Quantitative Analysis of Uncertainty and Variability in Environmental Policy Making

H. Christopher Frey, Ph.D.

AAAS/EPA Environmental Science and Engineering Fellow, Summer 1992

and

Research Associate
Center for Energy and Environmental Studies
Department of Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA 15213

September 1992

Note: H.C. Frey's current address is:

Department of Civil Engineering
North Carolina State University
Raleigh, NC 27695-7908
E-mail: frey@eos.ncsu.edu
Telephone: (919) 515-1155
Fax: (919) 515-7908

ACKNOWLEDGMENTS

I spent the summer of 1992 working as an American Association for the Advancement of Science (AAAS) and U.S. Environmental Protection Agency (EPA) Environmental Science and Engineering Fellow. This fellowship program is administered by AAAS under EPA sponsorship. My assignment was in the Exposure Assessment Group (EAG) of the Office of Health and Environmental Assessment, which is part of the Office of Research and Development (ORD) of EPA in Washington, DC. While in EAG, I worked closely with Paul White and enjoyed many long and thorough discussions with him of key concepts in uncertainty analysis. I benefited from interactions with many other individuals at EPA headquarters in ORD, the Office of Policy, Planning, and Evaluation, the Office of Air and Radiation, and the Office of Solid Waste. Thanks also to Max Henrion of Lumina Decision Systems for his helpful comments. I would like to thank the staff at AAAS for facilitating my stay at EPA and for providing a stimulating orientation and seminar program. In particular, thanks to Chris McPhaul, Claudia Sturges, and Patricia Curlin of AAAS, and to Karen Morehouse of EPA. Finally, while I am grateful for the input of many people at EPA and elsewhere, all of the text that follows is my sole responsibility.

DISCLAIMER

This draft report was prepared during the author's tenure as a Fellow under the American Association for the Advancement of Science/U.S. Environmental Protection Agency Environmental Science and Engineering Fellowship Program during the summer of 1992. The views expressed herein are entirely the author's and do not represent official policy of either the EPA or AAAS. The report is subject to further review and revision. Mention of trade names or commercial products does not constitute endorsement or recommendation.

ABSTRACT

Nearly all analyses of environmental problems are confronted with uncertainty. The implications of these uncertainties are particularly critical in the assessment and selection of regulatory options. Current practices within the regulatory community do not adequately deal with uncertainty. These practices, as embodied in EPA guideline documents, often place more emphasis on post-hoc qualitative approaches to caveating uncertainties. In contrast, a quantitative approach to uncertainty analysis is proposed as the most appropriate way to deal with uncertainties. The benefits of such an approach include more proper and explicit characterization of the state-of-knowledge about science and technology without the burden of simplifying, policy-motivated assumptions, concise communication of such information to decision-makers, and a capability to quantitatively identify modeling directions, data requirements, and research needs. General methodological issues of uncertainty analysis are discussed. These issues are illustrated based on a review of uncertainty analysis activities in exposure assessment. In exposure assessment and other environmental analyses, a distinction must be made between quantities which are uncertain, due to a lack of knowledge of some type, and those which are variable from one individual to another. A two-dimensional approach to Monte Carlo simulation is developed and illustrated to properly model these conceptual differences. The implications of this approach for data requirements, modeling, and communication of results to decision makers are discussed. With respect to uncertainty analysis in general, recommendations are made regarding the needs for appropriate training, data, exemplary case studies, peer review, and handbooks. Quantitative approaches to uncertainty analysis generate insights that help both analysts and decision-makers ask the right questions about environmental problems.

Table of Contents

List of Figures

List of Tables

1.0 INTRODUCTION

Nearly all environmental problems faced today entail some element of uncertainty. Estimates of emissions rates, fate and transport of pollutants, human exposure to pollutants, health effects, costs and benefits of regulatory options, and other attributes of environmental analyses are often fraught with uncertainty. Typically, only limited data are available to support an analysis, if in fact any data are available at all. Yet regulatory decisions must be made by the U.S. Environmental Protection Agency (EPA) and other agencies based on such information. These decisions should be founded on as complete an assessment of scientific and technical uncertainty as possible, to permit the identification of strategies which are robust even in the face of uncertainty, and to identify priorities for further research.

In developing estimates of the values of key quantities in environmental problems, a common approach is to assume a "best guess" point-value based on some combination of data and technical judgment. These judgments may be intended to represent neither undue optimism or pessimism, or they may be intended to incorporate a degree of conservatism. However, the basis for many assumptions, and the scope of thought that went into them, are often not explicitly documented in policy studies. Thus, the degree of confidence that a decision-maker should place in the estimates when evaluating regulatory alternatives is often not rigorously considered.

The most common approach to handling uncertainties is either to ignore them or to use simple "sensitivity" analysis. In sensitivity analysis, the value of one or a few model input parameters are varied, usually from "low" to "high" values, and the effect on a model output parameter is observed. Meanwhile, all other model parameters are held at their "nominal " values. In practical problems with many input variables which may be uncertain, the combinatorial explosion of possible sensitivity scenarios (e.g., one variable "high", another "low," and so on) becomes unmanageable. Furthermore, sensitivity analysis provides no insight into the *likelihood* of obtaining any particular result. Thus, while they indicate that a range of possible values may be obtained, sensitivity results do not provide any explicit indication of how a decision-maker should weigh each possible outcome.

A quantitative approach to uncertainty analysis is proposed as the most appropriate way to deal with uncertainty. Deterministic estimates, based on "best guess" point estimates, are often wrong or misleading in several ways: (1) they are often biased away from the mean values of the uncertainties they represent; (2) they provide no indication to decision makers regarding the magnitude of underlying uncertainties; and (3) they permit no indication of the key sources of uncertainty. Deterministic estimates are not based on complete and quantitative consideration of interactions among multiple, simultaneously uncertain variables, which are especially dangerous to overlook in the cases of skewed uncertainties and/or complex nonlinear models. Ignoring or surpressing uncertainty in environmental risk assessment often results in a misleading sense of confidence about numbers.

In contrast, quantitative estimates of uncertainty more properly characterize the state of knowledge affecting a regulatory decision, and more properly convey to decision makers the magnitude of the uncertainties in key quantities of interest (e.g., exposure levels, emission rates, risks, etc.). Furthermore, through simple extensions of traditional Monte Carlo techniques, it is possible to identify both the key sources of uncertainty which merit further research, as well as to identify uncertain factors which are unimportant to a given decision. The importance of the latter may be unappreciated unless one considers the number of often useless arguments that fixate on minutia, especially in emotionally-charged policy debates. In the process of identifying factors that really matter, quantitative uncertainty analysis can lead to more informed and focused policy debates.

1.1 Environmental Problem Domains With Uncertainties

Within the U.S. EPA, some of the key types of assessment activities subject to uncertainty include risk assessment, technology assessment, and economic analysis. Some of the features of these types of assessments are briefly summarized.

Risk assessment can be divided into a number of distinct activities. These include hazard identification, emissions characterization, exposure assessment, dose-response assessment, and risk characterization (Stine, 1992). Hazard identification is the stage at which a determination is made as to whether or not a biological or human health hazard exists for a range of biological process from a given chemical. Emissions characterization involves estimation of the sources and rates of emissions, based on process modeling, emission factor calculations, measurement of emissions at stacks and vents, or monitoring of ambient pollutant levels. Exposure assessment is generally concerned with estimating the transport of contaminants from emissions sources to bodily contact with an exposed population, including the intensity, frequency, and duration of contact and spatial and temporal patterns of the contaminant. The dose-response assessment involves estimating the relationship between exposure (or actual absorbed dose) and biological effects within individuals of an exposed population. Risk characterization is the final step in a risk assessment, and it involves developing numerical and qualitative descriptions of the types and magnitudes of adverse effects, the probabilities that the effects will occur, and a discussion of uncertainties and analytical assumptions.

Technology assessment refers here to the estimation of the performance, emissions, and cost of technologies which emit, control, or remediate discharges to the environment. Benefit/cost analysis refers to any analyses which attempt to estimate the cost and economic effects of regulatory strategies, and which attempt to quantify direct or indirect benefits and costs on some common basis.

Together, these three types of assessments often play a complimentary role in the analysis of regulatory options and in the decisions made by policy makers within EPA. Alternative technologies may be available to reduce the emissions of potentially hazardous pollutants. Thus, technology assessment, risk assessment, and benefit/cost analysis may all be factors in a regulatory analysis, each subject to various sources of uncertainty. This report deals with methodological aspects of how to deal with these uncertainties.

1.2 Scope of this Report

In the next section, motivations for doing uncertainty analysis with specific focus on the U.S. EPA are discussed. A snapshot of some current activities at EPA related to uncertainty analysis is given. General methodological aspects of uncertainty analysis are reviewed in Chapter 3. Aspects of uncertainty analysis specific to one problem domain, exposure assessment, are discussed in Chapter 4. In Chapter 5, an illustrative case study of a two-dimensional Monte Carlo approach to simulating both variability and uncertainty is presented. A hypothetical exposure model is used for this purpose. The implications of a quantitative approach to the analysis of variability and uncertainty are presented in Section 6.

2.0 MOTIVATIONS FOR UNCERTAINTY ANALYSIS

Analyses of many environmental problems involve uncertainties, which are often ignored or treated in a limited way using sensitivity analysis. However, sensitivity analysis suffers from shortcomings resulting from the difficulty in evaluating the effect of *simultaneous* variations in several parameters and the lack of insight into the *likelihood* of any particular result. These shortcomings are especially important for nonlinear models, in which the results may be sensitive to a given input variable only when other input variables take on certain values. These types of potentially complex interactions require a more integrated approach to the assessment of uncertainty.

A more robust approach is to represent uncertainties in model parameters using probability distributions. Using probabilistic simulation techniques, simultaneous uncertainties in any number of model input parameters can be propagated through a model to determine their combined effect on model outputs. The result of a probabilistic simulation includes both the possible range of values for model output parameters, and information about the likelihood of obtaining various results. This provides insights into the downside risks or potential pay-offs of alternative regulatory strategies. Statistical analysis on the input and output data can be used to identify trends (e.g., key input uncertainties affecting output uncertainties) without need to re-run the analysis. Thus, probabilistic analysis can be used as a research planning tool to identify the uncertainties in a problem that matter the most, thereby focusing research efforts where they are most needed. Probabilistic analysis may be referred to elsewhere as "range estimating" or "risk analysis".

The development of ranges and probability distributions for model input variables can be based on information available in published studies, statistical data analysis, and/or the judgments of technical experts with relevant problem domain experience. The approaches to developing judgments about probability distributions are similar in many ways to the approach one might take to pick a single "best guess" number for deterministic (point-estimate) analysis or to select a range of values to use in sensitivity analysis. However, the development of estimates of uncertainty usually requires more detailed thinking about possible outcomes and their relative likelihoods. This is an advantage for the analyst, because by thinking systematically about alternative outcomes, the analyst is more likely to uncover "surprises" that might otherwise have been overlooked.

2.1 Motivations for Uncertainty Analysis within EPA

The use of quantitative approaches to uncertainty analysis within EPA is currently very limited, due partly to the lack of widespread training in methodological aspects of uncertainty analysis. However, the development of in-house capabilities to perform uncertainty analysis is being pressured by the increasing use of uncertainty analysis by EPA contractors and other analysts outside the agency. EPA analysts need to be able to critically evaluate externally conducted studies which often become a basis for exposure, risk, and other types of estimates in policy making.

Furthermore, within EPA today, there are a number of institutional motivations for the increased used of quantitative uncertainty analysis. These motivations, which cut across disciplines and programs, are converging to require analysts at EPA to more fully and quantitatively characterize uncertainties in their assessments, and to communicate information about uncertainty to risk managers and other policy makers. A few of these motivations are studies by the National Academy of Sciences (NAS), practices outside EPA, requests for uncertainty analyses by the Office of Management and Budget (OMB), recommendations by the EPA Science Advisory Board (SAB), and a recent memorandum on risk characterization by EPA

Deputy Administrator F. Henry Habicht, II (the "Habicht memo'). In addition, there has been recent discussion about a possible presidential Executive Order which would require quantitative uncertainty analysis in the evaluation of regulatory options.

2.1.1 The 1983 NAS Report on Risk Assessment

A National Academy of Sciences (1983) study identified "pervasive uncertainty" as "the dominant analytic difficulty" in linking exposure to chemicals to chronic health risks. Sources of uncertainty include incomplete data, the types, probabilities, and health effects associated with exposure to a chemical agent, economic effects of regulatory action, and the extent of possible current and future exposures. The two key recommendations of the study were that: (1) regulatory agencies should establish a clear conceptual distinction between scientific assessment of risks and policy judgments in the choice of regulatory options; and (2) uniform guidelines should be developed for the risk assessment process.

It is often the case that, in a scientific assessment, there is no consensus on the appropriate "inference options" or plausible theories to employ. To cope with this type of uncertainty, various programs and offices within EPA have developed "guidelines" for how to perform scientific assessments, as well as "handbooks" of acceptable numerical values to use for key variables in risk assessment equations. A specific example of this approach is the "Risk Assessment Guidelines of 1986" (EPA, 1987) and the "Exposure Factors Handbook" (EPA, 1989). While such guidelines have the intended effect of streamlining the assumptions and approaches used by different analysts, they also have the effect of building policy assumptions into the risk assessment process. For example, the recommended approach to exposure and risk assessment, as well as the recommended values to use for specific parameters, are often based on conservative assumptions. A key concern is that appreciation for these conservative assumptions, as well as for other sources of uncertainty, is lost in the process of communicating assessments to regulators and then to the public.

Ironically, the 1983 NAS study tangentially addressed the notion of uncertainty analysis, but then dismissed it. The study suggests that it would be possible to develop risk assessments based on explicit consideration of alternative "plausible inference options." The study claims that such an assessment "could result in such a wide range of risk estimates that the analysis would not be useful to a regulator or to the public." If in fact such an estimate were obtained, it would reflect the scientific uncertainties in the risk assessment, and would leave for regulators the perhaps uncomfortable role of how to interpret the uncertain scientific information in the context of political, economic, and social considerations. In reality, the proper characterization and communication of uncertainty would disentangle many policy judgments from the assessment process. While the resulting estimates may look messy to a regulator, they are in fact more scientifically defensible than point-estimates based on policy-motivated assumptions.

A critique offered by the NAS study is that regulators might have the option of "ad hoc exercise of risk assessment policy decisions" if faced with a broad range of policy values from a risk assessment. However, what this argument ignores is that the risk manager will be required to justify the selection of any single value from the distribution of plausible outcomes and explain it to the public through the regulatory process. Furthermore, the regulator is typically also faced with uncertainties about key economic or other factors that must be included in the decision-making process. The regulator should be expected to explicitly consider all sources of uncertainty in choosing among regulatory options. The current process is more arbitrary, in that risk managers are often presented with only a point estimate (representing a single draw from the distribution of possible values) and a qualitative laundry list of caveats. In many typical case studies, these types of judgments, which are properly in the policy domain, are imbedded in the "scientific" assessment process through assessment guidelines or standard accepted approaches.

According to the NAS study, guidelines must, by their very nature, include a mixture of policy and science. As a practical matter, this is probably true — it may be impossible to completely separate policy judgments from "scientific assessment." Even purely scientific assessments entail some elements of subjectivity. However, it should be the goal of guidelines to minimize the policy content within assessments. The NAS report indicates that guidelines cannot effectively address case-specific factors such as quality of data or strength of evidence. However, it is the case that guidelines can (and should) promote consistent and defensible *methodological* approaches to these issues. For such ideal guidelines to be effective, assessors and analysts must be properly trained in the methodological techniques.

An observation of the NAS study panel was that "preparation of fully documented written risk assessments that explicitly define the judgments made and attendant uncertainties clarifies the agency decision-making process and aids the review process considerably" (p.148). The report goes on to state:

> *Conclusions based on a large number of sequential, discretionary choices necessarily entail a large, cumulative uncertainty. The degree of uncertainty may be masked to some extent, when, in the final form of an assessment, risk is presented with an associated measure of statistical significance. If they are to be most instructive to decision makers, assessments should provide some insight into qualitative characteristics of the data and interpretations that may impute more or less certainty to the final results. (p.165).*

The above statement is contradictory. First, any conclusion based on a series of discretionary choices regarding point estimates for model input assumptions completely sidesteps any quantitative treatment of uncertainty. Thus, it is more accurate to say that the degree of uncertainty is nearly completely masked. Second, it has rarely been the case that risk assessments are presented with any type of measure of statistical significance, although this is beginning to change. Third, while qualitative laundry lists of caveats may provide some vague indication of the degree of uncertainty of the estimate, they do not provide any quantitative indication of uncertainty. As a result, it is often difficult to identify whether the risks of exposure to one chemical are really higher or lower than the risks from another chemical, or to identify which uncertain factors in the assessment contribute most to uncertainty in the risk estimate. This is not meant to minimize the role of qualitative description of uncertainty. In some cases, one can do no better than that because of the scarcity of data to support any quantitative assessments or judgments. This is especially true in the case of dose-response models in risk assessment. However, it too often appears to be the case that a qualitative approach is used instead of a quantitative approach for convenience.

The use of guidelines and handbooks to justify the selection of point-estimates for input assumptions results in a tighter entanglement of policy within the assessment process, whereas a full display of uncertainties through probability distributions, where possible, provides a more scientifically defensible approach. Uniform guidelines should focus less on the specifics of what equations and parameter assumptions to use, and more on the methodology of how to properly conduct an assessment.

The NAS is currently conducting a new study intended to provide guidance to EPA on methodological approaches to risk assessment for air toxics. It is expected, however, that the recommendations of this study will be applied to exposure and risk assessment activities in general. The committee is reported to be considering the role of uncertainty analysis as part of exposure and risk assessment (Stine, 1992).

2.1.2 Practice Outside EPA

An increasing number of investigators outside EPA are employing quantitative approaches to uncertainty analysis in the context of exposure and risk assessment. Some examples of these are: Alberts and Weyman, 1991 (the same paper is authored by Alberts and Rouge, 1992); Bogen and Spear, 1987; Constantinou et al, 1992; Copeland et al, 1992; Eschenroeder and Faeder, 1988; Hattis et al., 1988; Johnson et al., 1992; Li et al., 1992; McKone and Bogen, 1991, 1992; McKone and Ryan, 1989; Morgan et al, 1985; Roseberry and Burmaster, 1992; Shlyakhter et al, 1992; Thompson and Burmaster, 1991; Thompson et al, 1992; Wilson et al. (1985), and Zankel et al, 1991. This listing is not intended to be all inclusive, but serves the purpose of indicating the trends related to exposure and risk assessment. A report by Rish and Marnicio (1988) identifies a number of earlier studies related to uncertainty in risk analysis.

Quantitative uncertainty analysis is also applied in other domains related to EPA's missions. These include technology assessment and economic analysis. Some examples of these types of applications are given by Apostolakis (1990), Frey (1991), Frey and Rubin (1991,1992), and Heslin and Hobbs (1991).

2.1.3 Requests by OMB

In several instances, OMB has requested EPA to explicitly consider uncertainties in developing estimates of the costs of environmental regulations. OMB has the authority to review proposed regulations with respect to economic impacts. In one example, an expert elicitation approach was used by EPA to estimate uncertainties in the costs and economic impacts of land disposal restrictions for newly listed wastes and contaminated debris (EPA, 1992). Presumably the intent of the request by OMB in this case was to assure that the rulemaking did not fall under the category of a "major rule." One of the criteria for a major rule is that the incremental annual effect to the economy is $100 million or more. The results of the analysis indicated a high probability that the economic effect of the rulemaking would be less than $100 million per year. Thus, the analysis in this case both characterized the degree of uncertainty in the economic analysis, and provided robust indication that, regardless of the uncertainty, the costs would be well below $100 million.

2.1.4 EPA Science Advisory Board

The SAB has been involved in peer review of recent EPA guidelines, including the 1992 Exposure Assessment Guidelines (EPA, 1992). The SAB (1992) was asked to comment specifically on the types of exposure estimators recommended by the guidelines, as well as the general discussion of uncertainty analysis in the guidelines. The SAB recommended a graphical representation of exposure estimators corresponding to various percentiles of a probability distribution of exposures. The SAB also commented that the presentation of uncertainty analysis in Chapter 6 of the guidelines was "comprehensive and scientifically correct," and further noted the "strong statement of the importance of uncertainty assessment in exposure assessment." The SAB commented favorably that the guidelines recognized the elements of "scientific judgment" required to perform uncertainty analysis. On the issue of presenting results, the SAB considered a standard format, but abandoned the approach due to concerns about stifling creativity.

In an earlier resolution on the use of mathematical models at EPA, the SAB (1989) recommended that:

> *Sensitivity and uncertainty analyses of environmental models and their predictions should be performed to provide decision-makers with an understanding of the level of confidence in model results, and to identify key areas for future study.*

The resolution also recommended that uncertainties in both model structure and model parameters should be evaluated to "determine possible effects on the ultimate regulatory decision." Furthermore, although not directly addressing the issue of uncertainty analysis, the SAB stated that "peer review at various levels is required to ensure proper model development and application."

From these two examples, it appears that the SAB is generally sympathetic to the notion of quantitative uncertainty analysis and its use within EPA.

2.1.5 Guidance Documents

Over the years, increasing attention has been given to uncertainty analysis in the formulation of guidance documents. In response to the 1983 NAS study, EPA program offices, and the Office of Research and Development (ORD), have developed a number of guidance documents dealing with accepted methods for performing exposure and risk assessments. Examples of these include the The Risk Assessment Guidelines of 1986 (EPA, 1987), the Superfund Exposure Assessment Manual (EPA, 1988), the Human Health Evaluation Manual for Superfund (EPA, 1989), and the ORD Guidelines for Exposure Assessment (EPA, 1992).

The 1986 Risk Assessment Guidelines were developed in direct response to the 1983 NAS study. This document includes guidance for carcinogen risk assessment, mutagenicity risk assessment, health assessment of suspect developmental toxicants, health risk assessment of chemical mixtures, and estimating exposures. Generally, the approaches to dealing with uncertainty can be divided into categories of qualitative "laundry lists" of caveats appended to the results, sensitivity analysis, and quantitative uncertainty analysis.

The 1988 Superfund Exposure Assessment Manual discusses sources of uncertainty in models and model parameters and approaches to deal with uncertainty. In some instances, statements in the manual are incomplete, while in others they are inaccurate. For example, it is stated without support that Monte Carlo analyses requires an assumption of statistical independence among input parameters. While this is the case in traditional Monte Carlo simulation, techniques exist to induce correlations among input variables (e.g., Iman and Conover, 1982). Moreover, other approaches are available to deal with correlations. These are discussed in Section 3.3 While many opponents of quantitative uncertainty analysis use excuses based on misconceptions about methodological approaches and capabilities, they overlook other more appropriate approaches to dealing with correlations.

An example of another misconception is found in the 1988 manual, which contains a statement that "assuming an input parameter distribution does not help to reduce uncertainty." There seems to be a notion in some of the EPA literature that the purpose of a quantitative uncertainty analysis is somehow to "reduce" uncertainty. This could not be farther from the truth. In reality, uncertainty analysis is an attempt to quantify the degree of confidence that an analyst has in the existing data and models, based on whatever information is currently available. In the face of sparce or non-existing data, one may elicit the judgments of problem domain experts in the form of subjective probability distribution functions (SPDF) using standard elicitation protocols (e.g., Morgan and Henrion, 1990). This hardly constitutes "assuming" an arbitrary distribution. While it is true, and often stated by critics of quantitative uncertainty analysis, that model results are only as good as model assumptions (or "garbage in, garbage out"), it is also the case that completely ignoring uncertainty when it is known that uncertainty exists yields misleading point-estimates. It is certainly better to make a good faith effort to quantify uncertainties, and caveat the judgments that went into the process, than to completely side-step uncertainty analysis justified on the basis of incoherent criticisms.

The 1988 manual states with some basis that "a quantitative appraisal of the uncertainty is the most preferably way to express the uncertainty," but that a qualitative approach is often "the most viable way to express the level of uncertainty." The difficulty of the latter approach is that it provides no clear indication of how good the estimates are, nor does it provide a strong basis for prioritizing future data collection or research.

The 1989 Superfund Human Health Evaluation Manual states that "estimating a probability distribution for exposures and risks can lead to a false sense of certainty about the analysis." The manual stresses that it is important to identify all sources of uncertainty which were not included in the quantitative analysis. While this is good advice, it also appears to be the case that both Superfund manuals tend to place more emphasis on qualitative approaches to dealing with uncertainty. This may be a program-specific recommendation, due to the large number of Superfund sites and the impracticability of developing detailed information for every site. Nonetheless, it is precisely for problems where data are limited that quantitative uncertainty analysis can provide an illuminating role, to help identify how robust conclusions about exposures and risks are, and to help target data gathering efforts.

The 1992 Exposure Assessment Guidelines represent current EPA thinking about approaches to characterizing variability and uncertainty. The guidelines detail the role of uncertainty analysis, types of uncertainty, and variability. Approaches to dealing with uncertainty are briefly summarized. In general, the guidelines provide a comprehensive overview of key concepts related to uncertainty analysis. However, there are few shortcomings in some of the technical discussion.

For example, the notion of "reducing" uncertainty by characterizing it appears in the 1992 guidelines. The guidelines state: "when the results are significantly affected by a particular parameter, the exposure assessor should try to reduce uncertainty by developing a description of the likely occurrence of particular values within the range." The guidelines then offer several suggestions for estimating a probability distribution for the parameter in questions. Fundamentally, the process of estimating uncertainty in a parameter is to fully describe the range and likelihood of values for the parameter, not to arbitrarily narrow the range or to somehow "reduce" uncertainty. Reductions in uncertainty can only come from gathering additional information. In reality, gathering additional information may or may not actually reduce the estimates of uncertainty, depending on how complete the prior understanding of uncertainty was.

Another shortcoming of the guidelines are statements regarding the use of Monte Carlo simulation. The guidelines indicate that it is cumbersome to determine the sensitivity of model output uncertainties to assumptions regarding uncertainties in model inputs. Furthermore, it is stated that Monte Carlo results provide no indication of which variables are the most important contributors to uncertainty. While these statements are true, they are misleading due to omission. Statistical techniques can be applied relatively easily to answer these questions. For example, various regression techniques can be used to identify the linear correlations between uncertainties in model outputs and model inputs. Such correlations can provide an indication of the relative importance of each input uncertainty in determining output uncertainty. Variations of regression analysis can be employed when nonlinearities exist in the model. Such techniques are discussed in Section 3.6.4. EPA guidelines often fall short of fully discussing the variety of techniques that are available to assist an analyst in dealing with perceived shortcomings of Monte Carlo simulation, such as correlated inputs, identification of key uncertainties, and so on. However, in many other respects, and especially in the case of the 1992 Exposure Assessment Guidelines, these documents often do present a concise summary of the philosophy and use of quantitative uncertainty analysis techniques.

2.1.6 1992 EPA Memorandum on Risk Characterization

The February 1992 "Habicht memo" (Habicht, 1992) provides guidance on the development and use of risk characterizations, and distinguishes the role of risk assessors from risk managers. The memo emphasizes the need to be "completely candid" about uncertainties in characterizing risks and justifying regulatory decisions.

The memo reinforces the idea that risk assessment and risk management should be distinct and separate activities. A key statement in this regard is that:

> *Matters such as risk assessment priorities, degree of conservatism, and acceptability of particular risk levels are reserved for decision-makers who are charged with making decisions regarding protection of public health.*

In contrast, risk assessors have the tasks of generating credible, scientifically defensible analyses, with clear indications of the uncertainties and limitations of the assessments.

An often heard, but ill-founded, criticism of uncertainty analysis on the part of some EPA analysts is that any rigorous treatment of uncertainty destroys the confidence one might otherwise have in a numerical point estimate. On this point, the memo is quite clear: "a balanced discussion of reliable conclusions and related uncertainties enhances, rather than detracts, from the overall credibility of each assessment." The goal here is to identify scientific conclusions in risk assessments and regulatory options in risk management which are robust in the face of uncertainties, and to identify areas in which additional information would be required to fill data gaps or to resolve uncertainties.

The memo also states that, "effective immediately," EPA policy requires that information on the range of exposures and multiple risk descriptors be presented in all exposure and risk characterizations. The intent is to provide an indication of the central ranges of exposure and/or risk, high end exposures and risks, and sensitive subgroups especially at risk. Individual exposure and risk is to be characterized using measures such as central tendency (mean or median) and high end (e.g., above the 90th percentile). Examples of different types of population risk descriptors to use include total number of cases of a health effect for a population over a certain time, the portion of the population within some range of risks, and the number of individuals in a population at or above some risk level. Highly susceptible subgroups, such as infants, pregnant women, the elderly, or other highly exposed or highly at risk groups, are to be identified. Furthermore, information of this type is to be retained at all stages of risk management, to "present a more complete picture of risk that corresponds to the range of different exposure conditions encountered by various populations exposed to most environmental chemicals."

While the memo does not directly address other types of analyses which decision-makers must integrate in selecting regulatory options, it does indicate that the same type of rigor in characterizing uncertainties in societal considerations and economic analyses is expected.

> *Risk management decisions involve numerous assumptions and uncertainties regarding technology, economics, and social factors, which need to be explicitly identified for decision makers and the public.*

Thus, the Habicht memo requires more attention to characterization of the variability in exposures and risks faced by different members of populations subject to exposure, as well as more attention to the uncertainties associated with such estimates. The memo indicates that techniques such as Monte Carlo analysis can be used for this purpose, but provides no rigid "cookbook" guidance on how to implement the new policy. This is, of course, the proper course, because the selection of appropriate techniques is often a problem-specific concern that should be left to the judgment of the analyst.

Of all of the motivations for quantitative approaches to analyses of variability and uncertainty which currently face EPA, the Habicht memo is perhaps the most salient. It provides the clearest directives, in terms of actual agency policy, in this regard.

2.2 Uncertainty Analysis Within EPA

A number of activities within EPA reflect both a long term interest in uncertainty analysis on the part of relatively few EPA analysts, as well as the current growing interest in uncertainty analysis in response to various motivations, such as the ones detailed above. A few examples are briefly discussed here. These include the development of uncertainty analysis software tools within EPA, the formation of an uncertainty discussion group at EPA headquarters in Washington, EPA-sponsored workshops on uncertainty analysis and related topics, and a limited number of in-house projects and sponsored studies involving some type of quantitative analysis of uncertainty.

2.2.1 Software Tools

Several software tools have been developed by EPA which have Monte Carlo simulation capabilities. Furthermore, a number of commercial or public domain programs are available. Of the EPA-supported tools, the most prominent s MOUSE, which is an acronym for Modular Oriented Uncertainty SystEm (Klee, 1992). MOUSE can be used to simulate uncertainties in models consisting of one or more equations. MOUSE has a number of built in capabilities, including probability distributions, graphics, sensitivity analysis, and so on, which facilitate the development and interpretation of Monte Carlo simulations. The author reports that MOUSE has been applied to a variety of environmental problems, including study of the migration of pollution plumes in streams, establishment of regulations for hazardous wastes in landfills, and estimation of pollution control costs. The manual for MOUSE also has a good introduction which motivates the need for uncertainty analysis and illustrates some of the key insights obtained.

Another software tool is the Statistical Simulation System developed by Smith (1992) in the Statistical Policy Branch of the Office of Policy Planning and Evaluation at EPA headquarters in Washington, DC. This software supports Monte Carlo simulations of relatively simple models, and offers ways to evaluate various statistics of probability distributions. A third tool is a Monte Carlo simulation shell which can be used in conjunction with a Fortran program. All three programs discussed here run on an IBM PC.

Other software is commercially available to support uncertainty analysis. Two of the most popular are @risk and Crystal Ball. The former is an add-on to the Lotus 1-2-3 spreadsheet software available for the IBM PC, while the latter is an add-on to the Microsoft Excel spreadsheet software available for the Macintosh family of computers. Demos is a Macintosh-based graphical environment for creating, analyzing, and communication probabilistic models for risk and policy analysis (Morgan and Henrion, 1990). Demos has been applied in a wide variety of environmental policy problems, including evaluation of the effects and costs of acid rain and mitigation strategies, assessment of advanced power generation and environmental control technologies, and exploration of the effects of global climate change.

Iman et al. (1984) have developed Fortran-based programs for generating samples using Monte Carlo or Latin Hypercube sampling and for analyzing modeling results using various linear regression techniques (Iman et al., 1985). These programs can be adopted to any modeling platform for which Fortran is available.

2.2.2 The Uncertainty Circle

During the summer of 1992, a group of analysts from various disciplines and EPA programs formed an "Uncertainty Circle." The Circle meets periodically (every month or so) to discuss issues related to the analysis of uncertainty and variability. Recent activities of the Circle have included development of an inventory of uncertainty analysis activities at EPA, seminars on methodological issues of uncertainty analysis, and discussion of practical aspects of performing and communicating probabilistic analyses.

2.2.3 Workshops

EPA has sponsored a small handful of workshops on various topics related to the analysis of uncertainty in environmental problems. For example, as part of the Research to Improve Health Risk Assessments (RIHRA) program (EPA, 1990), EPA invited a number of experts to participate in a workshop in 1988. In June of 1992, the Exposure Assessment Group co-sponsored a workshop on methodological aspects of uncertainty analysis as applied to exposure assessment. Additional workshops are expected in the near future, dealing with such issues as the characterization of uncertainties that are inputs to a model.

2.2.4 Projects

There have been relatively few examples of in-house EPA projects with a quantitative approach to uncertainty analysis using Monte Carlo techniques. One early example was a simplified case study by Walentowicz (1984), the purpose of which was to demonstrate methodological aspects of an uncertainty analysis in an exposure assessment problem. A recent study by the Office of Solid Waste (EPA, 1992b) featured a Monte Carlo simulation based on expert judgments regarding uncertainties. This work was actually performed by a contractor. A contractor also performed a probabilistic assessment of contaminant transport as part of an exposure assessment (Dean et al., 1987). Others within EPA are reportedly considering the use of uncertainty analysis techniques in-house. An example of a possible application would be for a benefit/cost analysis of a set of regulatory options. The Uncertainty Circle uncovered a small number of other projects involving uncertainty analysis using Monte Carlo techniques. Overall, however, within EPA headquarters in Washington, DC there is relatively little activity in this regard.

3.0 APPROACHES TO UNCERTAINTY ANALYSIS

This chapter describes generalizable aspects of uncertainty analysis with respect to philosophy and approaches.

3.1 Philosophy of Uncertainty Analysis

The "classical" approach in probability theory requires that estimates for probability distributions must be based on empirical data. Certainly, in a regulatory environment, this is the approach preferred wherever possible. However, in many practical cases, the available data may not be relevant to the problem at hand, or there may be few data points to support a statistical analysis. Thus, statistical manipulation of data may be an insufficient basis for estimating uncertainty in a real system of interest. As a result, some degree of judgment about the available data may be required. Furthermore, even the application of statistical techniques, such as goodness-of-fit (GOF) tests requires considerable judgment. For example, the analyst makes judgments about what types of parametric distributions are appropriate for to represent uncertainty in a given empirical quantity, even though the analyst may rely on the data and the GOF test to determine the values of the distribution parameters (e.g., mean, variance).

An alternative approach differs in how probability distributions are interpreted. In the so-called "Bayesian" view, the assessment of the probability of an outcome is based on a "degree of belief" that the outcome will occur, based on all of the relevant information an analyst currently has about the system. Thus, the probability distribution may be based on empirical data and/or other considerations, such as technically informed judgments or predictions. People with different information or theoretical beliefs may estimate different distributions for the same variable (Morgan and Henrion, 1990). The assessment of uncertainties requires one to think about all possible outcomes and their likelihoods, not just the "most likely" outcome.

3.2 A Taxonomy of Uncertainty and Variability

There are a number of distinct sources of uncertainty in analyses of environmental problems. These come under the general headings of model or structural uncertainty and parameter uncertainty. Several authors, including Morgan and Henrion (1990), Finkel (1990), and others, provide more detail regarding sources of uncertainty. Sources of uncertainty are also discussed in some EPA documents (e.g., EPA, 1992). A few key concepts are summarized here.

3.2.1 Model Uncertainty

The structure of mathematical models employed to represent scenarios and phenomenon of interest is often a key source of uncertainty, due to the fact that models are often only a simplified representation of a real-world system, and that the problem boundary encompassed by a model may be incomplete or incorrect. Significant approximations are often an inherent part of the assumptions upon which a model is built. Competing models may be available based on different scientific or technical assumptions. Model uncertainty may be small for a well-tested physical model, or may be more significant than uncertainties in the values of model input parameters. Furthermore, the limited spatial or temporal resolution (e.g., grid size) of many models is also a type of approximation that introduces uncertainty into model results. Different sources of model uncertainties, and how they may be evaluated, are be summarized as follows:

- *Model Structure:* Alternative sets of scientific or technical assumptions may be available for developing a model. The implications of these alternative foundations may be evaluated by constructing alternative models and comparing results from each alternative model. In some cases, it may be possible to parameterize alternative model structures into a higher order model, and to evaluate alternative models using traditional sensitivity analysis. An example is a general formulation of a dose-response model to include or exclude a threshold and to have either linear or nonlinear behavior. When there are alternative underlying assumptions between competing models, it may not always be possible to determine *a priori* which is more "correct." A typical approach is to report the key assumptions underlying each model, and the corresponding results. If the results from competing models lead to similar decisions, then one can be confident that the decision is robust even in the face of alternative theories. If, however, alternative model formulations lead to different conclusions, the judgment of an analyst or a decision maker may be required to choose the most plausible inference options for a given problem.

- *Model Detail:* Often, models are simplified for purposes of tractability. For example, simplifying assumptions may be made to convert a complex nonlinear model to a simpler linear model in a parameter space of interest. Uncertainty in the predictions of simplified models can sometimes be gleaned by comparison of their predictions to those of more detailed, inclusive models. In other cases, simple models are developed due to a lack of confidence or knowledge about what the actual model structure should be. In these cases, the simplified model is a signal that little is actually known or quantifiable about the phenomenon being modeled. Uncertainty about these models may be only qualitatively understood.
- *Validation:* Models for which extensive data are available, and which have been validated for a parameter space of interest can be evaluated quantitatively in terms of the accuracy and precision of their predictions. Uncertainties regarding models for which few data are available to test model predictions may require evaluation using expert judgments or may not be amenable to any quantitative characterization.
- *Extrapolation:* A key source of uncertainty is extrapolation. Models which are validated for one portion of a parameter space may be completely inappropriate for making predictions in other regions of the parameter space. For example, a dose-response model based on high-dose, short duration animal tests may be completely inaccurate for low-dose, long duration human exposures.
- *Model Resolution:* In numerical models, a spatial and/or temporal grid size must be assumed. The selection of the grid size involves a trade-off between computation time (hence, cost) and prediction accuracy. Standard techniques are often available to help select the appropriate grid size for a particular target accuracy. This type of model uncertainty is thus dealt with through the appropriate selection of model domain parameter values, or by comparing results based on different grid sizes.
- *Model Boundaries:* Any model may have limited boundaries in terms of time, space, number of chemical species, temperature range, types of pathways, and so on. The selection of a model boundary may be a type of simplification. Within the boundary of the model and parameter space of the problem, the model may be an accurate representation of the real-world phenomenon of interest. However, other overlooked phenomenon not included in the model may play a role in the scenario being modeled. The way to avoid this type of uncertainty is to think creatively about all of the factors that may come to bear within the context of a particular scenario.
- *Scenario Reasonableness:* Prior to using a model, an analyst must develop (explicitly or implicitly) a scenario for the problem of interest. A scenario is a set of assumptions about the nature of the problem to be analyzed. For example, in many environmental problems, assumptions are made about the source of pollutant emissions, the pathways of pollutant transport and deposition, the populations exposed to the pollutant, and the types of effects resulting from the exposure. Scenarios may be constructed to represent an actual environmental problem, or they may be constructed hypothetically based on policy motivations. In the latter case, for example, a scenario may focus on a hypothetical "porch potato" — an individual who spends their entire lifetime at the

point of maximum concentration of some pollutant. To the extent that the scenario fails to consider all factors affecting the key output variable (e.g., lifetime average daily dose), uncertainty will be introduced. Like the uncertainty associated with model boundaries, the uncertainty associated with the scenario can be addressed by imaginative thinking about all possible factors that come to bear in either the real-life or hypothetical problem.

3.2.2 Parameter Uncertainty

Morgan and Henrion (1990) have identified a number of different types of quantities used in models. These include:

- *Empirical:* measurable, at least in principle (e.g., pollutant concentration).
- *Defined Constants:* Some quantities whose values are accepted by convention, such as Planck's constant or the speed of light, are actually empirical quantities subject to measurement error, albeit small. Other quantities are defined by convention and are not uncertain. These include, for example, the mathematical constant π (pi).
- *Decision variables:* These are parameters over which a decision-maker exercises control, such as the maximum acceptable emission rate for a given emission source. A decision maker selects this value. Thus, it is not appropriate to treat this quantity probabilistically. Rather, the sensitivity of the result to different values of the decision variable(s) should be explored using sensitivity analysis.
- *Value parameters:* Represents the preferences or value judgments of a decision maker. Examples include the discount rate and parameters of utility functions used in decision analysis.
- *Model domain parameters:* these are parameters that are associated with a model, but not directly with the phenomenon the model represents. For example, the spatial or temporal grid size is a model domain parameter introduced in numerical models.

Of the types of quantities identified above, only empirical quantities are unambiguously subject to uncertainty. The other types of parameters represent quantities which are almost always more properly treated as point-estimates reflecting convention, the explicit preferences of a decision maker (broadly defined), or a descrete quantity by its nature (e.g., grid size). Thus, we focus here on identifying sources of uncertainty in empirical quantities. These include:

- *Random Error and Statistical Variation:* This type of uncertainty is associated with imperfections in measurement techniques. Statistical analysis of test data is thus one method for developing a representation of uncertainty in a variable.
- *Systematic Error:* The mean value of a measured quantity may not converge to the "true" mean value because of biases in measurements and procedures. Such biases may arise from imprecise calibration, faulty reading of meters, and inaccuracies in the assumptions used to infer the actual quantity of interest from the observed readings of other quantities ("surrogate" or "proxy" variables). Estimation of the possible magnitude of systematic error may be based on measurements of "spiked" or known samples, or on the judgment of an expert. For example, there is often systematic error involved in using small scale tests to estimate the values of quantities for large scale systems.
- *Variability:* Some quantities are variable over time, space, or some population of individuals (broadly defined) rather than for any individual event or component. Variability is modeled using a frequency, rather than a probability, distribution.

- *Inherent Randomness or Unpredictability:* Some quantities may be irreducibly random even in principle, the most obvious example being Heisenberg's Uncertainty Principle. However, this concept is often applied to quantities that are in principle measurable precisely but as a practical matter (due to cost, for example) are not.
- *Lack of Empirical Basis:* Lack of experience about or knowledge of a process or system is a source of uncertainty. This type of uncertainty cannot be treated statistically, because it requires predictions about something that has yet to be built, tested, or measured. This type of uncertainty can be represented using technical-based judgments about the range and likelihood of possible outcomes. These judgments may be based on a theoretical foundation or experience with analogous systems.
- *Dependence and Correlation:* When there is more than one uncertain quantity, it may be possible that the uncertainties may be statistically or functionally dependent. In such cases, failure to properly model the dependence between the quantities can lead to uncertainty in the result, in terms of improper prediction of the variance of output variables. Dependence among model input variables often arises because of model simplifications which fail to explicitly model the source of dependence between them. There are several ways to deal with dependence, which are discussed in more detail later. However, in general, it is recommended that the source of dependence be modeled explicitly wherever possible.
- *Disagreement:* Where there are limited data or alternative theoretical bases for modeling a system, experts may disagree on the interpretation of data or on their estimates regarding the range and likelihood of outcomes for empirical quantities. Disagreement is especially likely when there is a lack of empirical basis for estimating uncertainties or variabilities, and/or when characterizations of uncertainty are based on subjective probability distributions elicited from expert judgments. In cases of expert disagreement, it is usually best to explore separately the implications of the judgments of different experts to determine whether substantially different conclusions about the problem result. If not, then the results of the analysis are robust to the disagreements among the experts. If so, then one has to more carefully evaluated the sources of disagreement between the experts. In this situation, an analyst or decision maker may have to make a judgment regarding which experts are more plausible for the problem at hand.

3.2.3 Variability and Uncertainty

In many environmental problems, the distinction between uncertainty and variability is critically important. As noted above, variability is a heterogeneity between individual members of a population of some type, and is characterized using a frequency distribution. In principle, the characteristics of a specific individual in a population are knowable with certainty in most cases. Thus, the frequency distribution for the population reflects true differences between individuals. Knowing the frequency distribution for variability in the population can aid in determining whether the population should be disaggregated into smaller groups that are more nearly homogeneous. This type of information is important, for example, in identifying subgroups especially susceptible to specific health risks from exposures to a given chemical.

However, there may be uncertainty in the characteristics of specific individuals in the population, due to measurement error or other sources of uncertainty as described above. In these cases, there is a resulting uncertainty about the variability frequency distribution. For example, while individuals may be known to have different exposure levels levels to a certain pollutant, their health effects may be uncertain due to the limited applicability of dose-response models extrapolated from animal bioassay test results. Thus, the population distribution for health effects (e.g., excess cancers) may be both variable and uncertain.

To complicate matters further, however, it is possible for variability to be interpreted as uncertainty under certain conditions. For example, suppose we are interested in the exposure level faced by an individual selected at random from a population. If we select an individual at random, the probability of selecting an individual with a given exposure is the same as the relative frequency of all individuals in the population subject to the given exposure. Hence, in this case variability represents an *a priori* probability distribution for the exposure faced by a randomly selected individual. However, except for this special case, there is always a distinction between variability and uncertainty.

The implications of distributions for variable and uncertain quantities in terms of improving estimates are different. As indicated previously, knowing the frequency distribution for variability can guide the identification of significant subpopulations which merit more focused study. In contrast, knowing the uncertainty in the measurement of characteristics of interest for the population can aid in determining whether additional research or alternative measurement techniques are needed to reduce uncertainty. Thus, in order to sort out the differences between uncertainty and variability, it is desirable to separately characterize them.

3.3 Dependence and Correlation

One of the questions that often arises in the early stages of developing a simulation of uncertainty and variability is whether there is dependence between input variables to a model. In many cases, there are, due to the use of simplified models. In a complete model, the sources of dependence between variables would be explicitly modeled. In a simplified model, some quantities that may be more properly modeled as state variables are treated as if they are input (exogenous) variables. Thus, the analyst is faced with the task of approximating the dependence between input variables. For example, in a simplified process model, temperature and chemical reaction conversion rate may both be treated as input variables, whereas in reality the chemical reaction conversion rate is a function of temperature. The simplified model may be a convenience, or a necessity if insufficient information is available to support the development of a more detailed model. However, it is known that qualitatively the reaction conversion rate will tend to increase with temperature.

There are several approaches to dealing with dependence. These are:

- Model dependence explicitly
- Parameterization
- Stratification
- Bootstrap simulation
- Simulate correlations

Each of these approaches is discussed.

Modeling dependence explicitly involves the development of a more detailed modeled which captures the source of dependence between two quantities. Thus, in the previous example, a chemical kinetic reaction model in which reaction rate is a function of temperature would capture the dependence between temperature and conversion rate. Such a model formulation would also change conversion rate from an input variable to a model state variable.

Parameterization refers to grouping the input variables and treating the grouping as a new input variable. This treatment of dependence is useful when there is linear dependence between the variables. For example, air inhalation rate and body weight may both be variables in an exposure model. The inhalation rate is at least partly dependent on body weight. Thus, a new parameter, inhalation rate divided by body weight, may be used to capture this dependence.

Rather than try to model dependence, another approach attempts to reduce the effects of dependence. This approach, stratification, involves subdividing the problem by creating several subgroups or strata. For example, if the problem features a population of individuals with widely varying body weights and inhalation rates, the problem could be subdivided by body weight. Thus, there would be less variance in inhalation rate within the body weight subgroups. As a result, the effect of correlation or dependence within the strata would be smaller than for the population as a whole.

In the rare instances in which paired data are available for a set of model input variables, some or all of which are dependent, the paired data can be used directly employing a technique known as bootstrap simulation. In the bootstrap approach, the model is run for each pairing of data points. Suppose we have two input variables, body weight and inhalation rate, and 86 paired samples (representing different individuals) for these two variables. The model would be run 86 times with each set of data. Thus, any correlation between the data will be simulated using this approach. This approach can be extended to more than two variables if data are available.

An alternative to bootstrapping, which restricts the simulation sample size to the data sample size, is to use available paired data to estimate the correlations or covariance matrix among dependent quantities, and then to use specialized techniques to synthetically generated correlated variables as part of a Monte Carlo simulation. For example, there are easily applied techniques for inducing correlations between normal distributions in Monte Carlo simulations (Morgan and Henrion, 1990). There are more generalized techniques for inducing rank correlations among multiple variables (Iman et al, 1985).

In some cases, correlations may be known to exist between model input variables, but there may not be paired data sets available to support a bootstrap simulation or to use to estimate the correlations. There are two alternative approaches. One is to try to estimate the actual correlations through some type of expert elicitation process. However, it is generally more difficult to estimate correlation coefficients or covariance matrices than it is to make judgments about probability or frequency distributions (Morgan and Henrion, 1990). The other approach is to employ a generalized rank ordering pairing technique to explore the sensitivity of modeling results to alternative assumptions about correlations. In such an approach, high or low correlations may be compared to see what effect, if any, there is on model results.

An example of a study in which the sensitivity of results to correlations was evaluated is given by Frey (1991). A restricted pairing technique by Iman and Conover (1982) is included in a Fortran package developed by Iman and Shortencarier (1984) for simulating uncertainties using a variant of Monte Carlo simulation called Latin Hypercube sampling (described in a later section of this report). This technique was employed to evaluate alternative hypotheses regarding correlation structures in uncertainties for a study of coal gasification power plants. A simulation of correlations is illustrated in Figure 1. The figure shows uncorrelated and correlated simulations for two triangular distributions. For most cases evaluated, correlations had little effect on the resulting uncertainties in plant efficiency, emissions, and cost.

Correlations among input variables to a model may have little effect on modeling results in several cases. If all of the correlated random variables do not contribute significantly to uncertainty in key model outputs of interest, then correlations among them will tend to have little effect on model results. When only one of two or more dependent input variables contributes significantly to uncertainty in a model result, correlations will also have little effect on the result.

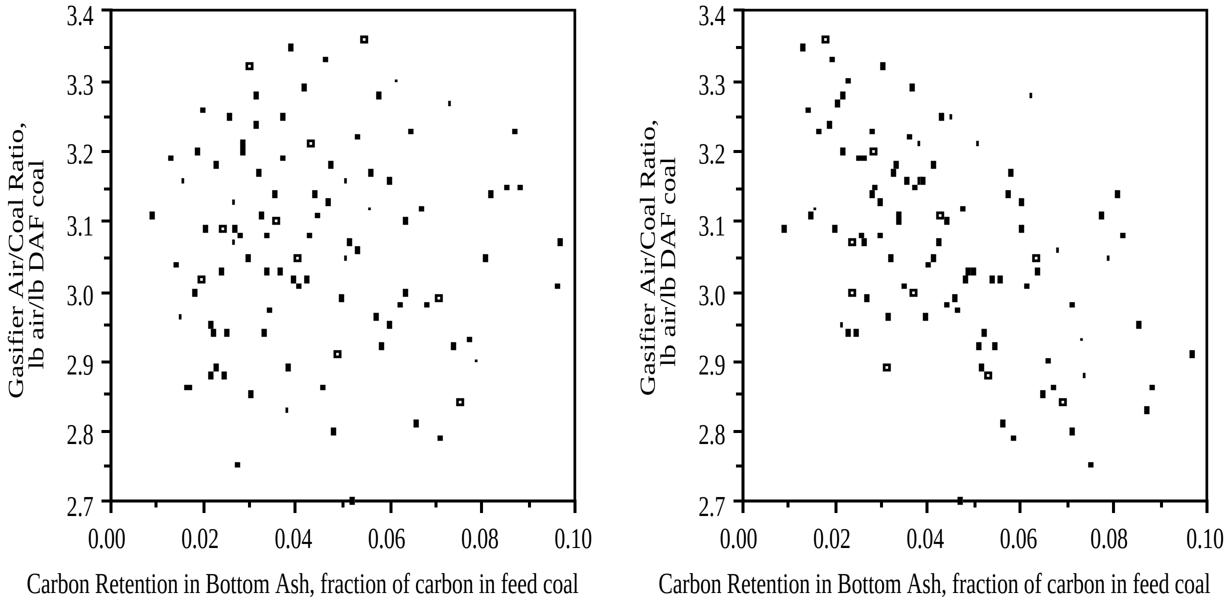


Figure 1. Simulation of Correlations between model input parameters for correlation coefficient $\rho = 0$ and -0.5 for two triangularly distributed quantities.

3.4 Encoding Uncertainties as Probability Distributions

There are two fundamental approaches for encoding uncertainty in terms of probability distributions. These include statistical estimation techniques and expert judgments. A combination of both methods may be appropriate in many practical situations. For example, a statistical analysis of measured test data may be a starting point for thinking about uncertainties in a hypothetical commercial scale system for a new process technology. One must then consider the effect that systematic errors, variability, or uncertainties about scaling-up the process might have on interpreting test results for commercial scale design applications (Frey and Rubin, 1992).

3.4.1 Statistical Techniques

Statistical estimation techniques involve estimating probability distributions from available data. The fit of data to a particular probability distribution function can be evaluated using various statistical tests. For example, the cumulative probability distribution of a set of data may be plotted on "probability" paper. If the data plot as a straight line, then the distribution is normal. Procedures for fitting probability distribution functions are discussed in many standard texts on probability and are not reviewed here. Rather, the focus of this discussion is on the situations where statistical analysis alone may be insufficient, because technical insights may be required to interpret whatever limited data are available.

3.4.2 Judgments about Uncertainties

In making judgments about a probability distribution for a quantity, there are a number of approaches (heuristics) that people use which psychologists have observed. Some of these can lead to biases in the probability estimate. Three of the most common are briefly summarized.[1]

---

[1] The discussion here is based on Morgan and Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, 1990.

1) *Availability*. The probability that experts assign to a particular possible outcome may be linked to the ease (availability) with which they can recall past instances of the outcome. For example, if tests have yielded high sorbent durability, it may be easier to imagine obtaining a high sorbent durability in the future than obtaining lower durabilities. Thus, one tends to expect experts to be biased toward outcomes they have recently observed or can easily imagine, as opposed to other possible outcomes that have not been observed in tests.

2) *Representativeness* has also been termed the "law of small numbers." People may tend to assume that the behavior they observe in a small set of data must be representative of the behavior of the system, which may not be completely characterized until substantially more data are collected. Thus, one should be cautious in inferring patterns from data with a small number of samples.

3) *Anchoring and adjustment* involves using a natural starting point as the basis for making adjustments. For example, an expert might choose to start with a "best guess" value, which represents perhaps a median or most likely (modal) value, and then make adjustments to the best guess to achieve "worst" and "best" outcomes as bounds. The "worst" and "best" outcomes may be intended to represent a 90 percent probability range for the variable. However, the adjustment from the central "best guess" value to the extreme values is often insufficient, with the result that the probability distribution is too tight and biased toward the central value. This phenomena is *overconfidence*, because the expert's judgment reflects less uncertainty in the variable than it should. The "anchor" can be any value, not just a central value. For example, if an expert begins with a "worst" case value, the entire distribution may be biased toward that value.

Judgments also may be biased for other reasons. One common concern is *motivational bias*. This bias may occur for reasons such as: a) a person may want to influence a decision to go a certain way; b) the person may perceive that they will be evaluated based on the outcome and might tend to be conservative in their estimates; c) the person may want to suppress uncertainty that they actually believe is present in order to appear knowledgeable or authoritative; and d) the expert has taken a strong stand in the past and does not want to appear to contradict themself by producing a distribution that lends credence to alternative views.

3.4.3 Designing an Elicitation Protocol

From studies of how well calibrated judgments about uncertainty are, it appears that the most frequent problem encountered is overconfidence (Morgan and Henrion, 1990). Knowledge about how most people make judgments about probability distributions can be used to design a procedure for eliciting these judgments. The appropriate procedure depends on the background of the expert and the quantity for which the judgment is being elicited. For example, if an expert has some prior knowledge about the shape of the distribution for the quantity, then it may be appropriate to ask him/her to think about extreme values of the distribution and then to draw the distribution. On the other hand, if a technical expert has little statistical background, it may be more appropriate to ask him/her a series of questions. For example, the expert might be asked the probability of obtaining a value less than or equal to some value x, and then the question would repeated for a few other values of x. The judgment can then be graphed by an elicitor, who would review the results of the elicitation with the expert to see if he/she is comfortable with the answers.

To overcome the typical problem of overconfidence, it is usual to begin by thinking about extreme high or low values before asking about central values of the distribution. In general, experts' judgments about uncertainties tend to improve when: (1) the expert is forced to consider how things could turn out differently than expected (e.g., high and low extremes); and (2) the expert is asked to list reasons for obtaining various outcomes. Keeney and von Winterfeldt (1991) provide a discussion of approaches they have employed in expert elicitations. Otway and von Winterfeldt (1992) review some case studies involving expert elicitations and discuss some advantages and shortcomings to formal approaches to expert elicitation.

While the development of expert judgments may be flawed in some respects, it does permit a more robust analysis of uncertainties in a process when limited data are available (Wolff et al, 1990). Furthermore, in many ways, the assessment of probability distributions is qualitatively no different than selecting single "best guess" values for use in a deterministic estimate. For example, a "best guess" value often represents a judgment about the single most likely value that one expects to obtain. The "best guess" value may be selected after considering several possible values. The types of heuristics and biases discussed above may play a similar role in selecting the value. Thus, even when only a single "best guess" number is used in an analysis, a seasoned engineer usually has at least a "sense" for "how good that number really is." This may be why engineers are often able to make judgments about uncertainties easily, because they implicitly make these types of judgments routinely.

3.5 Some Types of Probability Distributions

Examples of several types of probability distributions are shown in Figure 2 as both probability density functions (pdf's) and cumulative distribution functions (cdf's). The pdf is a graphical means of representing the relative likelihood or frequency with which values of a variable may be obtained. The pdf also clearly illustrates whether a probability distribution is symmetric or skewed. In a symmetric unimodal distribution, the mean (average), median (50th percentile), and mode (peak) coincide. In a positively skewed distribution (e.g., lognormal), the mean is greater than the median, and both are greater than the mode.

An alternative way to represent a probability distribution is the cdf. The cdf shows probability fractiles on the y-axis and the value of the distribution associated with each fractile on the x-axis. The cdf is a way to represent any probability distribution when there is information about various fractiles of the distribution (e.g., the values of the 5th, 50th and 95th percentiles).

A brief description of several types of probability distributions and their applications is given here:

- <u>Uniform:</u> Uniform probability of obtaining a value between upper and lower limits. Useful when an expert is willing to specify a finite range of possible values, but is unable to decide which values in the range are more likely to occur than others. The use of the uniform distribution is also a signal that the details about uncertainty in the variable are not known. Useful for screening studies.

- <u>Triangle</u>: Similar to uniform except a mode is also specified. Use when an expert is willing to specify both a finite range of possible values and a "most likely" (mode) value. The triangle distribution may be symmetric or skewed (as in Figure 2). Like the uniform, this distribution indicates that additional details about uncertainty are not yet known. The triangle distribution is excellent for screening studies and easy to obtain judgments for.

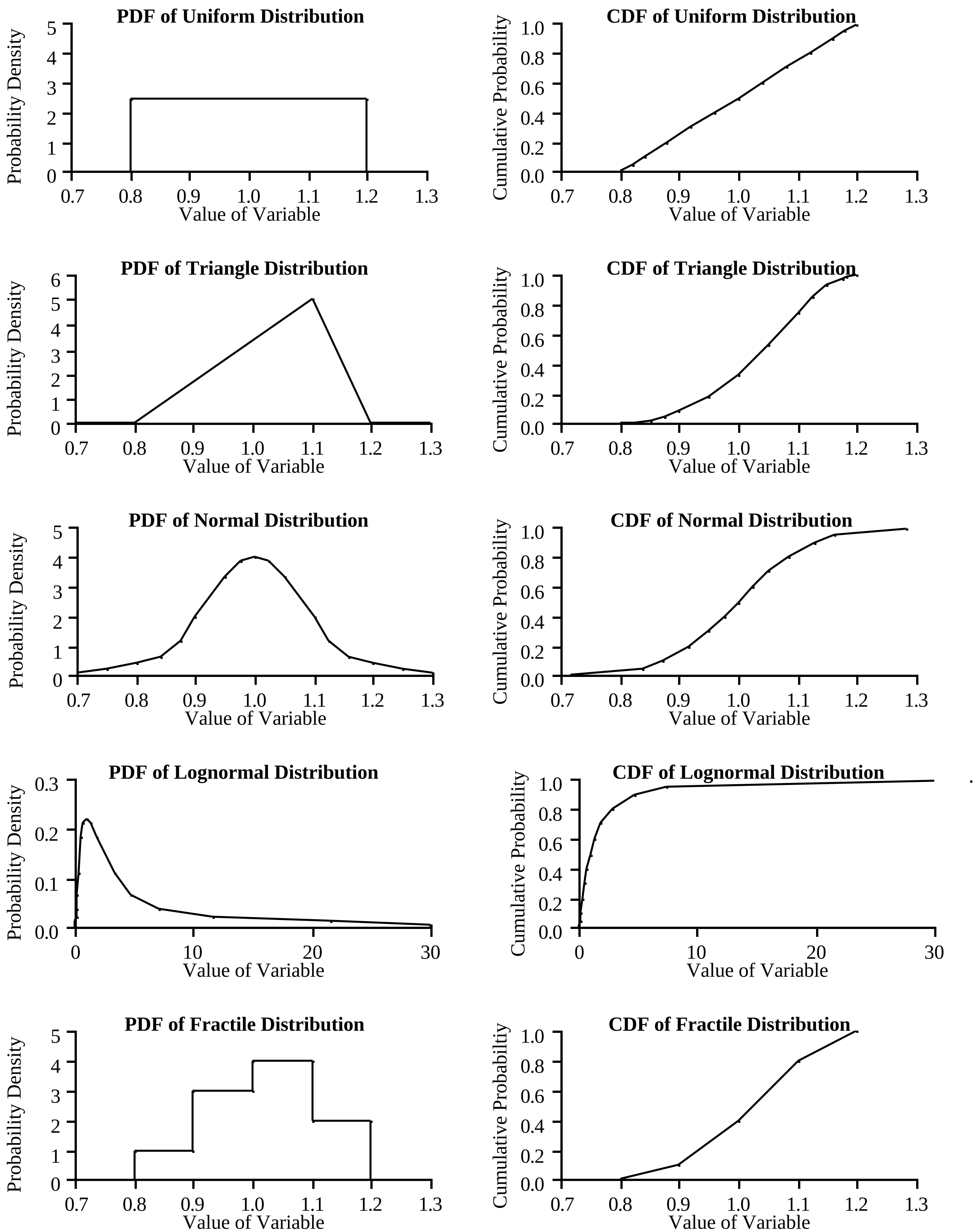


Figure 2. Some Types of Probability Distributions.

- Normal: A symmetric distribution with mean, mode, and median at the same point. Often assumed in statistical analysis as the basis for unbiased measurement errors. The normal distribution has infinite tails; however, over 99 percent of all values of the normal distribution lie within plus or minus three standard deviations of the mean. Thus, when used to represent uncertainty in physical quantities which much be greater than zero, the standard deviation should not be more than about 20 or 30 percent of the mean, or else the distribution must be truncated.

- Lognormal: A positively skewed distribution (it has a long tail to the right). This distribution is usually used to represent uncertainty in physical quantities which must be non-negative and are positively skewed, such as the size of an oil spill or the concentration of a pollutant. This distribution may be used when uncertainties are expressed on a multiplicative order-of-magnitude basis (e.g., factor of 2) or when there is a probability of obtaining extreme large values.

- Loguniform: A uniform distribution in log space (each decade has equal probability, not shown in Figure 2).

- Fractile: The finite range of possible values is divided into subintervals. Within each subinterval, the values are sampled uniformly according to a specified frequency for each subinterval. This distribution looks like a histogram and can be used to represent any arbitrary data or judgment about uncertainties in a parameter, when the parameter is continuous. Explicitly shows detail of the judgments about uncertainties.

- Chance: This is like the fractile distribution, except that it applies to discrete, rather than continuous, variables. An example of a discrete variable is the number of trains of equipment, which must be an integer (e.g., 30% chance of one train, 70% chance of two).

3.6 Probabilistic Modeling

In order to analyze uncertainties in environmental, a probabilistic modeling environment is required. A typical approach is the use of Monte Carlo simulation, as described by Ang and Tang (1984) and others. In Monte Carlo simulation, a model is run repeatedly, using different values for each of the uncertain input parameters each time. The values of each of the uncertain input parameters are generated based on the probability distribution for the parameter. If there are two or more uncertain input parameters, one value from each is sampled simultaneously in each repetition in the simulation. Over the course of a simulation, perhaps 20, 50, 100, or even more repetitions may be made. The result, then, is a set of sample values for each of the model output variables, which can be treated statistically as if they were an experimentally or empirical observed set of data.

Although the generation of sample values for model input parameters is probabilistic, the execution of the model for a given set of samples in a repetition is deterministic. The advantage of Monte Carlo methods, however, is that these deterministic simulations are repeated in a manner that yields important insights into the sensitivity of the model to variations in the input parameters, as well as into the likelihood of obtaining any particular outcome.
Monte Carlo methods also allow the modeler to use any type of probability distribution for which values can be generated on a computer, rather than to be restricted to forms which are analytically tractable.

3.6.1 Monte Carlo simulation

In random Monte Carlo simulation, a random number generator is used to generate uniformly distributed numbers between 0 and 1 for each uncertain variable. Note from Figure 2

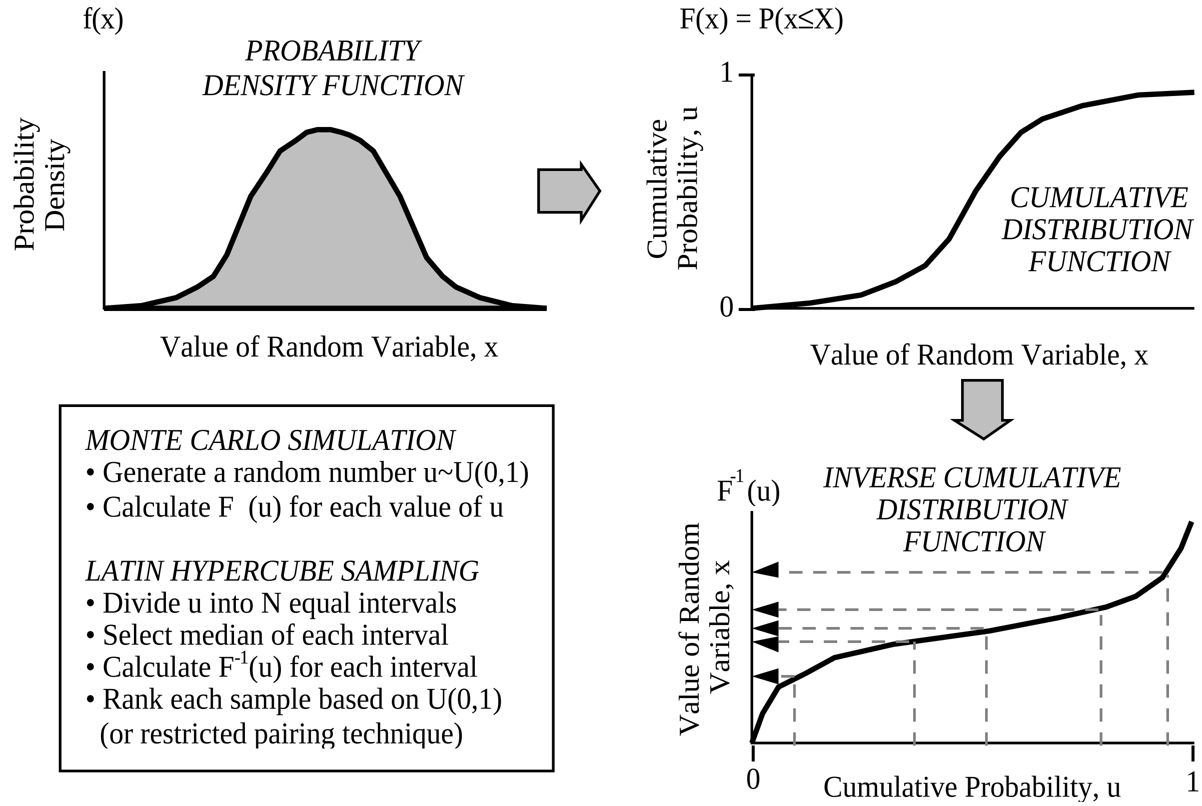


Figure 3. Monte Carlo Simulation.

that all cdf's have an ordinate axis ranging from zero to one. Thus, uniformly distributed random numbers are used to represent the fractile of the random variable for which a sample is to be generated. The sample values for the random variables are calculated using the inverse cdf functions based on the randomly generated fractiles. This approach is shown schematically in Figure 3.

Using Monte Carlo techniques, it is therefore possible to represent uncertainty in a model of a process technology by generating sample values for uncertain variables, and running the model repetitively. Instead of obtaining a single number for model outputs as in deterministic simulation, a set of samples is obtained. These can be represented as cdf's and summarized using typical statistics such as mean and variance.

3.6.2 Latin Hypercube Sampling

An alternative to random Monte Carlo simulation is Latin Hypercube Sampling (LHS). In LHS methods, the fractiles that are used as inputs to the inverse cdf are not randomly generated. Instead, the probability distribution for the random variable of interest is first divided into ranges of equal probability, and one sample is taken from each equal probability range. However, the ranking (order) of the samples is random over the course of the simulation, and the pairing of samples between two or more random input variables is usually treated as independent. In median LHS, one sample is taken from the median of each equal-probability interval, while in random LHS one sample is taken from random within each interval (Morgan and Henrion, 1990).

LHS methods guarantee that values from the entire range of the distribution will be sampled proportional to the probability density of the distribution. Because the distributions are sampled over the entire range of probable values in LHS, the number of samples required to adequately represent a distribution is less for LHS than for random Monte Carlo sampling. LHS is generally preferred over random Monte Carlo simulation (McKay, Conover, and Beckman, 1979; Iman and Helton, 1988; Morgan and Henrion, 1990). LHS is the technique employed later in this study. As noted earlier, restricted pairing techniques are available for the purpose of inducing correlations between variables in LHS (Iman et al, 1984).

3.6.3 Selecting Sample Size

The sample size corresponds to the number of repetitions used in the probabilistic simulation. The selection of sample size is usually constrained at the upper end by the limitations of computer software, hardware, and time, and at the lower end by the acceptable confidence interval for model results. In cases where the analyst is most interested in the central tendency of distributions for output variables, the sample size can often be relatively small. However, in cases were the analyst is interested in low probability outcomes at the tails of output variable distributions, large sample sizes may be needed. As sample size is increased, computer runtime, memory use, and disk use may become excessive. Therefore, it may be important to use no more samples than are actually needed for a particular application.

One approach to selecting sample size is to decide on an acceptable confidence interval for whatever fractile level is of most concern in the investigation (Morgan and Henrion, 1990). For example, we may wish to obtain a given confidence that the value of the p^{th} fractile will be bounded by the i^{th} and k^{th} fractiles. In a Monte Carlo simulation, we can use the following relations to estimate the required sample size:

$$i = mp - c\sqrt{mp(1 - p)} \tag{1}$$

$$k = mp + c\sqrt{mp(1 - p)} \tag{2}$$

The relations in Equations (1) and (2) yield a confidence interval for the p^{th} fractile if the sample size is known, where c is the standard deviation of the standard normal distribution associated with the confidence level of interest. To calculate the number of samples required, the expressions above can be rearranged to calculate the confidence interval $(Y_{p-\Delta p}, Y_{p+\Delta p})$ as follows:

$$m = p(1 - p)\left(\frac{c}{\Delta p}\right)^2 \tag{3}$$

For example, if we wish to be 90 percent confident that the value of the 90th percentile will be enclosed by the values of the 85th and 95th fractiles, then c would be 1.65 and m would be 98. However, another factor to consider in selecting sample size is whether a high degree of simulation accuracy is really needed. In screening studies based on a first-pass set of expert judgments, it may be unnecessary to obtain a high degree of confidence in specific fractiles of the output distribution, because initial estimates of uncertainty may be subject to considerable empirical uncertainty themselves.

The approach to selecting sample size described above is appropriate for use with the Monte Carlo simulation technique. In this work, LHS is employed as discussed previously. The approach to estimating the precision of modeling results based on confidence intervals will typically overestimate the required sample size needed with LHS.

3.6.4 Analyzing Results

Sample correlation coefficients are a simple but useful tool for identifying the linear correlations between uncertain variables. Other techniques are available in software packages such as one developed by Iman, Shortencarier, and Johnson (1985). These output analysis techniques are described here briefly.

A partial correlation coefficient (PCC) analysis is used to identify the degree to which correlations between output and input random variables may be linear, and it is estimated in conjunction with multi-variate linear regression analysis. In PCC analysis, the input variable most highly correlated the output variable of interest is assumed as the starting pointing for construction of a linear regression model. In the regression model, the output variable is treated as the dependent variable and the most highly correlated input variable is treated as a predictive variable. The partial correlation technique then searches for another input variable which is most highly correlated with the *residuals* of the regression model already containing the first input variable. The residual is the difference between the actual sample value of the dependent variable and the estimated sample values, using the linear regression model containing the first input variable. The process is repeated to add more variables in the analysis. The partial correlation coefficient is a measure of the unique linear relationship between the input and dependent variables that cannot be explained by variables already included in the regression model.

Standardized regression coefficients (SRC) can be used to measure the relative contribution of the uncertainty in the input variables on the uncertainty of the output variables. This analysis involves standardization of all the sample values for the model input variables and a multi-variate regression of an output variate based on the inputs. The regression coefficients for each input variate then indicate the relative importance of that variate as a factor determining the output. SRCs measure the shared contribution of the input to the output, because all of the simulation input uncertainties are included in the regression analysis simultaneously. The SRCs are the partial derivatives of the output variable with respect to each input variable. Because PCCs are a measure of the unique contribution of each parameter, and SRCs measure the shared contribution, they do not always lead to the same conclusions.

PCC and SRC analysis is limited to cases where the relationship between input and output variables is linear; however, by basing the regression analysis on the ranks of the samples for each variable, rather than on the values of the samples, the PCC and SRC techniques can be extended to non-linear cases. These techniques are known as partial rank correlation coefficients (PRCC) and standardized rank regression coefficients (SRRC) (Iman, Shortencarier, and Johnson, 1985).

While regression analysis of input and output sample vectors is an important tool for prioritizing input uncertainties that are most "sensitive," it is important to understand the limitations of partial correlation coefficients when using a given sample size. Edwards (1984) provides a clear discussion of tests of significance for correlation coefficients. When using partial correlation coefficients for output analysis, we are interested in testing the null hypothesis that the coefficient is equal to zero. For independent random variables, the t-test can be used and the value of t is calculated as follows:

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{m - n} \tag{4}$$

The degrees of freedom m-n is given by the number of samples m and the number of input variables n used in the regression analysis. The t statistic calculated in Equation (4) can then be compared to values in a table of the t-distribution for a given significance level and degrees of freedom. If the statistic calculated above is greater than the value from the table, the null hypothesis is regarded as sufficiently improbable that it can be rejected. As an example, for 100 samples, 50 independent variables used in a regression analysis, and a significance level of 0.01 for a one-sided test, an obtained value for r of greater than 0.322 or less than -0.322 would be grounds for rejection of the null hypothesis. Treatment of partial rank correlation coefficients is similar.

4.0 MONTE CARLO SIMULATION IN EXPOSURE ASSESSMENT

A specific problem domain in which there is increasing interest in and use of Monte Carlo techniques is exposure assessment. Exposure assessment is one step in the development of a risk assessment, as previously described. In this section, exposure assessment is defined in more detail. Methodological aspects of the use of Monte Carlo simulation in exposure assessment are described, followed by a brief discussion of some example applications.

4.1 Exposure Assessment

Per the EPA 1992 Exposure Assessment Guidelines (EPA, 1992), exposure is defined as contact of a chemical to the outer boundary of the body, which is the skin and openings into the body such as the mouth, nostrils, punctures, lesions. An exposure assessment is typically a quantitative evaluation of the intensity, frequency, and duration of the contact. As part of an exposure assessment, the rate at which the chemical crosses the boundary, the route or pathways (e.g., oral, dermal, respiratory), and the actual amount of chemical contacting and crossing the boundary are evaluated. In principle, exposure is a function of time. However, depending on whether the health effect to be evaluated is chronic or acute, attention may be focused on estimating average exposures or peak exposures, respectively. For exposure assessment of carcinogenic chemicals, an average exposure is often used. For developmental toxins, the effect of exposure may differ depending on the stage of development of the exposed individual. In some cases, exposure to very high contaminant levels may produce significant health effects even if the time-averaged exposure is low. In such cases, an exposure profile as a function of time is required.

While there are many different possible formulations for exposure models, based for example on the exposure route and the type of health effect resulting from the contact, one typical model will be considered here. For effects such as cancer, the biological response is usually described in terms of lifetime probabilities of developing cancer. Thus, although the exposure may occur only for a portion of a lifetime, the exposure (or potential dose) is often calculated based on a lifetime average. This approach is used in conjunction with risk assessments based on linear nonthreshold dose-response models. This type of exposure model is thus (for a description of nomenclature, please see Section 8):

$$\mathrm{LADD} = \frac{\overline{\mathrm{C}} \bullet \overline{\mathrm{IR}} \bullet \mathrm{ED}}{\mathrm{BW} \bullet \mathrm{LT}} \tag{5}$$

Here, the lifetime average daily dose is a function of average concentration, average ingestion rate (for an oral exposure route) or intake rate (for a respiratory route), exposure duration, body weight, and lifetime. Other exposure models are discussed in more detail in the Exposure Assessment Guidelines.

Information about exposure is often used with exposure-response relationships to estimate the probability of an adverse health effect occurring. In most cases, exposure is used as a basis for estimating the actual dose delivered to biological receptors in the body. In these cases, a dose-response model is employed to estimate health risks. In general, two types of risks are of interest to policy makers: individual risk and population risk. Individual risks are often calculated for one or more individuals in a population of exposed persons. Some key policy questions include: What are the characteristics of persons facing the highest exposure and/or risk levels? Can the people most highly susceptible to adverse health effects be identified? What is the average exposure and/or risk? Thus, assessors are often interested in characterizing exposures and/or risks for persons at the high end of the population distribution, defined by EPA as the 90th percentile or higher, or persons at the central part of the population distribution (e.g., the mean or median). Groups with different degrees of susceptibility can be treated as separate subpopulations, and it may be appropriate to employ different data and/or models for each group.

Population exposure levels and risks are used to answer questions regarding how many cases of a particular health effect may occur in a given population over a specified time period, how many persons are above a certain risk or exposure level, and how various subgroups compare to the overall population.

Ideally, there should be interaction between an exposure assessor and a decision-maker regarding the type of questions an assessment should address. Initially, a decision-maker may be interested in one type of "risk descriptor," such as risk to an average individual. However, in the course of performing an analysis an assessor may uncover important information regarding highly exposed individuals, or regarding the number of individuals subject to exposures at or above some reference level. Furthermore, based on the guidance in the Habicht memo, assessors at EPA are required to present more than one description of exposure and risk, to provide explicit indication of both central tendency and high end exposures.

There are a large number of factors to consider in developing an exposure assessment, such as purpose, scope, scenarios, level of detail, descriptors to use, data needs, data quality objectives, sampling plans, quality assurance, modeling strategy, model validation, approach to dealing with uncertainty, and presentation of results, to mention some of the key ones. The details regarding all of these factors are well beyond the scope of this report, but are described in the Exposure Assessment Guidelines. The remainder of this section will focus on methodological and practical aspects of the quantitative analysis of uncertainty and variability in exposure assessment.

4.2 Uncertainty Analysis in Exposure Assessment

Whitmore (1985) reviews a number of approaches for characterizing uncertainty in exposure assessments. These approaches focus primarily on procedures for estimation of distributions for uncertainty depending on the type of information available. The types of data available to support an exposure assessment may include: (1) measured exposure levels for a sample of population members; (2) measurements for a sample of population members of parameters that are inputs to an exposure assessment model; (3) estimated joint distributions for model input variables; (4) "limited" data for model input variables; and (5) data on concentration levels of a substance at fixed sites and regarding the geographic distribution of population members. Depending on the data available, the approaches to characterizing uncertainty include four basic categories: (1) confidence interval estimates for percentiles of the population exposure distribution; (2) estimation of measurement accuracy and precision; (3) use of goodness-of-fit techniques to model uncertainties with parametric distributions; and (4) evaluation of alternative parametric probability distribution models. The type of data available often depends on the resources devoted to the assessment, with more detailed measurements or surveys used to support more detailed assessments.

Whitmore discusses the role of expert judgments in formulating estimates for the joint distributions of all model input variables. With regard to correlations between variables, Whitmore points out that if the input variables are not highly correlated, or if the model is not sensitive to assumptions regarding correlation or independence, then it may be appropriate to treat the variables as independent. Thus, separate experts can be approached to estimate probability distributions for each variable. The type of information required of the expert might include the shape of the distribution, some indication of the lowest and highest values, the mean, median, and/or mode. Often, only a few of these factors are necessary to uniquely define a given distribution (assuming, for example, some type of parametric distribution). In other cases, the expert may be able to provide a histogram representing his/her judgment. More than one expert should be approached, and any differences among them should be evaluated through separate model runs, or should be reconciled through some type of feedback technique (e.g., the Delphi technique).

Whitmore addresses a critical issue of dealing with both variability and uncertainty in an exposure model. A suggested approach is the use of estimation techniques to estimate confidence intervals for the parameters of parametric probability distribution models. This approach is only possible, of course, when there are data (real or synthetic) to use in the estimation methods. For example, the confidence interval on the estimates of mean and standard deviation for a normal distribution would be used as a basis for estimating confidence intervals for any percentile of the normal distribution. Similarly, for a model consisting of multiple input quantities, some or all of which are treated probabilistically, the confidence intervals for the model result would be based on the joint confidence interval for all of the input distributions.

In cases where exposures are measured directly, confidence intervals can be developed based on actual sample data. A key aspect of such an effort is the quantification of measurement errors. In all cases, Whitmore also emphasizes the need for qualitative characterizations of sources of uncertainty that are not treated quantitatively, and for clear statements regarding modeling assumptions.

Rish and Marnicio (1988) reviewed several dozen studies with respect to uncertainty concepts, modeling frameworks for treatment of uncertainty, methodologies for dealing with uncertainty, software tools employed, and case study applications. In a related study, Rish (1988) develops a "comprehensive approach" to dealing with uncertainties in estimating and evaluating human health risks. Much of the discussion is of a general nature. Rish deals with issues such as comparison of alternative models, sensitivity of assumptions about correlation among model inputs, response surface methods, probability and decision trees, Monte Carlo simulation, and communication of results. This study is thus a useful overview of general aspects of uncertainty analysis, although it does not deal directly with issues in exposure assessment as does the work of Whitmore (1985).

White and Schaum (1989) discuss the use of Monte Carlo simulation in exposure assessments at EPA. This paper focuses on when to use Monte Carlo techniques and how to use and interpret them. Some of the issues discussed include alternative interpretations depending on the type of data used (e.g., empirically-derived or based on expert judgment), the type of problem (long-term vs. short term, large site vs. small site), and completeness of uncertainty characterizations.

A guide to uncertainty analysis in risk assessment by Finkel (1990) contains many concepts and approaches which are also appropriate for exposure assessment. These include a discussion of sources of uncertainty and variability in parameters, models, and decision-making, methods for communicating uncertainties, mechanics of uncertainty analysis, and implications of uncertainty analysis for risk management decision-making.

As discussed previously, the 1992 Exposure Assessment Guidelines (EPA, 1992) also contain a discussion of uncertainty analysis with respect to exposure assessment.

4.3 Example Applications

As previously indicated, there are a significant number of published exposure assessment studies which feature quantitative approaches to the analysis of variability and uncertainty. A few of these are discussed here to provide an indication of how Monte Carlo techniques are employed in this field, and to identify needs for methodological development.

Constantinou et al (1992) have employed Monte Carlo simulation to estimate the uncertainties in health risks from air toxics emissions from power plants. Their approach was to develop simplified response surface models based on sensitivity analysis of complex fate and transport, deposition, exposure and dose, and health risk models. For example, an atmospheric transport model can be simplified to a function of the chemical emission rate and a multiplicative constant. This constant is estimated based on meteorological data, topographical information, and source characteristics. The simplified model facilitates the analysis of uncertainty in chemical concentration in the air due to uncertainty in emission rates, assuming that all other parameters are fixed. More complex response surfaces can be constructed as appropriate.

Constantinou et al (1992) discuss several considerations in the selection of probability distributions to represent uncertainty in model input quantities. In the case where sufficient data are available, a classical statistical analysis may be performed. A probability distribution estimated this way may represent measurement error. When insufficient data are available for a statistical analysis, expert judgment is required to characterize the distribution. Even in the case where data are available, expert judgment plays a role in the selection of the parametric probability distribution model, even though goodness-of-fit techniques may be employed to estimate the parameters of the distribution based on limited sample size. In their case study, Constantinou employed statistical analysis, literature value ranges, and personal judgments to estimate probability distributions for twelve model parameters treated as statistically independent. Although not acknowledged in the papers, these twelve parameters represent a combination of uncertainty and variable quantities. For example, parameters such as inhalation rate, body weight, and exposure duration tend to be variable from one individual to another. The authors point out that a major source of uncertainty is the dose extrapolation method used to develop cancer potency factors, but the authors argue that this source of uncertainty cannot be evaluated quantitatively due to the lack of information to support the appropriateness of one method versus another.

The results reported by Constantinou et al (1992) indicate that deterministic ("point estimate") approaches to estimating risks for the particular case yielded a conservative estimate corresponding to the 93.5 percentile of the probabilistic risk assessment. Note, however, that the estimated distribution is a hybrid frequency/probability distribution based on a simulation of both variability and uncertainty on the same basis. Thus, the simulation does not properly distinguish between these two different types of quantities.

In a risk assessment of adverse health effects from exposure, via several pathways, of tetrachloroethylene (PCE) in groundwater, McKone and Bogen (1992) employ Monte Carlo simulation to estimate population risks. The authors address uncertainty in the exposure models, uncertainty in the dose-response models derived from animal data, and identify key contributors to overall uncertainty in population risk. Uncertainties and variabilities in model parameters were characterized based on empirical data, statistical analyses, other studies, and professional judgment. To account for correlations between parameters, such as intake rate and body weight, some model input quantities were grouped into new parameters, such as intake rate per unit body weight, and the variability in the new parameter was estimated. The authors used Crystal Ball to perform the Monte Carlo simulation. In analyzing the contribution to variance of the results, the key sources of variance in the population risk, in decreasing order, were variance in carcinogenic potency, PCE concentration, and the parameters of the exposure model. The authors comment that some factors are more properly treated as uncertainties, while others are variabilities. They suggest that the results could be reported graphically using separate dimensions for uncertainty and variability, although they do not actually do so.

A paper by Bogen and Spear (1987) and a book by Bogen (1990) consider in more detail the differences between uncertainty and variability. These studies do a thorough job of distinguishing between variability and uncertainty as inputs to an assessment. Although Bogen and Spear do not present a graphical representation of results as suggested by McKone and Bogen, they do present a number of probabilistic descriptors of risk to different individuals in the population. These include a randomly selected individual, a mean risk individual, a 95th percentile individual, and a "maximum risk" individual. Each individual represents a different sample from the variability in the population distribution. For each of these defined individuals, an estimate of uncertainty in risk is given. A key assumption in the analysis was that the uncertainties faced by all individuals are the same, such that there is no ambiguity regarding the ranking of their risk compared to other individuals. Interpreted differently, the underlying assumption is that the uncertainties are 100 percent correlated between all individuals. The validity of this assumption depends, of course, on the specific problem.

A review of these and other studies indicate that there is typically little attention paid to the notion of separating variability and uncertainty. However, these types of quantities have different implications for estimating population and individual exposures (and risks), and should not be mixed together without a clear statement as to the purpose of such a mixture. For example, in cases where there is relatively little uncertainty, combining uncertainty and variability in a simulation may lead to little error in estimating population characteristics (e.g., percentages of the population subject to various exposure levels). The next chapter presents an alternative approach to simulating uncertainty and variability in a model.

5.0 A GENERALIZABLE APPROACH TO SIMULATING VARIABILITY AND UNCERTAINTY

An approach to separately modeling variability and uncertainty was developed based on a two-dimensional Monte Carlo simulation. First, a taxonomy of Monte Carlo simulation options are described which motivate the need for considering uncertainty and variability separately. Then a generalized approach for Monte Carlo simulation of variability and uncertainty is described. The approach is illustrated by a case study of a hypothetical exposure model.

5.1 A Taxonomy of Simulation Options

In general, exposure (e.g., lifetime average daily dose) is a hybrid frequency/probability distribution, representing the frequency of different exposure levels due to the interindividual heterogeneity of the population of exposed individuals, and the probability of obtaining specific exposure levels for a given individual based on uncertainty in measurements or estimates. Thus, an exposure model may be written as:

$$E = E(\mathbf{V}, \mathbf{U}) \tag{6}$$

The estimate of exposure to the population of exposed individuals is a function of variability in parameters **V,** which have different values for each member of the population, and uncertainty in parameters **U** for which there is some lack of information about their true values. In general, it is possible for multiple uncertain quantities to be correlated with each other. It is also possible for the uncertainties, **U**, to be correlated or dependent on the variable quantities.

For a specific individual, the exposure equation may be written as:

$$E_i = E(\mathbf{v}_i, \mathbf{U}) \tag{7}$$

The specific values $\mathbf{v_i}$ of the empirical quantities **V** differ from one individual to another, representing the interindividual heterogeneity of the population of exposed individuals. However, in this formulation, the uncertainty in exposure faced by each individual is assumed to be the same; the realization of uncertain values is independent of each individual and, furthermore, the realization of uncertain values is the same for all individuals. An example of this type of problem would be exposure of a population of individuals to a contaminant in drinking water from a central supply system. Each individual will differ in terms of how much water they drink and over what exposure period they consume water from the contaminated supply system. However, given that the supply system equally distributes the contaminant to all users, all users are faced with the same value of the concentration. The measurements of the concentration may be uncertain, but all individuals face the same actual concentration level.

In a more general case, the realization of uncertainties may differ from one individual to another. This would be the case, for example, at a hazardous waste site, where different individuals may spend time at different parts of the site. Thus, individuals are exposed to different levels of contaminants, with different uncertainties regarding the actual amounts. Furthermore, the uncertainty in the measurements of a contaminant level at one part of the site will generally be independent of the uncertainty in the measurement at another location. Thus, the concentration level actually consists of a variable component and an uncertain component. Assuming that the variability and uncertainty in the concentration are appropriately disaggregated, the resulting general exposure equation for a single individual becomes:

$$E_i = E(\mathbf{v}_i, \mathbf{u}_i) \tag{8}$$

The uncertainty in measurements of concentration may be proportional to the level of concentration. For example, many measurement techniques tend to have less precision, on an absolute basis, as the magnitude of the quantity being measured increases. Thus, we may have the case where the uncertainty is conditional (or dependent upon) the specific values of the empirical (variable) quantities:

$$E_i = E(\mathbf{v}_i, \mathbf{u}_i(\mathbf{v}_i)) \tag{9}$$

For example, the variance in measurement of a concentration level is often proportional in some way to the actual value of the concentration. The concentration may vary from one site location to another, and thus be a source of variability in the exposure estimate.

An alternative way to view these simulation options is in terms of the effect that different relationships between variability and uncertainty have on the cumulative distribution function for the population of exposures to all individuals. In general, individuals can be rank ordered with regard to their exposures. Analogous to the discussion presented by Bogen and Spear (1987) for risks, we consider the following cases. In the first case, we assume that there are no variable quantities in the exposure model, implying that all individuals are homogeneous in terms of behavior. Thus, all individuals face the same exposure, which is uncertain:

$$E_i = E = E(\mathbf{U}) \tag{10}$$

The cumulative distribution function (cdf) of E, F(E≤e), is the *probability* distribution of exposure faced by all individuals. The function F(E≤e) varies from 0 to 1 and represents the fraction of the population with an exposure level less than or equal to some value e.

If there are no uncertain quantities in the exposure model, then each individual can be rank ordered unambiguously with respect to increasing exposure levels. The exposure model is written as:

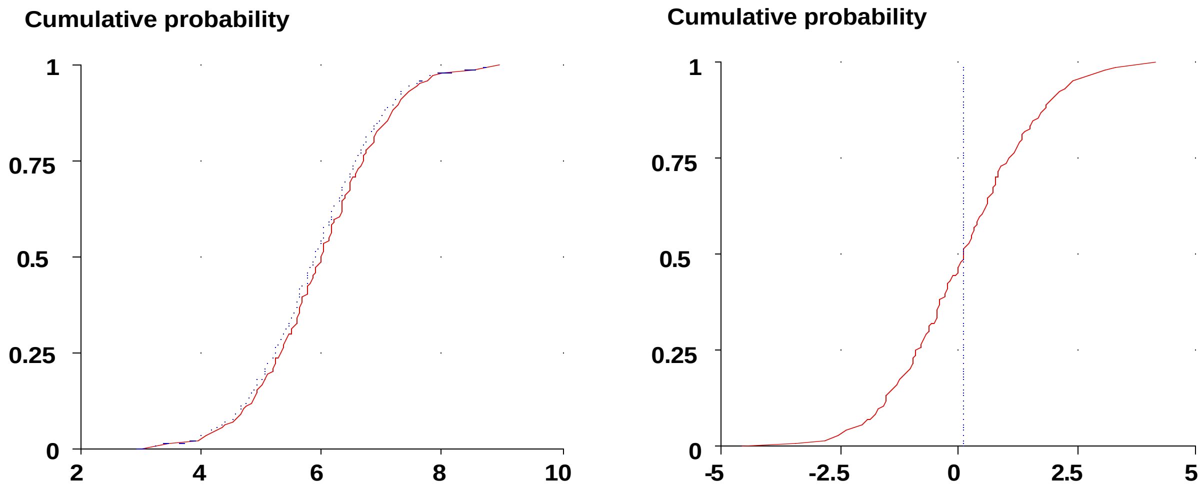
$$E = E(\mathbf{V}) \tag{11}$$

where E is a *frequency* distribution of the certain exposure levels faced by different individuals. We may rank order each of the n individuals with respect to their exposure levels such that:

$$F(E \le E_1) \le F(E \le E_2) \le \cdots \le F(E \le E_n) \tag{12}$$

If there is both variability and uncertainty in the exposure model, and if the uncertainties are systematically applicable to all individuals (e.i., if all individuals face the same rank ordering of realizations of the uncertain quantities), then the rank ordering of individuals is the same as the case when only variability in empirical quantities is simulated. However, for any given individual, there will be a range of values associated with the uncertainty in exposure. Thus, we may define two cumulative distribution functions. One, given by Equation (11), represents the frequency distribution of variability across all individuals, while the other, shown as Equation (12), represents the probability distribution of the uncertainty in exposure level for a specific individual:

$$E_i = E(\mathbf{v}_i, \mathbf{U}) \tag{13}$$



(a) Exposure Level

(b) Difference in Exposure Level assuming ρ=1 (narrow distribution) and ρ=0 (wide distribution)

Figure 4. Effect of assumptions regarding uncertainty on the differences in exposure between two individuals.

In the case where the uncertainties are not correlated with ρ=1 for all individuals, as assumed above, then the rank ordering of individuals will change depending on each realization of uncertainty. Thus, in addition to uncertainty regarding the actual exposure level faced by a given individual, there is also uncertainty regarding that individual's percentile within the population distribution.

To illustrate this concept, consider two individuals whose behaviors and exposures are nearly the same. The uncertainties in their exposures are modeled by the probability distributions in Figure 4(a). Note that there appears to be a slight difference in the means of the two distributions, although their standard deviations are nearly the same. In Figure 4(a), the difference in exposure level between the two individuals is modeled. If the uncertainty in the exposure levels is +100 percent correlated for the two individuals, then the individual with the higher mean exposure will always have a higher exposure. Therefore, this individual will always rank higher in the variability frequency distribution compared to the other individual. If, however, there is no correlation in the uncertainties in exposure to the two individuals, there is a 48 percent probability (in this case) that the individual with the lower mean could have a higher exposure. Thus, the rank ordering of the two individuals is uncertain. Forty-eight percent of the time, the individual with the lower mean exposure would actually have a higher rank than the other individual. This type of interaction would extend, of course, to all individuals within several standard deviations of the mean exposure values.

5.2 Two-Dimensional Monte Carlo Simulation

A two-dimensional approach to Monte Carlo simulation was developed to properly disaggregate and evaluate the consequences of variability and uncertainty, as well as to simulate their interactions. Thus, this approach is able to deal with the general case in which uncertainties may be independent from one individual to another and conditional on variable quantities. Because policy questions often focus on population characteristics such as the number of individuals at or above a given exposure level, or individual descriptors, such as the characteristics of persons at the high end of exposures and the amount of their exposures, it is critically important to disaggregate variability and uncertainty. However, it is also important to properly evaluate any interactions between them that might affect the rank ordering of

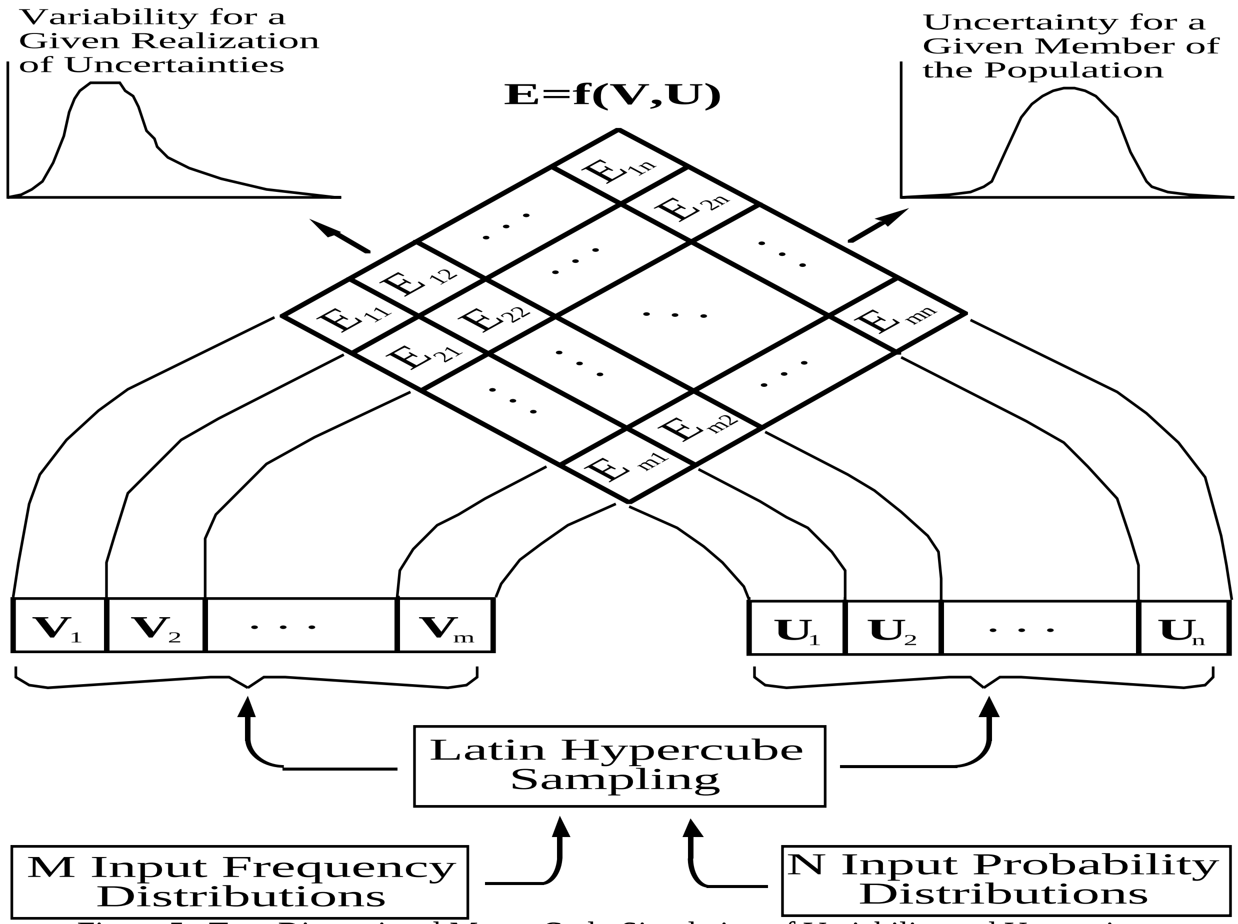


Figure 5. Two-Dimensional Monte Carlo Simulation of Variability and Uncertainty.

individuals. These types of interactions cannot be evaluated in traditional approaches to Monte Carlo simulation in exposure assessment. The two-dimensional simulation approach is shown schematically in Figure 5.

Given a general model E=f(**V**,**U**) as described above, the first step is to disaggregate the model variables into variable and uncertain components. For all variable quantities, frequency distributions must be specified. For all uncertain quantities, probability distributions must be specified. The development of both sets of distributions can be based on bootstrapping with actual data, goodness-of-fit tests of data and parametric distribution models, or technical judgments, as discussed in previous sections. A sampling scheme, such as random Monte Carlo simulation or Latin Hypercube sampling, is then employed to generate two sets of samples. For each of M variable quantities, the frequency distributions are simulated with a sample size of m. For each of the N uncertain quantities in the model, the corresponding probability distributions are simulated with a sample size of n. In principle, the sample sizes m and n for the variable and uncertain dimensions of the simulation need not be the same. However, it is preferable that these sample sizes be the same. Clearly, the precision of simulated distributions for uncertainties in exposure depends on the uncertainty sample size, n. Interactions among the multiple variable quantities can give rise to similar exposure levels. In cases where uncertainties between individuals are statistically independent, the rank ordering of the individuals can change, leading to uncertainty in both exposure levels and population fractiles. The precision of estimates of population fractiles depends on the variability sample size, m. Thus, the sample sizes of both dimensions can be equally important. In this case, the sample size for the two-dimensional simulation is $m \bullet n = m^2 = n^2$. Clearly, this can impose a potentially severe computational burden, depending on the required sample size.

The model is repetitively evaluated for each combination of samples from the variable and uncertain parameters. This is represented in Figure 5 by the matrix of values $E_{i,j}$, where i is an index from 1 to m of the sample values for the vector of variable quantities, and j is an index from 1 to n of the sample values for the vector of uncertain quantities. Any column of the matrix represents the frequency distribution for variability in exposure levels for the population of individuals for a given realization of uncertainties for each individual. Any row of the matrix represents the probability distribution for uncertainty in exposure level for a given member of the population.

In cases where the uncertainties are independent from one individual to another, the rank ordering of individuals is also uncertain. This important interaction is captured by the two dimensional approach to simulation, because for every sample of values for the vector of uncertain quantities, a separate frequency distribution for variability is simulated. For a given individual, each sample in the uncertainty dimension yields a separate estimate of the rank within the population. Therefore, there will be n estimates of the rank, or percentile, of each individual in the population. These estimates represent a probability distribution for the rank of the individual within the population. Thus, it is possible, using the two-dimensional approach to explicitly estimate not only the uncertainty in the exposure level faced by that individual, but also the associated uncertainty in their rank within the population. In this regard, the two-dimensional approach presented here is a systematic method to model and evaluate both the separate and interacting effects of variability and uncertainty in a model.

In contrast, the typical approaches used by investigators as described in Chapter 5 do not capture these interactions, nor do they properly convey the differences between variability and uncertainty. In the simulations by Constantinou et al (1992) and McKone and Bogen (1992), and others, variability and uncertainty were simulated in the same dimension. The resulting hybrid frequency/probability distribution for exposure is only meaningful if interpreted to represent an individual selected at random from the population. However, it is inaccurate to draw conclusions from such results regarding the rank ordering of individuals within the population, or the exposure level faced by an individual at a given fractile of the population.

Consider the taxonomy of simulation options in the previous section. In the case where there are only variable quantities in the model, the two-dimensional simulation simplifies to a one dimensional simulation, and the matrix $E_{i,j}$ simplifies to a vector E_i. The vector E_i represents the variability in exposures due to the variability in the model input parameters. In this one-dimensional simulation, it is assumed that the variability is known without uncertainty. In the other limiting case, if there is no variability from one individual to another (e.g., all individuals are the same), but there is uncertainty, the two-dimensional simulation again simplifies to one dimension. The matrix $E_{i,j}$ simplifies to a vector E_j, representing the uncertainty in exposures due to uncertainty in model inputs. However, the traditional one-dimensional approaches to modeling uncertainty and variability do not fall into either of these categories.

5.3 An Illustrative Case Study

The two-dimensional Monte Carlo simulation approach for dealing with variability and uncertainty is illustrated here via a hypothetical case study. Consider a simple scenario involving exposure to a chemical via ingestion of a contaminated fluid (e.g., drinking water). A simple exposure model would be:

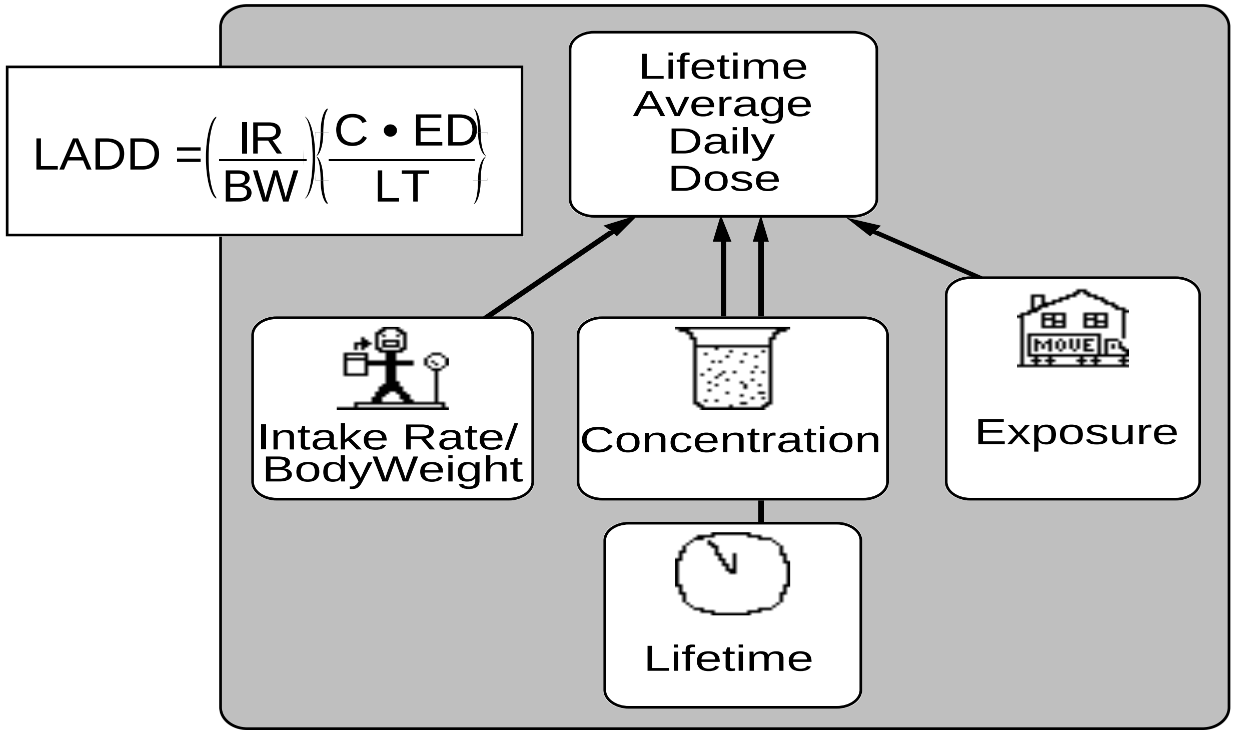


Figure 6. Influence diagram for a simple exposure model

$$\text{LADD} = \left(\frac{\text{IR}}{\text{BW}}\right)\left\{\frac{\text{C} \bullet \text{ED}}{\text{LT}}\right\} \tag{14}$$

In this formulation, intake rate and body weight are treated as a single parameter, because they are known to be correlated. This model is shown graphically as an influence diagram in Figure 6. For simplicity, it is assumed that all of each exposed individual's drinking water consumption is from the contaminated source.

5.3.1 Model Formulation

In order to evaluate the separate and interactive effects of variability and uncertainty, it is necessary to disaggregate all model inputs into separate variability and uncertainty components. In this example, it is assumed that variable and uncertain components of each parameter are additive, and the exposure equation is rewritten as:

$$\text{LADD} = \left\{\left(\frac{\text{IR}}{\text{BW}}\right)_v + \left(\frac{\text{IR}}{\text{BW}}\right)_u\right\}\left\{\frac{(\text{C}_v+\text{C}_u) \bullet (\text{ED}_v+\text{ED}_u)}{(\text{LT}_v+\text{LT}_u)}\right\} \tag{15}$$

The additivity assumption is based on the notion that there may be systematic and/or random error associated with the values of the model parameters. The disaggregated exposure model is illustrated graphically in Figure 7.

While the disaggregation of each model parameter into two components may seem arbitrary, in principle it should be possible to separately estimate the underlying variability and the uncertainty associated with measurements of the variable quantity. As an illustration, consider the simple case of measurement of a variable quantity. The distribution of measured values is given by:

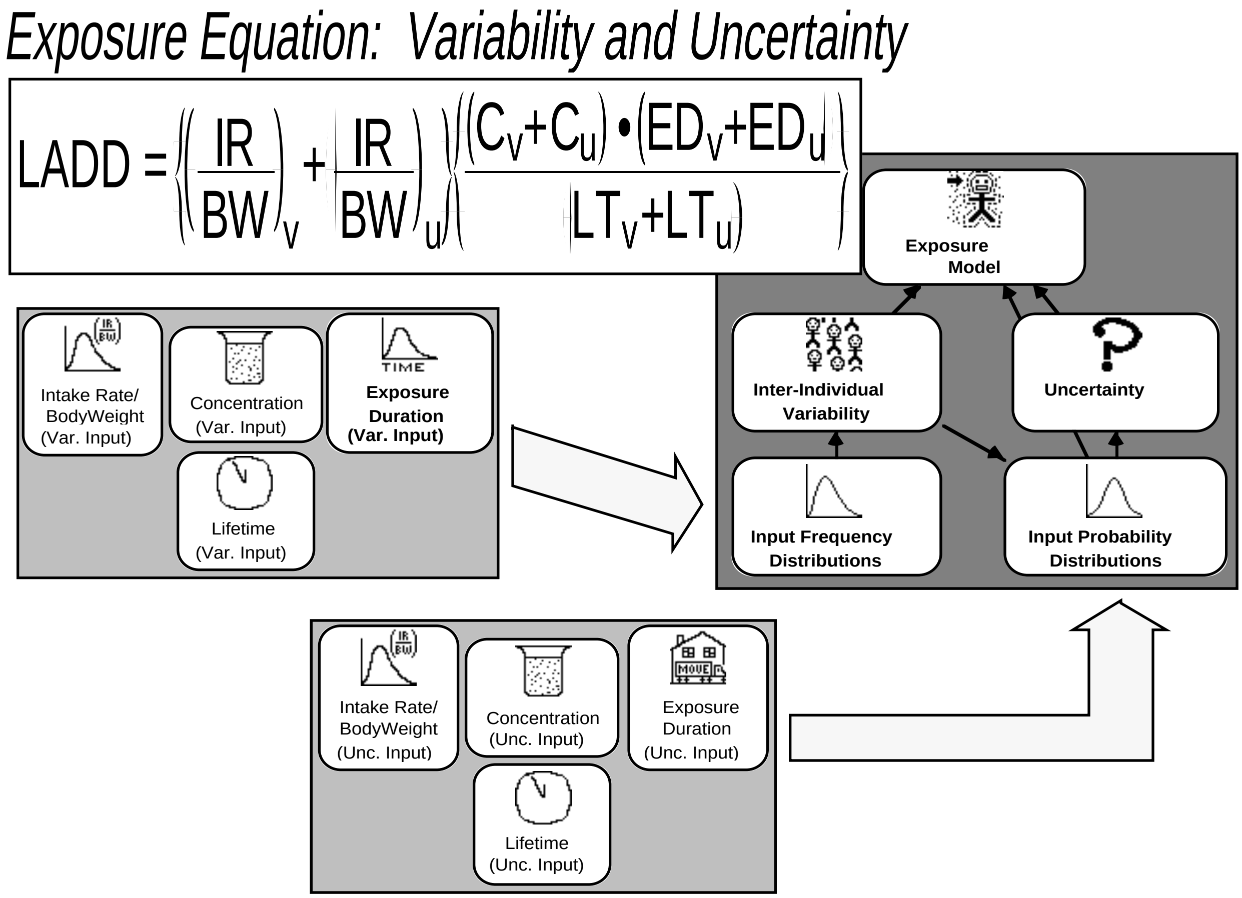


Figure 7. Influence diagram for exposure model with separate components for variability and uncertainty.

$$M = V + U \tag{16}$$

where V is the actual variability in the quantity being measured, and U is the uncertainty associated with the measurement technique. This uncertainty may be both random and biased. In principle, measurements made as part of EPA programs should be based on validated test methods, for which there is quantitative information about the distribution for U. Therefore, given samples for M and knowledge of the distribution of U, it is possible to estimate V:

$$V^{*} = M - U \tag{17}$$

In this manner, uncertainty and variability can in principle be disaggregated and used as inputs to a model. Alternatively, another approach for separating variability and uncertainty is based on estimating confidence intervals for the parameters of distributions intended to represent variability. Knowledge about the uncertainty in the parameters of the distribution can be used to simulate uncertainty about the frequency distribution. Whitmore (1985) touches upon this issue in his study.

5.3.2 Model Input Assumptions: Variability and Uncertainty

The data requirements for the disaggregated model include probabilistic estimates of both variability and uncertainty in each model parameter. To illustrate the modeling approach, hypothetical distributions have been assumed. However, to demonstrate that data are available

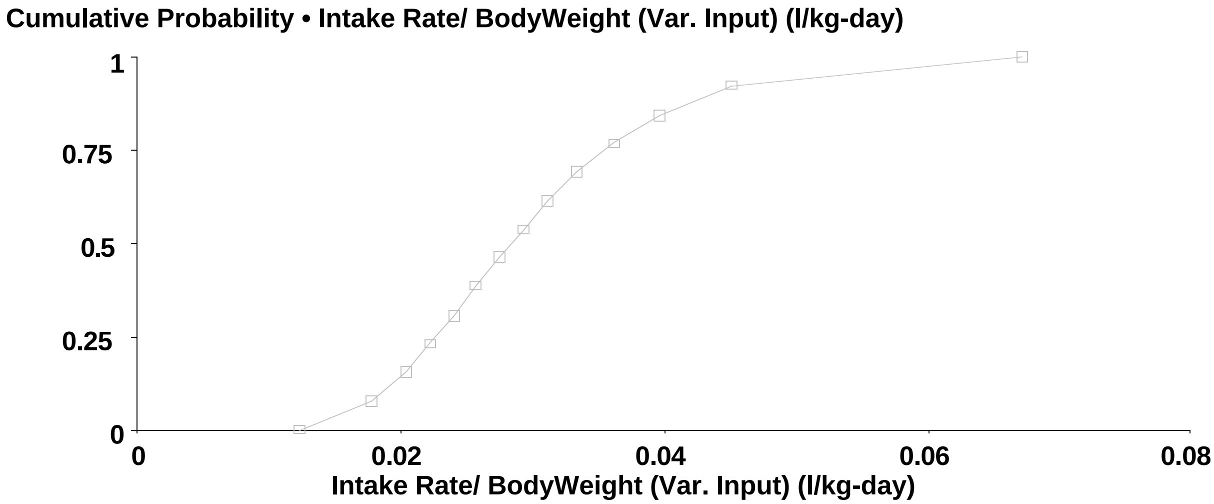


Figure 8. Model input frequency distribution for variability in intake rate per unit body weight.

to support such assessments, a few of these distributions are based directly on the exposure assessment literature.

5.3.2.1 Sources of Variability

The exposure assessment model consists of five parameters. All five of these are sources of variability in the assessment. Body weight, intake rate, exposure duration, and lifetime vary from one individual to another. Concentration of the chemical in the liquid stream which comes in contact with each individual may also be variable, due to the spatial or temporal distribution of individuals and the mechanism for transport and dilution of the contaminated stream. Each of these potential sources of variability are discussed in more detail.

Body weight and daily water intake rate are clearly variable from one individual to another, and also tend to be a function of developmental stage and sex. However, variability in body weight and daily intake rate are also correlated, with larger persons tending to drink more. McKone and Bogen (1992) have calculated a frequency distribution for the variability in the ratio of intake rate to body weight that is based on an assumed relationship between the two. The distribution for the ratio of intake rate to body weight represents the variability in intake rate that cannot be explained by body weight alone. The distribution estimated by McKone and Bogen is lognormal, with an arithmetic mean of 3×10^{-2} l/kg-day, and a geometric standard deviation of 1.4. This distribution is shown graphically in Figure 8.

Exposure duration is the length of time, in days, that an individual consumes contaminated water. To a first approximation, the exposure duration in this case is the amount of time an individual lives in a home connected to the contaminated water supply. Neglecting the possibility that the individual could change homes both connected to the same water supply, we model the exposure duration based on census survey data for the amount of time that people have lived in their current homes, per the Exposure Factors Handbook (EPA, 1989). In a more detailed assessment, it would be preferable to conduct a survey for the specific community exposed to the contaminated water supply, because the distribution of their exposure duration behavior may differ from the national distribution. However, there are two other shortcomings to the use of the data from the Exposure Factors Handbook.

One shortcoming is that the distribution is not completely specified, and the second is that the data are for a surrogate to the parameter we want to estimate. With regard to the first point, distributional data are presented in the Exposure Factors Handbook for individuals based

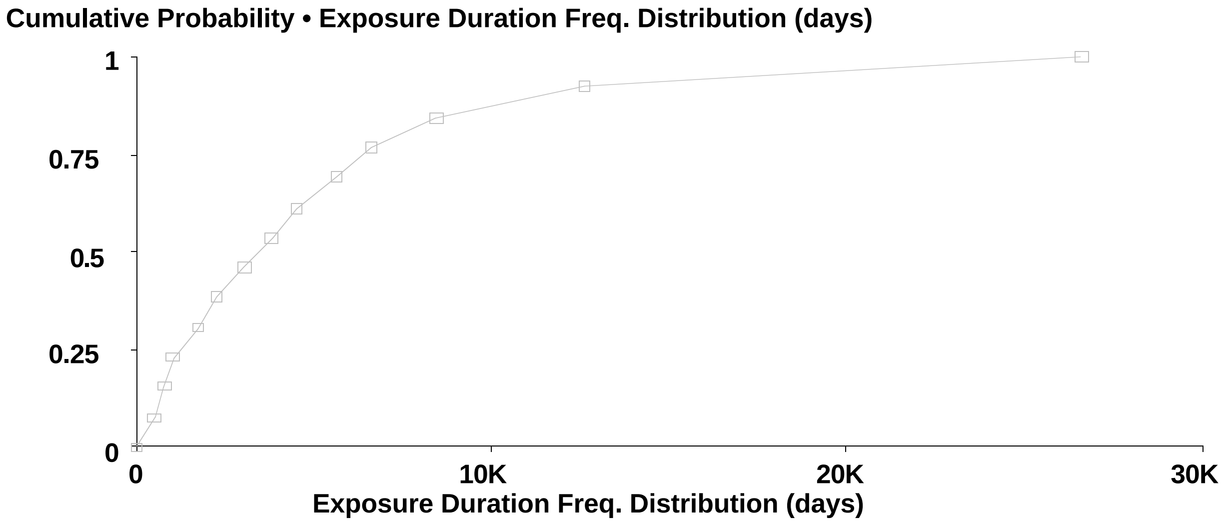


Figure 9. Model input frequency distribution for variability in exposure duration.

on the percentage of all respondents who have lived in their homes within certain time ranges. However, a number of respondents have lived in their home longer than the highest value report (33 years). Therefore, a bounding assumption was made that individuals would typically live in a home no longer than the average lifetime of an American, which is 75 years. Note that this assumption has no effect on the fractiles of the distribution below the 93rd percentile, but it does affect the mean of the distribution. The empirically-based distribution for exposure duration, including the bounding assumption, is shown in Figure 9. The second shortcoming is that the reported data are for the number of years lived in the current household at the time of the survey. This is not the same as the total number of years which a given respondent will actually live in their current household. Thus, the survey data will systematically underestimate the total time a resident will live in their current dwelling.[2] This is a source of uncertainty and is dealt with separately.

The lifetime of an individual is variable across the population. Some analysts have assigned distributions to this parameter, while others have used a fixed value of 75 years. From a policy perspective, it may be less controversial to use a fixed value. This in effect is a rights-based approach that implies that all individuals have a right to live at least as long as the national average. However, because the lifetime is used in this example as an averaging time, such an approach would also lead to underestimation of the lifetime average daily dose faced by members of the population who do not live as long as the national average. For illustrative purposes only, the variability in lifetime is assumed to be normally distributed with a standard deviation of five years, and a mean of 75 years. In principle, a more realistic distribution can be developed based on mortality data, but such an effort was beyond the scope of this illustrative case study.

Finally, the remaining source of variability is the concentration of the chemical contaminant in the drinking water. This quantity is site-specific, and thus a hypothetical assumption is made here. In the case of a municipal water supply system, the average

---

[2] The data are complete up to the 93rd percentile, with seven percent of those surveyed having lived in their current residence more than 33 years. Based on linear extrapolation of the data, the 50th percentile is approximately 9 years and the 90th percentile is 30 years. As an aside, the EPA Exposure Factors Handbook states that the "50th percentile represents the average length of time a typical homeowner will live in the same house." This statement is wrong on two counts: (1) the 50th percentile is the median, not the average (mean) of the distribution. Depending on the upper bound assumptions, the average is on the order of 15 years; (2) the data are not for the total length of time of residence in a house, as discussed.

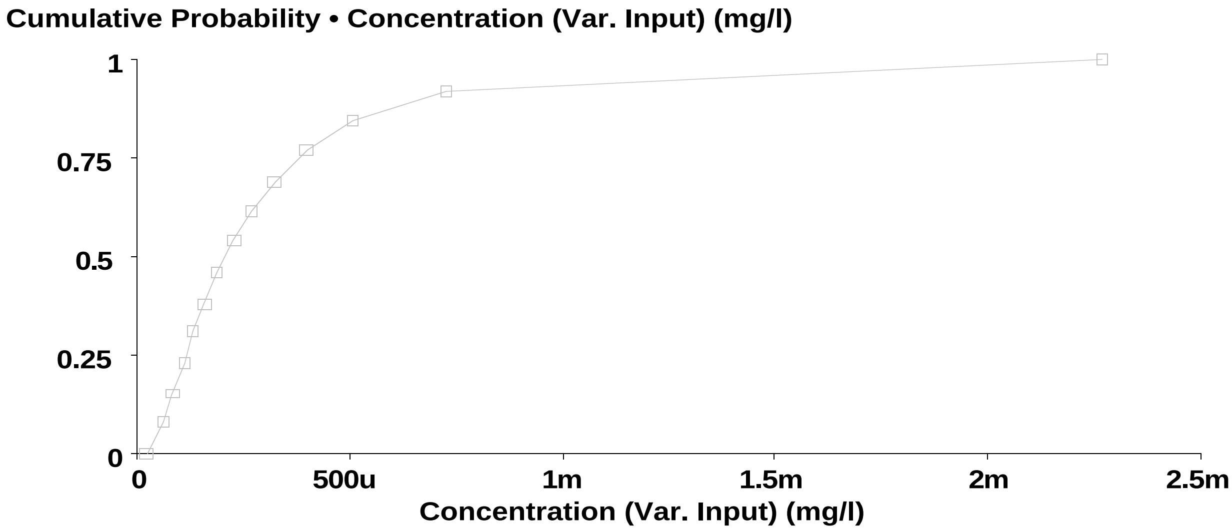


Figure 10. Model input frequency distribution for variability in contaminant concentration.

concentration may be very similar for all exposed individuals. In the case of individuals who draw their water from wells, the average concentration would be variable depending on location, due to dilution, diffusion, and transport of the contaminant within the aquifer. For the purposes of this case study, it is assumed that the variability in concentration is lognormally distributed, as shown in Figure 10.

The assumptions made here regarding sources of variability are summarized in Table 1.

Table 1. Illustrative Assumptions Regarding Sources of Variability

Parameter	Symbol	Units	Nominal Value	Distribution	Distribution Parameters	Comments
Intake Rate per Unit Body Weight	IR/BW	l/kg-day	$3x10^{-2}$	Lognormal	$\mu=3x10^{-4}$ GSD=1.4	McKone and Bogen (1992)
Concentration	C	mg/l	$3x10^{-4}$	Lognormal	$\mu=3x10^{-2}$ GSD=2.5	Hypothetical, depends on site.
Exposure Duration	ED	days	3,300	Empirical	Empirical	Based on EPA, 1989
Lifetime	LD	days	27,400	Normal	μ=27,400 σ=1825	Illustrative (see text)

5.3.2.2 Sources of Uncertainty

Of the four parameters in the illustrative exposure model, all four are modeled as sources of variability. Three of these are also modeled as sources of variability. It is assumed that there is not a significant source of uncertainty in the population distribution for lifetimes.

The estimate of variability in intake rate per unit body weight is based, presumably, on measurements. However, data regarding the measurement precision and accuracy are not readily available. In principle, such information could be used as described previously to explicitly separate uncertainty and variability. However, given the lack of such data and the illustrative nature of this case study, two simple cases where considered. Both cases assume that the

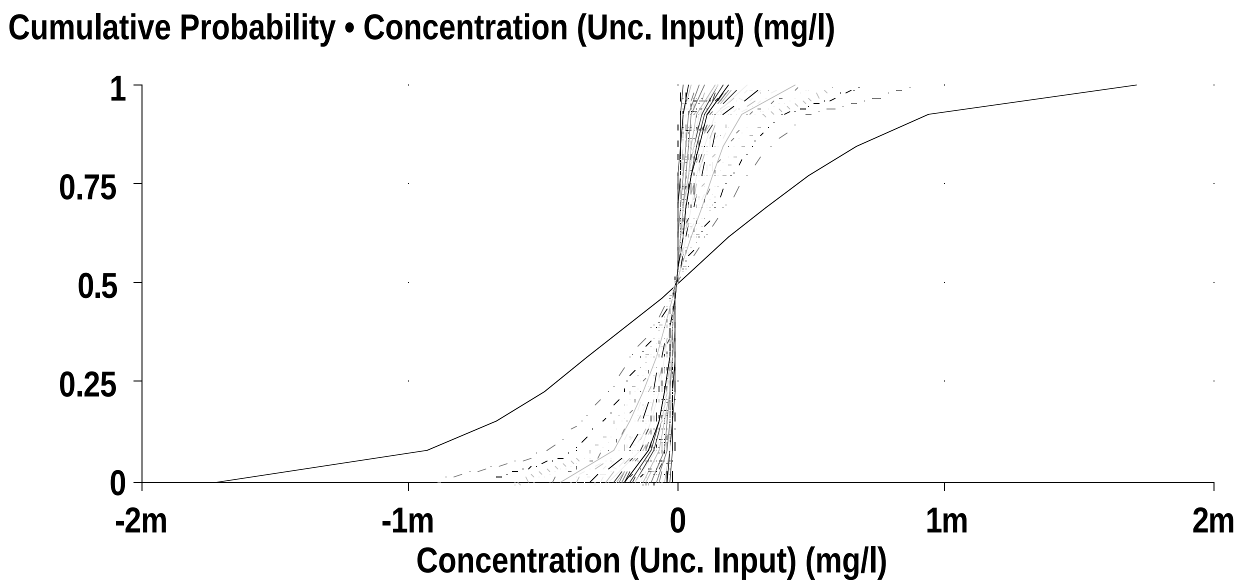


Figure 11. Family of Probability Distributions for Uncertainty in Concentration Associated with Each Individual in the Population.

lognormal distribution for variability represents the true variability, and that the measurement error is unbiased (mean equals zero) and normally distributed. In one case, a standard deviation of 10 percent of the "actual" value is assumed, and in an alternate case, a standard deviation of 30 percent is assumed. Similar illustrative assumptions were made regarding uncertainty in the measurements of concentration.

The estimate of variability for the exposure duration is believed to be biased toward low values, as previously described. Therefore, the uncertainty in the exposure duration was assumed to be biased, and is represented as a uniform distribution with a lower value equal to the sample from the variability distribution and an upper value of five years additional time in the current residence. An alternate sensitivity case was considered assuming up to 10 additional years in the current residence.

For all three sources of uncertainty, the shapes of the distributions were assumed to apply to all individuals, but to be statistically independent from one individual to another. Although the shapes of the distributions are the same for all individuals, the parameters of the distributions are not. For example, in the case of intake rate per unit body weight, it is assumed that the standard deviation is 10 (or 30) percent of the nominal value sampled from the variability frequency distribution. A similar assumption was made regarding uncertainty in concentration. Thus, the standard deviations of the uncertainty in intake rate per unit body weight and concentration were modeled as a function of the variability in those two parameters. As a result, there is a family of distributions for each of these uncertainty parameters associated with each individual in the population. The family of distributions is illustrated in Figure 11 for uncertainty in concentration.

The assumptions regarding sources of uncertainty are summarized in Table 2.

5.3.3. Running the Model

A two-dimensional simulation of the model in Equation (15) was performed using Latin Hypercube sampling. The software package Demos was used for this simulation. Demos functions were used to model the input frequency and probability distributions. Independence

Table 2. Illustrative Assumptions Regarding Sources of Uncertainty

Parameter	Symbol	Units	Nominal Value	Distribution	Distribution Parameters	Comments
Intake Rate per Unit Body Weight	IR/BW	l/kg-day	0	Normal	ν=0.1 ν=0.3	Unbiased measurement error, two cases
Concentration	C	mg/l	0	Normal	ν=0.1 ν=0.3	Unbiased measurement error, two cases
Exposure Duration	ED	days	0	Uniform	U(0,1826), U(0,3652)	Biased error, two cases
Lifetime	LD	days	n/a			No uncertainty

among uncertainties for each individual were simulated by using separate input uncertainty distributions for intake rate per unit body weight, exposure duration, and concentration for each individual in the simulated population. The sampled values from the variable and uncertain components of each model input were separately indexed into dimensions of variability and uncertainty, as illustrated in Figure 5. The exposure model was then evaluated for each combination of input values for the variable and uncertain components.

For the purposes of this illustrative case study, the selection of simulation sample sizes was arbitrary. The selection was based primarily on a trade-off between computational speed and obtaining reasonably stable results. Sample sizes for each dimension of 10, 20, 50, and 100 were evaluated. On this basis, a simulation of sample size of 50 for both the variable and uncertain dimensions was selected for the detailed case studies. Thus, a total of 2,500 exposure model evaluations were performed.

5.3.4 Interpreting and Presenting Results

The modeling results are presented in Figure 12. The figure consists of 50 separate cdf's representing the variability in exposure levels to the population of individuals for alternative realizations of uncertainties for each individual. In turn, each of these 50 cdf's is estimated based on simulation of 50 separate individuals. The results in Figure 12 are for the base case in which uncertainty in intake rate per unit body weight and concentration have a standard deviation of 10 percent of the variability samples for those parameters.

5.3.4.1 Uncertainty in Exposure Levels

For any given fractile of the simulation results in Figure 12, there is a probability distribution representing uncertainty in the exposure level. The uncertainties in exposure levels associated with five different fractiles of the distribution are shown in Figure 13. Each of these distributions is estimated based on the exposure levels associated with the given fractile for each of the 50 simulations of uncertainty in the exposure model. Note that the range of uncertainty tends to increase for the larger fractiles. This is a result of the positive skewness of all of the variability simulation results and the interactions with the variances in uncertainties for intake rate per unit body weight and concentration, which are proportional to the values of these quantities from the variability distributions.

These distributions provide a quantitative indication of the uncertainty in exposure level faced by a given fractile of the population. Furthermore, they can also be used to determine to what degree of confidence a given percentage of the population faces a specific exposure level.

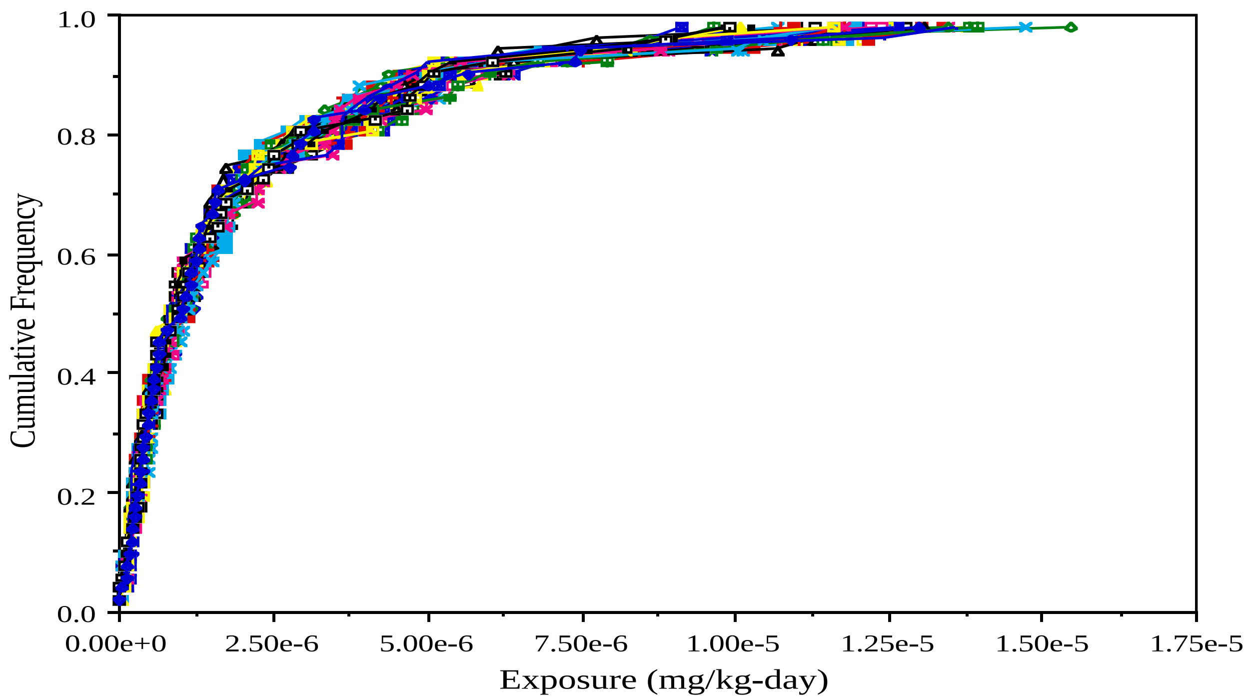


Figure 12. Results of a two-dimensional simulation of variability and uncertainty in a hypothetical exposure model.

For example, suppose that an exposure level of 7.5×10^{-6} mg/kg-day were of regulatory interest for health reasons. This exposure level corresponds to the 20th percentile of the probability distribution for uncertainty in exposure levels to the 95th percentile of the exposed population. Thus, the simulation results indicate that there is a 20 percent probability that 95 percent of the population would face an exposure less than this amount, due to the interactions between variability and uncertainty. Conversely, there is an 80 percent probability that five percent of the population would face a higher exposure level. If a higher exposure level were of concern, such as 1.0×10^{-6} mg/kg-day, there would be only a five percent probability that five percent of the population would have a higher exposure level. In this case, one could assert with 95 percent confidence that 95 percent of the population faces a lower exposure level.

This approach and interpretation differs significantly from typical approaches in which uncertainty and variability are simulated as part of the same dimension. In such simulations, uncertainty is usually treated as if it were variability, and results are reported in terms of the percent of the population at or above a given exposure level. However, these results can be erroneous if there are significant uncertainties involved, because the simulation of uncertainties with variability leads to bias. At the upper fractiles of the distribution, a one-dimensional simulation including both variability and simulation will tend to over-estimate exposure levels, while at the lower fractiles, such a simulation tends to under-estimate the exposure levels. The resulting estimates of exposure levels will not properly reflect either the uncertainty in the exposure level faced by a specific fractile of the population, nor will it account for interactions between uncertainty and variability that affect the rank ordering of individuals.

5.3.4.2 Uncertainty in Fractiles

For any given exposure level of the simulation results in Figure 12, there is a probability distribution representing uncertainty in the associated fractile of the population. This uncertainty is partly due to the different rank ordering of individuals within the population with respect to

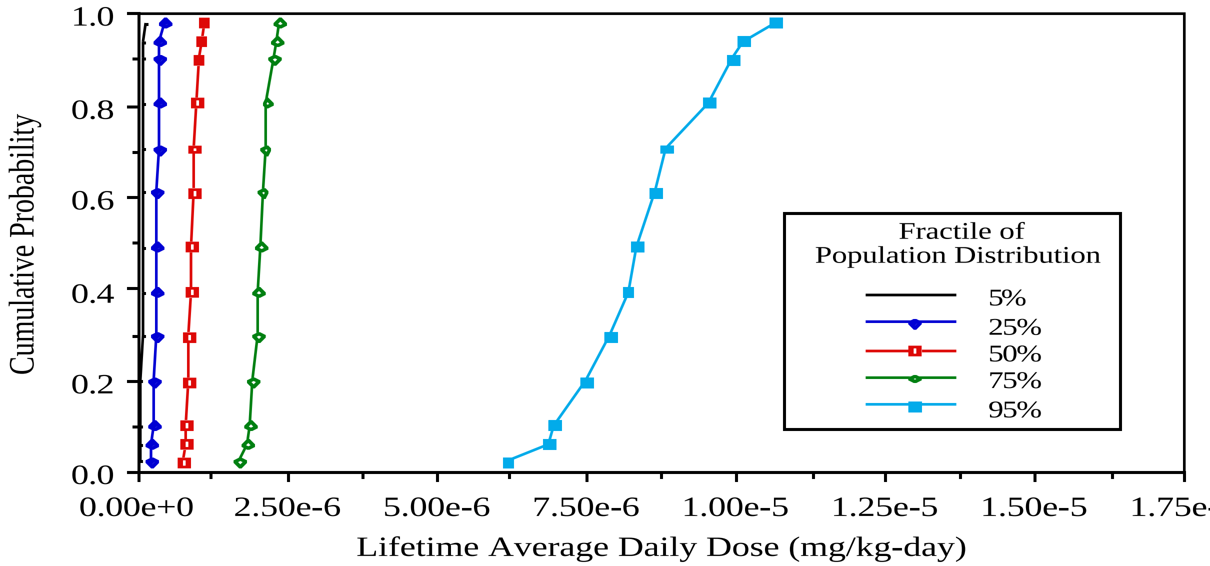


Figure 13. Uncertainty in exposure levels for specific fractiles of the population distribution.

exposure levels, based on alternative realizations (samples) of the uncertainties in the exposure model.

The uncertainties in fractiles associated with five different exposure levels of the population distribution are shown in Figure 14. Each of these distributions is estimated based on the fractile associated with the given exposure level for each of the 50 simulations of uncertainty in the exposure model. Note that while the number of samples used to estimate these cdf's is the same as the uncertainty sample size, n, the resolution of the estimates of the fractiles (the x-axis of the graph) depends on the variability sample size, m. In this case, because a sample size of m=50 was used, the fractiles are estimated in intervals of approximately 0.02.

These distributions provide a quantitative indication of the uncertainty in the fractile of the population faced with a given exposure level. Furthermore, they can be used to determine the probability that a given fractile of the population faces an exposure of less than or equal to a

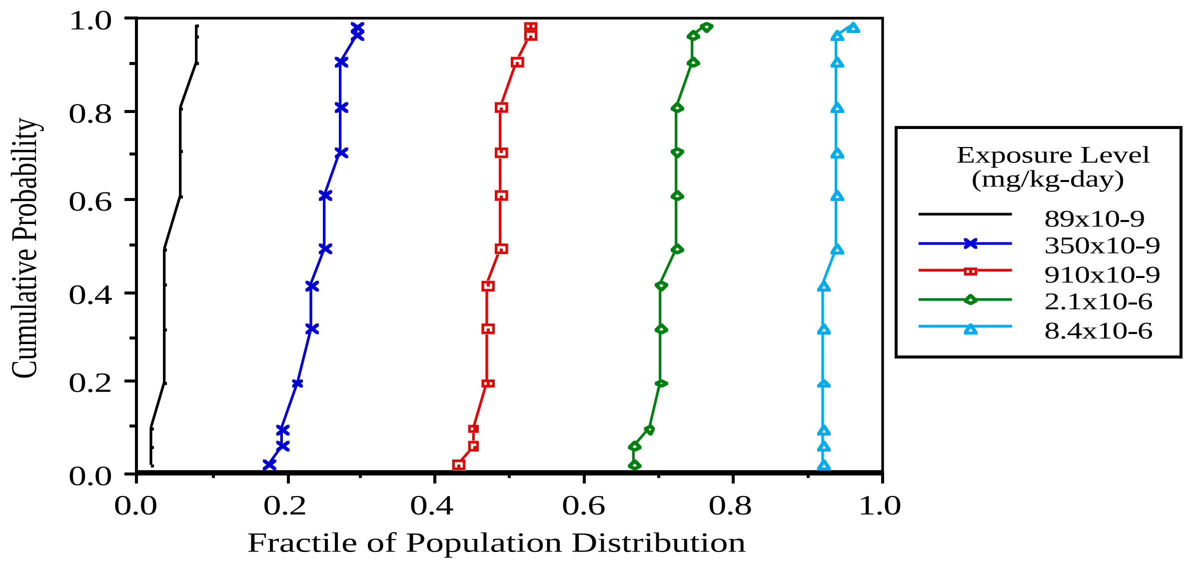


Figure 14. Uncertainty in fractiles for specific exposure levels of the population distribution.

certain value. For example, consider the probability distribution for the fractile of the population with an exposure of less than or equal to 2.1×10^{-6} mg/kg/day. The fifth percentile of this uncertainty distribution corresponds to a fractile of 0.67, while the 95th percentile corresponds to a fractile of 0.73. Thus, with 90 percent confidence, between 67 and 73 percent of the population have an exposure level of less than or equal to 2.1×10^{-6} mg/kg/day. An alternative interpretation is that there is a 95 percent probability that at least 67 percent of the population has an exposure level of less than or equal to 2.1×10^{-6} mg/kg/day.

The type of interpretation would depend on the policy question. If we are interested, for example, in the characteristics of a highly exposed individual, we might be more interested in the percentage of the population facing a higher exposure level such as 8.4×10^{-6} mg/kg/day. In the example of Figure 14, between 92 and 96 percent of the population would be expected to have an exposure this high, with a 98 percent confidence level. Conversely, between four and eight percent of the population would have an exposure greater or equal to this level. Whether this is an acceptable proportion of the population would have to be weighed with other factors, such as regulatory options for reducing exposures, their cost, and equity issues related to subpopulations subject to exposures.

5.3.4.3 A Graphical Summary of Results

In the previous two sections, uncertainties in the exposure levels for specific fractiles of the population, and uncertainties in the fractiles of the frequency distribution for a given level of exposure, have been presented using cumulative distribution functions. In this section, an alternate presentation format is considered.

An alternative approach for presenting the modeling results is illustrated in Figure 15. The figure is based on a simple model in which the simulated quantity E is a function of a variable quantity V and an uncertain quantity U. The variable quantity V is lognormally distributed, and the uncertain quantity U is normally distributed. In this simple example, the uncertainty is assumed to be +100 percent correlated for all individuals. Thus, the resulting estimate of E can be represented by a set of cdf's for the frequency distributions of variability in E, each based on an unambiguous percentile of the uncertainty distribution.

For selected percentiles of the frequency distributions, uncertainty distributions for exposure levels can be constructed. One way to represent these distributions graphically is through tukey plots. These are a convenient representation to use in cases of two-dimensional uncertainty or variability (Morgan and Henrion, 1990, p. 241). In Figure 15, tukey boxes (or "error bars") are shown graphically for five fractiles of the variability distribution. This type of graphical tool provides a method for communication information about both variability and uncertainty. While there have been no studies to determine if this is in fact the best approach for communicating two-dimensional results, discussions with several analysts and presentations reveal that it is at least a format that is understandable to many.

In the more detailed example described previously, uncertainties in both fractiles and exposure levels have been estimated and reported as probability distributions in Figures 13 and 14. A total of 10 distributions are shown in the two figures. These distributions can be represented by tukey plots. The results presented in Figure 12 are presented in Figure 16 with the alternative format.

The dark line in Figure 16 is the frequency distribution for a simulation based solely on variability, without regard to uncertainty. The dotted lines represent the approximate outer boundary of the simulation results given in Figure 12. The horizontal error bars represent the probability distributions for uncertainty in exposure levels given in Figure 13. The vertical error bars represent the probability distributions for uncertainty in fractiles given in Figure 14. The

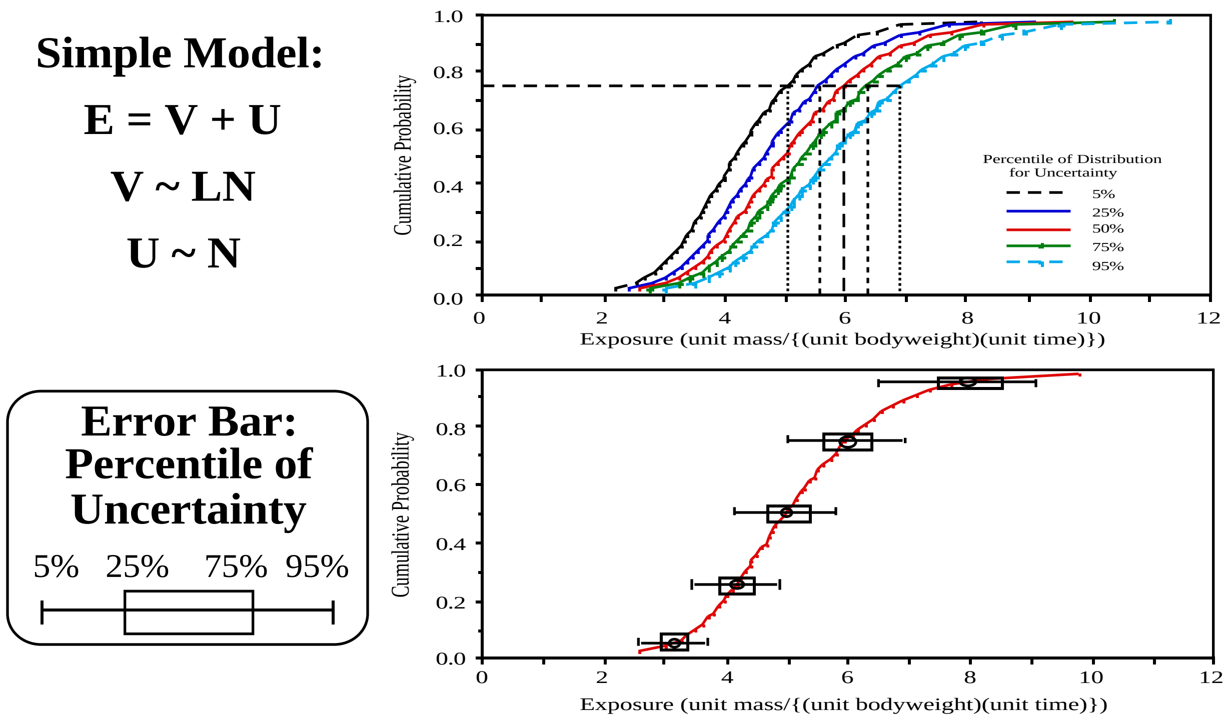


Figure 15. Use of error bars to represent uncertainty in exposure levels for specific fractiles of the population distribution: A simple example.

tukey bars illustrate the general features of both the range and shape of the uncertainties. For example, the ranges of uncertainty in exposure level become wider with increasing percentiles of the variability distribution. For the 95th percentile of the population, the uncertainty in exposure level is positively skewed, as indicated by the longer tail of the tukey bar between the 75th and 95th percentiles of the uncertainty compared to the 5th and 25th percentiles.

5.3.4.4 Relative Effects of Variability and Uncertainty

The relative magnitude of uncertainties and variabilities may differ from one problem to another. To illustrate the interactions between varying magnitudes of uncertainty and variability, an alternative case study was conducted with a higher magnitude of uncertainty. This case study assumed that the standard deviation of the uncertainty in intake rate per unit body weight and the concentration was 30 percent, instead of 10 percent, of the nominal value for each individual, and that the uncertainty in exposure duration ranged from 0 to 10 years, instead of 0 to 5 years. The effect of these assumptions about increased uncertainties on model results is shown in Figure 17.

The increased uncertainty has several consequences. One is that the overall distributions for variability have more variance and significantly longer tails than the base case. This is a direct result of the increased range of uncertainties and the assumption regarding independence of uncertainties from one individual to another. The latter factor leads to increased interactions involving re-ranking of individuals. The maximum exposure levels from the increased uncertainty case study exceed $4x10^{-5}$ mg/kg-day, in contrast to maximum values for the base case of less than $1.75x10^{-5}$ mg/kg-day. These higher values are the result of the increased skewness in the assumption regarding exposure duration, as well as of interactions among the increased range of uncertainties for intake rate per unit body weight and concentration.

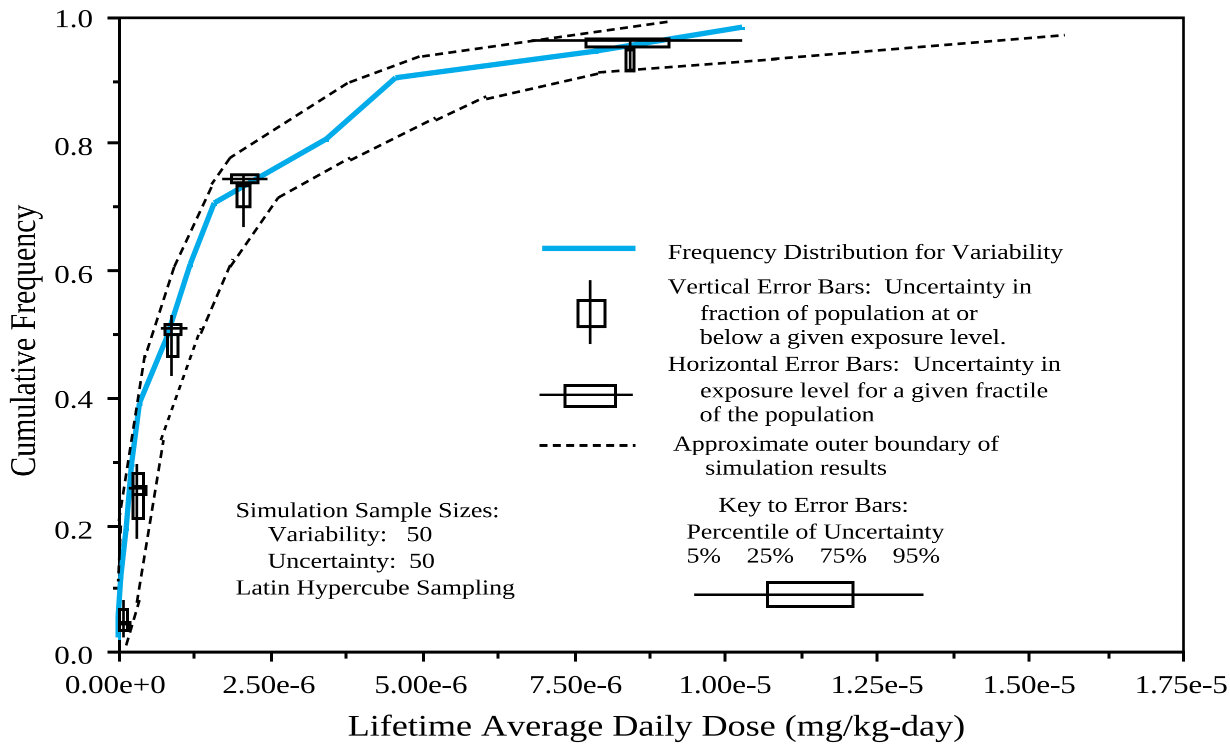


Figure 16. Results of a two-dimensional simulation of variability and uncertainty in a hypothetical exposure model: Representation of uncertainties using error bars.

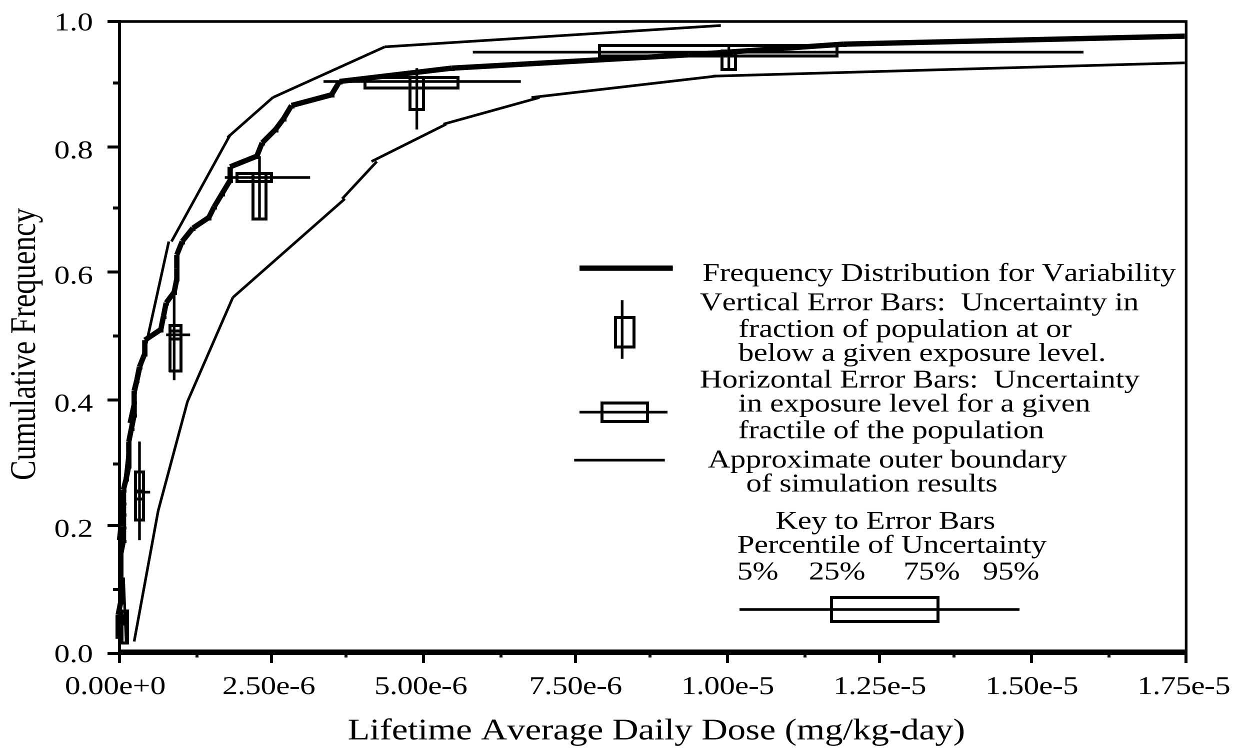


Figure 17. Results of a hypothetical case study with increased uncertainty.

The uncertainties in the exposure levels are larger in variance than for the base case, although the effect is more pronounced for the upper percentiles of the population. For example, the 95th percentile of the population faces a 90 percent confidence range of exposures from $6x10^{-6}$ to $1.6x10^{-5}$ mg/kg-day, in contrast to the range of $7x10^{-6}$ to $1.0x10^{-5}$ mg/kg-day for the base case. For the lower population percentiles, the exposure levels tend to be similar to the base case. For example, the 90 percent confidence range for the 25th percentile of the population is from $2.7x10^{-7}$ to $4.2x10^{-7}$ mg/kg-day in the base case, and $2.9x10^{-7}$ to $5.0x10^{-7}$ mg/kg-day in the increased uncertainty case.

The uncertainties in the fractiles of the population distribution associated with specific exposure levels also tend to be more uncertain than for the base case, indicating that there are increased effects associated with re-ranking of individuals in the population.

5.3.5 Prioritizing Data Needs

The results of the two-dimensional simulation provide distinct information regarding both variability and uncertainty in exposure to a population. The results of the simulation can also be used to identify priorities for future data collection and research. Statistical techniques can be employed to identify and prioritize the key sources of variability and uncertainty in the model. In this case study, rank correlation coefficients were used for this purpose. For each combination of model outputs and inputs, rank correlation coefficients were calculated. These coefficients measure the strength of the linear relationships between the rank values of the samples of the input and output distributions of interest.

5.3.5.1 Identifying Key Sources of Variability

The coefficients for rank correlation between the population estimate of variability in exposure and each of the four variability inputs is shown in Figure 18. The strongest correlation in exposure is with the exposure duration. Variability in concentration is also strongly correlated with the lifetime average daily dose. The correlations for the intake rate per unit body weight and the assumed distribution for lifetime are small. These results indicate that any additional effort to refine the estimates of variability should focus on exposure duration and concentration. The modeling results are not very sensitive to the current assumptions about intake rate per unit body weight and lifetime. As noted earlier, estimates of variability can be refined by stratifying the population under study into more homogeneous subgroups. Thus, these results suggest that there may be sensitive subpopulations characterized by high exposure durations and/or concentrations which merit more detailed study. Of course, the results here are hypothetical, but the types of insights regarding key sources of variability yielded by this methodological approach would be obtained in other studies.

5.3.5.2 Identifying Key Sources of Uncertainty

The evaluation of key sources of uncertainty is not as straight-forward in this case as was the evaluation of key sources of variability. The three uncertain quantities in the model may have varying importance from one individual to another. The rank correlations between the uncertainty in the lifetime average daily dose and the input uncertainties are shown in Figure 19 as a function of the mean exposure level for different individuals in the population.

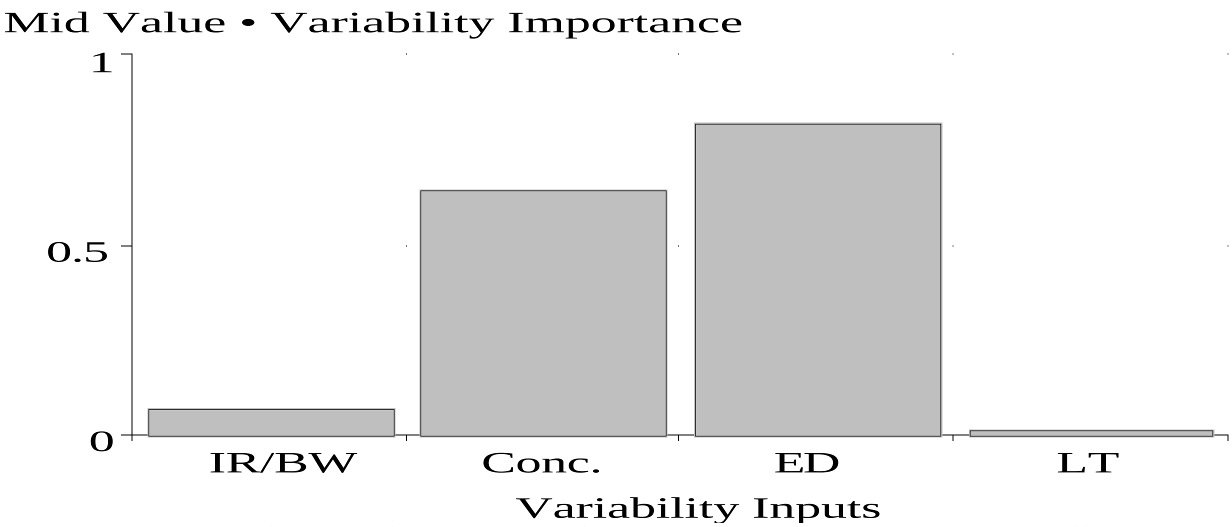


Figure 18. Identification of key sources of variability using rank correlation coefficients.

From Figure 19, it appears that uncertainty in exposure duration tends to have a high rank correlation for individuals with low mean exposure levels (i.e. less than 1.0x10^{-6} mg/kg-day). For these same individuals, the rank correlations with both concentration and exposure duration tend to be significantly lower. This result is more clearly indicated on a logarithmic scale, as shown in Figure 20.

As noted previously, the variance in the uncertainty for exposure duration is the same for all individuals. However, the variance in the uncertainty for concentration and ingestion rate per unit body weight were assumed to be proportional to the values of those parameters sampled from the variability frequency distributions. Individuals with low mean exposures are those who tend to have low intake rates per unit body weight or who tend to face low concentration levels. Therefore, for these individuals, there is less uncertainty, on an absolute basis, in their intake rates and concentrations. Thus, the uncertainty in exposure duration emerges as the key uncertainty for these individuals. For more highly exposed individuals, the results suggest that

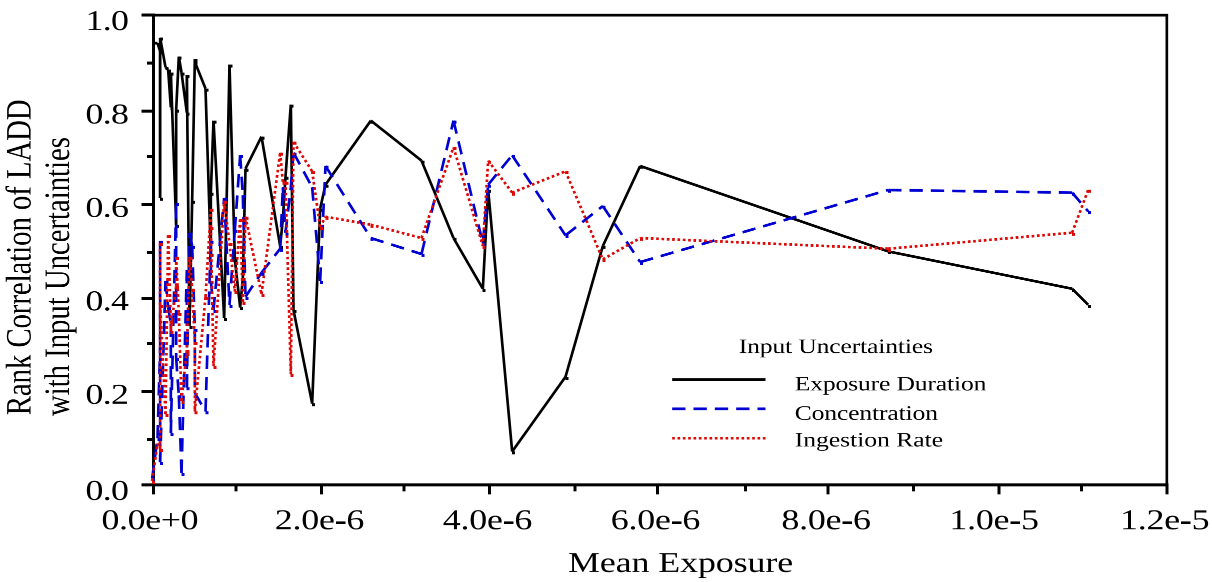


Figure 19. Rank correlations of LADD with model input uncertainties as a function of the mean LADD for the different individuals in the population.

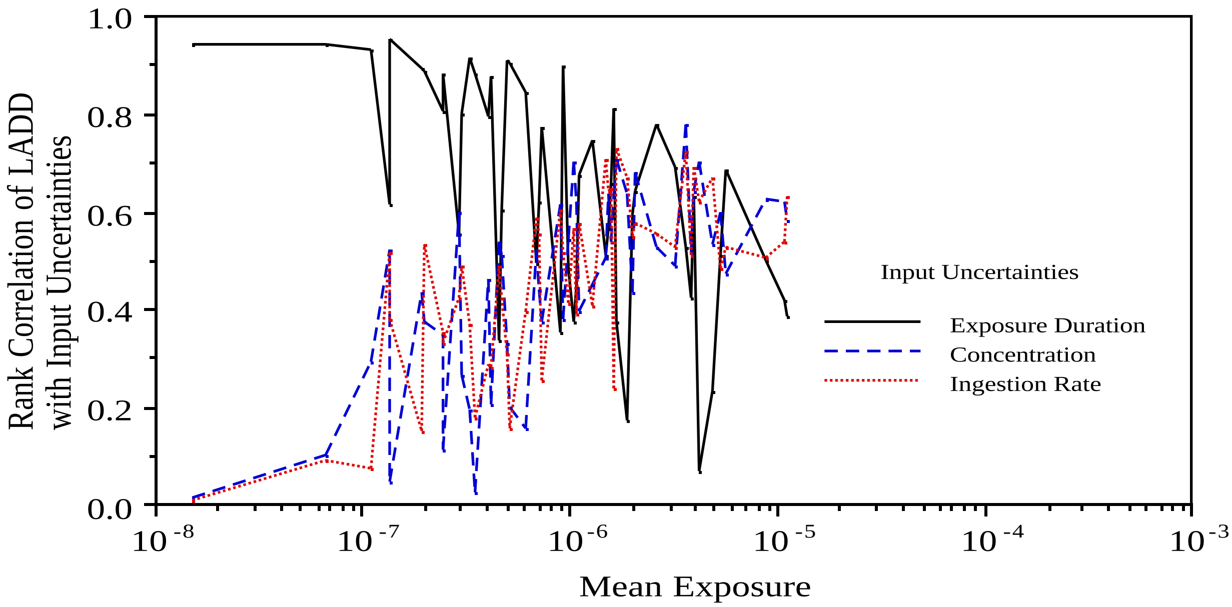


Figure 20. Rank correlations of LADD with model input uncertainties as a function of the mean LADD for the different individuals in the population (logarithmic scale).

all three sources of uncertainty are equally important.

The scatter in the curves for rank correlations associated with a specific input variable are the result of the interactions among variabilities and uncertainties. There are different ways that any given individual could have a specific exposure level. For example, one individual might have a high intake rate, but drink water from a well with a low concentration. Another individual might have a lower intake rate, but ingest more highly contaminated water. Thus, their exposure levels might be similar for different reasons. The alternative possibilities for pairing different values of the input uncertainties to get similar results is are manifested by the varying correlations for similarly exposed individuals.

The implications of this particular case study are that for the most highly exposed individuals, all sources of uncertainty need to be carefully considered. However, the results may differ depending on the magnitude of the input uncertainties. For example, consider the increased uncertainties case discussed previously. The rank correlation of each uncertain input parameter with the lifetime average daily dose is shown in Figure 21. The results for this case indicate that for the most highly exposed individuals in the population, uncertainty regarding intake rate per unit body weight and contaminant concentration are the key contributors to uncertainty. Thus, additional research in this case should be focused on reducing the uncertainty associated with measurement and survey techniques.

5.4 Implications of Two-Dimensional Approach

The two-dimensional approach to Monte Carlo simulation of variability and uncertainty allows the conceptual differences between these two to be properly modeled. The methodology facilitates more technically correct thinking about policy questions related to characterizing exposed populations, evaluated different exposure levels, prioritizing model directions (i.e. needs for stratification or identification of sensitive subpopulations), and prioritization of research needs. The principle drawback of this approach is its computational intensity. A key challenge for this method is the disaggregation of input data into separate variable and uncertain components.

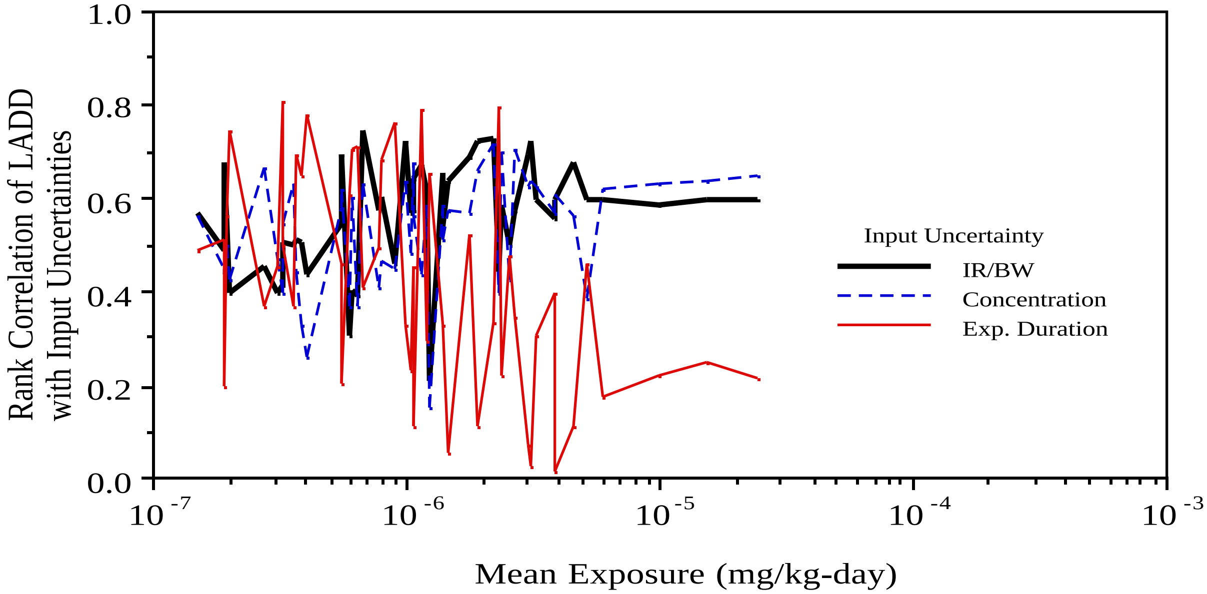


Figure 21. Rank correlations of LADD with model input uncertainties as a function of the mean LADD for the different individuals in the population: Increased Uncertainty Case.

The selection of sample size in the two-dimensional approach poses a unique challenge. The number of samples obtained for the frequency distributions representing variability depend on the sample size m, whereas the number of samples obtained for probability distributions (i.e., uncertainty in exposure for a given population fractile, and uncertainty in population fractile for a given exposure level) depend directly on the uncertainty sample size n. However, interactions among variabilities and uncertainties together influence the results. This is true in the cases where uncertainties are statistically independent from one individual to another. Thus, as illustrated in Figure 4, the rank ordering of individuals for any single estimate of the variability frequency distributions depends on the realized samples from both the variability and frequency inputs. This interaction suggests that the sample sizes for the two dimensions should be comparable.

Although the sample size of the estimated distributions for the uncertainty in population fractile associated with a given exposure level depends directly on the uncertainty sample size, the resolution of the fractile estimates depends on the variability sample size. Thus, if estimates of uncertainty in the fractile distributions are the key objective, it may be desirable to set the variability sample size to be larger than the uncertainty sample size.

The hypothetical case study has illustrated several ways of interpreting modeling results. Both population and individual descriptors of exposure are possible with this approach. For example, population questions often focus on the proportion of the population subject to exposure at or above a given level. As described previously, traditional approaches to Monte Carlo simulation do not properly answer this question, except in the rare cases when uncertainties are not important. When uncertainties do matter, it is necessary to provide an indication of the confidence associated with any estimate of the fractile of the population with an exposure less than or equal to some value.

The implications of the uncertainties regarding variability will depend on the specific problem. If there is, say, a 95 percent probability that a 95 percent of the population has exposures less than any level of concern, then regulatory action may not be required. If, however, there is a 95 percent probability that 10 percent of the population faces an unacceptably high exposure from a health risk standpoint, then regulatory action may be warranted or necessary. These examples represent extreme cases in which decisions about whether regulatory action are needed may be unambiguously clear even in the face of uncertainties. However, there may be cases in which the range of uncertainties encompasses an exposure level of concern. In such cases, additional efforts to obtain better data, and thus to reduce the uncertainties, may be warranted. In this regard, the types of analyses presented in Section 5.3.5 can be performed to identify priorities for additional research and study.

A key advantage of the two-dimensional simulation approach is the proper characterizations of interactions between uncertainty and variability in an assessment. This quantitative approach permits the implications of uncertainty and variability to be concisely summarized in graphical form, as discussed in Section 5.3.4.3. In the context of EPA assessments, such a summary would be consistent with the spirit of the Habicht memo.

While the benefits of the two-dimensional approach to Monte Carlo simulation have been illustrated, a critical issue is whether data are available or can be developed to support such analyses. The answer should be yes, because it is critically important in any type of survey or measurement to quantitatively understand and characterize the precision and accuracy of the techniques. If this type of information is available, as it should be, then it is possible to disaggregate the effects of variability and uncertainty, as previously discussed. If such data are not available, then the characterizations of uncertainty and variability may have to rely, at least initially, on expert judgments regarding subjective probability distributions. It is important to emphasize that subjective assessments can be very useful in helping to identify future data gathering and research priorities, even if such studies are problematic from a regulatory perspective. However, defensible approaches to expert elicitation are possible and can be (and have been) employed in the context of EPA studies.

6.0 DISCUSSION

The previous section has focused on a technical discussion of issues specific to the analysis of variability and uncertainty in exposure assessment. This section addresses institutional issues regarding needs and means for uncertainty analysis within EPA.

6.1 Educational Needs

Although there are many motivating forces for performing uncertainty analyses within EPA, these have not been sufficient to motivate the use of quantitative approaches to uncertainty analysis even in many areas where such an approach is warranted. Many believe that uncertainty analysis poses an additional drain on scarce resources of time. While uncertainty analysis does require an initial up-front investment in terms of effort to characterize uncertainties in model inputs and evaluate the results of probabilistic simulations, uncertainty analysis can reduce the overall time involved in an analysis by helping to identify and prioritize key uncertainties. This information can then be used to focus project resources on refining the analysis as necessary, resulting in potentially net savings of time and other resources.

However, more common criticisms to quantitative uncertainty analysis, and resistance to it on the part of some, are often based on a lack of understanding about what uncertainty analysis is, how it should be applied, and how it should be interpreted. Examples of misconceptions are found even in EPA guidance documents, as discussed in Section 2.1.5.

To facilitate the diffusion of quantitative uncertainty analysis within EPA, it is necessary to convey theory and examples regarding the fundamentals to those analysts who would be involved in such analyses. Furthermore, decision-makers who would be required to interpret such analyses would also require a grounding in the decision analytical aspects of uncertainty. These needs could be met through such means as short courses in fundamentals, as well as workshops in problem-specific areas (e.g., exposure assessment, benefit/cost analysis). Because methods for quantitative approaches to evaluating uncertainty and variability are evolving, an iterative approach to information exchange may be required. This would be facilitated by periodic updating of short courses, or the development of more advanced courses or workshops.

6.2 Data Needs

As noted previously, quantitative uncertainty analysis can be used to prioritize data collection. In addition, Monte Carlo techniques can be used to help design surveys. For example, questions about how many data points are needed for a given statistical precision can be evaluated using Monte Carlo simulation or related techniques.

While uncertainty analysis techniques can help prioritize and design data collection efforts, they also require information to characterize uncertainties. Quantitative uncertainty analysis may involve simulation of uncertainty in quantities that are directly measured, or estimation of uncertainty in quantities calculated in a model as a function of uncertain inputs. Some typical data needs might include paired data sets for quantities that are statistically dependent, sufficient and representative data to characterize distributions, quantitative characterization of errors in measurement techniques, and data unique to specific problems.

In reality, such data are not always available. In such cases, a number of techniques previously discussed may be employed to evaluate alternative hypotheses or assumptions regarding uncertainties. Critics of quantitative uncertainty analysis often characterize such efforts as "making up data." However, it is more appropriate and defensible to make a good faith effort to characterize what is known about uncertainties, even if some or all of the effort involves professional judgment, than to completely side-step and ignore uncertainties in a point-estimate analysis. Furthermore, what many critics of Monte Carlo techniques fail to either appreciate or acknowledge is the considerable amount of judgment that enters into the selection of single numbers for quantities in a deterministic analysis.

Currently, many efforts at dealing with uncertainty in EPA contexts are based on a post-hoc qualitative laundry-list approach to enumerating potential sources of uncertainty, without any quantitative evaluation. In the more rigorous approach to dealing with uncertainties quantitatively, uncertainty needs to be an integral part of an analysis. Collection of information and data to support an uncertainty analysis should begin early in problem analysis. Identification of data needs and evaluation of data gaps is an integral part of such an effort. Approaches to dealing with data gaps include sensitivity analysis and/or uncertainty analysis based on professional judgments. While such judgments may not always provide a comfortable basis for rule-making, they can be a useful basis for identifying data collection priorities.

As pressure mounts to do more quantitative uncertainty analysis at EPA and elsewhere, more attention to the development of appropriate data sets will be necessary. The focus for data collection efforts can be identified and prioritized based on preliminary screening studies of uncertainties. Thus, uncertainty analysis provides a useful tool for iteratively identifying the data needs required to support future assessments.

6.3 Other Needs

In order to bolster the credibility and facility of quantitative approaches to uncertainty analysis within EPA, a number of ingredients are required. These include exemplary case studies which demonstrate methodological approaches to uncertainty analysis in various problem domains, software tools that facilitate the types of analyses that are useful in specific problem domains, peer review of uncertainty analyses, and handbook-type publications specific to the needs of EPA programs.

While training in uncertainty analysis can be useful for many analysts, the development of exemplary case studies featuring analyses of real problems can provide perhaps a more meaningful introduction to methodology and the types of insights obtained from it. For example, it would be useful to develop a detailed exposure assessment case study for an actual problem in which variability and uncertainty are evaluated using the two-dimensional approach to Monte Carlo simulation.

One limitation to the use of quantitative uncertainty analysis is the availability of user-friendly software tools with the capabilities required for specific applications. A number of commercially available software tools are available, including @risk , Crystal Ball, and Demos, and all of these are used to some extent either at EPA or by EPA contractors. Furthermore, a number of software tools developed within EPA are available, with the most applicable one being MOUSE. For many applications, the capabilities of these software tools may be sufficient. In some cases, problem-specific modeling needs may motivate the development of new software tools. For example, if two-dimensional simulation of variability and uncertainty were to become an essential part of exposure assessment modeling, then a generalizable modeling environment with such a capability would be useful.

The use of probabilistic approaches to uncertainty analysis is in itself a field of study which transcends specific problem domains. Thus, the appropriate peer review of uncertainty analyses may pose a challenge. Scientists trained in a specific problem domain who may be eminent experts in a specific field may nonetheless not be optimally suited to reviewing and commenting on a quantitative uncertainty analysis featuring Monte Carlo techniques. Thus, attention should be given to the selection of review panels that include persons with appropriate background relevant to uncertainty analysis.

Methodological handbooks dealing with uncertainty analysis and issues specific to environmental problem domains would prove to be useful to EPA analysts. Such publications would summarize approaches for dealing with various types of problems in uncertainty analysis. However, because the approaches to characterizing uncertainties depend on the type of analysis being done, the level of detail, and the availability of data, among other factors, the development of any strict and binding guidelines should be avoided.

6.4 Separating Analysis and Decision-Making

The use of quantitative uncertainty analysis promotes the separation of assessment and decision-making functions. Uncertainty analyses performed to support decision-making do not have to be biased based on policy directives, such as regarding the use of conservative assumptions in some model parameters but not others, or by focusing on only certain characteristics of a frequency or uncertainty distribution to the exclusion of others. Instead, uncertainty analysis permits an evaluation of the full range of possible outcomes, leaving to a decision-maker the policy questions of how to interpret scientific information in the face of potentially conflicting economic or social objectives. Furthermore, the estimates of uncertainty reflect the state of the underlying science better than point estimates, and concisely convey to a decision maker information about the degree of confidence that can be placed in any point estimate. It is the risk manager, not the analyst, who ultimately should make decisions regarding what point estimates to use. Thus, the basis of the decision is more explicit than one based on a point-estimate analysis with buried policy assumptions.

7.0 CONCLUSIONS

Failure to fully consider the implications of uncertainties in environmental problems often yields the wrong answer and gives decision-makers a deceptive and incorrect sense of confidence about numbers. To properly characterize uncertainties in environmental analyses, results should be presented in the form of distributions wherever possible. Although quantitative analysis requires more critical attention to assumptions and inputs to an analysis, the result is a more defensible and meaningful estimate of both the range and likelihood of possible outcomes. Uncertainty analysis also involves a dynamic process of screening and iteration, in which priorities for research and data collection can be evaluated quantitatively. In this regard, uncertainty analysis provides a quantitative basis for caveating the results of a study, and for simultaneously identifying ways to improve the study in the future.

There is increasing discussion within EPA regarding the need for more rigorous approaches to dealing with uncertainties in analyses, and more complete communication of uncertainties to decision-makers. Such an approach is desirable, because it more explicitly separates the role of the analyst from that of the decision-maker, and provides the decision maker with the information required to evaluate the degree of confidence that can be placed in the results of an assessment.

While quantitative approaches to uncertainty analysis do not in themselves resolve debates, they do help focus them on both the concepts and quantities that really matter. In this report, an approach to evaluating the implications of variability and uncertainty in exposure assessments was developed based on two-dimensional Monte Carlo simulation. Uncertainty is a lack of knowledge of some type, whereas variability is a real difference in values among different individuals in a population. These important conceptual differences are muddled in traditional approaches to exposure assessment. The two-dimensional approach was applied to an example case study featuring a hypothetical exposure assessment. The results of the case study illustrated the benefits to separately simulating variability and uncertainty in terms of technical rigor, communication of results to decision-makers, identification of the characteristics of sensitive subpopulations, and identification of data collection needs.

Quantitative approaches to uncertainty analysis generate insights that help both analysts and decision makers ask the right questions about environmental problems.

8.0 NOMENCLATURE

BW	=	Body weight, kg
C	=	Concentration, mg/l
ED	=	Exposure Duration, days
IR	=	Intake Rate, l/day
LADD	=	Lifetime Average Daily Dose, mg/kg-day
LT	=	Lifetime, days
m	=	Sample size for simulation of variabilities
M	=	Number of quantities in a simulation which are variabilities
n	=	Sample size for simulation of uncertainties
N	=	Number of quantities in a simulation which are uncertainties
U	=	vector of uncertain quantities $\{U_1, U_2, U_3, ..., U_k\}$
$\mathbf{u_i}$ (or $\mathbf{U}_i$)	=	a vector (or matrix) of realizations (samples) for the uncertain quantities **U**, for a specific individual i.
$u_{i,j}$	=	one realization (or sample) j for the uncertainty quantity U_i
V	=	vector of variable (empirical) quantities $\{V_1, V_2, V_3, ..., V_j\}$
$\mathbf{v}_i$ (or $\mathbf{V}_i$)	=	vector of particular values $[v_{1,i}, v_{2,i}, v_{3,i}, ..., v_{j,i}]$ of **V** for a specific individual, where there are j (or m) variable quantities and one sample of each for an individual i.

9.0 LITERATURE CITED

Ang, A. H-S., and W.H.Tang (1984). *Probability Concepts in Engineering Planning and Design, Volume 2: Decision, Risk, and Reliability*. John Wiley and Sons, New York.

Alberts, M.T., and F.O. Weyman (1991), "Characterizing Uncertainty in the Risk Assessment Process," presented at the Electric Power Research Institute Conference on Managing Hazardous Air Pollutants: State of the Art, Washington, DC, November 11-14, 1991.

Alberts, M.T., and J.D. Rouge, "Stochastic Health Risk Assessments: Case Studies in California AB 2588 Air Toxics Hot Spots Alternative Analyses," Paper 92-95.03 presented at the 85th Annual Meeting and Exhibition of the Air and Waste Management Association, Kansas City, Missouri, June 21-26.

Apostolakis, G. (1990), "The Concept of Probability in Safety Assessments of Technological Systems," *Science*, Volume 250, December 7, pp. 1359-1364.

Bogen, K.T., and R.C. Spear (1987), "Integrating Uncertainty and Interindividual Variability in Environmental Risk Assessment," *Risk Analysis,* 7(4):427-436.

Bogen, K.T. (1990), *Uncertainty in Environmental Health Risk Assessment,* Garland Publishing, Inc., New York.

Constantinou, E., C. Seigneur, and T. Permutt (1992), "Uncertainty Analysis of Health Risk Estimates: Application to Power Plant Emissions," Paper 92-95.01 presented at the 85th Annual Meeting and Exhibition of the Air and Waste Management Association, Kansas City, Missouri, June 21-26.

Copeland, T.L, A.M. Holbrow, K. Conner, and D.J. Paustenbach (1992), "Use of Quantitative Uncertainty Analysis in Air Toxics Risk Assessment," Paper 92-166.04 presented at the 85th Annual Meeting and Exhibition of the Air and Waste Management Association, Kansas City, Missouri, June 21-26.

Dean, J.D., et al. (1987), Exposure Assessment for the Pesticide Aldecarb in Florida, USA, EPA/600/D-87/190, Prepard by Woodward-Clyde Consultants for the U.S. Environmental Protection Agency, Athens, GA, June.

Eschenroeder, A.Q., and E.J. Faeder, "A Monte Carlo Analysis of Health Risks from PCB-Contaminated Mineral Oil Transformer Fires," *Risk Analysis*, 8(2):291-297 (1988).

Edwards, A.L. (1984). *An Introduction to Linear Regression and Correlation*, Second Edition. W.H. Freeman and Company. New York.

Finkel, A.M. (1990), "Confronting Uncertainty in Risk Management: A Guide for Decision-Makers," Center for Risk Management, Resources for the Future, Washington, DC, January.

Frey, H.C. (1991), *Probabilistic Modeling of Innovative Clean Coal Technologies: Implications for Technology Evaluation and Research Planning*, Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA, May.

Frey, H.C., and E.S. Rubin (1991), Development and Application of a Probabilistic Evaluation Method for Advanced Process Technologies," DOE/MC/24248-3015, NTIS DE91002095, Prepared by Carnegie Mellon University for the U.S. Department of Energy Morgantown Energy Technology Center, Morgantown, WV, April.

Frey, H.C., and E.S. Rubin (1992), "Evaluate Uncertainties in Advanced Process Technologies," *Chemical Engineering Progress*, 88(5):63-70 (May)

Habicht, F.H. (1992), "Guidance on Risk Characterization for Risk Managers and Risk Assessors", Memorandum from F.H. Habicht, Deputy Administrator, to Assistant and Regional Administrators, U.S. Environmental Protection Agency, February 26, 1992.

Hattis, D., et al. (1990), "Uncertainties in Pharmocakinetic Modeling for Perchloroethylene: I. Comparison of Model Structure, Parameters, and Predictions for Low-Dose Metabolism Rates for Models Derived by Different Authors," *Risk Analysis*, 10(3):449-458.

Ibrekk, H., and M.G. Morgan (1987), "Graphical Communication of Uncertain Quantities to Nontechnical People," *Risk Analysis*, 7(4):519-529.

Iman, R.L., and W.J. Conover (1982), "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics*, B11(3):311-334.

Iman, R.L., and M.J. Shortencarier (1984), A Fortran 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use with Computer Moels, SAND83-2365, Sandia National Laboratory, Albuquerque, NM, January.

Iman, R.L., M.J. Shortencarier, and J.D. Johnson (1985). A Fortran 77 Program and User's Guide for the Calculation of Partial Correlation and Standardized Regression Coefficients, SAND85-0044, Sandia National Laboratory, Albuquerque, NM, June.

Iman, R.L., and J.C. Helton (1988), "An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models," *Risk Analysis*, 8(1):71-90.

Johnson et al. (1992), "Estimation of Carbon Monoxide Exposures and Associated Carboxyhemoglobin Levels in Denver Residents Using a Probabilistic Version of NEM," Paper 92-145.01 presented at the 85th Annual Meeting and Exhibition of the Air and Waste Management Association, Kansas City, Missouri, June 21-26.

Keeney, R.L, and D. von Winterfeldt (1991), "Eliciting Probabilities from Experts in Complex Technical Problems," *IEEE Transactions on Engineering Management,* 38(3):191-201.

Klee, A.J. (1992), "MOUSE: A Computerized Uncertainty Analysis System Operational Manual," U.S. Environmental Protection Agency, Cincinnati, OH, July.

Li, W-W., M.J. Firth, and R.H. Harris, "Health Risk Reduction Based on Monte Carlo Simulation and Air Dispersion Modeling at a Chemical Processing Facility," Paper 92-149.03 presented at the 85th Annual Meeting and Exhibition of the Air and Waste Management Association, Kansas City, Missouri, June 21-26.

McKay, M.D., W.J. Conover, and R.J. Beckman (1979), "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, 21(2):239-245.

McKone, T.E., and P. Barry Ryan (1989), "Human Exposure to Chemicals Through Food Chains: An Uncertainty Analysis," *Environmental Science and Technology*, 23(9):1154-1163.

McKone, T.E., and K.T. Bogen (1991), "Predicting the Uncertainties in Risk Assessment," *Environmental Science and Technology*, 25(10):1674-1681.

McKone, T.E., and K.T. Bogen (1992), "Uncertainties in Health-Risk Assessment: An Integrated Case Study Based on Tetrachloroethylene in California Groundwater," *Regulatory Toxicology and Pharmacology*, **15**:86-103.

Morgan, M.G., et al. (1985), "Uncertainty in Risk Assessment: A Case Study Involving Sulfur Transport and Health Effects," *Environmental Science and Technology*, 19(9):662-667.

Morgan, M.G., and M. Henrion (1990), *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, New York.

National Research Council (1983), *Risk Assessment in the Federal Government: Managing the Process*, National Academy Press, Washington, DC.

Otway, H., and D. von Winterfeldt (1992), "Expert Judgment in Risk Analysis and Management: Process, Context, and Pitfalls," *Risk Analysis*, 12(1):83-93.

Rish, W.R. (1988), "Approach to Uncertainty in Risk Analysis," ORNL/TM-10746, Prepared by Envirosphere Company for U.S. Environmental Protection Agency, August.

Rish, W.R. and R.J. Marnicio (1988), "Review of Studies Related to Uncertainty in Risk Analysis," ORNL/TM-10776, Prepared by Envirosphere Company for U.S. Environmental Protection Agency, August.

Roseberry, A.M., and D.E. Burmaster (1991), "Lognormal Distributions for Water Intake by Children and Adults," *Risk Analysis*, 12(1):99-104.

Science Advisory Board (1989), "Resolution on the Use of Mathematical Models by EPA for Regulatory Assessment and Decision-Making," EPA-SAB-EEC-89-012, U.S. Environmental Protection Agency, Washington, DC, January.

Science Advisory Board (1992), "An SAB Report: Review of Draft Final Exposure Assessment Guidelines," EPA-SAB-IAQC-92-015, U.S. Environmental Protection Agency, Washington, DC, January.

Shlyakhter, A.I., G. Goodman, and R. Wilson (1992), "Monte Carlo Simulation of Rodent Carcinogenicity Bioassays," *Risk Analysis*, 12(1):73-82.

Smith, W.P. (1992), "Statistical Simulation System," U.S. Environmental Protection Agency, Washington, DC.

Stine, D.D. (1992), "National Research Council Study on Risk Assessment of Hazardous Air Pollutants," Paper 92-95.05 presented at the 85th Annual Meeting and Exhibition of the Air and Waste Management Association, Kansas City, Missouri, June 21-26.

Thompson, K.M., and D.E. Burmaster (1991), "Parametric Distributions for Soil Ingested by Children," *Risk Analysis*, 11(2):339-342.

Thompson, K.M., D.E. Burmaster, and E.A.C. Crouch (1992), "Monte Carlo Techniques for Quantitative Uncertainty Analysis in Public Health Risk Assessments," *Risk Analysis*, 12(1):53-63.

U.S. Environmental Protection Agency (1987), "The Risk Assessment Guidelines of 1986," EPA/600/8-87/045, Washington, DC., August.

U.S. Environmental Protection Agency (1988), "Superfund Exposure Assessment Manual", EPA/540/1-88/001, Office of Emergency and Remedial Response, Washington, DC, April.

U.S. Environmental Protection Agency (1989), "Risk Assessment Guidance for Superfund, Volume 1: Human Health Evaluation Manual (Part A) Interim Final," EPA/540/1-89/002, Office of Emergency and Remedial Response, Washington, DC, December.

U.S. Environmental Protection Agency (1989b), Exposure Factors Handbook, EPA/600/8-89/043, Office of Health and Environmental Assessment, Washington, DC, July.

U.S. Environmental Protection Agency (1990), "Research to Improve Health Risk Assessments (RIHRA) Program", EPA/600/9-90/038, Office of Research and Development, Washington, DC, June.

U.S. Environmental Protection Agency (1992), "Guidelines for Exposure Assessment," Federal Register, 57(104):22888-22938, May 29.

U.S. Environmental Protection Agency (1992b), "Cost and Economic Impact Analysis of Land Disposal Restrictions for Newly Listed Wastes and Contaminated Debris (Phase 1 LDRs) Final Rule," Washington, DC, June 30.

Walentowicz, R. (1984), "Stochastic Processes Applied to Risk Analysis of TCDD Contaminated Soil," Exposure Assessment Group, U.S. Environmental Protection Agency, May 31.

White, P., and J. Schaum (1989), "Use of Monte Carlo Analysis in Exposure and Risk Assessments," Exposure Assessment Group, U.S. Environmental Protection Agency, Washington, DC, June 2.

Whitmore, R.W. (1985), "Methodology for Characterization of Uncertainty in Exposure Assessments," EPA/600/8-85/009, Prepared by Research Triangle Institute for U.S. Environmental Protection Agency, Washington, DC.

Wilson, R., E. Crouch, and L. Zeise (1985), "Uncertainty in Risk Assessment," *Banbury Report 19: Risk Quantitation and Regulatory Policy*, Cold Spring Harbor Laboratory.

Wolff, S.K., et al. (1990), "Selecting Experimental Data for Use in Quantitative Risk Assessment: An Expert-Judgment Approach," *Toxicology and Industrial Health*, 6(2):275-291.

Zankel, K.L., et al (1991), "Multipathway Risk Assessment at a Maryland Utility," presented at the Electric Power Research Institute Conference on Managing Hazardous Air Pollutants: State of the Art, Washington, DC, November 11-14, 1991.