1986

Sensitivity analysis approach for robust probabilistic risk assessment

Shahid Ahmed
Iowa State University

8615022

Ahmed, Shahid

SENSITIVITY ANALYSIS APPROACH FOR ROBUST PROBABILISTIC RISK ASSESSMENT

Iowa State University PH.D. 1986

Sensitivity analysis approach for robust probabilistic risk assessment

by

Shahid Ahmed

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Nuclear Engineering

Approved:

Members of the Committee:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1986

TABLE OF CONTENTS

Page

GENERAL INTRODUCTION

In recent years, Probabilistic Risk Assessment (PRA) has become an accepted tool to assess risk from different sources such as nuclear power plants (NPPs), dam failures, floods, earthquakes, civil aviation, regulation of pesticides, coal fuel cycles, liquified natural gas, and the chemical industry. The PRA techniques are being used to compare and reduce risk from various societal activities, and to manage risk in a cost-effective manner. Since no activity is risk free, PRA is helping to make a decision on acceptable risk level for the health and safety of the public from various man-made hazards. However, PRA is relatively a new field, and much development is needed in methods before realistic safety predictions can be made. This dissertation suggests new approaches in PRA in the areas of uncertainty propagation of component failure probabilities, point estimates of risk qualification, and sensitivity analysis to delineate critical components that contribute significantly to potential risk from nuclear power plants.

An area of utmost importance is a proper framework for making decisions pertaining to NPP risk based on a multitude of available information. Problems that require decision making in energy production in general, and NPPs in particular, are numerous. Some of the examples requiring decision analysis are: determination of acceptable risk limits to which the general public may be exposed from NPPs, determination of risk of nuclear proliferation due to various types of commercial nuclear power plants as compared to other routes of nuclear proliferation, and determination of the most desirable site (based on public risk, cost, and various site characteristics) for construction of a NPP. This dissertation provides a framework for such decision analyses based on trade-offs pertaining to various competing attributes.

Explanation of Dissertation Format

The dissertation is arranged in five sections; each section is complete in itself and documents the study of a new topic. These sections are written in the form of papers; the first four papers have been published, and the fifth one is in a form suitable for publication. The first two papers employ multi-attribute decision theory for nuclear power plant siting and proliferation issues. The last three papers are related to new quantitative techniques and investigations pertaining to probabilistic risk assessment.

Section I provides a framework for making siting decisions for a nuclear power plant considering a multitude of competing site characteristics such as man-made and natural hazards, economic impact, and various features of the land. A decision model and application for assessing risk of alternative proliferation routes is given in Section II, and a method and computer code for propagating uncertainty in PRA is provided in Section III. An investigation on the use of point-estimate quantification of risk associated with nuclear power plants is given in Section IV, and a sensitivity analysis approach for analyzing robustness of PRA results is provided in Section V.

Scope of Investigation

The investigation pertains to various topics in decision analysis and probabilistic risk assessment as it relates to Nuclear Power Plants. The topics of decision analysis and PRA are inherently related for solutions to safety problems in NPPs. Thus, the PRA methods are used to help in making decisions such as design modifications, changes in operating procedures, and relaxing or imposing regulatory requirements. To this end, PRA can be considered as a safety-related decision tool where the strengths and weaknesses of a complex system can be revealed through analytical procedures.

However, a formal decision analysis tool can be used for more generic problems, where multiple, competing factors or attributes are involved. The formal decision analysis explicitly considers the preferences and trade-offs of the decision-maker for the attributes of interest.

The dissertation addresses five topics related to nuclear power plants, all of which are directed to help either explicitly or implicitly in making engineering decisions. The method of formal decision analysis, namely, multi-attribute utility theory (MAUT), is used for problems in Sections I and II. The improvements in the methods of PRA that can be used for making safety related decisions are provided in Sections III, IV and V.

Significance of the Selected Topics

Five topics were selected here on the basis of their significance in making decisions for a variety of issues related to the nuclear industry.

The formal decision analysis model, namely MAUT, is used for making nuclear power plant siting decisions from among three alternatives. The methods actually used by the utility for making the siting decision did not involve formal decision analysis. However, the decisions resulting through formal decision analysis coincided with the ones actually made by the utility. The methodology for developing models for selecting sites is, of course, not limited to nuclear power facilities, and can be extended to other energy facilities as well. Such energy facilities may include fossil, nuclear fusion, and solar power plants.

The formal decision analysis was again applied to assess risk from alternate routes of nuclear proliferation. Various potential routes, including nuclear power plants such as high-temperature gas-cooled reactors (HTAR), pressurized water reactors (PWR), and boiling water reactors (BWR), were considered to evaluate the relative difficulty with which a terrorist group, or a nonnuclear country, could develop nuclear weapons. The formal decision analysis methodology again reinforces decisions made by the nuclear community in general, but based on other methods of qualitative and quantitative analysis.

Decisions to reduce risks associated with nuclear power plants are being made increasingly through the use of Probabilistic Risk Assessment methods. An important aspect influencing risk decisions is the uncertainty in the calculated risk. Since accidents in nuclear power plants resulting in release of large quantities of radioactive materials have never occurred, and since the probability of such accidents occurring during the lifetimes of the nuclear power plants is very small, the predicted risk is based on many assumptions, some of them very uncertain in nature. To calculate the uncertainty in risk calculations, the uncertainties in various assumptions and events involved are propagated through the risk model. In Section III, the Discrete Probability Distribution (DPD) method is developed and compared with other techniques of propagating uncertainties such as the Monte-Carlo method, and propagation of moments. The accuracy of the DPD method and the computation times are also compared and found to be favorable. The DPD method has the advantage that the failure probability distributions of the basic events can be obtained, using histograms, through the use of limited data and engineering judgment. The DPD method is "exact," in the sense that no sampling errors are involved. Further, the propagation of discrete probability distributions allows us not to make assumptions regarding certain types of probability distributions of the basic events without adequate validation.

Another issue that often arises for the PRA practitioners is the uncertainty in decisions pertaining to nuclear power plant risk management based on the point estimate calculations of risk. The interpretation of the point-estimate values of risk is a source of confusion, and is treated differently in different plant PRAs. The lack of a consistent approach has yielded results for different plant PRAs which cannot be compared. Thus, some plant PRAs used median values for the failure probabilities of the basic events, and propagated them to evaluate a point-estimate of the top event. These point-estimates were then compared with the mean or median values of the top-events for other plant PRAs. It is shown in Section IV that the point-estimate of the top-event evaluated using median values of the basic events may deviate significantly from both mean and median values of the top-event. It is also shown that such deviations depend upon the skewness of the basic event probability distributions as well as the complexity of the system logic model. The results obtained through the use of point-estimate evaluation is shown to be optimistic, i.e., the system will appear to be better than it actually is. The interpretations of the point-estimates of power plant risk will also greatly affect the validation procedures for nuclear power plant safety goals. The safety goals are expected to be issued by the U.S. Nuclear Regulatory Commission in the near future. The utilities will then have the burden of proving that the calculated nuclear power plant risks are within the specified safety goals. A major problem in validating the conformance of a nuclear power plant performance to risk goals is the interpretation of the point-estimates, and the rationale of the comparisons of such point-estimates to the safety goal. The work presented in Section IV will help in resolving this issue.

Some of the major decisions nuclear utilities are presently facing pertain to risk management. That is, having obtained an overall knowledge of the risk for a nuclear power plant, through PRA, we need to find the corrective measures that will significantly reduce the risk from the existing level. This requires a ranking of the basic events in order of their contributions to the overall plant risk (or top-event). A major problem is the uncertainty associated with such a ranking due to uncertainties in the probabilities of the basic events. The top event risk is very sensitive to the failure probability distributions of some of the basic events. The uncertainties in the basic event probability distributions are large because of the lack of sufficient data. Thus, a method must be devised to rank the basic events that account for the associated uncertainties. Further, it is highly desirable that such a ranking be robust, i.e., the ranking should not change significantly with assumptions pertaining to the basic event failure characteristics. This problem is addressed in Section V, where a sensitivity analysis approach is developed for measuring the ranking of the basic events in the face of data uncertainty.

SECTION I. A FORMAL METHODOLOGY FOR ACCEPTABILITY ANALYSIS OF ALTERNATE SITES FOR NUCLEAR POWER STATIONS

A formal methodology for acceptability analysis of alternate sites for nuclear power stations

Shahid Ahmed

A. A. Husseiny

Hang Youn Cho

From the Department of Nuclear Engineering, Iowa State University, Ames, IA, 50011, USA

Published in Nuclear Engineering and Design 51(1979):361-388

ABSTRACT

A formal methodology is developed for the selection of the best sites from among alternate suitable sites for a nuclear power station. The method is based on reducing the various variables affecting the decision to a single function that provides a metric for the level of site acceptability. The function accommodates well-known site selection criteria as well as other factors, such as public reactions to certain choices. The method is applied to the selection of a site from three acceptable alternate sites for Wolf Creek nuclear power station, Kansas.

INTRODUCTION

Siting of nuclear power plants requires extensive evaluation of various characteristics of local areas under consideration. Also, several socio-economical and political factors do, in the final analysis, enter into the decision process. In the presence of more than one suitable site for a plant, a trade-off is usually to be made between the myriad of factors affecting the choice.

The multidimensionality of the siting criteria makes the selection among alternatives a rather complex problem. Thus, decisions may be made on the basis of intuition or qualitative but rather incomplete assessment of the situation. Whenever quantitative analysis is emphasized, usually one of the multitude of factors affecting site selection is considered; for example, economics [1] or radioactive release to the public [2, 3]. In these cases, the margin of choice is narrowed by the fact that all appropriate sites have to meet siting regulations regarding public risk while factors that might vary with site, such as land cost, is often irrelevant to the selection.

Here, the multidimensionality of the site selection criteria is reduced to a univariate acceptability function through the use of integrated weights and multivariate decision theory. The utility of lottery wins and losses is paralleled in the present context by the construct of quantitative site acceptabilities which are functions of site characteristics. Comparison between alternate sites in this case is done by deciding on the option which has the maximum acceptability. The approach is versatile enough to accommodate for public attitude, political factors, impact on wildlife and the nonexpert opinion of action (or citizen) groups. Also, the decision makers can incorporate expert evaluation of certain aspects of the siting problem in the selection strategy for possible technical limitations within the utility company. Overall, the methodology presented here is systematic and self-consistent and its viability is demonstrated through practical applications to a case study.

QUANTITATIVE METHODS

Incorporation of the role of various decision-making and public groups in the site selection process was viewed upon from a pure management and social perception by Gros et al. [4]. The deployment of a 1000 MWe base-load unit on the New England coast was studied using a Paretian environmental approach. Specifically, the benefits and costs accruing to four separate groups are examined; namely, environmentalists, utility companies, regulatory agencies, and local groups. Consulting with individuals with a preference or utility functions for each of those were assessed over four proxy attributes: monetary costs, population within 15 miles of the site, temperature of water release after cooling, and capacity of the site measured in a number of 1000 MWe units. It can, however, be argued that only considering four attributes is far from sufficient to take into account the full dimensionality of the site selection among various alternatives. Also, if a large number of attributes are considered, the verification of the assumptions on which the Paretian model is based would become extremely difficult, if not impossible, since the number of restrictions on the model increases linearly with the number of attributes and over the number of separate groups considered. This is a useful preliminary work, but further analysis is required before it can be proven viable for the involved process of nuclear power plant siting.

Keeney and Nair [5] have also addressed the nuclear power plant siting problem in the perspective of multiattribute utility theory. The complexities encountered in siting decisions are identified with a theoretical exposition of the theory. The applicability of the approach is not given in true relationship to the nuclear power plant siting. Since the suggested approach has not been applied to a real situation, no mention is given as to how the large number of siting attributes would be handled, particularly since with the increase of the number of siting attributes the basic assumptions of the theory used can no longer be verified.

Hassan [6] has used a grading scale having a range from 1 to 5, which gives the relative importance of each site for a particular factor considered. Each factor is divided into subfactors which are further weighted on a scale from 0 to 10 to show their relative importance. A simple addition of the product of grade and weight is made for each alternative. The site with the highest product is the optimal choice. This approach is, in fact, equivalent to the multiconsequence risk indifference in utility theory [4]. The approach is limited and the multi-consequence risk aversion and multi-consequence risk seeking are altogether neglected.

Other quantitative methods of ranking selection criterion have been proposed [7-10]. The approach tends to overlook the technical details of each criterion and assigns a single weight which if in error would lead to cumulative gross errors in the final result. Although Hassan [6] has considered the various facets of each criterion through ranking of its subfactors, the weights assigned do not discriminate between the importance of different subfactors due to the large number of parameters involved in the analysis. The method could have been refined by analyzing the subfactors involved in each criterion separately with no regard to other considerations and then seeking a method for combining the intra-criteria weighting factors.

Since ranking methods may not resolve the conflicts between technologists and environmentalists, Fischer and Ahmed [11] used the Delphi method as a means to reconcile differences in opinion among concerned groups. An impact quotient is used to correlate between environmental quality with and without the plant through importance factors for each factor considered. To rank the sites in the order of increasing environmental costs, the additive utility is used to reflect the desirability of a particular alternative. However, additive utility fails short of accounting for risk aversion as well as risk seeking.

Beer [7] also suggested an "impact analysis matrix" which permits evaluation of interactions between proposed actions which may cause environmental impact and existing characteristics and conditions of the environment. This could give an overall subjective feeling about the impact of the power plant construction but does not indicate a quantitative approach towards decision on site selection among various alternatives.

SELECTION PRINCIPLES

It is logical to assume that the obvious and choice sites for nuclear power plants will be gradually consumed and future plants will have to be located on much more environmentally constrictive sites. Also, choice of alternatives will become more complex as the siting concepts advance as, for example, the case with offshore plants which are already a reality. Furthermore, future siting of nuclear plants is apt to face more stringent government (State and Federal) regulations. The reaction of citizen groups, environmentalists and the like to technological expansion may become severe and may lead to the development of sophisticated organizations that take part in decision making. Thus, if the utilities are to avoid financially costly delays in approval of a site or the utlimate rejection of the proposed site, the future site evaluation and selection must be done using a rational and systematic method of evaluation.

Fundamental to the selection process is the basic premise that there is no ideal site available. Perhaps a site that is suitable to a strict conservationist is unsuitable to a utility company if site characteristics appear to impose intolerable economic costs or requisite an engineering solution which is not practicable. Values of land and land reclamation for multipurpose utilization are different in the eyes of various "beholders." Therefore, the objectives of the site selection program and the method of evaluation of alternative sites must be well-planned prior to initiation of the program. A site selection program is successful if it provides adequate data on several alternative sites, tabulates these data in a logical order of importance and ranks sites on the basis of suitability under different design alternatives at each site. It is also important in site analysis as well as in conducting the selection decision to deal with the problem as a whole rather than complete devotion to a single criterion.

The following steps are recommended in the process of site selection.

1. At the onset, all areas that have no apparent potential for power plants be eliminated from consideration; for example, restricted areas such as cities and other densely populated areas, national parks, and areas where adequate water is considered highly impractical to develop.

2. The remaining potential sites should then be studied for further definition of siting possibilities employing the most critical or limiting criteria which, if could not be met individually, should cause rejection of a part or all of an individual site.

3. Specific sites should then be identified within the areas remaining after step 2 is completed. All site developmental and environmental criteria as well as other relevant restrictions should be employed as applicable in testing the suitability of sites at this stage.

4. By reducing the number of options at step 3, a qualitative comparison may be conducted. All characteristics common to the final group of alternative sites may not be considered any further and only those factors for which the level of acceptance differs among the group

should be considered for more analysis.

5. The methodology described here may then be used to decide among various specific sites as identified in step 3. The decision is arrived at by using acceptability functions which would be considered as a measure of importance of the attributes.

SITING MODEL

In modelling the power plant siting decision process, the analyst must obtain an objective function including the multiple attributes which describe the effectiveness of a decision. Such an objective function would indicate the relative ranking of consequences and identify the trade-offs among various levels of different attributes. In a risk-free environment, the optimal decision would be the one that maximizes the objective function.

But the power plant siting decision problem can involve uncertainties. For instance, the regulatory process which governs such decisions may change in some unpredictable fashion. This type of uncertainty should be considered in the modelling effort and the objective function should allow the uncertainty to be handled easily. One approach is to design an objective function such that the decision which maximizes the expected value of the objective function is the optimal decision. Such an objective function is the acceptability function used here to give a measure of the effectiveness of given site characteristics in meeting the criterion relevant to that site characteristics.

There are several ways of synthesizing acceptability functions. A direct approach would have the assessor to consider this multidimensional problem as a whole, assessing preferences for sets of attributes in the several years. This is an enormous time-consuming process, with no guarantee of consistency, since people have trouble visualizing trade-offs in more than one dimension. In order to have the assessment made in less time, the problem should be broken down into its simpler components and then reconstructed.

By assessing various trade-offs between specific quantities or qualities, the decision analyst can find an acceptability function that can serve as a guide in decision making. If the consequences chosen satisfy certain independence properties, the assessment problem is simplified. The two independence properties to be considered are preferential independence of attributes and acceptability independences.

Site Characteristics

Here, site characteristics refer to those attributes which meet the siting criteria within an acceptable range in addition to the well-known criteria which have been stated in the regulatory guides [12] and discussed in several places [13]. Several other factors are considered explicitly here; such as ease of construction, domestic water, and ecology. Public acceptance plays a major role in licensing. In many situations, this factor is considered as part of environmental, ecological or social impacts. However, resistance to siting of power plants in general and nuclear plants in particular in many locations has been a result of psychological discomfort and/or reluctance of small communities to social disturbances due to the influx of construction workers. Objection to extending transmission lines through the fields or to rising electric power cost have been reflected in the public attitude towards the activities of the utility companies.

The major site characteristics considered here are listed in Table 1.

Table 1. Site characteristics and selection criteria

No. (i)	Characteristics (Y_i)	Indifference probability (P_i)	Scaling factors (k_i)
1	Cooling water	0.083	-0.99
2	Soil	0.069	-0.894
3	Construction	0.059	-0.69
4	Services	0.054	-0.99
5	Domestic water	0.047	-0.933
6	Accessibility	0.056	-0.998
7	Compatibility of land use	0.049	-0.842
8	Ecology	0.064	-0.926
9	Demography	0.069	-0.978
10	Topography	0.058	-0.54
11	Geology	0.073	-0.927
12	Aircraft hazard	0.046	-0.856
13	Man-made hazard	0.048	-0.949
14	Natural phenomena	0.052	-0.998
15	Meteorology	0.046	-0.981
16	Public acceptance	0.056	-0.92
17	Economic impact	0.069	-0.962

The list is by no means exclusive and usually additional factors may be added whenever an issue of new concern arises. Some of the characteristics listed in Table 1 have been given more general names to include the more specific measures as sub-factors or sub-attributes, such as seismology, hydrology, aquatic biology, and terrestrial biology.

Since the characteristics Y_1, Y_2, ... Y_{17} represent a set $\{y_i\}$, i = 1, 2, ..., 17 of parameters which measure the achievement of the siting objective, then

$$Y = Y_1 * Y_2 * Y_3 \ \ldots \ * Y_{17} \tag{1}$$

may be identified as a 17-consequence space which includes all possible values of y_i with y_i being a specific outcome of y_i, i=1, 2, ..., 17. Actually, any member Y_i may have more than outcome y_{ij}, j=1, 2, ..., m which is the case if several sites are being considered. If, for example, the decision is made on the selection of a site from among three suitable sites, then

$$Y_i = (y_{i1}, y_{i2}, y_{i3}) , \tag{2}$$

when subscripts 1, 2 and 3 refer to the number designated to the site in question.

Site Acceptability Functions

For an outcome y_i of a characteristic Y_i, one can define an acceptability function, $U_i(y_i)$ which reflects the desirability level of having an outcome y_i for the characteristic Y_i. This function is obtained from detailed analysis of the factors affecting the value of the outcome y_i which assumes a value between the least desirable level y_{i*} and the most desirable level y_i^*. The acceptability function may be normalized to unity such that

$$U(y_i^*) = 1$$

and

$$U(y_{i*}) = 0 . \tag{3}$$

Although the acceptability $U(y_i)$ of an outcome y_i may be discrete, it is assumed here that all acceptability functions are continuous and are bounded by the values given in equation 3.

We may also define the outcome y_{i-} of the complement Y_{i-} of the characteristic Y_i as the 17-consequence space that includes all characteristics except the ith characteristic, that is

$$Y_{i-} = \{Y_1, Y_2, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_{17}\} . \tag{4}$$

Similarly, the complement of two characteristics Y_i and Y_j can be written as

$$Y_{ij-} = Y_i \times Y_2 \times \ldots \times Y_{i-1} \times Y_{i+1} \times \ldots \times Y_{j-1} \times Y_{j+1} \times \ldots \times Y_{17}, \tag{5}$$

where Y_{i-} is a member of Y_{i-} and y_{ij-} is a member of Y_{ij-}.

Supra-acceptability Function

If N sites are considered, then for each alternate site S_n, $n = 1, 2, \ldots, N$ one may define a supra-acceptability function μ, which combines the sum of all the acceptability functions corresponding to the consequences or outcomes of all the 17 characteristics weighted by an indifference probability P_i, $i = 1, 2, \ldots, 17$; namely

$$\mu = \sum_{i=1}^{17} P_i U_i(y_i) , \tag{6}$$

where

$$\sum_{i=1}^{17} P_i = 1 , \tag{7}$$

that is, if $U_i = U_i(y_i^*) = 1$, then $\mu = 1$.

The siting model gives an indifference probability distribution $\pi(P_i)$ representing which outcome may occur and the associated likelihood. Values of P_i will be defined and determined below.

The supra-acceptability function of the form given in equation 6 is in fact a best estimate of acceptance level when the decision makers are only interested in the equivalent quantitative value estimate of the level; that is, they are indifferent to risk and neither liberal risk seeking nor highly risk averse. In other situations, such as those presented below, a risk aversion preference may be followed. The physical interpretation of the three distinct attitudes towards risk in decision making may be interpreted through consideration of two alternative lotteries A and B.

In lottery A, Figure 1, the outcomes of the 17 characteristics in the consequence space are determined by 17 two-pronged lotteries, each giving a probability P_i at the best level and a probability level $(1 - P_i)$ at the worst level.

In lottery B, represented in Figure 2, the outcome values of the 17 characteristics in the consequence space for siting are determined by a single-pronged lottery, with a probability P_i of getting the best level for each consequence and a probability $(1 - P_i)$ of getting the worst level for each consequence.

Lottery A would be preferred to lottery B if the site selection preference for the assessor is to have a combination of best and worst characteristic values rather than all best or all worst characteristics. This represents a multivariate risk aversion. Multi-consequence risk indifference means the decision maker has no preference for either lottery, while a preference for lottery B over lottery A represents the liberal multi-consequence risk seeking.

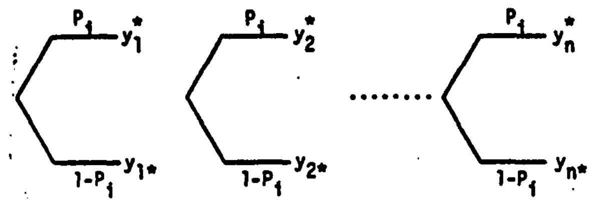


Figure 1. Lottery A

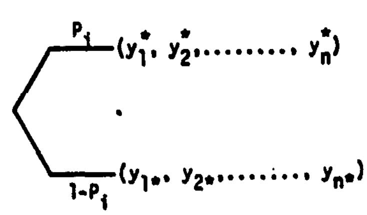


Figure 2. Lottery B

To illustrate the technique for assessing the indifference probability P_i, cooling water characteristic Y_1 may be taken as an example. A certain consequence with cooling water supply at its most acceptable level, and all the other 16 characteristics at their least acceptable level may be compared to a lottery yielding the consequence with all the 17 characteristics at their most acceptable level with probability P_i or the consequence with all the 17 characteristics least acceptable with probability $(1 - P_i)$. The object is to find a value of P_i such that the decision maker is indifferent between the lottery and the certain consequence. This utilizes the indifference probability procedure [13] which can be expressed, in general, as follows:

$$\{Y_1^*, Y_{1-*}\} \sim P_i\{Y_i^*\} + (1 - P_i)\{Y_{i*}\}, \tag{8}$$

where ~ refers to indifference between both sides of the relationship. The decision point is shown in Figure 3. From equations 3, 6 and 7:

$$u(y_1^*, y_{1-*}) = P_i \; . \tag{9}$$

Actually, the evaluation of a point of the most probable value of P_i depends on the acceptability pattern of the decision makers subject to regulation constraints.

The Delphi

To reconcile the differences among various experts regarding values of the indifference probabilities P_i, a Delphi questionnaire method is used. The Delphi method is an iterative questioning and answering

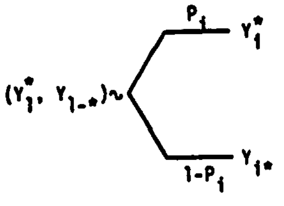


Figure 3. Lottery to find the difference probability P_i for the characteristics

procedure. The basic characteristics of the method are anonymity, feedback, and statistical response. A total of ten questionnaires have been distributed to a group of well-informed persons of different background and experience in the nuclear energy field. The results have shown a considerable degree of homogeneity in spite of the diversity between those who have responded. The final values of P_i which resulted from the Delphi analysis are contained in Table 1. A sample of the questionnaire is given in the Appendix.

Site Intra-characteristics

Each characteristic in the 17-consequences space comprises a subset of intra-characteristics, X_{ij}, that is:

$$Y_i = \{X_{ij}\} \ , \quad j = 1, 2, \ldots, J_i \ , \tag{10}$$

where j is the index of the jth intra-characteristic X_{ij} in the ith set Y_i and J_i is the total number of intra-characteristics within that set. A list of the intra-characteristics is given in Table 2 with corresponding measures and extreme values of the outcome x_j of X_{ij}, $j = 1, 2, \ldots, J_i$; namely, the best consequence x^*_{ij} and the worse x_{ij*}.

Many of the intra-characteristics listed have been assigned a subjective metric; nevertheless, appropriate units may be used when standards are set for each of those factors. An expert judgment is then required for such items from specialists in evaluating each site characteristic. The expert values can be alternatively transformed to subjective units. However, use of appropriate units would be more relevant since the transformation from characteristics to acceptability functions will remove the inconsistency among the selected measures.

Inter-acceptability Functions

For each intra-characteristic X_{ij} in the J-consequence space representing the characteristic set $\{Y_i\}$, an inter-acceptability function $u_{ij}(x_j)$ may be defined, such that

$$u_{ij}(x^*_{ij}) = 1$$

and

$$u_{ij}(x_{ij*}) = 0 \quad \text{for any j from 1 to } J_i \ . \tag{11}$$

These functions can be combined with appropriate integrated weight k_i and indifference probabilities p_i to give the acceptability of the

Table 2. Site intra-characteristics and acceptability bounds

No. (i)	Characteristic (Y_i)	Index (j)	Intra-characteristic (X_{ij})	Measure	Acceptability bounds: Best (x^*_{ij})	Acceptability bounds: Worst (x_{ij*})	Acceptability bounds: Indifference probability (P_{ij})
1	Cooling water	1	sources	ft^3/s(average)	20,000	1,000	0.723
		2	availability	average/max. flow rate	1	0	0.761
		3	distance	mile	0	4	0.615
		4	purity	subjective	100	0	0.5
		5	incremental water temperature at peak ambient water temperature period of the year	°F	0	5	0.554
2	Soil	1	classification	subjective	100	0	0.368
		2	characteristics	subjective	100	0	0.355
		3	properties	subjective	100	0	0.35
		4	bearing capacity	subjective	100	0	0.391
		5	fixation of radwaste	subjective	100	0	0.295
		6	ground water depth	subjective	100	0	0.608
3	Construction	1	land preparation	subjective	100	0	0.608
		2	labor	\$/h	0	25	0.675
4	Services	1	electricity	subjective	100	0	0.71
		2	gas	subjective	100	0	0.40
		3	drainage and sewage	subjective	100	0	0.582
		4	location of services	subjective	100	0	0.573
		5	capacities	subjective	100	0	0.664
		6	distribution lines	subjective	100	0	0.727
5	Domestic water	1	availability	subjective	100	0	0.636
		2	characteristics	subjective	100	0	0.636
		3	size, depth and pressure of mains	subjective	100	0	0.555

Table 2. Continued

No. (i)	Characteristic (Y_i)	Index (j)	Intra-characteristic (X_{ij})	Measure	Acceptability bounds Best (x^*_{ij})	Worst (x_{ij*})	Indifference probability (P_{ij})
		4	consumption	subjective	100	0	0.536
		5	distribution systems	subjective	100	0	0.518
		6	sources and location	subjective	100	0	0.527
6	Accessibility	1	land characteristics along route of transmission line network	subjective	100	0	0.67
		2	distance to transmission line intensities and substation	mile	0	50	0.77
		3	highway	subjective	100	0	0.73
		4	railroad	subjective	100	0	0.79
		5	waterways	subjective	100	0	0.75
7	Land use	1	agriculture	subjective	100	0	0.348
		2	special land use (parks, wildlife, refuge, etc.)	subjective	100	0	0.305
		3	residential characteristic	subjective	100	0	0.309
		4	public facilities	subjective	100	0	0.295
		5	sensitive industries (oil and gas pipe lines)	subjective	100	0	0.314
		6	weapons testing	subjective	100	0	0.314
8	Ecology	1	river or lake classification (irrigation, fishing, recreation, wild, etc.)	subjective	100	0	0.71
		2	endangered species	subjective	100	0	0.52
		3	biological conditions	subjective	100	0	0.63
9	Demography	1	existing population within 1 mile	persons	0	10,000	0.654

Table 2. Continued

No. (i)	Characteristic (Y_i)	Index (j)	Intra-characteristic (X_{ij})	Measure	Acceptability bounds: Best (x^*_{ij})	Worst (x_{ij*})	Indifference probability (P_{ij})
		2	expected population within 1 mile, year 2000	persons	0	10,000	0.569
		3	existing population within 20 miles	million persons	0	7	0.477
		4	estimated population within 20 miles	million persons	0	15	0.407
		5	proximity to nearest population center	miles	50	0	0.585
10	Topography	1	number of potential sites per region	sites/region	5	0	0.583
		2	terrain for power plant	subjective	100	0	0.608
11	Geology	1	mine hazard	subjective	100	0	0.423
		2	oil hazard	subjective	100	0	0.407
		3	seismology	subjective	0	100	0.631
		4	tectonic (geology dealing with faulting and folding)	subjective	0	100	0.577
12	Aircraft impact	1	proximity to airport	miles	10	0	0.522
		2	bombing ranges	miles	100	0	0.607
		3	low level training routes	miles	100	0	0.528
13	Man-made hazards	1	ship collision and explosion	subjective	0	100	0.575
		2	flammable gas and vapor clouds	subjective	0	100	0.575
		3	toxic chemicals	subjective	0	100	0.525
		4	fire	subjective	0	100	0.592
14	Natural phenomena	1	tornadoes	subjective	0	100	0.52
		2	high water (HW) level floods	foot	0	100	0.73
		3	HW from dam failure	foot	0	100	0.75
		4	HW from hurricanes	foot	0	100	0.70

Table 2. Continued

No. (i)	Characteristic (Y_i)	Index (j)	Intra-characteristic (X_{ij})	Measure	Acceptability bounds		
					Best (x^*_{ij})	Worst (x_{ij*})	Indifference probability (P_{ij})
		5	HW from tsunamies	foot	0	100	0.62
		6	HW from seiches	foot	0	100	0.52
		7	low water	foot	0	100	0.43
15	Meteorology	1	sustained wind velocity	mph	20	0	0.73
		2	diffusion (stability category) [14]	A to G	G	A	0.78
		3	inversion temperature	°C/100 m	2	-1.5	0.73
16	Public acceptance	1	occupational group (unions, merchants, corporations)	subjective	100	0	0.5
		2	citizen group (civic associations, environmentalists, etc.)	subjective	100	0	0.436
		3	media	subjective	100	0	0.45
		4	state laws	subjective	100	0	0.628
17	Economic impact	1	cost of cooling water supply related to the water intakes, pipelines, storage or water supply reservoirs and pumping facilities required to deliver a reliable supply of water to a plant at each identified site	M$	0	14	0.64
		2	land acquisition costs	$/acre	0	10,000	0.5
		3	site preparation costs, such as clearing, grading, flood protection and waterproofing, on-site access improvements and relocated road and bridge construction costs	M$	0	8	0.55

Table 2. Continued

No. (i)	Characteristic (Y_i)	Index (j)	Intra-characteristic (X_{ij})	Measure	Acceptability bounds: Best (x^*_{ij})	Acceptability bounds: Worst (x_{ij*})	Acceptability bounds: Indifference probability (P_{ij})
		4	foundation cost estimates such as cost of dewatering special foundation treatment excavation, foundation base, mats, pile foundation if required, and backfilling	M$	0	20	0.61

corresponding outcome y_i of Y_i, that is

$$Y(y_i) = \frac{1}{k_i} \{\Pi_{j=1}^{J_i}[1 + k_i p_{ij} u_{ij}(x_{ij})] - 1\} \text{ ,} \tag{12}$$

where p_{ij} is to be evaluated in a fashion similar to that used in determining the indifference probabilities P_i. The lottery used is shown in Figure 4, which can be represented by

$$\{x_{ij}^*, x_{ij-*}\} \sim p_{ij}\{x_{ij}^*\} + (1 - p_{ij})\{x_{ij*}\} \tag{13}$$

and therefore

$$U(y_i(x_{ij}^*, x_{ij-*})) = p_{ij} \tag{14}$$

from equations 6 and 11. Here, x_{ij-} represent the complement of x_j in the sense described above for y_i. The results have been obtained

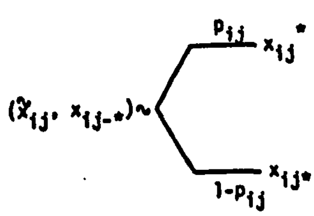


Figure 4. Lottery to determine the difference probability p_i for the intra-characteristics

using the Delphi approach and are listed in Table 2.

In reaching equation 12, the underlying assumption is that the decision maker is not indifferent to risk [15]. More specifically, a conservative risk aversion approach is used in obtaining the inter-acceptability functions or levels corresponding to intra-characteristic X_{ij}. This is due to the fact that the intra-characteristics must pass a stringent scrutiny based on measurements and analysis in order to pass the acceptance test. Once this is done properly, the need for further conservation in evaluating the site supra-acceptability from the more general characteristics Y_i can be disposed of and hence the multi-consequence risk function of equation 6 can be used.

For risk aversion, the following relationships [16] define the value of p_j and k of equation 12,

$$\sum_{j=1}^{J_{if}} p_{ij} > 1,$$

$$0 < p_{ij} \leq 1$$

and

$$-1 < k_i < 0 \; . \tag{15}$$

The scaling factors k_i are obtained from equation 12 by setting the outcomes intra-characteristics at their best consequence level; that is,

$$1 = k_i - \Pi_{j=1}^{J_i}(1 + k_i p_{ij}) \; , \tag{16}$$

the numerical values are then obtained by iteration of the roots of this J_i-order equation and the root that satisfies equation 15 is selected. The results are given in Table 1.

Equation 12 may be substituted into equation 6 to directly give the supra-acceptability function μ_k for site S_k, namely

$$\mu_i = \sum_{i=1}^{I=17} \frac{p_i}{k_j}\{\Pi_{j=1}^{J_i} [1 + k_i p_{ij} u_{ij}(x_{ij})] - 1\}. \qquad (17)$$

Once the acceptability functions are evaluated at x_{ij}, μ_k can be determined and the site to be selected is the one with the maximum supra-acceptability $\bar{\mu}_k$, that is

$$\bar{\mu}_k = \max_{p_i, k_i, p_{ij}} \{\mu_k\}. \qquad (18)$$

Validity of Acceptability and Characteristic Relationships

The multi-attribute utility theory reduces to simplistic functional form of representing the acceptability function under certain basic conditions [17-22]. The acceptability concept is theoretically sound, and the mathematical details are not involved. However, the difficulty comes when one tries to specify reasonable procedures for obtaining multi-attributed acceptability functions. In a real problem like ours, if the assumptions are verified, the functional form can be used to simplify the requisite assessments needed to specify the acceptability function.

The main assumptions used concern the concept of preferential independence and acceptability independence. The consequence pair $X_{ij} \times X_i$ is said to be preferentially independent of its complement $X_{ij\ell-}$ if one's preference order for consequences $(x_{ij}, x_{i\ell}, x_{ij\ell-})$, with $x_{ij\ell-}$ held fixed, does not depend on the fixed amount $x_{ij\ell-}$.

For example, consider the cooling water characteristic Y_1. One needs to verify whether the flow rate of the water source (X_{11}) and the distance of water source (X_{13}) are preferentially independent of other intra-characteristics in Y_1, that is, X_{12}, X_{14} and X_{15}. To do so, one starts by determining the flow rate x_{11} ft^3/s of a water source subject to the condition that

$$(x_{11}\ \text{ft}^3/\text{s};\ 3\ \text{miles}) \sim (3000\ \text{ft}^3/\text{s};\ 2\ \text{miles})\ . \tag{19}$$

That is, a water source of flow rate, x_{11} ft^3/s at a distance of 3 miles is indifferent to a water source of flow rate 3000 ft^3/s at a distance of 2 miles. A figure of x_{11} = 5000 ft^3/s may be obtained. However, the exact number is not important for verifying the assumptions, since the interest is in knowing if this number changes when other attributes are varied. Thus, the other attributes may be varied to any desirable magnitude and then x_{11} is re-evaluated using equation 19. The result remains the same regardless of the changes. In fact, the value of x_{11} is invariant for any trade-offs between water source and its distance from the power plant. Hence, it is concluded that water source and the distance are preferentially independent.

Other intra-characteristics in the 5-consequence space $\{Y_1\}$ are examined for independence. The availability (X_{12}) and incremental water temperature (X_{15}) are found to be preferentially independent. In fact, going through a number of combinations, we find that all intra-characteristics are preferentially independent of the remaining set of intra-characteristics in the J-consequence space for all Y_is. Similarly, the preferential independence for the characteristics in the 17-consequence space $\{Y_i\}$ are verified.

On the other hand, the conditions for acceptability independence depend upon the decision maker's preferences for lotteries involving uncertainty. One says that X_{1j}, for example, is inter-acceptability independent of X_{1j-} if one's preference order over lotteries on X_{1j}, represented as $(\tilde{x}_{1j}, x_{1j-})$ with X_{1i-} held fixed, does not depend on the fixed amount x_{1i-}. Here again, the same approach is applicable. Let us see, for example, if the incremental water temperature (°C), X_{15} is acceptability independent of X_{15-}. The other 4 intra-characteristics are set at reasonable magnitudes, and the conditional inter-acceptability function over X_{15} from 0°C to 5°C is assessed. It is found that 3.4°C was indifferent to a 50-50 lottery yielding either 0°C to 5°C. Then the values of the X_{15-} intra-characteristics are changed to less desirable magnitudes. Again, it is found that 3.4°C is indifferent to a 50-50 lottery yielding either 0°C to 5°C. This is verified for any fixed value of X_{15-}. Thus, it is found that relative preference for any lotteries and consequences involving uncertainties only about X_{15} would not depend on the other intra-characteristics of Y_1. The conclusion is, therefore, that X_{15} is inter-acceptability independent of the other four intra-characteristics. By going through identical procedures, it is verified that all the remaining intra-characteristics are also inter-acceptability independent. Similarly, the acceptability independence for the characteristics set $\{Y_i\}$ is verified.

Evaluation of Inter-acceptability Patterns

The single-consequence inter-acceptability function $u_{ij}(x_{ij})$ can be obtained by the technique demonstrated in the following example: let us take the intra-characteristic X_{15}, incremental water temperature, and determine the feasible range of values of x_{15} from the most preferred value to the least preferred value. Scaling the inter-acceptability function from zero to one, as in equation 11, u_{15} (0°C) = 1 and u_{15} (5°C) = 0. Next, consider the lottery with a probability of 0.5 of obtaining x^*_{15} and a probability of 0.5 of obtaining x_{15*}. Find a value $\tilde{x}_{15}$ such that the decision maker is indifferent to a choice between the lottery and the certain consequence:

$$\{\tilde{x}_{15}\} \sim 0.5\{x^*_{15}\} + 0.5\{x_{15*}\}. \tag{20}$$

Then, the inter-acceptability corresponding to the point $\tilde{x}_{15}$ in the intra-characteristic range from 0°C to 5°C is:

$$u_{15}(\tilde{x}_{15}) = 0.5\ . \tag{21}$$

A value of $\tilde{x}_{15}$ = 3.4°C is found; hence,

$$u_{15}(3.4°\mathrm{C}) = 0.5\ . \tag{22}$$

Repeating the process for the following lotteries:

$$\{4.3°\mathrm{C}\} \sim 0.5\{\tilde{x}_{15}\} + 0.5\{x_{15*}\} \tag{23}$$

and

$$\{2.2°\mathrm{C}\} \sim 0.5\{\tilde{x}_{15}\} + 0.5\{x^*_{15}\} \tag{24}$$

whence

$$u_{15}(4.3°\mathrm{C}) = 0.25 \tag{25}$$

and

$$u_{15}(2.2°\mathrm{C}) = 0.75 \; . \tag{26}$$

Other points can be obtained to define the inter-acceptability pattern, $u_{15}(x_{15})$. The process is repeated for all the intra-characteristics over the 17 siting consequence space. The resulting patterns are plotted in Figures 5 through 19. Each pattern may be fitted to a smooth curve of the form

$$u_{ij}(x_{ij}) = \alpha_{ij} + \beta_{ij} \exp(\gamma_{ij} x_{ij}) , \quad i = 1, 2 \ldots 17 \text{ and } j = 1, 2 \ldots J_i , \tag{27}$$

where α_{ij}, β_{ij}, and γ_{ij} are the inter-acceptability shape factors. The resulting equations for all the intra-characteristics are,

$$\begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{14} \\ u_{15} \end{bmatrix} = \begin{bmatrix} 1.225 \\ 1.225 \\ -0.125 \\ -0.1327 \\ 1.225 \end{bmatrix} \begin{bmatrix} -1.225 & 0 & 0 & 0 & 0 \\ 0 & -1.225 & 0 & 0 & 0 \\ 0 & 0 & 1.125 & 0 & 0 \\ 0 & 0 & 0 & 1.133 & 0 \\ 0 & 0 & 0 & 0 & -0.225 \end{bmatrix}$$

$$\times \exp - \begin{bmatrix} 0.8473 \times 10^{-4} & 0 & 0 & 0 & 0 \\ 0 & 1.695 & 0 & 0 & 0 \\ 0 & 0 & 0.5493 & 0 & 0 \\ 0 & 0 & 0 & 0.0214 & 0 \\ 0 & 0 & 0 & 0 & -0.3389 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \\ x_{15} \end{bmatrix} , \tag{28}$$

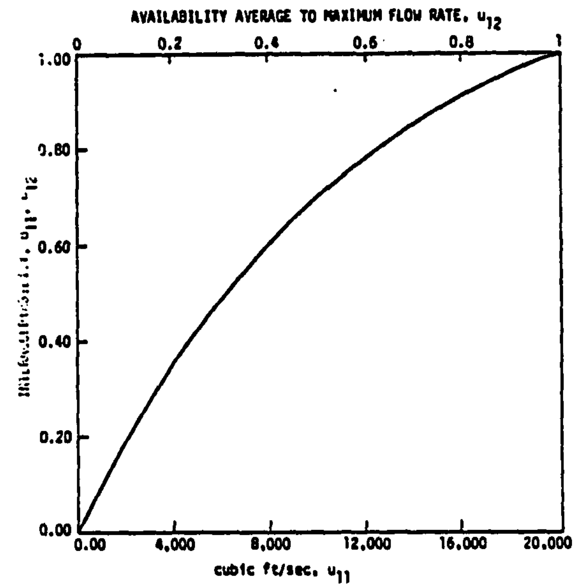


Figure 5. Inter-acceptability for the intra-characteristics x_{11}, x_{12}

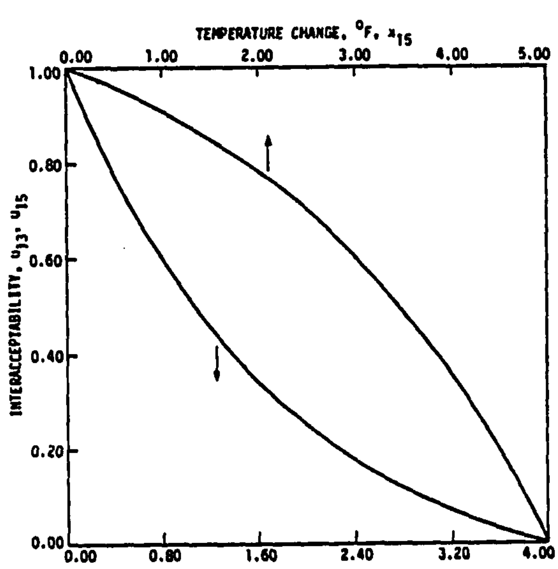


Figure 6. Inter-acceptability of cooling water intra-characteristics x_{13} and x_{15}

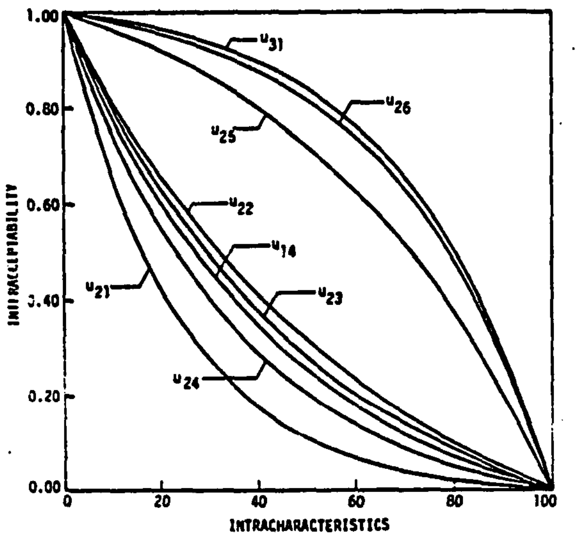


Figure 7. Inter-acceptability for the intra-characteristics x_{14}, x_{21}, x_{22}, x_{23}, x_{24}, x_{25}, x_{26}, and x_{31}

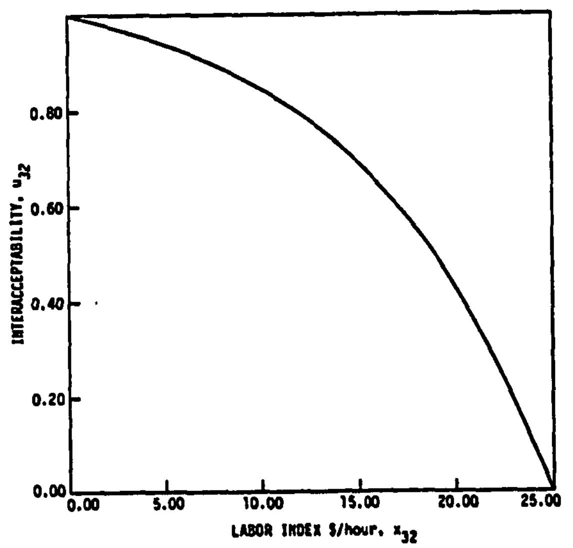


Figure 8. Inter-acceptability for labor index, x_{32}

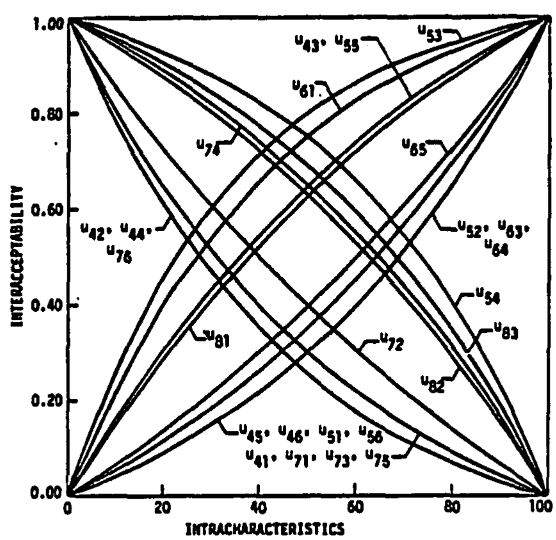


Figure 9. Inter-acceptability for the intra-characteristics of services, domestic water, accessibility, land use and ecology

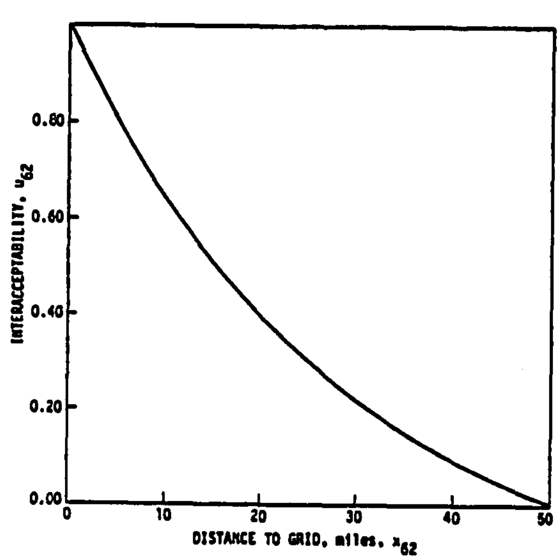


Figure 10. Inter-acceptability of distance to grid

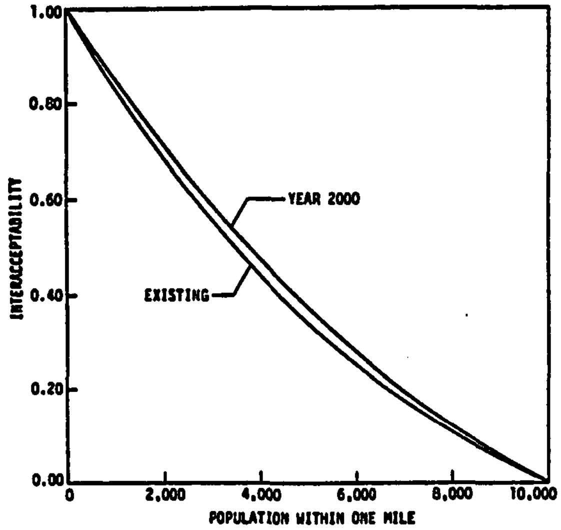


Figure 11. Inter-acceptability of existing population within 1 mile, presently and in the year 2000, x_{91}, x_{92}

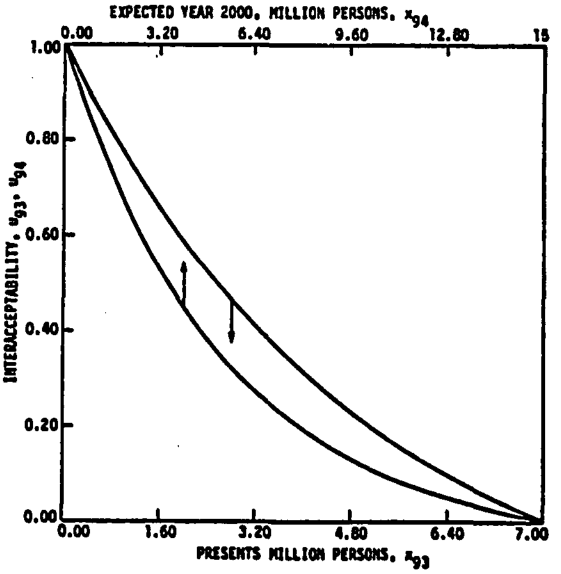


Figure 12. Population within 20 miles

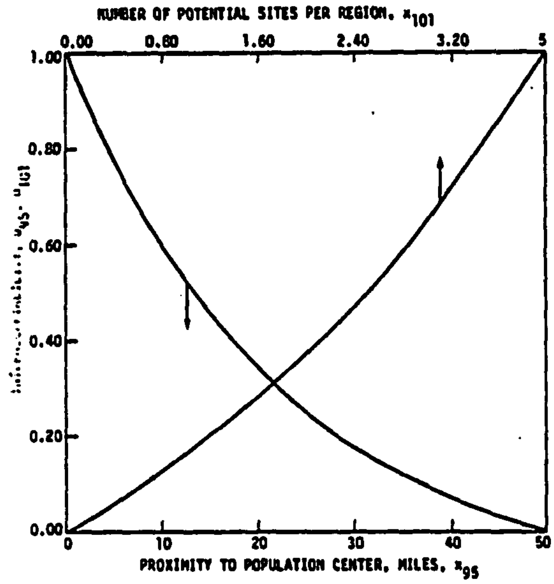


Figure 13. Proximity to population centers and number of potential sites per region

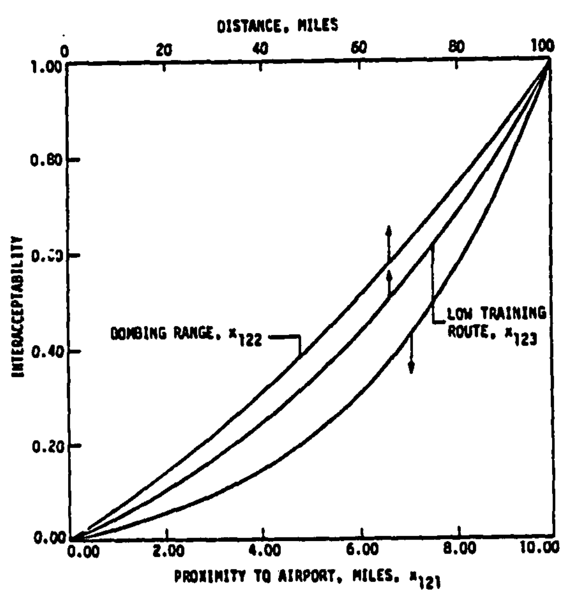


Figure 14. Aircraft impact

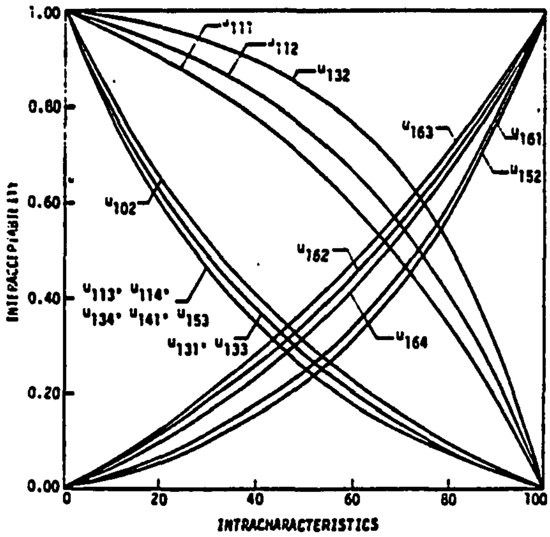


Figure 15. Inter-acceptability of intra-characteristics of x_{102}, geology, meteorology, man-made hazards, and public acceptance

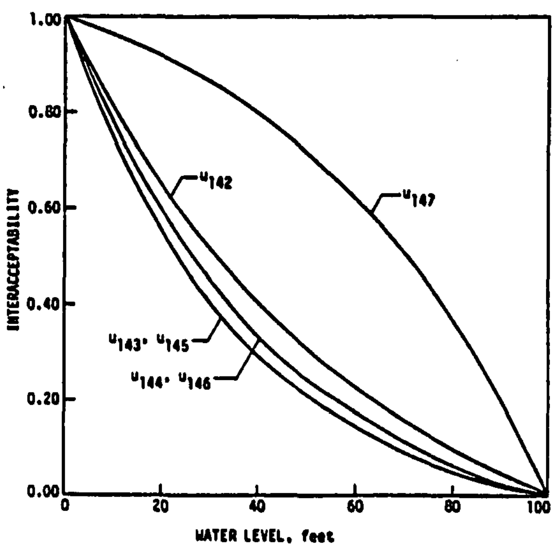


Figure 16. Water level

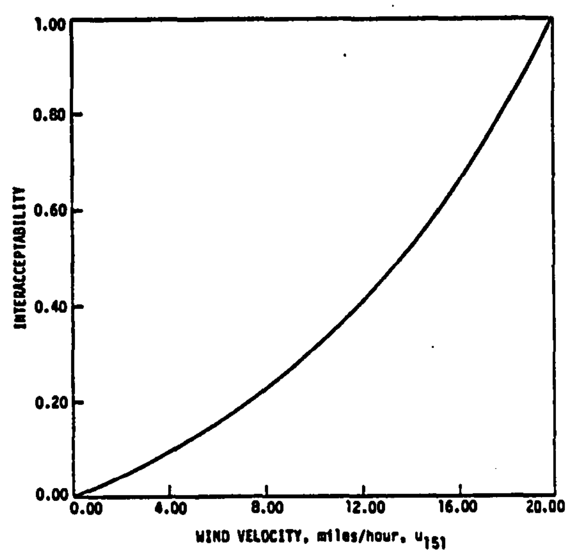


Figure 17. Wind velocity (mph)

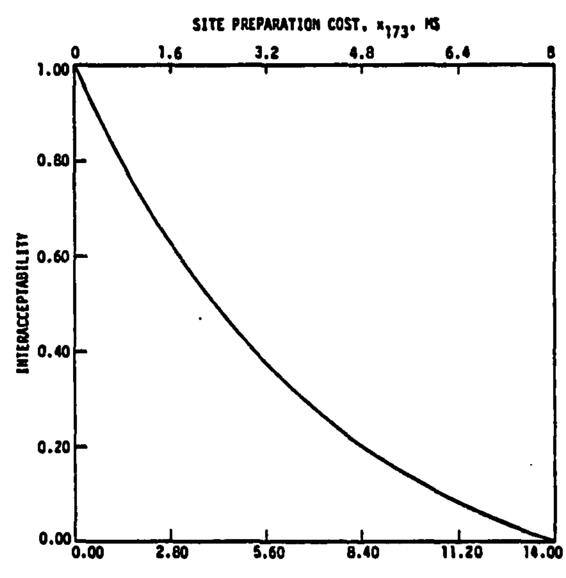


Figure 18. Cooling water and site preparation costs

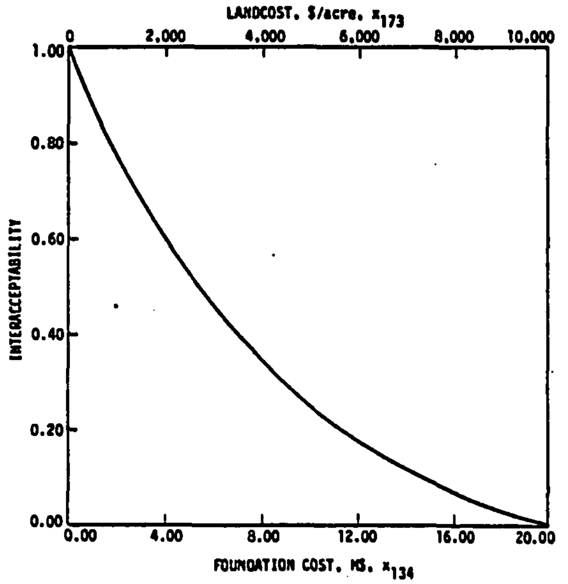


Figure 19. Foundation and land costs

$$\begin{bmatrix} u_{21} \\ u_{22} \\ u_{23} \\ u_{24} \\ u_{25} \\ u_{26} \end{bmatrix} = \begin{bmatrix} -0.016 \\ -0.2604 \\ -0.1782 \\ -0.067 \\ 1.178 \\ 1.051 \end{bmatrix} + \begin{bmatrix} 1.016 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.2604 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.178 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.067 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.178 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.051 \end{bmatrix}$$

$$\times \exp \begin{bmatrix} -0.0418 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.0157 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0189 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.0277 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0189 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0303 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \\ x_{23} \\ x_{24} \\ x_{25} \\ x_{26} \end{bmatrix}, \quad (29)$$

$$\begin{bmatrix} u_{31} \\ u_{32} \end{bmatrix} = \begin{bmatrix} 1.038 \\ 1.086 \end{bmatrix} - \begin{bmatrix} 0.038 & 0 \\ 0 & 0.086 \end{bmatrix} \exp \begin{bmatrix} 0.0332 & 0 \\ 0 & 0.1013 \end{bmatrix} \begin{bmatrix} x_{31} \\ x_{32} \end{bmatrix}, \quad (30)$$

$$\begin{bmatrix} u_{41} \\ u_{42} \\ u_{43} \\ u_{44} \\ u_{45} \\ u_{46} \end{bmatrix} = \begin{bmatrix} 0.225 \\ 0.133 \\ 1.408 \\ 0.133 \\ 0.408 \\ 0.408 \end{bmatrix} + \begin{bmatrix} 1.225 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.133 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1.408 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.133 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.408 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.408 \end{bmatrix}$$

$$\times \exp - \begin{bmatrix} 0.016 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0214 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0214 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0214 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.0124 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.0124 \end{bmatrix} \begin{bmatrix} x_{41} \\ x_{42} \\ x_{43} \\ x_{44} \\ x_{45} \\ x_{46} \end{bmatrix}, \quad (31)$$

$$\begin{bmatrix} u_{51} \\ u_{52} \\ u_{53} \\ u_{54} \\ u_{55} \\ u_{56} \end{bmatrix} = \begin{bmatrix} -0.225 \\ -0.225 \\ 1.067 \\ 1.125 \\ 1.408 \\ -0.408 \end{bmatrix} + \begin{bmatrix} 0.225 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.225 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1.067 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.125 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1.408 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0408 \end{bmatrix}$$

$$\times \exp \begin{bmatrix} 0.0169 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0169 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.0277 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0219 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.0124 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0124 \end{bmatrix} \begin{bmatrix} x_{51} \\ x_{52} \\ x_{53} \\ x_{54} \\ x_{55} \\ x_{56} \end{bmatrix}, \quad (32)$$

$$\begin{bmatrix} u_{61} \\ u_{62} \\ u_{63} \\ u_{64} \\ u_{65} \end{bmatrix} = - \begin{bmatrix} -1.125 \\ 0.225 \\ 0.225 \\ 0.225 \\ 0.8 \end{bmatrix} + \begin{bmatrix} -1.125 & 0 & 0 & 0 & 0 \\ 0 & 1.225 & 0 & 0 & 0 \\ 0 & 0 & 0.225 & 0 & 0 \\ 0 & 0 & 0 & 0.225 & 0 \\ 0 & 0 & 0 & 0 & 0.225 \end{bmatrix}$$

$$\times \exp \begin{bmatrix} -0.0219 & 0 & 0 & 0 & 0 \\ 0 & -0.0339 & 0 & 0 & 0 \\ 0 & 0 & 0.0169 & 0 & 0 \\ 0 & 0 & 0 & 0.0169 & 0 \\ 0 & 0 & 0 & 0 & 0.0081 \end{bmatrix} \begin{bmatrix} x_{61} \\ x_{62} \\ x_{63} \\ x_{64} \\ x_{65} \end{bmatrix} \quad (33)$$

$$\begin{bmatrix} u_{71} \\ u_{72} \\ u_{73} \\ u_{74} \\ u_{75} \\ u_{76} \end{bmatrix} = - \begin{bmatrix} 0.225 \\ 0.8 \\ 0.225 \\ -1.408 \\ 0.225 \\ 0.125 \end{bmatrix} + \begin{bmatrix} 1.225 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.255 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.408 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.225 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.125 \end{bmatrix}$$

$$\times \exp - \begin{bmatrix} 0.0169 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0081 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0169 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.0124 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0169 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0219 \end{bmatrix} \begin{bmatrix} x_{71} \\ x_{72} \\ x_{73} \\ x_{74} \\ x_{75} \\ x_{76} \end{bmatrix}, \quad (34)$$

$$\begin{bmatrix} u_{81} \\ u_{82} \\ u_{83} \end{bmatrix} = \begin{bmatrix} 1.563 \\ 1.563 \\ 1.253 \end{bmatrix} - \begin{bmatrix} 1.563 & 0 & 0 \\ 0 & 0.563 & 0 \\ 0 & 0 & 0.253 \end{bmatrix} \exp \begin{bmatrix} -0.0102 & 0 & 0 \\ 0 & 0.0102 & 0 \\ 0 & 0 & 0.016 \end{bmatrix} \begin{bmatrix} x_{81} \\ x_{82} \\ x_{83} \end{bmatrix}, \quad (35)$$

$$\begin{bmatrix} u_{91} \\ u_{92} \\ u_{93} \\ u_{94} \\ u_{95} \end{bmatrix} = - \begin{bmatrix} 0.361 \\ 0.527 \\ 0.259 \\ 0.056 \\ 0.125 \end{bmatrix} + \begin{bmatrix} 1.361 & 0 & 0 & 0 & 0 \\ 0 & 1.527 & 0 & 0 & 0 \\ 0 & 0 & 1.259 & 0 & 0 \\ 0 & 0 & 0 & 1.056 & 0 \\ 0 & 0 & 0 & 0 & 1.125 \end{bmatrix}$$

$$\times \exp - \begin{bmatrix} 0.000133 & 0 & 0 & 0 & 0 \\ 0 & 0.000106 & 0 & 0 & 0 \\ 0 & 0 & 0.1977 & 0 & 0 \\ 0 & 0 & 0 & 0.01837 & 0 \\ 0 & 0 & 0 & 0 & 0.0439 \end{bmatrix} \begin{bmatrix} x_{91} \\ x_{92} \\ x_{93} \\ x_{94} \\ x_{95} \end{bmatrix}, \quad (36)$$

$$\begin{bmatrix} u_{101} \\ u_{102} \end{bmatrix} = - \begin{bmatrix} 0.5625 \\ 0.253 \end{bmatrix} + \begin{bmatrix} 0.5625 & 0 \\ 0 & 1.253 \end{bmatrix} \exp \begin{bmatrix} 0.2554 \\ -0.016 \end{bmatrix} \begin{bmatrix} x_{101} \\ x_{102} \end{bmatrix}, \quad (37)$$

$$\begin{bmatrix} u_{111} \\ u_{112} \\ u_{113} \\ u_{114} \end{bmatrix} = \begin{bmatrix} 1.253 \\ 1.125 \\ -0.125 \\ -0.125 \end{bmatrix} + \begin{bmatrix} -0.253 & 0 & 0 & 0 \\ 0 & -0.125 & 0 & 0 \\ 0 & 0 & 1.125 & 0 \\ 0 & 0 & 0 & 1.125 \end{bmatrix}$$

$$\times \exp \begin{bmatrix} 0.016 & 0 & 0 & 0 \\ 0 & 0.0219 & 0 & 0 \\ 0 & 0 & -0.0219 & 0 \\ 0 & 0 & 0 & 0.0219 \end{bmatrix} \begin{bmatrix} x_{111} \\ x_{112} \\ x_{113} \\ x_{114} \end{bmatrix} \tag{38}$$

$$\begin{bmatrix} u_{121} \\ u_{122} \\ u_{123} \end{bmatrix} = -\begin{bmatrix} 0.0864 \\ 0.9339 \\ 0.3612 \end{bmatrix} + \begin{bmatrix} 0.0864 & 0 & 0 \\ 0 & 0.9339 & 0 \\ 0 & 0 & 2.3621 \end{bmatrix} \exp \begin{bmatrix} 0.2531 & 0 & 0 \\ 0 & 0.0073 & 0 \\ 0 & 0 & 0.0133 \end{bmatrix} \begin{bmatrix} x_{121} \\ x_{122} \\ x_{123} \end{bmatrix}, \tag{39}$$

$$\begin{bmatrix} u_{131} \\ u_{132} \\ u_{133} \\ u_{134} \end{bmatrix} = -\begin{bmatrix} 0.1782 \\ -1.038 \\ 0.178 \\ 0.125 \end{bmatrix} + \begin{bmatrix} 1.1782 & 0 & 0 & 0 \\ 0 & -0.038 & 0 & 0 \\ 0 & 0 & 1.178 & 0 \\ 0 & 0 & 0 & 1.125 \end{bmatrix}$$

$$\times \exp - \begin{bmatrix} 0.0189 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ -0.0332 & 0 & 0 \\ 0 & 0.0189 & 0 \\ 0 & 0 & 0.0219 \end{bmatrix} \begin{bmatrix} x_{131} \\ x_{132} \\ x_{133} \\ x_{134} \end{bmatrix} \tag{40}$$

$$\begin{bmatrix} u_{141} \\ u_{142} \\ u_{143} \\ u_{144} \\ u_{145} \\ u_{146} \\ u_{147} \end{bmatrix} = -\begin{bmatrix} 0.125 \\ 0.253 \\ 0.076 \\ 0.125 \\ 0.125 \\ 0.141 \\ -1.178 \end{bmatrix} + \begin{bmatrix} 1.125 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.253 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.076 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.125 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.125 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.141 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.178 \end{bmatrix}$$

$$\times \exp - \begin{bmatrix} 0.0219 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.016 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0265 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0219 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0219 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0209 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.0189 \end{bmatrix} \begin{bmatrix} x_{141} \\ x_{142} \\ x_{143} \\ x_{144} \\ x_{145} \\ x_{146} \\ x_{147} \end{bmatrix} \tag{41}$$

$$\begin{bmatrix} u_{151} \\ u_{152} \\ u_{153} \end{bmatrix} = -\begin{bmatrix} 0.2529 \\ 0.0864 \\ 0.125 \end{bmatrix} + \begin{bmatrix} 0.2529 & 0 & 0 \\ 0 & 0.0864 & 0 \\ 0 & 0 & 1.125 \end{bmatrix} \exp \begin{bmatrix} 0.08 & 0 & 0 \\ 0 & 0.0253 & 0 \\ 0 & 0 & -0.0119 \end{bmatrix} \begin{bmatrix} x_{151} \\ x_{152} \\ x_{153} \end{bmatrix}, \quad (42)$$

$$\begin{bmatrix} u_{161} \\ u_{162} \\ u_{163} \\ u_{164} \end{bmatrix} = -\begin{bmatrix} 0.125 \\ 0.3793 \\ 0.5625 \\ 0.3018 \end{bmatrix} + \begin{bmatrix} 0.125 & 0 & 0 & 0 \\ 0 & 0.3793 & 0 & 0 \\ 0 & 0 & 0.5625 & 0 \\ 0 & 0 & 0 & 0.3018 \end{bmatrix}$$

$$\times \exp \begin{bmatrix} 0.02197 & 0 & 0 & 0 \\ 0 & 0.0129 & 0 & 0 \\ 0 & 0 & 0.0102 & 0 \\ 0 & 0 & 0 & 0.0146 \end{bmatrix} \begin{bmatrix} x_{161} \\ x_{162} \\ x_{163} \\ x_{164} \end{bmatrix} \quad (43)$$

$$\begin{bmatrix} u_{171} \\ u_{172} \\ u_{173} \\ u_{174} \end{bmatrix} = -\begin{bmatrix} 0.178 \\ 0.125 \\ 0.178 \\ 0.125 \end{bmatrix} + \begin{bmatrix} 1.178 & 0 & 0 & 0 \\ 0 & 1.125 & 0 & 0 \\ 0 & 0 & 1.178 & 0 \\ 0 & 0 & 0 & 1.125 \end{bmatrix}$$

$$\times \exp -\begin{bmatrix} 0.1349 & 0 & 0 & 0 \\ 0 & 0.000219 & 0 & 0 \\ 0 & 0 & 0.2361 & 0 \\ 0 & 0 & 0 & 0.1099 \end{bmatrix} \begin{bmatrix} x_{171} \\ x_{172} \\ x_{173} \\ x_{174} \end{bmatrix} \quad (44)$$

APPLICATION

To illustrate for the use of the above formulation, an actual site selection undertaking is considered. The entire process of site evaluation and the selection of the most acceptable site is done using the methodology presented here. Then, the results are compared with the outcome of the decision made by the utility company which undertook that specific selection task using a different approach.

Problem Description

To cater the need for electrical power in eastern Kansas in 1981, Kansas Gas and Electric Company and Kansas City Power and Light Company proposed Wolf Creek Generating Station, Unit no. 1, an 1150 MWe nuclear power plant [23]. The U.S. Nuclear Regulatory Commission (NRC) granted the construction permit on October 1975. After preliminary screening, three alternative plant locations were selected for further evaluation, all of which depended on the John Redmond Reservoir, Figure 20, as a source of cooling water. The three sites would utilize a closed-cycle cooling pond scheme.

Site 1 is located approximately 3 miles north of the town of Ottumwa on the northeast shore of the reservoir. The plant site is situated on high ground between two drainages, one formed by Hickory Creek which flows into the reservoir and the other by a small unnamed stream which also flows into the reservoir. Connecting cooling ponds would be developed in each of these drainages for waste-heat dissipation.

Site 2 and the associated cooling lake will be located in Coffey

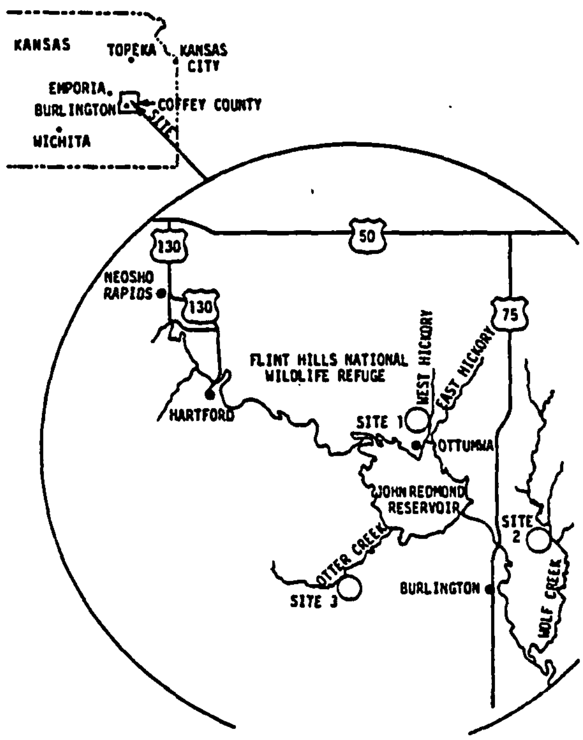


Figure 20. Site selection for nuclear power plant in John Redmond Reservoir region among three alternatives

County, which is located in eastern Kansas approximately 75 miles southwest of Kansas City, 53 miles south of Topeka and 90 miles east-northeast of Wichita, Kansas. This is the Wolf Creek site which was finally selected by the utility company.

Site 3 is situated south of the reservoir about 3.5 miles north of the Gridley and 7 miles west of Burlington. The cooling pond developed for this site would lie in the drainage basin of North Big Creek with an unnamed dam across the creek located about 11 miles upstream from its confluence with the Neosho River. The reactor site would be located near the upstream end of the cooling pond.

Land use, make-up water use, geological, ecological, meteorological and other environmental conditions at the three sites are similar. Transmission considerations are also similar, although the transmission line distance requirements vary by site. The most desirable site among the above mentioned sites has to be selected for the nuclear power plant construction.

The site characteristics and the intra-characteristics described earlier and contained in Table 1 are considered for each alternative site. The intra-characteristics levels or outcomes are assessed for each site from ref. [24]. These values have been placed in Table 3. It must be emphasized that the absolute values of the intra-characteristics levels are not important; rather, they should quantitatively represent the relative importance of the intra-characteristics. The inter-acceptabilities can be obtained from equations 28 through 44 for each intra-characteristic and then equation 12 may be used to calculate the

Table 3. Intra-characteristics levels for John Redmond sites 1, 2, and 3 for Wolf Creek Nuclear Power Station

Intra-characteristic (x_{ij})	Site 1	Site 2	Site 3
x_{11}	19.5	13.25	14.9
x_{12}	1	1	1
x_{13}	0.25	0.25	0.25
x_{14}	0	0	50
x_{15}	2.2	2.2	2.2
x_{21}	20	20	20
x_{22}	20	20	20
x_{23}	20	20	20
x_{24}	20	20	20
x_{25}	20	20	20
x_{26}	20	20	20
x_{31}	20	20	20
x_{32}	5	5	5
x_{41}	80	0	50
x_{42}	20	20	20
x_{43}	24	91	100
x_{44}	80	0	50
x_{45}	80	0	50
x_{46}	50	100	70
x_{51}	100	100	100
x_{52}	100	100	100
x_{53}	100	100	100
x_{54}	20	20	20
x_{55}	80	80	80
x_{56}	80	80	80
x_{61}	85	90	100
x_{62}	1.2	0.5	1.0
x_{63}	80	100	60
x_{64}	80	60	100
x_{65}	0	0	0
x_{71}	0	0	0
x_{72}	50	0	0
x_{74}	50	0	40
x_{75}	0	0	0
x_{76}	0	0	0
x_{81}	100	100	100
x_{82}	0	0	0
x_{83}	0	0	100
x_{91}	10	10	10
x_{92}	0	0	0
x_{93}	12	12	12
x_{94}	11	11	11

Table 3. Continued

Intra-characteristic (x_{ij})	Site 1	Site 2	Site 3
x_{95}	3	3	3.5
x_{101}	1	2	1
x_{102}	0	0	0
x_{111}	10	10	10
x_{112}	10	10	10
x_{113}	10	10	10
x_{114}	10	10	10
x_{121}	5	7.9	5
x_{122}	100	100	100
x_{123}	100	100	100
x_{131}	0	0	0
x_{132}	0	0	0
x_{133}	0	0	0
x_{134}	10	10	10
x_{141}	60	60	60
x_{142}	80	80	80
x_{143}	0	0	3
x_{144}	0	0	0
x_{145}	0	0	0
x_{146}	0	0	0
x_{147}	80	80	80
x_{151}	9	9	9
x_{152}	80	80	80
x_{153}	10	10	10
x_{161}	60	60	60
x_{162}	100	100	100
x_{163}	70	70	70
x_{164}	100	100	100
x_{171}	1	2.5	3
x_{172}	2600	2600	2600
x_{173}	3.4	2.8	2.9
x_{174}	32.1	23	24.4

acceptabilities of the 17 characteristics for each site. The results are given in Table 4. Finally, either equation 6 or 17 can be used to evaluate the supra-acceptability function for each site. A specially developed computer program is used to provide all necessary

Table 4. Acceptabilities of characteristics of the proposed three sites of Wolf Creek Nuclear Power Station

Characteristic (Y_i)	Site 1	Site 2	Site 3
Y_1	0.9747	0.9747	0.9488
Y_2	0.8504	0.8504	0.8504
Y_3	0.9654	0.9654	0.9654
Y_4	0.7918	1.00	0.9407
Y_5	0.9961	0.9961	0.9961
Y_6	0.9764	0.9854	0.9886
Y_7	0.9347	0.9998	0.9836
Y_8	0.999	0.999	0.8876
Y_9	0.9380	0.9380	0.9359
Y_{10}	0.6721	0.7548	0.6721
Y_{11}	0.9429	0.9429	0.9429
Y_{12}	0.8935	0.9320	0.8935
Y_{13}	0.9849	0.9849	0.9849
Y_{14}	0.9919	0.9919	0.9743
Y_{15}	0.8129	0.8129	0.8129
Y_{16}	0.9109	0.9109	0.9109
Y_{17}	0.7247	0.6777	0.6499

iterations and to compute the supra-acceptability in each case using the relationship

$$\mu_k = \sum_{i=1}^{I=17} \frac{p_i}{k_i} \{ \sum_{j=1}^{J_i} [1 + k_i p_{ij} (\alpha_{ij} + \beta_{ij} \exp \gamma_{ij} x_{ij}) - 1] \}, \tag{45}$$

which is obtained from equations 17 and 27. The results are given in Table 5. In the present case, the intra-characteristics having the same value for all sites may be dropped from the analysis; however, these are given here to demonstrate the technique.

According to equation 18, site 2 is the most acceptable taking in

Table 5. Supra-acceptability of the proposed 3 John Redmond sites

Site number (k)	Supra-acceptability (μ_k)
1	0.9014
2	0.9197
3	0.8970

consideration all the factors affecting the selection. This agrees with the choice made by the concerned utility company mentioned earlier.

CONCLUSION

The methodology developed gives a rational and self-consistent approach towards selection of the most desirable site for nuclear power plants, where more than one feasible site is available. A siting characteristic such as public acceptance, which includes attitude of environmentalists, regulatory bodies media and occupational groups, in addition to other engineering and economic characteristics, has been added to accommodate for their growing influence in site selection. A computer code is developed which takes into account the 17 siting characteristics and 75 intra-characteristics and selects the most desirable site for any number of alternates available. A two-round Delphi questionnaire is developed which reconciles the differences among various experts in arriving at the indifference probabilities. The technique is applied to a practical case where a group of utility companies had to select a site for a 1150 MWe nuclear power plant in Coffey County, Kansas, among three alternatives. Although the method used here for the site selection is entirely different, the results indicate that the site selected earlier, namely the Wolf Creek Generating Station, is the best one.

REFERENCES

1. G. G. Eichholz, Environmental Aspects of Nuclear Power (Ann Arbor Sci. Publ., Ann Arbor, 1977).

2. F. R. Farmer, Symp. on the Containment and Siting of Nuclear Power Plants (IAEA, Vienna, April 1967), p. 303.

3. M. Meleis and R. C. Erdmann, Nucl. Safety 13, 22 (1972).

4. J. G. Gros, R. Avenhaus, J. Linnerooth, P. D. Palmer and H. J. Otway, Symp. on the Siting of Nuclear Facilities (IAEA, Vienna, December 1974), p. 387.

5. R. L. Keeney and K. Nair, Proc. IEEE 63, 494 (1975).

6. E. M. A. Hassan, Symp. on the Siting of Nuclear Facilities (IAEA, Vienna, December 1974), p. 227.

7. L. P. Beer, Trans. ANS 19 (Suppl. 1) 12 (1974).

8. D. G. Jopling, Power Eng. 78, 56 (1974).

9. L. J. West and K. Kyar, Trans. ANS 19 (Suppl. 1), 8 (1974).

10. M. C. Cordaro and W. T. Malloy, Nucl. Technol. 23, 233 (1974).

11. J. A. Fischer and R. Ahmed, Trans. ANS 19 (Suppl. 1), 7 (1974).

12. "General Site Suitability Criteria for Nuclear Power Stations," Regulatory Guide 4.7 (US Atomic Energy Comm., Washington, DC, 1974).

13. J. J. DiNunno, Symp. on Siting of Nuclear Facilities (IAEA, Vienna, December 1974), p. 3.

14. A to G is converted into a scale from 0 to 100 for convenience, where A is 0 and G is 100.

15. J. G. Gros, A Paretian Environmental Approach to Power Plant Siting in New England (Harvard University, Cambridge, 1974).

16. R. L. Keeney, Tech. Rep. No. 10 (Operations Res. Center, MIT, Cambridge, Mass., 1972).

17. P. C. Fishburn, Operations Res. 13, 28 (1965).

18. P. C. Fishburn, Utility Theory for Decision Making (Wiley, New York 1970).

19. R. L. Keeney, Management Sci. 18, 22 (1972).

20. R. F. Meyer, Proc. 4th IFORS Conf., Venice, Italy, 1969.

21. R. A. Pollak, Econometrica 35, 485 (1967).

22. H. Raiffa, RM-5868-DOT/RC (Rand Corp., Santa Monica, CA, 1969).

23. Wolf Creek Generating Station Environmental Report, section 2.1 and 9.2.1.4.5.

24. Wolf Creek Generating Station Environmental Report, Table 9.3-1.

APPENDIX

First-round Questionnaire

<u>Please read these instructions carefully</u>.

You have been selected to participate in a study to develop a decision analysis technique for selection of site for nuclear power plants. This questionnaire is the first of a series of two questionnaires designed to gather information that only you and your fellow workers can provide. The answer you give to the questions could become the basis for possible improvement in the current technique of nuclear siting.

1. <u>Please read the instructions for each part carefully</u>, determine that you know exactly what is being asked, and give thorough considerations to <u>all</u> aspects of the question before answering.

2. Please do not discuss your answers with any other person engaged in the study until after the second questionnaire is completed. This is a critical requirement. Any violation could seriously distort the results.

3. You are allowed to use any records, files, or other source of information available to aid you in answering the questions; in fact, you are encouraged to do so. <u>The only exception is that you are not to discuss the questions or your answers with any other person taking the survey</u>.

Although your name has been identified with this particular answer sheet, you should understand that this is required only to compare answers between yourself and other participants as a group in the following questionnaires. No one but yourself and the person conducting the survey will know how you answered any particular question. You are asked, then, to give honest answers, not the answers that you think someone else would like to see. This study is an attempt to gather facts, not to falsely justify or condemn any particular policy or procedure. Please provide answers that you sincerely feel are as accurate as you can make them.

PART I

(1) Instructions for providing answers:

In the questions associated with this study, you will be asked to estimate a value from 0 to 10 for each factor considered for site selection. A 0 value means the least important factor to be considered in siting criteria while 10 should be assigned to a factor of highest importance. Any value between 0 to 10 can be picked for a particular factor according to its importance.

		Least important 0	1	2	3	4	5	6	7	8	9	Most important 10
(a)	Cooling water	0	1	2	3	4	5	6	7	8	9	10
(b)	Soil properties	0	1	2	3	4	5	6	7	8	9	10
(c)	Construction difficulties	0	1	2	3	4	5	6	7	8	9	10
(d)	Services availability	0	1	2	3	4	5	6	7	8	9	10
(e)	Domestic water	0	1	2	3	4	5	6	7	8	9	10
(f)	Accessibility	0	1	2	3	4	5	6	7	8	9	10
(g)	Present land use	0	1	2	3	4	5	6	7	8	9	10
(h)	Ecology	0	1	2	3	4	5	6	7	8	9	10
(i)	Demography	0	1	2	3	4	5	6	7	8	9	10
(j)	Topography	0	1	2	3	4	5	6	7	8	9	10

	Least important 0	1	2	3	4	5	6	7	8	9	Most important 10
(k) Geology	0	1	2	3	4	5	6	7	8	9	10
(l) Aircraft impact	0	1	2	3	4	5	6	7	8	9	10
(m) Man-made hazards (ship collision, explosions, flammable gases, etc.)	0	1	2	3	4	5	6	7	8	9	10
(n) Natural phenomena (tornadoes, flood, etc.)	0	1	2	3	4	5	6	7	8	9	10
(o) Meteorology	0	1	2	3	4	5	6	7	8	9	10
(p) Public acceptance	0	1	2	3	4	5	6	7	8	9	10
(q) Economical impact	0	1	2	3	4	5	6	7	8	9	10

PART II

We have further divided each factor considered for nuclear power plant site selection in sub-factors. Consider each factor to be independent of others. Give a number from 0 to 10 for each group of sub-factors according to its importance.

Factor	Sub-factors	Least important	A number from 0 to 10	Most important
(a) Cooling water	- Sources	0	1 2 3 4 5 6 7 8 9	10
	- Availability	0	1 2 3 4 5 6 7 8 9	10
	- Distance	0	1 2 3 4 5 6 7 8 9	10
	- Characteristics	0	1 2 3 4 5 6 7 8 9	10
	- Incremental water temperature °C at peak ambient water temp. period of year	0	1 2 3 4 5 6 7 8 9	10
(b) Soil	- Classification	0	1 2 3 4 5 6 7 8 9	10
	- Characteristics	0	1 2 3 4 5 6 7 8 9	10
	- Properties	0	1 2 3 4 5 6 7 8 9	10
	- Bearing capacity	0	1 2 3 4 5 6 7 8 9	10

Factor	Sub-factors	Least impor-tant	A number from 0 to 10	Most impor-tant
	- Fixation of rad-waste	0	1 2 3 4 5 6 7 8 9	10
	- Groundwater	0	1 2 3 4 5 6 7 8 9	10
(c) Construction	- Land properties	0	1 2 3 4 5 6 7 8 9	10
	- Labor	0	1 2 3 4 5 6 7 8 9	10
(d) Services	- Electricity	0	1 2 3 4 5 6 7 8 9	10
	- Gas	0	1 2 3 4 5 6 7 8 9	10
	- Drainage and sewerage	0	1 2 3 4 5 6 7 8 9	10
	- Location of services	0	1 2 3 4 5 6 7 8 9	10
	- Capacities	0	1 2 3 4 5 6 7 8 9	10
	- Distribution lines	0	1 2 3 4 5 6 7 8 9	10
(e) Domestic water	- Availability	0	1 2 3 4 5 6 7 8 9	10
	- Characteristics	0	1 2 3 4 5 6 7 8 9	10
	- Size, depth and pressure of mains	0	1 2 3 4 5 6 7 8 9	10
	- Consumption	0	1 2 3 4 5 6 7 8 9	10
	- Distribution systems	0	1 2 3 4 5 6 7 8 9	10
	- Sources and location	0	1 2 3 4 5 6 7 8 9	10
(f) Accessibility	- Land characteristics along route of trans-mission line network	0	1 2 3 4 5 6 7 8 9	10
	- Distance to transmission line interties and sub-station	0	1 2 3 4 5 6 7 8 9	10
	- Highway	0	1 2 3 4 5 6 7 8 9	10
	- Railroad	0	1 2 3 4 5 6 7 8 9	10
	- Waterways	0	1 2 3 4 5 6 7 8 9	10
(g) Land use	- Agriculture	0	1 2 3 4 5 6 7 8 9	10
	- Special land use (parks, wildlife refuge, etc.)	0	1 2 3 4 5 6 7 8 9	10
	- Residential character	0	1 2 3 4 5 6 7 8 9	10
	- Public facilities	0	1 2 3 4 5 6 7 8 9	10
	- Sensitive industries (oil	0	1 2 3 4 5 6 7 8 9	10
	and gas pipelines)	0	1 2 3 4 5 6 7 8 9	10
	- Weapons testing	0	1 2 3 4 5 6 7 8 9	10
	- Land values	0	1 2 3 4 5 6 7 8 9	10
(h) Ecology	- River classification (irrigation, fishing, recreational, wildlife, etc.)	0	1 2 3 4 5 6 7 8 9	10

Factor	Sub-factors	Least impor- tant	A number from 0 to 10									Most impor- tant
	- Endangered species	0	1	2	3	4	5	6	7	8	9	10
	- Biological conditions	0	1	2	3	4	5	6	7	8	9	10
(i) Demography	- Existing population within 1 mile	0	1	2	3	4	5	6	7	8	9	10
	- Population within 1 mile, 2020	0	1	2	3	4	5	6	7	8	9	10
	- Existing population within 20 miles	0	1	2	3	4	5	6	7	8	9	10
	- Population within 20 miles, 2020	0	1	2	3	4	5	6	7	8	9	10
	- Proximity to nearest population center	0	1	2	3	4	5	6	7	8	9	10
(j) Topography	- Number of potential sites per region	0	1	2	3	4	5	6	7	8	9	10
	- Terrain for power plant	0	1	2	3	4	5	6	7	8	9	10
(k) Geology	- Mine hazard	0	1	2	3	4	5	6	7	8	9	10
	- Oil hazard	0	1	2	3	4	5	6	7	8	9	10
	- Seismology	0	1	2	3	4	5	6	7	8	9	10
	- Tectonic	0	1	2	3	4	5	6	7	8	9	10
(l) Aircraft impact	- Proximity to airport	0	1	2	3	4	5	6	7	8	9	10
	- Bombing ranges	0	1	2	3	4	5	6	7	8	9	10
	- Low-level training routes	0	1	2	3	4	5	6	7	8	9	10
(m) Man-made hazards	- Ship collision and explosion	0	1	2	3	4	5	6	7	8	9	10
	- Flammable gases and vapor clouds	0	1	2	3	4	5	6	7	8	9	10
	- Toxic chemicals	0	1	2	3	4	5	6	7	8	9	10
	- Fire	0	1	2	3	4	5	6	7	8	9	10
(n) Natural phenomena	- Tornadoes	0	1	2	3	4	5	6	7	8	9	10
	- High water from floods	0	1	2	3	4	5	6	7	8	9	10
	- High water from dam failure	0	1	2	3	4	5	6	7	8	9	10
	- High water from hurricanes	0	1	2	3	4	5	6	7	8	9	10
	- High water from tsunamis	0	1	2	3	4	5	6	7	8	9	10
	- High water from seiches	0	1	2	3	4	5	6	7	8	9	10
	- Low water	0	1	2	3	4	5	6	7	8	9	10
(o) Meteorology	- Sustained wind velocity	0	1	2	3	4	5	6	7	8	9	10

Factor	Sub-factors	Least impor-tant	A number from 0 to 10	Most impor-tant
	- Diffusion	0	1 2 3 4 5 6 7 8 9	10
	- Inversion temperature	0	1 2 3 4 5 6 7 8 9	10
(p) Public acceptance	- Occupational group (unions, merchants, and corporations)	0	1 2 3 4 5 6 7 8 9	10
	- Citizen group (environmentalists, civic associations, etc.)	0	1 2 3 4 5 6 7 8 9	10
	- Influential individuals (ministers, publishers and editors)	0	1 2 3 4 5 6 7 8 9	10
	- State laws	0	1 2 3 4 5 6 7 8 9	10
(q) Economic impact	- The cost of cooling water supply related to the water intakes, pipelines, storage or water supply reservoirs, and pumping facility required to deliver a reliable supply of cooling water to a plant at each site identified	0	1 2 3 4 5 6 7 8 9	10
	- Land acquisition costs	0	1 2 3 4 5 6 7 8 9	10
	- Land preparation costs, such as clearing, site grading, flood protection, on-site access improvements, and relocated road and bridge construction costs	0	1 2 3 4 5 6 7 8 9	10
	- Foundation cost estimates, such as the cost of de-watering, special foundation treatment, excavating, foundation base mats, pile foundations, if required, and back-filling	0	1 2 3 4 5 6 7 8 9	10

Second-round Questionnaire

This is the second of a series of two questionnaires designed to gather information concerning a study to develop a decision analysis technique for selection of site for nuclear power plants. The same questions that appeared in the first questionnaire are repeated here. During the first round, you submitted the relative importance of the factors being considered for selection among various sites for a nuclear power plant. This questionnaire gives you the opportunity to revise any of these estimates <u>if</u> you feel they can be improved.

For each site selection criterion, the following information is presented: Your <u>first-round estimate</u>, the <u>average</u> of the estimates made by <u>all</u> participants, and the <u>central range</u> of the estimates made by <u>all</u> participants. The central range is chosen so that 25% of the estimates lie below the lower value and 25% of the estimates lie above the upper value. Therefore, the central range itself contains the middle 50% of the estimates made for that particular value.

You are asked to reconsider each of your previous answers, <u>possibly</u> revise them, and write your new answer in the space provided. For each estimate, if your <u>new</u> estimate lies <u>outside</u> the central range, you are asked to state <u>briefly but clearly</u>, in the space provided, the <u>major reason or reasons</u> why you feel the estimate should be lower (or higher) than those within the central range.

It should be pointed out that your "new" estimates <u>do not</u> have to be different from your first-round estimates. The purpose of this questionnaire is to give you a chance to reconsider your responses in the light of new information.

PART II

Factor	Sub-factor	Your first estimate	Group estimate: Average	Group estimate: Central range	Your new estimate	Major reason why you feel the estimate should be lower (or higher) than those within the central range
(a) Cooling water	-Sources -Availability -Distance -Characteristics -Incremental water temperature °C at peak ambient temp. period of year					
(b) Soil	-etc.[1]					
(c) Construction	-etc.					
(d) Services	-etc.					
(e) Domestic water	-etc.					
(f) Accessibility	-etc.					
(g) Land use	-etc.					
(h) Ecology	-etc.					
(i) Demography	-etc.					
(j) Topography	-etc.					
(k) Geology	-etc.					
(l) Aircraft impact	-etc.					
(m) Man-made hazards	-etc.					

[1]The reader is referred to the sub-factors given in Part II of the first-round questionnaire.

Factor	Sub-factor	Your first esti-mate	Group estimate: Aver-age	Group estimate: Central range	Your new esti-mate	Major reason why you feel the estimate should be lower (or higher) than those within the central range
(n) Natural phenomena	-etc.					
(o) Meteorology	-etc.					
(p) Public acceptance	-etc.					
(q) Economic impact	-etc.					

SECTION II. RISK ASSESSMENT OF ALTERNATIVE PROLIFERATION ROUTES

Risk assessment of alternative proliferation routes

Shahid Ahmed

A. A. Husseiny

From Nuclear Engineering Department, Iowa State University, Ames, Iowa 50011 and Technology International, Inc., 125 South Third Street, Sherman Place, Ames, Iowa 50010.

Published in Nuclear Technology 56(1982):507-515.

ABSTRACT

Multi-Attribute Decision Theory is applied to rank 11 alternative routes to nuclear proliferation in order of difficulty in acquiring nuclear weapons by nonnuclear countries. The method is based on reducing the various variables affecting the decision to a single function providing a measure for the proliferation route. The results indicate that the most difficult route to obtain atomic weapons is through nuclear power reactors, specifically the liquid-metal fast breeder reactor, heavy water Canada deuterium uranium reactor, and light water reactors such as boiling water and pressurized water reactors. The easiest routes are supercritical centrifuge isotope separation, laser isotope separation, and research reactor. However, the nonnuclear routes, also considered here for comparative purposes, that result in substantial damage to life and property are easier than any nuclear route.

INTRODUCTION

Worldwide depletion of existing energy resources and the vast industrial growth in developing countries have led many nations to consider nuclear electric power generation as one of the most viable energy alternatives. There is, however, a growing fear that fissile material from the power cycle may be diverted to the production of weapons. Several research projects have been directed to the assessment of the consequences of the spread of nuclear power plants and to the development of new fuel cycles which do not allow for undetected misuse of nuclear materials. One of the major efforts is the Nonproliferation Alternative Systems Assessment Program (NASAP). Concerns over nuclear material diversion have resulted in deferral of transfer of nuclear technology to nonnuclear countries.

The objective of this paper is to evaluate, in quantitative terms, various options of acquiring nuclear weapons material and their relative difficulty compared to the use of other destructive materials. A deterministic approach based on combinatorial multivariate preferences is used in the analysis that encompasses consideration of the various facets of the requirements and capabilities in obtaining nuclear weapons. The results would be helpful in planning for proliferation deterrence and for devision appropriate safeguard programs without depriving development countries of the benefits of peaceful uses of nuclear energy. A factual assessment of the requirements, methods, and difficulties in acquiring nuclear weapons is essential in assessing the impact of nuclear electric power programs in nonnuclear countries on nuclear proliferation. It is well-known that complex processes are required to produce weapons-grade materials, such as plutonium and highly enriched uranium. Also, sophisticated deployment capabilities are needed, which are nonexistent in nonnuclear countries. Nevertheless, the interest here is in understanding of the various routes of proliferation and the associated difficulties of acquiring weapons-grade materials.

A number of potential nuclear materials proliferation routes have been suggested [1]. For comparison, three nonnuclear destructive material proliferation options are included [1]. Those three options do not lead to nuclear explosions but are capable of causing significant damage to life and property using simpler technologies to those required for inducing similar damage by nuclear weapons. The merits of acquiring nuclear explosive capability are judged with respect to five factors:

1. resources
2. difficulty of nuclear route
3. cost and schedule
4. risks
5. weapons capability.

The resources needed to make a nuclear weapon are split into four acquisition factors--technical sophistication, facility requirements, instrumentation capability, and personnel requirements.

The multidimensionality of the attributes of each proliferation route complicates the process of relative ranking among various routes.

However, a decision based on intuition or qualitative assessment without due consideration of all the facets of the problem has the potential of being incomplete or inconsistent. Consequently, an approach based on the multiattribute utility theory is selected as a formal analytical methodology for selection of optimal strategies based on quantitative measures. Each measure or attribute is represented by a utility or objective function. The multiplicity of the objective functions is reduced to a univariate objective function through the use of trade-offs and multivariate decision theory. The expected value of the objective function is used as a guide to identify and rank the degree of difficulty of alternative routes. The multiattribute utility theory has been presented in the literature [2-5] and various applications were demonstrated in a variety of decision areas [6-10]. Also, the theory has the advantage of providing a tool for quantification of intangibles [11]. In fact, the approach presented here is versatile enough to accommodate political factors, expert opinion, and public participation. In the context of this paper, the decision-maker is assumed to be a developing country in which the planners have a tendency to risk money but otherwise are averse to risk of failure.

PROLIFERATION ROUTES

Potential proliferation routes generally include research facilities, which are commonly used in educational and research institutions, power reactors, which are normally operated by government commissions or by utility companies, and enrichment facilities, which may be operated with or without international license. In addition to those nuclear proliferation routes, there are accessible means of destruction that could be used by groups of individuals or regular armies and that are different from conventional weapons. The dispersal of biological or chemical material, radioactive material, or detonation of liquid natural gas (LNG) is compared with the difficult task of diversion of nuclear materials or the production of weapons-grade material. The alternative nuclear material proliferation routes considered here are listed below together with other potential arsenal routes:

1. isotope separation by centrifuge
2. research reactors
3. mass spectrograph isotope separation
4. small (clandestine) graphite pile
5. small (clandestine) heavy-water reactor
6. isotope separation by diffusion
7. laser isotope separation
8. Canada deuterium uranium (CANDU) reactor
9. high-temperature, gas-cooled reactor (HTGR)
10. light water reactors [such as the pressurized water reactor

(PWR) and the boiling water reactor (BWR)]

11. liquid-metal fast breeder reactor (LMFBR)

12. chemical/biological weapons

13. fission product dispersal weapons

14. LNG detonation.

ACQUISITION FACTORS

To evaluate the degree of difficulty of acquiring weapons through the alternate proliferation routes or other arsenal techniques, a number of acquisition factors are considered [1]. These are divided into five categories: resources, difficulty of route, cost and schedule, risks, and weapons capability, as shown in Table 1.

To acquire nuclear weapons-grade material, there are specific sources that must be present in the developing country interested in such material or otherwise the country has to seek those resources in other localities. A proliferation effort is likely to be wasted if the party interested in the diversion plans does not have the technological ability to design, select suitable materials to assemble, weld, cast, and forge the components required to make the weapon of interest and the associated deployment mechanism. In many developing countries, the technology is still in a stage that falls short of production of precise equipment or tools and the use of advanced techniques. Technological sophistication generally is gained through life experience and is greatly influenced by the degree of advancement of the local technology; although one can be trained on advanced techniques, the level of technological skill does not depend entirely on training, but is greatly shaped by social, cultural, and various human factors. An engineer who designs a fully automatic machine should have lived through the evolution of the design of the manual machine. A specific proliferation route may require more or less technological sophistication than others.

Table 1. Proliferation route acquisition factors

Category	Acquisition factors
Resources	Technological sophistication Facility requirements Instrument capability Personnel requirements
Difficulty	Availability of information Accessibility of fissile mass
Cost/schedule	Cost Schedule (years to completion)
Risks	Risks to personnel Risks to project detection
Weapons capability	Rate of weapons-grade fissile mass production Weapon reliability

Financial resources may allow a community to acquire a skill, but it could still fall short of having adequate facilities for diversion of nuclear material. Facility requirements depend, however, on the nature of the proliferation route. In certain situations, existence of specific material (other than nuclear material) or a special tool may affect the acquisition of weapons-grade material. High instrumentation capability is required for exploiting certain proliferation routes such as enrichment devices. Furthermore, each route has a different manpower demand.

The degree of difficulty of acquisition of nuclear weapons-grade material depends to a great extent on the availability of the information necessary for diversion or production. Whether such information is available or not, the degree of accessibility of fissile material has a great impact on the ability to provide the desired mass in the time needed for production.

Cost of acquisition of nuclear weapons-grade material from a given proliferation route and the schedule necessary to reach the desired goal would greatly differ from other routes. For example, using a mass spectrograph is less costly but takes more time than a centrifuge isotope separator.

A country alien to the production of nuclear weapons that attempts to acquire such capability without being noticed runs the risk of being detected and of undermining the safety of personnel engaged in the diversion activity. Also, successful attempt of acquiring weapons-grade fissile mass is affected by the rate of production. The final goal of weapons acquisition can only be achieved if the weapon produced is reliable enough to warrant viability of using it for the intended purpose. The weapons capability, even if acquired, would require sophisticated means of deployment.

The above-mentioned acquisition factors are the most important factors relevant to developing countries. Such factors, as defined, represent a list of independent items that do not greatly affect each other in evaluation of the various proliferation routes. That independence is important for the validity of the decision approach employed here.

MULTI-ATTRIBUTE DECISION MODEL

In modeling the decision process of alternative proliferation routes, the analyst must obtain an objective function including the multiple attributes which describe the effectiveness of a decision. Such an objective function would indicate the relative ranking of consequences and identify the trade-offs among various levels of the different attributes [2, 3]. In a risk-free environment, the optimal decision would be the one that maximizes the objective function. But the risk evaluation decision problem involves uncertainties. One approach is to synthesize an objective function such that the decision which maximizes the expected value of the objective function is the optimal decision. Such an objective function is usually called a utility function.

By asking simple questions about trade-offs between quantities, the decision analyst can find a utility function that can serve as a guide in decision-making. If the consequences chosen satisfy certain independence properties, the assessment problem is simplified. The two independent properties to be considered are preferential independence and utility independence [4].

If each consequence is utility independent of its complements, and each pair of consequences is preferentially independent of its complements, then the multi-consequence utility function, U, takes either of two special forms [4-6], namely, the pure product form

$$1 + KU = \prod_{i=1}^{N} [1 + Kk_i U_i(x_i)] \tag{1}$$

or the pure sum form

$$U = \sum_{i=1}^{N} k_i U_i(x_i) \ . \tag{2}$$

In both equations, U and $U_i(x_i)$ are utility functions scaled from zero for the worst state to one for the best state; the k_i values are appropriately selected weights for the given attributes with $0 < k_i \leq 1$, and $K > -1$. The integer N is the number of attributes and constant K is a scaling factor. The weight k_i is usually evaluated as a probability that reflects the decision-maker attitude.

If

$$\sum_{i=1}^{N} k_i = 1 \ ,$$

the utility function is of the pure sum form. If

$$\sum_{i=1}^{N} k_i \neq 1 \ ,$$

the pure product form is appropriate and a value of K must be obtained. The utility function U(x) for each consequence x_i is scaled from 0 to 1; and hence K is needed so that U may also be scaled from 0 to 1. The utility U should be 1 for the most preferred condition, when each of the functions $U_i(x_i)$ equals 0; therefore, K must satisfy the relationship

$$1 + K = \prod_{i-1}^{N} (1 - Kk_i) \ . \tag{3}$$

If

$$\sum_{i=1}^{N} k_i > 1 \ ,$$

the multi-consequence utility function exhibits multi-consequence risk aversion; if

$$\sum_{i=1}^{N} k_i < 1 ,$$

the utility function exhibits multi-consequence risk seeking; and if

$$\sum_{i=1}^{N} k_i = 1 ,$$

the utility function exhibits multi-consequence risk indifference.

The component utility functions can be assessed in a straightforward manner [7-10]. From Table 2, where we have chosen the range for each acquisition factor, we scale our utility function so that U_i (best) = 1 and U_i (worst) = 0. Thus, for risk, of project detection, x_{10} (see Table 2), for example, $U_{10}(0) = 1$ and $U_{10}(100) = 0$. That is, the utility of no risk of project detection is unity while a utility of zero is assigned for maximum risk project detection. To begin with, we find that 18% is indifferent to a lottery yielding 0 or 100%, each with probability 0.5. In other words, a 50-50 chance of detection is almost the same as when the probability of assured detection is 18%.

Therefore, the certainty equivalent for the lottery is

$$U_{10}(18) = 0.5\ U_{10}(0) + 0.5\ U_{10}(100) = 0.5 , \qquad (4)$$

which is the utility of 18% probability of risk detection.

Since 18 is less than the expected value E_{10} of the lottery, that is $E_{10}(18) = (0.5 \times 0 + 0.5 \times 100) = 50$, thus, the utility function exhibits an aversion to risk.

A moderate decision-maker would be indifferent to a return for

Table 2. Acquisition factors for alternate proliferation routes

Acquisition factors, x_i	Values of attributes for alternate acquisition routes					
	Centrifuge isotope separation, ^{235}U	Research reactors, plutonium	Mass spectrograph isotope separation, ^{235}U	Graphite pile, plutonium	Heavy water reactor, plutonium	Diffusion isotope separation, ^{235}U
Resources						
1. Technical sophistication, subjective	80	60	60	60	60	60
2. Facilities requirement, subjective	60	40	60	40	60	60
3. Instrumentation capability, subjective	60	60	80	80	80	80
4. Personnel requirement, persons	350	50	350	100	150	350
Difficulty of Route						
5. Information availability, subjective	60	20	20	20	20	60
6. Accessibility to fission material, subjective	40	60	40	60	60	40
Cost and Schedule						
7. Cost $10 [7]	10	1	10	1	10	10
8. Schedule (time to completion), years	15	0	15	5.5	5.5	5.5
Risks						
9. Risk to personnel, subjective	40	80	40	80	80	40
10. Risk to project detection, subjective	20	40	80	40	80	60
Weapon Capability						
11. Rate of weapons-grade fissile production, weapon/yr	25	0.1	1/18	1	1	3
12. Weapon reliability, subjective	20	40	20	40	40	20

[a]Indifference probability.
[b]First core.

Values of attributes for alternate acquisition routes										
Laser isotope separation, ^{235}U	CANDU reactor, plutonium	HTGR, ^{235}U	PWR/BWR plutonium	LMFBR, plutonium	Chemical/ biological, C/B	Fission product dispersal, radiotoxic isotope	LNG/Gas detonation, LNG/gas	Extreme values of acquisition factors: Worst	Extreme values of acquisition factors: Best	k_i[a]
100	80	80	80	100	20	20	20	100	0	0.3
40	80	80	80	100	20	20	20	100	0	0.26
60	100	60	100	100	20	20	20	100	0	0.111
30	6000	6000	6000	6000	10	10	8	10,000	0	0.22
100	20	20	20	20	40	20	40	100	0	0.15
40	100	40	100	100	20	80	20	100	0	0.08
1	100	100	100	100	1	1	1	100	0	0.12
15	10	10	10	15	0	0	0	20	0	0.115
40	80	40	100	100	80	60	40	100	0	0.015
20	60	80	80	80	20	100	20	100	0	0.18
23	8	20[b]	14	11	0	0	0	0	30	0.26
20	80	20	80	20	20	20	20	100	0	0.0618

certain that is equivalent to the expected value of the lottery. Such an attitude is normally represented by a straight line pattern that goes through U = 0 and U = 100. In contrast, a conservative pattern exhibits aversion to risk; thus, the decision-maker would accept the chances game only if its alternative is an outcome of higher risk than the expected value. The gambler, on the other hand, may prefer the lottery even if the alternative is of low risk for certain and hence a risk-taker would only feel indifferent to the lottery or the assured outcome if the alternative has a risk much higher than the expected risk from the lottery.

In a similar way, we can choose a few more values for the utility function; for example, we find 8% indifferent to a 0 and 18% lottery; and 35% indifferent to an 18 and 100% lottery; both at 50-50 chance, hence,

$$U_{10}(8) = 0.5\ U_{10}(0) + 0.5\ U_{10}(18) = 0.75 \qquad (5)$$

and

$$U_{10}(35) = 0.5\ U_{10}(18) + 0.5\ U_{10}(100) = 0.25\ . \qquad (6)$$

The empirically assessed points are then plotted to represent graphically the function dependence of U_i on x_i as shown in Figure 1; here i = 10.

This process is to be repeated for each attribute and a set of smoothed utility curves can thus be constructed. A country or a party having the intention to defy international restrictions through diversion of nuclear material would be a risk-taker and hence the utility functions associated with some acquisition factors would exhibit a

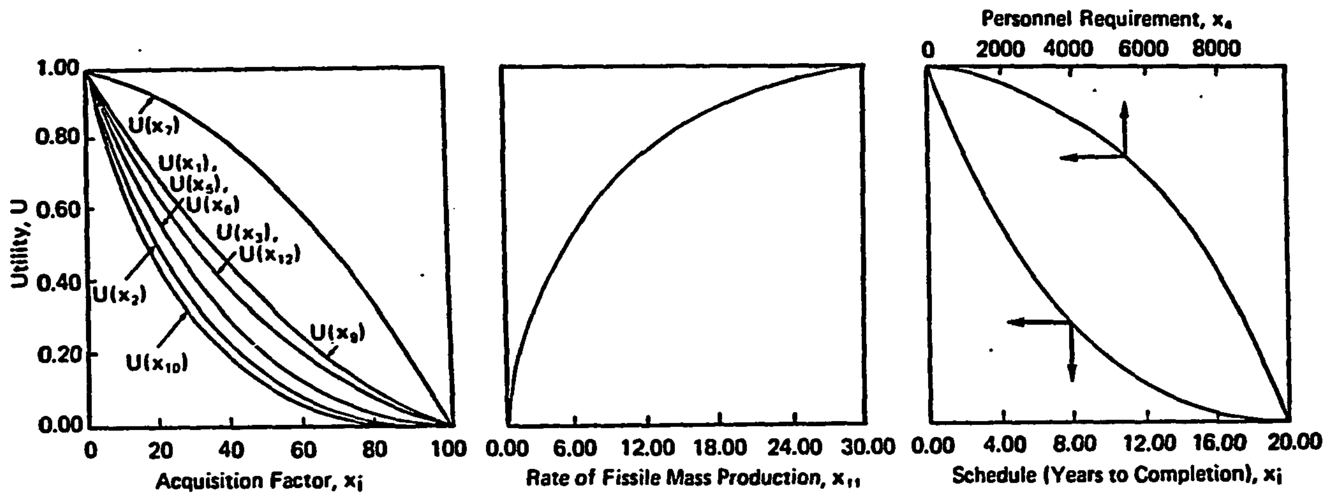


Figure 1. Utility functions for the acquisition factors considered for alternative proliferation routes

gambling trend. Such trend is depicted by slowly decreasing utility functions near the attribute value at which the corresponding utility is unity. The point representing that value of the attribute is almost the center of curvature of the utility curve. Those acquisition factors for which the expected utility functions are of the risk-taking type are personnel requirements, cost and rate of weapons-grade fissile mass production. Countries determined to acquire nuclear weapons are likely to take the risk of increased cost, take chances in meeting personnel requirements, and be patient for low production rates. The trend is reversed with issues such as risk of plan detection and facility requirements, where the decision-maker resorts to conservatism. Although a gambler may represent a direct threat because of impulsive action, a planner seeking diversion of nuclear fuel, who has a certain level of aversion to risk in some of his considerations, is the one who is likely to succeed. Figure 1 illustrates a conservative preference pattern in the utility functions associated with 9 of the 12 acquisition factors. The curvature of such functions is such that the focal point is on the opposite side of the point representing the attribute level of the highest utility. The most conservative attitude is toward the risk of project detection and the least conservative attitude is toward the risk to personnel.

The utility functions represent the general preference patterns of a decision-maker who is commonly a risk averse but willing to gamble with respect to cost, personnel requirement, and/or the rate at which the fossil mass is produced. Normally, the exact values of the utility function are dependent on the decision-maker. Methods of extracting the preferences of a given decision-maker are available, such as the Delphi. Other techniques can be used such as the random choice in which the preferences in such a sensitive situation can be inferred from answers to a well-designed set of questions which can be indirectly used to probe the actual preferences.

Generally, under the assumptions given here, the preference patterns will not greatly deviate from those shown in Figure 1 except in the degree of conservatism. The approach here is not greatly sensitive to the precision of the utility values for each attribute. The outcome of the analysis may change only if the preference pattern is inverted or totally assumed a different shape.

For computation purposes, each utility function may be fitted to an exponential form, that is,

$$U_1(x_1) = 1.067 \exp(-0.02773\, x_1) - 0.06667 \tag{7}$$

$$U_2(x_2) = 1.032 \exp(-0.03469\, x_2) - 0.03214 \tag{8}$$

$$U_3(x_3) = 1.225 \exp(-0.01695\, x_3) - 0.2250 \tag{9}$$

$$U_4(x_4) = 1.098 - 0.09796 \exp(0.0002417\, x_4) \tag{10}$$

$$U_5(x_5) = 1.225 \exp(-0.01695\, x_5) - 0.2250 \tag{11}$$

$$U_6(x_6) = 1.225 \exp(-0.01695\, x_6) - 0.2250 \tag{12}$$

$$U_7(x_7) = 1.2250 - 0.225 \exp(0.01695\, x_7) \tag{13}$$

$$U_8(x_8) = 1.038 \exp(-0.1658\, x_8) - 0.03765 \tag{14}$$

$$U_9(x_9) = 1.361 \exp(-0.01327\ x_9) - 0.3612 \tag{15}$$

$$U_{10}(x_{10}) = 1.021 \exp(-0.03892\ x_{10}) - 0.02083 \tag{16}$$

$$U_{11}(x_{11}) = 1.032 - 1.032 \exp(-0.1156\ x_{11}) \tag{17}$$

$$U_{12}(x_{12}) = 1.225 \exp(-0.01695\ x_{12}) - 0.225 \tag{18}$$

Often, utility functions can be fitted to exponential forms of the type given in equations 7 through 18. However, other forms may be used to provide the best fit for the data points. In simple computations, utility curves can very well serve the purpose.

As shown from the list of attributes given in Table 1, the measures of several attributes are treated as subjective. This is because decision-makers are normally inclined to use judgmental analysis based on their experience and the information they acquired. Also, decision-making is not an exact science, and the present situation is a case of decision under uncertainty. To accommodate for uncertainties, the utility functions have been synthesized through the lottery game described above by going through at least two different ways of preference assessment.

To combine the single-consequence utility functions obtained by the procedure above into a single, multi-consequence utility function of the pure sum or pure product form, a set of questions would be used to estimate the weights, k_i; such as "For what probability p is the decision-maker indifferent to the choice between

1. The situation with all consequences but the ith, at their least preferred values $[(X_j) = (X_{j*}), j \neq i$ where the

subscript * is used to refer to the least preferred value], and the ith consequence at its most preferred value ($x_i = x_j^*$ where superscript * is used to refer to the most preferred value).

2. An alternative with two possible results: all consequences at their most preferred values [$(x_i) = (x_i^*)$, $i = 1, \ldots, N$], with probability p, or all consequences at their least preferred values [$(x_i) = (x_{i*})$, $i = 1, \ldots, N$] with probability of $(1 - p)$?"

The utility value of the first alternative is k_i, since x_i is at its most preferred value [$U_i(x_i^*) = 1$], and all other consequences are at their least preferred values [$U_j(x_{j*}) = 0$, $j \neq i$]; the expected utility of the second alternative is p, since there is a chance p of obtaining the most preferred situation (which has a utility value of 1), and a chance $1 - p$ of obtaining the least preferred situation (which has a utility of 0). For the decision-maker to be indifferent to the choice between these two alternatives, the utility value of the first alternative must equal the expected utility of the second alternative. Therefore, $k_i = p$, and k_i is a positive number less than 1. It is, however, advisable in assessing the k_i values to order their magnitude. To do this, we set all 12 attributes given in Table 2 at their worst levels and ask the question, "If only one attribute, x_i, $i = 1, 2, \ldots, 12$, could be raised to its best level, which one would be preferred?" The response would be attribute x_1. This implies that k_1 must be assigned the largest value of the k_is. Had there been indifference between moving either x_i or x_j to its best level, then k_i would equal k_j. After several adjustments, the result is in the following order:

$$k_1 \quad k_2 = k_{11} > k_4 > k_{10} > k_5 > k_7$$
$$> k_8 > k_3 > k_6 > k_{12} > k_9 \,. \qquad (19)$$

The arrangement of the weights given in equation 19 shows that for a successful diversion of nuclear fuel, technological sophistication carries higher weight than availability of required facilities which is of the same degree of importance as the rate of production of weapons-grade fissile mass. The latter is more important in selecting a given proliferation route than availability of information, a factor that carries more weight than cost. This is an indication that the developing country is expected not to attach much weight to financial risk in diversion compared to the value of technological know-how and the required facilities and rate of production to achieve its goals. Nevertheless, finance of the selected alternative route for proliferation is rather more critical than the date of completion which is of higher weight than acquiring instrumentation capability to assure precision in production using specific routes. Accessibility to fission material comes next, although difficulty in acquisition of such material is important to the diversion plan. Such accessibility is surely more important than the reliability of the produced weapon. Finally, the risk to personnel carries the lowest weight due to the nature of the mission.

After the relative values are established among k_i, their numerical values are estimated using the lottery game described above and the results are listed in Table 2. The attribute levels of the degree to which a particular alternative contributes in evaluating the utility function are also given in Table 2, which contains the extreme values of the acquisition factors.

The attribute levels are estimated for each alternative based on the expected experience in typical developing countries which are considering the use of nuclear power.

A deterministic evaluation of the acquisition factors for each proliferation rank is performed based on the best engineering judgments, subjective evaluations, and current technical information. Most of the acquisition factors can only be evaluated in subjective units, in which case the best and the worst values are arbitrarily assigned as 0 and 100, respectively. In other cases, relevant units are used. The relative ranking rather than absolute levels is important in the decision model.

To evaluate multi-consequence utility function U for each alternative proliferation route, equation 1 is relevant since K = -0.797 from equation 3. A computer program yields the utility values for each proliferation route, which are given in Table 3. The higher the utility value for a proliferation route, the easier is the weapons acquisition. We notice that the easiest routes are the nonnuclear options, which are included for comparison with the proliferation routes, that is, LNG/gas detonation, chemical/biological dispersal, and fission product dispersal. Among the nuclear proliferation routes, the LMFBR is the hardest; next, in order of the degree of difficulty, are the CANDU reactor and

Table 3. Degree of difficulty for alternative routes to proliferation

Proliferation routes	Utility U	Relative ease of weapons acquisition
LNG/gas detonation	0.7332	Easiest
Chemical/biological dispersal	0.7311	
Fission product dispersal	0.6937	
Supercritical centrifuge isotope separation	0.6417	
Laser isotope separation	0.6352	
Research reactor	0.5907	
Clandestine graphite pile	0.5557	
Diffusion isotope separation	0.5244	
HTGR	0.5183	
Clandestine heavy water reactor	0.5150	
Mass spectrograph isotope separation	0.4909	
PWR/BWR	0.4534	
CANDU reactor	0.4199	
LMFBR	0.4123	Hardest

PWR/BWR. Easiest nuclear routes to proliferation are the supercritical centrifuge isotope separation, laser isotope separation, and the research reactor, in order of the degree of ease to weapons acquisition.

CONCLUSIONS

An objective assessment of the alternative routes to acquire nuclear weapons is essential to arrive at a coherent policy for proliferation deterrence. In this paper, we have used a quantitative decision model, based on the Multi-Attribute Utility Theory, to arrive at the relative ranking of 14 important routes. Eleven of the routes considered result in nuclear explosives; the rest are nonnuclear in nature but are capable of significant damage to life and property of the public. The results indicate that the most difficult route to acquire nuclear weapons for a nonnuclear country is through nuclear power reactors, specifically LMFBRs, heavy water CANDU reactors, and light water reactors (BWRs/PWRs). The easier routes are through supercritical centrifuge isotope separation, laser isotope separation, and research reactor. The easiest path available to incur substantial damage to life and property is, however, unrelated to nuclear explosives, such as LNG detonation, chemical/biological dispersal, and fission product dispersal.

The limitation of the study essentially emerges from the difficulty of validating the underlying assumptions of the model, namely the preferential and the utility independence. Deterministic values are assigned to the acquisition factors for the sake of simplicity. Uncertainty can be accounted for by using probability distributions and more involved computations. Further, a sensitivity analysis is suggested to assess the impact of different utility functions, and to screen out the dominant acquisition factors.

REFERENCES

1. Paul Wyatt and Howard Braun, "Evaluation of Alternative Proliferation Routes" (Westinghouse Electric Corporation, May 1977).

2. R. Schlaifer, Analysis of Decisions Under Uncertainty (McGraw-Hill Book Company, New York, 1969).

3. R. L. Keeney, "Multiple Utility Functions," Oper. Res., 22, 1, 101 (1974).

4. R. L. Keeney, "A Decision Analysis with Multiple Objectives: The Mexico City Airport," Bell J. Econ. Manag. Sci., 4, 1, 101 (1973).

5. H. Raiffa, Decision Analysis, Introductory Lectures on Choice Under Uncertainty (Addison-Wesley Publishing Co., Inc., Reading, Massachusetts, 1970).

6. Shahid Ahmed and A. A. Husseiny, "Selection of Energy Systems Based on Decision Analysis of Various Attributes," Proc. Symp. Simulation, Modeling, and Decision in Energy Systems, Montreal, Canada, June 1-2, 1978.

7. E. F. Wood, "Applying Multi-Attribute Utility Theory to Evaluation of Tisza River Basin Development Plans," International Institute of Applied Systems Analysis Conf. (International Institute of Applied Systems, Austria, 1976), Vol. 2.

8. Shahid Ahmed and A. A. Husseiny, "A Formal Methodology for Acceptability Analysis of Alternate Sites for Nuclear Power Stations," Nucl. Eng. Des., 51, 361 (1979).

9. Shahid Ahmed and A. A. Husseiny, "A Multivariate-Utility Approach for Selection of Energy Sources," Energy, 3, 669 (1978).

10. Shahid Ahmed and A. A. Husseiny, "Accommodation of Public Preference in Selection of Energy Alternatives," Trans. Am. Nucl. Soc., 30, 13 (1978).

11. J. J. Turnage and A. A. Husseiny, "Public Participation in Siting Decisions," Trans. Am. Nucl. Soc., 34, 689 (1980).

SECTION III. A METHOD FOR PROPAGATING UNCERTAINTY IN PROBABILISTIC RISK ASSESSMENT

A method for propagating uncertainty in probabilistic risk assessment

Shahid Ahmed

R. E. Clark

D. R. Metcalf

From The Kuljian Corporation, 3700 Science Center, Philadelphia, Pennsylvania 19104 (Ahmed) and University of Virginia Reactor Facility, Charlottesville, Virginia 22901 (Clark and Metcalf)

Published in Nuclear Technology 59 (1982): 238-245

ABSTRACT

A method is developed to propagate uncertainties in the basic event unavailabilities through a logic model to obtain the transient overpower event unavailability. The method consists of combining probability distributions in the discrete form without performing any sampling. The results are shown to be sufficiently accurate and contain no sampling errors; the computation time is considerably less compared to Monte Carlo simulation and histogram propagation. Uncertainty propagation methods are found to be sensitive to the spread of the basic event unavailability distributions; the proposed method produces results less conservative compared to those from propagation of moments or Monte Carlo simulation.

INTRODUCTION

In risk assessment of large technological systems, such as nuclear power plants (NPPs), uncertainty is involved at various stages of the analysis, ranging from the failure data analysis to the phenomenology of core melt. Here, we deal specifically with the uncertainty in the basic event failure data and present a method for propagating uncertainty to quantify system unavailability.

It is well-known that, because of the relatively short operating experience of NPPs, significant data bases of many component failure times are lacking and this is one important factor of many reflected in uncertainty in the failure rates. A realistic approach is to quantify uncertainty in the component failure rates in the form of probability distributions. A quantity of interest in risk assessment is the probability of failure of a standby safety system when it is demanded (also called "conditional unavailability") in the event of an abnormal transient. A system logic model, such as a fault tree or reliability block diagram, is constructed to determine various failure modes or cutsets of the system to quantify the system unavailability and the associated uncertainty. The uncertainties in the basic event unavailabilities are propagated to obtain the uncertainty in the transient overpower (TOP) event unavailability.

In this paper, a method is presented for propagating uncertainty in the discrete form. Here, the discrete probability distributions (DPDs) of the basic event unavailabilities are propagated as complete distributions, so that no sampling is necessary; this avoids sampling errors. The DPD method, originally proposed by Kaplan [1], is developed here for efficient use in probabilistic risk analysis (PRA). The method is compared with exact analytical results, as well as with other methods of uncertainty propagation, such as Monte Carlo simulation and the method of moments. The DPD method is also compared with the combination of probability distributions in the form of histograms developed by Colombo and Jaarsma [2].

A CASE FOR DISCRETE DISTRIBUTIONS IN PRA

Combining random variables using Monte Carlo requires random sampling from the distributions. In probabilistic risk assessment of NPPs, this may pose some problem, particularly because the uncertainty in the failure data of the basic events is large and the distributions positively skewed. Because the tail of the failure rate distribution of basic events can be spread by orders of magnitude, the resultant distribution through Monte Carlo would be sensitive to the random sampling at the tails of the distribution. Too many points picked up at higher values of probability distribution of competent failure rates (or component unavailability) would result in a very conservative final distribution. On the other hand, if not enough points are picked up at the tails, the results may be very optimistic.

One way to deal with this situation is to discretize the failure rate distributions of the basic events and perform Monte Carlo analyses of the discrete distributions. Discrete distributions have the following advantages over continuous probability distributions [3].

1. Since we consider the subjective probability framework to be appropriate for PRA, it is much more flexible to have subjective distributions in the form of discrete distributions. This is compatible with the limited failure data available from the nuclear industry. The data analysis techniques should not be more refined than the available data suggest. The restrictions imposed on the failure distribution data of a component by assigning a continuous probability distribution are

unrealistic since we do not have enough information to identify a continuous distribution. The best we can do is derive a failure rate (or failures per demand) distribution in the discrete form.

2. The basic failure rate distribution for the newly designed components can be formulated in the discrete form by pooling relevant information such as design, failure history of similar components, subjective assessment, and the best engineering judgment. This can be best obtained in the form of probability histograms. Methods of constructing subjective probability distributions are well-documented in the literature and have been successfully used to quantify expert opinion [4-8].

3. By formulating DPDs of component failure rates, we ensure that Monte Carlo sampling uses values that are realistic and not dependent on the properties of the distribution (such as large tails of log-normal distribution).

4. The human error probability data available are often in terms of a median value and lower and upper bounds [9]. One way to incorporate all the information is to fit a three-parameter continuous distribution [3]. A less complex but more realistic technique is to formulate a DPD that satisfies the conditions above.

Rather than using Monte Carlo sampling, however, the DPD method performs the complete combination of distributions. Here, instead of random sampling, probability distributions in the discrete form are propagated. This combination procedure is "exact"; i.e., no errors due to sampling are introduced.

DISCRETE PROPAGATION

The DPD Method

Operations like the product or sum of DPDs can be defined using basic axioms of the calculus of probability. Thus, if $f_x(X)$ and $f_y(Y)$ are probability distributions of independent random variables X and Y, so that

$$f_x(X) \equiv <x_1,p_1>,\ <x_2,p_2>,\ \ldots,\ <x_n,p_n>$$

and

$$f_y(Y) \equiv <y_1,q_1>,\ <y_2,q_2>,\ \ldots,\ <y_m,q_m>,$$

then

$$f_{z=x+y}(Z) = <x_1+y_1,p_1q_1>,\ <x_1+y_2,p_1q_2>,\ \ldots,\ <x_n+y_m,p_nq_m>$$

and

$$f_{z=xy}(Z) = <x_1y_1,p_1q_1>,\ <x_1y_2,p_1q_2>,\ \ldots,\ <x_ny_m,p_nq_m>,$$

where

$<>$ = cell of the histogram

x and y = values of the random variables X and Y

p and q = associated probabilities.

As an example, if we represent discrete distributions X and Y (as shown in Figure 1), then the probabilistic sum X + Y and multiplication X·Y obtained through the DPD method are also shown in Figure 1. The calculational procedure for obtaining the discrete distribution for the random variable (RV) X + Y is shown in Table 1.

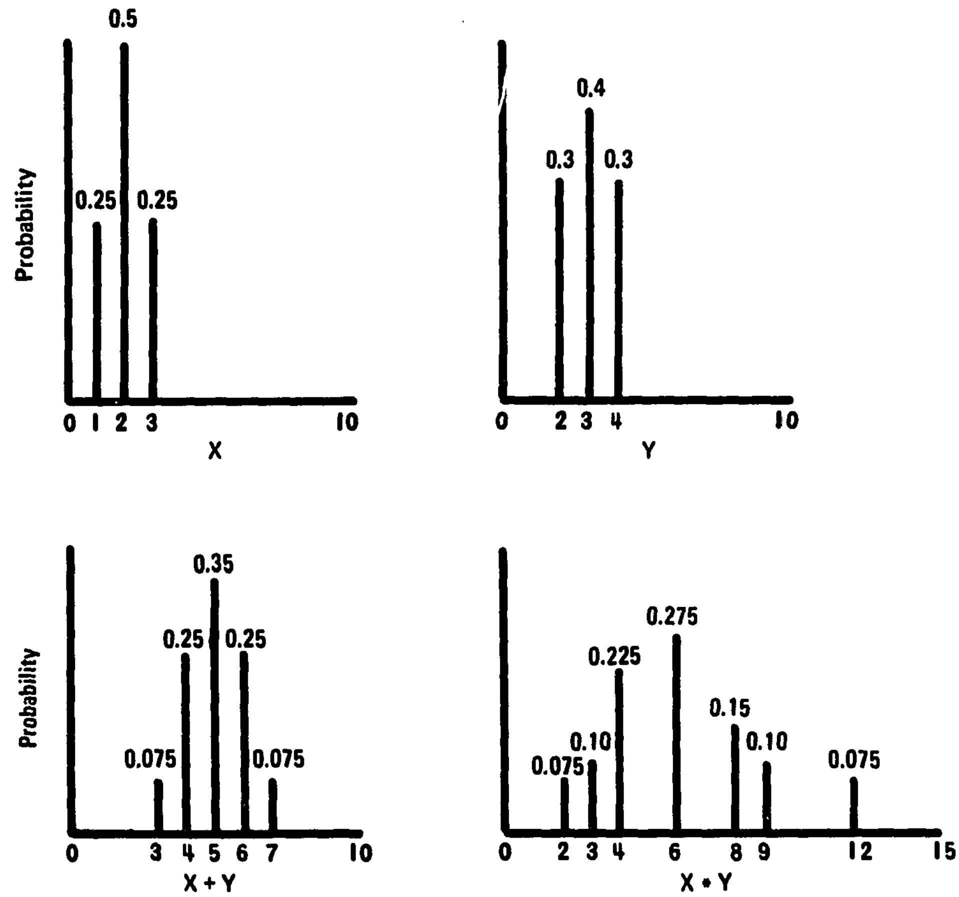


Figure 1. Addition and multiplication of random variable in discrete form

Table 1. Calculation of the probabilistic sum X + Y

All combinations		Reduced from	
Value	Probability	Value	Probability
3	(0.25)(0.3) = 0.075	3	0.075
4	(0.25)(0.4) = 0.10	4	0.10 + 0.15 = 0.25
5	(0.25)(0.3) = 0.075	5	0.075 + 0.20 + 0.75 = 0.35
4	(0.50)(0.3) = 0.15	6	0.15 + 0.10 = 0.25
5	(0.50)(0.4) = 0.20	7	0.075
6	(0.50)(0.3) = 0.15		
5	(0.25)(0.3) = 0.075		
6	(0.25)(0.4) = 0.10		
7	(0.25)(0.3) = 0.075		

Comparison with Analytical Results

To illustrate the results from the DPD method, let us combine a set of RVs for which the resultant distribution can be found analytically. For this purpose, we shall use the result that the product of n identical log-normal distributions, each with mean α and variance β^2, is also a log-normal distribution with

$$\alpha_p = \alpha^n \tag{1}$$

and

$$\beta_p^2 = (\beta^2 + \alpha^2)^n - (\alpha^2)^n \ , \tag{2}$$

where α_p and β_p^2 are the mean and variance of the product, respectively. Also, the parameters of the resultant log-normal distribution are given by

$$\sigma_p^2 = \ln \left(\frac{\beta_p^2}{\alpha_p^2} + 1\right) \tag{3}$$

and

$$\mu_p = \ln(\alpha_p) - \frac{1}{2}\sigma_p^2 . \qquad (4)$$

We consider five log-normal distributions with $\alpha = 10$ and $\beta^2 = 100$; the parameters of the products of log-normal distributions are obtained using equations 1 through 4 as $\sigma_p = 1.86$ and $\mu_p = 9.78$. The cumulative probability distribution obtained analytically is plotted in Figure 2.

The product of the log-normal distributions is next obtained using the DPD method. The log-normal distribution with $\alpha = 10$ and $\beta^2 = 100$ is discretized with equal probability intervals, as shown in Table 2. The DPD method is next applied to obtain the product of discrete distributions, using a computer code described in Section V, and the results are given in Table 3. Various percentile points obtained by the DPD method are also plotted in Figure 2 against the distribution obtained analytically. The results indicate that the DPD method is close to the exact distribution.

Comparison with Histogram Method and Monte Carlo Simulation

Combining RVs in the form of probability histograms was first proposed by Ingram et al. [10] and later developed by Colombo and Jaarsma [2]. Here, the DPD and histogram methods are compared in terms of accuracy and computer time; they are also compared with Monte Carlo simulation with varying sample sizes.

The example chosen is from Appendix II of Ref. 11, where six log-normal RVs are combined. The final expressions for the TOP event

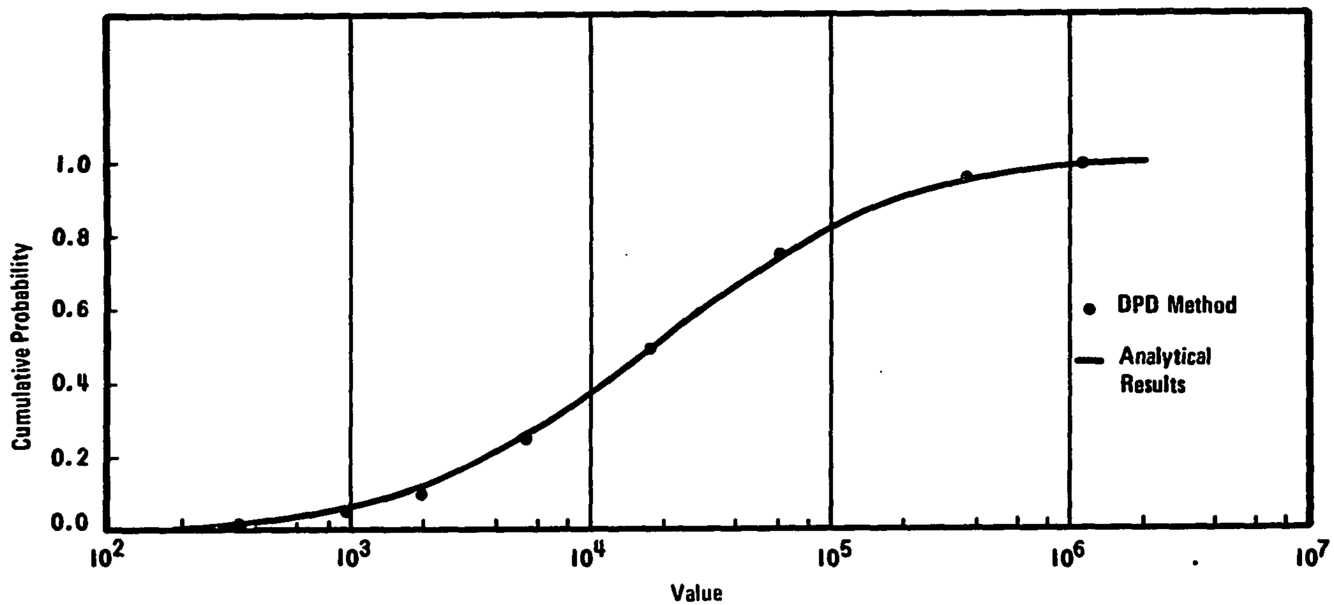


Figure 2. Comparison of product of five identical log-normal distributions with DPD method of uncertainty propagation

Table 2. Discretization of log-normal distribution with mean = 10 and variance = 100

Value	Probability	Cumulative probability
1.79	0.125	0.125
3.34	0.125	0.250
4.73	0.125	0.375
6.19	0.125	0.500
8.08	0.125	0.625
10.56	0.125	0.750
15.00	0.125	0.875
27.87	0.125	1.000

Table 3. Multiplication of five log-normal distributions

Cumulative percentile	Exact value	DPD method
1st	2.3×10^2	2.8×10^2
5th	8.3×10^2	9.1×10^2
10th	1.6×10^3	2.0×10^3
25th	5.0×10^3	5.4×10^3
50th	1.8×10^4	1.8×10^4
75th	6.2×10^4	6.0×10^4
95th	3.8×10^5	3.6×10^5
99th	1.3×10^6	1.1×10^6

unavailability are obtained for two cases--uncoupled and coupled--and are, respectively:

$$P(T) = X_1 + (X_2 + X_3)\cdot(X_4 + X_5) + X_6 + (X_3\cdot X_3\cdot X_5)^{1/2}$$

and

$$P_{coupled} = X_1 + (X_2 + X_3)^2 + X_6 + (X_3^3)^{1/2} \, .$$

For the DPD method, the component unavailabilities for X_1, X_2, ..., X_6 are each discretized into five equal probability intervals, as shown in Table 4. Results obtained through the DPD method for the coupled and uncoupled cases are given in Tables 5 and 6, respectively. The results of the Monte Carlo simulation for sample sizes 1200, 2600, 4800, and 6000 are evaluated by using the SAMPLE computer code [11]. Reactor Safety Study (RSS) evaluations of Monte Carlo simulation, as well as the results of the histogram method reported in Ref. 12, are also given in Tables 5 and 6.

The results indicate that both the DPD and histogram methods match the Monte Carlo simulation closely. Monte Carlo with increasing sample size is used to determine the pattern of function convergence. However, because of the sampling error, it is not obvious where the function converges. No assumptions, other than discrete approximation of continuous log-normal distributions, are made in the case of the DPD method, which shows lower values at higher percentile points. Thus, the DPD method not only avoids the sampling errors involved in Monte Carlo simulation, but also keeps the resultant function uncertainty from becoming unnecessarily large because of the large tails of log-normal distributions of the basic event unavailabities.

The histogram propagation method essentially provides the same results as the PDP method but requires a much longer computer time because it requires twice as many computations. Computer time increases

Table 4. Discrete distribution of component unavailability used to obtain TOP event unavailability by DPD method

Component	Probability				
	0.20	0.20	0.20	0.20	0.20
X_1	4.1×10^{-4}	7.1×10^{-4}	1.0×10^{-3}	1.4×10^{-3}	2.7×10^{-3}
X_2	1.2×10^{-2}	2.1×10^{-2}	3.0×10^{-2}	4.3×10^{-2}	8.1×10^{-2}
X_3	4.1×10^{-3}	7.1×10^{-3}	1.0×10^{-2}	1.4×10^{-2}	2.7×10^{-2}
X_4	1.2×10^{-2}	2.1×10^{-2}	3.0×10^{-2}	4.3×10^{-2}	8.1×10^{-2}
X_5	4.1×10^{-3}	7.1×10^{-3}	1.0×10^{-2}	1.4×10^{-2}	2.7×10^{-2}
X_6	1.7×10^{-3}	2.4×10^{-3}	3.0×10^{-3}	3.8×10^{-3}	5.5×10^{-3}

Table 5. TOP event unavailability obtained by DPD, histogram, and Monte Carlo methods for uncoupled case ($\times10^{-3}$)

Cumulative probability	Monte Carlo (sample sizes)					Histogram method	DPD method
	RSS[a]	This report					
	1200	1200	2600	4800	6000		
5th	4.54	4.41	4.50	4.48	4.47	4.38	4.54
10th	5.08	4.99	5.09	5.09	5.03	4.98	5.12
20th	5.86	5.86	5.86	5.86	5.83	5.82	5.93
50th	7.75	7.74	7.75	7.80	7.76	7.98	7.95
80th	10.60	10.32	10.25	10.48	10.45	11.40	10.50
90th	12.40	12.19	12.11	12.33	12.22	14.00	12.15
95th	14.40	13.99	13.99	14.26	14.08	17.10	13.65
CPU[b] time(s)		1.151	2.391	4.337	5.400	88.00[c]	2.781

[a]The RSS provides the upper end values of the confidence interval related to the sampling error.

[b]CPU = central processing unit.

[c]For both coupled and uncoupled cases.

Table 6. TOP event unavailability obtained by DPD, histogram, and Monte Carlo Methods for coupled case ($\times 10^{-3}$)

Cumulative probability	Monte Carlo (sample sizes) RSS[a] 1200	This report 1200	2600	4800	6000	Histo-gram method	DPD method
5th	4.2	4.33	4.26	4.25	4.23	4.14	4.20
10th		4.85	4.88	4.86	4.78	4.76	4.81
20th		5.66	5.71	5.73	5.67	5.64	5.73
50th	8.2	8.04	7.94	8.07	8.02	7.99	8.32
80th		12.02	12.02	12.01	11.94	12.00	12.67
90th		15.60	15.53	15.51	15.38	15.60	14.62
95th	20.0	19.30	19.43	19.51	19.43	20.20	17.01
CPU time(s)		1.147	2.388	4.330	5.409	88.00[b]	0.159

[a]The RSS provides the upper end values of the confidence interval related to the sampling error.

[b]For both coupled and uncoupled cases.

rapidly with larger sample size in Monte Carlo simulation. For a complex function, a large sample size is required to arrive at results with reasonably small sampling error. This situation is handled by considering only discrete points in the continuous distribution in the DPD and histogram methods. The approximation imposed does not compromise the accuracy of the results, as we noticed in the section Comparison with Analytical Results.

SENSITIVITY STUDY

The sensitivity of the spread of basic event unavailability distributions on the TOP event unavailability, pertaining to DPD and other uncertainty propagation techniques, is investigated in this section. Along with the DPD method, two other uncertainty propagation techniques, the method of moments and Monte Carlo simulation, are considered.

In the method of moments, if X and Y are two random variables, then the expectation of the sum of the random variables is the sum of the expectations of the random variables; that is,

$$\alpha_{(X+Y)} = \alpha_X + \alpha_Y \ . \tag{5}$$

Also, considering that the events X and Y are independent, the variance (second moment about the mean) of the sum of the random variables is the sum of the variances or

$$\beta^2_{(X+Y)} = \beta^2_X + \beta^2_Y \ . \tag{6}$$

Similarly, for the expectation of the product of two random variables,

$$\alpha_{(X\cdot Y)} = \alpha_X \cdot \alpha_Y \ , \tag{7}$$

again considering events X and Y as independent. In addition,

$$\beta^2_{(X\cdot Y)} = \alpha^2_X \beta^2_Y + \alpha^2_X \beta^2_Y + \beta^2_X \beta^2_Y \ . \tag{8}$$

Further, a log-normal fit to the mean and variance is evaluated analytically, and various percentile points, such as the 5th and 95th, are calculated. To investigate the sensitivity of the uncertainty propagation technique to the basic event uncertainty, a number of sample cases were studied, ranging from simple to complex unavailability expressions. To illustrate and compare the methods, let us consider the TOP event unavailability represented by

$$\mathrm{TOP} = X_1 + X_2X_3 + X_4X_5 + X_6 \,, \tag{9}$$

where X_1, X_2, ..., X_6 are the unavailabilities of the basic events. The data for the basic events are available in the form of median and range factors (RF);

$$\mathrm{RF} = \frac{\text{median}}{\text{5th percentile}} = \frac{\text{95th percentile}}{\text{median}} \;;$$

the median values for X_1, X_2, ..., X_6 are 0.001, 0.03, 0.01, 0.03, 0.01, and 0.003, respectively. Two cases for this example are given; in case 1, all basic events have RF = 3, and in case 2, all basic events have RF = 10. In the Monte Carlo method, sample sizes of 3000 and 1200 were run with log-normal distribution. To propagate uncertainty through the DPDs, the log-normal distributions of the basic events were discretized into five cells with probabilities 0.15, 0.20, 0.30, 0.20, and 0.15; the probability of each cell is associated at its mean cell value rather than the midpoint, which allows us to associate the probabilities at the tails of the distribution at the first and fifth intervals. The propagation of discrete distribution is also considered through Monte Carlo sampling for the sample size of 1200. For the method of moments, the mean and variance of the basic events were evaluated for log-normal distribution, and equations 5 through 8 were used. The results of this example are shown in Table 7.

Table 7. System unavailability by different uncertainty propagation techniques for sample case ($X_1 + X_2X_3 + X_4X_5 + X_6$) (where X_i, i = 1, 2, ..., 6 are the unavailabilities of the basic events)

Uncertainty methods	Percentile				Standard deviation (variance)
	5th	Median	Mean	95th	
Case 1 - Small Spread for Basic Events (RF = 3)					
Discrete	0.00247	0.00530	0.005936	0.0146	0.00269 (0.0000072)
Method of moments	0.00267	0.00527	0.005936	0.0132	0.00307 (0.0000094)
Monte Carlo					
Sample size 3000	0.00261	0.00531	0.005955	0.0113	0.00294 (0.0000086)
Sample size 1200	0.00260	0.00531	0.005996	0.0114	0.00305 (0.0000093)
Discrete size 1200	0.00271	0.00524	0.005788	0.0110	0.00250 (0.0000063)
Case 2 - Large Spread for Basic Events (RF = 10)					
Discrete	0.00187	0.00875	0.01490	0.04542	0.01464 (0.000214)
Method of moments	0.000836	0.00670	0.01491	0.05367	0.0296 (0.000876)
Monte Carlo					
Sample size 3000	0.00187	0.00848	0.01426	0.04407	0.01993 (0.000397)
Sample size 1200	0.00183	0.00874	0.01477	0.04493	0.02174 (0.000473)
Discrete size 1200	0.00197	0.00931	0.01402	0.04337	0.01375 (0.000189)

For less spread in the basic event distribution (RF = 3), there are no significant differences in the result. However, the differences are pronounced when the uncertainty increases (RF = 10). Here, we notice that the method of moments gives a much wider spread; the standard deviation is ~100% larger than for the discrete case. Also, the Monte Carlo using continuous log-normal gives a larger spread compared to the discrete case (standard deviation is ~35% larger), which is because the tail is truncated in the discretization scheme. With more complex unavailability expressions than considered here (equation 9), Monte Carlo results become more sensitive to sample size.

COMPUTER CODE PUFD

The computer code PUFD (propagation of uncertainty through finite probability distribution) is a general purpose computer program developed to perform uncertainty analysis by propagating distribution of the random variables in the discrete form [13]. The method is described in The DPD Method section. As opposed to the Monte Carlo approach, this program does not admit sampling on the variables, thus avoiding sampling errors. Most mathematical operations can be performed on the random variables. Various statistics of the resultant distribution are provided.

The PUFD program reads and evaluates a user-defined algebraic function of Boolean expression. Each independent variable in the function may have up to ten input values with associated probabilities. The program evaluates the function for every combination of independent variables and calculates the probability for that functional evaluation by taking the product of the associated probabilities.

A typical PUFD run might require several hundred thousand functional evaluations. It would not be practical to store each functional value. The program groups the probabilities associated with each functional evaluation into user-defined cells based on functional value. Consider the user input function $F = A \cdot B + C \cdot D + D \cdot E + B \cdot C \cdot D$ and the input variable values and associated probabilities

$$A \equiv a_1, p(a_1), a_2, p(a_2), \ldots, a_i, p(a_i) \;;\; i \le 10 \;,$$

$$B \equiv b_1, p(b_1), b_2, p(b_2), \ldots, b_j, p(b_j) \;;\; j \le 10 \;,$$

$$C \equiv c_1, p(c_1), c_2, p(c_2), \ldots, c_k, p(c_k) \;;\; k \le 10 ,$$

$$D \equiv d_1, p(d_1), d_2, p(d_2), \ldots, d_1, p(d_1) \;;\; 1 \le 10 ,$$

and

$$E \equiv e_1, p(e_1), e_2, p(e_2), \ldots, e_m, p(e_m) \;;\; m \le 10 ,$$

where the lower case letters represent the values of the respective random variables A, B, ..., E, and the $p(\cdot)$ is the probability of $(\cdot)$.

This case will require $i \cdot j \cdot k \cdot 1 \cdot m$ function evaluations. The first function value will be $f = a_1 \cdot b_1 + c_1 \cdot d_1 + d_1 \cdot e_1 + b_1 \cdot c_1 \cdot d_1$ and its associated probability will be $p = p(a_1) \cdot p(b_1) \cdot p(c_1) \cdot p(d_1) \cdot p(e_1)$.

The program will now search the mesh to find the cell in which to add the probability p. Assuming a mesh array M where M(0) is less than the minimum f and M(N) is greater than or equal to the maximum f, then a J exists so that $M(J-1) < f \le M(J)$. The program finds this J and increments the probability for cell J by p. This procedure is repeated for every combination of independent variables.

There are limitations on the number of random variables and the number of probability cells for each random variable. Up to ten independent random variables and ten probability values for each random variable are the upper limits of practicality. The cells for the final histogram result can be 100 at the maximum.

CONCLUSIONS

A method has been developed to propagate uncertainty in the form of DPDs (Ref. 5). The DPD method has been shown to be sufficiently accurate by comparing its results with analytical results. This method was compared with the Monte Carlo simulation and histogram propagation methods and produced similar results. A sensitivity analysis was performed with varying spreads of the basic event unavailability distributions and the DPD, Monte Carlo simulation, and method of moments compared for accuracy of results. The results indicate that the method of moments and Monte Carlo simulation give conservative results compared to the DPD method.

REFERENCES

1. S. Kaplan, "Handling Uncertainty in Reliability Decision Making," Workshop Proceedings: Power Plant Availability Engineering and Productivity Improvement, Research Project 771-6, Vol. 2, B. J. Garrick, Ed., Electric Power Research Institute (1979).

2. A. G. Colombo and R. J. Jaarsma, "A Powerful Numerical Method to Combine Random Variables," IEEE Trans. Reliab., R29, 2, 126 (June 1980).

3. S. Ahmed, D. R. Metcalf, and W. W. Weaver, "Probabilistic Risk Assessment Techniques for Nuclear Power Plants," BAW-1677 (Babcock & Wilcox Company, Lynchberg, Va., 1981).

4. C. Spetzler and C. A. S. Stael von Holstein, "Probability Encoding in Decision Analysis," Manage., 22 (Nov. 1975).

5. C. A. S. Stael von Holstein, "Assessment and Evaluation of Subjective Probability Distributions" (Economic Research Institute, Stockholm School of Economics, Stockholm, Sweden, 1970).

6. J. M. Hampton, P. G. Moore, and H. Thoms, "Subjective Probability and Its Measurement," J. R. Statistical Soc., 21, 136 (1973).

7. R. L. Winkler and A. G. Murphy, "Experiments in Laboratory and the Real World," Organ. Behav. Hum. Perform., 10, 252 (1973).

8. A. Tversky and D. Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," Science, 185, 1124 (1974).

9. A. D. Swain and H. E. Guttman, "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications," NUREG/CR-1278 (U.S. Nuclear Regulatory Commission, Washington, D.C., 1980).

10. G. E. Ingram, E. L. Welker, and C. R. Herrmann, "Designing for Reliability Based on Probabilistic Modeling Using Access Computer Systems," Proc. 7th Rel. and Main. Conference (San Francisco, California, 1968), p. 492.

11. "Reactor Safety Study," WASH-1400 (U.S. Nuclear Regulatory Commission, Washington, D.C., 1975).

12. A. G. Colombo and R. J. Jaarsma, "SCORE: A Computer Program for the Systematic Combination of Random Variables," EUR-6819EN (Commission of the European Communities, 1980).

13. S. Ahmed *et al.*, "PUFD--A Computer Program for the Propagation of Uncertainty Through Finite Probability Distribution," NPGD-TM-583 (Babcock & Wilcox Company, Lynchberg, Va., Aug. 1981).

SECTION IV. ON THE USE OF POINT ESTIMATES IN RISK ASSESSMENT OF NUCLEAR POWER PLANTS

On the use of point estimates in risk assessment of nuclear power plants

S. Ahmed

D. R. Metcalf

From Babcock & Wilcox Co., P.O. Box 1260, Lynchburg, Virginia 24505, and University of Virginia, Department of Nuclear Engineering and Engineering Physics, Reactor Facility, Charlottesville, Virginia 22901

Published in Reliability Engineering 6 (1983):211-227

ABSTRACT

Quantification of risk of large technological systems, such as nuclear power plants, is often done on the basis of point estimates. However, one source of confusion and controversy is the choice of point estimates for the basic event unavailabilities and the interpretation of the system unavailability thus obtained. It is shown through several examples that the system unavailability obtained by using median values of the basic event unavailabilities results in a very optimistic estimate and does not represent the central tendency of the TOP event probability distribution. On the other hand, the use of mean values of the basic event unavailability results in a conservative system unavailability.

INTRODUCTION

Point estimates of basic event unavailability are used in a number of recent plant risk studies [1-3] to evaluate system unavailability through fault tree or event tree modelling. Some basic concepts of point estimate and its relationship to the related probability distribution need to be clearly understood to derive a meaning from the calculated numerical value. This paper is devoted to clarifying various misconceptions on the use of point estimate values. Some suggestions are provided to quantify and propagate the point estimate of basic events to evaluate point estimate system unavailability.

One basic misconception is that the point estimate of the TOP event unavailability (system unavailability obtained, e.g., by probabilistic evaluations of a Boolean expression) obtained by using the median values of the basic events, is close to the median value. This point estimate is frequently interpreted as the median value and compared with the median unavailability of the corresponding system studied somewhere else, e.g., in Reactor Safety Study [4]. It is shown in this paper that the point estimate of the system unavailability thus obtained may be very optimistic, i.e., it may lie at the lower tail of the unavailability distribution of the TOP event, thus departing dramatically from the median unavailability obtained by propagating complete distributions.

Characteristics of the failure data for use in assessment are described in the next section. Quantifications of the TOP event unavailability by propagating uncertainty in the failure data and using various point estimates are described in the System Unavailability section. Several examples are given in the Examples section, and discussions and conclusions regarding the validity and usefulness of different point estimate evaluations are given in the Discussion and Conclusion section.

CHARACTERISTICS OF FAILURE DATA

Failure data available for quantification of the risk associated with large technological systems, such as nuclear power plants, are limited. For a plant-specific risk study, the limited failure history of its components leads to large uncertainty associated with the failure parameter of interest (failures per hour, or failures per demand). In the light of the limited data for the plant under study, failure data for similar components from elsewhere in the industry may be used to infer the failure rate of the specific component. In this case, plant-to-plant variability (due to different environmental conditions, manufacturers, and material properties) is a source of added uncertainty in the failure rate of the component. Lognormal distribution has been used for component failure rates in all plant risk studies to incorporate the uncertainty of orders of magnitude in the failure rate. The characteristics of lognormal distribution are that it is positively skewed, its higher tail may be orders of magnitude larger and the mean value is greater than the median, depending on the skewness of the distribution.

If μ and σ are the parameters of the lognormal probability density function,

$$f(\lambda) = \frac{1}{\sqrt{2\pi}\sigma\lambda} \exp\left[-(\ln\lambda - \mu)^2/2\sigma^2\right] \tag{1}$$

for the random variable λ, then it can be shown that

$$\mu = \ln(\lambda_{05} \cdot \lambda_{95})^{1/2} \tag{2}$$

and

$$\sigma = \frac{\ln \lambda_{95} - \mu}{1.645} \tag{3}$$

where λ_{05} and λ_{95} are the 5th and 95th percentile estimates, respectively, of the failures per unit time or the failures per demand of the component of interest. If we define the dispersion in the data for lognormal distribution in terms of the range factor (RF) as

$$RF = \frac{\lambda_{95}}{\lambda_{50}} = \frac{\lambda_{50}}{\lambda_{05}} \tag{4}$$

then,

$$\lambda_{50} = \sqrt{\lambda_{05}\lambda_{95}} \tag{5}$$

where λ_{50} is the median value. The mean is given by

$$\text{mean} = \lambda_{50} \cdot e^{\sigma^2/2} \tag{6}$$

There are various ways of estimating λ_{05} and λ_{95}; thus, using various data sources, one can assess the λ_{05} and λ_{95} directly [5]. However, in cases where it is more convenient to obtain λ_{50} and a RF, as in the Reactor Safety Study, λ_{50} and λ_{95} can be obtained from equation 4. Parameters of the lognormal distribution μ and σ are then evaluated from equations 2 and 3.

Mean and Median

As mentioned earlier, for a positively skewed distribution (such as a lognormal distribution), the mean is larger than the median. The relationship between the mean and the median depends only on the RF of the lognormal distribution and is given by

$$\frac{\text{mean}}{\text{median}} = \exp\left[\frac{\ln^2(\text{RF})}{5.412}\right] \tag{7}$$

Figure 1 is a plot of equation 7. We notice that the ratio of mean to median increases exponentially with the increase in RF, which represents the spread in the basic event unavailability. For further clarity, graphic representations of the locations of the mean values of the lognormal probability density functions for a fixed median value and varying RFs are given in Figure 2. Here, four distributions are shown on a lognormal plot with the median unavailability 5×10^{-4} and RFs of 3, 10, 25 and 50. We see clearly that with the increase in RF, the mean values depart increasingly from the median and shift toward the higher tail of the unavailability distribution. Therefore, for higher RFs, mean values are not representative of the central tendency of the lognormal distribution. On the other hand, median values, from their definition, do represent the central tendency of the distribution.

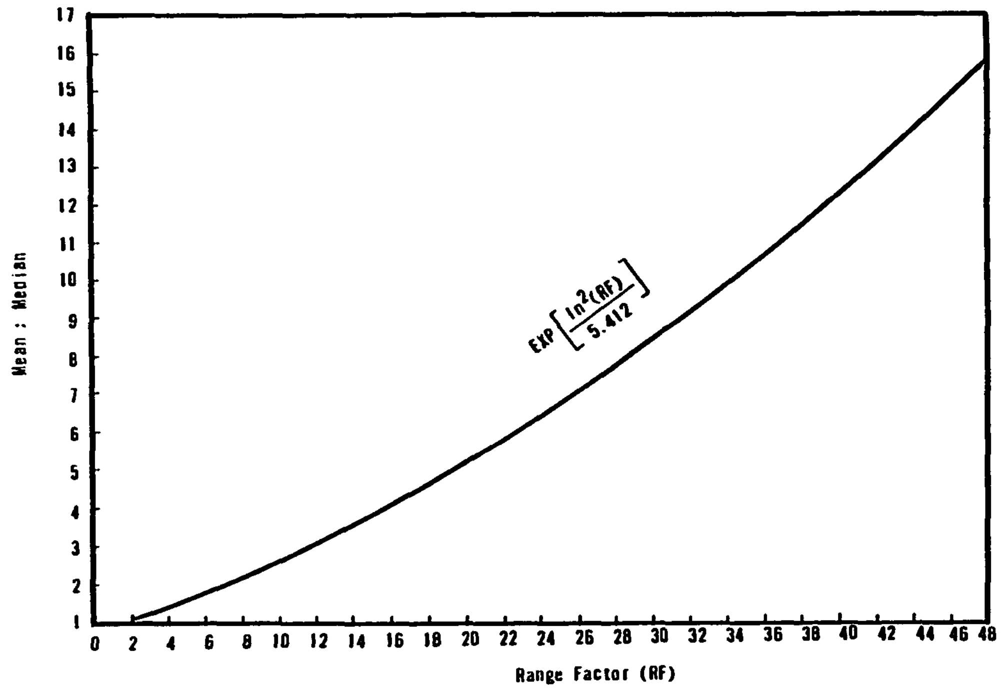


Figure 1. Relationship between mean and median of lognormal distribution with varying range factors

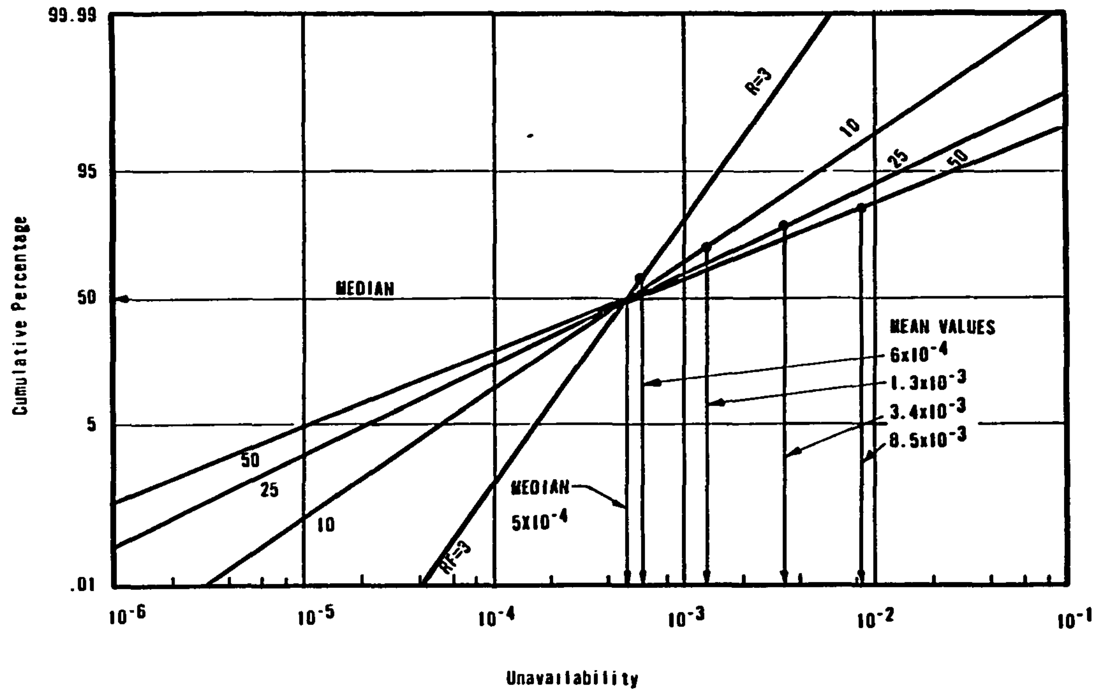


Figure 2. Mean values of lognormal distributions with fixed median and varying range factors

SYSTEM UNAVAILABILITY

To obtain the system unavailability, a logic model--such as a fault tree or reliability block diagram--is first constructed. A Boolean expression of the system failure is then obtained, which delineates qualitatively the various ways the system can fail. The Boolean expression can then be converted to a probability statement using the calculus of probability. For example, if X represents the event 'failure of valve x', and Y represents the event 'failure of valve y', and if x and y both have to fail for the system (represented by TOP) to fail, then the Boolean expression can be written as

$$TOP = X \cdot Y \ . \tag{8}$$

The corresponding probability statement is

$$\begin{aligned} P(TOP) &= P(X \cdot Y) \\ &= P(X)P(Y/X) \end{aligned} \tag{9}$$

If X and Y are independent, then

$$P(TOP) = P(X)P(Y) \ . \tag{10}$$

On the other hand, if the system fails when <u>either</u> X or Y fails, then

$$TOP = X + Y \tag{11}$$

and the corresponding probability statement is

$$P(TOP) = P(X + Y)$$

$$= P(X) + P(Y) - P(X \cdot Y) \quad . \tag{12}$$

Again, if X and Y are independent, then

$$P(TOP) = P(X) + P(Y) - P(X)P(Y) \quad . \tag{13}$$

In risk assessment, the failure probabilities are generally small and the product term in equation 13 is negligible. Thus, we can write

$$P(TOP) = P(X) + P(Y) \quad . \tag{14}$$

The Boolean expression obtained for a system failure such as that in a nuclear power plant may be very complex, and the corresponding probability statement in the exact form is difficult to obtain. Fortunately, simplifications like those represented in equation 14 introduce negligible error and can be used safely. We would now like to obtain the distribution P(TOP) by propagating the basic event distributions, such as P(X) and P(Y), in the equations above that will give us the uncertainty in the system unavailability [6]. For comparison, we shall also obtain the system unavailability P(TOP) by using point estimates of the basic event unavailability. Two different point estimates of P(TOP) will be obtained by using (1) expected or mean values of the basic event unavailabilities and (2) median values of the basic event unavailabilities. In case (1), it can be shown that, for equation 10,

$$E(TOP) = E(X)E(Y) \tag{15}$$

and for equation 14,

$$E(TOP) = E(X) + E(Y) \tag{16}$$

where E(*) represents the expected or the mean value of the variable (*). Thus, in a Boolean expression, the mean values used for the basic event unavailability can be propagated to obtain the mean value of the TOP event. Unfortunately, this simplicity is lost when median unavailability is used as the basic event point estimate. That is, we cannot propagate the median unavailability in a Boolean expression to obtain the median unavailability of the TOP event unless the Boolean expression is in the multiplicative form only.

EXAMPLES

Here, a number of Boolean expressions are chosen to represent the TOP event, and their unavailabilities are evaluated by propagating the basic event unavailabilities. The TOP event unavailability is obtained by propagating the basic event unavailability distribution by using Monte Carlo to include uncertainty. The SAMPLE computer code, which was also used in the RSS study, is used here for uncertainty propagation. Mean and median values of the basic events are also used to arrive at the point values of the system unavailability.

Example 1. Simple Boolean Equation and Small Uncertainty in Basic Event Unavailability

The following Boolean expression, with four minimal cut-sets, is chosen:

$$TOP = X_1 + X_2X_3 + X_4X_5 + X_6 \tag{17}$$

where X_1, X_2, ..., X_6 are the basic event unavailabilities, and TOP represents the system unavailability. The basic event unavailability data are obtained in terms of median and RF. For this example, the RF of each basic event is 3. The parameters of the lognormal distributions μ and σ are obtained from equations 2-4, and the mean is obtained from equation 6 (Table 1).

The probability distribution of the TOP event unavailability is obtained by propagating the uncertainty of the basic events; the various percentile points are as follows:

Table 1. Basic event unavailability parameters for Example 1

Component	Median	RF	μ	σ	Mean
X_1	0.001	3	-6.9078	0.6678	0.00125
X_2	0.03	3	-3.5066	0.6678	0.0375
X_3	0.01	3	-4.6052	0.6678	0.0125
X_4	0.03	3	-3.5066	0.6678	0.0375
X_5	0.01	3	-4.6052	0.6678	0.0125
X_6	0.003	3	-5.8091	0.6678	0.00375

5th percentile = 2.60×10^{-3}

median = 5.30×10^{-3}

mean = 5.96×10^{-3}

95th percentile = 1.10×10^{-2}

The point estimate of the TOP event obtained by using median values from Table 1 in (Boolean) equation 17 is 4.6×10^{-3}. The locations of the point estimate and the cumulative probability distribution of the TOP event for this example are shown in Figure 3. We notice that the point estimate actually represents the 36th percentile unavailability of the TOP event. The mean unavailability of the TOP event obtained by using the mean unavailability of the basic event is 5.94×10^{-3}, which, as expected from our discussions in the previous section, is about the same as obtained from the uncertainty propagation. The slight deviation in the mean obtained from Monte Carlo propagation may be attributed to the sampling error.

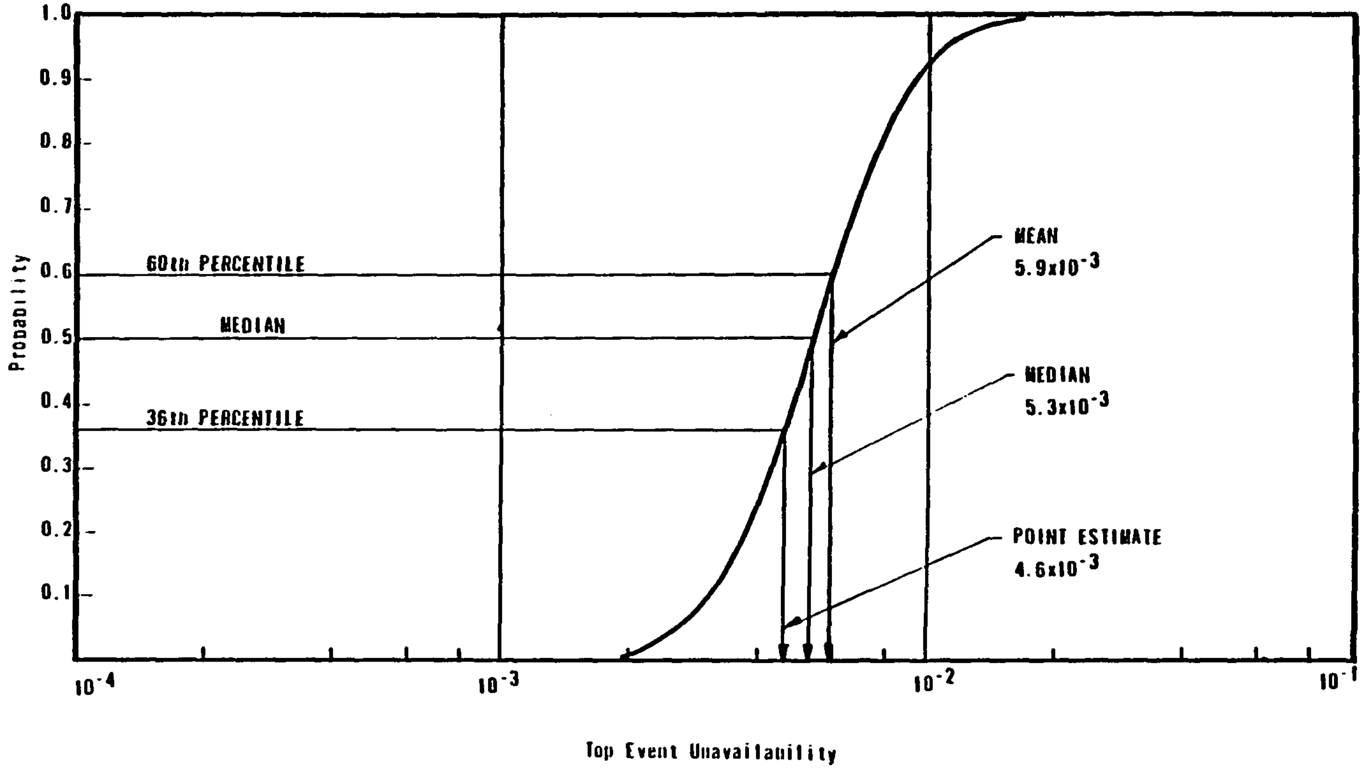


Figure 3. Cumulative probability distribution for TOP event unavailability for Example 1

Example 2. Simple Boolean Equation and Large Uncertainty in Basic Event Unavailability

The Boolean expression from Example 1 is used here except that the RF for each basic event unavailability is changed from 3 to 10. The parameters μ, σ, and the mean for the basic event unavailability distributions are given in Table 2. The TOP event unavailability distribution is obtained by using Monte Carlo simulation; the various percentile points are as follows:

$$\text{5th percentile} = 1.9 \times 10^{-3}$$
$$\text{median} = 8.5 \times 10^{-3}$$
$$\text{mean} = 1.4 \times 10^{-2}$$
$$\text{95th percentile} = 4.4 \times 10^{-2} \ .$$

The point estimate of the TOP event obtained by using median values for Table 2 in (Boolean) equation 17 is 4.6×10^{-3}. The location of the point estimate actually represents the 25th percentile unavailability of this example, as shown in Figure 4. We notice from the figure that the point

Table 2. Basic event unavailability parameters for Example 2

Component	Median	RF	μ	σ	Mean
X_1	0.001	10	-6.9078	1.3997	0.002663
X_2	0.03	10	-3.5066	1.3997	0.0799
X_3	0.01	10	-4.6052	1.3997	0.02663
X_4	0.03	10	-3.5066	1.3997	0.0799
X_5	0.01	10	-4.6052	1.3997	0.02663
X_6	0.003	10	-5.8091	1.3997	0.00799

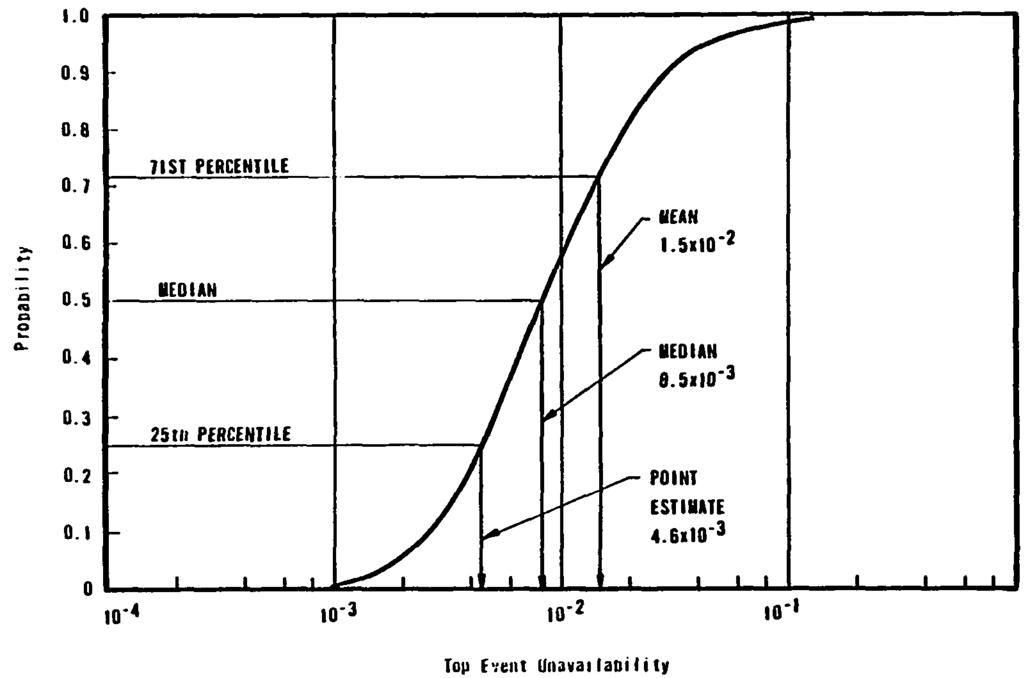


Figure 4. Cumulative probability distribution for TOP event unavailability for Example 2

estimate actually represents the 25th percentile unavailability of the TOP event, and the mean value represents 71st percentile of the TOP event unavailability.

Example 3. Moderately Complex Boolean Equation

The following Boolean expression, with 16 minimal cut-sets, is chosen:

$$\begin{aligned} TOP = {} & AX_2 + AK + AE + X_2X_3 + X_3K + DX_2 + DK + EX_3 + BX_1 \\ & + X_1X_4 + CX_1 + FB + BL + CL + FX_4 + X_4L \ . \end{aligned} \tag{18}$$

The parameters of the basic event unavailability distributions are given in Table 3. The various percentile points for the TOP event unavailability distribution obtained by Monte Carlo are as follows:

5th percentile = 2.2×10^{-4}

median = 8.9×10^{-4}

mean = 1.5×10^{-3}

95th percentile = 7.3×10^{-3}

The cumulative unavailability distribution of the TOP event is shown in Figure 5 along with the mean value and the point estimate obtained by using median values from Table 3. We notice that the point estimate and the mean represent the 15th and 72nd percentile values, respectively, of the TOP event unavailability distribution.

Table 3. Component unavailability for Boolean equation 18

Component	Median	RF	μ	σ	Mean
A	0.00553	8.10	-5.198	1.272	0.0124
X_2	0.00553	8.10	-5.198	1.272	0.0124
K	0.00553	8.10	-5.198	1.272	0.0124
E	0.00283	1.96	-5.868	0.409	0.00295
X_3	0.00553	8.10	-5.198	1.272	0.0124
D	0.00283	1.96	-5.868	0.409	0.00295
B	0.00553	8.10	-5.198	1.272	0.0124
X_1	0.00553	8.10	-5.198	1.272	0.0124
X_4	0.00553	8.10	-5.198	1.272	0.0124
C	0.00283	1.96	-5.868	0.409	0.00295
F	0.00283	1.96	-5.868	0.409	0.00295
L	0.00553	8.10	-5.198	1.272	0.0124

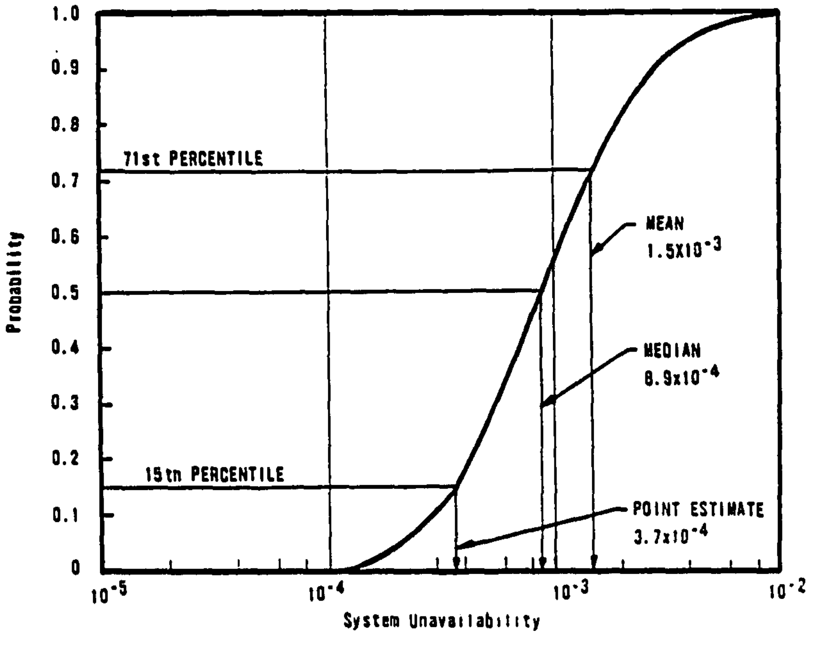


Figure 5. Cumulative probability distribution for TOP event unavailability for Example 3

Example 4. Complex Boolean Expression

The following Boolean expression, with 86 minimal cut-sets, is chosen:

$$
\begin{aligned}
TOP = {} & AB + AC + BC + AD + BD + CD + BE + CE + DE + AF \\
& + CF + DF + EF + AH + BH + DH + EH + FH + AI \\
& + BI + CI + EI + FI + HI + JD + JI + KC + KH + LB \\
& + LF + KL + MA + ME + JM + ND + NI + MN + PC \\
& + PH + LP + QB + QF + KQ + PQ + RA + RE + JR \\
& + NR + SB + SC + SD + SF + SH + SI + MS + RS + UA \\
& + UC + UD + UE + UH + UI + LU + QU + SU + VA \\
& + VB + VD + VE + VF + VI + KV + PV + SV + UV \\
& + WA + WB + WC + WE + WF + WH + JW + NW + SW + UW + VW \ . \qquad (19)
\end{aligned}
$$

The parameters of component unavailability distributions are shown in Table 4. The various percentile points for the TOP event unavailability distribution obtained by Monte Carlo are as follows:

Table 4. Component unavailability for Boolean equation 19

Component	Median	RF	μ	σ	Mean
A, B, C, D, E, F, H, I, J, K, L, M, S, U, V, W	0.00553	8.10	-5.198	1.272	0.0124
N, P, Q, R	0.00283	1.96	-5.868	0.409	0.00295

5th percentile = 2.3×10^{-3}

median = 7.7×10^{-3}

mean = 1.1×10^{-2}

95th percentile = 3.1×10^{-2} .

The cumulative unavailability distribution of the TOP event is shown in Figure 6 along with the mean value and the point estimate obtained by using median values from Table 4.

We notice from Figure 6 that the point estimate and the mean represent the 6th and 70th percentile values, respectively, of the TOP event unavailability.

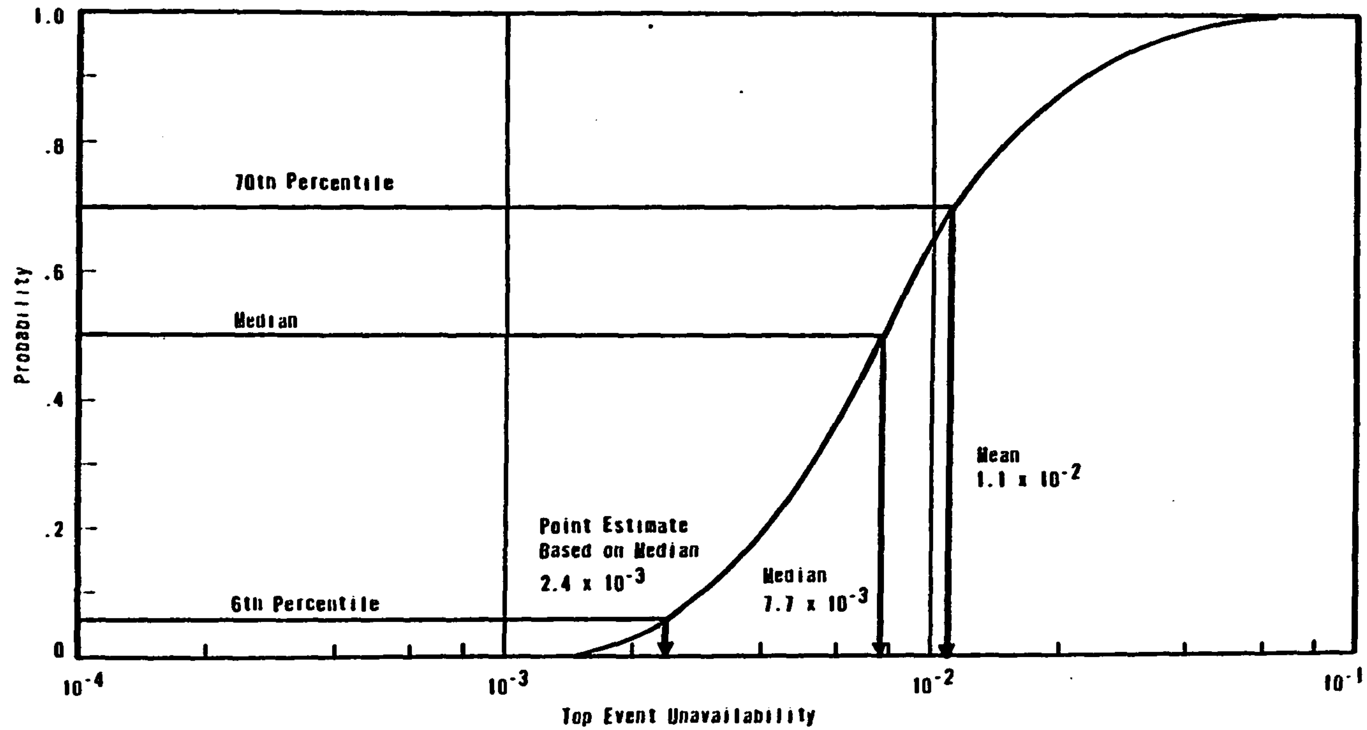


Figure 6. Cumulative probability distribution for TOP event unavailability for Example 4

DISCUSSION AND CONCLUSION

We would now determine whether the point estimates are representative of the central tendency of the TOP event distribution obtained by uncertainty propagation.

A measure of the central tendency is the cumulative probability associated with the point estimate values. The case studies presented in the Examples section range from simple Boolean expressions with four minimal cut-sets to a complex one with 86 minimal cut-sets. In all the cases, the point estimates based on the median values of the basic events are lower than the median obtained by uncertainty propagation. The point estimate shifts toward the lower tail of the distribution as the uncertainty in the basic event unavailability (represented by the range factor) increases or as the number of cut-sets increases. For the Boolean expression with 16 minimal cut-sets, the point estimate represents the 15th percentile of the TOP event distribution, and for the case with 86 cut-sets it represents only the 6th percentile.

The point estimate obtained by using values of the basic event is the same as that obtained by uncertainty propagation. However, the mean values are higher than the median and represent about 70th percentile unavailability for the case studies.

Based on the discussion above, we conclude that median values of the component failure data should not be used in risk assessments based on point estimates. The TOP event unavailability thus obtained can represent much lower values; i.e., the system would appear much better than it actually is. Therefore, special care is warranted in the choice and analysis of component failure data, which generally have positively skewed distribution with a mean value greater than the median. It is the mean value of the component failure that should be used for quantification of the TOP event unavailability. This gives the mean value of the system unavailability. When comparing it with the RSS, it is again the mean value that should be compared.

REFERENCES

1. Garcia, A. A. et al. Crystal River 2 Safety Study (Draft), SAI-002-80BE (Science Applications, Inc., Bethesda, Maryland, May 1980).

2. Carlson, D. D. et al. Reactor Safety Study Methodology Applications Program: Sequoyah 1 PWR Power Plant, NUREG/CR-1659 (US Nuclear Regulatory Commission, Washington, D.C., February 1981), No. 1 of four.

3. Kolb, G. J. et al. Reactor Safety Study Methodology Applications Program: Oconee 3 PWR Power Plant, NUREG/CR-1659 (US Nuclear Regulatory Commission, Washington, D.C., May 1981), No. 2 of four.

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

5. Ahmed, S., Metcalf, D. R. and Weaver, W. W. Probabilistic Risk Assessment Techniques for Nuclear Power Plants, BAW-1677 (Babcock & Wilcox, Lynchburg, Virginia, October 1981).

6. Ahmed, S., Metcalf, D. R. and Pegram, J. W. "Uncertainty propagation in probabilistic risk assessment: a comparative study." Trans. Am. Nucl. Soc., 38, 483 (1981).

SECTION V. SENSITIVITY ANALYSIS APPROACH FOR ROBUST PROBABLISTIC RISK ASSESSMENT

SECTION V. SENSITIVITY ANALYSIS APPROACH FOR ROBUST PROBABLISTIC RISK ASSESSMENT

Shahid Ahmed

from the

Franklin Research Center
Division of Arvin/Calspan
20th and Race Streets
Philadelphia, PA 19103

INTRODUCTION

In Sections III and IV, certain aspects of risk quantification and its associated uncertainty have been described. The uncertainty propagation method in Section III combines the uncertainties in the failure probabilistics of all the basic events to quantify the uncertainty in the Top event (or system failure) frequency. However, as we gain nuclear power plant operating experience, the component data base will improve, and the uncertainties in the component failure rates are expected to decrease. Thus, in the light of more experience, the Top event failure probability distribution will be modified.

Since experience with nuclear power plant (NPP) operations is limited, adequate failure histories of safety systems and the associated components are not available. The uncertainty in the component failure rates (as well as human error and common cause failure probabilities) is large. With large uncertainty in the basic parameters, the plant risk results obtained by using recent PRA methods may be less credible, because of lack of any sensitivity analysis of the uncertain parameters.

Lacking a satisfactory data base, however, we need to devise methods that will identify the components that contribute most to the Top event failure probability and to the associated uncertainties. A particular difficulty is encountered when a risk being evaluated has never been realized, e.g., large scale radioactivity release due to an accident in a nuclear power plant. In such cases, quantification of risk is highly dependent upon the adequacy of system modeling, the selection of all the significant ways in which accidents may happen, and the availability of failure data for the basic events (e.g., component failures and human errors).

One serious drawback in PRA for nuclear power plants (NPPs) is that the results cannot be verified due to the absence of events and thus of data (e.g., frequency of core-melt). For the results to be acceptable and trustworthy, a plant specific PRA should therefore, produce results that are robust (i.e., insensitive) to various modeling assumptions and to the data base. Although a number of plant PRAs have been performed, sensitivity analysis or robustness of the results are generally not addressed.

In this section we will first explore a Bayesian approach to analyzing the component failure rate data. The purpose of such an analysis will be to determine the sensitivity of the posterior failure rate distribution of a component based on the prior (or generic) failure rate distribution and the plant specific failure data.

Having developed the failure rate data base, we shall next develop a sensitivity analysis approach to identify and rank those components that contribute significantly to the failure probability and to the uncertainty associated with the failure probability. The basis of determining the relative ranking of the basic events pertaining to the NPP risk is the sensitivity of the Top event. The more sensitive the Top event is to a basic event or failure element, the greater is its risk contribution. A method is developed to quantify the sensitivity of the Top event by defining three different safety Significance Indices which assess the risk contribution of a basic event. The information obtained through the sensitivity analysis comprises a priority ranking of the basic events both with respect to the Mean value and the Variance of the failure probability.

The methodology produces, as an additional product, the Mean and Variance of the Top event. Furthermore, the methodology has the capability to incorporate the variations in the basic event unavailability distribution and to establish the impact of such variations on the Top event. These variations in the probability distribution can be introduced using the plant specific failure data through Bayesian updating model discussed on page 164. We will also devise an approach that will delineate the robustness of significance indices, and present a method to perform sensitivity analysis using component significance indices. Examples will be presented to demonstrate the applications of the method.

DATA ANALYSIS OF BASIC EVENTS

Introduction

Bayes' theorem is being extensively used in recent plant PRAs [1-4] where quantitative assessment of a number of basic parameters (e.g., basic event unavailabilities) are made using subjective judgment. Although the use of expert opinion is unavoidable in PRA studies, and subjective interpretation of probability is being suggested as an appropriate format for analyzing risk from NPPs [5-10], the robustness of the PRA results is of special interest in this context. The conclusions of a plant PRA are considered robust if they do not change with alterations in the subjective formulations of the data base and with physical modeling used in the study. If the PRA results are unstable, i.e., they change with the subjective assessment of various parameters, then the conclusions are suspect, and require further scrutiny to assure their validity.

The robustness criteria are of particular interest in preparing a failure rate data base for plant risk studies. In the absence of adequate failure data for a particular component (e.g., a valve) or an event (e.g., loss-of-offsite-power) for a specific NPP, a generic failure rate distribution, based on expert opinion and available field data, is subjectively assessed that encompasses different plants. This distribution, called the population variability curve of the component (explained further in the next section), is updated by plant experience data, to develop a plant specific failure rate distribution (or the posterior distribution). To qualify the posterior distribution as valid, the following robustness criteria are desirable:

i. The posterior distribution should be insensitive to the spread of the population variability curve.

ii. The posterior distribution should be insensitive to the type (lognormal, beta, or gamma) of population variability curve.

The failure rate data base for a plant PRA that has the above two characteristics is defined here to be "a robust data base."

Population Variability Curves

The major data sources available for component failure and human errors are U.S. Nuclear Regulatory Commission [11], IEEE Guide to the Collection and Presentation of Electrical, Electronic and Sensing Component Reliability Data for Nuclear Power Generation Stations [13], Swain and Guttman [14], and Ahmed et al. [15]. Other data sources can also be used to supplement and enhance the quality of data for specific needs. However, a clear interpretation related to the physical process of data accumulation from the data sources in use is necessary.

For a plant specific risk study, let us suppose that we are interested in the failure rates of a number of components. If the plant has been in operation for several years, we will have the failure history of the components. However, this failure history may be insufficient, or nonexistent, if the plant has not started operation. A prediction of the component failure rates therefore, would be incomplete, and sometimes impossible to make, on the basis of plant experience itself. However, we have failure data available from the operating history of other nuclear power plants, for components of similar type (either from the same manufacturer or different manufacturers), and working under similar or different environments. Also, data on similar components are available from other related industries, such as fossil fueled electric power plants. This raises the question of uncertainty due to inadequate plant specific data. To put things into perspective, let us consider the following hypothetical experiments.

Experiment I

Let us consider m similar components (e.g., a pump) from the same population. We can define the population in several ways, such as (1) pumps from the same manufacturers, (2) similar pumps from different manufacturers but operating in a similar environment, e.g., in a power plant where it is subject to similar test and maintenance procedure, and operated by the same group of personnel. Whatever is the definition of the population, let us suppose we have the history of the time of failure of each of the m pumps. We can now obtain the mean-time-between-failure (MTBF) for each pump. Let us consider here that the end of the life MTBF of each pump is the same, and a constant for the population under study. However, the observed data being incomplete (not end of life), MTBF for each pump will be different, and will form a probability distribution. Considering time to failure as exponentially distributed, the inverse of the MTBF is the parameter of the distribution which is also the failure rate or hazard function of the sample. This is then the generic failure rate distribution for the population, and our prior distribution for the population. This prior distribution can now be revised in the light of plant specific experience, i.e., the failure history of the pump for which we need the failure rate distribution.

Experiment II

Let us consider n different populations, each population having m similar components of interest (again, consider a particular type of pump) under study. The discussion in Experiment I is now valid for each of the n populations. We would now like to construct a generic distribution of the failure rate of the pump considering the failure history of all n populations. By extension of the logic in Experiment I, we would have the failure rate distribution of each population. It now remains to construct a composite prior distribution which accounts for the variability of n populations. Once this composite prior distribution is constructed, it can then be updated by the plant specific evidence of the pump failure, by employing Bayesian models.

However, the prior distribution that encompasses population variability is wider than the distribution of any one of the n populations.

The above two postulated experiments are fundamental in relating the physical process of data accumulation to the statistical analysis that would follow. Here, we shall consider WASH-1400 as a source to derive a rationale for data analysis.

The failure data base for the RSS included a number of sources that encompassed data from the nuclear industry (at that time the experience was 300 reactor-years) as well as other industries. Best estimate data was obtained for each source for the failure mode of the component of interest. Consider as an example, failure data for Motor Operated Valve (MOV) in Table III 2-1, Appendix III, WASH-1400 [11]. Point estimate data are given from different sources for various failure modes, such as failure to operate. From the data base, a lower 5th percentile and upper 95th percentile bound were assessed, and a median value obtained considering the parameter to be lognormally distributed.

The data base thus prepared for the risk study was considered to represent the uncertainty and the population variability of the parameter. To interpret the data we go back to Experiment II described earlier. Each of the data sources shown in the above cited table can be defined to represent a population, for which a point estimate represents a population, and for which a point estimate representing the central tendency is given. It is not clear if the individual point estimates represent the median or the mean of the population. However, a probability distribution based on the point estimates are, in fact, distributions of the mean or median of the parameter, and not of the parameter itself. In risk studies, the parameters of interest are the failure rate or failure per demand, and we would like to have a probability distribution which would include uncertainty due to insufficient experience data (Experiment I) as well as the population variability (Experiment II). The lower and the upper bounds assessed in the RSS are, therefore, those of the distribution of mean or median of the failure per hour or failure per demand.

It now remains to infer the failure rate distribution from the distribution of mean or median failure rate. Unfortunately, it is not so straightforward using the RSS data base. We do not know the failure history of the components for the populations, the size of each population, etc. However, we know that the distribution of the parameter is much more spread out than the mean of the parameter, as shown in Figure 1. Therefore, to obtain the desired distribution, we need to spread out the distributions given in the RSS. The question that arises is "how much should we spread the distributions given in the RSS so that it would reasonably represent the distribution of the population?" The answer to this question can be found empirically by looking into the following implications of spreading out the failure distributions:

1. RSS fits a lognormal distribution to the lower 5th percentile and the upper 95th percentile values. We shall investigate other fits such as gamma and beta distribution to notice the impact on the general shape of the distribution.
2. The distributions in the RSS will be spread out and new values will be assigned to the lower 5th percentile and the upper 95th percentile points. Various different spreads will be considered; the RSS 90 percent confidence interval will be considered to represent 70, 60, and 50 percent confidence intervals for the distributions representing population variability of the parameters. Thus, the 5th and 95th (5/95) percentile points in RSS will represent 15th and 85th (15/85) percentile points of the population variability curve to represent 70 percent confidence interval, 20th and 80th (20/80) percentile points to represent 60 percent confidence interval, and 25th and 75th percentile points (25/75) to represent 50 percent confidence interval. For each of these new confidence interval lognormal, gamma, and beta distribution fits will be considered in order to investigate the impact of the choice of the prior distributions.

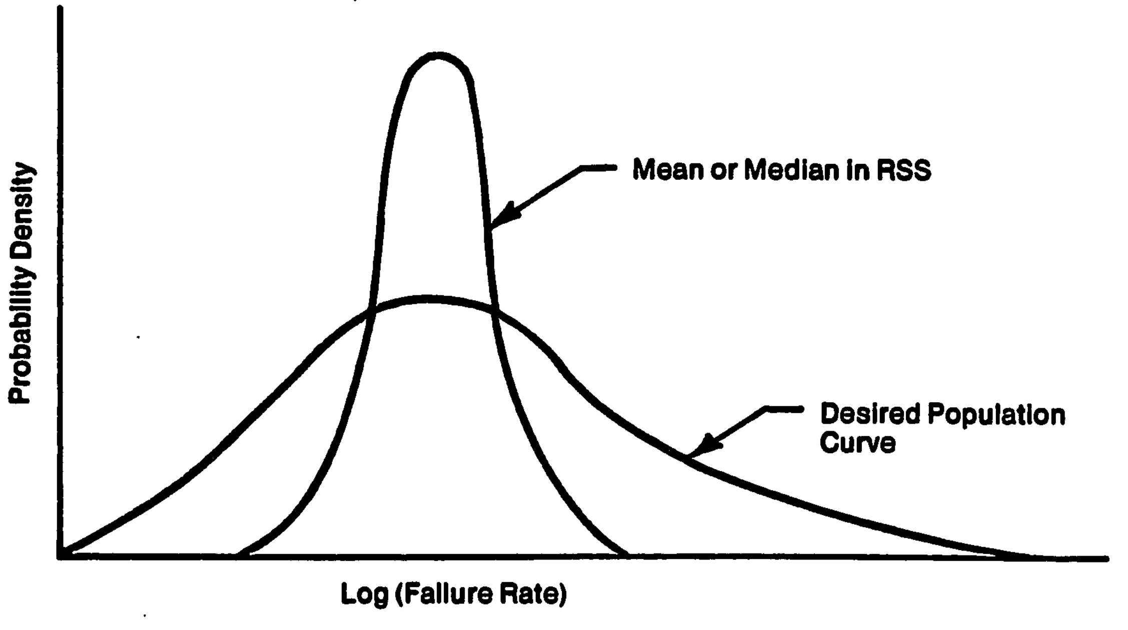


Figure 1. Distribution of the population failure rate vs. mean failure rate

3. Various prior distributions considered above will be updated by the plant specific information. Thus, the impact of the choice of the prior distribution representing various spreads upon the plant specific posterior distribution will be assessed.

Bayesian Models

Bayes theorem as an equation is written as

$$f(\lambda/E) = \frac{f\ (\lambda)\ L(E/\lambda)}{\int_0^\infty f(\lambda')\ L(E/\lambda')d\lambda'} \tag{1}$$

where:

$f(\lambda/E)$ = probability density function of λ given the evidence E, called the posterior distribution,

$f(\lambda)$ = probability density function of λ, called the prior distribution,

$L(E/\lambda)$ = probability of evidence E given the failure rate λ, called the likelihood function.

The likelihood function normally has two forms depending on the data base. If we have a data base of r failures in n trials, the likelihood function is the binomial distribution:

$$L(E/\lambda) = \left[{n \atop r}\right]\lambda^r\ (1-\lambda)^{n-r} \tag{2}$$

where:

$\left[{n \atop r}\right] = n!/[(n-r)!r!]$.

In the other case where the data base is given as r failures over a fixed operational time T, the likelihood function is

$$L(E/\lambda) = \frac{e^{-\lambda T} (\lambda T)^r}{r!} \qquad \text{(Poisson distribution)} \qquad (3)$$

The integration in the denominator of equation 1 can be done analytically if we use conjugate functions.

<u>Gamma prior distribution</u>

If the prior distribution is modeled as a gamma density function of the form

$$f(\lambda) = \frac{\beta^{\alpha} \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \qquad (4)$$

where the parameters β and α are fixed by knowing the 5 and 95 percentiles for λ; i.e., β and α are found by solving the two equations:

$$0.05 = \int_{o}^{\lambda_{05}} f(\lambda)d\lambda$$

and

$$0.95 = \int_{o}^{\lambda_{95}} f(\lambda)d\lambda$$

we now insert equations 3 and 4 into equation 1 to obtain the posterior function

$$f(\lambda/E) = \frac{(\beta+T)^{\alpha+r} \lambda^{\alpha+r-1} e^{-(\beta+T)\lambda}}{\Gamma(\alpha+r)} \qquad (5)$$

We note that the posterior density function given by equation 5 can be obtained from equation 4 by replacing β with β+T and α with α+r.

For the case of r failures in n trials where the failure rate (in failures per demand) is quite small, the binomial distribution can be approximated by the Poisson distribution. The final result can be obtained by replacing T in equation 3 with n to obtain the likelihood function

$$L(E/\lambda) = \frac{e^{-n\lambda}\,(n\lambda)^{r}}{r!} \tag{6}$$

where the units for λ are failures per demand. This result is then used in equation 1 where the limits on the integral in the denominator are zero to one. For cases where most of the distribution is clustered near zero, the value of the integral is mainly determined by integrating from zero to one. We can obtain (with some slight error) the same form as equation 5 if we extend the integration from zero to infinity. The final result for the posterior density function is

$$f(\lambda/E) = \frac{(\beta+r)^{\alpha+r}\;\;{}^{\alpha+r-1}\;e^{-(\beta+r)\lambda}}{\Gamma\,(\alpha+r)} \tag{7}$$

<u>Beta prior distribution</u>

A more exact model for the failures per demand case is to use the beta distribution for the prior distribution. The normal range for the variable is then zero to one and this is given by

$$f(\lambda) = \frac{\lambda^{a-1}\,(1-\lambda)^{b-1}}{B(a,b)}\,, \quad 0 \le \lambda \le 1 \tag{8}$$

where a and b are parameters to be determined in a manner similar to the determination of β and α for the gamma distribution. The denominator in equation 8 is defined as

$$B(a,b) = \frac{\Gamma(a)\ \ \Gamma(b)}{\Gamma(a+b)} \tag{9}$$

If equation 8 is now used in equation 1 along with the binomial likelihood function (equation 2), the integral in the denominator can be explicitly performed giving the posterior density function:

$$f(\lambda/E) = \frac{\lambda^{a+r-1}\ (1-\lambda)^{b+n-r-1}}{B(a+r,\ b+n-r)} \tag{10}$$

We note again that this posterior function, equation 10, can be obtained from equation 8 by replacing a with a+r and b with b+n-r.

Lognormal prior distribution

The lognormal prior density function is given by

$$f(\lambda) = \frac{1}{\sqrt{2\pi}\ 5\sigma\lambda} e^{-\frac{1}{2}\left[(\ln\ \lambda-\mu)/\sigma\right]^2} \tag{11}$$

One advantage of this density function is that the parameters λ and σ are easily determined from the λ_{05} and λ_{95} values as follows:

$$\mu = \frac{1}{2}\ \ln\ (\lambda_{05}\lambda_{95}) \tag{12}$$

and

$$\sigma = \frac{1}{3.29}\ \ln \frac{\lambda_{95}}{\lambda_{05}} \tag{13}$$

If we consider the case of λ failures per demand, then the likelihood function is given by equation 6, and the posterior density function is

$$f(\lambda/E) = \frac{\lambda^{r-1} e^{-\lambda T - 1/2 \, [\ln \lambda - \mu)/\sigma]^2}}{\int_0^\infty (\lambda')^{r-1} e^{-\lambda' T - 1/2 \, [\ln \lambda' - \mu)/\sigma]^2} d\lambda'} \tag{14}$$

An immediate disadvantage of this method is that the denominator in equation 14 cannot be performed explicitly but requires a numerical integration.

Example

A number of case studies were conducted to obtain generalized results for deriving plant specific population variability curves. For better understanding an example calculation will be shown here to illustrate the procedure.

The 5th and 95th percentile values from RSS for MOV failure to operate (3×10^{-4}/D and 3×10^{-3}/D, respectively) are used to fit lognormal, gamma and beta distributions. Those distributions are further expanded such that the 5th/95th percentile values represent 15/85, 20/80, and 25/75 percentile values for each type of distribution; these are the set of prior distributions shown in Figure 2, which are made plant specific by the use of plant specific history of MOV failures by using Bayes' theorem. The lognormal priors are sharply peaked while the gamma and beta priors are diffused. Also, the gamma and beta priors are almost identical. Postulated plant specific evidences are now used to ascertain the sensitivity of the posterior distributions. The class of evidences found useful are 5 failures in 1000 demands (5/1000), (1/100), and (1/10). The conjugate likelihood functions used to arrive at the plant specific posterior distributions are Poisson distribution for the lognormal and gamma priors, and binomial likelihood for beta priors. The Bayes' theorem formulation and the associated computer code BURD (Bayesian Updating of Reliability Data) used here are given in References [15] and [16].

The resultant posterior distributions related to lognormal and gamma priors are shown in Figures 3 through 6. The posterior distributions for the beta priors are almost identical with the gamma posteriors, and are not presented here. To facilitate comparison, the mean and median values for all the prior and posterior distributions shown in Figures 3 through 6 are provided in Table 1.

Results

Strong evidence

For a standby component the evidence can be said to be strong if the specific experience in the neighborhood of 1000 demands or more are available. As can be seen in Figure 3 the posterior distribution with strong evidence is insensitive (robust) with respect to the spread (5/95, 15/85, or 25/75) and type (lognormal, gamma, or beta) of the prior distribution. However, gamma prior with the conjugate Poisson likelihood function is found to be preferable for the following reasons:

- Mathematical convenience for Bayesian updating. Here the posterior is also a gamma distribution with revised parameters.
- Computational ease of computer evaluation of the parameters of gamma distribution given the 5th and 95th percentile values.

- Beta posterior distributions obtained by using conjugate beta priors and binomial likelihood functions are very similar to the conjugate gamma posteriors for the range of failure rates encountered in risk assessment (computationally parameters of beta distributions are difficult to obtain, given the 5th and 95th percentile values).
- Shape of gamma distributions are more flexible and diffused as compared to lognormal distributions which have sharper peaks for the same 5th and 95th percentile values.

Moderately strong evidence

The failure history for a component can be said to be moderately strong if the evidence (number of failures) in the neighborhood of 100 demands or more (but much less than 1000 demands) are available. As can be seen in Figure 4, the posterior distribution has the following two characteristics:

1. It is moderately sensitive (less robust as compared to the case of strong evidence) to the spread of the prior distribution.
2. It is less sensitive to the type of prior distribution.

As observed earlier, we notice that lognormal priors produce sharper and less diffused posterior distributions.

Weak evidence

The evidence is considered as weak if the observed experience for the component is of the order of 10 demands (e.g., 1 failure in 10 demands).

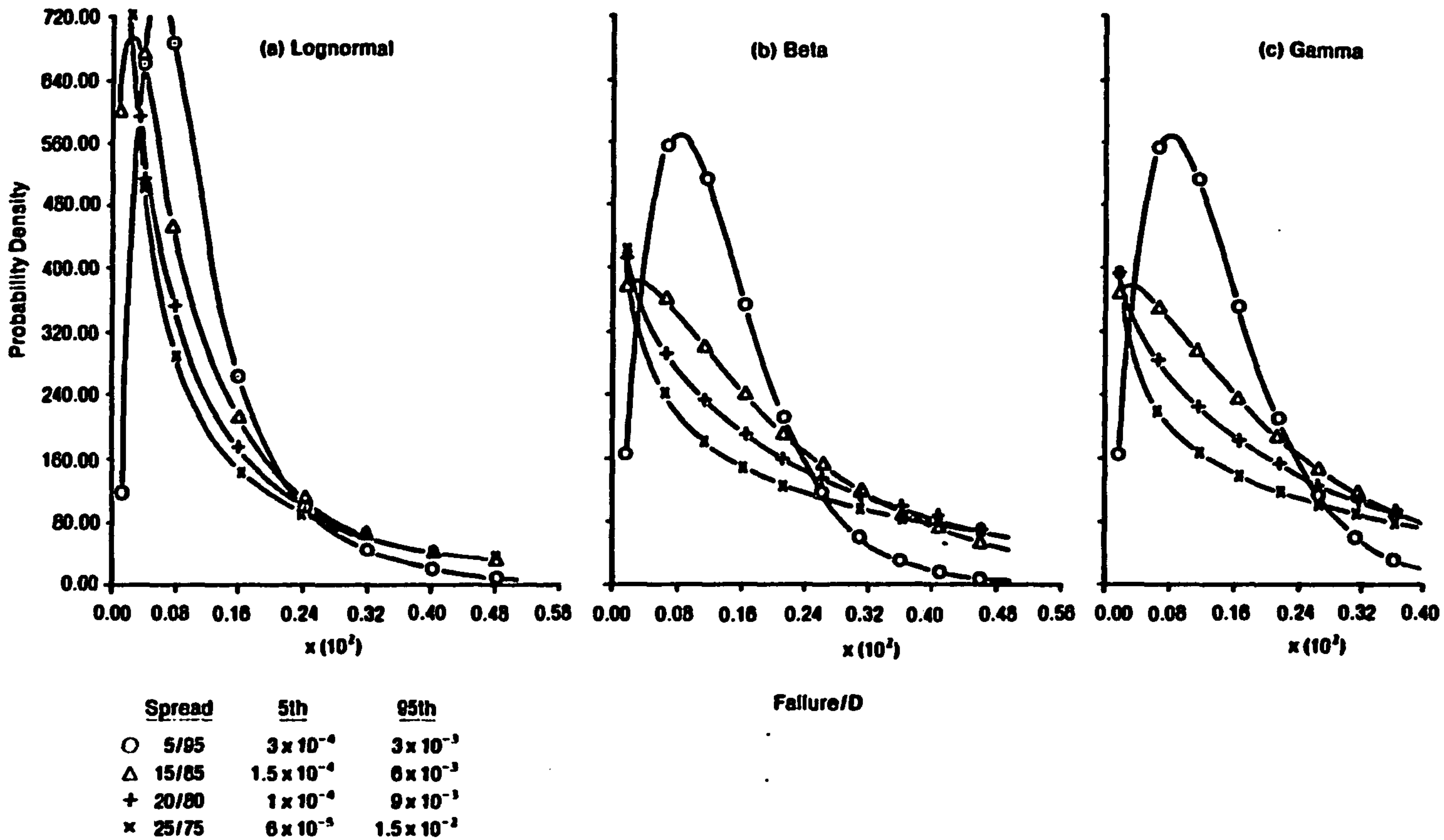
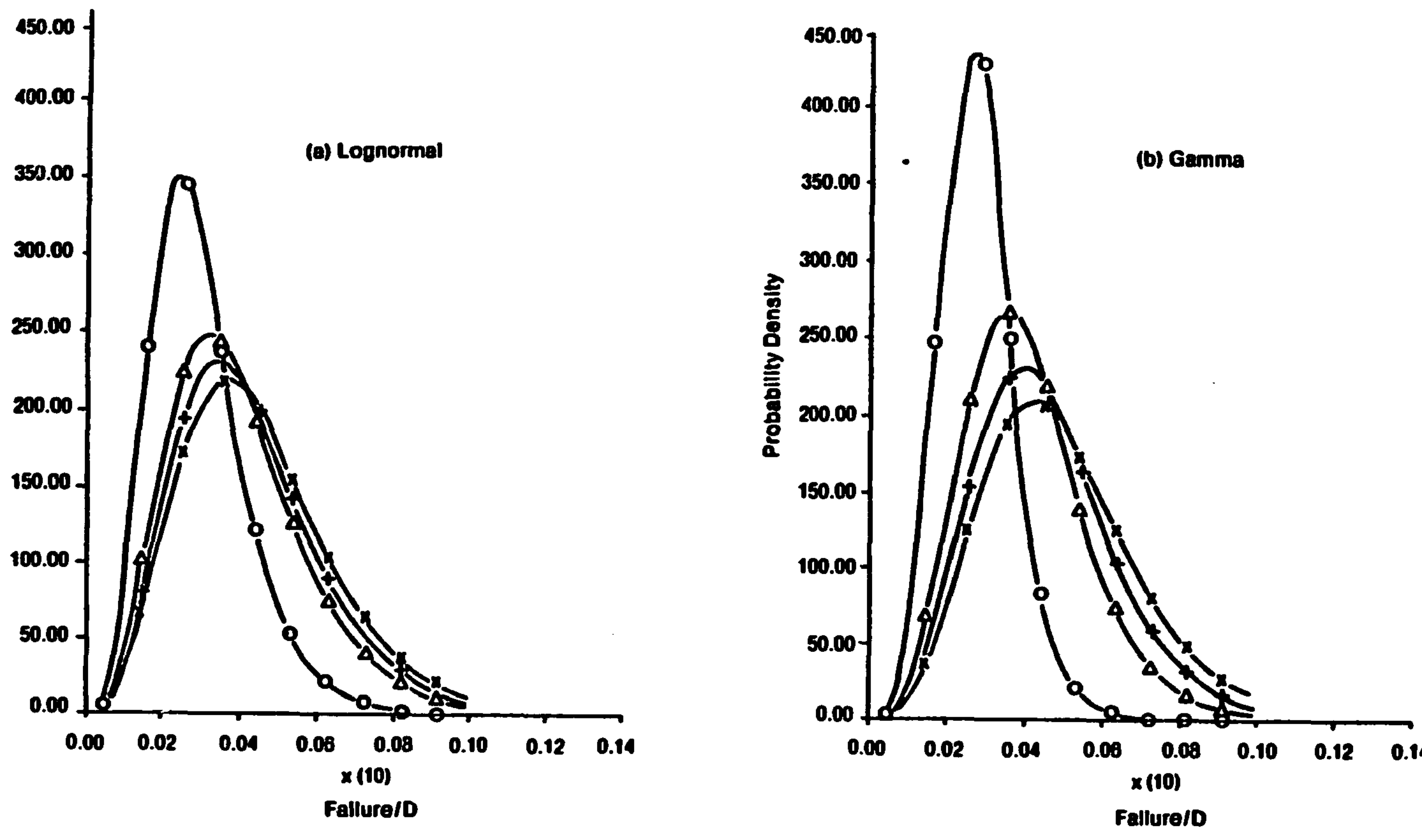


Figure 2. Prior distributions for motor operated valve failure per demand (D) for lognormal, beta, and gamma distribution with various spreads



Note: For legends see Fig. 2

Figure 3. Posterior distributions for MOV failure per demand (D) for strong evidence (5 failures in 1000 demands)

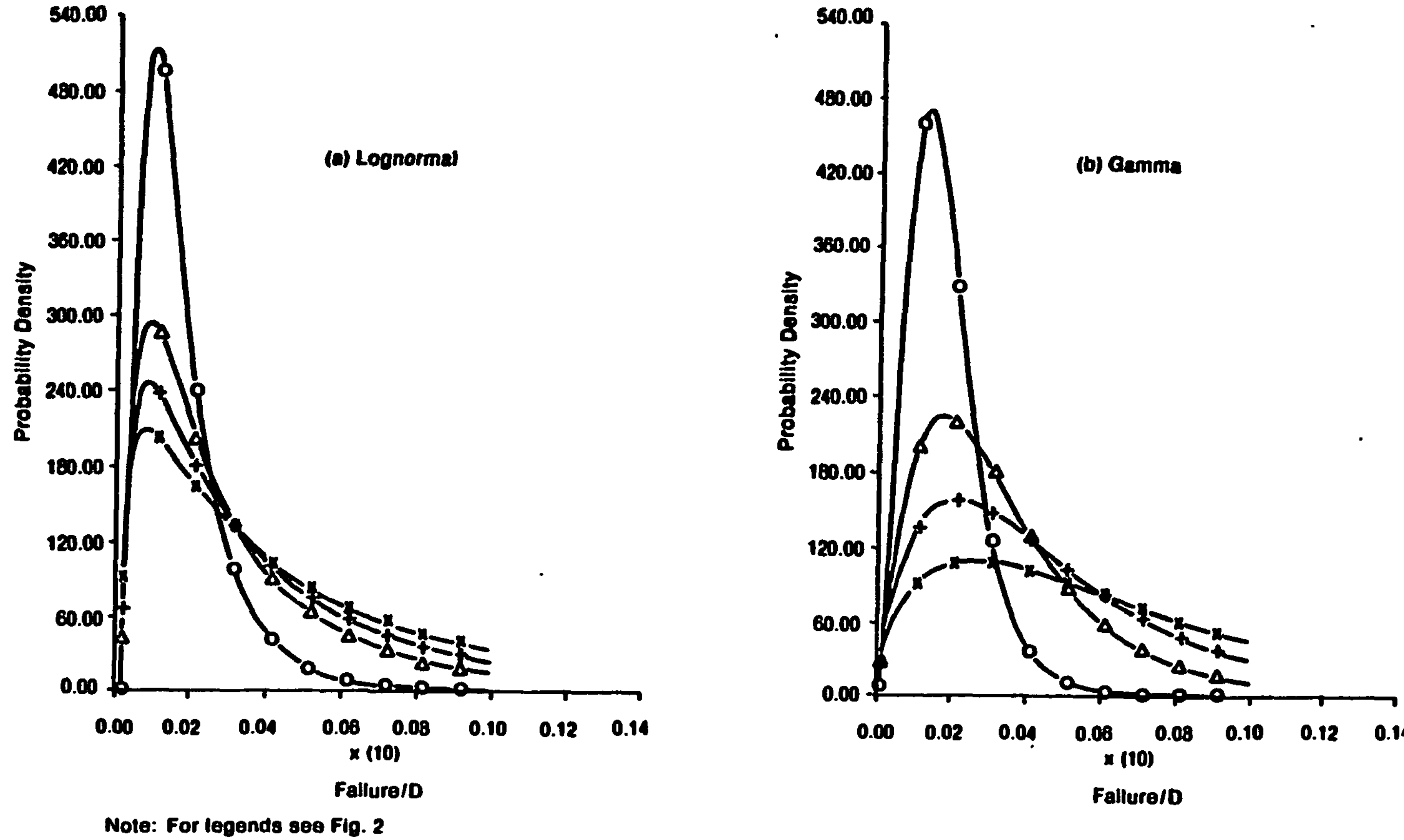
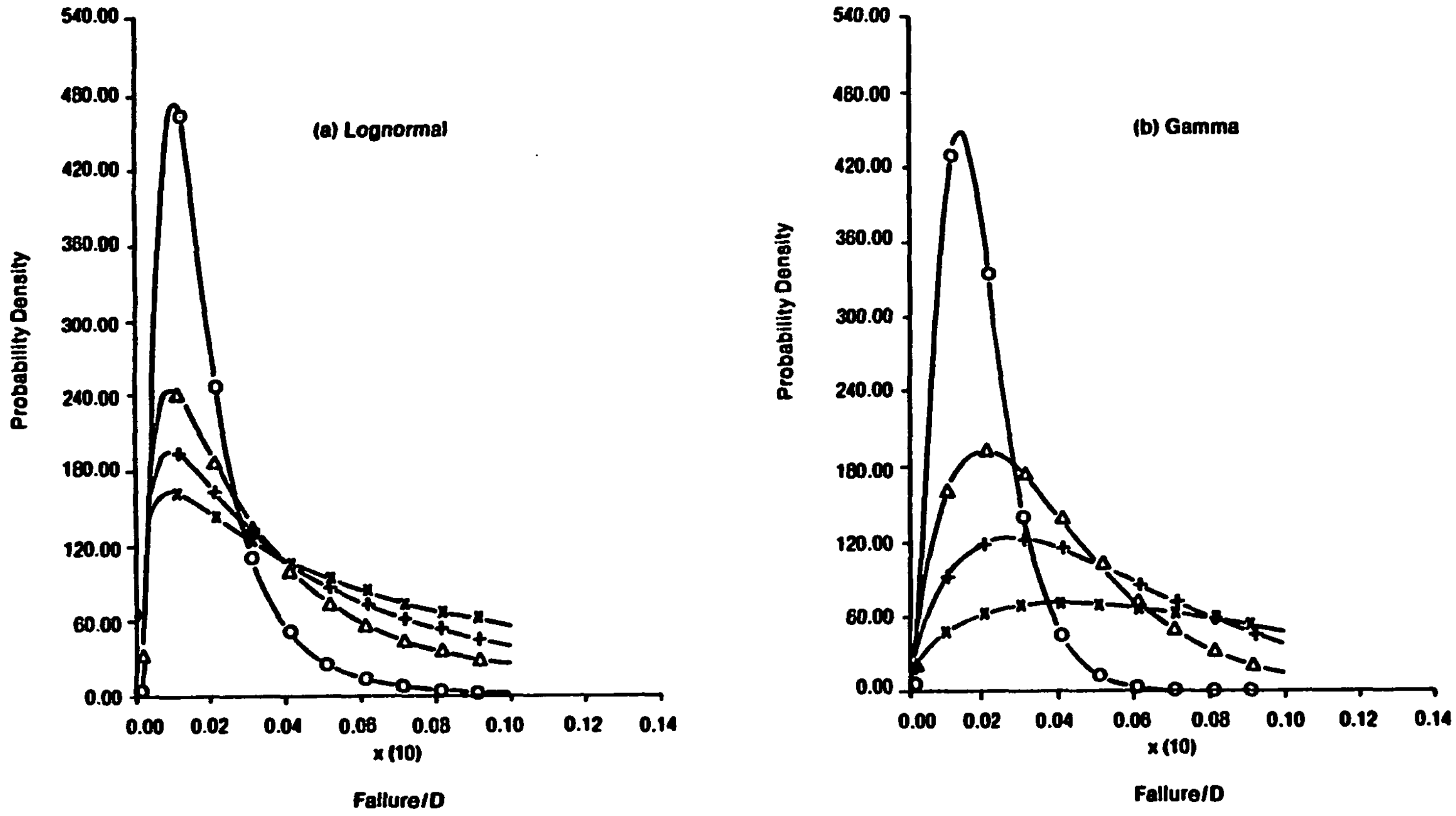
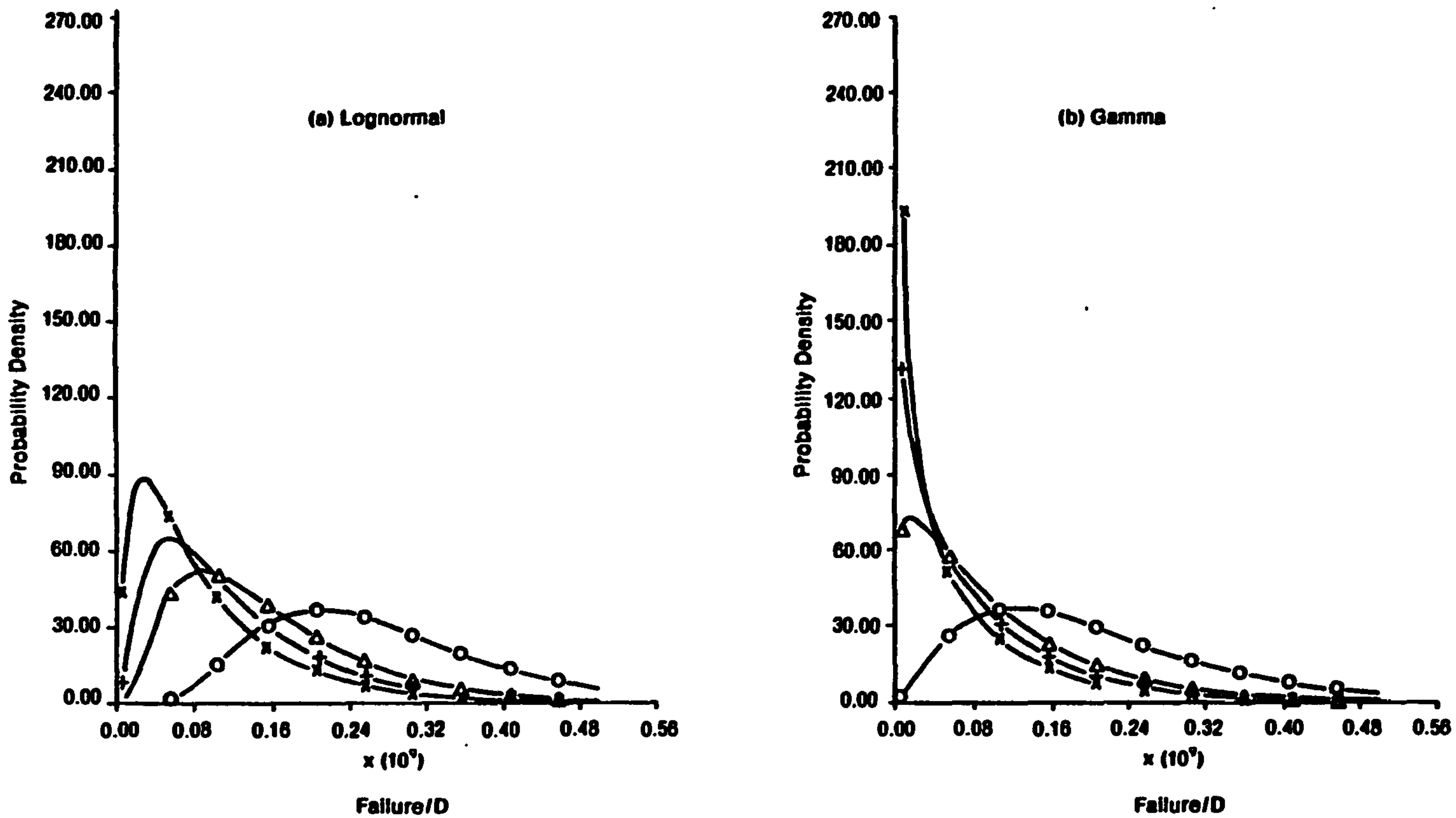


Figure 4. Posterior distributions for MOV failure per demand (D) for moderately strong evidence (1 failure in 100 demands)



Note: For legends see Fig. 2

Figure 5. Posterior distributions for MOV failure per demand (D) for weak evidence (1 failure in 10 demands)



Note: For legends see Fig. 2

Figure 6. Posterior distributions for MOV failure per demand (D) for the evidence (0 failures in 10,000 demands)

Table 1. Prior and posterior distribution mean and median[a] values for various types (lognormal, gamma, and beta) and spread of prior distribution for different plant specific evidences

	Lognormal[b]				
		Posterior			
Spread	Prior	5/1000[c]	1/100	1/10	0/10,000
5/95[d]	1.2×10^{-3} (9.5×10^{-4})	1×10^{-3} (8×10^{-3})	1.6×10^{-3} (1.2×10^{-3})	1.6×10^{-3} (1.4×10^{-3})	2.7×10^{-4} (2.5×10^{-4})
15/85[e]	1.8×10^{-3} (9.5×10^{-4})	4×10^{-3} (3.7×10^{-3})	3.3×10^{-3} (2.3×10^{-3})	5.5×10^{-3} (3.1×10^{-3})	1.6×10^{-4} (1.3×10^{-4})
20/80[f]	2.4×10^{-3} (9.5×10^{-4})	4.3×10^{-3} (4×10^{-3})	4.4×10^{-3} (2.9×10^{-3})	1.1×10^{-2} (5.3×10^{-3})	1.2×10^{-4} (1×10^{-4})
25/75[g]	4.9×10^{-3} (4.9×10^{-3})	4.5×10^{-3} (4.2×10^{-3})	5.6×10^{-3} (3.6×10^{-3})	2×10^{-2} (9.7×10^{-3})	9.5×10^{-5} (7.1×10^{-5})

[a]The median values are given in parentheses.

[b]Only prior is a lognormal distribution.

[c]Evidence 5 failures in 1000 demands.

[d]Prior 5th percentile = 3×10^{-4}/D and 95th percentile = 3×10^{-3}.

[e]Prior 5th percentile = 1.5×10^{-4}/D and 95th percentile = 6×10^{-3}.

[f]Prior 5th percentile = 1×10^{-4}/D and 95th percentile = 9×10^{-3}.

[g]Prior 5th percentile = 6×10^{-5}/D and 95th percentile = 1.5×10^{-2}.

	Gamma/Beta				
		Posterior			
Spread	Prior	5/1000[g]	1/100	1/10	0/10,000
5/95[c]	$1.3x10^{-3}$ ($1.2x10^{-3}$)	$6.3x10^{-4}$ ($6x10^{-4}$)	$1.8x10^{-3}$ ($1.6x10^{-3}$)	$1.9x10^{-3}$ ($1.7x10^{-3}$)	$2.1x10^{-4}$ ($1.8x10^{-4}$)
15/85[d]	$2.1x10^{-3}$ ($1.5x10^{-3}$)	$4x10^{-3}$ ($3.8x10^{-3}$)	$3.3x10^{-3}$ ($2.8x10^{-3}$)	$3.9x10^{-3}$ ($3.3x10^{-3}$)	$1.1x10^{-4}$ ($8x10^{-4}$)
20/80[e]	$2.9x10^{-3}$ ($1.9x10^{-3}$)	$4.5x10^{-3}$ ($4.2x10^{-3}$)	$4.6x10^{-3}$ ($3.8x10^{-3}$)	$6x10^{-3}$ ($4.9x10^{-3}$)	$8.4x10^{-5}$ ($5.5x10^{-5}$)
25/75[f]	$5.0x10^{-3}$ ($5x10^{-3}$)	$4.9x10^{-3}$ ($4.6x10^{-3}$)	$6.6x10^{-3}$ ($5.3x10^{-3}$)	$1x10^{-2}$ ($8.2x10^{-3}$)	$6.5x10^{-5}$ ($3.6x10^{-5}$)

It can be observed from Figure 5 that the posterior distribution is highly unstable and depends strongly upon the spread and the type of the population variability curve.

Conclusions

In some of the recent plant risk studies (e.g., Refs. 1, 2, and 4) the prior or generic distributions have been assessed using the available data sources and then they are spread out arbitrarily without studying the sensitivity of the spread or choice of the type of distribution. However, this study reveals that the failure data base for plant PRA from the Bayesian viewpoint can be developed without resorting to highly subjective arguments in formulating the component failure rate distributions. The criteria on acceptability should be the robustness of the posterior distribution, which is true when the experience data is significant. The development of the plant model (e.g., fault trees), therefore, is linked to the availability of component failure data, to the extent that the posterior failure rate distributions of the components are likelihood dominated. The results of the sensitivity analysis explored here will provide a basis for developing posterior distributions of the basic events for use in PRA.

SIGNIFICANCE INDICES FOR COMPONENT RANKING

To develop a ranking of critical components of a large system, we will define three significance indices and see if the significance indices are robust with respect to the basis event failure probability distributions.

In general, the failure probability or frequency of a system failure can be expressed in terms of the component failure probabilities as follows [17]:

$$P = \sum_{\alpha_1 \ldots \alpha_n}^{0,1} A_{[\alpha_1 \ldots \alpha_n]} P_1^{\alpha_1} \ldots P_n^{\alpha_n} \tag{15}$$

In equation 15, $P_1 \ldots P_n$ are n-component probabilities or frequencies, P is the Top-level (system-level) probability and $A_{[\alpha_1 \ldots \alpha_n]}$ are (2^n-1) coefficients with values 0, 1, or -1 which are characteristic of the fault tree considered. The expression in equation 15 corresponds directly to a fault tree or system block diagram.

Equation 15 as an algebraic representation offers certain advantages. It expresses for each probability $P_m (m=1 \ldots n)$ in terms of the Top-level probability P as:

$$P = A_m^{(0)} + A_m^{(1)} P_m \tag{16}$$

From equation 16 one can obtain:

$$\frac{\partial P}{\partial P_m} = A_m^{(1)} \tag{17a}$$

and

$$S_m = \frac{P_m}{P}\frac{\partial p}{\partial P_m} = \frac{A_m^{(1)}P_m}{P} = 1 - \frac{A_m^{(0)}}{P} \tag{17b}$$

where S_m is defined as the Significance Index of component m.

Equations 17a and 17b represent the sentivities of the Top-level tree (system failure event) with respect to the probabilities of the components. The parameters in equation 17b measure directly the percent changes in the Top-level probability due to a certain specified percentage change in the component probability. As such, equation 17b is used to define a measure of importance for the component under examination. Equation 17a can be used directly to calculate the changes in the Top-level probability as the component probability changes.

From equation 17a it can be seen that as $A_m^{(1)} \rightarrow 0$, i.e., component m appears in few sequences, $S_m \rightarrow 0$. Furthermore, from equation 17b as $P_m \rightarrow 0$, i.e., the element m approaches nonfailure, $S_m \rightarrow 0$. In addition, as $A_m \rightarrow 0$, $S_m \rightarrow 1$. This is the limit when component m appears in all sequences. It is, therefore, appropriate to define S_m as the significance index for the element m. It is to be noted from equations 17a and 17b that S_m is dimensionless and ranges from 0 (least significant) to 1 (most significant). S_m measures the importance of the element m in terms of the number of sequences it is associated with and the relative strength of such associations. Insofar as equations 16, 17a and 17b visibly exhibit the role that a specified element plays in the structure of the Top-level probability, the present approach can be judged to be an explicit formulation of the problem.

Historically, point estimate methods were used to calculate the frequency of a Top event, e.g., failure of a complex system. Since the experience base is not extensive, large uncertainties in the component failure rate estimates exist. It was, therefore, recognized that such failure rates needed to be modeled as probability distributions rather than point values. The modeled uncertainties can then be propagated to yield the uncertainty of the Top event.

The mean $\bar{P}$ and variance σ^2 for the Top-level probability are first obtained in terms of component mean $\bar{P}_m$ and variance σ_m^2. It is then possible to define three significance indices by the three partial derivatives as follows:

$$S_m^{(1)} = \frac{\bar{P}_m}{\bar{P}} \frac{\partial \bar{P}}{\partial \bar{P}_m}$$

$$S_m^{(2)} = \frac{\bar{P}_m}{2} \frac{\partial \sigma^2}{\partial \bar{P}_m}$$

$$S_m^{(3)} = \frac{\sigma_m^2}{\sigma^2} \frac{\partial \sigma^2}{\partial \sigma_m^2} \tag{18}$$

The three indices above measure, respectively, the percentage changes in the Top level: (1) mean with respect to unit percentage change in component mean, (2) variance with respect to unit percentage change in component mean, and (3) variance with respect to unit percentage change in component variance. If any one of these dimensionless indices has a relatively high value, the element can be judged to be potentially important.

Quantification of Significance Indices

Here, methods are developed to quantify the three Significance Indices. The equations for the Mean and Variance of the Top event are first formulated and then appropriate partial derivatives with respect to the basic event (or failure element) Mean and Variance are evaluated to quantify the Significance Indices.

Top event mean and variance

For a given Fault Tree, the Top level probability P is given in terms of component probabilities p_m (all considered point values for the time being) by equation 15 repeated here for convenience.

$$P = \sum_{\alpha_1 \ldots \alpha_n}^{0,1} A_{[\alpha_1 \ldots \alpha_n]} \, p_1^{\alpha_1} \ldots p_n^{\alpha_n} \tag{15}$$

Although equation 15 is strictly true for a Fault Tree, it can be used for an Event Tree or a Fault Tree-Event Tree mix by generalizing the Top event appropriately. The coefficients $A_{[\alpha_1 \ldots \alpha_n]}$ in equation 15 are characteristic values dependent on the logic of the Fault Tree. It can take on values of 0, 1 or -1; the nonzero values occur if and only if the set $\alpha_1 \ldots \alpha_n$ corresponds to a Sequence or Crossterm. For example, if {1,2} and {1,3} are two viable sequences with chain length 2 each, then

$$A_{[110000\ldots]} = A_{[101000]} = +1 \; .$$

Furthermore, the crossterm between these two sequences is obtained by the set theoretic union of the sequence elements. Thus, the crossterm corresponds to:

$$\{1,2\} \cup \{1,3\} = \{1,2,3\} .$$

It is seen that this term has a chain length 3. Additionally, the algebraic rules for compositions are such that the crossterms obtained by "uniting" two sequences have a value -1, those with three, +1, and so on. Thus, in the example considered:

$$A_{[111000\ldots]} = -1 .$$

As the chain length (including Sequences and Crossterms) grows bigger, the corresponding terms in equation 15 become smaller and involve a multiplication of a larger number of P's. Hence, it is efficient to truncate the process for a specified value of chain length N_{max}, defined to be the maximum number of failure elements that participate in an accident sequence for the Fault tree under consideration.

Since the component unavailabilities are more realistically represented by probability distributions rather than point estimates, it is necessary to consider their probability distribution function (PDF), $f_{m(i)}$, for values $P_{m(i)}$. The input probability distributions are discretized with the techniques outlined on Page 198. From equation 15 the first moment of the Top level probability is obtained as follows:

$$\bar{P} = \sum_{\alpha_1\ldots\alpha_n}^{0,1} f_{1(i_1)}\cdots f_{n(i_n)} P_{1(i_1)}^{\alpha_1}\cdots P_{n(i_n)}^{\alpha_n} A_{[\alpha_1\ldots\alpha_n]}$$

$$= \sum_{\alpha_1\ldots\alpha_n}^{0,1} A_{[\alpha_1\ldots\alpha_n]} \overline{P_1^{\alpha_1}}\,\overline{P_2^{\alpha_2}}\cdots\overline{P_n^{\alpha_n}} \tag{19}$$

Again, the second moment of the Top level probability is given by,

$$\overline{P^{(2)}} = \sum_{i_1 \ldots i_n}^{1 \to 11} \sum_{\alpha_1 \ldots \alpha_n}^{0,1} \sum_{\alpha'_n \ldots \alpha'_n}^{0,1} f_{1(i_1)} f_{2(i_2)} \ldots f_{n(i_n)} P_{1(i_1)}^{\alpha_1+\alpha'_1} \ldots P_n^{\alpha_n+\alpha'_n}$$

$$\times A_{[\alpha_1 \ldots \alpha_n]} \times A_{[\alpha'_1 \ldots \alpha'_n]}$$

$$= \sum_{\alpha_1 \ldots \alpha_n}^{0,1} \sum_{\alpha'_1 \ldots \alpha'_n}^{0,1} A_{[\alpha_1 \ldots \alpha_n]} A_{[\alpha'_1 \ldots \alpha'_n]} \overline{P_1^{\alpha_1+\alpha'_1}} \ldots \overline{P_n^{\alpha_n+\alpha'_n}}$$

(20)

Here, each probability distribution is discretized into 11 cells (the discretization procedure is described later in the subsection entitled, "Sensitivity Analysis Using Significance Indices").

From equations 19 and 20, the variance of the Top level probabilities is given by:

$$\sigma^2 = \sum_{\alpha_1 \ldots \alpha_n}^{0,1} \sum_{\alpha'_1 \ldots \alpha'_n}^{0,1} A_{[\alpha_1 \ldots \alpha_n]} A_{[\alpha'_1 \ldots \alpha'_n]}$$

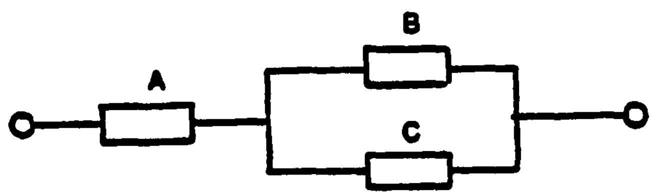
$$\times \left\{ \overline{P_1^{\alpha_1+\alpha'_1}} \ldots \overline{P_n^{\alpha_n+\alpha'_n}} - (\overline{P_1^{\alpha_1}}\ \overline{P_1^{\alpha'_1}}) \ldots (\overline{P_n^{\alpha_n}}\ \overline{P_n^{\alpha'_n}}) \right\}$$

(21)

The equations 19 and 21 can then be used to calculate the Mean and Variances of the Top level probability in terms of the first and second moments of the component and the coefficients $A_{[\alpha_1 \ldots \alpha_n]}$.

Calculation of significance indices

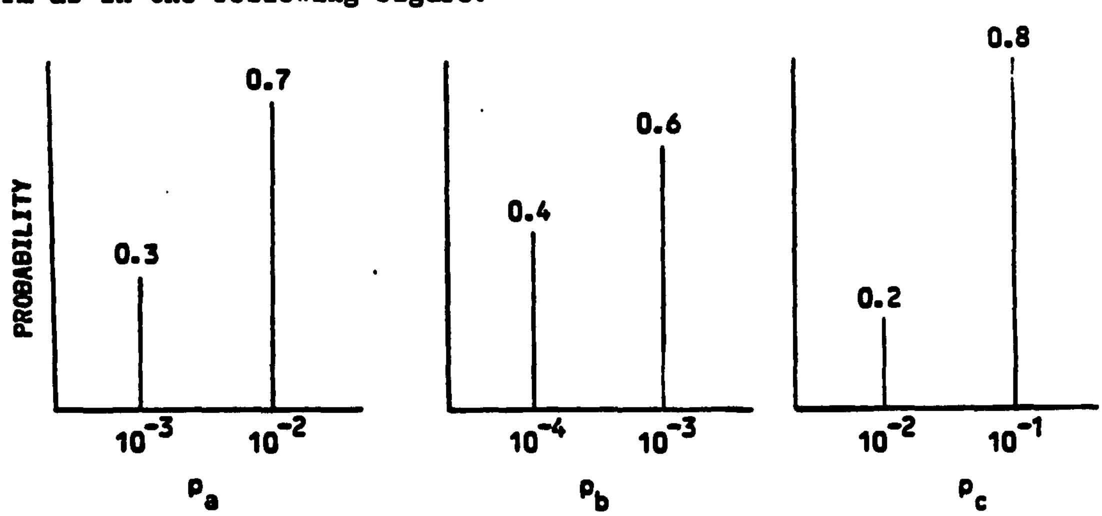
Let us consider a system X consisting of components A, B, and C, where B and C are in parallel and A in series as given by the following block diagram



Then, the probability of system failure can be written as

$$p_x = p_a + p_b p_c - p_a p_b p_c$$
$$\simeq p_a + p_b p_c$$

where p_a, p_b, and p_c are the probabilities of failures for A, B, C, respectively, and the product term $p_a p_b p_c$ is negligible. It is ~~also~~ assumed that the failure events are independent of each other. Let the distributions p_a, p_b, and p_c be given in the discrete form as in the following figure.



We can readily calculate the following expectation, to be used in later calculations.

$$\bar{p}_a = (10^{-3})(0.3) + (10^{-2})(0.7) = 7.3 \times 10^{-3} ,$$
$$\overline{p_a^2} = (10^{-3})^2\ (0.3) + (10^{-2})^2\ (0.7) = 7.03 \times 10^{-5} .$$

Similarly:

$$\overline{p_b} = 6.4 \times 10^{-4},\ \overline{p_b^2} = 6.04 \times 10^{-7}$$
$$\overline{p_c} = 8.2 \times 10^{-2},\ \overline{p_c^2} = 8.02 \times 10^{-3} .$$

Also, the variances are obtained as

$$\sigma_a^2 = \overline{p_a^2} - (\bar{p}_a)^2$$
$$= 7.03 \times 10^{-5} - (7.3 \times 10^{-3})^2$$
$$= 1.701 \times 10^{-7}$$
$$\sigma_b^2 = 1.944 \times 10^{-7}$$
$$\sigma_c^2 = 1.296 \times 10^{-3} .$$

For the system failure,

$$\bar{p}_x = \bar{p}_a + \bar{p}_b\bar{p}_c = 7.35 \times 10^{-3} \ \underline{\text{Mean}}$$
$$p_x^2 = (p_a + p_b p_c)^2 = p_a^2 + 2p_a p_b p_c + p_b^2 p_c^2$$
$$\overline{p_x^2} \simeq \overline{p_a^2} + 2\bar{p}_a\overline{p_b}\overline{p_c} + p_b^2 p_c^2$$
$$= 7.11 \times 10^{-5}$$
$$\sigma_x^2 = \overline{p_x^2} - \bar{p}_x^2$$
$$= 7.11 \times 10^{-5} - (7.35 \times 10^{-3})^2$$
$$= 1.71 \times 10^{-5} \ \underline{\text{Variance}} .$$

Then, the first significance index is calculated as follows:

$$\bar{p}_x = \bar{p}_a + \bar{p}_b\bar{p}_c$$

$$\frac{\partial\bar{p}_x}{\partial\bar{p}_a} = 1, \quad \frac{\partial\bar{p}_x}{\partial\bar{p}_b} = p_c = 8.2 \times 10^{-2}, \quad \frac{\partial\bar{p}_x}{\partial\bar{p}_c} = p_b = 6.4 \times 10^{-4}$$

$$S_a^{(1)} = \frac{\bar{p}_a}{\bar{p}_x}\frac{\partial\bar{p}_x}{\partial\bar{p}_a} = \frac{7.3 \times 10^{-3}}{7.35 \times 10^{-3}}(1) = 0.9932$$

$$S_b^{(1)} = \frac{\bar{p}_a}{\bar{p}_x}\frac{\partial\bar{p}_x}{\partial\bar{p}_b} = 8.14 \times 10^{-2}$$

$$S_c^{(1)} = \frac{\bar{p}_c}{\bar{p}_x}\frac{\partial\bar{p}_x}{\partial\bar{p}_c} = 7.14 \times 10^{-3}$$

To obtain the second and third significance indices, we first obtain the expression of the Top event variance as follows:

$$\begin{aligned}\sigma_x^2 &= \overline{p_x^2} - (\bar{p}_x)^2 \\ &= (\overline{p_a^2} + 2\bar{p}_a\bar{p}_b\bar{p}_c + \overline{p_b^2p_c^2}) - (\bar{p}_a + \bar{p}_b\bar{p}_c)^2.\end{aligned}$$

Writing the second moments in terms of variance and first-moment as:

$$\begin{aligned}\overline{p_a^2} &= \sigma_a^2 + (\bar{p}_a)^2 \\ \overline{p_b^2} &= \sigma_b^2 + (\bar{p}_b)^2 \\ \overline{p_c^2} &= \sigma_c^2 + (\bar{p}_c)^2,\end{aligned}$$

we can obtain the expression of the Top event variance as

$$\sigma_x^2 = \sigma_a^2 + (\sigma_b^2 + \bar{p}_b{}^2)(\sigma_c^2 + \bar{p}_c{}^2) - \bar{p}_b{}^2\bar{p}_c{}^2. \tag{21}$$

Then, the second significant indices can be obtained from equation 21 as follows:

$$\frac{\partial \sigma_x^2}{\partial \bar{p}_a} = 0$$

$$\frac{\partial \sigma_x^2}{\partial \bar{p}_b} = 2\,\bar{p}_b\,(\sigma_c^2 + \bar{p}_c^2) - 2\bar{p}_c^{\,2}\bar{p}_b = 1.659 \times 10^{-6}$$

$$\frac{\partial \sigma_x^2}{\partial \bar{p}_c} = (\sigma_b^2 + \bar{p}_b^{\,2})\,(2\bar{p}_c) - 2\bar{p}_b^{\,2}\,\bar{p}_c = 3.188 \times 10^{-8}$$

$$S_a^{(2)} = \frac{\bar{p}_a}{\sigma_x^2}\frac{\partial \sigma_x^2}{\partial \bar{p}_a} = 0$$

$$S_b^{(2)} = \frac{\bar{p}_b}{\sigma_x^2}\frac{\partial \sigma_x^2}{\partial \bar{p}_b} = 6.209 \times 10^{-5}$$

$$S_c^{(2)} = \frac{\bar{p}_c}{\sigma_x^2}\frac{\partial \sigma_x^2}{\partial \bar{p}_c} = 1.529 \times 10^{-4}$$

Further, the third significance indices are obtained from equation 20, as follows:

$$\frac{\partial \sigma_x^2}{\partial \sigma_a^2} = 1$$

$$\frac{\partial \sigma_x^2}{\partial \sigma_b^2} = (\sigma_c^2 + p_c^2) = 8.02 \times 10^{-3}$$

$$\frac{\partial \sigma_x^2}{\partial \sigma_c^2} = (\sigma_b^2 + p_b^2) = 6.04 \times 10^{-7}$$

$$S_a^{(3)} = \frac{\sigma_a^2}{\sigma_x^2}\;\frac{\partial \sigma_x^2}{\partial \sigma_a^2} = 0.9947$$

$$S_b^{(3)} = \frac{\sigma_b^2}{\sigma_x^2} \frac{\partial \sigma_x^2}{\partial \sigma_b^2} = 9.12 \times 10^{-5}$$

$$S_c^{(3)} = \frac{\sigma_c^2}{\sigma_x^2} \frac{\partial \sigma_x^2}{\partial \sigma_c^2} = 4.58 \times 10^{-5}$$

The first, second, and third significance indices are provided, for convenience, in Table 2.

Table 2. Significance Indices obtained for the example calculations for $p_x = p_a + p_b p_c$ where p_a, p_b, p_c, and p_x are the failure probability distributions for the components a, b, c, and the system x, respectively

Component	First Significance Index (1) S_m	Second Significance Index (2) S_m	Third Significance Index (3) S_m
A	0.9932	0	0.9947
B	8.14×10^{-2}	6.21×10^{-5}	9.12×10^{-5}
C	7.14×10^{-3}	1.53×10^{-4}	4.58×10^{-5}

It should be made clear that the Significance Indices do not distinguish between the importance of the component due to its location in the system (structural importance) and the importance of the component due to its high failure rate (reliability importance). Thus, a component having relatively low failure rate (or a good design) may have a high Significance Index if its location in the system does not have redundancy. In this case, it is not the design of the component that needs improvement; rather, redundancy is required to lower the Significance Index of the component. The Significance Index may also be high due to high failure rate of the component. In this case, design improvement as well as redundancy will help to reduce the Significance Index of the component. Further research is, therefore, required to apportion the Significance Indices developed here to distinguish between the structural importance and the reliability importance.

Robustness of Significance Indices

The Significance Indices of a component in a system depend upon the failure probability distribution of the component. Actually, as we will see in this section, the Significance Indices of the rest of the components in the system model may also vary due to the variation in the failure probability distribution of the component. This warrants us to investigate the impact of various assumptions in the failure probability distributions of the basic events on the ranking of the system components. Recalling that the ranking of the components is based on their Significance Indices, we would like to observe the modified Significance Indices based on various assumptions of the basic event probability distributions.

Considering that one objective of a plant PRA is to develop a ranking of components in order of their impact on plant safety (e.g., frequency of core melt), we would also like to analyze how sensitive is such a ranking on the assumptions of the basic event probability distributions. If the component ranking does not vary significantly, then the PRA results are defined here as being robust. The method for robustness analysis as related to the Significance Index is described here. Let us define an index:

$$D_{km} = \frac{\bar{p}_k}{S_m} \frac{\partial S_m}{\partial \bar{p}_k}$$

where:

D_{km} = significance of the mth element due to change induced in the kth element

$\bar{p}_k$ = mean probability of kth element

S_m = first significance index of mth element.

Since $\bar{S}_m = \frac{\bar{p}_m}{\bar{p}} \frac{\partial \bar{p}}{\partial \bar{p}_m}$

$$D_{km} = \frac{\bar{p}_k}{S_m} \frac{\partial}{\partial \bar{p}_k} \left[\frac{\bar{p}_m}{\bar{p}} \frac{\partial \bar{p}}{\partial \bar{p}_m}\right]$$

$$D_{km} = \frac{\bar{p}_k}{S_m} \left[\frac{\delta_{km}}{\bar{p}} \frac{\partial \bar{p}}{\partial \bar{p}_m} - \frac{\bar{p}_m}{\bar{p}^2} \frac{\partial \bar{p}}{\bar{p}_k} \frac{\partial \bar{p}}{\partial \bar{p}_m} + \frac{\bar{p}_m}{\bar{p}} \frac{\partial^2 \bar{p}}{\partial \bar{p}_m \partial \bar{p}_k}\right]$$

$$= \delta_{km} \frac{1}{S_m} \frac{\bar{p}_k}{\bar{p}} \frac{\partial \bar{p}}{\partial \bar{p}_m} - \frac{\bar{p}_k \bar{p}_m}{S_m \bar{p}^2} \frac{\partial \bar{p}}{\partial \bar{p}_k} \frac{\partial \bar{p}}{\partial \bar{p}_m} + \frac{\bar{p}_k \bar{p}_m}{S_m \bar{p}} \frac{\partial^2 \bar{p}}{\partial \bar{p}_m \partial \bar{p}_k}$$

where

δ_{km} is the Kronecker Delta function such that

$\delta_{km} = 1$ when $k = m$

$= 0$ when $k \neq m$

For $k = m$:

$$1\text{st Term} = \delta_{km} \frac{1}{S_m} \frac{\bar{p}_m}{\bar{p}} \frac{\bar{p}}{\bar{p}_m}$$

$$= \delta_{km}$$

$$= 1, \text{ by definition}$$

when $k \neq m$:

$$1\text{st Term} = 0$$

$$2\text{nd Term} = \left(\frac{\bar{p}_k}{\bar{p}} \frac{\partial \bar{p}}{\partial \bar{p}_k} \quad \frac{\bar{p}_m}{\bar{p}} \frac{\partial \bar{p}}{\partial \bar{p}_m}\right) \frac{1}{S_m} = S_k$$

Hence, $$D_{mk} = \delta_{km} - S_k + \frac{\bar{p}_k \bar{p}_m}{S_m \bar{p}} \frac{\partial^2 \bar{p}}{\partial \bar{p}_k \partial \bar{p}_m} \tag{22}$$

Thus,

$$D_{mk} = S_{km} - S_k \tag{23}$$

$$= 1 - S_k , \quad \text{when } k = m$$

$$= - S_k + \frac{\bar{p}_k \bar{p}_m}{S_m \bar{p}} \frac{\partial^2 \bar{p}}{\partial \bar{p}_k \partial \bar{p}_m} , \text{ when } k \neq m .$$

Calculations of robustness analysis

Let us consider a system with elements 1, 2, 3, 4, such that the mean system failure probability is given by

$$p = p_1 + p_2 p_3 + p_2 p_4$$

Let us assume that the first significance indices are

$$S_1 = 0.8, \quad S_2 = 0.8, \quad S_3 = 0.2, \quad \text{and } S_4 = 0.6 .$$

From the definition, we notice that

$$S_1 = \frac{\bar{p}_1}{\bar{p}} , \; S_2 = \frac{\bar{p}_2 \bar{p}_3 + \bar{p}_2 \bar{p}_4}{\bar{p}} , \; S_3 = \frac{\bar{p}_2 \bar{p}_3}{\bar{p}} , \text{ and } S_4 = \frac{\bar{p}_2 \bar{p}_4}{\bar{p}} .$$

Then, from equation 23 we write:

$$D_{11} = 1 - S_1 = 0.2$$

$$D_{12} = - S_2 = 0.8$$

$$D_{13} = - S_3 = 0.2$$

$$D_{14} = - S_4 = 0.6$$

$$D_{21} = - S_1 = - 0.8$$

$$D_{22} = - S_2 = 0.2$$

$$D_{23} = - S_3 + \frac{\bar{p}_2\bar{p}_3}{S_2\bar{p}} = - S_3 + \frac{S_3}{S_2} = 0.05$$

$$D_{24} = - S_4 + \frac{\bar{p}_2\bar{p}_4}{\bar{p}S_2} = 0.6 \quad + \frac{S_4}{S_2} = 0.15$$

$$D_{31} = - S_1 = - 0.8$$

$$D_{32} = - S_2 + \frac{\bar{p}_2\bar{p}_3}{\bar{p}S_3} = 0.2$$

$$D_{33} = - S_3 = - 0.2$$

$$D_{34} = - S_4 = - 0.6$$

$$D_{41} = - S_1 = - 0.8$$

$$D_{42} = - S_2 + \frac{\bar{p}_2\bar{p}_4}{\bar{p}S_4} = 0.2$$

$$D_{43} = - S_3 = - 0.2$$

$$D_{44} = 1 - S_4 = 1 - 0.6 = 0.4$$

If we introduce 100% change in component k, i.e., $(\partial\bar{p}_k/\bar{p}_k) = 1$, then

$$D_{km} = \frac{\bar{p}_k}{S_m}\frac{\partial S_m}{\partial \bar{p}_k} = \frac{(\partial S_m/S_m)}{(\partial \bar{p}_k/\bar{p}_k)} = \frac{\partial S_m}{S_m}$$

Hence, $\partial S_m = S_m D_{km}$.

Hence, the significance index for component m after introducing a change in component k is given by

$$S_m^{'} = S_m + S_m D_{km} = S_m\ (1 + D_{km})$$

Let us suppose that we induce 100% change in components 1, 2, 3, and 4, one at a time, and observe how the significance indices for each component change. We begin with a 100% change in the probability of component 1.

Then,

$$\frac{\Delta\bar{p}_1}{\bar{p}_1} = 1$$

and

$$S_1^{'} = S_1\ (1 + D_{11}) = 0.8\ (1 + 0.2) = 0.96$$

$$S_2^{'} = S_2\ (1 + D_{21}) = 0.8\ (1 - 0.8) = 0.16$$

$$S_3^{'} = S_3\ (1 + D_{31}) = 0.2\ (1 - 0.8) = 0.04$$

$$S_4^{'} = S_4\ (1 + D_{41}) = 0.6\ (1 - 0.8) = 0.12$$

Similarly, for $\Delta\bar{p}_2/\bar{p}_2 = 1$

$$S_1^{'} = 0.26, \quad S_2^{'} = 0.96, \quad S_3^{'} = 0.24, \quad S_4^{'} = 0.72$$

For $\Delta\bar{p}_3/\bar{p}_3 = 1$

$$S_1^{'} = 0.64, \quad S_2^{'} = 0.84, \quad S_3^{'} = 0.36, \quad S_4^{'} = 0.48$$

For $\Delta\bar{p}_4/\bar{p}_4 = 1$

$$S_1^{'} = 0.32, \quad S_2^{'} = 1, \quad S_3^{'} = 0.08, \quad S_4^{'} = 0.84$$

We can now observe the ranking of the components due to changes induced in their probability of failure. Thus, prior to any changes in the failure probabilities, the ranking was (1, 2, 3, 4) or (2, 1, 3, 4), obtained from the significance of component ranking upon changes in the component failure probabilities as given in Table 3. Notice that the component ranking does not change even if we induce significant changes in the component failure probabilities. The results of the reliability analysis (in this case the ranking of the components 1, 2, 3, 4) are, therefore, robust for this example, insofar as the dominant contributors to system unavailability remain the same. The corrective actions to reduce system unavailability (by reducing the probability of component failures), therefore, do not depend in this case upon the failure probabilities; rather, they are dependent upon the structure or the way the components are arranged in the system.

The approach presented above, therefore, provides a mechanism to analyze robustness of PRA results.

Table-3. Study of robustness for sample problem, $p = p_1 + p_2p_3 + p_2p_4$

100% Change In	First Significance Index for Elements 1, 2, 3, 4				
	S_1	S_2	S_3	S_4	Ranking
p_1	0.96	0.16	0.04	0.12	1, 2, 4, 3
p_2	0.26	0.96	0.24	0.72	2, 4, 1, 3
p_3	0.64	0.84	0.36	0.48	2, 1, 4, 3
p_4	0.32	1.0	0.08	0.84	2, 4, 1, 3
		Prior to Change Induced			1, 2, 4, 3 or 2, 1, 4, 3

Ranking After Introducing Changes in Component Probabilities

1st: Elements 1 and 2
2nd: Element 4
3rd: Element 3

Comparison with Other Measures of Importance

A number of importance measures have been defined in the literature pertaining to reliability and risk assessment. A brief comparison of the Significance Indices developed here will be made with other measures of importance.

The Significance Index 1, S_m^1, is an improvement over the Birnbaum measure of importance, and is the same as the Upgrading Function defined by Lambert [18]. Fussell importance [19] measures the percentage contribution of a component or a cut-set to the Top event probability but does not provide the improvement in the Top event probability due to certain improvement in a basic event. Other measures of importance of interest could be Barlow-Proschan basic event and cut-set importance [20] (or the steady state Barlow-Proschan importance [20]); they provide the expected number of Top event failures due to a basic event or cut-set failure in time 0 to t (or probability of Top event at steady state). From the point of view of applications, we find the Upgrading Function or S_m^1 to be the most appropriate measure of importance which provides fractional change in the Top event mean due to fractional change in a basic event or cut-set mean probability.

It should be noted that the set of three importance measures used here addresses first and second moments rather than the first moment only. As such, the proposed definition goes beyond identifying importance in terms of the point-estimates and covers the sensitivity of the Top event failure probability due to uncertainty in the component failure rates. It is interesting to note that Vicki Bier [21] at MIT has also defined, independently, an importance measure involving second moments, which is the same as the Significance Index 3. The Significance Indices 2 and 3 provide the informational needs for scatter in component failure data, i.e., they identify the basic events which significantly contribute to the uncertainty of the Top event.

Sensitivity Analysis Using Significance Indices

We have analyzed the robustness of the system reliability results to see how the ranking of the dominant components changes if the mean failure probability of the components is changed. For a simple example, we have observed that the ranking does not change. However, even if the ranking may not change, the system failure probability may be sensitive to the changes in the component failure probabilities. To analyze how sensitive is the system failure probability distribution upon the variations in the component failure probability distributions, the following approach is proposed.

Suppose for a complex system of interest the mean and variance of the system failure probability have been obtained through a system model (fault tree/event tree) and basic event (or component) failure probability data base. We would now like to observe the changes in the system failure probability mean and variance due to changes in each of the component failure probability distributions (specifically, the mean and variance).

Let us assume that the basic event failure probability distributions are lognormally distributed and the 5th and 95th percentile values are P_1 and P_{10}, respectively. Then, if the distribution on the logarithmic scale is discretized such that:

$$\log p_i = \log p_1 + (i-1)\,\frac{\log p_{10} - \log p_1}{9} \quad \text{for } i = 1, 2, \ldots 11, \tag{24}$$

Then, it can be shown that the cell probabilities are as follows:

Cell	Probability, f_2	Assigned the probability to
1	0.050	p_1
2	0.050	p_2
3	0.080	p_3
4	0.110	p_4
5	0.135	p_5
6	0.150	p_6
7	0.135	p_7
8	0.110	p_8
9	0.080	p_9
10	0.050	p_{10}
11	0.050	p_{11}

From equation 24 we can write:

$$p_i = p_1 (p_{10}/p_1)^{\frac{i-1}{9}}$$

Let $p_{10}/p_1 = r$, then $p_i = p_1 r^{(\frac{i-1}{9})}$

The mean and variance of the distribution can then be written as

$$\bar{p} = \sum_{i=1}^{11} p_i f_i = \sum_{i=1}^{11} p_1 r^{(\frac{i-1}{9})} f_i$$

$$\bar{p} = p_1 \sum_{i=1}^{11} r^{(\frac{i-1}{9})} f_i \tag{25}$$

and

$$\sigma^2 = \sum_{i=1}^{11} p_i^2 f_i - \bar{p}^2$$

$$\sigma^2 = \sum_{i=1}^{11} p_1^2 r^{\frac{2(i-1)}{9}} f_i - \bar{p}^2 \tag{26}$$

where $\bar{p}$ is given by equation 25.

Thus, given the 5th (p_1) and 95th (p_{10}) percentile values of a basic event failure probability, and assuming a lognormal distribution, the mean and variance can be obtained from equations 25 and 26, respectively.

Let us now induce a change in the basic event distribution through the following mechanism. Keeping the p_i's the same, we shall change f_i's such that

$$\Delta f_i = \frac{(i-6)}{5} f_i \tag{27}$$

This means that for p_{11} ($i = 11$), we are inducing a 100% increment in the probability. For $i = 7$ to 10, f_i's increase, $f_i = 0$ for $i = 6$, and f_i's decrease for $i = 1$ to 5, such that

$$\sum_{i=1}^{11} \Delta f_i = 0 ,$$

generating a bonafide probability distribution. The purpose of such an automated mechanism to induce variations in the failure probability distribution is to study the sensitivity of the system failure probability upon the high frequency tails of the basic event distributions (which stems from a lack of adequate data for some of the components).

As an example, let us consider a distribution with the 5th percentile value as p_1 and $(p_{10}/p_1) = r = 10$. Then, from equation 25 we can obtain

$$\text{mean } p_m = 4.45\ p_1$$

and variance, from equation 26 is

$$\sigma_m{}^2 = 9.04\ p_1^2.$$

After introducing changes in the distribution according to equation 27 the variations in the mean and variance are calculated, and are

$$\frac{\Delta p_m}{p_m} = 32.7\%$$

and

$$\frac{\Delta \sigma_m^2}{\sigma_m^2} = 18.3\%$$

The impact of such a distribution upon the Top event can now be calculated in terms of the significance indices.

From the definition, the percentage change in the Top event mean due to changes in the component mean is given by

$$\frac{\Delta \bar{p}}{\bar{p}} = S_m^{(1)} \left(\frac{\Delta \bar{p}_m}{\bar{p}_m}\right) \tag{28}$$

Also, percentage change in Top event variance due to percentage change in the basic event variance is given by

$$\frac{\Delta \sigma^2}{\sigma^2} = S_m^{(2)} \left(\frac{\Delta \bar{p}_m}{\bar{p}_m}\right) \tag{29}$$

We can also calculate the total percentage change in the Top event mean and variance due to changes in all or a group of components as

$$\frac{\Delta\bar{p}}{\bar{p}} = \sum_{j=1}^{k} S_{mj}^{(1)} \left(\frac{\Delta\bar{p}_j}{\bar{p}_j}\right) \tag{30}$$

$$\frac{\Delta\sigma^2}{\sigma^2} = \sum_{j=1}^{k} S_{mj}^{(2)} \left(\frac{\Delta\bar{p}_j}{\bar{p}_j}\right) \tag{31}$$

where

$S_{mj}^{(1)}$ = first significance index for component j, j=1, 2...k

$S_{mj}^{(2)}$ = second significance index for component j, j=1, 2...k

$\bar{p}_j$ = mean probability of failure for component j

$\Delta\bar{p}_j$ = changes induced in the mean probability of failure for component j

Thus, for the given example, if $S_m^{(1)} = 0.8$, and $S_m^{(2)} = 0.6$, then the percentage change in the Top event mean and variance are

$\left(\frac{\Delta p}{p}\right)$ = (0.8) (0.327) = 0.26 or 26%

$\left(\frac{\Delta\sigma^2}{\sigma^2}\right)$ = (0.6) (0.327) = 0.20 or 20%

Summary and Conclusions

The frequency of the sequence of events, leading to the top event, e.g., core degradation or simply a system unavailability, is sensitive to some of the basic events (e.g., random failure of components, human errors, and test and maintenance). We would like to identify those basic events and rank them in order of importance to their contribution to the top event. An important aspect of the component ranking is its sensitivity upon uncertainty characteristics of the basic event probabilities. We have investigated two related problems in arriving at a ranking of components for a system: (1) sensitivity of the basic event probability distribution under various modeling assumptions of the prior distribution; and (2) sensitivity of the ranking of the components upon uncertainty characteristics of the basic events.

It was found that when using subjective interpretation of probability and Bayes' theorem, full awareness of the interpretation and sensitivity of the posterior distribution upon the choice of prior distribution and likelihood function is essential. The plant model, such as fault trees, should be defined such that the posterior distribution of the basic events is likelihood-dominated. Some of the empirical observations in this paper will help in developing data bases that are relatively less sensitive to the choice of the prior distributions. In cases where the posterior distribution is unstable, a sensitivity analysis and adequate rationale for arriving at the posterior distribution must be developed. The posterior distribution becomes highly sensitive when the evidence is in the form of zero failures in n demands; in such cases, special care is warranted in the formulation of the prior distribution.

To develop ranking of basic events which contribute significantly to the mean and uncertainty of the top level event probability, we defined three measures of importance (or Significant Indices) which are:

1. Percent change in Top Level Mean induced due to a unit percentage change in the basic event Mean.
2. Percent change in Top Event Mean due to a unit percent change of the basic event Variance.
3. Percent change in Top Event Variance due to a unit percent change of the basic event Variance.

The inputs required to find the three safety indices are the logic models (Event Trees/Fault Trees or the Minimal Cut Sets), and the probability distribution of the unavailability of each basic event. The method consists of formulating appropriate multilinear expressions of the Mean and Variance of the Top event--in terms of the mean and variance of the basic events. We then obtain the partial derivatives of the Top Event Mean and Variance with respect to the Mean and Variance of the basic events which in turn yield the values of three Significance Indices for each basic event. The output of the method is a ranking of the basic events based on the corresponding Significance Indices.

A method is devised to analyze the sensitivity of the component ranking upon the basic event probability distribution model. Since the component ranking is based on the Significance Indices, variations in the Significance Indices of the components are obtained due to variations in the component failure probability distributions. If the new set of Significance Indices does not warrant a substantial change in the ranking of the components, then the PRA results are defined here as robust. If the component ranking is very sensitive to the basic event failure distribution models, then the PRA results are not considered as robust.

An important feature of the Significance Indices is the convenience of finding the percentage variations of the mean and variance of the top event due to single or multiple variations in the mean and variance of one or more basic events in the system model. Thus, the Significance Indices provide a simplistic method for performing sensitivity analysis of system failure probability due to improvement such as design changes or deterioration (due to component aging) of the components.

REFERENCES

1. Zion Units 1 and 2 Probabilistic Safety Study, Commonwealth Edison Company, 1981.

2. Indian Point Probabilistic Safety Study, Power Authority of the State of New York/Consolidated Edison Company of New York, Inc., 1982.

3. Oconee Probabilistic Safety Study, Nuclear Safety Analysis Center/ Duke Power Company (in progress).

4. Oyster Creek Probabilistic Safety Study.

5. S. Ahmed, D. R. Metcalf, and W. W. Weaver, Probabilistic Risk Assessment Techniques for Nuclear Power Plants, BAW-1677, Babcock & Wilcox Company, Nuclear Power Generation Division, Lynchburg, Virginia (October 1981).

6. G. W. Parry and P. W. Winter, Characterization and Evaluation of Uncertainty in Probabilistic Risk Analysis, Nuclear Safety 22, 28 (1981).

7. S. Kaplan and B. J. Garrick, On the Use of a Bayesian Reasoning in Safety and Reliability Decision--Three Examples, Nuclear Technology -44, 231, (1979).

8. S. Kaplan, B. J. Garrick, and P. Bieniarz, On the Use of Bayes' Theorem in Assessing the Frequency of Anticipated Transients, Nuclear Engineering and Design 65, 23 (1981).

9. G. Apostolakis, S. Kaplan, B. J. Garrick, and R. J. Duphily, Data Specialization for Plant Specific Risk Studies, Nuclear Engineering and Design 56, 321 (1980).

10. S. Ahmed, "Pitfalls in Rare-Event Data Analysis and Some Suggestions," Workshop on Low-Probability/High-Consequence Risk Analysis, Arlington, Virginia, June 15-17, 1982 (sponsored by NRC and Society for Risk Analysis, Oak Ridge National Laboratory).

11. U.S. Nuclear Regulatory Commission, Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, WASH-1400 (NUREG-75/014), Appendix III, (October 1975).

12. U.S. Nuclear Regulatory Commission, Data Summaries of Licensee Event Reports, NUREG/CR-1205, NUREG/CR-1362, NUREG/CR-1363 (1980).

13. IEEE Guide to the Collection and Presentation of Electrical, Electronic and Sensing Component Reliability Data for Nuclear Power Generation Stations, IEEE Std. 500-1977.

14. A. D. Swain and H. E. Guttman, Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications (Draft Report), NUREG/CR-1278 (1980).

15. S. Ahmed, D. R. Metcalf, R. E. Clark, and J. A. Jacobsen, BURD - A Computer Program for Bayesian Updating of Reliability Data, NPGD-TM-582, Babcock and Wilcox Company, Lynchburg, Virginia (August 1981).

16. S. Ahmed, R. E. Clark, and D. R. Metcalf, A Method for Propagating Uncertainty in Probabilistic Risk Assessment, Nuclear Technology, to be published (November 1982).

17. S. Ahmed, A. K. Bhattacharyya, and C. Kudtarkar, "Sensitivity Analysis for Fault/Tree/Event Tree Calculated Risk," Draft Report, Nuclear Safety Analysis Center, Electric Power Research Institute, Palo Alto, California, August 1982.

18. H. Lambert, "Fault Trees for Decision Making in Systems Analysis," Ph.D. Thesis, University of California, Berkeley, California, 1975.

19. J. B. Fussell, "How to Hand-Calculate System Reliability Characteristics," IEEE Trans. on Rel., R-24 (3), (1975).

20. R. E. Barlow and F. Proschan, "Importance of System Components and Fault Tree Analysis," Operations Research Center, University of California, Berkeley, California, Report ORC 74-3 (1974).

21. Vicki M. Bier, "A Measure of Uncertainty Importance for Components in Fault Trees," Ph.D. Thesis, MIT, January 1983.

SUGGESTIONS

In the course of performing the study in Section V, it became apparent that the sensitivity analysis approach can be expanded to investigate other issues in PRA. Two such areas of potential extension are dependent failure analysis and uncertainty propagation. The concepts of the extensions will be presented here as suggestions for further research. The two cases considered here are:

1. Sensitivity analysis approach to rank dependent failures.
2. Informational needs for uncertainty propagation.

Sensitivity Analysis Approach to Rank Dependent Failures

Although qualitative approaches to identify important Dependent Failures[1] through computer-aided manipulation of Fault Trees exist, a ranking among all the Common cause candidates is difficult to achieve. The currently available parametric quantification procedures can, at best, quantify isolated cases of Dependent Failure events. The limitations of scope and the data requirement do not allow the existing techniques to yield a quantitative ranking of the Common Cause candidates easily. However, the need obviously exists for such a ranking so that the significance of the CCFs can be understood and

---

[1]The terminologies Common Cause Failures (CCF) and Dependent Failures have been used here interchangeably.

assessed. In what follows, an approach is formally presented to quantify the sensitivity of the Top event on the Dependent Failure candidates utilizing techniques similar to the one in the main report. The method uses the system Fault Tree developed to a level such that component failure susceptibility due to Common Causes are identified and quantified. The identification part can be done with some of the available codes such as COMCAN (1), BACFIRE (2), and WAMCOM (3), with little or no modifications. The procedure is to obtain the combination of component failure elements or basic events (Minimal Cut Sets) which will fail the system if they occur together and which contain the common susceptibilities (Common Causes) that can fail various components simultaneously. The Boolean expression for the Top event failure thus generated can then be used to quantify the sensitivity of the Top event Mean frequency upon each of the Common Cause basic event Mean probability by defining an index that measures the sensitivity of the Top event upon the dependent failures as

$$D_{ji} = \frac{\partial \bar{P}}{\partial \bar{\varepsilon}_{ji}} \qquad (1)$$

where

$\bar{P}$ = mean value of the Top event probability

$\bar{\varepsilon}_{ji}$ = mean value of the component i failure due to common cause or susceptibility j (such as moisture, temperature, etc.) which also fail some other components.

D_{ji} = Susceptibility Index for common cause or susceptibility j acting on component i.

The failure probability of a component can be split up into two components: (1) random hardware failure, and (2) common cause failure as follows:

$$p_i = p_{i0} + \sum_j f_{ij} \quad ji \, , \tag{2}$$

where:

p_{i0} = probability of random failure of component i

f_{ij} = when the component i does not fail due to susceptibility j

= 1, when the component i fails due to susceptibility j

(with probability ε_j)

Now the failure probability or frequency of a Top event can be expressed in terms of the component probabilities as follows:

$$P = \sum_{\alpha_1 \ldots \alpha_n}^{0,1} A_{[\alpha_1 \ldots \alpha_n]} \, p_1^{\alpha_1} \ldots p_n^{\alpha_n} \tag{3}$$

In equation 3, $p_1 \ldots p_n$ are n-component probabilities or frequencies, P is the Top event probability, and $A_{[\alpha_1 \ldots \alpha_n]}$ coefficients with values 0, 1, or -1 which are characteristic of the fault tree considered. Using equations 2 and 3, we can obtain the Susceptibility Index of each component i due to susceptibility j.

<u>Example</u>

Let us consider that the failure of a specific system is given by the expression

$$\text{Top} = X_1 + X_2X_3 + X_4X_5 + X_6$$

where X_1, X_2,...X_6 are the basic events. Let us consider that the susceptibilities that may cause the basic events are moisture, temperature, vibration, and impact. Specifically, let us consider that moisture affects X_2, X_5, and X_6; temperature affects X_1, X_2, and X_5; vibration affects X_2, X_4, and X_6 and impacts affect X_1, X_3, and X_5. For simplicity, here we will consider that all equipment is in the same location so that all the applicable basic events are affected by a particular susceptibility. Then, we can write,

$$p_1 = p_{10} + \varepsilon_{1T} + \varepsilon_{1I}$$

where p_{10} is the probability of the basic event X_1 due to random failure, ε_{1T} is the probability of X_1 due to temperature T, and ε_{1I} is the probability of X_1 due to impact I. The total failure probability for X_1 is p_1. Similar expressions can be written for p_2, p_3,...p_6.

The dependent failure Sensitivity Index can now be obtained using equation 1. The results for this example are provided in Table 1. A ranking of D_{ji} can now be obtained based on the values of p_i.

The numerical values of p_i, the total failure probability of basic event i, can be extracted from such data sources as NPRDS, IEEE Std 500, LER equipment summaries, WASH-1400, and plant experience. These data sources generally do not separate the failure rates for random failures and dependent failures; rather the dependent and random failures are combined to provide total failure rates. Therefore, the failure data obtained from these data sources can be considered to be appropriate for use in the formulations presented.

Table 1. Dependent failure susceptibility index D_{ij} for the event i due to susceptibility j

Event i \ Susceptibility j	Temperature T	Vibration V	Moisture M	Impact I
x_1	1	0	0	1
x_2	p_3	p_3	p_3	0
x_3	0	0	0	p_2
x_4	0	p_5	0	0
x_5	p_4	0	p_4	p_4
x_6	0	1	1	0

The method looks at one common cause failure at a time and finds the sensitivity of the Top event failure probability to this source. The mathematical formulation states that the sensitivity is simply dependent on two factors: (a) the structure of the fault tree, and (b) the total probability of the basic events or component failure elements. The assumption used in this formulation is that the random failure probability of a component is not affected by any existing common cause susceptibility.

The method outlines the procedure for a preliminary analysis of the dependent failures in order to develop a ranking of the CCFs with minimal effort. An important favorable feature of the method is that the failure data requirements are not as demanding. In fact, in this model, the major data required are the total probability of individual failure (random failure and failure due to all the common causes that are being considered) of the component of interest from past observation. This is in contrast to the beta factor [4] or binomial failure rate [5] method of dependent failure quantification where the assessment of the parameters has to be based on multiple failure observations from the industry.

Informational Needs for Uncertainty Propagation

Introduction

Quantification of the probability of rare events such as failure of a safety system on demand, or the frequency of Core Melt due to an initiating event (such as Loss of Offsite Power) is associated with large uncertainty due to a number of reasons. These include, for example, the modeling uncertainty and the inadequacy of the failure history even at the component level. For a particular system failure model (e.g., a Fault Tree), the existing approaches calculate Top level uncertainty by propagating the uncertainties of the basic-event failure probabilities through the Fault Tree in one form or another, e.g., Monte-Carlo simulation, propagation of Moments, and propagation of Discrete Probability Distributions. The data analysis requirements for uncertainty propagation can be considered reduced by prejudging information needs obtained through sensitivity analysis.

The proposed PRA approach can identify the significant issues in the Top event uncertainty quantification, such as:

- How important are the uncertainties of the component unavailability due to random failures as compared to those of the Common Cause failures? How important are the uncertainties of the initiating events?

- How is the uncertainty of the frequency of the Top event related to the structure of the Fault Tree? Given a reasonably complex Fault Tree, what are the subset of the basic events whose uncertainties impact Top level uncertainty significantly?
- If the uncertainty of the Top event is sensitive to and dominated by only a few components (such as some of the Common Cause basic events) what computational procedure will adequately serve the purpoe of uncertainty quantification?
- If it is found that uncertainty propagation in the massive scale is unnecessary, what level of effort should be invested in formulating and structuring the data base?

Approach

The sensitivity analysis approach for risk assessment as developed in this study furnishes the information needs for uncertainty propagation. Before a full scale uncertainty propagation is pursued, it is proposed to identify the basic events which impact the Top event uncertainty significantly. This can be achieved by obtaining the ranking of the basic events based upon the Significance Indices $S_m^{(2)}$ and $S_m^{(3)}$ as defined by equation 18) in Section V. Significance index $S_m^{(2)}$ gives the percentage change in the Top event Variance induced by a percentage change in the basic event Mean.

Thus, $S_m^{(2)}$ identifies those components where the Mean values contribute significantly to the Top event Variance. Safety Significance index $S_m^{(3)}$ yields the percentage change in the Top event variance induced by a percentage change in the basic event Variance and in the process identifies those components uncertainties where Variances contribute significantly to the Top event Variance. Based on the ranking via these Significance Indices, the selection of the significant elements can follow in a straightforward fashion. The resources can then be concentrated on these major contributors in assessing their Mean values and Variances more accurately. The level of effort expended on insignificant element can be economized commensurately.

The methodology additionally provides a framework to determine the shape of the Top event distribution for various class of Fault Tree structure. For example, the uncertainty propagation for various mixes of AND and OR dominated Tree can be studied and generic conclusions derived. Such generic conclusions along with the knowledge of the dominant contributors to Top event uncertainty can potentially aid development of uncertainty propagation technique to a considerable extent.

REFERENCES

1. D. M. Rasmuson, N. H. Marshall, J. R. Wilson, and G. R. Burdick, COMCAN II - A Computer Program for Automated Common Cause Failure Analysis, TREE-1361, EG&G Idaho, Inc., Idaho Falls, Idaho, 1979.

2. C. L. Cate and J. B. Fussell, BACFIRE - A Computer Program for Common Cause Failure Analysis, NERS-7702, Nuclear Engineering Department, University of Tennessee, Knoxville, Tennessee, 1977.

3. B. F. Putney, WAMCOM, Common Cause Methodologies Using Large Fault Trees, SAI/SR-226-80-PA (Draft), Science Applications, Inc., Palo Alto, California, 1981.

4. K. N. Fleming, A Reliability Model for Common Mode Failure in Redundant Safety Systems, in Proceedings of the Sixth Annual Pittsburgh Conference on Modeling and Simulation, April 23-25, 1975, General Atomic Report GA-A 13284.

5. W. E. Vesely, Estimating Common Cause Failure Probability in Reliability and Risk Analyses: Marshall-Olkin Specializations, Proceedings, International Conference on Nuclear Systems Reliability Engineering and Risk Assessment, Gatlinburg, Tennessee, June, 1977.

SUMMARY AND CONCLUSIONS

Two basic approaches are explored in this dissertation to help the decision-maker in arriving at rational decisions pertaining to high technology energy systems. One approach deals with formal decision analysis using Multiahaibute Utility Theory (MAUT). The other approach uses Probabilistic Risk Assessment technology for risk management pertaining to high technology systems.

The MAUT is formulated for decisions pertaining to siting of nuclear power plants and assessing risk from various proliferation routes of acquiring nuclear weapons by nonnuclear countries and terrorist groups. The decision models, in both the cases, produced results consistent with the results obtained through other qualitative and quantitative approaches used in the nuclear industry, thus validating the MAUT formulations.

New quantitative approaches are presented to strengthen the Probabilistic Risk Assessment tools in predicting and reducing the risk associated with the high technology systems. The characteristics of point-estimate risk calculations are investigated. Also, an uncertainty propagation method based on the Discrete Propagation Distribution is explored as compared to the Monte Carlo method and Histogram method. Finally, a framework is presented to investigate the robustness of the PRA results based on three Significance Indices that denote the percentage variations of the Top event mean and variance due to variations induced in each of the basic events.

Further research is suggested, as an extension of the Significance Indices approach, for quantitative determinations of the Common Cause Failures and for informational needs for uncertainty propagation in PRA.

ACKNOWLEDGEMENT

I would like to express my deep appreciation to Professor D. M. Roberts for his continued guidance, encouragement, and efforts to coordinate the graduate program. I would also like to thank my committee members Dr. B. I. Spinrad, Dr. H. T. David, Dr. R. A. Danofsky, Dr. S. K. Adams, and Dr. S. P. Carfagno for their guidance, suggestions, and critiques on the dissertation. My special thanks are directed to Dr. S. P. Carfagno of the Franklin Research Center for his encouragement, his help and in providing the opportunity to pursue the Ph.D. program.

The development of special computer programs has been possible as a result of the valuable discussions, suggestions, critique, and help provided by a number of past and present colleagues and professors. Prominent among them are Dr. D. R. Metcalf, Dr. R. E. Clark, Dr. A. A. Husseiny, Dr. A. K. Bhattacharyya, and Dr. Hang Youn Cho.

I would also like to thank Ms. Carolyn Taylor and Ms. Carmen Schwartz for their excellent work in typing the manuscript.

Finally, I would like to express my appreciation to my wife Naheed for her patience and encouragement, without which the dissertation could not have been completed.