First-Order Reliability Analysis of Public Health Risk Assessment

Maged M. Hamed

Received April 11, 1996; revised September 30, 1996

This paper demonstrates a new methodology for probabilistic public health risk assessment using the first-order reliability method. The method provides the probability that incremental lifetime cancer risk exceeds a threshold level, and the probabilistic sensitivity quantifying the relative impact of considering the uncertainty of each random variable on the exceedance probability. The approach is applied to a case study given by Thompson et al. on cancer risk caused by ingestion of benzene-contaminated soil, and the results are compared to that of the Monte Carlo method. Parametric sensitivity analyses are conducted to assess the sensitivity of the probabilistic event with respect to the distribution parameters of the basic random variables, such as the mean and standard deviation. The technique is a novel approach to probabilistic risk assessment, and can be used in situations when Monte Carlo analysis is computationally expensive, such as when the simulated risk is at the tail of the risk probability distribution.

KEY WORDS: Risk assessment; first-order reliability method; Monte Carlo simulation; uncertainty analysis; probabilistic modeling.

1. INTRODUCTION

Probabilistic methods have recently been increasingly applied to public health risk assessment practices. The probabilistic approach is very appealing as it relaxes the need for using overly conservative point estimates for the risk parameters, and it separates, in a systematic fashion, the risk assessment and risk management steps. Furthermore, the probabilistic approach furnishes the risk assessor with much more objective information on the spectrum of risk to which the public are exposed.

The Monte Carlo simulation method (MCS) has been applied to numerous public health risk assessment analyses (e.g., Refs. 1–5). In this paper, an alternative probabilistic methodology is proposed and illustrated. The proposed probabilistic tool is based on first-order reliability method (FORM). FORM is a computationally efficient method with general applicability that provides probabilities of exceeding target risk values and probabilistic sensitivity measures. This will be illustrated below. In this work, both FORM and MCS analyses are conducted and compared. FORM is shown to require less computational effort than MCS, especially when the target threshold risk level is at the tail of the risk distribution.

It should be noted that the distinction between uncertainty and interindividual variability is not systematically addressed here, since the emphasis is on illustrating the methodology for use in risk analysis and highlighting its features rather than giving a detailed treatment of uncertainty/variability issues. Nevertheless, the proposed method is equally applicable to such cases, and is the topic of a current work by the author.

2. FIRST-ORDER RELIABILITY METHOD (FORM)

First- and second-order reliability methods (FORM and SORM) are popular in the analysis of structural systems with uncertain properties under dynamic loads.

\[ \text{FORM} \]

\[ \text{SORM} \]
The methods have been applied to modeling groundwater contaminant transport and remediation\(^{(10-11)}\) and recently to probabilistic public health risk assessment.\(^{(12)}\) Following is a short review of the basic features of the reliability methods. Madsen \textit{et al.}\(^{(13)}\) present a detailed review of the theory.

A scalar limit-state function, \(g(X)\), is formulated to define the performance of the system as a function of the random vector \(X\). The event of interest occurs when \(g(x) \leq 0\), where \(x\) is a specific realization of the random variables. In public health risk assessment problems, we are interested in the probability that the incremental lifetime cancer risk \(ILCR(X)\) in an exposure scenario exceeds a predetermined threshold \(ILCR_r\). The corresponding \(g\)-function is formulated as

\[
g(X) = ILCR_r - ILCR(X) \tag{1}
\]

The probability of exceeding the threshold risk level, termed the probability of failure, is given by

\[
P_f = P[g(X) \leq 0] = P[ILCR_r \leq ILCR(X)] \tag{2}
\]

where \(f_X(x)dx\) is the joint probability density function of \(X\) and the integration is carried over the failure domain. The estimation of the above \(n\)-fold integral is a formidable task, especially when the number of random variables considered is large. FORM and SORM are analytical methods to approximate the probability integral for basic variables with strictly increasing continuous joint cumulative distribution functions (CDFs), with knowledge of at least the first two moments. FORM and SORM are not based on randomly sampling from the random vector as in the case of the Monte Carlo method, instead they are based on transforming the problem into a nonlinear programming problem, as explained below.

FORM and SORM proceed in a multistep fashion to estimate the probability of failure.\(^{(14)}\) Initially, the random variables and the limit-state function are transformed using a nonlinear one-to-one mapping to the standard normal space of uncorrelated normally distributed variates of zero mean and unit variance, such that the integration density function becomes a standard normal integral. FORM approximation is then approximated at the design point, using a linear (first-order) approximation in FORM and a quadratic (second-order) approximation in SORM. In this paper, only FORM is used for the analysis, as it provided accurate results as explained in the application section. The first-order reliability index is given by the inner vector product

\[
\beta_{\text{FORM}} = \alpha^* \cdot u^* \tag{3}
\]

where \(\alpha^*\) is the unit vector normal to the limit-state surface at the design point in the standard space and directed toward the failure region. FORM approximation of the probability of failure is given by

\[
P_f \approx P_f^{{\text{FORM}}} = \Phi(-\beta_{\text{FORM}}) \tag{4}
\]

where \(\Phi\) is the standard normal cumulative distribution.

FORM analysis also provides sensitivity measures in the form of uncertainty importance factors.\(^{(16)}\) For independent variates, such as those considered herein, the alpha sensitivity for a random variable, \(\alpha_r\), is defined as the derivative of the first-order reliability index with respect to the corresponding variate in the standard normal space, and is given by

\[
\alpha_r = \frac{\partial \beta}{\partial u_r} \bigg|_{u^*} \tag{5}
\]

Since the \(\alpha\) vector is obtained in the optimization procedure used to determine the design point, the sensitivity information is obtained at no additional computational effort. In this work, \(100 \times \alpha_r\) is reported to express the uncertainty importance factor as a percentage, and it indicates the importance of modeling the random variable \(X_r\) as a distributed variable rather than a fixed value. Madsen\(^{(17)}\) provided an estimate of the ratio between the reliability index when considering the uncertainty of the
random variable to that assuming the median value as given by the omission factor:

\[
\frac{1}{\omega} = (1 - \alpha^2)^{0.5} = \frac{\beta_{\text{distribution}}}{\beta_{\text{median}}} \tag{6}
\]

where \(\beta_{\text{distribution}}\) is the reliability index estimated considering the uncertainty of the random variable \(X\), while \(\beta_{\text{median}}\) is the reliability index estimated when the variable is replaced by its median. Importance factors can be very helpful in guiding future data collection and in reducing the problem size, by replacing those variables with small importance factors by a deterministic point estimate.

2.1. Probabilistic Analysis Program

The reliability analysis and the Monte Carlo simulations presented in this work are conducted using PROBAN\(^{1(15)}\) which is a general purpose probability analysis program. PROBAN is a software package that is designed for sophisticated probabilistic analysis. It has a flexible input module, allowing for the definition of simple models as well as sophisticated models with complicated dependencies. It provides for a variety of methods aimed for different types of probability and distribution analyses, along with the estimation of sensitivity measures. PROBAN has an extensive distribution library that contains more than 20 probability distributions. Furthermore, the user can define her or his own probability distributions if needed. Thus, the user can create any marginal probability distribution or joint density of the input random variables, by assigning the required distribution to the respective variables. The program is executed on a SUN SPARCstation2 running SunOS 4.1.4.

3. APPLICATION

To illustrate the application of FORM to public health risk assessment, a case study presented by Thompson et al.\(^{1(1)}\) is analyzed. The probability of exceeding a predefined risk threshold due to the ingestion of benzene-contaminated soil is studied. Thompson et al.\(^{1(1)}\) calculate the incremental lifetime cancer risk (ILCR) as follows:

\[
\text{ILCR} = C_s \cdot S\text{IngR} \cdot R\text{BA} \cdot D\text{pW} \cdot W\text{pY} \cdot Y\text{pL} \cdot 10^{-6} \text{ kg/mg} \cdot BW \cdot D\text{inY} \cdot Y\text{inL} \tag{7}
\]

where the definition and distribution of parameters in the above equation are listed in Table I.

First, the probability of failure is estimated for a range of target ILCR values. In other words, a number of such target ILCR are selected and the FORM analysis is conducted for each of them to estimate the probability of exceeding the threshold risk level. The result is shown in Fig. 1a. The trend of decreasing probability of failure with increasing target ILCR should be intuitive, since it is more probable to exceed a smaller target ILCR for a given scenario. The corresponding reliability index is shown in Fig. 1b. The inverse relationship between the reliability index and probability of failure is evident. Negative reliability index corresponds to failure probability greater than 0.5, as indicated by Eq. (4). By def-
in which the cumulative distribution function for a target risk level $\text{ILCR}_i$ is given by

$$\text{CDF}(i) = P[\text{ILCR}(X) \leq \text{ILCR}_i]$$  \hspace{1cm} (8)

From (2) and (8), the CDF for ILCR is the complement of the probability of failure, that is

$$\text{CDF}(i) = 1 - P_r(i)$$  \hspace{1cm} (9)

Where $P_r(i)$ is the probability of failure at the $i$th target risk level. Figure 1c illustrates the CDF of risk obtained using the FORM analysis which is in excellent agreement with the results of Thompson et al. as shown in Fig. 1f in their paper.)

For this case study and for a target risk level which lies at the 50th percentile of the ILCR distribution, uncertainty importance factors for soil ingestion rate, soil concentration, cancer potency factor, and body weight were 37.4%, 34.6%, 26.2%, and 1.8%, respectively. It is then clear that if the risk assessor could be more certain about the soil ingestion rate, soil concentration, and cancer potency factor, s/he would reduce the uncertainty of the estimated risk. It is not as important to control the uncertainty of body weight. That is, to obtain a more accurate estimate of the probability of exceedance of the threshold risk level, available financial resources should be directed to reducing the uncertainty in the soil inges-
Reliability Analysis of Risk Assessment

3.1. Parametric Sensitivity Analysis

Another powerful feature of the first-order reliability analysis with PROBAN is the ability to conduct parametric sensitivity analyses. The sensitivity of the probability of failure with respect to the distribution parameters of the basic random variables can be evaluated. A few examples are shown here for illustrative purposes.

First, the effect of the mean value of each random variable on the resulting probability of failure is analyzed. The target risk level is, arbitrarily, chosen at the 50th percentile level. Figure 2 depicts the results, which show that the probability of failure increases with increasing the mean value of soil ingestion rate, soil concentration, and cancer potency factor. An increase in the mean of body weight, on the other hand, causes the probability of failure to decrease. An examination of Eq. (7) explains this result. Following the reliability theory terminology, the body weight, being in the denominator of the equation, can be thought of as a “resistance”...
term. An increase in a resistance term causes a decrease in the probability of exceeding the threshold risk level. Conversely, parameters appearing in the numerator are considered as "load" terms. Increasing any of these terms will result in an increase in the probability of failure.

To validate the reliability analysis, results are also obtained using Monte Carlo simulation analysis. Figure 2 shows the Monte Carlo results and its confidence intervals. The reliability results are in very good agreement with that of the Monte Carlo simulation method, and well within its confidence intervals. Every point on the dashed line represents 5000 Monte Carlo simulations.

Next, the standard deviation of each random variable was perturbed by 20% in either direction. For log-normally distributed variables, the standard deviation of the nonlogtransformed variable is perturbed. This is shown in Fig. 3. The intuitive result of increasing the probabilistic "importance" of any random variable as its level of uncertainty (characterized by its standard deviation) increases is observed. That is, the greater the standard deviation of a random variable, the greater the error introduced in estimating the probability of failure when the variable is replaced by a constant value.

The effect of standard deviations of the basic random variables on the probability of failure are then stud-
Figure 4 indicates that when the uncertainty of the random variables increase, the resulting probability of failure increases. This analysis was performed for a target risk level at the 95th percentile of the ILCR. Recall that as Eq. (9) indicates, the probability of failure is the complement of the CDF. Therefore, the y-axis in Fig. 4 is around the 5th percentile of failure probability (the complement of 95th percentile ILCR).

Interestingly, when the target risk level was taken as the 5th percentile of ILCR, the relationship is reversed (Fig. 5). Results in Figs. 4 and 5 indicate that the impact of increasing the level of uncertainty of the basic random variables on the probability of failure to meet a predefined ILCR level will depend on the location of the target threshold level in the ILCR distribution. When the target risk level is taken at the 50th percentile level, a trend similar to that of Fig. 5 was observed.

The computational efficiency of the first-order reliability analysis becomes evident when the probability of failure is small. For example, for each point of Fig. 4 approximately 120 evaluations of the risk equation (Eq. (7)) were needed for FORM. The same analysis was repeated using Monte Carlo simulation analysis. Enough realizations of the random vector were generated and the risk estimated, until the coefficient of variation of the simulated probability of failure is less than or equal to...
0.05. It was found that each point in the same figure required, on average, 7500 Monte Carlo simulations. Saving in computational time of about a factor of 50 was achieved on the SUN SPARCstation 2 platform.

4. CONCLUSIONS

Probabilistic analysis of public health risk assessment is gaining more recognition as a valid alternative to the use of conservative point estimates of the relevant parameters. This paper presented an example of the first-order reliability analysis as applied to public health risk assessment. The methodology is simple to follow and apply using available computer programs (such as PROBAN by Det Norske Veritas Research.) The method was applied to account for uncertainty in the parameters when estimating cancer risk due to the ingestion of contaminated soil. The probability of exceeding a predetermined threshold risk level was obtained, along with the probabilistic sensitivity of the probability with respect to various uncertain parameters in the form of importance factors. Parametric sensitivity analyses were conducted to study the effect of the mean and standard deviation of the random variables on the probability of failure and uncertainty importance factors. The reliability analysis was shown to be a plausible method for conducting probabilistic risk assessment, and can complement other,
more widely used, simulation methods such as Monte Carlo simulation.

ACKNOWLEDGMENTS

Editorial comments and valuable suggestions by Loren Hopkins are acknowledged.

REFERENCES