

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ES/ER/TM--35

DE93 002873

An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment

F. O. Hoffman
J. S. Hammonds

Date Issued—October 1992

Prepared by
Environmental Sciences Division
Oak Ridge National Laboratory
ESD Publication 3920

Prepared for
U.S. Department of Energy
Office of Environmental Restoration and Waste Management
under budget and reporting code EW 20

MARTIN MARIETTA ENERGY SYSTEMS, INC.
managing the
Oak Ridge K-25 Site
Oak Ridge Y-12 Plant
Oak Ridge National Laboratory
under contract DE-AC05-84OR21400
Paducah Gaseous Diffusion Plant
Portsmouth Gaseous Diffusion Plant
under contract DE-AC05-76OR00001

for the
U.S. DEPARTMENT OF ENERGY

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

CONTENTS

EXECUTIVE SUMMARY	v
1. INTRODUCTION	1
1.1 BACKGROUND ON EPA METHODS FOR RISK ASSESSMENT	1
1.2 AN OVERVIEW OF QUANTITATIVE UNCERTAINTY ANALYSIS	3
2. METHODS FOR UNCERTAINTY ANALYSIS	4
2.1 LIMITING THE SCOPE THROUGH SCREENING	4
2.2 GENERAL APPROACH TO UNCERTAINTY ANALYSIS	6
2.3 GUIDANCE ON THE SELECTION OF SUBJECTIVE PROBABILITY DISTRIBUTIONS FOR UNCERTAIN MODEL PARAMETERS	7
2.4 ANALYTICAL METHODS FOR UNCERTAINTY ANALYSIS	8
2.5 NUMERICAL METHODS FOR UNCERTAINTY ANALYSIS	11
2.6 ADVANTAGES OF AN UNCERTAINTY ANALYSIS	17
2.7 BRIEF INTRODUCTION TO UNCERTAINTY ANALYSIS FOR AN ASSESSMENT ENDPOINT THAT IS A STOCHASTIC VARIABLE	19
3. SUMMARY	20
REFERENCES	22
APPENDIX	A-1

EXECUTIVE SUMMARY

To compensate for the potential for overly conservative estimates of risk using standard U.S. Environmental Protection Agency methods, an uncertainty analysis should be performed as an integral part of each risk assessment. Uncertainty analyses allow one to obtain quantitative results in the form of confidence intervals that will aid in decision making and will provide guidance for the acquisition of additional data. To perform an uncertainty analysis, one must frequently rely on subjective judgment in the absence of data to estimate the range and a probability distribution describing the extent of uncertainty about a true but unknown value for each parameter of interest. This information is formulated from professional judgment based on an extensive review of literature, analysis of the data, and interviews with experts. Various analytical and numerical techniques are available to allow statistical propagation of the uncertainty in the model parameters to a statement of uncertainty in the risk to a potentially exposed individual. Although analytical methods may be straightforward for relatively simple models, they rapidly become complicated for more involved risk assessments. Because of the tedious efforts required to mathematically derive analytical approaches to propagate uncertainty in complicated risk assessments, numerical methods such as Monte Carlo simulation should be employed. The primary objective of this report is to provide an introductory guide for performing uncertainty analysis in risk assessments being performed for Superfund sites.

1. INTRODUCTION

When hazardous substances are inadvertently or purposefully introduced into the environment, the primary concern is what the effect will be on the environment and on human health. To provide some sort of answer to this question, a risk assessment is performed to quantify the potential detriment to exposed individuals and to evaluate the effectiveness of proposed remediation measures. The U. S. Environmental Protection Agency (EPA), in its current risk assessment guidance document for Superfund sites (EPA 1989), recommends that either a qualitative or quantitative uncertainty analysis accompany risk assessments so that the confidence in the results can be expressed. A discussion of uncertainty is critical to the full characterization of risk, because it provides a better understanding of the implications and limitations of the risk assessment (EPA 1991a). Uncertainty analysis is a valuable tool for prioritizing the contaminants and exposure pathways of concern to guide the acquisition of additional data to reduce uncertainty in risk predictions. An uncertainty analysis is especially necessary when balancing the cost against the benefits of remedial action options. The primary objective of this report is to address the issue of uncertainty in quantitative risk assessments and to present methods that can be used to perform an uncertainty analysis on risk estimates.

1.1 BACKGROUND ON EPA METHODS FOR RISK ASSESSMENT

The generic equations used in EPA methods for baseline risk assessment differ depending on whether the chemical is a noncarcinogen or a carcinogen. For noncarcinogens, a quantity called the Hazard Index (*HI*) is calculated using the following formula.

$$HI = \frac{C \times I \times ED \times EF}{BM \times AT \times RfD} , \quad (1.1)$$

where

C = concentration of the chemical in the contaminated medium,

I = estimated intake rate of the contaminant,

ED = exposure duration,

EF = exposure frequency,

BM = body mass,

AT = averaging time,

RfD = reference dose for the chemical of interest.

If the *HI* is below 1, it is highly unlikely that exposure to the chemical would lead to an adverse health effect. If the *HI* is greater than unity (1.0), remediation may be warranted.

For carcinogenic substances, a lifetime cancer risk (*CR*) is calculated from the following formula.

$$CR = \frac{C \times I \times ED \times EF \times SF}{BM \times AT}, \quad (1.2)$$

where

C = concentration in the contaminated medium,

I = estimated intake of the contaminated medium,

ED = exposure duration (30 years),

EF = exposure frequency (days per year divided by 365 days),

SF = slope factor (or cancer potency factor) for the chemical of interest,

BM = body mass,

AT = averaging time (70 years).

A primary difference in evaluating risk from noncarcinogens and evaluating risk from carcinogens is that there is always some risk associated with a carcinogen. For carcinogens, it is assumed that there is no threshold below which a risk is assumed to be zero; any exposure to a carcinogen results in some excess risk.

Performing a baseline risk assessment using EPA methods yields point estimates for the risk. These point estimates obscure the inherent uncertainty in the calculations. To better demonstrate the EPA methods, a brief example is provided.

Example 1

Situation. An inadvertent release of methyl mercury has contaminated a nearby lake. After further testing, a 95% upper confidence limit of the mean concentration of 1.57×10^{-1} mg/kg was obtained for the concentration in the fish. What would be the *HI* for adverse noncancer health effects to a maximally exposed individual? This is determined by asking: What is the *HI* for this concentration, and what is its interpretation?

Solution. A value of 70 kg (reference man) will be used for body mass. Assuming 2 fish meals per week, 50 weeks a year, and 230 grams per fish meal averaged over 1 year, one obtains a value of 0.065 kg/d for the daily ingestion rate. Finally, the RfD value of 3.00×10^{-4} mg/kg/d is obtained from the Health Effects Assessment Summary Tables (EPA 1991b). Using Eq. 1.1, one obtains an *HI* of 0.49. From this value, one would conclude that it is highly unlikely that exposure to methyl mercury in this case would warrant remedial action.

From this example, one can see that the EPA baseline risk assessment methods do not incorporate explicit estimates of uncertainty. This can lead to incorrect decisions based on risks that are either grossly overestimated or are associated with high and disproportionate amounts of uncertainty. Therefore, an uncertainty analysis should be performed in any risk assessment to quantify the degree of confidence about the estimate of risk.

It should be noted that, primarily, the type of uncertainty analysis presented in this report is for a true but unknown value [referred to in IAEA (1989) as a Type B uncertainty analysis]. In most EPA Superfund risk assessments, the objective is to obtain a reasonable estimate of the risk to a maximally exposed individual. In more advanced problems, one may be interested in the uncertainty about a true but unknown distribution of values such as the distribution of risk per individual within an exposed population [referred to in IAEA (1989) as a Type A uncertainty analysis]. This problem requires that confidence intervals (CIs) be obtained for the predicted distribution. A brief discussion of the procedure for including a Type A uncertainty analysis is presented at the end of Chap. 2.

1.2 AN OVERVIEW OF QUANTITATIVE UNCERTAINTY ANALYSIS

A quantitative uncertainty analysis requires knowledge about the potential range of values likely to encompass the true but unknown value for each parameter used in the equations employed for a reasonable estimate of the risk to a maximally exposed individual. These parameters include quantification of the source term, environmental transport of the contaminant in various environmental media, the factors used for the estimate of exposure, and the factors used to convert exposure to risk. In addition, knowledge about the potential for uncertainty resulting from use of the wrong model or equation is required.

To evaluate the magnitude of uncertainty about a risk estimate, statistical error propagation procedures should be utilized. The objective is to use these methods to obtain a CI about the risk estimate. This CI (usually described by a 90% or 95% CI) provides a quantitative region in which there is high confidence of bounding the true but unknown risk. Statistical error propagation procedures also provide the most appropriate means for prioritizing and ranking the assumptions and model subcomponents that dominate the uncertainty in the risk estimate. Such prioritization is crucial in guiding the needs for additional sampling and focused research.

The estimation of likely ranges and estimates of statistical distributions for each uncertain parameter requires a high level of expertise, because professional judgment must often be employed in the absence of data. One cannot simply look up the statistical information required for quantifying the uncertain parameters of interest in a reference manual of generic parameter values. Furthermore, the uncritical adoption of published statistical distributions of data for use in the uncertainty analysis is not recommended; these data may not apply to the conditions under consideration for risk assessment. Under no circumstance should a risk assessor treat an uncertain assumption or parameter as a constant simply because data are unavailable to define a range and distribution. In the absence of data, it may be necessary to contact experts outside Martin Marietta Energy Systems, Inc./Department of Energy (DOE) organizations to obtain the essential data and/or for assistance in judgmentally deriving uncertainty estimates. Where judgment is used to derive estimates of uncertainty, the assumptions and sources of information used must be documented in the analysis.

2. METHODS FOR UNCERTAINTY ANALYSIS

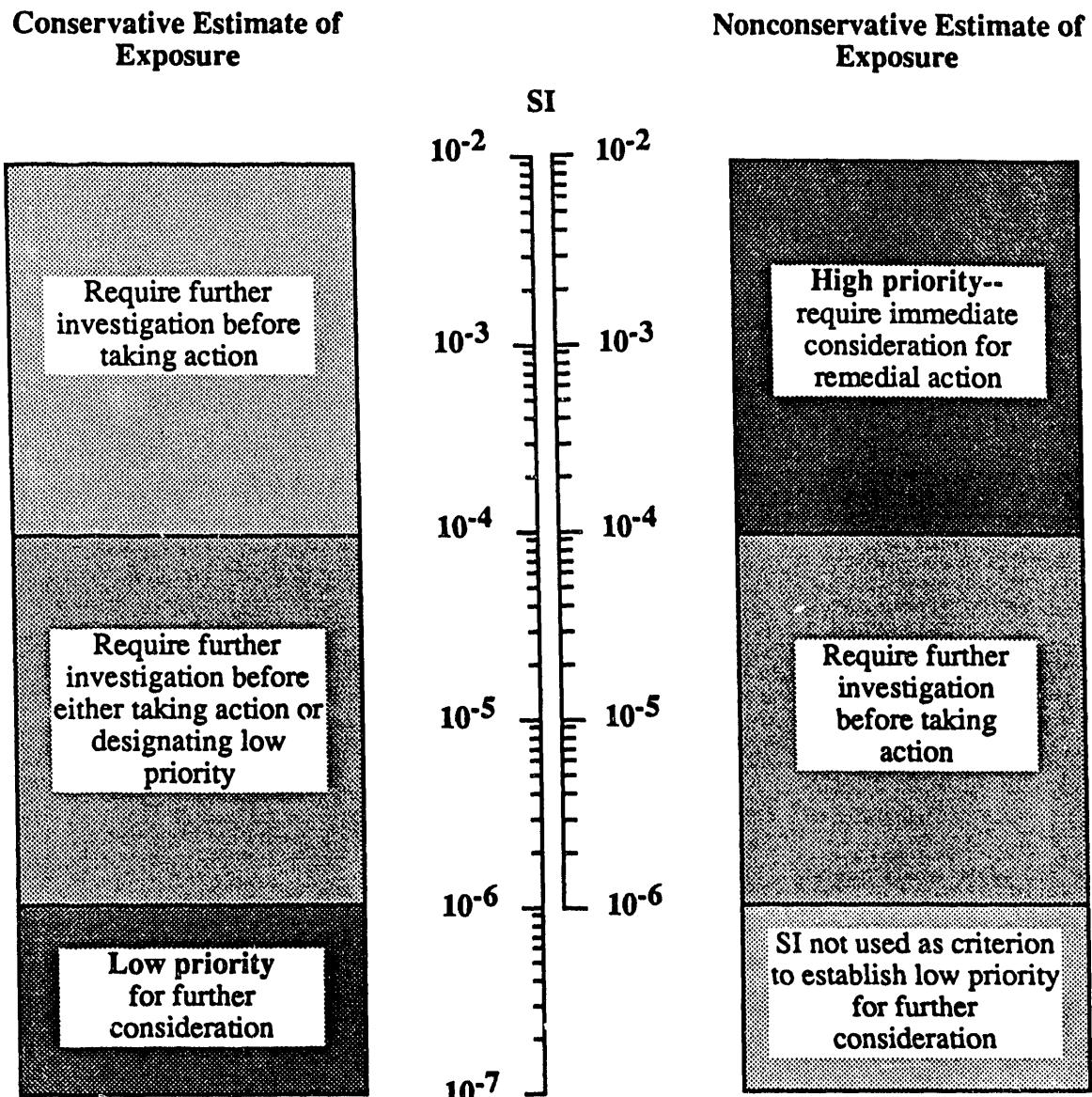
This section provides a description of methods and examples for performing statistical error propagation to quantify uncertainty in environmental and human-health-risk assessment. The methods involve analytical equations for simple models and numerical Monte Carlo approaches, which involve the use of computer technology to select values at random from a prescribed distribution, for more complex models. Other approaches to uncertainty analyses are referenced as well. The distributions describing uncertain parameters in the examples are provided for illustrative purposes only and should not be taken as a referenceable estimate of uncertainty for actual risk assessments.

2.1 LIMITING THE SCOPE THROUGH SCREENING

When taking into account the various exposure pathways for every hazardous substance found at a contaminated site, the risk assessment can become extremely lengthy and complicated. An uncertainty analysis on every parameter involved in a scenario such as this is impractical if not infeasible. Therefore, the first step in any risk assessment should be to narrow the scope of the problem.

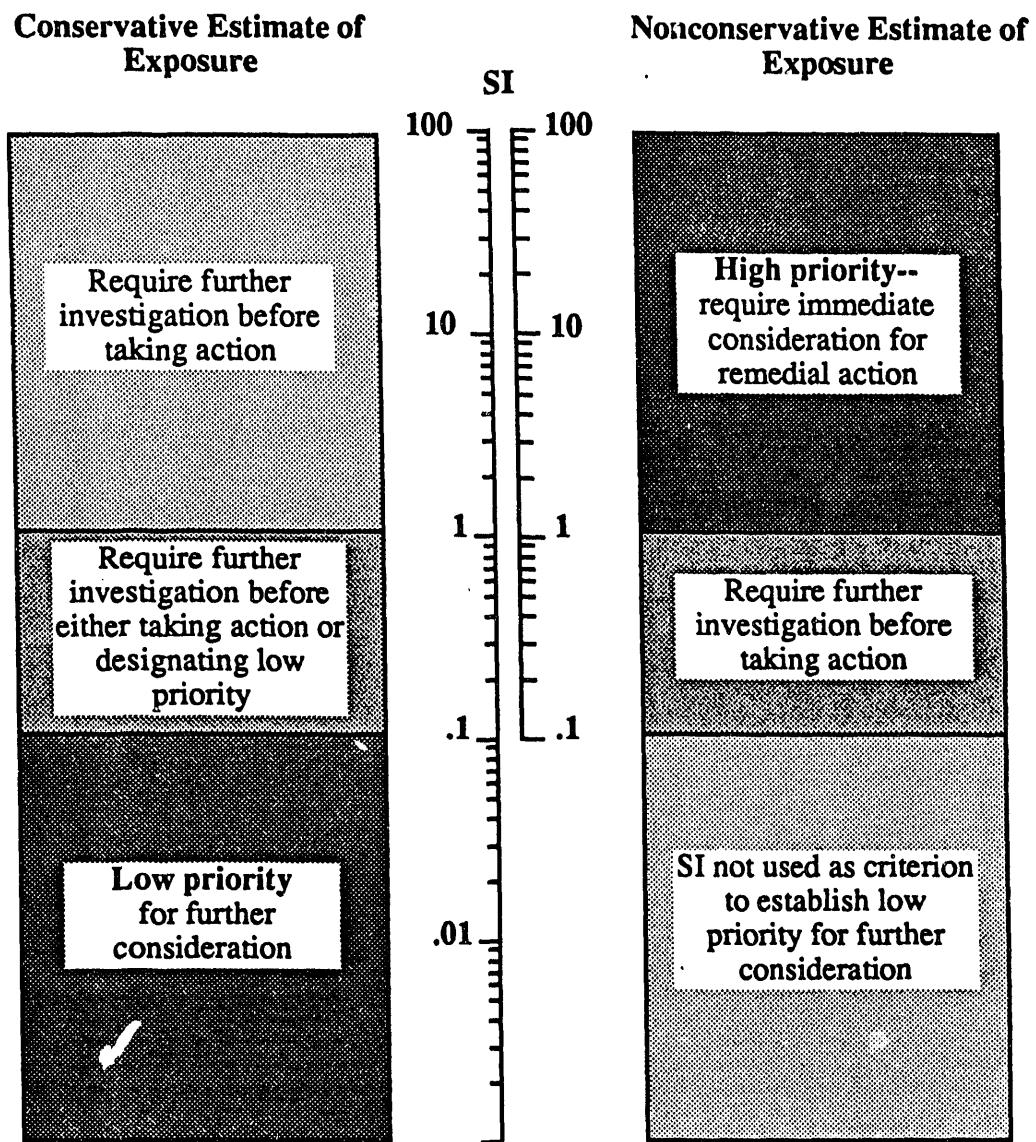
To limit the risk assessment problem, one must clearly define the objective(s) of the assessment and use a screening procedure to identify the contaminants and exposure pathways warranting a more detailed analysis (NCRP 1989, Hoffman et al. 1991, and Hoffman and Gardner 1983). Screening can be considered a first step in the approach to uncertainty analysis. Conservative assumptions used to produce a result that is not likely to underestimate the risk to a maximally exposed individual can be used to represent an upper bound estimate. Removing the conservatism in these assumptions to produce a result that is unlikely to overestimate the risk to a maximally exposed individual can be used to represent a lower bound estimate. The current EPA baseline risk assessment methods may be used as the conservative upper bound estimate of risk as long as the parameters are selected in a manner such that the actual risk will not be underestimated. This method is useful for rapidly identifying pathways and contaminants that may be given low priority for further investigation. A nonconservative (lower bound) screening calculation, on the other hand, is useful for rapidly identifying contaminants that warrant immediate consideration for remedial action. Examples of assumptions used in the general approaches for performing conservative and nonconservative screening are provided in Table 2.1.

To distinguish between low priority, potentially high priority, and high priority contaminants and pathways, one must establish baseline screening indexes. These should be chosen based on a distinction between risk levels that are clearly acceptable versus those that are clearly unacceptable. In most cases, this distinction may be influenced by risk levels associated with current regulatory standards. An example of previously chosen screening indexes for carcinogens is provided in Fig. 2.1, and a list for noncarcinogens is presented in Fig. 2.2 (Blaylock et al. 1991). Two primary sources of equations and parameter values used in screening calculations are NCRP (1989 In Press) for metals and radionuclides and Lyman et al. (1982) for organic chemicals. A brief example of the use of the described screening procedure is provided.



Screening index (SI) = exposure multiplied by a lifetime cancer slope factor.

Fig. 2.1 Criteria for conservative and nonconservative screening of carcinogens (Blaylock et al. 1991).



Screening index (SI) = exposure divided by reference dose factor (RfD).

Fig. 2.2 Criteria for conservative and nonconservative screening of noncarcinogens (Blaylock et al. 1991).

Table 2.1. An example of assumptions used for conservative and nonconservative screening (Hoffman et al. 1991)

Conservative screening	Nonconservative screening
Maximum concentration reported for a defined location	Average of detected values reported for a defined location
Models used to estimate concentrations in media that are not sampled or detected	Only measured concentrations in sediment, water, or fish are considered
Reasonable estimate of maximum diet and occupancy times assumed	Estimates of diet and occupancy times are generally a factor of 10 less than assumed for reasonable maximum
Human receptor exposed for 70 years	Probability of exposure period being less than 70 years considered in estimates of diet and occupancy times
Multiple pathway exposure considered	Multiple pathway exposure not considered
Exposure to dredged sediment considered separately from the consumption of water, fish, and irrigated agricultural produce	Dredging of sediment not considered; use of water for irrigation not considered
Calculated exposure should not underestimate actual maximum exposures	Calculated exposure should not overestimate potential maximum exposures
Screening approach most useful for identifying definitely low-priority contaminants	Screening approach most useful for identifying definitely and potentially high-priority contaminants

Example 2

Situation. Suppose that you have been contracted to perform a risk assessment for a contaminated site. When the results are received from the initial samples, you find that there are ~100 contaminants present at the site in four different environmental media, each giving rise to five to ten different exposure pathways. Because of time and money constraints, it would be questionable to perform an uncertainty analysis for each contaminant and exposure pathway represented at this site. What should you do?

Solution. Before embarking on a formal quantitative uncertainty analysis, you should narrow the scope of this problem. This is best accomplished by performing conservative and nonconservative screening calculations to aid in prioritizing pathways and contaminants warranting further investigation. Conservative screening will typically show that all but a few situations are of low priority with respect to potential health risk. These situations are contamination of sediment and fish with ^{137}Cs , As, Hg, and polychlorinated biphenyls (PCBs). Nonconservative screening may indicate that no situation warrants immediate action, but mercury and PCBs in fish are found to have potentially high priority

for further investigation because limits of concern are approached. Therefore, ^{137}Cs , As, Hg, and PCBs in sediment and fish are considered further within a formal quantitative uncertainty analysis. This analysis is carried out to first guide the acquisition of additional data and then to guide decisions for the need for remedial action.

2.2 GENERAL APPROACH TO UNCERTAINTY ANALYSIS

Once the scope of the problem has been refined to a smaller listing of contaminants and pathways, uncertainty analysis can be used to assess the extent of confidence in the estimate of health risk. In this report, it is assumed that uncertainty in the estimate of risk can be derived from an estimate of uncertainty in each of the parameters used in the risk-assessment equations. This approach is referred to as a "parameter uncertainty analysis" (IAEA 1989). Any additional uncertainty resulting from model structure should be represented either by alternative equations or by additional parameters added to the risk assessment model.

To perform parameter uncertainty analysis, one should use the following steps (IAEA 1989).

1. List all uncertain parameters (include additional parameters to represent uncertainty in model structure).
2. Specify the maximum conceivable range of potential values relevant for unknown parameters with respect to the endpoint.
3. Specify a subjective probability distribution (pdf) for values occurring within this range.
4. Account for dependencies and/or correlations among parameters.
5. Using either analytical or numerical procedures, propagate the joint probability density function of the uncertain model parameters to generate a subjective probability distribution of model predictions.
6. Derive quantitative statements of uncertainty in terms of a subjective CI interval for the true but unknown value [representing the prediction endpoint (i.e., excess cancer risk or HI)].
7. Rank the parameters contributing most to uncertainty in the model prediction.
8. Present and interpret the results of the analysis.

Steps 2 and 3 are usually obtained by using professional judgment based on an extensive review of available literature, collected data, and interviews with experts on the parameter of interest. In addition, one can incorporate correlations among the parameters by either specifying a correlation coefficient or changing the model structure to include the additional parameters that determine interdependencies among the original parameters of interest.

After the subjective probability distributions are analyzed, one obtains a subjective probability distribution for the risk, using one of the methods described in the next two sections. From this qualitative expression, one can formulate a quantitative description of the risk in the form of a subjective CI in which the true but unknown risk should lie.

The term "subjective CI" is used to denote that the probability distribution specified for the uncertain model parameters have been derived using subjective judgment in the absence of perfect data.

2.3 GUIDANCE ON THE SELECTION OF SUBJECTIVE PROBABILITY DISTRIBUTIONS FOR UNCERTAIN MODEL PARAMETERS

To perform an uncertainty analysis, one must assign subjective probability distributions for each of the uncertain parameters. The distributions reflect the degree of belief that the true but unknown value for a parameter lies within a specified range of values for the parameter. Where data are limited but uncertainty is relatively low (less than a factor of 10), a range may be used to specify a uniform distribution. If there is knowledge about a most likely value or midpoint, in addition to a range, a triangular distribution may be assigned. When the range of uncertainty exceeds a factor of 10, it is often prudent to assume either a log-uniform or log-triangular distribution. The assumption of normal, lognormal, or empirical distributions is usually dependent on the availability of relevant data. Many other distribution types are suitable for Monte Carlo analysis. A few of these other types are gamma, beta, Poisson, and Weibul and a variety of discrete distributions (Decisioneering 1991, Palisade Corporation 1991). In general, we have found that as long as the mean and variance of a distribution are held constant, the exact shape of the distribution of a parameter in a risk-assessment equation has minimal effect on the mean, variance, and general shape of the distribution of the model prediction (Gardner 1988). When dealing with several different distributions, it is more efficient to use Monte Carlo analysis than to use various analytical methods (algebraic equations). Analytical methods are options, however, when one has similar distributions for all of the parameters or when one has access to the formulas for the various distributions representing the uncertainties about the parameters.

2.4 ANALYTICAL METHODS FOR UNCERTAINTY ANALYSIS

For relatively simple equations, uncertainty analysis can be performed using analytical methods for statistical error propagation. Two types of analytical approaches frequently used for uncertainty analysis are variance propagation and moment matching (IAEA 1989). This section will discuss in more detail the technique of variance propagation for simple equations.

To best demonstrate an example of variance propagation, a simple additive model is desired. In the case of an additive model, the mean of the result is equal to the sum of the means of the parameters, and the variance of the result is equal to the sum of the variances of the parameters (IAEA 1989, Hoffman and Gardner 1983).

$$\mu_R = \sum_{i=1}^p \mu_i, \quad (2.1)$$

and

$$\sigma_R^2 = \sum_{i=1}^p \sigma_i^2, \quad (2.2)$$

where p is the number of parameters in the model.

In a series of summations of uncertain parameters, the result will tend to conform to a normal distribution even if the shapes of the distributions assigned to the model parameters are other than normal.

Note, however, that the basic form of EPA risk assessment models is a multiplicative chain of parameters for each contaminant and exposure pathway. Multiplicative models can be reduced to additive form by logarithmically transforming the variables. This is shown in equation form below.

$$Y = a \times b \times c. \quad (2.3)$$

$$\ln(Y) = \ln(a) + \ln(b) + \ln(c). \quad (2.4)$$

Therefore, the distribution of Y will tend to be approximately lognormal even when the parameters a , b , and c are assigned distribution shapes other than lognormal (Hoffman and Gardner 1983). For multiplicative calculations, one can find the median value (or geometric mean) for the risk simply by summing the means of the logarithms for the various parameters and exponentiating the sum. This relationship is presented in equation form by

$$X_{gR} = e^{\mu_R}, \quad (2.5)$$

where X_{gR} is the geometric mean of the result.

In addition, the standard deviation of the risk is found by taking the square root of the sum of the parameter variances obtained from the above equations and exponentiating (IAEA 1989, Hoffman and Gardner 1983). This formula is represented in the equation

$$S_{gR} = e^{\sqrt{\sigma_R^2}}, \quad (2.6)$$

where S_{gR} = the geometric standard deviation of the result.

The upper confidence limit is determined by multiplying the median value by the square (or some other power) of S_{gR} , the exponentiated standard deviation of logarithms. The lower confidence limit is obtained by dividing the median by the square (or some other power) of the exponentiated standard deviation of the logarithms. The exponential standard deviation of logarithms is often referred to as the geometric standard deviation (GSD). The use of the square of the GSD will lead to a 95% CI assuming that the distribution of the model prediction will be lognormal. Taking the GSD to a power of 1.65 will lead to a 90% CI for a lognormal distribution. The formula used to estimate mean and variance of logarithms for each uncertain parameter depends on the type of subjective probability distribution chosen to represent the uncertain parameter. Formulas for the mean and variance of logarithms of lognormal, log-uniform, and log-triangular distributions are provided in the Appendix.

Example 3

Situation. Let us assume, as in Example 1, that there has been an accidental spill of methyl mercury in a nearby lake. Using the technique of variance propagation, obtain a 90% CI on the HI to a maximally exposed individual. After reviewing the literature and available data and consulting with other experts, the subjective probability distributions shown in Table 2.2 are obtained for this problem.

Table 2.2. Information for Example 3

Parameter	Distribution	Minimum	Maximum	Mean	Standard deviation
Fish concentration (FC), mg/kg	Lognormal			7.10E-2	3.43E-2
Intake (I), kg/d	Log-uniform	2.00E-2	1.30E-1	6.50E-2	
Body mass (BM), kg	Log-triangle	4.50E+1	1.20E+2	7.00E+1	
RfD, mg/kg/d	Log-triangle	1.50E-4	3.00E-3	3.00E-4	

Note: The mean given for the intake, body mass, and RfD is the most likely value (mode). The distributions represent subjective confidence about the uncertainty associated with estimating a true but unknown value for each parameter.

Solution. The form of the equation used for this problem is:

$$HI = FC \times I \times (BM)^{-1} \times (RfD)^{-1} . \quad (2.7)$$

By log-transforming, this equation becomes:

$$\ln(HI) = \ln(FC) + \ln(I) - \ln(BM) - \ln(RfD) . \quad (2.8)$$

Therefore, the logarithmic mean and variance of the HI is found by using Eqs. 2.1 and 2.2. The mean equations given in the Appendix must be used to find the mean and variance of logarithms for each of the model parameters. Using Eqs. A.1 and A.2, one obtains a

mean of the logarithms for the fish concentration of -2.75 and finds that the variance of the logarithms for the fish concentration is 0.21.

Equations A.3 and A.4, given for the log-uniform distribution, will be used to find μ (the mean of logarithms) and σ (the variance of logarithms) for the intake. These values are calculated below.

$$\mu_I = \frac{\ln(0.02) + \ln(0.13)}{2} = -2.98 . \quad (2.9)$$

$$\sigma_I^2 = \frac{\left[\ln\left(\frac{0.13}{0.02}\right) \right]^2}{12} = 0.29 . \quad (2.10)$$

Equations A.5 and A.6, given for the log-triangular distribution, are the ones that are used for the body mass and RfD. To demonstrate, the μ and σ for the body mass is calculated below.

$$\mu_{BM} = \frac{1}{3}[\ln(70) + \ln(120) + \ln(45)] = 4.28 . \quad (2.11)$$

$$\sigma_{BM}^2 = \frac{1}{18} \{ [\ln(45)]^2 + [\ln(120)]^2 - [\ln(45)][\ln(120)] + [\ln(70)]^2 - [\ln(70)][\ln(45) + \ln(120)] \} = 0.04 . \quad (2.12)$$

This same process is performed for the RfD, from which one obtains a mean value of -7.58 and a variance of 0.41. One is now able to find the mean of logarithms and, thus, the geometric mean of the *HI*.

$$\mu_{HI} = -2.75 - 2.98 - 4.28 + 7.58 = -2.43 . \quad (2.13)$$

$$X_{g,HI} = e^{\mu_{HI}} = e^{2.43} = 0.09 . \quad (2.14)$$

The variance of the *HI* and, consequentially, the geometric standard deviation of the *HI* is found.

$$\sigma_{HI}^2 = 0.21 + 0.29 + 0.04 + 0.41 = 0.95 . \quad (2.15)$$

$$\sigma_{g,HI} = e^{\sqrt{\sigma_{HI}^2}} = e^{\sqrt{0.95}} = 2.65 . \quad (2.16)$$

The upper and lower confidence limits for a 90% subjective CI are calculated.

$$X_{95}^{HI} = X_g^{HI} \times S_g^{HI^{1.65}} = (0.09)(2.65)^{1.65} = 0.45 . \quad (2.17)$$

$$X_5^{HI} = \frac{X_g^{HI}}{S_g^{HI^{1.65}}} = \frac{(0.09)}{(2.65)^{1.65}} = 0.02 . \quad (2.18)$$

Therefore, there is high confidence (at a subjective level of 90%) that the *HI* should lie between 0.02 and 0.45.

As one can see, variance propagation is a straightforward process for simple additive and multiplicative models. For more complex calculations, variance propagation techniques are more difficult to derive analytically, and in some cases their derivation may not be practical. For current Superfund risk assessment, both variance propagation methods and numerical methods for propagating the uncertainty may be appropriate.

2.5 NUMERICAL METHODS FOR UNCERTAINTY ANALYSIS

To overcome problems encountered with variance propagation equations, various numerical methods are useful in performing an uncertainty analysis with the aid of a computer. Two such numerical techniques are Monte Carlo simulation (Hoffman and Gardner 1983, IAEA 1989) and deterministic uncertainty analysis (Worley 1987). Monte Carlo analysis usually employs two random selection processes: Simple Random Sampling (SRS) and Latin Hypercube Sampling (LHS) (IAEA 1989).

The random selection process known as SRS is conceptually straightforward. For each iteration a random number is chosen from within each distribution specified for an uncertain parameter. Even though this technique does work relatively well, it is less efficient than its counterpart, LHS. With LHS, fewer samples are required to obtain the same result achieved with SRS (IAEA 1989). In LHS, the distribution for each parameter is divided into sections. The number of sections depends on how many samples (or iterations) the assessor wants the simulation to take. Another feature of these divisions is that they each have the same area. During the simulation run, the numbers are selected at random within a section with only one random number being chosen from a specific section. In other words, once a random number has been selected from a section, that

section is disregarded from the rest of the analysis. In this manner the distributions are represented more efficiently; therefore, it takes less time to reach a stable mean and variance of the prediction endpoint.

Monte Carlo analysis may be performed in many ways. One may write one's own numerical code or use one of several currently available software packages. Several available Monte Carlo simulation programs are listed below.

MOUSE	Klee (1986)
TAM3	Kanyar and Nielsen (1989) and Gardner (1988)
PRISM	Gardner and Trabalka (1985) and Gardner et al. (1983)
Crystal Ball	Decisioneering, Inc. (1991)
@RISK	Palisade Corporation (1991)
ORMONTE	Williams and Hudson (1989)

The following example provides a more detailed description of a Monte Carlo simulation.

Example 4

Situation. Use the scenario presented in Example 3 to demonstrate the use of Monte Carlo simulation. With 90% subjective confidence, what is the risk to the maximally exposed individual? Please note that this example does not account for dependencies among parameters; this will be demonstrated in Example 6.

Solution. To begin an uncertainty analysis, one must describe the uncertainty about each variable with a subjective probability distribution. This is done through judgment after extensive review of all relevant data. The information presented in Table 2.2 is used as input for a Monte Carlo simulation for this problem.

When running a Monte Carlo technique, values are selected at random from each uncertain variable to produce a prediction. This procedure is repeated for a specified number of iterations and forms a distribution of predicted values. A sample of randomly selected values obtained by running 500 iterations of LHS for this problem is provided in Table 2.3.

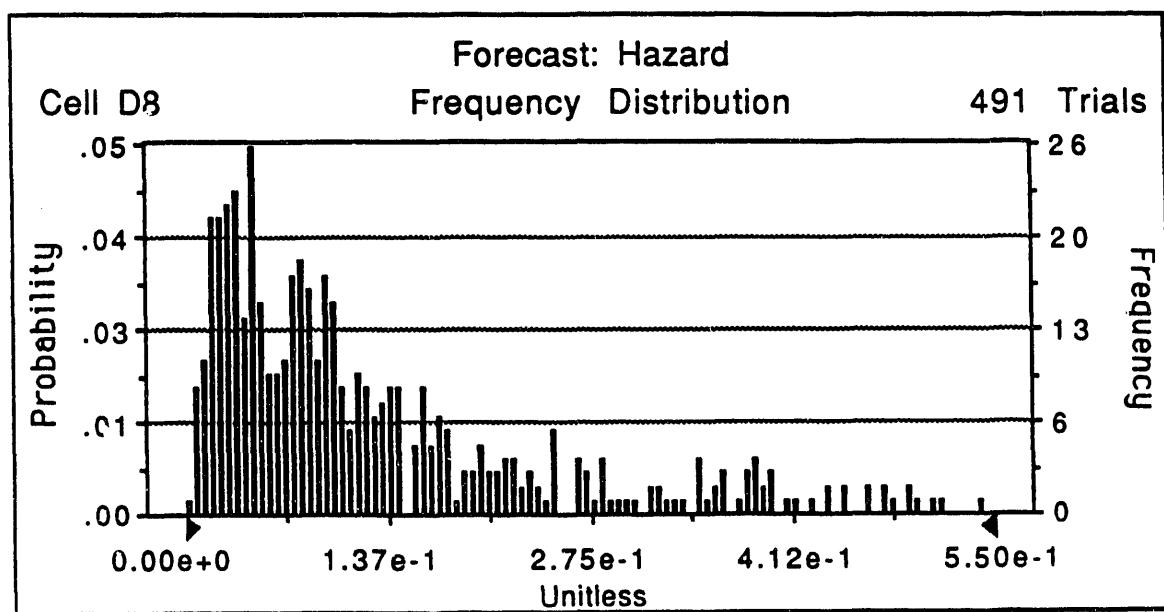
This process yields a subjective probability distribution for the *HI* from which a quantitative representation of the *HI* can be formed. Figure 2.3 contains the result for the risk after 500 iterations using LHS.

From this Monte Carlo simulation, a 90% CI of [1.70E-2, 4.17E-1] is obtained. This implies that after taking into account the uncertainties on the parameters, one is highly confident (at a subjective level of 90%) that the true *HI* should lie between 1.70E-2 and 4.17E-1.

Once familiarization with the Monte Carlo simulation software package has taken place, this technique becomes extremely quick to use. Even if a risk analysis becomes more complicated, the Monte Carlo technique does not. One reason that Monte Carlo

Table 2.3. A sample of random values obtained from 500 iterations of LHS for Example 4.

Sample number	Fish concentration (mg/kg)	Intake (kg/d)	Body mass (kg)	Reference dose (mg/kg/d)	Hazard index (unitless)
1	1.01E-01	3.40E-02	4.71E+01	5.25E-04	1.38E-01
2	1.14E-01	1.19E-01	7.68E+01	5.64E-04	3.15E-01
3	8.11E-02	1.05E-01	6.78E+01	1.76E-04	7.10E-01
4	6.51E-02	3.63E-02	7.50E+01	3.00E-04	1.05E-01
.
.
.
499	9.40E-02	9.21E-02	7.04E+01	2.71E-03	4.53E-02
500	8.60E-02	2.66E-02	8.15E+01	8.96E-04	3.13E-02

**Fig. 2.3 Subjective probability distribution of the hazard index for Example 4.**

calculations are more useful than other approaches to uncertainty analysis is that the variance propagation techniques can become complicated and time consuming for more involved risk analyses. Setting up simulations to run on the computer is much more efficient and accurate than performing hand calculations. The inputs required for Monte Carlo simulations are the subjective probability distributions and uncertainty bounds for each parameter. To come up with these subjective probability distributions and uncertainty bounds, one must use professional judgment after extensively reviewing the available literature and data. With the various input distributions, the Monte Carlo simulation program then provides a forecast of the risk in terms of a subjective probability distribution about which CIs for the risk can be obtained. A demonstration of this technique for a more complicated risk analysis situation is presented in Example 5.

Example 5

Situation. Let us assume that as the result of waste management practices, a mixture of contaminants is released inadvertently to the environment. Eventually, through various pathways, this contamination is transported to aquatic systems such as rivers and lakes where various fish and biota are exposed. After further investigation, it is discovered that the contaminants released were Aroclor-1254, Aroclor-1260, chlordane, and methyl mercury. Suppose that a fisherman catches some contaminated fish and eventually eats them. The assessment problem is as follows: What is the risk to the maximally exposed individual?

To perform this risk assessment, Eqs. 1.1 and 1.2 will be used. To quantify the uncertainty associated with each of the parameters introduced in these equations, one must derive (with the use of a considerable amount of judgment) subjective probability distributions from very limited sets of data and other relevant facts in the published literature. Once these distributions have been specified, one can utilize Monte Carlo techniques to obtain a probability distribution of the *HI*. From this propagated distribution, a subjective CI (90%) can be obtained for use in setting limits that are useful for decision making.

Table 2.4 contains values for the estimates of uncertainty on each of the parameters that would be used in an environmental risk assessment of Aroclor-1254, Aroclor-1260, chlordane, and methyl mercury in the fish potentially harvested from a contaminated fresh-water system.

Solution. The values given in Table 2.4 were used to find the median, the lower 5% subjective confidence limit, and the upper 95% subjective confidence limit for the noncarcinogen *HI* for chlordane and methyl mercury and for the cancer risk involved with the given concentrations of Aroclor-1254, Aroclor-1260, and chlordane in fish. These values (presented in Table 2.5) were obtained by using 500 iterations of the LHS Monte Carlo technique.

Table 2.4. Subjective probability distributions specified for the Monte Carlo Analysis of Example 5.

Chemical	Parameter	Subjective probability distribution	Minimum	Maximum	Mean (mode)	Standard deviation
Aroclor-1254	Fish concentration, mg/kg	Lognormal	4.00E-03	3.79E+00	5.34E-01	2.26E+00
	Intake, kg/d	Log-uniform	1.65E-02	8.25E-02		
	Body mass, kg	Log-triangle	4.50E+01	1.20E+02	7.00E+01	
	Slope factor, (mg/kg/d) ⁻¹	Triangle	0.00E+00	1.00E+01	7.70E+00	
	Fish concentration, mg/kg	Lognormal	3.19E-01	2.29E+00	9.75E-01	5.16E-1
	Intake, kg/d	Log-uniform	1.65E-02	8.25E-02		
Aroclor-1260	Body mass, kg	Log-triangle	4.50E+01	1.20E+02	7.00E+01	
	Slope factor, (mg/kg/d) ⁻¹	Triangle	0.00E+00	1.00E+01	7.70E+00	
	Fish concentration, mg/kg	Lognormal	3.96E-02	3.06E-01	1.27E-01	6.98E-02
	Intake, kg/d	Log-uniform	1.65E-02	8.25E-02		
	Body mass, kg	Log-triangle	4.50E+01	1.20E+02	7.00E+01	
	Slope factor, (mg/kg/d) ⁻¹	Triangle	0.00E+00	5.00E+00	1.30E+00	
Chlordane (carcinogen)	Fish concentration, mg/kg	Lognormal	3.96E-02	3.06E-01	1.27E-01	6.98E-02
	Intake, kg/d	Log-uniform	1.65E-02	8.25E-02		
	Body mass, kg	Log-triangle	4.50E+01	1.20E+02	7.00E+01	
	Slope factor, (mg/kg/d) ⁻¹	Triangle	0.00E+00	5.00E+00	1.30E+00	
	Fish concentration, mg/kg	Lognormal	3.96E-02	3.06E-01	1.27E-01	6.98E-02
	Intake, kg/d	Log-uniform	2.00E-02	1.30E-01		
Methyl mercury	Body mass, kg	Log-triangle	4.50E+01	1.20E+02	7.00E+01	
	Reference dose, mg/kg/d	Log-triangle	3.00E-05	1.90E-03	6.00E-05	
	Fish concentration, mg/kg	Lognormal	2.55E-02	1.57E-01	7.10E-02	3.43E-02
	Intake, kg/d	Log-uniform	2.00E-02	1.30E-01		
	Body mass, kg	Log-triangle	4.50E+01	1.20E+02	7.00E+01	
	Reference dose, mg/kg/d	Log-triangle	1.50E-04	3.00E-03	3.00E-04	

Table 2.5. Results obtained from Monte Carlo simulation using values from Table 2.4

Chemical	Type of result	Median	5% subjective confidence	95% subjective confidence
Aroclor-1254	Cancer risk	3.4E-04	6.0E-05	7.5E-03
Aroclor-1260	Cancer risk	2.5E-03	5.7E-04	8.5E-03
Chlordane	Cancer risk	1.0E-04	2.0E-05	4.4E-04
Total cancer risk		3.4E-03	8.4E-04	1.4E-02
Chlordane	Noncarcinogen <i>HI</i>	5.3E-01	6.7E-02	3.2E+00
Methyl mercury	Noncarcinogen <i>HI</i>	8.6E-02	1.7E-02	3.9E-01
Total <i>HI</i>		6.3E-01	1.3E-01	3.8E+00

As one can see from Table 2.5, the primary chemical contributing to the total cancer risk is Aroclor-1260, and the chemical contributing the majority of the total *HI* is chlordane. By performing a sensitivity analysis, one can determine which parameter has the most effect on the total result. This is done by simply holding a potentially sensitive parameter constant while varying the others. After this process is repeated for all of the parameters of interest, the different results obtained for each parameter are compared and ranked according to the parameters creating the biggest difference. These parameters are said to be the most sensitive parameters. Descriptions of alternative statistical approaches to sensitivity analysis using regressions of the randomly selected values of the uncertain parameters on the values produced for the model predictions can be found in IAEA Safety Series No. 100 (IAEA 1989).

For the total cancer risk, performing a sensitivity analysis shows that the amount of fish ingestion has the most effect, followed by the concentration of Aroclor-1254 in the fish. One might not expect the latter result, but the uncertainty involved with the Aroclor-1254 fish concentration is much greater than with the fish concentration for Aroclor-1260. For the total *HI*, a sensitivity analysis would show that the two parameters that are the most significant contributors to the total uncertainty are the amount of fish ingestion and the RfD for chlordane.

Example 6

Situation. The purpose of this example is to study the effect of the correlation between body mass and intake on the total cancer risk and the total *HI* for the situation given in Example 5. First, assume that a minimum correlation of 0.3 has been determined to exist between body mass and intake, and second, compare the results with those obtained with a correlation of 0.5 and 0.7.

Solution. In this case, rank correlations are used (Decisioneering, Inc. 1991) to account for interdependencies between body mass and intake. As can be seen from Table 2.6, where the results are produced from 500 iterations using LHS, the correlation

coefficients do not have a dramatic effect on the total risk. The values for the total cancer risk are virtually the same.

A slight difference is detected in the 95% upper confidence limits for the total *HI*. The values obtained with a correlation coefficient applied to the analysis are slightly lower than the value calculated without the correlation applied. One reason that the correlation does not have an obvious effect on the results is that the body mass is not a sensitive parameter.

Table 2.6. Results obtained for correlations between body mass and intake for the situation described in Example 6

	Rank correlation coefficient		
	0.3	0.5	0.7
Total cancer risk			
5%	8.9E-4	9.1E-4	9.6E-4
50%	3.4E-3	3.4E-3	3.4E-3
95%	1.3E-2	1.3E-2	1.3E-2
Total hazard index			
5%	1.2E-1	1.3E-1	1.2E-1
50%	6.8E-1	7.0E-1	7.0E-1
95%	3.4E+0	3.3E+0	3.3E+0

2.6 ADVANTAGES OF AN UNCERTAINTY ANALYSIS

One of the steps in a risk assessment is to rank the importance of the pathways and chemicals in terms of their potential contribution to the total risk. The first attempt at this is performed in the screening process. By screening, one obtains those pathways and chemicals that could be of potential concern. However, if the risk assessor attempts to rank the pathways and chemicals at this stage, the wrong conclusions may be reached, because the uncertainty involved is not necessarily equal among contaminants and exposure pathways. This is best demonstrated in the following example.

Example 7

Situation. Upon investigation of a potentially contaminated site, it was discovered that a nearby lake and the surrounding sediment were contaminated with methyl mercury and inorganic mercury, respectively. The 95% upper confidence limit value for the inorganic mercury in soil is found to be 700 mg/kg, and the 95% upper confidence limit for the concentration of methyl mercury in fish is 3.05×10^{-1} mg/kg. Which is the most hazardous pathway to the maximally exposed individual?

Solution. A summary of the values obtained for this example is provided in Table 2.7.

Table 2.7. Information for Example 7

Parameter	Distribution	Minimum	Maximum	Mean	Standard deviation
Fish concentration (<i>FC</i>), mg/kg	Lognormal			2.06E-1	4.22E-2
Intake of fish (<i>I_F</i>), kg/d	Log-uniform	2.00E-2	1.30E-1	6.50E-2	
Soil concentration (<i>SC</i>), mg/kg	Lognormal			3.11E+2	1.50E+2
Intake of soil (<i>I_S</i>), kg/d	Log-uniform	5.00E-5	2.00E-4	1.00E-4	
Exposure frequency (<i>EF</i>)	Log-uniform	2.70E-1	7.00E-1	7.00E-1	
Body mass (<i>BM</i>), kg	Log-triangle	4.50E+1	1.20E+2	7.00E+1	
Inorganic mercury RfD (RfD _{IM}), mg/kg/d	Log-uniform	3.00E-4	3.00E-2	3.00E-4	
Methyl mercury RfD (RfD _{MM}), mg/kg/d	Log-triangle	1.50E-4	3.00E-3	3.00E-4	

Note: The mean given for the intake, body mass, and RfD is the most likely value (mode).

The values for the *HIs* for the two pathways will be compared with each other for two situations: (1) by using EPA's generic equations and (2) by incorporating uncertainty analysis. The *HIs* for the fish and soil pathways follow.

$$HI_S = \frac{SC \times I_S \times EF}{BM \times RfD_{IM}} = \frac{(700)(1.00 \times 10^{-4})(0.7)}{(70)(3.00 \times 10^{-4})} = 2.33 . \quad (2.19)$$

$$HI_F = \frac{FC \times I_F}{BM \times RfD_{MM}} = \frac{(3.05 \times 10^{-1})(6.50 \times 10^{-2})}{(70)(3.00 \times 10^{-4})} = 0.94 . \quad (2.20)$$

From these calculations, one would conclude that the risk to the maximally exposed individual results from the soil-ingestion pathway. However, by incorporating the uncertainties for the parameters and using Monte Carlo simulation, one obtains different results. After a Monte Carlo simulation run of 500 iterations of LHS, the 95% upper confidence limit of the *HI* for the soil ingestion pathway is 0.72, and the 95% upper

confidence limit of the *HI* for the fish ingestion pathway is 1.20. This implies that the fish ingestion pathway is the source of most of the risk to the maximally exposed individual. The reversal of the ranking from that of the EPA calculations is primarily because of the large uncertainty in the RfD for inorganic mercury. If the uncertainty of this parameter had not been taken into account, an inaccurate conclusion and, possibly, an inappropriate course of action would have resulted.

The incorporation of uncertainty analysis portrays more confidence in the outcome of a risk assessment. An uncertainty analysis through a quantitative description provides better direction for further investigation. If uncertainty analysis is used in every risk assessment, money that would be more wisely spent on specific areas that need further study may be saved by preventing unwarranted remedial action.

2.7 BRIEF INTRODUCTION TO UNCERTAINTY ANALYSIS FOR AN ASSESSMENT ENDPOINT THAT IS A STOCHASTIC VARIABLE

All of the examples and the general subject of this report to this point have coincided with uncertainty about a true but unknown value [referred to in IAEA Safety Series No. 100 (1989) as "Type B" uncertainty]. However, some risk assessments may have an endpoint defined as a stochastic variable. An uncertainty analysis dealing only with stochastic variability is referred to as "Type A" uncertainty in IAEA Safety Series No. 100 (1989). An example that must include both types of uncertainty would be the estimation of the distribution of individual doses or risks within an exposed population group. The goal of this section, therefore, is to briefly describe the process of uncertainty analysis when the assessment endpoint is a stochastic variable and when there is lack of knowledge about the true distribution that describes this variable.

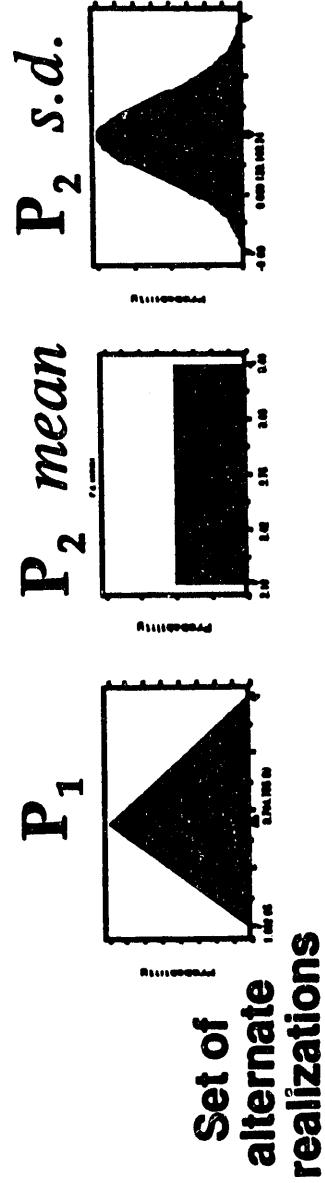
To perform this type of uncertainty analysis, one must first generate numerous alternative values for each of the uncertain deterministic quantities in the risk assessment equation. Deterministic quantities are those for which the true value is fixed with respect to the assessment endpoint. The true mean, variance, and shape of a stochastic variable are fixed with respect to the assessment endpoint and therefore are also considered as deterministic quantities. Alternative values for each deterministic quantity are achieved through Monte Carlo simulation as is illustrated in Fig. 2.4. In this figure, P_1 represents a true but unknown deterministic value, and P_2 reflects a stochastic variable for which there is lack of knowledge about the true but unknown mean and variance. An example would be for an assessment of the distribution of individual risks in a defined population of exposed persons in which P_1 is the true but unknown amount of released material and P_2 is the parameter that translates the amount released into the estimate of individual risk. After each alternative set of deterministic values have been simulated, the next step is to generate a realization of the stochastic endpoint for all sets of values. This will produce a probability distribution of the stochastic endpoint for each set of deterministic values.

In Fig. 2.5, one can see that "n" sets of values were obtained for P_1 and P_2 . The various values for the mean and standard deviation are used to simulate the corresponding normal distribution for P_2 . Next, the value of P_1 and the normal distribution of P_2 are combined to provide a simulation of the stochastic endpoint. This is repeated for each set of alternative values of P_1 and alternative distributions of P_2 .

Finally, to obtain confidence limits for the result, all of the alternative realizations of the stochastic endpoint must be plotted. To be computationally efficient, one may use Simple Random Sampling (SRS) with 59 samples for the deterministic uncertain quantities; it should be noted that the sampling procedure for the analysis of the stochastic variability may be different from the one used for the deterministic uncertain quantities. The 59 samples for the deterministic uncertain quantities will be sufficient to provide a 90% confidence interval or a 95% upper confidence limit at which the true value will not be underestimated (IAEA 1989). Figure 2.6 represents a cumulative probability plot of various alternative realizations of a stochastic endpoint which were obtained from 59 simple random samples. For the example presented in Fig. 2.6, one could say that with 90% confidence, the dose for the 95th percentile of the population is between 8 and 37, or one could say that the dose to the 95th percentile of the population will be less than or equal to 37 with 95% confidence.

When performing an uncertainty analysis where there is both stochastic variability and lack-of-knowledge uncertainty, correct interpretation of the results requires that these two sources of uncertainty be analyzed separately. One obtains various distributions representing the endpoint that is analogous to the various values obtained for the result in an uncertainty analysis where only true but unknown values are considered. The combination of these types of uncertainty analyses is facilitated using Monte Carlo simulation.

$$\text{Result} = P_1 \times P_2$$



$$1 \quad 4.73$$

$$2.34 \quad 0.08$$

$$2 \quad 3.17$$

$$2.47 \quad 0.16$$

$$n \quad 2.51$$

$$2.93 \quad 0.11$$

Fig. 2.4. A representation of the lack of knowledge: uncertainty for a parameter, P_1 , that has a single fixed but unknown value with respect to the assessment endpoint and a stochastic variable, P_2 , for which there is lack of knowledge about the true but unknown mean and standard deviation. For these deterministic quantities (i.e., P_1 and P_2 , $P_{2,\text{mean}}$, $P_{2,\text{s.d.}}$) alternative realizations of the true but unknown values are produced by randomly sampling from distributions that reflect lack of knowledge uncertainty for each of these quantities.

$$\text{Result} = P_1 \times P_2$$

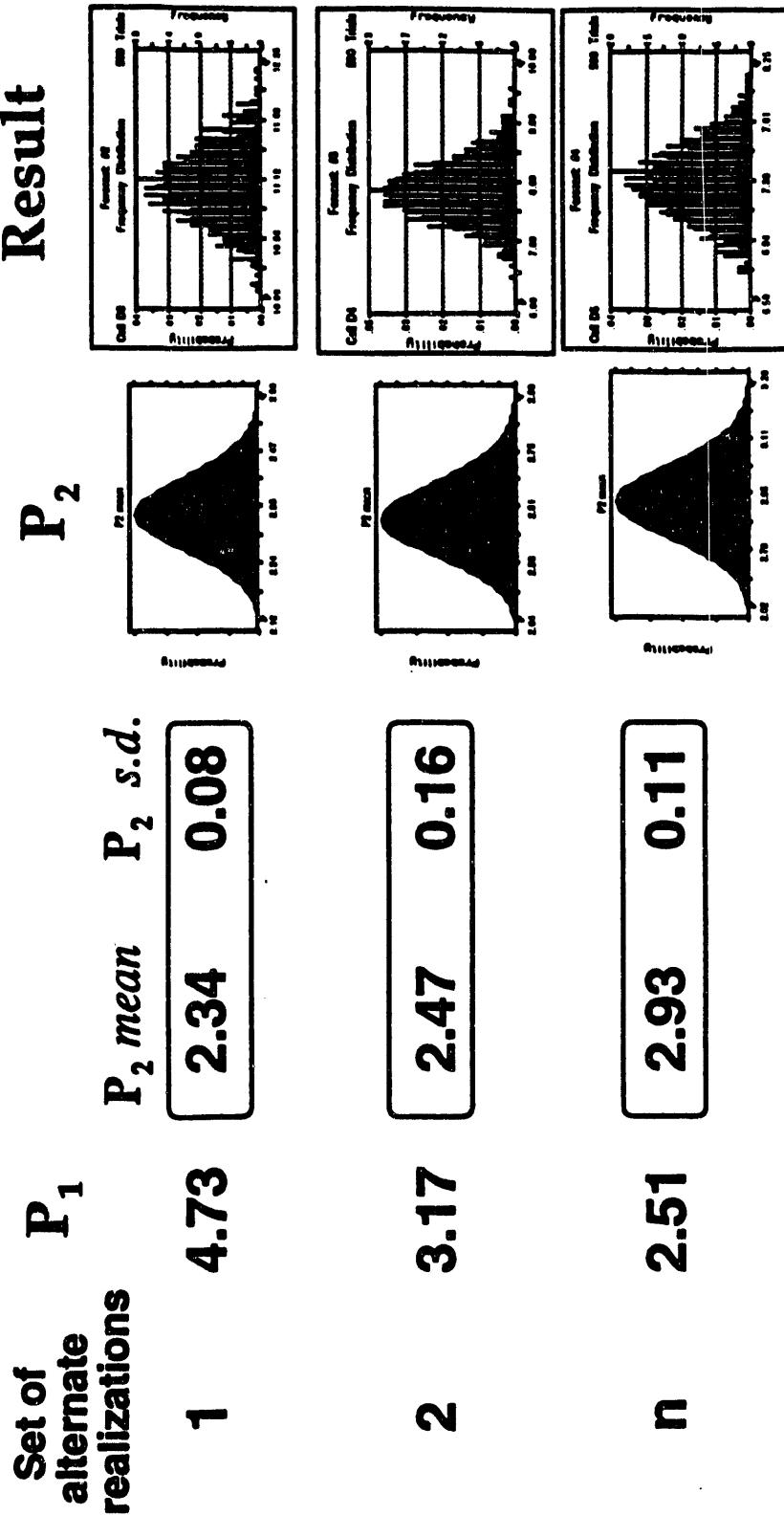


Fig. 2.5. Alternative values for a true but unknown value, P_1 , and alternative representations of a stochastic variable, P_2 , are combined to obtain numerous alternative realizations of stochastic variability for the assessment endpoint.

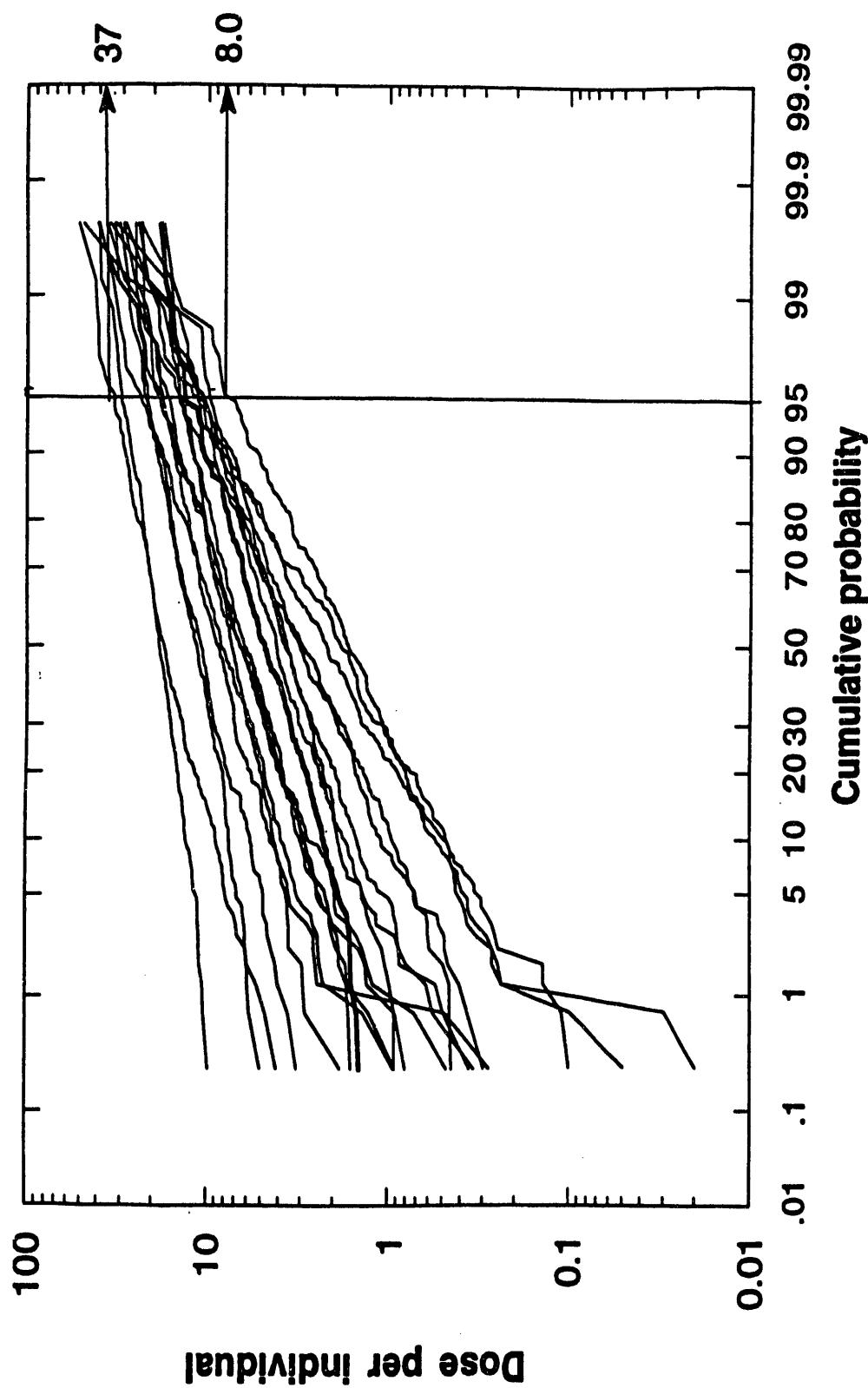


Fig. 2.6. A cumulative probability plot of alternative realizations of the stochastic endpoint. These alternative realizations represent lack-of-knowledge uncertainty about the true but unknown distribution. In this figure, an approximate 90% confidence interval is shown for doses occurring at the 95th percentile of the true but unknown distribution.

3. SUMMARY

The methods used to determine risk to an individual from ingestion of contaminated fish are straightforward. However, the methods currently recommended by EPA do not explicitly account for uncertainty and may tend to produce overly conservative estimates of risk by combining, through multiplication, several conservatively biased parameters. To compensate, an uncertainty analysis should be explicitly incorporated in risk assessments. As an alternative, the EPA's baseline methods should be more appropriately viewed as an initial screening tool. One may choose to perform the uncertainty analysis by using either analytical approaches (i.e., variance propagation techniques) or by using a Monte Carlo simulation package. Because variance propagation techniques can become complicated and time consuming, it is usually most efficient to use Monte Carlo methods.

You could say this report deals primarily with uncertainty about a true but unknown value. In more advanced problems, one may be interested in the uncertainty about a true but unknown distribution of values (i.e., the case for a distribution of risks per individual in a specified population of exposed individuals). This problem requires that CIs be obtained for the predicted distribution.

Incorporating uncertainty analysis into the risk assessment utilizes a major tool necessary in decision making. An uncertainty analysis will allow the risk assessor to rank the contaminants and pathways more accurately. In this manner, uncertainty analysis allows the assessor to see where further study is needed or where remediation must take place. Not only does uncertainty analysis allow one to rank pathways and contaminants, but it also provides a subjective probability distribution about which CIs can be formed to represent the risk.

This information can be used to guide decision analysis. For example, if a 5% lower confidence limit is above a regulatory standard of concern, then remediation is probably needed. If the 95% upper confidence limit is below the standard, remediation is probably not required. If the 95% upper confidence limit is above the standard but the 50th percentile is below the standard, further study should be recommended on those parameters that dominate the overall uncertainty. However, if the 50th percentile is above the standard, further study may still be recommended, but under some circumstances one may opt to proceed with remediation depending on the cost-effectiveness of measures for risk reduction.

Recognizing uncertainty provides more credibility to the risk assessment and is the first step in gaining more knowledge. We recognize that the most difficult task in quantitative uncertainty analysis is not associated with analytical or numerical methods for statistical error propagation in risk assessment equations; it is associated with the subjective judgment required to obtain subjective probability distributions for the uncertain model parameters. The extent of knowledge required to exercise this judgment often exceeds the capacity of any one individual. Thus, the judgment of several experts must often be solicited to estimate parameter uncertainty. Nevertheless, prior to

performing an uncertainty analysis for risk assessment, it may be useful to consider an ancient Arabic proverb that is cited in Finkel (1990).

“He who knows and knows he knows,
He is wise—follow him;

He who knows not and knows he knows not,
He is a child—teach him;

He who knows and knows not he knows,
He is asleep—wake him;

He who knows not and knows not he knows not,
He is a fool—shun him.”

REFERENCES

Beauchamp, John J. 1991. *Personal Communication*. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Blaylock, B. G., M. L. Frank, L. A. Hook, F. O. Hoffman, and C. J. Ford. 1992. *White Oak Creek Embayment Site Characterization and Contaminant Screening Report*. ORNL/ER-81. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Decisioneering, Inc. 1991. *Crystal Ball: User's Guide*. Denver, Colorado.

EPA (U.S. Environmental Protection Agency). 1989. *Interim Final: Risk Assessment Guidance for Superfund. Vol. I: Human Health Evaluation Manual*. OSWER Directive 9285.7-01a. EPA Office of Emergency and Remedial Response, Washington, D.C.

EPA (U.S. Environmental Protection Agency). 1991a. *Guidance for Risk Assessment*. Risk Assessment Council.

EPA (U.S. Environmental Protection Agency). 1991b. *Health Effects Assessment Summary Tables (HEAST)*. OERR 9200.6-303 (91-1). EPA Office of Emergency and Remedial Response, Washington, D.C.

Finkel, Adam M. 1990. *Confronting Uncertainty in Risk Management: A Guide for Decision-Makers*. Center for Risk Management, Resources for the Future, Washington, D.C.

Gardner, R. H., B. Rojder, and U. Bergstrom. 1983. *PRISM: A systematic method for determining the effect of parameter uncertainties on model predictions*. Studsvik Energiteknik AB report/NW-83/555, Nykoping Sweden.

Gardner, R. H. and J. R. Trabalka. 1985. *Methods of Uncertainty Analysis for a Global Carbon Dioxide Model*. DOE/OR/21400-4. Department of Energy Oak Ridge Operations.

Gardner, R. H. 1988. *TAM3: A Program Demonstrating Monte Carlo Sensitivity and Uncertainty Analysis*. Document prepared for the workshops of Biospheric Model Validation Study, BIOMOVS. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Hoffman, F. O. and R. H. Gardner. 1983. "Evaluation of Uncertainties in Radiological Assessment Models." Chapter 11 of *Radiological Assessment: A textbook on Environmental Dose Analysis*. J. E. Till and H. R. Meyer (eds.), NRC Office of Nuclear Reactor Regulation, Washington, D.C.

Hoffman, F. O., B. G. Blaylock, M. L. Frank, L. A. Hook, E. L. Etnier, and S. S. Talmage. 1991. *Preliminary Screening of Contaminants in the Off-Site Surface Water Environment Downstream of the U.S. Department of Energy Oak Ridge Reservation*. ORNL/ER-9. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

IAEA (International Atomic Energy Agency). 1989. *Evaluating the Reliability of Predictions Made Using Environmental Transfer Models*. IAEA Safety Series 100. Vienna, Austria.

Johnson, N. L., and S. Kotz. 1970. *Continuous Univariate Distributions*. Vol. 2. Houghton Mifflin Company, Boston, Massachusetts. p. 64-65.

Kanyar, Bela, and Sven P. Nielsen. 1989. *Users Guide for the Program TAMDYN*. Document prepared for the workshops of Biospheric Model Validation Study, BIOMOVS. Swedish National Institute for Radiation Protection, Stockholm, Sweden. Technical Report 4.

Klee, Albert J. 1986. *The MOUSE Manual*. U.S. Environmental Protection Agency, Cincinnati, Ohio.

Lyman, W. J., W. F. Reehl, and D. H. Rosenblatt (eds.). 1982. *Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Compounds*. McGraw Hill, New York.

NCRP (National Council on Radiation Protection and Measurements). 1989. *Screening Techniques for Determining Compliance with Environmental Standards, Releases of Radionuclides to the Atmosphere*. NCRP Commentary 3. Bethesda, Maryland.

NCRP (National Council on Radiation Protection and Measurements). In Press. Task Group 6. NCRP Scientific Committee 64.

Palisade Corporation. 1991. *@RISK: Risk Analysis and Simulation Add-In for Microsoft Excel; User's Guide*. Newfield, New York.

Williams, K. A. and C. R. Hudson II. 1989. *ORMONTE: An Uncertainty Analysis Code for Use with User-Developed Systems Models on Mainframe or Personal Computers; A User's Guide*. ORNL/TM-10714. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Worley, Brian A. 1987. *Deterministic Uncertainty Analysis*. ORNL-6428. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Appendix

FORMULAS FOR THE MEAN AND VARIANCE OF LOGARITHMS OF LOGNORMAL, LOG-UNIFORM, AND LOG-TRIANGULAR DISTRIBUTIONS

The following distributions are suggested for subjective probability distributions in analysis of multiplicative models.

Lognormal distribution (Hoffman and Gardner 1983):

$$\mu = \ln \left(\frac{\bar{x}}{\left[1 + \left(\frac{s}{\bar{x}} \right)^2 \right]^{0.5}} \right), \quad (\text{A.1})$$

and

$$\sigma^2 = \ln \left[1 + \left(\frac{s}{\bar{x}} \right)^2 \right], \quad (\text{A.2})$$

where \bar{x} is the arithmetic mean of the distribution and s is the standard deviation of the distribution.

Log-uniform distribution (Hoffman and Gardner 1983):

$$\mu = \frac{\ln(\min) - \ln(\max)}{2}, \quad (\text{A.3})$$

and

$$\sigma^2 = \frac{\left[\ln \left(\frac{\max}{\min} \right)^2 \right]}{12}. \quad (\text{A.4})$$

Asymmetrical log-triangular distribution (Beauchamp 1991, Johnson and Kotz 1970):

$$\mu = \frac{1}{3}[\ln(H^*) + \ln(b) + \ln(a)] , \quad (A.5)$$

and

$$\sigma^2 = \frac{1}{18} \{ [\ln(a)]^2 + [\ln(b)]^2 - [\ln(a)] + [\ln(b)] + [\ln(H^*)]^2 - [\ln(H^*)][\ln(a) + \ln(b)] \} , \quad (A.6)$$

where

H^* = the mode of the triangular distribution,
 b = the maximum of the triangular distribution,
 a = the minimum of the triangular distribution.

The following distributions are suggested for use as subjective probability distributions in analysis of additive models.

Normal distribution:

The mean value of the normal distribution is simply the value at the 50th percentile. With a normal distribution the median, mode, and mean are all the same. The variance of the normal distribution is the second central moment of the variable or the square of the standard deviation.

Uniform distribution:

$$\bar{x} = \frac{[\min + \max]}{2} , \quad (A.7)$$

and

$$s^2 = \frac{\left[\left(\frac{\max}{\min} \right)^2 \right]}{12} . \quad (A.8)$$

Asymmetrical triangular distribution (Beauchamp 1991, Johnson and Kotz 1970):

$$\bar{x} = \frac{1}{3}[H^* + b + a], \quad (A.9)$$

and

$$s^2 = \frac{1}{18}[(a)^2 + (b)^2 - (a)(b) + (H^*)^2 - (H^*)(a + b)]. \quad (A.10)$$

In addition to these suggested distributions, a few more distributions that one may use are Poisson, Weibul, gamma and beta distributions, custom designed, and any number of discrete distributions (Decisioneering 1991, Palisade Corporation 1991).

DISTRIBUTION

1. L. M. Anderson
2. A. Q. Armstrong
3. D. T. Bell
4. R. R. Bonczek
5. R. A. Brothers
6. J. S. Colley
7. M. F. P. DeLozier
8. D. M. Douthat
9. C. D. Goins
10. P. J. Halsey
11. R. Hull
12. A. D. King
13. R. C. Kramel
14. C. W. McGinn
15. D. B. Miller
16. P. D. Miller
17. P. T. Owen
18. T. S. Purucker
19. G. E. Rymer
20. R. A. Shaw
21. J. L. Skiles
22. M. E. Stack
23. D. R. Watkins
24. R. K. White
25. Central Research Library
- 26-30. ER Document Management Center
31. ORNL Laboratory Records
32. Office of Assistant Manager for Energy Research and Development, Department of Energy
Oak Ridge Field Office, P.O. Box 2001, Oak Ridge, Tennessee 37831-8600
33. J. S. Hammonds, SENES Oak Ridge, Inc., 677 Emory Valley Road, Suite C, Oak Ridge, TN 37830
34. F. O. Hoffman, SENES Oak Ridge, Inc., 677 Emory Valley Road, Suite C, Oak Ridge, TN 37830
35. D. W. Swindle, Radian Corporation, 120 South Jefferson Circle, Oak Ridge, TN 37830
- 36-37. R. L. Nace, Branch Chief, Nonenrichment Facilities, Oak Ridge Program Division, Office of Eastern Area Programs, Office of Environmental Restoration, EM-423, Trevion 2, U.S. Department of Energy, Washington, DC 20585
- 38-47. R. C. Sleeman, DOE Oak Ridge Field Office, P.O. Box 2001, Oak Ridge, TN 37831-8541
- 48-49. H. M. Thron, Chief, Enrichment Facilities, Oak Ridge Program Division, Office of Eastern Area Programs, Office of Environmental Restoration, EM-423, Trevion 2, U.S. Department of Energy, Washington, DC 20585
- 50-51. J. T. Sweeney, DOE Oak Ridge Field Office, P.O. Box 2001, Oak Ridge, TN 37831-8541
- 52-53. Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831

END

DATE
FILMED

3/29/93

