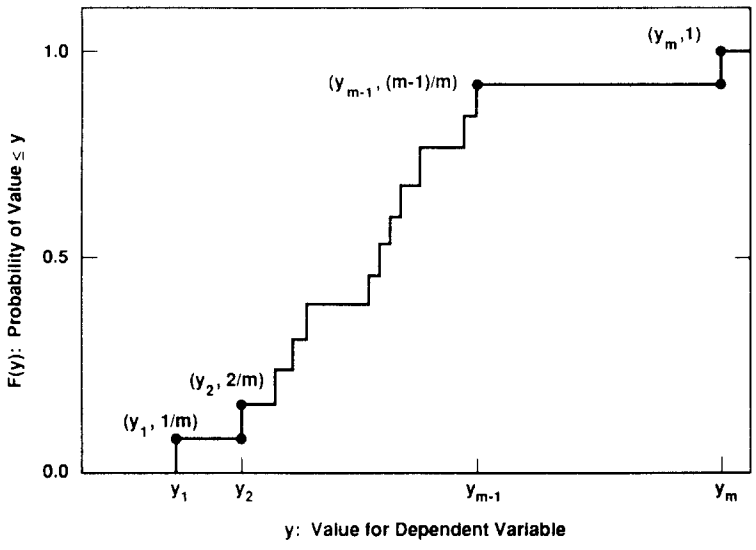

Fig. 6. Example of an estimated distribution function.

estimates obtained from independently generated samples. Further, when skewed distributions are under consideration, means and variances give a poor characterization for distribution shape. Basically, means and variances do not contain enough information to characterize highly skewed distributions adequately.

An estimated distribution function gives a better characterization of the uncertainty in an output variable than a mean and a variance. Distribution functions can be estimated from the relationship given in (20). An example estimated distribution function is shown in Fig. 6. The abscissa displays the values for the output variable, and the ordinate displays cumulative probability, which is the probability of obtaining a value equal to or less than a value on the abscissa. The step height is equal to the probability associated with the individual sample elements. If stratified sampling was being used, each observation would be assigned a weight that equalled the probability of the stratum from which it was obtained divided by the number of observations from the stratum (i.e., p_j/m_j in the notation used in conjunction with (51) and (52)).

Random sampling, stratified sampling and Latin hypercube sampling all yield unbiased estimates for distribution functions. When the restricted pairing technique developed by Iman & Conover[53] is used to control correlations within the sample, a small bias may be introduced. However, the amount of this bias does not appear to be significant.[53,92]

A distribution function readily displays the quantiles of a distribution. However, a distribution's mode (i.e. the subrange of a variable in which its probability is most concentrated) is more difficult to identify visually, although it can be done. Further, the mean is not apparent at all. Figure 7 shows an alternative uncertainty display that incorporates a distribution function, a density function and a mean into a single figure.[169] With respect to the generation of Fig. 6, the indicated mean was calculated with the original values for the variable rather than with the logarithms for these values; calculation of a mean for the logarithms is equivalent to taking a geometric rather than an arithmetic mean for the original, untransformed variable values. One advantage of the estimated distribution function is that it displays the results of every observation in an unaltered form. In contrast, the shape of the density function can be sensitivie to the gridding selected for use unless an appropriate smoothing procedure is used.[170]

As illustrated in Fig. 8, box plots[128] provide an alternative way to display the information in a distribution function. The endpoints of the boxes in Fig. 8 are formed by the lower and upper quartiles of the data, that is $x_{0.25}$ and $x_{0.75}$. The vertical line within the box represents the median, $x_{0.50}$. The sample mean is identified by the large dot. The bar on the right of the box extends to the minimum of $x_{0.75} + 1.5(x_{0.75} - x_{0.25})$ and the maximum observation. In a similar manner, the bar on the left of the box extends to the maximum of $x_{0.25} - 1.5(x_{0.75} - x_{0.25})$ and the minimum observation. The observations falling outside of these bars are shown with crosses. In symmetric distributions, these values would be considered outliers. Box plots contain the same information as a distribution function, but in a somewhat reduced form. Further, their flattened shape makes it convenient to compare different distributions. For comparison, the estimated distribution and density functions corresponding to the box plot for U-233 shown in Fig. 8 are given in Fig. 7.

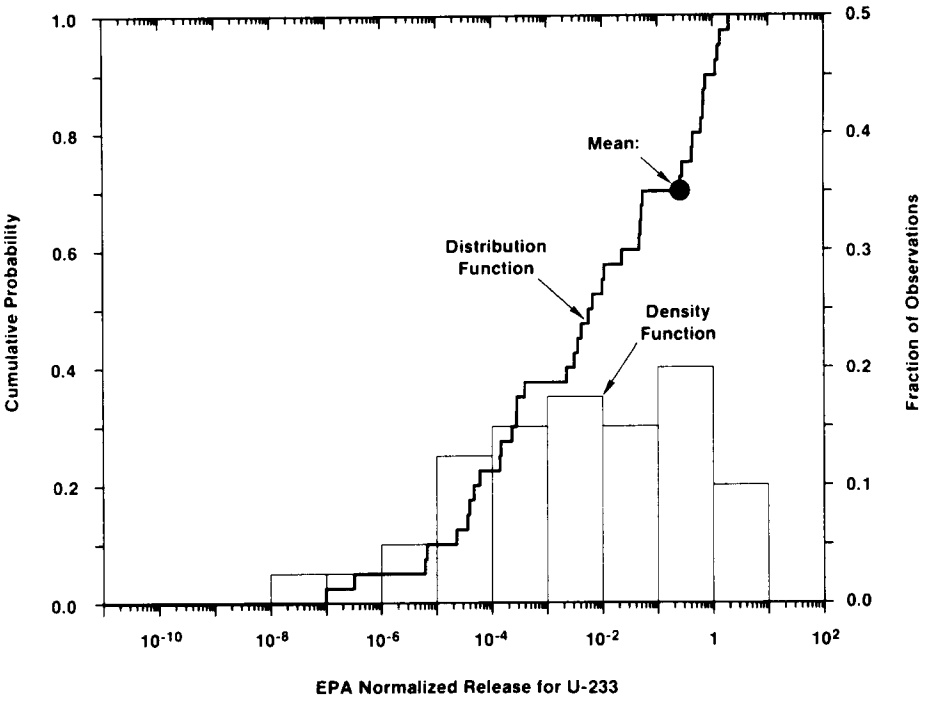

Fig. 7. Uncertainty display including estimated distribution function, density function, and mean for integrated release (normalized to EPA release limit) of U-233 to the accessible environment due to groundwater transport in fractures for scenario E1E2.

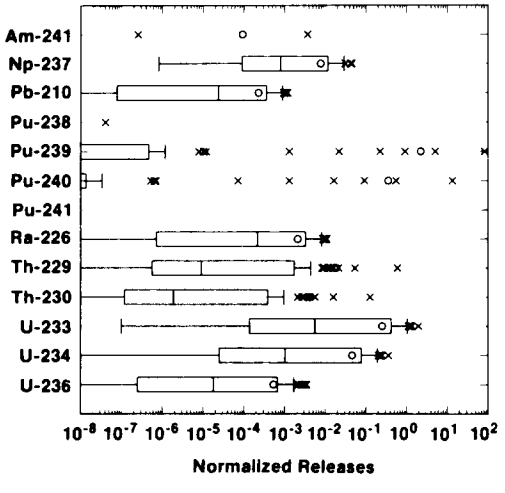

Fig. 8. Normalized releases to the accessible environment for scenario E1E2 due to groundwater transport in fractures.

Concern is often expressed with respect to the accuracy of the estimates for distribution functions obtained in Monte Carlo analyses. When random sampling is used, Kolmogorov–Smirnov bounds can be used to place confidence intervals about estimated distribution functions.[160] Other techniques also exist for use with random sampling.[171,172] When Latin hypercube sampling is used, replicated sampling can be used to place confidence intervals about estimated distribution functions.[67,173] Use of a technique called fast probability integration provides an alternative to Monte Carlo procedures for the calculation of the tails of distributions.[174–178] However, this technique does not appear to have been applied to a problem as complex as estimating the uncertainty in the results of a performance assessment.

An important property of Latin hypercube sampling is that the effects of different distribution assumptions for the input variables can be determined without rerunning the model under consideration. Specifically, once a Latin hypercube sample has been generated and the corresponding model evaluations have been performed, these evaluations can be used to estimate the means, moments and quantiles for model predictions that would result under changed distribution assumptions for the input variables (Ref. 44, Theorem 1). When a computationally demanding model is under study, this can result in significant computational savings. Examples of this procedure are given in Iman & Conover[44] and in Helton *et al.*[117] Further, Beckman & McKay investigate additional procedures for recalculating distributions without rerunning the model under study.[179]

3.5 Sensitivity analysis

The final step in a Monte Carlo study is sensitivity analysis. The generation of scatterplots is undoubtedly the simplest sensitivity analysis technique. As indicated in (21), this approach consists of generating plots of the points (x_{ij}, y_i), $i = 1, \ldots, m$, for each input variable x_j. Examples appear in Figs 9 and 10.

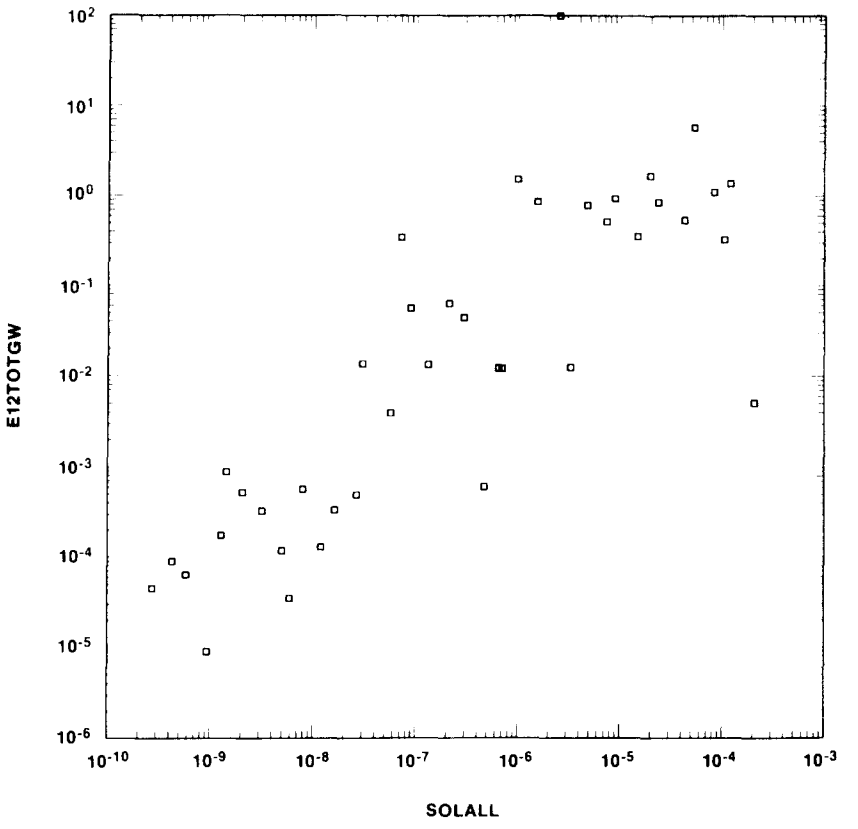

Fig. 9. Scatterplot for total normalized radionuclide release to the accessible environment due to groundwater transport in fractures for scenario E1E2 (E12TOTGW) versus solubility limit (SOLALL). Total radionuclide release involves the releases of all radionuclides converted to the EPA normalized release (Ref. 66), and the accessible environment is assumed to begin 3 km from the release point into the Culebra.

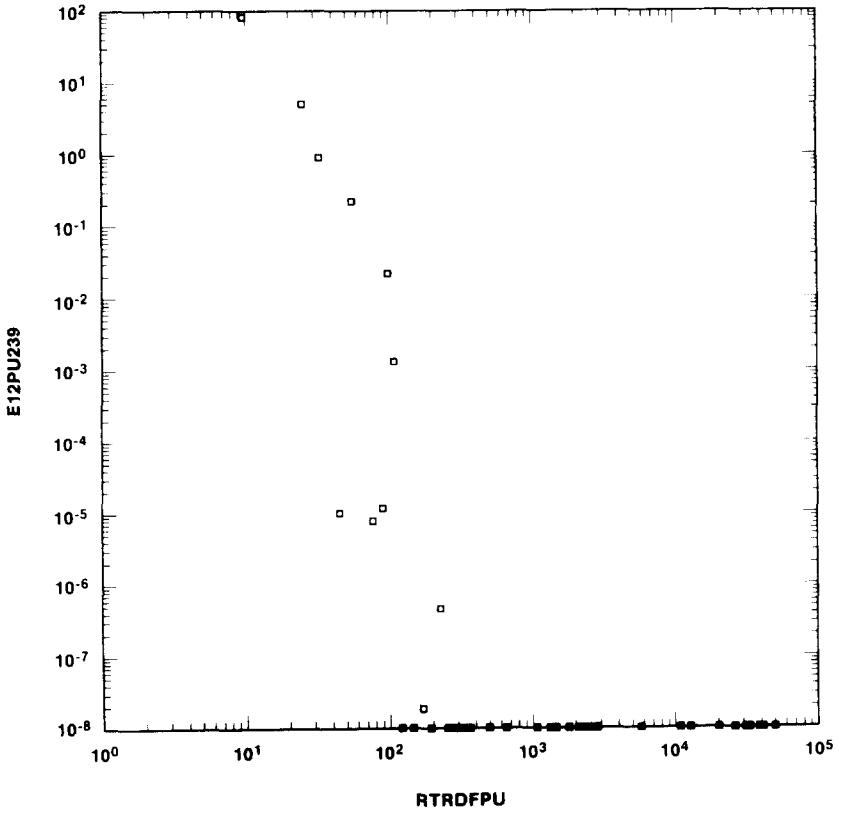

Fig. 10. Scatterplot for normalized release of Pu-239 to the accessible environment due to groundwater transport in fractures for scenario E1E2 (E12PU239) versus fracture retardation factor for plutonium (RTRDFPU).

Figure 9 shows a well-defined relationship between SOLALL (solubility limit) and total radionuclide release to the accessible environment. An interesting pattern appears in Fig. 10, where a threshold exists between RTRDFPU (retardation factor for plutonium) and the release of Pu-239. In contrast, the individual points will be randomly spread over the plot when there is no relationship between the input and the output variable.

Sometimes scatterplots alone will completely reveal the relationships between model input and model predictions. This is often the case when only one or two inputs completely dominate the outcome of the analysis. Further, scatterplots often reveal nonlinear relationships, thresholds and variable interactions that facilitate the understanding of model behavior and the planning of more sophisticated sensitivity studies. Iman & Helton[85] provide an example where the examination of scatterplots revealed a rather complex pattern of variable interactions. The examination of scatterplots is always a good starting point in a Monte Carlo sensitivity study. The examination of such plots when Latin hypercube sampling is used can be particularly revealing due to the full stratification over the range of each input variable.

Sensitivity analyses performed as part of Monte Carlo studies are often based on regression analysis. In this approach, a model of the form

$$y = b_0 + \sum_j b_j x_j \tag{62}$$

is developed from the mapping between analysis inputs and analysis results shown in (61), where the x_j are the input variables under consideration and the b_j are coefficients that must be determined. The coefficients b_j and other aspects of the construction of the regression model shown in (62) can be used to indicate the importance of the individual variables x_j with respect to the uncertainty in y.

The construction of the regression model in (62) is considered first. As shown in (61), there exists a sequence y_i, $i = 1, \ldots, m$, of values for the output variable. When expressed in the form of the model shown in (62), each y_i becomes

$$y_i = b_0 + \sum_j b_j x_{ij} + \varepsilon_i, \quad i = 1, \ldots, m \tag{63}$$

At this point, the b_j are still unknown. What is desired is to determine the b_j in some suitable manner. The method of least squares is widely used and will be employed here.[180,181] As a result of its extensive use, there exist a number of excellent textbooks on least squares regression analysis.[182–187] The purpose of the following discussion is to present just enough information to be able to describe some of the applications of regression-based techniques in sensitivity analysis. The indicated textbooks, as well as many others, provide far more information on regression analysis than can be presented here.

To determine the b_j, it is convenient to use the following matrix representation for the equalities in (63):

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \boldsymbol{\varepsilon} \tag{64}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1n} \\ \vdots & \vdots & & \vdots \\ 1 & x_{m1} & \cdots & x_{mn} \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} b_0 \\ \vdots \\ b_n \end{bmatrix} \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{bmatrix}$$

In the least squares approach, the intent is to determine the b_j such that the sum

$$S(\mathbf{b}) = \sum_{i=1}^{m} \left(y_i - b_0 - \sum_{j=1}^{m} b_j x_{ij} \right)^2 = (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b}) \tag{65}$$

is a minimum. The determination of the b_j in the least squares approach is just an exercise in calculus and is based on consideration of the first derivatives of $S(\mathbf{b})$ with respect to the individual b_j.[185]

This derivation leads to the following matrix equation that defines the coefficient vector $\mathbf{b}$ for which the sum $S(\mathbf{b})$ given in (65) is a minimum:

$$\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y} \tag{66}$$

For the analysis to produce a unique value for the coefficient vector $\mathbf{b}$, it is necessary that the matrix $\mathbf{X}^T\mathbf{X}$ be invertible. Then, $\mathbf{b}$ is given by

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{67}$$

The matrix $\mathbf{X}^T\mathbf{X}$ will always be invertible when the columns of $\mathbf{X}$ are linearly independent. This usually is the case in a Monte Carlo study in which the number sample elements (i.e. m) exceeds the number of independent variables (i.e. n).

The following identity holds for the least squares regression model and plays an important role in assessing the adequacy of such models:

$$\sum_i (y_i - \bar{y})^2 = \sum_i (\hat{y}_i - \bar{y})^2 + \sum_i (\hat{y}_i - y_i)^2 \tag{68}$$

where $\hat{y}_i$ denotes the estimate of y_i obtained from the regression model and $\bar{y}$ is the mean of the y_i.[185] Since

$$\sum_i (\hat{y}_i - y_i)^2 \tag{69}$$

provides a measure of variability about the regression line, the ratio

$$R^2 = \sum_i (\hat{y}_i - \bar{y}_i)^2 \Big/ \sum_i (y_i - \bar{y})^2 \tag{70}$$

provides a measure of the extent to which the regression model can match the observed data. Specifically, when the variation about the regression line is small (i.e. when $\sum_i (\hat{y}_i - y_i)^2$ is a small relative to $\sum_i (\hat{y}_i - \bar{y})^2$), then the corresponding R^2 value is close to 1, which indicates that the regression model is accounting for most of the variablity in the y_i. Conversely, an R^2 value close to zero indicates that the regression model is not very successful in accounting for the variability in the y_i. Another name for R^2 is the coefficient of multiple determination.

An important situation occurs when the rows of the matrix $\mathbf{X}$ (i.e. the variable values at which the model is evaluated) are selected so that $\mathbf{X}^T\mathbf{X}$ is a diagonal matrix. In this case, the columns of $\mathbf{X}$ are said to be orthogonal, and the estimated regression coefficients are given by

$$\begin{aligned}\mathbf{b} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \\ &= \begin{bmatrix} d_0 & 0 & \cdots & 0 \\ 0 & d_1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{m1} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}\end{aligned} \tag{71}$$

and so each element b_j of $\mathbf{b}$ is given by

$$b_j = \sum_{i=1}^{m} x_{ij} y_i / d_j = \sum_{i=1}^{m} x_{ij} y_i \Big/ \sum_{i=1}^{m} x_{ij}^2 \tag{72}$$

The important point to recognize is that the estimate of the regression coefficient b_j for the variable x_j depends only on the values for x_j in the design matrix $\mathbf{X}$ (i.e. $x_{1j}, \ldots, x_{mj}$). This is true regardless of the number of variables included in the regression. As long as the design is orthogonal, the addition or deletion of variables from the model will not change the regression coefficients. Further, when the design matrix $\mathbf{X}$ is orthogonal, the R^2 value for the regression can be expressed as

$$\begin{aligned}R^2 &= \sum_{i=1}^{m} (\hat{y}_i - \bar{y})^2 \Big/ \sum_{i=1}^{m} (y_i - \bar{y})^2 \\ &= R_1^2 + R_2^2 + \cdots + R_n^2\end{aligned} \tag{73}$$

where R_j^2 is the R^2 value that results from regressing y on only x_j (Ref. 5, eqn (III-74)). Thus, R_j^2 is equal to the contribution of x_j to R^2 when the design matrix $\mathbf{X}$ is orthogonal.

The regression model in (62) can be algebraically reformulated as

$$(y - \bar{y})/\hat{s} = \sum_j (b_j \hat{s}_j / \hat{s})(x_j - \bar{x}_j)/\hat{s}_j \tag{74}$$

where

$$\bar{y} = \sum_i y_i / m, \quad \hat{s} = \left[\sum_i (y_i - \bar{y})^2 / (m-1) \right]^{1/2}$$

$$\bar{x}_j = \sum_i x_{ij} / m, \quad \hat{s}_j = \left[\sum_i (x_{ij} - \bar{x}_j)^2 / (m-1) \right]^{1/2}$$

The coefficients $b_j \hat{s}_j / \hat{s}$ appearing in (74) are called standardized regression coefficients. When the x_j are independent, the absolute value of the standardized regression coefficients can be used to provide a measure of variable importance. Specifically, the coefficients provide a measure of importance based on the effect of moving each variable away from its expected value by a fixed fraction of its standard deviation while retaining all other variables at their expected values. Calculating standardized regression coefficients is equivalent to performing the regression analysis with the input and output variables normalized to mean zero and standard deviation one.

An example regression analysis is now given. The output variable is the release to the surface due to cuttings removal from a borehole, which is predicted by the model CUTTINGS shown in Fig. 2. The release is expressed as a normalized release as defined by the Environmental Protection Agency.[66] A regression model could be constructed with all 29 variables in Table 1. However, only the variables BHCSAREA and TIMINTR are used as input to CUTTINGS. The following regression model is obtained using these two variables and the previously indicated Latin hypercube sample (i.e. $n = 2$ and $m = 40$):

$$\begin{aligned}y = {} & 1{\cdot}2 \times 10^{-3} + 2{\cdot}8 \times 10^{-2} * \text{BHCSAREA} \\ & - 8{\cdot}0 \times 10^{-15} * \text{TIMINTR}\end{aligned} \tag{75}$$

The coefficients in the preceding model show the effect of a one unit change in an input variable on the output variable. Thus, increasing BHCSAREA increases the release while increasing TIMINTR decreases the release.

It is hard to assess variable importance from the regression coefficients in (75) because of the effects of units and distribution assumptions. In particular, the regression coefficient for TIMINTR is much smaller than the coefficient for BHCSAREA. Variable importance is more clearly shown by the following reformulation of (75) with standardized regression coefficients:

$$y = 0{\cdot}76 * \text{BHCSAREA} - 0{\cdot}50 * \text{TIMINTR} \tag{76}$$

The standardized regression coefficients in (76) provide a better characterization of variable importance than the unstandardized coefficients in (75). For perturbations equal to a fixed fraction of their standard deviation, the impact of BHCSAREA is approximately 50% larger than the impact of

TIMINTR. Both regression models have a R^2 value of 0·69 and thus can account for 69% of the variability in the dependent variable.

Regression analyses often perform poorly when the relationships between the input and output variables are nonlinear. This is not surprising since regression analysis is based on developing linear relationships between variables. The problems associated with poor linear fits to nonlinear data can often be avoided with the technique of rank regression.[188] Rank regression is a simple concept: data are replaced with their corresponding ranks and then the usual regression procedures are performed on these ranks. Specifically, the smallest value of each variable is assigned the rank 1, the next largest value is assigned the rank 2, and so on up to the largest value, which is assigned the rank m, where m denotes the number of observations. The analysis is then performed with these ranks being used as the values for the input and output variables.

As an example, repeating the regression analysis for the cuttings release shown in (75) and (76) with rank-transformed data leads to

$$y = 16{\cdot}0 + 0{\cdot}72 * \text{BHCSAREA} - 0{\cdot}54 * \text{TIMINTR} \tag{77}$$

When standardized regression coefficients are used, this equation becomes

$$y = 0{\cdot}70 * \text{BHCSAREA} - 0{\cdot}54 * \text{TIMINTR} \tag{78}$$

The R^2 value for these models is 0·84, which is an increase from the value of 0·69 obtained with untransformed variables. The R^2 value is increasing because the rank transformation tends to linearize nonlinear but monotonic relations. This linearization can be seen in Fig. 11 in the scatter-plots for TIMINTR using raw (i.e. untransformed) and rank-transformed data.

The ideas of correlation and partial correlation are useful concepts that often appear in sampling-based uncertainty/sensitivity studies. For a sequence of observations (x_i, y_i), $i = 1, \ldots, m$, the (sample) correlation r_{xy} between x and y is defined by

$$r_{xy} = \frac{\sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{m} (x_i - \bar{x})^2\right]^{1/2}\left[\sum_{i=1}^{m} (y_i - \bar{y})^2\right]^{1/2}} \tag{79}$$

where $\bar{x}$ and $\bar{y}$ are defined in conjunction with (74). The correlation coefficient r_{xy} provides a measure of the linear relationship between x and y.

The nature of the correlation coefficient r_{xy} is most readily understood by considering the regression

$$y = b_0 + b_1 x \tag{80}$$

The definition of r_{xy} in (79) is equivalent to the definition

$$r_{xy} = \text{sign}(b_1)(R^2)^{1/2} \tag{81}$$

where $\text{sign}(b_1) = 1$ if $b_1 \geq 0$, $\text{sign}(b_1) = -1$ if $b_1 < 0$, and R^2 is the coefficient of determination that results from regressing y on x. With respect to interpretation, the correlation coefficient r_{xy} provides a measure of the linear relationship between x and y, and the regression coefficient b_1 characterizes the effect that a unit change in x will have on y.

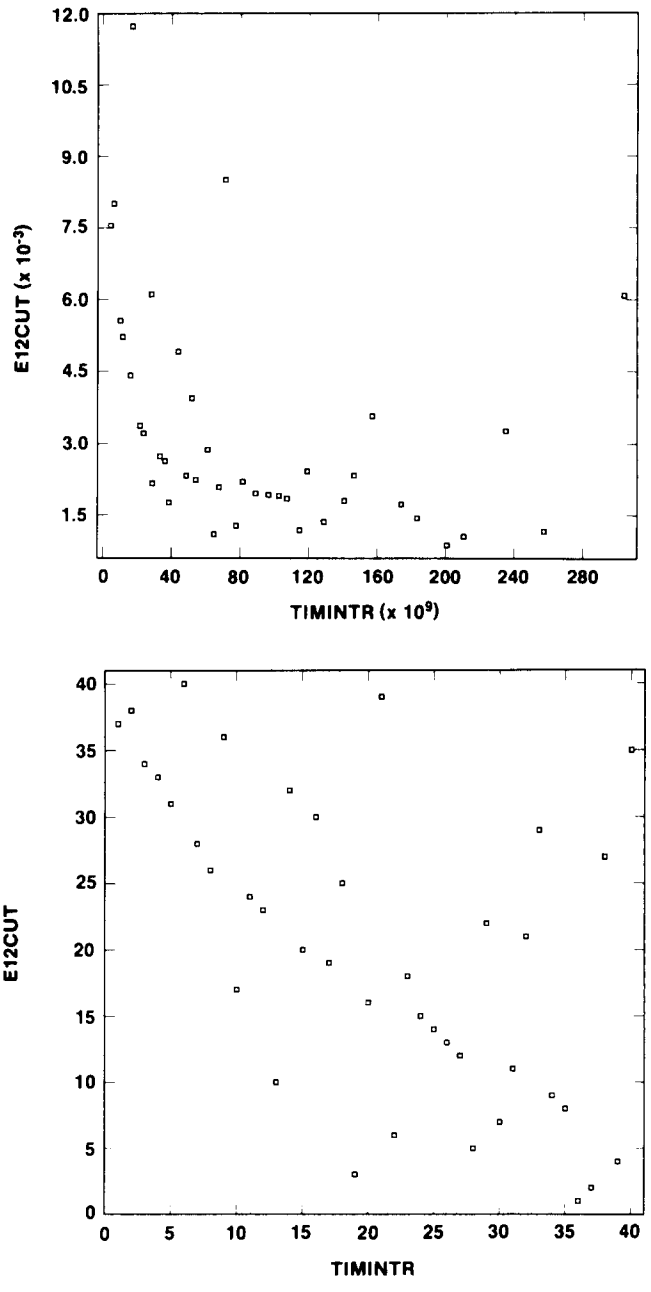

Fig. 11. Effect of a rank transformation. Upper frame shows the scatterplot for cuttings release (E12CUT) versus time of intrusion (TIMINTR) for raw (i.e. untransformed) data; lower frame shows the same scatterplot for rank-transformed data. The variable E12CUT involves the release of all radionuclides converted to the EPA normalized release (Ref. 66).

When more than one input variable is under consideration, partial correlation coefficients can be used to provide a measure of the linear relationships between the output variable y and the individual input variables. The partial correlation coefficient between y and an individual variable x_p is obtained from the use of a sequence of regression models. First, the following two regression models are constructed:

$$\hat{y} = b_0 + \sum_{j \neq p} b_j x_j \quad \text{and} \quad \hat{x}_p = c_0 + \sum_{j \neq p} c_j x_j \qquad (82)$$

Then, the results of the two preceding regressions are used to define the new variables $y - \hat{y}$ and $x_p - \hat{x}_p$. By definition, the partial correlation coefficient between y and x_p is the correlation coefficient between $y - \hat{y}$ and $x_p - \hat{x}_p$. Thus, the partial correlation coefficient provides a measure of the linear relationship between y and x_p with the linear effects of the other variables removed. The preceding provides a rather intuitive development of what a partial correlation coefficient is. A formal development of partial correlation coefficients and the relationships between partial correlation coefficients and standardized regression coefficients is provided by Iman *et al.*[126]

The partial correlation coefficient provides a measure of the strength of the linear relationship between two variables after a correction has been made for the linear effects of the other variables in the analysis, and the standardized regression coefficient measures the effect on the output variable that results from perturbing an input variable by a fixed fraction of its standard deviation. Thus, partial correlation coefficients and standardized regression coefficients provide related, but not identical, measures of variable importance. In particular, the partial correlation coefficient provides a measure of variable importance that tends to exclude the effects of other variables, the assumed distribution for the particular input variable under consideration, and the magnitude of the impact of an input variable on an output variable. In contrast, the value for a standardized regression coefficient is significantly influenced by both the distribution assigned to an input variable and the impact that this variable has on an output variable. However, when the input variables in an analysis are uncorrelated, an ordering of variable importance based on either the absolute value of standardized regression coefficients or the absolute value of partial correlation coefficients will yield the same ranking of variable importance, even though the standardized regression coefficients and partial correlation coefficients for individual variables may be quite different.[126]

Many output variables are functions of time or location. A useful way to present sensitivity results for such variables is with plots of partial correlation coefficients or standardized regression coefficients. An example of such a presentation is given in Fig. 12 for groundwater transport of U-233 along a path in the Culebra from a borehole to the accessible environment. As indicated in Fig. 2, the transport calculations for the Culebra were performed with the STAFF2D program.[124]

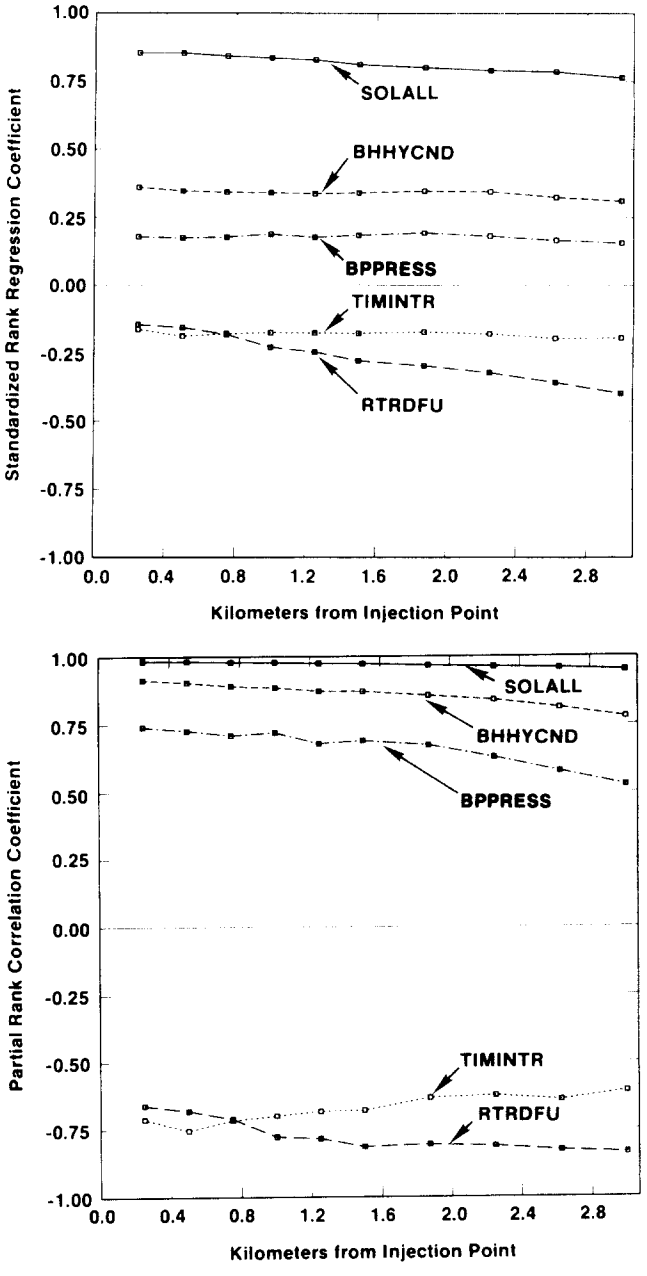

Fig. 12. Standardized rank regression coefficients and partial rank correlation coefficients for the integrated discharge of U-233 as a function of transport distance.

Figure 12 displays two sets of curves. The upper set contains standardized rank regression coefficients plotted as a function of transport distance. The lower set contains partial rank correlation coefficients (i.e. partial correlation coefficients calculated with rank-transformed data) as a function of transport distance.

For both sets of curves, the dependent variables are the total integrated discharge of U-233 over 10 000 years at fixed distances along the transport path, and each curve displays the values of the standardized rank regression coefficient or partial rank correlation coefficient relating these integrated discharges to a single input variable as a function of distance. Figure 12 displays curves for all variables that had a partial rank correlation coefficient of at least 0·75, in absolute value, at one or more points along the transport path.

When many input variables are involved, the direct construction of a regression model as shown in (62) containing all input variables may not be the best approach for several reasons. First, the large number of variables makes the regression model tedious to examine and unwieldy to display. Second, it is often the case that only a relatively small number of input variables have an impact on the output variable. As a result, there is no reason to include the remaining variables in the regression model. Third, correlated variables result in unstable regression coefficients (i.e. coefficients whose values are sensitive to the specific variables included in the regression model). When this occurs, the regression coefficients in a model containing all the input variables can give a misleading representation of variable importance. As a side point, if several input variables are highly correlated, consideration should be given to either removing all but one of the correlated variables or transforming the variables to correct for (i.e. remove) the correlations between them. Fourth, an overfitting of the data can result when variables are arbitrarily forced into the regression model. This phenomenon occurs when the regression model attempts to match the predictions associated with individual sample elements rather than match the trends shown by the sample elements collectively.

Stepwise regression analysis provides an alternative to constructing a regression model containing all the input variables. With this approach, a sequence of regression models is constructed. The first regression model contains the single input variable that has the largest impact on the output variable. The second regression model contains the two input variables that have the largest impact on the output variable: the input variable from the first step plus whichever of the remaining variables has the largest impact on the variation not accounted for by the first variable. The third regression model contains the three input variables that have the largest impact on the output variable: the two input variables from the second step plus whichever of the remaining variables has the largest impact on the variation not accounted for by the first two variables. Additional models in the sequence are defined in the same manner until the point is reached at which further models are unable to meaningfully increase the amount of the variation in the output variable that can be accounted for. Further, at each step of the process, the possibility exists for an already selected variable to be dropped out if it no longer has a significant impact on the uncertainty in the output variable; this only occurs when correlations exist between the input variables.

Several aspects of stepwise regression analysis provide insights on the importance of the individual variables. First, the order in which the variables are selected in the stepwise procedure provides an indication of their importance, with the most important variable being selected first, the next most important variable being selected second, and so on. Second, the R^2 values (see (70)) at successive steps of the analysis also provide a measure of variable importance by indicating how much of the variation in the dependent variable can be accounted for by all variables selected through each step. When the input variables are uncorrelated, the differences in the R^2 values for the regression models constructed at successive steps equals the fraction of the total variability in the output variable that can be accounted for by the individual input variables being added at each step (see (73)). Third, the absolute values of the standardized regression coefficients (see (74)) in the individual regression models provide an indication of variable importance. Further, the sign of a standardized regression coefficient indicates whether the input and output variable tend to increase and decrease together (a positive coefficient) or tend to move in opposite directions (a negative coefficient).

An important situation occurs when the input variables are uncorrelated. In this case, the orderings of variable importance based on order of entry into the regression model, size of the R^2 values attributable to the individual variables, the absolute values of the standardized regression coefficients, and the absolute values of the partial correlation coefficients are the same. In situations where the input variables are believed to be uncorrelated, one of the important applications of the previously discussed restricted pairing technique of Iman & Conover[53] is to ensure that the correlations between variables within a Latin hypercube or random sample are indeed close to zero. When variables are correlated, care must be used in the interpretation of the results of a regression analysis since the regression coefficients can change in ways that are basically unrelated to the importance of the individual variables as correlated variables are added to and deleted from the regression model.

When the stepwise technique is used to construct a regression model, it is necessary to have some criteria to stop the construction process. When there are many independent variables, there is usually no reason to let the construction process continue until all the variables have been used. It is also necessary to have some criteria to determine when a variable is no

longer needed and thus can be dropped from the regression model. As indicated earlier, this latter situation only occurs when the input variables are correlated.

The usual criterion for making the preceding decisions is based on whether or not the regression coefficient associated with an input variable appears to be significantly different from zero. Specifically, the t-test is used to determine the probability that a regression coefficient as large as or larger than the one constructed in the analysis would be obtained if, in reality, there was no relationship between the input and output variable, and, as a result, the apparent relationship that led to the constructed regression coefficient was due entirely to chance (Ref. 187, Section 7.5). The probability of exceeding a regression coefficient due to chance variation is often referred to as an α-value. The actual derivation of the α-value depends on assumptions involving normality and random variation that are not satisfied in sampling-based sensitivity studies for computer models since there is no variation in the predictions for a fixed set of input. However, the t-test and the associated α-value still constitute a useful criterion for adding or deleting variables from a regression model in a sensitivity study since they provide a measure of how viable the relationship between the input and output variable would appear to be in a study in which this relationship could possibly have arisen from random variation. Sensitivity studies often use an α-value of 0·01 or 0·02 to add a variable to a regression model and a somewhat larger value to drop a variable from the model.

As models involving more variables are developed in a stepwise regression analysis, the possibility exists of overfitting the data. Overfitting occurs when the regression model in essence 'chases' the individual observations rather than following an overall pattern in the data. For example, it is possible to obtain a good fit on a set of points by using a polynomial of high degree. However, in doing so, it is possible to overfit the data and produce a spurious model that makes poor predictions.

To protect against overfit, the Predicted Error Sum of Squares (PRESS) criterion can be used to determine the adequacy of a regression model.[189] For a regression model containing k variables and constructed from m observations, PRESS is computed in the following manner. For $i = 1, 2, \ldots, m$, the ith observation is deleted from the original set of m observations and then a regression model containing the original k variables is constructed from the remaining $m - 1$ observations. With this new regression model, the value $\hat{y}_k(i)$ is estimated for the deleted observation y_i. Then, PRESS is defined from the preceding predictions and the m original observations by

$$\text{PRESS}_k = \sum_{i=1}^{m} [y_i - \hat{y}_k(i)]^2 \tag{83}$$

The regression model having the smallest PRESS value is preferred when choosing between two competing models, as this is an indication of how well the basic pattern of the data has been fitted versus an overfit or an underfit. In addition to PRESS, there are also a number of other diagnostic tools that can be used to investigate the adequacy of regression models.[190,191]

It is very important to use scatterplots, PRESS values and other procedures to examine the reasonableness of regression models. This is especially true when regression models are used for sensitivity analysis. Such analyses often involve many input variables and large uncertainties in these variables. The appearance of spurious patterns is a possibility that must be checked for.

An example of stepwise regression analysis is now presented for the integrated release of U-233 to the accessible environment due to groundwater transport in fractures for scenario E1E2 shown in Fig. 1. This analysis uses the 29 input variables shown in Table 1 and the previously described Latin hypercube sample of size $m = 40$. Thus, the data available for analysis are of the form

$$[y_i, x_{i1}, x_{i2}, \ldots, x_{i,29}], \quad i = 1, \ldots, 40 \tag{84}$$

where y_i is the value for the integrated U-233 release to the accessible environment obtained with the ith sample element and x_{ij}, $j = 1, \ldots, 29$, are the values for the variables shown in Table 1 for the ith sample element. The distribution of the output variable y is shown in Fig. 7.

The analyses were tried with both raw and rank-transformed data. The analysis with raw data found one variable that resulted in a regression model with an α-value less than 0·02. Specifically, the variable BHHYCND was selected with an α-value of 0·0129 and led to a regression model with an R^2 value of only 0·15. In contrast, the analysis with rank-transformed data found five variables with α-values less than 0·02. Therefore, this example will be based on rank-transformed data. The individual steps in the analysis for rank-transformed data are now discussed.

The first step selects the input variable x_j that has the largest impact on the output variable y. Specifically, this is defined to be the variable that has the largest correlation, in absolute value, with y (see (79) and (81)). Thus, it is necessary to calculate the correlations between y and each of the 29 input variables under consideration. For illustration, Table 2 shows the 6×6 correlation matrix for y and the five

Table 2. Correlation matrix for variables selected in stepwise regression analysis for the integrated release of U-233 to the accessible environment due to groundwater transport in fractures for scenario E1E2. The correlations in this table are calculated with rank-transformed data

SOLALL	1·000 0					
TIMINTR	0·000 2	1·000 0				
BPPRES	−0·063 6	−0·027 6	1·000 0			
BHHYCND	−0·002 1	0·003 4	0·039 6	1·000 0		
RTRDFU	−0·048 8	0·004 7	0·093 2	0·055 7	1·000 0	
E12U233	0·777 5	−0·203 0	0·101 3	0·279 5	−0·346 9	1·000 0
	SOLALL	TIMINTR	BPPRES	BHHYCND	RTRDFU	E12U233

input variables ultimately selected in the stepwise regression, although the full correlation matrix would actually be $(29+1)\times(29+1)$. Each element in the correlation matrix is the correlation between the variables in the corresponding row and column. As examination of the correlation matrix in Table 2 shows, the variable SOLALL has the highest correlation with the U-233 release, which is denoted by E12U233. Thus, the first step in the analysis selects the variable SOLALL. A regression model relating y to SOLALL is then developed as shown in (67) with $n=1$ and $m=40$. The resultant rank regression model is

$$\hat{y} = 4{\cdot}56 + 0{\cdot}777 * \text{SOLALL} \tag{85}$$

which has an R^2 value of 0·604, an α-value of 0·0000, a standardized regression coefficient of 0·777 and a PRESS value of 2342. This model is summarized in Table 3.

Table 3. Results of stepwise regression analysis with rank-transformed data for the integrated release of U-233 to the accessible environment due to groundwater transport in fractures for scenario E1E2

Step[a]	Variables[b]	SRC[c]	α-Values[d]	R^2 values[e]	PRESS[f]
1	SOLALL	0·777	0·000 0	0·604	2342
2	SOLALL	0·762	0·000 0	0·700	1851
	RTRDFU	−0·310	0·001 5		
3	SOLALL	0·762	0·000 0	0·789	1375
	RTRDFU	−0·326	0·000 1		
	BHHYCND	0·299	0·000 4		
4	SOLALL	0·762	0·000 0	0·831	1184
	RTRDFU	−0·325	0·000 0		
	BHHYCND	0·300	0·000 1		
	TIMINTR	−0·203	0·006 2		
5	SOLALL	0·772	0·000 0	0·857	1059
	RTRDFU	−0·340	0·000 0		
	BHHYCND	0·294	0·000 1		
	TIMINTR	−0·198	0·004 3		
	BPPRES	0·165	0·016 2		

[a] Steps in the analysis.
[b] Variables selected at each step.
[c] Standardized regression coefficients (SRCs) for variables in the regression model at each step.
[d] α-Values for variables in the regression model at each step.
[e] R^2 value for the regression model at each step.
[f] Predicted error sum of squares (PRESS) value for the regression model at each step.

The second step selects the input variable x_j that has the largest impact on variation in the output variable y that cannot be accounted for by SOLALL, the variable selected in the first step. This selection is made by defining a new variable

$$\tilde{y} = y - \hat{y} = y - (4{\cdot}56 + 0{\cdot}777 * \text{SOLALL}) \tag{86}$$

where $\hat{y}$ is defined in (85), and then calculating the correlations between $\tilde{y}$ and the remaining variables. The variable with the largest correlation, in absolute value, with $\tilde{y}$ is selected as the second variable for inclusion in the model. In this example, the selected variable is RTRDFU. The regression model at this step will thus involve the two variables SOLALL and RTRDFU and is constructed as shown in (67) with $n=2$ and $m=40$. The resultant rank regression model is

$$\hat{y} = 11{\cdot}2 + 0{\cdot}762 * \text{SOLALL} - 0{\cdot}310 * \text{RTRDFU} \tag{87}$$

This model is summarized in Table 3.

The third step selects the input variable x_j that has the largest impact on the variation in the output variable y that cannot be accounted for by SOLALL and RTRDFU, the two variables from the second step. This selection is made by defining a new variable

$$\tilde{y} = y - \hat{y} = y - (11{\cdot}2 + 0{\cdot}762 * \text{SOLALL} - 0{\cdot}310 * \text{RTRDFU}) \tag{88}$$

where $\hat{y}$ is defined in (87). The variable with the largest correlation, in absolute value, with $\tilde{y}$ is selected as the third variable for inclusion in the model. In this example, the selected variable is BHHYCND. The regression model for this step will thus involve the three variables SOLALL, RTRDFU and BHHYCND. The resultant regression model is summarized in Table 3.

As shown in Table 3, the stepwise procedure then continues in the same manner through a total of five steps, until no more variables can be found with an α-value less than 0·02. At this point, the stepwise procedure stops.

At each step, the stepwise procedure also checks to see if any variable selected at a prior step now has an α-value that exceeds a specified level, 0·05 in the case of this analysis. If such a situation occurs, the variable

will be dropped from the analysis, with the possibility that it may be reselected at a later step as other variables are added and deleted from the model. This type of behavior only occurs when there are correlations between the input variables. As shown in the example correlation matrix in Table 2, the restricted pairing technique has been successful in keeping the correlations between the input variables close to zero. Thus, no variables meet the criterion to be dropped from the regression model once they have been selected at a prior step.

Another result of this lack of correlation is that the regression coefficients do not change significantly as additional variables are added to the regression model. As examination of Table 3 shows, the regression coefficients for a specific variable are essentially the same in all regression models containing that variable. Further, as indicated in (73), the R^2 values obtained for successive models can be subtracted to obtain the contribution to the variability in y due to the newly added variable. Thus, for example, SOLALL accounts for approximately 60% of the variability in y, while SOLALL and RTRDFU together account for approximately 70% of the variability. As a result, RTRDFU by itself accounts for approximately $70\% - 60\% = 10\%$ of the variability in y. Similar results hold for the other variables selected in the analysis.

Table 3 also reports the PRESS values for the regression models obtained at the individual steps in the analysis. A decreasing sequence of PRESS values indicates that the regression models are not overfitting the data on which they are based. An increase in the PRESS values suggests that a model is overfitting the data, and thus that the stepwise procedure should probably be stopped at the preceding step. As shown by the decreasing PRESS values in Table 3, the regression models in this analysis are probably not overfitting the data from which they were constructed.

Typically, a certain amount of discretion is involved in selecting the exact point at which to stop a stepwise regression analysis. Certainly, α-values and the behavior of PRESS values provide two criteria to consider in selecting a stopping point. Other criteria include the changes in the R^2 values as additional variables are added to the regression models and whether or not spurious variables are starting to enter the regression models. When only very small changes in R^2 values are taking place (e.g. $\leq 0{\cdot}01$), there is often little reason to continue the stepwise process. When α-values approach or exceed 0·01 and a large number of input variables are being considered, it is fairly common to start getting spurious variables in the regression. Such variables appear to have a small effect on the output variable which, in fact, is due to chance variation. In such situations, a natural stopping point may be just before spurious variables start being selected. Another possibility is to delete spurious variables from the regression model.

Table 4. Compact summary of stepwise regression analyses with rank-transformed data for the integrated release of U-233 to the accessible environment due to groundwater transport in fractures for scenario E1E2.

Step[a]	Var[b]	SRC[c]	R^{2d}
1	SOLALL	0·77	0·60
2	RTRDFU	−0·34	0·70
3	BHHYCND	0·29	0·79
4	TIMINTR	−0·20	0·83
5	BPPRES	0·16	0·86

[a] Steps in analysis.
[b] Variables listed in the order that they entered the analysis.
[c] Standardized regression coefficients (SRCs) for the variables in the final regression model.
[d] R^2 values that result with the entry of successive variables into the regression model.

When the input variables are uncorrelated, a display of the results of a stepwise regression analysis as shown in Table 3 contains a large amount of redundant information. A more compact display can be obtained by listing the variables in the order that they entered into the regression model, the R^2 values obtained with the entry of successive variables into the regression model, and the standardized regression coefficients for the variables contained in the final model. Table 4 shows what this summary looks like for the stepwise regression analysis presented in Table 3.

Monte Carlo analyses generate a mapping from analysis inputs to analysis results. Once this mapping is generated and saved, it can be explored with a wide variety of techniques. This section has discussed techniques based on scatterplots, regression, correlation, partial correlation, and stepwise regression. Other analysis techniques also exist.[192]

4. STOCHASTIC VERSUS SUBJECTIVE UNCERTAINTY

As indicated in the Introduction, there are two basic types of uncertainty present in a performance assessment: stochastic uncertainty and subjective uncertainty. To both facilitate the presentation of results and maintain a distinction between these two types of uncertainty, Kaplan & Garrick[11] have proposed that the results of a performance assessment be represented by a set R of the form

$$R = \{(E_i, fE_i, \mathbf{cE}_i), i = 1, \ldots, nE\} \quad (89)$$

where

E_i = a set of similar occurrences,
fE_i = frequency (or probability, as appropriate) at which occurrences in E_i take place,
$\mathbf{cE}_i$ = a vector of consequences associated with E_i

and

nE = number of sets E_i selected for consideration.

The frequencies fE_i represent stochastic uncertainty since each E_i is a set of occurrences that has a real possibility of taking place. Thus, the fE_i are properties of the system under study rather than a representation of the analysts' confidence in the information that they are using to perform the study. The division of all the possible occurrences into the E_i can be viewed as a form of importance sampling in which the E_i are the strata and the fE_i are frequencies at which the strata occur. When probabilities pE_i rather than frequencies fE_i are being used, then the pE_i are the strata probabilities.

The set R is the ultimate output variable of interest in a performance assessment. However, R can be difficult to display and examine when nE is large. As a result, the information contained in R is often summarized in exceedance frequency curves constructed from the pairs (fE_i, cE_i), $i = 1, \ldots, nE$, where cE_i is a specific consequence result contained in the vector $\mathbf{cE}_i$. An exceedance frequency curve displays the frequency at which individual consequence values are exceeded and is defined by the function

$$F(x) = \text{frequency at which } cE \text{ exceeds consequence value } x$$

$$= \sum_{j=i}^{nE} fE_j \tag{90}$$

where i is the smallest integer such that $cE_i > x$ and it is assumed that the cE_i have been ordered so that $cE_i \leq cE_{i+1}$ for $i = 1, \ldots, nE - 1$. When an exceedance frequency curve is used to summarize the results of a performance assessment, the output variable of interest is a function, and thus the presence of subjective uncertainties in the performance assessment leads to a distribtion of functions. When probabilities pE_i rather than frequencies fE_i are used, the construction process indicated in (90) produces a complementary cumulative distribution function (CCDF) rather than an exceedance frequency curve.

The presence of subjective uncertainties leads to a representation for R of the form

$$R(\mathbf{x}) = \{(E_i(\mathbf{x}), fE_i(\mathbf{x}), \mathbf{cE}_i(\mathbf{x})), \quad i = 1, \ldots, nE(\mathbf{x})\} \tag{91}$$

where $\mathbf{x}$ is a vector of imprecisely known quantities required in the performance assessment that produces R. Thus, there is subjective uncertainty in the value of $\mathbf{x}$.

As discussed in Section 3.1, ranges and distributions can be developed to characterize where the appropriate value for $\mathbf{x}$ is located. Uncertainty in $\mathbf{x}$ leads to uncertainty in R. Thus, rather than a single summary curve of the form indicated in (90), there will be a distribution of such curves for each consequence result contained in $\mathbf{cE}$. This distribution can be estimated by generating a sample

$$\mathbf{x}_j, \quad j = 1, \ldots, m \tag{92}$$

from $\mathbf{x}$ as described in Section 3.2. Propagation of this sample through the performance assessment as discussed in Section 3.3 produces a sequence of results of the form

$$R(\mathbf{x}_j) = \{(E_i(\mathbf{x}_j), fE_i(\mathbf{x}_j), \mathbf{cE}_i(\mathbf{x}_j)), \quad i = 1, \ldots, nE(\mathbf{x}_j)\} \tag{93}$$

for $j = 1, \ldots, m$.

Figure 13 provides an example of what the resultant distribution of curves might look like for a specific consequence result contained in $\mathbf{cE}$. Each curve summarizes results from a specific set $R(\mathbf{x}_j)$ and thus displays the effect of stochastic uncertainty. The distribution of curves arises from the uncertainty in the appropriate value to use for $\mathbf{x}$ and thus displays the effect of subjective uncertainty. Figure 13 presents uncertainty analysis results of the form discussed in Section 3.4, where the output variable is a function rather than a single number.

Sensitivity analysis techniques can be used to determine the importance of the individual input variables contained in $\mathbf{x}$ in causing the variation in the curves shown in Fig. 13. One way to do this is by selecting a sequence of consequence values $cE_1, cE_2, \ldots$ along the abscissa in Fig. 13. For each consequence value cE_k, there will be m estimates of the frequency at which cE_k is exceeded (i.e. the m exceedance frequencies that can be obtained by drawing a vertical line through cE_k and thus through the m exceedance frequency curves in Fig. 13). Partial correlation coefficients or standardized regression coefficients can be calculated between these exceedance frequencies and the input variables contained in $\mathbf{x}$. The resultant coefficients can then be plotted above the corresponding consequence values to produce a display of sensitivity analysis results of the form shown in Fig. 14. As can be seen in this figure, variables 1, 3 and 5 are important with respect to the exceedance frequencies for small consequence values and then decrease in importance with respect to the exceedance frequencies for large consequence values. The opposite pattern is shown by variables 2 and 4. Figure 14 presents sensitivity analysis results of the form discussed in Section 3.5 where the output variable is a function used to display the effects of stochastic uncertainty.

Additional discussion of the ideas presented in this section in the context of the EPA release limits for the geologic disposal of radioactive waste is available

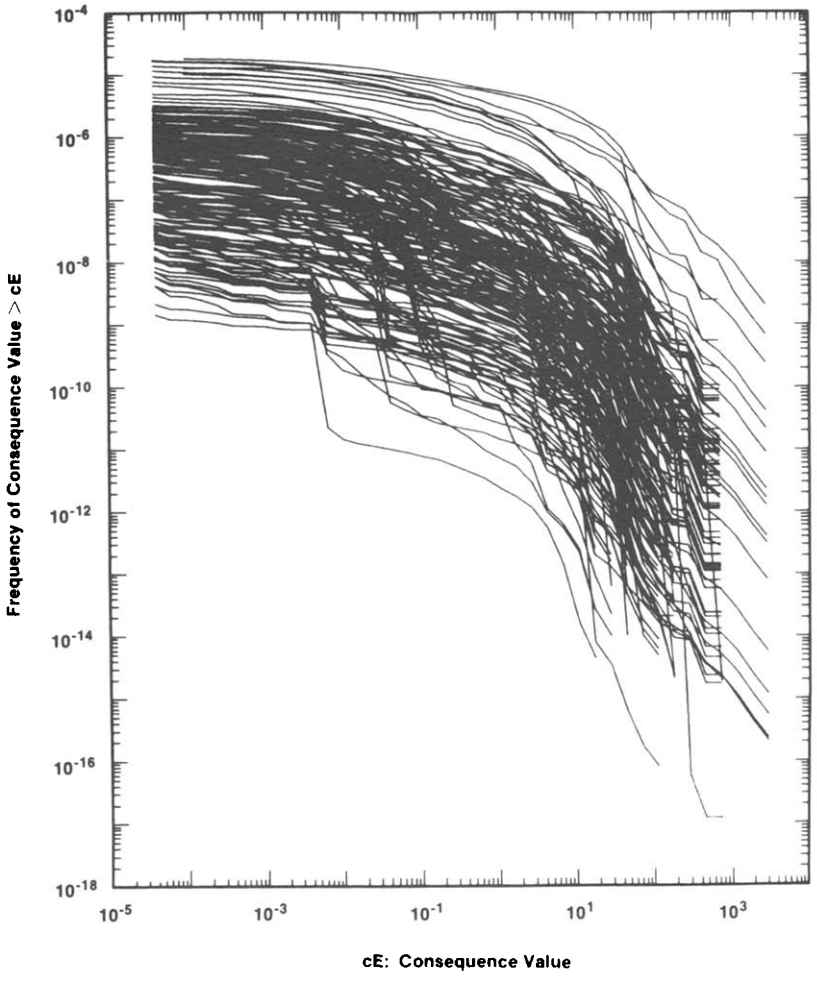

Fig. 13. Example distribution of exceedance frequency curves (Ref. 193, Fig. D.1).

elsewhere.[5,194] Extensive examples of analyses of the type described in this section are contained in the performance assessments conducted in support of the U.S. Nuclear Regulatory Commission's recent reassessment of the risk from commercial nuclear power plants (NUREG-1150)[68,69,193,195–199] and also in the 1991 WIPP performance assessment.[200]

5 DISCUSSION

The treatment of uncertainties in the analysis of complex systems is a topic of both importance and much current interest.[16,201–205] This paper has reviewed uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. The following techniques were considered: differential analysis, Monte Carlo analysis, response surface methodology, and Fourier amplitude sensitivity test. These approaches are quite general, and their development and application are not tied to performance assessment for radioactive waste disposal or any other specific area of study.

Each of the approaches has strengths and weaknesses. Thus, it would be a mistake to try to argue that one approach will always be superior to the others. Nonetheless, it is felt that Monte Carlo analysis is the most widely applicable approach for the analysis of problems in performance assessment and in other areas that involve complex models and large uncertaities. Desirable properties of Monte Carlo analysis include (1) full stratification over the range of each variable, (2) estimation of distribution functions without the use of intermediate models, (3) conceptual simplicity, (4) flexibility in adapting to specific analysis situations, including sequences of linked models, and (5) availability of a variety of sensitivity analysis techniques.

Regardless of the approach to uncertainty and sensitivity analysis selected for use, it is necessary to characterize the uncertainty in analysis inputs. When performed carefully, this is typically the most expensive part of an analysis. Complex analyses such as performance assessments are usually performed by teams of individuals. It is essential that everyone involved in characterizing the uncertainty in a variable

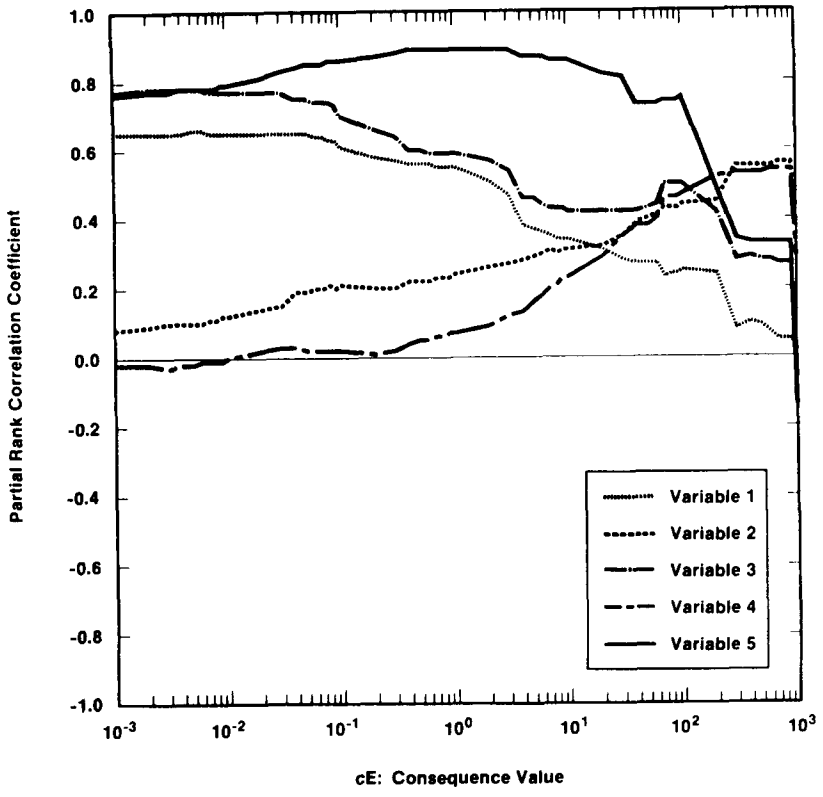

Fig. 14. Example sensitivity analysis for the exceedance frequency curves in Fig. 13 (Ref. 193, Fig. 5.1-8.)

clearly understand how that variable is used in the analysis. Otherwise, the distribution developed for the variable may bear little relationship to the actual uncertainties present in the analysis.

A common error is to assume that the observed variability in values measured for a quantity on one scale is the same as the uncertainty in the corresponding model input for use on a much larger scale. For example, the variability in measured values for distribution coefficients obtained from samples taken throughout a region is not the same as the uncertainty in the values for "effective" distribution coefficients used to model radionuclide transport over the entire region. The conversion from variability on one scale to uncertainty on another scale is an area of performance assessment for radioactive waste disposal that would benefit from additional investigation.

Uncertainty analysis and sensitivity analysis are closely linked. In particular, sensitivity analysis cannot be decoupled from uncertainty analysis since the impact of individual variables on model predictions depends on both their use in the model and on the possible values that they and other variables can take on. Thus, inappropriate characterizations of the uncertainty in analysis inputs can produce misleading results in both uncertainty analysis and sensitivity analysis. As long as the uncertainty/sensitivity analysis procedures selected for use are not truly inappropriate for the problem under study, uncertainty characterization is the part of the analysis that probably has the greatest impact on the quality of the final uncertainty/sensitivity results.

As uncertainties are characterized, it is important to maintain a distinction between stochastic uncertainty, which results from the fact that a system can behave in many different ways, and subjective uncertainty, which results from the existing state of knowledge with respect to the correctness of the assumptions used in an analysis. When this distinction is not maintained, it is difficult, if not impossible, to meaningfully interpret the results obtained in an uncertainty/sensitivity study conducted as part of a performance assessment. The ordered triple representation proposed by Kaplan & Garrick[11] provides both a conceptual structure for the organization of a performance assessment and a basis for the separation of stochastic and subjective uncertainty.

It is important to recognize that there are no truly global analyses that incorporate all sources of uncertainty. Rather, all analyses are conditional on assumptions made by the analysts that performed them. Of course, some analyses are more local than others. For example, a differential analysis based on a first-order Taylor series expansion is more local than a Monte Carlo study that stratifies across the range of each variable under consideration. However, the results of both studies are conditional on assumptions: models selected for use, input variables assumed to

have fixed values, input variables assumed to have uncertain values, distributions assigned to uncertain input variables, ... There are no analyses that consider everything. The best that can be hoped for is a consensus among the interested parties that an appropriate boundary has been placed around the problem under consideration. However, once the boundaries of a study are agreed upon, it is important that uncertainties be characterized and propagated in a way that is consistent with the starting assumptions. In particular, it is important to maintain a distinction between stochastic and subjective uncertainty as discussed in the preceding paragraph.

Most model inputs considered in uncertainty/sensitivity studies are real valued. However, more complex entities are also possible. For example, flow fields and time-dependent recharge rates are uncertain inputs to performance assessments that are functions. Typically, such inputs are incorporated into uncertainty/sensitivity studies by using real-valued surrogates (e.g. travel time or average recharge). As a result, the incorporation of such inputs into uncertainty/sensitivity studies is somewhat contrived. The possibility of more appropriate procedures for use with uncertain functions and other complex inputs merits consideration. However, to be helpful, such procedures would have to produce results that can be presented to and understood by fairly broad audiences within the technical community.

An important uncertainty in many analyses involves the choice between different models (i.e. model uncertainty). When only a few models are involved, model uncertainty can be studied with *ceteris paribus* procedures or incorporated into a formal study with an integer-valued variable that is used as a pointer to the individual models. However, modeling uncertainty remains an important component of performance assessment that is often characterized in an unsatisfying manner. Problems in the treatment of model uncertainty include both the probabilistic characterization of this uncertainty and the mechanics by which it can be incorporated into a performance assessment. Thus, procedures for the treatment of model uncertainty in performance assessment certainly merit additional consideration.

More sophisticated procedures to search for patterns would be helpful. After all, both scatterplots and regression-based techniques are rather simple procedures for identifying patterns. In many complex analyses, there are actually several different regimes of behavior that involve different relationships between the input and output variables. 'Intelligent' procedures for recognizing such regimes of behavior and then investigating sensitivity conditional on them are worth exploring. In concept, such procedures could both explore already existing calculations and also provide guidance on combinations of input values for which additional calculations should be performed.

A difficult, and unfortunately rather common, analysis situation involves 'censored' results. In this situation, many of the predicted values for the output variable of interest are zero. These zero values can obscure, or censor, the effects of important input variables. Often, specific subranges of one or a few input variables result in these zero values. In some analyses, the cause of these zero values can be determined by the examination of scatterplots. However, in other analyses, what is causing these zeros and what would be important in their absence are not readily discernible. The identification or development of procedures for studying results containing a large number of zeros in sampling-based sensitivity studies would be very beneficial. This type of analysis problem often arises in performance assessments for radioactive waste disposal due to both uncertainties in whether or not a release takes place from the engineered facility (i.e. the waste repository itself) and the complex transport mechanisms involved once such a release does take place.

A formal characterization of the bias introduced by the Iman/Conover restricted pairing technique[53] would be nice, as would additonal characterizations of the relative performance of random sampling and Latin hypercube sampling. However, it is unlikely that such characterizations would alter the widely held views that Latin hypercube sampling generally outperforms random sampling and that the use of the restricted pairing technique does not introduce a significant bias.

The two standard reasons for performing an uncertainty/sensitivity study are to determine if the uncertainty in a predicted quantity falls within a specified set of boundaries (e.g. below the EPA release limits for radionuclide releases from a waste repository to the accessible environment) and to provide guidance on the investment of research efforts to reduce uncertainty (e.g. should research efforts be concentrated on obtaining better characterizations for leach rates, or solubility limits, or distribution coefficients, or something else). However, sampling-based uncertainty/sensitivity analysis procedures also provide a powerful tool for model verification.

The models used in performance assessment for radioactive waste disposal and other involved processes are typically very complex with large numbers of both input and output variables. As a result, it is often difficult to determine whether the predictions made by a model (i.e. a computer program) are consistent with the mathematical relationships that are supposedly implemented within the model. Such models often contain numerical procedures that are too complicated to check by hand, with the result that individual calculations cannot be directly verified. Rather, a number of calculations

must be performed with different sets of input and the resultant patterns of change checked for reasonableness. One possibility is to perform this exploration with the *ceteris paribus* approach. However, these patterns are often hard to identify and typically only a very limited number of input variations can be considered. With a sampling-based approach, it is possible to (1) consider a large number of inputs, (2) generate sample elements that involve extensive combinations of these inputs, and (3) screen a large number of different output variables for the impacts of the input variables. Regression-based sensitivity analysis techniques can reveal relationships between input and output variables that cannot be detected by a simple visual screening of the results. The preceding procedure provides a very powerful tool for model verification and by itself is a sufficient justification for sampling-based sensitivity analysis.

ACKNOWLEDGEMENTS

This work was performed for Sandia National Laboratories and the United States Department of Energy under contract number DE-ACO4-76DP00789.

The support provided by D. R. (Rip) Anderson, S. G. Bertram-Howery, M. G. Marietta, R. P. Rechard, and other members of the 1990 Performance Assessment team for the Waste Isolation Pilot Plant is gratefully acknowledged. Particular appreciation is due J. W. Garner, A. P. Gilkey, R. D. McCurley, and D. K. Rudeen for performing the calculations for the example uncertainty/sensitivity analysis. Special recognition is extended to R. L. Iman for his work on many of the techniques described in this review. Extensive comments by anonymous reviewers contributed substantially to the quality of the review. The editorial support provided by the staff of Tech. Reps., Inc., with special thanks to P. N. Swift, is also gratefully appreciated. Finally, thanks are extended to E. J. Bonano for the invitation to write this paper.

REFERENCES

1. Bertram-Howery, S. G., Marietta, M. G., Rechard, R. P., Swift, P. N. Anderson, D. R., Baker, B., Bean, J., Beyeler, W., Brinster, K. F., Guzowski, R. V., Helton, J. C., McCurley, R. D., Rudeen, D. K. & Schreiber, J., *Preliminary Comparison with 40 CFR Part 191, Subpart B for the Waste Isolation Pilot Plant, December 1990.* SAND90-2347, Sandia National Laboratories, Albuquerque, NM, 1990.
2. Bonano, E. J., Davis, P. A., Shipers, L. R., Brinster, K. F., Beyeler, W. E., Updegraff, C. D., Shepherd, E. R., Tilton, L. M. & Wahi, K. K., *Demonstration of a Performance Assessment Methodology for High-Level Radioactive Waste Disposal in Basalt Formations.* NUREG/CR-4759, SAND86-2325, Sandia National Laboratories, Albuquerque, NM, 1989.
3. Campbell, J. E. & Cranwell, R. M., Performance Assessment of Radioactive Waste Repositories. *Science,* **239** (1988) 1389–92.
4. Cranwell, R. M., Campbell, J. E., Helton, J. C., Iman, R. L., Longsine, D. E., Ortiz, N. R., Runkle, G. E. & Shortencarier, M. J., *Risk Methodology for Geologic Disposal of Radioactive Waste: Final Report.* NUREG/CR-2452, SAND81-2573, Sandia National Laboratories, Albuquerque, NM, 1987.
5. Helton, J. C., Garner, J. W., McCurley, R. D. & Rudeen, D. K., *Sensitivity Analysis Techniques and Results for Performance Assessment at the Waste Isolation Pilot Plant.* SAND90-7103, Sandia National Laboratories, Albuquerque, NM, 1991.
6. Zimmerman, D. A., Wahi, K. K., Gutjahr, A. L. & Davis, P. A., *A Review of Techniques for Propagating Data and Parameter Uncertainties in High-Level Radioactive Waste Repository Performance Models.* NUREG/CR-5393, SAND89-1432, Sandia National Laboratories. Albuquerque, NM, 1990.
7. Doctor, P. G., Sensitivity and Uncertainty Analyses for Performance Assessment Modeling. *Engineering Geology,* **26** (1989) 411–29.
8. Bonano, E. J. & Cranwell, R. M., Treatment of Uncertainties in the Performance Assessment of Geologic High-Level Radioactive Waste Repositories. *Mathematical Geology,* **20** (1988) 543–65.
9. Kim, T. W., Chang, S. H. & Lee, B. H., Uncertainty and Sensitivity Analyses in Evaluating Risk of High Level Waste Repository. *Radioactive Waste Management and the Nuclear Fuel Cycle,* **10** (1988) 321–56.
10. Nuclear Energy Agency, *Uncertainty Analysis for Performance Assessments of Radioactive Waste Disposal Systems.* Organization for Economic Cooperation and Development, Paris, 1987.
11. Kaplan, S. & Garrick, B. J., On the Quantitative Definition of Risk. *Risk Analysis,* **1** (1981) 11–27.
12. Vesely, W. E. & Rasmuson, D. M., Uncertainties in Nuclear Probabilistic Risk Analyses. *Risk Analysis,* **4** (1984) 313–22.
13. Stern, E. & Tadmor, J., Uncertainty Bands in CCDF Risk Curves; Their Importance in Decision Making Processes. In *Proceedings of ANS/ENS International Topical Meeting on Probabilistic Safety Methods and Applications,* San Francisco, CA, USA, February 24–March 1, 1985. American Nuclear Society, La Grange Park, II, 1985, Vol. 1, pp. 22-1–22-9.
14. Paté-Cornell, M. E. Probability and Uncertainty in Nuclear Safety Decisions. *Nuclear Engineering and Design,* **93** (1986) 319–27.
15. Parry, G. W., On the Meaning of Probability in Probabilistic Safety Assessment. *Reliability Engineering and System Safety,* **23** (1988) 309–14.
16. International Atomic Energy Agency, *Evaluating the Reliability of Predictions Made Using Environmental Transfer Models.* Safety Series Report No. 100, IAEA Vienna, 1989.
17. Hahn, G. J. & Shapiro, S. S., *Statistical Models in Engineering.* Wiley, New York, 1967.
18. Tukey, J. W., *The Propagation of Errors, Fluctuations and Tolerances: Basic Generalized Formulas.* Report No. 10, Statistical Techniques Research Group, Princeton University, Princeton, NJ, 1957.
19. Tukey, J. W. *The Propagation of Errors, Fluctuations and Tolerances: Supplementary Formulas.* Report No.

11, Statistical Techniques Research Group, Princeton University, Princeton, NJ, 1957.
20. Tukey, J. W., *The Propagation of Errors, Fluctuations and Tolerances: An Exercise in Partial Differentiation.* Report No. 12. Statistical Techniques Research Group, Princeton University, Princeton, NJ, 1957.
21. Cacuci, D. G., Sensitivity Theory for Nonlinear Systems. I. Nonlinear Functional Analysis Approach. *Journal of Mathematical Physics*, **22** (1981) 2794–2802.
22. Cacuci, D. G., Sensitivity Theory for Nonlinear Systems. II. Extensions to Additional Classes of Responses. *Journal of Mathematical Physics*, **22** (1981) 2803–12.
23. Cacuci, D. G., Weber, C. F., Oblow, E. M. & Marable, J. H., Sensitivity Theory for General Systems of Nonlinear Equations. *Nuclear Science and Engineering*, **75** (1980) 88–110.
24. Dougherty, E. P. & Rabitz, H., A Computational Algorithm for the Green's Function Method of Sensitivity Analysis in Chemical Kinetics. *International Journal of Chemical Kinetics*, **11** (1979) 1237–48.
25. Dougherty, E. P., Hwang, J. T. & Rabitz, H. Further Developments and Applications of the Green's Function Method of Sensitivity Analysis in Chemical Kinetics. *Journal of Chemical Physics*, **71** (1979) 1794–1808.
26. Hwang, J. T., Dougherty, E. P., Rabitz, S. & Rabitz, H., The Green's Function Method of Sensitivity Analysis in Chemical Kinetics. *Journal of Chemical Physics*, **69** (1978) 5180–91.
27. Oblow, E. M., Pin, F. G. & Wright, R. Q., Sensitivity Analysis Using Computer Calculus: A Nuclear Waste Isolation Example. *Nuclear Science and Engineering*, **94** (1986) 46–65.
28. Pin, F. G., Worley, B. A., Oblow, E. M., Wright, R. Q. & Harper, W. V., An Automated Sensitivity Analysis Procedure for the Performance Assessment of Nuclear Waste Isolation Systems. *Nuclear and Chemical Waste Management*, **6** (1986) 255–63.
29. Worley, B. A. & Horwedel, J. E., *A Waste Package Performance Assessment Code with Automated Sensitivity-Calculation Capability.* ORNL/TM-9976, Oak Ridge National Laboratory, Oak Ridge, TN, 1986.
30. Oblow, E. M., *GRESS: Gradient Enhanced Software System, Version D, User's Guide.* ORNL/TM-9658, Oak Ridge National Laboratory, Oak Ridge, TN, 1985.
31. Griewank, A. & Corliss, G. (Eds), *Automatic Differentiation of Algorithms: Theory, Implementation, and Application.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1991.
32. Ronen, Y. (Ed.), *Uncertainty Analysis.* CRC Press, Boca Raton, FL, 1988.
33. Lewins, J. & Becker, M. (Eds), Sensitivity and Uncertainty Analysis of Reactor Performance Parameters. *Advances in Nuclear Science and Technology*, **14**. Plenum Press, New York, 1982.
34. Frank, P. M., *Introduction to System Sensitivity Theory.* Academic Press, New York, 1978.
35. Dickinson, R. P. & Gelinas, R. J., Sensitivity Analysis of Ordinary Differential Equation Systems—A Direct Method. *Journal of Computational Physics*, **21** (1976) 123–43.
36. Tomovic, R. & Vukobratovic, M., *General Sensitivity Theory.* Elsevier, New York, 1972.
37. McKay, M. D., Conover, W. J. & Beckman, R. J., A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, **21** (1979) 239–45.
38. Helton, J. C., Iman, R. L., Johnson, J. D. & Leigh, C. D., Uncertainty and Sensitivity Analysis of a Model for Multicomponent Aerosol Dynamics. *Nuclear Technology*, **73** (1986) 320–42.
39. Helton, J. C., Iman, R. L. & Brown, J. B., Sensitivity Analysis of the Asymptotic Behavior of a Model for the Environmental Movement of Radionuclides. *Ecological Modelling*, **28** (1985) 243–78.
40. Hendry, D. F., Monte Carlo Experimentation in Econometrics. In *Handbook of Econometrics*, ed. Z. Groiliches & M. D. Intriligator. Elsevier, New York, 1984, Vol. II, pp. 937–76.
41. Fedra, K., *Environmental Modelling under Uncertainty: Monte Carlo Simulation.* RR-83-28, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1983.
42. Gardner, R. H. & O'Neill, R. V., Parameter Uncertainty and Model Predictions: A Review of Monte Carlo Results. In *Uncertainty and Forecasting of Water Quality*, ed. M. B. Beck & G. Van Straten. Springer-Verlag, New York, 1983, pp. 245–57.
43. Iman, R. L. & Conover, W. J., *Sensitivity Analysis Techniques: Self-Teaching Curriculum.* NUREG/CR-2350, SAND81-1978, Sandia National Laboratories, Albuquerque, NM, 1982.
44. Iman, R. L. & Conover, W. J., Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment. *Communications in Statistics*, **A9** (1980) 1749–1842. Rejoinder to Comments, *ibid.*, 1863–74.
45. Iman, R. L., Helton, J. C. & Campbell, J. E., An Approach to Sensitivity Analysis of Computer Models, Part 1. Introduction, Input Variable Selection and Preliminary Variable Assessment. *Journal of Quality Technology*, **13** (1981) 174–83.
46. Iman, R. L., Helton, J. C. & Campbell, J. E., An Approach to Sensitivity Analysis of Computer Models, Part 2. Ranking of Input Variables, Response Surface Validation, Distribution Effect and Technique Synopsis. *Journal of Quality Technology*, **13** (1981) 232–40.
47. Schwarz, G. & Hoffman, F. O. Imprecision of Dose Predictions for Radionuclides Released to the Environment: An Application of a Monte Carlo Simulation Technique. *Environment International*, **4** (1980) 289–97.
48. Box, G. E. P. & Draper, N. R., *Empirical Model-Building and Response Surfaces.* Wiley, New York, 1987.
49. Kleijnen, J. P. C., *Statistical Tools for Simulation Practitioners.* Marcel Dekker, New York, 1987.
50. Myers, R. H., *Response Surface Methodology.* Allyn and Bacon, Boston, MA, 1971.
51. Sacks, J., Schiller, S. S. & Welch, W. J., Designs for Computer Experiments. *Technometrics*, **31** (1989) 41–7.
52. Sacks, J., Welch, W. J., Mitchell, T. J. & Wynn, H. P., Design and Analysis of Computer Experiments. *Statistical Science*, **4** (1989) 409–23.
53. Iman, R. L. & Conover, W. J., A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables. *Communications in Statistics*, **B11** (1982) 311–34.
54. Vaurio, J. K. & Mueller, C., Probabilistic Analysis of LMFBR Accident Consequences with Response Surface Techniques. *Nuclear Science and Engineering*, **65** (1978) 401–13.

55. Chu, T. L. & Apostolakis, G. E., Assessment of the Uncertainties Associated with the Core Uncovery Time in TMI-Type Accidents. *Reliability Engineering,* **8** (1984) 23–56.
56. Friedman, J. H., Multivariate Adaptive Regression Splines. *The Annals of Statistics,* **19** (1991) 1–67.
57. Olivi, L., *Response Surface Methodology Handbook for Nuclear Reactor Safety.* EUR 9600 EN, Commission of the Europoean Communities, Luxembourg, 1986.
58. Morton, R. H., Response Surface Methodology. *Mathematical Scientist,* **8** (1983) 31–52.
59. Mead, R. & Pike, D. J., A Review of Response Surface Methodology from a Biometric Viewpoint. *Biometrics,* **31** (1975) 803–51.
60. Kleijnen, J. P. C., *Statistical Techniques in Simulation,* Parts 1 and 2. Marcel Dekker, New York, 1974.
61. Liepmann, D. & Stephanopoulos, G., Development and Global Sensitivity Analysis of a Closed Ecosystem Model. *Ecological Modelling,* **30** (1985) 13–47.
62. McRae, G. J., Tilden, J. W. & Seinfeld, J. H., Global Sensitivity Analysis—A Computational Implementation of the Fourier Amplitude Sensitivity Test (FAST). *Computers and Chemical Engineering,* **5** (1981) 15–25.
63. Cukier, R. I., Levine, H. B. & Shuler, K. K., Nonlinear Sensitivity Analysis of Multiparameter Model Systems. *Journal of Computational Physics,* **26** (1978) 1–42.
64. Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G. & Schiably, J. H., Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients, I. Theory. *Journal of Chemical Physics,* **59** (1973) 3873–8.
65. Schaibly, J. H. & Shuler, K. E., Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients, II. Applications. *Journal of Chemical Physics,* **59** (1973) 3879–88.
66. U.S. Environmental Protection Agency, Environmental Standards for the Management and Disposal of Spent Nuclear Fuel, High-Level and Transuranic Radioactive Waste; Final Rule. 40 CFR Part 191, *Federal Register,* **50** (1985) 38066–89.
67. Iman, R. L., Statistical Methods for Including Uncertainties Associated with Geologic Isolation of Radioactive Waste Which Allow for a Comparison with Licensing Criteria. In *Proceedings of the Symposium on Uncertainties Associated with the Regulation of the Geologic Disposal of High Level Radioactive Waste,* Gatlinburg, TN, March 9–13, 1981, ed. D. C. Kocher. NUREG/CP-0022, CONF-810372, Oak Ridge National Laboratory, Oak Ridge, TN, 1982, pp. 145–57.
68. Gorham, E. D., *et al., Evaluation of Severe Accident Risks: Methodology for the Containment, Source Term, Consequence and Risk Integration Analysis.* NUREG/CR-4551, SAND86-1309, Vol. 1, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1993.
69. U.S. Nuclear Regulatory Commission, *Severe Accident Risks:* An Assessment for Five U.S. Nuclear Power Plants. NUREG-1150, U.S. Nuclear Regulatory Commission, Washington, D.C., 1990.
70. Helton, J. C., Griesmeyer, J. M., Haskin, F. E., Iman, R. L., Amos, C. N. & Murfin, W. B., Integration of the NUREG-1150 Analyses: Calculation of Risk and Propagation of Uncertainties. In *Proceedings of the Fifteenth Water Reactor Safety Research Information Meeting.* U.S. Nuclear Regulatory Commission, Washington, D.C., 1988, pp. 151–76.
71. Wu, Y.-T., Journel, A. G., Abramson, L. R. & Nair, P. K., *Uncertainty Evaluation Methods for Waste Package Performance Assessment.* NUREG/CR-5639, U.S. Nuclear Regulatory Commission, Washington, D.C., 1991.
72. Rish, W. R. & Marnicio, R. J., *Review of Studies Related to Uncertainty in Risk Analysis.* ORNL/TM-10776, Oak Ridge National Laboratory, Oak Ridge, TN, 1988.
73. Fischer, F. & Ehrhardt, J., *Uncertainty Analysis with a View Towards Applications in Accident Consequence Assessments.* KfK 3906, Kernforschungszentrum Karlsruhe, Karlsruhe, FRD, 1985.
74. Iman, R. L. & Helton, J. C., Overview of Methods for Uncertainty Analysis and Sensitivity Analysis in Probabilistic Risk Assessment. In *Proceedings of the ANS/ENS International Topical Meeting on Probabilistic Safety Methods and Applications,* San Francisco, CA, USA, 24 February–1 March 1985. American Nuclear Society, La Grange Park, IL, 1985, Vol. 1, pp. 15-1–15-11.
75. Hendrickson, R. G., *Survey of Sensitivity Analysis Methodology.* NBSIR 84-2814, U.S. National Bureau of Standards, Washington, D.C., 1984.
76. Rabitz, H., Kramer, M. & Dacol, D., Sensitivity Analysis in Chemical Kinetics. *Annual Reviews of Physical Chemistry,* **34** (1983) 410–61.
77. Cox, D. C. & Baybutt, P., Methods for Uncertainty Analysis: A Comparative Survey. *Risk Analysis,* **1** (1981) 251–8.
78. Rose, K. A. & Swartzman, G. S., *A Review of Parameter Sensitivity Methods Applicable to Ecosystem Models.* NUREG/CR-2016, University of Washington, Seattle, WA, 1981.
79. Tilden, J. W., Costanza, V., McRae, G. J. & Seinfeld, J. H., Sensitivity Analysis of Chemically Reacting Systems. In *Modelling of Chemical Reaction Systems,* ed. K. H. Ebert, P. Deuflhard & W. Jaeger. Springer-Verlag, New York, 1981, pp. 69–91.
80. Mazumdar, M., Marshall, J. A. & Chay, S. C., Propagation of Uncertainties in Problems of Structural Reliability. *Nuclear Engineering and Design,* **50** (1978) 163–7.
81. Mazumdar, M., *et al., Methodology Development for Statistical Evaluation of Reactor Safety Analyses.* EPRI-NP-194, Electric Power Research Institute, Palo Alto, CA, 1976.
82. Mazumdar, M., *et al., Review of the Methodology for Statistical Evaluation of Reactor Safety Analysis.* EPRI 309, Electric Power Research Institute, Palo Alto, CA, 1975.
83. Mishra, S. & Parker, J. C., Effects of Parameter Uncertainty on Predictions of Unsaturated Flow. *Journal of Hydrology,* **108** (1989) 19–33.
84. Doctor, P. G., Jackson, E. A. & Buchanan, J. A., *A Comparison of Uncertainty Analysis Methods Using a Groundwater Flow Model.* PNL-5649, Pacific Northwest Laboratory, Richland, WA, 1988.
85. Iman, R. L. & Helton, J. C., An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models. *Risk Analysis,* **8** (1988) 71–90.
86. Kim, T. W., Chang, S. H. & Lee, B. H., Comparative Study on Uncertainty and Sensitivity Analysis and Application to LOCA Model. *Reliability Engineering and System Safety,* **21** (1988) 1–26.
87. Maerker, R. E., *Comparison of Results Based on a Deterministic Versus a Statistical Sensitivity Analysis.*

ORNL/TM-10773, Oak Ridge National Laboratory, Oak Ridge, TN, 1988.
88. Seaholm, S. K., Ackerman, E. & Wu, S. C., Latin Hypercube Sampling and the Sensitivity Analysis of a Monte Carlo Epidemic Model. *International Journal of BioMedical Computing*, **23** (1988) 97–112.
89. Sykes, J. F. & Thomson, N. R., Parameter Identification and Uncertainty Analysis for Variably Saturated Flow. *Advances in Water Resources*, **11** (1988) 185–91.
90. Obray, C. D., Wright, C. C. & Baldwin, S. J., A Comparative Study of Monte-Carlo Simulation and Taylor Series Approaches to the Derivation of Uncertainty Estimates: A Case Study on Compact Heat Exchangers. In *Modelling Under Uncertainty*, ed. S. B. Jones & D. G. S. Davies, Institute of Physics Conference Series No. 80, Institute of Physics, Bristol, UK, 1986, pp. 243–52.
91. Downing, D.J., Gardner, R. H. & Hoffman, F. O., An Examination of Response-Surface Methodologies for Uncertainty Analysis in Assessment Models. *Technometrics*, **27** (1985) 151–63.
92. Iman, R. L. & Helton, J. C., *A Comparison of Uncertainty and Sensitivity Analysis Techniques for Computer Models*. NUREG/CR-3904, SAND84-1461, Sandia National Laboratories, Albuquerque, NM, 1985.
93. Jacobson, E. A., Freshley, M. D. & Dove, F. H., *Investigations of Sensitivity and Uncertainty in Some Hydrologic Models of Yucca Mountain and Vicinity*. SAND84-7212, Sandia National Laboratories, Albuquerque, NM, 1985.
94. Uliasz, M., Comparison of the Sensitivity Analysis Methods for Meteorological Models. *Zeitschrift fuer Meteorologie*, **6** (1985) 340–8.
95. Harper, W. V. & Gupta, S. K., *Sensitivity/Uncertainty Analysis of a Borehole Scenario Comparing Latin Hypercube Sampling and Deterministic Sensitivity Approaches*. BMI/ONWI-516, Battelle Memorial Institute, Columbus, OH, 1983.
96. Montgomery, R. H., Lee, V. D. & Reckhow, K. H., Predicting Variability in a Lake Ontario Phosphorus Model. *Journal of Great Lakes Research*, **9** (1983) 74–82.
97. Rose, K. A., A Simulation Comparison and Evaluation of Parameter Sensitivity Methods Applicable to Large Models. In *Analysis of Ecological Systems: State-of-the-Art in Ecological Modeling*, ed. W. K. Lauenroth *et al.*, Elsevier, New York, USA, 1982, pp. 129–40.
98. Ahmed, S., Metcalf, D. R. & Pegram, J. W., Uncertainty Propagation in Probabilistic Risk Assessment: A Comparative Study. *Nuclear Engineering and Design*, **68** (1981) 1–3.
99. Gardner, R. H., O'Neill, R. V., Mankin, J. B. & Carney, J. H., A Comparison of Sensitivity Analysis and Error Analysis Based on a Stream Ecosystem Model. *Ecological Modelling*, **12** (1981) 173–90.
100. Scavia, D., Powers, W. F., Canale, R. P. & Moody, J. L., Comparison of First-Order Error Analysis and Monte Carlo Simulation in Time-Dependent Lake Eutrophication Models. *Water Resources Research*, **17** (1981) 1051–9.
101. Cox, N. D., Comparison of Two Uncertainty Analysis Methods. *Nuclear Science and Engineering*, **64** (1977) 258–65.
102. Burns, J. R., Error Analysis of Nonlinear Simulations: Applications to World Dynamics. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-5** (1975) 331–40.
103. Student. On the Probable Error of a Mean. *Biometrika*, **6** (1908) 1–25.
104. Yule, G. U., Why Do We Sometimes Get Nonsense-Correlations Between Time-Series?—A Study in Sampling and the Nature of Time-Series. *Journal of the Royal Statistical Society*, **89** (1926) 1–64.
105. Orcutt, G. H. & Cochrane, D., A Sampling Study of the Merits of Autoregressive and Reduced Form Transformations in Regression Analysis. *Journal of the American Statistical Association*, **44** (1949) 356–72.
106. Teichrowe, D., A History of Distribution Sampling Prior to the Era of the Computer and Its Relevance to Simulations. *Journal of the American Statistical Association*, **60** (1965) 27–49.
107. Metropolis, N. & Ulam, S., The Monte Carlo Method. *Journal of the American Statistical Association*, **44** (1949) 335–41.
108. Kalos, M. H. & Whitlock, P. A., *Monte Carlo Methods*. Wiley, New York, 1986.
109. Rief, H., Gelbard, E. M., Schaefer, R. W. & Smith, K. S., Review of Monte Carlo Techniques for Analyzing Reactor Perturbations. *Nuclear Science and Engineering*, **92** (1986) 289–97.
110. Turner, J. E., Wright, H. A. & Hamm, R. N., A Monte Carlo Primer for Health Physicists. *Health Physics*, **48** (1985) 717–33.
111. Halton, J. H., A Retrospective and Prospective Survey of the Monte Carlo Method. *SIAM Review*, **12** (1970) 1–63.
112. Hammersley, J. M. & Handscomb, D. C., *Monte Carlo Methods*. Methuen, London, 1964.
113. Campbell, J. E., Dillon, R. T., Tierney, M. S., Davis, H. T., McGarth, P. E., Pearson, F. J., Shaw, H. R., Helton, J. C. & Donath, F. A., *Risk Methodology for Geologic Disposal of Radioactive Waste: Interim Report*. NUREG/CR-0458, SAND78-0029, Sandia National Laboratories, Albuquerque, NM, 1978.
114. Iman, R. L., Helton, J. C. & Campbell, J. E., *Risk Methodology for Geologic Disposal of Radioactive Waste: Sensitivity Analysis Techniques*. NUREG/CR-0394, SAND78-0912, Sandia National Laboratories, Albuquerque, NM, 1978.
115. Campbell, J. E., Iman, R. L. & Reeves, M., *Risk Methodology for Geologic Disposal of Radioactive Waste: Transport Model Sensitivity Analysis*. NUREG/CR-1377, SAND80-0644, Sandia National Laboratories, Albuquerque, NM, 1980.
116. Helton, J. C. & Iman, R. L., *Risk Methodology for Geologic Disposal of Radioactive Waste: Sensitivity Analysis of the Environmental Transport Model*. NUREG/CR-1636, Vol. 2, SAND79-1393, Sandia National Laboratories, Albuquerque, NM, 1980.
117. Helton, J. C., Brown, J. B. & Iman, R. L. *Risk Methodology for Geologic Disposal of Radioactive Waste: Asymptotic Properties of the Environmental Transport Model*. NUREG/CR-1636, Vol. 3, SAND79-1908, Sandia National Laboratories, Albuquerque, NM, 1981.
118. Cranwell, R. M. & Helton, J. C., Uncertainty Analysis Associated with Radioactive Waste Disposal: A Discussion Paper. In *Proceedings of the DOE Statistical Symposium*, October 29–31, 1980, Berkeley, CA. CONF-801045, Oak Ridge National Laboratory, Oak Ridge TN, 1981, pp. 186–96.
119. Cranwell, R. M. & Helton, J. C., Uncertainty Analysis for Geologic Disposal of Radioactive Waste. In

Proceedings of the Symposium on Uncertainties Associated with the Regulation of the Geologic Disposal of High-Level Radioactive Waste, Gatlinburg, TN, March 9–13, 1981, ed. D. C. Kocher. NUREG/CP-0022, CONF-810372, Oak Ridge National Laboratory, Oak Ridge, TN, 1982, pp. 131–43.

120. Rechard, R. P., Gilkey, A. P., Rudeen, D. K., Rath, J. S., Beyeler, W., Blaine, R., Iuzzolino, H. J. & McCurley, R. D., *User's Reference Manual for CAMCON: Compliance Assessment Methodology Controller Version 3.0.* SAND90-1983, Sandia National Laboratories, Albuquerque, NM, 1991.
121. Iman, R. L. & Shortencarier, M. J., *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, Sandia National Laboratories, Albuquerque, NM, 1984.
122. Voss, C. I., *SUTRA (Saturated–Unsaturated TRAnsport): A Finite-Element Simulation Model for Saturated–Unsaturated, Fluid-Density-Dependent Ground-Water Flow with Energy Transport or Chemically Reactive Single-Species Solute Transport.* U.S. Geological Survey National Center, Reston, VA, 1984.
123. Roache, P. J., Blaine, R. L. & Baker, B. L. *SECO 2DH User's Manual, Version 2.1.* SAND90-7096, Sandia National Laboratories, Albuquerque, NM, 1991.
124. Huyakorn, P. S., White, H. O. & Panday, S., *STAFF2D Solute Transport and Fracture Flow in Two Dimensions.* HydroGeologic, Inc., Herndon, VA, 1989.
125. Iman, R. L., Davenport, J. M., Frost, E. L. & Shortencarier, M. J., *Stepwise Regression with PRESS and Rank Regression (Program User's Guide).* SAND79-1472, Sandia National Laboratories, Albuquerque, NM, 1980.
126. Iman, R. L., Shortencarier, M. J. & Johnson, J. D., *A FORTRAN 77 Program and User's Guide for the Calculation of Partial Correlation and Standardized Regression Coefficients.* NUREG/CR-4122, SAND85-0044, Sandia National Laboratories, Albuquerque, NM, 1985.
127. Marietta, M. G., *et al., Assessment of Long-Term Safety of the Waste Isolation Pilot Plant, December 1991,* SAND93-1894, Sandia National Laboratories, Albuquerque, NM (in prep.).
128. Iman, R. L. & Conover, W. J., *A Modern Approach to Statistics.* Wiley, New York, 1983.
129. Winkler, R. L., *An Introduction to Bayesian Inference and Decision.* Holt, Reinhart and Winston, New York, 1972.
130. Levin, R. D. & Tribus, M. (Eds) *The Maximum Entropy Formalism.* The MIT Press, Cambridge, MA, 1978.
131. Tierney, M. S., *Constructing Probability Distributions of Uncertain Variables in Models of the Performance of the Waste Isolation Pilot Plant (WIPP).* SAND90-2510, Sandia National Laboratories, Albuquerque, NM, 1990.
132. Cook, I. & Unwin, S. D., Controlling Principles for Prior Probability Assignments in Nuclear Risk Assessment. *Nuclear Science and Engineering,* **94** (1986) 107–19.
133. Bonano, E. J., Hora, S. C., Keeney, R. L. & von Winterfeld, D., *Elicitation and Use of Expert Judgement in Performance Assessment for High-Level Radioactive Waste Repositories.* NUREG/CR-5411, SAND89-1821, Sandia National Laboratories, Albuquerque, NM, 1990.
134. Berger, J. O., *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York, 1985.
135. Meyer, M. A. & Booker, J. M., *Eliciting and Analyzing Expert Judgment: A Practical Guide.* Academic Press, New York, 1991.
136. Ortiz, N. R., Wheeler, T. A., Breeding, R. J., Hora, S. C., Meyer, M. A. & Keeney, R. L., Use of Expert Judgment in NUREG-1150. *Nuclear Engineering and Design,* **126** (1991) 313–31.
137. Hora, S. C. & Iman, R. L., Expert Opinion in Risk Analysis: The NUREG-1150 Methodology. *Nuclear Science and Engineering,* **102** (1989) 323–31.
138. Harper, F. T., *et al., Evaluation of Severe Accident Risks: Quantification of Major Input Parameters; Experts' Determination of In-Vessel Issues.* NUREG/CR-4551, SAND86-1309, Vol. 2, Part 1, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1990.
139. Harper, F. T., *et al., Evaluation of Severe Accident Risks: Quantification of Major Input Parameters; Experts' Determination of Containment Loads and Molten Core–Concrete Issues.* NUREG/CR-4551, SAND86-1309, Vol 2, Part 2, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1991.
140. Breeding, R. J., *et al., Evaluation of Severe Accident Risks: Quantification of Major Input Parameters; Experts' Determination of Structural Response Issues.* NUREG/CR-4551, SAND86-1309, Vol. 2, Part 3, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1992.
141. Harper, F. T., *et al., Evaluation of Severe Accident Risks: Quantification of Major Input Parameters; Experts' Determination of Source Term Issues.* NUREG/CR-4551, SAND86-1309, Vol 2, Part 4, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1992.
142. Harper, F. T., *et al., Evaluation of Severe Accident Risks: Quantification of Major Input Parameters; Supporting Material.* NUREG/CR-4551, SAND86-1309, Vol. 2, Part 5, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1993.
143. Electric Power Research Institute, *Probabilistic Seismic Hazard Evaluations at Nuclear Plant Sites in the Central and Eastern United States: Resolution of the Charleston Earthquake Issue.* EPRI-NP-6395D, EPRI, Palo Alto, CA, 1989.
144. Rechard, R. P., Sandha, J. S. & Iuzzolino, H. J., *Data Used in Preliminary Performance Assessment of the Waste Isolation Pilot Plant (1990).* SAND89-2408, Sandia National Laboratories, Albuquerque, NM, 1990.
145. Tierney, M. S., *Combining Scenarios in a Calculation of the Overall Probability Distribution of Cumulative Releases of Radioactivity from the Waste Isolation Pilot Plant, Southeastern New Mexico.* SAND90-0838, Sandia National Laboratories, Albuquerque, NM, 1991.
146. Freeze, R. A. & Cherry, J. A., *Groundwater.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1979.
147. Lappin, A. R., Hunter, R. L., Garber, D. P. & Davies, P. B. (Eds), *Systems Analysis, Long-Term Radionuclide Transport, and Dose Assessments, Waste Isolation Pilot Plant (WIPP), Southeastern New Mexico; March 1989.* SAND89-0462, Sandia National Laboratories, Albuquerque, NM, 1989.

148. Swift, P. N., Baker, B. L., Brinster, K. F., Marietta, M. G. & Roache, P. J., *Parameter and Boundary Conditions Sensitivity Studies as Related to Climate Variability and Scenario Screening for the Waste Isolation Pilot Plant.* SAND89-2029, Sandia National Laboratories, Albuquerque, NM, (in prep.).
149. McGrath, E. J., *et al. Techniques for Efficient Monte Carlo Simulation.* ORNL-RSIC-38, Vols. I–III. Oak Ridge National Laboratory, Oak Ridge, TN, 1975.
150. Steinberg, H. A., Generalized Quota Sampling. *Nuclear Science and Engineering,* **15** (1963) 142–5.
151. Raj, Des, *Sampling Theory.* McGraw-Hill, New York, 1968, pp. 206–9.
152. Khatib-Rahbar, M., Cazzoli, E., Lee, M., Nourbakhsh, H., Davis, R. & Schmidt, E., A Probabilistic Approach to Quantifying Uncertainties in the Progression of Severe Accidents. *Nuclear Science and Engineering,* **102** (1989) 219–59.
153. Anandalingam, G., Process Modeling under Uncertainty: Simulation Analysis of Indian Steel. *Journal of the Operational Research Society,* **38** (1987) 115–25.
154. Helton, J. C. & Johnson, J. D., An Uncertainty/Sensitivity Study for the Station Blackout Sequence at a Mark I Boiling Water Reactor. *Reliability Engineering and System Safety,* **26** (1989) 293–328.
155. Helton, J. C., Iman, R. L., Johnson, J. D. & Leigh, C. D., Uncertainty and Sensitivity Analysis of a Dry Containment Test Problem for the MAEROS Aerosol Model. *Nuclear Science and Engineering,* **102** (1989) 22–42.
156. Derwent, R. G., Treating Uncertainty in Models of the Atmospheric Chemistry of Nitrogen Compounds. *Atmospheric Environment,* **21** (1987) 1445–54.
157. Jaffe, P. R. & Ferrara, R. A., Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants. *Water Research,* **18** (1984) 1169–74.
158. Reed, K. L., Rose, K. A. & Whitmore, R. C., Latin Hypercube Analysis of Parameter Sensitivity in a Large Model of Outdoor Recreation Demand. *Ecological Modelling,* **24** (1984) 159–69.
159. Harris, C. M., Oil and Gas Supply Modeling Under Uncertainty: Putting DOE Midterm Forecasts in Perspective. *The Energy Journal,* **4** (1983) 53–65.
160. Conover, W. J., *Practical Nonparametric Statistics* (2nd edn). Wiley, New York, 1980.
161. Golub, G. & van Loan, C. F., *Matrix Computations.* Johns Hopkins University Press, Baltimore, MD, 1983.
162. Anderson, T. W., *An Introduction to Multivariate Statistical Analysis* (2nd edn). Wiley, New York, 1984.
163. Iman, R. L. & Davenport, J. M., Rank Correlation Plots for Use with Correlated Input Variables. *Communications in Statistics,* **B11** (1982) 335–60.
164. Rechard, R. P., *Review and Discussion of Code Linkage and Data Flow in Nuclear Waste Compliance Assessments.* SAND87-2833, Sandia National Laboratories, Albuquerque, NM, 1989.
165. Rechard, R. P., Iuzzolino, H. J., Rath, J. S., McCurley, R. D. & Rudeen, D. K., *User's Manual for CAMCON: Compliance Assessment Methodology Controller.* SAND88-1496, Sandia National Laboratories, Albuquerque, NM, 1989.
166. Campbell, J. E. & Longsine, D. E., Application of Generic Risk Assessment Software to Radioactive Waste Disposal. *Reliability Engineering and System Safety,* **30** (1990) 183–93.
167. Holmes, R. W., Uncertainty Analysis for the Long-Term Assessment of Uranium Mill Tailings. In *Uncertainty Analysis for Performance Assessments of Radioactive Waste Disposal Systems.* Organization for Economic Cooperation and Development, Paris, 1987, pp. 167–90.
168. Stein, M., Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics,* **29** (1987) 143–51.
169. Ibrekk, H. & Morgan, M. G., Graphical Communication of Uncertain Quantities to Nontechnical People. *Risk Analysis,* **7** (1987) 519–29.
170. Silverman, B. W., *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, New York, 1986.
171. Woo, G., A Quitting Rule for Monte Carlo Simulation of Extreme Risks. *Reliability Engineering and System Safety,* **31** (1991) 179–89.
172. Cheng, R. C. H. & Iles, T. C., Confidence Bounds for Cumulative Distribution Functions of Continuous Random Variables. *Technometrics,* **25** (1983) 77–86.
173. Iman, R. L. & Helton, J. C., The Repeatability of Uncertainty and Sensitivity Analyses for Complex Probabilistic Risk Assessments. *Risk Analysis,* **11** (1991) 591–606.
174. Wu, Y.-T., Millwater, H. R. & Cruse, T. A., Advanced Probabilistic Structural Method for Implicit Performance Functions. *AIAA Journal,* **28** (1990) 1663–9.
175. Wu, Y.-T., Demonstration of a New, Fast Probability Integration Method for Reliability Analysis. *Journal of Engineering for Industry,* **109** (1987) 24–28.
176. Wu, Y.-T. & Wirsching, P. H., New Algorithm for Structural Reliability. *Journal of Engineering for Industry,* **113** (1987) 1319–36.
177. Chen, X. & Lind, N. C., Fast Probability Integration by Three Parameter Tail Approximation. *Structural Safety,* **1** (1983) 169–76.
178. Rackwitz, R. & Fiessler, B., Structural Reliability Under Combined Random Load Sequences. *Journal of Computers and Structures,* **9** (1978) 489–94.
179. Beckman, R. J. & McKay, M. D., Monte Carlo Estimation under Different Distributions Using the Same Simulation. *Technometrics,* **29** (1987) 153–60.
180. Harter, H. L., Least Squares. In *Encyclopedia of Statistical Sciences,* **4,** ed. S. Kotz & N. L. Johnson. Wiley, New York, 1983, pp. 593–8.
181. Eisenhart, C., The Meaning of 'Least' in Least Squares. *Journal of the Washington Academy of Science,* **54** (1964) 24–33.
182. Myers, R. H. *Classical and Modern Regression with Applications.* Duxbury Press, Boston, MA, 1986.
183. Weisberg, S., *Applied Linear Regression* (2nd edn). Wiley, New York, 1985.
184. Seber, G. A. F., *Linear Regression Analysis.* Wiley, New York, 1977.
185. Draper, N. R. & Smith, H., *Applied Regression Analysis* (2nd edn). Wiley, New York, 1981.
186. Daniel, C. & Wood, F. S., *Fitting Equations to Data* (2nd edn). Wiley, New York, 1980.
187. Neter, J. & Wasserman, W., *Applied Linear Statistical Models.* Richard D. Irwin, Homewood, IL, 1974.
188. Iman, R. L. & Conover, W. J., The Use of the Rank Transform in Regression. *Technometrics,* **21** (1979) 499–509.
189. Allen, D. M., *The Prediction Sum of Squares as a Criterion for Selecting Predictor Variables.* Report No. 23, Department of Statistics, University of Kentucky, Lexington, KY, 1971.
190. Cook, D. R. & Weisberg, S., *Residuals and Influence in Regression.* Chapman and Hall, New York, 1982.
191. Belsley, D. A., Kuh, E. & Welsh, R. E., *Regression*

Diagnostics: Identifying Influential Data and Sources of Collinearity. Wiley, New York, 1980.
192. Saltelli, A. & Mariovet, J., Non-Parametric Statistics in Sensitivity Analysis for Model Output: A Comparison of Selected Techniques. *Reliability Engineering and System Safety*, **28** (1990) 229–53.
193. Breeding, R. J., Helton, J. C., Murfin, W. B. & Smith, L. N., *Evaluation of Severe Accident Risks: Surry Unit 1.* NUREG/CR-4551, SAND86-1309, Vol. 3, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1990.
194. Helton, J. C., Risk, Uncertainty in Risk and the EPA Release Limits for Radioactive Waste Disposal. *Nuclear Technology*, **101** (1993) 18–39.
195. Payne, A. C., Breeding, R. J., Jow, H.-N., Helton, J. C., Smith, L. N. & Shiver, A. W., *Evaluation of Severe Accident Risks: Peach Bottom Unit 2.* NUREG/CR-4551, SAND86-1309, Vol. 4, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1990.
196. Gregory, J. J., Breeding, R. J., Murfin, W. B., Helton, J. C., Higgins, S. J. & Shiver, A. W., *Evaluation of Severe Accident Risks: Sequoyah Unit 1.* NUREG/CR-4551, SAND86-1309, Vol. 5, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1990.
197. Brown, T. D., Breeding, R. J., Jow, H.-N., Helton, J. C., Higgins, S. J., Amos, C. N. & Shiver, A. W., *Evaluation of Severe Accident Risks: Grand Gulf Unit 1.* NUREG/CR-4551, SAND86-1309, Vol. 6, Rev. 1, Sandia National Laboratories, Albuquerque, NM, 1990.
198. Park, C. K., *et al.*, *Evaluation of Severe Accident Risks: Zion Unit 1.* NUREG/CR-4551, Vol. 7, Rev. 1, BNL/NUREG-5209, Brookhaven National Laboratory, Upton, NY, 1993.
199. Helton, J. C. & Breeding, R. J., Calculation of Reactor Accident Safety Goals. *Reliability Engineering and System Safety*, **39** (1993) 129–58.
200. *Preliminary Comparison with 40 CFR Part 191, Subpart B for the Waste Isolation Pilot Plant, December 1991.* SAND91-0893, Vols 1–4, Sandia National Laboratories, Albuquerque, NM, 1991.
201. Finkel, A. M., *Confronting Uncertainty in Risk Management*. Center for Risk Management, Resources for the Future, Washington, DC, 1990.
202. Rish, W. R., *Approach to Uncertainty in Risk Analysis*. ORNL/TM-10746, Oak Ridge National Laboratory, Oak Ridge, TN, 1988.
203. Draper, D., Hodges, J. S., Leamer, E. E., Morris, C. N. & Rubin, D. B., *A Research Agenda for Assessment and Propagation of Model Uncertainty*. N-2683-RC, The Rand Corporation, Santa Monica, CA, 1987.
204. Hodges, J. S., Uncertainty, Policy Analysis and Statistics. *Statistical Science*, **2** (1987) 259–75.
205. SENES Consultants Limited, *Uncertainty Analysis and Probabilistic Modeling*. INFO-0172, Atomic Energy Control Board, Ottawa, Canada, 1985.