NUREG/CR-4836
SAND87-0871
RX
Printed January 1988

Approaches to Uncertainty Analysis in Probabilistic Risk Assessment

Michael P. Bohn, Timothy A. Wheeler, Gareth W. Parry

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550
for the United States Department of Energy
under Contract DE-AC04-76DP00789

Prepared for
U. S. NUCLEAR REGULATORY COMMISSION

Available from

Superintendent of Documents
U.S. Government Printing Office
Post Office Box 37082
Washington, D.C. 20013-7082

and

National Technical Information Service
Springfield, VA 22161

NUREG/CR-4836
SAND87-0871
RX

APPROACHES TO UNCERTAINTY ANALYSIS IN PROBABILISTIC RISK ASSESSMENT

Michael P. Bohn
Timothy A. Wheeler
Gareth W. Parry*

January 1988

Sandia National Laboratories
Albuquerque, NM 87185
Operated by
Sandia Corporation
for the
U.S. Department of Energy

Prepared for
Division of Reactor and Plant Systems
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555
Under Memorandum of Understanding DOE 40-550-75
NRC FIN A1384

---

*NUS Corporation, 910 Clopper Road, Gaithersburg, Maryland, 20878

ABSTRACT

An integral part of any probabilistic risk assessment (PRA) is the performance of an uncertainty analysis to quantify the uncertainty in the point estimates of the risk measures considered. While a variety of classical methods of uncertainty analysis exist, application of these methods and developing new techniques consistent with existing PRA data bases and the need for expert (subjective) input has been an area of considerable interest since the pioneering Reactor Safety Study (WASH-1400) in 1975. This report presents the results of a critical review of existing methods for performing uncertainty analyses for PRAs, with special emphasis on identifying data base limitations on the various methods. Both classical and Baysian approaches have been examined. This work was funded by the U.S. Nuclear Regulatory Commission in support of its ongoing full-scope PRA of the LaSalle nuclear power station. Thus in addition to the review, this report contains recommendations for a suitable uncertainty analysis methodology for the LaSalle PRA.

CONTENTS

CONTENTS
(Continued)

CONTENTS
(Concluded)

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENT

This work was supported by the Division of Reactor and Plant Systems of the U.S. Nuclear Regulatory Commission. Valuable review and comments were provided by D. M. Rasmuson and M. A. Cunningham of the NRC staff. Within Sandia, R. G. Easterling, F. L. Spencer, and R. L. Iman provided generously of their time in providing background papers, discussion and comment based on their current research activities in uncertainty analysis techniques. We are especially in debt to Jane Booker and Mary Meyer of Los Alamos Laboratory (LASL) who prepared a summary of the LASL research on the use of expert opinion and guidelines for use of such information. This contribution is included as Appendix C. Finally, special thanks are due to Emily L. Preston and Bonnie Haynes for their patience and perseverance in preparation of this manuscript.

I. INTRODUCTION

This report presents the results of a critical review of methods for performing uncertainty analyses for Level 1 probabilistic risk assessments (PRAs). This work was performed in support of the Risk Methods Integration and Evaluation Program (RMIEP)[1] as part of the first full-scope probabilistic risk assessment being performed under the auspices of the US Nuclear Regulatory Commission since the Reactor Safety Study. This work was performed as part of two programs supporting RMIEP: the Dependent Failure Methodology Program and the PRA Methods Development Program, both of which are sponsored by the Division of Reactor System Safety, U.S. Nuclear Regulatory Commission. These programs have the goal of developing new risk assessment methods and integrating both new and existing methods into a uniform procedure for performing an in-depth PRA with consistent levels of analysis for external, internal and dependent failure scenarios.

As part of RMIEP, a PRA of the LaSalle Nuclear Generating Station is being performed. LaSalle is a Mark II BWR/6 commercial power station. The PRA will include an analysis of the uncertainty in the risk estimates calculated for LaSalle. This report presents recommendations for methods to be used in performing the RMIEP uncertainty analysis for LaSalle.

There are three major facets of an uncertainty analysis: (1) identification and characterization of the uncertainties in the inputs, (2) propagation of these uncertainties through the systems models to arrive at a measure of the uncertainty in the output, and (3) identification of the significant sources of uncertainty with respect to the result. In general, we are interested in the quantification of system fault trees and accident sequence equations. The inputs for which uncertainties are generally provided are the parameters which are used to derive estimates of basic event probabilities. The outputs are system unavailabilities, accident sequence frequencies, and core damage frequency.

Another type of input uncertainty that may be addressed is modeling uncertainty. Where it is not clear, for example, which assumptions about success criteria, or the effects on certain components of a particular environmental condition, are appropriate, alternative hypotheses can be explored by making different assumptions.

This report presents a summary of uncertainty analysis methods applicable to risk analyses. In the context of this report, risk refers to estimates of core damage frequency. The methods reviewed will be presented in two parts: methods for propagating the uncertainty in basic event model parameters (and also uncertainty due to alternative modeling hypotheses) through system and accident sequence models, and methods for characterizing the uncertainty in the basic event model parameters and modeling assumptions.

Input uncertainty characterization methods reviewed are:

Plant-Specific Data Analysis
Bootstrapping Method
Bayesian Update Approach
Use of Expert Opinion

The uncertainty propagation methods reviewed are:

Method of Moments
Method of Moments Using Tchebyschev's Inequality
Propagation of Discrete Probability Distributions
Maximus/Bounding Approach
Monte Carlo Methods
Latin Hypercube Sampling Methods

The uncertainty analysis method recommended for RMIEP is the Latin hypercube sampling technique which uses stratified sampling of the basic event probabilities, initiating event frequencies, and alternative modeling hypotheses for quantification of the top event uncertainty. Uncertainty in basic event probabilities is characterized using probability distributions derived in three different ways: when sufficient plant-specific data on individual components or component types exist, a Bayesian analysis using non-informative priors may be used; Bayesian updating of generic parameter estimates may be used for components for which only limited plant-specific data exist; expert opinion is used to characterize parameter uncertainty when data are limited or nonexistent. Expert opinion is, of necessity, the only method available for the quantification of uncertainty with respect to alternative modeling hypothesis. Expert opinion is used generally to develop uncertainty representations from literature surveys, Delphi surveys, and non-nuclear experience.

The approach recommended for RMIEP and discussed in this report is intended to cover all aspects of the uncertainty analysis of accident sequence frequencies and core damage frequency. Very similar (but somewhat generalized) techniques will be used for the uncertainty analysis of containment failures, source term, and health and consequence modeling calculations. These aspects are being considered in the NRC-sponsored Probabilistic Risk and Uncertainty Evaluation Program (PRUEP).[2] Appendix A presents a discussion of the interface between these two sets of calculations, and it is shown there that the data and modeling techniques used in the accident sequence/core melt uncertainty analysis are directly compatible with the accident progression and health and economic modeling process.

II. UNCERTAINTY METHODS FOR PROBABILISTIC RISK ASSESSMENT

In this section, methods used in PRA analyses for uncertainty are briefly reviewed. Top event uncertainty methods are treated first, then component parameter uncertainty specification techniques will be considered.

II.1 A Perspective on Uncertainty Analysis

Before discussing methods for performing uncertainty analyses, it is worthwhile to examine the goals of performing such an analysis, and to consider the sources of uncertainty we are trying to characterize. Clearly, the goal is to assess uncertainty in the estimate of the frequency and sensitivity measures of some top event given a characterization of the uncertainty in the input (independent) variables or assumptions. In the context of a PRA, the analyst is interested in assessing the uncertainty in computed risk measures (such as accident sequence frequency, core melt frequency, and plant damage state frequency) as a function of uncertainties in the interpretation of data and the parameter estimation process as well as uncertainties inherent in the modeling process itself. The ultimate goal, however, is to be able to make meaningful comparisons between measures of risk so that robust decisions regarding the margin of safety and allocation of resources to increase safety can be made. A crucial aspect of selecting any method of uncertainty analysis is to identify the comparisons to be made, and to assure that the selected uncertainty analysis method will provide the information needed to make those comparisons.

II.1.1 Uncertainty Analysis Goals

In the context of a PRA for a nuclear power plant, we ultimately wish to make the following comparisons and assessments after the PRA is completed:

1. Compare one plant's risk measures with those of other plants to assess relative level of safety.

2. Compare relative risk contributions due to the various sources of risk modeled in the PRA (random initiating events, fires, earthquakes, etc.) to identify any particular vulnerability which may exist at the plant in question.

3. Compare health and economic consequences of hypothesized accidents at the plant with similar consequence measures due to other sources of risk (e.g., aircraft crashes, chemical releases, etc.) to which the public is unavoidably subjected.

4. Compute both "best estimate" and "worst case" measures of risk and consequence to allow informed societal decisions as to "acceptable" level of risk to be made.

5. Compare the contributions of different component failures to the plant risk measures to determine those components for which an increase in reliability would result in the greatest reduction in plant risk and, thus, provide a basis for an effective allocation of resources.

6. Compare the "best estimate" accident-induced loads (e.g., pressures, thermal forces, earthquake ground accelerations, floor spectra, etc.) with the "conservative" deterministic loads used in the design of the plant to assess the "margin of safety" of the various components, structures, and safety systems and to identify any areas where changes in design procedures could result in a more uniform level of "margin of safety" throughout the plant.

Thus the goal of performing an uncertainty analysis is to qualify the conclusions made as a result of point estimate evaluations, so as to identify those conclusions which are robust when considering uncertainties in the analysis. It should also be used to identify where improving the state of knowledge can lead to maximum benefit with respect to an accurate assessment of risk.

In particular, an uncertainty analysis should identify those components or models whose uncertainty is "driving" the overall uncertainty in the plant risk measures, so as to identify any components or models whose uncertainty is preventing any desired comparisons (as discussed above) to be made, and hence identify those components or models for which a reduction in uncertainty (even without a change in our perception of their point estimate) would result in upgrading our decision-making capability.

II.1.2 Uncertainty and Physical Variability

It is important to distinguish between the concepts of uncertainty and variability. The nature of the events considered in a PRA (such as initiating events, component failures, operator actions, etc.) is such that they are treated as being random processes, and modeled through the use of probabilistic models. It is this use of probability which gives the PRA its name.

Sources of random variability are incorporated directly in the PRA models. As examples, the random nature of component failure times is incorporated by assuming a constant failure rate model for those components. The randomness of operator recovery times is incorporated by establishing a relationship between time and the probability of recovery. The randomness in the magnitudes of earthquakes in a given seismic region is

incorporated through the use of a relationship expressing frequency of occurrence as a function of magnitude. These sources of variability then directly affect the evaluation of the frequency of the events of interest, such as core melt frequency.

However, due to lack of data, or a lack of detailed understanding of the physical phenomena being modeled, the relationships that are used to describe the variability are not known precisely. This can be reflected in a lack of precision in the value of a component failure rate, or in the provision of alternative mathematical formalisms. This lack of knowledge is the uncertainty that is of interest here, and which leads to the lack of precision in the predictions of the PRA. An increased level of knowledge will not change the fact that a PRA is a probabilistic model, but it will give greater confidence in the predictions of that model.

II.1.3 The Representation of Uncertainty

There are two basic types of uncertainty to be addressed: the first is that associated with the values of the parameters of the input models, the second is that associated with the possibility of alternate modeling hypotheses.

In the case of parameter values that are estimated on the basis of data, there are two basic statistical approaches; the so-called classical approach, and the Bayesian approach. Given sufficient data, they both produce numerically compatible results, but due to differences of interpretation, propagation of uncertainty measures in the Bayesian approach is easier than in the classical approach.[3]

(a) Classical Statistical Measures of Uncertainty

For parameters that are assumed to have the same constant value for all members of the population being sampled, we can compute a statistical confidence interval, which is an interval about the computed parameter value that is "reasonable" considering the data from which the parameter estimate was derived. For example, consider a component with a history of x failures in n demands. Assuming that the demand failure rate is constant and that the Binomial distribution is an appropriate model for the occurrence of such failures, the point estimate of the demand failure rate is given by

$$\tilde{p} = \frac{x}{n} \qquad (1)$$

The cumulative binomial distribution function is given by

$$F(x_1,n,p) = \sum_{x=0}^{x_1} \binom{n}{x} p^x (1-p)^{n-x} \qquad (2)$$

which gives the probability of $X < x_1$ for a known value of p. From the component failure history, we know the number of failures x and number of demands n, so we can solve equation (2) for values of p such that the observed value of x lies in the central region of the distribution, i.e., would be likely to have been observed. Hence, if we sought values of p which would imply the measured value of (x,n) to lie in the central 95 percent of the distribution, we would solve the pair of equations

$$\sum_{0}^{x} \binom{n}{x} p_1^x (1-p_1)^{n-x} = 0.025$$

$$\sum_{0}^{x} \binom{n}{x} p_2^x (1-p_2)^{n-x} = 0.975 \qquad (3)$$

and obtain two values, p_1, p_2, between which we would (loosely speaking) expect the true value of p to lie with 95 percent confidence.*

Figure 1 gives the solution to this pair of equations for various values of x and n. The number attached to each curve is also n, the number of demands. The ordinate gives the 95 percent confidence intervals for any value of x/n. Similar curves could be constructed for any other desired level of confidence.

---

*More precisely, a $(1-\alpha)$ percent confidence interval is defined by (Ref. 4), "If, in a series of very many repeated experiments an interval such as the one calculated were obtained, we would in the long run be correct $(1-\alpha)$ 100 percent of the time in claiming that p is located in the interval."

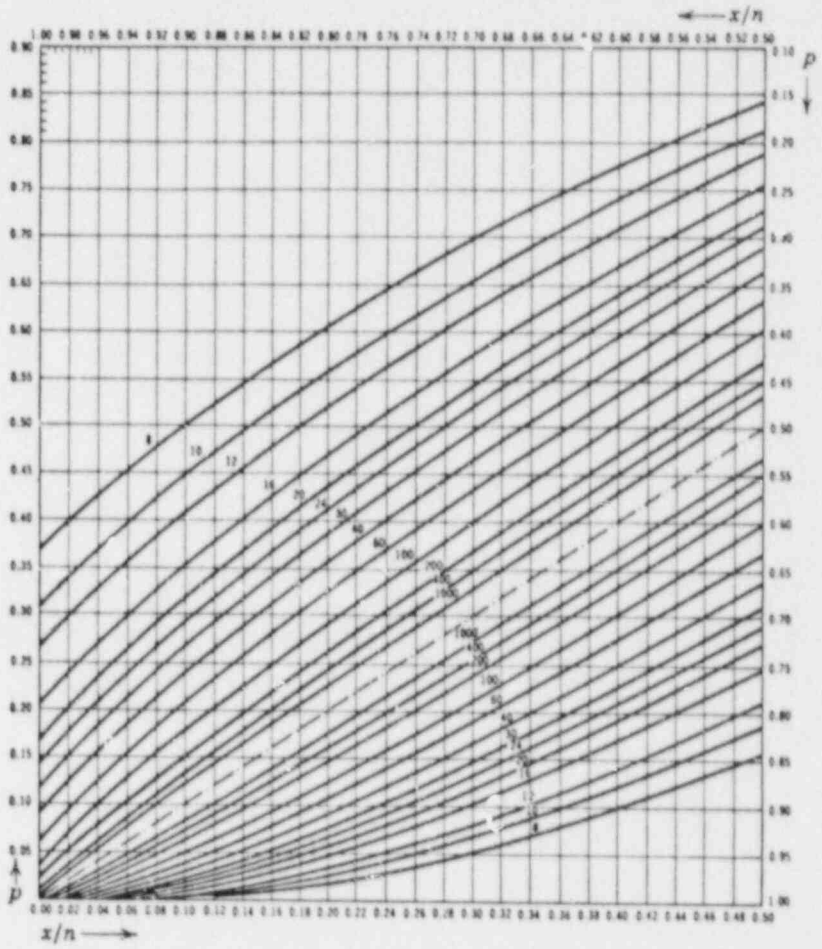

Figure 1. Chart Providing 95 Percent Confidence Limits for p in Binomial Sampling, Given a Sample Fraction x/n. Different curves represent different sample sizes n. Values on ordinate represent 95 percent confidence bounds on p. (Reference 4).

It is important to note from this figure that as n becomes large, the confidence limits approach the point estimate $p = x/n$. Thus, by accumulating more and more data, we can know the value of p as closely as desired (with 95 percent confidence). Thus, the confidence interval is a measure of the sampling uncertainty in determining a constant parameter from data.

Note also that p_1 and p_2 are not points on any distribution, for the failure rate is, by definition, a constant and not a random variable, and thus no distribution can be attached to it. In particular, it is not valid to assume p_1, p_2 to be the 5th and 95th percentiles of some (unknown) distribution for p.

Even though, as discussed earlier, the effect of variability arising from the inherent randomness of the processes modeled is generally incorporated directly into the frequency evaluation, it is useful to discuss here the representation of variability because it is a concept raised later in the consideration of pooling of data from different populations.

In dealing with a quantity which is a random variable with a known distribution, we state the variability in terms of a probability or tolerance interval taken directly off the probability distribution for the quantity. Thus, we can identify the 5th and 95th percentiles of the probability distribution which characterizes the random variability and state that there is a 90 percent probability that any randomly selected value lies between these two limits. When the parameters of the distribution are not known exactly, but are inferred from a set of n observations (such as the sample mean and sample standard deviation for a normally distributed random variable) we can compute a statistical tolerance interval, which is an interval "within which we can state with a given confidence of being correct that a prespecified portion of values from a sampled distribution are located" (Reference 4). The probability of being correct, of course, increases with increasing number of observations.

(b) Bayesian and Subjectivist Measures of Uncertainty

The subjectivist interpretation of probability as a measure of an analyst's degree of belief in some proposition allows the following approach to the characterization of a parameter value uncertainty. Suppose a probability density function $f(\lambda)$ exists, such that the integral

$$\int_0^{\lambda_c} f(\lambda)d\lambda = c$$

represents the analyst's degree of belief that the true value of the parameter is less than λ_c, then $(0, \lambda_c)$ is a Bayesian probability interval. While it has the same intent as the classical confidence interval, i.e., to represent the analyst's uncertainty in the value of the parameter, the mathematical formalism adopted does allow the combination of measures of uncertainty in different parameters in a straightforward manner. This is discussed in several references (see, for example, the review paper by Parry and Winter, Reference 3), and will not be elaborated on here.

There are often several parameters in a PRA model for which data do not exist from which to make estimates, and engineering judgment is used. If uncertainties on parameter values are to be characterized in a manner consistent with that for these parameters for which data do exist, the Bayesian framework provides a mechanism for doing so. By contrast, classical statistics does not provide a means of providing such estimates or uncertainty measures.

In addition, the subjectivist approach allows a probabilistic characterization of modeling uncertainty, whereas there is no equivalent classical statistical method. The value of performing such a characterization is a function of the purpose for which the PRA is being performed. Since the assignment of consistent probabilities is not trivial and is analyst-dependent, some have found the Bayesian approach to probability unacceptable.

II.1.4 Data Availability Constraints

The availability of data as to type, quantity, and applicability is the second crucial aspect in choosing an uncertainty analysis method for any given application. In the context of a PRA of a nuclear power plant, we are (in most cases) dealing with rare events. For example, random failures of components at any given plant are quite infrequent, and the analyst is usually forced to merge the experience of a number of plants to obtain a data base which covers the full spectrum of components needed in a PRA. Similarly, the occurrence of earthquakes is a rare event anywhere, while the occurrence of an earthquake in the vicinity of a plant (close enough to at least trip the plant) is even more rare. A similar situation exists with respect to pipe break frequencies (for LOCA initiators) and the occurrence of significant in-plant fires.

Given the sparseness of the data, the analyst is usually forced to merge the data from several plants together and to group similar (but not identical) components into generic classes so as to pool their data. Engineering judgment is required to determine the applicability of the data and to perform the aggregation of the different sources of data into the generic groupings.

Even in the case of the simplest type of data--component failure rate in time or per demand--true "data" in the sense of a set of measurements of the quantity in question does not exist. Typically, in the literature, one has a record of a number of components failing in a given span of years. From this the failure rate may be estimated by computing the ratio

$$\lambda = \frac{\text{Number of Component Failures}}{\text{Hours of Exposure or Number of Demands}} .$$

The data are often taken at a number of different plants and on components in different systems having different operating environments, test intervals and demand histories. Typically, the denomination is not known precisely and in many data sources engineering judgment has been used to determine reasonable average exposure times, demand histories, etc.

Thus, in addition to a component type's inherent (irreducible) variability in failure history* due to randomness in materials used in construction, tolerances in part sizes, etc., we also have variability which is due to data source differences:

- Plant-to-plant differences (type, location, climate, mode of plant operation, age of plant, vendor, etc.).
- In-plant differences (component manufacturer, age of component, location in plant, mode of usage during operation).
- Generic grouping.
- Mode of failure (electrical vs. binding vs. leakage, etc.)

This type of variability is often termed systematic, because we could (in theory) reduce or eliminate this variability given enough time and money for testing or monitoring the particular component of interest. However, this variability becomes a source of uncertainty when data from several sources are applied to the analysis of a particular plant, as is the case with RMIEP.

In practice, for failure data on nuclear power plant components, one must use the existing data with its systematic variability, and the questions become:

---

*It should be remembered that this variability is accommodated by the treatment of failures as random occurrences. The adoption of a constant failure rate or constant failure probability is equivalent to making an assumption about the appropriate probabilistic model to describe the failure history.

a. How to characterize the uncertainty in order to reflect the systematic variability.

b. How to reconcile generic data (from many sources) with more limited plant-specific data (if available).

The specific choices one makes in answering these questions will play an important and often limiting role in determining a suitable method for performing an uncertainty analysis. Some of the considerations involved are discussed in more detail in Mosleh[5] and Parry.[6]

One resolution, which has been adopted in several PRAs to date, is to define a generic group of components for which the failure rate with time or on demand are assumed to be a fixed value. Then, data from other plants are used to derive a probability distribution that reflects the plant-to-plant variation in the value of the parameter for the equivalent group of components. This probability distribution, which represents a measure of physical (systematic) variability can be regarded as a confidence interval measure for the parameter value for a particular plant of interest, under the assumption that the plant is one of the general population that has been sampled. Thus, for example, in the WASH-1400 study,[7] the data on valve failure rates from 10 different sources were plotted in cumulative distribution form to obtain the lognormal distribution shown in Figure 2.

When no plant-specific data are available, the generic probability distribution can be used unmodified. When plant-specific data are available, the generic distributions can, in the Bayesian method, be modified, using it as a prior distribution, and utilizing the plant-specific data to specify the likelihood (see later). This is discussed, for example, in Apostolakis et al.[8] and Kaplan.[9] This use of generic data has important implications for the propagation of uncertainty measures as discussed in Section III.1.3 - Incorporation of Correlation.

It may also be desirable to include other sources of variability in deriving the probability distribution. For example, the distribution may be chosen to reflect both plant-to-plant and system-to-system or component-to-component variability. This distribution should, however, be used to represent uncertainty in the component's failure characteristics only if there is no consistent system-to-system variability at all plants. If there is a significant consistent system-to-system variability, the components from different systems should not be grouped into the same population, but rather each system should be treated separately by constructing a plant-to-plant distribution for components of that system.

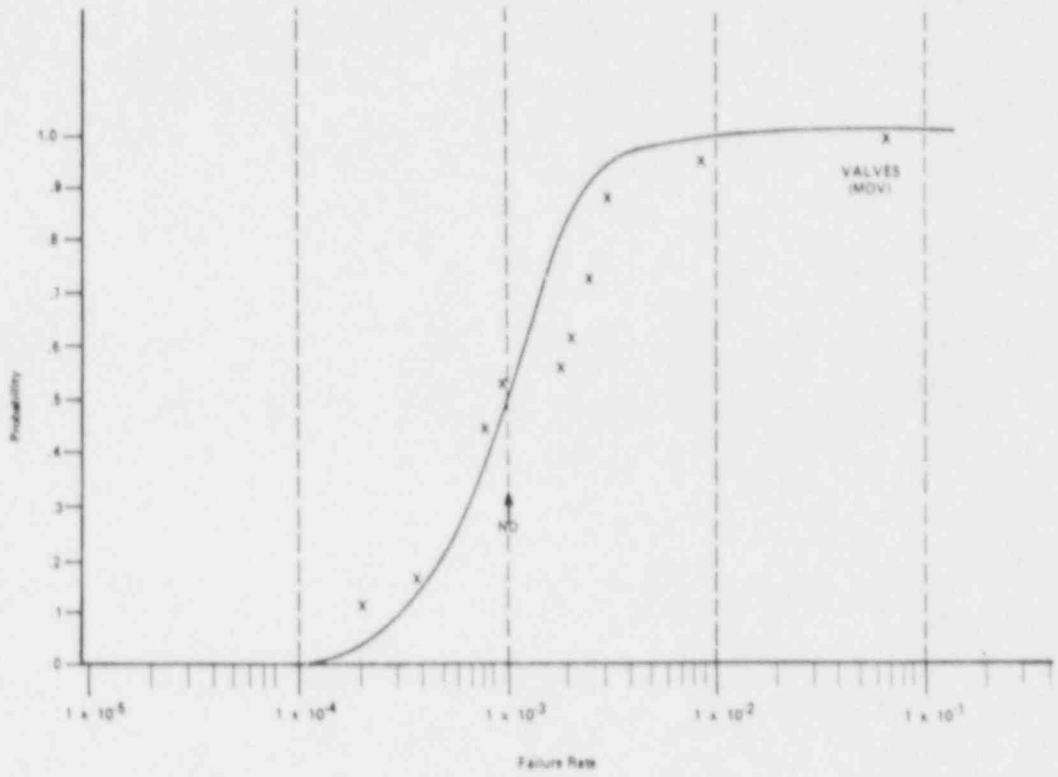

Figure 2. Lognormal Distribution--Valves (from Reference 7)

II.2 Methods for Propagating Component Uncertainty Through PRA Models

Several methods have been proposed for the propagation of uncertainty on parameter values through the PRA analysis, to provide uncertainty measures on the PRA results. Of the methods discussed below, the first is directed toward the classical statistical approach, the others are appropriate where a probability distribution is used to characterize uncertainty and thus can only be used in conjunction with a Bayesian approach.

II.2.1 Maximus/Bounding Approach

The Maximus/Bounding method by Spencer and Easterling (Reference 10) has two components to its propagation of parameter uncertainty through fault tree models: statistical confidence interval estimation (as discussed in Section II.1.2), and a bounding analysis. Data on component failures are divided into two categories. Data derived from designed experiments are classified as objective data. The contribution to top event uncertainty due to parameters based on objective data is modeled by a confidence interval estimate calculated using the Maximus technique. Parameter estimates that do not fit the objective category (e.g., expert opinion) are classified as subjective. Uncertainty due to subjective parameter estimates is exhibited by a bounding analysis and a seven point pictorial format.

The bounding analysis is accomplished by selecting a range of values for each subjectively defined parameter. The accident sequences are then quantified for the following three cases--all subjective parameters are fixed at their low estimate, nominal estimate, and high estimate. The uncertainty of the objectively estimated parameters is propagated through the sequence model (in each case) using the Maximus method for each set of subjective parameters.

The Maximus method of confidence interval calculation is an application of the "instrom-Madden theorem[11] and applies to failure data that are binomially distributed. In essence, for each series or parallel group of components, rules are prescribed from which "equivalent" binomial data (n-failures in m tests) can be derived. From this equivalent data, approximate binomial confidence limits for the accident sequence are derived in the usual way, using equations (3).

Note that, in application, a reliability block diagram of the system must be generated, as the method cannot be applied directly to fault trees or cutsets. Further, the algorithm applied to serial components is exact, while the algorithm applied to parallel components is heuristic, but has been tested over a wide variety of situations. Finally, it should

be noted that the results of the Maximus calculation depend on the set of values which have been selected for the subjective parameters.

After the accident sequence confidence intervals have been determined for each set of subjective parameters, the Maximus/Bounding approach provides three different perspectives of the top event uncertainty; subjective uncertainty, statistical uncertainty, and overall uncertainty. The subjective uncertainty is calculated by fixing all objective parameter values at their nominal estimates based on the data. The range of the top event value due to uncertainty of the subjective parameters is calculated by varying the subjective parameter values from their low values to their high values. The statistical confidence limits of the top event are calculated by fixing all subjective parameter values at their nominal estimates. A Maximus calculation is then performed using the objective parameter data to calculate a confidence interval. The overall uncertainty is estimated by setting the subjective parameter values fixed at two different values, their low value and their high value. At each set of fixed subjective values, a Maximus calculation is performed on the objective parameters. Recall from earlier discussion that the Maximus solution will vary for different values of the subjective parameters. An example of the three types of uncertainty intervals estimated by Maximus/Bounding is shown in Figure 3.

Comparisons of the Maximus/Bounding approach with a Bayes'/Monte Carlo approach (using assumed distributions) for a typical PRA systems model have shown[12] that the Maximus approach leads to wider confidence bounds (as expected) but that the upper bounds of both methods are in reasonable agreement.

The advantages of the Maximus/Bounding approach are:

1. No assumptions need be made concerning the distributional form of the basic event probabilities, since failure rates are assumed to be constant parameters.

2. Once the systems models are in series/parallel block diagram form, the method is straightforward and simple to apply.

Disadvantages of the method are:

1. The typical PRA systems model in fault tree format must be converted to a block diagram format.

2. A certain degree of arbitrariness exists when generating "equivalent" data for groups of like components.

3. The Maximus "rules" for parallel configurations cannot be proven with mathematical rigor, although extensive demonstration calculations substantiate their heuristic foundations.

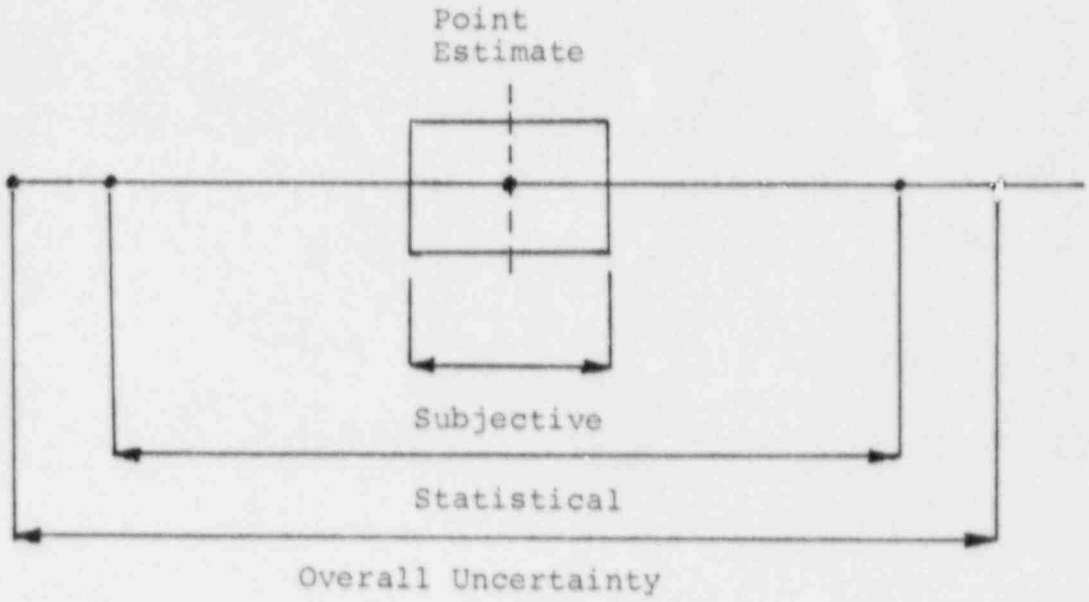

Figure 3. Illustration of Maximus/Bounding Intervals

II.2.2 Method of Moments

The Method of Moments propagates the basic event uncertainty through PRA models by generating lower order moments, such as the mean and variance, of the distribution of the top event from the moments of the basic event distributions. In essence, one is using a Taylor series expansion of the variation of the top event value with respect to variations in the independent variables (the basic events). The coefficients of the Taylor series are related to the moments of the distributions of the basic events.

In theory, if sufficient terms of the expansion are generated, the distribution of the top event can be determined completely. In practice, only a few lower ordered terms are generated. For example, if one has estimates of only the means and variances of the basic events, one can approximate the mean and variance of the distribution of the top event. However, with the Method of Moments, once the top event attributes (mean and variance) have been calculated, it is necessary to make an assumption about the form of its distribution so that other attributes, such as 95th and 5th quantiles, can be estimated. A more detailed treatment of this method is covered in a comparison study of various uncertainty analysis methods.[13] Techniques also exist that include the use of higher order moments as well.[14]

Advantageous of the Method of Moments are:

1. The mean and variance of the top event can be approximately determined even if only means and variances of the basic events are known, rather than their complete distributions.

2. For simple top event functions with derivatives that can be determined analytically, the Method of Moments is simple to apply.

Disadvantages are:

1. For functional relationships between the top event and the basic events which are significantly nonlinear, the use of only lower-ordered terms in the Taylor series expansion can lead to significant inaccuracies.

2. For complex functional relationships for which analytical derivatives cannot be easily obtained, some form of experimental design is required to numerically determine the coefficients of the Taylor-series expansion.

3. For large systems models of the type usually encountered in PRAs, this approach has been found to be unwieldy in application.

II.2.3 Method of Moments Using Tchebyschev's Inequality

In this technique, the mean and variance of the top event distribution are calculated from the basic event moments as in Section II.2.2. However, to arrive at uncertainty intervals, Tchebyschev's inequality is used instead of assuming a form for the top event's distribution. Tchebyschev's inequality is a theorem in probability theory which states, "for <u>any</u> distribution with finite mean and variance, at least $(1-k^{-2}) \cdot 100$ percent of the probability is in the range of $\pm k\sigma$ around the mean," (Reference 4). Thus, for example, this inequality implies that at least 88.9 percent of the probability lies in the interval of $\mu \pm 3\sigma$. Such tolerance intervals calculated with Tchebyschev's inequality tend to be very conservative.

II.2.4 Propagation of Discrete Probability Distributions

Propagation of Discrete Probability Distributions (DPD) is a technique in which the basic event distributions are discretized, and a discrete analog of the top event distribution is calculated. Each distribution is divided into some number of intervals, say n. A value for each basic event probability is chosen for each interval (usually the mean of the basic event probability in that interval), and the probability that the basic event probability occurs in each interval is calculated. Thus, the distribution of each basic event probability is discretized into n values, each value with a corresponding probability. The top event is evaluated n^k times where k is the number of independent variables. The result of each evaluation has an associated probability equal to the product of the probabilities of the independent variables. Thus, a discrete probability distribution can be constructed from these values.

The DPD technique is, in theory, a valid method for functions of many variables. However, the technique can become quite cumbersome with respect to computer storage, and the propagation must be performed in stages with intermediate re-discretization. Incorporation of correlation is then quite awkward. More discussion of this method can be found in Reference 15.

II.2.5 Monte Carlo Method

The Monte Carlo approach is the most fundamental approach to uncertainty analysis. Monte Carlo simulation consists of making repeated quantifications of the top event value. For each quantification or Monte Carlo run, each random variable basic event is sampled using a random number generator to select a value from the basic event's distribution. This procedure is repeated numerous times, and the various top event outcomes are sorted to obtain empirical estimates of the desired top event attributes such as mean, median, 95th, and 5th quantiles.

As more and more Monte Carlo samples are made for a top event equation, the precision of the empirical quantile estimates of the top event distribution improves. However, the rate of convergence to the true distribution tends to decrease as the number of samples increases.[13] Of course, the validity of the result is only as good as that of the input.

Advantages of the Monte Carlo method are:

1. Complete flexibility in the selection of basic event distributions.

2. Any specified precision of the top event attributes can be achieved (limited only by cost and roundoff-error).

3. The method is reasonably easy to implement.

Disadvantages of the Monte Carlo method are:

1. Probability distributions must be defined for all parameters.

2. Computer costs can become significant for large problems.

3. Sensitivity measures of the basic event uncertainties contribution to the top event uncertainty are not directly obtained.

II.2.6 Latin Hypercube Sampling

Latin hypercube sampling (LHS) was developed to improve upon the accuracy and precision of Monte Carlo simulations in estimating functions of multiple random variables.[16] LHS is a stratified sampling technique, where n different values are selected for each basic event. The values are selected by dividing each basic event's probability distribution into n intervals, each of equal probability. Within each interval, one value of each basic event's probability is randomly selected. The n values of a particular random variable are then combined (in a manner which keeps the pair-wise correlations near zero) with the n values of the other basic events. The result is an n x k matrix (k is the number of sampled basic events) where the ith row of the matrix contains specific values of each random variable to be used in the ith run or calculation of the top event value. The top event's distribution and quantiles are empirically estimated from the n runs.

LHS has the same advantages as associated with Monte Carlo methods. Additional advantages include:

1. The accuracy in modeling the top event uncertainty is improved with respect to direct Monte Carlo methods for the same number of samples.

2. Use of LHS ensures that the tails of the basic event distributions are included in the propagation of uncertainty.

Disadvantages of the LHS method may be:

1. Possible bias in result, particularly for small sample sizes.

2. If not suitably automated, this procedure can be cumbersome and time-consuming.

3. Being a relatively [illegible] echnique, all of its properties may not be fully un[illegible]od.

II.3 Method Recommended for RMIEP

Following a review of the various methods for uncertainty propagation, and in view of the recommendation that uncertainty in parameter values should be represented by di[illegible]butions (as discussed in II.1.3), an approach based on [illegible] hypercube sampling (LHS) and Monte Carlo simulation is re[illegible]ended for RMIEP. This approach was chosen for the following reasons:

1. This approach can conveniently accept any form of distribution for each basic event, and thus the effect of different distributional assumptions could be assessed.

2. This approach can be applied directly to the output of typical fault tree analysis codes (e.g., SETS, FTAP, etc.).

3. This approach allows a direct assessment of the effects of correlation between uncertainties for components in generic categories.

4. The approach has been found to be relatively simple and inexpensive for large systems models.

The choice of LHS as the method for propagation of uncertainty will allow flexibility in characterizing basic event uncertainty. Any method of basic event uncertainty characterization which establishes a distribution for the failure probability can be used in conjunction with LHS.

III. ELEMENTS OF RECOMMENDED APPROACH FOR UNCERTAINTY PROPAGATION

The Latin hypercube/Monte Carlo approach recommended for RMIEP will be implemented using the Top Event Matrix Analysis Code (TEMAC).[17] This computer code can accept the samples generated from either an LHS routine or a Monte Carlo simulation.

III.1 Propagating Uncertainty Through RMIEP Models

The following sections describe the codes which are recommended for use in the Latin Hypercube Sampling procedure for propagation of uncertainties through the RMIEP systems models. Code output capabilities and sensitivity measures computed are also described.

III.1.1 Latin Hypercube Sampling

The TEMAC code requires a set of sample vectors chosen in accordance with the stratified sampling approach. A Latin Hypercube Sampling (LHS) routine[16] is used to generate the input sample vectors by sampling basic event probabilities from their distributions. The output of the LHS will be read directly into TEMAC, which is designed to take simulations (either pure Monte Carlo or Latin hypercube samples) of parameter values and quantify a top event equation. The LHS and TEMAC codes permit flexibility in uncertainty modeling by allowing for different probability distributions to be used in the same computer run. This eliminates the necessity of assuming all random variables have the same type of distribution.

An introduction to the LHS technique, with a brief discussion of the theory of LHS and a simple example demonstration is included in Appendix B. The example and discussion are taken directly from Reference 16, the LHS user's guide.

III.1.2 Quantification of Top Event Attributes

Once the Latin hypercube sample matrix has been generated for the basic event probabilities, it is read into TEMAC, along with the top event equation. TEMAC evaluates the equation for each sample vector in the matrix.

TEMAC generates a variety of attributes of the top event's distribution as a result of the propagation of basic event parameter uncertainty through the model. These include estimates of the top event's mean, median, standard deviation, 5th and 95th quantiles and a nominal estimate of the top event's value calculated with the nominal estimates of the base events. Four importance measures of basic event contributions to the top event's value are calculated--partial derivative of the top event equation with respect to each basic event, risk reduction,

risk increase, and a measure of an event's uncertainty contribution to the variance of the top event distribution called uncertainty importance. For each measure, the basic events are ranked in order of contribution. These measures are briefly described in Section III.1.4.

III.1.3 Uncertainty Correlation in Sampling

It is recommended that uncertainty correlation be included in the LHS samples. The need for this is discussed below. When components are grouped together into a single generic category, the assumption is made that each component in that group has the same failure rate or probability. Thus suppose that an accident sequence ACC is given in terms of basic event probabilities as

$$ACC = a_1a_2 + a_1b_2 + a_2b_1 + b_1b_2 \quad , \tag{4}$$

where a_i is the failure unavailability of component i of the group identified by the letter a_i and similarly for b_i. Since each component within the group has the same unavailability, ACC can be rewritten as

$$ACC = a^2 + 2ab + b^2 \quad , \tag{5}$$

where a represents the unavailability of a component in generic group a and similarly for b.

In performing an uncertainty analysis, it is equation (5), must be used to evaluate the probability distribution on ACC. This gives different results from using equation (4), as can be seen from taking the expected value. In using the first expression, the a_i are treated as independent so that

$$E(ACC) = E^2(a) + 2E(a)E(b) + E^2(b) \quad , \tag{6}$$

whereas in the second it is

$$E(ACC) = E(a^2) + 2E(a)E(b) + E(b^2) \quad , \tag{7}$$

where E(x) is the expected value of x.

The necessity of including the so-called uncertainty correlation was discussed, in a Bayesian context, by Apostolakis and Kaplan

in Reference 18. They pointed out that the values of the a_1 and a_2 (and b_1 and b_2) must be regarded as being the same since the analysts state of knowledge about them is identical. Thus in a Monte Carlo analysis, when in one sample a random value from the probability distribution on a_1 is chosen, the same value must be used for the value of a_2, and so on. It is in this sense that the values are correlated.

Consideration of uncertainty correlation can be important when screening is performed. The fact that the expected value of ACC using equation (5) is greater than that using equation (4) means that extra care has to be taken when using point estimates directly in the equation for ACC as a screening method for identifying dominant cut sets. This is particularly important when there are more than two components from a single group in a cut set and/or the uncertainty on the value of the component availability is large, when

$$E(a^n) >> E^n(a) \quad .$$

It may thus be desirable, unless it is easy to identify in a general way the possibility of such cut sets, to accept some sort of second screening where, for example a^n is replaced by a whenever the former appears in a cut set, to guarantee saving the potentially troublesome cut sets. This would mean, in the current example, calculating the unavailability (for screening purposes) as

$$ACC = a + 2ab + b \quad .$$

Any cut sets that reviewed this second screening would also be included in a full uncertainty analysis, (expressed of course in their original form).

For the final quantification and uncertainty analysis, the TEMAC code can incorporate this uncertainty correlation directly, since the Latin hypercube sampling program permits the specification of full correlation between arbitrary numbers of variables in the top event Boolean equation. (In fact, these codes permit the specification of any degree of correlation less than full correlation also.) It is recommended, however, that full correlation be used for failure rates of components in the same generic category (and same failure mode).

III.1.4 Sensitivity and Importance Measures Computed in TEMAC

The TEMAC code calculates various measures of the sensitivity of the top event estimates to changes in the estimates of the basic events. These measures enable the basic events to be

ranked as to their contribution to the top event values and enhance interpretation and comparison of the PRA results.

In TEMAC, the mathematical representation of a Boolean equation consisting of m cut sets, each of which consists of k events, is written as

$$g(\underline{X}) = \sum_{j=1}^{m} X_1^{a_{1j}} X_2^{a_{2j}} \ldots X_k^{a_{kj}} \quad , \tag{8}$$

where $a_{ij} = 1$ if the i^{th} event, X_i, is contained in the j^{th} cut set, otherwise $a_{ij} = 0$.

The basic event X_i may be either a component failure rate described in probabilistic terms, or an initiating event representing a frequency of occurrence and expressed as a certain number of occurrences during some time period. Thus, $g(\underline{X})$ expresses the frequency of the top event.

The sensitivity and uncertainty measures available are computed (for <u>each</u> simulation) from this equation as described below.

<u>Partial Derivative</u>

The partial derivative is a measure of the sensitivity of a function to changes in a parameter's value at a specific point. The partial derivative of the top event $g(\underline{X})$ given with respect to X_i is

$$\partial g(\underline{X})/\partial X_i = \sum_{j=1}^{m} X_1^{a_{1j}} X_2^{a_{2j}} \ldots a_{ij} X_j^{a_{ij}-1} \ldots X_k^{a_{kj}} \quad . \tag{9}$$

Note that the partial derivative expressed above will never contain the variable X_i since a value of $a_{ij} = 0$ eliminates the term from the partial derivative and a value of $a_{ij} = 1$ generates an exponent of zero for X_i.

Risk Reduction*

Risk reduction is a measure of the maximum potential reduction in the top event value if a basic event failure rate is reduced to zero. Risk reduction identifies components for which improved reliability could yield significant improvement in risk. Improved concepts in system designs and operator procedures are types of insights possibly gained here.

The measure is calculated by setting the nominal value of an event, X_{ij}, equal to zero in $g(\underline{X})$ and subtracting the resulting value from the nominal evaluation of $g(\underline{X})$:

$$\text{Risk Reduction for } X_j = g(\underline{X}) - g(\underline{X}|X_j = 0) \quad . \qquad (10)$$

Risk Increase*

Risk increase is a measure of the maximum potential increase in top event frequency if a basic event's probability increases to one. Risk increase yields insights of critical components and procedures which, if allowed to become less reliable or effective, could result in a worsening of the risk profile at a plant. The importance of improvements in preventative measures, such as maintenance and testing, are types of insights derived from increase measures. The measure is calculated by:

$$\text{Risk Increase for } X_j = g(\underline{X}|X_j = 1) - g(X) \quad . \qquad (11)$$

Uncertainty Importance

Uncertainty importance, a new measure, is a measure of the significance of a basic event's uncertainty to the uncertainty of the top event. This measure gives insight into which basic event probabilities, due to their own uncertainty, are driving the uncertainty of the top event estimate. Areas of need for improved data analysis or increased data gathering can be identified by this measure. Uncertainty importance, UI_j, for event X_j, is calculated as:

$$UI_j = [Var(X_j)]^{1/2}(\partial g(X)/\partial X_j) \quad . \qquad (12)$$

*The risk increase and risk reduction measures in TEMAC are interval estimates. Ratio estimates can also be defined and are useful in certain applications. See Reference 19.

Note that each of these sensitivity and uncertainty measures are computed for each simulation, so that sample distributions are also available for each of the sensitivity measures as well as for the Boolean equation being evaluated.

The risk measures calculated by TEMAC provide a set of measures which enhance resolution of safety issues from various regulatory perspectives. Insights to design improvements, operational improvements, and critical factors for reactor integrity and accident consequences can be gained from the TEMAC output.

III.1.5 Inclusion of Modeling Uncertainty

As discussed in Section I.1, modeling uncertainty is that which arises because of the existence of plausible alternative models or hypotheses concerning the basic events being analyzed. Thus, for example, one might wish to include the uncertainty due to the fact that two alternative models for a reactor coolant pump seal failure exist, and it is not possible to exclude either model due to limited experience with the phenomena.

Incorporation of such modeling uncertainty can be directly accomplished with the Latin hypercube sampling routine and the TEMAC code, provided one can assign a relative likelihood to each alternative. Thus, for example, if a modeling alternative consists of two hypotheses A and B, and if hypothesis A is assigned a 10 percent likelihood and hypothesis B is assigned a 90 percent likelihood, then in choosing N vectors of samples for the TEMAC uncertainty analysis, the LHS routine will select 0.1N of the vectors utilizing hypothesis A, and 0.9N of the vectors utilizing hypothesis B. More generally, the LHS routine has the capability of considering multiple (more than two) alternatives for a given basic failure event, provided relative likelihoods are prescribed for each alternative (and, of course, these relative weightings sum to unity). By this method, the uncertainty in the top event will include both uncertainty in the basic event parameters as well as uncertainty due to alternative modeling parameters. In addition, for this simple case, regression modeling can be used to rank the various modeling issues with respect to their impact on the top event, and to rank the alternatives within each issue. An example application of this approach is given in Reference 20.

In the more general case, the LHS sampling code and the TEMAC code can, in principle, be used to incorporate modeling uncertainty into the distribution of the accident sequence frequency (or other top event of interest). To illustrate, consider a simple accident sequence involving these basic events:

$$ACC = C_1C_2C_3 \quad . \tag{13}$$

Each of the basic events has an associated uncertainty distribution. Modeling uncertainty can enter in if there is significant uncertainty with respect to the parameters of the associated distribution, say the mean. Thus, for example, due to lack of knowledge of the physics of the situation, the mean may be assured to take on one of two discrete values corresponding to two different failure modes. Further, based on experience on expert judgment we may associate a discrete probability with each of the two alternatives. The existence of two (or more) plausible alternatives (with associated probabilities) will result in greater uncertainty in the sequence frequency.

A second source of modeling uncertainty results when there is uncertainty in the Boolean expression for the accident sequence, due to uncertainty in the success criteria for safety systems. Thus a plausible alternate to the sequence might be

$$ACC = C_1 + C_2C_3 \qquad (14)$$

and again, we must associate some probability (degree of belief) with the validity of each of the plausible Boolean expressions. Again, this uncertainty in modeling the success criteria will increase uncertainty in the accident sequence.

These two types of modeling uncertainty can, in principle, be modeled using the LHS sampling code and the TEMAC code. Thus we can form the "super" sequence

$$ACC^* = A\ C_1C_2C_3 + (1-A)\ [C_1 + C_2C_3] \qquad (15)$$

and assert that the new variable A takes on the value 1.0 a specified percentage of the time, and a value of zero for the remainder of the time. In effect, the variable A becomes a switch characterized by two deterministic values with associated probabilities of occurrence.

To incorporate the modeling uncertainty associated with the variables C_1, C_2 and C_3 one must make multiple LHS sample code runs to generate the sample vectors for input to the TEMAC code. Since there are (in this case) two levels associated with each of three variables, there are $2^3 = 8$ distinct combinations to be considered. Each combination has an associated probability determined from the probability of occurrence of the basic event "levels" themselves. Thus if

$$\text{Prob}\ (C_1^{low}) = P_{11}$$

$$\text{Prob}\ (C_2^{\text{high}}) = P_{22}$$

$$\text{Prob}\ (C_3^{\text{high}}) = P_{32}$$

Then, assuming independence of the levels,

$$\text{Prob}\ (C_1^{\text{low}} \text{ and } C_2^{\text{high}} \text{ and } C_3^{\text{high}}) = P_{11}\ P_{22}\ P_{32}$$

If, however, the levels of the variables are not independent, then the degree of correlation must be specified. Thus, for example, if the "high levels" of variables C_2 and C_3 always occur together, then they are fully correlated (ρ=1) and

$$\text{Prob}\ (C_2^{\text{high}} \text{ and } C_3^{\text{high}}) = \text{Min}\ (P_{22}, P_{32})$$

rather than equal to $P_{22}\ P_{32}$ as in the independent case. Any degree of correlation can, in principle, be considered.

In any case, the probability of each distinct combination of levels can be computed, and the total number of LHS sample runs is subdivided in properties to the computed probabilities of occurrence of the various combinations.

As an example, consider each of the two accident sequences to be equally likely and the level of basic event C_1 to be independent of C_2 and C_3, but the events C_2 and C_3 to be fully (positively) correlated. Further, assume the levels of each of the three events to be equally likely. Thus

$$\text{Prob}\ (C_1^{\text{low}}) = 0.5$$

$$\text{Prob}\ (C_1^{\text{high}}) = 0.5$$

$$\text{Prob}\ (C_2^{\text{low}} \text{ and } C_3^{\text{low}}) = 0.5$$

$$\text{Prob}\ (C_2^{\text{high}} \text{ and } C_3^{\text{high}}) = 0.5$$

$$\text{Prob}\ (C_2^{\text{low}} \text{ and } C_2^{\text{high}}) = 0.0$$

and hence we have four distinct combinations

$$\text{Prob}\ (C_1^{low} \text{ and } C_2^{low} \text{ and } C_3^{low}) = 0.25$$

$$\text{Prob}\ (C_1^{low} \text{ and } C_2^{high} \text{ and } C_3^{high}) = 0.25$$

$$\text{Prob}\ (C_1^{high} \text{ and } C_2^{high} \text{ and } C_3^{high}) = 0.25$$

$$\text{Prob}\ (C_1^{high} \text{ and } C_2^{low} \text{ and } C_3^{low}) = 0.25$$

Thus, if we intended to generate 1000 LHS samples, we would generate samples of the basic events C_1, C_2, C_3 according to the relative proportive given in Equation 2. That is, 250 sample vectors with random variable C_1 sampled from its distribution with the low mean, but variables C_2 and C_3 sampled from their corresponding means set at their high mean value, etc.

Further, in this set of 250 sample vectors, the switch A would be treated as a variable, and set to 1.0 in 50 percent of the vectors and to zero in the remaining 50 percent of the vectors. Finally, the four sets of 250 sample vectors would be input to the TEMAC code, which would evaluate the "super-sequence" in Equation 1 for each sample vector, and generate an empirical distribution for the accident sequence frequency.

The above example illustrates how modeling uncertainty can be incorporated in the core damage frequency evaluations. In application, the problem is one of scale, for the number of "level" combinations into which the LHS samples must be partitioned is m^n where m is the number of levels associated with each variable and n is the number of variables. Clearly, a number of sensitivity runs should be performed prior to evaluating the empirical distribution of the top event to eliminate as many variables as possible. However, the generation of LHS samples for core damage accident sequences is relatively efficient and inexpensive, so that incorporation of modeling uncertainties using the LHS and TEMAC codes is quite feasible.

IV. DATA AND DATA ANALYSIS

As has been discussed earlier, the methods for performing uncertainty analyses are inextricably linked to the data available in the following, the data needed in performing a Level 1 PRA are described, and current sources of data are outlined.

IV.1 The Parameter Value Estimation Process

The development of generic basic event probability estimates and uncertainty is a vital component in performing any uncertainty analysis. Limited operating experience forces the analyst to rely on information compiled from other plants. The data development process is usually interative, with the initial accident sequence evaluations performed using a set of conservative, screening input data. The screening values are then reviewed and adjusted as appropriate for the development of a finalized set of parameter values. The process of developing the screening values and adjusting them to final values is discussed in this section.

IV.1.1 Component Definition and Classification of Failures

A vital link in developing parameter estimates is the coordination of the available information with the needs of the plant models. At the start of any PRA project, the analyst has a general idea of the level of detail and type of basic events he will incorporate into his system models. As the systems analysis progresses, the nature of the basic events--actual components modeled, failure modes, human errors of interest, recovery actions--become better defined and understood. However, the available data and generic estimates do not always correspond to the models an analyst may develop. The actual definition of component boundaries and the grouping of information into failure categories can vary between data sources. Component boundary and failure mode definitions are not always appropriate for the data analyst's needs. Problems in finding useful information on certain component failure modes can arise. While many of these component failure modes will be matched with well-documented, well-founded generic parameter estimates, usually several component failure modes will arise that represent areas where little or no data collection has occurred.

IV.1.2 Screening Quantification of Component Failures

Once the component boundaries and failure modes are defined, a set of conservatively estimated values are developed for the screening analysis based on a review of published data. A nominal estimate for each component failure mode is made. A conservative or high estimate is also made. In most cases, any difference between nominal and conservative values reflects differences in published values. In some cases, though, it

might reflect analyst judgment of uncertainty. The conservative values may become a foundation for modeling parameter uncertainty in the final basic event probability models. An initial screening analysis of the sequences is enformed using these conservative screening point estimates. For the screening analysis, no attempt is usually made to determine or select distributions for the parameters. This is undertaken during the finalizing of the parameter value estimates. However, care must be taken so that the screening estimates are sufficiently conservative that components which might contribute to the sequence uncertainty analysis are not screened out of the analysis due to their point estimate value. To check this, the analyst should examine those cutsets in the dominant accident sequences whose numerical value is close to the truncation value used in the screening process.

Incorporation of common cause failure events and associated screening values requires additional consideration in that an appropriate conservative model of common cause failure must be chosen. A rather complete treatment of these issues and recommendations for screening procedures is given in Reference 21.

IV.1.3 Finalizing of Parameter Estimates and Uncertainty

After the initial screening of the accident sequences has been completed, additional data is sought for the remaining basic event failure or unavailability frequencies, and probability distributions on these frequencies must be developed. The goal is to provide nominal values (usually means), which are best-estimate to the extent possible, and to characterize the uncertainty in these nominal values to reflect the full range of sampling, plant-to-plant and modeling assumption uncertainties inherent in the estimates.

Both site-specific and generic data are sought. The generic data sources are reexamined, and the generic failure frequency estimates are refined to the extent possible, based on a closer examination of the failure events in the data sources. The FRAC data attribute analysis code[22] developed at Los Alamos National Laboratory is a tool that can be used to develop component failure rates applicable to specific reactor types, specific component subgroups or specific failure modes--limited only by the degree of detail in the failure event reporting.

Site-specific data are sought to augment the generic data bases. In some cases, it may be possible to base a components failure or unavailability rate on site-specific data above. The more common situation is that the site-specific data is utilized in conjunction with the generic failure rate estimates. For failure modes for which little or no data exist, recourse to expert judgment must be made. In any case, probability distributions must be developed for the failure rates.

In the following sections, data needs, data sources and methods for developing uncertainty distributions from various combinations of site specific data, generic data and expert judgment are described.

IV.2 Data Needs and Data Types

For evaluation of accident sequences in the internal event analysis, point estimates and uncertainties are needed for the following:

a. Initiating Events
b. Component Unavailabilities
c. Recovery Terms
d. Common Mode Failure Model Parameters

Each of these are discussed below.

a. Initiating Events

A variety of LOCA and transient-inducing initiating events are identified for the plant under consideration. For example, a total of 30 initiating events are being treated for the LaSalle plant in RMIEP, as shown in Table 1. The choice of initiating events is based on a review of plant-specific occurrences plus an evaluation of initiating events occurring at similar plants which may be applicable to the site in question, and a detailed examination of plant support systems.

b. Component Unavailabilities

Components may be unavailable due to a wide variety of causes. In general, the component unavailability may be computed from (some or all) of the following expressions as appropriate:

Q_D	Demand Failure Unavailability
$\lambda_S \frac{T_t}{2}$	Standby Failure
$\lambda_R T_m$	Failure to Run
$\lambda_M T_R$	Unavailable Due to Unscheduled Maintenance
$\lambda_M P_{FM} \frac{T_t'}{2}$	Unavailable Due to Failure to Restore After Unscheduled Maintenance

Table 1

LaSalle Unit 2 Initiating Event Categories

Transients (General)

1. Turbine trip with turbine bypass available
2. Turbine trip with turbine bypass unavailable
3. Total main steam isolation valve closure
4. Loss of normal condenser vacuum
5. Total loss of feedwater
6. Trip of one feedwater or condensate pump
7. Inadvertent opening of a safety-relief valve (stuck)
8. Loss of offsite power

Transients (Special)

9. Loss of 125 Vdc bus
10. Loss of 4160 Vac bus
11. Loss of instrument air
12. Loss of drywell pneumatic
13. Loss of 100 drywell pneumatic
14. Complete loss of reactor vessel narrow range
15. Loss of Channels A and C of reactor vessel narrow range level instrumentation (false high level indications)

LOCA (Inside Containment)

16. Small LOCA inside containment ($\leq$ 0.005 square ft. for liquid, $\leq$ 0.1 square ft. for steam)
17. Medium LOCA inside containment (0.005 to 0.3 square ft. for liquid, 0.1 to 0.3 ft. for steam)

Table 1

LaSalle Unit 2 Initiating Event Categories (Concluded)

18. Large LOCA inside containment ($\geq$ 0.3 square ft.)

19. Reactor vessel rupture

LOCA (Outside Containment)

20. Steamline LOCA outside containment

LOCA (Interfacing System, Outside Containment)

21. RHR or LPCS LOCA outside containment without isolation

22. CRD LOCA outside drywell without isolation

ATWS

23. Turbine trip with turbine bypass available combined with reactor protection system failure

24. Turbine trip with turbine bypass unavailable combined with reactor protection system failure

25. Total main steam isolation valve closure combined with reactor protection system failure

26. Loss of normal condenser vacuum combined with reactor protection system failure

27. Total loss of feedwater combined with reactor protection system failure

28. Trip of one feedwater or condensate pump combined with reactor protection system failure

29. Inadvertent opening of a safety-relief valve (stuck) combined with reactor protection system failure

30. Loss of offsite power combined with reactor protection system failure

$\lambda_t T_{to}$	Unavailable Due to Test Outage
$\lambda_t P_{ft} T''_t$	Unavailable Due to Failure to Restore After Test

For any specific component, not all sources of unavailability are important, but this must be determined for each component. The parameters are defined as:

Q_D = Demand failure probability
λ_S = Standby failure rate
T_t = Average time between tests that would detect the standby failure
λ_r = Running failure rate
T_m = Mission time following demand that component must perform satisfactorily
λ_m = Frequency of unscheduled maintenance
T_r = verage repair time for components due to unscheduled maintenances
P_{fm} = Probability of failure to restore component to service following unscheduled maintenance
T'_t = Average time between tests which would detect the failure to restore error
λ_t = Frequency of scheduled testing
T_{to} = Average test outage time
P''_{ft} = Probability of failure to restore component to service following testing
T''_t = Average time between tests which would detect the failure to restore error

In general, we need both point estimates and uncertainties for each term above.

The demand, standby and running failure rates are determined using plant-specific component failure records and/or generic data. The process of combining plant-specific and generic data is described explicitly in Sections IV.4.2 and IV.4.3.

The frequency of scheduled testing is obtained directly from the plant test and maintenance procedures. No uncertainty is usually associated with this parameter.

The frequency of unscheduled maintenance is determined by examining plant documents such the control room log and maintenance request records or data from similar plants. These data will define a distribution on the frequency, again often taken in a lognormal form.

The mission time is determined for each accident sequence by identifying the time required for the component to function (following its demand) in order to mitigate the off-normal situation. Often, this is defined explicitly by success criteria for the safety system of which the component is a part. It is usually treated as a parameter with no uncertainty.

The average time to repair and average test outage times are determined from plant records (control room, log, etc.). These times are used to define distributions (often lognormal) from which average values and uncertainty measures can be taken.

The terms T_t, T_t', T_t'' define the times between scheduled testing or routine evaluation which would detect the components unavailability due to failure of restoration of the component. These times may or may not be the same. They are, however, identified from the plant test and maintenance procedures.

Finally, the probabilities of "failure to restore" are determined both from plant-specific histories of such occurrences, generic models and theoretical models. These human error probabilities are themselves treated as random variables, with uncertainty distributions assigned to each.

c. Recovery Terms

Recovery terms are appended to the (dominant) accident sequence cut sets to reflect the probability that the plant operators will either not correctly diagnose the accident condition and appropriate mitigating response, or will not effectively perform the actions required to effect the recovery. Models involve simulator data, generic grouping of recovery actions and expert judgment in choosing the "most likely" recovery action that the operator will select for a given accident scenario. Point estimates and uncertainties are usually provided by the model selected, hence no separate data gathering is involved. For the RMIEP PRA of the LaSalle plant, a new simulator-based recovery model has been developed (Reference 23). Thus the recovery terms used in this PRA will be quite plant specific, although indications are that the new model has wide generic applicability.

d. Common Cause Failure Model Parameters

Often, there are important cut sets that involve failure of like components in parallel service conditions, i.e., failure of both diesel generators or both emergency station battery banks. The common cause

failure of such important components is often explicitly modeled in the fault tree as a single basic event. The probabilities of these common mode failure basic events are computed using one of several empirically-based models, such as

1. The beta factor model[24]
2. The binomial failure rate model[25]
3. The multiple greek letter model[26]

Each of these models require data, usually in the form of the fraction of multiple component failures as a percentage of the total number of failures.

IV.3 Data Sources

There is a wide variety of component failure rate "data" existing in the literature, although the amount of actual raw data is quite limited. Available nuclear component reliability sources may be categorized as follows:

1. Collections of actual failure events
2. Statistical analyses of data
3. Generic failure rate data bases

The analyst must be aware that many sources of component failure characteristics are, in fact, just reanalyses of existing data and thus are not new or independent data sources. This situation must always be kept in mind when reviewing the literature for appropriate "generic" failure rates for use in a PRA. In this section, we will categorize and summarize the more important sources of reliability data for nuclear power plant components.

As mentioned above, the number of actual data collections (category 1) is relatively small. Historically, the most important in the United States are the Licensee Event Reports (LERs). Summaries of these reports and associated statistics for different component types are contained in reports generated at the Idaho National Engineering Laboratory. More recently, the In-Plant Reliability Data System Program at the Oak Ridge National Laboratory has been collecting and summarizing failure data from U.S. plants in a systematized format. A third source of data is the Nuclear Plant Reliability Data System (NPRDS) produced by the Institute for Nuclear Power Operations (INPO). In addition, there have been a number of special purpose data collections related to loss of offsite power, anticipated transients without scram and diesel generator reliability. The sources of actual failure event compilations are shown in Table 2.

The second category of reliability data sources consist of reports that have analyzed failure event data from one or more of the above sources, and produced data-based estimates of the

Table 2

Collections and Summaries of Actual Failure Events

	Title	Source	Reference
1.	Licensee Event Reports	USNRC	
2.	Licensee Event Report Summaries	Idaho National Engineering Laboratory	
	Pumps		NUREG/CR-1205
	Valves		NUREG/CR-1363
	Electrical Power		NUREG/CR-1362
	Circuit Breakers, Protective Relays		NUREG/CR-4126
	Initiating Events		NUREG/CR-3862
	Selected I&C Components		NUREG/CR-1740
	Control Rods and Drive Mechanisms		NUREG/CR-1331
3.	In-Plant Reliability Data System	Oak Ridge National Laboratory	
	Pumps		NUREG/CR-2886
	Valves		NUREG/CR-3154
	Electrical Power (Diesel Batteries, Chargers and Inverters)		NUREG/CR-1362
4.	Nuclear Plant Reliability Data System	Institute for Nuclear Power Operations	Quarterly Reports
5.	Reactor Safety Study Section III - LER Data for 1972-1973	USNRC	WASH-1400
6.	ATWS: A Reappraisal	Electric Power Research Institute	EPRI NP-2330
7.	Loss of Offsite Power at Nuclear Power Plants	Electric Power Research Institute	EPRI NP-2301
8.	Diesel Generator Reliability at Nuclear Power Plants	Electric Power Research Institute	EPRI NP-2433
9.	Classification and Analysis of Reactor Operating Experience Involving Dependent Events	Electric Power Research Institute	EPRI NP-3967
10.	PORV Failure Reduction Methods	Combustion Engineering	CEN-145

failure or unavailability rates for different components. Reports often differ as to assumptions with respect to number of demands, plant down-time or method of statistical analysis, and thus different reports can arrive at different failure rates using the same data base of failure events. Table 3 lists a number of such studies which have been found useful. They differ from the reports in category 1 in that sufficient information for reanalysis under different assumptions is not available.

The final category consists of compilations of "generic" component failure rates and associated estimates of uncertainty. These generic values are usually obtained by review of two or more category 1 or 2 sources, and may also include expert opinion on component failure rates or probabilities derived from other (non-nuclear) industrial experience. Table 4 lists the more important generic data bases in use today. The user is cautioned that these various generic data bases should never be construed as being independent, as in no case is this true. Further, considerable expert opinion has usually been used in choosing appropriate generic values.

IV.4 Data Analysis Methods Available

The techniques discussed earlier are methods of propagating the uncertainty of basic events through a system or accident sequence model to estimate the uncertainty of the top event. Methods of characterizing the uncertainty of the individual basic events of these models in terms of probability distributions are discussed here.

IV.4.1 Bootstrapping Method

This is an approach to generating a sampling distribution for one or more (constant) parameters in a failure model when only the results of a single sample are available. To illustrate this method, consider a binomial failure rate model, with data of x observed failures in n demands. Let p be the observed failure rate (x/n).

Random samples of the failure rate p are generated by randomly choosing a probability value P and solving for a value of y from the binomial probability distribution equation

$$P = \binom{n}{y} (p)^{y}(1-p)^{n-y} . \qquad (16)$$

Thus, we are using the observed values of n and p to generate random sample values of number of failures y, and thus random samples of the failure rate $p^* = y/n$. From these, a distribution on p can be generated.

Table 3

Statistical Analyses of Data

Title	Source	Reference
1. Probabilistic Safety Analysis of DC Power Requirements for Nuclear Power Plants	USNRC	NUREG-0666
2. Reliability Data Book	Swedish Nuclear Power Inspectorate	RKS85-25
3. Statistical Analysis of Nuclear Power Plant Pump Failure Rate Variability-Preliminary Results	Los Alamos National Laboratory	NUREG/CR-3650

In addition, items 2, 3, 5, 7, 8, 9, and 10 of Table 2 present analyses of reported data.

Table 4

Generic Failure Rate Data Bases

Title	Source	Reference
Reactor Safety Study	USNRC	WASH-1400
Interim Reliability and Evaluation Program (IREP) Procedures Guide	Sandia National Laboratories	NUREG/CR-2778
Accident Sequence Evaluation Program	Sandia National Laboratories	NUREG/CR-4550
Reliability Data Book	Swedish Nuclear Power Inspectorate	RKS85-25
Station Blackout Accident Analyses - TAP A-44	USNRC	NUREG/CR-3226

This bootstrapping process does yield a sampling distribution on the failure rate parameter which can be used in an uncertainty propagation procedure requiring parameter distributions as input. It has been found to lead to optimistic (conservative) results when dealing with components having few (or zero) failures in the data base used. The interpretation of the bootstrap-derived distributions may vary greatly depending on the sources of the data used in constructing the distributions. The method does not allow the impact of plant-specific data to be segregated and understood as is the case of the Bayesian approach and hence the latter is to be performed in those cases.

IV.4.2 Use of Plant-Specific Data in a Bayesian Analysis

In many older plants, data obtained from control room logs, maintenance request and event reports may be used to develop plant-specific estimates of failure rates, either for specific components or for a generic component category. Conversion of these point estimates into probability distributions can be accomplished utilizing Bayesian analysis in two ways depending on the extent of additional (industry wide) data available.

IV.4.2.1 Use of Non-Informative Priors

Ideally, sufficient data would exist so that statistical analysis of a specific component's actual failure record could be made so as to yield an experience-based estimate of a component's failure rate. A Bayesian analysis using a noninformative prior distribution (see IV.4.2.2) yields a characterization of the uncertainty in terms of a probability distribution. Such an analysis could be conducted on individual components, or on generic families of components, such as all check valves in a plant. All components or events of the same family would be modeled with the same distribution. Unfortunately, it is rare in nuclear power experience that plant-specific data exist in sufficient quantity to produce accurate parameter estimates i.e., narrow uncertainty intervals. However, the Bayesian technique should be employed by RMIEP whenever data is available in appropriate quantity.

IV.4.2.2 Bayesian Updating of Generic Data

Bayesian data analysis allows for incorporation of limited plant-specific data into past experience or knowledge as contained in generic data bases to yield improved understanding or updated distributions of basic event probabilities. Limited plant-specific data implies that some information on a specific component or family of components is available, but not enough for a full plant-specific data analysis as discussed in Section II.2.1. Bayesian updating utilizes Bayes' theorem which is discussed in Section IV.5.1.

In using Bayes' theorem, a basic event's probability may be modeled by a certain distribution based on past knowledge, perhaps from previous PRAs or other experience. This is referred to as the prior distribution of the parameter in question. Suppose also that there is some plant-specific data available for the basic event at hand. Bayes' theorem gives us a probabilistic relationship between the basic events prior distribution, the current data, and a new, or updated, distribution. This updated distribution is called the posterior distribution for the parameter, and represents a revision of the prior distribution based on the available plant-specific data. The posterior distribution is then used in the PRA models.

IV.4.3 Expert Opinion

Use of expert opinion or judgment is an approach to estimating failure characteristics based on tapping into unquantified experience. For example, one might seek failure rate data from personnel in related non-nuclear areas where accurate failure or repair records are not kept. Or one might seek to infer failure characteristics for one component (for which little or no actual data are available) from data on a component of similar design by appealing to the judgment of experts involved in the design and application of the two types of components. Or, in the case of simple mechanical failures, one might calculate a failure rate based on fatigue limit data for sub-pieces of the component in question. Here again, expert judgment is used to choose the most appropriate material data and calculational scheme.

The process of assembling expert opinion usually involves literature searches of past related experience and formal surveys of "experts" with applicable experience. Often, the results of formal surveys are combined, summarized and returned to the experts surveyed so that clarification or revisions are possible (as in a formal Delphi process). Generic lists of component failure probabilities or failure rates are a typical result of this method, such as printed in the Interim Reliability Evaluation Program (IREP) Procedures Guide[27] or the NUREG/CR-4550 document.[28] While expert opinion can yield good estimates of means and extreme values, assumptions must often be made regarding the shape of distributions. However, the assumptions on distribution shapes are frequently guided by expert opinion. Appendix C contains a summary paper prepared by Los Alamos National Laboratory which describes the results of an NRC-sponsored research program on the quantification of expert opinion, and presents recommendations for soliciting such information.

IV.5 Elements of the Bayesian Updating of Base Event Parameter Attributes

The Bayesian method allows for prior knowledge of a component's behavior to be revised with a limited amount of plant-specific

data. Experience from other power plants or data sources can be updated with the latest contribution to the data pool from the current component of interest. Use of Bayes' theorem to achieve this update ensures that the rules of probability theory are observed.

Bayes' theorem can be used to generate a distribution of a parameter's value for two different cases: (1) prior knowledge exists, and (2) no prior knowledge exists, or the analyst chooses not to use prior knowledge so as to avoid biasing the analysis. In the first case, a distribution of a parameter's value which reflects the analyst's beliefs about the parameter's possible values prior to collecting the current information is chosen. This distribution is called the prior distribution. Its form is not theoretically prescribed, but it is possible to select a prior distribution so that the mathematical calculations involved in the updating are not too complex. In the second case, the analyst chooses a "non-informative" prior distribution. In both cases, the prior distribution is used along with the current information on the parameter in Bayes' theorem to derive a new, updated distribution of the parameter's value. This new distribution is called the posterior distribution. The posterior distribution is used in whatever sampling technique is chosen for propagating component uncertainty through the system or sequence models.

IV.5.1 Bayes' Theorem

Bayes' theorem provides a medium by which some prior knowledge about a random variable can be updated using new data in a manner such that the rules of probability are observed. A prior probability distribution is defined for the parameter value in question. The parameter could be a demand failure rate for standby components, a time failure rate for standby components, or a time failure rate for components which must function for a certain period of time. The prior distribution will reflect the analysts opinion of the components behavior based on published analyses of similar failures at other plants. Licensee Event Reports (LERs), previous PRAs, and generic failure models are examples of sources of information used to define prior distributions. Any plant specific information is then used to update prior distributions.

Bayes' theorem can be expressed as

$$f(\lambda|\underline{x}) = \frac{L(x|\lambda)f(\lambda)}{\int L(\underline{x}|\lambda)f(\lambda)d\lambda} \qquad (17)$$

where

$L(\underline{x}|\lambda)$ = likelihood that, given a parameter value λ, a set of observations on some quantity x would result in the values given by a vector $\underline{x}$.

$f(\lambda|\underline{x})$ = posterior distribution on the values of the parameter λ, given that one has observed the values $\underline{x}$.

$f(\lambda)$ = prior distribution on the value of the parameter λ.

The form of the likelihood function is determined by the nature of the data, binomial (constant failure on demand model) or poisson (constant failure rate model). The form of the prior distribution is not necessarily prescribed. However, it is often advantageous to select a prior distribution which will simplify the mathematics of using Bayes' theorem. One such approach is to use conjugate priors, as discussed in Section II.1. Examples of applying Bayes' theorem to constant-failure-rate and constant failure on demand models are shown below.

IV.5.2 Constant Failure Rate Model

Many component failures are well described by a model that estimates the frequency at which failures will occur. The time period over which a component's failure probability is of interest may be short--the length of time needed to mitigate an accident condition, or long--the time between testing of standby components such as normally open valves. Furthermore, the rate at which a component will fail can be thought of as having a constant value. Unfortunately, the knowledge associated with the component's behavior is usually incomplete, and the exact value for the failure rate is not known. Data on failures of similar components in other systems and other nuclear plants may yield a probabilistic range for the value of the failure rate. This probability distribution is derived from information on other components of similar design, but different locations, environments, and maintenance histories. Even so, it serves as a basis to model the failure rate of a particular component. Bayes' theorem provides a mechanism by which limited plant-specific data on a component can be incorporated into a distribution characterized by a large body of data pooled from other plants.

IV.5.2.1 Bayes' Theorem with Conjugate Priors

Suppose that current experience has shown that a particular component type at a nuclear power plant has failed r times in a total time period of T. The likelihood function for this type of data is:

$$L(r|T,\lambda) = \frac{(\lambda T)^r}{r!} \exp(-\lambda T) \qquad (18)$$

The prior distribution for the component failure rate, λ, when combined with this likelihood function in Bayes' theorem, will yield a posterior distribution of λ of similar form as the prior is the gamma distribution. Suppose that past experience at other plants shows α failures in a total time period β. The prior distribution would be a gamma distribution with parameters α and β:

$$f(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda) \quad , \tag{19}$$

where $\Gamma(\alpha)$ is the gamma function. The resulting posterior distribution is also a gamma distribution with parameters $\alpha+T$ and $\beta+T$:

$$f(\lambda|\underline{x}) = \frac{(\beta+T)(\alpha+T)}{\Gamma(\alpha+T)} \lambda^{(\alpha+T-1)} \exp\left[-(\beta+T)\lambda\right] \quad . \tag{20}$$

The initial model of the failure rate distribution has thus been adjusted to account for the new information at hand.

While conjugate priors are attractive from a mathematical viewpoint, they are only useful when the existing data justifies the use of a distribution of conjugate form. Mathematical simplicity should not be the primary criterion of selection.

IV.5.2.2 Bayes' Theorem with Non-Informative Priors

If no additional information other than limited plant-specific data exists, or if the analyst does not wish to bias the analysis with outside information, a non-informative prior distribution can be chosen. A non-informative prior is merely a mechanism by which plant-specific data can be converted into a distributional form, and does not necessarily represent any engineering judgment or experience. As with other priors, mathematical simplicity can be enhanced with conjugate distributions.

A gamma prior distribution with parameters $\alpha = \beta = 0$ can be used as a conjugate non-informative prior* for the constant failure rate model, incorporating r failures in the time period T.

---

*It should be noted that the form of the appropriate non-informed prior does depend on the censoring scheme for sampling data as detailed in Reference 29.

$$f(\lambda) = 1/\lambda \quad . \tag{21}$$

The resulting posterior distribution is also a gamma distribution with parameters r and T.

IV.5.2.3 Bayes' Theorem with Lognormal Priors

In practice, mathematically convenient priors which are conjugates of the likelihood functions of the data are not necessarily an adequate characterization of past experience. More often, the lognormal distribution is used to represent uncertainty. This distribution is not a conjugate with Poisson data, but is very useful for many components. Although the mathematics of Bayes' theorem with lognormal priors is not as simple as with the conjugate priors, simple numerical integration techniques can be used to generate an empirical posterior density function, $f(\lambda|r,T)$ from Bayes' theorem:

$$f(\lambda|r,T) = \frac{\lambda^{r-1} e^{-\lambda T} \exp\left\{-\frac{1}{2}\left[\frac{\ell n\lambda-\mu}{s}\right]^2\right\}}{\int_0^\infty x^{r-1} e^{-xT} \exp\left\{-\frac{1}{2}\left[\frac{\ell nx-\mu}{\sigma}\right]^2\right\} dx} \tag{22}$$

where

μ = mean of the natural logarithms of the prior data

σ = the standard deviation of the natural logarithms of the prior data

r = number of failures observed

T = component exposure time

This expression is easily evaluated. Thus, it is seen that prior distributions need not fit conveniently into conjugate forms in order to benefit from applications of Bayes' theorem.

IV.5.3 Constant-Failure-Probability on Demand Model

Many component failure modes are characterized by a constant probability of failure given that the component is demanded to function in the appropriate mode. Examples of such component failure modes are pumps failing to start upon receiving initiation signals and valves which must change state for system operation. Just as with constant failure rate models, the knowledge of a component's failure probability on demand is often incomplete. Limited plant-specific data may not be

sufficient to estimate a component's behavior with a high degree of confidence. Information pooled from other sources can be used to derive a probability distribution indicative of the behavior of similar component failures. This probability distribution does not represent exactly the possible range of a particular component's failure probability, since plant-to-plant differences are incorporated into it. However, it does provide a foundation for generating a distribution using the limited plant-specific data.

IV.5.3.1 Bayes' Theorem with Conjugate Priors

Suppose the current information on a component's behavior is in the form of n failures in N demands. This is binomial data, and the likelihood function used is:

$$L(n|N,p) = \frac{N!}{n!(N-n)!} p^{n}(1-p)^{N-n} \qquad (23)$$

The conjugate prior for the failure probability, p, for this likelihood function is the beta distribution. If past experience has shown ℓ failures in m demands, the prior distribution is:

$$f(p) = \frac{\Gamma(\ell + m)}{\Gamma(\ell)\Gamma(m)} p^{\ell-1}(1-p)^{m-1} \qquad (24)$$

The resulting posterior distribution is also a beta distribution, with parameters $\ell+n$ and $m+N-n$.

IV.5.3.2 Bayes' Theorem with Non-Informative Priors

As with constant failure rate models, limited plant-specific data of constant probability on demand can be converted to a distribution even without the use or existence of other information. Mathematical simplicity can be enhanced by selecting a non-informed conjugate prior distribution for the available date.

The beta distribution with parameters 0 and 0 is a commonly used non-informative prior:

$$f(p) = \frac{1}{p(1-p)} \quad . \qquad (25)$$

The resulting posterior distribution is also a beta distribution with the parameters n and N-n.

IV.5.3.3 Bayes' Theorem with Lognormal Priors

Just as with constant failure rate models, constant demand probability information often can not justify the use of conjugate priors. The lognormal distribution has been used frequently to represent uncertainty in practice. Numerical integration techniques can be used to solve for an empirical posterior density function, f(p), from Bayes' theorem:

$$f(p|x,n) = \frac{p^{x-1}(1-p)^{n-x} \exp\left\{-\frac{1}{2}\left[\frac{\ell np-\mu}{\sigma}\right]^2\right\}}{\int_0^1 y^{x-1}(1-y)^{n-x} \exp\left\{-\frac{1}{2}\left[\frac{\ell ny-\mu}{\sigma}\right]^2\right\} dy} \qquad (26)$$

where

μ = mean of the natural logarithms of the prior data

σ = standard deviation of the natural logarithms of the data

x = number of failures observed

n = number of demands

Again, this expression is easily evaluated. Thus, again it is seen that there is no need to sacrifice accuracy of models for mathematical simplicity.

IV.5.4 Development of Prior Distributions from Subjective Opinion

In the above two examples, the parameters of the prior distributions were assumed to be from past information on component failures. However, methods exist for incorporating subjective opinion of component behavior in lieu of data.[30] For example, suppose our prior knowledge of a component's demand rate, p, is such that we believe there is only a five percent chance that the value of p will exceed some value, p*. A mean value of p is also assumed to be known. This pair of quantile estimates of p can be used to generate the parameters of an assumed probability distribution. Tables exist which relate these quantile estimates to appropriate values of parameters for distributions such as the beta and gamma.[31] Thus, an analyst's subjective opinion of an event's probability attributes such as mean and 95th quantile can be used to generate a prior distribution for the event's failure probability.

V. CONCLUSIONS

This report has presented the results of a brief review of available uncertainty analysis methods and described the Latin Hypercube sampling approach (as embodied in the TEMAC code) that is recommended for use in the RMIEP probabilistic risk analysis of the LaSalle plant. One important aspect of this approach, which was an important consideration in its selection, is its comprehensive sensitivity measure option. It should be remembered that uncertainty propagation is, of itself, of limited interest since it is essentially an integration function. The ability to perform sensitivity studies through the use of the importance measures and the ability to identify the parameter estimates whose uncertainties contribute most to the uncertainty in the end result are a key feature of the approach adopted here. In addition the method treats modeling uncertainty in a manner compatible with that proposed for the level 2 and 3 parts of the analysis. Thus the approach to be adopted has the potential for enhancing the understanding and analysis of the results of the Level I PRA both as a stand-alone analysis and as input to a Level 2 and 3 PRA.

The TEMAC code has recently been successfully applied in an uncertainty analysis of the Peach Bottom plant as part of the Accident Sequence Evaluation Program.[32] This application showed that this approach is both effective and relatively inexpensive with respect to computer cost. Based on the above considerations, the use of the TEMAC code and Latin hypercube sampling was deemed the optimal analysis tool for use in RMIEP.

VI. REFERENCES

1. Kolb, G. J.. "Overview and Status of the LaSalle Nuclear Plant Probabilistic Risk Assessment," Proceedings of the 8th International Structural Mechanics in Reactor Technology Conference, August 16-25, Brussels, Belgium, 1985.

2. Helton, J. C., et al., "Detegration of the NUREG-1150 Analysis: Calculation of Risk and Propagation of Uncertainties, paper prescribed at 14th Water Reactor Safety Information Meeting, October 26-30, 1987, Gaithersburg, MD.

3. Parry, G. W. and Winter, P. W., "Characterization and Evaluation of Uncertainty in Probabilistic Risk Analysis," Nuclear Safety, Vol. 22, No. 1, 1981.

4. Hahn, G. J. and Shapiro, S. S., Statistical Models in Engineering, (New York, NY: John Wiley & Sons, Inc., 1967).

5. Mosleh, A., "Hidden Sources of Uncertainty: Judgement in Collection and Analysis of Failure Data," Nuclear Engineering and Design, 93, pp. 187.

6. Parry, G. W., "Incompleteness in Data Bases: Impact on Parameter Estimation Uncertainty," Presented at 1984 Annual Meeting at the Society for Risk Analysis, Knoxville, TN; September 30-October 3, 1984.

7. U.S. Nuclear Regulatory Commission, "Reactor Safety Study: An Assessment of Accident Risks in US Nuclear Power Plants," WASH-1400 (NUREG 75/014) 1975.

8. Apostolakis, G., Kaplan, S., Garrick, B. J., and Duphily, R. J., "Data Specialization for Plant Specific Risk Studies," Nuclear Engineering and Design, 56, 1980.

9. Kaplan, S., "On a 'Two-Stage' Bayesian Procedure for Determining Failure Rates from Experimental Data," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No. 1, January 1983.

10. Spencer, F. W., and Easterling, R. G., Use of the Maximus Methodology for Confidence Bound Calculations in Fault Trees -- Trial Problem, Sandia National Laboratories, Albuquerque, NM, draft, 1984.

11. Lloyd, D. K., and Liporo, M., Reliability: Management, Methods and Mathematics, (NJ: Prentice Hall, 1962).

12. Hora, S. C., Iman, R. L., A Comparison of Maximus/ Bounding and Bayes/Monte Carlo for Fault Tree Uncertainty Analysis, SAND85-2939, Sandia National Laboratories, Albuquerque, NM, March, 1986.

13. Martz, H. F. et al, A Comparison of Methods for Uncertainty Analysis of Nuclear Power Plant Safety System Fault Tree Models, NUREG/CR-3263, Los Alamos National Laboratory, Los Alamos, NM, April 1983.

14. Jackson, P. S., Hockenbury, R. W., and Yeater, M. L., "Uncertainty Analysis of system Reliability and Availability Assessment," Nuclear Engineering Design, 68, May 29, 1981.

15. Kaplan, S., "On the Method of Discrete Probability Distributions in Risk and Reliability Calculation - Application to Seismic Risk Assessment," Risk Analysis 1, 189-196, 1981.

16. Iman, R. L. Shortencarier, M. J. A FORTRAN 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use with Computer Models, NUREG/CR-3624, SAND83-2365, Sandia National Laboratories, Albuquerque, NM, March 1984.

17. Iman, R. L., Shortencarier, M. J., A User's Guide for the Top Event Matrix Analysis Code (TEMAC), NUREG/CR-4598, SAND86-0960, Sandia National Laboratories, Albuquerque, NM, August 1986.

18. Apostolakis, G. and Kaplan, S., "Pitfalls in Risk Calculations," Reliability Engineering, 2, pp. 135-145, 1981.

19. Veseley, W., Measures of Risk Importance and Their Application, Idaho National Engineering Laboratory, Idaho Falls, ID, NUREG/CR-3385

20. Amos, C. N., et al, Evaluation of Severe Accident Risks and Potential for Risk Reduction: Peach Bottom, Unit 2, NUREG/CR-4551, Vol. 3, Sandia National Laboratories, Albuquerque, NM, February 1987.

21. Mosleh, A., et al., Procedures for Treating Common Cause Failures in Safety and Reliability Studies, U.S. Nuclear Regulatory Commission, Washington, DC, NUREG 4780, 1987.

22. Martz, H. F., Beckman, R. J. and McInteer, C. R., 1982, FRAC (Failure Rate Analysis Code): A Computer Program for Analysis of Variance of Failure Rates, Los Alamos National Laboratory, Los Alamos, NM, NUREG/CR-2433.

23. Whitehead, D. W., Weston, L. M., and Graves, N. L., Recovery Actions in PRA for the Risk Methods Integration and Evaluation Program (RMIEP); Vol 1: Development of the Data-Based Method; Vol 2: Application of the Data-Based Method, Sandia National Laboratories, Albuquerque, NM, NUREG/CR-4834, 1987.

24. Fleming, K. N., A Reliability Model for Common Mode Failures in Redundant Safety Systems, General Atomics Report, GA-A13284, 1975.

25. Atwood, C. L., Estimators for the Binomial Failure Rate Common Cause Model, Idaho National Engineering Laboratory, Idaho Falls, ID., NUREG/CR-1401, 1980.

26. Fleming, K. N. and Kalinowski, A. M., An Extension of the Beta Factor Method to Systems with High Level of Redundancy, Pickard, Lowe and Garrick Report, PLG-0289, 1983.

27. Carlson, D. D., et al., Interim Reliability Evaluation Program Procedures Guide, NUREG/CR-2728, SAND82-1100, Sandia National Laboratories, Albuquerque, NM, January 1983.

28. Drouin, M. T., et al., Analysis of Core Damage Frequency from Internal Events: Grand Gulf, Unit 1, NUREG/CR-4550, SAND86-2084, Volume 5, Sandia National Laboratories, Albuquerque, NM, February 1987.

29. Martz, H. F., Walker, R. A., Bayesian Reliability Analysis, (New York, NY: Wiley & Sons, 1982).

30. Cook, I. and Unwin, S. D., Controlling Principles for Prior Probability Assessments in Nuclear Risk Assessments, Sandia National Laboratories, Albuquerque, NM, NUREG/CR-4514, 1986.

31. Abramowitz, M. and I. Stegan, Handbook of Mathematical Functions, Dover, 1965.

32. Kolaczkowski, A. M., et al., Analysis of Core Damage Frequency from Internal Events: Peach Bottom, Unit 2, NUREG/CR-4550, SAND86-2084, Vol. 4, Sandia National Laboratories, Albuquerque, NM, October 1986.

Appendix A

Interface Between Core Damage Frequency Analysis (RMIEP), the Accident Progression and Containment Failure Analysis (PRUEP) and the Health and Economic Consequence Analysis

As has been indicated, the uncertainty and sensitivity analyses for the accident sequence and core damage frequency calculations will be performed using the TEMAC code and limited Latin hypercube sampling techniques. The analysis of accident sequence frequencies and core damage frequencies is often referred to as the "front end" analysis.

To complete the full scope PRA, it is necessary to characterize the accident progression following core damage, and to compute fission product radionuclide release, and estimate the health and economic consequences of each accident scenario. These calculations are typically referred to as the "back end" analysis, and an uncertainty and sensitivity analysis will also be performed here. Uncertainty analysis methods to be used in RMIEP for the LaSalle back end analysis are still evolving, and work is continuing into 1988.

It is currently planned that the RMIEP front end analysis will be performed and reported as an entity in itself. Then the back end analysis for RMIEP will be performed and reported as a separate entity. Given the planned sequential nature of the front and back end analyses, it was necessary to examine the interfaces between the front and back analyses to assure that no incompatibilities would be present, and that the data needed for the back end uncertainty calculations would be available in the proper form. This section discusses the information flow and compatibility between the front and back end analyses, and demonstrates that the uncertainty analysis data and methods to be used in the front end analysis are both compatible and complementary with those being used for the back end analysis. Areas are highlighted where careful attention to the interface is required.

Figure A-1 shows a block diagram representation of the four main steps in performing a full scope (Level III) PRA. The first block (accident sequence and core damage analysis) represents the front end calculations (the RMIEP analysis for LaSalle). The remaining three blocks together constitute the back end analysis (as will be analyzed in the PRUEP program for LaSalle). Each of these blocks of analysis and their interface requirements is discussed below.

Figure A-2 shows the input and output of the front end accident sequence analysis. The inputs consist of the plant system descriptions and operating procedures, plus the failure rate data which was discussed previously in Section IV. The output consists of the dominant cut sets for each of the accident sequences defined by the event trees, plus the frequencies and uncertainties of the accident sequences themselves. Although there are a variety of point estimate results, uncertainty analyses, and sensitivity studies which result from the front end analysis, there are really only two pieces of output which are input to the back end calculations, namely,

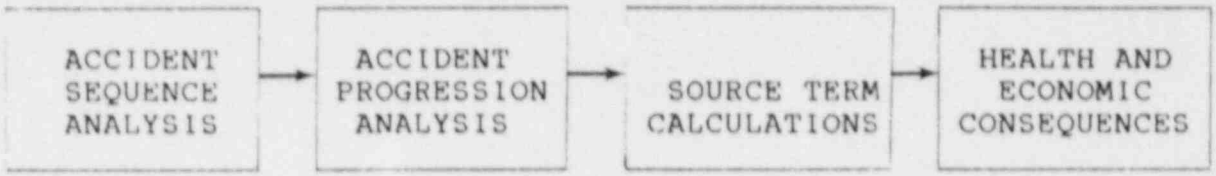

Figure A-1. Major Steps in a Level III PRA Analysis of Reactor Accidents

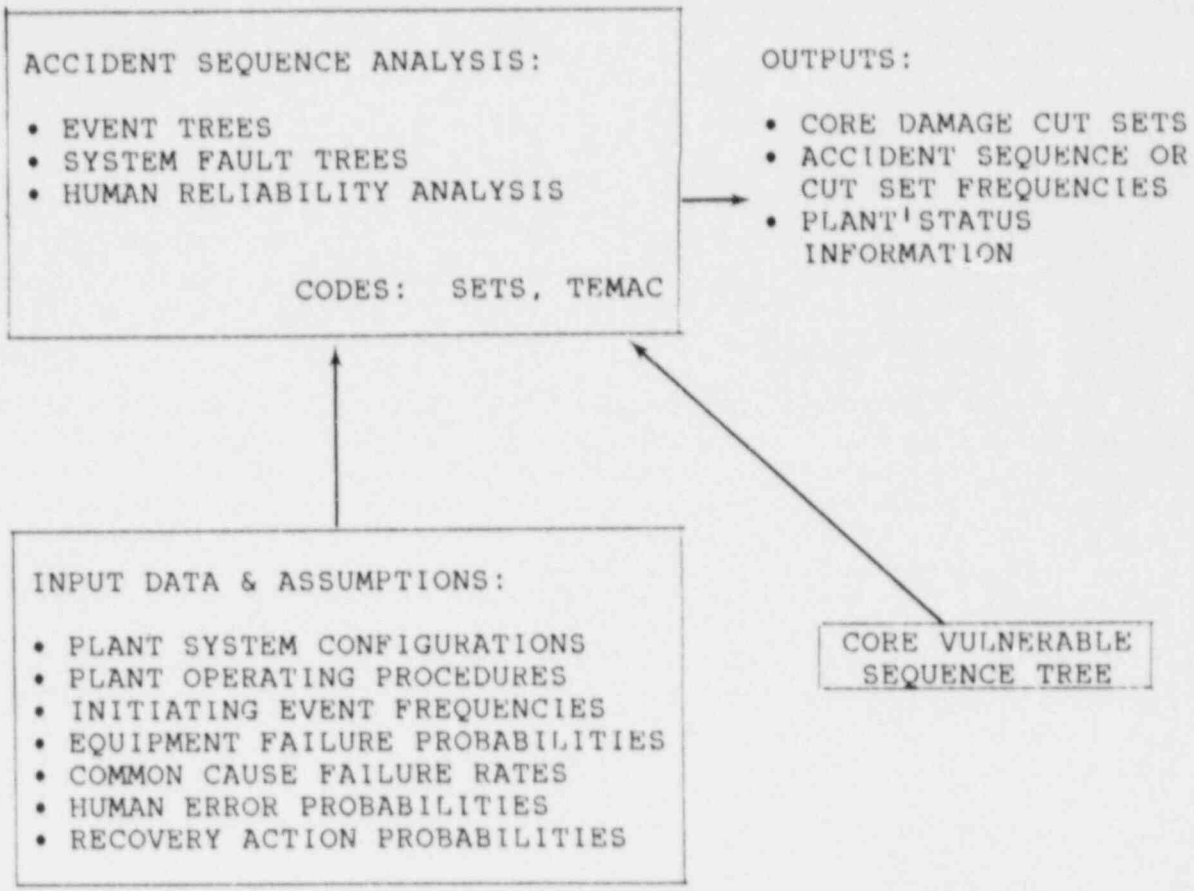

Figure A-2. Accident Sequence Analysis - Input and Output

- Core damage cut sets
- Plant safety systems status information implied by each cut set

Of course, all the component failure and test and maintenance outage data plus their associated uncertainty characterizations must also be available to the back end analysts. But the only front end result which is passed to the back end is a listing of the dominant core damage cut sets. Working together, the front end and the back end analysts will regroup these cut sets into plant damage states.

There is one other aspect of this interface which requires further description, and that has to do with those front end accident sequences which are denoted as "core vulnerable" rather than "core damage" states. In these accident sequences, the question of whether or not the core is damaged depends on how the accident progression affects the functioning of safety systems operating to prevent core damage. To determine which cut sets in the core vulnerable accident sequences lead to core damage, an additional small event tree is appended to these sequences. This tree contains phenomenological alternatives (i.e., does the containment fail, etc.). This tree and associated split fractions (probabilities) associated with each option will be provided by the back end analysis team prior to final accident sequence cut set determination by the front end team.

Figure A-3 shows the inputs and outputs to the accident progression calculations. The accident sequence cut sets provided by the front end analysis are first resorted into plant damage states. All cut sets assigned to a given plant damage state present the same initial and boundary conditions to the accident progression analysis, and thus are propagated together through the accident progression event tree. Cut sets from a given accident sequence may be assigned to different damage state bins. The definitions of the various damage states are determined by the first few (ten to 20) questions on the accident progression event tree (which is developed for each plant) and to some degree, with consideration of the important accident sequences identified in the front end analysis. An example of the questions which define the damage states (taken from the draft Grand Gulf PRA NUREG-1150) is given in Table A-1. Taken together, it was found on the basis of these questions that eight damage states could be defined which would cover all the dominant internal event cut sets. Definition of these plant damage states requires a significant amount of interaction with the internal event (front end) analysts to determine the plant systems status for each cut set.

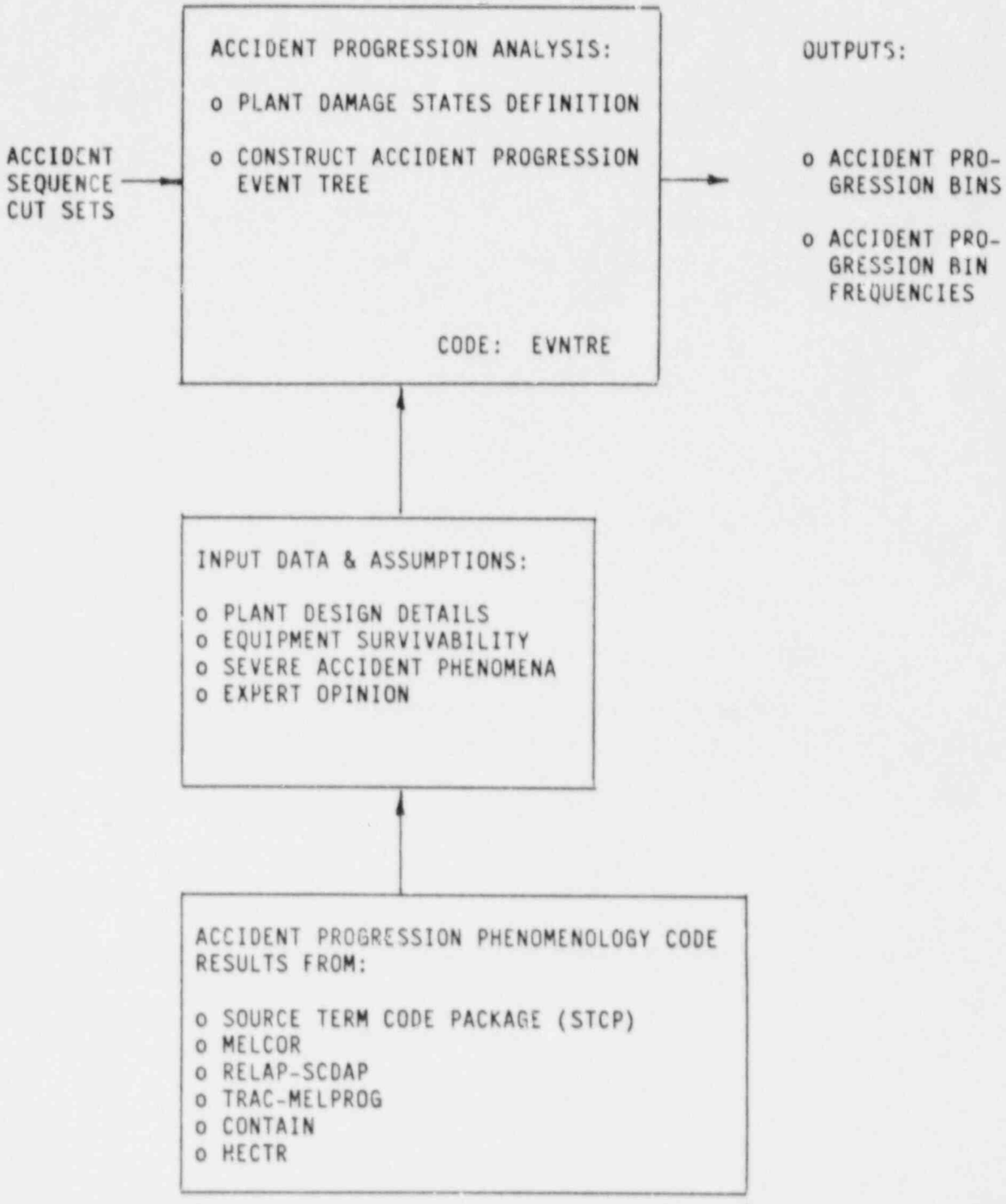

Figure A-3. Accident Progression Analysis - Input and Output

Table A-1

Initial Questions on the Grand Gulf Accident Progression Event Tree Used to Define Plant Damage States

	Questions	Prior Question Dependencies
1.	What is the initiating event?	None
2.	What is the level of preexisting leakage or isolation failure?	None
3.	What is the level of preexisting suppression pool bypass?	None
4.	Is there a station blackout (diesel generators fail)?	None
5.	Is DC power not available?	None
6.	For TC accident sequence, does SLC fail to inject?	None
7.	Does HPCS fail to inject?	None
8.	Does RCIC fail to inject?	None
9.	Does the CRD hydraulic system inject?	None
10.	Does the condensate system fail?	None
11.	Do the LPC and LPCI systems fail?	1,4,7,8
12.	Does RHR fail (heat exchangers not available)?	1,4,11
13.	Does the service water cross-tie to LPCI fail?	1,4,11
14.	Are the containment (wetwell) sprays failed?	1,4,11
15.	Is ADS inhibited or failed?	None
16.	Does the RPV remain pressurized?	1,5,7,8,15
17.	What type of sequence is this (summary of plant damage)?	1,7,8,16

The remainder of the accident progression analysis makes use of the full accident progression event tree (APET), which has on the order of 50 to 100 branch points at which different possibilities of accident progression are evaluated as a function of the plant systems status and the various phenomenological possibilities and their associated probabilities (split fractions). Although a logic tree having 50 to 100 branches has potentially millions of separate branches, these outcomes are grouped into accident progression bins. As with the plant damage states, each accident bin presents one unique set of initial and boundary conditions for the source term calculations of Block 3. Table A-2 is a (partial) list of typical BWR source term attributes. Each branch of the containment event tree is numerically evaluated, and each path through the tree corresponds to a particular set of values for the source term attributes. A vast majority of these accident paths are deleted from further consideration based on

a) negligible frequency of occurrence, and

b) negligible consequence of the resulting source terms

(At least one low consequence bin is always retained to represent those accident sequences with no containment failure, and hence, low consequence.) The remaining non-negligible accident paths are then grouped according to the source term attributes, and these groupings constitute the accident sequence bins. Typically, 50 to 200 individual accident bins result from these groupings for each plant damage state.

The accident progression event tree is a logical enumeration of all the initial and boundary conditions for the accident progression as well as all the potential accident paths identified by results of calculations using more detailed computer code calculations. The codes used in the phenomenological calculations are shown as an input on Figure A-3. The containment event tree is the vehicle for incorporating all the insights gained from these detailed and sometimes costly code analyses of accident paths into a model which can be efficiently evaluated for use in uncertainty and sensitivity studies.

Numerical evaluation of the accident progression event tree is performed using the computer code EVNTRE. Three important aspects of this code are that

a) Multiple branches are allowed at each node of the tree.

b) Branch probabilities may depend on paths taken through the tree.

c) Proper correlation between dependent probabilities can be incorporated.

Table A-2

Typical BWR Source Term Attributes Used
to Define Accident Progression Bins

Initiating Event Type

Time to Core Cooling Failure - <2 hr., <10 hr., >10 hr.

RPV Pressure at Vessel Breach -- High, Intermediate, Low

Containment Failure Time --

- Before Core Degradation
- During Core Degradation
- At Vessel Breach
- Late

Containment Leak Level -- Four sizes considered

Pool Bypass Level -- Three levels considered

Pool Temperature -- Subcooled, Saturated

Reactor Building Bypass -- No, Yes

Core Concrete Interaction -- No, Yes with Water, Yes, but dry

Containment Sprays -- Always, Never, Early Not Late, Late Not Early

The output of the accident progression analysis consist of frequencies of the accident progression bins.

Figure A-4 shows the input and output of the source term analysis block of the back end analysis. In this analysis, the flow of fission products released from the fuel is modeled and analyzed. The numerical analysis will be performed using the RELTRAC code which utilizes a set of time-dependent rate equations to model transport, suspension and deposition of fission products from the fuel, through the reactor vessel and primary piping, the containment and subsequent release to the environment. These rate equations are based on the results of core damage and fission product release calculations and experiments. The rates used in RELTRAC reflect insight and results from calculations using the more time consuming and detailed codes such as MELCOR. The use of the simpler rate equation formulation in RELTRAC results in a code which can be efficiently used in a sampling mode.

The RELTRAC code is run for each of the accident progression bins. The result (for each bin) is a vector of released radionuclide mass as a function of time for each of the nine radionuclide groups shown on Table A-3. These nine time histories (resulting from each accident progression bin) are then further collapsed to a vector of three numbers:

$$\begin{Bmatrix} \text{EEER} \\ \text{IEER} \\ \text{LR} \end{Bmatrix}$$

where

EEER = equivalent mass of I_2 released in the first two hours after warning

IEER = equivalent mass of I_2 released following two hours after warning

LR = equivalent mass of Cs released

The condensation from the vector of nine time histories of mass release to three constant values of release utilizes studies reported in NUREG/CR-4467 in which the relative importance of the various fission product elements and their isotopes were studied with respect to their radiobiological effect, and equivalences were established in terms of equivalent masses of iodine (early biological effect) and cesium (late biological effect).

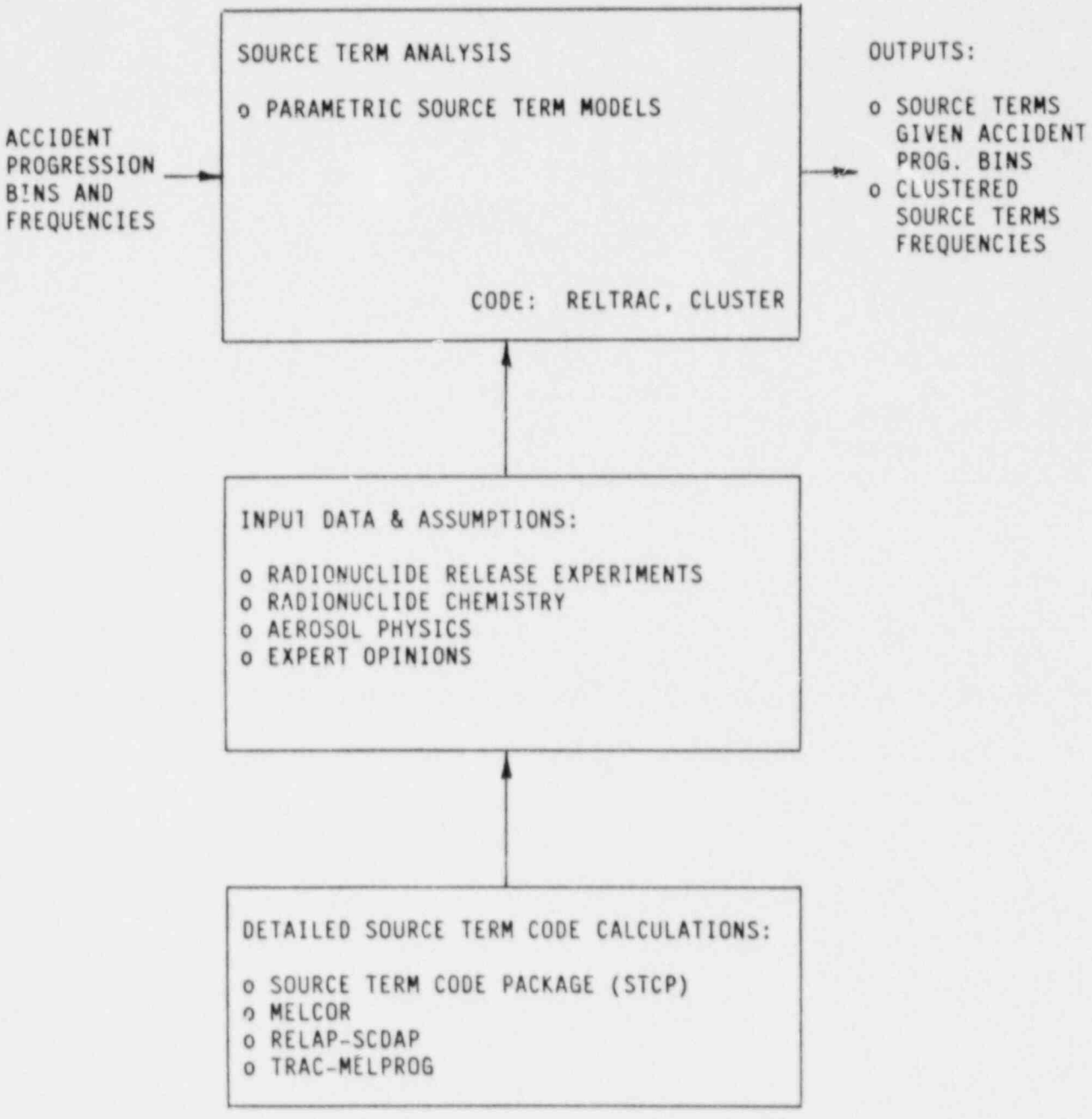
SOURCE TERM ANALYSIS
o PARAMETRIC SOURCE TERM MODELS
CODE: RELTRAC, CLUSTER
ACCIDENT PROGRESSION BINS AND FREQUENCIES
OUTPUTS:
o SOURCE TERMS GIVEN ACCIDENT PROG. BINS
o CLUSTERED SOURCE TERMS FREQUENCIES
INPUT DATA & ASSUMPTIONS:
o RADIONUCLIDE RELEASE EXPERIMENTS
o RADIONUCLIDE CHEMISTRY
o AEROSOL PHYSICS
o EXPERT OPINIONS
DETAILED SOURCE TERM CODE CALCULATIONS:
o SOURCE TERM CODE PACKAGE (STCP)
o MELCOR
o RELAP-SCDAP
o TRAC-MELPROG

Figure A-4

Table A-3

Radionuclide Groups

Group	Elements
1	Xe, Kr
2	I, Br
3	Cs, Rb
4	Te, Sb, Se
5	Sr
6	Ru, Rh, Pd, Mo, Tc
7	La, Zr, Nd, Eu, Nb, Pm, Pr
8	Ce, Pu, Np
9	Ba

Thus the source term analysis results in a release effect vector {EEER, IEER, LE} of three numbers for each accident progression bin identified as important in the accident progression analysis. When uncertainties in parameters, models and timing are included, a number of different vectors are obtained for each accident progression bin. These different vectors may be viewed as different observations for a Monte Carlo analysis. Thus, if we have 50 accident progression bins and 100 observations (i.e., resulting vectors) for each bin, then 5000 different release effect vectors result. Each one of these vectors could form the input to a single health/economic consequence calculation (using the MACCS code or equivalent). However, to reduce the number of such calculations, the release effect vectors are grouped into clusters, and a single health/economic consequence calculation is performed for each cluster. The cluster analysis is performed by considering each release effect vector as a point in three dimensional space with coordinates {EEER, EELR, LE}, and using the CLUSTER code to identify groups of points which are closely spatially related. Then, a single vector of fission products which characterizes the cluster is used as input to the consequence calculations.

The health and economic consequence calculation block is shown in Figure A-5. The input is the source term vector described above, as well as the site and environment data shown on the figure. The code to be used for the RMIEP application is the MELCOR Accident Consequence Code System (MACCS), which calculates estimates of population dose and health effects as well as associated costs. Repeating the MACCS calculations for each input source term vector yields distributions on the output values which characterize the uncertainty in the back end calculation of core damage accident consequences.

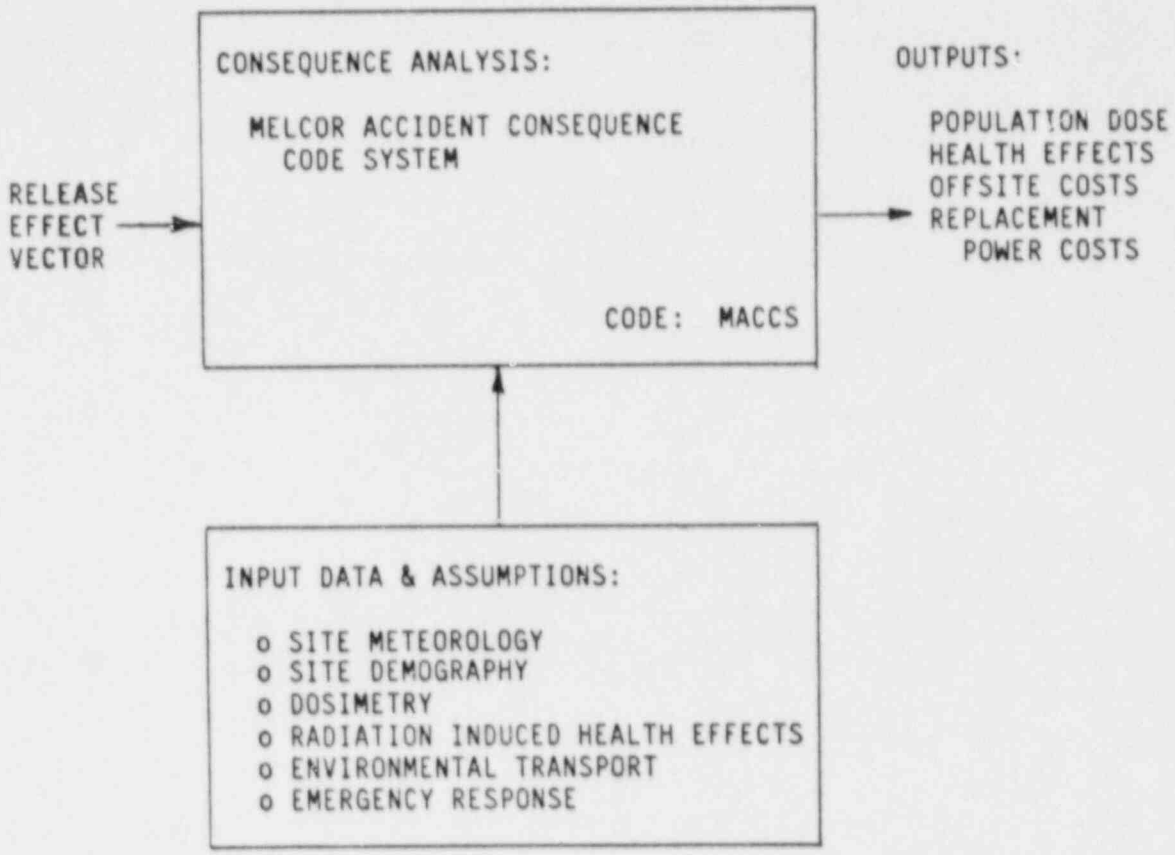

Figure A-5. Consequence Analysis - Input and Output

Appendix B

Latin Hypercube Sample Demonstration

Latin Hypercube Sample Demonstration

The following discussion and demonstration of Latin hypercube sampling are taken directly from the LHS users guide.

> Iman, R. L., Shortencarier, M. J., 1984, A FORTRAN 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use with Computer Models, NUREG/CR-3624, SAND83-2365, Sandia National Laboratories, Albuquerque, NM.

This appendix describes the latest version of a computer program for the generation of multivariate samples either completely at random or by a constrained randomization termed Latin hypercube sampling (LHS). This program has been developed at Sandia National Laboratories and replaces the previous program described in Iman, Davenport, and Zeigler (1980). Every attempt has been made to make the present program portable and user-friendly while, at the same time, expanding the capability of the program to include additional sampling distributions.

The situation addressed by the computer program is the following. There is a variable of interest, Y, that is a function of other variables X_1, X_2, ..., X_K. This function may be quite complicated, for example, a computer model. A question to be investigated is: How does Y vary when X's vary according to some assumed joint probability distribution? Related questions are: What is the expected value of Y? What is the 99th percentile of Y? etc.

A conventional approach to these questions is Monte Carlo. By sampling repeatedly from the assumed joint probability density function of the X's and evaluating Y for each sample, the distribution of Y, its mean, percentiles, etc., can be estimated. This is one option provided by the program for generating the X's. The program output, say for n Monte Carlo repetitions, is a set of k-dimensional vectors of input variables.

An alternative approach, which can yield more precise estimates, is to use a constrained sampling scheme. One such scheme, developed by McKay, Conover, and Beckman (1979), is Latin hypercube sampling (LHS). LHS selects n different values from each of k variables X_1, ..., X_K in the following manner. The range of each variable is divided into n nonoverlapping intervals on the basis of equal probability density in the interval. The n values thus obtained for X_1 are paired in a random manner (equally likely combinations) with n values of X_2. These n pairs are combined in a random manner with the n values of X_3 to form n triplets, and so on, until n k-tuplets are formed. This is the Latin hypercube sample. It is convenient

to think of the LHS, or a random sample of size n, as forming an n x k matrix of input where the ith row contains specific values of each of the k input variables to be used on the ith run of the computer model.

The LHS technique has been applied to many different computer models since 1975. The results of an application of LHS to a large computer model can be found in Steck, Iman, and Dahlgren (1976). A more detailed description of LHS with application to sensitivity analysis techniques can be found in Iman, Helton, and Campbell (1981, 1981b). A tutorial on LHS may be found in Iman and Conover (1982b). A comparison of LHS with other techniques is given in Iman and Helton (1983).

To help clarify how intervals are determined in the LHS, consider a simple example where it is desired to generate a LHS of a size n = 5 with two input variables. Let us assume that the first random variable X_1 has a normal distribution concentrated on the range from A to B. In this program, the following interpretations (not subject to change by the user without modifying the code) are given to A and B for both the normal and lognormal distributions, namely

$$P(X_1 \leq A) = .001 \text{ and } P(X_1 \leq B) = .001,$$

where P(E) denotes the probability of event E. That is, A is defined as the .001 quantile and B is defined as the .999 quantile of the distribution of X_1. Thus, $P(A \leq X_1 \leq B) = .998$, so both the normal and lognormal distributions are truncated slightly in the program. That is, the sampling procedure excludes values outside the interval [A, B]. These definitions of A and B imply that the mean of the normal distribution is given by $\mu = (A + B)/2$ and since for a standardized normal variable Z,

$$P(Z \leq -3.09) = .001,$$

it follows that the standard deviation of the desired truncated normal distribution is given (to a close approximation) by

$$\sigma = (B - \mu)/3.09 = (B - A)/6.18.$$

With the parameters μ and σ thus defined, the endpoints of the intervals are easily determined. The intervals for n = 5 are illustrated in Figure B-1 in terms of both the density function and the more easily used cumulative distribution function (cdf). If the distribution were not truncated, then the intervals in Figure B-1 would satisfy

$$P(A \leq X_1 \leq C) = P(F \leq X_1 \leq B) = .198$$

$$= P(C \leq X_1 \leq D) = P(D \leq X_1 \leq E) = P(E \leq X_1 \leq F) = .2.$$

To account for truncation requires dividing these probabilities by .998. Thus, for all practical purposes, the five intervals correspond to 20% probability.

We shall assume in this example that the second random variable, X_2, has a uniform distribution on the interval from G to H. The corresponding intervals used in the LHS for X_2 are given in Figure B-2 in terms of both the density function and the cdf.

The next step in obtaining the LHS is to pick specific values of X_1 and X_2 in each of their five respective intervals. This selection should be done in a random manner with respect to the density in each interval; that is, the selection should reflect the height of the density across the interval. For example, in the (A,C) interval for X_1, values close to C will have a higher probability of selection than will those values close to A. Next, the selected values of X_1 and X_2 are paired to form the required five input vectors. In the original concept of LHS as outlined in McKay, Conover, and Beckman

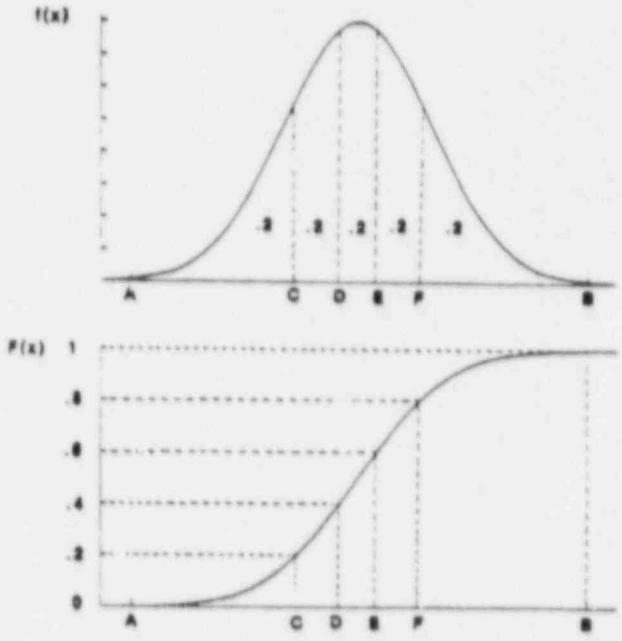

Figure B-1. Intervals Used with a LHS of Size n = 5 in Terms of the Density Function and Cumulative Distribution Function for a Normal Random Variable

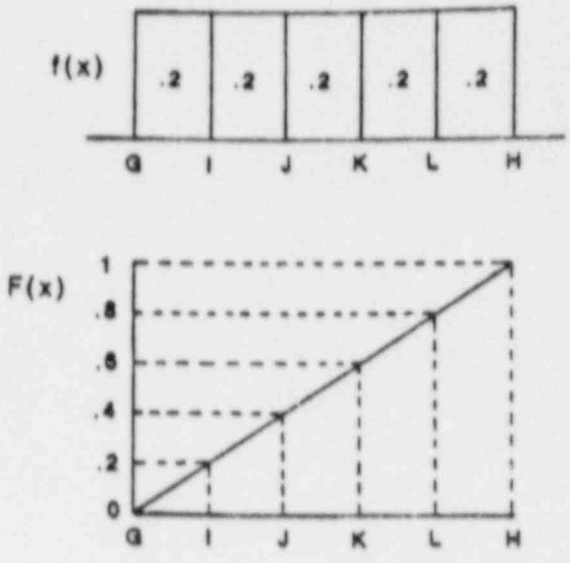

Figure B-2. Intervals Used with a LHS of Size n = 5 in Terms of the Density Function and Cumulative Distribution Function for a Uniform Random Variable

(1979), the pairing was done by associating a random permutation of the first n integers with each input variable. For purposes of illustration, in the present example consider two random permutations of the integers (1, 2, 3, 4, 5) as follows:

Permutation Set No. 1: (3, 1, 5, 2, 4)

Permutation Set No. 2: (2, 4, 1, 3, 5)

By using the respective position within these permutation sets as interval numbers for X_1 (Set 1) and X_2 (Set 2), the following pairing of intervals would be formed.

Computer Run No.	Interval No. Used for X_1	Interval No. Used for X_2
1	3	2
2	1	4
3	5	1
4	2	3
5	4	5

Thus, on computer run number 1, the input vector is formed by selecting the specific value of X_1 from interval number 3 (D to E) and pairing this value with the specific value of X_2 selected from interval number 2 (I to J), etc. Once the specific values of each variable are obtained to form the five input vectors, a two-dimensional representation of the LHS can be made such as given in Figure B-3.

Note in Figure B-3 that all of the intervals for X_1 have been sampled, and the same is true of X_2. In general, a set of n LHS points in k-dimensional Euclidean space contains one point in each of the intervals for each of the k variables.

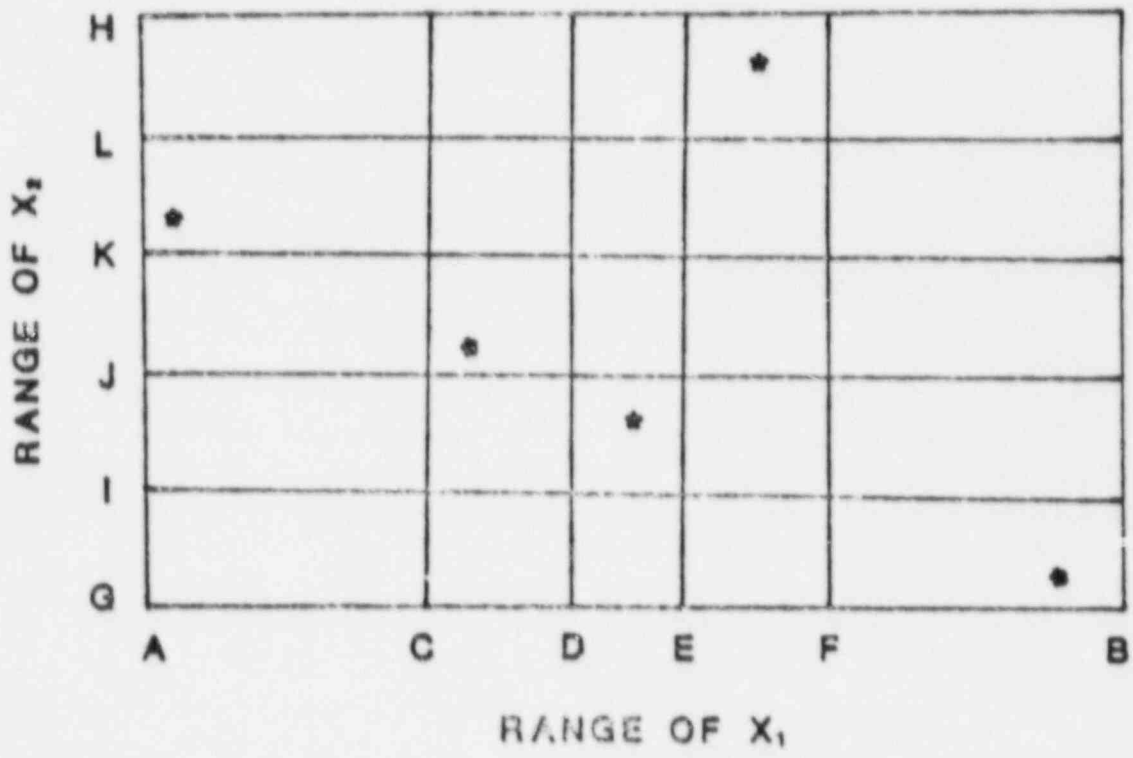

Figure B-3. A Two-Dimensional Representation of One Possible LHS of Size 5 Utilizing X_1 and X_2

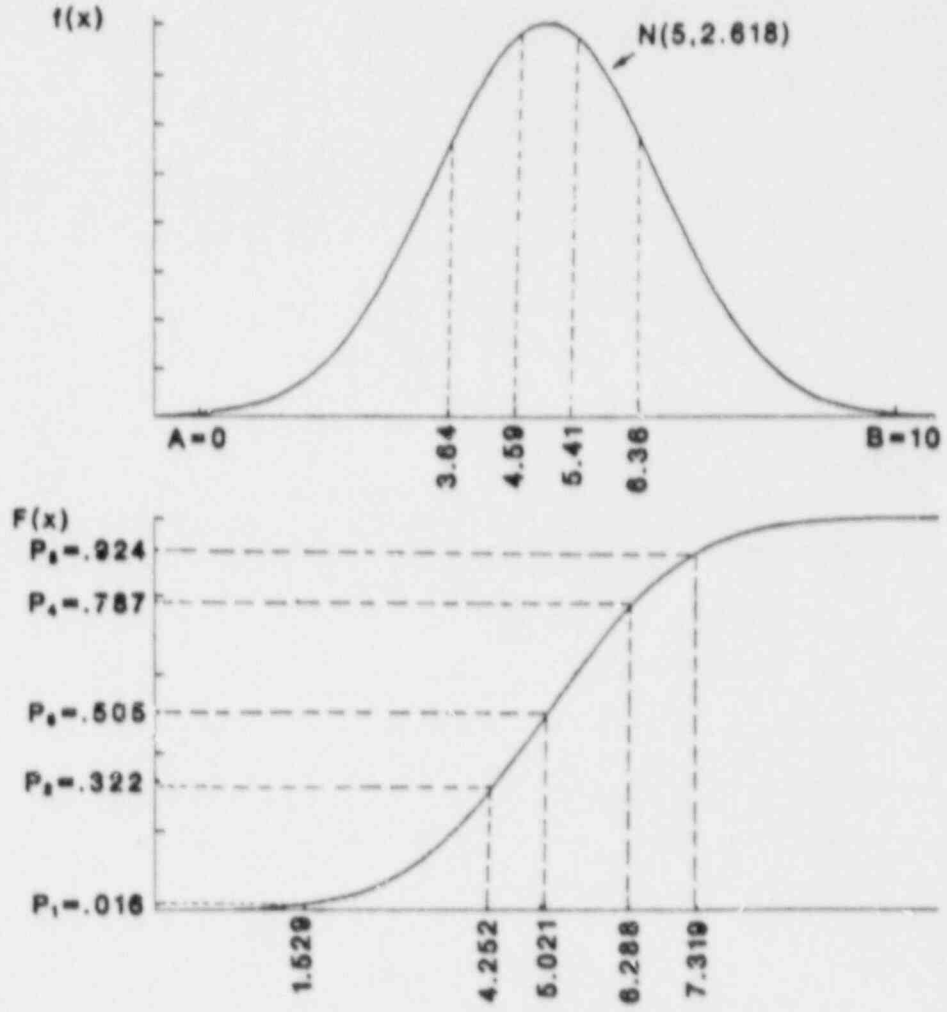

Figure B-4. Interval Endpoints Used with a LHS of Size 5 (top) and Specific Values of X Selected Through the Inverse of the Distribution Function (bottom)

To illustrate how the specific values of a variable are obtained in a LHS, consider the following example. Suppose it is desired to obtain a LHS of size n = 5 from a normal distribution on the range from 0.0 to 10.0. Recall that these two limits are taken to represent the lower and upper .001 quantiles, respectively. Therefore, the random variable has a mean of five and a variance of 2.618 as indicated in Figure 4. These points together with the density characteristics of the normal distribution allow for the definition of the equal probability interval endpoints. These endpoints are shown in Figure 4 in terms of a density function. The next step is to randomly select an observation within each of the intervals. This selection is not done uniformly within the intervals shown in Figure 4, but rather it is done relative to the distribution being sampled (in this case, the normal distribution). This means that the sampling is done uniformly on the vertical axis of the cdf as shown in Figure B-4.

Therefore, to get the specific values, n = 5 randomly selected uniform (0, 1) numbers (U_m, m = 1, 2, 3, 4, 5) are obtained to serve as probability levels. These probabilities are then scaled by

$$P_m = U_m(.2) + (m - 1)(.2) \qquad m = 1, 2, 3, 4, 5$$

This ensures that exactly one probability, P_m, will fall within each of the five intervals (0, .2), (.2, .4), (.4, .6), (.6, .8) and (.8, 1). The values P_m are used with the inverse normal distribution function to produce the specific values to be used in the LHS. Note that exactly one observation is taken from each interval shown in Figure B-4. The entire process is shown in Table B.1. Figure B-4 makes it clear that when obtaining a LHS, it is easier to work with the cdf for each variable. This is the approach used in the computer program, rather than defining the endpoints of the intervals on the x-axis.

The above illustration shows how one input variable having a normal distribution is sampled with LHS. This procedure is repeated for each input variable, each time working with the corresponding cumulative distribution function. If a random sample is desired, then it is not necessary to divide the vertical axis into n intervals of equal width. Rather, n random numbers between 0 and 1 are obtained and each is mapped through the inverse distribution function to obtain the specific values. The final step in the sampling process involves pairing the selected values.

It should be noted that even though two variables are sampled independently and paired randomly, the sample correlation coefficient of the n pairs of variables in either a random sample or a LHS will, in general, not equal zero, just due to sampling fluctuations. In order to obtain a sample in which the sample correlations more nearly match the assumed, or

Table B.1

One Possible Selection of Values for a LHS of Size 5
from a Normal Distribution on the Interval (0,10)

Interval Number m	Uniform (0,1) Random No. U_m	Scaled Probabilities Within the Interval Pm = Um(.2)+ (m-1)(.2)	Corresponding Standard Normal Value (z-score) From the Inverse Distribution	Corresponding N(5,2.618) Observation Within the Intervals
1	.080	.016	-2.144	1.529
2	.610	.322	-0.462	4.252
3	.525	.505	0.013	5.021
4	.935	.787	0.796	6.288
5	.620	.924	1.433	7.319

intended, correlations, Iman and Conover (1982a) proposed a method for restricting the way in which the variables are paired. The effect of this restriction on the statistical properties of the estimated distribution of Y, its mean and percentiles, is not known, but is felt to be small. The pairing of variables in the program can be done either randomly or by the restriction procedure through use of an input parameter, which is explained in the next section.

Additionally, the restricted pairing procedure of Iman and Conover can be used to induce a user-specified correlation among selected input variables through use of another input parameter explained in the next section. However, it should be pointed out that such induced correlations are based on the non-parametric technique known as rank correlation. Such a measure is used since it remains meaningful in the presence of non-normal distributions on the input variables.

As a final note if a correlation structure is not specified by the user, then the program computes a measure for detecting large pairwise correlations. This measure is known as the variance inflation factor (VIF) and is defined as the largest element on the diagonal of the inverse of the correlation matrix. As the VIF gets larger than 1, there may be some undesirably large pairwise correlations present. Marquardt and Snee (1975) deal with some very large VIFs ($> 2 \times 10^6$) and provide a very readable explanation on reasonable sizes of VIFs. Marquardt (1970) indicates that there can be serious collinearity (i.e., large pairwise correlations present) for VIF > 10. Thus, there is certainly no problem as long as the VIF is close to 1. The VIF appears as part of the computer printout when the user requests the correlation matrix to be printed, given that no correlation structure has been specified by the user.

References

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

Iman, R. L. and Conover, W. J. (1982b). Sensitivity Analysis Techniques: Self-Teaching Curriculum, NUREG/CR-2350, SAND81-1978, Sandia National Laboratories, Albuquerque, New Mexico 87185.

Iman, R. L. and Davenport, J. M. (1982). An Interative Algorithm to Produce a Positive Definite Correlation Matrix from an "Approximate Correlation Matrix (With a Program User's Guide), SAND81-1376, Sandia National Laboratories, Albuquerque, New Mexico 87185.

Iman, R. L., Davenport, J. M., and Zeigler, D. K. (1980). Latin Hypercube Sampling (Program User's Guide), SAND79-1473, Sandia National Laboratories, Albuquerque, New Mexico 87185.

Iman, R. L. and Helton, J. C. (1983). A Comparison of Uncertainty and Sensitivity Analysis Techniques for Computer Models, NUREG/CR-3904, SAND84-1401, Sandia National Laboratories, Albuquerque, New Mexico, 87175.

Iman, R. L., Helton, J. C., and Campbell, J. E. (1981a). "An Approach to Sensitivity Analysis of Computer Models, Part 1. Introduction, Input Variable Selection and Preliminary Variable Assessment," Journal of Quality Technology, 13(3), 174-183.

Iman, R. L., Helton, J. C., and Campbell, J. E. (1981b). "An Approach to Sensitivity Analysis of Computer Models, Part 2. Ranking of Input Variables, Response Surface Validation, Distribution Effect and Technique Synopsis," Journal of Quality Technology, 13(4), 232-240.

Marquardt, D. W. and Snee, R. D. (1975). "Ridge Regression in Practice," The American Statistician, 29(1), 3-20.

Marquardt, D. W. (1970). "Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation," Technometrics, 12(3), 591-612.

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

Steck, G. P., Iman, R. L., and Dahlgren, D. A. (1976). Probabilistic Analysis of LOCA, Annual Report for FY1976, SAND76-0535 (NUREG76-6513), Sandia National Laboratories, Albuquerque, New Mexico.

Appendix C

Elicitation of Expert Opinion--Some Guidelines

ELICITATION OF EXPERT OPINION--SOME GUIDELINES

J. M. Booker and M. A. Meyer
Los Alamos National Laboratory

This appendix outlines some guidelines and issues of importance to eliciting expert opinion data. It is intended to provide a brief overview. However, it offers references for those interested in more detailed information.

Some of the concepts discussed below are 1) the nature of expert opinion data; 2) means of selecting and motivating experts; 3) basic methods for eliciting expert opinion and tradeoffs in selecting one; 4) selection of a response mode and means of measuring dispersion; 5) how much data to elicit and record; and 6) reducing or countering the occurrence of bias in elicitation methods, response modes, or requests for problem-solving data; and 7) analyzing the information.

EXPERT OPINION DATA

Expert opinion data consists of expert's subjective judgments. Typically, it has been gathered where experimental data is sparse or unavailable such as in risk analyses or reliability assessments. The use of expert opinion data is wide spread in Probabilistic Risk Assessments (PRA) applications and ranges from those formally elicited to those which are done as part of the PRA process. Most people think of the formal use of expert opinion such as in the requesting of probabilities of rare events. However, expert opinion also includes the less formally elicited judgments that are made by engineers and analysts at various stages of the PRA. Some examples are: which systems will be included; what will be used as a lower bound on screening the cut sets; what features will be included in the accident sequence or fault tree; who will serve on the review board of the PRA; who will be on the PRA team; and what data will be used in the PRA?.

Until recently, many scientific disciplines viewed expert opinion, as being lower in quality than data gathered by observations or measurements. This view had been most common in fields, such as engineering, where even measured data that was probabilistic in nature ("fuzzy" or uncertain) was not readily accepted. In these environments, information that was not deterministic or totally quantitative in nature tended to be regarded as other than "real data" or as unanalyzable. This philosophy has been changing and now even the data purists seem to be more amenable to utilizing all the information available, including that from human beings.

The current trend is one of acceptance of "fuzzy" or probabilistic data, whether from measurements or observations, and of concentration on better, more consistent ways of handling it. Expert opinion fits into this viewpoint because it is observational data, can be both quantitative and qualitative, and has a probabilistic structure. Like any other type of data, care is needed in its elictation, analysis, and interpretation. The "care" that is needed is in the use of consistent and valid procedures for handling this data. This appendix focuses on the use of such procedures in the elicitation process.

Although expert opinion ranges from those formally to informally elicited, none of these judgments need be gathered in an ad-hoc fashion. All

varieties of expert judgment can benefit from careful elicitation. Adherrance to simple quidelines on eliciting and recording is usually sufficient to providing well-documented and consistent procedures for handling expert data in PRAs. Some of these suggestions are given below.

SELECTING AND MOTIVATING EXPERTS

An expert is anyone especially knowledgeable in the field and at the level of knowledge being elicited. For example, one level of knowledge could be core damage frequency for a group of plants ; and a more detailed level could be the risk of direct containment heating at the Surry Plant. Expertise can be of two types: normative and substantive. Substantive expertise is having knowledge in the field in question such as on nuclear reactors or seismology. Normative expertise is based on having knowledge related to the response mode (i.e., the form in which the experts are asked to encode their judgments). Substantive expertise is no guarantee of normative expertise. For example, because the experts are asked frequently about the probabilities of events, an expert could have substantive knowledge of the field at the necessary level of detail required but not normative knowledge of the rules of probabilities.

In selecting experts, normative expertise, substantive expertise, and knowledge at the needed level of detail are always important. In addition, such factors as the reputations and diversity of the experts can become important. Selecting experts who are well known and respected among their peers or the public can lend greater credibility to the study as well as provide motivation for other experts to participate. Diversity among experts is desirable when trying to counter anticipated correlation or dependency among experts (Seaver, 1976).

Motivating experts can be done by communicating the intrinsic benefits of participation or by offering renumeration. The authors recommend motivating through communication because paying the experts is costly, may attract one type of participant (Gordon, 1980:118), or may detrimentally affect their view of the study (Baron and Byrne, 1974:122). Some aspects of the study which could be presented as motivating to the experts are: the chance for recognition; the opportunity to contribute to the development or improvement of a process; and the experiencing of something new or, at least, different from their usual work routine.

In general, how the aspects of the study are communicated to the expert will affect his desire to participate. Usually, more individuals will respond positively to a request delivered in person than by mail. For this reason, it is recommended that the experts be contacted, initially, in person or via the telephone. Guidelines abstracted from communications theory (Stroud 1981; Gorden, 1980) and the authors' interviewing experiences suggest that particular items of information be communicated and in the following order: 1) why the experts are being contacted (the purpose of the study and why they are needed); 2) who is conducting or sponsoring the study 3) how much time and effort their participation will involve; 4) how they were selected; 5) What tasks will be required; 6) how confidentiality will be handled; 7) whether participation is voluntary or required; 7) what the product of the study will be and whether they will have access to the results (Meyer and Booker, 1987b).

ELICITATION METHODS

There are about four methods commonly used for elicitation: the staticized (also called the "nominal") group, the Delphi, the structured interactive group, and the interactive group. These vary in terms of the degree of interaction that they allow between the experts and how the expert's

estimates are combined. For example, the staticized group method allows no interaction either by mail or in person between the experts. In this case, estimates tend to be mathematically aggregated. The Delphi method involves limited interactions between the experts via the mail. The experts' estimates and reasoning are made anonymous and then circulated to the experts to allow them to consider this data in revising their estimates. This process can be continued until consensus, if it is desired, is achieved. The structured interactive method permits the face-to-face meeting of experts but structures their interactions. For example, the experts may be allowed to speak to one another only at specific times during their meetings. Their estimates may be aggregated mathemathically or through achieving consensus during the elicitation sessions. The interactive group method consists of a group of experts who are interacting in a spontaneous fashion in a face-to-face situation. This method employs few controls and, thus, resembles the meetings that commonly take place in work settings. Estimates obtained from this method can be combined by either method, although mathematical aggregation is recommended. In particular, simple aggregation schemes, such as medians, are suggested (Martz et al., 1985). (For a more complete description of these elicitation methods, see Meyer et al., 1982).

At the basis of these four methods lie two views of interaction between experts. One view holds that interactions are to be avoided or limited because of their affect on the judgment process. For example, the staticized group method avoids any inter-expert interaction, while the Delphi method permits only limited and controlled interaction. The Delphi method, developed by Rand Corporation, was designed to prevent bias arising from group dynamics in a face-to-face setting. One common problem arising from group dynamics is the group think phenomenon, also called the follow-the-leader effect (Janis, 1972). To avoid having experts consciously or unconsciously adjust their judgments to agree with those of the leader(s), the Delphi keeps experts physically separated. Their judgments are sent to a central location where any distinguishing features are removed. The newly anonymous judgments are then recirculated to all the experts, who are allowed to change their previous estimates.

At the other end of the spectrum, are approaches like the interactive group methods that encourage group interaction in the belief that the best quality answers come from the interaction of knowledgeable persons in a face-to-face setting (Seaver, 1976). Some interactive methods, like the structured interactive group, employ structure to minimize the effects of group dynamics. A few examples of this structuring would be: having everyone state their view, one at a time, while no one else is allowed to comment; having the natural group leaders give their thoughts last; and having the individuals record their judgments as a means of encouraging them to anchor to their own judgments. The interactive group method can be structured to any degree as time permits.

Elicitation methods are usually selected on the basis of the analyst's belief as to whether interaction between the experts is desirable or not. The analyst frequently makes this decision because the literature has not been conclusive as to which method produces the best, the least biased, results. The authors favor a structured interactive group method because of the synergistic effect of discussion on the judgment process. Techniques, such as mentioned in Reducing or Countering Bias, can be used to mitigate the biases that occur in an interactive group setting.

Any of the basic elicitation methods can be structured to different degrees. For example, the interviewing of the experts in the staticized group can be done informally or in a structured manner. The structured manner could

involve presenting each expert with the assumptions or definitions that they are to use in solving the problem and then monitoring their use of these. As a general rule, the more structure imposed on the method, the more time required.

Generally, it is best to gather expert opinion in a structured and controlled manner for several reasons: 1) to know what considerations have entered into the experts' estimates, 2) to defend the gathered data, and 3) to update or modify this data (Meyer, 1984).

Some considerations that enter into the experts' estimates are their definition of the problem; their interpretation of directions; their assumptions or boundary conditions; and the calculations or heuristics used to solve the problem.

Expert data may require defense against critics of the elicitation and/or the conclusions. The method of elictation needs to be considered valid by current theory in communications, cognition, and psychology. Often, if the elicitation is properly conducted, the experts will feel positively enough about the study to defend it.

Frequently, there is a need to update or modify the expert data (e.g., if circumstances have changed). If the elictation method has been structured properly, there is a greater chance of knowing which factors relating to the estimate have been affected by the change and therefore, how to modify the data acordingly. It is recognized that most expert opinion is biased, one way or another. The challenge is to know which way so that a counter-bias can be employed (Payne, 1951; Meyer, 1984).

RESPONSE MODE AND DISPERSION MEASURES

Response mode: The response mode is the form in which the expert is requested to give his judgment. Probabilities, odds, logs, intervals, paired comparisons, and ratings (Seaver and Stillwell 1983; Comer et al., 1984) are a few of the commonly used response modes. Selection of a response mode involves two issues: obtaining the experts' judgments in a form that can be analyzed or used in a model, and avoiding bias. For example, the analyst may prefer probabilities and request the expert's judgments in this form. If the data is to serve as input to a model, the experts' judgments will need to be given in a compatible form or to be converted later. Associated with these considerations is the question of whether the expert can encode his thoughts into the requested response mode.

Typically, the focus in expert opinion has been on the practical issue of obtaining the experts' judgments in the needed form rather than on the experts' ability to encode their thoughts into that form. The later involves the problem of bias. Bias can be defined as occurring when: 1) the expert's underlying judgments or reporting of these is altered by the elicitation process; and 2) when the expert's estimates do not follow normative statistical or logical rules (Meyer, 1987). The response mode can lead to either type of bias. For example, the expert may not be able to express his exact thinking in the requested response mode -- In the translation from his judgment to an acceptable response, his judgment may be misrepresented. Then too, experts can give estimates that do not follow the mathematical or logical rules governing their use. For example, probabilities for mutually exlusive, exhaustive events should sum to 1.0 and the expert's probabilities may not.

Some of the bias may stem from the expert's unfamiliarity with the response mode. Generally, substantive experts (those with expertise in their field) are not accustomed to couching their judgments in the form of logarithms, intervals, percentiles, or even probabilities. Their substantive expertise does not guarantee normative expertise (expertise in the use of the

response mode). Another contributor to bias may be the way in which the human mind works.

Hogarth (1980) has attributed many biases to our "limited, sequential, information processing." The human mind is not a miniature calculator. Because of its limited memory space for information processing, the mind resorts to short cuts, heuristics, to solve problems. These heuristics can result in a skewing, biasing, of the answer from what it should be mathematically. For example, in solving a complex problem an individual may form an initial impression of what the answer should be and then consider components of the problem one at a time. In solving each part of the problem, the individual may mentally adjust up or down from his last impression. This heuristic, anchoring and adjustment, tends to produce a final estimate which is biased towards the individual's initial impression. (See Hogarth 1980 for other biases stemming from the way in which we think.)

It is advisable to avoid the use of response modes with which individuals are reputed to have difficulty, such as percentiles and probabilities, in favor of simple linear scales. Difficult response modes are defined here to mean those that lead frequently to estimates that do not follow normative statistical or logical rules. A continuous linear scale is a number line with defined endpoints. These endpoints should represent extreme values and be labeled with text or numbers. The scale also includes tick marks or intermediate values. The expert is instructed to mark his answer at or between any of the delineations on the scale.

The experts find these scales convenient to use and the results of their use are easily interpreted for analysis. Some advantages of this scale are: 1) it requires very little definition from the interviewer; 2) it does not restrict the range of response possibilities like a multiple choice response mode; 3) it provides the opportunity for responses in the extremes values; 4) it relies on linear thinking, a mode that most humans use naturally (Kahneman and Tversky, 1982); and 5) it is easily converted to numerical, continuous variables for analysis.

Dispersion Measures: In addition to the single estimate (a central probability measure), experts are often asked to provide a dispersion measure such as percentiles, variances, ranges, or error bars. These dispersion measures are prone to bias (Kahneman and Tversky, 1982) because the experts are not able to provide the measures that match the prescribed defintions. For example, the expert may be asked to provide the 5th and 95th percentiles of his distribution for the probability of an event. He will provide two estimates that he thinks correspond to the 5th and 95th percentiles; however, those two estimates will more likely correspond to approximately the 33rd and 67th percentiles. Humans, in general, tend to underestimate uncertainties. Even in studies where experts were trained or were knowledgeable of this bias, the results were still only marginally improved and uncertainty remained underestimated (Lichtenstein et al., 1982). In addition, humans also tend to be poor at estimating variances or standard deviations (Kahneman and Tversky, 1982). Humans also have difficulties in estimating extreme values such as the absolute maximum or minimum values (Martz et al., 1985).

Therefore, asking experts for specifically defined measures of dispersion will yield estimates that cannot be accurately interpreted. The only interpretation that can be made is that they will underestimate the true uncertainty. This bias also occurs if the experts are asked to estimate the dispersion of a distribution of multiple experts' estimates. This occurs because each expert is likely to perceive others as having the same values as he does.

One solution is to elicit some specified measure (e.g., the 5th and 95th percentiles), and then to double or triple the range of values knowing that the provided ranges will underestimate the true uncertainty. Then the question becomes how much adjustment is needed. There is no clear answer to this question because there is no way of knowing what is actually represented by the ranges given.

A second solution is to elicit only the central tendancies and to use multiple experts to form a single distribution. The problems with this are twofold: 1) a large number of expert estimates is required; 2) the resulting distribution is usually multimodal which leads to the question of how to interpret the mixture of distributions.

A third solution is to elicit the central tendancies and to allow the expert to volunteer an undefined range of values around it. This undefined range could also be requested; however, in such a case, the expert would need a definition of the range being requested and the analyst is faced again with the problems in solution one. The volunteered dispersion also lacks a specific interpretation; however, it can be interpreted as defining a set of values that the expert felt like specifying. This set of values could then be used as multiple estimates from the same expert allowing the expansion of a usually sparse data set, or it could be interpreted as some middle quantiles (e.g., 40th and 60th percentiles) for formulating expert prior or likelihood distributions.

The specified and volunteered dispersion measures both have the same interpretation problem when using them to formulate expert distributions. The one advantage of the volunteered measure over the specified one is that the volunteered measure is free from any biases imposed by the interviewer in requesting and defining it. The volunteered estimates fit into the elictation philosophy of attempting to minimize bias.

HOW MUCH DATA TO ELICIT AND RECORD

A variety of data can be elicited and recorded: the estimate with a dispersion measure, if elicited; data on how the expert arrived at his answer (e.g., information considered and heuristics used in solving the problem); and some of the characteristics of the expert's professional background (e.g., education, work experiences, and experience with similar types of problems). The question of how much data to gather arises when the researcher weighs the possible benefits of gathering more data than the estimate against the additional time that this would take. Many researchers do not know how an expert's reasoning could be used or how this qualitative data could be analyzed by itself or in combination with the expert's estimates. For these reasons. data on how the experts arrived at an estimate has generally not been gathered in the field of expert opinion. The few exceptions to this rule can be found in studies where the expert has asked that one of his considerations, such as a caveat, be recorded to clarify his answer (Benjamin et al., 1986). In addition, data on the experts' background is either not elicited or not printed to protect the experts' identities. For example, in an earlier seismic study, not only were the expert's names not linked to their judgments but their names were not identified (Baecher, 1979).

The authors suggest that data on how the experts solve the problem be elicited and recorded. Each expert should be observed and asked to name any information that they are using in solving the problem. For example, the expert may say that he is considering the conditions of state or the definitions given in the background to the problem. In addition, the expert should be asked to explain how he is solving the problem. His description

might include some equations, calculations, or more general heuristics, such as mentioned in the Response Mode section.

Recording the problem-solving data offers such benefits as providing guidance and justification for aggregation and for the elimination of outliers. Aggregations are frequently performed to combine multiple, and often differing estimates, into one estimate. A bimodal distribution of estimates is common but if the analyst does not know why the experts are split into two camps, he is handicapped in his decision on how to aggregate (Booker and Meyer, 1985). An aggregation scheme can be selected on the basis of the problem-solving information rather than guesswork, if problem-solving data is elicited. In a Bayesian approach to aggregation, the decision maker has to decide how to combine the expert's estimates with each other and with his judgments. Knowledge of the expert's problem-solving methods will allow the decison maker to judge which estimates he wishes to weight most highly (Meyer and Booker, 1987a).

In the treatment of outliers, the "aberrant observations should be recorded and the reason for deleting them explained" (Tietjen, 1986). The problem-solving data can show whether the experts who supplied the suspected outliers solved the problem in a different manner and whether their estimates can be eliminated. If the suspected outlier appears to be legitimate data (e.g., the expert did not appear to use unwarranted assumptions or unacceptable definitions), then this evidence can be stated. In this case, the outliers could be accomodated by means of a robust estimator.

In addition, it may be a good idea to record some background information such as, the experts' years of experience, formal education, previous work experiences related to the problem area, years since the expert worked on a similar problem, and colleagues with whom the expert has been associated.

The recording of this magnitude of data can be made easier and more thorough by the use of a note-taking observer or a tape recorder. The means of recording can be chosen by asking which will be the most convenient, unobtrusive, and reliable in the particular setting. Whatever means is used should have as its goal the recording of data in as exact manner as possible. Filtering, reducing and modeling of the information gathered can be done following the interview; however, the results from such screening should always be traceable back to the original data.

REDUCING OR COUNTERING BIAS IN ELICITATION

As a general rule, bias in expert opinion can be handled by anticipating where it is likely to occur, designing the means to counteract it, and then monitoring the elicitation process. This bias countering plan can be applied to the different parts of elicitation, such as the elicitation method, the response mode/dispersion measure, and the problem-solving interview. As mentioned earlier, bias can occur when: 1) the subject's underlying thoughts or reporting of these has been altered by some aspect of the elicitation; or 2) when the subject's thinking does not follow normative statistical or logical rules.

For example, in selecting an elicitation method, one could anticipate the occurrence of both types of bias. If, for instance, an interactive group method was being considered for use, one could expect the possibility of group think. With the group think phenomena, the expert's thinking is altered as he consciously or unconsciously acquieses to what he believes to be the group opinion. Then too, in an interactive group method, the experts' failure to follow logical or statistical rules could surface. For example, if the experts were meeting for several sessions, they would be prone to inconsistency in their thinking. In particular, the definitions or conditions

that the group has agreed to use are likely to under go subtle changes and reinterpretation in the minds of each expert. Because the expert's judgments are based on such information, these changes can affect their answers.

The elicitation method can be designed to counter the anticipated biases. For example, in the case of the interactive group method, group think could be countered in several ways. One way would be to have the experts record their thoughts and then, one at a time, present these to the group. (Other ways are given in the section, Elicitation Methods, or in Meyer, 1984.) To counter inconsistency, in part, one could conduct a review of the definitions or conditions at the beginning of every session.

In attempting to counter bias in the use of response modes and dispersion measures, selection is the key. As mentioned in the Response Mode section, people do not use the modes and measures, statistically or logically, as they should. Humans not natural Bayesians nor are they good assessors of uncertainty. We recommend the use of simple response modes, such as linear scales, which are not as prone to bias.

If the expert is to be interviewed on how he solved the problem, the phrasing and timing of these questions is critical in avoiding bias. The major bias to occur in this form of elicitation is the influencing of the expert's thoughts or his descriptions or these. To avoid introducing this source of bias, the authors recommend several techniques: verbal protocal from educational psychology (Belkin et al., 1986), the ethnographic method from cultural anthropology (Spradley, 1979), and a type of verbal probe from psychology (Ericsson and Simon, 1980). These techniques work best in individual interviews but they can be adapted to interactive group situations.

With the verbal protocal method, the expert is instructed to think aloud as he progresses through the problem. The verbal protocal is believed to deliver the data that the subject currently possesses in his short-term memory (Ericsson and Simon, 1980). Because the interviewer does not ask questions but simply records the subject's verbalizations, there is little chance of his affecting the subject's thoughts. The method of verbal protocal need not be used, if sufficient information can be obtained using the ethnographic method and verbal probe, described below. A disadvantage of the verbal protocal method is that it can slow the expert's progress on solving the problem. Then too, it does not work on very complex problems because the expert stops thinking aloud to concentrate.

The ethnographic method involves restating the expert's words into questions (Meyer and Booker, 1987a). Thus, it does not insert the interviewer's thoughts and bias the expert's account. The ethnographic method can used to pursue, in greater depth, elements that the expert mentioned in his thinking aloud.

The verbal probe is used immediately after the expert has given his solution to gather additional data on how the expert solved the problem. This probe is asked immediately because the needed information is only available in an individual's short-term memory for a brief period of time. Thereafter, the individual is likely to be able to retrieve only portions of what he had in his short-term memory and to have to fill in the gaps by guessing (Ericsson and Simon, 1980). A simple phrasing of the probe is used, such as "Why did you give this answer," as opposed to a check-list phrasing. An example of a check-list phrasing would be, "Did you consider factors X, Y, and Z?" The simple phrasing is preferrable because it is less prone to bias. With the check-list phrasing, the subject is more likely to say that he considered these factors when he did not and to believe that he should consider these in his future problems. The authors recommend that the verbal probe be used to obtain data on the expert's problem solving, even if no other method is used.

It takes very little time and can be used easily in group settings. Asking the expert to explain his answer immediately after he has given it will provide enough data for the receiving the benefits mentioned in Elication Methods. If the explanation is unclear or too general, the ethnographic method can be used to gather additional information.

In addition to designing or structuring elicitation methods to counter bias, it is advisable to monitor the sessions. For example, during an interactive group session, one could watch for any indication of a group think phenomena. The lack of any dissenting views could be an indication of a group think situation. By monitoring the elicitation sessions, one can gauge the presence of bias and employ further countermeasures, if necessary.

ANALYZING THE DATA

There are several analytical problem areas worthy of mention. These are discussed in Meyer and Booker (1987c) and include recommendations on: 1) how to handle the large amounts of data gathered; 2) how to handle the qualitative/quantitative data mixture; 3) how to formulate models at the desired level of generality; 4) how to deal with the correlation among experts, if any exists; and 5) how to aggregate estimates or distributions from multiple experts. Whether the approach taken is a Bayesian or classical one, these problems are likely to be present. Each is discussed briefly below.

Many multivariate statistical techniques exist for handling the large amounts of data. Cluster and correlation analyses can provide means for identifying relationships between the variables and the response data.

Non-parametric and rank procedures are suggested for handling the qualitative/quantitative data mixture to avoid the making of assumptions about data distributions (which are not usually normal or multivariate normal).

The level of generality is an important assumption made in these analyses. All analysts deal with this problem in designing and analyzing experiments. Usually, the experiments are designed to gather the specific information required to answer the research questions. In expert opinion, the information gathered cannot always be done in such a controlled environment. This is why the authors recommend that the analyst gather all possible information first and screen, filter, and model what is needed later. It is at this point that the level of generality must be decided. How finely will the information be screened? What granularity will be assumed? It should be noted that the conclusions drawn only apply to the level of generality used in the modeling process. Conclusions can be different for different levels (Meyer and Booker, 1987c).

An example of the importance of the level of generality can be found in the issue of correlation among experts. At a general level of problem solving, correlation among experts was found (Meyer and Booker, 1987a). However, at a more detailed level of analysis, this source of correlation was not evident (Meyer and Booker, 1987c). Inter-expert correlation seems to exist only at certain levels of modeling the problem-solving processes. Thus, the correlation issue may not be as prevasive or troublesome as previously thought.

If correlation among experts is not a problem, then combining expert estimates or distributions together becomes an easier process. Weighting factors or distributions can be assumed and used in many aggregation schemes (see Martz et al., 1985). The aggregation can be done in a Bayesian framework using prior and likelihood distributions for data and for experts, or by the formulation of a single distribution from the multiple experts. Decision

analysis techniques can be used in either framework to incorporate the effects of the decision maker (Booker and Bryson, 1982).

SUMMARY

The benefits of using the suggested techniques for elictation outweigh the efforts required to implement them. 1) A consistent and structured methodology for elicitation can be established and used. 2) As situations change, old information can be updated if it has been properly documented. 3) Information can be compared from different studies. 4) Defense against critics can be given with confidence. 5) Futher avenues of research and development can be discovered and implemented thus, adding to the field of knowledge and improving the methodology.

REFERENCES

Most of these references are available, upon request, from the authors: J. M. Booker or M. A. Meyer at Los Alamos National Laboratory, P.O. Box 1663, MS F600, Los Alamos. NM, 87545, USA.

Baecher, G.B. (1979), "Correlations Among Experts' Opinions," Unpublished manuscript, Massachusettes Institute of Technology, Boston, MA.

Baron, R.A. and Bryne, D. (1974), Social Psychology: Understanding Human Interaction, Allyn and Bacon Inc., Boston, MA.

Belkin, N.J., Brooks, H.M., and Daniels, P.J. (1986), "Knowledge Elicitation Using Discourse Analysis," Proceedings from the Knowledge Acquisition for Knowledge-Based Systems Workshop, Banff, Canada, pp. 3,0-3,16.

Benjamin, A.S., Kunsman, D.M., Williams, D.C., Boyd, G.J., and Murfin, W.B. (1986), "Evaluation of Severe Accident Risks and the Potential for Risk Reduction: Surry Power Station, Unit 1," SAND86-1309, Sandia National Laboratory, Albuquerque, NM.

Booker, J.M. and Bryson, M.C. (1985), "Decision Analysis with Applications to Project Management: State-of-the-Art Survey and Bibliography," American Journal of Mathematical and Management Sciences, 5, pp. 1-62.

Booker, J.M. and Meyer, M.A. (1985), "Sources and Effects of Correlation of Expert Opinion," LA-UR-85-1879, Los Alamos National Laboratory, Los Alamos, NM.

Cleaves, D.A. (1986), "Cognitive Biases and Corrective Techniques: Proposals for Improving Elicitation Procedures for Knowledge-Based Systems," Proceedings from Knowledge Acquisition for Knowledge-Based Systems Workshop, Banff, Canada, pp. 9,0-9,11.

Comer, M.K, Seaver, D.A., Stillwell, W.G. and Gaddy, C.D. (1984), "Generating Human Reliability Estimates Using Expert Judgment," NUREG/CR-3688, SAND84-7115, Sandia National Laboratories, Albuquerque, NM.

Ericsson, K.A. and Simon, H.A. (1980), "Verbal Reports as Data," Psychological Review, 87, 3, pp. 215-250.

Gorden, R.L. (1980), Interviewing: Strategy, Techniques, and Tactics, Dorsey Press, Homewood, ILL.

Hogarth, R. (1980), Judgment and Choice: The Psychology of Decisions, Wiley-Interscience, Chicago, IL.

Janis, I.L. (1972), Victims of Group Think: A Psychological Study of Foreign Policy Decisions and Fiascos, Houghton Mifflin, Boston, MA.

Kahneman, D. and Tversky, A. (1982), "Subjective Probability: A Judgment of Representativeness," D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgment Under Uncertainty: Heuristics and Biases, Cambridge University Press, NY., pp. 32-47.

Lichtenstein, S., Fischhoff, B. and Phillips, L.D. (1982), "Calibration of Probabilities: The State of the Art to 1980", D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgment Under Uncertainty: Heuristics and Biases, Cambridge University Press, NY., pp. 306-334.

Martz, H.F., Bryson, M.C. and Waller, R.A. (1985), "Eliciting and Aggregating Subjective Judgments--Some Experimental Results," Proceedings for the Tenth Annual SAS Users Group International Conference, SAS Institute Inc., Cary, NC.

Meyer, M.A. (1984), "Human Factors Affecting Subjective Judgments," LA-UR-84-3176, Los Alamos National Laboratory, Los Alamos, NM.

Meyer, M.A. and Booker, J.M. (1987a), "Sources of Correlation Between Experts: Empirical Results from Two Extremes," LA-10918-MS, NUREG/CR-4814, Los Alamos National Laboratory, Los Alamos, NM.

Meyer, M.A. and Booker, J.M. (1987b), "Selecting and Motivating Experts," A Guide for the Elicitation and Analysis of Expert Opinion, draft for book, Los Alamos National Laboratory, Los Alamos, NM.

Meyer, M.A. and Booker, J.M. (1987c), "Problems with Expert Data in Studies of Inter-Expert Correlation," Invited paper to be presented at the Joint Statistical Meetings, Aug 17-20, 1987, San Francisco, CA.

Meyer, M.A. (1987) "Handling Human Bias in Acquiring Knowledge for Knowledge-Based Systems," LA-UR-87-1221, Los Alamos National Laboratory, Los Alamos, NM.

Meyer, M.A., Peaslee, A.T. and Booker, J.M. (1982), "Group Consensus Method and Results," LA-9584-MS, Los Alamos National Laboratory, Los Alamos, NM.

Payne, S. (1951), The Art of Asking Questions, Princeton University Press, Princeton, NJ.

Seaver, D.A. (1976), "Assessments of Group Preferences and Group Uncertainty for Decision Making," Social Science Research Institute, University of Southern California, Los Angeles, CA.

Seaver, D.A. and Stillwell W.G. (1983), "Procedures for Using Expert Judgment to Estimate Human Error Probabilities in Nuclear Power Plant Operations," NUREG/CR-2743, SAND82-7052, Sandia National Laboratories, Albuquerque, NM.

Spradley, J.P., (1979), The Ethnographic Interview, Holt, Rinehart and Winston, New York, NY.

Stroud, M. (1981), "Communication Strategies: Strategies to Improve the Transfer of Knowledge," Communication Strategies for Advanced Technologies, Albuquerque, NM.

Tietjen, G.L. (1986), A Topical Dictionary of Statistics, Chapman and Hall, New York, NY.

Distribution

US Government Printing Office
Receiving Branch (Attn: NRC Stock)
8610 Cherry Lane
Laurel, MD 20707
320 copies for RX

James Abel [10]
Commonwealth Edison Co.
35 1st National West
Chicago, Illinois 60690

B. B. Agrawal
US Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
5650 Nicholson Lane
Rockville, MD 20852

H. Banon
EQE Inc.
3300 Irvine Ave., Suite 345
Newport Beach, CA 92660

Patrick Baranowsky
US Nuclear Regulatory Commission
Office of Nuclear Reactor Regulation
Washington, DC 20555

Robert A. Bari
Brookhaven National Laboratories
Building 130
Upton, NY 11973

Richard J. Barrett
US Nuclear Regulatory Commission
Office of Nuclear Reactor Regulation
Division of Radiation Protection
and Emergency Preparedness
Washington, DC 20555

Doyle Batt
NRC Risk Analysis Unit
EG&G Idaho, Inc.
P. O. Box 1625
Idaho Falls, ID 83415

Barbara Bell
Battelle Columbus Laboratories
505 King Avenue
Columbus, OH 43201

R. C. Bertucio
Energy Incorporated
1851 So. Central Place, Suite 201
Kent, WA 98031

Dennis Bley
Pickard, Lowe & Garrick
2260 University Drive
Newport Beach, CA 92660

Gary Boyd
Safety & Reliability Optimization
Services
624 Glen Willow Dr.
Knoxville, TN 37922

Robert J. Budnitz
Future Resources Associates
734 Alameda
Berkeley, CA 94707

Gary Burdick
Division of Reactor and Plant Systems
US Nuclear Regulatory Commission
5650 Nicholson Lane
Rockville, MD 20852

N. G. Cathey
Idaho National Engineering Lab.
P. O. Box 1625
Idaho Falls, ID 83415

Frank Coffman
US Nuclear Regulatory Commission
Division of Reactor and Plant Systems
Office of Nuclear Regulatory Research
Washington, DC 20555

George Crane
1570 E. Hobble Creek Dr.
Springville, Utah 84663

Garth Cummings
Lawrence Livermore Laboratory
L-91, Box 808
Livermore, CA 94526

Mark A. Cunningham
US Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555

Peter Cybulskis
Battelle Columbus Division
505 King Avenue
Columbus, OH 43201

Richard S. Denning
Battelle Columbus Division
505 King Avenue
Columbus, OH 43201

Ed Dougherty
Science Applications International Corp.
P. O. Box 2501
Oak Ridge, TN 37831

Adel El-Bassioni
US Nuclear Regulatory Commission
Office of Nuclear Reactor Regulation
Division of Radiation Protection and
Emergency Preparedness
Washington, DC 20555

Karl Fleming
Pickard, Lowe & Garrick
2260 University Drive
Newport Beach, CA 92660

J. B. Fussell
JBF Associates, Inc.
1630 Downtown West Boulevard
Knoxville, TN 37919

Norman L. Graves
2114 Sparrow Ct.
West Richland, WA 99352

Lon N. Haney
Human Factors Organization
EG&G Idaho, Inc.
P. O. Box 1625
Idaho Falls, ID 83415

G. William Hannaman
NUS Corporation
16835 West Bernardo Drive
Suite 202
San Diego, CA 92127

Steven Hodge
Oak Ridge National Laboratories
P. O. Box Y
Oak Ridge, TN 37831

Peter Humphreys
US Atomic Energy Authority
Wigshaw Lane, Culcheth
Warrington, Cheshire
United Kingdom, WA3 4NE

Brian Ives
UNC Nuclear Industries
P. O. Box 490
Richland, WA 99352

Alan Kolaczkowski
Science Applications Int. Corp.
2109 Air Park Rd. SE
Albuquerque, NM 87106

Jim Kolanowski
Commonwealth Edison Co.
35 1st National West
Chicago, Illinois 60690

George Klopp
Commonwealth Edison Co.
P. O. Box 767, Room 35W
Chicago, Illinois 60690

Dave Lappa
Lawrence Livermore National Labs
7000 E Avenue
Livermore, CA 94550

Josette Larchier-Boulanger
Electricte de France
Direction des Etudes Et Recherches
30, Rue de Conde
65006 Paris, France

Peter Lohnberg
Expresswork International, Inc.
1740 Technology Drive
San Jose, CA 95110

Herbert Massin
Commonwealth Edison Co.
35 1st National West
Chicago, Illinois 60690

Andrew S. McClymont
IT-Delian Corporation
1340 Saratoga-Sunnyvale Rd.
Suite 206
San Jose, CA 95129

Joseph A. Murphy
Division of Reactor Accident Analysis
US Nuclear Regulatory Commission
5650 Nicholson Lane
Rockville, MD 20852

Gareth Parry
NUS Corporation
910 Clopper Rd.
Gaithersburg, MD 20878

Blake Putney
Science Applications Int. Corp.
5150 El Camino Real, Suite C31
Los Altos, CA 94022

David Pyatt
US Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Washington, DC 20555

Dale Rasmuson [10]
Division of Reactor and Plant Systems
US Nuclear Regulatory Commission
5650 Nicholson Lane, NL-005
Rockville, MD 20852

M. K. Ravindra
EQE Inc.
3300 Irvine Ave., Suite 345
Newport Beach, CA 92660

Tom Ryan
Division of Reactor and Plant Systems
US Nuclear Regulatory Commission
5650 Nicholson Lane
Rockville, MD 20852

Paul Shemanski
US Nuclear Regulatory Commission
Office of Nuclear Reactor Regulation
Washington, DC 20555

Stephen Sholly
MHB Technical Associates
1723 Hamilton Ave.
Suite K
San Jose, CA 95125

Desmond Stack
Los Alamos National Laboratory
Group Q-6, Mail Stop K556
Los Alamos, NM 87545

Alan D. Swain
712 Sundown Pl. SE
Albuquerque, NM 87108

Steve Topp
EI DuPont DeNemours & Co. Inc.
Savannah River Laboratory
Aiken, SC 29808-0001

Stephen D. Unwin
Department of Nuclear Energy
Division of Safety and Risk Evaluation
Brookhaven National Laboratory
Building 130, Brookhaven Lane
Upton, NY 11973

William E. Vesely
Science Applications Int. Corp.
2929 Kenny Rd., Suite 245
Columbus, Ohio 43221

Jim E. Wells
Lawrence Livermore National Labs
7000 E Avenue
Livermore, CA 94550

3141 S. A. Landenberger (5)
3151 W. L. Garner
6400 D. J. McCloskey
6410 N. R. Ortiz
6412 A. L. Camp
6412 W. R. Cramond
6412 S. L. Daniel
6412 D. M. Kunsman
6412 K. J. Maloney
6412 L. A. Miller
6412 A. C. Payne, Jr.
6412 T. T. Sype
6412 T. A. Wheeler
6412 D. W. Whitehead
6413 E. Gorham-Bergeron
6413 F. T. Harper
6415 J. C. Helton
6415 R. L. Iman
6440 D. D. Dahlgren
6447 M. P. Bohn (25)
6513 D. D. Carlson
8024 P. W. Dean

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

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

SEE INSTRUCTIONS ON THE REVERSE

1 REPORT NUMBER (Assigned by TIDC, add Vol. No., if any)

NUREG/CR-4836

SAND87-0871

2 TITLE AND SUBTITLE

Approaches to Uncertainty Analysis in Probabilistic Risk Assessment

3 LEAVE BLANK

4 DATE REPORT COMPLETED

MONTH: December YEAR: 1987

5 AUTHOR(S)

Michael P. Bohn, Timothy A. Wheeler, Gareth W. Parry

6 DATE REPORT ISSUED

MONTH: December YEAR: 1987

7 PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

Sandia National Laboratories
Albuquerque, New Mexico 87185

8 PROJECT/TASK/WORK UNIT NUMBER

9 FIN OR GRANT NUMBER

FIN A1384

10 SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

Division of Reactor and Plant Systems
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555

11a TYPE OF REPORT

Final

b PERIOD COVERED (Inclusive dates)

12 SUPPLEMENTARY NOTES

13 ABSTRACT (200 words or less)

An integral part of any probabilistic risk assessment (PRA) is the performance of an uncertainty analysis to quantify the uncertainty in the point estimates of the risk measures considered. While a variety of classical methods of uncertainty analysis exist, application of these methods and developing new techniques consistent with existing PRA data bases and the need for expert (subjective) input has been an area of considerable interest since the pioneering Reactor Safety Study (WASH-1400) in 1974. This report presents the results of a critical review of existing methods for performing uncertainty analyses for PRAs, with special emphasis on identifying data base limitations on the various methods. Both classical and Bayesian approaches have been examined. This work was funded by the U.S. Nuclear Regulatory Commission in support of its ongoing full-scope PRA of the LaSalle nuclear power station. Thus, in addition to the review, this report contains recommendations for a suitable uncertainty analysis methodology for the LaSalle PRA.

14 DOCUMENT ANALYSIS - a KEYWORDS/DESCRIPTORS

b IDENTIFIERS/OPEN ENDED TERMS

15 AVAILABILITY STATEMENT

Unlimited

16 SECURITY CLASSIFICATION

(This page) U

(This report) U

17 NUMBER OF PAGES

18 PRICE

☆ U.S. GOVERNMENT PRINTING OFFICE: 1985—573-049/61038